Submitted to the Annals of Statistics

SUPPLEMENT TO "ASYMPTOTIC NORMALITY FOR EIGENVALUE STATISTICS OF A GENERAL SAMPLE COVARIANCE MATRIX WHEN $P/N \to \infty$ AND APPLICATIONS"

BY JIAXIN QIU[1,a], ZENG LI[2,b] AND JIANFENG YAO[3,c]

[1]Department of Statistics and Actuarial Science, The University of Hong Kong, [a]qiujx@connect.hku.hk

[2]Department of Statistics and Data Science, Southern University of Science and Technology, [b]liz9@sustech.edu.cn

[3]School of Data Science, The Chinese University of Hong Kong (Shenzhen), [c]jeffyao@cuhk.edu.cn

CONTENTS

This supplementary document contains some technical lemmas and their proofs, proofs of Lemmas 6.1 – 6.2, equations (15), (35), (36), Theorem 4.2, and Proposition 4.1. We also report some additional simulation results in this document.

S1. Some technical lemmas.

LEMMA S1.1 (Bai and Silverstein (2010), Lemma B.26). *Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ nonrandom matrix and $\mathbf{x} = (X_1, \ldots, X_n)'$ be a random vector of independent entries. Assume that $\mathbb{E}X_i = 0$, $\mathbb{E}|X_i|^2 = 1$ and $\mathbb{E}|X_i|^\ell \leqslant \nu_\ell$. Then, for any $k \geqslant 1$,*

$$\mathbb{E}|\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathrm{tr}\mathbf{A}|^k \leqslant C_k \Big[\big\{ \nu_4 \mathrm{tr}(\mathbf{A}\mathbf{A}^*) \big\}^{k/2} + \nu_{2k} \mathrm{tr}(\mathbf{A}\mathbf{A}^*)^{k/2} \Big],$$

where C_k is a constant depending on k only.

LEMMA S1.2 (Pan and Zhou (2011), Lemma 5). *Let* $\mathbf{A}$ *be a* $p \times p$ *deterministic complex matrix with zero diagonal elements. Let* $\mathbf{x} = (X_1, \ldots, X_p)'$ *be a random vector with i.i.d. real entries. Assume that* $\mathbb{E}X_i = 0$, $\mathbb{E}|X_i|^2 = 1$. *Then, for any* $k \geqslant 2$,

$$\mathbb{E}|\mathbf{x}'\mathbf{A}\mathbf{x}|^k \leqslant C_k \big\{\mathbb{E}|X_1|^k\big\}^2 \big(\mathrm{tr}\mathbf{A}\mathbf{A}^*\big)^{k/2}, \tag{S1.1}$$

where C_k *is a constant depending on* k *only.*

LEMMA S1.3 (Burkholder's inequality, Burkholder (1973)). *Let* $\{X_i\}$ *be a complex martingale difference sequence withe respect to the increasing* σ*-field* $\{\mathcal{F}_i\}$. *Then for* $k \geqslant 2$, *the following inequality*

$$\mathbb{E}\Big|\sum_i X_i\Big|^k \leqslant C_k \mathbb{E}\Big[\sum_i \mathbb{E}\big\{|X_i|^2 \mid \mathcal{F}_{i-1}\big\}\Big]^{k/2} + C_k \mathbb{E}\sum_i |X_i|^k$$

holds, where C_k *is a constant depending on* k *only.*

LEMMA S1.4 (Martingale CLT, Billingsley (2008)). *Suppose for each* n, $\{Y_{nk}\}_{1\leqslant k\leqslant r_n}$ *is a real martingale difference sequence with respect to the* σ*-field* $\{\mathcal{F}_{nk}\}$ *having second moments. If as* $n \to \infty$,

$$\sum_{k=1}^{r_n} \mathbb{E}(Y_{nk}^2 \mid \mathcal{F}_{n,k-1}) \xrightarrow{p} \sigma^2,$$

where σ^2 *is a positive constant, and for each* $\varepsilon > 0$,

$$\sum_{k=1}^{r_n} \mathbb{E}\Big(Y_{nk}^2 \mathbb{1}_{\{|Y_{nk}|\geqslant\varepsilon\}}\Big) \to 0,$$

then

$$\sum_{k=1}^{r_n} Y_{nk} \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

LEMMA S1.5 (Billingsley (1968), Theorem 12.3). *The sequence* $\{X_n\}$ *is tight if it satisfies these two conditions:*

(i) *The sequence* $\{X_n(0)\}$ *is tight.*
(ii) *There exist constants* $\gamma \geqslant 0$ *and* $\alpha > 1$ *and a non-decreasing, continuous function* F *on* $[0,1]$ *such that*

$$\mathbb{P}\Big(\big|X_n(t_2) - X_n(t_1)\big| \geqslant \lambda\Big) \leqslant \frac{1}{\lambda^\gamma}\big|F(t_2) - F(t_1)\big|^\alpha$$

holds for all t_1, t_2 *and* n *and all positive* λ.

LEMMA S1.6 (Bai and Silverstein (2004), Lemma 2.3). *Let* f *be analytic in* D, *a connected open set of* $\mathbb{C}$, *satisfying* $|f(z)| \leqslant M$ *for any* $z \in D$, *then, on any set bounded by a contour interior to* D, $f'(z)$ *is bounded.*

Lemma S1.7 is about the asymptotic expression of $\mathbb{Z}_k$, which is used in Section S2.16 to derive the finite-dimensional convergence of $M_n^{(1)}(z)$.

LEMMA S1.7. *For $z_1, z_2 \in \mathbb{C}^+$,*

$$\mathbb{Z}_k := \frac{1}{n(pb_p)^2}\mathrm{tr}\Big\{\mathbb{E}_k\mathbf{M}_k^{(1)}(z_1)\cdot\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\Big\} = \frac{\frac{k}{n}m(z_1)m(z_2)}{1-\frac{k}{n}m(z_1)m(z_2)} + o_{L_1}(1).$$

Lemmas S1.8 and S1.9 are used in the proof of Lemma S1.7.

LEMMA S1.8. *For $\vartheta_i(z)$ and $\zeta_i(z)$ defined in Lemma S1.7, we have*

$$\mathbb{E}\Big|\vartheta_i(z) - \frac{m(z)}{z}\Big|^4 \to 0, \qquad \mathbb{E}\Big|\zeta_i(z) + zm(z)\Big|^4 \to 0, \qquad \text{as } n\to\infty.$$

LEMMA S1.9. *Let $\mathbf{B}$ be any matrix independent of $\mathbf{x}_i$.*

$$\mathbb{E}\big|\mathbf{x}_i'\mathbf{M}_{ki}\mathbf{B}\mathbf{x}_i\big|^2 \leqslant Kp^2n^2\mathbb{E}\|\mathbf{B}\|^2, \tag{S1.2}$$

$$\mathbb{E}\big|\mathbf{x}_i'\boldsymbol{\Sigma}_p\mathbf{M}_{ki}\mathbf{B}\boldsymbol{\Sigma}_p\mathbf{x}_i\big|^2 \leqslant Kp^2n^2\mathbb{E}\|\mathbf{B}\|^2. \tag{S1.3}$$

Lemmas S1.10 and S1.11 are used in Sections S2.16 and 6.4.

LEMMA S1.10. *For $z\in\mathbb{C}_1$, we have*

$$|\beta_k(z)| \leqslant 1/v_1, \qquad |\beta_k^{\mathrm{tr}}(z)| \leqslant 1/v_1,$$

$$\Big|1+\frac{1}{npb_p}\mathrm{tr}\mathbf{M}_k^{(s)}(z)\Big| \leqslant 1+\frac{1}{v_1^s}, \qquad s=1,2,$$

$$\big|\beta_k\{1+\mathbf{q}_k'\mathbf{D}_k^2(z)\mathbf{q}_k\}\big| \leqslant \frac{1}{v_1}. \tag{S1.4}$$

LEMMA S1.11. *Under the assumption $p\wedge n\to\infty,\ p/n\to\infty$ and truncation, for $z\in\mathbb{C}_1$,*

$$\mathbb{E}|\gamma_{ks}|^2 \leqslant \frac{K}{n}, \qquad \mathbb{E}|\gamma_{ks}|^4 \leqslant K\left(\frac{1}{n^2}+\frac{n}{p^2}\right),$$

$$\mathbb{E}|\eta_k|^2 \leqslant \frac{K}{n}, \qquad \mathbb{E}|\eta_k|^4 \leqslant K\frac{\delta_n^4}{n} + K\left(\frac{1}{n^2}+\frac{n}{p^2}\right).$$

Lemmas S1.12 , S1.13 and S1.14 are used in Section 6.5 to derive the convergence of the non-random part $M_n^{(2)}(z)$. They are proved following the strategy in Bao (2015).

LEMMA S1.12. *Under the assumption $p\wedge n\to\infty,\ p/n\to\infty$, for $z\in\mathbb{C}_1$, we have*

$$\mathrm{Var}(m_n) = O\left(\frac{1}{n^2}\right). \tag{S1.5}$$

LEMMA S1.13. *Under the assumption $p\wedge n\to\infty,\ p/n\to\infty$, for $z\in\mathbb{C}_1$ and $1\leqslant\ell\leqslant n$,*

$$\mathbb{E}\Big|D_{\ell\ell}+\frac{1}{z+\mathbb{E}m_n}\Big|^2 = O\left(\frac{1}{n}\right)+O\left(\frac{n}{p}\right).$$

LEMMA S1.14. *Under the assumption $p\wedge n\to\infty,\ p/n\to\infty$, for $z\in\mathbb{C}_1$ and $1\leqslant\ell\leqslant p$,*

$$\mathbb{E}\Big|\widetilde{D}_{\ell\ell}+\frac{1}{a_p\sqrt{p/(nb_p)}+z+\mathbb{E}m_n}\Big|^2 = O\left(\left(\frac{n}{p}\right)^3\right)+O\left(\frac{n}{p^2}\right), \tag{S1.6}$$

where $D_{\ell\ell}$ is the ℓ-th diagonal entry of the matrix

$$\widetilde{\mathbf{D}} = (\widetilde{D}_{ij})_{p\times p} = \left(\mathbf{\Sigma}_p^{1/2}\mathbf{Y}\mathbf{Y}'\mathbf{\Sigma}_p^{1/2} - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_p - z\mathbf{I}_p\right)^{-1}.$$

Lemmas S1.15 and S1.16 provide some derivatives of F_{jk} and $\widehat{F}_{jk}$ for applying the generalized Stein's equation in Section 6.5.

Recall that

$$\mathbf{D} := (\mathbf{A}_n - z\mathbf{I}_n)^{-1}, \qquad \mathbf{E} := \mathbf{\Sigma}_p\mathbf{Y}\mathbf{D}\mathbf{Y}'\mathbf{\Sigma}_p = (E_{ij})_{p\times p}, \qquad \mathbf{F} := \mathbf{\Sigma}_p\mathbf{Y}\mathbf{D} = (F_{ij})_{p\times n}.$$

LEMMA S1.15 (Derivatives of F_{jk}).

$$\frac{\partial F_{jk}}{\partial Y_{jk}} = (\mathbf{\Sigma}_p)_{jj}D_{kk} - E_{jj}D_{kk} - F_{jk}^2;$$

$$\frac{\partial^2 F_{jk}}{\partial Y_{jk}^2} = -6(\mathbf{\Sigma}_p)_{jj}F_{jk}D_{kk} + 6E_{jj}F_{jk}D_{kk} + 2F_{jk}^3;$$

$$\frac{\partial^3 F_{jk}}{\partial Y_{jk}^3} = -6(\mathbf{\Sigma}_p)_{jj}^2D_{kk}^2 + 36(\mathbf{\Sigma}_p)_{jj}F_{jk}^2D_{kk} + 12(\mathbf{\Sigma}_p)_{jj}E_{jj}D_{kk}^2 - 36E_{jj}F_{jk}^2D_{kk} - 6E_{jj}^2D_{kk}^2 - 6F_{jk}^4;$$

$$\begin{aligned}\frac{\partial^4 F_{jk}}{\partial Y_{jk}^4} &= 120(\mathbf{\Sigma}_p)_{jj}^2F_{jk}D_{kk}^2 - 240(\mathbf{\Sigma}_p)_{jj}F_{jk}^3D_{kk} - 240(\mathbf{\Sigma}_p)_{jj}E_{jj}F_{jk}D_{kk}^2 + 240E_{jj}F_{jk}^3D_{kk} \\ &\quad + 120E_{jj}^2F_{jk}D_{kk}^2 + 24F_{jk}^5;\end{aligned}$$

$$\begin{aligned}\frac{\partial^5 F_{jk}}{\partial Y_{jk}^5} &= -120F_{jk}^6 - 1800E_{jj}F_{jk}^4D_{kk} - 1800E_{jj}^2F_{jk}^2D_{kk}^2 - 120E_{jj}^3D_{kk}^3 + 1800(\mathbf{\Sigma}_p)_{jj}F_{jk}^4D_{kk} \\ &\quad + 3600(\mathbf{\Sigma}_p)_{jj}E_{jj}F_{jk}^2D_{kk}^2 + 360(\mathbf{\Sigma}_p)_{jj}E_{jj}^2D_{kk}^3 - 1800(\mathbf{\Sigma}_p)_{jj}^2F_{jk}^2D_{kk}^2 \\ &\quad - 360(\mathbf{\Sigma}_p)_{jj}^2E_{jj}D_{kk}^3 + 120(\mathbf{\Sigma}_p)_{jj}^3D_{kk}^3.\end{aligned}$$

Recall that

$$\widehat{\mathbf{E}} = \mathbf{\Sigma}_p\mathbf{Y}\mathbf{D}\mathbf{Y}'\mathbf{\Sigma}_p^2 = (\widehat{E}_{ij})_{p\times p}, \qquad \widehat{\mathbf{F}} := \mathbf{\Sigma}_p^2\mathbf{Y}\mathbf{D} = (\widehat{F}_{ij})_{p\times n}.$$

LEMMA S1.16 (Derivatives of $\widehat{F}_{jk}$).

$$\frac{\partial \widehat{F}_{jk}}{\partial Y_{jk}} = (\mathbf{\Sigma}_p^2)_{jj}D_{kk} - \widehat{E}_{jj}D_{kk} - F_{jk}\widehat{F}_{jk};$$

$$\frac{\partial^2 \widehat{F}_{jk}}{\partial Y_{jk}^2} = -2(\mathbf{\Sigma}_p)_{jj}\widehat{F}_{jk}D_{kk} - 4(\mathbf{\Sigma}_p^2)_{jj}F_{jk}D_{kk} + 2F_{jk}^2\widehat{F}_{jk} + 4\widehat{E}_{jj}F_{jk}D_{kk} + 2E_{jj}\widehat{F}_{jk}D_{kk};$$

$$\begin{aligned}\frac{\partial^3 \widehat{F}_{jk}}{\partial Y_{jk}^3} &= -6(\mathbf{\Sigma}_p^2)_{jj}(\mathbf{\Sigma}_p)_{jj}D_{kk}^2 - 6F_{jk}^3\widehat{F}_{jk} - 18\widehat{E}_{jj}F_{jk}^2D_{kk} - 18E_{jj}F_{jk}\widehat{F}_{jk}D_{kk} - 6E_{jj}\widehat{E}_{jj}D_{kk}^2 \\ &\quad + 18(\mathbf{\Sigma}_p)_{jj}F_{jk}\widehat{F}_{jk}D_{kk} + 6(\mathbf{\Sigma}_p)_{jj}\widehat{E}_{jj}D_{kk}^2 + 18(\mathbf{\Sigma}_p^2)_{jj}F_{jk}^2D_{kk} + 6(\mathbf{\Sigma}_p^2)_{jj}E_{jj}D_{kk}^2;\end{aligned}$$

$$\begin{aligned}\frac{\partial^4 \widehat{F}_{jk}}{\partial Y_{jk}^4} &= 24F_{jk}^4\widehat{F}_{jk} + 96\widehat{E}_{jj}F_{jk}^3D_{kk} + 144E_{jj}F_{jk}^2\widehat{F}_{jk}D_{kk} + 96E_{jj}\widehat{E}_{jj}F_{jk}D_{kk}^2 + 24E_{jj}^2\widehat{F}_{jk}D_{kk}^2 \\ &\quad - 144(\mathbf{\Sigma}_p)_{jj}F_{jk}^2\widehat{F}_{jk}D_{kk} - 96(\mathbf{\Sigma}_p)_{jj}\widehat{E}_{jj}F_{jk}D_{kk}^2 - 48(\mathbf{\Sigma}_p)_{jj}E_{jj}\widehat{F}_{jk}D_{kk}^2 + 24(\mathbf{\Sigma}_p)_{jj}^2\widehat{F}_{jk}D_{kk}^2 \\ &\quad - 96(\mathbf{\Sigma}_p^2)_{jj}F_{jk}^3D_{kk} - 96(\mathbf{\Sigma}_p^2)_{jj}E_{jj}F_{jk}D_{kk}^2 + 96(\mathbf{\Sigma}_p)_{jj}(\mathbf{\Sigma}_p^2)_{jj}F_{jk}D_{kk}^2.\end{aligned}$$

Lemmas S1.17 and S1.18 provide the derivatives of some quantities with respect to Y_{jk}, which can be used to obtain the derivatives of F_{jk} (Lemma S1.15) and $\widehat{F}_{jk}$ (Lemma S1.16).

LEMMA S1.17. *For any $\alpha, j \in \{1,2,\ldots,p\}$ and $\beta, k \in \{1,2,\ldots,n\}$, we have*

$$\frac{\partial D_{\alpha\beta}}{\partial Y_{jk}} = -F_{j\alpha}D_{\beta k} - F_{j\beta}D_{\alpha k};$$

$$\frac{\partial F_{\alpha\beta}}{\partial Y_{jk}} = (\mathbf{\Sigma}_p)_{\alpha j}D_{k\beta} - E_{j\alpha}D_{\beta k} - F_{j\beta}F_{\alpha k};$$

$$\frac{\partial (E_{jj}D_{kk})}{\partial Y_{jk}} = 2(\mathbf{\Sigma}_p)_{jj}F_{jk}D_{kk} - 4E_{jj}F_{jk}D_{kk}.$$

LEMMA S1.18. *For any $\alpha, j \in \{1,2,\ldots,p\}$ and $\beta, k \in \{1,2,\ldots,n\}$, we have*

$$\frac{\partial \widehat{F}_{\alpha\beta}}{\partial Y_{jk}} = (\mathbf{\Sigma}_p^2)_{\alpha j}D_{k\beta} - \widehat{E}_{j\alpha}D_{\beta k} - F_{j\beta}\widehat{F}_{\alpha k};$$

$$\frac{\partial (\widehat{E}_{jj}D_{kk})}{\partial Y_{jk}} = (\mathbf{\Sigma}_p)_{jj}\widehat{F}_{jk}D_{kk} + (\mathbf{\Sigma}_p^2)_{jj}F_{jk}D_{kk} - E_{jj}\widehat{F}_{jk}D_{kk} - 3\widehat{E}_{jj}F_{jk}D_{kk}.$$

Lemmas S1.19 and S1.20 are used in Section S3 to prove equations (35) and (36).

LEMMA S1.19. *For $z \in \mathcal{C}_\ell \cup \mathcal{C}_r$, we have*

$$|\beta_k|\mathbb{1}_{U_n} \leqslant K, \qquad |\varepsilon_k| \leqslant K, \qquad \mathbb{E}|\gamma_{k2}|^4\mathbb{1}_{U_n} = O(n^{-2}), \qquad \mathbb{E}|\mu_k|^4\mathbb{1}_{U_n} = O(n^{-1}).$$

LEMMA S1.20. *For $z \in \mathcal{C}_\ell \cup \mathcal{C}_r$, we have*

$$\mathbb{E}\left|M_n^{(1)}(z)\mathbb{1}_{U_n}\right|^2 \leqslant K.$$

Lemma S1.21 is used in the proof of Lemma 6.2.

