Hypothesis Testing for Block-structured Correlation for High-dimensional Variables

Shurong Zheng[1], Xuming He[2] and Jianhua Guo[1]

[1] *Northeast Normal University, China and* [2] *University of Michigan, USA*

Supplementary Material 1

The first supplementary material consists of the proofs of Lemma 1 and Theorem 1-3, and Tables 3-4-5-6. Lemma 1, Theorem 1-3 and the simulation settings of Tables 3-4-5-6 are in the main paper.

$$H_0 : \mathbf{\Sigma} = \text{diag}(\mathbf{\Sigma}_{11}, \ldots, \mathbf{\Sigma}_{KK}). \tag{1}$$

The one-sided rejection region for H_0 at the nominal level α is

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n : T_n - \hat{\mu} > \hat{\sigma}_0 q_{1-\alpha}\}, \tag{2}$$

where q_α is the α-th quantile of the standard normal distribution. To check the sensitivity of the threshold $s^*(n,p)$ and any scaled version of T_{n0}, we consider the rejection region

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n : T_n(c_1, c_2) - \hat{\mu} > \hat{\sigma}_0 q_{1-\alpha}\}, \tag{3}$$

where $\hat{\mu}$ is in (4), $\hat{\sigma}_0$ is just before (2.4) of the main paper and $T_n(c_1, c_2) = T_{n1} + c_1 \cdot T_{n0}(c_2)$, $\quad T_n = T_n(1,1)$ with $T_{n1} = \text{tr}[\mathbf{S}_n - \text{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})]^2$,

$$\begin{aligned} T_{n0}(c_2) &= p^2 \delta_{\{\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2 (\hat{\theta}_{\ell_1\ell_2})^{-1} > s^*(n,p,c_2)\}}, \\ s^*(n,p,c_2) &= c_2 \cdot [4 + (\log\log n - 1)^2](\log p_0 - 0.25 \log\log p_0) + q. \end{aligned}$$

Lemma 1. *Under Assumption [A]-[B], and under H_0 specified by (1), we have*

$$\frac{T_{n1}-\mu}{\sigma} \to N(0,1) \quad \text{and} \quad \frac{T_{n1}-\hat{\mu}}{\sigma_0} \to N(0,1),$$

where

$$\begin{aligned}
\mu &= \frac{(n^2-n-1)[(\mathrm{tr}\boldsymbol{\Sigma})^2 - \sum_{k=1}^{K}(\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2]}{n(n-1)^2}, \\
\hat{\mu} &= \frac{(n^2-n-1)[(\mathrm{tr}\mathbf{S}_n)^2 - \sum_{k=1}^{K}(\mathrm{tr}\mathbf{S}_{kk})^2]}{n(n-1)^2}, \\
\sigma_0^2 &= 4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2)^2 - 4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2)^2, \\
\sigma^2 &= \sigma_0^2 + 4n^{-3}\sum_{k=1}^{K}(\mathrm{tr}\boldsymbol{\Sigma}_{kk} - \mathrm{tr}\boldsymbol{\Sigma})^2\left[2\,\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2 + \beta_w\sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_{\ell k})^2\right], \\
\beta_w &= \mathrm{E}(w_{ji}^4) - 3.
\end{aligned} \tag{4}$$

Here $\mathbf{e}_\ell$ is a p-dimensional vector with the ℓ-th element being one and other elements being zeros and $\mathbf{e}_{\ell k}$ is a p_k-dimensional vector with the ℓ-th element being one and other elements being zeros.

Theorem 1. *Under Assumptions [A]-[B], and under H_0 specified by (1), if*

$$\liminf_{n\to\infty} \inf_{(i,j)\in A_0} \mathrm{var}[(x_{1i}-\mathrm{E}x_{1i})(x_{1j}-\mathrm{E}x_{1j})][\mathrm{var}(x_{1i})\mathrm{var}(x_{1j})]^{-1/2} > 0,$$

$s^*(n,p) - 4\log p_0 \to +\infty$, and $\sup_{1\le\ell\le p} \mathrm{E}\exp(t_0|x_{\ell 1}|^{m_0}) < \infty$ for some constants $t_0 > 0$ and $0 < m_0 \le 2$, we have

$$\hat{\sigma}_0^{-1}(T_n - \hat{\mu}) \to N(0,1).$$

Theorem 2. *Under Assumptions [A]-[B], we have*

$$\sigma_1^{-1}(T_{n1} - \hat{\mu} - \mu_1) \to N(0,1)$$

where $\mu_1 = (n^2-n+2)\mathrm{tr}\mathbf{A}/(n-1)^2$, $\sigma_1^2 = \sigma_0^2 + 4[2n^{-1}\mathrm{tr}\mathbf{A}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\mathbf{A}\mathbf{e}_\ell)^2]$. Here $\mathbf{e}_\ell$

is the p-dimensional vector with the ℓth element being one and other elements being zeros and $\beta_w = \mathrm{E}w_{ij}^4 - 3$.

Theorem 3. *Under Assumptions [A]-[B] and $\boldsymbol{\Sigma}^2 = \mathrm{diag}(\boldsymbol{\Sigma}_{11}^2, \ldots, \boldsymbol{\Sigma}_{KK}^2) + \mathbf{A}$,*

(1). We have $\beta_{T_n}(\mathbf{A}) \geq \alpha$ when n is large enough; Especially, when $\mathrm{tr}\mathbf{A} > \epsilon_0 > 0$ for any positive constant ϵ_0, we have $\beta_{T_n}(\mathbf{A}) > \alpha$ for sufficiently large n;

(2). If $\mathrm{tr}\mathbf{A}$ tends to infinity or $P(\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2(\hat{\theta}_{\ell_1\ell_2})^{-1} > s^*(n,p))$ converges to one, then we have $\beta_{T_n}(\mathbf{A}) \to 1$ as $n \to \infty$.

S1 Tables 3-4-5-6

S2 Proofs of Lemma 1 and Theorem 1-2-3

Define $\mathbf{r}_i = n^{-1/2}\mathbf{w}_i$, $\mathbf{w}_i = (w_{1i}, \ldots, w_{pi})^\top$, $\mathbf{r}_{ik} = n^{-1/2}\mathbf{w}_{ik}$, $\mathbf{w}_{ik} = (w_{\tilde{p}_{k-1}+1,i}, \ldots, w_{\tilde{p}_k,i})^\top$ with $\tilde{p}_0 = 0$ and $\tilde{p}_k = p_1 + \ldots + p_k$ for $k = 1, \ldots, K,\ i = 1, \ldots, n$. Then $\mathbf{r}_i = (\mathbf{r}_{i1}^\top, \ldots, \mathbf{r}_{iK}^\top)^\top$ and $\mathbf{w}_i = (\mathbf{w}_{i1}^\top, \ldots, \mathbf{w}_{iK}^\top)^\top$ for $i = 1, \ldots, n$. We have

$$(n-1)^2 n^{-2}\mathrm{tr}(\mathbf{S}_n^2) = \mathrm{tr}[(\sum_{i=1}^n \boldsymbol{\Sigma}^{1/2}\mathbf{r}_i\mathbf{r}_i^\top\boldsymbol{\Sigma}^{1/2})^2] + n^2(\bar{\mathbf{r}}^\top\boldsymbol{\Sigma}\bar{\mathbf{r}})^2 - 2n\bar{\mathbf{r}}^\top\boldsymbol{\Sigma}\sum_{i=1}^n \mathbf{r}_i\mathbf{r}_i^\top\boldsymbol{\Sigma}\bar{\mathbf{r}},$$

where $\bar{\mathbf{r}} = n^{-1}\sum_{i=1}^n \mathbf{r}_i$. By Lemma S.2.1 and S.2.2 from the supplementary file 2, letting ϵ be a very small positive number, we have $n^2(\bar{\mathbf{r}}^\top\boldsymbol{\Sigma}\bar{\mathbf{r}})^2 = (n-1)n^{-3}(\mathrm{tr}\boldsymbol{\Sigma})^2 + o_p(n^{-(1-\epsilon)})$, and

$$n\bar{\mathbf{r}}^\top\boldsymbol{\Sigma}\sum_{i=1}^n \mathbf{r}_i\mathbf{r}_i^\top\boldsymbol{\Sigma}\bar{\mathbf{r}} = (n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 + (n-1)n^{-2}\mathrm{tr}(\boldsymbol{\Sigma}^2) + o_p(n^{-(1-\epsilon)}).$$

Table 3: Empirical test sizes and empirical powers (in percentage) of comparison of three methods with with $(p_1, \ldots, p_K) = (p/K, \ldots, p/K)$ and $K = 2, 3$ for Gamma variables. The vector $(\theta_1, \theta_2, \theta_3)$ specifies the Σ matrix. The rejection region is given in (2). When a test is not applicable, the corresponding entries are marked $-$.

$(\theta_1, \theta_2, \theta_3)$	n	Methods	p=180	360	900	180	360	900
				$K = 2$			$K = 3$	
			Empirical test sizes					
$(0, 0, 0)$	150	FDS	4.86	4.90	4.44	5.12	4.99	4.54
		BHPZ	4.46	—	—	5.22	4.90	—
		YHN	4.94	5.36	5.48	5.30	5.29	5.06
	300	FDS	4.92	4.82	4.81	4.92	5.02	4.84
		BHPZ	4.76	4.94	—	5.38	5.14	—
		YHN	4.84	4.92	5.10	5.02	5.22	4.90
			Empirical powers					
$(1, 0, 0)$	150	FDS	33.98	22.44	13.36	52.32	33.08	19.24
		BHPZ	5.89	—	—	8.02	5.02	—
		YHN	8.82	7.42	5.88	12.88	7.88	6.38
	300	FDS	95.56	90.78	81.34	99.58	99.02	95.42
		BHPZ	8.34	5.86	—	14.17	7.26	—
		YHN	13.28	8.76	6.12	22.12	11.93	7.16
$(0, 1, 0)$	150	FDS	59.82	47.54	31.44	79.76	67.61	47.70
		BHPZ	9.48	—	—	21.44	7.40	—
		YHN	8.44	6.92	5.76	10.30	7.78	6.10
	300	FDS	99.08	98.46	96.04	99.98	99.96	99.86
		BHPZ	31.06	11.00	—	74.36	27.98	—
		YHN	11.40	8.00	5.84	16.78	10.36	6.62
$(0, 0, 1)$	150	FDS	75.24	98.62	100	83.14	99.28	100
		BHPZ	7.40	—	—	11.42	6.62	—
		YHN	77.30	98.86	100	84.34	99.39	100
	300	FDS	99.38	100	100	99.74	100	100
		BHPZ	14.78	8.02	—	34.00	19.76	—
		YHN	99.50	100	100	99.78	100	100

Table 4: Empirical test sizes and powers (in percentage) for comparison of four methods with $n = 200$, $(p_1, \ldots, p_K) = (p/K, \ldots, p/K)$ and $K = 2, 3$ for Gaussian variables. The vector $(\theta_1, \theta_2, \theta_3)$ specifies the Σ matrix. The rejection region is given in (3).

(c_1, c_2)	$(\theta_1, \theta_2, \theta_3)$	Methods	$p = 60$	120	180	60	120	180
			$K = 2$			$K = 3$		
			Empirical test sizes					
(0.001, 1)	(0, 0, 0)	FDS	5.61	5.48	5.65	5.77	5.73	5.20
		CLRT	5.15	5.36	5.38	5.26	5.49	5.29
		BHPZ	5.20	5.08	4.88	4.86	5.29	5.15
		YHN	5.32	5.36	5.54	5.55	5.58	4.87
(5, 1)	(0, 0, 0)	FDS	5.61	5.48	5.65	5.77	5.73	5.20
		CLRT	5.15	5.36	5.38	5.26	5.49	5.29
		BHPZ	5.20	5.08	4.88	4.86	5.29	5.15
		YHN	5.32	5.36	5.54	5.55	5.58	4.87
			Empirical powers					
(0.001, 1)	(1, 0, 0)	FDS	87.63	77.34	70.30	98.20	93.21	88.66
		CLRT	19.54	9.78	7.07	38.47	14.27	8.51
		BHPZ	17.27	9.08	6.69	35.03	14.41	9.75
		YHN	27.55	13.91	9.60	52.16	22.80	14.83
(5, 1)	(1, 0, 0)	FDS	87.77	77.34	70.30	98.21	93.21	88.66
		CLRT	19.54	9.78	7.07	38.47	14.27	8.51
		BHPZ	17.27	9.08	6.69	35.03	14.41	9.75
		YHN	27.55	13.91	9.60	52.16	22.80	14.83

Table 5: Empirical test sizes and powers (in percentage) for comparison of four methods with $n = 200$, $(p_1, \ldots, p_K) = (p/K, \ldots, p/K)$ and $K = 2, 3$ for Gaussian variables. The vector $(\theta_1, \theta_2, \theta_3)$ specifies the Σ matrix. The rejection region is given in (3).

