On eigenvalues of sample covariance matrices based on high-dimensional compositional data*

Qianqian Jiang,[†] Jiaxin Qiu,[‡] Zeng Li[§]

(v1.0: 2023-Dec-22; v2.0: 2024-Dec-17.)

Abstract

This paper studies the asymptotic spectral properties of the sample covariance matrix for high-dimensional compositional data, including the limiting spectral distribution, the limit of extreme eigenvalues, and the central limit theorem for linear spectral statistics. All asymptotic results are derived under the high-dimensional regime where the data dimension increases to infinity proportionally with the sample size. The findings reveal that the limiting spectral distribution is the well-known Marčenko-Pastur law. The largest (or smallest non-zero) eigenvalue converges almost surely to the left (or right) endpoint of the limiting spectral distribution, respectively. Moreover, the linear spectral statistics demonstrate a Gaussian limit. Based on our CLT result, we investigate a test problem on the population covariance structure of the basis data. Simulation experiments demonstrate the accuracy of theoretical results.

Contents

1 Introduction

In recent years, there has been increasing interest in the analysis of high-dimensional compositional data (HCD), which arise in various fields including genomics, ecology, finance, and social sciences. Compositional data refers to observations whose sum is a constant, such as proportions or percentages. HCD often involve a large number of variables or features measured for each sample, posing unique challenges for analysis. In the field of genomics, HCD analysis plays a crucial role in studying the composition and abundance of microbial communities, such as the human gut microbiome. Understanding the microbial composition and its relationship with health and disease has significant implications for personalized medicine and therapeutic interventions.

Statistical inference in HCD involves microbial mean tests, covariance matrix structural tests, and linear regression hypothesis testing. These inferences are intricately linked to the statistical properties of the sample covariance matrix. Mean tests typically utilize sum-of-squares-type and maximum-type statistics for dense and sparse alternative hypotheses, respectively. **?** extended the maximum test framework by **?** for compositional data. However, there's a gap in having a suitable sum-of-squares-type statistic for dense alternatives in HCD mean tests. Many sum-of-squares-type

*arXiv: 2312.14420

[†]Southern University of Science and Technology, email: jqq172515@gmail.com

[‡]The University of Hong Kong, email: jxqiu@hku.hk

[§]Southern University of Science and Technology, email: liz9@sustech.edu.cn. Corresponding author.

statistics, like Hotelling's T^2-statistic, rely on the sample covariance matrix. For bacterial species correlation, **?** introduced the permutation-renormalization bootstrap (ReBoot), directly calculating correlations from compositional components. Shuffling is suggested due to compositional data's closure constraint, introducing negative correlations. Yet, compositional data's unique properties require an additional normalization step within the same sample post-shuffling, potentially impacting the theoretical validity of permutation and resampling methods. Additionally, resampling increases computational complexity for p-value calculation and confidence interval construction. To address these challenges, **?** developed a covariance matrix element hypothesis testing method, allowing control over false discovery proportion (FDP) and false discovery rate (FDR). All these studies are closely related to the sample covariance matrix of HCD.

Current research predominantly focuses on sparse compositional data. In dense scenarios, researchers often turn to the spectral properties of sample covariance matrices. Despite this, there is a notable gap in the field of random matrices where specific attention to structures resembling compositional data, where row sum of the data matrix is constant, is lacking. Statistical inference for HCD encounters challenges arising not only from constraints but also from high dimensionality. Recognizing the crucial role of spectral theory in sample covariance matrices is also vital for addressing statistical challenges associated with high-dimensional data. Importantly, while previous research on statistical inference for HCD has overlooked studies under the spectral theory of sample covariance matrices, our work takes on these challenges from a Random Matrix Theory perspective. Existing literature extensively covers spectral properties of large-dimensional sample covariance matrices, but most results rely on independent component data structure, i.e. $\mathbf{Z} = \mathbf{\Gamma}\mathbf{X}$, where $\mathbf{\Gamma}$ is determined, and $\mathbf{X}$ has independent and identically distributed (i.i.d.) components. Seminal works by **?** and **?** established the limiting spectral distribution (LSD) of the sample covariance matrix $n^{-1}\mathbf{X}\mathbf{X}^{'}$, where $\mathbf{X}$ is an i.i.d. data matrix with zero mean, leading to the well-known Marčenko-Pastur law. Subsequent research by **?** and **?** extended these findings to the sample covariance matrix $n^{-1}\mathbf{X}\mathbf{\Sigma}\mathbf{X}^{'}$ for data with a linear dependence structure. **?** extended to the general separable product form $n^{-1}\mathbf{A}^{1/2}\mathbf{X}\mathbf{B}\mathbf{X}^{'}\mathbf{A}^{1/2}$, where $\mathbf{A}$ is nonnegative definite, and $\mathbf{B}$ is Hermitian. Another important area of interest is the investigation of extreme eigenvalues. **?** explored the fluctuation of the extreme eigenvalues of the sample covariance matrix $n^{-1}\mathbf{X}\mathbf{X}^{'}$, proving that the standardized largest eigenvalue follows the Tracy-Widom law. Related extensions include sample covariance matrices with linear dependence structures (**?**), Kendall rank correlation coefficient matrices (**?**), among others. Considerable attention has also been given to the study of linear functionals of eigenvalues. **?** established the Central Limit Theorem (CLT) for the Linear Spectral Statistics (LSS) of the sample covariance matrix $n^{-1}\mathbf{A}^{1/2}\mathbf{X}\mathbf{X}^{'}\mathbf{A}^{1/2}$, later extended to sample correlation coefficient matrices (**?**), and separable product matrices (**?**). To summarize, existing results in spectral theory of large dimensional sample covariance matrix predominantly rely on independent component data structure which, unfortunately, HCD does not fit in.

Specifically, current second-order limit theorems do not apply to HCD, making the exploration of spectral theory for HCD with distinct constraints crucial. This paper delves into spectral theory for sample covariance matrices of HCD, including LSD, extreme eigenvalues, and CLT for LSS. Analyzing HCD faces challenges due to compositional data's specific dependence structure, making existing techniques for i.i.d. observations less applicable. However, we can assume that HCD are generated from unobservable basis data, while the underlying basis data follow independent component model structure. In this way, spectral analysis of the sample covariance matrix of HCD can be approached through the basis data. In fact, the structure of the sample covariance matrix of HCD is similar to that of the Pearson sample correlation matrix in basis data. Therefore, we leverage the analysis methods of the spectral theory of the Pearson sample correlation matrix to study the spectral theory of the sample covariance matrix of HCD. In the field of random matrices, research on the spectral theory of the Pearson sample correlation matrix based on independent data is relatively mature. **?** demonstrated that the LSD of sample correlation matrix for i.i.d data is the well-known Marčenko-Pastur law. **?** derive the CLT for LSS of the Pearson sample correlation matrix. The derivation of spectral theory for the sample covariance matrix of HCD can benefit from methods in this context. The LSD of the sample covariance matrix for HCD in Theorem **??** is established following the strategy in **?**, and we further investigate the extreme eigenvalues in Proposition **??**. The proof strategy of CLT for LSS in Theorem **??** follows the methodologies outlined in **?** for the sample covariance matrix and **?** for the sample correlation matrix. However, due to the dependence inherent in HCD, certain tools from these works cannot be directly applied to the sample covariance matrix of HCD. In response, we introduce new techniques. Specifically, we establish concentration inequalities for compositional data. One of the central ideas of the paper, grounded in concentration phenomena, permeates the entire proof (details in Section **??** and Section **??**), where we develop three crucial technique lemmas (see Lemmas **??** – **??**) essential for the proof. Finally, it is noteworthy that the mean and variance-covariance in Theorem **??** differ from those in **?**, and additional terms are present in both the mean and variance-covariance.

The paper is organized as follows. Section **??** investigates the LSD and extreme eigenvalues of the sample covariance matrix for HCD. Section **??** establishes the CLT for LSS of the sample covariance matrix for HCD. Section **??** studies a test problem on the population covariance structure of the basis data based on our CLT result. Section **??** reports numerical studies. Section **??** presents the sketch of proof of our CLT for LSS. Auxiliary lemmas and technical proofs are relegated to the supplementary material.

Before moving forward, we introduce some notations that will be used throughout this paper. We adopt the convention of using regular letters for scalars and boldface letters for vectors or matrices. For any matrix $\mathbf{A}$, we denote its (i,j)-th entry by A_{ij}, its transpose by $\mathbf{A}'$, its trace by $\mathrm{tr}(\mathbf{A})$, its j-th largest eigenvalue by $\lambda_j(\mathbf{A})$, its spectral norm by $\|\mathbf{A}\| = \sqrt{\lambda_1(\mathbf{A}\mathbf{A}')}$. For a set of random variables $\{X_n\}_{n=1}^{\infty}$ and a corresponding set of nonnegative real numbers $\{a_n\}_{n=1}^{\infty}$, we write $X_n = O_P(a_n)$ if, for any $\varepsilon > 0$, there exists a constant $C > 0$ and $N > 0$ such that $\mathbb{P}(|X_n/a_n| \geq C) \leq \varepsilon$ for all $n \geq N$. We write $X_n = o_P(a_n)$ if $\lim_{n\to\infty} \mathbb{P}(|X_n/a_n| \geq \varepsilon) = 0$ for any $\varepsilon > 0$. Furthermore, we write $X_n \xrightarrow{a.s.} X$ ($X_n \xrightarrow{P} X$, $X \xrightarrow{D} X$ respectively) if X_n converges almost surely (in probability, in distribution, respectively) to X. We denote by C and K constants that may vary from line to line.