LEMMA S1.21. *Under the assumption $p \wedge n \to \infty$, $p/n \to \infty$, for $z \in \mathbb{C}_1$, we have*

$$\mathbb{E}\left|\frac{1}{npb_p}\mathrm{tr}\left\{\mathbf{M}_k^{(1)}(z)\right\} - m(z)\right|^2 \leqslant \frac{Kn}{p} + \frac{K}{n^2}.$$

S2. Proofs of lemmas. This section contains proofs of Lemmas S1.7 – S1.21, Lemmas 6.1 – 6.2.

S2.1. *Proof of Lemma S1.7.*

PROOF. Let $\{\mathbf{e}_i, i = 1\ldots, k-1, k+1, \ldots, n\}$ be the $(n-1)$-dimensional unit vectors with the i-th (or $(i-1)$-th) element equal to one and the remaining equal to zero according as $i < k$ (or $i > k$). Write $\mathbf{X}_k = \mathbf{X}_{ki} + \mathbf{x}_i \mathbf{e}_i'$. Let $\mathbf{I}_{(i)}$ be $n \times n$ diagonal matrix with all 1's on the diagonal except the i-th element being zero, and

$$\mathbf{D}_{ki,r}^{-1} = \mathbf{D}_k^{-1} - \mathbf{e}_i \mathbf{h}_i' = \frac{1}{\sqrt{npb_p}}\Big(\mathbf{X}_{ki}' \boldsymbol{\Sigma}_p \mathbf{X}_k - pa_p \mathbf{I}_{(i)}\Big) - z\mathbf{I}_{n-1},$$

$$\mathbf{D}_{ki}^{-1} = \mathbf{D}_k^{-1} - \mathbf{e}_i \mathbf{h}_i' - \mathbf{r}_i \mathbf{e}_i' = \frac{1}{\sqrt{npb_p}}\Big(\mathbf{X}_{ki}' \boldsymbol{\Sigma}_p \mathbf{X}_{ki} - pa_p \mathbf{I}_{(i)}\Big) - z\mathbf{I}_{n-1},$$

$$\mathbf{h}_i' = \frac{1}{\sqrt{npb_p}} \mathbf{x}_i' \boldsymbol{\Sigma}_p \mathbf{X}_{ki} + \frac{1}{\sqrt{npb_p}}\Big(\mathbf{x}_i' \boldsymbol{\Sigma}_p \mathbf{x}_i - pa_p\Big)\mathbf{e}_i', \qquad \mathbf{r}_i = \frac{1}{\sqrt{npb_p}} \mathbf{X}_{ki}' \boldsymbol{\Sigma}_p \mathbf{x}_i,$$

$$\zeta_i = \frac{1}{1+\vartheta_i}, \qquad \vartheta_i = \mathbf{h}_i' \mathbf{D}_{ki,r}(z)\mathbf{e}_i, \qquad \mathbf{M}_{ki} = \boldsymbol{\Sigma}_p \mathbf{X}_{ki} \mathbf{D}_{ki}(z) \mathbf{X}_{ki}' \boldsymbol{\Sigma}_p.$$

We have some crucial identities,

$$\mathbf{X}_{ki}\mathbf{e}_i = \mathbf{0}, \qquad \mathbf{e}_i' \mathbf{D}_{ki,r} = \mathbf{e}_i' \mathbf{D}_{ki} = -\frac{\mathbf{e}_i'}{z}, \tag{S2.1}$$

where $\mathbf{0}$ is a p-dimensional vector with all the elements equal to 0. By using (S2.1) and some frequently used formulas about the inverse of matrices, we obtain two useful identities:

$$\begin{aligned} \mathbf{D}_k - \mathbf{D}_{ki,r} &= -\mathbf{D}_{ki,r}(\mathbf{D}_k^{-1} - \mathbf{D}_{ki,r}^{-1})\mathbf{D}_k = -\mathbf{D}_{ki,r}(\mathbf{e}_i \mathbf{h}_i')\mathbf{D}_k \\ &= -\mathbf{D}_{ki,r}(\mathbf{e}_i \mathbf{h}_i')(\zeta_i \mathbf{D}_{ki,r}) = -\zeta_i \mathbf{D}_{ki,r}(\mathbf{e}_i \mathbf{h}_i')\mathbf{D}_{ki,r} \end{aligned} \tag{S2.2}$$

and

$$\begin{aligned} \mathbf{D}_{ki,r} - \mathbf{D}_{ki} &= -\mathbf{D}_{ki}(\mathbf{D}_{ki,r}^{-1} - \mathbf{D}_{ki}^{-1})\mathbf{D}_{ki,r} = -\mathbf{D}_{ki}(\mathbf{r}_i \mathbf{e}_i')\mathbf{D}_{ki,r} \\ &= -\mathbf{D}_{ki}\Big(\frac{1}{\sqrt{npb_p}} \mathbf{X}_{ki}' \boldsymbol{\Sigma}_p \mathbf{x}_i \mathbf{e}_i'\Big)\mathbf{D}_{ki} = \frac{1}{z\sqrt{npb_p}} \mathbf{D}_{ki} \mathbf{X}_{ki}' \boldsymbol{\Sigma}_p \mathbf{x}_i \mathbf{e}_i'. \end{aligned} \tag{S2.3}$$

Using (S2.2) and (S2.3), for $i < k$, we obtain the following decomposition of $\mathbb{E}_k \mathbf{M}_k^{(1)}(z)$,

$$\begin{aligned} \mathbb{E}_k \mathbf{M}_k^{(1)}(z) &= \mathbb{E}_k\Big\{\boldsymbol{\Sigma}_p(\mathbf{X}_{ki} + \mathbf{x}_i \mathbf{e}_i')\mathbf{D}_k(\mathbf{X}_{ki} + \mathbf{x}_i \mathbf{e}_i')'\boldsymbol{\Sigma}_p\Big\} \\ &= \mathbb{E}_k\Big(\boldsymbol{\Sigma}_p \mathbf{X}_{ki} \mathbf{D}_k \mathbf{X}_{ki}' \boldsymbol{\Sigma}_p + \boldsymbol{\Sigma}_p \mathbf{X}_{ki} \mathbf{D}_k \mathbf{e}_i \mathbf{x}_i' \boldsymbol{\Sigma}_p \\ &\qquad + \boldsymbol{\Sigma}_p \mathbf{x}_i \mathbf{e}_i' \mathbf{D}_k \mathbf{X}_{ki}' \boldsymbol{\Sigma}_p + \boldsymbol{\Sigma}_p \mathbf{x}_i \mathbf{e}_i' \mathbf{D}_k \mathbf{e}_i \mathbf{x}_i' \boldsymbol{\Sigma}_p\Big) \\ &= \mathbb{E}_k \mathbf{M}_{ki} - \mathbb{E}_k\Big\{\frac{\zeta_i(z)}{znpb_p} \mathbf{M}_{ki} \mathbf{x}_i \mathbf{x}_i' \mathbf{M}_{ki}\Big\} + \mathbb{E}_k\Big\{\frac{\zeta_i(z)}{z\sqrt{npb_p}} \mathbf{M}_{ki}\Big\}\mathbf{x}_i \mathbf{x}_i' \boldsymbol{\Sigma}_p \\ &\qquad + \boldsymbol{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbb{E}_k\Big\{\frac{\zeta_i(z)}{z\sqrt{npb_p}} \mathbf{M}_{ki}\Big\} - \mathbb{E}_k\{\zeta_i(z)/z\}\boldsymbol{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \boldsymbol{\Sigma}_p \\ &:= \mathbf{B}_1(z) + \mathbf{B}_2(z) + \mathbf{B}_3(z) + \mathbf{B}_4(z) + \mathbf{B}_5(z). \end{aligned} \tag{S2.4}$$

Write

$$\mathbf{D}_k^{-1} = \sum_{i=1(\neq k)}^{n} \mathbf{e}_i \mathbf{h}_i' - z\mathbf{I}_{n-1}.$$

Multiplying $\mathbf{D}_k$ on the right-hand side, we have

$$z\mathbf{D}_k = -\mathbf{I}_{n-1} + \sum_{i=1(\neq k)}^{n} \mathbf{e}_i \mathbf{h}_i' \mathbf{D}_k.$$

Multiplying $\mathbf{\Sigma}_p \mathbf{X}_k$ on the left-hand side, $\mathbf{X}_k' \mathbf{\Sigma}_p$ on the right-hand side, we get

$$z\mathbf{M}_k^{(1)}(z) = -\mathbf{\Sigma}_p \mathbf{X}_k \mathbf{X}_k' \mathbf{\Sigma}_p + \sum_{i=1(\neq k)}^{n} \mathbf{\Sigma}_p \mathbf{X}_k \mathbf{e}_i \mathbf{h}_i' \mathbf{D}_k \mathbf{X}_k' \mathbf{\Sigma}_p.$$

Thus,

$$\begin{aligned} z\mathbb{E}_k\big(\mathbf{M}_k^{(1)}(z)\big) &= -\mathbb{E}_k\big(\mathbf{\Sigma}_p \mathbf{X}_k \mathbf{X}_k' \mathbf{\Sigma}_p\big) + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k(\mathbf{\Sigma}_p \mathbf{X}_k \mathbf{e}_i \mathbf{h}_i' \mathbf{D}_k \mathbf{X}_k' \mathbf{\Sigma}_p) \\ &= -\mathbf{\Sigma}_p \mathbb{E}_k\Bigg(\sum_{i=1(\neq k)}^{n} \mathbf{x}_i \mathbf{x}_i'\Bigg)\mathbf{\Sigma}_p + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k\Big\{\zeta_i \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{h}_i' \mathbf{D}_{ki,r}(\mathbf{X}_{ki}' + \mathbf{e}_i \mathbf{x}_i')\mathbf{\Sigma}_p\Big\} \\ &= -(n-k)\mathbf{\Sigma}_p^2 - \sum_{i<k}\Big(\mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p\Big) \\ &\quad + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k\bigg(\frac{\zeta_i}{\sqrt{npb_p}} \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p \mathbf{X}_{ki} \mathbf{D}_{ki,r} \mathbf{X}_{ki}' \mathbf{\Sigma}_p\bigg) \\ &\quad + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k\Big(\zeta_i \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{h}_i' \mathbf{D}_{ki,r} \mathbf{e}_i \mathbf{x}_i' \mathbf{\Sigma}_p\Big) \\ &= -(n-k)\mathbf{\Sigma}_p^2 - \sum_{i<k}\Big(\mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p\Big) + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k\bigg(\frac{\zeta_i}{\sqrt{npb_p}} \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{M}_{ki}\bigg) \\ &\quad + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k\Big(\zeta_i \vartheta_i \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p\Big). \end{aligned} \tag{S2.5}$$

Applying (S2.4) and (S2.5) to $\mathbb{E}_k \mathbf{M}_k^{(1)}(z_2)$ (for $i < k$) and $z_1 \mathbb{E}_k \mathbf{M}_k^{(1)}(z_1)$, we get the following decomposition:

$$\begin{aligned} z_1 \mathbb{Z}_k &= \frac{z_1}{n(pb_p)^2} \operatorname{tr}\Big\{\mathbb{E}_k \mathbf{M}_k^{(1)}(z_1) \cdot \mathbb{E}_k \mathbf{M}_k^{(1)}(z_2)\Big\} \\ &= \frac{1}{n(pb_p)^2} \operatorname{tr}\Bigg[\bigg\{-(n-k)\mathbf{\Sigma}_p^2 - \sum_{i<k}\Big(\mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p\Big) + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k\bigg(\frac{\zeta_i}{\sqrt{npb_p}} \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{M}_{ki}\bigg) \\ &\qquad + \sum_{i=1(\neq k)}^{n} \mathbb{E}_k\Big(\zeta_i \vartheta_i \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p\Big)\bigg\} \times \mathbb{E}_k \mathbf{M}_k^{(1)}(z_2)\Bigg] \\ &= C_1(z_1, z_2) + C_2(z_1, z_2) + C_3(z_1, z_2) + C_4(z_1, z_2), \end{aligned} \tag{S2.6}$$

where

$$C_1(z_1,z_2) = -\frac{n-k}{n(pb_p)^2}\mathrm{tr}\Big\{\mathbf{\Sigma}_p^2 \cdot \mathbb{E}_k \mathbf{M}_k^{(1)}(z_2)\Big\},$$

(S2.7)

$$C_2(z_1,z_2) = -\frac{1}{n(pb_p)^2}\sum_{i<k}\mathbf{x}_i'\mathbf{\Sigma}_p\Big\{\sum_{j=1}^{5}\mathbf{B}_j(z_2)\Big\}\mathbf{\Sigma}_p\mathbf{x}_i = \sum_{j=1}^{5} C_{2j},$$

$$C_3(z_1,z_2) = \frac{1}{n(pb_p)^2}\sum_{i<k}\mathbb{E}_k\left[\frac{\zeta_i(z_1)}{\sqrt{npb_p}}\mathbf{x}_i'\mathbf{M}_{ki}(z_1)\Big\{\sum_{j=1}^{5}\mathbf{B}_j(z_2)\Big\}\mathbf{\Sigma}_p\mathbf{x}_i\right]$$

$$\text{(S2.8)} \qquad + \frac{1}{n(pb_p)^2}\sum_{i>k}\mathbb{E}_k\left[\frac{\zeta_i(z_1)}{\sqrt{npb_p}}\mathbf{x}_i'\mathbf{M}_{ki}(z_1)\Big\{\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\Big\}\mathbf{\Sigma}_p\mathbf{x}_i\right] = \sum_{j=1}^{6} C_{3j},$$

$$C_4(z_1,z_2) = \frac{1}{n(pb_p)^2}\sum_{i<k}\mathbb{E}_k\left[\zeta_i(z_1)\vartheta_i(z_1)\mathbf{x}_i'\mathbf{\Sigma}_p\Big\{\sum_{j=1}^{5}\mathbf{B}_j(z_2)\Big\}\mathbf{\Sigma}_p\mathbf{x}_i\right]$$

$$\text{(S2.9)} \qquad + \frac{1}{n(pb_p)^2}\sum_{i>k}\mathbb{E}_k\left[\zeta_i(z_1)\vartheta_i(z_1)\mathbf{x}_i'\mathbf{\Sigma}_p\Big\{\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\Big\}\mathbf{\Sigma}_p\mathbf{x}_i\right] = \sum_{j=1}^{6} C_{4j}.$$

Now we estimate all the terms in (S2.6). We will show that these terms are negligible as $n\to\infty$, expect C_{25}, C_{33}, C_{45} defined in (S2.7) – (S2.9).

For $C_1(z_1,z_2)$, we have

$$\mathbb{E}|C_1(z_1,z_2)| = \frac{n-k}{n(pb_p)^2}\Big|\mathrm{tr}\Big\{\mathbf{\Sigma}_p^2\cdot\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\Big\}\Big| = O\left(\frac{1}{p^2}\right)\cdot O(np) = O\left(\frac{n}{p}\right),$$

where the second equality follows from the fact $\big|\mathrm{tr}\big\{\mathbf{\Sigma}_p^2\cdot\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\big\}\big| = O(np)$, which can be verified by using the similar argument in the proof of Lemma S1.21.

Applying Lemma S1.8 and inequality (S1.3) with $\mathbf{B} = \mathbf{I}_p$, we have

$$\begin{aligned}\mathbb{E}|C_{21}| &\leqslant \frac{1}{n(pb_p)^2}\sum_{i<k}\mathbb{E}\big|\mathbf{x}_i'\mathbf{\Sigma}_p\cdot\mathbb{E}_k\mathbf{M}_{ki}(z_2)\cdot\mathbf{\Sigma}_p\mathbf{x}_i\big| \\ &\leqslant \frac{1}{n(pb_p)^2}\sum_{i<k}\Big(\mathbb{E}\big|\mathbf{x}_i'\mathbf{\Sigma}_p\cdot\mathbb{E}_k\mathbf{M}_{ki}(z_2)\cdot\mathbf{\Sigma}_p\mathbf{x}_i\big|^2\Big)^{1/2} \leqslant \frac{Kn}{p}.\end{aligned}$$

Applying Lemma S1.8 and inequality (S1.3) with $\mathbf{B} = \mathbf{\Sigma}_p$, we have

$$\begin{aligned}\mathbb{E}|C_{22}| &\leqslant \frac{1}{n(pb_p)^2}\sum_{i<k}\mathbb{E}\left|\mathbf{x}_i'\mathbf{\Sigma}_p\cdot\mathbb{E}_k\left\{\frac{\zeta_i(z_2)}{z_2 npb_p}\mathbf{M}_{ki}(z_2)\mathbf{x}_i\mathbf{x}_i'\mathbf{M}_{ki}(z_2)\right\}\cdot\mathbf{\Sigma}_p\mathbf{x}_i\right| \\ &= \frac{K}{n^2(pb_p)^3}\sum_{i<k}\mathbb{E}\big|\mathbf{x}_i'\mathbf{M}_{ki}(z_2)\mathbf{\Sigma}_p\mathbf{x}_i\big|^2 \leqslant \frac{Kn}{p}.\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}\mathbb{E}|C_{23}| = \mathbb{E}|C_{24}| &\leqslant \frac{1}{n(pb_p)^2}\sum_{i<k}\mathbb{E}\left|\mathbf{x}_i'\mathbf{\Sigma}_p\cdot\mathbb{E}_k\left\{\frac{\zeta_i(z_2)}{z_2\sqrt{npb_p}}\mathbf{M}_{ki}(z_2)\right\}\mathbf{x}_i\mathbf{x}_i'\mathbf{\Sigma}_p\cdot\mathbf{\Sigma}_p\mathbf{x}_i\right| \\ &\leqslant \frac{K}{np^2\sqrt{np}}\sum_{i<k}\mathbb{E}\big|\mathbf{x}_i'\mathbf{\Sigma}_p\mathbf{M}_{ki}(z_2)\mathbf{x}_i\cdot\mathbf{x}_i'\mathbf{\Sigma}_p^2\mathbf{x}_i\big|\end{aligned}$$

$$\leqslant \frac{K}{np^2\sqrt{np}} \sum_{i<k} \left\{ \mathbb{E}\left| \mathbf{x}_i' \mathbf{\Sigma}_p \mathbf{M}_{ki}(z_2) \mathbf{x}_i \right|^2 \right\}^{1/2} \cdot \left\{ \mathbb{E}\left| \mathbf{x}_i' \mathbf{\Sigma}_p^2 \mathbf{x}_i \right|^2 \right\}^{1/2} \leqslant K\sqrt{\frac{n}{p}}.$$

Applying Lemma S1.8 and inequality (S1.2) with $\mathbf{B} = \mathbb{E}_k \mathbf{M}_{ki}(z_2) \mathbf{\Sigma}_p$, we have

$$\mathbb{E}|C_{31}| = \frac{1}{n(pb_p)^2} \sum_{i<k} \mathbb{E}\left| \mathbb{E}_k \left\{ \frac{\zeta_i(z_1)}{\sqrt{npb_p}} \mathbf{x}_i' \mathbf{M}_{ki}(z_1) \cdot \mathbb{E}_k \mathbf{M}_{ki}(z_2) \cdot \mathbf{\Sigma}_p \mathbf{x}_i \right\} \right|$$
$$\leqslant \frac{K}{np^2\sqrt{np}} \sum_{i<k} \mathbb{E}\left| \mathbf{x}_i' \mathbf{M}_{ki}(z_1) \cdot \mathbb{E}_k \mathbf{M}_{ki}(z_2) \cdot \mathbf{\Sigma}_p \mathbf{x}_i \right| \leqslant K\sqrt{\frac{n}{p}}.$$

