Error-Correction Factor Models for High-dimensional Cointegrated Time Series

Yundong Tu, Qiwei Yao and Rongmao Zhang

Peking University, London School of Economics and Zhejiang University

Supplementary Material

Lemmas and Technical Proofs

Lemma 5. *Under Condition 1 or conditions of Theorem 3, we have*

$$\frac{1}{n}\sum_{t=1}^{n-1}(\widehat{\mathbf{A}}_2'\mathbf{y}_t\mathbf{y}_t'\widehat{\mathbf{A}}_2 - \mathbf{A}_2'\mathbf{y}_t\mathbf{y}_t'\mathbf{A}_2) = o_p(1). \tag{S0.1}$$

Proof. We first show the case with fixed p. Since $\{\mathbf{x}_{t2}, \mathbf{f}_t, \boldsymbol{\varepsilon}_t\}$ is α mixing with mixing coefficients α_m satisfying

$$\sum_{m=1}^{\infty}\alpha_m^{1-1/\gamma} < \infty, \tag{S0.2}$$

it follows that $\{\nabla\mathbf{y}_t\}$ is a α mixing process with mixing coefficients satisfying (S0.2). Thus, by Theorem 3.2.3 of Lin and Lu (1997), there exists a p-dimensional Gaussian process $\mathbf{g}(t)$ such that

$$\mathbf{y}_{[nt]}/\sqrt{n} \xrightarrow{d} \mathbf{g}(t), \text{ on } D[0,1]. \tag{S0.3}$$

From (S0.3) and the continuous mapping theorem, it follows that

$$\frac{1}{n^2}\sum_{t=1}^{n}\mathbf{y}_t\mathbf{y}_t' \xrightarrow{d} \int_0^1 \mathbf{g}(t)\mathbf{g}'(t)\,dt. \tag{S0.4}$$

Further, by $\mathrm{E}||\mathbf{x}_{t2}||^{2\gamma} < \infty$ for some $\gamma > 1$, we have

$$\max_{1\le t\le n} ||\mathbf{x}_{t2} - \mathrm{E}\mathbf{x}_{t2}||/\sqrt{n} = o_p(1), \text{ and } \frac{1}{n}\sum_{t=1}^{n}||\mathbf{x}_{t2} - \mathrm{E}\mathbf{x}_{t2}|| = O_p(1). \tag{S0.5}$$

Combining (S0.3) and (S0.5) (see Lemma 7 of ZRY) yields

$$\frac{1}{n^{3/2}}||\sum_{t=1}^{n}\mathbf{y}_t\mathbf{x}_{t2}'||_2 = o_p(1). \tag{S0.6}$$

On the other hand, by $\nabla\mathbf{x}_{t1} = \mathbf{A}_1'\nabla\mathbf{y}_t$, we know $(\nabla\mathbf{x}_{t1}, \mathbf{x}_{t2})$ is also α mixing with mixing coefficients satisfying (S0.2). As a result, by the proof of Theorem 1 in ZRY,

$$||\widehat{\mathbf{A}}_2 - \mathbf{A}_2||_2 = O_p(1/n). \tag{S0.7}$$

By (S0.4), (S0.6) and (S0.7), we have

$$\begin{aligned}
& ||\frac{1}{n}\sum_{t=1}^{n-1}(\widehat{\mathbf{A}}_2'\mathbf{y}_t\mathbf{y}_t'\widehat{\mathbf{A}}_2 - \mathbf{A}_2'\mathbf{y}_t\mathbf{y}_t'\mathbf{A}_2)||_2 \\
= \;& ||(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\frac{\sum_{t=1}^{n-1}\mathbf{y}_t(\mathbf{A}_2'\mathbf{y}_t)'}{n} + \frac{\sum_{t=1}^{n-1}(\mathbf{A}_2'\mathbf{y}_t)\mathbf{y}_t'}{n}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2) \\
& +(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\frac{\sum_{t=1}^{n-1}\mathbf{y}_t\mathbf{y}_t'}{n}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)||_2 \\
= \;& ||(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\frac{\sum_{t=1}^{n-1}\mathbf{y}_t\mathbf{x}_{t2}'}{n} + \frac{\sum_{t=1}^{n-1}\mathbf{x}_{t2}\mathbf{y}_t'}{n}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2) + (\widehat{\mathbf{A}}_2 - \mathbf{A}_2)\frac{\sum_{t=1}^{n-1}\mathbf{y}_t\mathbf{y}_t'}{n}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'||_2 \\
= \;& o_p(1).
\end{aligned} \tag{S0.8}$$

Next, consider the case $p = o(n^c)$. Let $\boldsymbol{\varsigma}_t$ be a k-dimensional $I(1)$ process such that $\nabla\boldsymbol{\varsigma}_t = \mathbf{v}_t$. By Remark 2 of ZRY, we know that Condition 3 (i) and Remark 3 of ZRY hold for $\boldsymbol{\varsigma}_t$. Let $\mathbf{M}_1, \mathbf{M}_2$ be $k\times(p-r)$ and $k\times r$

matrices such that $\mathbf{M}$ given in (i) of Condition 3 satisfying $\mathbf{M}' = (\mathbf{M}_1, \mathbf{M}_2)$. Let $\mathbf{F}(t) = (F^1(t), \cdots, F^k(t))'$ be defined as in ZRY and $\bar{\boldsymbol{\varsigma}} = \frac{1}{n}\sum_{t=1}^n \boldsymbol{\varsigma}_t$, then

