# Automated Estimation of Heavy-tailed Vector Error Correction Models

Feifei Guo, Shiqing Ling and Zichuan Mi

*Beijing Institute of Technology*

*Hong Kong University of Science and Technology*

*Shanxi University of Finance and Economics*

**Supplementary Material**

## S1 Preliminary result

Denote $\{\mathbf{J}_i\}$ as the coefficient series in the representation $\Delta\mathbf{Y}_{t-j} = \sum_{i=0}^{\infty}\mathbf{J}_i\boldsymbol{\varepsilon}_{t-j-i}$, see Johansen (1995). We give the following lemma and its proof.

**Lemma S1.** *Suppose that the conditions of Theorem 4 are satisfied. Then*

$$(a).\ a_n^{-2}[F'_{\pi,n}(1),\cdots,F'_{\pi,n}(m)]' \to_d \mathbf{Q}^*\mathbf{T}_{1,\pi_o}\mathbf{R}_2\mathbf{T}_{2,\pi_o},$$

*where* $\mathbf{Q}^*$ *is the normalized left eigenvector matrix of eigenvalues of* $\mathbf{\Pi}_o$, $\mathbf{T}_{1,\pi_o} = \mathbf{I}_m - \boldsymbol{\alpha}_o(\boldsymbol{\alpha}_o'\boldsymbol{\alpha}_o)^{-1}\boldsymbol{\alpha}_o'$, $\mathbf{T}_{2,\pi_o} = (\boldsymbol{\beta}_{o,\perp}'\boldsymbol{\alpha}_{o,\perp})^{-1}\boldsymbol{\beta}_{o,\perp}'$ *and* $\mathbf{R}_2$ *is defined* *in Lemma 2. Furthermore,*

$$(b). \quad \tilde{a}_n^{-1}[F_{b,n}(1)', \cdots, F_{b,n}(p)']' \to_d \mathbf{R}_{1\mathcal{S}}^* - \mathbf{R}_1\boldsymbol{\Gamma}_{11}^{-1}\boldsymbol{\Gamma}_{11\mathcal{S}}^*$$

$$\text{when } \alpha \in (1,2) \text{ or } \alpha = 1 \text{ and } \tilde{L}(n) \to 0,$$

$$(c). \quad na_n^{-2}[F_{b,n}(1)', \cdots, F_{b,n}(p)']' \to_d \mathbf{R}_2\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\Gamma}_{21\mathcal{S}}^* - \mathbf{R}_2\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\Gamma}_{21}\boldsymbol{\Gamma}_{11}^{-1}\boldsymbol{\Gamma}_{11\mathcal{S}}^*$$

$$\text{when } \alpha \in (0,1) \text{ or } \alpha = 1 \text{ and } \tilde{L}(n) \to \infty,$$

*where* $\boldsymbol{\Gamma}_{11\mathcal{S}}^* = \sum_{l=0}^{\infty} \mathbf{A}_l\mathbf{S}_1\mathbf{J}_l'$, *and* $\mathbf{R}_{1\mathcal{S}}^*$ *and* $\boldsymbol{\Gamma}_{21\mathcal{S}}^*$ *are defined as* $\mathbf{R}_1^*$ *and* $\boldsymbol{\Gamma}_{21}^*$ *in Lemma 3 with* $\mathbf{B}_i$ *replaced by* $\mathbf{J}_i$*, respectively.*

**Proof .** We rewrite $a_n^{-2}F_{\pi,n}(k)$ as

$$\begin{aligned} a_n^{-2}F_{\pi,n}(k) &= \frac{\mathbf{P}_n(k)'}{a_n^2}\sum_{t=1}^{n}(\Delta\mathbf{Y}_t - \widehat{\boldsymbol{\Pi}}_n\mathbf{Y}_{t-1} - \sum_{j=1}^{p}\widehat{\mathbf{B}}_{n,j}\Delta\mathbf{Y}_{t-j})\mathbf{Y}_{t-1}' \\ &= \frac{\mathbf{Q}_n(k)}{a_n^2}[\sum_{t=1}^{n}\boldsymbol{\varepsilon}_t\mathbf{Y}_{t-1}' - (\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o})\mathbf{Q}_{\mathcal{S}}^{-1}\sum_{t=1}^{n}\mathbf{Z}_{\mathcal{S},t-1}\mathbf{Y}_{t-1}'] + o_P(1), \end{aligned}$$

where the second equality is by $\mathbf{P}_n = \mathbf{Q}_n^{-1}$ and $\mathbf{Q}_n$ is normalized. By Lemma 2,

$$\begin{aligned} a_n^{-2}\sum_{t=1}^{n}\boldsymbol{\varepsilon}_t\mathbf{Z}_{t-1}' &= \left[a_n^{-2}\sum_{t=1}^{n}\boldsymbol{\varepsilon}_t\mathbf{Y}_{t-1}'\boldsymbol{\beta}_o,\ a_n^{-2}\mathbf{R}_{2n}\right] \\ &\to_d \left[\mathbf{0}_{m\times r_o}, [\int_0^1 \mathbf{P}(r)d\mathbf{P}'(r)]'\boldsymbol{\psi}'[\mathbf{I}_d, \mathbf{0}]'\right], \end{aligned} \tag{S1.1}$$

and

$$\frac{\mathbf{D}_{n,\mathcal{S}}\sum_{t=1}^{n}\mathbf{Z}_{\mathcal{S},t-1}\mathbf{Z}'_{t-1}}{a_n^2} \rightarrow_d \begin{pmatrix} \mathbf{0}_{(mp+r_o)\times r_o} & \mathbf{0} \\ \mathbf{0} & [\mathbf{I}_d,\mathbf{0}]\boldsymbol{\psi}[\int_0^1 \mathbf{P}(r)\mathbf{P}'(r)dr]\boldsymbol{\psi}'[\mathbf{I}_d,\mathbf{0}]' \end{pmatrix}.$$

Using the expression $(\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o})\mathbf{Q}_{\mathcal{S}}^{-1}\mathbf{D}_{n,\mathcal{S}}^{-1}$ in the proof of Theorem 4, we obtain

$$\begin{aligned} & a_n^{-2}(\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o})\mathbf{Q}_{\mathcal{S}}^{-1}\sum_{t=1}^{n}\mathbf{Z}_{\mathcal{S},t-1}\mathbf{Z}'_{t-1} \\ &= a_n^{-2}(\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o})\mathbf{Q}_{\mathcal{S}}^{-1}\mathbf{D}_{n,\mathcal{S}}^{-1}\mathbf{D}_{n,\mathcal{S}}\sum_{t=1}^{n}\mathbf{Z}_{\mathcal{S},t-1}\mathbf{Z}'_{t-1} \\ &\rightarrow_d \left[\mathbf{0}, \boldsymbol{\alpha}_o(\boldsymbol{\alpha}'_o\boldsymbol{\alpha}_o)^{-1}\boldsymbol{\alpha}'_o[\int_0^1 \mathbf{P}(r)d\mathbf{P}'(r)]'\boldsymbol{\psi}'[\mathbf{I}_d,\mathbf{0}]'\right]. \end{aligned} \tag{S1.2}$$

The results in (S1.1) and (S1.2) imply that

$$\begin{aligned} a_n^{-2}F_{\pi,n}(k) =& \frac{\mathbf{Q}_n(k)}{a_n^2}\left[\sum_{t=1}^{n}\boldsymbol{\varepsilon}_t\mathbf{Z}'_{t-1} - (\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o})\mathbf{Q}_{\mathcal{S}}^{-1}\sum_{t=1}^{n}\mathbf{Z}_{\mathcal{S},t-1}\mathbf{Z}'_{t-1}\right]\mathbf{Q}'^{-1} \\ &\rightarrow_d \mathbf{Q}^*(k)[\mathbf{I}_m - \boldsymbol{\alpha}_o(\boldsymbol{\alpha}'_o\boldsymbol{\alpha}_o)^{-1}\boldsymbol{\alpha}'_o]\mathbf{R}_2(\boldsymbol{\beta}'_{o,\perp}\boldsymbol{\alpha}_{o,\perp})^{-1}\boldsymbol{\beta}'_{o,\perp}, \end{aligned}$$

where $\mathbf{Q}^*(k)$ is the $k$-th row vector of $\mathbf{Q}^*$ and $\mathbf{Q}^*$ is the normalized left eigenvector matrix of eigenvalues of $\boldsymbol{\Pi}_o$ and $\mathbf{R}_2$ is defined in Lemma 2. It is easy to show the joint weak convergence of $(a_n^{-2}F_{\pi,n}(k), k = 1, \cdots, m)$. Thus, (a) holds. We next show the claim of (b) and (c). When $\alpha \in (1, 2)$, or $\alpha = 1$ and $\tilde{L}(n) \to 0$, we rewrite $\tilde{a}_n^{-1} F_{b,n}(j)$ as

$$\tilde{a}_n^{-1} F_{b,n}(j) = \tilde{a}_n^{-1}[\sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \Delta \mathbf{Y}'_{t-j} - (\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o}) \mathbf{Q}_{\mathcal{S}}^{-1} \sum_{t=1}^{n} \mathbf{Z}_{\mathcal{S},t-1} \Delta \mathbf{Y}'_{t-j}].$$

Using the arguments in the proof of Theorem 4,

$$\begin{aligned}
&\frac{1}{\tilde{a}_n}[\sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \Delta \mathbf{Y}'_{t-j} - (\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o}) \mathbf{Q}_{\mathcal{S}}^{-1} \sum_{t=1}^{n} \mathbf{Z}_{\mathcal{S},t-1} \Delta \mathbf{Y}'_{t-j}] \\
&= \frac{1}{\tilde{a}_n} \sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \Delta \mathbf{Y}'_{t-j} - \frac{1}{\tilde{a}_n} \sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \mathbf{Z}'_{1\mathcal{S},t-1} (\sum_{t=1}^{n} \mathbf{Z}_{1\mathcal{S},t-1} \mathbf{Z}'_{1\mathcal{S},t-1})^{-1} \sum_{t=1}^{n} \mathbf{Z}_{1\mathcal{S},t-1} \Delta \mathbf{Y}'_{t-j} + o_p(1) \\
&\to_d \mathbf{R}^*_{1\mathcal{S}} - \mathbf{R}_1 \boldsymbol{\Gamma}_{11}^{-1} \boldsymbol{\Gamma}^*_{11\mathcal{S}}, \qquad \text{(S1.3)}
\end{aligned}$$

where $\mathbf{R}^*_{1\mathcal{S}} = \sum_{i=0}^{\infty} \mathbf{S}_{i+j+1} \mathbf{J}'_i$ and $\boldsymbol{\Gamma}^*_{11\mathcal{S}} = \sum_{l=0}^{\infty} \mathbf{A}_l \mathbf{S}_1 \mathbf{J}'_{l+1-j}$.

