Limiting Behavior of Eigenvalues in High-Dimensional MANOVA via RMT

Zhidong Bai[1], Kwok Pui Choi[2] and Yasunori Fujikoshi[3]

[1] *School of Mathematics and Statistics and KLAS, Northeast Normal University, Changchun, Jilin 130024, China*

[2] *Department of Statistics and Applied Probability, National University of Singapore, 6 Science Drive 2, 117546, Singapore*

[3] *Department of Mathematics, Graduate School of Science, Hiroshima University, Higashi Hiroshima, Hiroshima 739-8626, Japan*

Abstract

In this paper we derive the asymptotic joint distributions of the eigenvalues under the null case in MANOVA model and multiple discriminant analysis when both the dimension and the sample size are large. Our results are obtained by random matrix theory (RMT) without assuming normality in the populations. It is worth pointing out that the null distributions of the eigenvalues and related test statistics are asymptotically robust against departure from normality in high-dimensional situations. Some new formulas in RMT are also presented.

AMS 2000 Subject Classification: primary 62H10; secondary 62E20

Key Words and Phrases: Asymptotic distribution, eigenvalues, discriminant analysis, high-dimensional case, MANOVA, nonnormality, RMT, test statistics.

Abbreviated title: Limiting Behavior in MANOVA with High-Dimension

1 Introduction

It is both basic and important to study the distributions of the eigenvalues in a one-way multivariate analysis of variance (MANOVA) model. Suppose there are $q+1$ groups and $\{\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}\}$ represents a random sample of p-vectors from the i-th group, which has mean vector $\boldsymbol{\mu}_i$ and common covariance matrix $\boldsymbol{\Sigma}$. Various inferential procedures are based on the matrices

$$\mathbf{S}_b = \frac{1}{n}\sum_{i=1}^{q+1} n_i(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})', \quad \mathbf{S}_e = \frac{1}{n}\sum_{i=1}^{q+1}\sum_{j=1}^{n_i}(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)',$$

where

$$\bar{\mathbf{x}}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{x}_{ij}, \;\; \bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{q+1} n_i\bar{\mathbf{x}}_i, \;\; n = \sum_{i=1}^{q+1} n_i.$$

The matrices $n\mathbf{S}_b$ and $n\mathbf{S}_e$ are called the matrices of sums of squares and products due to between-groups and within-groups, respectively. Let $\mathbf{S}_t = \mathbf{S}_b + \mathbf{S}_e$, then $n\mathbf{S}_t$ is called the matrix of sums of squares and products due to total variation. These matrices are also used in canonical discriminant analysis, which is a statistical procedure designed to discriminate between several different groups in terms of a few discriminant functions.

The large-sample asymptotic distributions of the eigenvalues of $\mathbf{S}_b\mathbf{S}_e^{-1}$ or $\mathbf{S}_b\mathbf{S}_t^{-1}$ were derived under normality by Hsu (1941) and Anderson (1951). A gap in the proof of the main Theorem in Hsu (1941) was corrected by Bai (1985). Amemiya (1990) characterised the limiting distributions of the roots of a general determinantal equation. Hsu's results were extended to asymptotic expansions by a number of authors, see, Sugiura (1976), Fujikoshi (1977), Muirhead (1978, 1982), Glynn and Muirhead (1978), etc. For extension to elliptical case with different covariance matrices, see Seo, Kanda and Fujikoshi (1994).

As is well known, the accuracy of large sample approximations deteriorates as the dimension increases (see, for instance, Bai and Saranadasa (1996)). As an alternative approach to overcome this shortcoming, it has been considered to derive asymptotic distributions of the eigenvalues in a high-dimensional situation where the dimension p, the sample size n and the number of groups $q+1$ are large. The high-dimensional asymptotic results when $p/n \to c \in$

$(0,1)$ were obtained (Fujikoshi et al. (2008)) by assuming that the population eigenvalues are simple. Johnstone (2008) derived the limiting distribution of the largest eigenvalue when $q/n \to \tilde{c} \in (0,1)$ in addition to $p/n \to c \in (0,1)$ under the assumption that all the population eigenvalues are zero, i.e., under the null case. All these results were obtained under normality assumption of the populations.

In this paper we derive asymptotic joint distributions of the eigenvalues under the null case in MANOVA model and multiple discriminant analysis when both the dimension and the sample size are large. Our results are obtained without assuming normality via random matrix theory (RMT). There are several asymptotic results based on RMT, see, for instances, Bai, Jiang, Yao and Zheng (2009), Bai, Liu and Wong (2011), Zheng (2012), etc. In this paper, we show that by applying RMT, it is possible to derive the asymptotic distributions of the eigenvalues without assuming normality. In the course of derivation, we obtain some new limit theorems which may be of independent interests. It is worth pointing out that our asymptotic results coincide with those under normality assumption. Therefore some MANOVA tests, including likelihood ratio criterion, Lawley-Hotelling criterion and Bartlett-Nanda-Pillai criterion, are robust against departure from normality when the dimension and the sample size are large.

The organization of the paper is as follows. In Section 2 we state our main results (Theorem 2.4 and Corollary 2.1) on the limiting joint distributions of the nonzero eigenvalues of $\mathbf{S}_b\mathbf{S}_t^{-1}$ and $\mathbf{S}_b\mathbf{S}_e^{-1}$ followed by application to null robustness of some multivariate tests. In the same section, we also highlight some limit theorems which are of independent interests. The proof of Theorem 2.4 is given in Section 3. The proofs of Theorems 2.1 and 2.2 are given in Sections 4 and 5 respectively.

2 Statements of Main Results

Throughout this article, $\{y_{jk} : 1 \le j \le p, 1 \le k \le n\}$ denotes a double array of *i.i.d.* random variables with mean 0, variance 1 and finite 4th moment. Let $\mathbf{y}_k = (y_{1k}, \dots, y_{pk})'$ and $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_n]$. Write $\bar{\mathbf{y}} = \frac{1}{n}\sum_{k=1}^n \mathbf{y}_k$, and $\mathbf{S} = \frac{1}{n}\mathbf{Y}\mathbf{Y}'$. For $1 \le j \le p$, let $\mathbf{Z}_j$ denote the $(p-1)\times n$ matrix obtained by removing the j-th row, denoted by $\mathbf{z}_j'$, from $\mathbf{Y}$. Similarly, for $1 \leq k \leq n$, we let $\mathbf{Y}_k$ denote the $p \times (n-1)$ matrix obtained by removing the k-th column, $\mathbf{y}_k$, from $\mathbf{Y}$. We define $\mathbf{S}_k = \frac{1}{n}\mathbf{Y}_k\mathbf{Y}_k'$ for $1 \leq k \leq n$. Similarly, for $1 \leq k \neq \ell \leq n$, $\mathbf{Y}_{k\ell}$ denotes the matrix after removing the k-th and ℓ-th columns from $\mathbf{Y}$, and $\mathbf{S}_{k\ell} = \frac{1}{n}\mathbf{Y}_{k\ell}\mathbf{Y}_{k\ell}'$. Let the minimum and the maximum eigenvalues of $\mathbf{S}$ be denoted by $\mathrm{ch}_{\min}(\mathbf{S})$ $\mathrm{ch}_{\max}(\mathbf{S})$ respectively.

We use $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{D} X$ to denote X_n converges to X in probability, and in distribution respectively. We use K to denote an absolute constant which may change from one use to another. Let $\mathbf{I}_n$ denote the $n \times n$ identity matrix, and $\mathbf{1}_n$ the n-vector of all components 1. Let $\mathbf{0}$ and $\mathbf{O}$ respectively represent a zero vector and a zero matrix of appropriate order, which is often clear in the context of their appearances. Throughout this article, vectors are column vectors equipped with the Euclidean norm. The norm of a matrix $\mathbf{B}$ is taken to be the spectral norm, $\|\mathbf{B}\| = \{\mathrm{ch}_{\max}(\mathbf{B}\mathbf{B}')\}^{1/2}$.

We consider an $n \times q$ matrix $\mathbf{A} = [a_{ij}]_{1\leq i\leq n, 1\leq j\leq q}$ satisfying the following properties:

$$\mathbf{A}'\mathbf{A} = \mathbf{I}_q, \tag{2.1}$$

$$\mathbf{A}'\mathbf{1}_n = \mathbf{0}, \tag{2.2}$$

$$\max_{1\leq i\leq n, 1\leq j\leq q} |a_{ij}| = O(n^{-1/2}). \tag{2.3}$$

We first state the following theorems which are of independent interest. We then apply Theorem 2.3 to deduce the limiting behavior of the eigenvalues in the MANOVA setting under the null case.

Theorem 2.1 *Let $p \times n$ random matrix $\mathbf{Y}$ be as described at the beginning of this section. Suppose $n \times q$ matrix $\mathbf{A}$ satisfies* (2.1) *and* (2.3)*. Suppose q is fixed, $p, n \to \infty$ in such a way that $p/n \to c \in (0,1)$, then we have*

$$\sqrt{n}\left\{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A} - \frac{p}{n}\mathbf{I}_q\right\} \xrightarrow{D} \mathbf{W},$$

where $\mathbf{W} = [w_{ij}]$ is a symmetric Gaussian Wigner matrix: for $1 \leq i \leq q$, $w_{ii} \sim N(0, 2c(1-c))$; and for $1 \leq i < j \leq q$, $w_{ij} \sim N(0, c(1-c))$. Moreover, for $1 \leq i_1 \leq j_1 \leq q$, $1 \leq i_2 \leq j_2 \leq q$ and $(i_1, j_1) \neq (i_2, j_2)$,

$$\mathrm{Cov}(w_{i_1j_1}, w_{i_2j_2}) = 0.$$

Theorem 2.2 *Let $p \times n$ random matrix $\mathbf{Y}$ and p-vector $\bar{\mathbf{y}}$ be as described at the beginning of this section. Let $n \times q$ matrix $\mathbf{A}$ satisfy* (2.1)–(2.3)*. Suppose q is fixed, $p, n \to \infty$ in such a way that $p/n \to c \in (0,1)$, then we have*

$$\sqrt{n}\left\{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}' - n\bar{\mathbf{y}}\bar{\mathbf{y}}')^{-1}\mathbf{Y}\mathbf{A} - \mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A}\right\} \xrightarrow{P} \mathbf{O}.$$

Theorem 2.3 below follows immediately from Theorems 2.1 and 2.2. The proof of Theorem 2.1 will be given in Section 4 and that of Theorem 2.2 in Section 5.

Theorem 2.3 *Under the conditions stated in Theorem* 2.2*, we have*

$$\sqrt{n}\left\{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}' - n\bar{\mathbf{y}}\bar{\mathbf{y}}')^{-1}\mathbf{Y}\mathbf{A} - \frac{p}{n}\mathbf{I}_q\right\} \xrightarrow{D} \mathbf{W},$$

where $\mathbf{W}$ is a symmetric Gaussian Wigner matrix as specified in Theorem 2.1*.*

Let $\mathbf{\Sigma}$ be a $p \times p$ positive definite matrix and $\boldsymbol{\mu}_i$ a p-vector for $i = 1, \ldots, q+1$. Define

$$\mathbf{x}_k = \mathbf{\Sigma}^{1/2}\mathbf{y}_k + \boldsymbol{\mu}_i, \quad \text{for} \quad N_{i-1} < k \le N_i, \tag{2.4}$$

where $N_0 = 0, N_i = n_1 + \cdots + n_i$, $i = 1, \ldots, q+1$. We regard the observations of the random sample from the i-th group are of the form $\mathbf{x}_k$, $N_{i-1} < k \le N_i$, which has mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\mathbf{\Sigma}$. Further, the matrices $\mathbf{S}_e$ and $\mathbf{S}_b$ defined in the Introduction section can be expressed in terms of $\mathbf{y}_k, k = 1, 2, \ldots, n$ as follows:

$$\begin{aligned}
\mathbf{S}_e &= \mathbf{\Sigma}^{1/2}\left\{\frac{1}{n}\sum_{i=1}^{q+1}\sum_{k=N_{i-1}+1}^{N_i}(\mathbf{y}_k - \bar{\mathbf{y}}_i)(\mathbf{y}_k - \bar{\mathbf{y}}_i)'\right\}\mathbf{\Sigma}^{1/2} \\
&= \mathbf{\Sigma}^{1/2}(\mathbf{S} - \alpha_{n1}\bar{\mathbf{y}}_1\bar{\mathbf{y}}_1' - \cdots - \alpha_{n,q+1}\bar{\mathbf{y}}_{q+1}\bar{\mathbf{y}}_{q+1}')\mathbf{\Sigma}^{1/2}, \\
\mathbf{S}_b &= \mathbf{\Sigma}^{1/2}\left\{\frac{1}{n}\sum_{i=1}^{q+1} n_i(\bar{\mathbf{y}}_i - \bar{\mathbf{y}} + \boldsymbol{\xi}_i)(\bar{\mathbf{y}}_i - \bar{\mathbf{y}} + \boldsymbol{\xi}_i)'\right\}\mathbf{\Sigma}^{1/2},
\end{aligned} \tag{2.5}$$

where $\alpha_{ni} := \frac{n_i}{n}$, $\bar{\mathbf{y}}_i = \frac{1}{n_i}\sum_{k=N_{i-1}+1}^{N_i}\mathbf{y}_k$, $\bar{\mathbf{y}} = \frac{1}{n}\sum_{k=1}^{n}\mathbf{y}_k = \frac{1}{n}\sum_{i=1}^{q+1}n_i\bar{\mathbf{y}}_i$, $\boldsymbol{\xi}_i = \mathbf{\Sigma}^{-1/2}(\boldsymbol{\mu}_i - \bar{\boldsymbol{\mu}})$, $\bar{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{q+1}n_i\boldsymbol{\mu}_i$. Note that $\sum_{i=1}^{q+1}n_i\boldsymbol{\xi}_i = \mathbf{0}$. Let $\mathbf{S}_t = \mathbf{S}_b + \mathbf{S}_e$.

Matrices $n\mathbf{S}_b$ and $n\mathbf{S}_e$ are commonly referred to as the matrices of sums of squares and products due to between-groups, and within-groups respectively. We impose a natural assumption:

$$\alpha_{ni} = \frac{n_i}{n} \to \alpha_i > 0, \quad i = 1, 2, \ldots, q+1. \tag{2.6}$$

Theorem 2.4 *Suppose $n - q - 1 \geq p \geq q \geq 1$. Let $n\mathbf{S}_b$ and $n\mathbf{S}_t$ be the matrices of sums of squares and products due to between-groups and total variation, based on the $\mathbf{x}_k$'s observations respectively. Let $d_1 > \cdots > d_q$ be the nonzero eigenvalues of $\mathbf{S}_b\mathbf{S}_t^{-1}$. Suppose that $\boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_{q+1}$. Put*

$$\tilde{d}_i = \frac{\sqrt{n}}{\sqrt{2c(1-c)}}\,(d_i - p/n), \quad i = 1, \ldots, q. \tag{2.7}$$

Suppose q is fixed, $p, n \to \infty$ in such a way that $p/n \to c \in (0,1)$, then the limiting joint density function of the normalized eigenvalues $(\tilde{d}_1, \tilde{d}_2, \ldots, \tilde{d}_q)$ is given by $f(y_1, \ldots, y_q; q)$ where

$$f(y_1, \ldots, y_q; q) = \frac{\pi^{q(q-1)/4}}{2^{q/2}\Gamma_q(\frac{1}{2}q)} \exp\left(-\frac{1}{2}\sum_{i=1}^{q} y_i^2\right) \prod_{1\leq i<j\leq q} (y_i - y_j), \tag{2.8}$$

and the multivariate gamma function $\Gamma_q(t) = \pi^{q(q-1)/4}\prod_{i=1}^{q}\Gamma[t-(i-1)/2]$.