2 Main Results

2.1 Preliminaries and notations

Let $\mathbf{X}_n = (x_{ij})_{n\times p}$ denote the $n \times p$ observed data matrix, and each row $(x_{i1}, \ldots, x_{ip})$ represents compositions that lie in the $(p-1)$-dimensional simplex $\mathcal{S}^{p-1} = \{(y_1, \ldots, y_p) : \sum_{j=1}^{p} y_j = 1, y_j \geq 0\}$. We assume that the compositional variables arise from a vector of latent variables, which we call the basis. Let $\mathbf{W}_n = (w_{ij})_{n\times p}$ denote the $n \times p$ matrices of unobserved bases, where w_{ij}'s are positive and i.i.d. with mean $\mu > 0$ and variance σ^2. The observed compositional data is generated via the normalization

$$x_{ij} = \frac{w_{ij}}{\sum_{\ell=1}^{p} w_{i\ell}}, \qquad 1 \leq i \leq n,\ 1 \leq j \leq p. \tag{1}$$

The unbiased sample covariance matrix of $\mathbf{X}_n$ is defined by $\mathbf{S}_{n,N} = \frac{1}{N}\mathbf{X}_n'\mathbf{C}_n\mathbf{X}_n$, where $\mathbf{C}_n = \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$, $\mathbf{1}_n$ is a n-dimensional vector of all ones, and $N = n-1$ is the adjusted sample size. Since $\sum_{\ell=1}^{p} w_{i\ell} =$

$p\mu(1+\varepsilon_i)$ with $\sup_i \varepsilon_i = o_P(1)$, we rescale $\mathbf{S}_{n,N}$ as

$$\mathbf{B}_{p,N} = p^2\mathbf{S}_{n,N} = \frac{1}{N}(p\mathbf{X}_n)'\mathbf{C}_n(p\mathbf{X}_n).$$

For any $p \times p$ Hermitian matrix $\mathbf{B}_p$ with eigenvalues $\{\lambda_i\}_{i=1}^p$, its *empirical spectral distribution* (ESD) is defined by

$$F^{\mathbf{B}_p}(x) = \frac{1}{p}\sum_{i=1}^{p} I_{\{\lambda_i(\mathbf{B}_p)\leq x\}},$$

where $I_{\{\cdot\}}$ denotes the indicator function. If $F^{\mathbf{B}_p}(x)$ converges to a non-random limit $F(x)$ as $p \to \infty$, we call $F(x)$ the *limiting spectral distribution* of $\mathbf{B}_p$. The LSD of $\mathbf{B}_p$ is described in terms of its Stieltjes transform. The Stieltjes transform of any cumulative distribution function G is defined by

$$m_G(z) = \int \frac{1}{\lambda - z}\,\mathrm{d}G(\lambda), \qquad z \in \mathbb{C}^+ := \{z : \Im(z) > 0\}.$$

Many classes of statistics related to the eigenvalues of $\mathbf{B}_{p,N}$ are important for multivariate inference, particularly functionals of the ESD. To explore this, for any function f defined on $[0,\infty)$, we consider the *linear spectral statistics* (LSS) of $\mathbf{B}_{p,N}$ given by

$$\int f(x)\,\mathrm{d}F^{\mathbf{B}_{p,N}}(x) = \frac{1}{p}\sum_{i=1}^{p} f\Big(\lambda_i\Big),$$

where λ_i, $i = 1,\ldots,p$, are eigenvalues of $\mathbf{B}_{p,N}$.

In this paper, we study the asymptotic spectral properties of $\mathbf{B}_{p,N}$, including the LSD (see Theorem **??**), the behavior of extreme eigenvalues (see Proposition **??**), and the CLT for LSS (see Theorem **??**).

2.2 Limiting spectral distribution and extreme eigenvalues

Analyzing HCD poses challenges due to its unique dependence structure, making existing techniques for i.i.d. observations less applicable. To overcome this difficulty, we assume that the compositional data is generated from basis data and the basis data follows the commonly used independent component structure. Specifically, the unbiased sample covariance matrix of $\mathbf{X}_n$ is defined by

$$\mathbf{S}_{n,N} = \frac{1}{N}\mathbf{X}_n^{'}\mathbf{C}_n\mathbf{X}_n = \frac{1}{N}\mathbf{W}_n^{'}\mathbf{\Lambda}_n\mathbf{C}_n\mathbf{\Lambda}_n\mathbf{W}_n,$$

where $\mathbf{X}_n = \mathbf{\Lambda}_n\mathbf{W}_n$, and $\mathbf{\Lambda}_n = \mathrm{diag}(1/\sum_{j=1}^p w_{1j},\ldots,1/\sum_{j=1}^p w_{nj})$. Here, we assume that $\mathbf{W}_n$ has i.i.d. components w_{ij}, with $\mathbb{E}(w_{ij}) = \mu > 0$, $\mathrm{Var}(w_{ij}) = \sigma^2$. Recall that the Pearson sample correlation matrix for $\mathbf{W}_n$ is

$$\mathbf{R}_n = \frac{1}{n}\widetilde{\mathbf{X}}_n^{'}\mathbf{C}_n\widetilde{\mathbf{X}}_n = \frac{1}{n}\widetilde{\mathbf{\Lambda}}_p\mathbf{W}_n^{'}\mathbf{C}_n\mathbf{W}_n\widetilde{\mathbf{\Lambda}}_p,$$

where $\widetilde{\mathbf{X}}_n = \mathbf{W}_n\widetilde{\mathbf{\Lambda}}_p$, $\widetilde{\mathbf{\Lambda}}_p = \mathrm{diag}(\sqrt{n}\|\mathbf{w}_1 - \overline{\overline{\mathbf{w}}}_1\|_2^{-1},\ldots,\sqrt{n}\|\mathbf{w}_p - \overline{\overline{\mathbf{w}}}_p\|_2^{-1})$, and $\overline{\overline{\mathbf{w}}}_j = n^{-1}\sum_{i=1}^n w_{ij}\mathbf{1}_n$ with $\mathbf{1}_n$ being an n-dimensional vector whose entries are all 1's. It can be seen that the normalizing matrix $\mathbf{\Lambda}_n$ of $\mathbf{S}_{n,N}$ is very similar to $\widetilde{\mathbf{\Lambda}}_p$ of $\mathbf{R}_n$. The former uses $(\sum_{j=1}^p w_{ij})^{-1}$ for normalization, while the latter utilizes $\sqrt{n}\|\mathbf{w}_j - \overline{\overline{\mathbf{w}}}_j\|_2^{-1}$. This allows us to leverage the techniques from the spectral theory of the Pearson sample correlation matrix in studying the asymptotic spectral properties of the sample covariance matrix for HCD.

Before diving into the LSS of $\mathbf{B}_{p,N}$, we first explore its LSD and extreme eigenvalues. Specifically, suppose the following assumptions hold,

Assumption 2.1. *$\{w_{ij} > 0, i = 1,\dots,n, j = 1,\dots,p\}$ are i.i.d. real random variables with $\mathbb{E}w_{11} = \mu > 0$, $\mathrm{Var}(w_{11}) = \sigma^2$ and $\mathbb{E}(w_{11} - \mu)^4 < \infty$. For notational simplicity, we write $\lambda = \sigma^2/\mu^2$ throughout the paper.*

Assumption 2.2. *$c_N = p/N$ tends to a positive $c > 0$ as $p, N \to \infty$.*

Theorem 2.3. *Under Assumptions **??** and **??**, with probability one, the ESD of $\mathbf{B}_{p,N}$ converges weakly to a deterministic probability distribution with a density function*

$$f_{\mathsf{MP}}(x) = \frac{1}{2\pi c\lambda x}\sqrt{[(b-x)(x-a)]_+}, \tag{2}$$

and a point mass $1 - 1/c$ at $x = 0$ if $c > 1$, where $a := \lambda(1-\sqrt{c})^2$ and $b := \lambda(1+\sqrt{c})^2$.

Proof of Theorem **??**. Let $\mathbf{Y}_n = p\mathbf{C}_n\mathbf{\Lambda}_n\mathbf{W}_n/\sqrt{N}$ and $\check{\mathbf{Y}}_n = \mathbf{C}_n\mathbf{W}_n/(\sqrt{N}\mu)$. Note that the LSD of $\check{\mathbf{Y}}_n'\check{\mathbf{Y}}_n$ is the well-known Marčenko-Pastur law with the density function given by (**??**). From Theorem A.47 of **?** and our Proposition **??**, it suffices to prove that

$$\|\mathbf{Y}_n - \check{\mathbf{Y}}_n\| \xrightarrow{a.s.} 0. \tag{3}$$

By Lemma **??**, we have $\max_{1\le i\le n}|\sum_{j=1}^{p} w_{ij}/(p\mu) - 1| \xrightarrow{a.s.} 0$, which implies that $\|p\mu\mathbf{\Lambda}_n - \mathbf{I}_n\| \xrightarrow{a.s.} 0$. Moreover, we get from Theorem 2.9 in **?** that $\|\mathbf{W}_n/(\mu\sqrt{N})\|$ is bounded almost surely. Hence, we have

$$\|\mathbf{Y}_n - \check{\mathbf{Y}}_n\| = \left\|\mathbf{C}_n(p\mu\mathbf{\Lambda}_n - \mathbf{I}_n)\tfrac{\mathbf{W}_n}{\mu\sqrt{N}}\right\| \le \|p\mu\mathbf{\Lambda}_n - \mathbf{I}_n\|\cdot\left\|\tfrac{\mathbf{W}_n}{\mu\sqrt{N}}\right\| \xrightarrow{a.s.} 0.$$

This completes the proof. $\square$

The LSD $F^c(x)$ has a Dirac mass $1 - 1/c$ at the origin when $c > 1$. We see that $m(\bar{z}) = \overline{m(z)}$. For each $z \in \mathbb{C}^+$, by Theorem **??**, the Stieltjes transform $m(z) = m_{F^c}(z)$ is the unique solution to $m = \{\lambda(1-c-czm) - z\}^{-1}$ in the set $\{m \in \mathbb{C} : \frac{1-c}{z} + \underline{m}(z) \in \mathbb{C}^+\}$. Define $\underline{m}(z)$ as the Stieltjes transform of the companion LSD $\underline{F}^c(x) = (1-c)\delta_0 + cF^c(x)$, where δ_0 is the point distribution at zero. Then $\underline{m}(z)$ is the unique solution in $\{\underline{m} \in \mathbb{C} : \frac{1-c}{z} + \underline{m}(z) \in \mathbb{C}^+\}$ to the equation:

$$z = -\frac{1}{\underline{m}(z)} + \frac{c\lambda}{1+\lambda\underline{m}(z)}, \qquad z \in \mathbb{C}^+. \tag{4}$$

Proposition 2.4. *Under Assumptions **??** and **??**, we have*

$$\lambda_{\max}(\mathbf{B}_{p,N}) \xrightarrow{a.s.} \lambda(1+\sqrt{c})^2 \quad \text{and} \quad \lambda_{\min}(\mathbf{B}_{p,N}) \xrightarrow{a.s.} \lambda(1-\sqrt{c})^2, \tag{5}$$

*where $\lambda_{\max}(\mathbf{B}_{p,N})$ is the largest eigenvalue of $\mathbf{B}_{p,N}$, and $\lambda_{\min}(\mathbf{B}_{p,N})$ is the smallest non-zero eigenvalue of $\mathbf{B}_{p,N}$. Furthermore, for any $\ell > 0$, $\eta_1 > \lambda(1+\sqrt{c})^2$ and $0 < \eta_2 < \lambda(1-\sqrt{c})^2 I_{\{0<c<1\}}$, under condition (**??**), we have*

$$\mathbb{P}\Big(\lambda_{\max}(\mathbf{B}_{p,N}) \ge \eta_1\Big) = o(n^{-\ell}) \quad \text{and} \quad \mathbb{P}\Big(\lambda_{\min}(\mathbf{B}_{p,N}) \le \eta_2\Big) = o(n^{-\ell}). \tag{6}$$

Proof of Proposition **??**. The convergence (**??**) is an immediate consequence of Equation (2.7) in **?**, Theorem 1.4 in **?**, Equation (**??**) and Lemma **??**. The proof of (**??**) is postponed to the supplementary material. $\square$

Remark 1. The results of extreme eigenvalues are useful in locating eigenvalues of the population covariance matrix and proving the CLT for LSS. Proposition **??** demonstrates that, with probability one, there are no eigenvalues of $\mathbf{B}_{p,N}$ outside the support of the LSD under Assumptions **??** – **??**. These results are crucial for applying the Cauchy integral formula (see Equation (**??**)) and proving tightness.