When $\alpha \in (0,1)$, or $\alpha = 1$ and $\tilde{L}(n) \to \infty$, we rewrite $n a_n^{-2} F_{b,n}(j)$ as

$$n a_n^{-2} F_{b,n}(j) = n a_n^{-2}[\sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \Delta \mathbf{Y}'_{t-j} - (\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o}) \mathbf{Q}_{\mathcal{S}}^{-1} \sum_{t=1}^{n} \mathbf{Z}_{\mathcal{S},t-1} \Delta \mathbf{Y}'_{t-j}].$$

Then by Lemma 2 and Theorem A.1 in She and Ling (2020),

$$\begin{aligned}
&n a_n^{-2}[\sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \Delta \mathbf{Y}'_{t-j} - (\widehat{\boldsymbol{\theta}}_{\mathcal{S},n} - \boldsymbol{\theta}_{\mathcal{S},o}) \mathbf{Q}_{\mathcal{S}}^{-1} \sum_{t=1}^{n} \mathbf{Z}_{\mathcal{S},t-1} \Delta \mathbf{Y}'_{t-j}] \\
&= n a_n^{-2} \sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \mathbf{Z}'_{2,t-1} (\sum_{t=1}^{n} \mathbf{Z}_{2,t-1} \mathbf{Z}'_{2,t-1})^{-1} \sum_{t=1}^{n} \mathbf{Z}_{2,t-1} \Delta \mathbf{Y}'_{t-j} \\
&\quad - n a_n^{-2} \sum_{t=1}^{n} \boldsymbol{\varepsilon}_t \mathbf{Z}'_{2,t-1} (\sum_{t=1}^{n} \mathbf{Z}_{2,t-1} \mathbf{Z}'_{2,t-1})^{-1} \sum_{t=1}^{n} \mathbf{Z}_{2,t-1} \mathbf{Z}'_{1\mathcal{S},t-1} \\
&\qquad \times (\sum_{t=1}^{n} \mathbf{Z}_{1\mathcal{S},t-1} \mathbf{Z}'_{1\mathcal{S},t-1})^{-1} \sum_{t=1}^{n} \mathbf{Z}_{1\mathcal{S},t-1} \Delta \mathbf{Y}'_{t-j} + o_p(1) \\
&\to_d \mathbf{R}_2 \boldsymbol{\Gamma}_{22}^{-1} \boldsymbol{\Gamma}^*_{21\mathcal{S}} - \mathbf{R}_2 \boldsymbol{\Gamma}_{22}^{-1} \boldsymbol{\Gamma}_{21} \boldsymbol{\Gamma}_{11}^{-1} \boldsymbol{\Gamma}^*_{11\mathcal{S}}, \qquad \text{(S1.4)}
\end{aligned}$$

where $\boldsymbol{\Gamma}^*_{21\mathcal{S}} = \mathbf{R}_2' \sum_{i=0}^{\infty} \mathbf{J}_i' + [\mathbf{I}_d, \mathbf{0}] \sum\limits_{i=0}^{\infty} \sum\limits_{j=0}^{i} \boldsymbol{\psi}_j \mathbf{S}_1 \mathbf{J}_i'$. From the results in (S1.3) and (S1.4), it is easy to show the joint weak convergence of $(F_{b,n}(j), j = 1, \cdots, p)$. Thus, (b) and (c) holds. This completes the proof. $\square$

## S2 Other Simulation Results

Table S2.1 and S2.2 report simulation results of estimated parameters in model (5.1) when $r_o = 0$ and $r_o = 2$, i.e., $\boldsymbol{\Pi}_o$ is specified as follows:

$$\boldsymbol{\Pi}_o = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} -0.5 & 0.1 \\ 0.2 & -0.4 \end{bmatrix}.$$

Table S2.3 and S2.4 report simulation results of estimated parameters in model (5.2) when $r_o = 1$ and $r_o = 2$, i.e., $\boldsymbol{\Pi}_o$ is specified as follows:

$$\boldsymbol{\Pi}_o = \begin{bmatrix} -0.5 & -0.25 & 0.5 \\ 0.1 & 0.05 & -0.1 \\ 0.2 & 0.1 & -0.2 \end{bmatrix} \; and \; \begin{bmatrix} -0.5 & -0.2 & 0.7 \\ 0.1 & -0.3 & 0.2 \\ 0.2 & 0.2 & -0.4 \end{bmatrix}.$$

## S3 Proofs of Lemma1 and Lemma 3

**Proof of Lemma 1.** For any $k \in \mathcal{S}_\phi$, by Lemma 3(b) we know that

$$\|\phi_k(\widehat{\boldsymbol{\Pi}}_{1st})\|^\omega \to_p \|\phi_k(\boldsymbol{\Pi}_o)\|^\omega > 0,$$

which implies

$$n^{\tau_1} \delta_{r,n} = \frac{n^{\tau_1} \lambda^*_{r,k,n}}{\|\phi_k(\widehat{\boldsymbol{\Pi}}_{1st})\|^\omega} \to_p 0.$$

Table S2.1: Finite sample properties of the shrinkage estimates of model (5.1)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $r_o = 0$ and n=400 | | | | | |
| | | $\Pi_{11}$ | | $\Pi_{21}$ | | $\Pi_{12}$ | | $\Pi_{22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.00057 | 0.01008 | 0.00070 | 0.01392 | 0.00147 | 0.01787 | 0.00012 | 0.00927 |
| | OLS | 0.00925 | 0.06551 | 0.01801 | 0.23447 | 0.00168 | 0.05989 | 0.01297 | 0.06494 |
| | Oracle | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| $\alpha = 1.3$ | Lasso | 0.00094 | 0.00528 | 0.00036 | 0.00569 | 0.00022 | 0.00822 | 0.00096 | 0.00530 |
| | OLS | 0.00171 | 0.00957 | 0.00041 | 0.01723 | 0.00002 | 0.00927 | 0.00222 | 0.0080 |
| | Oracle | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| | | $B_{1,11}$ | | $B_{1,21}$ | | $B_{1,12}$ | | $B_{1,22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.07673 | 0.35755 | 0.00898 | 0.43545 | 0.00092 | 0.33002 | 0.08249 | 0.35720 |
| | OLS | 0.02417 | 0.24134 | 0..01307 | 0.35951 | 0.01968 | 0.37631 | 0.02887 | 0.23954 |
| | Oracle | 0.00972 | 0.24020 | 0.00462 | 0. 83921 | 0.01156 | 0.33349 | 0.00224 | 0.23641 |
| $\alpha = 1.3$ | Lasso | 0.00454 | 0.11952 | 0.00523 | 0.12750 | 0.00847 | 0.19159 | 0.00673 | 0.12012 |
| | OLS | 0.00922 | 0.06087 | 0.00423 | 0.10971 | 0.00238 | 0.11584 | 0.00676 | 0.05712 |
| | Oracle | 0.00351 | 0.05071 | 0.00154 | 0.09728 | 0.00480 | 0.08263 | 0.00559 | 0.04968 |
| | | $B_{2,11}$ | | $B_{2,21}$ | | $B_{2,12}$ | | $B_{2,22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.00015 | 0.00179 | 0.00004 | 0.00476 | 0.00001 | 0.00067 | 0.00012 | 0.00179 |
| | OLS | 0.00279 | 0.16952 | 0.00091 | 0.23405 | 0.00715 | 0.29294 | 0.00190 | 0.17617 |
| | Oracle | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| $\alpha = 1.3$ | Lasso | 0.00037 | 0.10930 | 0.00399 | 0.11840 | 0.00671 | 0.15588 | 0.00027 | 0.10523 |
| | OLS | 0.00271 | 0.10241 | 0.00246 | 0.14440 | 0.00984 | 0.17708 | 0.00289 | 0.10518 |
| | Oracle | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| | | $B_{3,11}$ | | $B_{3,21}$ | | $B_{3,12}$ | | $B_{3,22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.04564 | 0.25509 | 0.01345 | 0.33630 | 0.01318 | 0.25153 | 0.05585 | 0.25624 |
| | OLS | 0.03252 | 0.17117 | 0.02544 | 0.20248 | 0.00431 | 0.26082 | 0.01603 | 0.16745 |
| | Oracle | 0.04540 | 0.57318 | 0.00568 | 0.65972 | 0.00266 | 0.46912 | 0.03398 | 0.57466 |
| $\alpha = 1.3$ | Lasso | 0.00024 | 0.07089 | 0.00061 | 0.08057 | 0.00147 | 0.12696 | 0.00842 | 0.07755 |
| | OLS | 0.01098 | 0.08961 | 0.00173 | 0.16535 | 0.00017 | 0.12253 | 0.00778 | 0.08903 |
| | Oracle | 0.00664 | 0.05216 | 0.00328 | 0.08345 | 0.00016 | 0.07124 | 0.00238 | 0.05370 |

Note: The oracle estimate is simply the OLS estimate with the restriction that $\mathbf{\Pi}_o = \mathbf{0}$ and $\mathbf{B}_{o,2} = \mathbf{0}$.

