Invariance principle and CLT for the spiked eigenvalues of large-dimensional Fisher matrices and applications

Dandan Jiang[1], Zhiqiang Hou[2], Zhidong Bai[3] and Runze Li[4]

[1]*Xi'an Jiaotong University,* [2]*Shandong University of Finance and Economics,*

[3]*Northeast Normal University and* [4]*Pennsylvania State University at University Park*

Supplementary Material

S1. Truncation and centralization procedures.

Let $\delta(A)$ be the indication of set A. Let $\hat{x}_{ij} = x_{ij}\delta(|x_{ij}| < \eta_{n_1}\sqrt{n_1})$, $\tilde{x}_{ij} = (\hat{x}_{ij} - \mathrm{E}\hat{x}_{ij})/\sigma_{n_1}$ and $\hat{y}_{ij} = y_{ij}\delta(|y_{ij}| < \eta_{n_2}\sqrt{n_2})$, $\tilde{y}_{ij} = (\hat{y}_{ij} - \mathrm{E}\hat{y}_{ij})/\sigma_{n_2}$, where $\sigma_{n_1}^2 = \mathrm{E}\,|\hat{x}_{ij} - \mathrm{E}\hat{x}_{ij}|^2$ and $\sigma_{n_2}^2 = \mathrm{E}\,|\hat{y}_{ij} - \mathrm{E}\hat{y}_{ij}|^2$. By the related techniques of the proofs in Supplement B of Jiang and Bai (2021), we can show that it is equivalent to replace the entries of x_{ij}, y_{ij} with their corresponding truncated and centralized variables by Assumption 4. In addition, the convergence rates of arbitrary moments of $\tilde{x}_{ij}$ and $\tilde{y}_{ij}$ are the same as the ones depicted in Lemma C.1 in Jiang and Bai (2021). Therefore, it is reasonable to consider the generalized spiked Fisher matrix

$\mathbf{F} = \mathbf{S}_1\mathbf{S}_2^{-1}$, which is generated from the entries truncated at $\eta_{n_1}\sqrt{n_1}$ for x_{ij} and $\eta_{n_2}\sqrt{n_2}$ for y_{ij}, centralized and renormalized. For simplicity, we assume that $|x_{ij}| < \eta_{n_1}\sqrt{n_1}, |y_{ij}| < \eta_{n_2}\sqrt{n_2}$, $\mathrm{E}x_{ij} = \mathrm{E}y_{ij} = 0, \mathrm{E}|x_{ij}|^2 = \mathrm{E}|y_{ij}|^2 = 1$ for the real case and Assumption 4 is satisfied. But for the complex case, the truncation and renormalization cannot reserve the requirement of $\mathrm{E}x_{ij}^2 = \mathrm{E}y_{ij}^2 = 0$. However, one may prove that $\mathrm{E}x_{ij}^2 = o(n_1^{-1})$ and $\mathrm{E}y_{ij}^2 = o(n_2^{-1})$. By the truncation procedures, the Proposition 1 still holds in probability without the bounded fourth-moment assumption if the Assumption 4 is satisfied.

S2. The proof of Theorem 1

Let $(\mathbf{X}, \mathbf{Y})$ and $(\mathbf{Z}, \mathbf{W})$ be two independent pairs of random samples satisfying Assumptions 1–5, where $\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_{n_1})$, $\mathbf{Y} = (\mathbf{y}_1, \cdots, \mathbf{y}_{n_2})$, $\mathbf{Z} = (\mathbf{z}_1, \cdots, \mathbf{z}_{n_1})$ and $\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_{n_2})$. Denote $\mathbf{X}_k = (\mathbf{x}_1, \cdots, \mathbf{x}_k, \mathbf{z}_{k+1}, \cdots, \mathbf{z}_{n_1})$ and $\mathbf{Y}_k = (\mathbf{y}_1, \cdots, \mathbf{y}_k, \mathbf{w}_{k+1}, \cdots, \mathbf{w}_{n_2})$ with convention $\mathbf{X} = \mathbf{X}_{n_1}$, $\mathbf{Z} = \mathbf{X}_0$, $\mathbf{Y} = \mathbf{Y}_{n_2}$ and $\mathbf{W} = \mathbf{Y}_0$. Further, let $\mathbf{X}_{k0} = (\mathbf{x}_1 \cdots, \mathbf{x}_{k-1}, \mathbf{z}_{k+1} \cdots \mathbf{z}_{n_1})$ is the overlapping part between $\mathbf{X}_{k-1}$ and $\mathbf{X}_k$, and $\mathbf{Y}_{k0} = (\mathbf{y}_1 \cdots, \mathbf{y}_{k-1}, \mathbf{w}_{k+1} \cdots \mathbf{w}_{n_2})$ is the notation of the overlapping part between $\mathbf{Y}_{k-1}$ and $\mathbf{Y}_k$. $\mathbf{\Omega}_M(\lambda, \mathbf{X}, \mathbf{Y})$ defined in (2.12) is simply denoted as $\mathbf{\Omega}_M(\mathbf{X}, \mathbf{Y})$ or $\mathbf{\Omega}_M$, as the case may be. Note that the difference between $\mathbf{X}_k$ and $\mathbf{X}_{k0}$

lies in the kth column. That is, $\mathbf{x}_k$ in $\mathbf{X}_k$, and the difference between $\mathbf{X}_{k-1}$ and $\mathbf{X}_{k0}$ is also in the kth column, (that is, $\mathbf{z}_k$ in $\mathbf{X}_{k-1}$). Similarly, we obtain that the difference between $\mathbf{Y}_k$ and $\mathbf{Y}_{k0}$ is $\mathbf{y}_k$ in $\mathbf{Y}_k$, and the difference between $\mathbf{Y}_{k-1}$ and $\mathbf{Y}_{k0}$ is $\mathbf{w}_k$ in $\mathbf{Y}_{k-1}$.

To prove Theorem 1, it is equivalent to show that, for any $M \times M$ symmetric matrix $\boldsymbol{\Xi}$, $\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,j}(\mathbf{X},\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{W})\right\}} \to 0$ holds for all $\boldsymbol{\Omega}_{M,j}(\mathbf{X},\mathbf{Y}), j = 1, \cdots, 5$. Thus, it is sufficient to show that

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,j}(\mathbf{X},\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y})\right\}} \to 0; \tag{S2.1}$$

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{W})\right\}} \to 0. \tag{S2.2}$$

Consider the following differences

$$\boldsymbol{\Omega}_{M,j}(\mathbf{X},\mathbf{Y}) - \boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}) = \sum_{k=1}^{n_1} \left\{\boldsymbol{\Omega}_{M,j}(\mathbf{X}_k,\mathbf{Y}) - \boldsymbol{\Omega}_{M,j}(\mathbf{X}_{k-1},\mathbf{Y})\right\}$$
$$= \sum_{k=1}^{n_1} \Big[\left\{\boldsymbol{\Omega}_{M,j}(\mathbf{X}_k,\mathbf{Y}) - \boldsymbol{\Omega}_{M,j}(\mathbf{X}_{k0},\mathbf{Y})\right\} - \left\{\boldsymbol{\Omega}_{M,j}(\mathbf{X}_{k-1},\mathbf{Y}) - \boldsymbol{\Omega}_{M,j}(\mathbf{X}_{k0},\mathbf{Y})\right\}\Big]$$

and

$$\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}) - \boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{W}) = \sum_{k=1}^{n_2} \left\{\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_k) - \boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_{k-1})\right\} \tag{S2.3}$$
$$= \sum_{k=1}^{n_2} \Big[\left\{\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_k) - \boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_{k0})\right\} - \left\{\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_{k-1}) - \boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_{k0})\right\}\Big].$$

Thus, we focus on two terms: $\boldsymbol{\Omega}_{M,j}(\mathbf{X}_k,\mathbf{Y}) - \boldsymbol{\Omega}_{M,j}(\mathbf{X}_{k0},\mathbf{Y})$ and $\boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_k) - \boldsymbol{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_{k0})$ because we can get similar results for the other two terms

$\mathbf{\Omega}_{M,j}(\mathbf{X}_{k-1},\mathbf{Y}) - \mathbf{\Omega}_{M,j}(\mathbf{X}_{k0},\mathbf{Y})$ and $\mathbf{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_{k-1}) - \mathbf{\Omega}_{M,j}(\mathbf{Z},\mathbf{Y}_{k0})$, only with $\mathbf{x}_k, \mathbf{y}_k$ substituted by $\mathbf{z}_k, \mathbf{w}_k$, respectively.

Denote

$$\tilde{\mathbf{F}}_{k_0} = \frac{1}{n_1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_{k_0}\mathbf{X}_{k_0}^*\mathbf{U}_2\mathbf{D}_2^{\frac{1}{2}}\mathbf{Q}^{-\frac{1}{2}};$$
$$\underline{\tilde{\mathbf{F}}}_{k0} = \frac{1}{n_1}\mathbf{X}^*\mathbf{U}_2\mathbf{D}_2^{\frac{1}{2}}(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_{k0}\mathbf{Y}_{k0}^*\mathbf{V}_2)^{-1}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X};$$
$$\beta_{kx} = 1 - 1/n_1\mathbf{x}_k^*\mathbf{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I} - \tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{\Gamma}^{\frac{1}{2}*}\mathbf{x}_k;\ \beta_{k_0} = 1 - 1/n_1\mathrm{tr}\{\mathbf{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I} - \tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{\Gamma}^{\frac{1}{2}*}\};$$
$$\epsilon_{kx} = \beta_{kx} - \beta_{k_0};$$

where $\mathbf{\Gamma}^{1/2} = \mathbf{U}_2\mathbf{D}_2^{1/2}\mathbf{Q}^{-1/2}$ and β_{kz}, ε_{kz} are similarly defined as β_{kx} and ε_{kx} with $\mathbf{x}_k$ replaced by $\mathbf{z}_k$.

Furthermore, we are frequently taking expectations about the kth sample when the remain samples are given in this proof. After taking the expectations, we need to evaluate the resulting quantities which usually relate to $\|(\lambda\mathbf{I} - \tilde{\mathbf{F}}_{k0})^{-1}\|$, $\|(\lambda\mathbf{I} - \underline{\tilde{\mathbf{F}}}_{k0})^{-1}\|$, $\|n_2^{-1/2}\mathbf{Y}_{k0}\|$, $\|n_1^{-1/2}\mathbf{X}_{k0}\|$, where $\|\cdot\|$ is the L_2 norm. To avoid the overall expectation of the resulting quantities brows up by the fact of huge values on the small probability, we need to introduce some events to bound the quantities listed above as done in Jiang and Bai (2021). As shown in Lemma C.5 of Jiang and Bai (2021), the probability of the event that each quantity above is larger than a large given value B is as small as n^{-t} for any given $t > 0$. Following procedures given in Jiang and Bai (2021), one may prove that the final expectations are as small as needed. For brevity, we simply suppress the notations of those events in the proofs as below.