We define $\tilde{\zeta}_i(z)$ and $\widetilde{\mathbf{M}}_{ki}(z)$ as the analogues of $\zeta_i(z)$ and $\mathbf{M}_{ki}(z)$, respectively, using $\{\mathbf{x}_1, \ldots, \mathbf{x}_k, \tilde{\mathbf{x}}_{k+1}, \ldots, \tilde{\mathbf{x}}_n\}$, where $\tilde{\mathbf{x}}_{k+1}, \ldots, \tilde{\mathbf{x}}_n$ are i.i.d. copies of $\mathbf{x}_{k+1}, \ldots, \mathbf{x}_n$ and independent of $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Then,

$$\begin{aligned}
&\mathbb{E}|C_{32}| \\
&= \frac{1}{n(pb_p)^2} \sum_{i<k} \mathbb{E}\left| \mathbb{E}_k \left[\frac{\zeta_i(z_1)}{\sqrt{npb_p}} \mathbf{x}_i' \mathbf{M}_{ki}(z_1) \cdot \mathbb{E}_k \left\{ \frac{\zeta_i(z_2)}{z_2 npb_p} \mathbf{M}_{ki}(z_2) \mathbf{x}_i \mathbf{x}_i' \mathbf{M}_{ki}(z_2) \right\} \cdot \mathbf{\Sigma}_p \mathbf{x}_i \right] \right| \\
&= \frac{1}{n(pb_p)^2} \sum_{i<k} \mathbb{E}\left| \mathbb{E}_k \left[\frac{\zeta_i(z_1)}{\sqrt{npb_p}} \mathbf{x}_i' \mathbf{M}_{ki}(z_1) \cdot \mathbb{E}_k \left\{ \frac{\tilde{\zeta}_i(z_2)}{z_2 npb_p} \widetilde{\mathbf{M}}_{ki}(z_2) \mathbf{x}_i \mathbf{x}_i' \widetilde{\mathbf{M}}_{ki}(z_2) \right\} \cdot \mathbf{\Sigma}_p \mathbf{x}_i \right] \right| \\
&\leqslant \frac{K}{n^2p^3\sqrt{np}} \sum_{i<k} \mathbb{E}\left| \left[\mathbf{x}_i' \mathbf{M}_{ki}(z_1) \widetilde{\mathbf{M}}_{ki}(z_2) \mathbf{x}_i \cdot \mathbf{x}_i' \widetilde{\mathbf{M}}_{ki}(z_2) \mathbf{\Sigma}_p \mathbf{x}_i \right] \right| \\
&\leqslant \frac{K}{n^2p^3\sqrt{np}} \sum_{i<k} \left\{ \mathbb{E}\left| \mathbf{x}_i' \mathbf{M}_{ki}(z_1) \widetilde{\mathbf{M}}_{ki}(z_2) \mathbf{x}_i \right|^2 \right\}^{1/2} \left\{ \mathbb{E}\left| \mathbf{x}_i' \widetilde{\mathbf{M}}_{ki}(z_2) \mathbf{\Sigma}_p \mathbf{x}_i \right|^2 \right\}^{1/2} \\
&\overset{(S1.2)}{\leqslant} K\sqrt{\frac{n}{p}}.
\end{aligned}$$

Similarly, we have

$$\mathbb{E}|C_{3j}| \leqslant K\frac{n}{p}, \qquad j = 4, 5, 6.$$

Applying Lemma S1.8 and inequality (S1.3) with $\mathbf{B} = \mathbf{I}_{n-1}$, we obtain

$$\mathbb{E}|C_{4j}| \leqslant K\frac{n}{p}, \qquad j = 1, 2, 3, 4, 6.$$

Moreover, by using Lemmas S1.8 – S1.9 and Lemma S1.21, we obtain the following limits:

$$\begin{aligned}
C_{25} &= -\frac{1}{n(pb_p)^2} \sum_{i<k} \left(\mathbf{x}_i' \mathbf{\Sigma}_p \left[-\mathbb{E}_k \left\{ \frac{\zeta_i(z_2)}{z_2} \right\} \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p \right] \mathbf{\Sigma}_p \mathbf{x}_i \right) \\
&= -\frac{1}{n(pb_p)^2} m(z_2) \sum_{i<k} \left(\mathbf{x}_i' \mathbf{\Sigma}_p^2 \mathbf{x}_i \right)^2 \\
&= -\frac{k}{n} m(z_2) + o_{L_1}(1),
\end{aligned}$$

$$C_{45} = \frac{1}{n(pb_p)^2} \sum_{i<k} \mathbb{E}_k \left(\zeta_i(z_1) \vartheta_i(z_1) \mathbf{x}_i' \mathbf{\Sigma}_p \left[-\mathbb{E}_k \left\{ \frac{\zeta_i(z_2)}{z_2} \right\} \mathbf{\Sigma}_p \mathbf{x}_i \mathbf{x}_i' \mathbf{\Sigma}_p \right] \mathbf{\Sigma}_p \mathbf{x}_i \right)$$

$$= \frac{1}{n(pb_p)^2} \sum_{i<k} \mathbb{E}_k \Big[-m^2(z_1)m(z_2)\big(\mathbf{x}_i'\mathbf{\Sigma}_p^2\mathbf{x}_i\big)^2 \Big] + o_{L_1}(1)$$

$$= -\frac{k}{n} m^2(z_1)m(z_2) + o_{L_1}(1),$$

and

$$C_{33} = \frac{1}{n(pb_p)^2} \sum_{i<k} \mathbb{E}_k \left[\frac{\zeta_i(z_1)}{\sqrt{npb_p}} \mathbf{x}_i'\mathbf{M}_{ki}(z_1) \left\{ \mathbb{E}_k \frac{\zeta_i(z_2)}{z_2\sqrt{npb_p}} \mathbf{M}_{ki}(z_2) \right\} \mathbf{x}_i\mathbf{x}_i'\mathbf{\Sigma}_p^2\mathbf{x}_i \right]$$

$$= \frac{1}{n^2p^2b_p^2} z_1 m(z_1)m(z_2) \left\{ \sum_{i<k} \mathbf{x}_i'\mathbb{E}_k\mathbf{M}_{ki}(z_1)\mathbb{E}_k\mathbf{M}_{ki}(z_2)\mathbf{x}_i \right\} + o_{L_4}(1)$$

$$= \frac{k}{n} m(z_1)m(z_2)z_1\mathbb{Z}_k + o_{L_1}(1).$$

From above estimates, we have

$$z_1\mathbb{Z}_k = -\frac{k}{n}m(z_2) - \frac{k}{n}m^2(z_1)m(z_2) + \frac{k}{n}m(z_1)m(z_2)z_1\mathbb{Z}_k + o_{L_1}(1)$$

$$= \frac{k}{n}z_1 m(z_1)m(z_2) + \frac{k}{n}z_1 m(z_1)m(z_2)\mathbb{Z}_k + o_{L_1}(1),$$

which is equivalent to

$$\mathbb{Z}_k = \frac{\frac{k}{n}m(z_1)m(z_2)}{1 - \frac{k}{n}m(z_1)m(z_2)} + o_{L_1}(1).$$

□

S2.2. *Proof of Lemma S1.8.*

PROOF. This lemma can be proved by using similar arguments in Section 5.2.2 of Chen and Pan (2015). □

S2.3. *Proof of Lemma S1.9.*

PROOF. Note that $\mathbf{M}_{ki}$ and $\mathbf{x}_i$ are independent. By using Lemma S1.1, we have

$$\mathbb{E}\big|\mathbf{x}_i'\mathbf{M}_{ki}\mathbf{B}\mathbf{x}_i - \mathrm{tr}\mathbf{M}_{ki}\mathbf{B}\big|^2 \leqslant K\big\{\nu_4\mathbb{E}\mathrm{tr}\big(\mathbf{M}_{ki}\mathbf{B}\overline{\mathbf{B}}\overline{\mathbf{M}}_{ki}\big)\big\} \leqslant Knp^2\|\mathbf{B}\|^2, \tag{S2.10}$$

where we use the fact that

$$\begin{aligned}\big|\mathrm{tr}\big(\mathbf{M}_{ki}\mathbf{B}\overline{\mathbf{B}}\overline{\mathbf{M}}_{ki}\big)\big| &= \big|\mathrm{tr}\big(\mathbf{\Sigma}_p\mathbf{X}_{ki}\mathbf{D}_{ki}\mathbf{X}_{ki}'\mathbf{\Sigma}_p\mathbf{B}\overline{\mathbf{B}}\mathbf{\Sigma}_p\mathbf{X}_{ki}\overline{\mathbf{D}}_{ki}\mathbf{X}_{ki}'\mathbf{\Sigma}_p\big)\big| \\ &= \big|\mathrm{tr}\big(\mathbf{D}_{ki}^{1/2}\mathbf{X}_{ki}'\mathbf{\Sigma}_p\mathbf{B}\overline{\mathbf{B}}\mathbf{\Sigma}_p\mathbf{X}_{ki}\overline{\mathbf{D}}_{ki}\mathbf{X}_{ki}'\mathbf{\Sigma}_p^2\mathbf{X}_{ki}\mathbf{D}_{ki}^{1/2}\big)\big| \\ &\leqslant n\cdot\|\mathbf{D}_{ki}^{1/2}\mathbf{X}_{ki}'\mathbf{\Sigma}_p^{1/2}\|\cdot\|\mathbf{\Sigma}_p^{1/2}\|\cdot\|\mathbf{B}\overline{\mathbf{B}}\|\cdot\|\mathbf{\Sigma}_p^{1/2}\| \\ &\qquad \times\|\mathbf{\Sigma}_p^{1/2}\mathbf{X}_{ki}\overline{\mathbf{D}}_{ki}\mathbf{X}_{ki}'\mathbf{\Sigma}_p^{1/2}\|\cdot\|\mathbf{\Sigma}_p\|\cdot\|\mathbf{\Sigma}_p^{1/2}\mathbf{X}_{ki}\mathbf{D}_{ki}^{1/2}\| \\ &= n\cdot\|\mathbf{\Sigma}_p\|^2\cdot\|\mathbf{B}\|^2\cdot\|\mathbf{\Sigma}_p^{1/2}\mathbf{X}_{ki}\mathbf{D}_{ki}\mathbf{X}_{ki}'\mathbf{\Sigma}_p^{1/2}\|^2 \\ &= n\cdot\|\mathbf{\Sigma}_p\|^2\cdot\|\mathbf{B}\|^2\cdot\|\mathbf{D}_{ki}\mathbf{X}_{ki}'\mathbf{\Sigma}_p\mathbf{X}_{ki}\|^2 \\ &= n\cdot\|\mathbf{\Sigma}_p\|^2\cdot\|\mathbf{B}\|^2\cdot\|\sqrt{npb_p}(\mathbf{I}_{n-1}+z\mathbf{D}_{ki})+pa_p\mathbf{I}_{(i)}\mathbf{D}_{ki}\|^2 \\ &\leqslant Knp^2\|\mathbf{B}\|^2. \qquad \text{(S2.11)}\end{aligned}$$

By (S2.10) and the c_r-inequality, we have

$$\mathbb{E}\big|\mathbf{x}_i'\mathbf{M}_{ki}\mathbf{B}\mathbf{x}_i\big|^2 \leqslant K\Big\{\mathbb{E}\big|\mathbf{x}_i'\mathbf{M}_{ki}\mathbf{B}\mathbf{x}_i - \mathrm{tr}\mathbf{M}_{ki}\mathbf{B}\big|^2 + \mathbb{E}\big|\mathrm{tr}\mathbf{M}_{ki}\mathbf{B}\big|^2\Big\} \leqslant Kp^2n^2\mathbb{E}\|\mathbf{B}\|^2,$$

which completes the proof of (S1.2). By using the same argument, we get (S1.3). □

S2.4. *Proof of Lemma S1.10.*

PROOF. The proof of Lemma S1.10 exactly follows Chen and Pan (2015), so is omitted. □

S2.5. *Proof of Lemma S1.11.*

PROOF. By Lemma S1.1 and taking $\mathbf{B} = \mathbf{I}_p$ in the inequality (S2.11), we have

$$\mathbb{E}|\gamma_{k2}|^2 \leqslant \frac{K}{n^2p^2}\mathrm{tr}\big(\mathbf{M}_k^{(s)}\overline{\mathbf{M}}_k^{(s)}\big) \leqslant \frac{K}{n}.$$

Similarly, we can prove that $\mathbb{E}|\eta_k|^2 \leqslant K/n$.

Now, we prove the bounds for the 4-th moments of γ_{ks} and η_{ks}. Let $\mathbf{H}$ be $\mathbf{M}_k^{(s)}$ with all diagonal elements replaced by zeros, then we have

$$\mathbb{E}|\mathbf{x}_k'\mathbf{H}\mathbf{x}_k|^4 \leqslant K(\mathbb{E}X_{11}^4)^2\mathbb{E}\big(\mathrm{tr}\mathbf{H}\mathbf{H}^*\big)^2 \leqslant K\mathbb{E}\big(\mathrm{tr}\mathbf{M}_k^{(s)}\overline{\mathbf{M}}_k^{(s)}\big)^2 \leqslant Kn^2p^4. \tag{S2.12}$$

The first inequality follows from Lemma S1.2, and the last inequality follows from (S2.11).

Let $\mathbb{E}_j(\cdot)$ denote the conditional expectation with respect to $(X_{1k}, X_{2k}, \ldots, X_{jk})$, and let $m_{jj}^{(s)}$ denote the j-th diagonal entry of $\mathbf{M}_k^{(s)}$, where $j = 1, 2, \ldots, p$. Since $\mathbb{E}_{j-1}(X_{jk}^2 - 1)m_{jj}^{(s)} = 0$, then $(X_{jk}^2 - 1)m_{jj}^{(s)}$ can be expressed as a martingale difference

$$(X_{jk}^2 - 1)m_{jj}^{(s)} = (\mathbb{E}_j - \mathbb{E}_{j-1})\Big\{(X_{jk}^2 - 1)m_{jj}^{(s)}\Big\}. \tag{S2.13}$$

Applying the Burkholder's inequality (Lemma S1.3) to (S2.13) yields that

$$\begin{aligned}
&\mathbb{E}\Big|\sum_{j=1}^{p}(X_{jk}^2 - 1)m_{jj}^{(s)}\Big|^4 \\
&\leqslant K\mathbb{E}\bigg(\sum_{j=1}^{p}\mathbb{E}_{j-1}\Big|(X_{jk}^2 - 1)m_{jj}^{(s)}\Big|^2\bigg)^2 + K\mathbb{E}\bigg(\sum_{j=1}^{p}\Big|(X_{jk}^2 - 1)m_{jj}^{(s)}\Big|^4\bigg)^2 \\
&\leqslant K\bigg(\sum_{j=1}^{p}\mathbb{E}|X_{11}|^4\big|m_{jj}^{(s)}\big|^2\bigg) + K\sum_{j=1}^{p}\mathbb{E}|X_{11}|^8\mathbb{E}\big|m_{jj}^{(s)}\big|^4 \\
&\leqslant Kn^5p^2 + Kn^3p^3, \qquad \text{(S2.14)}
\end{aligned}$$

where we use the fact that, with $\mathbf{e}_j$ be the j-th p-dimensional standard basis vector and $\mathbf{y}$ be an $(n-1)$-dimensional random vector with $\mathbb{E}y_i = 0$ and $\mathbb{E}y_i^2 = 1$,

$$\begin{aligned}
\mathbb{E}\big|m_{jj}^{(s)}\big|^4 &= \mathbb{E}\Big|\mathbf{e}_j'\boldsymbol{\Sigma}_p\mathbf{X}_k\mathbf{D}_k^s\mathbf{X}_k'\boldsymbol{\Sigma}_p\mathbf{e}_j\Big|^4 \\
&\leqslant v_1^{-4s}\mathbb{E}\big\|\mathbf{e}_j'\boldsymbol{\Sigma}_p\mathbf{X}_k\big\|^8 = v_1^{-4s}(\boldsymbol{\Sigma}_p^2)_{jj}^4\mathbb{E}\big\|\mathbf{y}\big\|^8 \leqslant Kn^4 + Kn^2p, \qquad \text{(S2.15)}
\end{aligned}$$

where $(\mathbf{\Sigma}_p^2)_{jj} = \sum_\ell (\mathbf{\Sigma}_p)_{j\ell}^2$ is the j-th diagonal elements of $\mathbf{\Sigma}_p^2$. By Rayleigh-Ritz Theorem, we know that $(\mathbf{\Sigma}_p^2)_{jj} \leqslant \lambda_{\max}(\mathbf{\Sigma}_p^2) \leqslant K$. Combining (S2.12) and (S2.14) yields that

$$\begin{aligned}\mathbb{E}|\gamma_{ks}|^4 &\leqslant \frac{1}{(npb_p)^4}\mathbb{E}\bigg|\sum_{j=1}^p (X_{jk}^2 - 1)m_{jj}^{(s)} + \mathbf{x}_k'\mathbf{H}\mathbf{x}_k\bigg|^4 \\ &\leqslant \frac{K}{n^4p^4}\mathbb{E}\bigg|\sum_{j=1}^p (X_{jk}^2 - 1)m_{jj}^{(s)}\bigg|^4 + \frac{K}{n^4p^4}\mathbb{E}|\mathbf{x}_k'\mathbf{H}\mathbf{x}_k|^4 \\ &\leqslant K\left(\frac{1}{n^2} + \frac{n}{p^2}\right).\end{aligned}$$

Moreover, by Lemma S1.1, we have

$$\mathbb{E}|\eta_k|^4 \leqslant \frac{K}{n^2p^2}\mathbb{E}\big|\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{x}_k - pa_p\big|^4 + K\mathbb{E}\big|\gamma_{k1}\big|^4 \leqslant \frac{K\delta_n^4}{n} + K\left(\frac{1}{n^2} + \frac{n}{p^2}\right).$$

This completes the proof of the lemma. □

S2.6. *Proof of Lemma S1.12.*

PROOF. By the identity $m_n - \mathbb{E}m_n = -\sum_{k=1}^n \big(\mathbb{E}_{k-1}\, m_n - \mathbb{E}_k\, m_n\big)$, we have

$$\mathrm{Var}(m_n) = \sum_{k=1}^n \mathbb{E}\left|\mathbb{E}_{k-1}\, m_n - \mathbb{E}_k\, m_n\right|^2 + 2\sum_{1\leqslant s<t\leqslant 1} \mathbb{E}\left(\mathbb{E}_{s-1}\, m_n - \mathbb{E}_s\, m_n\right)\left(\mathbb{E}_{t-1}\, m_n - \mathbb{E}_t\, m_n\right).$$

Since each term in the second sum on the RHS of the above identity is zero, we write

$$\begin{aligned}\mathrm{Var}(m_n) &= \sum_{k=1}^n \mathbb{E}\left|\mathbb{E}_{k-1}\, m_n - \mathbb{E}_k\, m_n\right|^2 \\ &= \sum_{k=1}^n \mathbb{E}\left|\mathbb{E}_{k-1}\left(m_n - \mathbb{E}_{(k)}\, m_n\right)\right|^2 \\ &\leqslant \sum_{k=1}^n \mathbb{E}\left|m_n - \mathbb{E}_{(k)}\, m_n\right|^2,\end{aligned}$$

where $\mathbb{E}_{(k)}(\cdot)$ denotes the expectation w.r.t. the σ-field generated by $\mathbf{x}_k$. To prove (S1.5), it suffices to show

$$\mathbb{E}\left|m_n - \mathbb{E}_{(k)}\, m_n\right|^2 = O\left(\frac{1}{n^3}\right), \qquad 1 \leqslant k \leqslant n. \tag{S2.16}$$

Now we deal with the case $k = 1$, and the remaining cases are analogous and omitted.