Table S2.2: Finite sample properties of the shrinkage estimates of model (5.1)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $r_o = 2$ and n=400 | | | | | |
| | | $\Pi_{11}$ | | $\Pi_{21}$ | | $\Pi_{12}$ | | $\Pi_{22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.03127 | 0.14316 | 0.02279 | 0.11743 | 0.01567 | 0.07149 | 0.01586 | 0.07139 |
| | OLS | 0.00512 | 0.05476 | 0.01647 | 0.10598 | 0.00127 | 0.04812 | 0.00748 | 0.05061 |
| | Oracle | 0.00056 | 0.05796 | 0.00229 | 0.08451 | 0.00179 | 0.05583 | 0.00224 | 0.04834 |
| $\alpha = 1.3$ | Lasso | 0.02081 | 0.05484 | 0.00749 | 0.06985 | 0.00035 | 0.04580 | 0.02442 | 0.05694 |
| | OLS | 0.00063 | 0.05589 | 0.00641 | 0.08858 | 0.00081 | 0.04241 | 0.00430 | 0.04424 |
| | Oracle | 0.00275 | 0.05252 | 0.00408 | 0.07291 | 0.00197 | 0.04270 | 0.00198 | 0.04147 |
| | | $B_{1,11}$ | | $B_{1,21}$ | | $B_{1,12}$ | | $B_{1,22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.01946 | 0.07179 | 0.00247 | 0.04293 | 0.00428 | 0.03732 | 0.09407 | 0.07346 |
| | OLS | 0.00245 | 0.06992 | 0.00416 | 0.12533 | 0.01047 | 0.07991 | 0.00334 | 0.06969 |
| | Oracle | 0.00211 | 0.07049 | 0.00156 | 0.09715 | 0.00264 | 0.07259 | 0.00106 | 0.06706 |
| $\alpha = 1.3$ | Lasso | 0.01760 | 0.06866 | 0.00408 | 0.08581 | 0.00190 | 0.06510 | 0.01227 | 0.07422 |
| | OLS | 0.00014 | 0.06714 | 0.00369 | 0.10401 | 0.00058 | 0.06960 | 0.00231 | 0.06591 |
| | Oracle | 0.00069 | 0.06497 | 0.00233 | 0.07556 | 0.00451 | 0.06008 | 0.00061 | 0.05181 |
| | | $B_{2,11}$ | | $B_{2,21}$ | | $B_{2,12}$ | | $B_{2,22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| | OLS | 0.00950 | 0.07158 | 0.00755 | 0.14329 | 0.00847 | 0.08178 | 0.01025 | 0.07254 |
| | Oracle | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| $\alpha = 1.3$ | Lasso | 0.01183 | 0.04105 | 0.00042 | 0.05063 | 0.00024 | 0.05168 | 0.00983 | 0.03680 |
| | OLS | 0.00037 | 0.05777 | 0.00484 | 0.12853 | 0.00187 | 0.06523 | 0.00128 | 0.05834 |
| | Oracle | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| | | $B_{3,11}$ | | $B_{3,21}$ | | $B_{3,12}$ | | $B_{3,22}$ | |
| | | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\alpha = 0.2$ | Lasso | 0.09178 | 0.08305 | 0.00056 | 0.04556 | 0.00526 | 0.04938 | 0.08057 | 0.08425 |
| | OLS | 0.00848 | 0.08109 | 0.01558 | 0.12233 | 0.00197 | 0.11189 | 0.01183 | 0.07611 |
| | Oracle | 0.00643 | 0.07765 | 0.00438 | 0.11304 | 0.00155 | 0.08754 | 0.00454 | 0.07381 |
| $\alpha = 1.3$ | Lasso | 0.01522 | 0.07394 | 0.00701 | 0.08625 | 0.00438 | 0.07636 | 0.01761 | 0.07545 |
| | OLS | 0.00546 | 0.05863 | 0.00128 | 0.10052 | 0.00064 | 0.06551 | 0.00311 | 0.05709 |
| | Oracle | 0.00184 | 0.05511 | 0.00247 | 0.08196 | 0.00278 | 0.05941 | 0.00606 | 0.05378 |

Note: The oracle estimate is simply the OLS estimate with the restriction that $\mathbf{B}_{o,2} = \mathbf{0}$.

Table S2.3: Finite sample properties of the shrinkage estimates of model (5.2)

| $r_o = 1$ and n=400 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.6$ | | | | | | $\alpha = 1.3$ | | | | | |
| | Lasso | | OLS | | Oracle | | Lasso | | OLS | | Oracle | |
| | Bais | Std | Bais | Std | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\Pi_{11}$ | 0.0697 | 0.1051 | 0.0113 | 0.0703 | 0.0011 | 0.0116 | 0.0919 | 0.0654 | 0.0197 | 0.0536 | 0.0019 | 0.0125 |
| $\Pi_{21}$ | 0.0184 | 0.0765 | 0.0023 | 0.0498 | 0.0193 | 0.0527 | 0.0245 | 0.0460 | 0.0055 | 0.0430 | 0.0050 | 0.0357 |
| $\Pi_{31}$ | 0.0278 | 0.0608 | 0.0122 | 0.0816 | 0.0031 | 0.0775 | 0.0374 | 0.0513 | 0.0009 | 0.0663 | 0.0025 | 0.0547 |
| $\Pi_{12}$ | 0.0873 | 0.1083 | 0.0135 | 0.0740 | 0.0044 | 0.0128 | 0.00695 | 0.0663 | 0.0195 | 0.0438 | 0.0024 | 0.0127 |
| $\Pi_{22}$ | 0.0184 | 0.0784 | 0.0038 | 0.0524 | 0.0193 | 0.0537 | 0.0344 | 0.0465 | 0.0040 | 0.0473 | 0.0011 | 0.0358 |
| $\Pi_{32}$ | 0.0275 | 0.0627 | 0.0090 | 0.0848 | 0.0038 | 0.0763 | 0.0367 | 0.0517 | 0.0008 | 0.0671 | 0.0632 | 0.0548 |
| $\Pi_{13}$ | 0.0703 | 0.1056 | 0.0117 | 0.0730 | 0.0144 | 0.0121 | 0.0603 | 0.0656 | 0.0191 | 0.0439 | 0.0020 | 0.0126 |
| $\Pi_{23}$ | 0.0165 | 0.0739 | 0.0008 | 0.0505 | 0.0194 | 0.0532 | 0.0244 | 0.0460 | 0.0054 | 0.0431 | 0.0046 | 0.0358 |
| $\Pi_{33}$ | 0.0290 | 0.0615 | 0.0078 | 0.0839 | 0.0030 | 0.0766 | 0.0369 | 0.0517 | 0.0023 | 0.0667 | 0.0088 | 0.0546 |
| $B_{1,11}$ | 0.1258 | 0.0717 | 0.0015 | 0.0516 | 0.0277 | 0.0899 | 0.0758 | 0.0618 | 0.0046 | 0.0412 | 0.0069 | 0.0407 |
| $B_{1,21}$ | 0.0056 | 0.0271 | 0.0237 | 0.0454 | 0.0138 | 0.0450 | 0.0045 | 0.0333 | 0.0026 | 0.0432 | 0.0011 | 0.0326 |
| $B_{1,31}$ | 0.0048 | 0.0308 | 0.0785 | 0.0649 | 0.0066 | 0.0547 | 0.0087 | 0.0492 | 0.0042 | 0.0623 | 0.0080 | 0.0474 |
| $B_{1,12}$ | 0.0137 | 0.0611 | 0.0274 | 0.0980 | 0.0254 | 0.0989 | 0.0318 | 0.0740 | 0.0278 | 0.0647 | 0.0048 | 0.0501 |
| $B_{1,22}$ | 0.0985 | 0.0782 | 0.0136 | 0.0807 | 0.0315 | 0.0762 | 0.0689 | 0.0694 | 0.0139 | 0.0648 | 0.0122 | 0.0594 |
| $B_{1,32}$ | 0.0068 | 0.0498 | 0.1019 | 0.1143 | 0.0180 | 0.0800 | 0.0220 | 0.0693 | 0.0023 | 0.0981 | 0.0027 | 0.0656 |
| $B_{1,13}$ | 0.0047 | 0.0456 | 0.0067 | 0.0685 | 0.0432 | 0.0951 | 0.0169 | 0.0610 | 0.0183 | 0.0517 | 0.0019 | 0.0484 |
| $B_{1,23}$ | 0.0022 | 0.0319 | 0.0205 | 0.0568 | 0.0015 | 0.0183 | 0.0078 | 0.0400 | 0.0036 | 0.0502 | 0.0050 | 0.0191 |
| $B_{1,33}$ | 0.1074 | 0.0724 | 0.0846 | 0.0960 | 0.0246 | 0.0721 | 0.0596 | 0.0632 | 0.0049 | 0.0738 | 0.0068 | 0.0522 |
| $B_{2,11}$ | 0.0000 | 0.0000 | 0.0096 | 0.0428 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0093 | 0.0379 | 0.0000 | 0.0000 |
| $B_{2,21}$ | 0.0000 | 0.0000 | 0.0060 | 0.0401 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0043 | 0.0412 | 0.0000 | 0.0000 |
| $B_{2,31}$ | 0.0000 | 0.0000 | 0.0034 | 0.0565 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0010 | 0.0585 | 0.0000 | 0.0000 |
| $B_{2,12}$ | 0.0000 | 0.0000 | 0.0017 | 0.0783 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0134 | 0.0594 | 0.0000 | 0.0000 |
| $B_{2,22}$ | 0.0000 | 0.0000 | 0.0055 | 0.0602 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0070 | 0.0600 | 0.0000 | 0.0000 |
| $B_{2,32}$ | 0.0000 | 0.0000 | 0.0014 | 0.0971 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0896 | 0.0000 | 0.0000 |
| $B_{2,13}$ | 0.0000 | 0.0000 | 0.0080 | 0.0525 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0084 | 0.0421 | 0.0000 | 0.0000 |
| $B_{2,23}$ | 0.0000 | 0.0000 | 0.0034 | 0.0424 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0029 | 0.0437 | 0.0000 | 0.0000 |
| $B_{2,33}$ | 0.0000 | 0.0000 | 0.0083 | 0.0646 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0064 | 0.0625 | 0.0000 | 0.0000 |
| $B_{3,11}$ | 0.1185 | 0.0832 | 0.0092 | 0.0510 | 0.0189 | 0.0506 | 0.0881 | 0.0572 | 0.0013 | 0.0368 | 0.0020 | 0.0369 |
| $B_{3,21}$ | 0.0080 | 0.0315 | 0.0102 | 0.0407 | 0.0156 | 0.0311 | 0.0099 | 0.0337 | 0.0014 | 0.0402 | 0.0021 | 0.0309 |
| $B_{3,31}$ | 0.0012 | 0.0323 | 0.0002 | 0.0637 | 0.0405 | 0.0797 | 0.0055 | 0.0462 | 0.0020 | 0.0596 | 0.0033 | 0.0585 |
| $B_{3,12}$ | 0.0057 | 0.0473 | 0.0059 | 0.0604 | 0.0164 | 0.0571 | 0.0101 | 0.0455 | 0.0073 | 0.0507 | 0.0052 | 0.0493 |
| $B_{3,22}$ | 0.1218 | 0.0762 | 0.0274 | 0.0646 | 0.0326 | 0.1008 | 0.0799 | 0.0503 | 0.0076 | 0.0488 | 0.0039 | 0.0626 |
| $B_{3,32}$ | 0.0010 | 0.0387 | 0.0176 | 0.0988 | 0.0277 | 0.0888 | 0.0045 | 0.0517 | 0.0036 | 0.0775 | 0.0126 | 0.0729 |
| $B_{3,13}$ | 0.0052 | 0.0324 | 0.0058 | 0.0416 | 0.0014 | 0.0483 | 0.0152 | 0.0422 | 0.0127 | 0.0355 | 0.0054 | 0.0433 |
| $B_{3,23}$ | 0.0040 | 0.0230 | 0.0052 | 0.0453 | 0.0167 | 0.0631 | 0.0029 | 0.0285 | 0.0004 | 0.0382 | 0.0103 | 0.0511 |
| $B_{3,33}$ | 0.1094 | 0.0728 | 0.0167 | 0.0649 | 0.0051 | 0.0234 | 0.0831 | 0.0532 | 0.0106 | 0.0510 | 0.0037 | 0.0237 |