$$\begin{aligned} & ||\frac{1}{n^2}\sum_{t=1}^n (\mathbf{x}_{t1} - \bar{\mathbf{x}}_1)(\mathbf{x}_{t1} - \bar{\mathbf{x}}_1)' - \mathbf{M}_1' \int_0^1 \mathbf{F}(t)\mathbf{F}'(t)\, dt \mathbf{M}_1||_2 \\ = & ||\mathbf{M}_1' \left(\frac{1}{n^2}\sum_{t=1}^n (\boldsymbol{\varsigma}_t - \bar{\boldsymbol{\varsigma}})(\boldsymbol{\varsigma}_t - \bar{\boldsymbol{\varsigma}})' - \int_0^1 \mathbf{F}(t)\mathbf{F}'(t)\, dt \right) \mathbf{M}_1||_2 = o_p(1). \end{aligned} \quad \text{(S0.9)}$$

By Remark 3 of ZRY, we have $\lambda_{\min}\left(\int_0^1 \mathbf{F}(t)\mathbf{F}'(t)\, dt\right) \geq 1/k$ in probability. Since $c_1 \leq \lambda_{\min}(\mathbf{M}) \leq \lambda_{\max}(\mathbf{M}) \leq c_2$, it follows $\lambda_{\min}\left(\mathbf{M}_1' \int_0^1 \mathbf{F}(t)\mathbf{F}'(t)\, dt \mathbf{M}_1'\right) \geq 1/k$ in probability. Further, for any given $j \geq 0$,

$$\begin{aligned} & ||\frac{1}{n}\sum_{t=1}^{n-j} (\mathbf{x}_{t+j,2} - \bar{\mathbf{x}}_2)(\mathbf{x}_{t2} - \bar{\mathbf{x}}_2)' - \mathrm{Cov}(\mathbf{x}_{t+j,2}, \mathbf{x}_{t2})||_2 \\ = & ||\mathbf{M}_2' \Big(\frac{1}{n}\sum_{t=1}^n [(\mathbf{v}_{t+j} - \bar{\mathbf{v}})(\mathbf{v}_t - \bar{\mathbf{v}})' - \mathrm{Cov}(\mathbf{v}_{t+j}, \mathbf{v}_t)] \Big) \mathbf{M}_2||_2 = o_p(1), \text{ and} \end{aligned} \quad \text{(S0.10)}$$

$$\begin{aligned} ||\frac{1}{n^{3/2}}\sum_{t=1}^{n-j} (\mathbf{x}_{t+j,1} - \bar{\mathbf{x}}_2)(\mathbf{x}_{t2} - \bar{\mathbf{x}}_2)'||_2 &= ||\mathbf{M}_1' \left(\frac{1}{n^{3/2}}\sum_{t=1}^n (\boldsymbol{\varsigma}_{t+j} - \bar{\boldsymbol{\varsigma}})(\mathbf{v}_t - \bar{\mathbf{v}})' \right) \mathbf{M}_2||_2 \\ &= O_p(k/n^{1/2}), \end{aligned} \quad \text{(S0.11)}$$

where $\mathbf{v}_t$ is given in (i) of Condition 3.

By (S0.9)–(S0.11), similar to the proof of Theorem 3 in ZRY, it can be shown that when $k = o(n^{1/2-1/\eta})$,

$$||\widehat{\mathbf{A}}_2 - \mathbf{A}_2||_2 = O_p(p^{1/2}k/n). \quad \text{(S0.12)}$$

Similar to (S0.9), there exists a k-dimensional Gaussian process $\mathbf{w}(t)$ such that

$$||\frac{1}{n^2}\sum_{t=1}^n \mathbf{y}_t\mathbf{y}_t' - \mathbf{A}_1\mathbf{M}_1' \int_0^1 \mathbf{w}(t)\mathbf{w}'(t)\, dt \mathbf{M}_1\mathbf{A}_1'||_2 = o_p(1) \quad \text{(S0.13)}$$

and similar to (S0.11), we can show (S0.6) holds provided $k/n^{1/2} \to 0$ as $n \to \infty$. Thus, by (S0.12) and (S0.13), we also have (S0.8) and complete the proof of Lemma 5. □

Lemma 6. *Under Condition 1,*

$$||\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{y}'_{t-1}(\widehat{\mathbf{A}}_2-\mathbf{A}_2)||_2 = o_p(1),$$

and under the conditions of Theorem 3,

$$||\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{y}'_{t-1}(\widehat{\mathbf{A}}_2-\mathbf{A}_2)||_2 = O_p(p^{1/2}k^2/n^{1/2}). \tag{S0.14}$$

Proof. When p is fixed, similar to (S0.6), we have

$$\frac{1}{n^{3/2}}||\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{y}'_{t-1}||_2 = o_p(1).$$

As a result, it follows from (S0.7) that

$$||\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{y}'_{t-1}(\widehat{\mathbf{A}}_2-\mathbf{A}_2)||_2 = o_p(1). \tag{S0.15}$$

When p tends to infinity as $n \to \infty$, using the same idea as in (S0.11), we can show

$$\frac{1}{n^{3/2}}||\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{y}'_{t-1}||_2 = O_p(k/n^{1/2}). \tag{S0.16}$$

Thus, by (S0.12) and $p \le k = o(n^{1/2})$, it follows that

$$||\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{y}'_{t-1}(\widehat{\mathbf{A}}_2-\mathbf{A}_2)||_2 = O_p(p^{1/2}k^2/n^{1/2}).$$

