Supplementary Material for "Large Global Volatility Matrix Analysis Based on Observation Structural Information"

Sung Hoon Choi* Donggyu Kim†

July, 2024

Abstract

This supplement contains all the remaining proofs, the detailed explanation of the Double-POET (Choi and Kim, 2023) estimation procedure and its asymptotic theory, data generating process for simulation study, and additional tables for empirical study.

S Appendix

Let $\lambda_{\min}(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A})$ denote the minimum and maximum eigenvalues of matrix $\mathbf{A}$, respectively. In addition, we denote by $\|\mathbf{A}\|_F$, $\|\mathbf{A}\|_2$ (or $\|\mathbf{A}\|$ for short), $\|\mathbf{A}\|_1$, $\|\mathbf{A}\|_\infty$, and $\|\mathbf{A}\|_{\max}$ the Frobenius norm, operator norm, l_1-norm, l_∞-norm and elementwise norm, which are defined, respectively, as $\|\mathbf{A}\|_F = \mathrm{tr}^{1/2}(\mathbf{A}'\mathbf{A})$, $\|\mathbf{A}\|_2 = \lambda_{\max}^{1/2}(\mathbf{A}'\mathbf{A})$, $\|\mathbf{A}\|_1 = \max_j \sum_i |a_{ij}|$, $\|\mathbf{A}\|_\infty = \max_i \sum_j |a_{ij}|$, and $\|\mathbf{A}\|_{\max} = \max_{i,j} |a_{ij}|$. When $\mathbf{A}$ is a vector, the maximum norm is denoted as $\|\mathbf{A}\|_\infty = \max_i |a_i|$, and both $\|\mathbf{A}\|$ and $\|\mathbf{A}\|_F$ are equal to the Euclidean norm. We denote $\mathrm{diag}(\mathbf{A}_1, \dots, \mathbf{A}_n)$ with the diagonal block entries as $\mathbf{A}_1, \dots, \mathbf{A}_n$.

*Department of Economics, University of Connecticut, Storrs, CT 06269, USA. E-mail: sung_hoon.choi@uconn.edu.

†Department of Economics, University of California, Riverside, CA 92521, USA. Email: donggyu.kim@ucr.edu.

S.1 Rank Choice

To implement S-POET, we need to determine the rank k^*_{sq} and the number of factors, which are unknown in practice. We note that each (s,q)th off-diagonal partitioned block $\mathbf{R}_{h,sq}$ in (2.6) is a low-rank matrix, and each rank is less than or equal to the number of global factors (i.e., $k^*_{sq} \le k$). Thus, to determine the rank and number of global factors, we can use the data-driven methods proposed by Ahn and Horenstein (2013); Bai and Ng (2002); Onatski (2010). For example, the rank k^*_{sq} can be determined by finding the largest singular value gap such that $\max_{i \le \bar{k}_{sq}}(\widehat{\xi}_i - \widehat{\xi}_{i+1})$, where $\bar{k}_{sq} = \min\{p_s, p_q\}$. In this paper, to consistently estimate k, we employ the modified version of the eigenvalue ratio method, introduced by Choi and Kim (2023), based on $\widehat{\mathbf{\Sigma}}_h$.

S.2 Double-POET procedure

We decompose the covariance matrix of the sth continent as follows:

$$\mathbf{\Sigma}^s = \mathbf{\Sigma}^s_g + \mathbf{\Sigma}^s_l + \mathbf{\Sigma}^s_u.$$

Then, each component as well as $\mathbf{\Sigma}^s$ can be estimated by the Double-POET procedure (Choi and Kim, 2023) as follows:

1. Given a sample covariance matrix, $\widehat{\mathbf{\Sigma}}^s$, using T observations, let $\{\widehat{\delta}^s_i, \widehat{v}^s_i\}_{i=1}^p$ be the eigenvalues and eigenvectors of $\widehat{\mathbf{\Sigma}}^s$ in decreasing order. We compute

$$\widehat{\mathbf{\Sigma}}^{s,\mathcal{D}}_g = \widehat{\mathbf{V}}^s \widehat{\mathbf{\Gamma}}^s \widehat{\mathbf{V}}^{s\prime},$$

where $\widehat{\mathbf{\Gamma}}^s = \mathrm{diag}(\widehat{\delta}^s_1, \dots, \widehat{\delta}^s_k)$ and $\widehat{\mathbf{V}}^s = (\widehat{v}^s_1, \dots, \widehat{v}^s_k)$.

2. Define $\widehat{\mathbf{\Sigma}}^{l,s}_E$ as each $p_l \times p_l$ diagonal block of $\widehat{\mathbf{\Sigma}}^s_E = \widehat{\mathbf{\Sigma}}^s - \widehat{\mathbf{\Sigma}}^{s,\mathcal{D}}_g$. For the lth block, let $\{\widehat{\kappa}^{l,s}_i, \widehat{\eta}^{l,s}_i\}_{i=1}^{p_l}$ be the eigenvalues and eigenvectors of $\widehat{\mathbf{\Sigma}}^{l,s}_E$ in decreasing order. Then, we

compute

$$\widehat{\boldsymbol{\Sigma}}_l^{s,\mathcal{D}} = \widehat{\boldsymbol{\Phi}}^s \widehat{\boldsymbol{\Psi}}^s \widehat{\boldsymbol{\Phi}}^{s\prime},$$

where $\widehat{\boldsymbol{\Psi}}^s = \mathrm{diag}(\widehat{\boldsymbol{\Psi}}^1, \ldots, \widehat{\boldsymbol{\Psi}}^{L_s})$ for $\widehat{\boldsymbol{\Psi}}^l = \mathrm{diag}(\widehat{\kappa}_1^{l,s}, \ldots, \widehat{\kappa}_{r_l}^{l,s})$, and the block diagonal matrix $\widehat{\boldsymbol{\Phi}}^s = \mathrm{diag}(\widehat{\boldsymbol{\Phi}}^1, \ldots, \widehat{\boldsymbol{\Phi}}^{L_s})$ for $\widehat{\boldsymbol{\Phi}}^l = (\widehat{\eta}_1^{l,s}, \ldots, \widehat{\eta}_{r_l}^{l,s})$ for $l = 1, 2, \ldots, L_s$, where L_s is the number of countries in continent s.

3. Let $\widehat{\boldsymbol{\Sigma}}_u^s = \widehat{\boldsymbol{\Sigma}}^s - \widehat{\boldsymbol{\Sigma}}_g^{s,\mathcal{D}} - \widehat{\boldsymbol{\Sigma}}_l^{s,\mathcal{D}}$ be the principal orthogonal complement. We apply the adaptive thresholding method on $\widehat{\boldsymbol{\Sigma}}_u^s = (\widehat{\sigma}_{u,ij})_{p\times p}$ following Bickel and Levina (2008) and Fan, Liao, and Mincheva (2013). Specifically, define $\widehat{\boldsymbol{\Sigma}}_u^{s,\mathcal{D}}$ as the thresholded error covariance matrix estimator:

$$\widehat{\boldsymbol{\Sigma}}_u^{s,\mathcal{D}} = (\widehat{\sigma}_{u,ij}^{s,\mathcal{D}})_{p_s \times p_s}, \qquad \widehat{\sigma}_{u,ij}^{s,\mathcal{D}} = \begin{cases} \widehat{\sigma}_{u,ij}, & i = j \\ s_{ij}(\widehat{\sigma}_{u,ij}) I(|\widehat{\sigma}_{u,ij}| \geq \tau_{ij}), & i \neq j \end{cases},$$

where an entry-dependent threshold $\tau_{ij} = \tau(\widehat{\sigma}_{u,ii}\widehat{\sigma}_{u,jj})^{1/2}$ and $s_{ij}(\cdot)$ is a generalized thresholding function (e.g., hard or soft thresholding; see Cai and Liu, 2011; Rothman, Levina, and Zhu, 2009). The thresholding constant is determined by $\tau \asymp \omega_T$, where ω_T is defined in Theorem 3.1.

4. The final estimator of $\boldsymbol{\Sigma}^s$ is then defined as

$$\widehat{\boldsymbol{\Sigma}}^{s,\mathcal{D}} = \widehat{\boldsymbol{\Sigma}}_g^{s,\mathcal{D}} + \widehat{\boldsymbol{\Sigma}}_l^{s,\mathcal{D}} + \widehat{\boldsymbol{\Sigma}}_u^{s,\mathcal{D}}.$$

By using the proof of Theorem 3.1 of Choi and Kim (2023) and Assumption 3.1, we can obtain the following results: for each continent $s \in \{1, \ldots, S\}$,

$$\|\widehat{\boldsymbol{\Sigma}}_g^{s,\mathcal{D}} - \boldsymbol{\Sigma}_g^s\|_{\max} = O_P(p^{\frac{5}{2}(1-a_1)}\sqrt{\log p/T} + 1/p^{\frac{5a_1}{2}-\frac{3}{2}-c}), \tag{S.1}$$

$$\|\widehat{\boldsymbol{\Sigma}}_l^{s,\mathcal{D}} - \boldsymbol{\Sigma}_l^s\|_{\max} = O_P(\omega_T), \tag{S.2}$$

$$\|\widehat{\boldsymbol{\Sigma}}_u^{s,\mathcal{D}} - \boldsymbol{\Sigma}_u^s\|_{\max} = O_P(\omega_T), \tag{S.3}$$

where $\omega_T = p^{\frac{5}{2}(1-a_1)+\frac{5}{2}c(1-a_2)}\sqrt{\log p/T} + 1/p^{\frac{5}{2}a_1-\frac{3}{2}+c(\frac{5}{2}a_2-\frac{7}{2})} + m_p/\sqrt{p^{c(5a_2-3)}}$.

S.3 Double-POET Using Lower-Frequency Data

To capture the global factor, local factor, and idiosyncratic components, we can apply the Double-POET method. However, when considering international stocks, practitioners commonly use lower-frequency data to minimize the impact of different observation time points. Let $\widehat{\mathbf{\Sigma}}_h = T^{-\alpha}\sum_{t=1}^{T^\alpha}(y_t - \bar{y})(y_t - \bar{y})'$ be the sample covariance matrix using d-day return data. Then, $d^{-1}\widehat{\mathbf{\Sigma}}_h$ is used for the initial pilot estimator for covariance matrix $\mathbf{\Sigma}$, since $\widehat{\mathbf{\Sigma}}_h$ is the amplified estimator by d, which slowly grows (see Remark S.1). Let $\widehat{\mathbf{\Gamma}} = \mathrm{diag}(\widehat{\delta}_1, \dots, \widehat{\delta}_k)$ and $\widehat{\mathbf{V}} = (\widehat{v}_1, \dots, \widehat{v}_k)$ be the leading eigenvalues and their corresponding eigenvectors of $d^{-1}\widehat{\mathbf{\Sigma}}_h$. Next, let $\widehat{\mathbf{\Sigma}}_E^l$ be the lth $p_l \times p_l$ diagonal block of $\widehat{\mathbf{\Sigma}}_E = d^{-1}\widehat{\mathbf{\Sigma}}_h - \widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}'$. Let $\widehat{\mathbf{\Psi}}^l = \mathrm{diag}(\widehat{\kappa}_1^l, \dots, \widehat{\kappa}_{r_l}^l)$ and $\widehat{\mathbf{\Phi}}^l = (\widehat{\eta}_1^l, \dots, \widehat{\eta}_{r_l}^l)$ be the leading eigenvalues and their corresponding eigenvectors of $\widehat{\mathbf{\Sigma}}_E^l$. Let $\widehat{\mathbf{\Psi}} = \mathrm{diag}(\widehat{\mathbf{\Psi}}^1, \dots, \widehat{\mathbf{\Psi}}^L)$, $\widehat{\mathbf{\Phi}} = \mathrm{diag}(\widehat{\mathbf{\Phi}}^1, \dots, \widehat{\mathbf{\Phi}}^L)$, and $\widehat{\mathbf{\Sigma}}_u = d^{-1}\widehat{\mathbf{\Sigma}}_h - \widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \widehat{\mathbf{\Phi}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{\Phi}}'$. Then, the Double-POET estimator is defined as follows:

$$\widehat{\mathbf{\Sigma}}^{\mathcal{D}} = \widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' + \widehat{\mathbf{\Phi}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{\Phi}}' + \widehat{\mathbf{\Sigma}}_u^{\mathcal{D}},$$

where $\widehat{\mathbf{\Sigma}}_u^{\mathcal{D}}$ is the thresholded error covariance matrix estimator based on $\widehat{\mathbf{\Sigma}}_u = (\widehat{\sigma}_{u,ij})_{p\times p}$ (Bickel and Levina, 2008; Fan, Liao, and Mincheva, 2013):

$$\widehat{\mathbf{\Sigma}}_u^{\mathcal{D}} = (\widehat{\sigma}_{u,ij}^{\mathcal{D}})_{p\times p}, \qquad \widehat{\sigma}_{u,ij}^{\mathcal{D}} = \begin{cases} \widehat{\sigma}_{u,ij}, & i = j \\ s_{ij}(\widehat{\sigma}_{u,ij})I(|\widehat{\sigma}_{u,ij}| \geq \tau_{ij}), & i \neq j \end{cases},$$

where an entry-dependent threshold $\tau_{ij} = \tau(\widehat{\sigma}_{u,ii}\widehat{\sigma}_{u,jj})^{1/2}$ and $s_{ij}(\cdot)$ is a generalized thresholding function such as hard thresholding ($s_{ij}(x) = x$), soft thresholding ($s_{ij}(x) = \mathrm{sgn}(x)(|x| - \tau_{ij})$, where $\mathrm{sgn}(\cdot)$ is the sign function) and the adaptive lasso (see Rothman, Levina, and Zhu, 2009). The thresholding constant is determined by $\tau \asymp \omega_{T^\alpha}$, where ω_{T^α} is defined in Theorem S.1.