The proof of Theorem 2.4 will be given in the next section. Since the nonzero eigenvalues of $\mathbf{S}_b\mathbf{S}_e^{-1}$, $\ell_1 > \cdots > \ell_q > 0$, and those of $\mathbf{S}_b\mathbf{S}_t^{-1}$, $d_1 > \cdots > d_q > 0$ are related as follows:

$$\ell_i = \frac{d_i}{1-d_i}, \quad d_i = \frac{\ell_i}{1+\ell_i}, \qquad i = 1, \ldots, q, \tag{2.9}$$

Theorem 2.4 implies the following corollary.

Corollary 2.1 *Let $\ell_1 > \cdots > \ell_q$ be the nonzero eigenvalues of $\mathbf{S}_b\mathbf{S}_e^{-1}$, and put*

$$\tilde{\ell}_i = \sqrt{n(1-c)^3/(2c)}\ \{\ell_i - p/(n-p)\}, \quad i = 1, \ldots, q. \tag{2.10}$$

Then under the same conditions as in Theorem 2.4, the limiting density function of the normalized eigenvalues $(\tilde{\ell}_1, \ldots, \tilde{\ell}_q)$ is given by $f(y_1, \ldots, y_q; q)$ in (2.8).

A significant implication of our results is the support of null robustness for nonnormality of some multivariate tests even when the dimension and the sample size are large. For example, testing the hypothesis $H_0 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_{q+1}$ in MANOVA model, we have the following three test statistics:

$$\begin{aligned}
T_1 &= -\log \frac{|\mathbf{S}_e|}{|\mathbf{S}_t|} = -\log \prod_{i=1}^{q}(1+\ell_i)^{-1} = -\sum_{i=1}^{q} \log(1-d_i), \\
T_2 &= \mathrm{tr}\mathbf{S}_b\mathbf{S}_e^{-1} = \sum_{i=1}^{q} \ell_i = \sum_{i=1}^{q} \frac{d_i}{1-d_i}, \\
T_3 &= \mathrm{tr}\mathbf{S}_b\mathbf{S}_t^{-1} = \sum_{i=1}^{q} \frac{\ell_i}{1+\ell_i} = \sum_{i=1}^{q} d_i.
\end{aligned}$$

The test statistic T_1 is based on the likelihood ratio test. The test statistics T_2 and T_3 are called Lawley-Hotelling criterion and Bartlett-Nanda-Pillai criterion, respectively (see, for example, Anderson (2003)).

Here we consider high-dimensional asymptotic null distributions of T_1, T_2 and T_3 under an asymptotic frame work $p/n \to c \in (0,1)$. Assuming normality in the populations, Wakaki et al. (2014) showed the following asymptotic results:

$$\frac{1}{\sigma}\widetilde{T}_i \xrightarrow{D} N(0,1), \quad i = 1,2,3, \tag{2.11}$$

where $\sigma = \sqrt{2q(1+p/m)}$, $m = n - p + q$ and the normalized test statistics $\tilde{T}_i, i = 1,2,3$ are defined by

$$\begin{aligned}
\widetilde{T}_1 &= \sqrt{p}\left(1+\frac{m}{p}\right)\left\{T_1 - q\log\left(1+\frac{p}{m}\right)\right\}, \\
\widetilde{T}_2 &= \sqrt{p}\left(\frac{m}{p}T_2 - q\right), \\
\widetilde{T}_3 &= \sqrt{p}\left(1+\frac{p}{m}\right)\left\{\left(1+\frac{m}{p}\right)T_3 - q\right\}.
\end{aligned} \tag{2.12}$$

These results can be expressed as

$$\widetilde{T}_i \xrightarrow{D} N(0, \frac{2q}{1-c}), \quad i = 1,2,3, \tag{2.13}$$

since $\sigma^2 \to 2q/(1-c)$. Our results show that these results continue to hold without assuming normality, but under a rather natural additional assumption (2.6).

3 Proof of Theorem 2.4

Note that the eigenvalues of the matrix $\mathbf{S}_b\mathbf{S}_t^{-1}$ are independent of Σ provided that $\boldsymbol{\mu}_i$'s are changed to $\boldsymbol{\xi}_i$'s. Hence, without loss of generality, we shall assume that $\boldsymbol{\Sigma} = \mathbf{I}_p$.

Let $\mathbf{y}_i$'s be the random vectors generating the observations $\mathbf{x}_i$'s as depicted in (2.4). Write

$$\begin{aligned}
\mathbf{Y}_b &= \left[\sqrt{\alpha_{n1}}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}), \cdots, \sqrt{\alpha_{n,q+1}}(\bar{\mathbf{y}}_{q+1} - \bar{\mathbf{y}})\right], \\
\Xi_b &= \left[\sqrt{\alpha_{n1}}\boldsymbol{\xi}_1, \cdots, \sqrt{\alpha_{n,q+1}}\boldsymbol{\xi}_{q+1}\right].
\end{aligned}$$

Hence $\mathbf{S}_b = (\mathbf{Y}_b + \Xi_b)(\mathbf{Y}_b + \Xi_b)'$ and the non-zero eigenvalues of $\mathbf{S}_b\mathbf{S}_e^{-1}$ are the same as those of $(\mathbf{Y}_b + \Xi_b)'\mathbf{S}_e^{-1}(\mathbf{Y}_b + \Xi_b)$. Note that we can rewrite

$$\mathbf{Y}_b = \mathbf{Y}\left(\frac{1}{\sqrt{n}}\mathbf{G} - \frac{1}{n}\mathbf{1}_n\mathbf{g}'\right),$$

where $\mathbf{g} = (\sqrt{\alpha_{n1}}, \ldots, \sqrt{\alpha_{n,q+1}})'$, $\mathbf{G} = \left[\frac{\mathbf{e}_1}{\sqrt{n_1}}, \ldots, \frac{\mathbf{e}_{q+1}}{\sqrt{n_{q+1}}}\right]$, and for $1 \le i \le q+1$, $\mathbf{e}_i' = (\mathbf{0}'_{N_{i-1}}, \mathbf{1}'_{n_i}, \mathbf{0}'_{n-N_i})$.

Let $\mathbf{F}$ be a $(q+1)\times q$ matrix such that $[\mathbf{F}, \mathbf{g}]$ is a $(q+1)\times(q+1)$ orthogonal matrix. It follows that

$$\mathbf{F}'\mathbf{g} = \mathbf{0}, \quad \mathbf{F}'\mathbf{F} = \mathbf{I}_q, \quad \text{and} \quad \mathbf{F}\mathbf{F}' + \mathbf{g}\mathbf{g}' = \mathbf{I}_{q+1}.$$

As $(\mathbf{Y}_b + \Xi_b)\mathbf{g} = \mathbf{0}$, we can rewrite $\mathbf{S}_b$ as

$$\begin{aligned}
\mathbf{S}_b &= (\mathbf{Y}_b + \Xi_b)[\mathbf{F}, \mathbf{g}]\left\{(\mathbf{Y}_b + \Xi_b)[\mathbf{F}, \mathbf{g}]\right\}' \\
&= \frac{1}{n}\left(\mathbf{Y}\mathbf{A} + \sqrt{n}\Xi_b\mathbf{F}\right)\left(\mathbf{Y}\mathbf{A} + \sqrt{n}\Xi_b\mathbf{F}\right)',
\end{aligned}$$

where

$$\mathbf{A} = \left(\mathbf{G} - n^{-1/2}\mathbf{1}_n\mathbf{g}'\right)\mathbf{F} = \mathbf{G}\mathbf{F}. \tag{3.14}$$

Under the assumption of Theorem 2.4 (the null case), $\Xi_b = \mathbf{O}$ together with $\boldsymbol{\Sigma} = \mathbf{I}_p$,

$$\mathbf{S}_t = \frac{1}{n}\sum_{k=1}^{n}(\mathbf{y}_k - \bar{\mathbf{y}})(\mathbf{y}_k - \bar{\mathbf{y}})' = \frac{1}{n}(\mathbf{Y}\mathbf{Y}' - n\bar{\mathbf{y}}\bar{\mathbf{y}}'),$$

and

$$\mathbf{S}_b = \mathbf{Y}_b\mathbf{Y}_b' = \frac{1}{n}(\mathbf{YA})(\mathbf{YA})',$$

the eigenvalues $d_1 > \cdots > d_q$ of $\mathbf{S}_b\mathbf{S}_t^{-1}$ are the same as those of

$$\mathbf{A}'\mathbf{Y}'(\mathbf{YY}' - n\bar{\mathbf{y}}\bar{\mathbf{y}}')^{-1}\mathbf{YA}.$$

To complete our proof, it remains to show

(i) verify that $\mathbf{A}$ satisfies (2.1)–(2.3); and

(ii) deduce Theorem 2.4 from the symmetric Gaussian Wigner matrix.

Since $\mathbf{A}'\mathbf{A} = (\mathbf{F}'\mathbf{G}')\mathbf{GF} = \mathbf{F}'\mathbf{I}_{q+1}\mathbf{F} = \mathbf{I}_q$ and $\mathbf{A}'\mathbf{1}_n = \sqrt{n}\mathbf{F}'\mathbf{g} = \mathbf{0}$, conditions (2.1) and (2.2) hold. Writing $\mathbf{A} = [a_{ij}]$ and $\mathbf{F} = [f_{ij}]$, then we have for $N_{k-1} < i \le N_k, k = 1, \ldots, q+1$,

$$|a_{ij}| = |f_{kj}/\sqrt{n_k}| \le 1/\sqrt{n_k}$$

which implies (2.3). Part (ii) follows from Theorem 13.3.1 in Anderson (2003). ■

4 Proof of Theorem 2.1

Before proceeding to the proof of Theorem 2.1, we apply a truncation technique (see page 70 of Bai and Silverstein (2010)) and hence without loss of generality, we may assume that $|y_{jk}| \le \eta_n\sqrt{n}$, for a sequence $\eta_n \downarrow 0$. For the rest of this paper, we shall assume that the basic random variables are truncated at $\eta_n\sqrt{n}$ and then re-normalized.

By Bai and Silverstein (2004), for any positive constant $\mu < a = (1-\sqrt{c})^2$ and any given $\ell > 0$, after truncation, we have

$$\mathrm{P}(\mathrm{ch}_{\min}(\mathbf{S}) < \mu) = o(n^{-\ell}). \tag{4.1}$$

Define $\mathcal{B} = 1_{\{\mathrm{ch}_{\min}(\mathbf{S}) \ge \mu\}}$ which converges to 1 a.s. as $n \to \infty$. Throughout this paper, all functions under the expectation sign are assumed having the variable $\mathcal{B}$ inserted though not explicitly indicated, for otherwise, the expectation may not exist.

By the same reasoning, the estimation (4.1) holds when $\mathbf{S}$ is replaced by $\mathbf{S}_k$ or $\mathbf{S}_{jk}$. In the paper, we shall replace $\mathcal{B}$ by $\mathcal{B}_k$ or $\mathcal{B}_{jk}$ as and when needed. Most of the time, we do this implicitly.

Before proceeding to the proof of Theorem 2.1, we shall first introduce further notations and collect some useful preliminary results in the following subsection 4.1.

4.1 Preliminary lemmas

For $1 \le k \le n$ and $1 \le j \le p$, recall $\mathbf{y}_k$ and $\mathbf{z}'_j$ denote the k-th column and j-th row of $\mathbf{Y}$ respectively; and $\mathbf{Y}_k$ and $\mathbf{Z}_j$ are the matrices after removing the k-th column and the j-th row of $\mathbf{Y}$.

Define, for $1 \le j \le p$,

$$\begin{aligned}
B_j &= \left\{\mathbf{z}'_j\mathbf{z}_j - \mathbf{z}'_j\mathbf{Z}'_j(\mathbf{Z}_j\mathbf{Z}'_j)^{-1}\mathbf{Z}_j\mathbf{z}_j\right\}^{-1}, \\
\alpha_j &= n - \operatorname{tr}\left\{\mathbf{Z}'_j(\mathbf{Z}_j\mathbf{Z}'_j)^{-1}\mathbf{Z}_j\right\} = n - p + 1, \\
W_j &= \mathbf{z}'_j\left\{\mathbf{I}_n - \mathbf{Z}'_j(\mathbf{Z}_j\mathbf{Z}'_j)^{-1}\mathbf{Z}_j\right\}\mathbf{z}_j - \alpha_j \\
\beta_j &= 1/\alpha_j = 1/(n-p+1);
\end{aligned}$$

and for $1 \le k \le n$,

$$\begin{aligned}
C_k &= \left\{1 + \mathbf{y}'_k(\mathbf{Y}_k\mathbf{Y}'_k)^{-1}\mathbf{y}_k\right\}^{-1}, \\
M_k &= \operatorname{tr}(\mathbf{Y}_k\mathbf{Y}'_k)^{-1}, \\
V_k &= \mathbf{y}'_k(\mathbf{Y}_k\mathbf{Y}'_k)^{-1}\mathbf{y}_k - M_k, \\
\bar{C}_k &= 1/(1+M_k).
\end{aligned}$$

With these notations, we have

$$B_j = \beta_j - \beta_j B_j W_j = \beta_j - \beta_j^2 W_j + \beta_j^2 B_j W_j^2, \quad 1 \le j \le p; \tag{4.2}$$
$$C_k = \bar{C}_k - \bar{C}_k C_k V_k = \bar{C}_k - \bar{C}_k^2 V_k + \bar{C}_k^2 C_k V_k^2, \quad 1 \le k \le n. \tag{4.3}$$

Lemma 4.1 *For any $n \times n$ positive definite matrix $\mathbf{B}$ and n-vector $\mathbf{r}$, we have*

$$
\begin{aligned}
(\mathbf{B} + \mathbf{r}\mathbf{r}')^{-1} &= \mathbf{B}^{-1} - \frac{\mathbf{B}^{-1}\mathbf{r}\mathbf{r}'\mathbf{B}^{-1}}{1 + \mathbf{r}'\mathbf{B}^{-1}\mathbf{r}}, \\
(\mathbf{B} + \mathbf{r}\mathbf{r}')^{-1}\mathbf{r} &= \mathbf{r}'(\mathbf{B} + \mathbf{r}\mathbf{r}')^{-1} = \frac{\mathbf{B}^{-1}\mathbf{r}}{1 + \mathbf{r}'\mathbf{B}^{-1}\mathbf{r}}.
\end{aligned}
$$

Proof. The second identity follows immediately from the first identity,which in turn can be verified directly. ■

Lemma 4.2 *Let $r \geq 1$ be an integer.*

(a) *We have*

$$\lim_{n\to\infty} \left\{ \frac{1}{n} \mathbf{y}_k' \mathbf{S}_k^{-r} \mathbf{y}_k - \frac{1}{n} \mathrm{tr} \mathbf{S}_k^{-r} \right\} = 0 \ a.s.$$

uniformly in $1 \leq k \leq n$.