Thus, we have Lemma 6. □

Lemma 7. *Let* $\mathbf{\Sigma} = \mathrm{E}\{(\mathbf{f}'_{t-1},\cdots,\mathbf{f}'_{t-s})'(\mathbf{f}'_{t-1},\cdots,\mathbf{f}'_{t-s})\}$. *Under Condition 1*

, for any given positive integer s,

$$\frac{1}{n}\Big[\sum_{t=s+1}^{n}(\widehat{\mathbf{f}}'_{t-1},\cdots,\widehat{\mathbf{f}}'_{t-s})'(\widehat{\mathbf{f}}'_{t-1},\cdots,\widehat{\mathbf{f}}'_{t-s})-M\Big]\xrightarrow{p}\mathbf{\Sigma} \tag{S0.17}$$

and under the condition of Theorem 3, in probability

$$\frac{1}{n}\Big[\sum_{t=s+1}^{n}(\widehat{\mathbf{f}}'_{t-1},\cdots,\widehat{\mathbf{f}}'_{t-s})'(\widehat{\mathbf{f}}'_{t-1},\cdots,\widehat{\mathbf{f}}'_{t-s})-M\Big]>0, \tag{S0.18}$$

where $\mathbf{A}>0$ *means that* $\mathbf{A}$ *is a positive definition matrix.*

Proof. By some elementary computation, we have

$$\begin{aligned}\widehat{\mathbf{f}}_t &= [\mathbf{f}_t+\mathbf{B}'\boldsymbol{\varepsilon}_t]+[(\widehat{\mathbf{B}}-\mathbf{B})'(\mathbf{B}\mathbf{f}_t+\boldsymbol{\varepsilon}_t)]+[\widehat{\mathbf{B}}'(\mathbf{D}-\widehat{\mathbf{D}})\mathbf{x}_{t2}]+[\widehat{\mathbf{B}}'\widehat{\mathbf{D}}(\mathbf{A}_2-\widehat{\mathbf{A}}_2)'\mathbf{y}_{t-1}]\\ &\equiv \sum_{i=1}^{4}\boldsymbol{\zeta}_{t,i}. \end{aligned} \tag{S0.19}$$

Next, we first show (S0.17) holds for fixed p. By (S0.33) (see below), we have

$$||\widehat{\mathbf{B}}-\mathbf{B}||_2=O_p(n^{-1/2}), \tag{S0.20}$$

which gives

$$||\frac{1}{n}\sum_{t=s+1}^{n}(\boldsymbol{\zeta}'_{t-1,2},\cdots,\boldsymbol{\zeta}'_{t-s,2})'(\boldsymbol{\zeta}'_{t-1,2},\cdots,\boldsymbol{\zeta}'_{t-s,2})||_2=o_p(1). \tag{S0.21}$$

Similarly, by (S0.29) (see below) and (S0.7), we have

$$\sum_{i=3}^{4}||\frac{1}{n}\sum_{t=s+1}^{n}(\boldsymbol{\zeta}'_{t-1,i},\cdots,\boldsymbol{\zeta}'_{t-s,i})'(\boldsymbol{\zeta}'_{t-1,i},\cdots,\boldsymbol{\zeta}'_{t-s,i})||_2=o_p(1). \tag{S0.22}$$

On the other hand, by law of large numbers for α-mixing process, we get

$$\frac{1}{n}\Big[\sum_{t=s+1}^{n}(\boldsymbol{\zeta}'_{t-1,1},\cdots,\boldsymbol{\zeta}'_{t-s,1})'(\boldsymbol{\zeta}'_{t-1,1},\cdots,\boldsymbol{\zeta}'_{t-s,1})-M\Big]\xrightarrow{p}\mathbf{\Sigma}. \tag{S0.23}$$

Combining (S0.21)–(S0.23) yields that

$$
\begin{aligned}
& \frac{1}{n}\sum_{t=s}^{n}[(\widehat{\mathbf{f}}_{t-1})',\cdots,(\widehat{\mathbf{f}}_{t-s})']'[(\widehat{\mathbf{f}}_{t-1})',\cdots,(\widehat{\mathbf{f}}_{t-s})'] \\
= & \frac{1}{n}\sum_{t=s+1}^{n}(\sum_{i=1}^{4}\boldsymbol{\zeta}'_{t-1,i},\cdots,\sum_{i=1}^{4}\boldsymbol{\zeta}'_{t-s,i})'(\sum_{i=1}^{4}\boldsymbol{\zeta}'_{t-1,i},\cdots,\sum_{i=1}^{4}\boldsymbol{\zeta}'_{t-s,i}) \\
= & \frac{1}{n}\sum_{t=s+1}^{n}(\boldsymbol{\zeta}'_{t-1,1},\cdots,\boldsymbol{\zeta}'_{t-s,1})'(\boldsymbol{\zeta}'_{t-1,1},\cdots,\boldsymbol{\zeta}'_{t-s,1})+o_p(1)\xrightarrow{p}\boldsymbol{\Sigma}
\end{aligned}
$$

and (S0.17) follows.