Define

$$\mathcal{E}_{k1} = \{|\epsilon_{kx}| \geq b\} \cup \{|\epsilon_{kz}| \geq b\},$$

where b is a given small positive constant.

For $\mathbf{\Omega}_{M,1}$, it is obvious that $\mathbf{\Omega}_{M,1}(\mathbf{X}_k, \mathbf{Y}) - \mathbf{\Omega}_{M,1}(\mathbf{Z}, \mathbf{Y}) = 0$, then (S2.1) holds for $\mathbf{\Omega}_{M,1}$. Consider (S2.2) for $\mathbf{\Omega}_{M,1}$, we have

$$\mathbf{\Omega}_{M,1}(\mathbf{Z}, \mathbf{Y}_k) - \mathbf{\Omega}_{M,j}(\mathbf{Z}, \mathbf{Y}_{k0}) = \frac{\sqrt{p}}{n_2}\mathbf{V}_1^*(\mathbf{Y}_k\mathbf{Y}_k^* - \mathbf{Y}_{k0}\mathbf{Y}_{k0}^*)\mathbf{V}_1 = \frac{\sqrt{p}}{n_2}\mathbf{V}_1^*\mathbf{y}_k\mathbf{y}_k^*\mathbf{V}_1.$$

Similarly,

$$\begin{aligned}
&\mathbf{\Omega}_{M,1}(\mathbf{Z}, \mathbf{Y}_{k-1}) - \mathbf{\Omega}_{M,j}(\mathbf{Z}, \mathbf{Y}_{k0}) \\
=&\frac{\sqrt{p}}{n_2}\mathbf{V}_1^*(\mathbf{Y}_{k-1}\mathbf{Y}_{k-1}^* - \mathbf{Y}_{k0}\mathbf{Y}_{k0}^*)\mathbf{V}_1 = \frac{\sqrt{p}}{n_2}\mathbf{V}_1^*\mathbf{w}_k\mathbf{w}_k^*\mathbf{V}_1.
\end{aligned}$$

Since $\mathbf{y}_k$ and $\mathbf{w}_k$ have the same expectations and the second moments, and their fourth moments satisfy the Assumption 4, then by (S2.3) and Taylor expansion, we have

$$\begin{aligned}
&\mathrm{E}e^{i\mathrm{tr}\left\{\mathbf{\Xi}\mathbf{\Omega}_{M,1}(\mathbf{Z},\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\mathbf{\Xi}\mathbf{\Omega}_{M,1}(\mathbf{Z},\mathbf{W})\right\}} = \sum_{k=1}^{n_2}\mathrm{E}e^{i\mathrm{tr}\left\{\mathbf{\Xi}\mathbf{\Omega}_{M,1}(\mathbf{Z},\mathbf{Y}_{k0})\right\}} \\
&\cdot\left[\mathrm{E}_k e^{i\mathrm{tr}\left[\mathbf{\Xi}\{\mathbf{\Omega}_{M,1}(\mathbf{Z},\mathbf{Y}_k)-\mathbf{\Omega}_{M,1}(\mathbf{Z},\mathbf{Y}_{k0})\}\right]} - \mathrm{E}_k e^{i\mathrm{tr}\left[\mathbf{\Xi}(\mathbf{\Omega}_{M,1}\{\mathbf{Z},\mathbf{Y}_{k-1})-\mathbf{\Omega}_{M,1}(\mathbf{Z},\mathbf{Y}_{k0})\}\right]}\right]
\end{aligned}$$

$$=\sum_{k=1}^{n_2}\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,1}(\mathbf{Z},\mathbf{Y}_{k0})\right\}}\left\{\mathrm{E}_k e^{i\mathrm{tr}(\boldsymbol{\Xi}\frac{\sqrt{p}}{n_2}\mathbf{V}_1^*\mathbf{y}_k\mathbf{y}_k^*\mathbf{V}_1)}-\mathrm{E}_k e^{i\mathrm{tr}(\boldsymbol{\Xi}\frac{\sqrt{p}}{n_2}\mathbf{V}_1^*\mathbf{w}_k\mathbf{w}_k^*\mathbf{V}_1)}\right\}\to 0.$$

Therefore, (S2.2) is also true for $\boldsymbol{\Omega}_{M,1}$. Here and hereafter, $\mathrm{E}_k(\cdot)=\mathrm{E}(\cdot|\mathbf{Z},\mathbf{Y}_{k0})$ in the proof of (S2.2); otherwise, $\mathrm{E}_k(\cdot)=\mathrm{E}(\cdot|\mathbf{X}_{k0},\mathbf{Y})$ when it is involved in the proof of (S2.1).

For $\boldsymbol{\Omega}_{M,2}$, let $\tilde{\mathbf{F}}_k=\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\mathbf{X}_k^*\mathbf{U}_2\mathbf{D}_2^{\frac{1}{2}}\mathbf{Q}^{-\frac{1}{2}}/n_1$, then

$$\boldsymbol{\Omega}_{M,2}(\mathbf{X}_k,\mathbf{Y})-\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})=\frac{\sqrt{p}\lambda}{n_2}\left\{\mathrm{tr}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_k)^{-1}-\mathrm{tr}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\right\}\mathbf{I}_M$$
$$-\frac{\sqrt{p}\lambda}{n_2^2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left\{(\lambda\mathbf{I}-\tilde{\mathbf{F}}_k)^{-1}-(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\right\}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_1.$$

It follows that

$$\tilde{\mathbf{F}}_k=\tilde{\mathbf{F}}_{k_0}+\frac{1}{n_1}\boldsymbol{\Gamma}^{\frac{1}{2}*}\mathbf{x}_k\mathbf{x}_k^*\boldsymbol{\Gamma}^{\frac{1}{2}} \tag{S2.4}$$

Then by the relationship (S2.4) and Lemma 6.9 in Bai and Silverstein (2010), we obtain that

$$\boldsymbol{\Omega}_{M,2}(\mathbf{X}_k,\mathbf{Y})-\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})=\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{kx}}\left[\left\{\frac{1}{n_1}\mathbf{x}_k^*\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-2}\boldsymbol{\Gamma}^{\frac{1}{2}*}\mathbf{x}_k\right\}\mathbf{I}_M\right.$$
$$\left.-\mathbf{V}_1^*\mathbf{Y}\left\{\frac{1}{n_2}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0})^{-1}\frac{1}{n_1}\boldsymbol{\Gamma}^{\frac{1}{2}*}\mathbf{x}_k\mathbf{x}_k^*\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\right\}\mathbf{Y}^*\mathbf{V}_1\right],$$

where by Lemma C.3 in Jiang and Bai (2021), β_{kx} and β_{k_0} satisfy that

$$\beta_{k0}+\frac{1}{\lambda\underline{m}(\lambda)}\to 0;\qquad \frac{1}{\lambda\underline{m}(\lambda)}\neq 0,\ \text{if } \lambda>0;$$
$$\varepsilon_{kx}=\beta_{kx}-\beta_{k0}\to 0;\qquad \mathrm{E}_k\varepsilon_{kx}^2\le o(n_1^{-1}\log n_1);\quad \mathrm{E}_k\varepsilon_{kx}^4=o(n_1^{-1}).$$

Here $m(\lambda)$ is the Stieltjes transform of the LSD of the matrix $\tilde{\mathbf{F}}_{k_0}$ and $\underline{m}(\lambda) = -(1-c_1)/\lambda + c_1 m(\lambda)$. Denote

$$\begin{aligned}
\boldsymbol{\tau}_{kx} =& \frac{1}{n_1}\Big[\mathbf{x}_k^*\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-2}\boldsymbol{\Gamma}^{\frac{1}{2}*}\mathbf{x}_k - \mathrm{tr}\big\{\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-2}\boldsymbol{\Gamma}^{\frac{1}{2}*}\big\}\Big]\mathbf{I}_M \\
& - \mathbf{V}_1^*\mathbf{Y}\big\{\frac{1}{n_2}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0})^{-1}\frac{1}{n_1}\boldsymbol{\Gamma}^{\frac{1}{2}*}(\mathbf{x}_k\mathbf{x}_k^*-\mathbf{I}_p)\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\big\}\mathbf{Y}^*\mathbf{V}_1 \\
\boldsymbol{\tau}_{k0} =& \frac{1}{n_1}\mathrm{tr}\big\{\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-2}\boldsymbol{\Gamma}^{\frac{1}{2}*}\big\}\mathbf{I}_M \\
& - \mathbf{V}_1^*\mathbf{Y}\big\{\frac{1}{n_2}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0})^{-1}\frac{1}{n_1}\boldsymbol{\Gamma}^{\frac{1}{2}*}\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\big\}\mathbf{Y}^*\mathbf{V}_1
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\boldsymbol{\Omega}_{M,2}(\mathbf{X}_k,\mathbf{Y}) - \boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y}) \\
=& \frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx}) - \frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}^2}(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx} + \frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}^2\beta_{kx}}(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}^2
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k-1},\mathbf{Y}) - \boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y}) \\
=& \frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz}) - \frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}^2}(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz} + \frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}^2\beta_{kz}}(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz}^2,
\end{aligned}$$

where $\boldsymbol{\tau}_{kz}$ is defined similarly to $\boldsymbol{\tau}_{kx}$ with $\mathbf{x}_k$ replaced by $\mathbf{z}_k$.