Remark 2. For the special case when $p = n$, the matrix $\mathbf{X}_n = (w_{ij}/\sum_{\ell=1}^{n} w_{i\ell})_{n\times n}$ is a *random Markov matrix*. The work of Bordenave et al. **?** provided key insights into the first-order properties of both eigenvalues and singular values of the $n \times n$ matrix $\mathbf{X}_n$, including the limiting distribution of its singular values and the convergence of its extreme singular values. Our first-order results (see Theorem **??** and Proposition **??**) can be viewed as an extension of their findings regarding singular values when $p, n \to \infty$ and $p/n \to c \in (0, \infty)$.

In contrast to the scope of Bordenave et al. **?**, our work focus on a different setting and investigates a different aspect of $\mathbf{X}_n$. Specifically, we examine the centered version $\mathbf{C}_n\mathbf{X}_n$ without requiring $p = n$ and we focus on second-order fluctuations of LSS. Beyond first-order limits, we derive the CLT for LSS of the sample covariance matrix of $\mathbf{X}_n$ (see Theorem **??** in the following Section **??**).

2.3 CLT for LSS

We focus on linear functionals of eigenvalues of $\mathbf{B}_{p,N}$, i.e. $\frac{1}{p}\sum_{i=1}^{p} f(\lambda_i)$. Naturally, it converges to the functional integration of the LSD of $\mathbf{B}_{p,N}$, that is, $\int f(x)\,\mathrm{d}F^c(x)$. In this section, we explore the second-order fluctuation of $\frac{1}{p}\sum_{i=1}^{p} f(\lambda_i)$, which describes how such LSS converges to its first-order limit. Define

$$G_{p,N}(f) = p\int f(x)\,\mathrm{d}\{F^{\mathbf{B}_{p,N}}(x) - F^{c_N}(x)\},$$

where $F^{c_N}(x)$ substitutes c_N for c in $F^c(x)$, the LSD of $\mathbf{B}_{p,N}$. We show that under Assumptions **??** – **??** and the analyticity of f, the rate of $\int f(x)\,\mathrm{d}\{F^{\mathbf{B}_{p,N}}(x) - F^{c_N}(x)\}$ approaching zero is essentially $1/n$, and $G_{p,N}(f)$ convergence weakly to a Gaussian variable.

Before presenting the main result, we recall some notations. We denote $\lambda = \sigma^2/\mu^2$. Let $m(z)$ and $\underline{m}(z)$ be the Stieltjes transforms of the LSD $F^c(x)$ and the companion LSD $\underline{F}^c(x)$, respectively. Furthermore, we define $m'(z)$ as the derivative of $m(z)$ with respect to z throughout the rest of this paper. The main result is stated in the following theorem.

Theorem 2.5. *Under Assumptions **??** and **??**, let $f_1, f_2, \ldots, f_k$ be functions on $\mathbb{R}$ and analytic on an open interval containing*

$$\left[\lambda(1-\sqrt{c})^2,\ \lambda(1+\sqrt{c})^2\right].$$

Then, the random vector $(G_{p,N}(f_1), \ldots, G_{p,N}(f_k))'$ forms a tight sequence in p and converges weakly to a Gaussian vector $(X_{f_1}, \ldots, X_{f_k})$ with mean function

$$\begin{aligned}\mathbb{E}X_f &= \frac{c\lambda^2}{2\pi i}\oint_{\mathcal{C}} \frac{f(z)\underline{m}^3(z)\{1+\lambda\underline{m}(z)\}}{[\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)]^2}\,\mathrm{d}z \\ &\quad - \frac{1}{2\pi i}\oint_{\mathcal{C}} \frac{f(z)\underline{m}^2(z)\{1+\lambda\underline{m}(z)\}\{z(h_1+\lambda)m(z)+\lambda\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z \\ &\quad - \frac{c}{2\pi i}\oint_{\mathcal{C}} \frac{f(z)z^2\underline{m}^3(z)\{1+\lambda\underline{m}(z)\}\{(\alpha_1+\alpha_2)m^2(z)+2\lambda^2 m'(z)\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z,\end{aligned}$$

and covariance function

$$\operatorname{Cov}(X_f, X_g) = -\frac{1}{2\pi^2}\oint_{\tilde{\mathcal{C}}_1}\oint_{\mathcal{C}_2}\frac{f(z_1)g(z_2)}{\{\underline{m}(z_1)-\underline{m}(z_2)\}^2}\,\mathrm{d}\underline{m}(z_1)\,\mathrm{d}\underline{m}(z_2)$$
$$-\frac{c(\alpha_1+\alpha_2)}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\frac{f(z_1)g(z_2)}{\{1+\lambda\underline{m}(z_1)\}^2\{1+\lambda\underline{m}(z_2)\}^2}\,\mathrm{d}\underline{m}(z_1)\,\mathrm{d}\underline{m}(z_2),$$

where $\alpha_1 = \mathbb{E}(w_{11}/\mu-1)^4 - 3\lambda^2$, $\alpha_2 = -4\lambda\mathbb{E}w_{11}^3/\mu^3 + 4\lambda^3 + 12\lambda^2 + 4\lambda$, $h_1 = -2\mathbb{E}w_{11}^3/\mu^3 + 3\lambda^2 + 5\lambda + 2$. *The contours* $\mathcal{C}, \mathcal{C}_1, \mathcal{C}_2$ *are closed and taken in the positive direction in the complex plane, each enclosing the support of* $F^c(x)$*, i.e.,* $[\lambda(1-\sqrt{c})^2, \lambda(1+\sqrt{c})^2]$.

Remark 3. We restrict attention to functions f which are analytic in a region of the complex plane containing the support of $F^c(x)$. As demonstrated in **?**, the analyticity requirement for f in the CLT can be relaxed by representing the LSS with the help of Helffer–Sjöstrand's formula instead of the Cauchy integral formula. For now, we focus on analytic cases because analytic functions are sufficient to achieve our current statistical objectives.

Applying Theorem **??** to three polynomial functions, we obtain the following corollary. The proof of Theorem **??** is postponed to Section **??**, and detailed calculations in these applications are postponed to the supplementary material.

Corollary 2.6. *Under Assumptions* **??** *and* **??***, let* $f_r = x^r$ *for* $r = 1, 2, 3$*, we have*

$$G_p(f_1) = \operatorname{tr}(\mathbf{B}_{p,N}) - p\lambda \xrightarrow{D} \mathcal{N}(\mu_1, V_1),$$
$$G_p(f_2) = \operatorname{tr}(\mathbf{B}_{p,N}^2) - p(1+c_N)\lambda^2 \xrightarrow{D} \mathcal{N}(\mu_2, V_2),$$
$$G_p(f_3) = \operatorname{tr}(\mathbf{B}_{p,N}^3) - p(1+3c_N+c_N^2)\lambda^3 \xrightarrow{D} \mathcal{N}(\mu_3, V_3),$$

where $c_N = p/N$*, and*

$$\mu_1 = h_1, \quad \mu_2 = (1+c)\lambda^2 + 2(1+c)\lambda h_1 + c(\alpha_1+\alpha_2),$$
$$\mu_3 = (2+6c+3c^2)\lambda^3 + 3(1+3c+c^2)\lambda^2 h_1 + 3c(1+c)\lambda(\alpha_1+\alpha_2),$$
$$V_1 = 2c\lambda^2 + c(\alpha_1+\alpha_2),$$
$$V_2 = 4c(2+c)(1+2c)\lambda^4 + 4c(1+c)^2\lambda^2(\alpha_1+\alpha_2),$$
$$V_3 = 6c(1+6c+3c^2)(3+6c+c^2)\lambda^6 + 9c(1+3c+c^2)^2\lambda^4(\alpha_1+\alpha_2).$$

3 Application

Let $\mathbf{x} = (\frac{w_1}{w}, \frac{w_1}{w}, \dots, \frac{w_p}{w})'$ be a p-dimensional random vector, where $w = \sum_{i=1}^p w_i$. We aim to test the sphericity of the basis data $\mathbf{w} = (w_1, w_2, \dots, w_p)'$ based on the samples $\{\mathbf{x}_i\}_{i=1}^n$, namely the hypothesis $H_0 : \operatorname{Cov}(\mathbf{w}) = \sigma^2\mathbf{I}_p$ where σ^2 is unspecified.