Now, we turn to show the case with p varying with n. Since $p=o(n^{1/2})$, (S0.23) still holds. Note that $\frac{1}{n}\sum_{t=s}^{n}(\boldsymbol{\zeta}'_{t-1,i},\cdots,\boldsymbol{\zeta}'_{t-s,i})'(\boldsymbol{\zeta}'_{t-1,i},\cdots,\boldsymbol{\zeta}'_{t-s,i})\geq\mathbf{0}$ for $i=1,\cdots,4$. For the proof of (S0.18), it is enough to show for all $1\leq i\neq j\leq 4$,

$$
||\frac{1}{n}\sum_{t=s+1}^{n}(\boldsymbol{\zeta}'_{t-1,i},\cdots,\boldsymbol{\zeta}'_{t-s,i})'(\boldsymbol{\zeta}'_{t-1,j},\cdots,\boldsymbol{\zeta}'_{t-s,j})||_2=o_p(1). \qquad \text{(S0.24)}
$$

We only give $i=1,j=4$ in details, other cases can be shown similarly. Since $\mathbf{y}_t=\mathbf{A}\mathbf{x}_t$, it follows from (2.1) that

$$
\boldsymbol{\zeta}_{t,1}=\mathbf{B}'(\nabla\mathbf{y}_t-\mathbf{D}\mathbf{x}_{t-1,2})=\mathbf{B}'\mathbf{A}\mathbf{e}_t-\mathbf{B}'(\mathbf{D}+\mathbf{A}_2)\mathbf{x}_{t-1,2}=\mathbf{B}'\mathbf{A}\mathbf{M}\mathbf{v}_t-\mathbf{B}'(\mathbf{D}+\mathbf{A}_2)\mathbf{M}_2'\mathbf{v}_{t-1}.
$$

Thus, by the fact that for any $-s-1\leq j\leq s+1$,

$$
||\sum_{t=1}^{n}\sum_{s=1}^{t}\mathbf{v}_s\mathbf{v}_{t+j}||_2=O_p(kn) \qquad \text{(S0.25)}
$$

and (S0.12), we have the left-hand side of (S0.24) is of order $O_p(p^{1/2}k^2/n)=o_p(1)$, where (S0.25) holds because the components of $\mathbf{v}_t$ are independent. Thus, we have (S0.18) and complete the proof of Lemma 7. □

Proof of Theorem 1. Let $\mathbf{b}_i, i=1,\cdots,p$ be the rows of $\mathbf{B}$. Lemmas 5 and 6 implies that for any $1 \le i \le p$,

$$\begin{aligned}\sqrt{n}(\widehat{\mathbf{d}}_i - \mathbf{d}_i) &= \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\mathbf{b}_i\mathbf{f}_t + \varepsilon_t^i)\mathbf{y}'_{t-1}\mathbf{A}_2\right)\left(\frac{1}{n}\sum_{i=1}^{n}(\mathbf{A}'_2\mathbf{y}_{t-1})(\mathbf{A}'_2\mathbf{y}_{t-1})'\right)^{-1} + o_p(1) \\ &= \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\mathbf{b}_i\mathbf{f}_t + \varepsilon_t^i)\mathbf{x}'_{t-1,2}\right)\left(\frac{1}{n}\sum_{i=0}^{n-1}\mathbf{x}_{t2}\mathbf{x}'_{t2}\right)^{-1} + o_p(1). \qquad \text{(S0.26)}\end{aligned}$$

Since $\{\mathbf{x}_{t2}\}$ is α mixing with mixing coefficients satisfying (S0.2), it follows that

$$\frac{1}{n}\sum_{i=0}^{n-1}\mathbf{x}_{t2}\mathbf{x}'_{t2} \xrightarrow{p} \mathrm{E}(\mathbf{x}_{t2}\mathbf{x}'_{t2}) =: \mathbf{\Pi}. \qquad \text{(S0.27)}$$

On the other hand, by central limit theory (CLT) for α-mixing process $\{(\mathbf{b}_i\mathbf{f}_t + \varepsilon_t^i)\mathbf{x}'_{t-1,2},\ 1 \le i \le p\}$, there exists a $pr \times pr$ matrix $\mathbf{\Lambda}$ such that

$$\frac{1}{\sqrt{n}}\left(\sum_{t=1}^{n}(\mathbf{b}_1\mathbf{f}_t + \varepsilon_t^1)\mathbf{x}'_{t-1,2}, \cdots, \sum_{t=1}^{n}(\mathbf{b}_p\mathbf{f}_t + \varepsilon_t^p)\mathbf{x}'_{t-1,2}\right) \xrightarrow{d} N(0, \mathbf{\Lambda}). \qquad \text{(S0.28)}$$

Thus, by (S0.27) and (S0.28), we have

$$\sqrt{n}(\mathrm{vech}(\widehat{\mathbf{D}}) - \mathrm{vech}(\mathbf{D})) \xrightarrow{d} N(0, \mathbf{\Pi}^{-1}\mathbf{\Lambda}\mathbf{\Pi}^{-1}). \qquad \text{(S0.29)}$$

Further, by (S0.29) and (S0.7), it is easy to show that

$$||\widehat{\mathbf{C}} - \mathbf{C}||_2 = ||(\widehat{\mathbf{D}} - \mathbf{D})\mathbf{A}'_2 + \widehat{\mathbf{D}}(\widehat{\mathbf{A}}'_2 - \mathbf{A}'_2)||_2 = O_p(n^{-1/2}).$$