Assumption S.1. *Let $d = T^{1-\alpha}$ for $\alpha \in (0,1)$. The sample covariance matrix using d-day return data, $\widehat{\mathbf{\Sigma}}_h = T^{-\alpha}\sum_{t=1}^{T^\alpha}(y_t - \bar{y})(y_t - \bar{y})'$, satisfies*

$$\|d^{-1}\widehat{\mathbf{\Sigma}}_h - \mathbf{\Sigma}_h\|_{\max} = O_P(\sqrt{\log p/T^\alpha}).$$

Remark S.1. Assumption S.1 is similar to Assumption 3.1(iii) in Choi and Kim (2023). However, to match the scale of $\mathbf{\Sigma}_h$, the sample covariance matrix using d-day return data, $\widehat{\mathbf{\Sigma}}_h$, needs to be divided by d. To illustrate this point, consider the case when $p = 1$ and a collection of T i.i.d. random variables, $\{y_1, \ldots, y_T\}$, where y_t is a log-return defined as $y_t = \log x_t - \log x_{t-1}$ and x_t is the asset price at time t. Assume that y_t has a mean of zero and a variance of σ^2. We can obtain lower-frequency data by summing daily log-returns for each d window size, and this is equivalent to sub-sampling based on the price data. The variance of the resulting d-day return data is $d \times \sigma^2$. Therefore, we can compare the estimator $\widehat{\sigma}_h/d$ with the true variance σ^2, where $\widehat{\sigma}_h = T^{-\alpha}\sum_{t=1}^{T^\alpha}(y_t - \bar{y})^2$ using d-day log-returns. Using this fact and Assumption 3.1(iii), we can impose the above element-wise convergence condition. However, Structured-POET does not require this assumption because it can remove the scale issue by using the correlation matrix and recovering with daily-based variance estimator $\widehat{\mathbf{D}}$ in Section 2. In the simulation study, we used $d^{-1}\widehat{\mathbf{\Sigma}}_h$ for the initial sample covariance matrix.

Similar to the proofs of Choi and Kim (2023), we can show that Double-POET yields the following convergence rates.

Theorem S.1. *Suppose that $m_p = o(p^{c(5a_2-3)/2})$ and Assumptions 3.1 and S.1 hold. Let $\omega_{T^\alpha} = p^{\frac{5}{2}(1-a_1)+\frac{5}{2}c(1-a_2)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5}{2}a_1-\frac{3}{2}+c(\frac{5}{2}a_2-\frac{7}{2})} + m_p/\sqrt{p^{c(5a_2-3)}}$. If $m_p\omega_{T^\alpha}^{1-q} = o(1)$, we have*

$$\|\widehat{\mathbf{\Sigma}}^{\mathcal{D}} - \mathbf{\Sigma}\|_{\max} = O_P\left(\omega_{T^\alpha}\right), \quad \text{(S.4)}$$

$$\|(\widehat{\mathbf{\Sigma}}^{\mathcal{D}})^{-1} - \mathbf{\Sigma}^{-1}\|_2 = O_P\left(m_p\omega_{T^\alpha}^{1-q} + p^{\frac{c}{2}(1-a_2)}\omega_{T^\alpha} + p^{3(1-a_1)}\left(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\right) + \frac{1}{p^{3a_1-2-c}}\right). \quad \text{(S.5)}$$

In addition, if $a_1 > \frac{3}{4}$ and $a_2 > \frac{3}{4}$, we have

$$
\begin{aligned}
\|\widehat{\mathbf{\Sigma}}^{\mathcal{D}} - \mathbf{\Sigma}\|_{\Sigma} =& O_P\Big(m_p \omega_{T^\alpha}^{1-q} + p^{\frac{11}{2} - 5a_1 + 5c(1-a_2)} \big(\frac{\log p}{T^\alpha} + \frac{1}{T^{2(1-\alpha)\beta}} \big) \\
&+ \frac{1}{p^{5a_1 - \frac{7}{2} - c(7-5a_2)}} + \frac{m_p^2}{p^{5ca_2 - 3c - \frac{1}{2}}} \Big).
\end{aligned} \tag{S.6}
$$

Remark S.2. For simplicity, consider $m_p = O(1)$, $a_1 = 1$, and $a_2 = 1$, and ignore the log order terms. Define the optimal $\alpha^* = \frac{2\beta}{1+2\beta}$ (see Remark 3.2). With $\alpha = \alpha^*$, we have

$$
\begin{aligned}
\|\widehat{\mathbf{\Sigma}}^{\mathcal{D}} - \mathbf{\Sigma}\|_{\Sigma} &= O_P\left(\left(\frac{1}{T^{\frac{\beta}{1+2\beta}}} + \frac{1}{p^{1-c}} + \frac{1}{p^c} \right)^{1-q} + \frac{\sqrt{p}}{T^{\frac{2\beta}{1+2\beta}}} + \frac{1}{p^{\frac{3}{2} - 2c}} + \frac{1}{p^{2c - \frac{1}{2}}} \right), \\
\|\widehat{\mathbf{\Sigma}}^{\mathcal{S}} - \mathbf{\Sigma}\|_{\Sigma} &= O_P\left(\left(\frac{1}{\sqrt{T}} + \frac{1}{p^{1-c}} + \frac{1}{p^c} \right)^{1-q} + \frac{1}{T^{\frac{\beta}{1+2\beta}}} + \frac{\sqrt{p}}{T^{\frac{2\beta}{1+2\beta}}} + \frac{1}{p^{\frac{3}{2} - 2c}} + \frac{1}{p^{2c - \frac{1}{2}}} \right),
\end{aligned}
$$

where $\widehat{\mathbf{\Sigma}}^{\mathcal{D}}$ is the Double-POET estimator defined in Section S.2 of the online supplement. Specifically, when $q \neq 0$, Structured-POET achieves a faster convergence rate under the relative Frobenius norm. This is because utilizing all observations enhances the estimation accuracy of each block diagonal matrix. However, when $q = 0$, the convergence rates of both estimators are the same. This is because the estimation error of the correlations between continents dominates the benefit mentioned above. Importantly, we note that this does not mean that their estimation errors are exactly the same. In fact, based on our simulation study, we can conjecture that Structured-POET has smaller convergence rates than Double-POET for $q = 0$. That is, the relative ratio of the convergence rate of Structured-POET with respect to that of Double-POET may be less than 1. Unfortunately, due to the complex upper bound calculations used to handle high-dimensional matrices, we cannot theoretically show this statement for $q = 0$. We leave this for a future study. Similarly, under the spectral norm for the inverse matrix, the convergence rate of Structure-POET can be faster than that of Double-POET when $q \neq 0$.

S.4 Proof of Theorem 3.1

We first provide useful lemmas below. Let $\{\delta_i, v_i\}_{i=1}^{p}$ be the eigenvalues and their corresponding eigenvectors of $\mathbf{\Sigma}$ in decreasing order. Let $\{\bar{\delta}_i, \bar{v}_i\}_{i=1}^{k}$ and $\{\widetilde{\delta}_i, \widetilde{v}_i\}_{i=1}^{k}$ be the leading eigenvalues and eigenvectors of $\mathbf{B}\mathbf{B}'$ and $\widetilde{\mathbf{\Sigma}}_g$, respectively, where $\widetilde{\mathbf{\Sigma}}_g = (\widetilde{\mathbf{\Sigma}}_g^{\mathcal{D}} + \widehat{\mathbf{D}}^{\frac{1}{2}}\widehat{\mathbf{\Theta}}\widehat{\mathbf{D}}^{\frac{1}{2}})$. Define $\mathbf{\Sigma}_E = \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Sigma}_u$ and let $\mathbf{\Sigma}_E^l = \mathbf{\Lambda}^l\mathbf{\Lambda}^{l\prime} + \mathbf{\Sigma}_u^l$ be the lth diagonal block of $\mathbf{\Sigma}_E$. For each country l, let $\{\kappa_i^l, \eta_i^l\}_{i=1}^{p_l}$ be the eigenvalues and eigenvectors of $\mathbf{\Sigma}_E^l$ in decreasing order, and $\{\bar{\kappa}_i^l, \bar{\eta}_i^l\}_{i=1}^{r_l}$ for $\mathbf{\Lambda}^l\mathbf{\Lambda}^{l\prime}$.

By Weyl's theorem, we have the following lemma under the pervasive conditions.

Lemma S.1. *Under Assumption 3.1(i), we have*

$$|\delta_i - \bar{\delta}_i| \le \|\mathbf{\Sigma}_E\| \text{ for } i \le k, \qquad |\delta_i| \le \|\mathbf{\Sigma}_E\| \text{ for } i > k,$$

and, for $i \le k$, $\bar{\delta}_i/p^{a_1}$ is strictly bigger than zero for all p. In addition, for each national group l, we have

$$|\kappa_i^l - \bar{\kappa}_i^l| \le \|\mathbf{\Sigma}_u^l\| \text{ for } i \le r_l, \qquad |\kappa_i^l| \le \|\mathbf{\Sigma}_u^l\| \text{ for } i > r_l,$$

and, for $i \le r_l$, $\bar{\kappa}_i^l/p_l^{a_2}$ is strictly bigger than zero for all p_l.

The following lemma presents the individual convergence rate of leading eigenvectors using Lemma S.1 and the l_∞ norm perturbation bound theorem of Fan, Wang, and Zhong (2018).

Lemma S.2. *Under Assumption 3.1(i), we have the following results.*

(i) We have

$$\max_{i \le k} \|\bar{v}_i - v_i\|_\infty \le C \frac{\|\mathbf{\Sigma}_E\|_\infty}{p^{3(a_1 - \frac{1}{2})}}.$$

(ii) For each national group l, we have

$$\max_{i \le r_l} \|\bar{\eta}_i^l - \eta_i^l\|_\infty \le C \frac{\|\boldsymbol{\Sigma}_u^l\|_\infty}{p_l^{3(a_2 - \frac{1}{2})}}.$$

Proof. (i) Let $\mathbf{B} = (\tilde{b}_1, \dots, \tilde{b}_k)$. Then, for $i \le k$, $\bar{\delta}_i = \|\tilde{b}_i\|^2 \asymp p^{a_1}$ from Lemma S.1 and $\bar{v}_i = \tilde{b}_i / \|\tilde{b}_i\|$. Hence, $\|\bar{v}_i\|_\infty \le \|\mathbf{B}\|_{\max} / \|\tilde{b}_i\| \le C/\sqrt{p^{a_1}}$. In addition, for $\widetilde{\mathbf{V}} = (\bar{v}_1, \dots, \bar{v}_k)$, the coherence $\mu(\widetilde{\mathbf{V}}) = p \max_i \sum_{j=1}^k \widetilde{V}_{ij}^2 / k \le C p^{1-a_1}$, where $\widetilde{V}_{ij}$ is the (i, j) entry of $\widetilde{\mathbf{V}}$. Thus, by Theorem 1 of Fan, Wang, and Zhong (2018), we have

$$\max_{i \le k} \|\bar{v}_i - v_i\|_\infty \le C p^{2(1-a_1)} \frac{\|\boldsymbol{\Sigma}_E\|_\infty}{\bar{\gamma}\sqrt{p}},$$

where the eigengap $\bar{\gamma} = \min\{\bar{\delta}_i - \bar{\delta}_{i+1} : 1 \le i \le k\}$ and $\delta_{k+1} = 0$. By the similar argument, we can show the result (ii). □

Lemma S.3. *Let $\mathbf{R}_0 = (\mathbf{R}_{0,sq})_{S \times S}$, where $\mathbf{R}_{0,sq}$ is the (s, q)th off-diagonal partitioned block matrix for $s, q \in \{1, \dots, S\}$. Under Assumption 3.1, for $s \ne q$, we have*

$$\|\widehat{\boldsymbol{\Theta}}_{sq} - \mathbf{R}_{0,sq}\|_{\max} = O_P\left(p^{\frac{5}{2}(1-a_1)} \big(\sqrt{\log p / T^\alpha} + 1/T^{(1-\alpha)\beta}\big)\right).$$