(b) *Moreover, uniformly over $1 \leq k \leq n$ on the set $\cup_{k=1}^{n} \{\mathrm{ch}_{\min}(\mathbf{S}_k) \geq \mu\}$ where $0 < \mu < a = (1 - \sqrt{c})^2$,*

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{tr} \mathbf{S}_k^{-r} \overset{a.s.}{=} c \int_a^b x^{-r} \frac{\sqrt{(b-x)(x-a)}}{2\pi c x} dx := c_r.$$

Here $a = (1 - \sqrt{c})^2$ and $b = (1 + \sqrt{c})^2$. In particular, $c_1 = c/(1 - c)$ and $c_2 = c/(1 - c)^3$.

Proof. By (9.7.9) of Bai and Silverstein (2009), we have $P(\|\mathbf{S}_k^{-1}\| > \mu_2) = o(n^{-t})$ for any fixed $t > 0$, where $\mu_2 > a^{-1} = (1 - \sqrt{c})^{-2}$. For any $\varepsilon > 0$, we apply Lemma 9.1 of Bai and Silverstein (2009) to obtain

$$
\begin{aligned}
& P\left(\left| \frac{1}{n} \mathbf{y}_k' \mathbf{S}_k^{-r} \mathbf{y}_k - \frac{1}{n} \mathrm{tr} \mathbf{S}_k^{-r} \right| \geq \varepsilon \right) \\
\leq \ & (n\varepsilon)^{-m} \mathrm{E} \left\{ \left| \mathbf{y}_k' \mathbf{S}_k^{-r} \mathbf{y}_k - \mathrm{tr} \mathbf{S}_k^{-r} \right|^m I(\|\mathbf{S}_k^{-1}\| \leq \mu_2) \right\} + P(\|\mathbf{S}_k^{-1}\| > \mu_2) \\
\leq \ & (n\varepsilon)^{-m} \nu n^m (n\eta_n^4)^{-1} (40 b^2 \mu_2^r \eta_n^2)^m + o(n^{-t}) \\
= \ & o(n^{-t})
\end{aligned}
$$

for any $t > 0$ by choosing $m = [\log n]$. Part (a) follows by choosing $t > 1$ and Borel-Cantelli lemma.

(b) First equality follows from Marčenko-Pastur (MP) law. Indeed, by Bai and Silverstein (2009), the spectral distribution of $\mathbf{S} = \frac{1}{n}\mathbf{Y}\mathbf{Y}'$ tends to the MP law a.s., and its largest and smallest eigenvalues tend to $b = (1+\sqrt{c})^2$ and $a = (1-\sqrt{c})^2$ a.s. respectively. The same is true for the matrix $\mathbf{S}_k = \frac{1}{n}\mathbf{Y}_k\mathbf{Y}_k'$ for each k. Hence

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\mathrm{tr}\mathbf{S}_k^{-r} &= \lim_{n\to\infty} \frac{p}{n}\left(\frac{1}{p}\mathrm{tr}\mathbf{S}_k^{-r}\right) \\ &= c\int_a^b x^{-r}\frac{\sqrt{(b-x)(x-a)}}{2\pi c x}dx.\end{aligned}$$

That $c_1 = c/(1-c)$ and $c_2 = c/(1-c)^3$ can be proved by contour integration or series expansion. ■

Lemma 4.3 *Let* $\mathbf{B}$ *and* $\mathbf{C}$ *be* $n \times n$ *matrices. Let* $\mathbf{y} = (y_1, \ldots, y_n)'$ *be a random n-vector with* y_i*'s being independent, mean 0, variance 1, and finite 4-th moments. Let* $\kappa := \max_{1\le i\le n} \mathrm{E}(y_i^4) - 3$.

(a) *If, in addition,* y_i*'s are assumed to be identically distributed, then*

$$\mathrm{E}\left\{(\mathbf{y}'\mathbf{B}\mathbf{y} - \mathrm{tr}\mathbf{B})(\mathbf{y}'\mathbf{C}\mathbf{y} - \mathrm{tr}\mathbf{C})\right\} = \mathrm{tr}(\mathbf{B}\mathbf{C}) + \mathrm{tr}(\mathbf{B}\mathbf{C}') + \kappa\sum_{i=1}^n b_{ii}c_{ii}. \quad (4.4)$$

(b) *Suppose* $\mathbf{B}$ *is symmetric, then*

$$\mathrm{E}\left\{(\mathbf{y}'\mathbf{B}\mathbf{y} - \mathrm{tr}\mathbf{B})^2\right\} \le (\kappa+2)\mathrm{tr}(\mathbf{B}^2), \quad (4.5)$$

$$\mathrm{E}\left\{(\mathbf{y}'\mathbf{B}\mathbf{y} - \mathrm{tr}\mathbf{B})^4\right\} = o(n^3), \quad (4.6)$$

where we assume further that the norm of $\mathbf{B}$ *is bounded, and that* $|y_i| \le \eta_n\sqrt{n}$ *for some sequence* $\eta_n \downarrow 0$ *in (4.6).*

Proof. (a) We have

$$\mathrm{E}\left\{(\mathbf{y}'\mathbf{B}\mathbf{y} - \mathrm{tr}\mathbf{B})(\mathbf{y}'\mathbf{C}\mathbf{y} - \mathrm{tr}\mathbf{C})\right\} = \mathrm{E}(\mathbf{y}'\mathbf{B}\mathbf{y}\mathbf{y}'\mathbf{C}\mathbf{y}) - \mathrm{tr}\mathbf{B}\,\mathrm{tr}\mathbf{C}$$

$$
\begin{aligned}
&= \sum_{1\le i,j,k,l\le n} b_{ij}c_{kl}\mathrm{E}(y_iy_jy_ky_l) - \mathrm{tr}\mathbf{B}\,\mathrm{tr}\mathbf{C} \\
&= \sum_{i=1}^n b_{ii}c_{ii}\mathrm{E}(y_1^4) + \sum_{i=1}^n \sum_{k=1,k\neq i}^n b_{ii}c_{kk} + \sum_{i=1}^n \sum_{j=1,j\neq i}^n b_{ij}c_{ij} \\
&\quad + \sum_{i=1}^n \sum_{j=1,j\neq i}^n b_{ij}c_{ji} - \mathrm{tr}\mathbf{B}\,\mathrm{tr}\mathbf{C} \\
&= \mathrm{tr}(\mathbf{BC}) + \mathrm{tr}(\mathbf{BC}') + \left\{\mathrm{E}(y_1^4) - 3\right\}\sum_{i=1}^n b_{ii}c_{ii}.
\end{aligned}
$$

This proves (4.4).

The method of proof above can be easily modified to prove (4.5) by taking $\mathbf{B} = \mathbf{C}$ and that $\mathbf{B}$ is assumed to be symmetric.

To prove (4.6), we apply $(a+b)^4 \le 8(a^4+b^4)$ to obtain

$$
\mathrm{E}\left\{(\mathbf{y}'\mathbf{B}\mathbf{y} - \mathrm{tr}\mathbf{B})^4\right\} \le 8\left[\mathrm{E}\left\{\sum_{i=1}^n b_{ii}(y_i^2-1)\right\}^4 + \mathrm{E}\left(2\sum_{1\le i<j\le n} b_{ij}y_iy_j\right)^4\right].
$$

The first term can be bounded as follows

$$
\begin{aligned}
&\mathrm{E}\left[\left\{\sum_{i=1}^n b_{ii}(y_i^2-1)\right\}^4\right] \\
&= \sum_{i=1}^n b_{ii}^4\mathrm{E}\left\{(y_i^2-1)^4\right\} + 6\sum_{1\le i<j\le n} b_{ii}^2b_{jj}^2\mathrm{E}\left\{(y_i^2-1)^2\right\}\,\mathrm{E}\left\{(y_j^2-1)^2\right\} = o(n^3),
\end{aligned}
$$

and the second term

$$
\begin{aligned}
&\mathrm{E}\left\{\left(\sum_{1\le i<j\le n} b_{ij}y_iy_j\right)^4\right\} \\
&= \sum_{1\le i<j\le n} b_{ij}^4\mathrm{E}(y_i^4)\,\mathrm{E}(y_j^4) + 3\sum_{\substack{1\le i<j\le n, 1\le k<\ell\le n \\ (i,j)\neq(k,\ell)}} b_{ij}^2b_{k\ell}^2\mathrm{E}(y_i^2y_j^2y_k^2y_\ell^2) = o(n^3).
\end{aligned}
$$

Further details to obtain the last equality can be found in Lemma B.26 in Bai and Silverstein (2009). ■

Lemma 4.4 *As $p, n \to \infty$ such that $c_{n,p} := p/n \to c \in (0,1)$, we have*

$$\sqrt{n}\left[\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_1\right\} - c_{n,p}\right] = O(n^{-1/2}), \tag{4.7}$$

$$\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_2\right\} = O(n^{-1}). \tag{4.8}$$

Proof. Defining $\varepsilon_1 = \frac{1}{n}(\mathbf{y}_1'\mathbf{S}_1^{-1}\mathbf{y}_1 - \mathrm{tr}\mathbf{S}_1^{-1})$ and using Lemma 4.1, we have

$$\begin{aligned}
\sqrt{n}\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_1 - c_{n,p}\right\} &= \sqrt{n}\mathrm{E}\left\{\frac{\mathbf{y}_1'(\mathbf{Y}_1\mathbf{Y}_1')^{-1}\mathbf{y}_1}{1+\mathbf{y}_1'(\mathbf{Y}_1\mathbf{Y}_1')^{-1}\mathbf{y}_1} - c_{n,p}\right\} \\
&= \sqrt{n}\mathrm{E}\left\{\frac{(1-c_{n,p})\frac{1}{n}\mathbf{y}_1'\mathbf{S}_1^{-1}\mathbf{y}_1 - c_{n,p}}{1+\frac{1}{n}\mathbf{y}_1'\mathbf{S}_1^{-1}\mathbf{y}_1}\right\} \\
&= \sqrt{n}\mathrm{E}\left\{\frac{(1-c_{n,p})\frac{1}{n}\mathrm{tr}\mathbf{S}_1^{-1} - c_{n,p}}{1+\frac{1}{n}\mathrm{tr}\mathbf{S}_1^{-1}} + \frac{\varepsilon_1}{(1+\frac{1}{n}\mathrm{tr}\mathbf{S}_1^{-1})(1+\frac{1}{n}\mathbf{y}_1'\mathbf{S}_1^{-1}\mathbf{y}_1)}\right\} \\
&= R_{n1} + R_{n2}.
\end{aligned}$$

To complete the proof of (4.7), we shall show below that $R_{n1}, R_{n2} = O(n^{-1/2})$.

From Bai and Silverstein (2004), we have

$$\begin{aligned}
\frac{1}{1-\frac{p}{n-1}} &= \int x^{-1} dF_{\frac{p}{n-1}}(x) \\
&= \int x^{-1} dF^{\mathbf{S}_1}(x) + \int x^{-1} d\left(F_{\frac{p}{n-1}}(x) - F^{\mathbf{S}_1}(x)\right) \\
&= \frac{1}{p}\mathrm{tr}(\mathbf{S}_1^{-1}) + O_P(n^{-1}).
\end{aligned}$$

Here F_y denotes the Marčenko-Pastur law with parameter y, and $F^{\mathbf{S}_1}$ the empirical spectral distribution of $\mathbf{S}_1$. So

$$|R_{n1}| \le \sqrt{n}\mathrm{E}\left|c_{n,p}(1-c_{n,p})\frac{1}{p}\mathrm{tr}\mathbf{S}_1^{-1} - c_{n,p}\right| = O(n^{-1/2}).$$

Next we show that $R_{n2} = O(n^{-1/2})$. We decompose

$$\begin{aligned}
R_{n2} &= \sqrt{n}\mathrm{E}\left\{\frac{\varepsilon_1}{(1+\frac{1}{n}\mathrm{tr}\mathbf{S}_1^{-1})(1+\frac{1}{n}\mathbf{y}_1'\mathbf{S}_1^{-1}\mathbf{y}_1)}\right\} \\
&= \sqrt{n}\mathrm{E}\left\{\frac{\varepsilon_1}{(1+\frac{1}{n}\mathrm{tr}\mathbf{S}_1^{-1})^2}\right\} - \sqrt{n}\mathrm{E}\left\{\frac{\varepsilon_1^2}{(1+\frac{1}{n}\mathrm{tr}\mathbf{S}_1^{-1})^2(1+\frac{1}{n}\mathbf{y}_1'\mathbf{S}_1^{-1}\mathbf{y}_1)}\right\} \\
&= -\sqrt{n}\mathrm{E}\left\{\frac{\varepsilon_1^2}{(1+\frac{1}{n}\mathrm{tr}\mathbf{S}_1^{-1})^2(1+\frac{1}{n}\mathbf{y}_1'\mathbf{S}_1^{-1}\mathbf{y}_1)}\right\}.
\end{aligned}$$

Hence

$$|R_{n2}| \le \sqrt{n}\mathrm{E}(\varepsilon_1^2) \le \frac{\kappa+2}{\sqrt{n}}\mathrm{E}\left\{\frac{1}{n}\mathrm{tr}\left(\mathbf{S}_1^{-2}\right)\right\} = O(n^{-1/2}),$$

where we used Lemmas 4.3 and 4.2 in the second inequality and the last equality respectively. This proves (4.7).

To prove (4.8), we recall the definitions of $C_k, \bar{C}_k$ and V_k at the beginning of Subsection 4.1. We need an additional notation:

$$C_{jk} := \left\{1 + \mathbf{y}_k'(\mathbf{Y}_{jk}\mathbf{Y}_{jk}')^{-1}\mathbf{y}_k\right\}^{-1}, \quad 1 \le j \ne k \le n. \tag{4.9}$$

We remark that $C_{jk} \ne C_{kj}$, and $0 < C_{jk} < 1$ a.s.