Now, we are in position to prove $\mathrm{E}e^{i t \mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{X},\mathbf{Y})\}} - \mathrm{E}e^{i t \mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{Z},\mathbf{Y})\}} \to 0$ for any $M\times M$ symmetric matrix $\boldsymbol{\Xi}$. By the notations introduced before, we have

$$\mathrm{E}e^{i t \mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{X},\mathbf{Y})\}} - \mathrm{E}e^{i t \mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{Z},\mathbf{Y})\}} = \sum_{k=1}^{n_1}\mathrm{E}e^{i t \mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})\}} \tag{S2.5}$$

$$
\cdot \Big[\mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_k,\mathbf{Y})-\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})\}\big]} - \mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k-1},\mathbf{Y})-\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})\}\big]}\Big]
$$

$$
= \sum_{k=1}^{n_1} \mathrm{E} e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})+\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}\boldsymbol{\tau}_{k0}\}\big]} \Big[\mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kx}}{\sqrt{n_2}\beta_{k0}} - \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}}{\sqrt{n_2}\beta_{k0}^2} + \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}^2}{\sqrt{n_2}\beta_{k0}^2\beta_{kx}}\}\big]}
$$

$$
- \mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kz}}{\sqrt{n_2}\beta_{k0}} - \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz}}{\sqrt{n_2}\beta_{k0}^2} + \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz}^2}{\sqrt{n_2}\beta_{k0}^2\beta_{kz}}\}\big]}\Big]
$$

$$
= \sum_{k=1}^{n_1} \mathrm{E} e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})+\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}\boldsymbol{\tau}_{k0}\}\big]}
$$

$$
\Big[\mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kx}}{\sqrt{n_2}\beta_{k0}} - \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}}{\sqrt{n_2}\beta_{k0}^2}\}\big]} - \mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kz}}{\sqrt{n_2}\beta_{k0}} - \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz}}{\sqrt{n_2}\beta_{k0}^2}\big]}\Big]
$$

$$
+ \sum_{k=1}^{n_1} \mathrm{E} e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})+\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}\boldsymbol{\tau}_{k0}\}\big]}
$$

$$
\Big[\mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kx}}{\sqrt{n_2}\beta_{k0}} - \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}}{\sqrt{n_2}\beta_{k0}^2}\}\big]} \big[e^{i\mathrm{tr}\big\{\frac{\boldsymbol{\Xi}\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}^2}{\sqrt{n_2}\beta_{k0}^2\beta_{kx}}\big\}} - 1\big]
$$

$$
- \mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kz}}{\sqrt{n_2}\beta_{k0}} - \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz}}{\sqrt{n_2}\beta_{k0}^2}\}\big]} \big[e^{i\mathrm{tr}\big\{\frac{\boldsymbol{\Xi}\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz}^2}{\sqrt{n_2}\beta_{k0}^2\beta_{kz}}\big\}} - 1\big]\Big].
$$

Similar to the proof of Theorem 3.2 in Jiang and Bai (2021), we can show that the last summation is $o(1)$. Take its first term as an example,

$$
\Bigg| \sum_{k=1}^{n_1} \mathrm{E} e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})+\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}\boldsymbol{\tau}_{k0}\}\big]}
$$

$$
\mathrm{E}_k e^{i\mathrm{tr}\big[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kx}}{\sqrt{n_2}\beta_{k0}} - \frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}}{\sqrt{n_2}\beta_{k0}^2}\}\big]} \big[e^{i\mathrm{tr}\big\{\frac{\boldsymbol{\Xi}\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}^2}{\sqrt{n_2}\beta_{k0}^2\beta_{kx}}\big\}} - 1\big]\Bigg|
$$

$$
\le \sum_{k=1}^{n_1} \mathrm{E}\left[K_0 \mathrm{E}_k\{2\delta(|\varepsilon_{kx}| \ge \delta_k)\} + \frac{\tilde{K}_0}{\sqrt{n_2}} \mathrm{E}_k|\mathrm{tr}\boldsymbol{\Xi}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}^2|\right]
$$

$$
\le \sum_{k=1}^{n} K_0 \mathrm{E}_k\varepsilon_k^4 + \frac{\tilde{K}_0}{\sqrt{n_2}}\{\mathrm{E}_k|(\mathrm{tr}\boldsymbol{\Xi}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_k)|^2 \mathrm{E}\varepsilon_k^4\}^{1/2} = o(1),
$$

where we have used the facts that $\mathrm{E}_k\varepsilon_{kx}^4 = o(n_1^{-1}), \mathrm{E}_k|\mathrm{tr}\boldsymbol{\Xi}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_k)|^2 \le K_0$ with $K_0, \tilde{K}_0$ being some suitable constants valued different at different appearances. The Chebyshev's inequality is also applied when $|\varepsilon_{kx}| \geq \delta_k$ and then the Taylor expansion and Cauchy-Schwartz inequality are used to the case of $|\varepsilon_{kx}| < \delta_k$. The δ_k should not be too small, like the half of the non-zero limit of β_{k0}; then, we have

$$|\beta_{kx}| \geq |\beta_{k0}| - |\varepsilon_{kx}|$$

and moreover $|\beta_{kx}| \geq 1/2|\beta_{k0}|$, which is bounded from below. Then its last term can be handled in the same way. Therefore, we have

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{X},\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{Z},\mathbf{Y})\right\}} = \sum_{k=1}^{n_1}\mathrm{E}e^{i\mathrm{tr}\left[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})+\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}\boldsymbol{\tau}_{k0}\}\right]}$$
$$\cdot\left[\mathrm{E}_k e^{i\mathrm{tr}\left[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kx}}{\sqrt{n_2}\beta_{k0}}-\frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kx})\varepsilon_{kx}}{\sqrt{n_2}\beta_{k0}^2}\}\right]} - \mathrm{E}_k e^{i\mathrm{tr}\left[\boldsymbol{\Xi}\{\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kz}}{\sqrt{n_2}\beta_{k0}}-\frac{\sqrt{c_2}\lambda(\boldsymbol{\tau}_{k0}+\boldsymbol{\tau}_{kz})\varepsilon_{kz}}{\sqrt{n_2}\beta_{k0}^2}\right]}\right] + o(1)$$

By the same strategy, we may show that

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{X},\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,2}(\mathbf{Z},\mathbf{Y})\right\}} = \sum_{k=1}^{n_1}\mathrm{E}e^{i\mathrm{tr}\left[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})+\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}\boldsymbol{\tau}_{k0}\}\right]} \tag{S2.6}$$
$$\cdot\left[\mathrm{E}_k e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}(\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kx}}{\sqrt{n_2}\beta_{k0}}-\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{k0}\varepsilon_{kx}}{\sqrt{n_2}\beta_{k0}^2})\right\}} - \mathrm{E}_k e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}(\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kz}}{\sqrt{n_2}\beta_{k0}}-\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{k0}\varepsilon_{kz}}{\sqrt{n_2}\beta_{k0}^2})\right\}}\right] + o(1)$$

Since

$$\left[\mathrm{tr}\left\{\boldsymbol{\Xi}(\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{kx}}{\sqrt{n_2}\beta_{k0}}-\frac{\sqrt{c_2}\lambda\boldsymbol{\tau}_{k0}\varepsilon_{kx}}{\sqrt{n_2}\beta_{k0}^2})\right\}\right]^2$$
$$=\frac{c_2}{n_2}\left[\frac{1}{\beta_{k0}^2}\left\{\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{kx})\right\}^2 - \frac{2}{\beta_{k0}^3}\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{kx})\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{k_0}\varepsilon_{kx}) + \frac{1}{\beta_{k0}^4}\left\{\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{k_0})\right\}^2\varepsilon_{kx}^2\right]$$

and noted that

$$\mathrm{E}_k\Big\{\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_k)\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{k0})\Big\} = o(1); \quad \mathrm{E}_k\big\{\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{k0})\big\}^2 = o(n_1^{-1}\log n_1).$$

So we only focus on the term $\big\{\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{kx})\big\}^2$.

Let $\mathbf{R}_1 = \lambda\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0})^{-2}\boldsymbol{\Gamma}^{\frac{1}{2}*}$ and $\mathbf{R}_2 = \frac{\sqrt{\lambda}}{\sqrt{n_2}}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0})^{-1}\boldsymbol{\Gamma}^{\frac{1}{2}*}$, then we have

$$\begin{aligned}
&\mathrm{E}_k\big\{\mathrm{tr}(\boldsymbol{\Xi}\lambda\boldsymbol{\tau}_{kx})\big\}^2\\
=&\mathrm{E}_k\Big[\mathrm{tr}\{\frac{1}{n_1}\boldsymbol{\Xi}(\mathbf{x}_k^*\mathbf{R}_1\mathbf{x}_k-\mathrm{tr}\mathbf{R}_1)\}\Big]^2+\mathrm{E}_k\Big[\mathrm{tr}\{\frac{1}{n_1}\boldsymbol{\Xi}\mathbf{V}_1^*\mathbf{Y}\mathbf{R}_2(\mathbf{x}_k\mathbf{x}_k^*-\mathbf{I}_p)\mathbf{R}_2^*\mathbf{Y}^*\mathbf{V}_1\}\Big]^2\\
&-2\mathrm{E}_k\Big[\mathrm{tr}\{\frac{1}{n_1}\boldsymbol{\Xi}(\mathbf{x}_k^*\mathbf{R}_1\mathbf{x}_k-\mathrm{tr}\mathbf{R}_1)\}\mathrm{tr}\{\frac{1}{n_1}\boldsymbol{\Xi}\mathbf{V}_1^*\mathbf{Y}\mathbf{R}_2(\mathbf{x}_k\mathbf{x}_k^*-\mathbf{I}_p)\mathbf{R}_2^*\mathbf{Y}^*\mathbf{V}_1\}\Big]
\end{aligned}$$

By equation (1.15) in Bai and Silverstein (2004), we can further show that

$$\begin{aligned}
&\mathrm{E}_k\Big[\mathrm{tr}\{\frac{1}{n_1}\boldsymbol{\Xi}(\mathbf{x}_k^*\mathbf{R}_1\mathbf{x}_k-\mathrm{tr}\mathbf{R}_1)\}\Big]^2 = \frac{1}{n_1^2}\{\mathrm{tr}(\boldsymbol{\Xi})\}^2\mathrm{E}_k(\mathbf{x}_k^*\mathbf{R}_1\mathbf{x}_k-\mathrm{tr}\mathbf{R}_1)^2\\
=&\frac{1}{n_1^2}\{\mathrm{tr}(\boldsymbol{\Xi})\}^2\Big\{(\mathrm{E}|x_{11}|^4-|\mathrm{E}x_{11}^2|^2-2)\sum_t r_{1,tt}^2+2\mathrm{tr}(\mathbf{R}_1^2)\Big\} = o(n_1^{-1}\log n_1),
\end{aligned}$$

where $r_{1,ts}$ is the (t,s)th element of the matrix $\mathbf{R}_1$.

Let $\boldsymbol{\Delta} = \mathbf{R}_2^*\mathbf{Y}^*\mathbf{V}_1\boldsymbol{\Xi}\mathbf{V}_1^*\mathbf{Y}\mathbf{R}_2$, then

$$\begin{aligned}
&\mathrm{E}_k\Big[\mathrm{tr}\{\frac{1}{n_1}\boldsymbol{\Xi}(\mathbf{x}_k^*\mathbf{R}_1\mathbf{x}_k-\mathrm{tr}\mathbf{R}_1)\}\mathrm{tr}\{\frac{1}{n_1}\boldsymbol{\Xi}\mathbf{V}_1^*\mathbf{Y}\mathbf{R}_2(\mathbf{x}_k\mathbf{x}_k^*-\mathbf{I}_p)\mathbf{R}_2^*\mathbf{Y}^*\mathbf{V}_1\}\Big]\\
=&\frac{1}{n_1^2}\mathrm{tr}(\boldsymbol{\Xi})(\mathbf{x}_k^*\mathbf{R}_1\mathbf{x}_k-\mathrm{tr}\mathbf{R}_1)(\mathbf{x}_k^*\boldsymbol{\Delta}\mathbf{x}_k-\mathrm{tr}\boldsymbol{\Delta})\\
=&\frac{1}{n_1^2}\mathrm{tr}(\boldsymbol{\Xi})\Big\{(\mathrm{E}|x_{11}|^4-|\mathrm{E}x_{11}^2|^2-2)\sum_t r_{1,tt}\Delta_{tt}+2\mathrm{tr}(\mathbf{R}_1\boldsymbol{\Delta})\Big\} = o(n_1^{-1}\log n_1),
\end{aligned}$$

where Δ_{ts} is the (t,s)th element of the matrix $\mathbf{\Delta}$.