Denote $\widetilde{\mathbf{Y}} = (\widetilde{Y}_{ij})_{p\times n} := \mathbf{\Sigma}_p^{1/2}\mathbf{Y}$ where $\mathbf{Y} = (npb_p)^{-1/4}\mathbf{X}$, and let $\widetilde{\mathbf{y}}_k$ be the k-th column of $\widetilde{\mathbf{Y}}$. Let $\widetilde{\mathbf{Y}}_k$ be the $p \times (n-1)$ matrix extracted from $\widetilde{\mathbf{Y}}$ by removing $\widetilde{\mathbf{y}}_k$, then the matrix model (1) can be written as

$$\mathbf{A}_n = \begin{pmatrix} \widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} & (\widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{y}}_1)' \\ \widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{y}}_1 & \widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{Y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{n-1} \end{pmatrix}.$$

With notations $\mathbf{A}_k = \widetilde{\mathbf{Y}}_k'\widetilde{\mathbf{Y}}_k - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{n-1}$ and $\mathbf{D}_k = (\mathbf{A}_k - z\mathbf{I}_n)^{-1}$, we have

$$\begin{aligned}
&\mathsf{tr}\mathbf{D} - \mathsf{tr}\mathbf{D}_1\\
&= \frac{1 + (\widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{y}}_1)'\Big(\widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{Y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{n-1} - z\mathbf{I}_{n-1}\Big)^{-2}(\widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{y}}_1)}{\Big(\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} - z\Big) - (\widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{y}}_1)'\Big(\widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{Y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{n-1} - z\mathbf{I}_{n-1}\Big)^{-1}(\widetilde{\mathbf{Y}}_1'\widetilde{\mathbf{y}}_1)}\\
&= \frac{1 + \widetilde{\mathbf{y}}_1'\Big[\widetilde{\mathbf{Y}}_1\widetilde{\mathbf{Y}}_1'\Big(\widetilde{\mathbf{Y}}_1\widetilde{\mathbf{Y}}_1' - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{n-1} - z\mathbf{I}_{n-1}\Big)^{-2}\Big]\widetilde{\mathbf{y}}_1}{\Big(\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} - z\Big) - \widetilde{\mathbf{y}}_1'\Big\{\widetilde{\mathbf{Y}}_1\widetilde{\mathbf{Y}}_1'\Big(\widetilde{\mathbf{Y}}_1\widetilde{\mathbf{Y}}_1' - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{n-1} - z\mathbf{I}_{n-1}\Big)^{-1}\Big\}\widetilde{\mathbf{y}}_1}\\
&=: \frac{1+U}{V},
\end{aligned}$$

where the second "=" comes from the identity

$$\mathbf{B}(\mathbf{AB} - \alpha\mathbf{I})^{-n}\mathbf{A} = \mathbf{BA}(\mathbf{BA} - \alpha\mathbf{I})^{-n}.$$

Moreover, with notations U and V, we can write $D_{11} = 1/V$ and

$$\begin{aligned}
\mathbb{E}\left|m_n - \mathbb{E}_{(1)} m_n\right|^2 &= \frac{1}{n^2}\mathbb{E}\Big|(\mathsf{tr}\mathbf{D} - \mathsf{tr}\mathbf{D}_1) - \mathbb{E}_{(1)}(\mathsf{tr}\mathbf{D} - \mathsf{tr}\mathbf{D}_1)\Big|^2 \qquad (\because \mathbb{E}_{(1)}\mathsf{tr}\mathbf{D}_1 = \mathsf{tr}\mathbf{D}_1)\\
&= \frac{1}{n^2}\mathbb{E}\left|\frac{1+U}{V} - \mathbb{E}_{(1)}\left(\frac{1+U}{V}\right)\right|^2\\
&\leqslant \frac{2}{n^2}\left\{\mathbb{E}\left|\frac{1}{V} - \mathbb{E}_{(1)}\left(\frac{1}{V}\right)\right|^2 + \mathbb{E}\left|\frac{U}{V} - \mathbb{E}_{(1)}\left(\frac{U}{V}\right)\right|^2\right\}.
\end{aligned}$$

By the same arguments as those on Page 196 of Bao (2015), it is sufficient to prove that

$$\text{(S2.17)} \qquad \mathbb{E}_{(1)}\left|U - \mathbb{E}_{(1)} U\right|^2 = O\left(\frac{1}{n}\right), \qquad \mathbb{E}_{(1)}\left|V - \mathbb{E}_{(1)} V\right|^2 = O\left(\frac{1}{n}\right).$$

For simplicity of presentation, we define

$$\mathbf{H}^{[\ell]} = \left(H_{jk}^{[\ell]}\right)_{p\times p} := \widetilde{\mathbf{Y}}_1\widetilde{\mathbf{Y}}_1'\Big(\widetilde{\mathbf{Y}}_1\widetilde{\mathbf{Y}}_1' - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{n-1} - z\mathbf{I}_{n-1}\Big)^{-\ell}, \qquad \ell = 1, 2.$$

Then, we write

$$\text{(S2.18)} \qquad U - \mathbb{E}_{(1)} U = \sum_{i\neq j} H_{ij}^{[2]}\widetilde{Y}_{i1}\widetilde{Y}_{j1} + \sum_{i=1}^{p} H_{ii}^{[2]}\big(\widetilde{Y}_{i1}^2 - \mathbb{E}\,\widetilde{Y}_{i1}^2\big),$$

$$\text{(S2.19)} \qquad V - \mathbb{E}_{(1)} V = \widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} - \sum_{i\neq j} H_{ij}^{[1]}\widetilde{Y}_{i1}\widetilde{Y}_{j1} - \sum_{i=1}^{p} H_{ii}^{[1]}\big(\widetilde{Y}_{i1}^2 - \mathbb{E}\,\widetilde{Y}_{i1}^2\big).$$

Now we proceed to prove (S2.17). From (S2.19), we have

$$\begin{aligned}
&\mathbb{E}_{(1)}|V - \mathbb{E}_{(1)} V|^2\\
&\leqslant K\left\{\mathbb{E}_{(1)}\left|\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\right|^2 + \mathbb{E}_{(1)}\left|\sum_{i\neq j} H_{ij}^{[1]}\widetilde{Y}_{i1}\widetilde{Y}_{j1}\right|^2 + \mathbb{E}_{(1)}\left|\sum_{i=1}^{p} H_{ii}^{[1]}\Big(\widetilde{Y}_{i1}^2 - \mathbb{E}\,\widetilde{Y}_{i1}^2\Big)\right|^2\right\}
\end{aligned}$$

(S2.20)

$$= K\left\{\mathbb{E}\left|\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\right|^2 + \sum_{i\neq j}\left|H_{ij}^{[1]}\right|^2\mathbb{E}\left(\widetilde{Y}_{i1}^2\widetilde{Y}_{j1}^2\right) + \sum_{i=1}^{p}\left|H_{ii}^{[1]}\right|^2\mathbb{E}\left(\widetilde{Y}_{i1}^2 - \mathbb{E}\,\widetilde{Y}_{i1}^2\right)^2\right\}.$$

After some straightforward calculations, we obtain some estimates:

$$\mathbb{E}\,\widetilde{Y}_{i1}^2 = O\left(\frac{1}{\sqrt{np}}\right), \quad \mathbb{E}\,\widetilde{Y}_{i1}^4 = O\left(\frac{1}{np}\right), \quad \mathbb{E}\left(\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\right)^2 = O\left(\frac{1}{n}\right). \tag{S2.21}$$

Combining (S2.21) and (S2.20), we obtain

$$\mathbb{E}_{(1)}\left|V - \mathbb{E}_{(1)}V\right|^2 \leqslant \frac{K}{n} + \frac{K}{np}\mathsf{tr}\left|\mathbf{H}^{[1]}\right|^2. \tag{S2.22}$$

Similarly, we can show that

$$\mathbb{E}_{(1)}\left|U - \mathbb{E}_{(1)}U\right|^2 \leqslant \frac{K}{np}\mathsf{tr}\left|\mathbf{H}^{[2]}\right|^2. \tag{S2.23}$$

To get (S2.17), it suffices to show that

$$\mathsf{tr}\left|\mathbf{H}^{[\ell]}\right|^2 = O(p), \qquad \ell = 1, 2.$$

Let $\{\mu_i^{(k)}, i = 1, \ldots, n-1\}$ be eigenvalues of $\mathbf{A}_k$, then the eigenvalues of $\mathbf{H}^{[\ell]}$ ($\ell = 1, 2$) are

$$\frac{\{\mu_i^{(1)} + a_p\sqrt{p/(nb_p)}\}^2}{\left|\mu_i^{(1)} - z\right|^{2\ell}}, \qquad i = 1, 2, \ldots, n-1,$$

and a zero eigenvalue with algebraic multiplicity $(p - n + 1)$. Using the fact $\mu_i^{(1)} \geqslant -a_p\sqrt{p/(nb_p)}$, we conclude that

$$\mathsf{tr}\left|\mathbf{H}^{[\ell]}\right|^2 = \sum_{i=1}^{n-1}\frac{\{\mu_i^{(1)} + a_p\sqrt{p/(nb_p)}\}^2}{\left|\mu_i^{(1)} - z\right|^{2\ell}} = O(p), \qquad \ell = 1, 2.$$

This completes the proof of the lemma. □

S2.7. *Proof of Lemma S1.13.*

PROOF. We only provide the estimation of D_{11}, since others are analogous. Note that

$$D_{11} = V^{-1} = \left(\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} - z - \widetilde{\mathbf{y}}_1'\mathbf{H}^{[1]}\widetilde{\mathbf{y}}_1\right)^{-1}.$$

Let $\mathbf{v}_i^{(1)} = \left(v_{i1}^{(1)}, \ldots, v_{ip}^{(1)}\right)$, $(i = 1, 2, \ldots, n-1)$ be the unit eigenvector of $\mathbf{A}_1$ corresponding to the eigenvalue $\mu_i^{(1)}$, and let

$$w_i^{(1)} = \frac{\sqrt{np}a_p}{\sqrt{b_p}}\left|\widetilde{\mathbf{y}}_1'\mathbf{v}_i^{(1)}\right|^2.$$

Applying spectral decomposition to $\mathbf{H}^{[1]}$ yields

$$D_{11} = \left\{\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} - z - \sum_{i=1}^{n-1}\left(\frac{\mu_i^{(1)} + \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}}{\mu_i^{(1)} - z}\right)\left|\widetilde{\mathbf{y}}_1'\mathbf{v}_i^{(1)}\right|^2\right\}^{-1}$$

$$= \left\{ \widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} - z - \frac{1}{\sqrt{np}}\frac{\sqrt{b_p}}{a_p}\sum_{i=1}^{n-1}\frac{\left(\mu_i^{(1)} + \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\right)w_i^{(1)}}{\mu_i^{(1)} - z} \right\}^{-1}$$

$$=: \left(-z - m_n(z) + h_1\right)^{-1}, \tag{S2.24}$$

where

$$h_1 = \left\{ m_n - \frac{1}{n}\sum_{i=1}^{n-1}\left(\frac{\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\mu_i^{(1)} + 1}{\mu_i^{(1)} - z}\right)\right\}$$
$$+ \left\{ \widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} - \frac{1}{n}\sum_{i=1}^{n-1}\left(\frac{\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\mu_i^{(1)} + 1}{\mu_i^{(1)} - z}\right)(w_i^{(1)} - 1)\right\}.$$

By (S2.24), we obtain

$$\left| D_{11} + \frac{1}{z + \mathbb{E}\, m_n} \right| = \left| \frac{\mathbb{E}\, m_n - m_n + h_1}{\left(-z - m_n + h_1\right)\left(z + \mathbb{E}\, m_n\right)} \right| \leqslant K\left|(\mathbb{E}\, m_n - m_n) + h_1\right|,$$

which implies that

$$\mathbb{E}\left| D_{11} + \frac{1}{z + \mathbb{E}\, m_n} \right|^2$$
$$\leqslant K\left\{ \mathbb{E}\left|\mathbb{E}\, m_n - m_n\right|^2 + \mathbb{E}\left| m_n - \left(\frac{1}{n}\sum_{i=1}^{n-1}\frac{1}{\mu_i^{(1)} - z}\right) - \sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\left(\frac{1}{n}\sum_{i=1}^{n-1}\frac{\mu_i^{(1)}}{\mu_i^{(1)} - z}\right)\right|^2 \right.$$
$$\left. + \mathbb{E}\left| \frac{1}{n}\sum_{i=1}^{n-1}\left(\frac{\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\mu_i^{(1)} + 1}{\mu_i^{(1)} - z}\right)(w_i^{(1)} - 1)\right|^2 + \mathbb{E}\left|\widetilde{\mathbf{y}}_1'\widetilde{\mathbf{y}}_1 - \sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\right|^2 \right\}$$
$$=: K(\mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV})$$

$$= O\left(\frac{1}{n^2}\right) + \left[O\left(\frac{1}{n^2}\right) + O\left(\frac{n}{p}\right)\right] + O\left(\frac{1}{n}\right) + O\left(\frac{1}{n}\right) \tag{S2.25}$$

$$= O\left(\frac{1}{n}\right) + O\left(\frac{n}{p}\right). \tag{S2.26}$$

Below we explain (S2.25) in more detail:

(I) Follows from Lemma S1.12.
(II) Use the fact

$$\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\left|\frac{1}{n}\sum_{i=1}^{n-1}\frac{\mu_i^{(1)}}{\mu_i^{(1)} - z}\right| = O\left(\sqrt{\frac{n}{p}}\right)$$

and

$$\left| m_n - \frac{1}{n}\sum_{i=1}^{n-1}\frac{1}{\mu_i^{(1)} - z}\right| = \left|\frac{1}{n}\mathrm{tr}\mathbf{D} - \frac{1}{n}\mathrm{tr}\mathbf{D}_k\right| \overset{\text{(S2.39)}}{=} O\left(\frac{1}{n}\right).$$

(III) Use (S2.21).
(IV) Analogous to the estimation of $\mathbb{E}\left|V-\mathbb{E}_{(1)}V\right|^2$.

□

S2.8. *Proof of Lemma S1.14.*

PROOF. We only provide the estimation of $\widetilde{D}_{11}$, since the others are analogous.

Let $\widetilde{\mathbf{r}}_k'$ be k-th row of $\widetilde{\mathbf{Y}}=\mathbf{\Sigma}_p^{1/2}\mathbf{Y}$ and let $\mathbf{B}_k$ be the $(p-1)\times n$ matrix extracted from $\widetilde{\mathbf{Y}}$ by deleting $\widetilde{\mathbf{r}}_k'$.

With notations defined above, we can write

$$\widetilde{\mathbf{A}}=\begin{pmatrix}\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{r}}_1-\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}} & \widetilde{\mathbf{r}}_1'\mathbf{B}_1' \\ \mathbf{B}_1\widetilde{\mathbf{r}}_1 & \mathbf{B}_1\mathbf{B}_1'-\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_{p-1}\end{pmatrix}.$$

Denote

$$\widetilde{\mathbf{A}}_k=\mathbf{B}_k'\mathbf{B}_k-\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_n,\qquad k=1,\ldots,n,$$

and

$$W=\widetilde{\mathbf{r}}_1'\mathbf{B}_1'\mathbf{B}_1\left(\mathbf{B}_1'\mathbf{B}_1-\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}\mathbf{I}_n-z\mathbf{I}_n\right)^{-1}\widetilde{\mathbf{r}}_1.$$

Let $\{\widetilde{\mu}_i^{(k)},\, i=1,\ldots,n\}$ be the eigenvalues of $\widetilde{\mathbf{A}}_k$, and let $\widetilde{\mathbf{v}}_i^{(1)}=(\widetilde{v}_{i1}^{(1)},\ldots,\widetilde{v}_{ip}^{(1)})$, $i=1,2,\ldots,n$, be the unit eigenvector of $\widetilde{\mathbf{A}}_1$ corresponding to the eigenvalue $\widetilde{\mu}_i^{(1)}$, and set

$$\widetilde{w}_i^{(1)}=\frac{\sqrt{np}a_p}{\sqrt{b_p}}|\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{v}}_i^{(1)}|^2,$$

then we have

$$W=\frac{1}{n}\sum_{i=1}^n\left(\frac{\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\widetilde{\mu}_i^{(1)}+1}{\widetilde{\mu}_i^{(1)}-z}\right)\widetilde{w}_i^{(1)},$$

and

$$\widetilde{D}_{11}=\left(\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{r}}_1-\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}-z-W\right)^{-1}=:\left(-\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}-z-m_n+\widetilde{h}_1\right)^{-1},$$

where

$$\widetilde{h}_1=\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{r}}_1+m_n-W$$

(S2.27)

$$=\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{r}}_1+m_n-\frac{1}{n}\sum_{i=1}^n\left(\frac{\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\widetilde{\mu}_i^{(1)}+1}{\widetilde{\mu}_i^{(1)}-z}\right)-\frac{1}{n}\sum_{i=1}^n\left(\frac{\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\widetilde{\mu}_i^{(1)}+1}{\widetilde{\mu}_i^{(1)}-z}\right)(\widetilde{w}_i^{(1)}-1).$$

We define the set of events

$$\Omega_0=\left\{\left|\mathbb{E}\,m_n-m_n+\widetilde{h}_1\right|\geqslant\frac{1}{2}\sqrt{\frac{p}{n}}\right\},$$

then the inequality

$$\left|\left(\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z+\mathbb{E}\,m_n\right)\left(\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z+m_n-\widetilde{h}_1\right)\right|\geqslant K\frac{p}{n}$$

holds on Ω_0. Thus we obtain

$$\begin{aligned}
&\mathbb{E}\left|\widetilde{D}_{11}+\frac{1}{a_p\sqrt{p/(nb_p)}+z+\mathbb{E}\,m_n}\right|^2\\
&\quad\leqslant\mathbb{E}\left|\frac{\mathbb{E}\,m_n-m_n+\widetilde{h}_1}{\left(a_p\sqrt{p/(nb_p)}+z+\mathbb{E}\,m_n\right)\left(a_p\sqrt{p/(nb_p)}+z+m_n-\widetilde{h}_1\right)}\right|^2\\
&\quad\leqslant K\left\{\left(\frac{n}{p}\right)^2\cdot\mathbb{P}(\Omega_0^c)+\frac{n}{p}\cdot\mathbb{P}(\Omega_0)\right\}\cdot\mathbb{E}\left|\mathbb{E}\,m_n-m_n+\widetilde{h}_1\right|^2,
\end{aligned}$$

where we use the inequality

$$\left|\left(\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z+\mathbb{E}\,m_n\right)\left(\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z+m_n-\widetilde{h}_1\right)\right|\geqslant K\sqrt{\frac{p}{n}},\tag{S2.28}$$

that holds on the full set Ω. The inequality (S2.28) follows from the facts

$$\begin{aligned}
&\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z+m_n-\widetilde{h}_1\\
&=\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z-\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{r}}_1+\widetilde{\mathbf{r}}_1'\mathbf{B}_1'\mathbf{B}_1\left(\mathbf{B}_1'\mathbf{B}_1-\sqrt{\tfrac{p}{n}}\tfrac{a_p}{\sqrt{b_p}}\mathbf{I}_n\right)^{-1}\widetilde{\mathbf{r}}_1\\
&=\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z+\widetilde{\mathbf{r}}_1'\left\{\mathbf{B}_1'\mathbf{B}_1\left(\mathbf{B}_1'\mathbf{B}_1-\sqrt{\tfrac{p}{n}}\tfrac{a_p}{\sqrt{b_p}}\mathbf{I}_n\right)^{-1}-\mathbf{I}_n\right\}\widetilde{\mathbf{r}}_1\\
&=\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z+\frac{1}{\sqrt{np}}\frac{\sqrt{b_p}}{a_p}\sum_{i=1}^{n}\left(\frac{\widetilde{\mu}_i^{(1)}+\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}}{\widetilde{\mu}_i^{(1)}-z}-1\right)\widetilde{w}_i^{(1)}\\
&=\left(\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z\right)\left(1+\frac{1}{\sqrt{np}}\frac{\sqrt{b_p}}{a_p}\sum_{i=1}^{n}\frac{\widetilde{w}_i^{(1)}}{\widetilde{\mu}_i^{(1)}-z}\right)\\
&=:\left(\sqrt{\frac{p}{n}}\frac{a_p}{\sqrt{b_p}}+z\right)(1+S),
\end{aligned}$$

and

$$|1+S|\geqslant K\sqrt{\frac{n}{p}}.$$

We now proceed to complete the proof of (S1.6). Note that we have

$$\mathbb{P}(\Omega_0)\leqslant\frac{4n}{p}\mathbb{E}\left|\mathbb{E}\,m_n-m_n+\widetilde{h}_1\right|^2,$$

thus it is sufficient to prove that

$$\mathbb{E}\left|\mathbb{E}\,m_n-m_n+\widetilde{h}_1\right|^2=O\left(\frac{1}{n}\right)+O\left(\frac{n}{p}\right).\tag{S2.29}$$