(c_1, c_2)	$(\theta_1, \theta_2, \theta_3)$	Methods	$p = 60$	120	180	60	120	180
				$K = 2$			$K = 3$	
			Empirical test sizes					
(1, 0.5)	(0, 0, 0)	FDS	20.53	29.74	39.05	22.40	32.66	41.39
		CLRT	5.15	5.36	5.38	5.26	5.49	5.29
		BHPZ	5.20	5.08	4.88	4.86	5.29	5.15
		YHN	5.32	5.36	5.54	5.55	5.58	4.87
(1, 2)	(0, 0, 0)	FDS	5.43	5.35	5.55	5.60	5.55	5.12
		CLRT	5.15	5.36	5.38	5.26	5.49	5.29
		BHPZ	5.20	5.08	4.88	4.86	5.29	5.15
		YHN	5.32	5.36	5.54	5.55	5.58	4.87
			Empirical powers					
(1, 0.5)	(1, 0, 0)	FDS	99.10	98.47	97.91	99.97	99.93	99.93
		CLRT	19.54	9.78	7.07	38.47	14.27	8.51
		BHPZ	17.27	9.08	6.69	35.03	14.41	9.75
		YHN	27.55	13.91	9.60	52.15	22.80	14.82
(1, 2)	(1, 0, 0)	FDS	40.26	20.03	13.34	64.40	31.24	19.74
		CLRT	19.54	9.78	7.07	38.47	14.27	8.51
		BHPZ	17.27	9.08	6.69	35.03	14.41	9.75
		YHN	27.55	13.91	9.60	52.15	22.80	14.82

Table 6: Empirical test sizes and powers (in percentage) for comparison of four methods with $n = 200$, $(p_1, \ldots, p_K) = (p/K, \ldots, p/K)$ and $K = 2, 3$ for Gaussian variables. The vector $(\theta_1, \theta_2, \theta_3)$ specifies the Σ matrix. The rejection region is given in (2).

n	$(\theta_1, \theta_2, \theta_3)$	Methods	$p = 6$	12	18	6	12	18
			$K = 2$			$K = 3$		
			Empirical test sizes					
600	(0, 0, 0)	FDS	6.65	6.28	5.92	6.82	6.12	5.87
		CLRT	6.51	6.13	5.65	6.68	5.90	5.67
		BHPZ	6.46	6.09	5.59	6.69	5.93	5.50
		YHN	6.57	5.97	5.65	6.72	5.92	5.64
750	(0, 0, 0)	FDS	6.36	6.22	5.84	6.46	6.12	6.36
		CLRT	6.48	5.99	5.81	6.49	5.84	6.19
		BHPZ	6.45	5.99	5.72	6.46	5.82	6.23
		YHN	6.35	6.04	5.79	6.39	6.00	6.19
1000	(0, 0, 0)	FDS	6.54	6.07	6.05	6.54	5.87	6.36
		CLRT	6.29	5.86	5.96	6.49	5.69	6.10
		BHPZ	6.26	5.83	5.90	6.39	5.67	6.21
		YHN	6.51	6.01	5.91	6.59	5.87	6.21

Thus, we have

$$\mathrm{tr}(\mathbf{S}_n^2)=\frac{n^2}{(n-1)^2}\mathrm{tr}[(\sum_{i=1}^n\mathbf{\Sigma}^{1/2}\mathbf{r}_i\mathbf{r}_i^\top\mathbf{\Sigma}^{1/2})^2]-\frac{n+1}{n(n-1)^2}(\mathrm{tr}\mathbf{\Sigma})^2-\frac{2}{n-1}\mathrm{tr}(\mathbf{\Sigma}^2)+o_p(n^{-(1-\epsilon)}).$$

Because $\mathrm{tr}\mathbf{S}_n=n(n-1)^{-1}(\sum_{i=1}^n\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i-n\bar{\mathbf{r}}^\top\mathbf{\Sigma}\bar{\mathbf{r}})$, we have $\mathrm{tr}\mathbf{S}_n=n(n-1)^{-1}\sum_{i=1}^n\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i-(n-1)^{-1}\mathrm{tr}\mathbf{\Sigma}+o_p(n^{-(1-\epsilon)})$ by Lemma S.2.1 from the supplementary file 2. As shown in Bai and Silverstein (2004) (p. 559-560),

$$\mathrm{tr}[(\sum_{i=1}^n\mathbf{\Sigma}^{1/2}\mathbf{r}_i\mathbf{r}_i^\top\mathbf{\Sigma}^{1/2})^q]-\mathrm{tr}[(\sum_{i=1}^n\mathbf{\Sigma}^{1/2}\tilde{\mathbf{r}}_i\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}^{1/2})^q]=o_p(n^{-1/4}),\ q=1,2$$

where $\tilde{\mathbf{r}}_i=n^{-1/2}\tilde{\mathbf{w}}_i$, $\tilde{\mathbf{w}}_i=(\tilde{w}_{1i},\ldots,\tilde{w}_{pi})^\top$,

$$\tilde{w}_{\ell i}=[\mathrm{Var}(w_{\ell i}\delta_{\{|w_{\ell i}|\le\sqrt{n}\eta_n\}})]^{-1/2}(w_{\ell i}\delta_{\{|w_{\ell i}|\le\sqrt{n}\eta_n\}}-\mathrm{E}w_{\ell i}\delta_{\{|w_{\ell i}|\le\sqrt{n}\eta_n\}}),$$

$|\tilde{w}_{\ell i}|\le c\sqrt{n}\eta_n$, $\mathrm{E}\tilde{w}_{\ell i}=0$, $\mathrm{E}(\tilde{w}_{\ell i}^2)=1$ and $\mathrm{E}(\tilde{w}_{\ell i}^4)<\infty$ for $\ell=1,\ldots,p$ and $i=1,\ldots,n$ with $\eta_n\downarrow 0$, $n^{1/4}\eta_n\to\infty$ and c being a positive constant. For simplicity, we shall rename the variables $\tilde{w}_{\ell i}$ simply as $w_{\ell i}$ and proceed by assuming that $|w_{\ell i}|\le\sqrt{n}\eta_n$, $\mathrm{E}w_{\ell i}=0$, $\mathrm{E}(w_{\ell i}^2)=1$ and $\mathrm{E}(w_{\ell i}^4)<\infty$ with $\eta_n\downarrow 0$ and $n^{1/4}\eta_n\to\infty$. Let $\mathbf{B}_n=\sum_{i=1}^n\mathbf{\Sigma}^{1/2}\mathbf{r}_i\mathbf{r}_i^\top\mathbf{\Sigma}^{1/2}$, then

$$\mathrm{tr}(\mathbf{S}_n^2)=\frac{n^2}{(n-1)^2}\mathrm{tr}(\mathbf{B}_n^2)-\frac{n+1}{n(n-1)^2}(\mathrm{tr}\mathbf{\Sigma})^2-\frac{2}{n-1}\mathrm{tr}(\mathbf{\Sigma}^2)+o_p(n^{-1/4}).\tag{S2.1}$$

Similarly, let $\mathbf{B}_n=\sum_{i=1}^n\mathbf{\Sigma}_{kk}^{1/2}\mathbf{r}_{ik}\mathbf{r}_{ik}^\top\mathbf{\Sigma}_{kk}^{1/2}$, then $\mathrm{tr}\mathbf{S}_{kk}=n(n-1)^{-1}\sum_{i=1}^n\mathbf{r}_{ik}^\top\mathbf{\Sigma}_{kk}\mathbf{r}_{ik}-(n-1)^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk}+o_p(1)$ and

$$\mathrm{tr}(\mathbf{S}_{kk}^2)=\frac{n^2}{(n-1)^2}\mathrm{tr}(\mathbf{B}_{kk}^2)-\frac{n+1}{n(n-1)^2}(\mathrm{tr}\mathbf{\Sigma}_{kk})^2-\frac{2}{n-1}\mathrm{tr}(\mathbf{\Sigma}_{kk}^2)+o_p(n^{-1/4}),\tag{S2.2}$$

where $o_p(n^{-1/4})$ is uniform for $k = 1, \ldots, K$.

S2.1 Part I of Lemma 1 and its proof

Lemma 2. *Under Assumption [A]-[B] and under* $H_0 : \mathbf{\Sigma} = \mathrm{diag}(\mathbf{\Sigma}_{11}, \ldots, \mathbf{\Sigma}_{KK})$, *we have* $\sigma^{-1}(T_{n1} - \mu) \to N(0, 1)$, *where the quantities* μ *and* σ *are given in Lemma 1 in the main paper.*

Proof of Lemma 2. First note that $T_{n1} = \mathrm{tr}[\mathbf{S}_n - \mathrm{diag}(\mathbf{S}_{11}, \ldots, \mathbf{S}_{KK})]^2 = \mathrm{tr}(\mathbf{S}_n^2) - \sum_{k=1}^K \mathrm{tr}(\mathbf{S}_{kk}^2)$. By (S2.1) and (S2.2), we have

$$T_{n1} = \frac{n^2}{(n-1)^2}[\mathrm{tr}(\mathbf{B}_n^2) - \sum_{k=1}^K \mathrm{tr}(\mathbf{B}_{kk}^2)] - \frac{n+1}{n(n-1)^2}[(\mathrm{tr}\mathbf{\Sigma})^2 - \sum_{k=1}^K (\mathrm{tr}\mathbf{\Sigma}_{kk})^2] - \frac{2}{n-1}[\mathrm{tr}(\mathbf{\Sigma}^2) - \sum_{k=1}^K \mathrm{tr}(\mathbf{\Sigma}_{kk}^2)] + o_p(n^{-1/4}). \tag{S2.3}$$

Under H_0, we have

$$T_{n1} = \frac{n^2}{(n-1)^2}[\mathrm{tr}(\mathbf{B}_n^2) - \sum_{k=1}^K \mathrm{tr}(\mathbf{B}_{kk}^2)] - \frac{n+1}{n(n-1)^2}[(\mathrm{tr}\mathbf{\Sigma})^2 - \sum_{k=1}^K (\mathrm{tr}\mathbf{\Sigma}_{kk})^2] + o_p(n^{-1/4}).$$

That is, the central limit theorem for T_{n1} can be obtained by establishing the central limit theorem for $[\mathrm{tr}(\mathbf{B}_n^2) - \sum_{k=1}^K \mathrm{tr}(\mathbf{B}_{kk}^2)]$. We need to compute the mean μ and the variance σ^2 of the statistic T_{n1}. The asymptotic normality is due to the fact that $\{\mathrm{E}_j(\mathrm{tr}\mathbf{B}_n^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_n^2), j = 1, \ldots, n\}$ and $\{\mathrm{E}_j(\mathrm{tr}\mathbf{B}_{kk}^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_{kk}^2), j = 1, \ldots, n\}$ for $k = 1, \ldots, K$ are two martingale difference sequences, where we use E_j as the conditional expectation given $\mathbf{x}_1, \ldots, \mathbf{x}_j$.

Lemma S.2.3 from the supplementary file 2 shows that these martingale difference sequences satisfy the Lindeberg's conditions, that is,

$$\sum_{j=1}^{n} \mathrm{E}([\mathrm{E}_j(\mathrm{tr}\mathbf{B}_n^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_n^2)]^2 \delta_{\{|\mathrm{E}_j(\mathrm{tr}\mathbf{B}_n^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_n^2)| \geq \epsilon\}}) = O(\eta_n^4), \quad \text{(S2.4)}$$

$$\sum_{j=1}^{n} \mathrm{E}([\mathrm{E}_j(\mathrm{tr}\mathbf{B}_{kk}^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_{kk}^2)]^2 \delta_{\{|\mathrm{E}_j(\mathrm{tr}\mathbf{B}_{kk}^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_{kk}^2)| \geq \epsilon\}}) = O(\eta_n^4), \text{(S2.5)}$$

for any $\epsilon > 0$ where $O(\eta_n^4)$ is uniform for $k = 1, ..., K$. For simplicity, $\mathrm{E}_j(\mathrm{tr}\mathbf{B}_n^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_n^2)$ is often written as $(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_n^2)$ in this paper.

To compute the mean and the variance, we take the following two steps.

Step 1 computes the mean

$$\mu = \frac{n^2}{(n-1)^2}\mathrm{E}[\mathrm{tr}(\mathbf{B}_n^2) - \sum_{k=1}^{K} \mathrm{tr}(\mathbf{B}_{kk}^2)] - \frac{n+1}{n(n-1)^2}[(\mathrm{tr}\boldsymbol{\Sigma})^2 - \sum_{k=1}^{K}(\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2].$$

We have

$$\mathrm{E}[\mathrm{tr}(\mathbf{B}_n^2)] = n^{-1}[2\mathrm{tr}\boldsymbol{\Sigma}^2 + \beta_w \sum_{j=1}^{p}(\mathbf{e}_j^\top \boldsymbol{\Sigma}\mathbf{e}_j)^2] + n^{-1}(\mathrm{tr}\boldsymbol{\Sigma})^2 + (n-1)n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^2),$$

$$\mathrm{E}[\mathrm{tr}(\mathbf{B}_{kk}^2)] = n^{-1}[2\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2 + \beta_w \sum_{j=1}^{p_k}(\mathbf{e}_{jk}^\top \boldsymbol{\Sigma}_{kk}\mathbf{e}_{jk})^2] + n^{-1}(\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2 + (n-1)n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^2),$$

for $k = 1, \ldots, K$. Then under H_0, we have

$$\mu = \frac{n^2-n-1}{n(n-1)^2}(\mathrm{tr}\boldsymbol{\Sigma})^2 - \frac{n^2-n-1}{n(n-1)^2}\sum_{k=1}^{K}(\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2.$$

Step 2 shows that $\sigma^2 = \sigma_{00} + \sum_{k=1}^{K}\sigma_{kk} - 2\sum_{k=1}^{K}\sigma_{0k}$ converges in probability, where $\sigma_{00} = \sum_{j=1}^{n}\mathrm{E}_{j-1}[(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_n^2)]^2$, $\sigma_{kk} = \sum_{j=1}^{n}\mathrm{E}_{j-1}[(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_{kk}^2)]^2$, $\sigma_{0k} = \sum_{j=1}^{n}\mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_n^2)][(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_{kk}^2)]\}$

for $k = 1, \cdots, K$. To do so, we have

$$
\begin{aligned}
& (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathrm{tr} \mathbf{B}_n^2 \\
= \;& 2(n-j)n^{-1}(\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_j^\top \boldsymbol{\Sigma}^2 \mathbf{r}_j + (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j \mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j \\
& +2 \sum_{\ell \le j-1} (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_\ell \mathbf{r}_\ell^\top \boldsymbol{\Sigma} \mathbf{r}_j \\
= \;& 2(n-j)n^{-1}(\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_j^\top \boldsymbol{\Sigma}^2 \mathbf{r}_j + (\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1} \mathrm{tr} \boldsymbol{\Sigma})^2 - \mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1} \mathrm{tr} \boldsymbol{\Sigma})^2] \\
& +2(n^{-1} \mathrm{tr} \boldsymbol{\Sigma})(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}) + 2 \sum_{\ell \le j-1} (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_\ell \mathbf{r}_\ell^\top \boldsymbol{\Sigma} \mathbf{r}_j, \\
& (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathrm{tr} \mathbf{B}_{11}^2 \\
= \;& 2(n-j)n^{-1}(\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11}^2 \mathbf{r}_{j1} + (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} \mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} \\
& +2 \sum_{\ell \le j-1} (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{\ell 1} \mathbf{r}_{\ell 1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} \\
= \;& 2(n-j)n^{-1}(\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11}^2 \mathbf{r}_{j1} \\
& +(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11})^2 - \mathrm{E}[(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11})^2] \\
& +2(n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11})(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11}) + 2 \sum_{\ell \le j-1} (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{\ell 1} \mathbf{r}_{\ell 1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1},
\end{aligned}
$$

where

$$
\begin{aligned}
(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j)^2 - \mathrm{E}(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j)^2 &= (\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1} \mathrm{tr} \boldsymbol{\Sigma})^2 - \mathrm{E}(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1} \mathrm{tr} \boldsymbol{\Sigma})^2 \\
& \quad +2(n^{-1} \mathrm{tr} \boldsymbol{\Sigma})(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}), \\
(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1})^2 - \mathrm{E}(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1})^2 &= (\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11})^2 - \mathrm{E}(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11})^2 \\
& \quad +2(n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11})(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1} \mathrm{tr} \boldsymbol{\Sigma}_{11}).
\end{aligned}
$$