Moreover,

$$\mathrm{E}_k\big[\mathrm{tr}\{\frac{1}{n_1}\mathbf{\Xi}\mathbf{V}_1^*\mathbf{Y}\mathbf{R}_2(\mathbf{x}_k\mathbf{x}_k^*-\mathbf{I}_p)\mathbf{R}_2^*\mathbf{Y}^*\mathbf{V}_1\}\big]^2=\frac{1}{n_1^2}(\mathbf{x}_k^*\mathbf{\Delta}\mathbf{x}_k-\mathrm{tr}\mathbf{\Delta})^2$$
$$=\frac{1}{n_1^2}\big\{(\mathrm{E}|x_{11}|^4-|\mathrm{E}x_{11}^2|^2-2)\sum_t\Delta_{tt}^2+2\mathrm{tr}(\mathbf{\Delta}^2)\big\}=o(n_1^{-1}\log n_1)+\frac{2}{n_1^2}\mathrm{tr}(\mathbf{\Delta}^2),$$

where $2n_1^{-2}\mathrm{tr}(\mathbf{\Delta}^2)$ is bounded as shown below. In fact, let $\mathbf{e}_t$ be a p-dimensional unit vector with the tth element equal to 1 and others equal to 0, we obtain

$$\frac{2}{n_1^2}\mathrm{tr}(\mathbf{\Delta}^2)=\frac{2}{n_1^2}\sum_{t,s}\Delta_{ts}^2=\frac{2}{n_1^2}\sum_{t,s}(\mathbf{e}_t^*\mathbf{R}_2^*\mathbf{Y}^*\mathbf{V}_1\mathbf{\Xi}\mathbf{V}_1^*\mathbf{Y}\mathbf{R}_2\mathbf{e}_s)^2=O(1).$$

Therefore,

$$\mathrm{E}_k\big\{\mathrm{tr}(\mathbf{\Xi}\lambda\boldsymbol{\tau}_{kx})\big\}^2=\frac{2}{n_1^2}\mathrm{tr}(\mathbf{\Delta}^2)+o(n_1^{-1}\log n_1).$$

Because $\mathbf{X}$ and $\mathbf{Z}$ satisfy Assumptions 1–5, the bounded $2n_1^{-2}\mathrm{tr}(\mathbf{\Delta}^2)$ is the same as $\mathbf{x}_k$ instead by $\mathbf{z}_k$. Furthermore, recall (S2.6), we have

$$\mathrm{E}e^{i\mathrm{tr}\big\{\mathbf{\Xi}\mathbf{\Omega}_{M,2}(\mathbf{X},\mathbf{Y})\big\}}-\mathrm{E}e^{i\mathrm{tr}\big\{\mathbf{\Xi}\mathbf{\Omega}_{M,2}(\mathbf{Z},\mathbf{Y})\big\}}=\sum_{k=1}^{n_1}\mathrm{E}e^{i\mathrm{tr}\big[\mathbf{\Xi}\{\mathbf{\Omega}_{M,2}(\mathbf{X}_{k0},\mathbf{Y})+\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}\beta_{k0}}\boldsymbol{\tau}_{k0}\}\big]}$$
$$\Bigg[\mathrm{E}_k e^{\frac{i\sqrt{c_2}}{\sqrt{n_2}\beta_{k0}}\mathrm{tr}\big\{\mathbf{\Xi}(\lambda\boldsymbol{\tau}_{kx}-\frac{\lambda\boldsymbol{\tau}_{k0}\varepsilon_{kx}}{\beta_{k0}})\big\}}-1-\frac{i\sqrt{c_2}}{\sqrt{n_2}\beta_{k0}}\mathrm{tr}\big\{\mathbf{\Xi}(\lambda\boldsymbol{\tau}_{kx}-\frac{\lambda\boldsymbol{\tau}_{k0}\varepsilon_{kx}}{\beta_{k0}})\big\}+\frac{c_2\mathrm{tr}\mathbf{\Delta}^2}{n_2\beta_{k0}^2n_1^2}$$
$$-\mathrm{E}_k e^{\frac{i\sqrt{c_2}}{\sqrt{n_2}\beta_{k0}}\mathrm{tr}\big\{\mathbf{\Xi}(\lambda\boldsymbol{\tau}_{kz}-\frac{\lambda\boldsymbol{\tau}_{k0}\varepsilon_{kz}}{\beta_{k0}})\big\}}+1+\frac{i\sqrt{c_2}}{\sqrt{n_2}\beta_{k0}}\mathrm{tr}\big\{\mathbf{\Xi}(\lambda\boldsymbol{\tau}_{kz}-\frac{\lambda\boldsymbol{\tau}_{k0}\varepsilon_{kz}}{\beta_{k0}})\big\}-\frac{c_2\mathrm{tr}\mathbf{\Delta}^2}{n_2\beta_{k0}^2n_1^2}\Bigg]+o(1)$$
$$\to 0.$$

Therefore (S2.1) holds for $\mathbf{\Omega}_{M,2}$.

Consider (S2.2) for $\mathbf{\Omega}_{M,2}$. Since we only care about the trace, and there exists an orthogonal $(p-M)\times n$ matrix $\mathbf{U}_{0n} = (\mathbf{U}_{0n-1}, \mathbf{U}_{01})$ such that it is equivalent to focus on $\tilde{\mathbf{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_k) - \tilde{\mathbf{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_{k0})$ as below when we study $\mathbf{\Omega}_{M,2}(\mathbf{Z},\mathbf{Y}_k) - \mathbf{\Omega}_{M,2}(\mathbf{Z},\mathbf{Y}_{k0})$.

$$\tilde{\mathbf{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_k) - \tilde{\mathbf{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_{k0}) = \frac{\sqrt{p}\lambda}{n_2}\Big[\big\{\mathrm{tr}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{ky})^{-1} - \mathrm{tr}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\big\}\mathbf{I}_M$$
$$-\big\{\mathbf{V}_1^*\mathbf{Y}_k\mathbf{U}_{0n}^*(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{ky})^{-1}\mathbf{U}_{0n}\mathbf{Y}_k^*\mathbf{V}_1 - \mathbf{V}_1^*\mathbf{Y}_{k0}\mathbf{U}_{0n-1}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{U}_{0n-1}^*\mathbf{Y}_{k0}^*\mathbf{V}_1\big\}\Big],$$

where

$$\tilde{\mathbf{F}}_{ky} = \frac{1}{n_1}(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_k\mathbf{Y}_k^*\mathbf{V}_2)^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}\mathbf{X}^*\mathbf{U}_2\mathbf{D}_2^{\frac{1}{2}}(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_k\mathbf{Y}_k^*\mathbf{V}_2)^{-\frac{1}{2}};$$
$$\tilde{\mathbf{F}}_{ky0} = \frac{1}{n_1}(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_{k0}\mathbf{Y}_{k0}^*\mathbf{V}_2)^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}\mathbf{X}^*\mathbf{U}_2\mathbf{D}_2^{\frac{1}{2}}(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_{k0}\mathbf{Y}_{k0}^*\mathbf{V}_2)^{-\frac{1}{2}}.$$

Since $\mathbf{Y}_k\mathbf{Y}_k^* = \mathbf{Y}_{k0}\mathbf{Y}_{k0}^* + \mathbf{y}_k\mathbf{y}_k^*$, by the (6.1.11) in Bai and Silverstein (2010), we have $\tilde{\mathbf{F}}_{ky} = \tilde{\mathbf{F}}_{ky0}/c_{ky}$, where

$$c_{ky} = 1 + \frac{1}{n_1}\mathbf{y}_k^*\mathbf{V}_2(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_{k0}\mathbf{Y}_{k0}^*\mathbf{V}_2)^{-1}\mathbf{V}_2^*\mathbf{y}_k; \tag{S2.7}$$
$$c_{k0} = 1 + \frac{1}{n_1}\mathrm{tr}(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_{k0}\mathbf{Y}_{k0}^*\mathbf{V}_2)^{-1} \text{ and } c_{ky} \to c_{k0}.$$

Then

$$\tilde{\mathbf{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_k) - \tilde{\mathbf{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_{k0}) \tag{S2.8}$$
$$=\frac{\sqrt{c_2}\lambda}{\sqrt{n_2}}\Big[\big\{\mathrm{tr}(\lambda\mathbf{I}-\frac{\tilde{\mathbf{F}}_{ky0}}{c_{ky}})^{-1} - \mathrm{tr}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{ky0})^{-1}\big\}\mathbf{I} - \mathbf{V}_1^*\mathbf{y}_k\mathbf{U}_{01}^*(\lambda\mathbf{I}-\frac{\tilde{\mathbf{F}}_{ky0}}{c_{ky}})^{-1}\mathbf{U}_{01}\mathbf{y}_k^*\mathbf{V}_1$$
$$-\mathbf{V}_1^*\mathbf{Y}_{k0}\mathbf{U}_{0n-1}^*(\lambda\mathbf{I}-\frac{\tilde{\mathbf{F}}_{ky0}}{c_{ky}})^{-1}\mathbf{U}_{01}\mathbf{y}_k^*\mathbf{V}_1 - \mathbf{V}_1^*\mathbf{y}_k\mathbf{U}_{01}^*(\lambda\mathbf{I}-\frac{\tilde{\mathbf{F}}_{ky0}}{c_{ky}})^{-1}\mathbf{U}_{0n-1}\mathbf{Y}_{k0}^*\mathbf{V}_1$$

$$-\mathbf{V}_1^*\mathbf{Y}_{k0}\mathbf{U}_{0n-1}^*\Big\{\big(\lambda\mathbf{I}-\frac{\tilde{\mathbf{F}}_{ky0}}{c_{ky}}\big)^{-1}-\big(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{ky0}\big)^{-1}\Big\}\mathbf{U}_{0n-1}\mathbf{Y}_{k0}^*\mathbf{V}_1\Big].$$

Similarly, $\tilde{\boldsymbol{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_{k-1})-\tilde{\boldsymbol{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y}_{k0})$ has the same expression as the one in (S2.8), except that $\mathbf{y}_k$ and c_{ky} are replaced by $\mathbf{w}_k$ and c_{kw}, where c_{kw} is similarly defined as c_{ky} with $\mathbf{w}_k$ instead of $\mathbf{w}_k$ and satisfies that $c_{kw}\to c_{k0}$, too. Because $\mathbf{y}_k$ and $\mathbf{w}_k$ have the same mean and variance, by the same techniques in (S2.5), it can be proved that

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\tilde{\boldsymbol{\Omega}}_{M,2}(\mathbf{Z},\mathbf{Y})\right\}}-\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\tilde{\boldsymbol{\Omega}}_{M,2}(\mathbf{Z},\mathbf{W})\right\}}\to 0.$$

Therefore, the equation (S2.2) holds for $\boldsymbol{\Omega}_{M,2}$.