Note: The oracle estimate refers to the RLSE with $r_o = 1$ and the restriction that $\mathbf{B}_{o,2} = \mathbf{0}$.

Table S2.4: Finite sample properties of the shrinkage estimates of model (5.2)

| | $r_o = 2$ and n=400 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.6$ | | | | | | $\alpha = 1.3$ | | | | | |
| | Lasso | | OLS | | Oracle | | Lasso | | OLS | | Oracle | |
| | Bais | Std | Bais | Std | Bais | Std | Bais | Std | Bais | Std | Bais | Std |
| $\Pi_{11}$ | 0.1106 | 0.0930 | 0.0204 | 0.0839 | 0.0065 | 0.0114 | 0.1089 | 0.0744 | 0.0130 | 0.0446 | 0.0020 | 0.0319 |
| $\Pi_{21}$ | 0.0211 | 0.0568 | 0.0019 | 0.0402 | 0.0007 | 0.0761 | 0.0157 | 0.0357 | 0.0073 | 0.0479 | 0.0007 | 0.0646 |
| $\Pi_{31}$ | 0.0621 | 0.0462 | 0.0354 | 0.0813 | 0.0076 | 0.0813 | 0.0301 | 0.0384 | 0.0013 | 0.0748 | 0.0004 | 0.0850 |
| $\Pi_{12}$ | 0.0417 | 0.0959 | 0.0066 | 0.0534 | 0.0278 | 0.0875 | 0.0169 | 0.0532 | 0.0049 | 0.0394 | 0.0079 | 0.0604 |
| $\Pi_{22}$ | 0.0942 | 0.0580 | 0.0017 | 0.0452 | 0.0380 | 0.1071 | 0.0680 | 0.0336 | 0.0058 | 0.0410 | 0.0148 | 0.0756 |
| $\Pi_{32}$ | 0.0175 | 0.0533 | 0.0071 | 0.0706 | 0.0213 | 0.0728 | 0.0258 | 0.0311 | 0.0071 | 0.0661 | 0.0249 | 0.0694 |
| $\Pi_{13}$ | 0.0670 | 0.1150 | 0.0137 | 0.1017 | 0.0353 | 0.0895 | 0.0829 | 0.1198 | 0.0168 | 0.0696 | 0.0056 | 0.0655 |
| $\Pi_{23}$ | 0.0133 | 0.1049 | 0.0037 | 0.0723 | 0.0376 | 0.1194 | 0.0051 | 0.0590 | 0.0036 | 0.0763 | 0.0156 | 0.0830 |
| $\Pi_{33}$ | 0.0891 | 0.0843 | 0.0411 | 0.1217 | 0.0239 | 0.1073 | 0.0650 | 0.0617 | 0.0098 | 0.1089 | 0.0246 | 0.1230 |
| $B_{1,11}$ | 0.1362 | 0.0399 | 0.0163 | 0.0527 | 0.0099 | 0.0684 | 0.0915 | 0.0569 | 0.0075 | 0.0450 | 0.0005 | 0.0464 |
| $B_{1,21}$ | 0.0048 | 0.0544 | 0.0043 | 0.0383 | 0.0191 | 0.0646 | 0.0421 | 0.0314 | 0.0185 | 0.0484 | 0.0082 | 0.0653 |
| $B_{1,31}$ | 0.0059 | 0.0564 | 0.0198 | 0.0578 | 0.0241 | 0.0857 | 0.0059 | 0.0287 | 0.0145 | 0.0753 | 0.0091 | 0.0836 |
| $B_{1,12}$ | 0.0127 | 0.0843 | 0.0164 | 0.0542 | 0.0210 | 0.0745 | 0.0038 | 0.0480 | 0.0037 | 0.0521 | 0.0038 | 0.0614 |
| $B_{1,22}$ | 0.0966 | 0.0425 | 0.0072 | 0.0553 | 0.0086 | 0.0620 | 0.0744 | 0.0622 | 0.0093 | 0.0508 | 0.0015 | 0.0482 |
| $B_{1,32}$ | 0.0124 | 0.0734 | 0.0068 | 0.0706 | 0.0155 | 0.0979 | 0.0089 | 0.0395 | 0.0002 | 0.0792 | 0.0032 | 0.0874 |
| $B_{1,13}$ | 0.0071 | 0.0785 | 0.0040 | 0.1099 | 0.0407 | 0.1086 | 0.0072 | 0.0614 | 0.0207 | 0.0729 | 0.0146 | 0.1023 |
| $B_{1,23}$ | 0.0117 | 0.0890 | 0.0099 | 0.0659 | 0.0195 | 0.0893 | 0.0374 | 0.0486 | 0.0374 | 0.0797 | 0.0235 | 0.0875 |
| $B_{1,33}$ | 0.1032 | 0.0410 | 0.0471 | 0.0952 | 0.0353 | 0.1132 | 0.0617 | 0.0590 | 0.0211 | 0.0904 | 0.0072 | 0.1192 |
| $B_{2,11}$ | 0.0000 | 0.0000 | 0.0028 | 0.0398 | 0.0000 | 0.0000 | 0.0001 | 0.0081 | 0.0075 | 0.0428 | 0.0000 | 0.0000 |
| $B_{2,21}$ | 0.0000 | 0.0000 | 0.0171 | 0.0355 | 0.0000 | 0.0000 | 0.0002 | 0.0088 | 0.0280 | 0.0451 | 0.0000 | 0.0000 |
| $B_{2,31}$ | 0.0000 | 0.0000 | 0.0146 | 0.0489 | 0.0000 | 0.0000 | 0.0008 | 0.1267 | 0.0156 | 0.0660 | 0.0000 | 0.0000 |
| $B_{2,12}$ | 0.0000 | 0.0000 | 0.0049 | 0.0577 | 0.0000 | 0.0000 | 0.0007 | 0.0105 | 0.0041 | 0.0512 | 0.0000 | 0.0000 |
| $B_{2,22}$ | 0.0000 | 0.0000 | 0.0249 | 0.0589 | 0.0000 | 0.0000 | 0.0017 | 0.0106 | 0.0023 | 0.0520 | 0.0000 | 0.0000 |
| $B_{2,32}$ | 0.0000 | 0.0000 | 0.0069 | 0.0657 | 0.0000 | 0.0000 | 0.0008 | 0.0162 | 0.0013 | 0.0832 | 0.0000 | 0.0000 |
| $B_{2,13}$ | 0.0000 | 0.0000 | 0.0201 | 0.0841 | 0.0000 | 0.0000 | 0.0014 | 0.0132 | 0.0137 | 0.0606 | 0.0000 | 0.0000 |
| $B_{2,23}$ | 0.0000 | 0.0000 | 0.0293 | 0.0561 | 0.0000 | 0.0000 | 0.0006 | 0.0142 | 0.0459 | 0.0658 | 0.0000 | 0.0000 |
| $B_{2,33}$ | 0.0000 | 0.0000 | 0.0122 | 0.0778 | 0.0000 | 0.0000 | 0.0016 | 0.0200 | 0.0199 | 0.0860 | 0.0000 | 0.0000 |
| $B_{3,11}$ | 0.0971 | 0.0451 | 0.0194 | 0.0515 | 0.0245 | 0.0636 | 0.0699 | 0.0486 | 0.0099 | 0.0396 | 0.0131 | 0.0587 |
| $B_{3,21}$ | 0.0109 | 0.0946 | 0.0014 | 0.0334 | 0.0158 | 0.0941 | 0.0326 | 0.0281 | 0.0334 | 0.0408 | 0.0147 | 0.0816 |
| $B_{3,31}$ | 0.0118 | 0.0986 | 0.0228 | 0.0493 | 0.0134 | 0.0956 | 0.0073 | 0.0259 | 0.0016 | 0.0605 | 0.0119 | 0.0851 |
| $B_{3,12}$ | 0.0022 | 0.0127 | 0.0343 | 0.0470 | 0.0296 | 0.0926 | 0.0533 | 0.0404 | 0.0100 | 0.0487 | 0.0095 | 0.0813 |
| $B_{3,22}$ | 0.0753 | 0.0504 | 0.0116 | 0.0484 | 0.0159 | 0.0819 | 0.1011 | 0.0379 | 0.0105 | 0.0486 | 0.0058 | 0.0792 |
| $B_{3,32}$ | 0.0030 | 0.0950 | 0.0603 | 0.0666 | 0.0437 | 0.1057 | 0.0022 | 0.0334 | 0.0009 | 0.0765 | 0.0187 | 0.0831 |
| $B_{3,13}$ | 0.0096 | 0.0980 | 0.0033 | 0.0444 | 0.0120 | 0.0735 | 0.0032 | 0.0280 | 0.0056 | 0.0497 | 0.0010 | 0.0739 |
| $B_{3,23}$ | 0.0061 | 0.0796 | 0.0197 | 0.0429 | 0.0299 | 0.0864 | 0.0021 | 0.0261 | 0.0196 | 0.0526 | 0.0053 | 0.0733 |
| $B_{3,33}$ | 0.1308 | 0.0491 | 0.0007 | 0.0594 | 0.0224 | 0.0253 | 0.0861 | 0.0379 | 0.0016 | 0.0766 | 0.0047 | 0.0374 |