Applying Lemma 4.1 twice in the first equality below, we obtain

$$\begin{aligned} \left|\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_2\right\}\right| &= \left|\mathrm{E}\left\{C_1 C_{12}\mathbf{y}_1'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-1}\mathbf{y}_2\right\}\right| \\ &= \left|\mathrm{E}\left\{(C_1 - \bar{C}_1)C_{12}\mathbf{y}_1'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-1}\mathbf{y}_2\right\}\right| \\ &\le \left[\mathrm{E}\left\{(C_1 - \bar{C}_1)^2\right\}\ \mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-1}\mathbf{y}_2\right\}^2\right]^{1/2}. \end{aligned}$$

We first prove that $\mathrm{E}\left\{(C_1 - \bar{C}_1)^2\right\} = \mathrm{E}\left[\{\mathbf{y}_1'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-1}\mathbf{y}_2\}^2\right] = O(n^{-1})$. By (4.3) and $0 < C_1 < 1$, we have

$$\mathrm{E}\left\{(C_1 - \bar{C}_1)^2\right\} = \mathrm{E}\left(C_1^2\bar{C}_1^2V_1^2\right) \le \mathrm{E}\left(V_1^2\right) \le K\mathrm{E}\left[\mathrm{tr}\left\{(\mathbf{Y}_1\mathbf{Y}_1')^{-2}\right\}\right] = O(n^{-1}),$$

where we applied Lemma 4.3 in the second inequality and Lemma 4.2 in the last equality. By Lemma 4.2 again, we have

$$\begin{aligned} \mathrm{E}\left[\left\{\mathbf{y}_1'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-1}\mathbf{y}_2\right\}^2\right] &= \mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-1}\mathbf{y}_2\,\mathbf{y}_2'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-1}\mathbf{y}_1\right\} \\ &= \mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}_{12}\mathbf{Y}_{12}')^{-2}\mathbf{y}_1\right\} \\ &= n^{-2}\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{S}_{12})^{-2}\mathbf{y}_1\right\} = O(n^{-1}). \end{aligned}$$

So $|\mathrm{E}\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_2\}| = O(n^{-1})$, which is (4.8). This proves Lemma 4.4. ■

4.2 Proof

Our proof of Theorem 2.1 consists of three steps. In step 1, we compute the mean, $\mathbf{M} = \mathrm{E}\{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A}\}$, and prove that $\sqrt{n}\left(\mathbf{M} - \frac{p}{n}\mathbf{I}_q\right) \to \mathbf{O}$. In step 2, we prove the convergence of $\mathbf{V} := \sqrt{n}\,\{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A} - \mathbf{M}\}$ to a Gaussian Wigner matrix using martingale decomposition. Finally, in step 3, we compute the covariance of any two entries in the Wigner matrix.

Step 1. We shall first compute m_{ij}, the (i,j)-th entry of $\mathbf{M}$ and then deduce that

$$\sqrt{n}\left(\mathbf{M} - \frac{p}{n}\mathbf{A}'\mathbf{A}\right) = \sqrt{n}\,(\mathbf{M} - c_{n,p}\mathbf{I}_q) \to \mathbf{O} \tag{4.10}$$

where $c_{n,p} = p/n$. Recall $\mathbf{a}_k$ denotes the k-th column of $\mathbf{A}$, we have

$$\begin{aligned} m_{ij} &= \mathbf{a}_i'\mathrm{E}\left\{\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\right\}\mathbf{a}_j = \sum_{1\le r,s\le n} a_{ri}a_{sj}\mathrm{E}\left\{\mathbf{y}_r'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_s\right\} \\ &= \mathbf{a}_i'\mathbf{a}_j\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_1\right\} + \{(\mathbf{1}_n'\mathbf{a}_i)(\mathbf{1}_n'\mathbf{a}_j) - \mathbf{a}_i'\mathbf{a}_j\}\ \mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_2\right\} \\ &= \mathbf{a}_i'\mathbf{a}_j\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_1\right\} + O(n^{-1}) \end{aligned}$$

by using (4.8). Applying (4.7) and (4.8) in Lemma 4.4, we have

$$\sqrt{n}\,(m_{ii} - c_{n,p}) = \sqrt{n}\left[\mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_1\right\} - c_{n,p} + O(n^{-1})\right] \to 0.$$

This proves (4.10), and completes step 1.

Step 2. We shall prove that $\mathbf{V} \xrightarrow{D} \mathbf{W}$ via martingale decomposition. With $\mathrm{E}_k, \mathbf{D}_k$ and $\widetilde{\mathbf{D}}_k$ defined below, we outline our approach as follows: Rewrite $\mathbf{V}$ as

$$\begin{aligned} \mathbf{V} &= \sqrt{n}\sum_{k=1}^{p}(\mathrm{E}_k - \mathrm{E}_{k-1})(\mathbf{D}_k) \\ &= \sqrt{n}\sum_{k=1}^{p}\mathrm{E}_k(\widetilde{\mathbf{D}}_k) + o_P(1), \end{aligned} \tag{4.11}$$

and then apply martingale central limit theorem to show that $\sqrt{n}\sum_{k=1}^{n}\mathrm{E}_k(\widetilde{\mathbf{D}}_k)$ converges to a Gaussian Wigner matrix.

First recall $\mathbf{Z}_k$ is the $(p-1)\times n$ matrix obtained from removing the k-th row $\mathbf{z}_k'$ from $\mathbf{Y}$. Let E_k denote the conditional expectation given $\{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$, and E_0 is the usual expectation. We apply $(\mathrm{E}_k - \mathrm{E}_{k-1})\,\{\mathbf{A}'\mathbf{Z}_k'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{Z}_k\mathbf{A}\} = \mathbf{0}$ to obtain

$$\mathbf{V} = \sqrt{n}\sum_{k=1}^{p}(\mathrm{E}_k - \mathrm{E}_{k-1})\left\{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A} - \mathbf{A}'\mathbf{Z}_k'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{Z}_k\mathbf{A}\right\}$$

$$= \sqrt{n}\sum_{k=1}^{p}(\mathrm{E}_k - \mathrm{E}_{k-1})(\mathbf{D}_k),$$

where

$$\mathbf{D}_k = B_k\left(\mathbf{A}'\mathbf{P}_k\mathbf{z}_k\mathbf{z}_k'\mathbf{P}_k\mathbf{A}\right) \tag{4.12}$$

and $\mathbf{P}_k = \mathbf{I}_n - \mathbf{Z}_k'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{Z}_k$. In the last equality, we used the following result

$$\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y} = \mathbf{Z}_k'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{Z}_k + B_k\mathbf{P}_k\mathbf{z}_k\mathbf{z}_k'\mathbf{P}_k.$$

We introduce

$$\mathbf{Q}_k = \mathbf{A}'\mathbf{P}_k(\mathbf{z}_k\mathbf{z}_k' - \mathbf{I}_n)\mathbf{P}_k\mathbf{A}.$$

With B_k, W_k and β_k defined at the beginning of Subsection 4.1, we define

$$\widetilde{\mathbf{D}}_k = \beta_k\mathbf{Q}_k.$$

Note that $\mathbf{D}_k = B_k(\mathbf{Q}_k + \mathbf{A}'\mathbf{P}_k\mathbf{A})$. Applying (4.2), we have

$$\begin{aligned}\mathbf{D}_k - \widetilde{\mathbf{D}}_k &= -\beta_k B_k W_k\mathbf{Q}_k + B_k\mathbf{A}'\mathbf{P}_k\mathbf{A}\\ &= -\beta_k B_k W_k\mathbf{Q}_k + \beta_k\mathbf{A}'\mathbf{P}_k\mathbf{A} - \beta_k B_k W_k\mathbf{A}'\mathbf{P}_k\mathbf{A}.\end{aligned}$$

Note also $(\mathrm{E}_k - \mathrm{E}_{k-1})(\beta_k\mathbf{A}'\mathbf{P}_k\mathbf{A}) = \mathbf{0}$. Since $\beta_k = \frac{1}{n-p+1} \approx \frac{1}{(1-c)n}$, we have

$$\begin{aligned}&\mathrm{E}\left\{\|(\mathrm{E}_k - \mathrm{E}_{k-1})(\mathbf{D}_k - \widetilde{\mathbf{D}}_k)\|^2\right\}\\ &\le 2\mathrm{E}(\beta_k^2B_k^2W_k^2\|\mathbf{Q}_k\|^2) + 2\mathrm{E}(\beta_k^2B_k^2W_k^2\|\mathbf{A}'\mathbf{P}_k\mathbf{A}\|^2)\\ &\le Kn^{-4}\left\{\mathrm{E}(W_k^2\|\mathbf{Q}_k\|^2) + \mathrm{E}(W_k^2)\right\} + o(n^{-t}) = o(n^{-2}).\end{aligned}$$

In the second inequality, we used the fact that $\|\mathbf{A}'\mathbf{P}_k\mathbf{A}\| \le 1$; and considered two complementary events: $|B_k| \le 1/n$ and $|B_k| > 1/n$ with the latter event giving rise to $o(n^{-t})$. In the last equality, we applied Lemma 4.3 to bound $\mathrm{E}(W_k^2) \le K\mathrm{E}\left(\mathrm{tr}(\mathbf{P}_k^2)\right) = K\mathrm{E}\left(\mathrm{tr}(\mathbf{P}_k)\right) \le Kn$; and Lemma 4.5 below to obtain $\mathrm{E}(W_k^2\|\mathbf{Q}_k\|^2) = o(n^2)$. So

$$\mathrm{E}\left\{\left\|\sqrt{n}\sum_{k=1}^{p}(\mathrm{E}_k - \mathrm{E}_{k-1})(\mathbf{D}_k - \widetilde{\mathbf{D}}_k)\right\|^2\right\} = o(1),$$

proving (4.11).

Let $\mathrm{E}^{(k)}$ denote the expectation with respect to $\mathbf{z}_k$. Lemma 4.6 below shows that $\mathrm{E}^{(k)}\left(\|\mathbf{Q}_k\|^4\right) = o(n)$. Consequently

$$\sum_{k=1}^{p} \mathrm{E}\left(\|\sqrt{n}\widetilde{\mathbf{D}}_k\|^4\right) = n^2 \sum_{k=1}^{p} \beta_k^4 \mathrm{E}\left(\|\mathbf{Q}_k\|^4\right) = o(1).$$

In other words, the sequence of martingale differences $\{\sqrt{n}\mathrm{E}_k(\widetilde{\mathbf{D}}_k)\}$ satisfies the Lyapunov condition, and thus $\mathbf{V} \xrightarrow{D} \mathbf{W}$. The matrix norm used for the above proof is, for convenience purposes, the Euclidean norm which differs from the spectral norm by a fixed factor since the order of the matrix under consideration is fixed.

Step 3. This step concerns the computation of the covariance of a pair of entries in the Wigner matrix, $\mathbf{W}$. Let unit n-vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ denote 4 columns of $\mathbf{A}$, either identical or orthogonal. We need to compute

$$\begin{aligned}
& \frac{1}{n(1-c)^2} \sum_{k=1}^{p} \mathrm{E}_{k-1}\left[\mathrm{E}_k\left\{\mathbf{a}'\mathbf{P}_k(\mathbf{z}_k\mathbf{z}_k' - \mathbf{I}_n)\mathbf{P}_k\mathbf{b}\right\} \mathrm{E}_k\left\{\mathbf{c}'\mathbf{P}_k(\mathbf{z}_k\mathbf{z}_k' - \mathbf{I}_n)\mathbf{P}_k\mathbf{d}\right\}\right] \\
= & \frac{1}{n(1-c)^2} \sum_{k=1}^{p} \mathrm{E}_{k-1}\big[\left\{\mathbf{z}_k'\mathrm{E}_k(\mathbf{P}_k\mathbf{b}\mathbf{a}'\mathbf{P}_k)\mathbf{z}_k - \mathrm{tr}\mathrm{E}_k(\mathbf{P}_k\mathbf{b}\mathbf{a}'\mathbf{P}_k)\right\} \\
& \quad \times \left\{\mathbf{z}_k'\mathrm{E}_k(\mathbf{P}_k\mathbf{d}\mathbf{c}'\mathbf{P}_k)\mathbf{z}_k - \mathrm{tr}(\mathrm{E}_k\mathbf{P}_k\mathbf{d}\mathbf{c}'\mathbf{P}_k)\right\}\big] \\
= & \frac{1}{n(1-c)^2} \sum_{k=1}^{p} \mathrm{E}_{k-1}[\mathrm{tr}\left\{\mathrm{E}_k(\mathbf{P}_k\mathbf{b}\mathbf{a}'\mathbf{P}_k)\, \mathrm{E}_k(\mathbf{P}_k\mathbf{d}\mathbf{c}'\mathbf{P}_k)\right\} \\
& \quad + \mathrm{tr}\left\{\mathrm{E}_k(\mathbf{P}_k\mathbf{b}\mathbf{a}'\mathbf{P}_k)\, \mathrm{E}_k(\mathbf{P}_k\mathbf{c}\mathbf{d}'\mathbf{P}_k)\right\}] + R_n \\
= & \frac{1}{n(1-c)^2} \sum_{k=1}^{p} \mathrm{E}_k\left\{\mathrm{tr}(\mathbf{P}_k\mathbf{b}\mathbf{a}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{d}\mathbf{c}'\breve{\mathbf{P}}_k) + \mathrm{tr}(\mathbf{P}_k\mathbf{b}\mathbf{a}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{c}\mathbf{d}'\breve{\mathbf{P}}_k)\right\} + R_n \\
= & \frac{1}{n(1-c)^2} \sum_{k=1}^{p} \mathrm{E}_k\{(\mathbf{a}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{d})\,(\mathbf{b}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{c}) + (\mathbf{a}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{c})\,(\mathbf{b}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{d})\} + R_n \\
= & \frac{1}{(1-c)^2}\left\{I(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}) + I(\mathbf{a},\mathbf{b},\mathbf{d},\mathbf{c})\right\} + R_n,
\end{aligned}$$

where we applied Lemma 4.3(a) in the second equality. Here

$$I(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}) = \frac{1}{n}\sum_{k=1}^{p} \mathrm{E}_k\{(\mathbf{a}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{d})(\mathbf{b}'\mathbf{P}_k\breve{\mathbf{P}}_k\mathbf{c})\}, \tag{4.13}$$

$$R_n = \frac{\mathrm{E}(y_{11}^4)}{n(1-c)^2} \sum_{k=1}^{p} \mathrm{E}_{k-1}\left(\sum_{i=1}^{n} \mathbf{a}'\mathbf{P}_k\mathbf{e}_i\, \mathbf{b}'\mathbf{P}_k\mathbf{e}_i\, \mathbf{c}'\breve{\mathbf{P}}_k\mathbf{e}_i\, \mathbf{d}'\breve{\mathbf{P}}_k\mathbf{e}_i\right),$$

and $\breve{\mathbf{P}}_k$ is the projection matrix onto the space spanned by the vectors $\mathbf{z}_1, \cdots, \mathbf{z}_{k-1}, \breve{\mathbf{z}}_{k+1}, \cdots, \breve{\mathbf{z}}_p$. Here $\breve{\mathbf{z}}_j$'s are *i.i.d.* copies of $\mathbf{z}_j$'s.