We first compute σ_{01}, and the calculations of $\{\sigma_{0k}, k = 2, \ldots, K\}$ can be similarly obtained.

$$
\begin{aligned}
\sigma_{01} &= \sum_{j=1}^{n} \mathrm{E}_{j-1}[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n^2][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{11}^2]\\
&= \sum_{j=1}^{n} \mathrm{E}_{j-1}\Big\{(\mathrm{E}_j-\mathrm{E}_{j-1})\Big[\frac{2(n-j)}{n}\mathbf{r}_j^\top\mathbf{\Sigma}^2\mathbf{r}_j+\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j+2\sum_{\ell\le j-1}\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_\ell\mathbf{r}_\ell^\top\mathbf{\Sigma}\mathbf{r}_j\Big]\\
&\quad (\mathrm{E}_j-\mathrm{E}_{j-1})\Big[\frac{2(n-j)}{n}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}+\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}+2\sum_{\ell\le j-1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{\ell1}\mathbf{r}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\Big]\Big\}\\
&= (S2.6)+(S2.7)+(S2.8),
\end{aligned}
$$

where (S2.6)-(S2.8) are given as follows.

$$
\begin{aligned}
&\sum_{j=1}^{n} 2(n-j)n^{-1}\mathrm{E}_{j-1}\big\{(\mathbf{r}_j^\top\mathbf{\Sigma}^2\mathbf{r}_j-n^{-1}\mathrm{tr}\mathbf{\Sigma}^2)(\mathrm{E}_j-\mathrm{E}_{j-1})\big[2(n-j)n^{-1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}\\
&\qquad +\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}+2\sum_{\ell\le j-1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{\ell1}\mathbf{r}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\big]\big\}, \qquad \text{(S2.6)}
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{\mathrm{E}_j-\mathrm{E}_{j-1})(\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j)(\mathrm{E}_j-\mathrm{E}_{j-1})\big[2(n-j)n^{-1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}\\
&\qquad +\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}+2\sum_{\ell\le j-1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{\ell1}\mathbf{r}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\big]\big\}, \qquad \text{(S2.7)}
\end{aligned}
$$

$$
\begin{aligned}
&2\sum_{j=1}^{n} \mathrm{E}_{j-1}\big\{\mathrm{E}_j-\mathrm{E}_{j-1})(\sum_{\ell\le j-1}\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_\ell\mathbf{r}_\ell^\top\mathbf{\Sigma}\mathbf{r}_j)(\mathrm{E}_j-\mathrm{E}_{j-1})\big[2(n-j)n^{-1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}\\
&\qquad +\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}+2\sum_{\ell\le j-1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{\ell1}\mathbf{r}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\big]\big\}. \qquad \text{(S2.8)}
\end{aligned}
$$

As verified in the supplementary file, we have

$$
\begin{aligned}
(S2.6) &= 2(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell1}]\\
&\quad +2[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^4)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})^2]+O_p(\eta_n^2),
\end{aligned}
$$

$$
\begin{aligned}
(S2.7) &= 4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})(n^{-1}\mathrm{tr}\mathbf{\Sigma})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^2)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell 1})^2] \\
&\quad +4(n^{-1}\mathrm{tr}\mathbf{\Sigma})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell 1}\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell 1}]+O_p(\eta_n^2), \\
(S2.8) &= 2(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell 1})(\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell 1})] \\
&\quad +2[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^4)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell 1})^2]+4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2+O_p(\eta_n^2).
\end{aligned}
$$

Thus under H_0, we have

$$
\begin{aligned}
\sigma_{01} &= (S2.6)+(S2.7)+(S2.8) \\
&= 4[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^4)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell 1})^2]+4[n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^2)]^2 \\
&\quad +(4n^{-1}\mathrm{tr}\mathbf{\Sigma}+4n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell 1}\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell 1}] \\
&\quad +4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})(n^{-1}\mathrm{tr}\mathbf{\Sigma})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^2)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell 1})^2]+O_p(\eta_n^2).
\end{aligned}
$$

Similarly, for $k=2,\ldots,K$, under H_0, we have

$$
\begin{aligned}
\sigma_{0k} &= 4[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{kk}^4)+\beta_w n^{-1}\sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top\mathbf{\Sigma}_{kk}^2\mathbf{e}_{\ell k})^2]+4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk}^2)^2 \\
&\quad +(4n^{-1}\mathrm{tr}\mathbf{\Sigma}+4n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{kk}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_k}\mathbf{e}_{\ell k}^\top\mathbf{\Sigma}_{kk}^2\mathbf{e}_{\ell k}\mathbf{e}_{\ell k}^\top\mathbf{\Sigma}_{kk}\mathbf{e}_{\ell k}] \\
&\quad +4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})(n^{-1}\mathrm{tr}\mathbf{\Sigma})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{kk}^2)+\beta_w n^{-1}\sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top\mathbf{\Sigma}_{kk}\mathbf{e}_{\ell k})^2]+O_p(\eta_n^2),
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{00} &= \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})(\mathrm{tr}\mathbf{B}_n^2)]^2 \\
&= 4[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^4) + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top \boldsymbol{\Sigma}^2 \mathbf{e}_\ell)^2] \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^2) + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top \boldsymbol{\Sigma}\mathbf{e}_\ell)^2] \\
&\quad +4[n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^2)]^2 + 8(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^3) + \beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top \boldsymbol{\Sigma}^2\mathbf{e}_\ell \mathbf{e}_\ell^\top \boldsymbol{\Sigma}\mathbf{e}_\ell] + o_p(1) \\
\sigma_{kk} &= \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}^2]^2 \\
&= 4n^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^4) + \beta_w \sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top \boldsymbol{\Sigma}_{kk}^2 \mathbf{e}_{\ell k})^2] \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2 n^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^2) + \beta_w \sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top \boldsymbol{\Sigma}_{kk}\mathbf{e}_{\ell k})^2] + 4[n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^2)]^2 \\
&\quad +8(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk})n^{-1}[2\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^3) + \beta_w \sum_{\ell=1}^{p_k}\mathbf{e}_{\ell k}^\top \boldsymbol{\Sigma}_{kk}^2 \mathbf{e}_{\ell k}\mathbf{e}_\ell^\top \boldsymbol{\Sigma}_{kk}\mathbf{e}_\ell] + O_p(\eta_n^2).
\end{aligned}
$$

Putting things together, we have under H_0,

$$
\begin{aligned}
\sigma^2 &= \sigma_{00} + \sum_{k=1}^{K}\sigma_{kk} - 2\sum_{k=1}^{K}\sigma_{0k} \\
&= 4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^2) + \beta_w n^{-1}\sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top \boldsymbol{\Sigma}_{kk}\mathbf{e}_{\ell k})^2] \\
&\quad +4[n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^2)]^2 - 4\sum_{k=1}^{K}[n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^2)]^2 + O_p(K\eta_n^2).
\end{aligned}
$$

S2.2 Proofs of Theorem 1, Part II of Lemma 1 and Theorem 2

Proof of Theorem 2

Under H_0, we have $(T_{n1} - \mu)/\sigma \to N(0,1)$. But

$$\mu = \frac{n^2-n-1}{n(n-1)^2}(\mathrm{tr}\boldsymbol{\Sigma})^2 - \frac{n^2-n-1}{n(n-1)^2}\sum_{k=1}^{K}(\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2$$

is unknown. We now replace $\mathrm{tr}\boldsymbol{\Sigma}$ and $\mathrm{tr}\boldsymbol{\Sigma}_{kk}$ by $\mathrm{tr}\mathbf{S}_n$ and $\mathrm{tr}\mathbf{S}_{kk}$ in μ, and establish the asymptotic distribution of

$$T_{n1} - \hat{\mu} = \mathrm{tr}\mathbf{S}_n^2 - \sum_{k=1}^{K}\mathrm{tr}\mathbf{S}_{kk}^2 - \hat{\mu}$$

where $\hat{\mu} = \frac{n^2-n-1}{n(n-1)^2}[(\mathrm{tr}\mathbf{S}_n)^2 - \sum_{k=1}^{K}(\mathrm{tr}\mathbf{S}_{kk})^2]$. By (S2.1) and (S2.2), we have

$$\begin{aligned}
& T_{n1} - \hat{\mu} \\
= \; & \frac{n^2}{(n-1)^2}[\mathrm{tr}(\mathbf{B}_n^2) - \sum_{k=1}^{K}\mathrm{tr}(\mathbf{B}_{kk}^2)] \\
& -\frac{n}{n-1}\frac{n^2-n-1}{(n-1)^3}[(\mathrm{tr}\mathbf{B}_n - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 - \sum_{k=1}^{K}(\mathrm{tr}\mathbf{B}_{kk} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2] \\
& -\frac{n+1}{n(n-1)^2}[(\mathrm{tr}\boldsymbol{\Sigma})^2 - \sum_{k=1}^{K}(\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2] - \frac{2}{n-1}[\mathrm{tr}(\boldsymbol{\Sigma}^2) - \sum_{k=1}^{K}\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^2)] + o_p(n^{-1/4}).
\end{aligned}$$

That is, the central limit theorem for $T_{n1} - \hat{\mu}$ can be obtained by establishing the central limit theorem for $(\mathrm{tr}\mathbf{B}_n^2 - \sum_{k=1}^{K}\mathrm{tr}\mathbf{B}_{kk}^2, \mathrm{tr}\mathbf{B}_n, \mathrm{tr}\mathbf{B}_{11}, \ldots, \mathrm{tr}\mathbf{B}_{KK})$. The asymptotic normality is due to the fact that the sequences $\{(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_n^2), j = 1, \ldots, n\}$, $\{(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_n), j = 1, \ldots, n\}$, $\{(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_{kk}), j = 1, \ldots, n\}$ and $\{(\mathrm{E}_j - \mathrm{E}_{j-1})(\mathrm{tr}\mathbf{B}_{kk}^2), j = 1, \ldots, n\}$ for $k = 1, \ldots, K$ are martingale difference sequences and Lindeberg-type conditions are satisfied by Lemma S.2.3 from the supplementary file 2. Then we have

$$\sigma_1^{-1}\{T_{n1} - \hat{\mu} - \mu_1\} \to N(0,1),$$

where $\mu_1 = n^2(n-1)^{-2}\mathrm{E}[\mathrm{tr}(\mathbf{B}_n^2) - \sum_{k=1}^K \mathrm{tr}(\mathbf{B}_{kk}^2)] - (n+1)n^{-1}(n-1)^{-2}[(\mathrm{tr}\mathbf{\Sigma})^2 - \sum_{k=1}^K (\mathrm{tr}\mathbf{\Sigma}_{kk})^2] - 2(n-1)^{-1}[\mathrm{tr}(\mathbf{\Sigma}^2) - \sum_{k=1}^K \mathrm{tr}(\mathbf{\Sigma}_{kk}^2)] - \mu$ and

$$\begin{aligned}
\sigma_1^2 \;=\;& \sigma_{00A} + \sum_{k=1}^K \sigma_{kkA} - 2\sum_{k=1}^K \sigma_{0kA} - 4(n^{-1}\mathrm{tr}\mathbf{\Sigma})\sigma_{000A} + 4(n^{-1}\mathrm{tr}\mathbf{\Sigma})\sum_{k=1}^K \sigma_{00kA} \\
&+4\sum_{k=1}^K (n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})\sigma_{0kkA} - 4\sum_{k=1}^K (n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})\sigma_{kkkA} + 4(n^{-1}\mathrm{tr}\mathbf{\Sigma})^2\sigma_{0000A} \\
&+4\sum_{k=1}^K (n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})^2\sigma_{kkkkA} - 8\sum_{k=1}^K (n^{-1}\mathrm{tr}\mathbf{\Sigma})(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})\sigma_{00kkA},
\end{aligned}$$

if the following terms converge in probability

$\sigma_{00A} = \sum_{\ell=1}^n \mathrm{E}_{\ell-1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n^2]^2$,

$\sigma_{0kA} = \sum_{\ell=1}^n \mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n^2][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}^2]\}$,

$\sigma_{kkA} = \sum_{\ell=1}^n \mathrm{E}_{\ell-1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}^2]^2$,

$\sigma_{0000A} = \sum_{\ell=1}^n \mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n]\}$,

$\sigma_{000A} = \sum_{\ell=1}^n \mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n^2][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n]\}$,

$\sigma_{kkkkA} = \sum_{\ell=1}^n \mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}]^2$,

$\sigma_{kkkA} = \sum_{\ell=1}^n \mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}^2][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}]\}$,

$\sigma_{0kkA} = \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n^2][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}]\}$,

$\sigma_{00kA} = \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}^2]\}$,

$\sigma_{00kkA} = \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}[(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_{kk}][(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{B}_n]$.