Note: The oracle estimate refers to the RLSE with $r_o = 2$ and the restriction that $\mathbf{B}_{o,2} = \mathbf{0}$.

On the other hand, for any $k \in \mathcal{S}_\phi^c$, by Lemma 3 (c), $\|n\phi_k(\widehat{\mathbf{\Pi}}_{1st})\|^\omega \to_d \|\tilde{\phi}_{o,k}\|^\omega = O_p(1)$, which implies that

$$n^{1-\frac{2}{\alpha}}\tilde{L}(n)^{-1}\lambda_{r,k,n} = \frac{n^{1-\frac{2}{\alpha}}n^\omega\tilde{L}(n)^{-1}\lambda^*_{r,k,n}}{\|n\phi_k(\widehat{\mathbf{\Pi}}_{1st})\|^\omega} \to_p \infty.$$

Similarly, for any $j \in \mathcal{S}_B$, by Lemma 3 (a) $\|\widehat{\mathbf{B}}_{1st,j}\|^\omega \to_p \|\mathbf{B}_{o,j}\|^\omega > 0$, then we have

$$n^{\tau_1}\delta_{b,n} = \frac{n^{\tau_1}m^\omega\lambda^*_{b,j,n}}{\|\widehat{\mathbf{B}}_{1st,j}\|^\omega} \to_p 0.$$

For any $j \in \mathcal{S}_B^c$, when $\alpha \in (1,2)$ or $\alpha = 1$ and $\tilde{L}(n) \to 0$ by Lemma 3 (a) $\|n^{\frac{1}{\alpha}}\widehat{\mathbf{B}}_{1st,j}\|^\omega = O_p(1)$, then we obtain

$$n^{1-\frac{1}{\alpha}}\tilde{L}(n)^{-1}\lambda_{b,j,n} = \frac{n^{1-\frac{1}{\alpha}}m^\omega n^{\frac{\omega}{\alpha}}\tilde{L}(n)^{-1}\lambda^*_{b,j,n}}{\|n^{\frac{1}{\alpha}}\widehat{\mathbf{B}}_{1st,j}\|^\omega} \to_p \infty,$$

and when $\alpha \in (0,1)$ or $\alpha = 1$ and $\tilde{L}(n) \to \infty$, $\|n\widehat{\mathbf{B}}_{1st,j}\|^\omega = O_p(1)$

$$n^{2-\frac{2}{\alpha}}\tilde{L}(n)^{-1}\lambda_{b,j,n} = \frac{n^{2-\frac{2}{\alpha}}m^\omega n^\omega\tilde{L}(n)^{-1}\lambda^*_{b,j,n}}{\|n\widehat{\mathbf{B}}_{1st,j}\|^\omega} \to_p \infty.$$

□

**Proof of Lemma 3.** (a) When $\alpha \in (1,2)$ or $\alpha = 1$ and $\tilde{L}(n) \to 0$, we have

$$\begin{aligned}
[(\widehat{\mathbf{\Pi}}_{1st}, \widehat{\mathbf{B}}_{1st}) - (\mathbf{\Pi}_o, \mathbf{B}_o)]\mathbf{Q}_B^{-1}\mathbf{D}_{n,B}^{-1} &= (\frac{1}{\tilde{a}_n}\mathbf{R}_{1n}\ \ \frac{1}{a_n^2}\mathbf{R}_{2n}) \begin{pmatrix} \frac{1}{a_n^2}\mathbf{S}_{11n} & \frac{\tilde{a}_n}{a_n^4}\mathbf{S}'_{21n} \\ \frac{1}{n\tilde{a}_n}\mathbf{S}_{21n} & \frac{1}{na_n^2}\mathbf{S}_{22n} \end{pmatrix}^{-1} \\
&= \Big[\{\frac{\mathbf{R}_{1n}}{\tilde{a}_n} - \frac{1}{\tilde{a}_n}\mathbf{R}_{2n}\mathbf{S}_{22n}^{-1}\mathbf{S}_{21n}\}\{\frac{\mathbf{S}_{11n}}{a_n^2} - \frac{1}{a_n^2}\mathbf{S}'_{21n}\mathbf{S}_{22n}^{-1}\mathbf{S}_{21n}\}^{-1}, \\
&\quad \{\frac{\mathbf{R}_{2n}}{a_n^2} - \frac{1}{a_n^2}\mathbf{R}_{1n}\mathbf{S}_{11n}^{-1}\mathbf{S}'_{21n}\}\{\frac{\mathbf{S}_{22n}}{na_n^2} - \frac{1}{na_n^2}\mathbf{S}_{21n}\mathbf{S}_{11n}^{-1}\mathbf{S}'_{21n}\}^{-1}\Big] \\
&\equiv [L_{1n}, L_{2n}].
\end{aligned}$$

By Lemma 2, we have

$$\begin{aligned}
L_{1n} &= \left\{\frac{1}{\tilde{a}_n}\mathbf{R}_{1n} - \frac{1}{n\tilde{a}_n}\mathbf{R}_{2n}(\frac{\mathbf{S}_{22n}}{na_n^2})^{-1}\frac{\mathbf{S}_{21n}}{a_n^2}\right\}\left\{\frac{1}{a_n^2}\mathbf{S}_{11n} - \frac{1}{a_n^2}\mathbf{S}_{21n}'(\frac{\mathbf{S}_{22n}}{na_n^2})^{-1}\frac{\mathbf{S}_{21n}}{na_n^2}\right\}^{-1} \\
&= \left\{\frac{1}{\tilde{a}_n}\mathbf{R}_{1n} + o_p(1)\right\}\left\{\frac{1}{a_n^2}\mathbf{S}_{11n} + o_p(1)\right\}^{-1} \to_d \mathbf{R}_1^*\mathbf{\Gamma}_{11}^{*-1},
\end{aligned}$$

where the second equality is by $\frac{1}{n\tilde{a}_n}\mathbf{R}_{2n} = o_p(1)$ and $\frac{1}{na_n^2}\mathbf{S}_{23n} = o_p(1)$. Similarly, we have

$$\begin{aligned}
L_{2n} &= \left\{\frac{1}{a_n^2}\mathbf{R}_{2n} - \frac{1}{a_n^2}\mathbf{R}_{1n}(\frac{\mathbf{S}_{11n}}{a_n^2})^{-1}\frac{\mathbf{S}_{21n}'}{a_n^2}\right\}\left\{\frac{1}{na_n^2}\mathbf{S}_{22n} - \frac{1}{na_n^2}\mathbf{S}_{21n}(\frac{\mathbf{S}_{11n}}{a_n^2})^{-1}\frac{\mathbf{S}_{21n}'}{a_n^2}\right\}^{-1} \\
&= \left\{\frac{1}{a_n^2}\mathbf{R}_{2n} + o_p(1)\right\}\left\{\frac{1}{na_n^2}\mathbf{S}_{22n} + o_p(1)\right\}^{-1} \to_d \mathbf{R}_2\mathbf{\Gamma}_{22}^{-1},
\end{aligned}$$

where the second equality is by $\frac{1}{a_n^2}\mathbf{R}_{1n} = o_p(1)$ and $\frac{1}{na_n^2}\mathbf{S}_{21n} = o_p(1)$. Hence we obtain

$$[(\widehat{\mathbf{\Pi}}_{1st}, \widehat{\mathbf{B}}_{1st}) - (\mathbf{\Pi}_o, \mathbf{B}_o)]\mathbf{Q}_B^{-1}\mathbf{D}_{n,B}^{-1} \to_d [\mathbf{R}_1^*\mathbf{\Gamma}_{11}^{*-1}, \mathbf{R}_2\mathbf{\Gamma}_{22}^{-1}].$$