Applying (S2.27) gives us

$$\mathbb{E}\left|\mathbb{E}\,m_n - m_n + \widetilde{h}_1\right|^2$$

$$\leqslant K\left\{\mathbb{E}\left|\mathbb{E}\,m_n - m_n\right|^2 + \mathbb{E}\left|m_n - \left(\frac{1}{n}\sum_{i=1}^{n-1}\frac{1}{\widetilde{\mu}_i^{(1)} - z}\right) - \sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\left(\frac{1}{n}\sum_{i=1}^{n-1}\frac{\widetilde{\mu}_i^{(1)}}{\widetilde{\mu}_i^{(1)} - z}\right)\right|^2\right.$$

(S2.30)

$$\left. + \mathbb{E}\left|\frac{1}{n}\sum_{i=1}^{n-1}\left(\frac{\sqrt{\frac{n}{p}}\frac{\sqrt{b_p}}{a_p}\widetilde{\mu}_i^{(1)} + 1}{\widetilde{\mu}_i^{(1)} - z}\right)\left(\widetilde{w}_i^{(1)} - 1\right)\right|^2 + \mathbb{E}\left(\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{r}}_1\right)^2\right\}.$$

Combining the similar method used for (S2.26) with (S2.30) and the fact

$$\begin{aligned}
\mathbb{E}\left(\widetilde{\mathbf{r}}_1'\widetilde{\mathbf{r}}_1\right)^2 &= \mathbb{E}\left\{\sum_{j=1}^{n}\Big(\sum_{i=1}^{p}(\mathbf{\Sigma}_p^{1/2})_{1i}Y_{ij}\Big)^2\right\}^2 \\
&= \mathbb{E}\left\{\sum_{j=1}^{n}\Big(\sum_{i=1}^{p}(\mathbf{\Sigma}_p^{1/2})_{1i}Y_{ij}\Big)^4 + \sum_{j_1\neq j_2}\Big(\sum_{i=1}^{p}(\mathbf{\Sigma}_p^{1/2})_{1i}Y_{ij_1}\Big)^2\Big(\sum_{i=1}^{p}(\mathbf{\Sigma}_p^{1/2})_{1i}Y_{ij_2}\Big)^2\right\} \\
&= O(n/p),
\end{aligned}$$

we obtain (S2.29). □

S2.9. *Proofs of Lemmas S1.15 and S1.16.*

PROOF. The derivatives in these two lemmas can be derived by using the chain rule and Lemmas S1.17 and S1.18 repeatedly, and the details are omitted here. □

S2.10. *Proof of Lemma S1.17.*

TABLE S.1
Derivatives of $(Y_{rs}Y_{\ell t})$ *w.r.t.* Y_{jk}

$\frac{\partial(Y_{rs}Y_{\ell t})}{\partial Y_{jk}}$	$r=\ell=j$	$r\neq j,\ell\neq j$	$r=j,\ell\neq j$	$r\neq j,\ell=j$
$s=t=k$	$2Y_{jk}$	0	$Y_{\ell k}$	Y_{rk}
$s\neq k,t\neq k$	0	0	0	0
$s=k,t\neq k$	Y_{jt}	0	$Y_{\ell t}$	0
$s\neq k,t=k$	Y_{js}	0	0	Y_{rs}

PROOF. (1) By using the chain rule and derivatives shown in Table S.1, we have

$$\begin{aligned}
\frac{\partial D_{\alpha\beta}}{\partial Y_{jk}} &= \sum_{1\leqslant s\leqslant t\leqslant p}\frac{\partial D_{\alpha\beta}}{\partial A_{st}}\cdot\frac{\partial A_{st}}{\partial Y_{jk}} \qquad \left[\frac{\partial A_{st}}{\partial Y_{jk}} := \frac{\partial(\mathbf{Y}'\mathbf{\Sigma}_p\mathbf{Y})_{st}}{\partial Y_{jk}}\right] \\
&= \sum_{s=1}^{p}\frac{\partial D_{\alpha\beta}}{\partial A_{ss}}\cdot\frac{\partial A_{ss}}{\partial Y_{jk}} + \sum_{1\leqslant s<t\leqslant p}\frac{\partial D_{\alpha\beta}}{\partial A_{st}}\cdot\frac{\partial A_{st}}{\partial Y_{jk}} \\
&= \sum_{s=1}^{p}\left(-D_{\alpha s}D_{t\beta}\right)\cdot\sum_{r,\ell}\left\{(\mathbf{\Sigma}_p)_{r\ell}\frac{\partial\left(Y_{rs}Y_{\ell s}\right)}{\partial Y_{jk}}\right\}
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{s<t}\big(-D_{\alpha s}D_{t\beta}-D_{\alpha t}D_{s\beta}\big)\cdot\sum_{r,\ell}\Big\{(\mathbf{\Sigma}_p)_{r\ell}\frac{\partial\big(Y_{rs}Y_{\ell t}\big)}{\partial Y_{jk}}\Big\}\\
&=\big(-D_{\alpha k}D_{k\beta}\big)\cdot\Big\{2(\mathbf{\Sigma}_p)_{jj}Y_{jk}+\sum_{\ell\neq j}(\mathbf{\Sigma}_p)_{j\ell}Y_{\ell k}+\sum_{r\neq j}(\mathbf{\Sigma}_p)_{rj}Y_{rk}\Big\}\\
&\quad+\sum_{k<t}\big(-D_{\alpha k}D_{t\beta}-D_{\alpha t}D_{k\beta}\big)\cdot\Big\{(\mathbf{\Sigma}_p)_{jj}Y_{jt}+\sum_{\ell\neq j}(\mathbf{\Sigma}_p)_{j\ell}Y_{\ell t}\Big\}\\
&\quad+\sum_{s<k}\big(-D_{\alpha s}D_{k\beta}-D_{\alpha k}D_{s\beta}\big)\cdot\Big\{(\mathbf{\Sigma}_p)_{jj}Y_{js}+\sum_{r\neq j}(\mathbf{\Sigma}_p)_{rj}Y_{rs}\Big\}\\
&=\sum_{s=1}^{p}\big(-D_{\alpha s}D_{k\beta}-D_{\alpha k}D_{s\beta}\big)\Big(\sum_{r=1}^{p}(\mathbf{\Sigma}_p)_{rj}Y_{rs}\Big)\\
&=\sum_{s,r}\Big\{-\Big((\mathbf{\Sigma}_p)_{jr}Y_{rs}D_{s\alpha}\Big)D_{\beta k}-\Big((\mathbf{\Sigma}_p)_{jr}Y_{rs}D_{s\beta}\Big)D_{\alpha k}\Big\}\\
&=-F_{j\alpha}D_{\beta k}-F_{j\beta}D_{\alpha k},
\end{aligned}
$$

where the third equality follows from the formula (II. 18) in Khorunzhy, Khoruzhenko and Pastur (1996);

(2)

$$
\begin{aligned}
\frac{\partial F_{\alpha\beta}}{\partial Y_{jk}}&=\frac{\partial}{\partial Y_{jk}}\sum_{s,t}\Big((\mathbf{\Sigma}_p)_{\alpha s}Y_{st}D_{t\beta}\Big)=\sum_{s,t}(\mathbf{\Sigma}_p)_{\alpha s}\Big(\frac{\partial Y_{st}}{\partial Y_{jk}}\cdot D_{t\beta}+Y_{st}\cdot\frac{\partial D_{t\beta}}{\partial Y_{jk}}\Big)\\
&=(\mathbf{\Sigma}_p)_{\alpha j}D_{k\beta}-\sum_{s,t}(\mathbf{\Sigma}_p)_{\alpha s}Y_{st}\big(F_{jt}D_{\beta k}+F_{j\beta}D_{tk}\big)\\
&=(\mathbf{\Sigma}_p)_{\alpha j}D_{k\beta}-E_{j\alpha}D_{\beta k}-F_{j\beta}F_{\alpha k};
\end{aligned}
$$

(3)

$$
\begin{aligned}
\frac{\partial E_{jj}}{\partial Y_{jk}}&=\frac{\partial}{\partial Y_{jk}}\sum_{r}\big(\mathbf{\Sigma}_p\mathbf{Y}\mathbf{D}\big)_{jr}\big(\mathbf{Y}'\mathbf{\Sigma}_p\big)_{rj}\\
&=\sum_r\frac{\partial F_{jr}}{\partial Y_{jk}}\cdot\big(\mathbf{Y}'\mathbf{\Sigma}_p\big)_{rj}+\sum_r F_{jr}\cdot\frac{\partial\big(\mathbf{Y}'\mathbf{\Sigma}_p\big)_{rj}}{\partial Y_{jk}}\\
&=\sum_r\Big((\mathbf{\Sigma}_p)_{jj}D_{kr}-E_{jj}D_{rk}-F_{jr}F_{jk}\Big)\cdot\big(\mathbf{Y}'\mathbf{\Sigma}_p\big)_{rj}+(\mathbf{\Sigma}_p)_{jj}F_{jk}\\
&=2(\mathbf{\Sigma}_p)_{jj}F_{jk}-2E_{jj}F_{jk},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial(E_{jj}D_{kk})}{\partial Y_{jk}}&=\frac{\partial E_{jj}}{\partial Y_{jk}}\cdot D_{kk}+\frac{\partial D_{kk}}{\partial Y_{jk}}\cdot E_{jj}\\
&=\Big(2(\mathbf{\Sigma}_p)_{jj}F_{jk}-2E_{jj}F_{jk}\Big)\cdot D_{kk}-2F_{jk}D_{kk}\cdot E_{jj}\\
&=2(\mathbf{\Sigma}_p)_{jj}F_{jk}D_{kk}-4E_{jj}F_{jk}D_{kk}.
\end{aligned}
$$

The proof of lemma is complete. □

S2.11. *Proof of Lemma S1.18.*

PROOF. 1.

$$\frac{\partial \widehat{F}_{\alpha\beta}}{\partial Y_{jk}} = \frac{\partial}{\partial Y_{jk}} \sum_{s,t} \Big((\mathbf{\Sigma}_p^2)_{\alpha s} Y_{st} D_{t\beta} \Big) = \sum_{s,t} (\mathbf{\Sigma}_p^2)_{\alpha s} \left(\frac{\partial Y_{st}}{\partial Y_{jk}} \cdot D_{t\beta} + Y_{st} \cdot \frac{\partial D_{t\beta}}{\partial Y_{jk}} \right)$$
$$= (\mathbf{\Sigma}_p^2)_{\alpha j} D_{k\beta} - \sum_{s,t} (\mathbf{\Sigma}_p^2)_{\alpha s} Y_{st} \big(F_{jt} D_{\beta k} + F_{j\beta} D_{tk} \big)$$
$$= (\mathbf{\Sigma}_p^2)_{\alpha j} D_{k\beta} - \widehat{E}_{j\alpha} D_{\beta k} - F_{j\beta} \widehat{F}_{\alpha k};$$

2.

$$\frac{\partial E_{jr}}{\partial Y_{jk}} = \frac{\partial}{\partial Y_{jk}} \sum_{\ell} F_{j\ell} \big(\mathbf{Y}' \mathbf{\Sigma}_p \big)_{\ell r} = \sum_{\ell} \frac{\partial F_{j\ell}}{\partial Y_{jk}} \cdot \big(\mathbf{Y}' \mathbf{\Sigma}_p \big)_{\ell r} + \sum_{\ell} F_{j\ell} \cdot \frac{\partial \big(\mathbf{Y}' \mathbf{\Sigma}_p \big)_{\ell r}}{\partial Y_{jk}}$$
$$= \sum_{\ell} \Big((\mathbf{\Sigma}_p)_{jj} D_{k\ell} - E_{jj} D_{\ell k} - F_{j\ell} F_{jk} \Big) \cdot \big(\mathbf{Y}' \mathbf{\Sigma}_p \big)_{\ell r} + (\mathbf{\Sigma}_p)_{jr} F_{jk}$$
$$= (\mathbf{\Sigma}_p)_{jj} F_{rk} + (\mathbf{\Sigma}_p)_{jr} F_{jk} - E_{jj} F_{rk} - F_{jk} E_{jr},$$

$$\frac{\partial (\widehat{E}_{jj} D_{kk})}{\partial Y_{jk}} = \left(\frac{\partial}{\partial Y_{jk}} \sum_{r} E_{jr} (\mathbf{\Sigma}_p)_{rj} \right) \cdot D_{kk} + \widehat{E}_{jj} \cdot \big(-2 F_{jk} D_{kk} \big)$$
$$= D_{kk} \sum_{r} (\mathbf{\Sigma}_p)_{rj} \Big((\mathbf{\Sigma}_p)_{jj} F_{rk} + (\mathbf{\Sigma}_p)_{jr} F_{jk} - E_{jj} F_{rk} - F_{jk} E_{jr} \Big) - 2 \widehat{E}_{jj} F_{jk} D_{kk}$$
$$= (\mathbf{\Sigma}_p)_{jj} \widehat{F}_{jk} D_{kk} + (\mathbf{\Sigma}_p^2)_{jj} F_{jk} D_{kk} - E_{jj} \widehat{F}_{jk} D_{kk} - 3 \widehat{E}_{jj} F_{jk} D_{kk}.$$

The proof of the lemma is complete. □

S2.12. *Proof of Lemma S1.19.*

PROOF. The proofs of the first two inequalities are analogous with that of Lemma 4.2 and (4.35) in Chen and Pan (2015), it is then omitted. As follows we prove the remaining two inequalities.

When the event U_n happens, inequality (S2.11) holds and then the proof of the second inequality in Lemma S1.11 holds for $z \in \mathcal{C}_\ell \cup \mathcal{C}_r$, thus, we have

$$\mathbb{E}|\gamma_{k2}|^4 \mathbb{1}_{U_n} \leqslant K \left(\frac{1}{n^2} + \frac{1}{np} \right). \tag{S2.31}$$

Moreover, by (S2.51) and Burkholder's inequality (Lemma S1.3),

$$\mathbb{E} \left| \frac{1}{npb_p} \mathrm{tr} \mathbf{M}_k^{(1)} - \frac{1}{npb_p} \mathbb{E} \mathrm{tr} \mathbf{M}_k^{(1)} \right|^4 \mathbb{1}_{U_n}$$
$$\leqslant \frac{K}{n^4} \left(\frac{a_p}{b_p} + z \sqrt{\frac{n}{pb_p}} \right)^4 \mathbb{E} \left| \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) (\mathrm{tr} \mathbf{D} - \mathrm{tr} \mathbf{D}_j) \right|^4 \mathbb{1}_{U_n}$$
$$\leqslant \frac{K}{n^4} \left(\frac{a_p}{b_p} + z \sqrt{\frac{n}{pb_p}} \right)^4 \left\{ \sum_{j=1}^{n} \mathbb{E} |\mathrm{tr} \mathbf{D} - \mathrm{tr} \mathbf{D}_j|^4 + \mathbb{E} \left(\sum_{j=1}^{n} \mathbb{E}_k |\mathrm{tr} \mathbf{D} - \mathrm{tr} \mathbf{D}_j|^2 \right)^2 \right\} \mathbb{1}_{U_n}$$
$$\leqslant \frac{K}{n^2}, \tag{S2.32}$$

where the last "$\leqslant$" follows from the fact that $|\mathrm{tr}\mathbf{D} - \mathrm{tr}\mathbf{D}_j|\mathbb{1}_{U_n}$ is bounded.

Combining Lemma S1.1, (S2.31) and (S2.32), we obtain

$$\mathbb{E}|\mu_k|^4\mathbb{1}_{U_n} \leqslant K\left(\frac{1}{n} + \frac{1}{n^2} + \frac{1}{np}\right).$$

The proof of lemma is complete. □

S2.13. *Proof of Lemma S1.20.*

PROOF. Let

$$\varepsilon_k = \frac{1}{z + (npb_p)^{-1}\mathrm{tr}\mathbf{M}_k^{(1)}},$$
$$\mu_k = \frac{1}{\sqrt{npb_p}}(\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{x}_k - pa_p) - \gamma_{k1} - \left(\frac{1}{npb_p}\mathrm{tr}\mathbf{M}_k^{(1)} - \frac{1}{npb_p}\mathbb{E}\mathrm{tr}\mathbf{M}_k^{(1)}\right).$$

We have the decomposition:

$$M_n^{(1)}(z) = \mathrm{tr}\mathbf{D} - \mathbb{E}\mathrm{tr}\mathbf{D} = \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k-1})\mathrm{tr}(\mathbf{D} - \mathbf{D}_k)$$
$$= \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k-1})\big(-\beta_k(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k)\big)$$
$$= \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k-1})\Big\{-(\beta_k - \varepsilon_k)(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k) - \varepsilon_k\gamma_{k2}\Big\} \tag{S2.33}$$
$$= \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k-1})\Big\{-(\varepsilon_k^2\mu_k + \beta_k\varepsilon_k^2\mu_k^2)(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k) - \varepsilon_k\gamma_{k2}\Big\} \tag{S2.34}$$
$$= \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k-1})\Big\{-\varepsilon_k^2\mu_k\Big(1 + \frac{1}{npb_p}\mathrm{tr}\mathbf{M}_k^{(2)}\Big) - \varepsilon_k^2\mu_k\gamma_{k2}$$
$$- \beta_k\varepsilon_k^2\mu_k^2(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k) - \varepsilon_k\gamma_{k2}\Big\}$$
$$=: \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k-1})(u_{k1} + u_{k2} + u_{k3} + u_{k4}) \tag{S2.35}$$

Below we explain (S2.33) and (S2.34) in more details:

- (S2.33) follows from

$$(\mathbb{E}_k - \mathbb{E}_{k-1})\varepsilon_k(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k) = (\mathbb{E}_k - \mathbb{E}_{k-1})\varepsilon_k(\mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k)$$
$$= (\mathbb{E}_k - \mathbb{E}_{k-1})\varepsilon_k\Big(\mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k - \frac{1}{npb_p}\mathrm{tr}\mathbf{M}_k^{(2)}\Big) = (\mathbb{E}_k - \mathbb{E}_{k-1})\varepsilon_k\gamma_{k2}.$$

- (S2.34) follows from the identity $\beta_k = \varepsilon_k + \beta_k\varepsilon_k\mu_k = \varepsilon_k + \varepsilon_k^2\mu_k + \beta_k\varepsilon_k^2\mu_k^2$.