The first step is to compute μ_1. Because $\mathrm{E}[\mathrm{tr}(\mathbf{B}_n^2)] = n^{-1}[2\mathrm{tr}\mathbf{\Sigma}^2 + \beta_w \sum_{j=1}^{p}(\mathbf{e}_j^\top \mathbf{\Sigma}\mathbf{e}_j)^2] + n^{-1}(\mathrm{tr}\mathbf{\Sigma})^2 + (n-1)n^{-1}\mathrm{tr}(\mathbf{\Sigma}^2)$ and $\mathrm{E}[\mathrm{tr}(\mathbf{B}_{kk}^2)] = n^{-1}[2\mathrm{tr}\mathbf{\Sigma}_{kk}^2 + \beta_w \sum_{j=1}^{p_k}(\mathbf{e}_{jk}^\top \mathbf{\Sigma}_{kk}\mathbf{e}_{jk})^2] + n^{-1}(\mathrm{tr}\mathbf{\Sigma}_{kk})^2 + (n-1)n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{kk}^2)$ for $k = 1, \dots, K$, thus we have

$$\begin{aligned}\mu_1 &= n^2(n-1)^{-2}\mathrm{E}[\mathrm{tr}(\mathbf{B}_n^2) - \sum_{k=1}^{K}\mathrm{tr}(\mathbf{B}_{kk}^2)] - (n+1)n^{-1}(n-1)^{-2}[(\mathrm{tr}\mathbf{\Sigma})^2 - \sum_{k=1}^{K}(\mathrm{tr}\mathbf{\Sigma}_{kk})^2] \\ &\quad -2(n-1)^{-1}[\mathrm{tr}(\mathbf{\Sigma}^2) - \sum_{k=1}^{K}\mathrm{tr}(\mathbf{\Sigma}_{kk}^2)] - \mu = \frac{n^2-n+2}{(n-1)^2}\mathrm{tr}\mathbf{A}\end{aligned}$$

where $\mathbf{A} = \mathbf{\Sigma}^2 - \mathrm{diag}(\mathbf{\Sigma}_{11}^2, \dots, \mathbf{\Sigma}_{KK}^2)$.

The second step is to compute σ_1^2. Let $\mathbf{\Sigma}_{(kk)}$ is the $p \times p$ dimensional matrix with the kth diagonal block being $\mathbf{\Sigma}_{kk}$ and other entries being zeros. The detailed proofs of σ_{00A}, σ_{0kA}, σ_{kkA}, σ_{0000A}, σ_{000A}, σ_{kkkkA}, σ_{kkkA}, σ_{0kkA}, σ_{00kA} and σ_{00kkA} are similar for $k = 1, \dots, K$. Moreover, the proof of σ_{01A} is similar to σ_{01}. Therefore, we do not give the details of the proofs of σ_{01A}. We have

$$\sigma_{01A} = \sum_{j=1}^{n} \mathrm{E}_{j-1}[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}\mathbf{S}_n^2][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}\mathbf{S}_{11}^2] = (S2.6) + (S2.7) + (S2.8)$$

where under the alternative hypothesis,

$$
\begin{aligned}
(S2.6) &= 2[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(11)}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell)(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}^2\mathbf{e}_\ell)] \\
&\quad +2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(11)}[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(11)} + \beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}\mathbf{e}_\ell] + O_p(\eta_n^4), \\
(S2.7) &= 4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_{(11)}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}^2\mathbf{e}_\ell\big) \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(11)})(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\big[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_{(11)} + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell)(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}\mathbf{e}_\ell)\big] + O_p(\eta_n^4), \\
(S2.8) &= 2[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(11)}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell)(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}^2\mathbf{e}_\ell)] \\
&\quad +2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(11)})\big[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(11)} + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell)(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}\mathbf{e}_\ell)\big] + O_p(\eta_n^4).
\end{aligned}
$$

Therefore, under the alternative hypothesis, we have

$$
\begin{aligned}
\sigma_{01A} &= 4[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(11)}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell)(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}^2\mathbf{e}_\ell)] \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(11)}^3 + \beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}^2\mathbf{e}_\ell) \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(11)}(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(11)} + \beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}\mathbf{e}_\ell) \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(11)})(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(11)}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell)(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(11)}\mathbf{e}_\ell)] \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_{(11)})^2 + O_p(\eta_n^2).
\end{aligned}
$$

Similarly, for $k=1,\dots,K$, under the alternative hypothesis, we have

$$\begin{aligned}
\sigma_{0kA} &= \sum_{j=1}^{n} \mathrm{E}_{j-1}[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{S}_n^2][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{S}_{kk}^2]\\
&= 4\big[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(kk)}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell)(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(kk)}^2\mathbf{e}_\ell)\big]+4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(kk)}^2)^2\\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(kk)}^3+\beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(kk)}\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(kk)}^2\mathbf{e}_\ell\big)\\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(kk)})\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(kk)}+\beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(kk)}\mathbf{e}_\ell\big)\\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(kk)})(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\big[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{(kk)}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(kk)}\mathbf{e}_\ell)^2\big]+O_p(\eta_n^2),
\end{aligned}$$

$$\begin{aligned}
\sigma_{00A} &= \sum_{\ell=1}^{n}\mathrm{E}_{\ell-1}[(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_n^2]^2\\
&= 4n^{-1}[2\mathrm{tr}\boldsymbol{\Sigma}^4+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell)^2]+4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2n^{-1}\big[2\mathrm{tr}\boldsymbol{\Sigma}^2+\beta_w\sum_{\ell}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell)^2\big]\\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2)^2+8(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})n^{-1}\big[2\mathrm{tr}\boldsymbol{\Sigma}^3+\beta_w\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell\big]+O_p(\eta_n^2),
\end{aligned}$$

$$\begin{aligned}
\sigma_{kkA} &= \sum_{\ell=1}^{n}\mathrm{E}_{\ell-1}[(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_{kk}^2]^2\\
&= 4n^{-1}[2\mathrm{tr}\boldsymbol{\Sigma}_{kk}^4+\beta_w\sum_{\ell=1}^{p_k}(\mathbf{e}_{\ell k}^\top\boldsymbol{\Sigma}_{kk}^2\mathbf{e}_{\ell k})^2]+4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk})^2n^{-1}\big[2\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2+\beta_w\sum_{\ell}^{p_k}(\mathbf{e}_{\ell k}^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_{\ell k})^2\big]\\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2)^2+8(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk})n^{-1}\big[2\mathrm{tr}\boldsymbol{\Sigma}_{kk}^3+\beta_w\sum_{\ell=1}^{p_k}\mathbf{e}_{\ell k}^\top\boldsymbol{\Sigma}_{kk}^2\mathbf{e}_{\ell k}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_\ell\big]+O_p(\eta_n^2),\\
\sigma_{0000A} &= \sum_{\ell=1}^{n}\mathrm{E}_{\ell-1}\{[(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_n][(\mathrm{E}_\ell-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_n]\}\\
&= 2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^2)+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell)^2+O_p(\eta_n^2),
\end{aligned}$$

$$\begin{aligned}
\sigma_{000A} &= \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}\{[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_n^2][(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_n]\} \\
&= 2\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^3+\beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell\big) \\
&\quad +2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell)^2\big)+O_p(\eta_n^2), \\
\sigma_{kkkkA} &= \sum_{j=1}^{n} \mathrm{E}_{j-1}\{[(\mathrm{E}_{j}-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{S}_{kk}]^2 = 2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{kk}^2)+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_\ell)^2+O_p(\eta_n^2), \\
\sigma_{kkkA} &= \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}\{[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_{kk}^2][(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_{kk}]\} \\
&= 2n^{-1}(2\mathrm{tr}\boldsymbol{\Sigma}_{kk}^3+\beta_w\sum_{\ell=1}^{p}\mathbf{e}_{k\ell}^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_{k\ell}\mathbf{e}_{k\ell}^\top\boldsymbol{\Sigma}_{kk}^2\mathbf{e}_{k\ell}) \\
&\quad +2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk}[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_{k\ell}^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_{k\ell})^2]+O_p(\eta_n^2),
\end{aligned}$$

$$\begin{aligned}
\sigma_{0kkA} &= \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}\{[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_n^2][(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_{kk}]\} \\
&= 2\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}_{(kk)}+\beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}^2\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(kk)}\mathbf{e}_\ell\big) \\
&\quad +2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_\ell)^2\big)+O_p(\eta_n^2),
\end{aligned}$$

$$\begin{aligned}
\sigma_{00kA} &= \sum_{\ell=1}^{n} \mathrm{E}_{\ell-1}\{[(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_n][(\mathrm{E}_{\ell}-\mathrm{E}_{\ell-1})\mathrm{tr}\mathbf{S}_{kk}^2]\} \\
&= 2\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_{(kk)}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_\ell^\top\boldsymbol{\Sigma}\mathbf{e}_\ell\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{(kk)}^2\mathbf{e}_\ell\big) \\
&\quad +2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk})\big(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{kk}^2+\beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\boldsymbol{\Sigma}_{kk}\mathbf{e}_\ell)^2\big)+O_p(\eta_n^2),
\end{aligned}$$

$$\begin{aligned}\sigma_{00kkA} &= \sum_{j=1}^{n} \mathrm{E}_{j-1}[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}\mathbf{S}_{kk}][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}\mathbf{S}_n] \\ &= 2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{kk}^2) + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top \mathbf{\Sigma}_{kk}\mathbf{e}_\ell)^2 + O_p(\eta_n^2).\end{aligned}$$

Then we have

$$\begin{aligned}\sigma_1^2 &= \sigma_{00A} + \sum_{k=1}^{K}\sigma_{kkA} - 2\sum_{k=1}^{K}\sigma_{0kA} - 4(n^{-1}\mathrm{tr}\mathbf{\Sigma})\sigma_{000A} + 4(n^{-1}\mathrm{tr}\mathbf{\Sigma})\sum_{k=1}^{K}\sigma_{00kA} \\ &\quad +4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})\sigma_{0kkA} - 4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})\sigma_{kkkA} + 4(n^{-1}\mathrm{tr}\mathbf{\Sigma})^2\sigma_{0000A} \\ &\quad +4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})^2\sigma_{kkkkA} - 8\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\mathbf{\Sigma})(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk})\sigma_{00kkA} \\ &= 4(n^{-1}\mathrm{tr}\mathbf{\Sigma}^2)^2 - 4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk}^2)^2 + 4[2n^{-1}\mathrm{tr}\mathbf{A}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\mathbf{A}\mathbf{e}_\ell)^2] + O_p(K\eta_n^2).\end{aligned}$$

with $\mathbf{A} = \mathbf{\Sigma}^2 - \mathrm{diag}(\mathbf{\Sigma}_{11}^2, \ldots, \mathbf{\Sigma}_{KK}^2)$. Thus $\sigma_1^{-1}(T_{n1} - \hat{\mu} - \mu_1) \to N(0,1)$.

The proof of Theorem 2 is now complete.

Proof of Part II of Lemma 1

Under H_0, $\mu_1 = 0$ and $\sigma_0^2 = \sigma_1^2 = 4(n^{-1}\sum_{k=1}^{K}\mathrm{tr}\mathbf{\Sigma}_{kk}^2)^2 - 4\sum_{k=1}^{K}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk}^2)^2$. Then under H_0 and by Theorem 2, we have $\sigma_0^{-1}(T_{n1} - \hat{\mu}) \to N(0,1)$.

The proof of Lemma 1 is complete.

Proof of Theorem 1

We have $(n-2)^{-1}[\mathrm{tr}\mathbf{S}_n^2 - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_n)^2] - n^{-1}\mathrm{tr}\mathbf{\Sigma}^2 = o_p(1)$ and $(n-2)^{-1}[\mathrm{tr}\mathbf{S}_{kk}^2 - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_{kk})^2] - n^{-1}\mathrm{tr}\mathbf{\Sigma}_{kk}^2 = o_p(1)$, $k = 1, \ldots, K$. Thus under H_0, we have

$$\hat{\sigma}_0^{-1}(T_{n1} - \hat{\mu}) \to N(0,1) \tag{S2.9}$$

where $\hat{\sigma}_0^2 = 4(n-2)^{-2}\{\sum_{k=1}^K[\mathrm{tr}\mathbf{S}_{kk}^2 - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_{kk})^2]\}^2 - 4(n-2)^{-2}\sum_{k=1}^K[\mathrm{tr}\mathbf{S}_{kk}^2 - (n+2)^{-1}(\mathrm{tr}\mathbf{S}_{kk})^2]^2$.

Moreover, Xiao and Wu (2013) presented that $\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2\hat{\theta}_{\ell_1\ell_2}^{-1} - 4\log p_0 + \log\log p_0$ converges to a type I extreme value distribution under H_0. Then if the threshold $s^*(n,p)$ is taken to satisfy $s^*(n,p) - 4\log p_0 \to +\infty$, then $P(\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2\hat{\theta}_{\ell_1\ell_2}^{-1} > s^*(n,p)) \to 0$ under H_0. That is, $T_n - T_{n1} = o_p(1)$ under H_0. By (S2.9), we have

$$\hat{\sigma}_0^{-1}(T_n - \hat{\mu}) \to N(0,1).$$

The proof of Theorem 1 is complete.

S2.3 Proof of Theorem 3

When tr$\mathbf{A}$ tends to infinity, σ_1 also converges and $\mu_1 \to +\infty$. Then

$$\frac{\sigma_0 q_{1-\alpha} - \mu_1}{\sigma_1} \to -\infty.$$

Thus we have $\beta_{T_n}(\mathbf{A}) \to 1$. Moreover, if $P(\max_{(\ell_1,\ell_2)\in A_0} n(s_{\ell_1\ell_2})^2\hat{\theta}_{\ell_1\ell_2}^{-1} > s^*(n,p)) \to 1$, then $T_{n0} \to \infty$ in probability as $n \to \infty$. Then the power function will tend to one.

The proof of Lemma 3 is complete.

Supplementary material 2

This supplementary material consists of three lemmas and the detailed proofs of (S2.6)-(S2.8). These proofs are conducted under Assumption [A]-[B].

S.2.1 Lemma S.2.1-S.2.4 and their proofs

Let $\mathbf{r}_i = n^{-1/2}\mathbf{w}_i$ and ϵ be a very small positive number.