When $\alpha \in (0,1)$ or $\alpha = 1$ and $\tilde{L}(n) \to \infty$, we have

$$\begin{aligned}
&[(\widehat{\mathbf{\Pi}}_{1st}, \widehat{\mathbf{B}}_{1st}) - (\mathbf{\Pi}_o, \mathbf{B}_o)]\mathbf{Q}_B^{-1}\mathbf{D}_{n,B}^{-1} \\
&= \Big[\{\frac{1}{\tilde{a}_n}\mathbf{R}_{1n} - \frac{1}{\tilde{a}_n}\mathbf{R}_{2n}\mathbf{S}_{22n}^{-1}\mathbf{S}_{21n}\}\{\frac{1}{n\tilde{a}_n}\mathbf{S}_{11n} - \frac{1}{n\tilde{a}_n}\mathbf{S}_{21n}'\mathbf{S}_{22n}^{-1}\mathbf{S}_{21n}\}^{-1}, \\
&\quad \{\frac{1}{a_n^2}\mathbf{R}_{2n} - \frac{1}{a_n^2}\mathbf{R}_{1n}\mathbf{S}_{11n}^{-1}\mathbf{S}_{21n}'\}\{\frac{1}{na_n^2}\mathbf{S}_{22n} - \frac{1}{na_n^2}\mathbf{S}_{21n}\mathbf{S}_{11n}^{-1}\mathbf{S}_{21n}'\}^{-1}\Big] \\
&\equiv [L_{1n}^*, L_{2n}].
\end{aligned}$$

By Lemma 2, we have $L_{2n} \to_d \mathbf{R}_2\mathbf{\Gamma}_{22}^{-1}$ and

$$\begin{aligned}
L_{1n}^* &= \left\{\frac{1}{\tilde{a}_n}\mathbf{R}_{1n}\frac{n\tilde{a}_n}{a_n^2} - \frac{1}{a_n^2}\mathbf{R}_{2n}(\frac{\mathbf{S}_{22n}}{na_n^2})^{-1}\frac{\mathbf{S}_{21n}}{n\tilde{a}_n}\frac{n\tilde{a}_n}{a_n^2}\right\}\left\{\frac{\mathbf{S}_{11n}}{n\tilde{a}_n}\frac{n\tilde{a}_n}{a_n^2} - \frac{\mathbf{S}_{21n}'}{na_n^2}(\frac{\mathbf{S}_{22n}}{na_n^2})^{-1}\frac{\mathbf{S}_{21n}}{n\tilde{a}_n}\frac{n\tilde{a}_n}{a_n^2}\right\}^{-1} \\
&= \left\{o_p(1) - \frac{1}{a_n^2}\mathbf{R}_{2n}(\frac{\mathbf{S}_{22n}}{na_n^2})^{-1}\frac{\mathbf{S}_{21n}}{a_n^2}\right\} \times \left\{\frac{1}{a_n^2}\mathbf{S}_{11n} + o_p(1)\right\}^{-1} \to_d -\mathbf{R}_2\mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{21}\mathbf{\Gamma}_{11}^{*-1}.
\end{aligned}$$

Hence we have

$$[(\widehat{\mathbf{\Pi}}_{1st}, \widehat{\mathbf{B}}_{1st}) - (\mathbf{\Pi}_o, \mathbf{B}_o)]\mathbf{Q}_B^{-1}\mathbf{D}_{n,B}^{-1} \to_d [-\mathbf{R}_2\mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{21}\mathbf{\Gamma}_{11}^{*-1}, \mathbf{R}_2\mathbf{\Gamma}_{22}^{-1}].$$

(b) This result follows directly by (a) and the continuous mapping theorem.

(c) Define $\mathbf{S}_n(\mu) = \frac{\mu}{n}\mathbf{I}_m - \widehat{\mathbf{\Pi}}_{1st}$. Then

$$\begin{aligned}
&|\mathbf{S}_n(\mu)| \times |(\boldsymbol{\beta}_o, \boldsymbol{\beta}_{o,\perp})|^2 \\
&= \left| \begin{pmatrix} \boldsymbol{\beta}_o' \\ \boldsymbol{\beta}_{o,\perp}' \end{pmatrix} \mathbf{S}_n(\mu) \begin{pmatrix} \boldsymbol{\beta}_o, & \boldsymbol{\beta}_{o,\perp} \end{pmatrix} \right| \\
&= \left| \begin{pmatrix} \boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o & \boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_{o,\perp} \\ \boldsymbol{\beta}_{o,\perp}'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o & \boldsymbol{\beta}_{o,\perp}'\mathbf{S}_n(\mu)\boldsymbol{\beta}_{o,\perp} \end{pmatrix} \right| \\
&= |\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o| \times |\boldsymbol{\beta}_{o,\perp}'\{\mathbf{S}_n(\mu) - \mathbf{S}_n(\mu)\boldsymbol{\beta}_o[\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o]^{-1}\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\}\boldsymbol{\beta}_{o,\perp}|.
\end{aligned}$$

Let $\mu_k^* = n\phi_k(\widehat{\mathbf{\Pi}}_{1st})(k = r_o + 1, \ldots, m)$. Then $\mu_k^*$ is a solution of the equation

$$|\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o| \times |\boldsymbol{\beta}_{o,\perp}'\{\mathbf{S}_n(\mu) - \mathbf{S}_n(\mu)\boldsymbol{\beta}_o[\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o]^{-1}\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\}\boldsymbol{\beta}_{o,\perp}| = 0.$$

Then, we can invoke the results in (a) to show

$$\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o = \frac{\mu}{n}\boldsymbol{\beta}_o'\boldsymbol{\beta}_o - \boldsymbol{\beta}_o'(\widehat{\mathbf{\Pi}}_{1st} - \mathbf{\Pi}_o)\boldsymbol{\beta}_o - \boldsymbol{\beta}_o'\mathbf{\Pi}_o\boldsymbol{\beta}_o \to_p \boldsymbol{\beta}_o'\boldsymbol{\alpha}_o\boldsymbol{\beta}_o'\boldsymbol{\beta}_o. \tag{S3.1}$$

From Assumption 1 (b), we have

$$|\boldsymbol{\beta}_o'\boldsymbol{\alpha}_o\boldsymbol{\beta}_o'\boldsymbol{\beta}_o| = |\boldsymbol{\beta}_o'\boldsymbol{\alpha}_o| \times |\boldsymbol{\beta}_o'\boldsymbol{\beta}_o| \neq 0.$$

Thus, $\mu_k^*$ are asymptotically the solutions of the following determinantal equation

$$|\boldsymbol{\beta}_{o,\perp}'\{\mathbf{S}_n(\mu) - \mathbf{S}_n(\mu)\boldsymbol{\beta}_o[\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o]^{-1}\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\}\boldsymbol{\beta}_{o,\perp}| = 0,$$

where

$$\boldsymbol{\beta}_{o,\perp}'\mathbf{S}_n(\mu)\boldsymbol{\beta}_{o,\perp} = \frac{\mu}{n}\mathbf{I}_{m-r_o} - \boldsymbol{\beta}_{o,\perp}'(\widehat{\boldsymbol{\Pi}}_{1st} - \boldsymbol{\Pi}_o)\boldsymbol{\beta}_{o,\perp}, \tag{S3.2}$$

$$\boldsymbol{\beta}_{o,\perp}'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o = \frac{\mu}{n}\boldsymbol{\beta}_{o,\perp}'\boldsymbol{\beta}_o - \boldsymbol{\beta}_{o,\perp}'\widehat{\boldsymbol{\Pi}}_{1st}\boldsymbol{\beta}_o \to -\boldsymbol{\beta}_{o,\perp}'\boldsymbol{\alpha}_o\boldsymbol{\beta}_o'\boldsymbol{\beta}_o, \tag{S3.3}$$

$$\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_{o,\perp} = -\boldsymbol{\beta}_o'(\widehat{\boldsymbol{\Pi}}_{1st} - \boldsymbol{\Pi}_o)\boldsymbol{\beta}_{o,\perp}. \tag{S3.4}$$

Using (S3.1)-(S3.4), we can show that

$$\begin{aligned}&\boldsymbol{\beta}_{o,\perp}'\{\mathbf{S}_n(\mu) - \mathbf{S}_n(\mu)\boldsymbol{\beta}_o[\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o]^{-1}\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\}\boldsymbol{\beta}_{o,\perp} \\ &= \frac{\mu}{n}\mathbf{I}_{m-r_o} - \boldsymbol{\beta}_{o,\perp}'[\mathbf{I}_m - \boldsymbol{\alpha}_o(\boldsymbol{\beta}_o'\boldsymbol{\alpha}_o)^{-1}\boldsymbol{\beta}_o' + o_p(1)](\widehat{\boldsymbol{\Pi}}_{1st} - \boldsymbol{\Pi}_o)\boldsymbol{\beta}_{o,\perp}.\end{aligned} \tag{S3.5}$$

Note that

$$n(\widehat{\boldsymbol{\Pi}}_{1st} - \boldsymbol{\Pi}_o)\boldsymbol{\beta}_{o,\perp} \to_d \mathbf{R}_2\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\alpha}_{o,\perp}'\boldsymbol{\beta}_{o,\perp}. \tag{S3.6}$$