Next, we show (b) of Theorem 1. Observe that

$$\widehat{\mathbf{v}}_t = \nabla\mathbf{y}_t - \widehat{\mathbf{D}}\widehat{\mathbf{A}}'_2\mathbf{y}_{t-1} = (\nabla\mathbf{y}_t - \mathbf{D}\mathbf{x}_{t-1,2}) - (\widehat{\mathbf{D}} - \mathbf{D})[(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\mathbf{y}_{t-1} + \mathbf{x}_{t-1,2}] - \mathbf{D}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\mathbf{y}_{t-1},$$

which means that

$$
\begin{aligned}
& \frac{1}{n}\sum_{t=1}^{n-j}[\widehat{\mathbf{v}}_{t+j}\widehat{\mathbf{v}}_t' - \mathrm{E}(\nabla\mathbf{y}_{t+j} - \mathbf{D}\mathbf{x}_{t+j-1})(\nabla\mathbf{y}_t - \mathbf{D}\mathbf{x}_{t-1})'] \\
= \ & \frac{1}{n}\sum_{t=1}^{n-j}[(\nabla\mathbf{y}_{t+j} - \mathbf{D}\mathbf{x}_{t+j-1})(\nabla\mathbf{y}_t - \mathbf{D}\mathbf{x}_{t-1})' - \mathrm{E}(\nabla\mathbf{y}_{t+j} - \mathbf{D}\mathbf{x}_{t+j-1})(\nabla\mathbf{y}_t - \mathbf{D}\mathbf{x}_{t-1})'] \\
& +(\widehat{\mathbf{D}} - \mathbf{D})\left(\frac{1}{n}\sum_{t=1}^{n-j}[(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\mathbf{y}_{t+j-1} + \mathbf{x}_{t+j-1,2}][(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\mathbf{y}_{t-1} + \mathbf{x}_{t-1,2}]'\right)(\widehat{\mathbf{D}} - \mathbf{D})' \\
& +\mathbf{D}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\left(\frac{1}{n}\sum_{t=1}^{n-j}\mathbf{y}_{t+j-1}\mathbf{y}_{t-1}'\right)(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)\mathbf{D}' \\
& -\frac{1}{n}\sum_{t=1}^{n-j}(\nabla\mathbf{y}_{t+j} - \mathbf{D}\mathbf{x}_{t+j-1,2})\{[\mathbf{y}_{t-1}'(\widehat{\mathbf{A}}_2 - \mathbf{A}_2) + \mathbf{x}_{t-1,2}'](\widehat{\mathbf{D}} - \mathbf{D})' + \mathbf{y}_{t-1}'(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)\mathbf{D}'\} \\
& -\frac{1}{n}\sum_{t=1}^{n-j}\{(\widehat{\mathbf{D}} - \mathbf{D})[(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\mathbf{y}_{t+j-1} + \mathbf{x}_{t+j-1,2}] + \mathbf{D}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\mathbf{y}_{t+j-1}\}(\nabla\mathbf{y}_t - \mathbf{D}\mathbf{x}_{t-1,2})' \\
& +\frac{1}{n}\sum_{t=1}^{n-j}[(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'\mathbf{y}_{t+j-1}\mathbf{y}_{t-1}' + \mathbf{x}_{t+j-1,2}\mathbf{y}_{t-1}'](\widehat{\mathbf{A}}_2 - \mathbf{A}_2)\mathbf{D}' \\
& +\frac{1}{n}\sum_{t=1}^{n-j}\mathbf{D}(\widehat{\mathbf{A}}_2 - \mathbf{A}_2)'[\mathbf{y}_{t+j-1}\mathbf{y}_{t-1}'(\widehat{\mathbf{A}}_2 - \mathbf{A}_2) + \mathbf{y}_{t+j-1}\mathbf{x}_{t-1,2}'](\widehat{\mathbf{D}} - \mathbf{D})'. \qquad \text{(S0.30)}
\end{aligned}
$$

By (S0.7), (S0.29) and the law of large numbers, we have that the spectral norm of the last six terms of the right-hand side in (S0.30) is $O_p(n^{-1})$. And by CLT of α mixing process, for any given j, the first term of the right-hand side of (S0.30) is $O_p(n^{-1/2})$. Similarly, we can show

$$\Big|\Big|\frac{1}{n}\sum_{t=1}^{n-j}\bar{\mathbf{v}}\widehat{\mathbf{v}}_t'\Big|\Big|_2 = O_p(n^{-1}). \qquad \text{(S0.31)}$$

Thus,

$$||\widehat{\mathbf{\Sigma}}_v(j) - \mathbf{\Sigma}_v(j)||_2 = O_p(n^{-1/2}), \qquad \text{(S0.32)}$$

where $\mathbf{\Sigma}_v(j) = \mathrm{E}(\nabla\mathbf{y}_{t+j} - \mathbf{D}\mathbf{x}_{t+j-1})(\nabla\mathbf{y}_t - \mathbf{D}\mathbf{x}_{t-1})'$. Since j_0 is fixed, it follows from (S0.32) that

$$||\widehat{\mathbf{W}} - \sum_{j=1}^{j_0} \mathbf{\Sigma}_v(j)\mathbf{\Sigma}_v'(j)||_2 = O_p(n^{-1/2}). \tag{S0.33}$$

Note that $D(\mathcal{M}(\widehat{\mathbf{B}}), \mathcal{M}(\mathbf{B})) = O_p(||\widehat{\mathbf{W}} - \sum_{j=1}^{j_0} \mathbf{\Sigma}_v(j)\mathbf{\Sigma}_v'(j)||_2)$ (see for example, Chang, Guo and Yao (2015)), we have (b) of Theorem 1 as desired.