Proof. For $s \ne q$, let the singular value decomposition be $\mathbf{R}_{0,sq} = \mathbf{U}\boldsymbol{\Xi}\mathbf{W}' = \sum_{i=1}^{k_{sq}^*} \xi_i u_i w_i'$, where k_{sq}^* is the rank of $\mathbf{R}_{0,sq}$, the singular values are $\xi_1 \ge \xi_2 \ge \cdots \ge \xi_{k_{sq}^*} > 0$, and the matrices $\mathbf{U} = (u_1, \dots, u_{k_{sq}^*})$, $\mathbf{W} = (w_1, \dots, w_{k_{sq}^*})$ consist of the singular vectors. By Lipschitz condition, $\|\mathbf{R}_h - \mathbf{R}_0\|_{\max} = O(1/T^{(1-\alpha)\beta})$, and Assumption 3.1 (iii), we have

$$\|\widehat{\mathbf{R}}_h - \mathbf{R}_0\|_{\max} = O_P\left(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\right). \tag{S.7}$$

Note that $\mathbf{R}_{0,sq}$ is k_{sq}^*-rank matrix for $s \ne q \in \{1, \dots, S\}$. By Weyl's inequality, we have

$$|\widehat{\xi}_i - \xi_i| \le \|\widehat{\mathbf{R}}_{h,sq} - \mathbf{R}_{0,sq}\|_F = O_P\left(p\big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\big)\right). \tag{S.8}$$

By Theorem 1 of Fan, Wang, and Zhong (2018), we have

$$\|\widehat{u}_i - u_i\|_\infty \le Cp^{2(1-a_1)}\frac{\|\widehat{\mathbf{R}}_{h,sq} - \mathbf{R}_{0,sq}\|_\infty}{p^{a_1}\sqrt{p}} \le Cp^{2(1-a_1)}\frac{\|\widehat{\mathbf{R}}_{h,sq} - \mathbf{R}_{0,sq}\|_{\max}}{p^{a_1-1}\sqrt{p}}$$
$$= O_P\left(p^{\frac{5}{2}-3a_1}\left(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\right)\right). \tag{S.9}$$

Similarly, we can obtain the same rate for $\|\widehat{w}_i - w_i\|_\infty$. Note that $\|\mathbf{U}\mathbf{\Xi}^{\frac{1}{2}}\|_{\max} = O_P(1)$. By (S.8) and (S.9), we have

$$\|\widehat{\mathbf{U}}\widehat{\mathbf{\Xi}}^{\frac{1}{2}} - \mathbf{U}\mathbf{\Xi}^{\frac{1}{2}}\|_{\max} \le \|\widehat{\mathbf{U}}(\widehat{\mathbf{\Xi}}^{\frac{1}{2}} - \mathbf{\Xi}^{\frac{1}{2}})\|_{\max} + \|(\widehat{\mathbf{U}} - \mathbf{U})\mathbf{\Xi}^{\frac{1}{2}}\|_{\max}$$
$$= O_P\left(p^{\frac{5}{2}(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta})\right) = o(1).$$

Then, we have $\|\widehat{\mathbf{U}}\|_{\max} = O_P(1/\sqrt{p^{a_1}})$. Similarly, we can obtain $\|\widehat{\mathbf{W}}\|_{\max} = O_P(1/\sqrt{p^{a_1}})$. Therefore, we have

$$\|\widehat{\mathbf{\Theta}}_{sq} - \mathbf{R}_{0,sq}\|_{\max} \le \|\widehat{\mathbf{U}}(\widehat{\mathbf{\Xi}} - \mathbf{\Xi})\widehat{\mathbf{W}}'\|_{\max} + \|(\widehat{\mathbf{U}} - \mathbf{U})\mathbf{\Xi}(\widehat{\mathbf{W}} - \mathbf{W})'\|_{\max} + 2\|(\widehat{\mathbf{U}} - \mathbf{U})\mathbf{\Xi}\mathbf{W}'\|_{\max}$$
$$= O_P(p^{-a_1}\|\widehat{\mathbf{\Xi}} - \mathbf{\Xi}\|_{\max} + \sqrt{p^{a_1}}\|\widehat{\mathbf{U}} - \mathbf{U}\|_{\max}) = O_P\big(p^{\frac{5}{2}(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta})\big).$$

□

Proof of Theorem 3.1.

Consider (3.2). Similar to the proofs of (S.24), we can show $\|(\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1} - \mathbf{\Sigma}_E^{-1}\| = O_P(m_p\omega_T^{1-q} + p^{\frac{c}{2}(1-a_2)}\omega_T)$. Let $\widehat{\mathbf{H}} = \widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}}\widetilde{\mathbf{V}}_g'(\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1}\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}}$ and $\widetilde{\mathbf{H}} = \widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'\mathbf{\Sigma}_E^{-1}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}$. Using the Sherman-Morrison-Woodbury formula, we have

$$\|(\widehat{\mathbf{\Sigma}}^{\mathcal{S}})^{-1} - \mathbf{\Sigma}^{-1}\| \le \|(\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1} - \mathbf{\Sigma}_E^{-1}\| + \Delta,$$

where $\Delta = \|(\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1}\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}}(\mathbf{I}_k + \widehat{\mathbf{H}})^{-1}\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}}\widetilde{\mathbf{V}}_g'(\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1} - \mathbf{\Sigma}_E^{-1}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}(\mathbf{I}_k + \widetilde{\mathbf{H}})^{-1}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'\mathbf{\Sigma}_E^{-1}\|$. Then, the right hand side can be bounded by following terms:

$$L_1 = \|((\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1} - \mathbf{\Sigma}_E^{-1})\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}(\mathbf{I}_k + \widetilde{\mathbf{H}})^{-1}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'\mathbf{\Sigma}_E^{-1}\|,$$
$$L_2 = \|\mathbf{\Sigma}_E^{-1}(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}} - \widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}})(\mathbf{I}_k + \widetilde{\mathbf{H}})^{-1}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'\mathbf{\Sigma}_E^{-1}\|,$$
$$L_3 = \|\mathbf{\Sigma}_E^{-1}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}((\mathbf{I}_k + \widehat{\mathbf{H}})^{-1} - (\mathbf{I}_k + \widetilde{\mathbf{H}})^{-1})\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'\mathbf{\Sigma}_E^{-1}\|.$$

By Weyl's inequality, we have $\lambda_{\min}(\mathbf{\Sigma}_E) > c$ since $\lambda_{\min}(\mathbf{\Sigma}_u) > c$ and $\lambda_{\min}(\mathbf{\Lambda}\mathbf{\Lambda}') = 0$. Hence, $\|\mathbf{\Sigma}_E^{-1}\| = O_P(1)$. Note that $\|\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\| = O_P(p^{\frac{a_1}{2}})$. By Lemma 7.1, we have $\|\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}} - \widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\|_{\max} = O_P(p^{5(1-a_1)}\big(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}\big) + 1/p^{5a_1-4-c})$. Then, we have

$$\begin{aligned}\|\widehat{\mathbf{H}} - \widetilde{\mathbf{H}}\| &\le \|(\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}}\widetilde{\mathbf{V}}_g' - \widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}')(\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1}(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}} - \widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}})\| \\ &\quad + \|(\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}}\widetilde{\mathbf{V}}_g' - \widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}')(\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1}\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\| + \|\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'((\widehat{\mathbf{\Sigma}}_E^{\mathcal{S}})^{-1} - \mathbf{\Sigma}_E^{-1})\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\| \\ &= O_P\left(p^{a_1}m_p\omega_T^{1-q} + p^{\frac{11}{2}-\frac{9}{2}a_1}\left(\sqrt{\frac{\log p}{T^{\alpha}}} + \frac{1}{T^{(1-\alpha)\beta}}\right) + \frac{1}{p^{\frac{9}{2}(a_1-1)-c}}\right).\end{aligned}$$

Since $\lambda_{\min}(\mathbf{I}_k + \widetilde{\mathbf{H}}) \ge \lambda_{\min}(\widetilde{\mathbf{H}}) \ge \lambda_{\min}(\mathbf{\Sigma}_E^{-1})\lambda_{\min}^2(\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}) \ge Cp^{a_1}$, we have $\|(\mathbf{I}_k + \widetilde{\mathbf{H}})^{-1}\| = O_P(1/p^{a_1})$. Then, $L_1 = O_P(m_p\omega_T^{1-q})$. In addition, $L_2 = O_P(p^{\frac{-a_1}{2}}\|\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g^{\frac{1}{2}} - \widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\|) = O_P(p^{\frac{11}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{11}{2}a_1-\frac{9}{2}-c})$ and $L_3 = O_P(p^{a_1}\|(\mathbf{I}_k + \widehat{\mathbf{H}})^{-1} - (\mathbf{I}_k + \widetilde{\mathbf{H}})^{-1}\|) = O_P(p^{-a_1}\|\widehat{\mathbf{H}} - \widetilde{\mathbf{H}}\|) = O_P(m_p\omega_T^{1-q} + p^{\frac{11}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{11}{2}a_1-\frac{9}{2}-c})$. Thus, we have

$$\Delta = O_P(m_p\omega_T^{1-q} + p^{\frac{11}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{11}{2}a_1-\frac{9}{2}-c}). \tag{S.10}$$

Therefore, we have

$$\|(\widehat{\mathbf{\Sigma}}^{\mathcal{S}})^{-1} - \mathbf{\Sigma}^{-1}\| = O_P\left(m_p\omega_T^{1-q} + p^{\frac{c}{2}(1-a_2)}\omega_T + p^{\frac{11}{2}(1-a_1)}\left(\sqrt{\frac{\log p}{T^{\alpha}}} + \frac{1}{T^{(1-\alpha)\beta}}\right) + \frac{1}{p^{\frac{11}{2}a_1-\frac{9}{2}-c}}\right). \tag{S.11}$$

Consider (3.3). We derive the rate of convergence for $\|\widehat{\mathbf{\Sigma}}^{\mathcal{S}} - \mathbf{\Sigma}\|_{\Sigma}$. The SVD decomposition of $\mathbf{\Sigma}$ is

$$\mathbf{\Sigma} = (\mathbf{V}_{p\times k}\ \ \mathbf{\Phi}_{p\times r}\ \ \mathbf{\Omega}_{p\times(p-k-r)}) \begin{pmatrix} \mathbf{\Gamma}_{k\times k} & & \\ & \mathbf{\Psi}_{r\times r} & \\ & & \mathbf{\Theta}_{(p-k-r)\times(p-k-r)} \end{pmatrix} \begin{pmatrix} \mathbf{V}' \\ \mathbf{\Phi}' \\ \mathbf{\Omega}' \end{pmatrix}.$$

Note that $\mathbf{\Omega}$ is used to denote the precision matrix in Section 2. Moreover, since all the eigenvalues of $\mathbf{\Sigma}$ are strictly bigger than 0, for any matrix $\mathbf{A}$, we have $\|\mathbf{A}\|_\Sigma^2 = O_P(p^{-1})\|\mathbf{A}\|_F^2$. Then, we have

$$\begin{aligned} \|\widehat{\mathbf{\Sigma}}^{\mathcal{S}} - \mathbf{\Sigma}\|_\Sigma \leq p^{-1/2}\Big(&\|\mathbf{\Sigma}^{-1/2}(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}')\mathbf{\Sigma}^{-1/2}\|_F \\ &+ \|\mathbf{\Sigma}^{-1/2}(\widetilde{\mathbf{\Sigma}}_l^{\mathcal{D}} - \mathbf{\Lambda}\mathbf{\Lambda}')\mathbf{\Sigma}^{-1/2}\|_F + \|\mathbf{\Sigma}^{-1/2}(\widetilde{\mathbf{\Sigma}}_u^{\mathcal{D}} - \mathbf{\Sigma}_u)\mathbf{\Sigma}^{-1/2}\|_F\Big) \\ =:&\ \Delta_G + \Delta_L + \Delta_S. \end{aligned}$$

By using the fact that S is fixed and proofs of (3.5) in Choi and Kim (2023), we can obtain

$$\begin{aligned} \Delta_L = O_P\Big(& p^{\frac{5}{2}(1-a_1)+c(1-a_2)}\sqrt{\frac{\log p}{T}} + \frac{1}{p^{\frac{5}{2}a_1-\frac{3}{2}-2c+ca_2}} + \frac{m_p}{p^{ca_2}} \\ &+ p^{\frac{11}{2}-5a_1+5c(1-a_2)}\frac{\log p}{T} + \frac{1}{p^{5a_1-\frac{7}{2}-c(7-5a_2)}} + \frac{m_p^2}{p^{5ca_2-3c-\frac{1}{2}}}\Big) \end{aligned} \tag{S.12}$$

and

$$\Delta_S = O_P(p^{-1/2}\|\widetilde{\mathbf{\Sigma}}_u^{\mathcal{D}} - \mathbf{\Sigma}_u\|_F) = O_P(\|\widetilde{\mathbf{\Sigma}}_u^{\mathcal{D}} - \mathbf{\Sigma}_u\|_2) = O_P(m_p\omega_T^{1-q}). \tag{S.13}$$