Note that $\mathbb{E}(\mathbf{B}_{p,N}) = \frac{p\nu_2}{p-1}\mathbf{G}_p$, where $\mathbf{G}_p := \mathbf{I}_p - p^{-1}\mathbf{1}_p\mathbf{1}_p'$. We consider the following Frobenius-norm-type test statistic

$$T := \frac{1}{p}\left\|\mathbf{B}_{p,N} - \frac{p\nu_2}{p-1}\mathbf{G}_p\right\|^2.$$

We reject H_0 when T is too large. The test statistic is linked to particular forms of LSS of $\mathbf{B}_{p,N}$ by taking $f(x) = x^k$, $k = 1,2$, that is,

$$T = \frac{1}{p}\mathrm{tr}(\mathbf{B}_{p,N}^2) - \frac{2\nu_2}{p-1}\mathrm{tr}(\mathbf{B}_{p,N}) + \frac{p\nu_2^2}{p-1}.$$

Using the CLT for $\mathrm{tr}(\mathbf{B}_{p,N})$ and $\mathrm{tr}(\mathbf{B}_{p,N}^2)$ under H_0 (Corollary **??**) and the Delta method, the limiting null distribution of T is obtained:

Theorem 3.1. *Suppose that Assumptions **??** and **??** hold, under H_0, we have*

$$p(T-\mu_T) \xrightarrow{D} \mathcal{N}(0,\, \sigma_T^2),$$

where

$$\mu_T = \lambda^2 c_N + \frac{\mu_2}{p} - \frac{2\lambda h_1 + \lambda^2}{p-1}, \quad \sigma_T^2 = 4\lambda^2 V_1 - 4\lambda V_{12} + V_2, \quad V_{12} = 2\lambda c(1+c)(2\lambda^2 + \alpha_1 + \alpha_2), \tag{7}$$

*h_1, α_1, α_2 are defined in Theorem **??**, μ_2, V_1, V_2 are defined in Corollary **??**.*

The detailed proof is postponed to the supplementary material. In practical applications, we replace μ_T with its finite sample counterpart $\hat{\mu}_T$. To eliminate correlation between T and $\hat{\mu}_T$, we split the data $\mathbf{X}$ into two parts:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^{(1)}_{n_1\times p} \\ \mathbf{X}^{(2)}_{n_2\times p} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{pmatrix}, \quad \frac{p}{n_1} \to c_1 \in (0,\infty), \quad \frac{p}{n_2} \to c_2 \in (0,\infty),$$

where $\mathbf{X}^{(1)}$ is used to calculate the test statistic, and $\mathbf{X}^{(2)}$ is used to compute $\hat{\mu}_T$. The new test statistic is defined by

$$\widetilde{T} = \frac{1}{p}\left\| \mathbf{B}^{(1)}_{p,N_1} - \frac{p\nu_2}{p-1}\mathbf{G}_p \right\|^2, \quad \mathbf{B}^{(1)}_{p,N_1} = \frac{p^2}{N_1}\left(\mathbf{X}^{(1)}\right)'\mathbf{C}_{n_1}\mathbf{X}^{(1)}, \quad N_1 = n_1 - 1.$$

The estimate $\hat{\mu}_T$ is obtained by replacing N in μ_T with $N_2 = n_2 - 1$, and substituting the terms λ, $\mathbb{E}(w_{11}/\mu)^3$, and $\mathbb{E}(w_{11}/\mu - 1)^4$ with their finite sample counterparts as follows:

$$\begin{aligned} &\hat{\lambda} = \frac{1}{pn_2}\sum_{i=1}^{n_2}\sum_{j=1}^{p}\left(pX_{ij}^{(2)} - 1\right)^2, \quad \widehat{\mathbb{E}(w_{11}/\mu)^3} = \frac{1}{pn_2}\sum_{i=1}^{n_2}\sum_{j=1}^{p}\left(pX_{ij}^{(2)}\right)^3, \\ &\widehat{\mathbb{E}(w_{11}/\mu - 1)^4} = \frac{1}{pn_2}\sum_{i=1}^{n_2}\sum_{j=1}^{p}\left(pX_{ij}^{(2)} - 1\right)^4. \end{aligned} \tag{8}$$

Using these newly defined notations, we derive the following CLT for $\widetilde{T}$:

Theorem 3.2. *Suppose that Assumptions **??** and **??** hold, under H_0, we have*

$$p(\widetilde{T} - \hat{\mu}_T) \xrightarrow{D} \mathcal{N}(\mu_\lambda, \sigma^2_{\widetilde{T}} + \sigma^2_\lambda),$$

*where $\sigma^2_{\widetilde{T}}$ is defined similarly to σ_T^2 in (**??**), with the limiting value c replaced by c_1, and*

$$\mu_\lambda = -2c_1\lambda h_1, \quad \sigma_\lambda^2 = 4\lambda^2 c_1^2 c_2 \left\{ \mathbb{E}\left(\frac{w_{11}}{\mu} - 1\right)^4 - \lambda^2 + h_2 - 2\lambda h_1 \right\}.$$

*Here, h_1 and h_2 are defined in Lemma **??**.*

The detailed proof of this theorem is postponed to the supplementary material. Based on Theorem **??**, the procedure for testing H_0 is:

$$\text{Reject } H_0 \text{ if } p(T-\hat{\mu}_T)-\hat{\mu}_\lambda > z_\alpha\sqrt{\hat{\sigma}^2_{\tilde{T}}+\hat{\sigma}^2_\lambda}, \tag{9}$$

where z_α represents the upper α-quantile of the standard normal distribution at the nominal level α. The terms $\hat{\mu}_\lambda$, $\hat{\sigma}^2_{\tilde{T}}$, and $\hat{\sigma}^2_\lambda$ are finite-sample estimates of μ_λ, $\sigma^2_{\tilde{T}}$, and σ^2_λ, respectively. These estimates are obtained by replacing λ, $\mathbb{E}(w_{11}/\mu)^3$, and $\mathbb{E}(w_{11}/\mu-1)^4$ with their finite-sample counterparts, as defined in (**??**).

4 Numerical experiments

4.1 Limiting spectral distribution

In this section, simulation experiments are conducted to verify the LSD of the sample covariance matrix $\mathbf{B}_{p,N}$ from compositional data, as stated in Theorem **??**. Compositional data $\{x_{ij}\}_{1\le i\le n, 1\le j\le p}$ is generated by the normalization $x_{ij}=w_{ij}/\sum_{\ell=1}^{p} w_{i\ell}$. We generate basis data w_{ij} from three populations, drawing histograms of eigenvalues of $\mathbf{B}_{p,N}$ and comparing them with theoretical densities. Specifically, three types of distributions for w_{ij} are considered:

1. w_{ij} follows the exponential distribution with rate parameter 5;
2. w_{ij} follows the truncated standard normal distribution lying within the interval $(0,10)$, denoted by $\mathrm{TN}(0,1;0,10)$, where the first two parameters (0 and 1) represent the mean and variance of the standard normal distribution;
3. w_{ij} follows the Poisson distribution with parameter 10.

The dimension and sample size pair, (p,n), is set to $(500,500)$ or $(500,800)$. We display histograms of eigenvalues of $\mathbf{B}_{p,N}$ generated by three populations under various (p,n) combinations and compare them with their respective limiting densities in Figures **??** – **??**. The figures reveal that all histograms align with their theoretical limits, affirming the accuracy of our theoretical results.

(a) Exponential(5) (b) TN(0,1;0,10) (c) Poisson(10)

Figure 1: Histograms of sample eigenvalues of $\mathbf{B}_{p,N}$ with $(p,n)=(500,500)$. The curves are density functions of their corresponding limiting spectral distribution.

(a) Exponential(5) (b) TN(0,1;0,10) (c) Poisson(10)

Figure 2: Histograms of sample eigenvalues of $\mathbf{B}_{p,N}$ with $(p,n)=(500,800)$. The curves are density functions of their corresponding limiting spectral distribution.

4.2 CLT for LSS

In this section, we implement some simulation studies to examine finite-sample properties of some LSS for $\mathbf{B}_{p,N}$ by comparing their empirical means and variances with theoretical limiting values, as stated in Corollary **??**.

First, we compare the empirical mean and variance of $G_{p,N}(x^r)$, $r = 1, 2, 3$, with their corresponding theoretical limits in Corollary **??**. Two types of data distribution of w_{ij} are consider:

1. w_{ij} follows the exponential distribution with rate parameter 5;
2. w_{ij} follows the Chi-squared distribution with degree of freedom 1.

Empirical mean and variance of $G_{p,N}(x^r)$, are calculated for various combinations of (p,n) with $p/n = 3/4$ or $p/n = 1$. For each pair of (p,n), 2000 independent replications are used to obtain the empirical values. Tables **??** – **??** report the empirical results for Exp(5) population and $\chi^2(1)$ population, respectively. As shown in Tables **??** – **??**, the empirical mean and variance of $G_{p,N}(x^r)$ closely match their respective theoretical limits under all scenarios. To verify the asymptotic normality of LSS, we draw the histogram of normalized LSS, $\overline{G}_{p,N}(x^r) = \{G_{p,N}(x^r) - \mu_r\}/\sqrt{V_r}$, $r = 1, 2, 3$, where μ_r and V_r are defined in Corollary **??**, and compare them with the standard normal density. Figures **??** and **??** depict the histograms of $\overline{G}_{p,N}(x^r)$ for Exp(5) population with $p/n = 1$ and $\chi^2(1)$ population with $p/n = 3/4$, respectively. The histograms for the cases of Exp(5) population with $p/n = 3/4$ and $\chi^2(1)$ population with $p/n = 1$ exhibit similar patterns and are omitted for brevity. It can be seen from Figures **??** – **??** that all the histograms conform to the standard normal density, which fully supports our theoretical results.

Table 1: Empirical mean and variance of $G_{p,N}(x^r)$, $r = 1, 2, 3$, with $w_{ij} \sim$ Exp(5).

			$G_{p,N}(x)$		$G_{p,N}(x^2)$		$G_{p,N}(x^3)$	
	p/n	n	mean	var	mean	var	mean	var
Emp	3/4	100	-2.01	2.63	-4	36.54	-7.82	463.32
		200	-1.99	2.93	-3.85	39.73	-7.23	485.05
		300	-1.93	3.03	-3.57	40.3	-6.32	483.76
		400	-2.04	2.95	-3.98	38.78	-7.67	460.01
Theo			**-2**	**3**	**-3.75**	**39**	**-6.81**	**457**
Emp	1	100	-1.91	3.61	-3.83	64.09	-6.56	1064.75
		200	-1.96	3.89	-3.96	68.37	-6.91	1090.14
		300	-2.01	3.97	-4.06	68.7	-7.16	1082.72
		400	-1.98	3.71	-3.99	64.22	-7.07	1010.09
Theo			**-2**	**4**	**-4**	**68**	**-7**	**1050**

(a) $\overline{G}_{p,N}(x)$ (b) $\overline{G}_{p,N}(x^2)$ (c) $\overline{G}_{p,N}(x^3)$

Figure 3: Histograms of normalized LSS $\overline{G}_{p,N}(x^r)$, $r = 1, 2, 3$, with $w_{ij} \sim$ Exp(5) and $p = n = 400$. The curves are density functions of the standard normal distribution.

(a) $\overline{G}_{p,N}(x)$ (b) $\overline{G}_{p,N}(x^2)$ (c) $\overline{G}_{p,N}(x^3)$

Figure 4: Histograms of normalized LSS $\overline{G}_{p,N}(x^r)$, $r = 1, 2, 3$, with $w_{ij} \sim \chi^2(1)$ and $(p,n) = (300, 400)$. The curves are density functions of the standard normal distribution.