Thirdly, for $\boldsymbol{\Omega}_{M,3}$, we consider the equation (S2.1) firstly. Since for each (k,k)th block of $\boldsymbol{\Omega}_{M,3}$, if we choose $\lambda=\psi_{n,k}$, then

$$[\boldsymbol{\Omega}_{M,3}(\lambda,\mathbf{X},\mathbf{Y})]_{[k,k]}=\Big[\frac{K_0\lambda}{\sqrt{n_1}}\Big\{\mathrm{tr}(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}})^{-1}\mathbf{I}-\mathbf{U}_1^*\mathbf{X}\big(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}\big)^{-1}\mathbf{X}^*\mathbf{U}_1\Big\}\Big]_{[k,k]},$$

where $\underline{\tilde{\mathbf{F}}}=\frac{1}{n_1}\mathbf{X}^*\mathbf{U}_2\mathbf{D}_2^{\frac{1}{2}}\mathbf{Q}^{-1}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}$. We find that the latter term

$$\tilde{\boldsymbol{\Omega}}_{M,3}(\lambda,\mathbf{X},\mathbf{Y})=\frac{\lambda}{\sqrt{n_1}}\Big\{\mathrm{tr}(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}})^{-1}\mathbf{I}-\mathbf{U}_1^*\mathbf{X}\big(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}\big)^{-1}\mathbf{X}^*\mathbf{U}_1\Big\}$$

to be in line with the conclusions of Theorem 3.2 in Jiang and Bai (2021), as well as we set the $\boldsymbol{\Gamma}=\mathbf{U}_2\mathbf{D}_2^{1/2}\mathbf{Q}^{-1}\mathbf{D}_2^{1/2}\mathbf{U}_2^*$ in their result. Thus, we get

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{X},\mathbf{Y})\right\}}-\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y})\right\}}\to 0.$$

Furthermore, the equation (S2.1) holds for each (k,k)th block of $\boldsymbol{\Omega}_{M,3}$ with $\lambda=\psi_{n,k}$.

We continue to prove the equation (S2.2) for $\boldsymbol{\Omega}_{M,3}$. As known that

$$\begin{aligned}&\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y}_k)-\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y}_{k0})\\ =&\frac{\lambda}{\sqrt{n_1}}\Big[\big\{\mathrm{tr}(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{ky})^{-1}-\mathrm{tr}(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{k0})^{-1}\big\}\mathbf{I}-\mathbf{U}_1^*\mathbf{X}\big\{(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{ky})^{-1}-(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{k0})^{-1}\big\}\mathbf{X}^*\mathbf{U}_1\Big],\end{aligned}$$

where $\underline{\tilde{\mathbf{F}}}_{ky}=1/n_1\mathbf{X}^*\mathbf{U}_2\mathbf{D}_2^{\frac{1}{2}}(\frac{1}{n_2}\mathbf{V}_2^*\mathbf{Y}_k\mathbf{Y}_k^*\mathbf{V}_2)^{-1}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}$. Since $\mathbf{Y}_k\mathbf{Y}_k^*=\mathbf{Y}_{k0}\mathbf{Y}_{k0}^*+\mathbf{y}_k\mathbf{y}_k^*$, we also have $\underline{\tilde{\mathbf{F}}}_{ky}=\underline{\tilde{\mathbf{F}}}_{k0}/c_{ky}$, where c_{ky} is defined in (S2.7). Then

$$\begin{aligned}&\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y}_k)-\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y}_{k0})\\ =&\frac{\lambda}{\sqrt{n_1}}\Big[\big\{\mathrm{tr}(\lambda\mathbf{I}-\frac{\underline{\tilde{\mathbf{F}}}_{k0}}{c_{ky}})^{-1}-\mathrm{tr}(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{k0})^{-1}\big\}\mathbf{I}-\mathbf{U}_1^*\mathbf{X}\big\{(\lambda\mathbf{I}-\frac{\underline{\tilde{\mathbf{F}}}_{k0}}{c_{ky}})^{-1}-(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{k0})^{-1}\big\}\mathbf{X}^*\mathbf{U}_1\Big].\end{aligned}\tag{S2.9}$$

Similarly,

$$\begin{aligned}&\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y}_{k-1})-\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y}_{k0})\\ =&\frac{\lambda}{\sqrt{n_1}}\Big[\big\{\mathrm{tr}(\lambda\mathbf{I}-\frac{\underline{\tilde{\mathbf{F}}}_{k0}}{c_{kw}})^{-1}-\mathrm{tr}(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{k0})^{-1}\big\}\mathbf{I}-\mathbf{U}_1^*\mathbf{X}\big\{(\lambda\mathbf{I}-\frac{\underline{\tilde{\mathbf{F}}}_{k0}}{c_{kw}})^{-1}-(\lambda\mathbf{I}-\underline{\tilde{\mathbf{F}}}_{k0})^{-1}\big\}\mathbf{X}^*\mathbf{U}_1\Big],\end{aligned}$$

where c_{kw} is similarly defined as c_{ky} with $\mathbf{y}_k$ replaced by $\mathbf{w}_k$ and satisfies that $c_{kw}\to c_{k0}$, too. Because $\mathbf{y}_k$ and $\mathbf{w}_k$ have the same mean and variance, by the same techniques of Theorem 3.2 in Jiang and Bai (2021), it can be proved that

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{Y})\right\}}-\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\tilde{\boldsymbol{\Omega}}_{M,3}(\mathbf{Z},\mathbf{W})\right\}}\to 0.$$

Therefore, the equation (S2.2) holds for each (k,k)th block of $\boldsymbol{\Omega}_{M,3}$ with $\lambda=\psi_{n,k}$.

For the fourth part,

$$
\begin{aligned}
&\boldsymbol{\Omega}_{M,4}(\mathbf{X}_k,\mathbf{Y})-\boldsymbol{\Omega}_{M,4}(\mathbf{X}_{k0},\mathbf{Y}) \qquad\qquad \text{(S2.10)}\\
=&\frac{\sqrt{p}}{n_1n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left(\lambda\mathbf{I}-\tilde{\mathbf{F}}_k\right)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\mathbf{X}_k^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\\
&-\frac{\sqrt{p}}{n_1n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0}\right)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_{k0}\mathbf{X}_{k0}^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\\
=&\frac{\sqrt{c_1}}{\sqrt{n_1}}\left\{\frac{1}{n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left(\lambda\mathbf{I}-\tilde{\mathbf{F}}_k\right)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\mathbf{X}_k^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\right.\\
&\left.-\frac{1}{n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0}\right)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\mathbf{X}_k^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\right\}\\
&+\frac{\sqrt{c_1}}{\sqrt{n_1}}\left\{\frac{1}{n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0}\right)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\mathbf{X}_k^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\right.\\
&\left.-\frac{1}{n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0}\right)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_{k0}\mathbf{X}_{k0}^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\right\}.
\end{aligned}
$$

Consider the former part in (S2.10), we recall the Lemma 6.9 in Bai and Silverstein (2010) and obtain

$$
\begin{aligned}
&\frac{\sqrt{c_1}}{\sqrt{n_1}n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left\{(\lambda\mathbf{I}-\tilde{\mathbf{F}}_k)^{-1}-(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\right\}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\mathbf{X}_k^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\\
=&\frac{\sqrt{c_1}}{\sqrt{n_1}\beta_{kx}}\mathbf{V}_1^*\mathbf{Y}\left\{\frac{1}{n_2}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0})^{-1}\right.\\
&\left.\cdot\frac{1}{n_1}\boldsymbol{\Gamma}^{\frac{1}{2}*}\mathbf{x}_k\mathbf{x}_k^*\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\right\}\mathbf{X}_k^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\\
=&\frac{\sqrt{c_1}}{\sqrt{n_1}\beta_{kx}}\mathrm{tr}\left\{\frac{1}{n_2}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0})^{-1}\right.\\
&\left.\cdot\frac{1}{n_1}\boldsymbol{\Gamma}^{\frac{1}{2}*}\mathbf{x}_k\mathbf{x}_k^*\boldsymbol{\Gamma}^{\frac{1}{2}}(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k_0})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{X}_k\right\}\mathrm{Cov}(\mathbf{V}_1^*\mathbf{Y},\mathbf{U}_1^*\mathbf{X}_k)\mathbf{D}_2^{\frac{1}{2}}+o(1)\\
\to&\mathbf{0}_M.
\end{aligned}
$$

By $\mathbf{X}_k\mathbf{X}_k^* = \mathbf{X}_{k0}\mathbf{X}_{k0}^* + \mathbf{x}_k\mathbf{x}_k^*$, the later part in (S2.10) is

$$\frac{\sqrt{c_1}}{\sqrt{n_1}}\left(\frac{1}{n_2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\left(\lambda\mathbf{I}-\tilde{\mathbf{F}}_{k0}\right)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^*\mathbf{x}_k\mathbf{x}_k^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\right) \to \mathbf{0}_M.$$

Therefore, for any $M \times M$ symmetric matrix $\boldsymbol{\Xi}$, we have

$$\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,4}(\mathbf{X},\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,4}(\mathbf{Z},\mathbf{Y})\right\}} = \sum_{k=1}^{n}\left[\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,4}(\mathbf{X}_k,\mathbf{Y})\right\}} - \mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,4}(\mathbf{X}_{k-1},\mathbf{Y})\right\}}\right]$$

$$= \sum_{k=1}^{n}\mathrm{E}e^{i\mathrm{tr}\left\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,4}(\mathbf{X}_{k0},\mathbf{Y})\right\}}\left[\mathrm{E}_k e^{i\mathrm{tr}\left[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,4}(\mathbf{X}_k,\mathbf{Y})-\boldsymbol{\Omega}_{M,4}(\mathbf{X}_{k0},\mathbf{Y})\}\right]} - \mathrm{E}_k e^{i\mathrm{tr}\left[\boldsymbol{\Xi}\{\boldsymbol{\Omega}_{M,4}(\mathbf{X}_{k-1},\mathbf{Y})-\boldsymbol{\Omega}_{M,4}(\mathbf{X}_{k0},\mathbf{Y})\}\right]}\right]$$

$$\to 0.$$

Apply the similar techniques in (S2.9) and (S2.10), we also have

$$\mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,4}(\mathbf{Z},\mathbf{Y})\}} - \mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,4}(\mathbf{Z},\mathbf{W})\}} \to 0.$$

The fifth, since the $\boldsymbol{\Omega}_{M,5}$ is the transpose of $\boldsymbol{\Omega}_{M,4}$, it automatically arrives at

$$\mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,5}(\mathbf{X},\mathbf{Y})\}} - \mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,5}(\mathbf{Z},\mathbf{Y})\}} \to 0; \mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,5}(\mathbf{Z},\mathbf{Y})\}} - \mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M,5}(\mathbf{Z},\mathbf{W})\}} \to 0.$$

To conclude, we obtain the equations (S2.1) and (S2.2) are valid for the $\boldsymbol{\Omega}_{M,j}, j = 1, 2, \cdots, 5$. In a word, we finally get

$$\mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M}(\mathbf{X},\mathbf{Y})\}} - \mathrm{E}e^{i\mathrm{tr}\{\boldsymbol{\Xi}\boldsymbol{\Omega}_{M}(\mathbf{Z},\mathbf{W})\}} \to 0$$

for any $M \times M$ symmetric matrix $\boldsymbol{\Xi}$. The invariance principle theorem for generalized spiked Fisher matrices is completed.