We have

$$\Delta_G = p^{-1/2}\left\| \begin{pmatrix} \mathbf{\Gamma}^{-\frac{1}{2}}\mathbf{V}' \\ \mathbf{\Psi}^{-\frac{1}{2}}\mathbf{\Phi}' \\ \mathbf{\Theta}^{-\frac{1}{2}}\mathbf{\Omega}' \end{pmatrix} (\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}') \left(\mathbf{V}\mathbf{\Gamma}^{-\frac{1}{2}}\ \ \mathbf{\Phi}\mathbf{\Psi}^{-\frac{1}{2}}\ \ \mathbf{\Omega}\mathbf{\Theta}^{-\frac{1}{2}}\right) \right\|_F$$

$$\le p^{-1/2}\big(\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}')\mathbf{V}\mathbf{\Gamma}^{-1/2}\|_F + \|\mathbf{\Psi}^{-1/2}\mathbf{\Phi}'(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}')\mathbf{\Phi}\mathbf{\Psi}^{-1/2}\|_F$$
$$+ \|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}')\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F + 2\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}')\mathbf{\Phi}\mathbf{\Psi}^{-1/2}\|_F$$
$$+ 2\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}')\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F + 2\|\mathbf{\Psi}^{-1/2}\mathbf{\Phi}'(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{B}\mathbf{B}')\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F\big)$$
$$=: \Delta_{G1} + \Delta_{G2} + \Delta_{G3} + 2\Delta_{G4} + 2\Delta_{G5} + 2\Delta_{G6}.$$

In order to find the convergence rate of relative Frobenius norm, we consider the above terms separately. Note that $\mathbf{\Gamma} = \mathrm{diag}(\delta_1, \ldots, \delta_k)$ and $\mathbf{V} = (v_1, \ldots, v_k)$. For Δ_{G1}, we have

$$\Delta_{G1} \le p^{-1/2}\left(\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g' - \mathbf{V}\mathbf{\Gamma}\mathbf{V}')\mathbf{V}\mathbf{\Gamma}^{-1/2}\|_F + \|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\mathbf{V}\mathbf{\Gamma}\mathbf{V}' - \mathbf{B}\mathbf{B}')\mathbf{V}\mathbf{\Gamma}^{-1/2}\|_F\right)$$
$$=: \Delta_{G1}^{(a)} + \Delta_{G1}^{(b)}.$$

We bound the two terms separately. We have

$$\Delta_{G1}^{(a)} \le p^{-1/2}\big(\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widetilde{\mathbf{V}}_g - \mathbf{I})\widetilde{\mathbf{\Gamma}}_g(\widetilde{\mathbf{V}}_g'\mathbf{V} - \mathbf{I})\mathbf{\Gamma}^{-1/2}\|_F + 2\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widetilde{\mathbf{V}}_g - \mathbf{I})\widetilde{\mathbf{\Gamma}}_g\mathbf{\Gamma}^{-1/2}\|_F$$
$$+ \|(\mathbf{\Gamma}^{-1/2}(\widetilde{\mathbf{\Gamma}}_g - \mathbf{\Gamma})\mathbf{\Gamma}^{-1/2}\|_F\big) =: I + II + III.$$

By Lemmas S.2 and 7.1, we obtain $\|\mathbf{V}'\widetilde{\mathbf{V}}_g - \mathbf{I}\|_F = \|\mathbf{V}'(\widetilde{\mathbf{V}}_g - \mathbf{V})\|_F \le \|\widetilde{\mathbf{V}}_g - \mathbf{V}\|_F = O_P(p^{\frac{11}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{11}{2}a_1 - \frac{9}{2} - c})$. Then, $II = O_P(p^{5-\frac{11}{2}a_1}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{11}{2}a_1 - 4 - c})$ and I is of smaller order. In addition, we have $III \le \|\mathbf{\Gamma}^{-1/2}(\widetilde{\mathbf{\Gamma}}_g - \mathbf{\Gamma})\mathbf{\Gamma}^{-1/2}\| = O_P(p^{\frac{7}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{7}{2}a_1 - \frac{5}{2} - c} + 1/p^{a_1 - ca_2})$ by Lemma 7.1. Thus, $\Delta_{G1}^{(a)} = O_P(p^{\frac{7}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{7}{2}a_1 - \frac{5}{2} - c} + 1/p^{a_1 - ca_2})$. Similarly, we have

$$\Delta_{G1}^{(b)} \le p^{-1/2}\big(\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widetilde{\mathbf{V}} - \mathbf{I})\widetilde{\mathbf{\Gamma}}(\widetilde{\mathbf{V}}'\mathbf{V} - \mathbf{I})\mathbf{\Gamma}^{-1/2}\|_F + 2\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widetilde{\mathbf{V}} - \mathbf{I})\widetilde{\mathbf{\Gamma}}\mathbf{\Gamma}^{-1/2}\|_F$$
$$+ \|(\mathbf{\Gamma}^{-1/2}(\widetilde{\mathbf{\Gamma}} - \mathbf{\Gamma})\mathbf{\Gamma}^{-1/2}\|_F\big) =: I' + II' + III'.$$

By $\sin\theta$ theorem, $\|\mathbf{V}'\widetilde{\mathbf{V}} - \mathbf{I}\| = \|\mathbf{V}'(\widetilde{\mathbf{V}} - \mathbf{V})\| \le \|\widetilde{\mathbf{V}} - \mathbf{V}\| = O(\|\mathbf{\Sigma}_E\|/p^{a_1})$. Then, we have

$II' = O(1/p^{a_1 - ca_2})$ and I' is of smaller order. By Lemma S.1, we have $III' = O(1/p^{a_1 - ca_2})$. Thus, $\Delta_{G1}^{(b)} = O(1/p^{a_1 - ca_2})$. Then, we obtain

$$\Delta_{G1} = O_P\left(p^{\frac{7}{2}(1-a_1)}\Big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{\frac{7}{2}a_1 - \frac{5}{2} - c}} + \frac{1}{p^{a_1 - ca_2}}\right). \tag{S.14}$$

For Δ_{G3}, we have

$$\Delta_{G3} \le p^{-1/2}\|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'\widetilde{\mathbf{V}}_g\widetilde{\mathbf{\Gamma}}_g\widetilde{\mathbf{V}}_g'\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F + p^{-1/2}\|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}\widetilde{\mathbf{V}}'\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F =: \Delta_{G3}^{(a)} + \Delta_{G3}^{(b)}.$$

By Lemmas S.2 and 7.1, we have

$$\begin{aligned}\|\mathbf{\Omega}'\widetilde{\mathbf{V}}_g\|_F =& \|\mathbf{\Omega}'(\widetilde{\mathbf{V}}_g - \mathbf{V})\|_F = O(\sqrt{p}\|\widetilde{\mathbf{V}}_g - \mathbf{V}\|_{\max}) \\ &= O_P(p^{\frac{11}{2}(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{11}{2}a_1 - \frac{9}{2} - c}).\end{aligned}$$

Since $\|\widetilde{\mathbf{\Gamma}}_g\| = O_P(p^{a_1})$, we have

$$\Delta_{G3}^{(a)} \le p^{-1/2}\|\mathbf{\Theta}^{-1}\|\|\mathbf{\Omega}'\widetilde{\mathbf{V}}_g\|_F^2\|\widetilde{\mathbf{\Gamma}}_g\| = O_P(p^{\frac{21}{2} - 10a_1}(\log p/T^\alpha + 1/T^{2(1-\alpha)\beta}) + 1/p^{10a_1 - \frac{17}{2} - 2c}).$$

Similarly, $\Delta_{G3}^{(b)} = O_P(1/p^{5a_1 - \frac{7}{2} - 2c})$ because $\|\mathbf{\Omega}'\widetilde{\mathbf{V}}\|_F = O(\sqrt{p}\|\widetilde{\mathbf{V}} - \mathbf{V}\|_{\max}) = O_P(1/p^{3a_1 - 2 - c})$ by Lemma S.2. Then, we obtain

$$\Delta_{G3} = O_P\left(p^{\frac{21}{2} - 10a_1}\Big(\frac{\log p}{T^\alpha} + \frac{1}{T^{2(1-\alpha)\beta}}\Big) + \frac{1}{p^{10a_1 - \frac{17}{2} - 2c}}\right).$$

Similarly, we can show that the terms Δ_{G2}, Δ_{G4}, Δ_{G5} and Δ_{G6} are dominated by Δ_{G1} and Δ_{G3}. Therefore, we have

$$\begin{aligned}\Delta_G = O_P\Big(& p^{\frac{7}{2}(1-a_1)}\Big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{\frac{7}{2}a_1 - \frac{5}{2} - c}} + \frac{1}{p^{a_1 - ca_2}} \\ &+ p^{\frac{21}{2} - 10a_1}\Big(\frac{\log p}{T^\alpha} + \frac{1}{T^{2(1-\alpha)\beta}}\Big) + \frac{1}{p^{10a_1 - \frac{17}{2} - 2c}}\Big).\end{aligned} \tag{S.15}$$

Combining the terms Δ_G, Δ_L and Δ_S together, we complete the proof of (3.3). $\square$

S.5 Proof of Theorem S.1

We provide useful technical lemmas below.

Lemma S.4. *Under Assumptions 3.1 and S.1, for $i \leq k$, we have*

$$|\widehat{\delta}_i/\delta_i - 1| = O_P\left(p^{1-a_1}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta})\right),$$
$$\|\widehat{v}_i - v_i\|_\infty = O_P\left(\frac{1}{p^{3(a_1-\frac{1}{2})-c}} + p^{\frac{5}{2}-3a_1}\left(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\right)\right).$$

Proof. By Lipschitz condition and Assumption S.1, we have

$$\|d^{-1}\widehat{\mathbf{\Sigma}}_h - \mathbf{\Sigma}\|_{\max} = O_P(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}). \tag{S.16}$$

Then, we can obtain the first statement by Weyl's theorem. We have

$$d^{-1}\widehat{\mathbf{\Sigma}}_h = \mathbf{B}\mathbf{B}' + \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Sigma}_u + (d^{-1}\widehat{\mathbf{\Sigma}}_h - \mathbf{\Sigma}) = \mathbf{B}\mathbf{B}' + \mathbf{\Sigma}_E + (d^{-1}\widehat{\mathbf{\Sigma}}_h - \mathbf{\Sigma}).$$

We can treat $\mathbf{B}\mathbf{B}'$ as a low rank matrix and the remaining terms as a perturbation matrix. Note that $\|\mathbf{\Sigma}_E\|_\infty = O(p^c)$. By Theorem 1 of Fan, Wang, and Zhong (2018), Lemma S.2, Assumption 3.1 and (S.16), we have

$$\begin{aligned}\|\widehat{v}_i - v_i\|_\infty &\leq Cp^{2(1-a_1)}\frac{\|\mathbf{\Sigma}_E\|_\infty}{p^{a_1}\sqrt{p}} + Cp^{2(1-a_1)}\frac{\|d^{-1}\widehat{\mathbf{\Sigma}}_h - \mathbf{\Sigma}\|_{\max}}{p^{a_1-1}\sqrt{p}} \\ &= O_P\left(\frac{1}{p^{3(a_1-\frac{1}{2})-c}} + p^{\frac{5}{2}-3a_1}\left(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\right)\right).\end{aligned}$$

$\square$

Lemma S.5. *Under Assumptions 3.1 and S.1, for $i \le r_l$, we have*

$$|\widehat{\kappa}_i^l/\kappa_i^l - 1| = O_P\left(p^{\frac{5}{2}(1-a_1)+c(1-a_2)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5a_1}{2}-\frac{3}{2}-2c+ca_2}\right),$$

$$\|\widehat{\eta}_i^l - \eta_i^l\|_\infty = O_P\left(p^{\frac{5}{2}(1-a_1)+c(\frac{5}{2}-3a_2)}\left(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\right) + \frac{1}{p^{\frac{5a_1}{2}-\frac{3}{2}+c(3a_2-\frac{7}{2})}} + \frac{m_p}{p^{3c(a_2-\frac{1}{2})}}\right).$$

Proof. We have

$$\|\mathbf{\Sigma}_E\| \le \|\mathbf{\Lambda}\mathbf{\Lambda}'\| + \|\mathbf{\Sigma}_u\| \le \|\mathbf{\Lambda}\mathbf{\Lambda}'\| + O(m_p) = O(p^{ca_2}).$$

Let $\mathbf{B}\mathbf{B}' = \widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}\widetilde{\mathbf{V}}'$, where $\widetilde{\mathbf{\Gamma}} = \mathrm{diag}(\bar{\delta}_1, \dots, \bar{\delta}_k)$ and their corresponding leading k eigenvectors $\widetilde{\mathbf{V}} = (\bar{v}_1, \dots, \bar{v}_k)$. Also, we let $\mathbf{\Gamma} = \mathrm{diag}(\delta_1, \dots, \delta_k)$ and the corresponding eigenvectors $\mathbf{V} = (v_1, \dots, v_k)$ of covariance matrix $\mathbf{\Sigma}$. Note that $\|\mathbf{B}\|_{\max} = \|\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{1/2}\|_{\max} = O(1)$. By Lemmas S.1-S.2, we have

$$\begin{aligned}\|\mathbf{V}\mathbf{\Gamma}^{\frac{1}{2}} - \widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\|_{\max} &\le \|\mathbf{B}\widetilde{\mathbf{\Gamma}}^{-\frac{1}{2}}(\mathbf{\Gamma}^{\frac{1}{2}} - \widetilde{\mathbf{\Gamma}}^{\frac{1}{2}})\|_{\max} + \|(\mathbf{V} - \widetilde{\mathbf{V}})\mathbf{\Gamma}^{\frac{1}{2}}\|_{\max} \\ &\le C\frac{\|\mathbf{\Sigma}_E\|}{p^{a_1}} + C\frac{\|\mathbf{\Sigma}_E\|_\infty}{\sqrt{p^{5a_1-3}}} = o(1).\end{aligned}$$