Table 2: Empirical mean and variance of $G_{p,N}(x^r)$, $r = 1, 2, 3$, with $w_{ij} \sim \chi^2(1)$.

	p/n	n	$G_{p,N}(x)$ mean	$G_{p,N}(x)$ var	$G_{p,N}(x^2)$ mean	$G_{p,N}(x^2)$ var	$G_{p,N}(x^3)$ mean	$G_{p,N}(x^3)$ var
Emp	3/4	100	-5.79	15.53	-24.19	888.99	-97.31	46790.03
		200	-5.96	16.74	-24.39	920.63	-96.17	45375.75
		300	-5.94	16.6	-23.75	882.92	-90.59	42487.68
		400	-5.88	17.51	-22.68	912.28	-81.2	42922.06
Theo			**-6**	**18**	**-23**	**918**	**-83**	**41806.12**
Emp	1	100	-5.92	20.81	-26.15	1563.02	-102.73	107846.2
		200	-5.98	23.01	-25.15	1639.95	-90.25	105467.9
		300	-5.81	21.82	-23.16	1526.34	-74.54	96864.11
		400	-6.13	23.18	-25.41	1599.96	-90.31	99475.82
Theo			**-6**	**24**	**-24**	**1600**	**-80**	**96000**

4.3 Covariance testing for basis data

Numerical simulations are conducted to determine the empirical size and power of our proposed test statistic. We apply the procedure (**??**) to test H_0. The nominal level is set to be $\alpha = 0.05$. To evaluate the finite sample performance of the test statistic, data are generated from different model scenarios for various (p, n) combinations. For each pair of (p, n), we conduct 2000 independent replications. To assess empirical size, we consider two scenarios for the data matrix $\mathbf{W}_n = (w_{ij})_{n\times p}$:

- Exponential model: $w_{ij} \sim \text{Exp(rate} = 5)$ i.i.d. for $1 \le i \le n$, $1 \le j \le p$;
- Chi-squared model: $w_{ij} \sim \chi^2(5)$ i.i.d. for $1 \le i \le n$, $1 \le j \le p$.

To examine empirical power, we consider two alternatives. An $n\times p$ matrix $\mathbf{Z}_n$ with i.i.d. components is generated from either the Exp(5) or $\chi^2(5)$ models, then the data matrix $\mathbf{W}_n = \mathbf{Z}_n\Sigma$ is constructed with a covariance matrix Σ defined as follows:

- $\Sigma = \Sigma_1$ is diagonal. The first diagonal element is ψ_1, and the remaining elements are all 1. We vary $\psi_1 = 4, 5, 6, 7, 8$.
- $\Sigma = \Sigma_2$ is tridiagonal. The main diagonal elements are all 2, and the lower/upper diagonal elements are ψ_2. We vary $\psi_2 = 0.1, 0.2, 0.3, 0.4, 0.5$.

The simulation results are reported in Tables **??** – **??**. The empirical size values are generally close to the nominal level ($\alpha = 0.05$). The empirical power increase consistently as ψ_1 or ψ_2 increase. For large values of p, n, ψ_1, ψ_2, the power approaches one, indicating that the test performs well in distinguishing between the null and alternative hypotheses as sample size and signal strength increase.

Table 3: Empirical size and power over 2000 replications when the covariance matrix is Σ_1.

			Size	Power				
c	p	n	$\psi_1 = 1$	$\psi_1 = 4$	$\psi_1 = 5$	$\psi_1 = 6$	$\psi_1 = 7$	$\psi_1 = 8$
			Exp(5) *model*					
0.5	100	200	0.1	0.876	0.9345	0.974	0.9885	0.9985
	200	400	0.073	0.9845	0.995	0.9995	1	1
	300	600	0.076	0.998	0.9995	1	1	1
	400	800	0.063	1	1	1	1	1
1	100	100	0.0905	0.5175	0.6045	0.69	0.74	0.806
	200	200	0.077	0.6995	0.8245	0.891	0.921	0.954
	300	300	0.0785	0.8165	0.922	0.963	0.983	0.986
	400	400	0.0615	0.8815	0.9655	0.988	0.995	0.9975
2	100	50	0.1105	0.284	0.319	0.347	0.398	0.4255
	200	100	0.0795	0.3235	0.452	0.4945	0.5805	0.5925
	300	150	0.064	0.379	0.5515	0.631	0.695	0.749
	400	200	0.07	0.4215	0.6065	0.691	0.808	0.8315
			$\chi^2(5)$ *model*					
0.5	100	200	0.0755	0.7455	0.843	0.888	0.898	0.903
	200	400	0.0635	0.9675	0.9955	1	1	0.9995
	300	600	0.055	0.997	1	1	1	1
	400	800	0.05	1	1	1	1	1
1	100	100	0.0705	0.272	0.3235	0.3525	0.3815	0.416
	200	200	0.0735	0.4345	0.624	0.752	0.7875	0.8225
	300	300	0.0565	0.5875	0.814	0.914	0.9475	0.972
	400	400	0.054	0.7005	0.906	0.979	0.996	0.995
2	100	50	0.083	0.1005	0.127	0.1445	0.1275	0.1405
	200	100	0.064	0.1145	0.1845	0.2255	0.278	0.2885
	300	150	0.0565	0.153	0.231	0.3245	0.414	0.4545
	400	200	0.067	0.159	0.2985	0.4205	0.5655	0.625

5 Proof of Theorem ??

In this section, we first present the difference between the CLT for centralized sample covariance $\mathbf{B}_p^0$ and unbiased sample covariance $\mathbf{B}_{p,N}$ by substitution principle in Section **??**, where

$$\mathbf{B}_p^0 = p^2\mathbf{S}_n^0 = \frac{p^2}{n}(\mathbf{X}_n - \mathbb{E}\mathbf{X}_n)'(\mathbf{X}_n - \mathbb{E}\mathbf{X}_n) = \frac{1}{n}\mathbf{Y}_n'\mathbf{Y}_n,$$

$$\mathbf{B}_{p,N} = p^2\mathbf{S}_{n,N} = \frac{p^2}{N}\mathbf{X}_n'\mathbf{C}_n\mathbf{X}_n,$$

and $\mathbf{Y}_n = (y_{ij})_{n\times p}$, $y_{ij} = \frac{w_{ij}}{\overline{w}_i} - 1$ and $\overline{w}_i = \frac{1}{p}\sum_{j=1}^{p} w_{ij}$. By substituting the adjusted sample size $N = n-1$ for the actual sample size n in the centering term, the unbiased sample covariance matrix $\mathbf{B}_{p,N}$ and

Table 4: Empirical size and power over 2000 replications when the covariance matrix is Σ_2.

			Size	Power				
c	p	n	$\psi_2 = 0$	$\psi_2 = 0.1$	$\psi_2 = 0.2$	$\psi_2 = 0.3$	$\psi_2 = 0.4$	$\psi_2 = 0.5$
			Exp(5) *model*					
0.5	100	200	0.1055	0.149	0.397	0.8165	0.9845	1
	200	400	0.075	0.197	0.7895	0.999	1	1
	300	600	0.0635	0.29	0.9775	1	1	1
	400	800	0.066	0.395	0.999	1	1	1
1	100	100	0.1035	0.1135	0.1885	0.332	0.527	0.741
	200	200	0.076	0.1245	0.2775	0.6135	0.923	0.993
	300	300	0.063	0.1245	0.432	0.87	0.9985	1
	400	400	0.0645	0.138	0.5665	0.9765	1	1
2	100	50	0.098	0.1125	0.129	0.1695	0.2335	0.319
	200	100	0.068	0.0875	0.1385	0.2445	0.361	0.5615
	300	150	0.061	0.0885	0.165	0.3035	0.5685	0.789
	400	200	0.062	0.0955	0.1975	0.434	0.7385	0.938
			$\chi^2(5)$ *model*					
0.5	100	200	0.0605	0.1505	0.4535	0.8915	0.996	1
	200	400	0.064	0.2085	0.88	1	1	1
	300	600	0.051	0.3195	0.994	1	1	1
	400	800	0.055	0.45	1	1	1	1
1	100	100	0.0815	0.0975	0.18	0.356	0.6255	0.812
	200	200	0.064	0.1125	0.3035	0.72	0.966	0.9985
	300	300	0.0585	0.1135	0.487	0.947	1	1
	400	400	0.056	0.155	0.6865	0.9965	1	1
2	100	50	0.079	0.0865	0.1325	0.184	0.234	0.342
	200	100	0.065	0.086	0.134	0.242	0.431	0.6305
	300	150	0.0595	0.0965	0.174	0.3715	0.6445	0.877
	400	200	0.06	0.099	0.222	0.5215	0.8355	0.973

the centralized sample covariance $\mathbf{B}_p^0$ share the same CLT (see, Section **??**). The general strategy of the main proof of Theorem **??** is explained in the following and four major steps of the general strategy are presented in Section **??**.

The general strategy of the proof follows the method established in **?** and **?**, with necessary adjustments for handling the sample covariance matrix of HCD, where conventional tools are not directly applicable. Our novel techniques play a pivotal role in overcoming these challenges. To begin with, we follow the strategy in **?** to establish the LSD of $\mathbf{B}_{p,N}$ in Theorem **??**. Then, we develop Proposition **??** to find the extreme eigenvalues of $\mathbf{B}_{p,N}$. Notably, these extreme eigenvalues are highly concentrated around two edges of the support, a crucial aspect for applying the Cauchy integral formula (**??**) and proving tightness. Given that compositional data $x_{ij} = w_{ij}/\sum_{\ell=1}^{p} w_{i\ell}$ are not i.i.d., dealing with the CLT for LSS of the unbiased sample covariance matrix $\mathbf{B}_{p,N}$ presents challenges. To address this, we employ the substitution principle (**?**) to reduce the problem to the CLT for LSS of the centralized sample covariance $\mathbf{B}_p^0$. By substituting the adjusted sample size $N = n-1$ for the actual sample size n in the centering term, both the unbiased sample covariance matrix $\mathbf{B}_{p,N}$ and the centralized sample covariance $\mathbf{B}_p^0$ share the same CLT (see Section **??**). We then leverage the independence of samples to further study the CLT for LSS of $\mathbf{B}_p^0$. Specifically, we exploit the independence of samples to establish independence for $\mathbf{r}_i = \frac{1}{\sqrt{n}}\Big(\frac{w_{i1}}{\overline{w}_i}-1,\ldots,\frac{w_{ip}}{\overline{w}_i}-1\Big)', i=1,2,\ldots,n$, and express $\mathbf{B}_p^0$ as $\mathbf{B}_p^0 = \frac{1}{n}\mathbf{Y}_n'\mathbf{Y}_n = \sum_{i=1}^n \mathbf{r}_i\mathbf{r}_i'$. The ultimate goal is to establish the CLT for LSS of $\mathbf{B}_p^0$.