By Lemma 4.7 below, we have $R_n = O_P(n^{-1})$. Therefore, we only need to compute the limit for $I(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$. We proceed to remove successively row by row starting from the last row until we reach the k-th row. We may assume $k < p$. For $k < j \le p$, let $\mathbf{Z}_{kj}$ denote the $(j-2) \times n$ matrix consisting of the rows $\mathbf{z}_1', \ldots, \mathbf{z}_{k-1}', \mathbf{z}_{k+1}', \ldots, \mathbf{z}_{j-1}'$ and $\mathbf{P}_{kj} = \mathbf{I}_n - \mathbf{Z}_{kj}'(\mathbf{Z}_{kj}\mathbf{Z}_{kj}')^{-1}\mathbf{Z}_{kj}$. Define

$$
\begin{aligned}
B_{kj} &= 1/(\mathbf{z}_j'\mathbf{P}_{kj}\mathbf{z}_j), \\
\beta_{kj} &= 1/\mathrm{tr}\mathbf{P}_{kj} = 1/(n-j+2), \\
W_{kj} &= \mathbf{z}_j'\mathbf{P}_{kj}\mathbf{z}_j - \mathrm{tr}\mathbf{P}_{kj}.
\end{aligned}
$$

Note that $\mathbf{P}_k$ is a projection matrix onto the orthogonal complement of the space spanned by the rows of $\mathbf{Z}_k$, we have

$$
\begin{aligned}
\mathbf{P}_k &= \mathbf{P}_{kp} - B_{kp}\mathbf{P}_{kp}\mathbf{z}_p\mathbf{z}_p'\mathbf{P}_{kp} \\
&= (1-\beta_{kp})\mathbf{P}_{kp} - \beta_{kp}\mathbf{R}_{k1} + \beta_{kp}\mathbf{R}_{k2},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{R}_{k1} &= \mathbf{P}_{kp}(\mathbf{z}_p\mathbf{z}_p' - \mathbf{I}_n)\mathbf{P}_{kp}, \\
\mathbf{R}_{k2} &= B_{kp}W_{kp}\mathbf{P}_{kp}\mathbf{z}_p\mathbf{z}_p'\mathbf{P}_{kp}.
\end{aligned}
$$

Substituting this into (4.13) and noting that $\mathbf{z}_p$ is independent of $\breve{\mathbf{P}}_k$, we obtain

$$
\begin{aligned}
& I(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) \\
= \; & \frac{1}{n}\sum_{k=1}^{p}\left(\frac{n-p+1}{n-p+2}\right)^2 \mathrm{E}_k\mathbf{a}'\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{b}'\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{c} + \frac{1}{n}\sum_{k=1}^{p}\sum_{j=1}^{6}\Delta_{kj}, \quad (4.14)
\end{aligned}
$$

where

$$
\begin{aligned}
\Delta_{k1} &= (n-p+2)^{-2}\,\mathrm{E}_k(\mathbf{a}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{b}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{c}), \\
\Delta_{k2} &= (n-p+1)/(n-p+2)^2\,\mathrm{E}_k(\mathbf{a}'\mathbf{R}_{k2}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{b}'\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{c}), \\
\Delta_{k3} &= (n-p+1)/(n-p+2)^2\,\mathrm{E}_k(\mathbf{a}'\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{b}'\mathbf{R}_{k2}\breve{\mathbf{P}}_k\mathbf{c}), \\
\Delta_{k4} &= -(n-p+2)^{-2}\,\mathrm{E}_k(\mathbf{a}'\mathbf{R}_{k2}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{b}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{c}), \\
\Delta_{k5} &= -(n-p+2)^{-2}\,\mathrm{E}_k(\mathbf{a}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{b}'\mathbf{R}_{k2}\breve{\mathbf{P}}_k\mathbf{c}), \\
\Delta_{k6} &= (n-p+2)^{-2}\,\mathrm{E}_k(\mathbf{a}'\mathbf{R}_{k2}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{b}'\mathbf{R}_{k2}\breve{\mathbf{P}}_k\mathbf{c}).
\end{aligned}
$$

By Lemma 4.7, we have $\Delta_{kj} = O_P(n^{-2})$, for $j = 1, 2, 3, 4, 5, 6$. Expanding $\breve{\mathbf{P}}_k$ in a similar way, we obtain

$$\begin{aligned} & I(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) \\ = \;& \frac{1}{n} \sum_{k=1}^{p} \left(\frac{n-p+1}{n-p+2} \right)^4 \mathrm{E}_k(\mathbf{a}'\mathbf{P}_{kp}\breve{\mathbf{P}}_{kp}\mathbf{d}\mathbf{b}'\mathbf{P}_{kp}\breve{\mathbf{P}}_{kp}\mathbf{c}) + O_P(n^{-2}), \end{aligned} \tag{4.15}$$

Repeating this step successively for $j = p-1, \ldots, k+1$, we obtain

$$\begin{aligned} & I(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) \\ = \;& \frac{1}{n} \sum_{k=1}^{p} \left(\frac{n-p+1}{n-j+2} \right)^4 \mathrm{E}_k(\mathbf{a}'\mathbf{P}_{kj}\breve{\mathbf{P}}_{kj}\mathbf{d}\mathbf{b}'\mathbf{P}_{kj}\breve{\mathbf{P}}_{kj}\mathbf{c}) + O_P\left(\frac{p-j+1}{n^2} \right) \\ = \;& \frac{1}{n} \sum_{k=1}^{p} \left(\frac{n-p+1}{n-k+2} \right)^4 \mathrm{E}_k(\mathbf{a}'\mathbf{P}_{kk}\breve{\mathbf{P}}_{kk}\mathbf{d}\mathbf{b}'\mathbf{P}_{kk}\breve{\mathbf{P}}_{kk}\mathbf{c}) + O_P(n^{-1}) \\ = \;& \frac{1}{n} \sum_{k=1}^{p} \left(\frac{n-p+1}{n-k+2} \right)^4 \mathbf{a}'\mathbf{P}_{kk}\mathbf{d}\, \mathbf{b}'\mathbf{P}_{kk}\mathbf{c} + O_P(n^{-1}), \end{aligned} \tag{4.16}$$

where $\mathbf{P}_{kk} = \mathbf{I}_n - \mathbf{Z}_{kk}'(\mathbf{Z}_{kk}\mathbf{Z}_{kk}')^{-1}\mathbf{Z}_{kk}$ and $\mathbf{Z}_{kk}$ is $(k-1) \times n$ consisting of rows $\mathbf{z}_1', \cdots, \mathbf{z}_{k-1}'$. The last step in (4.16) follows from the fact that $\breve{\mathbf{P}}_{kk} = \mathbf{P}_{k,k}$ that is a projection matrix.

By Lemma 4.8, we have

$$\begin{aligned} I(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) \;&=\; \frac{\mathbf{a}'\mathbf{d}\mathbf{b}'\mathbf{c}}{n} \sum_{k=1}^{p} \frac{(n-k+1)^2}{n^2} \left(\frac{n-p+1}{n-k+2} \right)^4 + o_p(1) \\ &\to\; \mathbf{a}'\mathbf{d}\mathbf{b}'\mathbf{c} \int_0^c \frac{(1-c)^4}{(1-t)^2} dt = c(1-c)^3 \mathbf{a}'\mathbf{d}\mathbf{b}'\mathbf{c}. \end{aligned} \tag{4.17}$$

So the limiting variance of a diagonal element of $\mathbf{V}$ is $2I(\mathbf{a}, \mathbf{a}, \mathbf{a}, \mathbf{a})/(1-c)^2 = 2c(1-c)$; and that of an off-diagonal element is $\{I(\mathbf{a}, \mathbf{b}, \mathbf{a}, \mathbf{b}) + I(\mathbf{a}, \mathbf{b}, \mathbf{b}, \mathbf{a})]\}/(1-c)^2 = c(1-c)$ as $\mathbf{a} \perp \mathbf{b}$. Similarly, the limiting covariance of any two diagonal elements corresponds to $2I(\mathbf{a}, \mathbf{a}, \mathbf{c}, \mathbf{c})/(1-c)^2 = 0$ as $\mathbf{a} \perp \mathbf{c}$; and that of a diagonal and an off-diagonal elements is $\{I(\mathbf{a}, \mathbf{a}, \mathbf{c}, \mathbf{d}) + I(\mathbf{a}, \mathbf{a}, \mathbf{d}, \mathbf{c})\}/(1-c)^2 = 0$ as $\mathbf{a}, \mathbf{c}$ and $\mathbf{d}$ are mutually orthogonal. Similarly, the limiting covariance of any two off-diagonal elements is also 0. This completes the proof of Theorem 2.1.

4.3 Lemmas 4.5–4.8 and proofs

We shall state and prove Lemmas 4.5-4.8 which were used in the proof of Theorem 2.1 in Section 4.2.

Lemma 4.5 *We have* $\mathrm{E}(W_k^2\|\mathbf{Q}_k\|^2) = o(n^2)$.

Proof. By Cauchy-Schwartz, we have $\mathrm{E}(W_k^2\|\mathbf{Q}_k\|^2) \leq \{\mathrm{E}(W_k^4)\ \mathrm{E}(\|\mathbf{Q}_k\|^4)\}^{1/2}$. Noting that $W_k = \mathbf{z}_k'\mathbf{P}_k\mathbf{z}_k - \mathrm{tr}(\mathbf{P}_k)$, and applying Lemma 4.3(b), we obtain $\mathrm{E}(W_k^4) = o(n^3)$. By Lemma 4.6 below, $\mathrm{E}(\|\mathbf{Q}_k\|^4) = o(n)$. It follows that $\mathrm{E}(W_k^2\|\mathbf{Q}_k\|^2) = o(n^2)$. ■

Lemma 4.6 *We have*

$$\sum_{i=1}^{n} \mathrm{E}\left(\|\mathbf{A}'\mathbf{P}_k\mathbf{e}_i\|^4\right) = O(n^{-1}), \tag{4.18}$$

and

$$\mathrm{E}^{(k)}(\|\mathbf{Q}_k\|^4) \leq K\left\{\|\mathbf{A}'\mathbf{P}_k\mathbf{A}\| + \sum_{i=1}^{n}\|\mathbf{A}'\mathbf{P}_k\mathbf{e}_i\|^8\ \mathrm{E}(y_{11}^8)\right\} = o(n). \tag{4.19}$$

Proof. Write the i-th row of $\mathbf{A}$ as $\mathbf{a}_i'$ and the remainder of $\mathbf{A}$ with $\mathbf{a}_i'$ removed by $\mathbf{A}_i$. Write the i-th row of $\mathbf{Z}_k$ as $\mathbf{y}_{ki}$ and remainder of $\mathbf{Z}_k$ with $\mathbf{y}_{ki}$ removed as $\mathbf{Z}_{ki}$. Then,

$$\begin{aligned}\|\mathbf{A}'\mathbf{P}_k\mathbf{e}_i\|^2 &= \|\mathbf{a}_i(1-\mathbf{y}_{ki}'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{y}_{ki}) + \mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{y}_{ki}\|^2 \\ &\leq 2\|\mathbf{a}_i(1-\mathbf{y}_{ki}'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{y}_{ki})\|^2 + 2\|\mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{y}_{ki}\|^2.\end{aligned}$$

By Lemma 4.1 and condition (2.3), we have

$$\|\mathbf{a}_i(1-\mathbf{y}_{ki}'(\mathbf{Z}_k\mathbf{Z}_k')^{-1}\mathbf{y}_{ki})\|^2 = \|\mathbf{a}_i\|^2\frac{1}{(1+\mathbf{y}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1}\mathbf{y}_{ki})^2} \leq K/n.$$

Note first that $\mathrm{ch}_{\min}(\mathbf{Z}_{ki}\mathbf{Z}_{ki}') \geq \mu n$, so

$$\begin{aligned}\mathrm{E}\left(\|\mathbf{A}'\mathbf{P}_k\mathbf{e}_i\|^4\right) &\leq K\left(n^{-2} + \mathrm{E}\frac{\|\mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1}\mathbf{y}_{ki}\|^4}{(1+\mathbf{y}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1}\mathbf{y}_{ki})^4}\right) \\ &\leq K\left(n^{-2} + \mathrm{E}\|\mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1}\|^4\mathrm{E}(y_{11}^4)\right) \\ &\leq K\left(n^{-2} + \mu^{-2}n^{-2}\mathrm{E}\|\mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1/2}\|^4\right) = O(n^{-2}).\end{aligned}$$

And hence (4.18) follows.

Similarly

$$
\begin{aligned}
\mathrm{E}\|\mathbf{A}'\mathbf{P}_k\mathbf{e}_i\|^8 &\le K\left(n^{-4}+\mathrm{E}\frac{\|\mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1}\mathbf{y}_{ki}\|^8}{(1+\mathbf{y}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1}\mathbf{y}_{ki})^8}\right)\\
&\le K\left(n^{-4}+\mathrm{E}\|\mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1}\|^8\mathrm{E}(y_{11}^8)\right)\\
&\le K\left(n^{-4}+\mu^{-4}n^{-2}\eta_n^4\mathrm{E}\|\mathbf{A}_i'\mathbf{Z}_{ki}'(\mathbf{Z}_{ki}\mathbf{Z}_{ki}')^{-1/2}\|^8\right)=o(n^{-2}),
\end{aligned}
$$

which implies the equality in (4.19). It remains to prove the first inequality in (4.19). We proceed as follows

$$
\begin{aligned}
\mathrm{E}^{(k)}(\|\mathbf{Q}_k\|^4) &\le K\left(\mathrm{E}\|\mathbf{A}'\mathbf{P}_k\mathbf{A}\|^4+\mathrm{E}\|\mathbf{A}'\mathbf{P}_k\mathbf{z}_k\mathbf{z}_k'\mathbf{P}_k\mathbf{A}\|^4\right)\\
&\le K\left\{1+\sum_{i=1}^n\|\mathbf{A}'\mathbf{P}_k\mathbf{e}_i\|^8\ \mathrm{E}(y_{11}^8)\right\},
\end{aligned}
$$

where we used the fact that $\|\mathbf{A}'\mathbf{P}_k\mathbf{A}\|\le 1$ and applied Rosenthal inequality (see Theorem 3 in Rosenthal (1970)) to $\mathbf{A}'\mathbf{P}_k\mathbf{z}_k$ in the last inequality. This proves the inequality in (4.19), and so completes the proof of the lemma. ■

Lemma 4.7 *We have $R_n=O_P(n^{-1})$ and $\Delta_{kj}=O_P(n^{-2})$ for $1\le k\le p, 1\le j\le 6$.*

Proof. Applying Lemma 4.6, we have

$$
\begin{aligned}
&\mathrm{E}\left|\mathrm{E}_{k-1}\left(\sum_{i=1}^n\mathbf{a}'\mathbf{P}_k\mathbf{e}_i\ \mathbf{b}'\mathbf{P}_k\mathbf{e}_i\ \mathbf{c}'\mathbf{P}_k\mathbf{e}_i\ \mathbf{d}'\mathbf{P}_k\mathbf{e}_i\right)\right|\\
\le\ & K\left(\sum_{i=1}^n\mathrm{E}|\mathbf{a}'\mathbf{P}_k\mathbf{e}_i|^4+\sum_{i=1}^n\mathrm{E}|\mathbf{b}'\mathbf{P}_k\mathbf{e}_i|^4+\sum_{i=1}^n\mathrm{E}|\mathbf{c}'\mathbf{P}_k\mathbf{e}_i|^4+\sum_{i=1}^n\mathrm{E}|\mathbf{d}'\mathbf{P}_k\mathbf{e}_i|^4\right)\\
=\ & O(n^{-1}).
\end{aligned}
$$

So $R_n=O_P(n^{-1})$.