By Lemma S1.19, we have

$$\mathbb{E}|u_{ki}|^4\mathbb{1}_{U_n} = O(n^{-1}), \qquad i = 1, 2, 3, 4,$$

which, together with the decomposition (S2.35) and Burkholder's inequality (Lemma S1.3), implies that

$$\mathbb{E}\big|M_n^{(1)}(z)\mathbb{1}_{U_n}\big|^2 \leqslant K.$$

This completes the proof. □

S2.14. *Proof of Lemma S1.21.*

PROOF. Using Lemma S1.1, we have

$$\mathbb{E}(\mathbf{x}_i'\mathbf{\Sigma}_p^2\mathbf{x}_i - pb_p)^2 \leqslant K\nu_4 \mathsf{tr}(\mathbf{\Sigma}_p^4) \leqslant K \cdot p\|\mathbf{\Sigma}_p^4\| \leqslant Kp. \tag{S2.36}$$

Note that $\mathsf{tr}(\mathbf{A}^*\mathbf{B})$ is the inner product of $\mathsf{vec}(\mathbf{A})$ and $\mathsf{vec}(\mathbf{B})$ for any $n \times m$ matrices $\mathbf{A}$ and $\mathbf{B}$. It follows from the Cauchy–Schwarz inequality that

$$\big|\mathsf{tr}(\mathbf{A}^*\mathbf{B})\big|^2 \leqslant \mathsf{tr}(\mathbf{A}^*\mathbf{A}) \cdot \mathsf{tr}(\mathbf{B}^*\mathbf{B}). \tag{S2.37}$$

By using (S2.37), we have

$$\begin{aligned}
&\mathbb{E}\bigg|\frac{1}{npb_p}\mathsf{tr}\mathbf{M}_k^{(1)}(z) - \frac{1}{n}\mathsf{tr}\mathbf{D}_k(z)\bigg|^2 \\
&= \frac{1}{(npb_p)^2}\mathbb{E}\left|\mathsf{tr}\Big\{\mathbf{D}_k(z)\big(\mathbf{X}_k'\mathbf{\Sigma}_p^2\mathbf{X}_k - pb_p\mathbf{I}_{n-1}\big)\Big\}\right|^2 \\
&\leqslant \frac{1}{(npb_p)^2}\mathbb{E}\left\{\mathsf{tr}\big(\mathbf{D}_k(\bar{z})\mathbf{D}_k(z)\big) \cdot \mathsf{tr}\big(\mathbf{X}_k'\mathbf{\Sigma}_p^2\mathbf{X}_k - pb_p\mathbf{I}_{n-1}\big)^2\right\} \\
&\leqslant \frac{1}{(npb_p)^2}\mathbb{E}\left\{n\big\|\mathbf{D}_k(\bar{z})\mathbf{D}_k(z)\big\| \cdot \mathsf{tr}\big(\mathbf{X}_k'\mathbf{\Sigma}_p^2\mathbf{X}_k - pb_p\mathbf{I}_{n-1}\big)^2\right\} \\
&\leqslant \frac{1}{n(pb_pv_1)^2}\mathbb{E}\left\{\mathsf{tr}\big(\mathbf{X}_k'\mathbf{\Sigma}_p^2\mathbf{X}_k - pb_p\mathbf{I}_{n-1}\big)^2\right\}.
\end{aligned}$$

Indeed, by using (S2.36) and the fact $\mathbb{E}(\mathbf{x}_i'\mathbf{\Sigma}_p^2\mathbf{x}_j)^2 = \mathsf{tr}(\mathbf{\Sigma}_p^4)$ for $i \neq j$, we have

$$\begin{aligned}
\mathbb{E}\Big\{\mathsf{tr}\big(\mathbf{X}_k'\mathbf{\Sigma}_p^2\mathbf{X}_k - pb_p\mathbf{I}_{n-1}\big)^2\Big\} &= \sum_{i\neq k}\mathbb{E}(\mathbf{x}_i'\mathbf{\Sigma}_p^2\mathbf{x}_i - pb_p)^2 + \sum_{i\neq j, i\neq k, j\neq k}\mathbb{E}(\mathbf{x}_i'\mathbf{\Sigma}_p^2\mathbf{x}_j)^2 \\
&\leqslant (n-1)\cdot pK + (n-1)(n-2)\cdot pK.
\end{aligned}$$

Thus we have

$$\mathbb{E}\bigg|\frac{1}{npb_p}\mathsf{tr}\mathbf{M}_k^{(1)}(z) - \frac{1}{n}\mathsf{tr}\mathbf{D}_k(z)\bigg|^2 \leqslant \frac{Kn}{p}. \tag{S2.38}$$

Moreover, by (S2.48) and (S1.4), we have

$$\bigg|\frac{1}{n}\mathsf{tr}\mathbf{D}(z) - \frac{1}{n}\mathsf{tr}\mathbf{D}_k(z)\bigg| \overset{\text{(S2.48)}}{=} \frac{1}{n}\Big|\beta_k(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k)\Big| \overset{\text{(S1.4)}}{\leqslant} \frac{1}{nv_1}, \tag{S2.39}$$

which, together with (S2.38) and the fact that $m_n(z) \xrightarrow{a.s.} m$, completes the proof. □

S2.15. *Proof of Lemma 6.1.*

PROOF. In general, this proof extends the result of Chen and Pan (2015). We denote the non-diagonal part of $\mathbf{A}_n = (\mathbf{X}'\boldsymbol{\Sigma}_p\mathbf{X} - pa_p\mathbf{I}_n)/\sqrt{npb_p}$ as

$$\mathbf{B}_n = \mathbf{A}_n - \mathsf{diag}(\mathbf{A}_n) = \frac{1}{\sqrt{npb_p}}\big(\mathbf{X}'\boldsymbol{\Sigma}_p\mathbf{X} - \mathsf{diag}(\mathbf{X}'\boldsymbol{\Sigma}_p\mathbf{X})\big),$$

where we use the notation $\mathsf{diag}(\mathbf{A})$ to denote the diagonal matrix of $\mathbf{A}$ (replacing all off-diagonal entries with zero). Then we have

$$\max_{1\leqslant j\leqslant n}|\lambda_j^{\mathbf{A}_n}| = \|\mathbf{A}_n\| \leqslant \|\mathbf{B}_n\| + \|\mathsf{diag}(\mathbf{A}_n)\| = \|\mathbf{B}_n\| + \max_{1\leqslant i\leqslant n}|A_{ii}|, \tag{S2.40}$$

where $\|\cdot\|$ denotes the spectral norm and A_{ii} is the (i,i)-th entry of $\mathbf{A}_n$.

Note that

$$\begin{aligned}\max_{1\leqslant i\leqslant n}|A_{ii}| &= \frac{1}{\sqrt{npb_p}}\max_{1\leqslant i\leqslant n}\left|\sum_{s=1}^{p}\sum_{t=1}^{p}(\boldsymbol{\Sigma}_p)_{st}X_{si}X_{ti} - pa_p\right| \\ &= \frac{1}{\sqrt{npb_p}}\max_{1\leqslant i\leqslant n}\left|\sum_{s=1}^{p}(\boldsymbol{\Sigma}_p)_{ss}(X_{si}^2-1) + \sum_{s\neq t}(\boldsymbol{\Sigma}_p)_{st}X_{si}X_{ti}\right| \\ &\leqslant \frac{1}{\sqrt{npb_p}}\max_{1\leqslant i\leqslant n}\left|\sum_{s=1}^{p}(\boldsymbol{\Sigma}_p)_{ss}(X_{si}^2-1)\right| + \frac{1}{\sqrt{npb_p}}\max_{1\leqslant i\leqslant n}\left|\sum_{s\neq t}(\boldsymbol{\Sigma}_p)_{st}X_{si}X_{ti}\right|. \end{aligned}\tag{S2.41}$$

From (S2.40) and (S2.41), it suffices to prove that, for any $\varepsilon > 0$,

$$\Pr\big(\|\mathbf{B}_n\| \geqslant \eta + \varepsilon\big) = o(n^{-1}), \tag{S2.42}$$

$$\Pr\left(\frac{1}{\sqrt{npb_p}}\max_{1\leqslant i\leqslant n}\left|\sum_{s=1}^{p}(\boldsymbol{\Sigma}_p)_{ss}(X_{si}^2-1)\right| \geqslant \varepsilon\right) = o(n^{-1}), \tag{S2.43}$$

and

$$\Pr\left(\frac{1}{\sqrt{npb_p}}\max_{1\leqslant i\leqslant n}\left|\sum_{s\neq t}(\boldsymbol{\Sigma}_p)_{st}X_{si}X_{ti}\right| \geqslant \varepsilon\right) = o(n^{-1}). \tag{S2.44}$$

Since $(\boldsymbol{\Sigma}_p)_{ss} \leqslant \|\boldsymbol{\Sigma}_p\|$, (S2.43) follows from inequality (9) in Chen and Pan (2012).

Next, we consider (S2.42). From the well-known Courant-Fischer theorem, we have

$$\begin{aligned}\|\mathbf{B}_n\|^2 &= \max_{\|\mathbf{z}\|=1}\|\mathbf{B}_n\mathbf{z}\|^2 = \frac{1}{npb_p}\max_{\|\mathbf{z}\|=1}\left\|\Big[\mathbf{X}'\boldsymbol{\Sigma}_p\mathbf{X} - \mathsf{diag}(\mathbf{X}'\boldsymbol{\Sigma}_p\mathbf{X})\Big]\mathbf{z}\right\|^2 \\ &= \frac{1}{npb_p}\max_{\|\mathbf{z}\|=1}\sum_{i=1}^{n}\left(\sum_{j\neq i}(\mathbf{x}_i'\boldsymbol{\Sigma}_p\mathbf{x}_j)z_j\right)^2 \\ &\leqslant \frac{\|\boldsymbol{\Sigma}_p\|^2}{b_p}\max_{\|\mathbf{z}\|=1}\frac{1}{np}\sum_{i=1}^{n}\left(\sum_{j\neq i}(\mathbf{x}_i'\mathbf{x}_j)z_j\right)^2 \leqslant \left(\frac{\eta}{2}\right)^2\|\widehat{\mathbf{B}}_n\|^2, \end{aligned}\tag{S2.45}$$

where $\widehat{\mathbf{B}}_n = \frac{1}{\sqrt{np}}\big(\mathbf{X}'\mathbf{X} - \mathsf{diag}(\mathbf{X}'\mathbf{X})\big)$. For any sequence of positive numbers $k = k_p \to \infty$ and $\varepsilon > 0$, we have

$$\mathbb{P}\big(\|\widehat{\mathbf{B}}_n\| \geqslant 2+\varepsilon\big) \leqslant \frac{\mathbb{E}\,\|\widehat{\mathbf{B}}_n\|^{2k}}{(2+\varepsilon)^{2k}} \leqslant \frac{\mathbb{E}\,\sum_{j=1}^{n}\big(\lambda_j^{\widehat{\mathbf{B}}_n}\big)^{2k}}{(2+\varepsilon)^{2k}} = \frac{\mathbb{E}\,\mathsf{tr}(\widehat{\mathbf{B}}_n^{2k})}{(2+\varepsilon)^{2k}}. \tag{S2.46}$$

With an appropriate choice of the sequence $\{k = k_p\}$ and some sophisticated combinational techniques, (Chen and Pan, 2012, p. 1413 - 1418) proved that for any $\varepsilon > 0$,

$$(2+\varepsilon)^{-2k}\mathbb{E}\,\mathrm{tr}(\widehat{\mathbf{B}}_n^{2k}) = o(n^{-1}),$$

which, together with (S2.45) (S2.46), implies (S2.42).

Finally, we consider (S2.44). For and $\varepsilon > 0$ and $\delta > 0$, we have

$$\begin{aligned}
&\Pr\left(\frac{1}{\sqrt{npb_p}}\max_{1\leqslant i\leqslant n}\left|\sum_{s\neq t}(\mathbf{\Sigma}_p)_{st}X_{si}X_{ti}\right| > \varepsilon\right)\\
&\leqslant n\Pr\left(\left|\sum_{s\neq t}(\mathbf{\Sigma}_p)_{st}X_{s1}X_{t1}\right| > \varepsilon\sqrt{npb_p}\right)\\
&\leqslant n\cdot(\varepsilon\sqrt{npb_p})^{-(4+\delta)}\cdot\mathbb{E}\left|\sum_{s\neq t}(\mathbf{\Sigma}_p)_{st}X_{s1}X_{t1}\right|^{4+\delta}\\
&\leqslant n\cdot(\varepsilon\sqrt{npb_p})^{-(4+\delta)}\cdot K\cdot[\mathrm{tr}(\mathbf{\Sigma}_p^2)]^{2+\delta/2} \qquad \text{(S2.47)}\\
&= o(n^{-1}),
\end{aligned}$$

the estimation (S2.47) follows from Lemma S1.2 with $k = 4+\delta$. This completes the proof. □

S2.16. *Proof of Lemma 6.2.* As explained in the main text, we first decompose the random part $M_n^{(1)}(z)$ as a sum of martingale difference, which is given in (S2.54). Then, we apply the martingale CLT (Lemma S1.4) to obtain the asymptotic distribution of $M_n^{(1)}(z)$.

Step 1: Martingale difference decomposition of $M_n^{(1)}(z)$.

First, we introduce some notations. Define

$$\mathbf{X}_k = (\mathbf{x}_1,\ldots,\mathbf{x}_{k-1},\mathbf{x}_{k+1},\ldots,\mathbf{x}_n), \qquad \mathbf{A}_k = \frac{1}{\sqrt{npb_p}}\left(\mathbf{X}_k'\mathbf{\Sigma}_p\mathbf{X}_k - pa_p\mathbf{I}_{n-1}\right),$$

$$\mathbf{D} = (\mathbf{A}_n - z\mathbf{I}_n)^{-1}, \qquad \mathbf{D}_k = (\mathbf{A}_k - z\mathbf{I}_{n-1})^{-1}, \qquad \mathbf{M}_k^{(s)} = \mathbf{\Sigma}_p\mathbf{X}_k\mathbf{D}_k^s\mathbf{X}_k'\mathbf{\Sigma}_p,\ s=1,2,$$

$$a_{kk}^{\mathsf{diag}} = A_{kk} - z = \frac{1}{\sqrt{npb_p}}\left(\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{x}_k - pa_p\right) - z, \qquad \mathbf{q}_k' = \frac{1}{\sqrt{npb_p}}\left(\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{X}_k\right)$$

$$\beta_k = \frac{1}{-a_{kk}^{\mathsf{diag}} + \mathbf{q}_k'\mathbf{D}_k\mathbf{q}_k}, \qquad \beta_k^{\mathsf{tr}} = \frac{1}{z + (npb_p)^{-1}\mathrm{tr}\mathbf{M}_k^{(1)}},$$

$$\gamma_{ks} = -\frac{1}{npb_p}\mathrm{tr}\mathbf{M}_k^{(s)} + \mathbf{q}_k'\mathbf{D}_k^s\mathbf{q}_k,\ s=1,2, \qquad \eta_k = \frac{1}{\sqrt{npb_p}}\left(\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{x}_k - pa_p\right) - \gamma_{k1},$$

$$\ell_k = -\beta_k\beta_k^{\mathsf{tr}}\eta_k\left(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\right).$$

Note that a_{kk}^{diag} is the k-th diagonal element of $\mathbf{D}^{-1}$ and $\mathbf{q}_k'$ is the vector from the k-th row of $\mathbf{D}^{-1}$ by deleting the k-th element. By applying Theorem A.5 in Bai and Silverstein (2010), we obtain the equality

$$\mathrm{tr}\mathbf{D} - \mathrm{tr}\mathbf{D}_k = -\frac{1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k}{-a_{kk}^{\mathsf{diag}} + \mathbf{q}_k'\mathbf{D}_k\mathbf{q}_k} = -\beta_k(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k). \qquad \text{(S2.48)}$$

Straightforward calculation gives:

$$\beta_k - \beta_k^{\mathsf{tr}} = \beta_k \beta_k^{\mathsf{tr}} \eta_k, \tag{S2.49}$$

and

$$(\mathbb{E}_k - \mathbb{E}_{k-1})\beta_k^{\mathsf{tr}}(1 + \mathbf{q}_k' \mathbf{D}_k^2 \mathbf{q}_k) = \mathbb{E}_k\big(\beta_k^{\mathsf{tr}}\gamma_{k2}\big), \qquad \mathbb{E}_{k-1}\big(\beta_k^{\mathsf{tr}}\gamma_{k2}\big) = 0, \tag{S2.50}$$

where $\mathbb{E}_k(\cdot)$ is the expectation with respect to the σ-field generated by the first k columns of $\mathbf{X}$.

By the definition of $\mathbf{D}_k$, we obtain a basic identity:

$$\mathbf{D}_k \mathbf{X}_k' \mathbf{\Sigma}_p \mathbf{X}_k = pa_p \mathbf{D}_k + \sqrt{npb_p}(\mathbf{I}_{n-1} + z\mathbf{D}_k). \tag{S2.51}$$

If $\mathbf{\Sigma}_p = \mathbf{I}_p$, it is straightforward to derive that the limit of $\mathsf{tr}\big(\mathbf{M}_k^{(1)}(z)\big)/(npb_p)$ is $m(z)$ by using (S2.51). However, when $\mathbf{\Sigma}_p \neq \mathbf{I}_p$, we need a more detailed estimate (see Lemma S1.21).

Applying (S2.48) – (S2.50), we have the following decomposition:

$$M_n^{(1)}(z) = \mathsf{tr}\mathbf{D} - \mathbb{E}\mathsf{tr}\mathbf{D} = \sum_{k=1}^{n}(\mathbb{E}_k - \mathbb{E}_{k-1})\big(\mathsf{tr}\mathbf{D} - \mathsf{tr}\mathbf{D}_k\big)$$

$$= -\sum_{k=1}^{n}(\mathbb{E}_k - \mathbb{E}_{k-1})\beta_k\Big(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\Big)$$

$$\overset{(S2.49)}{=} \sum_{k=1}^{n}(\mathbb{E}_k - \mathbb{E}_{k-1})(-\beta_k\beta_k^{\mathsf{tr}}\eta_k)\Big(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\Big) - \sum_{k=1}^{n}(\mathbb{E}_k - \mathbb{E}_{k-1})\beta_k^{\mathsf{tr}}\Big(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\Big)$$

$$= \sum_{k=1}^{n}\Big[(\mathbb{E}_k - \mathbb{E}_{k-1})\ell_k - \mathbb{E}_k\big(\beta_k^{\mathsf{tr}}\gamma_{k2}\big)\Big]. \tag{S2.52}$$

By using (S2.49), we can split ℓ_k as

$$\ell_k = -(\beta_k^{\mathsf{tr}} + \beta_k\beta_k^{\mathsf{tr}}\eta_k)\beta_k^{\mathsf{tr}}\eta_k\Big(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\Big)$$

$$= -\big[(\beta_k^{\mathsf{tr}})^2\eta_k + \beta_k(\beta_k^{\mathsf{tr}})^2\eta_k^2\big]\Big(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\Big)$$

$$= -(\beta_k^{\mathsf{tr}})^2\eta_k\Big(1 + \frac{1}{npb_p}\mathsf{tr}\mathbf{M}_k^{(2)}\Big) - (\beta_k^{\mathsf{tr}})^2\eta_k\gamma_{k2} - \beta_k(\beta_k^{\mathsf{tr}})^2\eta_k^2\Big(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\Big)$$

$$= \ell_{k1} + \ell_{k2} + \ell_{k3}, \tag{S2.53}$$

where $\ell_{k1} = -(\beta_k^{\mathsf{tr}})^2\eta_k\big(1 + \mathsf{tr}\mathbf{M}_k^{(2)}/(npb_p)\big)$, $\ell_{k2} = -(\beta_k^{\mathsf{tr}})^2\eta_k\gamma_{k2}$, $\ell_{k3} = -\beta_k(\beta_k^{\mathsf{tr}})^2\eta_k^2\big(1 + \mathbf{q}_k'\mathbf{D}_k^2\mathbf{q}_k\big)$.

By Lemma S1.10 and Lemma S1.11, it is not difficult to verify that

$$\mathbb{E}\bigg|\sum_{k=1}^{n}(\mathbb{E}_k - \mathbb{E}_{k-1})\ell_{k2}\bigg|^2 = o(1), \qquad \mathbb{E}\bigg|\sum_{k=1}^{n}(\mathbb{E}_k - \mathbb{E}_{k-1})\ell_{k3}\bigg|^2 = o(1).$$

These estimates, together with (S2.52) and (S2.53), imply that

$$M_n^{(1)}(z) = \sum_{k=1}^{n}\mathbb{E}_k\bigg\{-(\beta_k^{\mathsf{tr}})^2\eta_k\Big(1 + \frac{1}{npb_p}\mathsf{tr}\mathbf{M}_k^{(2)}\Big) - \beta_k^{\mathsf{tr}}\gamma_{k2}\bigg\} + o_P(1)$$

$$=: \sum_{k=1}^{n} Y_k(z) + o_P(1), \tag{S2.54}$$

where $Y_k(z)$ is a sequence of martingale differences.
Step 2: Application of martingales CLT to (S2.54).