Hence, we have $\|\mathbf{V}\mathbf{\Gamma}^{\frac{1}{2}}\|_{\max} = O(1)$ and $\|\mathbf{V}\|_{\max} = O(1/\sqrt{p^{a_1}})$. By this fact and the results from Lemmas S.1-S.4, we have

$$\begin{aligned}\|\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}\widetilde{\mathbf{V}}' - \mathbf{V}\mathbf{\Gamma}\mathbf{V}'\|_{\max} &\le \|\widetilde{\mathbf{V}}(\widetilde{\mathbf{\Gamma}} - \mathbf{\Gamma})\widetilde{\mathbf{V}}'\|_{\max} + \|(\widetilde{\mathbf{V}} - \mathbf{V})\mathbf{\Gamma}(\widetilde{\mathbf{V}} - \mathbf{V})'\|_{\max} + 2\|\mathbf{V}\mathbf{\Gamma}(\widetilde{\mathbf{V}} - \mathbf{V})'\|_{\max} \\ &= O(p^{-a_1}\|\widetilde{\mathbf{\Gamma}} - \mathbf{\Gamma}\|_{\max} + \sqrt{p^{a_1}}\|\widetilde{\mathbf{V}} - \mathbf{V}\|_{\max}) = O(1/p^{\frac{5a_1}{2}-\frac{3}{2}-c}), \\ \|\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{V}\mathbf{\Gamma}\mathbf{V}'\|_{\max} &\le \|\widehat{\mathbf{V}}(\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma})\widehat{\mathbf{V}}'\|_{\max} + \|(\widehat{\mathbf{V}} - \mathbf{V})\mathbf{\Gamma}(\widehat{\mathbf{V}} - \mathbf{V})'\|_{\max} + 2\|\mathbf{V}\mathbf{\Gamma}(\widehat{\mathbf{V}} - \mathbf{V})'\|_{\max} \\ &= O_P(p^{-a_1}\|\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}\|_{\max} + \sqrt{p^{a_1}}\|\widehat{\mathbf{V}} - \mathbf{V}\|_{\max}) \\ &= O_P(p^{\frac{5}{2}(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5a_1}{2}-\frac{3}{2}-c}).\end{aligned}$$

Thus, we have

$$\|\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}'\|_{\max} = O_P(p^{\frac{5}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5a_1}{2}-\frac{3}{2}-c}). \tag{S.17}$$

Then, we have

$$\begin{aligned}\|\widehat{\mathbf{\Sigma}}_E - \mathbf{\Sigma}_E\|_{\max} &\leq \|\widehat{\mathbf{\Sigma}}_h - \mathbf{\Sigma}\|_{\max} + \|\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}'\|_{\max} \\ &= O_P(p^{\frac{5}{2}(1-a_1)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5a_1}{2}-\frac{3}{2}-c}). \end{aligned}\tag{S.18}$$

Therefore, the first statement is followed by (S.18) and the Weyl's theorem.

We decompose the sample covariance matrix $\widehat{\mathbf{\Sigma}}_E^l$ for each group l as follows:

$$\widehat{\mathbf{\Sigma}}_E^l = \mathbf{\Lambda}^l\mathbf{\Lambda}^{l\prime} + \mathbf{\Sigma}_u^l + (\widehat{\mathbf{\Sigma}}_E^l - \mathbf{\Sigma}_E^l).$$

Then, by Theorem 1 of Fan, Wang, and Zhong (2018), Lemma S.2 and (S.18), we have

$$\begin{aligned}\|\widehat{\eta}_i^l - \eta_i^l\|_{\infty} &\leq C p_l^{2(1-a_2)} \frac{\|\mathbf{\Sigma}_u^l + (\widehat{\mathbf{\Sigma}}_E^l - \mathbf{\Sigma}_E^l)\|_{\infty}}{p_l^{a_2}\sqrt{p_l}} \\ &\leq C p_l^{2(1-a_2)} \frac{\|\mathbf{\Sigma}_u^l\|_{\infty}}{p_l^{a_2}\sqrt{p_l}} + C p_l^{2(1-a_2)} \frac{\|\widehat{\mathbf{\Sigma}}_E^l - \mathbf{\Sigma}_E^l\|_{\max}}{p_l^{a_2-1}\sqrt{p_l}} \\ &= O_P\left(p^{\frac{5}{2}(1-a_1)+c(\frac{5}{2}-3a_2)}\Big(\sqrt{\frac{\log p}{T^{\alpha}}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{\frac{5a_1}{2}-\frac{3}{2}+c(3a_2-\frac{7}{2})}} + \frac{m_p}{p^{3c(a_2-\frac{1}{2})}}\right).\end{aligned}$$

□

Proof of Theorem S.1. We first consider (S.4). We have

$$\|\widehat{\mathbf{\Phi}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{\Phi}}' - \mathbf{\Lambda}\mathbf{\Lambda}'\|_{\max} = \max_l \|\widehat{\mathbf{\Phi}}^l\widehat{\mathbf{\Psi}}^l\widehat{\mathbf{\Phi}}^{j\prime} - \mathbf{\Lambda}^l\mathbf{\Lambda}^{l\prime}\|_{\max}.$$

For each group l, let $\mathbf{\Lambda}^l\mathbf{\Lambda}^{l\prime} = \widetilde{\mathbf{\Phi}}^l\widetilde{\mathbf{\Psi}}^l\widetilde{\mathbf{\Phi}}^{l\prime}$, where $\widetilde{\mathbf{\Psi}}^l = \mathrm{diag}(\bar{\kappa}_1^l, \dots, \bar{\kappa}_{r_l}^l)$ and the corresponding eigenvectors $\widetilde{\mathbf{\Phi}}^l = (\bar{\eta}_1, \dots, \bar{\eta}_{r_l})$. In addition, let $\mathbf{\Psi}^l = \mathrm{diag}(\kappa_1^l, \dots, \kappa_{r_l}^l)$ and $\mathbf{\Phi}^l = (\eta_1, \dots, \eta_{r_l})$

to be the leading eigenvalues and the corresponding eigenvectors of $\mathbf{\Sigma}_E^l$, respectively. Then, we have

$$\begin{aligned}\|\mathbf{\Phi}^l\mathbf{\Psi}^{l\frac{1}{2}} - \widetilde{\mathbf{\Phi}}^l\widetilde{\mathbf{\Psi}}^{l\frac{1}{2}}\|_{\max} &\le \|\mathbf{\Lambda}^l\widetilde{\mathbf{\Psi}}^{l-\frac{1}{2}}(\mathbf{\Psi}^{l\frac{1}{2}} - \widetilde{\mathbf{\Psi}}^{l\frac{1}{2}})\|_{\max} + \|(\mathbf{\Phi}^l - \widetilde{\mathbf{\Phi}}^l)\mathbf{\Psi}^{l\frac{1}{2}}\|_{\max} \\ &\le \frac{\|\mathbf{\Sigma}_u^l\|}{p_l^{a_2}} + \frac{\|\mathbf{\Sigma}_u^l\|_\infty}{\sqrt{p_l^{5a_2-3}}} = o(1). \end{aligned} \tag{S.19}$$

Since $\|\mathbf{\Lambda}^l\|_{\max} = \|\widetilde{\mathbf{\Phi}}^l\widetilde{\mathbf{\Psi}}^{l\frac{1}{2}}\|_{\max} = O(1)$, $\|\mathbf{\Phi}^l\mathbf{\Psi}^{l\frac{1}{2}}\|_{\max} = O(1)$ and $\|\mathbf{\Phi}^l\|_{\max} = O(1/\sqrt{p_l^{a_2}})$. Using this fact and results from Lemmas S.1, S.2 and S.5, we can show

$$\begin{aligned}&\|\widetilde{\mathbf{\Phi}}^l\widetilde{\mathbf{\Psi}}^l\widetilde{\mathbf{\Phi}}^{l\prime} - \mathbf{\Phi}^l\mathbf{\Psi}^l\mathbf{\Phi}^{l\prime}\|_{\max} \le O(p_l^{-a_2}\|\widetilde{\mathbf{\Psi}}^l - \mathbf{\Psi}^l\|_{\max} + \sqrt{p_l^{a_2}}\|\widetilde{\mathbf{\Phi}}^l - \mathbf{\Phi}^l\|_{\max}) = O(m_p/\sqrt{p^{c(5a_2-3)}}), \\ &\|\widehat{\mathbf{\Phi}}^l\widehat{\mathbf{\Psi}}^l\widehat{\mathbf{\Phi}}^{l\prime} - \mathbf{\Phi}^l\mathbf{\Psi}^l\mathbf{\Phi}^{l\prime}\|_{\max} \le O_P(p_l^{-a_2}\|\widehat{\mathbf{\Psi}}^l - \mathbf{\Psi}^l\|_{\max} + \sqrt{p_l^{a_2}}\|\widehat{\mathbf{\Phi}}^l - \mathbf{\Phi}^l\|_{\max}) \\ &\qquad = O_P\left(p^{\frac{5}{2}(1-a_1)+\frac{5}{2}c(1-a_2)}\Big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{\frac{5}{2}a_1-\frac{3}{2}+c(\frac{5}{2}a_2-\frac{7}{2})}} + \frac{m_p}{\sqrt{p^{c(5a_2-3)}}}\right).\end{aligned}$$

By using these rates, we obtain

$$\begin{aligned}&\|\widehat{\mathbf{\Phi}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{\Phi}}' - \mathbf{\Lambda}\mathbf{\Lambda}'\|_{\max} \\ &\quad = O_P\left(p^{\frac{5}{2}(1-a_1)+\frac{5}{2}c(1-a_2)}\Big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{\frac{5}{2}a_1-\frac{3}{2}+c(\frac{5}{2}a_2-\frac{7}{2})}} + \frac{m_p}{\sqrt{p^{c(5a_2-3)}}}\right).\end{aligned} \tag{S.20}$$

By (S.16), (S.17) and (S.20), we then have

$$\begin{aligned}\|\widehat{\mathbf{\Sigma}}_u - \mathbf{\Sigma}_u\|_{\max} &\le \|d^{-1}\widehat{\mathbf{\Sigma}}_h - \mathbf{\Sigma}\|_{\max} + \|\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}'\|_{\max} + \|\widehat{\mathbf{\Phi}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{\Phi}}' - \mathbf{\Lambda}\mathbf{\Lambda}'\|_{\max} \\ &= O_P\left(p^{\frac{5}{2}(1-a_1)+\frac{5}{2}c(1-a_2)}\Big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{\frac{5}{2}a_1-\frac{3}{2}+c(\frac{5}{2}a_2-\frac{7}{2})}} + \frac{m_p}{\sqrt{p^{c(5a_2-3)}}}\right).\end{aligned} \tag{S.21}$$

By definition, $\|\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}} - \widehat{\boldsymbol{\Sigma}}_u\|_{\max} = \max_{ij} |s_{ij}(\widehat{\sigma}_{ij}) - \widehat{\sigma}_{ij}| \le \max_{ij} \tau_{ij} = O_P(\tau)$. Then, we have

$$\|\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}} - \boldsymbol{\Sigma}_u\|_{\max} = O_P(\tau + \omega_{T^\alpha}) = O_P(\omega_{T^\alpha}), \tag{S.22}$$

when τ is chosen as the same order of $\omega_{T^\alpha} = p^{\frac{5}{2}(1-a_1)+\frac{5}{2}c(1-a_2)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5}{2}a_1 - \frac{3}{2} + c(\frac{5}{2}a_2 - \frac{7}{2})} + m_p/\sqrt{p^{c(5a_2-3)}}$. Therefore, by the results of (S.17), (S.20) and (S.22), we have

$$\|\widehat{\boldsymbol{\Sigma}}^{\mathcal{D}} - \boldsymbol{\Sigma}\|_{\max} \le \|\widehat{\mathbf{V}}\widehat{\boldsymbol{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}'\|_{\max} + \|\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}'\|_{\max} + \|\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}} - \boldsymbol{\Sigma}_u\|_{\max} = O_P(\omega_{T^\alpha}).$$

Consider (S.5). Similar to the proofs of Theorem 2.1 in Fan, Liao, and Mincheva (2011), we can show $\|\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}} - \boldsymbol{\Sigma}_u\|_2 = O_P(m_p\omega_{T^\alpha}^{1-q})$. In addition, since $\lambda_{\min}(\boldsymbol{\Sigma}_u) > c_1$ and $m_p\omega_{T^\alpha}^{1-q} = o(1)$, the minimum eigenvalue of $\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}}$ is strictly bigger than 0 with probability approaching 1. Then, we have

$$\|(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_u^{-1}\|_2 \le \lambda_{\min}(\boldsymbol{\Sigma}_u)^{-1}\|\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}} - \boldsymbol{\Sigma}_u\|_2 \lambda_{\min}(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1} = O_P(m_p\omega_{T^\alpha}^{1-q}). \tag{S.23}$$