Lemma S.2.1. *Under Assumptions [A]-[B], we have*

$$n\bar{\mathbf{r}}^\top\mathbf{\Sigma}\bar{\mathbf{r}} = n^{-1}\mathrm{tr}\mathbf{\Sigma} + o_p(n^{-(0.5-\epsilon)}).$$

Proof. We have

$$n\bar{\mathbf{r}}^\top\mathbf{\Sigma}\bar{\mathbf{r}} = 2n^{-1}\sum_{i<j}\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_j + n^{-1}\sum_{i=1}^n\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i.$$

First, we have $\mathrm{E}(n^{-1}\sum_{i<j}\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_j) = 0$ and

$$\begin{aligned} & \mathrm{E}(n^{-1}\sum_{i<j}\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_j)^2 \\ = \; & (n-1)n^{-1}\mathrm{E}(\mathbf{r}_1^\top\mathbf{\Sigma}\mathbf{r}_2\mathbf{r}_2^\top\mathbf{\Sigma}\mathbf{r}_1) + n^{-2}\sum_{i<j<k<\ell}\mathrm{E}(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_k^\top\mathbf{\Sigma}\mathbf{r}_\ell) \\ & +2n^{-2}\sum_{i<j<k}\mathrm{E}(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_k) \\ \leq \; & n^{-2}\mathrm{tr}(\mathbf{\Sigma}^2) = o(n^{-2(0.5-\epsilon)}), \end{aligned}$$

for any small positive number ϵ. That is,

$$n^{-1}\sum_{i<j}\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_j = o_p(n^{-(0.5-\epsilon)}).$$

Second, we have $\mathrm{E}(n^{-1}\sum_{i=1}^n\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i) = n^{-1}\mathrm{tr}\mathbf{\Sigma}$ and

$$\begin{aligned} \mathrm{Var}(n^{-1}\sum_{i=1}^n\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i) &= n^{-1}\mathrm{E}[(\mathbf{r}_1^\top\mathbf{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\mathbf{\Sigma})^2] \\ &= n^{-2}[2\mathrm{tr}(\mathbf{\Sigma}^2) + \beta_w\sum_{j=1}^p(\mathbf{e}_j^\top\mathbf{\Sigma}\mathbf{e}_j)^2] = o(n^{-2(0.5-\epsilon)}), \end{aligned}$$

where the second equality is from (1.15) of Bai and Silverstein (2004). That is,

$$n^{-1}\sum_{i=1}^n\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma} = o_p(n^{-(0.5-\epsilon)}).$$

Thus we have

$$n\bar{\mathbf{r}}^\top\mathbf{\Sigma}\bar{\mathbf{r}} = n^{-1}\mathrm{tr}\mathbf{\Sigma} + o_p(n^{-(0.5-\epsilon)}).$$

Lemma S.2.2. *Under Assumptions [A]-[B], we have*

$$n\bar{\mathbf{r}}^T\mathbf{\Sigma}\sum_{i=1}^{n}\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Sigma}\bar{\mathbf{r}} = (n^{-1}\mathrm{tr}\mathbf{\Sigma})^2 + (n-1)n^{-2}\mathrm{tr}(\mathbf{\Sigma}^2) + o_p(1).$$

Proof. We have

$$\begin{aligned} n\bar{\mathbf{r}}^T\mathbf{\Sigma}\sum_{i=1}^{n}\mathbf{r}_i\mathbf{r}_i^\top\mathbf{\Sigma}\bar{\mathbf{r}} &= n^{-1}\sum_{i,j,\ell\ unequal}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_\ell + n^{-1}\sum_{i,j\ unequal}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i \\ &\quad +2n^{-1}\sum_{i,j\ unequal}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j + n^{-1}\sum_{i=1}^{n}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i. \end{aligned}$$

Step 1. We have

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i &= n^{-1}\sum_{i=1}^{n}(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})^2 \\ &\quad +2n^{-1}\sum_{i=1}^{n}(n^{-1}\mathrm{tr}\mathbf{\Sigma})(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma}) + (n^{-1}\mathrm{tr}\mathbf{\Sigma})^2. \end{aligned}$$

Because $n^{-1}\sum_{i=1}^{n}\mathrm{E}[(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})^2] = n^{-2}[2\mathrm{tr}(\mathbf{\Sigma}^2) + \beta_w\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}\mathbf{e}_j)^2] = o(n^{-(1-\epsilon)})$, then we have $n^{-1}\sum_{i=1}^{n}(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})^2 = o(n^{-(1-\epsilon)})$. Because $n^{-1}\sum_{i=1}^{n}\mathrm{E}(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma}) = 0$ and

$$\mathrm{Var}[n^{-1}\sum_{i=1}^{n}(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})] = n^{-3}[2\mathrm{tr}(\mathbf{\Sigma}^2) + \beta_w\sum_{j=1}^{p}(\mathbf{e}_j^T\mathbf{\Sigma}\mathbf{e}_j)^2] = o(n^{-2(1-\epsilon)}),$$

then we have $n^{-1}\sum_{i=1}^{n}(\mathbf{r}_i^\top\mathbf{\Sigma}\mathbf{r}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma}) = o_p(n^{-(1-\epsilon)})$. Thus,

$$n^{-1}\sum_{i=1}^{n}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i = (n^{-1}\mathrm{tr}\mathbf{\Sigma})^2 + o_p(n^{-(1-\epsilon)}). \tag{S.2.1}$$

Step 2. We have $n^{-1}\sum_{i,j,\ell\ unequal}\mathrm{E}(\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Sigma}\mathbf{r}_\ell)=0$ and

$$
\begin{aligned}
& \mathrm{E}(n^{-1}\sum_{i,j,\ell\ unequal}\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Sigma}\mathbf{r}_\ell)^2 \\
= & \ 2n^{-2}\sum_{i,j,\ell,k\ unequal}\mathrm{E}[\mathrm{tr}(\boldsymbol{\Sigma}^{1/2}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Sigma}\mathbf{r}_\ell\mathbf{r}_\ell^T\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^T\boldsymbol{\Sigma}^{1/2})] \\
& +2n^{-2}\sum_{i,j,\ell\ unequal}\mathrm{E}[\mathrm{tr}(\boldsymbol{\Sigma}^{1/2}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Sigma}\mathbf{r}_\ell\mathbf{r}_\ell^T\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Sigma}^{1/2})] \\
& +4n^{-2}\sum_{i,j,\ell\ unequal}\mathrm{E}[\mathrm{tr}(\boldsymbol{\Sigma}^{1/2}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Sigma}\mathbf{r}_\ell\mathbf{r}_\ell^T\boldsymbol{\Sigma}\mathbf{r}_i\mathbf{r}_\ell^T\boldsymbol{\Sigma}^{1/2})] \\
\leq & \ 2n^{-2}\mathrm{tr}(\boldsymbol{\Sigma}^4)+2n^{-1}\mathrm{E}[(\mathbf{r}_1^T\boldsymbol{\Sigma}^2\mathbf{r}_1)^2]+4\mathrm{E}[\mathrm{tr}(\boldsymbol{\Sigma}^{1/2}\mathbf{r}_1\mathbf{r}_1^T\boldsymbol{\Sigma}^2\mathbf{r}_2\mathbf{r}_2^T\boldsymbol{\Sigma}\mathbf{r}_1\mathbf{r}_2^T\boldsymbol{\Sigma}^{1/2})] \\
\leq & \ 2n^{-2}\mathrm{tr}(\boldsymbol{\Sigma}^4)+2n^{-3}[2\mathrm{tr}(\boldsymbol{\Sigma}^4)+\beta_w\sum_{j=1}^p(\mathbf{e}_j^T\boldsymbol{\Sigma}^2\mathbf{e}_j)^2+(\mathrm{tr}\boldsymbol{\Sigma}^2)^2] \\
& +8\mathrm{E}[(\mathbf{r}_1^T\boldsymbol{\Sigma}^2\mathbf{r}_2\mathbf{r}_2^T\boldsymbol{\Sigma}^2\mathbf{r}_1)]+8\mathrm{E}[(\mathbf{r}_1^T\boldsymbol{\Sigma}\mathbf{r}_2\mathbf{r}_2^T\boldsymbol{\Sigma}\mathbf{r}_1)^2] \\
= & \ 2n^{-2}\mathrm{tr}(\boldsymbol{\Sigma}^4)+2n^{-3}[2\mathrm{tr}(\boldsymbol{\Sigma}^4)+\beta_w\sum_{j=1}^p(\mathbf{e}_j^T\boldsymbol{\Sigma}^2\mathbf{e}_j)^2+(\mathrm{tr}\boldsymbol{\Sigma}^2)^2] \\
& +8n^{-2}\mathrm{tr}(\boldsymbol{\Sigma}^4)+24n^{-2}\mathrm{E}(\mathbf{r}_2^T\boldsymbol{\Sigma}^2\mathbf{r}_2)^2+8n^{-2}\beta_w\sum_{j=1}^p\mathrm{E}(\mathbf{e}_j^T\boldsymbol{\Sigma}\mathbf{r}_2)^4 \\
= & \ 10n^{-2}\mathrm{tr}(\boldsymbol{\Sigma}^4)+(2n^{-3}+24n^{-4})[2\mathrm{tr}(\boldsymbol{\Sigma}^4)+\beta_w\sum_{j=1}^p(\mathbf{e}_j^T\boldsymbol{\Sigma}^2\mathbf{e}_j)^2+(\mathrm{tr}\boldsymbol{\Sigma}^2)^2] \\
& +24n^{-4}\beta_w\sum_{j=1}^p(\mathbf{e}_j^T\boldsymbol{\Sigma}^2\mathbf{e}_j)^2+8n^{-4}\beta_w^2\sum_{j=1}^p\sum_{\ell=1}^p(\mathbf{e}_j^T\boldsymbol{\Sigma}\mathbf{e}_\ell)^4 \\
= & \ o_p(n^{-2(0.5-\epsilon)}).
\end{aligned}
$$

Then we have

$$
n^{-1}\sum_{i,j,\ell\ unequal}(\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\boldsymbol{\Sigma}\mathbf{r}_\ell)=o_p(n^{-(0.5-\epsilon)}). \tag{S.2.2}
$$

Step 3. We have $n^{-1}\sum_{i,j\ unequal} \mathrm{E}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i = (n-1)n^{-2}\mathrm{tr}(\mathbf{\Sigma}^2)$ and

$$\begin{aligned}
& n^{-2}\mathrm{E}(\sum_{i,j\ unequal} \mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i)^2 \\
= \ & 2n^{-2}\sum_{i,j\ unequal}(\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i)^2 + n^{-2}\sum_{i,j,k,\ell\ unequal}(\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i)(\mathbf{r}_\ell^T\mathbf{\Sigma}\mathbf{r}_k\mathbf{r}_k^T\mathbf{\Sigma}\mathbf{r}_\ell) \\
& +4n^{-2}\sum_{i,j,\ell\ unequal}(\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i)(\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_\ell\mathbf{r}_\ell^T\mathbf{\Sigma}\mathbf{r}_i) \\
= \ & 6n^{-3}(n-1)\mathrm{E}[(\mathbf{r}_1^T\mathbf{\Sigma}^2\mathbf{r}_1)^2] + 2n^{-3}(n-1)\beta_w\sum_{j=1}^p \mathrm{E}(\mathbf{e}_j^T\mathbf{\Sigma}\mathbf{r}_1)^4 \\
& +n^{-5}(n-1)(n-2)(n-3)[\mathrm{tr}(\mathbf{\Sigma}^2)]^2 \\
& +4n^{-5}(n-1)(n-2)[2\mathrm{tr}(\mathbf{\Sigma}^4) + \beta_w\sum_{j=1}^p(\mathbf{e}_j^T\mathbf{\Sigma}^2\mathbf{e}_j)^2 + (\mathrm{tr}\mathbf{\Sigma}^2)^2] \\
\leq \ & 6n^{-5}(n-1)\beta_w\sum_{j=1}^p(\mathbf{e}_j^T\mathbf{\Sigma}^2\mathbf{e}_j)^2 + 2n^{-5}(n-1)\beta_w^2\sum_{j=1}^p\sum_{\ell=1}^p(\mathbf{e}_j^T\mathbf{\Sigma}\mathbf{e}_\ell)^4 \\
& +n^{-5}(n-1)(n-2)(n-3)[\mathrm{tr}(\mathbf{\Sigma}^2)]^2 \\
& +2n^{-5}(n-1)(2n-1)[2\mathrm{tr}(\mathbf{\Sigma}^4) + \beta_w\sum_{j=1}^p(\mathbf{e}_j^T\mathbf{\Sigma}^2\mathbf{e}_j)^2 + (\mathrm{tr}\mathbf{\Sigma}^2)^2].
\end{aligned}$$

Then we have

$$\begin{aligned}
& \mathrm{Var}(n^{-1}\sum_{i,j\ unequal}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i) \\
= \ & n^{-2}\mathrm{E}\Big[\Big(\sum_{i,j\ unequal}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i\Big)^2\Big] - \Big(n^{-1}\sum_{i,j\ unequal}\mathrm{E}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i\Big)^2 \\
= \ & o(n^{-2(0.5-\epsilon)}).
\end{aligned}$$

That is,

$$n^{-1}\sum_{i,j\ unequal}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^T\mathbf{\Sigma}\mathbf{r}_i - (n-1)n^{-2}\mathrm{tr}(\mathbf{\Sigma}^2) = o_p(n^{-(0.5-\epsilon)}). \tag{S.2.3}$$

Step 4. By (1.8) of Bai and Silverstein (2004), there exists $\eta_n \downarrow 0$ satisfying $n^{1/4}\eta_n \to \infty$ and $\eta_n^{-4}\mathrm{E}[w_{11}^4\delta(|w_{11}| \geq \eta_n\sqrt{n})] \to 0$. Then let $\hat{\mathbf{r}}_i$ be the truncated version of $\mathbf{r}_i$, that is, $\hat{\mathbf{r}}_i^T = n^{-1/2}\hat{\mathbf{w}}_i$ with $\hat{\mathbf{w}}_i = (\hat{w}_{1i}, \ldots, \hat{w}_{pi})^\top$ and $\hat{w}_{\ell i} = w_{\ell i}\delta_{\{|w_{\ell i}| \leq \sqrt{n}\eta_n\}}$. Then we have