Now, we turn to show (c). By (S0.19), we get

$$\begin{aligned}
& \sum_{t=s+1}^{n} [\widehat{\mathbf{f}}_{t-1}', \cdots, \widehat{\mathbf{f}}_{t-s}']'[\widehat{\mathbf{f}}_t - \sum_{i=1}^{s} \mathbf{E}_i \widehat{\mathbf{f}}_{t-i}]' \\
= & \sum_{t=s+1}^{n} \left[\mathbf{f}_{t-1}' + \boldsymbol{\varepsilon}_{t-1}'\mathbf{B}, \cdots, \mathbf{f}_{t-s}' + \boldsymbol{\varepsilon}_{t-s}'\mathbf{B}\right]' [\mathbf{e}_t' + \boldsymbol{\varepsilon}_t'\mathbf{B}] \\
& - \sum_{t=s+1}^{n} \left[\mathbf{f}_{t-1}' + \boldsymbol{\varepsilon}_{t-1}'\mathbf{B}, \cdots, \mathbf{f}_{t-s}' + \boldsymbol{\varepsilon}_{t-s}'\mathbf{B}\right]' [\sum_{i=1}^{s} \boldsymbol{\varepsilon}_{t-i}'\mathbf{B}\mathbf{E}_i'] \\
& + \sum_{t=s+1}^{n} \left[\mathbf{f}_{t-1}' + \boldsymbol{\varepsilon}_{t-1}'\mathbf{B}, \cdots, \mathbf{f}_{t-s}' + \boldsymbol{\varepsilon}_{t-s}'\mathbf{B}\right]' [\sum_{j=2}^{4} (\boldsymbol{\zeta}_{t,j} - \sum_{i=1} \mathbf{E}_i \boldsymbol{\zeta}_{t-i,j})]' \\
& + \sum_{t=s+1}^{n} \sum_{j=2}^{4} [\boldsymbol{\zeta}_{t-1,j}', \cdots, \boldsymbol{\zeta}_{t-s,j}']'[\mathbf{e}_t + \mathbf{B}'\boldsymbol{\varepsilon}_t - \sum_{i=1}^{s} \mathbf{E}_i \mathbf{B}'\boldsymbol{\varepsilon}_{t-i} + \sum_{j=2}^{4} (\boldsymbol{\zeta}_{t,j} - \sum_{i=1} \mathbf{E}_i \boldsymbol{\zeta}_{t-i,j})]' \\
=: & \sum_{i=1}^{4} \Delta_{ni}.
\end{aligned} \tag{S0.34}$$

By (S0.7), (S0.20) and (S0.29), we can show that for any given positive integer s,

$$||\Delta_{n3}||_2 + ||\Delta_{n4}||_2 = O_p(\sqrt{n}). \tag{S0.35}$$

On the other hand, since for any $1 \le i, j \le s$ and $l \ne i$, $\text{vech}\{(\mathbf{f}_{t-i} + \mathbf{B}\boldsymbol{\varepsilon}_{t-i})(\mathbf{e}_t' + \boldsymbol{\varepsilon}_t'\mathbf{B})$, $\mathbf{f}_{t-i}\boldsymbol{\varepsilon}_{t-j}'\mathbf{B}$, $\mathbf{B}'\mathbf{v}_{t-i}\boldsymbol{\varepsilon}_{t-l}'\mathbf{B}\}$ is a α mixing process with finite 2γ-moment and mixing coefficients satisfying (S0.2), it follows from the CLT of α mixing process (see for example Corollary 3.2.1 of Lin and Lu) that for some matrix Γ_1,

$$\frac{1}{\sqrt{n}}\sum_{t=s+1}^{n}\text{vech}\{(\mathbf{f}_{t-i}+\mathbf{B}\boldsymbol{\varepsilon}_{t-i})(\mathbf{e}_t'+\boldsymbol{\varepsilon}_t'\mathbf{B}),\ \mathbf{f}_{t-i}\boldsymbol{\varepsilon}_{t-j}'\mathbf{B},\ \mathbf{B}'\mathbf{v}_{t-i}\boldsymbol{\varepsilon}_{t-l}'\mathbf{B}\}\xrightarrow{d}N(0,\Gamma_1) \tag{S0.36}$$

Set $\Omega=\Big[\sum_{t=s+1}^{n}(\widehat{\mathbf{f}}_{t-1}',\cdots,\widehat{\mathbf{f}}_{t-s}')'(\widehat{\mathbf{f}}_{t-1}',\cdots,\widehat{\mathbf{f}}_{t-s}')-M\Big]$. By the definition of $\widehat{\mathbf{E}}_i$, $i=1,2,\cdots,s$, we have

$$\begin{pmatrix}\widehat{\mathbf{E}}_1'-\mathbf{E}_1'\\ \vdots\\ \widehat{\mathbf{E}}_s'-\mathbf{E}_s'\end{pmatrix}=\Omega^{-1}\left[\begin{pmatrix}\sum_{t=s}^{n}\widehat{\mathbf{f}}_{t-1}(\widehat{\mathbf{f}}_t-\sum_{i=1}^{s}\mathbf{E}_i\widehat{\mathbf{f}}_{t-i})'\\ \vdots\\ \sum_{t=s}^{n}\widehat{\mathbf{f}}_{t-s}(\widehat{\mathbf{f}}_t-\sum_{i=1}^{s}\mathbf{E}_i\widehat{\mathbf{f}}_{t-i})'\end{pmatrix}+M\begin{pmatrix}\mathbf{E}_1'\\ \vdots\\ \mathbf{E}_s'\end{pmatrix}\right] \tag{S0.37}$$

Thus, by Lemma 7 and (S0.34)–(S0.36), we have conclusion (c) and complete the proof of Theorem 1. □

Next, we first develop bounds for the estimated eigenvalues $\widehat{\lambda}_j$, $j=1,2,\cdots p$.