Define $\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}} = \widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' + \widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}}$. We first show that $\|(\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_E^{-1}\| = O_P(p^{\frac{c}{2}(1-a_2)}\omega_{T^\alpha} + m_p\omega_{T^\alpha}^{1-q})$. Let $\widehat{\mathbf{J}} = \widehat{\boldsymbol{\Psi}}^{\frac{1}{2}}\widehat{\boldsymbol{\Phi}}'(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1}\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}}$ and $\widetilde{\mathbf{J}} = \widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}'\boldsymbol{\Sigma}_u^{-1}\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}$. Using the Sherman-Morrison-Woodbury formula, we have

$$\|(\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_E^{-1}\| \le \|(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_u^{-1}\| + \Delta_{1'},$$

where $\Delta_{1'} = \|(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1}\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}}(\mathbf{I}_r + \widehat{\mathbf{J}})^{-1}\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}}\widehat{\boldsymbol{\Phi}}'(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_u^{-1}\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}(\mathbf{I}_r + \widetilde{\mathbf{J}})^{-1}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}'\boldsymbol{\Sigma}_u^{-1}\|$. Then, the right hand side can be bounded by the following terms:

$$L_{1'} = \|((\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_u^{-1})\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}(\mathbf{I}_r + \widetilde{\mathbf{J}})^{-1}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}'\boldsymbol{\Sigma}_u^{-1}\|,$$
$$L_{2'} = \|\boldsymbol{\Sigma}_u^{-1}(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}} - \widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}})(\mathbf{I}_r + \widetilde{\mathbf{J}})^{-1}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}'\boldsymbol{\Sigma}_u^{-1}\|,$$

$$L_{3'} = \|\boldsymbol{\Sigma}_u^{-1}\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}((\mathbf{I}_r + \widehat{\mathbf{J}})^{-1} - (\mathbf{I}_r + \widetilde{\mathbf{J}})^{-1})\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}'\boldsymbol{\Sigma}_u^{-1}\|.$$

By Lemma S.5, $\|\boldsymbol{\Phi}^l\boldsymbol{\Psi}^{l\frac{1}{2}} - \widehat{\boldsymbol{\Phi}}^l\widehat{\boldsymbol{\Psi}}^{l\frac{1}{2}}\|_{\max} \leq \|\boldsymbol{\Lambda}^l\widehat{\boldsymbol{\Psi}}^{l-\frac{1}{2}}(\boldsymbol{\Psi}^{l\frac{1}{2}} - \widehat{\boldsymbol{\Psi}}^{l\frac{1}{2}})\|_{\max} + \|(\boldsymbol{\Phi}^l - \widehat{\boldsymbol{\Phi}}^l)\boldsymbol{\Psi}^{l\frac{1}{2}}\|_{\max} = O_P(\omega_{T^\alpha})$, and by (S.19) and (S.23), we then have

$$\|\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\| \leq \max_l \|\widetilde{\boldsymbol{\Phi}}^l\widetilde{\boldsymbol{\Psi}}^{l\frac{1}{2}}\| = O_P(\sqrt{p^{ca_2}}),$$
$$\|\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}} - \widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\| \leq \max_l \sqrt{p^c}\|\widehat{\boldsymbol{\Phi}}^l\widehat{\boldsymbol{\Psi}}^{l\frac{1}{2}} - \widetilde{\boldsymbol{\Phi}}^l\widetilde{\boldsymbol{\Psi}}^{l\frac{1}{2}}\|_{\max} = O_P\left(\sqrt{p^c}\omega_{T^\alpha}\right),$$

and

$$\begin{aligned}
\|\widehat{\mathbf{J}} - \widetilde{\mathbf{J}}\| &\leq \|(\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}}\widehat{\boldsymbol{\Phi}}' - \widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}')(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1}(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}} - \widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}})\| \\
&\quad + \|(\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}}\widehat{\boldsymbol{\Phi}}' - \widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}')(\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1}\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\| + \|\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\widetilde{\boldsymbol{\Phi}}'((\widehat{\boldsymbol{\Sigma}}_u^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_u^{-1})\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\| \\
&= O_P(p^{\frac{c}{2}(1+a_2)}\omega_{T^\alpha} + p^{ca_2}m_p\omega_{T^\alpha}^{1-q}).
\end{aligned}$$

Since $\lambda_{\min}(\mathbf{I}_r + \widetilde{\mathbf{J}}) \geq \lambda_{\min}(\widetilde{\mathbf{J}}) \geq \lambda_{\min}(\boldsymbol{\Sigma}_u^{-1})\lambda_{\min}^2(\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}) \geq Cp^{ca_2}$, we have $\|(\mathbf{I}_r + \widetilde{\mathbf{J}})^{-1}\| = O_P(1/p^{ca_2})$. Then, $L_{1'} = O_P(m_p\omega_{T^\alpha}^{1-q})$ by (S.23). In addition, $L_{2'} = O_P(p^{-ca_2/2}\|\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}^{\frac{1}{2}} - \widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\|) = O_P(p^{\frac{c}{2}(1-a_2)}\omega_{T^\alpha})$ and $L_{3'} = O_P(p^{ca_2}\|(\mathbf{I}_r+\widehat{\mathbf{J}})^{-1}-(\mathbf{I}_r+\widetilde{\mathbf{J}})^{-1}\|) = O_P(p^{-ca_2}\|\widehat{\mathbf{J}}-\widetilde{\mathbf{J}}\|) = O_P(p^{\frac{c}{2}(1-a_2)}\omega_{T^\alpha} + m_p\omega_{T^\alpha}^{1-q})$. Thus, we have

$$\Delta_{1'} = O_P(p^{\frac{c}{2}(1-a_2)}\omega_{T^\alpha} + m_p\omega_{T^\alpha}^{1-q}), \tag{S.24}$$

which yields $\|(\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_E^{-1}\| = O_P(p^{\frac{c}{2}(1-a_2)}\omega_{T^\alpha} + m_p\omega_{T^\alpha}^{1-q})$.

Let $\widehat{\mathbf{H}} = \widehat{\boldsymbol{\Gamma}}^{\frac{1}{2}}\widehat{\mathbf{V}}'(\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}})^{-1}\widehat{\mathbf{V}}\widehat{\boldsymbol{\Gamma}}^{\frac{1}{2}}$ and $\widetilde{\mathbf{H}} = \widetilde{\boldsymbol{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'\boldsymbol{\Sigma}_E^{-1}\widetilde{\mathbf{V}}\widetilde{\boldsymbol{\Gamma}}^{\frac{1}{2}}$. Using the Sherman-Morrison-Woodbury formula again, we have

$$\|(\widehat{\boldsymbol{\Sigma}}^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}^{-1}\| \leq \|(\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_E^{-1}\| + \Delta_{2'},$$

where $\Delta_{2'} = \|(\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}})^{-1}\widehat{\mathbf{V}}\widehat{\boldsymbol{\Gamma}}^{\frac{1}{2}}(\mathbf{I}_k + \widehat{\mathbf{H}})^{-1}\widehat{\boldsymbol{\Gamma}}^{\frac{1}{2}}\widehat{\mathbf{V}}'(\widehat{\boldsymbol{\Sigma}}_E^{\mathcal{D}})^{-1} - \boldsymbol{\Sigma}_E^{-1}\widetilde{\mathbf{V}}\widetilde{\boldsymbol{\Gamma}}^{\frac{1}{2}}(\mathbf{I}_k + \widetilde{\mathbf{H}})^{-1}\widetilde{\boldsymbol{\Gamma}}^{\frac{1}{2}}\widetilde{\mathbf{V}}'\boldsymbol{\Sigma}_E^{-1}\|$. By

Weyl's inequality, we have $\lambda_{\min}(\mathbf{\Sigma}_E) > c$ since $\lambda_{\min}(\mathbf{\Sigma}_u) > c$ and $\lambda_{\min}(\mathbf{\Lambda}\mathbf{\Lambda}') = 0$. Hence, $\|\mathbf{\Sigma}_E^{-1}\| = O_P(1)$. By Lemmas S.1-S.4, we have $\|\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}^{\frac{1}{2}} - \widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}^{\frac{1}{2}}\|_{\max} = O_P(p^{\frac{5}{2}(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5}{2}a_1 - \frac{3}{2} - c})$. Similar to the proof of (S.10), we can show $\Delta_{2'} = O_P(m_p \omega_{T^\alpha}^{1-q} + p^{3(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{3a_1-2-c})$. Therefore, we have $\|(\widehat{\mathbf{\Sigma}}^{\mathcal{D}})^{-1} - \mathbf{\Sigma}^{-1}\| = O_P(m_p \omega_{T^\alpha}^{1-q} + p^{\frac{c}{2}(1-a_2)}\omega_{T^\alpha} + p^{3(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{3a_1-2-c})$.

Consider (S.6). We derive the rate of convergence for $\|\widehat{\mathbf{\Sigma}}^{\mathcal{D}} - \mathbf{\Sigma}\|_\Sigma$. The SVD decomposition of $\mathbf{\Sigma}$ is

$$\mathbf{\Sigma} = \begin{pmatrix}\mathbf{V}_{p\times k} & \mathbf{\Phi}_{p\times r} & \mathbf{\Omega}_{p\times(p-k-r)}\end{pmatrix} \begin{pmatrix} \mathbf{\Gamma}_{k\times k} & & \\ & \mathbf{\Psi}_{r\times r} & \\ & & \mathbf{\Theta}_{(p-k-r)\times(p-k-r)} \end{pmatrix} \begin{pmatrix} \mathbf{V}' \\ \mathbf{\Phi}' \\ \mathbf{\Omega}' \end{pmatrix}.$$

Note that $\mathbf{\Omega}$ is used to denote the precision matrix in Section 2. Moreover, since all the eigenvalues of $\mathbf{\Sigma}$ are strictly bigger than 0, for any matrix $\mathbf{A}$, we have $\|\mathbf{A}\|_\Sigma^2 = O_P(p^{-1})\|\mathbf{A}\|_F^2$. Then, we have

$$\begin{aligned}
\|\widehat{\mathbf{\Sigma}}^{\mathcal{D}} - \mathbf{\Sigma}\|_\Sigma \leq p^{-1/2}\Big(&\|\mathbf{\Sigma}^{-1/2}(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}')\mathbf{\Sigma}^{-1/2}\|_F \\
&+ \|\mathbf{\Sigma}^{-1/2}(\widehat{\mathbf{\Phi}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{\Phi}}' - \mathbf{\Lambda}\mathbf{\Lambda}')\mathbf{\Sigma}^{-1/2}\|_F + \|\mathbf{\Sigma}^{-1/2}(\widehat{\mathbf{\Sigma}}_u^{\mathcal{D}} - \mathbf{\Sigma}_u)\mathbf{\Sigma}^{-1/2}\|_F\Big) \\
=:& \Delta_{G'} + \Delta_{L'} + \Delta_{S'}
\end{aligned}$$

and

$$\Delta_{S'} = O_P(p^{-1/2}\|\widehat{\mathbf{\Sigma}}_u^{\mathcal{D}} - \mathbf{\Sigma}_u\|_F) = O_P(\|\widehat{\mathbf{\Sigma}}_u^{\mathcal{D}} - \mathbf{\Sigma}_u\|_2) = O_P(m_p \omega_{T^\alpha}^{1-q}).$$

We have

$$\begin{aligned}
\Delta_{G'} &= p^{-1/2}\left\| \begin{pmatrix} \mathbf{\Gamma}^{-\frac{1}{2}}\mathbf{V}' \\ \mathbf{\Psi}^{-\frac{1}{2}}\mathbf{\Phi}' \\ \mathbf{\Theta}^{-\frac{1}{2}}\mathbf{\Omega}' \end{pmatrix} (\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}') \begin{pmatrix}\mathbf{V}\mathbf{\Gamma}^{-\frac{1}{2}} & \mathbf{\Phi}\mathbf{\Psi}^{-\frac{1}{2}} & \mathbf{\Omega}\mathbf{\Theta}^{-\frac{1}{2}}\end{pmatrix} \right\|_F \\
&\leq p^{-1/2}\big(\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}')\mathbf{V}\mathbf{\Gamma}^{-1/2}\|_F + \|\mathbf{\Psi}^{-1/2}\mathbf{\Phi}'(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}')\mathbf{\Phi}\mathbf{\Psi}^{-1/2}\|_F
\end{aligned}$$

$$+ \|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}')\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F + 2\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}')\mathbf{\Phi}\mathbf{\Psi}^{-1/2}\|_F$$

$$+ 2\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}')\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F + 2\|\mathbf{\Psi}^{-1/2}\mathbf{\Phi}'(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{B}\mathbf{B}')\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F\Big)$$

$$=: \Delta_{G1'} + \Delta_{G2'} + \Delta_{G3'} + 2\Delta_{G4'} + 2\Delta_{G5'} + 2\Delta_{G6'}.$$