$\mathrm{E}\hat{w}_{11} \to 0$, $\mathrm{E}\hat{w}_{11}^2 \to 1$ and $\mathrm{Var}(\hat{w}_{11}) \to 1$ as $n \to \infty$. Let $\hat{\boldsymbol{\mu}} = n^{-1/2}(\mathrm{E}\hat{w}_{11})\mathbf{1}_p$ where $\mathbf{1}_p$ is the p-dimensional vector with all entries being ones. Because $\mathrm{E}w_{11} = 0$, then we have

$$|\mathrm{E}\hat{w}_{11}| = |\mathrm{E}[w_{11}\delta(|w_{11}| > \eta_n\sqrt{n})]| \le \eta_n^{-3}n^{-3/2}\mathrm{E}[w_{11}^4\delta(|w_{11}| > \eta_n\sqrt{n})] = o(n^{-3/2}).$$

That is

$$\hat{\boldsymbol{\mu}}^\top\hat{\boldsymbol{\mu}} = n^{-1}(\mathrm{E}\hat{w}_{11})^2\mathbf{1}_p^\top\mathbf{1}_p \le o(n^{-1/2}).$$

Because

$$\begin{aligned}
& P(n^{-1}\sum_{i,j\ unequal}\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j \ne n^{-1}\sum_{i,j\ unequal}\hat{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\mathbf{r}}_i\hat{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\mathbf{r}}_j) \\
\le\ & P(\text{for some } \ell, \mathrm{i}, \hat{w}_{\ell i} \ne w_{\ell i}) \\
\le\ & \sum_{\ell=1}^{p}\sum_{i=1}^{n} P(|w_{\ell i}| \ge \eta_n\sqrt{n}) \\
\le\ & (\eta_n\sqrt{n})^{-4}np\mathrm{E}[w_{11}^4\delta(|w_{11}| \ge \eta_n\sqrt{n})] \\
=\ & (p/n)\eta_n^{-4}\mathrm{E}[w_{11}^4\delta(|w_{11}| \ge \eta_n\sqrt{n})] \to 0
\end{aligned}$$

where the third inequality is from the Chebyshev inequality and the last equality is from (1.8) of Bai and Silverstein (2004), then we have

$$n^{-1}\sum_{i,j\ unequal}\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}\mathbf{r}_j = n^{-1}\sum_{i,j\ unequal}\hat{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\mathbf{r}}_i\hat{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\mathbf{r}}_j + o_p(1). \qquad \text{(S.2.4)}$$

Let $\tilde{\mathbf{r}}_i = (\hat{\mathbf{r}}_i - \hat{\boldsymbol{\mu}})/\sqrt{\mathrm{Var}(\hat{w}_{11})}$, then

$$\begin{aligned}
\hat{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\mathbf{r}}_i &= \mathrm{Var}(\hat{w}_{11})\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\tilde{\mathbf{r}}_i + 2\sqrt{\mathrm{Var}(\hat{w}_{11})}\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\mu}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\mu}} \\
&= \mathrm{Var}(\hat{w}_{11})(\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\tilde{\mathbf{r}}_i - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}) + n^{-1}\mathrm{Var}(\hat{w}_{11})\mathrm{tr}\boldsymbol{\Sigma} \\
&\quad +2\sqrt{\mathrm{Var}(\hat{w}_{11})}\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\boldsymbol{\mu}} + o(n^{-1/2}),
\end{aligned}$$

and $\hat{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\mathbf{r}}_j = \mathrm{Var}(\hat{w}_{11})\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\tilde{\mathbf{r}}_j + \sqrt{\mathrm{Var}(\hat{w}_{11})}\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\hat{\boldsymbol{\mu}} + \sqrt{\mathrm{Var}(\hat{w}_{11})}\tilde{\mathbf{r}}_j^T\boldsymbol{\Sigma}\hat{\boldsymbol{\mu}} + o(n^{-1/2})$. Because

$$n^{-1}\sum_{i,j\ unequal}\mathrm{E}\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\tilde{\mathbf{r}}_j = 0, \quad \mathrm{E}(n^{-1}\sum_{i,j\ unequal}\tilde{\mathbf{r}}_i^T\boldsymbol{\Sigma}\tilde{\mathbf{r}}_j)^2 \le n^{-2}\mathrm{tr}(\boldsymbol{\Sigma}^2) = o(n^{-2(0.5-\epsilon)}),$$

we have $n^{-1}\sum_{i,j\ unequal}\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_j = o_p(n^{-(0.5-\epsilon)})$. Because

$$n^{-1}\sum_{i,j\ unequal}\mathrm{E}\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}} = 0, \quad \mathrm{E}(n^{-1}\sum_{i,j\ unequal}\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}})^2 \le \hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}^2\hat{\boldsymbol{\mu}} = o(n^{-1/2}),$$

we have $n^{-1}\sum_{i,j\ unequal}\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}} = o_p(n^{-1/2})$. Because $n^{-1}\sum_{i,j\ unequal}\mathrm{E}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}})^2 = \hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}^2\hat{\boldsymbol{\mu}} = o(n^{-1/2})$, we have $n^{-1}\sum_{i,j\ unequal}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}})^2 = o_p(n^{-1/2})$. Because

$$n^{-1}\sum_{i,j\ unequal}\mathrm{E}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_j) = 0, \mathrm{E}(n^{-1}\sum_{i,j\ unequal}\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_j)^2 \le (\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}^2\hat{\boldsymbol{\mu}})^2 = o(n^{-1}),$$

we have $n^{-1}\sum_{i,j\ unequal}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_j) = o_p(n^{-1/2})$. Because

$$\begin{aligned}
&\mathrm{E}(n^{-1}\sum_{i,j\ unequal}\tilde{\mathbf{r}}_j^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}\hat{\boldsymbol{\mu}})^2 \\
= \;& n^{-2}\sum_{i,j,\ell\ unequal}\mathrm{E}\tilde{\mathbf{r}}_j^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_\ell\tilde{\mathbf{r}}_\ell^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_j + n^{-2}\sum_{i,j,\ unequal}\mathrm{E}\tilde{\mathbf{r}}_j^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_i\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_j \\
&+n^{-2}\sum_{i,j,\ unequal}\mathrm{E}\tilde{\mathbf{r}}_j^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_j\tilde{\mathbf{r}}_j^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_i \\
\le \;& n^{-2}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}^4\hat{\boldsymbol{\mu}} + \mathrm{E}\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_1\tilde{\mathbf{r}}_1^\top\mathbf{\Sigma}^2\tilde{\mathbf{r}}_1 + \mathrm{E}(\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}\tilde{\mathbf{r}}_2)^2\tilde{\mathbf{r}}_1^\top\mathbf{\Sigma}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_2 = o(n^{-1/2}),
\end{aligned}$$

we have $n^{-1}\sum_{i,j\ unequal}\tilde{\mathbf{r}}_j^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}\hat{\boldsymbol{\mu}} = o_p(n^{-1/2})$. Because

$$\begin{aligned}
&\mathrm{E}(n^{-1}\sum_{i,j\ unequal}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_j)^2 \\
= \;& n^{-2}\sum_{i,j,\ell\ unequal}\mathrm{E}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_j\tilde{\mathbf{r}}_j^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_\ell(\tilde{\mathbf{r}}_\ell^T\mathbf{\Sigma}\tilde{\mathbf{r}}_\ell - n^{-1}\mathrm{tr}\mathbf{\Sigma}) \\
&+n^{-2}\sum_{i,j\ unequal}\mathrm{E}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_j\tilde{\mathbf{r}}_j^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_i(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma}) \\
&+n^{-2}\sum_{i,j\ unequal}\mathrm{E}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_j\tilde{\mathbf{r}}_i^\top\mathbf{\Sigma}\tilde{\mathbf{r}}_j(\tilde{\mathbf{r}}_j^T\mathbf{\Sigma}\tilde{\mathbf{r}}_j - n^{-1}\mathrm{tr}\mathbf{\Sigma}) \\
\le \;& [\mathrm{E}(\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}\tilde{\mathbf{r}}_1 - n^{-1}\mathrm{tr}\mathbf{\Sigma})^2]^2 + \mathrm{E}(\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}^2\tilde{\mathbf{r}}_2)^2 + n^{-1}\mathrm{E}(\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}\tilde{\mathbf{r}}_1 - n^{-1}\mathrm{tr}\mathbf{\Sigma})^2\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}^2\tilde{\mathbf{r}}_1 \\
&+[\mathrm{E}(\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}\tilde{\mathbf{r}}_1 - n^{-1}\mathrm{tr}\mathbf{\Sigma})^2]^2 + \mathrm{E}(\tilde{\mathbf{r}}_1^T\mathbf{\Sigma}\tilde{\mathbf{r}}_2)^4 = o(n^{-2(0.5-\epsilon)}),
\end{aligned}$$

by (1.15) of Bai and Silverstein (2004) and (9.9.6) of Bai and Silverstein (2010), we have

$$n^{-1}\sum_{i,j\ unequal}(\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_i - n^{-1}\mathrm{tr}\mathbf{\Sigma})\tilde{\mathbf{r}}_i^T\mathbf{\Sigma}\tilde{\mathbf{r}}_j = o_p(n^{-(0.5-\epsilon)}).$$

Thus we have

$$n^{-1}\sum_{i,j\ unequal}\hat{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\mathbf{r}}_i\hat{\mathbf{r}}_i^T\mathbf{\Sigma}\hat{\mathbf{r}}_j=o(n^{-(0.5-\epsilon)}). \tag{S.2.5}$$

By (S.2.4) and (S.2.5), we have

$$n^{-1}\sum_{i,j\ unequal}\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_j=o_p(1). \tag{S.2.6}$$

By (S.2.1), (S.2.2), (S.2.3) and (S.2.6), we have

$$n\bar{\mathbf{r}}^T\mathbf{\Sigma}\sum_{i=1}^n\mathbf{r}_i\mathbf{r}_i^T\mathbf{\Sigma}\bar{\mathbf{r}}=(n^{-1}\mathrm{tr}\mathbf{\Sigma})^2+(n-1)n^{-2}\mathrm{tr}(\mathbf{\Sigma}^2)+o_p(1). \tag{S.2.7}$$

That is, the proof of Lemma S.2.2 is complete.

Lemma S.2.3. *Under Assumptions [A]-[B] with* $|w_{\ell i}|\le\sqrt{n}\eta_n$, $\mathrm{E}w_{\ell i}=0$, $\mathrm{E}(w_{\ell i}^2)=1$ *and* $\mathrm{E}(w_{\ell i}^4)<\infty$ *with* $\eta_n\downarrow 0$ *and* $n^{1/4}\eta_n\to\infty$, *we have*

$$\sum_{j=1}^n\mathrm{E}([(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n]^2\delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n|\ge\epsilon\}})=O(\eta_n^4),$$

$$\sum_{j=1}^n\mathrm{E}([(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n^2]^2\delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n^2|\ge\epsilon\}})=O(\eta_n^4),$$

$$\sum_{j=1}^n\mathrm{E}([(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}]^2\delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}|\ge\epsilon\}})=O(\eta_n^4),$$

and

$$\sum_{j=1}^n\mathrm{E}([(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}^2]^2\delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}^2|\ge\epsilon\}})=O(\eta_n^4).$$

Proof. We have $\mathrm{tr}\mathbf{B}_n=\sum_{i=1}^n\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i$ and $\mathrm{E}(\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i-n^{-1}\mathrm{tr}\mathbf{\Sigma})^4\le Cn^{-1}\eta_n^4\|\mathbf{\Sigma}\|^4=O(\eta_n^4n^{-1})$ by (9.9.6) of Bai and Silverstein (2010) where C is a constant independent of n and p. Then we have

$$\sum_{j=1}^n\mathrm{E}([(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n]^2\delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n|\ge\epsilon\}})\le n\mathrm{E}(\mathbf{r}_i^T\mathbf{\Sigma}\mathbf{r}_i-n^{-1}\mathrm{tr}\mathbf{\Sigma})^4/\epsilon^2=O(\eta_n^4).$$

Similarly, we have

$$\sum_{j=1}^{n} \mathrm{E}([(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}^2]^2 \delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}^2|\geq\epsilon\}}) = O(\eta_n^4).$$

$\mathrm{E}_j(\mathrm{tr}\mathbf{B}_n^2) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{B}_n^2)$ can be expressed by

$$\begin{aligned}
\mathrm{E}_j\mathrm{tr}(\mathbf{B}_n^2) - \mathrm{E}_{j-1}\mathrm{tr}(\mathbf{B}_n^2) &= 2(n-j)n^{-1}[\mathbf{r}_j^\top\boldsymbol{\Sigma}^2\mathbf{r}_j - n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^2)] \\
&\quad +[\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j - \mathrm{E}(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j)] \\
&\quad +2\sum_{k\leq j-1}[\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}\mathbf{r}_j - n^{-1}(\mathbf{r}_j^\top\boldsymbol{\Sigma}^2\mathbf{r}_j)].
\end{aligned}$$