Using (S3.5), (S3.6) and the equality

$$\boldsymbol{\alpha}_o(\boldsymbol{\beta}_o'\boldsymbol{\alpha}_o)^{-1}\boldsymbol{\beta}_o' + \boldsymbol{\beta}_{o,\perp}(\boldsymbol{\alpha}_{o,\perp}'\boldsymbol{\beta}_{o,\perp})^{-1}\boldsymbol{\alpha}_{o,\perp}' = \mathbf{I}_m, \tag{S3.7}$$

and since $\boldsymbol{\beta}_{o,\perp}'\boldsymbol{\beta}_{o,\perp} = \mathbf{I}_{m-r_o}$ we deduce that

$$\begin{aligned}&n\boldsymbol{\beta}_{o,\perp}'[\mathbf{I}_m - \boldsymbol{\alpha}_o(\boldsymbol{\beta}_o'\boldsymbol{\alpha}_o)^{-1}\boldsymbol{\beta}_o' + o_p(1)](\widehat{\boldsymbol{\Pi}}_{1st} - \boldsymbol{\Pi}_o)\boldsymbol{\beta}_{o,\perp} \\ &\to_d (\boldsymbol{\alpha}_{o,\perp}'\boldsymbol{\beta}_{o,\perp})^{-1}\boldsymbol{\alpha}_{o,\perp}'\mathbf{R}_2\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\alpha}_{o,\perp}'\boldsymbol{\beta}_{o,\perp},\end{aligned}$$

which implies

$$|n\boldsymbol{\beta}_{o,\perp}'\{\mathbf{S}_n(\mu) - \mathbf{S}_n(\mu)\boldsymbol{\beta}_o[\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\boldsymbol{\beta}_o]^{-1}\boldsymbol{\beta}_o'\mathbf{S}_n(\mu)\}\boldsymbol{\beta}_{o,\perp}| \to_d |\mu\mathbf{I}_{m-r_o} - \boldsymbol{\alpha}_{o,\perp}'\mathbf{R}_2\boldsymbol{\Gamma}_{22}^{-1}|. \tag{S3.8}$$

The result in (c) follows from (S3.8) and by continuous mapping. □

# S4 Proofs of Theorem 1 and Theorem 2

**Proof of Theorem 1.** Let $\mathbf{W}_{t-1} = [\mathbf{Y}'_{t-1}, \Delta\mathbf{X}'_{t-1}]'$, $\Delta\mathbf{Y} = (\Delta\mathbf{Y}_1, \ldots, \Delta\mathbf{Y}_n)_{m\times n}$ and $\mathbf{W}_{-1} = (\mathbf{W}_0, \ldots, \mathbf{W}_{n-1})_{m\times n}$. We can write

$$\begin{aligned} L_n(\mathbf{\Pi}, \mathbf{B}) &= \sum_{t=1}^{n} \|\Delta\mathbf{Y}_t - \boldsymbol{\theta}\mathbf{W}_{t-1}\|^2 \\ &= [vec(\Delta\mathbf{Y}) - (\mathbf{W}'_{-1}\otimes\mathbf{I}_m)vec(\boldsymbol{\theta})]'[vec(\Delta\mathbf{Y}) - (\mathbf{W}'_{-1}\otimes\mathbf{I}_m)vec(\boldsymbol{\theta})], \end{aligned}$$

where $\boldsymbol{\theta} = (\mathbf{\Pi}, \mathbf{B})$. Denote

$$\begin{aligned} V_n(\boldsymbol{\theta}) =& [vec(\Delta\mathbf{Y}) - (\mathbf{W}'_{-1}\otimes\mathbf{I}_m)vec(\boldsymbol{\theta})]'[vec(\Delta\mathbf{Y}) - (\mathbf{W}'_{-1}\otimes\mathbf{I}_m)vec(\boldsymbol{\theta})] \\ &+ n\sum_{j=1}^{p}\lambda_{b,j,n}\|\mathbf{B}_j\| + n\sum_{k=1}^{m}\lambda_{r,k,n}\|\Phi_{n,k}(\mathbf{\Pi})\|. \end{aligned} \tag{S4.1}$$

Set $\widehat{\boldsymbol{\theta}}_n = (\widehat{\mathbf{\Pi}}_n, \widehat{\mathbf{B}}_n)$ and $\boldsymbol{\theta}_o = (\mathbf{\Pi}_o, \mathbf{B}_o)$. Define an infeasible estimator $\widetilde{\boldsymbol{\theta}}_n = (\mathbf{\Pi}_{n,f}, \mathbf{B}_o)$. Then

$$(\widetilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_o)\mathbf{Q}_B^{-1}\mathbf{D}_{n,B}^{-1} = (\mathbf{\Pi}_{n,f} - \mathbf{\Pi}_o, \mathbf{0})\mathbf{Q}_B^{-1}\mathbf{D}_{n,B}^{-1} = O_p(1), \tag{S4.2}$$

where the last equality holds by (7.3). By definition of $\widehat{\boldsymbol{\theta}}_n$, $V_n(\widehat{\boldsymbol{\theta}}_n) \le V_n(\widetilde{\boldsymbol{\theta}}_n)$. Furthermore, by (S4.1), we can show that

$$\begin{aligned} & vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^{n}\mathbf{W}_{t-1}\mathbf{W}'_{t-1}\otimes\mathbf{I}_m)vec((\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n) + 2vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'vec(\sum_{t=1}^{n}\mathbf{W}_{t-1}\boldsymbol{\varepsilon}'_t) \\ & + 2vec((\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^{n}\mathbf{W}_{t-1}\mathbf{W}'_{t-1}\otimes\mathbf{I}_m)vec(\boldsymbol{\theta}_o - \widetilde{\boldsymbol{\theta}}_n) \le d_n, \end{aligned} \tag{S4.3}$$

where $d_n = n\sum_{j=1}^{p}\lambda_{b,j,n}[\|\mathbf{B}_{o,j}\| - \|\widehat{\mathbf{B}}_{n,j}\|] + n\sum_{k=1}^{m}\lambda_{r,k,n}[\|\Phi_{n,k}(\mathbf{\Pi}_{n,f})\| - \|\Phi_{n,k}(\widehat{\mathbf{\Pi}}_n)\|]$.

When $r_o = m$, $\mathbf{W}_t$ is a stationary process. By Theorem A.2 in She and Ling (2020) and (S4.2), we have

$$c_{1n} \equiv \|a_n^{-2}\sum_{t=1}^{n}\mathbf{W}_{t-1}\boldsymbol{\varepsilon}'_t\| = O_p[n^{-\frac{1}{\alpha}}\tilde{L}(n)^{-1}], \tag{S4.4}$$

$$c_{2n} \equiv m\|a_n^{-2}\sum_{t=1}^{n}\mathbf{W}_{t-1}\mathbf{W}'_{t-1}\|\|\boldsymbol{\theta}_o - \widetilde{\boldsymbol{\theta}}_n\| = O_p[n^{-\frac{1}{\alpha}}\tilde{L}(n)^{-1}]. \tag{S4.5}$$

Let $\mu_{1n,min}$ be the smallest eigenvalue of $a_n^{-2}\sum_{t=1}^n \mathbf{W}_{t-1}\mathbf{W}'_{t-1}$, which is positive w.p.a.1. By (S4.3)-(S4.5), we get

$$\mu_{1n,min}\|\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n\|^2 - 2(c_{1n}+c_{2n})\|\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n\| - a_n^{-2}d_n \le 0. \tag{S4.6}$$

By (7.1), (7.2), $\delta_{r,n} = o_p(1)$ and $\delta_{b,n} = o_p(1)$, we have $a_n^{-2}d_n = o_p(1)$. So, it is straightforward to deduce that $\|\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n\| = o_p(1)$ from (S4.6). The consistency of $\widehat{\boldsymbol{\theta}}_n$ follows from the triangle inequality and the consistency of $\widetilde{\boldsymbol{\theta}}_n$.

When $0 \le r_o < m$, we first consider the case when $\alpha \in (1,2)$, or $\alpha = 1$ and $\tilde{L}(n) \to 0$. Denote $\mathbf{B}_n = (\mathbf{D}_{n,B}\mathbf{Q}_B)^{-1}$, and $\mathbf{H}_B = diag(\frac{\tilde{a}_n}{a_n}\mathbf{I}_{r_o+mp}, n^{-\frac{1}{2}}a_n\mathbf{I}_{m-r_o})$. By Lemma 2, we can deduce

$$\mathbf{H}_B^{-1}\mathbf{B}_n^{-1}\sum_{t=1}^n \mathbf{W}_{t-1}\mathbf{W}'_{t-1}\mathbf{B}_n^{-1'}\mathbf{H}_B^{-1} \to_d \begin{pmatrix} \sum_{l=0}^\infty \mathbf{B}_l\mathbf{S}_1\mathbf{B}'_l & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}_{22} \end{pmatrix}. \tag{S4.7}$$

By Lemma 2 and (S4.2), we have

$$e_{1,n} \equiv \|\mathbf{H}_B^{-1}\mathbf{B}_n^{-1}\sum_{t=1}^n \mathbf{W}_{t-1}\boldsymbol{\varepsilon}'_t\| = O_p(n^{\frac{1}{\alpha}-\frac{1}{2}}), \tag{S4.8}$$

$$e_{2,n} \equiv m\|\mathbf{H}_B^{-1}\mathbf{B}_n^{-1}\sum_{t=1}^n \mathbf{W}_{t-1}\mathbf{W}'_{t-1}\mathbf{B}_n^{-1'}\mathbf{H}_B^{-1}\|\|(\widetilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_o)\mathbf{B}_n\mathbf{H}_B\| = O_p(n^{\frac{1}{\alpha}-\frac{1}{2}}). \tag{S4.9}$$

Let $\mu_{2n,min}$ be the smallest eigenvalue of $\mathbf{H}_B^{-1}\mathbf{B}_n^{-1}\sum_{t=1}^n \mathbf{W}_{t-1}\mathbf{W}'_{t-1}\mathbf{B}_n^{-1'}\mathbf{H}_B^{-1}$, which is positive w.p.a.1. Then

$$vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^n \mathbf{W}_{t-1}\mathbf{W}'_{t-1}\otimes\mathbf{I}_m)vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n) \ge \mu_{2n,min}\|(\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n)\mathbf{B}_n\mathbf{H}_B\|^2. \tag{S4.10}$$