Symmetry consideration shows that it suffices to prove that $\Delta_{kj}=O(n^{-2})$ for $j=1,2,4,$ and 6.

As $\mathrm{E}|\mathrm{E}_k(\mathbf{a}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{d}\ \mathbf{b}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{c})| \le \left\{\mathrm{E}(\mathbf{a}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{d})^2 \times \mathrm{E}(\mathbf{b}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{c})^2\right\}^{1/2}$, and that $\mathbf{a}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{d} = \mathbf{z}_p'(\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{a}'\mathbf{P}_{kp})\mathbf{z}_p - \mathrm{tr}(\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{a}'\mathbf{P}_{kp})$. Applying Lemma 4.3(b),

$$\mathrm{E}(\mathbf{a}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{d})^2 \le K\mathrm{E}\left\{\mathrm{tr}\left((\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{d}\mathbf{a}'\mathbf{P}_{kp})^2\right)\right\} = K\mathrm{E}\left\{(\mathbf{a}'\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{d})^2\right\} \le K.$$

Similarly, $\mathrm{E}(\mathbf{b}'\mathbf{R}_{k1}\breve{\mathbf{P}}_k\mathbf{c})^2$ is bounded. It follows that $\Delta_{k1} = O_P(n^{-2})$.

To show that $\Delta_{k6} = O_P(n^{-2})$, by Cauchy-Schwarz inequality, it is sufficient to show that $\mathrm{E}|\mathbf{a}'\mathbf{R}_{k2}\breve{\mathbf{P}}_k\mathbf{d}|^2$ is bounded. From the expression of $\mathbf{R}_{k2}$ and Cauchy-Schwarz inequality, it suffices to show that $\mathrm{E}(|\mathbf{a}'\mathbf{P}_{kp}\mathbf{z}_p|^8), \mathrm{E}(|\mathbf{z}_p'\mathbf{P}_{kp}\breve{\mathbf{P}}_k d|^8)$ and $\mathrm{E}(|B_{kp}W_{kp}|^4)$ are bounded. The first two are 8-moments of linear combinations of *i.i.d.* random variables. The last is bounded above by $Kn^{-4}\mathrm{E}(W_{kp}^4)$, which is bounded by Lemma 4.3.

The proof for $\Delta_{k4} = O(n^{-2})$, which is similar but simpler, will be omitted.

First note that the factor $\mathbf{b}'\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{c}$ is independent of $\mathbf{z}_p$ and bounded, we just treat it as constant. Next we note that $\mathbf{z}_p$ is contained only in W_{kp} as a centralized quadratic form of $\mathbf{z}_p$, and two linear combinations $\mathbf{z}_p'\mathbf{P}_{kp}\breve{\mathbf{P}}_k\mathbf{d}$ and $\mathbf{a}'\mathbf{P}_{kp}\mathbf{z}_p$. Integrating $\mathbf{z}_p$ first leads to consideration of the following general form

$$\begin{aligned}\mathrm{E}\left\{(\mathbf{z}_p'\mathbf{Q}\mathbf{z}_p - \mathrm{tr}\mathbf{Q})\mathbf{a}'\mathbf{z}_p\mathbf{b}'\mathbf{z}_p\right\} &= \mathrm{E}\left\{(\mathbf{z}_p'\mathbf{Q}\mathbf{z}_p - \mathrm{tr}\mathbf{Q})(\mathbf{z}_p'\mathbf{B}\mathbf{z}_p - \mathrm{tr}\mathbf{B})\right\} \\ &= \mathrm{tr}(\mathbf{Q}\mathbf{B}) + \mathrm{tr}(\mathbf{Q}\mathbf{B}') + \left(\mathrm{E}(y_{11}^4) - 3\right)\sum_{i=1}^n q_{ii}b_{ii}\end{aligned}$$

where we applied Lemma 4.3(a) with $\mathbf{B} = \mathbf{a}\mathbf{b}'$. Here q_{ii} and b_{ii} denote the (i,i)-th entries of $\mathbf{Q}$ and $\mathbf{B}$ respectively. Note that $\mathrm{tr}(\mathbf{Q}\mathbf{B}) = \mathbf{b}'\mathbf{Q}\mathbf{a}$ and $\mathrm{tr}(\mathbf{Q}\mathbf{B}') = \mathbf{a}'\mathbf{Q}\mathbf{b}$ are bounded because $\|\mathbf{Q}\|$ is bounded, and that q_{ii} is bounded. Also, since $b_{ii} = a_ib_i$, so

$$|\sum q_{ii}b_{ii}| \le K\sum|a_ib_i| \le K\|a\|\|b\| \le K.$$

Finally, we can extract a factor of n^{-1} from the factor B_{kp} in Δ_{k2}, and thus $\Delta_{k2} = O(n^{-2})$. ■

Lemma 4.8 *Recall* $\mathbf{a}$ *and* $\mathbf{d}$ *are any two columns of* $\mathbf{A}$, *either identical or orthogonal. Recall also that* $\mathbf{P}_{kk} = \mathbf{I}_n - \mathbf{Z}'_{kk}(\mathbf{Z}_{kk}\mathbf{Z}'_{kk})^{-1}\mathbf{Z}_{kk}$ *where* $\mathbf{Z}_{kk}$ *is the* $(k-1)\times n$ *matrix consisting of the rows* $\mathbf{z}'_1, \ldots, \mathbf{z}'_{k-1}$. *Then,*

$$\mathbf{a}'\mathbf{P}_{kk}\mathbf{d} = \frac{n-k+1}{n}\ \mathbf{a}'\mathbf{d} + o_P(1).$$

Proof. Let $\mathbf{U}$ denote the $(k-1)\times n$ matrix obtained from $\mathbf{Y}$ by keeping the top $k-1$ rows. Represent $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ where $\mathbf{u}_r = (y_{1r}, \ldots, y_{k-1,r})'$ for $1 \le r \le n$. Recall $c_{n,k-1} = (k-1)/n$.

$$\begin{aligned}
\mathbf{a}'\mathbf{P}_{kk}\mathbf{d} &= \mathbf{a}'\mathbf{d} - \mathbf{a}'\mathbf{Z}'_{kk}(\mathbf{Z}_{kk}\mathbf{Z}'_{kk})^{-1}\mathbf{Z}_{kk}\mathbf{d} \\
&= (1-c_{n,k-1})\mathbf{a}'\mathbf{d} - \sum_{i=1}^{n} a_i d_i[\mathbf{u}'_i(\mathbf{U}\mathbf{U}')^{-1}\mathbf{u}_i - c_{n,k-1}] \\
&\quad - \sum_{1\le i\ne j\le n} a_i d_j \mathbf{u}'_i(\mathbf{U}\mathbf{U}')^{-1}\mathbf{u}_j \\
&= \frac{n-k+1}{n}\mathbf{a}'\mathbf{d} + o_P(1),
\end{aligned}$$

where we applied Lemma 5.1. ■

5 Proof of Theorem 2.2

We outline the proof of Theorem 2.2 here leaving the technical details to Section 5.1. By Lemma 4.1,

$$\begin{aligned}
&\sqrt{n}\left\{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}' - n\bar{\mathbf{y}}\bar{\mathbf{y}}')^{-1}\mathbf{Y}\mathbf{A} - \mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A}\right\} \\
&= \sqrt{n}\mathbf{A}'\mathbf{Y}'\left\{(\mathbf{Y}\mathbf{Y}' - n\bar{\mathbf{y}}\bar{\mathbf{y}}')^{-1} - (\mathbf{Y}\mathbf{Y}')^{-1}\right\}\mathbf{Y}\mathbf{A} \\
&= n^{3/2}\frac{\mathbf{A}'\mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}\bar{\mathbf{y}}\bar{\mathbf{y}}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A}}{1 - n\bar{\mathbf{y}}'(\mathbf{Y}\mathbf{Y}')^{-1}\bar{\mathbf{y}}}.
\end{aligned}$$

Theorem 2.2 follows immediately from (5.1) and (5.2) below

$$n\bar{\mathbf{y}}'(\mathbf{Y}\mathbf{Y}')^{-1}\bar{\mathbf{y}} = \bar{\mathbf{y}}'\mathbf{S}^{-1}\bar{\mathbf{y}} \xrightarrow{P} c < 1, \tag{5.1}$$

$$n^{3/4}\bar{\mathbf{y}}'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{A} = n^{-1/4}\bar{\mathbf{y}}'\mathbf{S}^{-1}\mathbf{Y}\mathbf{A} \xrightarrow{P} \mathbf{0}. \tag{5.2}$$

Since

$$\bar{\mathbf{y}}'\mathbf{S}^{-1}\bar{\mathbf{y}} = \frac{1}{n^2}\sum_{k=1}^{n}\mathbf{y}_k'\mathbf{S}^{-1}\mathbf{y}_k + \frac{2}{n^2}\sum_{1\le j<k\le n}\mathbf{y}_j'\mathbf{S}^{-1}\mathbf{y}_k,$$

we apply Lemma 4.1 and Lemma 5.1 below to obtain

$$\frac{1}{n^2}\sum_{k=1}^{n}\mathbf{y}_k'\mathbf{S}^{-1}\mathbf{y}_k = \frac{1}{n^2}\sum_{k=1}^{n}\frac{\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k}{1+\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k} \xrightarrow{a.s.} \frac{c/(1-c)}{1+c/(1-c)} = c.$$

This proves (5.1).

Since q is fixed, and each of the q columns of $\mathbf{A}$ is orthogonal to $\mathbf{1}_n$, therefore (5.2) follows from Lemma 5.2 below.

Lemma 5.1 *Let* $\mathbf{a}$ *and* $\mathbf{d}$ *denote any two columns of* $\mathbf{A}$*, either identical or orthogonal. Let* a_k *and* d_k *denote their* k*-th entries respectively. We have*

$$\sum_{k=1}^{n} a_k d_k\left(\frac{1}{n}\mathbf{y}_k'\mathbf{S}^{-1}\mathbf{y}_k - c_{n,p}\right) = o_P(1), \tag{5.3}$$

$$\sum_{1\le j<k\le n} a_j d_k\left(\frac{1}{n}\mathbf{y}_j'\mathbf{S}^{-1}\mathbf{y}_k\right) = O_P(n^{-1/2}). \tag{5.4}$$

Lemma 5.2 *Let* $\mathbf{a}$ *be a unit* n*-vector satisfying* $\mathbf{1}_n'\mathbf{a} = 0$*. Then*

$$n^{-1/4}\bar{\mathbf{y}}'\mathbf{S}^{-1}\mathbf{Y}\mathbf{a} \xrightarrow{P} 0.$$

5.1 Proofs of Lemmas 5.1 and 5.2

As outlined above, the proof of Theorem 2.2 will be complete if we can prove Lemmas 5.1 and 5.2. Recall that $\mathbf{y}_k$ denotes the k-th column of $\mathbf{Y}$, and $\mathbf{Y}_k$ the $p\times(n-1)$ matrix after removing the k-th column from $\mathbf{Y}$. Similarly, $\mathbf{Y}_{jk}$ is the $p\times(n-2)$ matrix after removing the j-th and the k-th columns of $\mathbf{Y}$. Denote $\mathbf{S} = \frac{1}{n}\mathbf{Y}\mathbf{Y}'$, $\mathbf{S}_k = \frac{1}{n}\mathbf{Y}_k\mathbf{Y}_k'$ and $\mathbf{S}_{jk} = \frac{1}{n}\mathbf{Y}_{jk}\mathbf{Y}_{jk}'$.

Proof of Lemma 5.1 Applying Lemma 4.2 for $r=1$, we have

$$\frac{1}{n}\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k - c_{n,p} = \left(\frac{1}{n}\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k - \frac{1}{n}\mathrm{tr}\mathbf{S}_k^{-1}\right) + \left(\frac{1}{n}\mathrm{tr}\mathbf{S}_k^{-1} - c_{n,p}\right) \to 0 \ a.s.$$

uniformly in $1 \le k \le n$. Hence (5.3) follows.