We have

$$\sum_{j=1}^{n}(n-j)^4 n^{-4}\mathrm{E}\left[(\mathbf{r}_1^\top\boldsymbol{\Sigma}^2\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2)^4\right] \leq C\sum_{j=1}^{n}(n-j)^4 n^{-5}\eta_n^4, \tag{S.2.8}$$

which is from Lemma 9.1 of Bai and Silverstein (2004) and C is a constant not dependent on p or n. Moreover, we have

$$\begin{aligned}
&\sum_{j=1}^{n}\mathrm{E}[(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j - \mathrm{E}\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j)^4] \\
= \;& n\mathrm{E}[(\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1 - \mathrm{E}\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1)^4] \\
\leq \;& Cn\mathrm{E}[(\mathbf{r}_1^T\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^8] + nO(n^{-4}) + Cn(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^4\mathrm{E}[(\mathbf{r}_1^T\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^4] \\
\leq \;& O(\eta_n^{12}) + O(n^{-3}) + O(n^{-4}) \qquad\qquad \text{(S.2.9)}
\end{aligned}$$

where $(\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 = n^{-2}[2\mathrm{tr}(\boldsymbol{\Sigma}^2) + \beta_w\sum_{j=1}^{p}(\mathbf{e}_j^T\boldsymbol{\Sigma}\mathbf{e}_j)^2]$, the last inequality is from (9.9.6) of Bai and Silverstein (2010) and

$$\begin{aligned}
&(\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1)^2 - \mathrm{E}[(\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1)^2] \\
= \;& (\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 - \mathrm{E}[(\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2] + 2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}) \\
= \;& (\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 + 2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_1^\top\boldsymbol{\Sigma}\mathbf{r}_1 - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}) + O(n^{-1}).
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
&\sum_{j=1}^{n}\mathrm{E}\{[\sum_{k\le j-1}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}\mathbf{r}_j]^4\}\\
=\;&\sum_{j=1}^{n}\mathrm{E}\{[\sum_{k\le j-1}(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}\mathbf{r}_j-n^{-1}\mathbf{r}_k^\top\boldsymbol{\Sigma}^2\mathbf{r}_k)]^4\}\\
\le\;& C\eta_n^4 n^{-1}\sum_{j=1}^{n}\mathrm{E}(\|\sum_{k\le j-1}\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}\|^4)\\
\le\;& C\eta_n^4\mathrm{E}(\|\sum_{k=1}^{n}\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}\|^4)\le C\eta_n^4\|\boldsymbol{\Sigma}\|^8\mathrm{E}(\|\sum_{k=1}^{n}\mathbf{r}_k\mathbf{r}_k^\top\|^4)\\
\le\;& 2C\eta_n^4\|\boldsymbol{\Sigma}\|^2(1+\sqrt{y_n})^8=O(\eta_n^4) \qquad \text{(S.2.10)}
\end{aligned}$$

where the second inequality is from (9.9.6) of Bai and Silverstein (2010), $\|\sum_{k\le j-1}\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}\|$ is the spectral norm of the random matrix $\sum_{k\le j-1}\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}$, that is, the maximum eigenvalue of $\sum_{k\le j-1}\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}$ and the last inequality is from (4.2) of Yin, Bai and Krishnaiah (1988). From (S.2.8)-(S.2.9)-(S.2.10), we have

$$\begin{aligned}
&\sum_{j=1}^{n}\mathrm{E}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n^2]^2\delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_n^2|\ge\epsilon\}}\}\\
\le\;& C\sum_{j=1}^{n}\mathrm{E}[2(n-j)n^{-1}(\mathbf{r}_j^\top\boldsymbol{\Sigma}^2\mathbf{r}_j-n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2)]^4+C\sum_{j=1}^{n}\mathrm{E}[\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j-\mathrm{E}(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j)]^4\\
&+C\sum_{j=1}^{n}\mathrm{E}\{[\sum_{k\le j-1}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\boldsymbol{\Sigma}\mathbf{r}_j]^4\}=O(\eta_n^4)+O(n^{-3})+O(n^{-4})=O(\eta_n^4).
\end{aligned}$$

Similarly, we have

$$\sum_{j=1}^{n}\mathrm{E}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}^2]^2\delta_{\{|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{B}_{kk}^2|\ge\epsilon\}}\}=O(\eta_n^4)+O(n^{-3})+O(n^{-4})=O(\eta_n^4).$$

The proof of Lemma S.2.3 is complete.

Lemma S.2.4. *Under Assumptions [A]-[B] with* $|w_{\ell i}|\le\sqrt{n}\eta_n$, $\mathrm{E}w_{\ell i}=0$, $\mathrm{E}(w_{\ell i}^2)=1$

and $\mathrm{E}(w_{\ell i}^4) < \infty$ *with* $\eta_n \downarrow 0$ *and* $n^{1/4}\eta_n \to \infty$, *we have*

$$
\begin{aligned}
(S2.6) &= 2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{11}^3) + \beta_w n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{e}_{\ell 1}\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell 1}] \\
&\quad +2[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{11}^4) + \beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{e}_{\ell 1})^2] + O_p(\eta_n^2), \\
(S2.7) &= 4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{11}^2) + \beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell 1})^2] \\
&\quad +4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{11}^3) + \beta_w n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell 1}\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{e}_{\ell 1}] + O_p(\eta_n^2), \\
(S2.8) &= 2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{11}^3) + \beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell 1})(\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{e}_{\ell 1})] \\
&\quad +2[2n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_{11}^4) + \beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{e}_{\ell 1})^2] + 4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11}^2)^2 + O_p(\eta_n^2),
\end{aligned}
$$

where $O_p(\eta_n^2)$ *is uniform for* $k = 1, ..., K$.

Proof. We have

$$
\begin{aligned}
(S2.6) &= (S.2.11) + (S.2.12) + (S.2.13), \\
(S2.7) &= (S.2.14) + (S.2.15) + (S.2.16), \\
(S2.8) &= (S.2.17) + (S.2.18) + (S.2.19),
\end{aligned}
$$

where

$$4\sum_{j=1}^{n}\frac{(n-j)^2}{n^2}\mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_j^\top\boldsymbol{\Sigma}^2\mathbf{r}_j][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{r}_{j1}]\} \tag{S.2.11}$$

$$2\sum_{j=1}^{n}\frac{(n-j)}{n}\mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_j^\top\boldsymbol{\Sigma}^2\mathbf{r}_j][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}]\} \tag{S.2.12}$$

$$4\sum_{j=1}^{n}\frac{(n-j)}{n}\sum_{k\le j-1}\mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_j^\top\boldsymbol{\Sigma}^2\mathbf{r}_j][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}]\} \tag{S.2.13}$$

$$2\sum_{j=1}^{n}\frac{(n-j)}{n}\mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{r}_{j1}]\} \tag{S.2.14}$$

$$\sum_{j=1}^{n}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}]\} \tag{S.2.15}$$

$$2\sum_{j=1}^{n}\sum_{k\leq j-1}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_j][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}]\} \tag{S.2.16}$$

$$4\sum_{j=1}^{n}\frac{(n-j)}{n}\sum_{k\leq j-1}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\mathbf{\Sigma}\mathbf{r}_j][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}]\} \tag{S.2.17}$$

$$2\sum_{j=1}^{n}\sum_{k\leq j-1}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\mathbf{\Sigma}\mathbf{r}_j][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}]\} \tag{S.2.18}$$

$$4\sum_{j=1}^{n}\sum_{k\leq j-1}\sum_{\ell\leq j-1}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\mathbf{\Sigma}\mathbf{r}_k\mathbf{r}_k^\top\mathbf{\Sigma}\mathbf{r}_j][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{\ell 1}\mathbf{r}_{\ell 1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}]\} \tag{S.2.19}$$

Detailed proof of (S.2.11):

$$\begin{aligned}(S.2.11) &= 4\sum_{j=1}^{n}\frac{(n-j)^2}{n^2}\mathrm{E}_{j-1}[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\mathbf{\Sigma}^2\mathbf{r}_j(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}]\\ &= \frac{4n(1+O(n^{-1}))}{3}\mathrm{E}\{[\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}-n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^2)]^2\}\\ &= \frac{4}{3n}[2\mathrm{tr}(\mathbf{\Sigma}_{11}^4)+\beta_w\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell 1})^2]+O(n^{-1})\end{aligned}$$

where the last equality is from (1.15) of Bai and Silverstein (2004).

Detailed proof of (S.2.12):

$$\begin{aligned}&(S.2.12)\\ =\ & 2\sum_{j=1}^{n}\frac{(n-j)}{n}\mathrm{E}_{j-1}[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\mathbf{\Sigma}^2\mathbf{r}_j(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}]\\ =\ & n(1+O(n^{-1}))\mathrm{E}[\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}-\mathrm{E}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}][(\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1})^2-\mathrm{E}(\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1})^2]\\ =\ & n(1+O(n^{-1}))\mathrm{E}[(\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}-\mathrm{E}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1})(\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})^2]\\ & +2n(1+O(n^{-1}))(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})\mathrm{E}[(\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1}-\mathrm{E}\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{j1})(\mathbf{r}_{j1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})]\ \text{(S.2.20)}\end{aligned}$$

where the last equality is from the following equality

$$\begin{aligned} & (\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1})^2 - \mathrm{E}[(\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1})^2] \\ = & (\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1} - n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})^2 - \mathrm{E}[(\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1} - n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})^2] \\ & +2(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})(\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1} - n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}). \end{aligned}$$

By (9.9.6) of Bai and Silverstein (2010), we have

$$n\mathrm{E}[(\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}^2\mathbf{r}_{j1} - \mathrm{E}\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}^2\mathbf{r}_{j1})(\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1} - n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})^2] \le C_0\|\mathbf{\Sigma}_{11}\|^2\|\mathbf{\Sigma}_{11}^2\|\eta_n^2 = O(\eta_n^2) \tag{S.2.21}$$

where C_0 is a constant. By (1.15) of Bai and Silverstein (2004), we have

$$\begin{aligned} & \mathrm{E}[(\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}^2\mathbf{r}_{j1} - \mathrm{E}\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}^2\mathbf{r}_{j1})(\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1} - n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})] \\ = & n^{-2}[2\mathrm{tr}(\mathbf{\Sigma}_{11}^3) + \beta_w \sum_{\ell=1}^{p} \mathbf{e}_{\ell 1}^\top \mathbf{\Sigma}_{11}^2 \mathbf{e}_{\ell 1}\mathbf{e}_{\ell 1}^\top \mathbf{\Sigma}_{11}\mathbf{e}_{\ell 1}] = O(n^{-1}). \end{aligned} \tag{S.2.22}$$

By (S.2.20)-(S.2.21)-(S.2.22), we have

$$(S.2.12) = 2(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3) + \beta_w n^{-1}\sum_{\ell=1}^{p_1} \mathbf{e}_{\ell 1}^\top \mathbf{\Sigma}_{11}^2 \mathbf{e}_{\ell 1}\mathbf{e}_{\ell 1}^\top \mathbf{\Sigma}_{11}\mathbf{e}_{\ell 1}] + O(\eta_n^2).$$

Moreover, the detailed proof of (S.2.14) is similar to the proof of (S.2.12).

Detailed proof of (S.2.13):

$$\begin{aligned} & (S.2.13) \\ = & 4\sum_{j=1}^{n}\sum_{k\le j-1}\frac{(n-j)}{n}\mathrm{E}_{j-1}[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_j^\top \mathbf{\Sigma}^2\mathbf{r}_j(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1}] \\ = & 4\sum_{j=1}^{n}\sum_{k\le j-1}\frac{(n-j)}{n}\mathrm{E}_{j-1}\{[\mathbf{r}_j^\top \mathbf{\Sigma}^2\mathbf{r}_j - n^{-1}\mathrm{tr}(\mathbf{\Sigma}^2)](\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1} - n^{-1}\mathbf{r}_{k1}^\top \mathbf{\Sigma}_{11}^2\mathbf{r}_{k1})\} \\ = & 4\sum_{j=1}^{n}\sum_{k\le j-1}\frac{(n-j)}{n}\mathrm{E}_{j-1}\{[\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}^2\mathbf{r}_{j1} - n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^2)](\mathbf{r}_{j1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{j1} - n^{-1}\mathbf{r}_{k1}^\top \mathbf{\Sigma}_{11}^2\mathbf{r}_{k1})\} \\ = & 4\sum_{j=1}^{n}\sum_{k\le j-1}\frac{(n-j)}{n^3}(2\mathbf{r}_{k1}^\top \mathbf{\Sigma}_{11}^4\mathbf{r}_{k1} + \beta_w\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell}^\top \mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell}\mathbf{e}_{\ell}^\top \mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top \mathbf{\Sigma}_{11}\mathbf{e}_{\ell}) \end{aligned} \tag{S.2.23}$$

where the last equality is from (1.15) of Bai and Silverstein (2004). It is clear that $\sum_{j=1}^{n}\sum_{k\le j-1}(n-j)n^{-3}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}^{4}\mathbf{r}_{k1}$ is the weighted sum of independent random variables $\{\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}^{4}\mathbf{r}_{k1}, k=1,...,n\}$ with $\mathrm{E}\Big[\sum_{j=1}^{n}\sum_{k\le j-1}\frac{n-j}{n^3}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}^{4}\mathbf{r}_{k1}\Big]=(3n)^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^{4}+O(n^{-1})$ and $\mathrm{var}\Big[\sum_{j=1}^{n}\sum_{k\le j-1}\frac{n-j}{n^3}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}^{4}\mathbf{r}_{k1}\Big]=O(n^{-1})$. That is

$$\sum_{j=1}^{n}\sum_{k\le j-1}(n-j)n^{-3}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}^{4}\mathbf{r}_{k1}=(3n)^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^{4}+O_p(n^{-1/2}). \tag{S.2.24}$$

It is clear that $\sum_{j=1}^{n}(n-j)n^{-3}\sum_{k\le j-1}\sum_{\ell=1}^{p}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}\mathbf{e}_{\ell}$ is the weighted sum of the independent random variables $\{\sum_{\ell=1}^{p}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}\mathbf{e}_{\ell}, k=1,...,n\}$ with

$$\mathrm{E}\Big[\sum_{j=1}^{n}(n-j)n^{-3}\sum_{k\le j-1}\sum_{\ell=1}^{p}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}\mathbf{e}_{\ell}\Big]=(6n)^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell})^{2}+O(n^{-1})$$
$$\mathrm{var}\Big[\sum_{j=1}^{n}(n-j)n^{-3}\sum_{k\le j-1}\sum_{\ell=1}^{p}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}\mathbf{e}_{\ell}\Big]=O(n^{-1}).$$

That is,

$$\sum_{j=1}^{n}(n-j)n^{-3}\sum_{k\le j-1}\sum_{\ell=1}^{p}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell}\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^{\top}\mathbf{\Sigma}_{11}\mathbf{e}_{\ell}=(6n)^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell1})^{2}+O_p(n^{-1/2}). \tag{S.2.25}$$

By (S.2.23)-(S.2.24)-(S.2.25), we have

$$(S.2.13)=\frac{2}{3n}[2\mathrm{tr}\mathbf{\Sigma}_{11}^{4}+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_{\ell}^{\top}\mathbf{\Sigma}_{11}^{2}\mathbf{e}_{\ell})^{2}]+O_p(n^{-1/2}).$$

Moreover, the detailed proof of (S.2.17) is similar to the proof of (S.2.13).