In order to find the convergence rate of relative Frobenius norm, we consider the above terms separately. For $\Delta_{G1'}$, we have

$$\Delta_{G1'} \le p^{-1/2}\left(\|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}' - \mathbf{V}\mathbf{\Gamma}\mathbf{V}')\mathbf{V}\mathbf{\Gamma}^{-1/2}\|_F + \|\mathbf{\Gamma}^{-1/2}\mathbf{V}'(\mathbf{V}\mathbf{\Gamma}\mathbf{V}' - \mathbf{B}\mathbf{B}')\mathbf{V}\mathbf{\Gamma}^{-1/2}\|_F\right)$$
$$=: \Delta_{G1'}^{(a)} + \Delta_{G1'}^{(b)}.$$

We bound the two terms separately. We have

$$\Delta_{G1'}^{(a)} \le p^{-1/2}\big(\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widehat{\mathbf{V}} - \mathbf{I})\widehat{\mathbf{\Gamma}}(\widehat{\mathbf{V}}'\mathbf{V} - \mathbf{I})\mathbf{\Gamma}^{-1/2}\|_F + 2\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widehat{\mathbf{V}} - \mathbf{I})\widehat{\mathbf{\Gamma}}\mathbf{\Gamma}^{-1/2}\|_F$$
$$+ \|(\mathbf{\Gamma}^{-1/2}(\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma})\mathbf{\Gamma}^{-1/2}\|_F\big) =: I + II + III.$$

By Lemma S.4, $\|\mathbf{V}'\widehat{\mathbf{V}} - \mathbf{I}\|_F = \|\mathbf{V}'(\widehat{\mathbf{V}} - \mathbf{V})\|_F \le \|\widehat{\mathbf{V}} - \mathbf{V}\|_F = O_P(p^{3(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{3a_1-2-c})$. Then, II is of order $O_P(p^{3(1-a_1)-\frac{1}{2}}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{3(a_1-\frac{1}{2})-c})$ and I is of smaller order. In addition, we have $III \le \|\mathbf{\Gamma}^{-1/2}(\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma})\mathbf{\Gamma}^{-1/2}\| = O_P(p^{1-a_1}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}))$ by Lemma S.4. Thus, $\Delta_{G1'}^{(a)} = O_P(p^{1-a_1}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{3(a_1-\frac{1}{2})-c})$. Similarly, we have

$$\Delta_{G1'}^{(b)} \le p^{-1/2}\big(\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widetilde{\mathbf{V}} - \mathbf{I})\widetilde{\mathbf{\Gamma}}(\widetilde{\mathbf{V}}'\mathbf{V} - \mathbf{I})\mathbf{\Gamma}^{-1/2}\|_F + 2\|\mathbf{\Gamma}^{-1/2}(\mathbf{V}'\widetilde{\mathbf{V}} - \mathbf{I})\widetilde{\mathbf{\Gamma}}\mathbf{\Gamma}^{-1/2}\|_F$$
$$+ \|(\mathbf{\Gamma}^{-1/2}(\widetilde{\mathbf{\Gamma}} - \mathbf{\Gamma})\mathbf{\Gamma}^{-1/2}\|_F\big) =: I' + II' + III'.$$

By $\sin\theta$ theorem, $\|\mathbf{V}'\widetilde{\mathbf{V}} - \mathbf{I}\| = \|\mathbf{V}'(\widetilde{\mathbf{V}} - \mathbf{V})\| \le \|\widetilde{\mathbf{V}} - \mathbf{V}\| = O(\|\mathbf{\Sigma}_E\|/p^{a_1})$. Then, we have $II' = O(1/p^{a_1-ca_2})$ and I' is of smaller order. By Lemma S.1, we have $III' = O(1/p^{a_1-ca_2})$.

Thus, $\Delta_{G1'}^{(b)} = O(1/p^{a_1 - ca_2})$. Then, we obtain

$$\Delta_{G1'} = O_P\left(p^{1-a_1}\Big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{3(a_1-\frac{1}{2})-c}} + \frac{1}{p^{a_1-ca_2}}\right). \tag{S.25}$$

For $\Delta_{G3'}$, we have

$$\Delta_{G3'} \le p^{-1/2}\|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'\widehat{\mathbf{V}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{V}}'\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F + p^{-1/2}\|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'\widetilde{\mathbf{V}}\widetilde{\mathbf{\Gamma}}\widetilde{\mathbf{V}}'\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F =: \Delta_{G3'}^{(a)} + \Delta_{G3'}^{(b)}.$$

By Lemma S.4, we have

$$\|\mathbf{\Omega}'\widehat{\mathbf{V}}\|_F = \|\mathbf{\Omega}'(\widehat{\mathbf{V}} - \mathbf{V})\|_F = O(\sqrt{p}\|\widehat{\mathbf{V}} - \mathbf{V}\|_{\max}) = O_P(p^{3(1-a_1)}(\sqrt{\log p/T^\alpha} + 1/T^{(1-\alpha)\beta}) + 1/p^{3a_1-2-c}).$$

Since $\|\widehat{\mathbf{\Gamma}}\| = O_P(p^{a_1})$, we have

$$\Delta_{G3'}^{(a)} \le p^{-1/2}\|\mathbf{\Theta}^{-1}\|\|\mathbf{\Omega}'\widehat{\mathbf{V}}\|_F^2\|\widehat{\mathbf{\Gamma}}\| = O_P(p^{11/2-5a_1}(\log p/T^\alpha + 1/T^{2(1-\alpha)\beta}) + 1/p^{5a_1-7/2-2c}).$$

Similarly, $\Delta_{G3'}^{(b)} = O_P(1/p^{5a_1-7/2-2c})$ because $\|\mathbf{\Omega}'\widetilde{\mathbf{V}}\|_F = O(\sqrt{p}\|\widetilde{\mathbf{V}} - \mathbf{V}\|_{\max}) = O_P(1/p^{3a_1-2-c})$ by Lemma S.2. Then, we obtain

$$\Delta_{G3'} = O_P\left(p^{\frac{11}{2}-5a_1}\Big(\frac{\log p}{T^\alpha} + \frac{1}{T^{2(1-\alpha)\beta}}\Big) + \frac{1}{p^{5a_1-\frac{7}{2}-2c}}\right).$$

Similarly, we can show that the terms $\Delta_{G2'}$, $\Delta_{G4'}$, $\Delta_{G5'}$ and $\Delta_{G6'}$ are dominated by $\Delta_{G1'}$ and $\Delta_{G3'}$. Therefore, we have

$$\Delta_{G'} = O_P\left(p^{1-a_1}\Big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\Big) + \frac{1}{p^{a_1-ca_2}} + p^{\frac{11}{2}-5a_1}\Big(\frac{\log p}{T^\alpha} + \frac{1}{T^{2(1-\alpha)\beta}}\Big) + \frac{1}{p^{5a_1-\frac{7}{2}-2c}}\right). \tag{S.26}$$

Similarly, we consider

$$\begin{aligned}
\Delta_{L'} &= p^{-1/2}\left\| \begin{pmatrix} \boldsymbol{\Gamma}^{-\frac{1}{2}}\mathbf{V}' \\ \boldsymbol{\Psi}^{-\frac{1}{2}}\boldsymbol{\Phi}' \\ \boldsymbol{\Theta}^{-\frac{1}{2}}\boldsymbol{\Omega}' \end{pmatrix} (\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}') \left(\mathbf{V}\boldsymbol{\Gamma}^{-\frac{1}{2}} \;\; \boldsymbol{\Phi}\boldsymbol{\Psi}^{-\frac{1}{2}} \;\; \boldsymbol{\Omega}\boldsymbol{\Theta}^{-\frac{1}{2}}\right) \right\|_F \\
&\le p^{-1/2}\big(\|\boldsymbol{\Gamma}^{-1/2}\mathbf{V}'(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\mathbf{V}\boldsymbol{\Gamma}^{-1/2}\|_F + \|\boldsymbol{\Psi}^{-1/2}\boldsymbol{\Phi}'(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Phi}\boldsymbol{\Psi}^{-1/2}\|_F \\
&\quad + \|\boldsymbol{\Theta}^{-1/2}\boldsymbol{\Omega}'(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Omega}\boldsymbol{\Theta}^{-1/2}\|_F + 2\|\boldsymbol{\Gamma}^{-1/2}\mathbf{V}'(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Phi}\boldsymbol{\Psi}^{-1/2}\|_F \\
&\quad + 2\|\boldsymbol{\Gamma}^{-1/2}\mathbf{V}'(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Omega}\boldsymbol{\Theta}^{-1/2}\|_F + 2\|\boldsymbol{\Psi}^{-1/2}\boldsymbol{\Phi}'(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Omega}\boldsymbol{\Theta}^{-1/2}\|_F\big) \\
&=: \Delta_{L1'} + \Delta_{L2'} + \Delta_{L3'} + 2\Delta_{L4'} + 2\Delta_{L5'} + 2\Delta_{L6'}.
\end{aligned}$$

For $\Delta_{L2'}$, similar to the proof of (S.26), we have

$$\begin{aligned}
\Delta_{L2'} &\le p^{-1/2}\left(\|\boldsymbol{\Psi}^{-1/2}\boldsymbol{\Phi}'(\widehat{\boldsymbol{\Phi}}\widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Phi}}' - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\Phi}')\boldsymbol{\Phi}\boldsymbol{\Psi}^{-1/2}\|_F + \|\boldsymbol{\Psi}^{-1/2}\boldsymbol{\Phi}'(\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\Phi}' - \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Phi}\boldsymbol{\Psi}^{-1/2}\|_F\right) \\
&=: \Delta_{L2'}^{(a)} + \Delta_{L2'}^{(b)}.
\end{aligned}$$

We have

$$\begin{aligned}
\Delta_{L2'}^{(a)} &\le p^{-1/2}\big(\|\boldsymbol{\Psi}^{-1/2}(\boldsymbol{\Phi}'\widehat{\boldsymbol{\Phi}} - \mathbf{I})\widehat{\boldsymbol{\Psi}}(\widehat{\boldsymbol{\Phi}}'\boldsymbol{\Phi} - \mathbf{I})\boldsymbol{\Psi}^{-1/2}\|_F + 2\|\boldsymbol{\Psi}^{-1/2}(\boldsymbol{\Phi}'\widehat{\boldsymbol{\Phi}} - \mathbf{I})\widehat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1/2}\|_F \\
&\quad + \|(\boldsymbol{\Psi}^{-1/2}(\widehat{\boldsymbol{\Psi}} - \boldsymbol{\Psi})\boldsymbol{\Psi}^{-1/2}\|_F\big) =: I + II + III.
\end{aligned}$$

By Lemma S.5, we have $\|\widehat{\boldsymbol{\Phi}}^l - \boldsymbol{\Phi}^l\|_F \le \sqrt{p_l r_l}\|\widehat{\boldsymbol{\Phi}}^l - \boldsymbol{\Phi}^l\|_{\max} = O_P\big(p^{\frac{5}{2}(1-a_1)+3c(1-a_2)}(\sqrt{\log p/T^{\alpha}} + 1/T^{(1-\alpha)\beta}) + 1/p^{\frac{5a_1}{2}-\frac{3}{2}+c(3a_2-4)} + m_p/p^{c(3a_2-2)}\big)$. Because $\widehat{\boldsymbol{\Phi}}$ and $\boldsymbol{\Phi}$ are block diagonal matrices, we have

$$\begin{aligned}
\|\widehat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}\|_F^2 &= \sum_{l=1}^{L}\|\widehat{\boldsymbol{\Phi}}^l - \boldsymbol{\Phi}^l\|_F^2 \\
&= O_P\left(p^{1-c}\big(p^{5(1-a_1)+6c(1-a_2)}(\frac{\log p}{T^{\alpha}} + \frac{1}{T^{2(1-\alpha)\beta}}) + \frac{1}{p^{5a_1-3+2c(3a_2-4)}} + \frac{m_p^2}{p^{2c(3a_2-2)}}\big)\right).
\end{aligned}$$