For $1 \le \ell \le n$, let E_ℓ denote the conditional expectation given $\{\mathbf{y}_1, \ldots, \mathbf{y}_\ell\}$. Denote the unconditional expectation by E_0. Consider

$$\begin{aligned}
&\frac{1}{n} \sum_{1 \le j < k \le n} a_j d_k \mathbf{y}_j' \mathbf{S}^{-1} \mathbf{y}_k \\
=\ & \frac{1}{n} \sum_{\ell=1}^{n} (\mathrm{E}_\ell - \mathrm{E}_{\ell-1}) \left[\sum_{1 \le j < k \le n} a_j d_k \mathbf{y}_j' \mathbf{S}^{-1} \mathbf{y}_k \right] + \mathrm{E} \left(\frac{1}{n} \sum_{1 \le j < k \le n} a_j d_k \mathbf{y}_j' \mathbf{S}^{-1} \mathbf{y}_k \right) \\
=\ & I_1 + I_2 + I_3 + \mathrm{E} \left(\frac{1}{n} \sum_{1 \le j < k \le n} a_j d_k \mathbf{y}_j' \mathbf{S}^{-1} \mathbf{y}_k \right),
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{n} \sum_{1 \le \ell < k \le n} a_\ell d_k (\mathrm{E}_\ell - \mathrm{E}_{\ell-1}) \mathbf{y}_\ell' \mathbf{S}^{-1} \mathbf{y}_k, \\
I_2 &= \frac{1}{n} \sum_{1 \le j < \ell \le n} a_j d_\ell (\mathrm{E}_l - \mathrm{E}_{\ell-1}) \mathbf{y}_j' \mathbf{S}^{-1} \mathbf{y}_\ell, \\
I_3 &= \frac{1}{n} \sum_{\ell=1}^{n} \sum_{\substack{1 \le j < k \le n \\ \ell \ne j,k}} a_j d_k (\mathrm{E}_\ell - \mathrm{E}_{\ell-1}) \mathbf{y}_j' \mathbf{S}^{-1} \mathbf{y}_k.
\end{aligned}$$

Lemma 5.1 will be proved if we can show that

$$\mathrm{E}(I_1^2) = O\left(n^{-1}\right), \tag{5.5}$$

$$\mathrm{E}(I_2^2) = O\left(n^{-1}\right), \tag{5.6}$$

$$\mathrm{E}(I_3^2) = O\left(n^{-1}\right), \tag{5.7}$$

$$\mathrm{E} \left(\frac{1}{n} \sum_{1 \le j < k \le n} a_j d_k \mathbf{y}_j' \mathbf{S}^{-1} \mathbf{y}_k \right) = O\left(n^{-1}\right). \tag{5.8}$$

Using $\mathbf{S}^{-1} = \mathbf{S}_\ell^{-1} - \frac{1}{n} C_\ell \mathbf{S}_\ell^{-1} \mathbf{y}_\ell \mathbf{y}_\ell' \mathbf{S}_\ell^{-1}$, we split $I_1 = I_{11} + I_{12}$, where

$$\begin{aligned}
\mathrm{E}(I_{11}^2) &= \mathrm{E} \left\{ \left| \frac{1}{n} \sum_{1 \le \ell < k \le n} a_\ell d_k (\mathrm{E}_\ell - \mathrm{E}_{\ell-1}) \mathbf{y}_\ell' \mathbf{S}_\ell^{-1} \mathbf{y}_k \right|^2 \right\} \\
&= \frac{1}{n^2} \sum_{\ell=1}^{n-1} a_\ell^2 \mathrm{E} \left\{ \left\| \mathbf{S}_\ell^{-1} \sum_{k=\ell+1}^{n} d_k \mathbf{y}_k \right\|^2 \right\}
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{\mu^2 n^2}\sum_{\ell=1}^{n-1} a_\ell^2 \mathrm{E}\left\{\left\|\sum_{k=\ell+1}^{n} d_k \mathbf{y}_k\right\|^2\right\} \\
&= \frac{1}{\mu^2 n^2}\sum_{\ell=1}^{n-1}\sum_{k=\ell+1}^{n} a_\ell^2 d_k^2 \mathrm{E}(\mathbf{y}_k'\mathbf{y}_k) \\
&\leq \frac{1}{\mu^2 n^2}\sum_{\ell=1}^{n-1}\sum_{k=\ell+1}^{n} a_\ell^2 d_k^2 p = O(n^{-1}),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}(I_{12}^2) &= \mathrm{E}\left|\frac{1}{n^2}\sum_{1\leq \ell<k\leq n} a_\ell d_k(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\left\{C_\ell \mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\mathbf{y}_\ell\mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\mathbf{y}_k\right\}\right|^2 \\
&= \frac{1}{n^4}\sum_{\ell=1}^{n-1} a_\ell^2 \mathrm{E}\left|\sum_{k=\ell+1}^{n}(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\left\{C_\ell \mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\mathbf{y}_\ell\mathbf{y}_\ell'\mathbf{S}_\ell^{-1}(d_k\mathbf{y}_k)\right\}\right|^2 \\
&\leq \frac{1}{n^4}\sum_{\ell=1}^{n-1} a_\ell^2\left(\mathrm{E}|\mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\mathbf{y}_\ell|^4 \mathrm{E}\left\{\left\|\mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\sum_{k=\ell+1}^{n} d_k\mathbf{y}_k\right\|^4\right\}\right)^{1/2} \\
&\leq \frac{1}{n^4}\sum_{\ell=1}^{n-1} a_\ell^2\left(O(n^4)\cdot O\left(\mathrm{E}\left\{\left\|\sum_{k=\ell+1}^{n} d_k\mathbf{y}_k\right\|^4\right\}\right)\right)^{1/2} = O(n^{-1}),
\end{aligned}
$$

Here, we used the facts that

$$
\begin{aligned}
&\mathrm{E}\left(|\mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\mathbf{y}_\ell|^4\right) \leq \mu^{-4}\mathrm{E}\left(\|\mathbf{y}_\ell\|^8\right) \\
\leq\ & K\Big(\sum_{i=1}^{p}\mathrm{E}(y_{i\ell}^8) + \sum_{i=1}^{p}\mathrm{E}(y_{i\ell}^6)\sum_{i=1}^{p}\mathrm{E}(y_{i\ell}^2) + \Big(\sum_{i=1}^{p}\mathrm{E}(y_{i\ell}^4)\Big)^2 + \sum_{i=1}^{p}\mathrm{E}(y_{i\ell}^4)\Big(\sum_{i=1}^{p}\mathrm{E}(y_{i\ell}^2)\Big)^2 \\
&+\Big(\sum_{i=1}^{p}\mathrm{E}(y_{i\ell}^2)\Big)^4\Big) = O(n^4),
\end{aligned}
$$

and for any nonrandom vector $\boldsymbol{\alpha}$,

$$
\mathrm{E}|\boldsymbol{\alpha}'\mathbf{y}_\ell|^4 \leq K(\boldsymbol{\alpha}'\boldsymbol{\alpha})^2 = K\|\boldsymbol{\alpha}\|^4.
$$

This proves (5.5). In a similar way, we can prove (5.6) .

To prove (5.7), first note that $(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\left\{\mathbf{y}_j'\mathbf{S}_\ell^{-1}\mathbf{y}_k\right\} = 0$ for $\ell \neq j, k$. Introducing $\mathbf{T}_\ell = \sum_{\substack{1\le j<k\le n \\ j,k\neq \ell}} a_j d_k \mathbf{y}_k \mathbf{y}_j'$, we have

$$
\begin{aligned}
\mathrm{E}(I_3^2) &= \frac{1}{n^2}\sum_{\ell=1}^n \mathrm{E}\left\{\left|\sum_{\substack{1\le j<k\le n \\ j,k\neq \ell}} (\mathrm{E}_\ell - \mathrm{E}_{\ell-1}) a_j d_k \mathbf{y}_j'(\mathbf{S}^{-1} - \mathbf{S}_\ell^{-1})\mathbf{y}_k\right|^2\right\} \\
&= \frac{1}{n^4}\sum_{\ell=1}^n \mathrm{E}\left|\sum_{\substack{1\le j<k\le n \\ j,k\neq \ell}} (\mathrm{E}_\ell - \mathrm{E}_{\ell-1}) C_\ell a_j d_k \mathbf{y}_j'(\mathbf{S}_\ell^{-1}\mathbf{y}_\ell\mathbf{y}_\ell'\mathbf{S}_\ell^{-1})\mathbf{y}_k\right|^2 \\
&= \frac{1}{n^4}\sum_{\ell=1}^n \mathrm{E}\left|(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\left\{C_\ell \mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\mathbf{T}_\ell\mathbf{S}_\ell^{-1}\mathbf{y}_\ell\right\}\right|^2 \\
&\le \frac{2}{n^4}\sum_{\ell=1}^n \Big[\mathrm{E}\left|(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\left\{C_\ell\left(\mathbf{y}_\ell'\mathbf{S}_\ell^{-1}\mathbf{T}_\ell\mathbf{S}_\ell^{-1}\mathbf{y}_\ell - \mathrm{tr}\left(\mathbf{S}_\ell^{-1}\mathbf{T}_\ell\mathbf{S}_\ell^{-1}\right)\right)\right\}\right|^2 \\
&\quad + \mathrm{E}\left|(\mathrm{E}_\ell - \mathrm{E}_{\ell-1})\left\{(C_\ell - (1-c_{n,p}))\mathrm{tr}\left(\mathbf{S}_\ell^{-1}\mathbf{T}_\ell\mathbf{S}_\ell^{-1}\right)\right\}\right|^2\Big] \\
&\le \frac{2}{n^4}\sum_{l=1}^n \Big[\mathrm{E}\left\{\left|\mathbf{y}_l'\mathbf{S}_l^{-1}\mathbf{T}_\ell\mathbf{S}_l^{-1}\mathbf{y}_l - \mathrm{tr}\left(\mathbf{S}_l^{-1}\mathbf{T}_\ell\mathbf{S}_l^{-1}\right)\right|^2\right\} \\
&\quad + \mathrm{E}\left\{(C_l - (1-c_{n,p}))^2\mathrm{tr}^2\left(\mathbf{S}_\ell^{-1}\mathbf{T}_\ell\mathbf{S}_\ell^{-1}\right)\right\}\Big] \\
&\le \frac{2}{n^4}\sum_{\ell=1}^n \left[\mathrm{E}\left\{\mathrm{tr}\left(\mathbf{S}_\ell^{-1}\mathbf{T}_\ell\mathbf{S}_\ell^{-1}\right)^2\right\} + O(n^{-1})\mathrm{E}\left\{\mathrm{tr}^2\left(\mathbf{S}_\ell^{-1}\mathbf{T}_\ell\mathbf{S}_\ell^{-1}\right)\right\}\right] \\
&\le \frac{4}{\mu^4 n^4}\sum_{\ell=1}^n \mathrm{E}(\mathrm{tr}\mathbf{T}_\ell^2) \le \frac{Knp^2}{\mu^4 n^4} = O(n^{-1}).
\end{aligned}
$$

We have used the fact that

$$
(\mathrm{tr}(\mathbf{A}))^2 = \left(\sum_{i=1}^p A_{ii}\right)^2 \le n\sum_{i=1}^p A_{ii}^2 \le n\ \mathrm{tr}(\mathbf{A}\mathbf{A}')
$$

which implies that

$$
\begin{aligned}
\left(\mathrm{tr}\left\{\mathbf{S}_l^{-1}\mathbf{T}_\ell\mathbf{S}_l^{-1}\right\}\right)^2 &\le n\ \mathrm{tr}\left\{\mathbf{S}_l^{-1}\mathbf{T}_\ell\mathbf{S}_l^{-2}\mathbf{T}_\ell\mathbf{S}_l^{-1}\right\} \\
&\le \frac{n}{\mu^4}\ \mathrm{tr}\left\{\left(\sum_{\substack{j<k \\ j,k\neq l}} a_j d_k \mathbf{y}_j\mathbf{y}_k'\right)\left(\sum_{\substack{j<k \\ j,k\neq l}} a_j d_k \mathbf{y}_k\mathbf{y}_j'\right)\right\}.
\end{aligned}
$$

This proves (5.7).

To prove (5.8) and hence complete the proof Lemma 5.1, we apply condition (2.3) and Lemma 4.4 as follows

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{n}\sum_{1\le j<k\le n} a_j d_k \mathbf{y}_j'\mathbf{S}^{-1}\mathbf{y}_k\right) &= \sum_{1\le j<k\le n} a_j d_k \mathrm{E}\left(\frac{1}{n}\mathbf{y}_1'\mathbf{S}^{-1}\mathbf{y}_2\right) \\
&= \sum_{1\le j<k\le n} a_j d_k \mathrm{E}\left\{\mathbf{y}_1'(\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{y}_2\right\} \\
&= O(n^{-1}).
\end{aligned}$$

■

Proof of Lemma 5.2. Denote the k-th element of $\mathbf{a}$ by a_k, and $\mathbf{a}_k$ the vector with a_k removed. Recall $C_k = (1 + \frac{1}{n}\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k)^{-1}$. Note that $\mathrm{E}(\mathrm{tr}\mathbf{S}_k^{-1})$ is independent of k, hence we write

$$\tilde{b} = \left\{1 + \frac{1}{n}\mathrm{E}(\mathrm{tr}\mathbf{S}_k^{-1})\right\}^{-1}.$$

Let

$$\widetilde{V}_k = \frac{1}{n}\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k - \frac{1}{n}\mathrm{E}(\mathrm{tr}\mathbf{S}_k^{-1}).$$

Recall also that we implicitly write $\mathcal{B}_k$ under the expectation sign in the definition of $\tilde{b}$ and $\widetilde{V}_k$. With these notations, and the identities

$$C_k = \tilde{b} - \tilde{b}C_k\widetilde{V}_k, \quad C_k\left(\frac{1}{n}\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k\right) = 1 - C_k, \quad \text{and} \quad \sum_{k=1}^{n} a_k = 0,$$

we have

$$\begin{aligned}
n^{-1/4}\bar{\mathbf{y}}'\mathbf{S}^{-1}\mathbf{Y}\mathbf{a} &= n^{-5/4}\sum_{k=1}^{n} C_k\left(\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{Y}_k\mathbf{a}_k + a_k\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k\right) \\
&= n^{-5/4}\sum_{k=1}^{n}(\tilde{b} - \tilde{b}C_k\widetilde{V}_k)\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{Y}_k\mathbf{a}_k - n^{-1/4}\sum_{k=1}^{n} a_k C_k \\
&= n^{-5/4}\tilde{b}\sum_{k=1}^{n}\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{Y}_k\mathbf{a}_k - n^{-5/4}\tilde{b}\sum_{k=1}^{n} C_k\widetilde{V}_k\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{Y}_k\mathbf{a}_k \\
&\quad + n^{-1/4}\tilde{b}\sum_{k=1}^{n} a_k C_k\widetilde{V}_k. \qquad (5.9)
\end{aligned}$$

The last sum in (5.9) converges to 0 in probability as $n \to \infty$ since

$$n^{-1/4} \sum_{k=1}^{n} \mathrm{E}\left\{|\tilde{b} a_k C_k \widetilde{V}_k|\right\} \le n^{-1/4} \left\{\sum_{k=1}^{n} \mathrm{E}(\widetilde{V}_k^2)\right\}^{1/2} \le K n^{-1/4}$$

where we applied Lemma 5.3 in the last inequality.

To show the second sum in (5.9) converges to 0 in probability, we consider

$$\begin{aligned} & n^{-5/4} \sum_{k=1}^{n} \mathrm{E}\left\{|\tilde{b} C_k \widetilde{V}_k \mathbf{y}_k' \mathbf{S}_k^{-1} \mathbf{Y}_k \mathbf{a}_k|\right\} \\ \le\ & n^{-5/4} \sum_{k=1}^{n} \left[\mathrm{E}(\widetilde{V}_k^2)\ \mathrm{E}\left\{(\mathbf{y}_k' \mathbf{S}_k^{-1} \mathbf{Y}_k \mathbf{a}_k)^2\right\}\right]^{1/2} \\ \le\ & \frac{K}{\sqrt{n}} \times \sqrt{n} \left[n^{-5/2} \sum_{k=1}^{n} \mathrm{E}\left\{(\mathbf{y}_k' \mathbf{S}_k^{-1} \mathbf{Y}_k \mathbf{a}_k)^2\right\}\right]^{1/2} \\ =\ & O(n^{-1/4}). \end{aligned}$$

We applied Lemma 5.3 and Cauchy-Schwartz inequality in the second inequality, and (5.11) below in the last step.