S3. Detailed derivations of $\mathbf{B}_1(l_{p,j})$ and $\mathbf{B}_2(l_{p,j})$.

By the formula $\mathbf{A}^{-1}-\mathbf{B}^{-1}=\mathbf{A}^{-1}(\mathbf{B}-\mathbf{A})\mathbf{B}^{-1}$ for any two invertible square matrices $\mathbf{A}$ and $\mathbf{B}$, we obtain that

$$\begin{aligned}
\mathbf{B}_1(l_{p,j}) =& \frac{\psi_{n,k}^2}{n_2^2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_1 \\
&-\frac{l_{p,j}^2}{n_2^2}\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(l_{p,j}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_1 \\
=& \frac{\psi_{n,k}^2}{n_2}\mathbf{V}_1^*\mathbf{Y}\Big[\frac{1}{n_2}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\Big\{(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}-(l_{p,j}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\Big\}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\Big]\mathbf{Y}^*\mathbf{V}_1 \\
&-\frac{l_{p,j}^2-\psi_{n,k}^2}{n_2}\mathbf{V}_1^*\mathbf{Y}\Big\{\frac{1}{n_2}\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(l_{p,j}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}\Big\}\mathbf{Y}^*\mathbf{V}_1 \\
=& \frac{1}{\sqrt{p}}\gamma_{kj}\psi_{n,k}^3\frac{1}{n_2}\mathrm{tr}\Big[(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\big\{(\psi_{n,k}+\frac{1}{\sqrt{p}}\gamma_{kj}\psi_{n,k})\mathbf{I}-\tilde{\mathbf{F}}\big\}^{-1}\Big]\mathbf{I}_M \\
&-\frac{1}{\sqrt{p}}\gamma_{kj}2\psi_{n,k}^2\frac{1}{n_2}\mathrm{tr}(l_{p,j}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{I}_M+o(\frac{\psi_{n,k}}{\sqrt{p}}) \\
=& \frac{1}{\sqrt{p}}\gamma_{kj}\Big\{c_2\psi_{n,k}^3m_2(\psi_{n,k})+2c_2\psi_{n,k}^2m(\psi_{n,k})\Big\}\mathbf{I}_M+o(\frac{\psi_{n,k}}{\sqrt{p}}); \quad \text{(S3.1)}
\end{aligned}$$

$$\begin{aligned}
\mathbf{B}_2(l_{p,j}) =& \frac{\psi_{n,k}}{n_1}\mathbf{D}_1^{\frac{1}{2}}\mathbf{U}_1^*\mathbf{X}\big(\psi_{n,k}\mathbf{I}-\underline{\tilde{\mathbf{F}}}\big)^{-1}\mathbf{X}^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}-\frac{l_{p,j}}{n_1}\mathbf{D}_1^{\frac{1}{2}}\mathbf{U}_1^*\mathbf{X}\big(l_{p,j}\mathbf{I}-\underline{\tilde{\mathbf{F}}}\big)^{-1}\mathbf{X}^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}} \\
=& \frac{\psi_{n,k}}{n_1}\mathbf{D}_1^{\frac{1}{2}}\mathbf{U}_1^*\mathbf{X}\Big\{(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-1}-(l_{p,j}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-1}\Big\}\mathbf{X}^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}} \\
&-\frac{l_{p,j}-\psi_{n,k}}{n_1}\mathbf{D}_1^{\frac{1}{2}}\mathbf{U}_1^*\mathbf{X}(l_{p,j}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-1}\mathbf{X}^*\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}} \\
=& \frac{1}{\sqrt{p}}\gamma_{kj}\psi_{n,k}^2\frac{1}{n_1}\mathrm{tr}\big\{(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-2}\big\}\mathbf{D}_1 \\
&-\frac{1}{\sqrt{p}}\gamma_{kj}\psi_{n,k}\frac{1}{n_1}\mathrm{tr}\big\{(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-1}\big\}\mathbf{D}_1+o(\frac{\psi_{n,k}}{\sqrt{p}}) \\
=& \frac{1}{\sqrt{p}}\gamma_{kj}\Big\{\psi_{n,k}^2\underline{m}_2(\psi_{n,k})+\psi_{n,k}\underline{m}(\psi_{n,k})\Big\}\mathbf{D}_1+o(\frac{\psi_{n,k}}{\sqrt{p}}). \qquad \text{(S3.2)}
\end{aligned}$$

The proof is completed.

S4. Proof of Theorem 2

Proof. As shown in Section 2 in article body, every sample spiked eigenvalue of $\mathbf{F}$, $l_{p,j}, j \in \mathcal{J}_k, k = 1, \cdots, K$, satisfies the equation (2.14). Furthermore, since $\psi_{n,k}$ satisfies the equation (2.13), it means that the population spiked eigenvalues α_u in the uth diagonal block of $\mathbf{D}_1$ makes $\psi_{n,k} + c_2\psi_{n,k}^2 m(\psi_{n,k}) + \psi_{n,k}\underline{m}(\psi_{n,k})\alpha_u$ keep away from 0, if $u \neq k$; and satisfies $\psi_{n,k} + c_2\psi_{n,k}^2 m(\psi_{n,k}) + \psi_{n,k}\underline{m}(\psi_{n,k})\alpha_k = 0$. For nonzero limit of spiked eigenvalue, $\psi_{n,k}$, each kth diagonal block of the equation (2.14) is multiplied $p^{1/4}$ by rows and columns, respectively. Then, by Lemma 4.1 in Bai, et al. (1991), it follows from (2.14) that

$$\Big|\gamma_{kj}\psi_{n,k}\big\{1+c_2\psi_{n,k}^2 m_2(\psi_{n,k})+2c_2\psi_{n,k}m(\psi_{n,k})+\alpha_k\psi_{n,k}\underline{m}_2(\psi_{n,k})+\alpha_k\underline{m}(\psi_{n,k})\big\}\mathbf{I}_{m_k}$$
$$+\psi_{n,k}\big[\mathbf{\Omega}_M(\psi_{n,k},\mathbf{X},\mathbf{Y})\big]_{kk} + o(\psi_{n,k})\Big| = 0, \tag{S4.1}$$

where $[\ \cdot\ ]_{kk}$ is the kth diagonal block of a matrix corresponding to the indices $\{i, j \in \mathcal{J}_k\}$. According to the Skorokhod strong representation in Skorokhod (1956); Hu and Bai (2014), it follows that the convergence of $\mathbf{\Omega}_M(\psi_{n,k}, \mathbf{X}, \mathbf{Y})$ and (2.14) can be achieved simultaneously in probability 1 by choosing an appropriate probability space.

Recall ϕ_k in (3.2), the equation (S4.1) arrives at $\left|\gamma_{kj}\phi_k\mathbf{I}_{m_k} + \left[\mathbf{\Omega}_{\psi_k}\right]_{kk} + o(1)\right| = 0$. Thus, it is obvious that γ_{kj} asymptotically satisfies that

$$\left|\gamma_{kj} \cdot \phi_k\mathbf{I}_{m_k} + \left[\mathbf{\Omega}_{\psi_k}\right]_{kk}\right| = 0, \tag{S4.2}$$

where $\mathbf{\Omega}_{\psi_k}$ is an $M \times M$ Hermitian matrix, being the limiting distribution of $\mathbf{\Omega}_M(\psi_{n,k}, \mathbf{X}, \mathbf{Y})$.

Therefore, by the equation (S4.2), the m_k-dimensional real vector $\{\gamma_{kj}, j \in \mathcal{J}_k\}$ converges weakly to the distribution of the m_k eigenvalues of the Gaussian random matrix $-\left[\mathbf{\Omega}_{\psi_k}\right]_{kk}/\phi_k$ for each distinct spiked eigenvalue. The distribution of $\mathbf{\Omega}_{\psi_k}$ is detailed in Corollary 1. Then, the CLT for each distinct spiked eigenvalue of a generalized spiked Fisher matrix is established. □

S5. Proof of Corollary 1

Proof. When Assumption 5 is valid, by Theorem 1, we can derive the limiting distribution of $\mathbf{\Omega}_M(\psi_{n,k}, \mathbf{X}, \mathbf{Y})$ under the Gaussian assumption of the entries from $\mathbf{X}$ and $\mathbf{Y}$. For $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2)$, $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2)$ defined in (2.5), let $\mathbf{U}_{s,\cdot i}$ and $\mathbf{V}_{s,\cdot i}$ be the ith column of $\mathbf{U}_s$ and $\mathbf{V}_s$, $s = 1, 2$, respectively. Denote $\boldsymbol{\xi}_{s,i} = \mathbf{X}^*\mathbf{U}_{s,\cdot i}$ and $\boldsymbol{\eta}_{s,i} = \mathbf{Y}^*\mathbf{V}_{s,\cdot i}$ where $s = 1, 2$; $i = 1, \cdots M$ for $s = 1$ and $i = M + 1, \cdots p$ for $s = 2$. Thus $\boldsymbol{\xi}_{s,i}, \boldsymbol{\eta}_{s,i}$ can be viewed as the independent random sample matrix with zero mean and identity covariance matrix under Gaussian assumption.