By the Cauchy integral formula, we have

$$\int f(x)\,\mathrm{d}G(x) = -\frac{1}{2\pi i}\oint_{\mathcal{C}} f(z)m_G(z)\,\mathrm{d}z \tag{10}$$

valid for any c.d.f G and any analytic function f on an open set containing the support of G, where $\oint_{\mathcal{C}}$ is the contour integration in the anti-clockwise direction. In our case, $G(x) = p(F^{\mathbf{B}_p^0}(x) - F^{c_n}(x))$. Therefore, the problem of finding the limiting distribution reduces to the study of $M_p(z)$ defined as follows:

$$\begin{aligned}
M_p(z) &= p\Big\{m_p(z) - m_p^0(z)\Big\} = n\Big\{\underline{m}_p(z) - \underline{m}_p^0(z)\Big\},\\
m_p(z) &= m_{F^{\mathbf{B}_p^0}}(z) = \frac{1}{p}\mathrm{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1}, \qquad m_p^0(z) = m_{F^{c_n}}(z),\\
\underline{m}_p(z) &= m_{F^{\underline{\mathbf{B}}_p^0}}(z) = \frac{1}{p}\mathrm{tr}(\underline{\mathbf{B}}_p^0 - zI_n)^{-1}, \qquad \underline{m}_p^0(z) = \underline{m}_{F^{c_n}}(z),\\
\underline{\mathbf{B}}_p^0 &= \frac{p^2}{n}(\mathbf{X}_n - \mathbb{E}\mathbf{X}_n)(\mathbf{X}_n - \mathbb{E}\mathbf{X}_n)'.
\end{aligned}$$

Note that the support of $F^{\mathbf{B}_{p,N}}$ is random. Fortunately, we have shown that the extreme eigenvalues of $\mathbf{B}_{p,N}$ are highly concentrated around two edges of the support of the limiting MP law $F^c(x)$ (see, Theorem **??**, Proposition **??**). Then the contour $\mathcal{C}$ can be appropriately chosen. Moreover, as in **?**, by Proposition **??**, we can replace the process $\{M_p(z), z\in\mathcal{C}\}$ by a slightly modified process $\{\widehat{M}_p(z), z\in\mathcal{C}\}$. Below we present the definitions of the contour $\mathcal{C}$ and the modified process $\widehat{M}_p(z)$. Let x_r be any number greater than $\lambda(1+\sqrt{c})^2$. Let x_l be any negative number if $\lambda(1-\sqrt{c})^2 = 0$. Otherwise we choose $x_l \in (0, \lambda(1-\sqrt{c})^2)$. Now let $\mathcal{C}_u = \{x + iv_0 : x\in[x_l, x_r]\}$. Then we define $\mathcal{C}^+ := \{x_l + iv : v\in[0,v_0]\}\cup\mathcal{C}_u\cup\{x_r+iv : v\in[0,v_0]\}$, and $\mathcal{C} = \mathcal{C}^+\cup\overline{\mathcal{C}^+}$. Now we define the subsets $\mathcal{C}_n$ of $\mathcal{C}$ on which $M_p(\cdot)$ equals to $\widehat{M}_p(\cdot)$. Choose sequence $\{\varepsilon_n\}$ decreasing to zero satisfying for some $\alpha\in(0,1)$, $\varepsilon_n \ge n^{-\alpha}$. Let

$$\mathcal{C}_l = \begin{cases}\{x_l + iv : v\in[n^{-1}\varepsilon_n, v_0]\}, & \text{if } x_l > 0,\\ \{x_l + iv : v\in[0,v_0]\}, & \text{if } x_l < 0,\end{cases}$$

and $\mathcal{C}_r = \{x_r + iv : v\in[n^{-1}\varepsilon, v_0]\}$ for any $v_0 > 0$. Then $\mathcal{C}_n = \mathcal{C}_l\cup\mathcal{C}_u\cup\mathcal{C}_r$. For $z = x+iv$, we define

$$\widehat{M}_p(z) = \begin{cases} M_p(z), & \text{for } z\in\mathcal{C}_n\\ M_p(x_r + in^{-1}\varepsilon_n), & \text{for } x = x_r, v\in[0, n^{-1}\varepsilon_n], \text{ and if } x_l > 0\\ M_p(x_l + in^{-1}\varepsilon_n), & \text{for } x = x_l, v\in[0,n^{-1}\varepsilon_n],\end{cases}$$

Most of the paper will deal with proving the following proposition.

Proposition 5.1. *Under Assumption* **??** *and* **??**, $\widehat{M}_p(\cdot)$ *converges weakly to a two-dimensional Gaussian process* $M(\cdot)$ *for* $z \in \mathcal{C}$, *with mean*

$$\mathbb{E}M(z) = \frac{\underline{m}^2(z)\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\Bigg[\{z(h_1+\lambda)m(z)+\lambda\}$$
$$+ cz^2\underline{m}(z)\Big\{(\alpha_1+\alpha_2)m^2(z)+2\lambda^2 m'(z)\Big\} - \frac{c\lambda^2\underline{m}(z)}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\Bigg], \tag{11}$$

and covariance function

$$\mathrm{Cov}(M(z_1),M(z_2)) = 2\left[\frac{\underline{m}'(z_1)\underline{m}'(z_2)}{\{\underline{m}(z_1)-\underline{m}(z_2)\}^2} - \frac{1}{(z_1-z_2)^2}\right] + \frac{c(\alpha_1+\alpha_2)\underline{m}'(z_1)\underline{m}'(z_2)}{\{1+\lambda\underline{m}(z_1)\}^2\{1+\lambda\underline{m}(z_2)\}^2}. \tag{12}$$

Now we explain how Theorem **??** follows from the above proposition. As in **?**, with probability one, $|\int f(z)\{M_p(z)-\widehat{M}_p(z)\}\mathrm{d}z| \to 0$ as $n\to\infty$. Combining this observation with (**??**), Theorem **??** follows from Proposition **??**. To prove Proposition **??**, we decompose $M_p(z)$ into a random part $M_p^{(1)}(z)$ and a deterministic part $M_p^{(2)}(z)$ for $z \in \mathcal{C}_n$, that is, $M_p(z) = M_p^{(1)}(z) + M_p^{(2)}(z)$, where

$$M_p^{(1)}(z) = p\Big\{m_p(z) - \mathbb{E}m_p(z)\Big\} \quad \text{and} \quad M_p^{(2)}(z) = p\Big\{\mathbb{E}m_p(z) - m_p^0(z)\Big\}.$$

The random part contributes to the covariance function and the deterministic part contributes to the mean function. By Theorem 8.1 in **?**, the proof of Proposition **??** is then complete if we can verify the following four steps:

Step 1 Truncation.

Step 2 Finite-dimensional convergence of $M_p^{(1)}(z)$ in distribution on $\mathcal{C}_n$ to a centered multivariate Gaussian random vector with covariance function given by (**??**).

Step 3 Tightness of the $M_p^{(1)}(z)$ for $z \in \mathcal{C}_n$.

Step 4 Convergence of the non-random part $M_p^{(2)}(z)$ to (**??**) on $z \in \mathcal{C}_n$.

The proof of these steps is presented in the coming sections. Before that, we introduce the substitution principle and crucial lemmas in Sections **??** and **??** respectively. The former explains the reduction of problem of the CLT for LSS of $\mathbf{B}_{p,N}$ to that of $\mathbf{B}_p^0$, while the latter provides essential lemmas for these four steps in proving the CLT for LSS of $\mathbf{B}_p^0$.

5.1 Substitution principle

By the Cauchy integral formula, we have

$$G_{p,N}(f) = -\frac{1}{2\pi i}\oint_{\mathcal{C}} f(z)\Big\{\mathrm{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - pm_N^0(z)\Big\}\mathrm{d}z$$

valid for any function f analytic on an open set containing the support of $G_{p,N}$, where

$$m_N^0(z) \equiv m_{F^{c_N}}(z) = \frac{1}{\lambda(1-c_N-c_N z m_N^0) - z},$$

with $c_N = p/N$. To obtain the asymptotic distribution of $G_{p,N}(f)$, it is necessary to find the asymptotic distribution of $\mathrm{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - pm_N^0(z)$. To achieve this, we derive the following Lemma **??** whose proof is postponed to the supplementary material.

Lemma 5.2. *Under Assumptions* **??** *and* **??**, *as* $n \to \infty$,

$$\operatorname{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - pm_N^0(z) = \operatorname{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - pm_n^0(z) + o_P(1).$$

By Lemma **??**, the asymptotic distribution of $G_{p,N}(f)$ is identical to that of

$$G_p^0(f) = -\frac{1}{2\pi i}\oint_{\mathcal{C}} f(z)\Big\{\operatorname{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - pm_n^0(z)\Big\}\mathrm{d}z,$$

where $c_n = p/n$, $m_n^0(z) = m_{F^{c_n}}(z)$ (note that we denote $m_n^0(z)$ as $m_p^0(z)$ in other sections except this subsection).

5.2 Some important lemmas

Before delving into the proof of the CLT for LSS, it is crucial to introduce three pivotal lemmas, representing novel contributions to this paper, that unveil concentration phenomena. Lemma **??** is crafted to estimate essential parameters, facilitating the derivation of estimates of any order. Concerning ν_2 and ν_{12}, the terms h_1/p and h_2/p emerge as non-negligible due to the multiplication by p in the CLT. To address these parameters, we establish that the probability of the event $B_p^c(\varepsilon)$ decays polynomially to 0 and leverage Taylor expansion on the event $B_p(\varepsilon) = \{\omega : |\overline{w}_i - \mu| \le \varepsilon, \overline{w}_i = \sum_{j=1}^p w_{ij}/p\}$ to handle the issue of dependence. The proof of the CLT for LSS relies on two pivotal steps: the moment inequality for random quadratic forms and the precise estimation of the expectation of the product of two random quadratic forms. Lemma **??** establishes the former step, essential for converting them into the corresponding traces, while Lemma **??** establishes the latter step, enabling the application of CLT for martingale differences. Both Lemma **??** and Lemma **??** heavily hinge on the estimation of parameters ν_2, ν_4, and ν_{12} in Lemma **??**. The proof of Lemmas **??** – **??** are postponed to the supplementary material.