Lemma 8. *Let λ_j, $j=1,\cdots,p$ be the eigenvalues of $\mathbf{W}_v$. Under Condition 1 or conditions of Theorem 3,*

$$|\widehat{\lambda}_m-\lambda_m|=O_p(pn^{-1/2})\quad\textit{and}\quad|\widehat{\lambda}_{m+1}|=O_p(pn^{-1/2}). \tag{S0.38}$$

Proof. By (b) of Theorem 1 and (b) of Theorem 3, we have for any $1\le i\le p$,

$$|\widehat{\lambda}_i-\lambda_i|\le||\widehat{\mathbf{W}}_v-\mathbf{W}_v||_2=O_p(pn^{-1/2})\quad\text{and}\quad\lambda_{m+1}=\cdots=\lambda_p=0.$$

This gives Lemma 8 as desired. □

Proof of Theorem 2. It is enough to show that

$$\lim_{n\to\infty}P\{\tilde{m}<m\}=0. \tag{S0.39}$$

Suppose $\tilde{m}<m$ is true, then by Lemma 8, there exists a positive constant

c_1 such that

$$\lim_{n\to\infty} P\{\widehat{\lambda}_{\widetilde{m}+1}/\widehat{\lambda}_{\widetilde{m}} \geq c_1\} = 1, \quad \text{and} \quad \lim_{n\to\infty} P\{\widehat{\lambda}_{m+1}/\widehat{\lambda}_m < c_1/2\} = 1.$$

This implies that

$$\lim_{n\to\infty} P\{\widehat{\lambda}_{\widetilde{m}+1}/\widehat{\lambda}_{\widetilde{m}} > \widehat{\lambda}_{m+1}/\widehat{\lambda}_m\} = 1,$$

which contradicts the definition of $\widetilde{m}$. Thus, (S0.39) holds. □

Proof of Theorem 3. Since $p = o(n^{1/2})$ and $\{\mathbf{x}_{t2}\}$ is a α mixing process with mixing coefficients satisfying (S0.2), it follows that (S0.27) also holds for this case. Further, note that for any $1 \leq i \leq p$ and $1 \leq j \leq r$, applying CLT of mixing process to $\{(\mathbf{b}_i\mathbf{f}_t + \varepsilon_t^i)x_{t-1,2}^j\}$, which is a α mixing process with coefficients satisfying (3.2), we get

$$|\sum_{t=1}^{n}(\mathbf{b}_i\mathbf{f}_t + \varepsilon_t^i)x_{t-1,2}^j| = O_p(\sqrt{n}),$$

which implies

$$||\frac{1}{n}\sum_{t=1}^{n}(\mathbf{B}\mathbf{f}_t + \boldsymbol{\varepsilon}_t)\mathbf{x}_{t-1,2}'||_2 = O_p(n^{-1/2}(pr)^{1/2}). \tag{S0.40}$$

Thus, by Lemmas 5 and 6,

$$\begin{aligned}
||\widehat{\mathbf{D}} - \mathbf{D}||_2 &= \left\| \left(\frac{1}{n}\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{y}_{t-1}'\widehat{\mathbf{A}}_2\right)\left(\frac{1}{n}\sum_{i=1}^{n}\widehat{\mathbf{A}}_2'\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\widehat{\mathbf{A}}_2\right)^{-1} - \mathbf{D}\right\|_2 \\
&= \left\| \left(\frac{1}{n}\sum_{t=1}^{n}\nabla\mathbf{y}_t\mathbf{x}_{t-1,2}'\right)\left(\frac{1}{n}\sum_{i=0}^{n-1}\mathbf{x}_{t-1,2}\mathbf{x}_{t-1,2}'\right)^{-1} - \mathbf{D}\right\|_2 + O_p(p^{1/2}k^2/n) \\
&= \left\| \left(\frac{1}{n}\sum_{t=1}^{n}(\mathbf{B}\mathbf{f}_t + \boldsymbol{\varepsilon}_t)\mathbf{x}_{t-1,2}'\right)\left(\frac{1}{n}\sum_{i=0}^{n-1}\mathbf{x}_{t-1,2}\mathbf{x}_{t-1,2}'\right)^{-1}\right\|_2 + O_p(p^{1/2}k^2/n) \\
&= O_p(n^{-1/2}(pr)^{1/2} + p^{1/2}k^2/n), \qquad \text{(S0.41)}
\end{aligned}$$

this combining with (S0.12) yields

$$||\widehat{\mathbf{C}}-\mathbf{C}||_2 = ||(\widehat{\mathbf{D}}-\mathbf{D})\mathbf{A}_2' + \widehat{\mathbf{D}}'(\widehat{\mathbf{A}}_2'-\mathbf{A}_2')||_2 = O_p(n^{-1/2}(pr)^{1/2} + p^{1/2}k^2/n). \quad \text{(S0.42)}$$

Thus, (a) of Theorem 3 follows from (S0.41) and (S0.42).