Let ω_{ij} be the (i,j)th element of $\mathbf{\Omega}_M$, then by the Lemma 2.7 in Bai and Silverstein (1998) and the expression of $\mathbf{\Omega}_M$, we have $\lim_{min(n_1,n_2,p)\to 0} \mathrm{E}(\omega_{ij}) \to 0$. The $\mathbf{\Omega}_{M,i}(\lambda, \mathbf{X}, \mathbf{Y}), i = 1, \cdots, 5$ defined in (2.12) are briefly denoted as $\mathbf{\Omega}_{M,i}$ if no confusion. Then we have

$$\mathrm{Cov}\big(\mathbf{\Omega}_M(\psi_{n,k}, \mathbf{X}, \mathbf{Y})\big) = \mathrm{Cov}\big(\mathbf{\Omega}_{M,1}\big) + \mathrm{Cov}\big(\mathbf{\Omega}_{M,2}\big) + 2\mathrm{Cov}\big(\mathbf{\Omega}_{M,1}, \mathbf{\Omega}_{M,2}\big) \tag{S5.1}$$

$$+\mathrm{Cov}\big(\mathbf{\Omega}_{M,3}\big) + \mathrm{Cov}\big(\mathbf{\Omega}_{M,4}\big) + \mathrm{Cov}\big(\mathbf{\Omega}_{M,5}\big) + \mathrm{Cov}\big(\mathbf{\Omega}_{M,4}, \mathbf{\Omega}_{M,5}\big) + \mathrm{Cov}\big(\mathbf{\Omega}_{M,5}, \mathbf{\Omega}_{M,4}\big)$$

For the first term in (S5.1), we have $\mathrm{Cov}\big(\mathbf{\Omega}_{M,1}\big) = c_2/n_2 \mathrm{Cov}(\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_1 - n_2\mathbf{I}_M)$. Then, by the equation (1.15) in Bai and Silverstein (2004), it is obvious that the variance of diagonal elements of $\mathbf{\Omega}_{M,1}$, denoted as $\omega_{1,jj}$, is

$$\mathrm{Var}(\omega_{1,jj}) = \frac{c_2}{n_2}\mathrm{E}(\boldsymbol{\eta}_{1,j}^*\mathbf{I}_{n_2}\boldsymbol{\eta}_{1,j} - n_2)^2 \to 2c_2. \tag{S5.2}$$

and the variance of off-diagonal elements of $\mathbf{\Omega}_{M,1}$, denoted as $\omega_{1,j_1j_2}, j_1 \neq j_2$, is

$$\mathrm{Var}(\omega_{1,j_1j_2}) = \frac{c_2}{n_2}\mathrm{E}(\boldsymbol{\eta}_{1,j_1}^*\mathbf{I}_{n_2}\boldsymbol{\eta}_{1,j_2} - n_2)^2 \to c_2, \qquad j_1 \neq j_2$$

for the real case. For the complex case, $\mathrm{Var}(\omega_{1,j_1j_2}) \to c_2$ for all j_1, j_2, including the case of $j_1 = j_2$.

Similar calculations as below are all based on the equation (1.15) in Bai and Silverstein (2004). The second term in (S5.1) is

$$\begin{aligned}&\mathrm{Cov}\big(\mathbf{\Omega}_{M,2}\big)\\=&\frac{c_2\psi_{n,k}^2}{n_2}\mathrm{Cov}\Big[\mathbf{V}_1{}^*\mathbf{Y}\frac{\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2{}^*\mathbf{Y}}{n_2}\mathbf{Y}^*\mathbf{V}_1-\mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{I}_M\Big]\end{aligned}$$

Then the variance of diagonal element of $\mathbf{\Omega}_{M,2}$, denoted as $\omega_{2,jj}$, is

$$\begin{aligned}&\mathrm{Var}(\omega_{2,jj})\\=&\frac{c_2\psi_{n,k}^2}{n_2}\mathrm{E}\Big[\boldsymbol{\eta}_{1,j}^*\frac{\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}}{n_2}\boldsymbol{\eta}_{1,j}-\mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\Big]^2\\=&2\frac{c_2}{n_2}\psi_k^2\mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-2}\to 2c_2^2\psi_k^2m_2(\psi_k)\end{aligned}\tag{S5.3}$$

and the one of off-diagonal element of $\mathbf{\Omega}_{M,2}$, denoted as ω_{2,j_1j_2}, $j_1\neq j_2$, is

$$\mathrm{Var}(\omega_{2,j_1j_2})\to c_2^2\psi_k^2m_2(\psi_k),\qquad j_1\neq j_2$$

for the real case. For the complex case, $\mathrm{Var}(\omega_{2,j_1j_2})=c_2^2\psi_k^2m_2(\psi_k)$ for all j_1,j_2, including the case of $j_1=j_2$.

The third term in (S5.1) is

$$\begin{aligned}&2\mathrm{Cov}\big(\mathbf{\Omega}_{M,1},\mathbf{\Omega}_{M,2}\big)=-\frac{2c_2\psi_{n,k}}{n_2}\mathrm{Cov}\Big(\big(\mathbf{V}_1^*\mathbf{Y}\mathbf{Y}^*\mathbf{V}_1-n_2\mathbf{I}_M\big),\\&\big[\mathbf{V}_1{}^*\mathbf{Y}\frac{\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2{}^*\mathbf{Y}}{n_2}\mathbf{Y}^*\mathbf{V}_1-\mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{I}_M\big]\Big).\end{aligned}$$

Then the diagonal element of $2\mathrm{Cov}(\mathbf{\Omega}_{M,1},\mathbf{\Omega}_{M,2})$ has the following variance

$$\mathrm{Var}(\omega_{12,jj})=-\frac{2c_2\psi_{n,k}}{n_2}\mathrm{E}\Big\{\big(\boldsymbol{\eta}_{1,j}^*\mathbf{I}_{n_2}\boldsymbol{\eta}_{1,j}-n_2\big)$$

$$\cdot\big[\boldsymbol{\eta}_{1,j}^{*}\frac{\mathbf{Y}^{*}\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^{*}\mathbf{Y}}{n_2}\boldsymbol{\eta}_{1,j}-\mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\big]\Big\}$$

$$=-\frac{4c_2\psi_{n,k}}{n_2}\mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\to 4c_2^2\psi_k m(\psi_k). \tag{S5.4}$$

For the off-diagonal element of $2\mathrm{Cov}(\boldsymbol{\Omega}_{M,1},\boldsymbol{\Omega}_{M,2})$, we have $\mathrm{Var}(\omega_{12,j_1j_2})\to 2c_2^2\psi_k m(\psi_k),\ j_1\neq j_2$ for the real case. For the complex case, $\mathrm{Var}(\omega_{12,j_1j_2})=2c_2^2\psi_k m(\psi_k)$ for all j_1,j_2, including the case of $j_1=j_2$.

The fourth term in (S5.1) is obtained that

$$\mathrm{Cov}\big(\boldsymbol{\Omega}_{M,3}\big)=\frac{c_1}{n_1}\mathrm{Cov}\Big(\mathbf{D}_1^{\frac{1}{2}}\mathbf{U}_1^{*}\mathbf{X}\big(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}}\big)^{-1}\mathbf{X}^{*}\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}-\mathrm{tr}(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-1}\mathbf{D}_1\Big).$$

Then the variance of diagonal elements of $\boldsymbol{\Omega}_{M,3}$, denoted as $\omega_{3,jj}$, is

$$\mathrm{Var}(\omega_{3,jj})=\frac{c_1}{n_1}\alpha_k^2\mathrm{E}\big\{\boldsymbol{\xi}_{1,j}^{*}(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-1}\boldsymbol{\xi}_{1,j}-\mathrm{tr}(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-1}\big\}^2$$

$$=2\frac{c_1}{n_1}\alpha_k^2\mathrm{tr}(\psi_{n,k}\mathbf{I}_{n_1}-\underline{\tilde{\mathbf{F}}})^{-2}\to 2c_1\alpha_k^2\underline{m}_2(\psi_k) \tag{S5.5}$$

and the one for the off-diagonal element of $\boldsymbol{\Omega}_{M,3}$ has the variance that $\mathrm{Var}(\omega_{3,j_1j_2})\to c_1\alpha_k^2\underline{m}_2(\psi_k), j_1\neq j_2$ for the real case. For the complex case, $\mathrm{Var}(\omega_{2,j_1j_2})=c_1\alpha_k^2\underline{m}_2(\psi_k)$ for all j_1,j_2, including the case of $j_1=j_2$.

Finally, consider the rest terms in (S5.1). Since $\boldsymbol{\Omega}_{M,5}$ is the transpose of $\boldsymbol{\Omega}_{M,4}$, then it holds for the diagonal elements of the rest terms that

$$\mathrm{Cov}\big(\boldsymbol{\Omega}_{M,4}\big)+\mathrm{Cov}\big(\boldsymbol{\Omega}_{M,5}\big)+\mathrm{Cov}\big(\boldsymbol{\Omega}_{M,4},\boldsymbol{\Omega}_{M,5}\big)+\mathrm{Cov}\big(\boldsymbol{\Omega}_{M,5},\boldsymbol{\Omega}_{M,4}\big)$$

$$=\frac{4c_2}{n_1^2n_2}\mathrm{Cov}\Big(\mathbf{V}_1^{*}\mathbf{Y}\mathbf{Y}^{*}\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}\big(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}}\big)^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{U}_2^{*}\mathbf{X}\mathbf{X}^{*}\mathbf{U}_1\mathbf{D}_1^{\frac{1}{2}}\Big)$$

Then, for the diagonal elements of rest part, denoted as $\omega_{4,jj}$, the variance is

$$\begin{aligned}\mathrm{Var}(\omega_{4,jj}) = 4\frac{c_2}{n_1^2 n_2}\mathrm{E}\Big\{&\boldsymbol{\eta}_{1,j}^* \mathbf{Y}^* \mathbf{V}_2 \mathbf{Q}^{-\frac{1}{2}} \big(\psi_{n,k}\mathbf{I} - \tilde{\mathbf{F}}\big)^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{D}_2^{\frac{1}{2}} \mathbf{U}_2^* \mathbf{X}\mathbf{X}^* \mathbf{U}_1 \mathbf{D}_1^{\frac{1}{2}} \\ &\cdot \mathbf{D}_1^{\frac{1}{2}} \mathbf{U}_1^* \mathbf{X}\mathbf{X}^* \mathbf{U}_2 \mathbf{D}_2^{\frac{1}{2}} \mathbf{Q}^{-\frac{1}{2}} \big(\psi_{n,k}\mathbf{I} - \tilde{\mathbf{F}}\big)^{-1} \mathbf{Q}^{-\frac{1}{2}} \mathbf{V}_2^* \mathbf{Y} \boldsymbol{\eta}_{1,j}\Big\} \\ = \frac{4c_2\alpha_k}{n_1}&\mathrm{tr}\big\{(\psi_{n,k}\mathbf{I} - \tilde{\mathbf{F}})^{-1}\tilde{\mathbf{F}}(\psi_{n,k}\mathbf{I} - \tilde{\mathbf{F}})^{-1}\big\} \to 4c_1c_2\alpha_k m_3(\psi_k) \end{aligned} \tag{S5.6}$$

and the variance of off-diagonal elements of the rest, denoted as $\omega_{4,j_1j_2}, j_1 \neq j_2$, is $\mathrm{Var}(\omega_{4,j_1j_2}) \to 2c_1c_2\alpha_k m_3(\psi_k),\ j_1 \in \mathcal{J}_{k_1} \neq j_2 \in \mathcal{J}_{k_2}$ for the real case. For the complex case, $\mathrm{Var}(\omega_{4,j_1j_2}) = 2c_1c_2\alpha_k m_3(\psi_k)$ for all j_1, j_2, including the case of $j_1 = j_2$.