Lemma 5.3. *Suppose that* $\mathbf{w} = \big(w_1, \dots, w_p\big)'$ *has i.i.d. entries with* $\mathbb{E}w_1 = \mu$, $\mathbb{E}(w_1 - \mu)^2 = \sigma^2$, *and* $\mathbb{E}\big|w_1 - \mu\big|^4 < \infty$, *let* $\overline{w} = p^{-1}\sum_{j=1}^p w_j$, $\lambda = \sigma^2/\mu^2$, *then*

$$\begin{aligned}
\nu_2 &:= \mathbb{E}\Big(\frac{w_1}{\overline{w}} - 1\Big)^2 = \lambda + \frac{h_1}{p} + o(p^{-1}),\\
\nu_{12} &:= \mathbb{E}\Big(\frac{w_1}{\overline{w}} - 1\Big)^2\Big(\frac{w_2}{\overline{w}} - 1\Big)^2 = \lambda^2 + \frac{h_2}{p} + o(p^{-1}),\\
\nu_4 &:= \mathbb{E}\Big(\frac{w_1}{\overline{w}} - 1\Big)^4 = \mathbb{E}\Big(\frac{w_1}{\mu} - 1\Big)^4 + o(1),
\end{aligned}$$

where

$$h_1 = -2\frac{\mathbb{E}w_1^3}{\mu^3} + 3\lambda^2 + 5\lambda + 2, \quad h_2 = -8\lambda\frac{\mathbb{E}w_1^3}{\mu^3} + 10\lambda^3 + 22\lambda^2 + 8\lambda.$$

Lemma 5.4. *Suppose that* $\mathbf{w} = \big(w_1, \dots, w_p\big)'$ *has i.i.d. entries with* $\mathbb{E}w_1 = \mu$, $\mathbb{E}(w_1 - \mu)^2 = \sigma^2$, *for any* $p \times p$ *matrix* $\mathbf{A}$ *and* $q \ge 2$, *then there is a positive constant* K_q *depending on* q *such that*

$$\begin{aligned}
&\mathbb{E}\Big|\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n}\operatorname{tr}\mathbf{A}\Big|^q\\
&\le K_q\Big\{n^{-q}\Big[\Big\{\mathbb{E}\big|w_1 - \mu\big|^4\operatorname{tr}(\mathbf{A}\mathbf{A}')\Big\}^{q/2} + \mathbb{E}|w_1 - \mu|^{2q}\operatorname{tr}(\mathbf{A}\mathbf{A}')^{q/2}\Big] + n^q\mathbb{P}\Big(B_p^c(\varepsilon)\Big)\|\mathbf{A}\|^q + n^{-q}\|\mathbf{A}^q\|h_1^q\Big\},
\end{aligned}$$

where $\mathbf{r} = n^{-1/2}(w_1/\overline{w} - 1, \ldots, w_p/\overline{w} - 1)'$, h_1 *is defined in Lemma* **??**, $B_p(\varepsilon) = \{\omega : |\overline{w} - u| \le \varepsilon, \overline{w} = \sum_{j=1}^{p} w_j/p\}$, *and*

$$\mathbb{P}(B_p^c(\varepsilon)) \le C\varepsilon^{-kq_1}\Big(\sigma^{kq_1} p^{-kq_1/2} + p^{-kq_1+1}\mathbb{E}\left|w_1 - \mu\right|^{kq_1}\Big), \tag{13}$$

in which $\varepsilon, k, q_1, C > 0$ *are constants. Furthermore, if* $\mathbb{E}\left|w_1 - \mu\right|^4 < \infty$, $|w_j - \mu| < \delta_n\sqrt{n}$ *for all* $j = 1, \ldots, p$, *and* $\|\mathbf{A}\|$ *is bounded, then, for any* $q \ge 2$,

$$\mathbb{E}\left|\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\right|^q \le K_q n^{-1}\delta_n^{2q-4}.$$

Lemma 5.5. *Suppose that* $\mathbf{w} = \big(w_1, \ldots, w_p\big)'$ *has i.i.d. entries with* $\mathbb{E}w_1 = \mu$, $\mathbb{E}(w_1 - \mu)^2 = \sigma^2$, $\mathbf{A}$ *and* $\mathbf{B}$ *are* $p \times p$ *matrices , if* $\mathbb{E}\left|w_1 - \mu\right|^4 < \infty$, $\|\mathbf{A}\|$ *and* $\|\mathbf{B}\|$ *are bounded, then*

$$\begin{aligned}
&\mathbb{E}\Big(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big)\Big(\mathbf{r}'\mathbf{B}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{B}\Big) \\
&= \frac{\nu_4 - 3\nu_{12}}{n^2}\sum_{i=1}^{p} A_{ii}B_{ii} + \frac{\nu_{12}}{n^2}\Big\{\mathrm{tr}(\mathbf{A}\mathbf{B}') + \mathrm{tr}(\mathbf{A}\mathbf{B})\Big\} + \frac{\nu_{12} - \nu_2^2}{n^2}\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B} + o(n^{-1}).
\end{aligned}$$

5.3 CLT for LSS of the centralized sample covariance matrix B_p^0

5.3.1 Step 1: Truncation

We begin the proof of Proposition **??** with the replacement of the entries of $\mathbf{W}_n$ with truncated variables. We can choose a positive sequence of $\{\delta_n\}$ such that

$$\delta_n \to 0, \quad \delta_n n^{1/4} \to \infty, \quad \delta_n^{-4}\mathbb{E}w_{11}^4 I_{\{|w_{11}-\mu| \ge \delta_n\sqrt{n}\}} \to 0.$$

Let $\widehat{\mathbf{B}}_p^0 = \frac{p^2}{n}(\widehat{\mathbf{X}}_n - \mathbb{E}\widehat{\mathbf{X}}_n)'(\widehat{\mathbf{X}}_n - \mathbb{E}\widehat{\mathbf{X}}_n)$, where the (i,j)-th entry of $\widehat{\mathbf{X}}_n$ is normalized using truncated variables $\hat{w}_{ij} = w_{ij}I_{\{|w_{ij}-\mu|<\delta_n\sqrt{n}\}}$, as described in (**??**). We then have

$$\begin{aligned}
\mathbb{P}(\mathbf{B}_p^0 \ne \widehat{\mathbf{B}}_p^0) &\le \mathbb{P}\Big(\cup_{i\le n, j\le p}\Big\{|w_{ij} - \mu| \ge \delta_n\sqrt{n}\Big\}\Big) \le np \cdot \mathbb{P}\Big(|w_{ij} - \mu| \ge \delta_n\sqrt{n}\Big) \\
&\le K\delta_n^{-4}\int_{\{|w_{ij}-\mu|\ge\delta_n\sqrt{n}\}} |w_{11}|^4 = o(1).
\end{aligned}$$

Let $\widehat{G}_p^0(x)$ be $G_p^0(x)$ with $\mathbf{B}_p^0$ replaced by $\widehat{\mathbf{B}}_p^0$, then $\mathbb{P}(\widehat{G}_p^0(x) \ne G_p^0(x)) \le \mathbb{P}(\mathbf{B}_p^0 \ne \widehat{\mathbf{B}}_p^0) = o(1)$. In view of the above, we obtain

$$\int f_j(x)\,\mathrm{d}G_p^0(x) = \int f_j(x)\,\mathrm{d}\widehat{G}_p^0(x) + o_P(1).$$

To simplify notations, we below still use w_{ij} instead of $\hat{w}_{ij}$, and assume that

$$|w_{ij} - \mu| < \delta_n\sqrt{n}, \quad \mathbb{E}w_{ij} = \mu > 0, \quad \mathrm{Var}(w_{ij}) = \sigma^2, \quad \mathbb{E}|w_{ij} - \mu|^4 < \infty. \tag{14}$$

5.3.2 Step 2: Finite dimensional convergence of $M_p^{(1)}(z)$ in distribution

Lemma 5.6. *Under Assumptions* **??** *and* **??**, *as* $p \to \infty$, *for any set of* r *points* $\{z_1, z_2, \ldots, z_r\} \bigcup \mathcal{C}$, *the random vector* $\Big(M_p^{(1)}(z_1), \ldots, M_p^{(1)}(z_r)\Big)$ *converges weakly to a* r*-dimensional centered Gaussian distribution with covariance function* (**??**).

We now proceed to the proof of this lemma. By the fact that a random vector is multivariate normally distributed if and only if every linear combination of its components is normally distributed, we need only show that, for any positive integer r and any complex sequence $\{\alpha_j\}_{j=1}^r$, the sum $\sum_{j=1}^r \alpha_j M_p^{(1)}(z_j)$ converges weakly to a Gaussian random variable. To this end, we first approximate $M_p^{(1)}(z)$ by a sum of martingale difference, which is given in (**??**). Then, we apply the martingale CLT (Theorem 35.12 in **?**) to obtain the asymptotic distribution of $M_p^{(1)}(z)$. Details of these two steps are provided in the following two parts.

Part 1: Martingale difference decomposition of $M_p^{(1)}(z)$.

First, we introduce some notations. In the following proof, we assume that $v = \Im z \geq v_0 > 0$. Moreover, for $j = 1, 2, \dots, n$, let

$$\mathbf{r}_j = \frac{1}{\sqrt{n}}\left(\frac{w_{j1}}{\overline{w}_j} - 1, \dots, \frac{w_{jp}}{\overline{w}_j} - 1\right)', \quad \mathbf{D}(z) = \mathbf{B}_p^0 - z\mathbf{I}_p, \quad \mathbf{D}_j(z) = \mathbf{D}(z) - \mathbf{r}_j\mathbf{r}_j',$$
$$\beta_j(z) = \frac{1}{1 + \mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{r}_j}, \quad \overline{\beta}_j(z) = \frac{1}{1 + n^{-1}\nu_2 \mathrm{tr}\mathbf{D}_j^{-1}(z)}, \quad b_p(z) = \frac{1}{1 + n^{-1}\nu_2 \mathbb{E}\mathrm{tr}\mathbf{D}_1^{-1}(z)},$$
$$\varepsilon_j(z) = \mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{r}_j - \frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_j^{-1}(z), \quad \tilde{\varepsilon}_j(z) = \mathbf{r}_j'\mathbf{D}_j^{-2}(z)\mathbf{r}_j - \frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_j^{-2}(z) = \frac{\mathrm{d}}{\mathrm{d}z}\varepsilon_j(z).$$

By Lemma **??**, we have, for any $q \geq 2$,

$$\mathbb{E}|\varepsilon_j(z)|^q \leq \frac{K}{v^{2q}} n^{-1}\delta_n^{2q-4} \quad \text{and} \quad \mathbb{E}|\tilde{\varepsilon}_j(z)|^q \leq \frac{K}{v^{2q}} n^{-1}\delta_n^{2q-4}. \tag{15}$$