Then, II is of order $O_P(p^{\frac{5}{2}(1-a_1)+c(\frac{5}{2}-3a_2)}(\sqrt{\log p/T^{\alpha}}+1/T^{(1-\alpha)\beta})+1/p^{\frac{5}{2}a_1-\frac{3}{2}+c(3a_2-\frac{7}{2})}+m_p/p^{3c(a_2-\frac{1}{2})})$ and I is of smaller order. Also, $III=O_P(p^{\frac{5}{2}(1-a_1)+c(1-a_2)}(\sqrt{\log p/T^{\alpha}}+1/T^{(1-\alpha)\beta})+1/p^{\frac{5}{2}a_1-\frac{3}{2}-2c+ca_2})$ by Lemma S.5. Thus, $\Delta_{L2'}^{(a)}=O_P(p^{\frac{5}{2}(1-a_1)+c(1-a_2)}(\sqrt{\log p/T^{\alpha}}+1/T^{(1-\alpha)\beta})+1/p^{\frac{5}{2}a_1-\frac{3}{2}-2c+ca_2}+m_p/p^{3c(a_2-\frac{1}{2})})$. Similarly, we have

$$\begin{aligned}\Delta_{L2'}^{(b)} &\leq p^{-1/2}\big(\|\mathbf{\Psi}^{-1/2}(\mathbf{\Phi}'\widetilde{\mathbf{\Phi}}-\mathbf{I})\widetilde{\mathbf{\Psi}}(\widetilde{\mathbf{\Phi}}'\mathbf{\Phi}-\mathbf{I})\mathbf{\Psi}^{-1/2}\|_F+2\|\mathbf{\Psi}^{-1/2}(\mathbf{\Phi}'\widetilde{\mathbf{\Phi}}-\mathbf{I})\widetilde{\mathbf{\Psi}}\mathbf{\Psi}^{-1/2}\|_F\\ &\quad+\|(\mathbf{\Psi}^{-1/2}(\widetilde{\mathbf{\Psi}}-\mathbf{\Psi})\mathbf{\Psi}^{-1/2}\|_F\big)=:I'+II'+III'.\end{aligned}$$

By $\sin\theta$ theorem, $\|\mathbf{\Phi}'\widetilde{\mathbf{\Phi}}-\mathbf{I}\|\leq\|\widetilde{\mathbf{\Phi}}-\mathbf{\Phi}\|\leq\max_{j\leq L}\|\widetilde{\mathbf{\Phi}}^{l}-\mathbf{\Phi}^{l}\|\leq O(m_p/p^{ca_2})$. Then, we have $II'=O(m_p/p^{ca_2})$ and I' is of smaller order. By Lemma S.1, we have $III'=O(m_p/p^{ca_2})$. Thus, $\Delta_{L2'}^{(b)}=O(m_p/p^{ca_2})$. Then, we obtain

$$\Delta_{L2'}=O_P\left(p^{\frac{5}{2}(1-a_1)+c(1-a_2)}\big(\sqrt{\frac{\log p}{T^{\alpha}}}+\frac{1}{T^{(1-\alpha)\beta}}\big)+\frac{1}{p^{\frac{5}{2}a_1-\frac{3}{2}-2c+ca_2}}+\frac{m_p}{p^{ca_2}}\right). \tag{S.27}$$

For $\Delta_{L3'}$, we have

$$\Delta_{L3'}\leq p^{-1/2}\|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'\widehat{\mathbf{\Phi}}\widehat{\mathbf{\Psi}}\widehat{\mathbf{\Phi}}'\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F+p^{-1/2}\|\mathbf{\Theta}^{-1/2}\mathbf{\Omega}'\widetilde{\mathbf{\Phi}}\widetilde{\mathbf{\Psi}}\widetilde{\mathbf{\Phi}}'\mathbf{\Omega}\mathbf{\Theta}^{-1/2}\|_F=:\Delta_{L3'}^{(a)}+\Delta_{L3'}^{(b)}.$$

Since $\|\widehat{\mathbf{\Psi}}\|=O_P(p^{ca_2})$, we have

$$\begin{aligned}\Delta_{L3'}^{(a)} &\leq p^{-1/2}\|\mathbf{\Theta}^{-1}\|\|\mathbf{\Omega}'(\widehat{\mathbf{\Phi}}-\mathbf{\Phi})\|_F^2\|\widehat{\mathbf{\Psi}}\|\\ &=O_P\left(p^{\frac{11}{2}-5a_1+5c(1-a_2)}\big(\frac{\log p}{T^{\alpha}}+\frac{1}{T^{2(1-\alpha)\beta}}\big)+\frac{1}{p^{5a_1-\frac{7}{2}-c(7-5a_2)}}+\frac{m_p^2}{p^{5ca_2-3c-\frac{1}{2}}}\right).\end{aligned}$$

Similarly, by Lemma S.2, $\Delta_{L3'}^{(b)}=O_P(m_p^2/p^{5ca_2-3c-1/2})$ because $\|\widetilde{\mathbf{\Phi}}^{l}-\mathbf{\Phi}^{l}\|_F\leq\sqrt{p_l r_l}\|\widetilde{\mathbf{\Phi}}^{l}-\mathbf{\Phi}^{l}\|_{\max}=O(m_p/p^{c(3a_2-2)})$ and $\|\mathbf{\Omega}'\widetilde{\mathbf{\Phi}}\|_F^2\leq\|\widetilde{\mathbf{\Phi}}-\mathbf{\Phi}\|_F^2=\sum_{l=1}^{L}\|\widetilde{\mathbf{\Phi}}^{l}-\mathbf{\Phi}^{l}\|_F^2=O(m_p^2/p^{3c(2a_2-1)-1})$. Similarly, we can show $\Delta_{L1'}$, $\Delta_{L4'}$, $\Delta_{L5'}$ and $\Delta_{L6'}$ are dominated by $\Delta_{L2'}$ and $\Delta_{L3'}$. Therefore, we have

$$\Delta_{L'} = O_P\Big(p^{\frac{5}{2}(1-a_1)+c(1-a_2)}\big(\sqrt{\frac{\log p}{T^\alpha}} + \frac{1}{T^{(1-\alpha)\beta}}\big) + \frac{1}{p^{\frac{5}{2}a_1-\frac{3}{2}-2c+ca_2}} + \frac{m_p}{p^{ca_2}}$$
$$+ p^{\frac{11}{2}-5a_1+5c(1-a_2)}\big(\frac{\log p}{T^\alpha} + \frac{1}{T^{2(1-\alpha)\beta}}\big) + \frac{1}{p^{5a_1-\frac{7}{2}-c(7-5a_2)}} + \frac{m_p^2}{p^{5ca_2-3c-\frac{1}{2}}}\Big). \quad \text{(S.28)}$$

Combining the terms $\Delta_{G'}$, $\Delta_{L'}$ and $\Delta_{S'}$ together, we complete the proof of (S.6). □

S.6 Data Generating Process for Simulation Study

We considered the true covariance as $\mathbf{\Sigma} = \mathbf{B}\mathbf{B}' + \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Sigma}_u$, where each row of $\mathbf{B}$ was drawn from $\mathcal{N}(\mu_B, \mathbf{I}_k)$, where each element of μ_B is i.i.d. Uniform$(-0.5, 0.5)$; for $\mathbf{\Lambda} = \mathrm{diag}(\Lambda^1, \dots, \mathbf{\Lambda}^l)$, each row of $\mathbf{\Lambda}^l$ for each l was drawn from $\mathcal{N}(\mu_{\mathbf{\Lambda}^l}, \mathbf{I}_{r_l})$, where each element of μ_{Λ^l} is i.i.d. Uniform$(-0.3, 0.3)$. We generated $\mathbf{\Sigma}_u$ as follows. Let $\mathbf{D}_u = \mathrm{diag}(d_1, \dots, d_p)$, where each $\{d_i\}$ was generated independently from Uniform$(0.5, 1.5)$. Let $\pi = (\pi_1, \dots, \pi_p)'$ be a sparse vector, where each π_i was drawn from $\mathcal{N}(0, 1)$ with probability $\frac{0.5}{\sqrt{p}\log p}$, and $\pi_i = 0$ otherwise. Then, we set $\mathbf{\Sigma}_u = \mathbf{D}_u + \pi\pi' - \mathrm{diag}\{\pi_1^2, \dots, \pi_p^2\}$. In the simulation, we generated $\mathbf{\Sigma}_u$ until it was positive definite.

Let $\mathbf{D}$ be the diagonal matrix consisting of the diagonal elements of $\mathbf{\Sigma}$. We then obtained the true correlation matrix $\mathbf{R} = \mathbf{D}^{-\frac{1}{2}}\mathbf{\Sigma}\mathbf{D}^{-\frac{1}{2}} = (\rho_{0,ij})_{p\times p}$. Next, we set $\mathbf{R}_h = (\rho_{h,ij})_{p\times p}$, where $\rho_{h,ij} = \mathrm{sgn}(\rho_{0,ij})(|\rho_{0,ij}| + 0.5h^\beta)$ if i and j belong to different continent groups, for $h = \frac{0.5}{d}$ and $\beta = 0.75$, and $\rho_{h,ij} = \rho_{0,ij}$ if i and j are in the same continent group. Let $\{\gamma_i, v_i\}_{i=1}^k$ be the leading eigenvalues and eigenvectors of $\widetilde{\mathbf{\Sigma}}_g = \mathbf{D}^{\frac{1}{2}}\mathbf{R}_h\mathbf{D}^{\frac{1}{2}} - \mathbf{\Lambda}\mathbf{\Lambda}' - \mathbf{\Sigma}_u$. Then, we obtained $\mathbf{B}_h = \mathbf{V}\mathbf{\Gamma}^{\frac{1}{2}}$, where $\mathbf{\Gamma} = \mathrm{diag}(\gamma_1, \dots, \gamma_k)$ and $\mathbf{V} = (v_1, \dots, v_k)$. We note that $\mathbf{B}_h$ represents the non-synchronized structure. Thus, we generated non-synchronized observations by

$$y_t = \mathbf{B}_h G_t + \mathbf{\Lambda} F_t + u_t,$$

where $G_t = \mathbf{\Upsilon} G_{t-1} + v_t$, $F_t = \bar{\mathbf{\Upsilon}} F_{t-1} + \bar{v}_t$, and $u_t = \mathbf{\Sigma}_u^{1/2}\widetilde{u}_t$, where $\widetilde{u}_t = \widetilde{\mathbf{\Upsilon}}\widetilde{u}_{t-1} + \epsilon_t$, with

$k \times k$, $r \times r$, $p \times p$ diagonal matrices $\mathbf{\Upsilon}$, $\bar{\mathbf{\Upsilon}}$, and $\widetilde{\mathbf{\Upsilon}}$, respectively. Each diagonal element of $\mathbf{\Upsilon}$, $\bar{\mathbf{\Upsilon}}$, and $\widetilde{\mathbf{\Upsilon}}$ was generated from Uniform(0,0.7), and υ_t, $\bar{\upsilon}_t$, and ϵ_t were drawn from $\mathcal{N}(0, \mathbf{I}_k)$, $\mathcal{N}(0, \mathbf{I}_r)$, and $\mathcal{N}(0, \mathbf{I}_p)$, respectively.

S.7 Additional Tables for Empirical Study

Table S.1: Distributions of the number of firms

America		Asia		Europe	
United States (US)	221	China (CN)	100	United Kingdom (GB)	100
Canada (CA)	100	Japan (JP)	100	France (FR)	100
Brazil (BR)	100	Hong Kong (HK)	100	Germany (DE)	100
Mexico (MX)	48	India (IN)	100	Switzerland (CH)	100
Chile (CL)	31	Korea (KR)	100	Sweden (SE)	100
				Total	1500

Table S.2: Out-of-sample Sharpe ratios and returns (multiplied by 10^4) for the full period from 2018 to 2022

	SamCov_{W}	POET_{W}	D-POET_{W}	$\text{SamCov}_{\text{2D}}$	POET_{2D}	$\text{D-POET}_{\text{2D}}$
Sharpe ratio	0.0477	0.0521	0.0580	0.0637	0.0579	0.0658
Return	6.255	6.129	6.771	7.795	6.716	7.410
	SamCov_{D}	POET_{D}	D-POET_{D}	S-POET_{W}	$\text{S-POET}_{\text{2D}}$	
Sharpe ratio	0.0712	0.0656	0.0663	0.0841	0.0696	
Return	7.978	7.315	7.373	8.965	7.417	

References

AHN, S. C., AND A. R. HORENSTEIN (2013): "Eigenvalue ratio test for the number of factors," *Econometrica*, 81(3), 1203–1227.

BAI, J., AND S. NG (2002): "Determining the number of factors in approximate factor models," *Econometrica*, 70(1), 191–221.

BICKEL, P. J., AND E. LEVINA (2008): "Covariance regularization by thresholding," *The Annals of Statistics*, 36(6), 2577–2604.

CAI, T., AND W. LIU (2011): "Adaptive thresholding for sparse covariance matrix estimation," *Journal of the American Statistical Association*, 106(494), 672–684.

CHOI, S. H., AND D. KIM (2023): "Large volatility matrix analysis using global and national factor models," *Journal of Econometrics*, 235(2), 1917–1933.

FAN, J., Y. LIAO, AND M. MINCHEVA (2011): "High dimensional covariance matrix estimation in approximate factor models," *The Annals of Statistics*, 39(6), 3320.

——— (2013): "Large covariance estimation by thresholding principal orthogonal complements," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 75(4), 603–680.

FAN, J., W. WANG, AND Y. ZHONG (2018): "An l_∞ eigenvector perturbation bound and its application to robust covariance estimation," *Journal of Machine Learning Research*, 18(207), 1–42.

ONATSKI, A. (2010): "Determining the number of factors from empirical distribution of eigenvalues," *The Review of Economics and Statistics*, 92(4), 1004–1016.

ROTHMAN, A. J., E. LEVINA, AND J. ZHU (2009): "Generalized thresholding of large covariance matrices," *Journal of the American Statistical Association*, 104(485), 177–186.