Next we show that the first sum in (5.9) tends to 0 in probability. That is,

$$n^{-5/4} \sum_{k=1}^{n} \mathbf{y}_k' \mathbf{S}_k^{-1} \mathbf{Y}_k \mathbf{a}_k \xrightarrow{P} 0. \tag{5.10}$$

It suffices to show that its second moment tends to 0. The expectation of the sum of the square terms converges to 0 as follows:

$$\begin{aligned} n^{-5/2} \sum_{k=1}^{n} \mathrm{E}\left\{\left(\mathbf{y}_k' \mathbf{S}_k^{-1} \mathbf{Y}_k \mathbf{a}_k\right)^2\right\} &= n^{-5/2} \sum_{k=1}^{n} \mathbf{a}_k' \mathrm{E}\left(\mathbf{Y}_k' \mathbf{S}_k^{-1} \mathbf{y}_k \mathbf{y}_k' \mathbf{S}_k^{-1} \mathbf{Y}_k\right) \mathbf{a}_k \\ &= n^{-5/2} \sum_{k=1}^{n} \mathbf{a}_k' \mathrm{E}\left(\mathbf{Y}_n' \mathbf{S}_n^{-2} \mathbf{Y}_n\right) \mathbf{a}_k \\ &= n^{-5/2} \sum_{k=1}^{n} \mathrm{E}\left\{(\mathbf{Y}_n \mathbf{a}_k)' \mathbf{S}_n^{-2} (\mathbf{Y}_n \mathbf{a}_k)\right\} \\ &\le n^{-5/2} \sum_{k=1}^{n} \mathrm{E}\left\{\|\mathbf{Y}_n \mathbf{a}_k\|^2 / \mathrm{ch}_{\min}(\mathbf{S}_n)^2\right\} \\ &\le \frac{2}{(1-\sqrt{c})^2 n^{5/2}} \sum_{k=1}^{n} \mathrm{E}\left(\|\mathbf{Y}_n \mathbf{a}_k\|^2\right) \\ &= O(n^{-1/2}). \end{aligned} \tag{5.11}$$

In the second step above, we used the fact that $\mathrm{E}(\mathbf{Y}_k'\mathbf{S}_k^{-2}\mathbf{Y}_k) = \mathrm{E}(\mathbf{Y}_n'\mathbf{S}_n^{-2}\mathbf{Y}_n)$; and in the last step, $\mathrm{E}(\|\mathbf{Y}_n\mathbf{a}_k\|^2) \le (n-1)\|\mathbf{a}_k\|^2 \le (n-1)\|\mathbf{a}\|^2 \le n$.

Next we show that the expectation of the sum of the cross-product terms:

$$n^{-5/2} \sum_{1\le j\ne k\le n} \mathrm{E}\left\{\left(\mathbf{y}_j'\mathbf{S}_j^{-1}\mathbf{Y}_j\mathbf{a}_j\right)\left(\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{Y}_k\mathbf{a}_k\right)\right\} \to 0. \tag{5.12}$$

To this end, we use the expansion in Lemma 4.1

$$\mathbf{S}_j^{-1} = \mathbf{S}_{jk}^{-1} - \frac{\frac{1}{n}\mathbf{S}_{jk}^{-1}\mathbf{y}_k\mathbf{y}_k'\mathbf{S}_{jk}^{-1}}{1+\frac{1}{n}\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_k}.$$

Write

$$\begin{aligned}
\widetilde{V}_{jk} &= \frac{1}{n}\left\{\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_k - \mathrm{E}(\mathrm{tr}\mathbf{S}_{jk}^{-1})\right\},\\
C_{jk} &= \frac{1}{1+\frac{1}{n}\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_k} \text{ (note } C_{jk}\ne C_{kj}),\\
b_1 &= \frac{1}{1+\frac{1}{n}\mathrm{E}(\mathrm{tr}\mathbf{S}_{jk}^{-1})},
\end{aligned}$$

where C_{jk} has been defined earlier in (4.9). Note also $C_{jk} = b_1 - b_1C_{jk}\widetilde{V}_{jk}$.

Let $\mathbf{a}_{jk}$ denote the $(n-2)$-vector after removing the j-th and the k-th components of $\mathbf{a}$. With these notations, we have

$$\begin{aligned}
\mathbf{y}_j'\mathbf{S}_j^{-1}\mathbf{Y}_j\mathbf{a}_j &= \mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_j\mathbf{a}_j - \frac{1}{n}C_{jk}\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{Y}_j\mathbf{a}_j\\
&= \mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_{jk}\mathbf{a}_{jk} + a_k\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k - \frac{1}{n}C_{jk}\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{Y}_j\mathbf{a}_j.
\end{aligned}$$

Similarly,

$$\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{Y}_k\mathbf{a}_k = \mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{Y}_{jk}\mathbf{a}_{jk} + a_j\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j - \frac{1}{n}C_{kj}\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_k\mathbf{a}_k.$$

The first term in the expansion of $\mathbf{y}_j'\mathbf{S}_j^{-1}\mathbf{Y}_j\mathbf{a}_j$ is independent of $\mathbf{y}_k$ and thus we have

$$n^{-5/2} \sum_{1\le j\ne k\le n} \mathrm{E}\left\{\left(\mathbf{y}_j'\mathbf{S}_j^{-1}\mathbf{Y}_j\mathbf{a}_j\right)\left(\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{Y}_k\mathbf{a}_k\right)\right\}$$

$$
\begin{aligned}
&= n^{-5/2} \sum_{1\le j\neq k\le n} \mathrm{E}\Big\{ \left(a_k \mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k - \frac{1}{n}C_{jk}\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{Y}_j\mathbf{a}_j\right) \\
&\quad \times \left(a_j \mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j - \frac{1}{n}C_{kj}\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_k\mathbf{a}_k\right) \Big\} \\
&\le 2n^{-5/2} \left\{ \sum_{1\le j\neq k\le n} \left[a_k^2 \mathrm{E}\left\{ \left(\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k\right)^2\right\} + \frac{1}{n^2}\mathrm{E}\left\{\left(\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{Y}_j\mathbf{a}_j\right)^2\right\}\right]\right\}^{1/2} \\
&\quad \times \left\{ \sum_{1\le j\neq k\le n} \left[a_j^2 \mathrm{E}\left\{ \left(\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j\right)^2\right\} + \frac{1}{n^2}\mathrm{E}\left\{\left(\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_k\mathbf{a}_k\right)^2\right\}\right]\right\}^{1/2} \\
&= 2n^{-5/2} \sum_{1\le j\neq k\le n} \left[a_j^2 \mathrm{E}\left\{ \left(\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j\right)^2\right\} + \frac{1}{n^2}\mathrm{E}\left\{\left(\mathbf{y}_k'\mathbf{S}_{jk}^{-1}\mathbf{y}_j\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_k\mathbf{a}_k\right)^2\right\}\right] \\
&\le 2n^{-5/2} \sum_{j\neq k} \left[a_j^2 \mathrm{E}\left(\mathrm{tr}\mathbf{S}_{jk}^{-2}\right) + \frac{1}{n^2}\mathrm{E}\left\{\mathbf{y}_j'\mathbf{S}_{jk}^{-2}\mathbf{y}_j(\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_j\mathbf{a}_j)^2\right\}\right] \\
&\le Kn^{-1/2}.
\end{aligned}
$$

In the last inequality, we applied $\mathrm{tr}(\mathbf{S}_{jk}^{-2}) \le \mu^{-2}p$ to bound the first term; and the second term is bounded as follows:

$$
\begin{aligned}
&\mathrm{E}\left\{\mathbf{y}_j'\mathbf{S}_{jk}^{-2}\mathbf{y}_j(\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_j\mathbf{a}_j)^2\right\} \\
&\le 2\mathrm{E}\left\{\mathbf{y}_j'\mathbf{S}_{jk}^{-2}\mathbf{y}_j(\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_{jk}\mathbf{a}_{jk})^2\right\} + 2a_k^2\mathrm{E}\left\{\mathbf{y}_j'\mathbf{S}_{jk}^{-2}\mathbf{y}_j(\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{y}_k)^2\right\} \\
&= 2\mathrm{E}\left\{\mathrm{tr}(\mathbf{S}_{jk}^{-2})(\mathbf{a}_{jk}'\mathbf{Y}_{jk}'\mathbf{S}_{jk}^{-2}\mathbf{Y}_{jk}\mathbf{a}_{jk})\right\} + 2\mathrm{E}\left\{\left(\mathbf{y}_j'\mathbf{S}_{jk}^{-2}\mathbf{y}_j - \mathrm{tr}(\mathbf{S}_{jk}^{-2})\right)(\mathbf{y}_j'\mathbf{S}_{jk}^{-1}\mathbf{Y}_{jk}\mathbf{a}_{jk})^2\right\} \\
&\quad +2a_k^2\mathrm{E}\left\{(\mathbf{y}_j'\mathbf{S}_{jk}^{-2}\mathbf{y}_j)^2\right\} \\
&\le K(n+n^2a_k^2).
\end{aligned}
$$

This proves (5.12) and hence Lemma 5.2. ■

Lemma 5.3 *Recall* $\tilde{V}_k = \frac{1}{n}\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k - \frac{1}{n}\mathrm{E}\left(\mathrm{tr}\mathbf{S}_k^{-1}\right)$. *We have the following bounds*

$$
\mathrm{E}(\tilde{V}_k^{2\ell}) \le \begin{cases} \frac{K}{n} & \text{if} \quad \ell = 1, \\ o(n^{-1}) & \text{if} \quad \ell > 1. \end{cases} \tag{5.13}
$$

Proof. Since there exists a positive constant K_ℓ such that $(a+b)^{2\ell} \le K_\ell(a^{2\ell}+b^{2\ell})$, we have

$$
\mathrm{E}(\tilde{V}_k^{2\ell}) \le \frac{K_\ell}{n^{2\ell}}\mathrm{E}\left\{|\mathbf{y}_k'\mathbf{S}_k^{-1}\mathbf{y}_k - \mathrm{tr}\mathbf{S}_k^{-1}|^{2\ell}\right\} + \frac{K_\ell}{n^{2\ell}}\mathrm{E}\left\{|\mathrm{tr}\mathbf{S}_k^{-1} - \mathrm{E}(\mathrm{tr}\mathbf{S}_k^{-1})|^{2\ell}\right\}.
$$

To bound the first term, we apply Lemma 9.1 of Bai and Silverstein (2010). To bound the second term, we apply $z = iv$ for $v = an^{-1/2}, v_y = 1 - \sqrt{c} + \sqrt{v}$ with $a > 1$ in Lemmas 8.20 and 8.21 in Bai and Silverstein (2010). ∎

Acknowledgements

The first author is partially supported by NSFChina 11571067 and PCSIRT; the second author by the Singapore Ministry of Education Academic Research Fund R-155-000-141-112; and the third author by the Ministry of Education, Science, Sports, and Culture, a Grant-in-Aid for Scientific Research (C), #25330038, 2013-2015.

References

[1] Anderson, T. W. (1951). The asymptotic distribution of certain characteristic roots and vectors. *Proc. 2nd Berkeley Symp. on Math. Statist. and Prob.*, 103–130.

[2] Anderson, T.W. (2003). *Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley, Hoboken, New Jersey.

[3] Amemiya, Y. (1990). A note on the limiting distribution of certain characteristic roots. *Statist. & Probab. Lett.*, **9**(5), 465–470.

[4] Bai, Z.D. (1985). A note on the limiting distribution of the eigenvalues of a kind of random matrices. *J. Math. Research and Exposition*, **5**(2), 113–118.

[5] Bai, Z.D., Jiang, D., Yao, J.F. and Zheng, S. (2009). Corrections to LRT on large-dimensioanl covariance matrix by RMT. *Ann. Statist.*, **37**, 3822–3840.

[6] Bai, Z.D., Liu, H.X. and Wong, W.K. (2011). Asymptotic properties of eigenmatrices of a large sample covariance matrix. *Ann. Appl. Probab.*, **21**, 1994–2015.

[7] Bai, Z.D. and Saranadasa, H. (1996). Effect of high dimension: by an example of a two sample problem. *Statistica Sinica*, **6**, 311–329

[8] Bai, Z.D. and Silverstein, J.W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.*, **32**, 535–605.

[9] Bai, Z.D. and Silverstein, J.W. (2009). *Spectral Analysis of Large Dimensional Random Matrices, 2nd ed.*. Springer.

[10] Fujikoshi, Y. (1977). Asymptotic expansions for the distributions of the latent roots in MANOVA and the canonical correlations. *J. Multivariate Anal.*, **7**, 386–396.

[11] Fujikoshi, Y. (2015). High-dimensional asymptotic distributions of characteristic roots in multivariate linear model MANOVA and canonical correlations. In preparation.

[12] Fujikoshi, Y., Himeno, T. and Wakaki, H. (2008). Asymptotic results in canonical discriminant analysis when the dimension is large compared to the sample. *J. Statist. Plann. Inference*, **138**, 3457–3466.

[13] Fujikoshi, Y. and Sakurai, T. (2009). High-dimensional asymptotic expansions for the distributions of canonical correlations. *J. Multivariate Anal.*, **100**, 231–241.

[14] Fujikoshi, Y., Ulyanov, V. and Shimizu, R. (2010). *Multivariate Analysis: High-Dimensional and Large-Sample Approximations*. Wiley, Hoboken, New Jersey.

[15] Glynn, W. J. (1980). Asymptotic representations of the densities of canonical correlations and latent roots in MANOVA when the population parameters have arbitrary multiplicities. *Ann. Statist.*, **8**, 468–477.

[16] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and its Applications*. Academic Press, New York.

[17] Hsu, P. L. (1941). On the limiting distribution of roots of a determinantal equation. *J. London Math. Soc.*, **16**, 183–194.

[18] Johnstone, I. M. (2008). Multivariate analysis and Jacobi ensembles: largest eigenvalue, Tracy-Widom limits and rate of convergence. *Ann. Statist.*, **36**, 2638–2716.

[19] Muirhead, R. J. (1978), Latent roots and matrix variates. *Ann. Statist.*, **6**, 5–33.

[20] Muirhead, R. J. (1982), *Aspect of Multivariate Statistical Theory*, Wiley, New York.

[21] Muirhead, R. J. and Watermaux, C.M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate nonnormal populations. *Biometrika*, **67**, 31–43.

[22] Rosenthal, H. P. (1970). On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.*, **8**, 273–303.

[23] Seo, T., Kanda, T. and Fujikoshi, Y. (1994). Asymptotic distributions of the sample roots in MANOVA models under nonnormality. *J. Japan Statist. Soc.*, **24**, 133-140.

[24] Sugiura, N. (1976). Asymptotic expansions of the distributions of the latent roots and latent vectors of the Wishart and multivariate F matrices. *J. Multivariate Anal.*, **6**, 500–525.

[25] Wakaki, H., Fujikoshi, Y. and Ulyanov, V. (2014). Asymptotic expansions of the distributions of MANOVA test statistics when the dimension is large. *Hiroshima Math. J.*, **44**, 247–259..

[26] Zheng, S. (2012). Central limit theorems for linear spectral statistics of large-dimensional F-matrices. *Ann. l'Institut Henri Poincareé-Probabilites et Statistiques*, **47**, 444–476.