To prove finite-dimensional convergence of $M_n^{(1)}(z), z \in \mathbb{C}_1$, we need only to consider the limit of the following martingale difference decomposition:

$$\sum_{j=1}^{r} a_j M_n^{(1)}(z_j) = \sum_{j=1}^{r} a_j \sum_{k=1}^{n} Y_k(z_j) + o(1) = \sum_{k=1}^{n} \left(\sum_{j=1}^{r} a_j Y_k(z_j) \right) + o(1),$$

where $\{a_j\}$ is any complex sequence and r is any positive integer. We apply the martingale CLT (Lemma S1.4) to this martingale difference decomposition of $\sum_{j=1}^{r} a_j M_n^{(1)}(z_j)$. To this end, we need to check two conditions:

Condition 1. For any $\varepsilon > 0$,

$$\sum_{k=1}^{n} \mathbb{E}\left(\left| \sum_{j=1}^{r} a_j Y_k(z_j) \right|^2 \mathbb{1}_{\{\sum_{j=1}^{r} a_j Y_k(z_j) \geqslant \varepsilon\}} \right) = o(1).$$

Condition 2. For $z_1, z_2 \in \mathbb{C}_1$, the sum

$$\Lambda_n(z_1, z_2) := \sum_{k=1}^{n} \mathbb{E}_{k-1}\big(Y_k(z_1) Y_k(z_2)\big) \tag{S2.55}$$

converges in probability to $\Lambda(z_1, z_2)$ defined in (33).

First, we verify Condition 1. By Lemma S1.10 and Lemma S1.11, we have

$$\mathbb{E}|Y_j(z)|^4 \leqslant K \frac{\delta_n^4}{n} + K\left(\frac{1}{n^2} + \frac{n}{p^2} \right),$$

which implies that, for each $\varepsilon > 0$,

$$\sum_{k=1}^{n} \mathbb{E}\left(\left| \sum_{j=1}^{r} a_j Y_k(z_j) \right|^2 \mathbb{1}_{\{\sum_{j=1}^{r} a_j Y_k(z_j) \geqslant \varepsilon\}} \right) \leqslant \frac{1}{\varepsilon^2} \sum_{k=1}^{n} \mathbb{E}\left| \sum_{j=1}^{r} a_j Y_k(z_j) \right|^4 = o(1),$$

thus, Condition 1 is satisfied.

Then, we verify Condition 2. Note that

$$-\big(\beta_k^{\mathsf{tr}}\big)^2 \eta_k \Big(1 + \frac{1}{npb_p} \mathsf{tr} \mathbf{M}_k^{(2)} \Big) - \beta_k^{\mathsf{tr}} \gamma_{k2} = \frac{\partial}{\partial z} \Big\{ \beta_k^{\mathsf{tr}}(z) \eta_k(z) \Big\},$$

thus, we can rewrite $\Lambda_n(z_1, z_2)$ as

$$\Lambda_n(z_1, z_2) = \frac{\partial^2}{\partial z_2 \partial z_1} \sum_{k=1}^{n} \mathbb{E}_{k-1}\Big[\mathbb{E}_k\big\{ \beta_k^{\mathsf{tr}}(z_1) \eta_k(z_1) \big\} \cdot \mathbb{E}_k\big\{ \beta_k^{\mathsf{tr}}(z_2) \eta_k(z_2) \big\} \Big]. \tag{S2.56}$$

It is enough to consider the limit of

$$\sum_{k=1}^{n} \mathbb{E}_{k-1}\Big[\mathbb{E}_k\big\{ \beta_k^{\mathsf{tr}}(z_1) \eta_k(z_1) \big\} \cdot \mathbb{E}_k\big\{ \beta_k^{\mathsf{tr}}(z_2) \eta_k(z_2) \big\} \Big]. \tag{S2.57}$$

By equation (4), Lemma S1.21, and the dominated convergence theorem, we conclude that

$$\mathbb{E}\Big| \beta_k^{\mathsf{tr}}(z) + m(z) \Big|^2 = o(1). \tag{S2.58}$$

Combining (S2.57) and (S2.58) yields that

$$\begin{aligned}
&\sum_{k=1}^{n}\mathbb{E}_{k-1}\Big[\mathbb{E}_k\big\{\beta_k^{\mathsf{tr}}(z_1)\eta_k(z_1)\big\}\cdot\mathbb{E}_k\big\{\beta_k^{\mathsf{tr}}(z_2)\eta_k(z_2)\big\}\Big]\\
&=m(z_1)m(z_2)\sum_{k=1}^{n}\mathbb{E}_{k-1}\big\{\mathbb{E}_k\eta_k(z_1)\cdot\mathbb{E}_k\eta_k(z_2)\big\}+o_P(1)\\
&=: m(z_1)m(z_2)\widetilde{\Lambda}_n(z_1,z_2)+o_P(1). \qquad \text{(S2.59)}
\end{aligned}$$

In view of (S2.55) – (S2.59), it suffices to derive the limit of $\widetilde{\Lambda}_n(z_1,z_2)$, which further gives the limit of (S2.55).

Since $\mathbb{E}_k\{\eta_k(z)\}=(1/\sqrt{npb_p})\big(\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{x}_k-pa_p\big)-\mathbb{E}_k\big[\gamma_{k1}(z)\big]$, we have

$$\mathbb{E}_{k-1}\big\{\mathbb{E}_k\eta_k(z_1)\cdot\mathbb{E}_k\eta_k(z_2)\big\}=\frac{1}{n}\bigg\{\frac{\tilde{b}_p}{b_p}(\nu_4-3)+2\bigg\}+A_1^{(k)}+A_2^{(k)}+A_3^{(k)}, \qquad \text{(S2.60)}$$

where

$$\begin{aligned}
A_1^{(k)}&=\mathbb{E}_{k-1}\big\{\mathbb{E}_k\gamma_{k1}(z_1)\cdot\mathbb{E}_k\gamma_{k1}(z_2)\big\},\\
A_2^{(k)}&=-\mathbb{E}_{k-1}\bigg\{\frac{1}{\sqrt{npb_p}}\big(\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{x}_k-pa_p\big)\cdot\mathbb{E}_k\gamma_{k1}(z_1)\bigg\},\\
A_3^{(k)}&=-\mathbb{E}_{k-1}\bigg\{\frac{1}{\sqrt{npb_p}}\big(\mathbf{x}_k'\mathbf{\Sigma}_p\mathbf{x}_k-pa_p\big)\cdot\mathbb{E}_k\gamma_{k1}(z_2)\bigg\}.
\end{aligned}$$

First, we show that $A_2^{(k)}$ and $A_3^{(k)}$ are negligible. Denote $\mathbf{M}_k^{(1)}(z)=\big(m_{ij}^{(1)}(z)\big)_{p\times p}$, using the independence between $\mathbf{x}_k$ and $\mathbf{M}_k^{(1)}$, we have

$$\begin{aligned}
A_2^{(k)}&=\frac{-1}{(npb_p)^{3/2}}\mathbb{E}_{k-1}\Bigg[\bigg(\sum_{i,j}(\mathbf{\Sigma}_p)_{ij}X_{ik}X_{jk}-pa_p\bigg)\\
&\qquad\times\bigg\{\sum_{i\neq j}X_{ik}X_{jk}\mathbb{E}_k m_{ij}^{(1)}+\sum_{i=1}^{p}(X_{ik}^2-1)\mathbb{E}_k m_{ii}^{(1)}\bigg\}\Bigg]\\
&=\frac{-1}{(npb_p)^{3/2}}\mathbb{E}_{k-1}\bigg\{\sum_{i\neq j}(\mathbf{\Sigma}_p)_{ij}X_{ik}^2X_{jk}^2\mathbb{E}_k m_{ij}^{(1)}+\sum_{i=1}^{p}(\mathbf{\Sigma}_p)_{ii}X_{ik}^2(X_{ik}^2-1)\mathbb{E}_k m_{ii}^{(1)}\bigg\}\\
&=\frac{-1}{(npb_p)^{3/2}}\bigg\{\sum_{i\neq j}(\mathbf{\Sigma}_p)_{ij}\mathbb{E}_k m_{ij}^{(1)}+(\nu_4-1)\sum_{i=1}^{p}(\mathbf{\Sigma}_p)_{ii}\mathbb{E}_k m_{ii}^{(1)}\bigg\}\\
&=\frac{-1}{\sqrt{npb_p}}\mathbb{E}_k\bigg\{\frac{1}{npb_p}\mathsf{tr}\Big(\mathbf{\Sigma}_p\mathbf{M}_k^{(1)}\Big)-\frac{\nu_4-2}{npb_p}\sum_{i=1}^{p}(\mathbf{\Sigma}_p)_{ii}m_{ii}^{(1)}\bigg\}. \qquad \text{(S2.61)}
\end{aligned}$$

As for the first term in the bracket of (S2.61), we can estimate it by using a similar argument as in the proof of Lemma S1.21. Replacing pb_p and $\mathbf{M}_k^{(1)}$ in the proof of Lemma S1.21 with $\mathsf{tr}(\mathbf{\Sigma}_p^3)$ and $\mathbf{\Sigma}_p\mathbf{M}_k^{(1)}$, we can prove that

$$\mathbb{E}\bigg|\frac{1}{n\mathsf{tr}(\mathbf{\Sigma}_p^3)}\mathsf{tr}\Big(\mathbf{\Sigma}_p\mathbf{M}_k^{(1)}\Big)-\frac{1}{n}\mathsf{tr}\mathbf{D}_k\bigg|^2\leqslant\frac{Kn}{p}.$$

Moreover, by the fact $\frac{b_p^2}{a_p} \leqslant \mathsf{tr}(\mathbf{\Sigma}_p^3) \leqslant Kp$, the first inequality of which follows from Cauchy–Schwarz inequality, we conclude that

$$\frac{1}{npb_p}\mathsf{tr}\Big(\mathbf{\Sigma}_p\mathbf{M}_k^{(1)}\Big) = \frac{\mathsf{tr}(\mathbf{\Sigma}_p^3)}{pb_p}\cdot\frac{1}{n\mathsf{tr}(\mathbf{\Sigma}_p^3)}\mathsf{tr}\Big(\mathbf{\Sigma}_p\mathbf{M}_k^{(1)}\Big) = O_P(1).$$

As for the second term in the bracket of (S2.61), we have

$$\frac{1}{npb_p}\sum_{i=1}^{p}(\mathbf{\Sigma}_p)_{ii}a_{ii}^{(1)} \leqslant \frac{\|\mathbf{\Sigma}_p\|}{npb_p}\sum_{i=1}^{p}a_{ii}^{(1)} = \frac{\|\mathbf{\Sigma}_p\|}{npb_p}\mathsf{tr}\mathbf{M}_k^{(1)} = O_P(1).$$

Thus, the term in the square bracket of (S2.61) is bounded in probability. Therefore, we conclude that $\Big|\sum_{k=1}^n A_2^{(k)}\Big| \to 0$. Similarly, we can show that $\Big|\sum_{k=1}^n A_3^{(k)}\Big| \to 0$.

Now we consider $A_1^{(k)}$ with the notation $\mathbf{M}_k^{(1)}(z) = \big(m_{ij}^{(1)}(z)\big)_{p\times p}$,

$$\begin{aligned}
A_1^{(k)} = \frac{1}{(npb_p)^2}\mathbb{E}_{k-1}\Bigg[&\Bigg\{\sum_{i\neq j}X_{ik}X_{jk}\mathbb{E}_k m_{ij}^{(1)}(z_1) + \sum_{i=1}^{p}(X_{ik}^2-1)\mathbb{E}_k m_{ii}^{(1)}(z_1)\Bigg\} \\
&\times\Bigg\{\sum_{i\neq j}X_{ik}X_{jk}\mathbb{E}_k m_{ij}^{(1)}(z_2) + \sum_{i=1}^{p}(X_{ik}^2-1)\mathbb{E}_k m_{ii}^{(1)}(z_2)\Bigg\}\Bigg] \\
= \frac{1}{(npb_p)^2}\mathbb{E}_{k-1}&\Bigg\{2\sum_{i\neq j}X_{ik}^2X_{jk}^2\mathbb{E}_k m_{ij}^{(1)}(z_1)\mathbb{E}_k m_{ij}^{(1)}(z_2) + \sum_{i=1}^{p}(X_{ik}^2-1)^2\mathbb{E}_k m_{ii}^{(1)}(z_1)\mathbb{E}_k m_{ii}^{(1)}(z_2)\Bigg\} \\
= \frac{1}{(npb_p)^2}&\Bigg\{2\sum_{i,j}\mathbb{E}_k m_{ij}^{(1)}(z_1)\mathbb{E}_k m_{ij}^{(1)}(z_2) + (\nu_4-3)\sum_{i=1}^{p}\mathbb{E}_k m_{ii}^{(1)}(z_1)\mathbb{E}_k m_{ii}^{(1)}(z_2)\Bigg\} \\
= \frac{2}{(npb_p)^2}&\mathsf{tr}\Big(\mathbb{E}_k\mathbf{M}_k^{(1)}(z_1)\cdot\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\Big) + o_P(1),
\end{aligned}$$

where the last step follows from

$$\begin{aligned}
&\mathbb{E}\Bigg|\sum_{i=1}^{p}\mathbb{E}_k m_{ii}^{(1)}(z_1)\cdot\mathbb{E}_k m_{ii}^{(1)}(z_2)\Bigg|^2 \leqslant p\cdot\sum_{i=1}^{p}\mathbb{E}\Big|\mathbb{E}_k m_{ii}^{(1)}(z_1)\cdot\mathbb{E}_k m_{ii}^{(1)}(z_2)\Big|^2 \\
&\leqslant p\cdot\sum_{i=1}^{p}\Big(\mathbb{E}\Big|\mathbb{E}_k m_{ii}^{(1)}(z_1)\Big|^4\Big)^{1/2}\cdot\Big(\mathbb{E}\Big|\mathbb{E}_k m_{ii}^{(1)}(z_2)\Big|^4\Big)^{1/2} \\
&\leqslant p\cdot\sum_{i=1}^{p}\Big(\mathbb{E}\Big|m_{ii}^{(1)}(z_1)\Big|^4\Big)^{1/2}\cdot\Big(\mathbb{E}\Big|m_{ii}^{(1)}(z_2)\Big|^4\Big)^{1/2} \leqslant K(n^4p^2+n^2p^3).
\end{aligned}$$

By the above estimates, we obtain

$$\widetilde{\Lambda}_n(z_1,z_2) = \frac{2}{(npb_p)^2}\sum_{k=1}^{n}\mathsf{tr}\Big(\mathbb{E}_k\mathbf{M}_k^{(1)}(z_1)\cdot\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\Big) + \Big\{\frac{\tilde{b}_p}{b_p}(\nu_4-3)+2\Big\} + o_P(1),$$

$$= \frac{2}{n}\sum_{k=1}^{n}\mathbb{Z}_k + \frac{\tilde{b}_p}{b_p}(\nu_4-3) + 2 + o_P(1), \tag{S2.62}$$

where

$$\mathbb{Z}_k = \frac{1}{n(pb_p)^2}\mathsf{tr}\Big(\mathbb{E}_k\mathbf{M}_k^{(1)}(z_1)\cdot\mathbb{E}_k\mathbf{M}_k^{(1)}(z_2)\Big).$$

In Lemma S1.7, we derive the asymptotic expression of $\mathbb{Z}_k$. This asymptotic expression ensures that

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{Z}_k \to \int_0^1 \frac{tm(z_1)m(z_2)}{1-tm(z_1)m(z_2)}\mathrm{d}t = -1 - \frac{\log\big(1-m(z_1)m(z_2)\big)}{m(z_1)m(z_2)}. \tag{S2.63}$$

By (S2.56), (S2.59), (S2.62) and (S2.63), we have

$$\widetilde{\Lambda}_n(z_1,z_2) \xrightarrow{p} \frac{\omega}{\theta}(\nu_4-3) - \frac{2\log\big(1-m(z_1)m(z_2)\big)}{m(z_1)m(z_2)}.$$

Therefore,

$$\begin{aligned}\Lambda_n(z_1,z_2) &\xrightarrow{p} \frac{\partial^2}{\partial z_1 \partial z_2}\left\{\frac{\omega}{\theta}(\nu_4-3)m(z_1)m(z_2) - 2\log\big(1-m(z_1)m(z_2)\big)\right\} \\ &= m'(z_1)m'(z_2)\left\{\frac{\omega}{\theta}(\nu_4-3) + 2\big(1-m(z_1)m(z_2)\big)^{-2}\right\}.\end{aligned}$$

The verification of Condition 2 is then complete.

S3. Proofs of equations (35) and (36). This section contains proofs of equations (35) and (36).

PROOF. We only consider the case $j=\ell,0$. (The case $j=r$ is similar to $j=\ell$.) From Proposition 6.1, we have $\mathbb{E}|M(z)|^2 = \Lambda(z,\bar{z}) = O(1)$. Figure S.1 shows the decomposition of the contour $\mathcal{C} = \mathcal{C}_\ell \cup \mathcal{C}_r \cup \mathcal{C}_u \cup \mathcal{C}_0$. Let $\|\mathcal{C}_j\|$ denote the length of $\mathcal{C}_j$, $j=\ell,0$, then

$$\int_{\mathcal{C}_0} \mathbb{E}|M(z)|^2\,\mathrm{d}z \leqslant |\Lambda(z,\bar{z})|\cdot\|\mathcal{C}_0\| = |\Lambda(z,\bar{z})|\cdot(2\xi_n/n) \to 0$$

and

$$\lim_{v_1\downarrow 0}\int_{\mathcal{C}_\ell}\mathbb{E}|M(z)|^2\,\mathrm{d}z \leqslant \lim_{v_1\downarrow 0}|\Lambda(z,\bar{z})|\cdot\|\mathcal{C}_\ell\| = \lim_{v_1\downarrow 0}|\Lambda(z,\bar{z})|\cdot 2(v_1-\xi_n/n) = 0.$$

Thus, the estimate (36) holds for $z\in\mathcal{C}_j,\ j=\ell,0$.

We choose the event $U_n = \{\max_{j\leqslant n}|\lambda_j^{\mathbf{A}_n}| < \eta+\varepsilon\}$ with $\varepsilon = (u_1-\eta)/2$. By Lemma 6.1, the probability of U_n^c decays to zero faster than n^{-1}, that is,

$$\mathbb{P}(U_n^c) = o(n^{-1}).$$

When the event U_n happens, for any $z\in\mathcal{C}_0$, we have $|m_n(z)|\leqslant 2/(u_1-\eta)$ and $|m(z)|\leqslant 1$. Thus, we have

$$\begin{aligned}\int_{\mathcal{C}_0}\mathbb{E}\big|M_n(z)\mathbb{1}_{U_n}\big|^2\,\mathrm{d}z &= \int_{\mathcal{C}_0}\mathbb{E}\big|n[m_n(z)-m(z)-\mathcal{X}_n(m)]\mathbb{1}_{U_n}\big|^2\,\mathrm{d}z \\ &\leqslant n\left(\frac{2}{u_1-\eta}+1+o(1)\right)^2\|\mathcal{C}_0\| \\ &= 2\left(\frac{2}{u_1-\eta}+1+o(1)\right)^2\xi_n,\end{aligned}$$

since $\xi_n\to 0$ as $n\to\infty$, we conclude that (35) is true for $z\in\mathcal{C}_0$.