Detailed proof of (S.2.15):

$$\begin{aligned}(S.2.15) \quad &= \sum_{j=1}^{n} \mathrm{E}_{j-1}[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j \mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j (\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} \mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1}] \\ &= n\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 (\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})^2] \\ &\quad - n\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2]\mathrm{E}[(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})^2] \\ &\quad + 2n(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 (\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})] \\ &\quad + 2n(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})^2] \\ &\quad + 4n(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})]. \quad (\mathrm{S.2.26})\end{aligned}$$

By (1.15) of Bai and Silverstein (2004), we have

$$\begin{aligned}& n\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2]\mathrm{E}[(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})^2] \quad (\mathrm{S.2.27}) \\ = \;& n^{-1}[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top \boldsymbol{\Sigma} \mathbf{e}_\ell)^2][2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top \boldsymbol{\Sigma} \mathbf{e}_{\ell 1})^2] = O(n^{-1}),\end{aligned}$$

and

$$\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})] = n^{-2}[2\mathrm{tr}\boldsymbol{\Sigma}_{11}^2 + \beta_w \sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top \boldsymbol{\Sigma}_{11} \mathbf{e}_{\ell 1})^2]. \quad (\mathrm{S.2.28})$$

By (9.9.6) of Bai and Silverstein (2010), we have

$$\begin{aligned}& n\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 (\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})^2] \le \eta_n^4 \cdot C_0 \|\boldsymbol{\Sigma}\|^2 \|\boldsymbol{\Sigma}_{11}\|^2 = O(\eta_n^4), \\ & n\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2 (\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})] \le \eta_n^2 \cdot C_0 \|\boldsymbol{\Sigma}\|^2 \|\boldsymbol{\Sigma}_{11}\| = O(\eta_n^2), \\ & n\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})^2] \le \eta_n^2 \cdot C_0 \|\boldsymbol{\Sigma}\|^2 \|\boldsymbol{\Sigma}_{11}\| = O(\eta_n^2). \quad (\mathrm{S.2.29})\end{aligned}$$

By (1.15) of Bai and Silverstein (2004), we have

$$n\mathrm{E}[(\mathbf{r}_j^\top \boldsymbol{\Sigma} \mathbf{r}_j - n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_{j1}^\top \boldsymbol{\Sigma}_{11} \mathbf{r}_{j1} - n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})] = 2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top \boldsymbol{\Sigma}_{11} \mathbf{e}_{\ell 1})^2. \quad (\mathrm{S.2.30})$$

Then by (S.2.26)-(S.2.27)-(S.2.28)-(S.2.29)-(S.2.30), we have

$$(S.2.15) = 4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})[2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11}^2 + \beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell 1}^\top \boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell 1})^2] + O(\eta_n^2).$$

Detailed proof of (S.2.16):

$$\begin{aligned}
&(S.2.16)\\
= \ & 2\sum_{j=1}^{n}\sum_{k\le j-1}\mathrm{E}_{j-1}[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}]\\
= \ & 2\sum_{j=1}^{n}\sum_{k\le j-1}\mathrm{E}_{j-1}[(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j-n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2(\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{r}_{k1})]\\
&+4(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})\sum_{j=1}^{n}\sum_{k\le j-1}\mathrm{E}_{j-1}[(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j-n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{r}_{k1})]
\end{aligned} \quad (S.2.31)$$

By (9.9.6) of Bai and Silverstein (2010), we have

$$\begin{aligned}
&\left|\sum_{j=1}^{n}\sum_{k\le j-1}\mathrm{E}_{j-1}[(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j-n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2(\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{r}_{k1})]\right|\\
\le \ & \sum_{j=1}^{n}\left|\mathrm{E}_{j-1}[(\mathbf{r}_j^\top\boldsymbol{\Sigma}\mathbf{r}_j-n^{-1}\mathrm{tr}\boldsymbol{\Sigma})^2(\mathbf{r}_{j1}^\top\boldsymbol{\Sigma}_{11}\sum_{k\le j-1}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\sum_{k\le j-1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}^2\mathbf{r}_{k1})]\right|\\
\le \ & \sum_{j=1}^{n} n^{-1}\eta_n^2\cdot C_0\|\boldsymbol{\Sigma}\|^2\|\boldsymbol{\Sigma}_{11}\sum_{k\le j-1}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\|\\
\le \ & \eta_n^2\cdot C_0\|\boldsymbol{\Sigma}\|^2\|\boldsymbol{\Sigma}_{11}\sum_{k=1}^{n}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\boldsymbol{\Sigma}_{11}\|\\
\le \ & \eta_n^2\cdot C_0\|\boldsymbol{\Sigma}\|^2\|\boldsymbol{\Sigma}_{11}\|^2\|\sum_{k=1}^{n}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top\|\\
= \ & \eta_n^2\cdot C_0\|\boldsymbol{\Sigma}\|^2\|\boldsymbol{\Sigma}_{11}\|^2\lambda_{\max}(\sum_{k=1}^{n}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top)\\
= \ & \eta_n^2\cdot C_0\|\boldsymbol{\Sigma}\|^2\|\boldsymbol{\Sigma}_{11}\|^2[(1+\sqrt{y_n})^2+o_{a.s.}(1)] = O_{a.s.}(\eta_n^2)
\end{aligned} \quad (S.2.32)$$

where $\lambda_{\max}(\sum_{k=1}^{n}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top) = (1+\sqrt{y_n})^2 + o_{a.s.}(1)$ is the maximum eigenvalue of the random matrix $\sum_{k=1}^{n}\mathbf{r}_{k1}\mathbf{r}_{k1}^\top$ by Yin, Bai and Krishnaiah (1988). Similar to the proofs of (S.2.24) and (S.2.25), we have

$$2n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}^{3}\mathbf{r}_{k1}=n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11}^{3}+O_p(n^{-1/2})$$

and

$$n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}=0.5n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}^{2}\mathbf{e}_{\ell1}+O_p(n^{-1/2}).$$

Thus we have

$$\begin{aligned}
&\sum_{j=1}^{n}\sum_{k\le j-1}\mathrm{E}_{j-1}[(\mathbf{r}_{j}^{\top}\boldsymbol{\Sigma}\mathbf{r}_{j}-n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(\mathbf{r}_{j1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}^{2}\mathbf{r}_{k1})]\\
=&\sum_{j=1}^{n}\sum_{k\le j-1}\mathrm{E}_{j-1}[(\mathbf{r}_{j1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11})(\mathbf{r}_{j1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}-n^{-1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}^{2}\mathbf{r}_{k1})]\\
=&\ n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}\mathrm{E}\left[2\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}^{3}\mathbf{r}_{k1}+\beta_w\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{\ell1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1}\right]\\
=&\ 0.5n^{-1}(2\mathrm{tr}\boldsymbol{\Sigma}_{11}^{3}+\beta_w\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}^{2}\mathbf{e}_{\ell1})+O_p(n^{-1}).
\end{aligned}\tag{S.2.33}$$

By (S.2.31)-(S.2.32)-(S.2.33), we have

$$(S.2.16)=2(n^{-1}\mathrm{tr}\boldsymbol{\Sigma})(2n^{-1}\mathrm{tr}\boldsymbol{\Sigma}_{11}^{3}+\beta_w n^{-1}\sum_{\ell=1}^{p}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^{\top}\boldsymbol{\Sigma}_{11}^{2}\mathbf{e}_{\ell1})+O_p(\eta_n^2).$$

The detailed proofs of (S.2.18) is similar to the proofs of (S.2.16).

Detailed proofs of (S.2.19):

$$\begin{aligned}
(S.2.19)&=4\sum_{j=1}^{n}\mathrm{E}_{j-1}\sum_{k\le j-1}\sum_{i\le j-1}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j}^{\top}\boldsymbol{\Sigma}\mathbf{r}_{k}\mathbf{r}_{k}^{\top}\boldsymbol{\Sigma}\mathbf{r}_{j}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{r}_{j1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{i1}\mathbf{r}_{i1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{j1}\\
&=4\sum_{j=1}^{n}\mathrm{E}_{j-1}\sum_{k\le j-1}\sum_{i\le j-1}[(\mathbf{r}_{j}^{\top}\boldsymbol{\Sigma}\mathbf{r}_{k})^2-n^{-1}\mathbf{r}_{k}^{\top}\boldsymbol{\Sigma}^{2}\mathbf{r}_{k}][(\mathbf{r}_{j1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{i1})^2-n^{-1}\mathbf{r}_{i1}^{\top}\boldsymbol{\Sigma}^{2}\mathbf{r}_{i1}]\\
&=4\sum_{j=1}^{n}\mathrm{E}_{j-1}\sum_{k\le j-1}\sum_{i\le j-1}[(\mathbf{r}_{j1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{k1})^2-n^{-1}\mathbf{r}_{k1}^{\top}\boldsymbol{\Sigma}_{11}^{2}\mathbf{r}_{k1}][(\mathbf{r}_{j1}^{\top}\boldsymbol{\Sigma}_{11}\mathbf{r}_{i1})^2-n^{-1}\mathbf{r}_{i1}^{\top}\boldsymbol{\Sigma}^{2}\mathbf{r}_{i1}]
\end{aligned}$$

$$= 4n^{-2}\sum_{j=1}^{n}\mathrm{E}_{j-1}\sum_{k\le j-1}\sum_{i\le j-1}\left[2(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{i1})^2+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^2(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{i1})^2\right]$$

$$= 4n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}\left[2(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1})^2+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^4\right]$$

$$+4n^{-2}\sum_{j=1}^{n}\sum_{1\le k\ne i\le j-1}\left[2(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{i1})^2+\beta_w\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^2(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{r}_{i1})^2\right] \quad \text{(S.2.34)}$$

where the fourth equality is from (1.15) of Bai and Silverstein (2004). Because

$$\begin{cases} n^{-2}\sum_{j=1}^n\sum_{k\le j-1}\mathrm{E}(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2=0.5n^{-2}[2\mathrm{tr}\mathbf{\Sigma}_{11}^4+\beta_w\sum_{\ell=1}^p(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})^2]=O(n^{-1}),\\ n^{-2}\sum_{j=1}^n\sum_{k\le j-1}\mathrm{E}|\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)|\le 0.5\{\mathrm{E}[(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2]\}^{1/2}=O(n^{-1/2}),\end{cases}$$

leads to

$$\begin{cases} n^{-2}\sum_{j=1}^n\sum_{k\le j-1}\mathrm{E}(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2=O_p(n^{-1}),\\ n^{-2}\sum_{j=1}^n\sum_{k\le j-1}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)=O_p(n^{-1/2}),\end{cases}$$

then we have

$$\begin{aligned} n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1})^2 &= n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2+0.5(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2\\ &\quad +2n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{k1}-n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)\\ &= 0.5(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2+O_p(n^{-1/2}). \end{aligned} \quad \text{(S.2.35)}$$

Because $n^{-2}\sum_{j=1}^n\sum_{k\le j-1}\sum_{\ell=1}^p(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^4=n^{-2}\sum_{j=1}^n\sum_{k\le j-1}\sum_{\ell=1}^p(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_\ell\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^2$,

similar to the proof of $(S.2.35)$, we have

$$n^{-2}\sum_{j=1}^{n}\sum_{k\le j-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^4=O_p(n^{-1/2}). \quad \text{(S.2.36)}$$

Because

$$\begin{aligned} &4n^{-2}\sum_{j=1}^{n}\sum_{1\le k\ne i\le j-1}\mathrm{E}\left[2(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{i1})^2+\beta_w\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^2(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{i1})^2\right]\\ =\;&\frac{4}{3}[2n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^4+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})^2]+O(n^{-1}), \end{aligned}$$

and

$$\begin{cases} n^{-4}\mathrm{var}[\sum_{j=1}^{n}\sum_{1\le k\neq i\le j-1}(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{i1})^2]=O(n^{-1}), \\ n^{-4}\mathrm{var}[\sum_{j=1}^{n}\sum_{1\le k\neq i\le j-1}\sum_{\ell=1}^{p}(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^2(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{i1})^2]=O(n^{-1}), \end{cases}$$

then we have

$$\begin{aligned} & 4n^{-2}\sum_{j=1}^{n}\sum_{1\le k\neq i\le j-1}\left[2(\mathbf{r}_{k1}^\top\mathbf{\Sigma}_{11}^2\mathbf{r}_{i1})^2+\beta_w\sum_{\ell=1}^{p_1}(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{k1})^2(\mathbf{e}_\ell^\top\mathbf{\Sigma}_{11}\mathbf{r}_{i1})^2\right] \\ = & \frac{4}{3}[2n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^4+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})^2]+4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2+O_p(n^{-1/2}). \end{aligned} \quad \text{(S.2.37)}$$

By (S.2.34)-(S.2.35)-(S.2.36)-(S.2.37), we have

$$(S.2.19)=\frac{4}{3}[2n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^4+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})^2]+4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2+O_p(n^{-1/2}).$$

Thus, we have

$$\begin{aligned} (S2.6) &= 2(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell1}] \\ &\quad +2[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^4)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})^2]+O_p(\eta_n^2), \\ (S2.7) &= 4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})(n^{-1}\mathrm{tr}\mathbf{\Sigma})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^2)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell1})^2] \\ &\quad +4(n^{-1}\mathrm{tr}\mathbf{\Sigma})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell1}\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1}]+O_p(\eta_n^2), \\ (S2.8) &= 2(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11})[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^3)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}\mathbf{e}_{\ell1})(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})] \\ &\quad +2[2n^{-1}\mathrm{tr}(\mathbf{\Sigma}_{11}^4)+\beta_w n^{-1}\sum_{\ell=1}^{p_1}(\mathbf{e}_{\ell1}^\top\mathbf{\Sigma}_{11}^2\mathbf{e}_{\ell1})^2]+4(n^{-1}\mathrm{tr}\mathbf{\Sigma}_{11}^2)^2+O_p(\eta_n^2). \end{aligned}$$

The proof of Lemma S.2.4 is complete.

Bibliography

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

Bai, Z. D. and Silverstein, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices (2nd Edition)*. Springer, New York.

Xiao, H. and Wu, W. B. (2013). Asymptotic theory for maximum deviations of sample covariance matrix estimates. *Stochastic Processes and their Applications*, **123**, 2899–2920.

Yin, Y. Q., Bai, Z. D. and Krishnaiah, P. R. (1988). On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probab. Theory Relat. Fields*, **78**, 509–521.