Let

$$\theta_k = c_2 + c_2^2\psi_k^2 m_2(\psi_k) + 2c_2^2\psi_k m(\psi_k) + c_1\alpha_k^2 \underline{m}_2(\psi_k) + 2c_1c_2\alpha_k m_3(\psi_k).$$

Therefore, combining all the equations (S5.2)–(S5.6), it is concluded that $\boldsymbol{\Omega}_M(\psi_{n,k}, \mathbf{X}, \mathbf{Y})$ converges weakly to an $M \times M$ Hermitian matrix $\boldsymbol{\Omega}_{\psi_k}$, where $\theta_k^{-1/2}\,[\boldsymbol{\Omega}_{\psi_k}]_{kk}$ is GOE for the real case, with the entries above the diagonal being i.i.d. $\mathcal{N}(0,1)$ and the entries on the diagonal being i.i.d.$\mathcal{N}(0,2)$. For the complex case, the $\theta_k^{-1/2}\,[\boldsymbol{\Omega}_{\psi_k}]_{kk}$ is GUE, whose diagonal entries are i.i.d. real $\mathcal{N}(0,1)$, and the off diagonal entries are i.i.d. complex $\mathcal{CN}(0,1)$. □

S6. Proof of Remark 1

When Assumption 5 is invalid, but is weakened to the Assumption 5′, we reconsider each $\mathbf{\Omega}_{M,i}$ and find that only the first three terms, $\mathbf{\Omega}_{M,1}$, $\mathbf{\Omega}_{M,2}$, $\mathbf{\Omega}_{M,3}$, give rise to the additional items associated with the fourth moment.

The first item is about $\mathbf{\Omega}_{M,1}$. By the formula (1.15) of Bai and Silverstein (2004), for the variance of diagonal elements of $\mathbf{\Omega}_{M,1}$, we have

$$\mathrm{Var}(\omega_{1,jj}) = \frac{c_2}{n_2}\mathrm{E}(\boldsymbol{\eta}_{1,j}^*\mathbf{I}_{n_2}\boldsymbol{\eta}_{1,j} - n_2)^2 \to (1+q)c_2 + \beta_{y,jjjj}c_2.$$

For the variance of off-diagonal elements of $\mathbf{\Omega}_{M,1}$, $\mathrm{Var}(\omega_{1,j_1j_2}) \to c_2 + \beta_{y,j_1j_2j_1j_2}c_2$, $j_1 \neq j_2$. For the covariances, we have $\mathrm{Cov}(\omega_{1,i_1j_1}, \omega_{1,i_2j_2}) = \beta_{y,i_1j_1i_2j_2}c_2$, where $q = 1$ for real case and $q = 0$ for complex, and these $\beta_{y,\cdot}$'s are defined in Remark 1. So the first additive item involved with the fourth moment is $\beta_{y,\cdot}c_2$.

The second term is related to $\mathbf{\Omega}_{M,2}$. For the variance of diagonal elements of $\mathbf{\Omega}_{M,2}$, we have

$$\mathrm{Var}(\omega_{2,jj}) = \frac{c_2\psi_{n,k}^2}{n_2}\mathrm{E}\Big\{\boldsymbol{\eta}_{1,j}^*\frac{\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}}{n_2}\boldsymbol{\eta}_{1,j} - \mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\Big\}^2$$
$$\to (1+q)c_2^2\psi_k^2 m_2(\psi_k) + \beta_{y,jjjj}c_2^3\psi_k^2 m^2(\psi_k).$$

For the variance of off-diagonal elements of $\mathbf{\Omega}_{M,2}$, it arrives at $\mathrm{Var}(\omega_{2,j_1j_2}) \to c_2^2\psi_k^2 m_2(\psi_k) + \beta_{y,j_1j_2j_1j_2}c_2^3\psi_k^2 m^2(\psi_k)$, $j_1 \neq j_2$. For the covariances, we have

$\mathrm{Cov}(\omega_{2,i_1j_1},\omega_{2,i_2j_2}) = \beta_{y,i_1j_1i_2j_2}c_2^3\psi_k^2 m^2(\psi_k)$. So the second additive item involved with the fourth moment is $\beta_{y,\cdot}c_2^3\psi_k^2 m^2(\psi_k)$.

The third term is given by the covariance matrix $2\mathrm{Cov}(\mathbf{\Omega}_{M,1},\mathbf{\Omega}_{M,2})$. For its diagonal elements, we have

$$\begin{aligned}\mathrm{Var}(\omega_{12,jj}) =& -\frac{2c_2\psi_{n,k}}{n_2}\mathrm{E}\Big\{\big(\boldsymbol{\eta}_{1,j}^*\mathbf{I}_{n_2}\boldsymbol{\eta}_{1,j}-n_2\big)\\ &\cdot\big[\boldsymbol{\eta}_{1,j}^*\frac{\mathbf{Y}^*\mathbf{V}_2\mathbf{Q}^{-\frac{1}{2}}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\mathbf{Q}^{-\frac{1}{2}}\mathbf{V}_2^*\mathbf{Y}}{n_2}\boldsymbol{\eta}_{1,j}-\mathrm{tr}(\psi_{n,k}\mathbf{I}-\tilde{\mathbf{F}})^{-1}\big]\Big\}\\ \to& 2(1+q)c_2^2\psi_k m(\psi_k)+\beta_{y,jjjj}2c_2^2\psi_k m(\psi_k).\end{aligned}$$

For the off-diagonal elements of $2\mathrm{Cov}(\mathbf{\Omega}_{M,1},\mathbf{\Omega}_{M,2})$, it follows that $\mathrm{Var}(\omega_{12,j_1j_2}) \to 2c_2^2\psi_k m(\psi_k)+\beta_{y,j_1j_2j_1j_2}2c_2^2\psi_k m(\psi_k)$, when $j_1\neq j_2$. For its covariances, $\mathrm{Cov}(\omega_{12,i_1j_1},\omega_{12,i_2j_2}) = \beta_{y,i_1j_1i_2j_2}2c_2^2\psi_k m(\psi_k)$. The third additive item involved with the fourth moment is $\beta_{y,\cdot}2c_2^2\psi_k m(\psi_k)$.

The fourth term comes from $\mathbf{\Omega}_{M,3}$. From the detailed proof of Corollary 3.1 in Jiang and Bai (2021), we may obtain the following results. For the variance of diagonal elements of $\mathbf{\Omega}_{M,3}$,

$$\mathrm{Var}(\omega_{3,jj}) \to (1+q)c_1\alpha_k^2\underline{m}_2(\psi_k)+\beta_{x,jjjj}c_1\alpha_k^2\underline{m}^2(\psi_k).$$

For the variance of off-diagonal elements of $\mathbf{\Omega}_{M,3}$,

$$\mathrm{Var}(\omega_{3,j_1j_2}) \to c_1\alpha_k^2\underline{m}_2(\psi_k)+\beta_{x,j_1j_2j_1j_2}\cdot c_1\alpha_k^2\underline{m}^2(\psi_k),\ j_1\neq j_2.$$

For its covariances,

$$\mathrm{Cov}(\omega_{3,i_1j_1}, \omega_{3,i_2j_2}) = \beta_{x,i_1j_1i_2j_2} c_1 \alpha_k^2 \underline{m}^2(\psi_k).$$

So the fourth additive item involved with the fourth moment is $\beta_{x,\cdot} c_1 \alpha_k^2 \underline{m}^2(\psi_k)$.

To sum up, the whole additive item involved with the fourth moment is

$$\beta_{x,\cdot}\nu_1 + \beta_{y,\cdot}\nu_2,$$

where $\nu_1 = c_1\alpha_k^2\underline{m}^2(\psi_k)$ and $\nu_2 = c_2 + 2c_2^2\psi_k m(\psi_k) + c_2^3\psi_k^2 m^2(\psi_k) = c_2\big(1 + c_2\psi_k m(\psi_k)\big)^2$. Then, under the assumptions of Remark 1, the limiting distribution of $\mathbf{\Omega}_M(\psi_{n,k}, \mathbf{X}, \mathbf{Y})$ turns to an $M \times M$ Hermitian matrix $\mathbf{\Omega}_{\psi_k} = (\omega_{ij})$, which has the independent Gaussian entries of mean zero and variance

$$\mathrm{Cov}(\omega_{i_1,j_1}, \omega_{i_2,j_2}) = \begin{cases} (q+1)\theta_k + \beta_{x,iiii}\nu_1 + \beta_{y,iiii}\nu_2, & i_1 = j_1 = i_2 = j_2 = i; \\ \theta_k + \beta_{x,ijij}\nu_1 + \beta_{y,ijij}\nu_2, & i_1 = i_2 = i \neq j_1 = j_2 = j; \\ \beta_{x,i_1j_1i_2j_2}\nu_1 + \beta_{y,i_1j_1i_2j_2}\nu_2, & \text{other cases}, \end{cases}$$

where θ_k is defined in (2.15). Thus, the conclusion of Remark 1 is obtained.

References

Bai, Z. D., Miao, B.Q. and Rao, C. Radbakrisbna. (1991). Estimation of directions of arrival of signals: Asymptotic results. *Advances in Spectrum Analysis and Array Processing*, **I**, 327–347.

Bai, Z. D. and Silverstein, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *The Annals of Probability*, **26**, 316–345.

Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probability*, **32**, 553–605.

Bai, Z. D. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices.* Springer Series in Statistics, Springer-Verlag, New York, ISSN: 0172–7397.

Hu, J. and Bai, Z. D. (2014). Estimation of directions of arrival of signals: Asymptotic results. *Science China Mathematics*, **57** (11), DOI: 10.1007/s11425-014-4855-6.

Jiang, D. D. and Bai, Z. D. (2021). Generalized Four Moment Theorem and an Application to CLT for Spiked Eigenvalues of Large-dimensional Covariance Matrices. *Bernoulli*, **27(1)**, 274–294.

Jiang, D. D, Hou, Z. Q. and Hu, J. (2021). The limits of the sample spiked eigenvalues for a high-dimensional generalized Fisher matrix and its applications. *Journal of Statistical Planning and Inference*, **3**, 208–217.

Skorokhod, A. V. (1956). Limit theorems for stochastic processes. *Theory Probab Appl.*, **1**, 261–290.