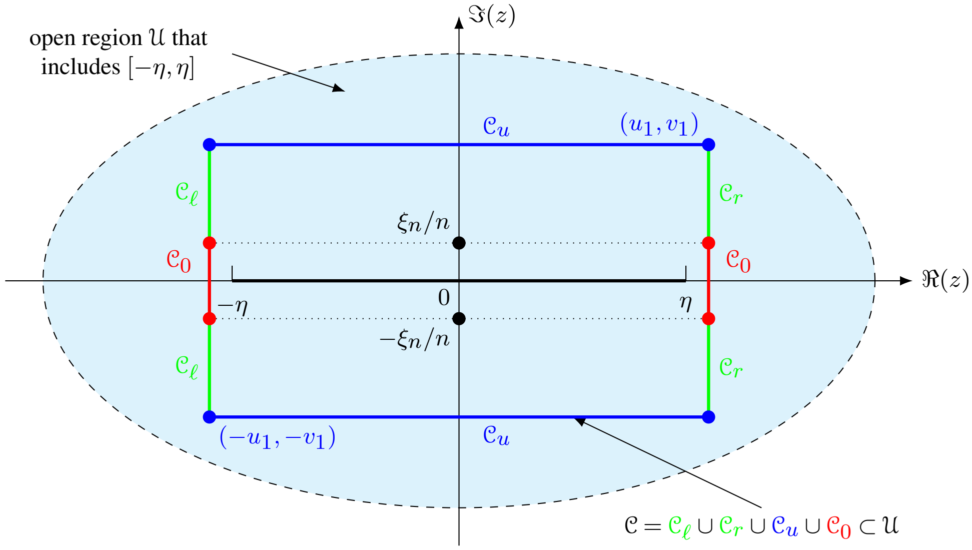


Fig S.1: Open region $\mathcal{U}$ and decomposition of the closed contour $\mathcal{C}$

Recall that we decompose $M_n(z)$ into a random part $M_n^{(1)}(z)$ and a deterministic part $M_n^{(2)}(z)$ for $z \in \mathcal{C}$, where

$$M_n^{(1)}(z) = n\big[m_n(z) - \mathbb{E}m_n(z)\big], \qquad M_n^{(2)}(z) = n\big[\mathbb{E}m_n(z) - m(z) - \mathcal{X}_n(m(z))\big],$$

thus

$$\text{(S3.1)} \qquad \int_{\mathcal{C}_\ell} \mathbb{E}\Big|M_n(z)\mathbb{1}_{U_n}\Big|^2 \mathrm{d}z \leqslant K \int_{\mathcal{C}_\ell} \mathbb{E}\Big|M_n^{(1)}(z)\mathbb{1}_{U_n}\Big|^2 \mathrm{d}z + K \int_{\mathcal{C}_\ell} \Big|M_n^{(2)}(z)\mathbb{1}_{U_n}\Big|^2 \mathrm{d}z.$$

By Lemma S1.20, we have

$$0 \leqslant \int_{\mathcal{C}_\ell} \mathbb{E}\Big|M_n^{(1)}(z)\mathbb{1}_{U_n}\Big|^2 \mathrm{d}z \leqslant K\|\mathcal{C}_\ell\| \to 0, \qquad \text{as } n \to \infty,\ v_1 \downarrow 0.$$

Similarly, we have

$$\int_{\mathcal{C}_\ell} \Big|M_n^{(2)}(z)\mathbb{1}_{U_n}\Big|^2 \mathrm{d}z \to 0, \qquad \text{as } n \to \infty,\ v_1 \downarrow 0.$$

Plugging these estimation into (S3.1) implies that (35) is true for $z \in \mathcal{C}_j$, $j = \ell$. □

S4. Proofs in applications. This section contains proofs of equation (15), Theorem 4.2, and Proposition 4.1.

S4.1. *Proof of equation* (15).

PROOF. By Lemma 2.2 in Wang and Yao (2013), under the high-dimensional setting $c_n = p/n \to c$ as $p \to \infty$, we have

$$n\begin{pmatrix} \frac{1}{p}\mathrm{tr}(\mathbf{S}_n^2) - \Big(1 + \frac{p}{n} + \frac{c(\nu_4 - 2)}{p}\Big) \\ \frac{1}{p}\mathrm{tr}(\mathbf{S}_n) - 1 \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{1}{c^2}\mathbf{H}\right),$$

where

$$\mathbf{H} = \begin{pmatrix} 4c^2 + 4(\nu_4 - 1)(1+c)^2 c & 2(\nu_4 - 1)(1+c)c \\ 2(\nu_4 - 1)(1+c)c & (\nu_4 - 1)c \end{pmatrix}.$$

Define the function $f(x,y) = x - 2y + 1 - py^2/n + p/n$, then $W = f\big(\mathrm{tr}(\mathbf{S}_n^2)/p, \mathrm{tr}(\mathbf{S}_n)/p\big)$, and

$$\frac{\partial f}{\partial x}\Big(1 + \frac{p}{n} + \frac{c(\nu_4 - 2)}{p}, 1\Big) = 1$$

$$\frac{\partial f}{\partial y}\Big(1 + \frac{p}{n} + \frac{c(\nu_4 - 2)}{p}, 1\Big) = -2\Big(1 + \frac{p}{n}\Big)$$

$$f\Big(1 + \frac{p}{n} + \frac{c(\nu_4 - 2)}{p}, 1\Big) = \frac{p}{n} + \frac{c(\nu_4 - 2)}{p}.$$

By the delta method, we obtain

$$n\bigg(W - f\Big(1 + \frac{p}{n} + \frac{c(\nu_4 - 2)}{p}, 1\Big)\bigg) \xrightarrow{d} \mathcal{N}(0, \lim D),$$

where

$$D = \begin{pmatrix} \frac{\partial f}{\partial x}(1 + \frac{p}{n} + \frac{c(\nu_4-2)}{p}, 1) \\ \frac{\partial f}{\partial y}(1 + \frac{p}{n} + \frac{c(\nu_4-2)}{p}, 1) \end{pmatrix}' \Big(\frac{1}{c^2}\mathbf{H}\Big) \begin{pmatrix} \frac{\partial f}{\partial x}(1 + \frac{p}{n} + \frac{c(\nu_4-2)}{p}, 1) \\ \frac{\partial f}{\partial y}(1 + \frac{p}{n} + \frac{c(\nu_4-2)}{p}, 1) \end{pmatrix} \to 4$$

as $n \to \infty$. Thus,

$$nW - p - (\nu_4 - 2) \xrightarrow{d} \mathcal{N}(0, 4),$$

and the proof of (15) is complete. □

S4.2. *Proof of Theorem 4.2.*

PROOF. Recall that $\mathbf{S}_n = \mathbf{Y}\mathbf{Y}'/n$, $a_p = \mathrm{tr}(\mathbf{\Sigma}_p)/p$ and $b_p = \mathrm{tr}(\mathbf{\Sigma}_p^2)/p$. Let $\mathbf{A}_n = \frac{1}{\sqrt{npb_p}}(\mathbf{Y}'\mathbf{Y} - pa_p\mathbf{I}_n) = \frac{1}{\sqrt{npb_p}}(\mathbf{X}'\mathbf{\Sigma}_p\mathbf{X} - pa_p\mathbf{I}_n)$. By some elementary calculations, we obtain two identities:

$$\mathrm{tr}(\mathbf{S}_n) = \sqrt{\frac{pb_p}{n}}\mathrm{tr}(\mathbf{A}_n) + pa_p, \qquad \mathrm{tr}(\mathbf{S}_n^2) = \frac{pb_p}{n}\mathrm{tr}(\mathbf{A}_n^2) + \frac{2pa_p}{n}\sqrt{\frac{pb_p}{n}}\mathrm{tr}(\mathbf{A}_n) + \frac{(pa_p)^2}{n}.$$

Then W can be written as

$$W = \frac{b_p}{n}\mathrm{tr}(\mathbf{A}_n^2) - \frac{2}{p}\sqrt{\frac{pb_p}{n}}\mathrm{tr}(\mathbf{A}_n) - \frac{b_p}{n^2}\big[\mathrm{tr}(\mathbf{A}_n)\big]^2 + \frac{p}{n} - 2a_p + 1.$$

Li and Yao (2016) derived the limiting joint distribution of $\big(\mathrm{tr}(\mathbf{A}_n^2)/n, \mathrm{tr}(\mathbf{A}_n)/n\big)$ (see their Lemma 3.1) as follows:

$$\text{(S4.1)} \qquad n\begin{pmatrix} \frac{1}{n}\mathrm{tr}(\mathbf{A}_n^2) - 1 - \frac{1}{n}\big(\frac{\omega}{\theta}(\nu_4 - 3) + 1\big) \\ \frac{1}{n}\mathrm{tr}(\mathbf{A}_n) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & \frac{\omega}{\theta}(\nu_4 - 3) + 2 \end{pmatrix}\right).$$

Define the function

$$g(x,y) = b_p x - \frac{2n}{p}\sqrt{\frac{pb_p}{n}}y - b_p y^2 + \frac{p}{n} - 2a_p + 1,$$

then $W = g\big(\mathrm{tr}(\mathbf{A}_n^2)/n, \mathrm{tr}(\mathbf{A}_n)/n\big)$, we have

$$\frac{\partial g}{\partial x}\Big(1+\frac{1}{n}\big(\frac{\omega}{\theta}(\nu_4-3)+1\big),0\Big) = b_p,$$
$$\frac{\partial g}{\partial y}\Big(1+\frac{1}{n}\big(\frac{\omega}{\theta}(\nu_4-3)+1\big),0\Big) = -\frac{2n}{p}\sqrt{\frac{pb_p}{n}},$$
$$g\Big(1+\frac{1}{n}\big(\frac{\omega}{\theta}(\nu_4-3)+1\big),0\Big) = b_p + \frac{b_p}{n}\Big(\frac{\omega}{\theta}(\nu_4-3)+1\Big) + \frac{p}{n} - 2a_p + 1.$$

By (S4.1), we have

$$n\bigg(W - g\Big(1+\frac{1}{n}\big(\frac{\omega}{\theta}(\nu_4-3)+1\big),0\Big)\bigg) \xrightarrow{d} \mathcal{N}(0, \lim A),$$

where

$$A = \begin{pmatrix} \frac{\partial g}{\partial x}(1+\frac{1}{n}(\frac{\omega}{\theta}(\nu_4-3)+1),0) \\ \frac{\partial g}{\partial x}(1+\frac{1}{n}(\frac{\omega}{\theta}(\nu_4-3)+1),0) \end{pmatrix}' \begin{pmatrix} 4 & 0 \\ 0 & \frac{\omega}{\theta}(\nu_4-3)+2 \end{pmatrix} \begin{pmatrix} \frac{\partial g}{\partial x}(1+\frac{1}{n}(\frac{\omega}{\theta}(\nu_4-3)+1),0) \\ \frac{\partial g}{\partial x}(1+\frac{1}{n}(\frac{\omega}{\theta}(\nu_4-3)+1),0) \end{pmatrix} \to 4\theta^2.$$

Thus,

$$n\Big(W - b_p - \frac{b_p}{n}\Big(\frac{\omega}{\theta}(\nu_4-3)+1\Big) + 2a_p - 1 - \frac{p}{n}\Big) \xrightarrow{d} \mathcal{N}(0, 4\theta^2),$$

that is,

$$nW - p - \theta\Big(\frac{\omega}{\theta}(\nu_4-3)+1\Big) + n(2\gamma-1-\theta) \xrightarrow{d} \mathcal{N}(0, 4\theta^2).$$

The proof of Theorem 4.2 is complete. □

S4.3. *Proof of Proposition 4.1.*

PROOF. For the test based on statistic W, by Theorem 4.1 and 4.2, we have

$$\begin{aligned}
\beta(H_1) &= \mathbb{P}\bigg(\frac{1}{2}\Big(nW - p - (\nu_4-2)\Big) \geqslant z_\alpha \,\Big|\, H_1\bigg) \\
&= \mathbb{P}\bigg(nW - p - \theta\Big(\frac{\omega}{\theta}(\nu_4-3)+1\Big) + n(2\gamma-1-\theta) \\
&\qquad\qquad \geqslant 2z_\alpha - \theta\Big(\frac{\omega}{\theta}(\nu_4-3)+1\Big) + n(2\gamma-1-\theta) + (\nu_4-2) \,\Big|\, H_1\bigg) \\
&= 1 - \Phi\bigg(\frac{1}{2\theta}\Big\{2z_\alpha - \omega(\nu_4-3) - \theta + n(2\gamma-1-\theta) + (\nu_4-2)\Big\}\bigg),
\end{aligned}$$

since $2\gamma - 1 \leqslant \gamma^2 \leqslant \theta$, Proposition 4.1 follows. □

S5. Additional simulation results. This section contains some additional simulation results of the paper. The simulation settings are the same as those in Section 5.1 of the main paper except $p = n^{2.5}$.

TABLE S.2
Empirical mean and variance of $\overline{G}_n(f_i)$, $i = 1, 2, 3$ *from* 5000 *replications. Theoretical mean and variance are* 0 *and* 1, *respectively. Dimension* $p = n^{2.5}$.

n	$\boldsymbol{\Sigma}_p = \boldsymbol{\Sigma}_A$ mean	var	$\boldsymbol{\Sigma}_p = \boldsymbol{\Sigma}_B$ mean	var	$\boldsymbol{\Sigma}_p = \boldsymbol{\Sigma}_C$ mean	var	$\boldsymbol{\Sigma}_p = \boldsymbol{\Sigma}_D$ mean	var
50	-0.0024	1.0064	0.0087	0.9873	0.0008	1.0101	-0.0063	0.9999
100	0.0021	1.0039	0.0242	0.9834	-0.0185	0.9992	-0.023	0.9877
150	-0.0067	1.0312	0.0208	0.9798	0.0191	0.9923	0.0068	0.9977
200	0.0081	0.9752	-0.0271	0.9767	-0.0012	0.9817	0.0042	0.9924
$\overline{G}_n(f_1)$	**Gaussian**							
50	0.0064	0.9928	-0.0172	1.0451	0.0064	1.0145	0.0248	1.0085
100	0.0204	0.9853	-0.0105	0.9678	0.0201	1.0295	-0.0036	1.0107
150	0.0156	1.0115	-0.0024	0.9977	0.0143	0.9766	-0.0002	1.0046
200	0.0091	0.9842	-0.0201	0.9863	-0.0087	1.0251	0.0107	0.9621
$\overline{G}_n(f_1)$	**Non-Gaussian**							
50	0.0036	1.0309	-0.0089	1.0246	-0.0024	1.0002	-0.0032	1.0283
100	-0.0101	0.9941	-0.002	1.0386	-0.0238	0.9857	0.0002	1.023
150	-0.0131	1.0129	0.0031	0.9589	0.0012	0.9781	0.0106	1.0162
200	0.0199	0.998	-0.0177	1.0273	-0.0151	1.0115	0.0132	1.0278
$\overline{G}_n(f_2)$	**Gaussian**							
50	-0.0077	1.1137	-0.0114	1.1008	0.0085	1.1179	0.0116	1.0816
100	0.016	1.022	0.0128	1.0405	-0.0093	1.0207	-0.0118	1.0573
150	-0.0159	1.0203	-0.016	1.0384	0.0174	1.0538	0.0067	0.9585
200	-0.0017	1.0158	-0.0159	1.0257	-0.0038	1.0273	0.0192	1.0463
$\overline{G}_n(f_2)$	**Non-Gaussian**							
50	0.0049	1.0541	0.0208	1.0843	0.0443	1.0284	0.0107	1.0174
100	0.0115	1.06	0.0475	1.0313	-0.0048	1.0689	0.0051	1.0027
150	0.0059	1.0479	0.0317	1.0151	0.0309	1.0301	0.0184	1.0596
200	0.0209	0.9741	-0.0142	1.0178	-0.0019	0.9765	0.0304	1.0195
$\overline{G}_n(f_3)$	**Gaussian**							
50	0.0511	1.129	0.0423	1.1733	0.0584	1.1513	0.0861	1.1456
100	0.0485	1.0639	0.0198	1.0319	0.0488	1.0817	0.0252	1.0579
150	0.0288	1.0518	0.0189	1.0393	0.032	1.0263	0.0348	1.1
200	0.0128	1.0134	0.0055	1.0179	0.0215	1.0218	0.0386	0.9725
$\overline{G}_n(f_3)$	**Non-Gaussian**							

REFERENCES

BAI, Z. and SILVERSTEIN, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probability* **32**. https://doi.org/10.1214/aop/1078415845

BAI, Z. and SILVERSTEIN, J. W. (2010). *Spectral analysis of large dimensional random matrices*, 2nd ed. Springer.

BAO, Z. (2015). On asymptotic expansion and central limit theorem of linear eigenvalue statistics for sample covariance matrices when $N/M \to 0$. *Theory of Probability & Its Applications* **59** 185–207. https://doi.org/10.1137/s0040585x97t987089

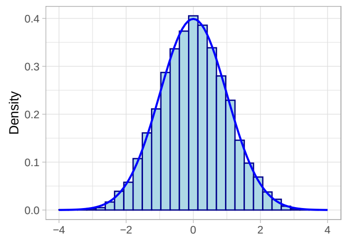
BILLINGSLEY, P. (1968). *Convergence of probability measures*. New York: Wiley.

BILLINGSLEY, P. (2008). *Probability and measure*. John Wiley & Sons.

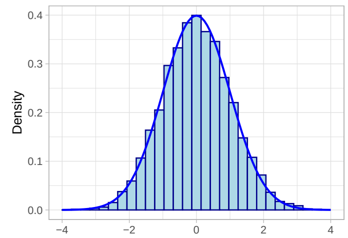
BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *The Annals of Probability* **1** 19–42. https://doi.org/10.1214/aop/1176997023

CHEN, B. and PAN, G. (2012). Convergence of the largest eigenvalue of normalized sample covariance matrices when p and n both tend to infinity with their ratio converging to zero. *Bernoulli* **18**. https://doi.org/10.3150/11-bej381

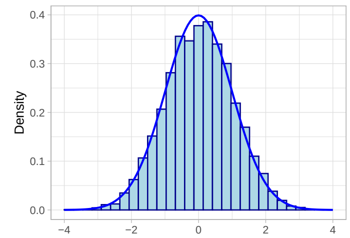
CHEN, B. and PAN, G. (2015). CLT for linear spectral statistics of normalized sample covariance matrices with the dimension much larger than the sample size. *Bernoulli* **21** 1089–1133. https://doi.org/10.3150/14-bej599



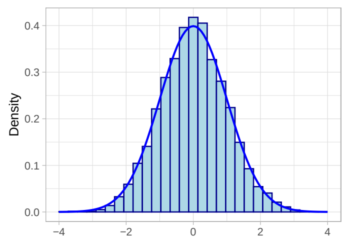
(a) $f_1 = x$, Gaussian data



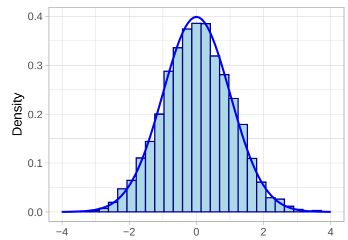
(b) $f_2 = x^2$, Gaussian data



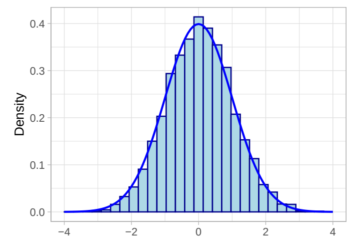
(c) $f_3 = x^3$, Gaussian data



(d) $f_1 = x$, non-Gaussian data



(e) $f_2 = x^2$, non-Gaussian data



(f) $f_3 = x^3$, non-Gaussian data

Fig S.2: Histograms of $\overline{G}_n(f_i), i = 1, 2, 3$ from 5000 replications under the case (D) with $(p, n) = (200^{2.5}, 200)$. The curves are density functions of standard normal distribution.

KHORUNZHY, A. M., KHORUZHENKO, B. A. and PASTUR, L. A. (1996). Asymptotic properties of large random matrices with independent entries. *Journal of Mathematical Physics* **37** 5033–5060. https://doi.org/10.1063/1.531589

LI, Z. and YAO, J. (2016). Testing the sphericity of a covariance matrix when the dimension is much larger than the sample size. *Electronic Journal of Statistics* **10** 2973–3010. https://doi.org/10.1214/16-ejs1199

PAN, G. and ZHOU, W. (2011). Central limit theorem for Hotelling's T^2 statistic under large dimension. *The Annals of Applied Probability* **21** 1860–1910. https://doi.org/10.1214/10-aap742

WANG, Q. and YAO, J. (2013). On the sphericity test with large-dimensional observations. *Electronic Journal of Statistics* **7** 2164–2192. https://doi.org/10.1214/13-ejs842