Note that

$$|vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'vec(\sum_{t=1}^n \mathbf{W}_{t-1}\boldsymbol{\varepsilon}'_t)| \le \|(\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n)\mathbf{B}_n\mathbf{H}_B\|e_{1,n}, \tag{S4.11}$$

$$|vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^{n} \mathbf{W}_{t-1}\mathbf{W}_{t-1}' \otimes \mathbf{I}_m) vec(\boldsymbol{\theta}_o - \widetilde{\boldsymbol{\theta}}_n)| \leq \|(\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n)\mathbf{B}_n\mathbf{H}_B\| e_{2,n}. \quad \text{(S4.12)}$$

From (S4.3) and (S4.10)-(S4.12), we have

$$\mu_{2n,min}\|(\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n)\mathbf{B}_n\mathbf{H}_B\|^2 - 2(e_{1,n} + e_{2,n})\|(\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n)\mathbf{B}_n\mathbf{H}_B\| \leq d_n. \quad \text{(S4.13)}$$

By the definition of the penalty function and Lemma 3, we find

$$d_n \leq n \sum_{j \in \mathcal{S}_B} \lambda_{b,j,n} \|\mathbf{B}_{o,j}\| + n \sum_{k \in \mathcal{S}_\phi} \lambda_{r,k,n} \|\Phi_{n,k}(\boldsymbol{\Pi}_{n,f})\| = O_p(n\delta_{b,n} + n\delta_{r,n}).$$

From (S4.13), we have

$$(\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n)\mathbf{B}_n\mathbf{H}_B = O_p(n^{\frac{1}{\alpha}-\frac{1}{2}} + n^{\frac{1}{2}}\delta_{d,n}^{\frac{1}{2}} + n^{\frac{1}{2}}\delta_{r,n}^{\frac{1}{2}}), \quad \text{(S4.14)}$$

which together with (S4.2) and the definition of $\mathbf{B}_n$ and $\mathbf{H}_B$ implies

$$\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_o = o_p(1), \quad \text{(S4.15)}$$

which implies the consistency of $\widehat{\boldsymbol{\theta}}_n$.

When $\alpha \in (0,1)$ or $\alpha = 1$ and $\tilde{L}(n) \to \infty$, using the similar argument, we can show the consistency of $\widehat{\boldsymbol{\theta}}_n$. □

**Proof of Theorem 2.** By the triangle inequality and (7.2), we have

$$\begin{aligned}
&|\sum_{k=1}^{m} \lambda_{r,k,n}[\|\Phi_{n,k}(\boldsymbol{\Pi}_{n,f})\| - \|\Phi_{n,k}(\widehat{\boldsymbol{\Pi}}_n)\|]| + |\sum_{j=1}^{p} \lambda_{b,j,n}[\|\mathbf{B}_{o,j}\| - \|\widehat{\mathbf{B}}_{n,j}\|]| \\
&\quad \leq |\sum_{k=1}^{r_0} \lambda_{r,k,n}[\|\Phi_{n,k}(\boldsymbol{\Pi}_{n,f})\| - \|\Phi_{n,k}(\widehat{\boldsymbol{\Pi}}_n)\|]| + |\sum_{j \in \mathcal{S}_B} \lambda_{b,j,n}[\|\mathbf{B}_{o,j}\| - \|\widehat{\mathbf{B}}_{n,j}\|]| \\
&\quad \leq c(\delta_{b,n} + \delta_{r,n})\|\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n\|, \qquad \text{(S4.16)}
\end{aligned}$$

where $c$ is some positive constant. Using (S4.16) and invoking (S4.3), we get

$$\begin{aligned}
& vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^{n} \mathbf{W}_{t-1}\mathbf{W}_{t-1}' \otimes \mathbf{I}_m) vec((\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n) \\
& \quad + 2vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)' vec(\sum_{t=1}^{n} \mathbf{W}_{t-1}\boldsymbol{\varepsilon}_t') \\
& \quad + 2vec((\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^{n} \mathbf{W}_{t-1}\mathbf{W}_{t-1}' \otimes \mathbf{I}_m) vec(\boldsymbol{\theta}_o - \widetilde{\boldsymbol{\theta}}_n) \\
& \leq nc(\delta_{b,n} + \delta_{r,n}) \|\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n\|. \qquad \text{(S4.17)}
\end{aligned}$$

When $r_o = m$, we use (S4.17) to obtain

$$\mu_{1n,min} \|\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n\|^2 - 2(c_{1n} + c_{2n} + n a_n^{-2} \delta_n) \|\widehat{\boldsymbol{\theta}}_n - \widetilde{\boldsymbol{\theta}}_n\| \leq 0, \qquad \text{(S4.18)}$$

where $\delta_n = \delta_{r,n} + \delta_{b,n}$, $\mu_{1n,min}$, $c_{1n}$ and $c_{1n}$ are defined in (S4.6). (S4.18) and (7.3) lead to

$$\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_o = O_p(n^{-\frac{1}{\alpha}} \tilde{L}(n)^{-1} + n^{1-\frac{2}{\alpha}} \tilde{L}(n)^{-1} \delta_n).$$

When $0 \leq r_o < m$ and $\alpha \in (1,2)$ or $\alpha = 1$ and $\tilde{L}(n) \to 0$, denote

$$\begin{aligned}
& \nu_1 = a_n \|[(\boldsymbol{\Pi}_{n,f} - \widehat{\boldsymbol{\Pi}}_n)\bar{\boldsymbol{\beta}},\ \mathbf{B}_o - \widehat{\mathbf{B}}_n]\|,\ \nu_2 = n^{\frac{1}{2}} a_n \|(\boldsymbol{\Pi}_{n,f} - \widehat{\boldsymbol{\Pi}}_n)\bar{\boldsymbol{\beta}}_\perp\|, \\
& \mathbf{H}_B^{-1}\mathbf{B}_n^{-1} \sum_{t=1}^{n} \mathbf{W}_{t-1}\boldsymbol{\varepsilon}_t' = [\mathbf{e}_{1,1}', \mathbf{e}_{1,2}']',\ \text{and} \\
& \mathbf{H}_B^{-1}\mathbf{B}_n^{-1} \sum_{t=1}^{n} \mathbf{W}_{t-1}\mathbf{W}_{t-1}'\mathbf{B}_n^{-1'}\mathbf{H}_B^{-1}[(\widetilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_o)\mathbf{B}_n\mathbf{H}_B]' = [\mathbf{e}_{2,1}', \mathbf{e}_{2,2}']',
\end{aligned}$$

where $\mathbf{e}_{1,2}$ and $\mathbf{e}_{2,2}$ are the last $m - r_o$ rows of the corresponding matrices. By Lemma 2 and (S4.2), we have

$$\mathbf{e}_{1,1} = O_p(\frac{\tilde{a}_n}{a_n}), \mathbf{e}_{1,2} = O_p(n^{-\frac{1}{2}} a_n), \mathbf{e}_{2,1} = O_p(\frac{\tilde{a}_n}{a_n}), \mathbf{e}_{2,2} = O_p(n^{-\frac{1}{2}} a_n). \qquad \text{(S4.19)}$$

Note that

$$vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^{n} \mathbf{W}_{t-1}\mathbf{W}_{t-1}' \otimes \mathbf{I}_m) vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n) \geq \mu_{2n,min}(\nu_1^2 + \nu_2^2),$$

$$|vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'vec(\sum_{t=1}^{n}\mathbf{W}_{t-1}\boldsymbol{\varepsilon}_t')| \leq \nu_1\|\mathbf{e}_{1,1}\| + \nu_2\|\mathbf{e}_{1,2}\|,$$

and

$$|vec(\widetilde{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\theta}}_n)'(\sum_{t=1}^{n}\mathbf{W}_{t-1}\mathbf{W}_{t-1}' \otimes \mathbf{I}_m)vec(\boldsymbol{\theta}_o - \widetilde{\boldsymbol{\theta}}_n)| \leq \nu_1\|\mathbf{e}_{2,1}\| + \nu_2\|\mathbf{e}_{2,2}\|.$$

Then, we can deduce that

$$\mu_{2n,min}(\nu_1^2+\nu_2^2) - 2(\|\mathbf{e}_{1,1}\| + \|\mathbf{e}_{2,1}\|)\nu_1 - 2(\|\mathbf{e}_{1,2}\| + \|\mathbf{e}_{2,2}\|)\nu_2 \leq cna_n^{-1}\delta_n\nu_1 + cn^{\frac{1}{2}}a_n^{-1}\delta_n\nu_2,$$

which together with (S4.19) implies that $\nu_1 = O_p(\tilde{a}_n/a_n + na_n^{-1}\delta_n)$ and $\nu_2 = O_p(n^{-\frac{1}{2}}a_n + n^{\frac{1}{2}}a_n^{-1}\delta_n)$. So, we get

$$a_n^2/\tilde{a}_n[(\widehat{\mathbf{\Pi}}_n - \mathbf{\Pi}_o)\bar{\boldsymbol{\beta}},\ \widehat{\mathbf{B}}_n - \mathbf{B}_o] = O_p(1 + n\tilde{a}_n^{-1}\delta_n),$$

$$n(\widehat{\mathbf{\Pi}}_n - \mathbf{\Pi}_o)\bar{\boldsymbol{\beta}}_\perp = O_p(1 + na_n^{-2}\delta_n).$$

When $0 \leq r_o < m$ and $\alpha \in (0,1)$ or $\alpha = 1$ and $\tilde{L}(n) \to \infty$, using the same arguement as previous, we can show that

$$(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_o)\mathbf{B}_n = O_p(1),$$

which finishes the proof.

□