Next, we show (b). It is easy to see that

$$\|\frac{1}{n^2}\sum_{t=1}^{n-j}\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\|_2 = O_p(p). \quad \text{(S0.43)}$$

Thus, by (S0.12), (S0.41) and (iii) of Condition 3, it can be shown that $||\cdot||_2$ norm of the last six terms of the right-hand side in (S0.30) are of order $o(pn^{-1/2})$, provided $k = o(n^{1/2})$ and $p = O(n^{1/4})$. On the other hand, applying CLT of α mixing process to the first term of the right-hand side of (S0.30), we get for any given j, this term is of order $O_p(pn^{-1/2})$. Similarly, we can show $n^{-1}\sum_{t=1}^{n-j}\bar{\mathbf{v}}\bar{\mathbf{v}}_t' = O_p(n^{-1/2}p)$. Thus,

$$||\widehat{\boldsymbol{\Sigma}}_v(j)-\boldsymbol{\Sigma}_v(j)||_2 = O_p(n^{-1/2}p). \quad \text{(S0.44)}$$

Since j_0 is fixed, it follows from (S0.44) that

$$||\widehat{\mathbf{W}} - \sum_{j=1}^{j_0}\boldsymbol{\Sigma}_v(j)\boldsymbol{\Sigma}_v'(j)||_2 = O_p(n^{-1/2}p). \quad \text{(S0.45)}$$

Note that $D(\mathcal{M}(\widehat{\mathbf{B}}),\mathcal{M}(\mathbf{B})) = O_p(||\widehat{\mathbf{W}} - \sum_{j=1}^{j_0}\boldsymbol{\Sigma}_v(j)\boldsymbol{\Sigma}_v'(j)||_2)$ (see for example, Chang, Guo and Yao (2015)), we have (b) of Theorem 3 as desired.

In the following, we give the proof of (c). Let Δ_{ni}, $i = 1,2,3,4$ be defined as in (S0.34). By conclusions (a), (b) of Theorem 3 and (S0.12), we can show that

$$||\Delta_{n3}+\Delta_{n4}||_2 = O_p\left(n^{1/2}(pr)^{1/2}[n^{-1/2}(pr)^{1/2} + p^{1/2}k^2/n + pn^{-1/2}] + p^{1/2}k^2\right). \quad \text{(S0.46)}$$

On the other hand, applying CLT of α mixing to the elements of vech$\{(\mathbf{f}_{t-i}+\mathbf{B}\boldsymbol{\varepsilon}_{t-i})(\mathbf{e}_t'+\boldsymbol{\varepsilon}_t'\mathbf{B})$, $\mathbf{f}_{t-i}\boldsymbol{\varepsilon}_{t-j}'\mathbf{B}$, $\mathbf{B}'\mathbf{v}_{t-i}\boldsymbol{\varepsilon}_{t-l}'\mathbf{B}$, $l\neq i$, $1\leq i,j\leq s\}$, we get

$$||\Delta_{n1}+\Delta_{n2}-M||_2 = O_p((pmn)^{1/2}). \quad \text{(S0.47)}$$

Combining equations (S0.46)–(S0.47) with Lemma 7 and $p = o(n^{1/2})$ yield

$$||(\mathbf{E}_1, \cdots, \mathbf{E}_s)||_2 = O(p^{1/2}k^2n^{-1} + pm^{1/2}n^{-1/2}), \quad \text{(S0.48)}$$

this gives (c) and completes the proof of Theorem 3. □

Proof of Theorem 4. By Lemma 8, Theorem 4 can be shown similarly as for Theorem 2. Therefore, we omit the detailed proofs. □

Proof of Remark 1. Since the proofs are similar, we only show the case with fixed p in details. It follows from the definition of $\widehat{m}$ that

$$\sum_{j=\widehat{m}+1}^{p} \widehat{\lambda}_j + \widehat{m}\omega_n \leq \sum_{j=m}^{p} \widehat{\lambda}_{p+1-j} + m\omega_n. \quad \text{(S0.49)}$$

Suppose that $\widehat{m} > m$, it follows from (S0.49) that

$$(\widehat{m} - m)\omega_n \leq \sum_{j=m+1}^{\widehat{m}} \widehat{\lambda}_j \leq (\widehat{m} - m)\widehat{\lambda}_{m+1}. \quad \text{(S0.50)}$$

Since $\omega_n/n^{-1/2} \to \infty$, it follows from Lemma 8 that equation (S0.50) holds with probability zero. This gives that

$$\lim_{n\to\infty} P\{\widehat{m} > m\} = 0. \quad \text{(S0.51)}$$

On the other hand, if $\widehat{m} < m$, equation (S0.49) yields

$$(m - \widehat{m})\widehat{\lambda}_m \leq \sum_{j=\widehat{m}+1}^{m} \widehat{\lambda}_j \leq (m - \widehat{m})\omega_n. \quad \text{(S0.52)}$$

Lemma 8 implies $\widehat{\lambda}_m \geq \lambda_m/2 > 0$. Thus, by (S0.52) and $\omega_n \to 0$ as $n \to \infty$, we have

$$\lim_{n\to\infty} P\{\widehat{m} < m\} = 0. \quad \text{(S0.53)}$$

Equation (S0.51) together with (S0.53) give the consistency of $\widehat{m}$ as desired.

□

References

Lin, Z. and Lu, C. (1997). *Limit Theory on Mixing Dependent Random Variables*. Kluwer Academic Publishers, New York.

Zhang, R. M., Robinosn, P. and Yao, Q. (2015). Identifying Cointegration by Eigenanalysis. *A Manuscript.*