It is easy to see that

$$\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z) = -\mathbf{D}_j^{-1}(z)\mathbf{r}_j\mathbf{r}_j'\mathbf{D}_j^{-1}(z)\beta_j(z), \tag{16}$$

where we use the formula that $\mathbf{A}_1^{-1} - \mathbf{A}_2^{-1} = \mathbf{A}_2^{-1}(\mathbf{A}_2 - \mathbf{A}_1)\mathbf{A}_1^{-1}$ holds for any two invertible matrices $\mathbf{A}_1$ and $\mathbf{A}_2$. Note that $|\beta_j(z)|$, $|\overline{\beta}_j(z)|$ and $|b_n(z)|$ are bounded by $\frac{|z|}{v}$. Let $\mathbb{E}_j(\cdot)$ denote conditional expectation with respect to the σ-field generated by $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_j\}$, where $j = 1, 2, \dots, n$. By convention, we use $\mathbb{E}_0 = \mathbb{E}$ to denote expectation. By using (**??**), we write

$$M_p^{(1)}(z) = \sum_{j=1}^n \mathrm{tr}\Big\{(\mathbb{E}_j - \mathbb{E}_{j-1})\mathbf{D}^{-1}(z)\Big\} = -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z)\mathbf{r}_j'\mathbf{D}_j^{-2}(z)\mathbf{r}_j.$$

From the identity $\beta_j(z) = \overline{\beta}_j(z) - \beta_j(z)\overline{\beta}_j(z)\varepsilon_j(z) = \overline{\beta}_j(z) - \overline{\beta}_j^2(z)\varepsilon_j(z) + \overline{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z)$ and the definition of $\tilde{\varepsilon}_j(z)$, we obtain that

$$\begin{aligned}
&(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z)\mathbf{r}_j'\mathbf{D}_j^{-2}(z)\mathbf{r}_j \\
&= (\mathbb{E}_j - \mathbb{E}_{j-1})\left[\left\{\overline{\beta}_j(z) - \overline{\beta}_j^2(z)\varepsilon_j(z) + \overline{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z)\right\}\left\{\tilde{\varepsilon}_j(z) + \frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_j^{-2}(z)\right\}\right] \\
&= -Y_j(z) + \mathbb{E}_{j-1}Y_j(z) - (\mathbb{E}_j - \mathbb{E}_{j-1})\left[\overline{\beta}_j^2(z)\left\{\varepsilon_j(z)\tilde{\varepsilon}_j(z) - \beta_j(z)\varepsilon_j^2(z)\mathbf{r}_j'\mathbf{D}_j^{-2}(z)\mathbf{r}_j\right\}\right],
\end{aligned}$$

where

$$Y_j(z) := -\mathbb{E}_j\left\{\overline{\beta}_j(z)\tilde{\varepsilon}_j(z) - \overline{\beta}_j^2(z)\varepsilon_j(z)\frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_j^{-2}(z)\right\},$$

and the second equality follows from $(\mathbb{E}_j - \mathbb{E}_{j-1})\overline{\beta}_j(z)\mathrm{tr}\mathbf{D}_j^{-2}(z) = 0$. By using (**??**), we have

$$\mathbb{E}\left|\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\overline{\beta}_j^2(z)\varepsilon_j(z)\tilde{\varepsilon}_j(z)\right|^2 \leq 4\sum_{j=1}^{n}\mathbb{E}\left|\overline{\beta}_j^2(z)\varepsilon_j(z)\tilde{\varepsilon}_j(z)\right|^2 = o(1),$$

here we use the the martingale difference property of $(\mathbb{E}_j - \mathbb{E}_{j-1})\overline{\beta}_j^2(z)\varepsilon_j(z)\tilde{\varepsilon}_j(z)$. Thus, $\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\overline{\beta}_j^2(z)\varepsilon_j(z)\tilde{\varepsilon}_j(z) \xrightarrow{P} 0$. By the same argument, we have

$$\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\overline{\beta}_j^2(z)\beta_j(z)\mathbf{r}_j'\mathbf{D}_j^{-2}(z)\mathbf{r}_j\varepsilon_j^2(z) \xrightarrow{P} 0.$$

The estimates above imply that

$$M_p^{(1)}(z) = \sum_{j=1}^{n}\{Y_j(z) - \mathbb{E}_{j-1}Y_j(z)\} + o_P(1), \tag{17}$$

where $\{Y_j(z) - \mathbb{E}_{j-1}Y_j(z)\}_{j=1}^{n}$ is a sequence of martingale difference.
Part 2: Application of martingales CLT to (**??**).

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

$$\sum_{i=1}^{r}\alpha_i M_p^{(1)}(z_i) = \sum_{j=1}^{n}\sum_{i=1}^{r}\alpha_i\Big\{Y_j(z_i) - \mathbb{E}_{j-1}Y_j(z_i)\Big\} + o_P(1), \tag{18}$$

where $\Im(z_i) \neq 0$, and $\{\alpha_i\}_{i=1}^{r}$ are constants. We apply the martingale CLT (**?**, Theorem 35.12) to this martingale difference decomposition (**??**). To this end, we need to check two conditions:

$$\sum_{j=1}^{n}\mathbb{E}\left|\sum_{i=1}^{r}\alpha_i\{Y_j(z_i) - \mathbb{E}_{j-1}Y_j(z_i)\}\right|^2 I_{\{|\sum_{i=1}^{r}\alpha_i(Y_j(z_i)+\mathbb{E}_{j-1}Y_j(z_i))|\geq\varepsilon\}} \to 0, \tag{19}$$

$$\sum_{j=1}^{n}\mathbb{E}_{j-1}\Big[\Big\{Y_j(z_1) - \mathbb{E}_{j-1}Y_j(z_1)\Big\}\Big\{Y_j(z_2) - \mathbb{E}_{j-1}Y_j(z_2)\Big\}\Big] \xrightarrow{P} (\mathbf{??}). \tag{20}$$

First, we verify (**??**). By Lemma **??**, we obtain

$$\mathbb{E}|Y_j(z)|^4 \leq K\mathbb{E}|\varepsilon_j(z)|^4 = o(p^{-1}),$$

which, together with Jensen's inequality, implies that

$$\mathbb{E}|\mathbb{E}_{j-1}Y_j(z)|^4 \leq \mathbb{E}(\mathbb{E}_{j-1}|Y_j(z)|^4) = \mathbb{E}|Y_j(z)|^4 = o(p^{-1}).$$

It follows from the above two equations that

$$\text{LHS of } (\mathbf{??}) \leq \frac{K}{\varepsilon^2}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{i=1}^{r}\alpha_i Y_j(z_i)\right|^4 + \frac{K}{\varepsilon^2}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{i=1}^{r}\alpha_i\mathbb{E}_{j-1}Y_j(z_i)\right|^4 \to 0.$$

Then, we verify (**??**). Since $Y_j(z) = -\mathbb{E}_j \frac{\mathrm{d}}{\mathrm{d}z}\Big\{\overline{\beta}_j(z)\varepsilon_j(z)\Big\}$, we have

$$\begin{aligned}\text{LHS of } (\textbf{??}) &= \frac{\partial^2}{\partial z_1 \partial z_2}\Big\{\mathcal{Y}_1(z_1,z_2) - \mathcal{Y}_2(z_1,z_2)\Big\}, \qquad (21)\\ \mathcal{Y}_1(z_1,z_2) &:= \sum_{j=1}^{n} \mathbb{E}_{j-1}\Big[\mathbb{E}_j\Big\{\overline{\beta}_j(z_1)\varepsilon_j(z_1)\Big\}\mathbb{E}_j\Big\{\overline{\beta}_j(z_2)\varepsilon_j(z_2)\Big\}\Big],\\ \mathcal{Y}_2(z_1,z_2) &:= \sum_{j=1}^{n} \mathbb{E}_{j-1}\Big\{\overline{\beta}_j(z_1)\varepsilon_j(z_1)\Big\}\mathbb{E}_{j-1}\Big\{\overline{\beta}_j(z_2)\varepsilon_j(z_2)\Big\}.\end{aligned}$$

Thus, it is enough to consider the limits of $\mathcal{Y}_i(z_1,z_2), i = 1,2$, which are provided in the following lemma.

Lemma 5.7. *Under Assumptions* **??** *and* **??**, *we have*

$$\frac{\partial^2}{\partial z_1 \partial z_2}\mathcal{Y}_1(z_1,z_2) \xrightarrow{P} (\textbf{??}), \qquad \frac{\partial^2}{\partial z_1 \partial z_2}\mathcal{Y}_2(z_1,z_2) \xrightarrow{P} 0.$$

The proof of Lemma **??** is postponed to the supplementary material. This lemma and Equation (**??**) complete the proof of (**??**).

5.3.3 Step 3: Tightness of $M_p^{(1)}(z)$

To prove tightness of $M_p^{(1)}(z)$, it is sufficient to prove the moment condition of **?**, i.e.,

$$\sup_{n; z_1, z_2 \in \mathcal{C}_n} \frac{\mathbb{E}|M_p^{(1)}(z_1) - M_p^{(1)}(z_2)|^2}{|z_1 - z_2|^2}$$

is finite. Its proof exactly follows **?**, and is postponed to the supplementary material.

5.3.4 Step 4: Convergence of $M_p^{(2)}(z)$

Recalling that $M_p^{(2)}(z) = p\{\mathbb{E}m_p(z) - m_p^0(z)\} = n\{\mathbb{E}\underline{m}_p(z) - \underline{m}_p^0(z)\}$. From the identity

$$\frac{1}{\mathbb{E}\underline{m}_p(z)} + z - \frac{c_n\lambda}{1+\lambda\mathbb{E}\underline{m}_p(z)} = \frac{1}{\mathbb{E}\underline{m}_p(z)}\left\{1 - c_n + z\mathbb{E}\underline{m}_p(z) + \frac{c_n}{1+\lambda\mathbb{E}\underline{m}_p(z)}\right\},$$

we have

$$\mathbb{E}\underline{m}_p(z) = \left\{-z + \frac{\lambda c_n}{1+\lambda\mathbb{E}\underline{m}_p(z)} + \frac{A_p(z)}{\mathbb{E}\underline{m}_p(z)}\right\}^{-1},$$

where $A_p(z) := \frac{c_n}{1+\lambda\mathbb{E}\underline{m}_p(z)} + zc_n\mathbb{E}m_p(z)$. From this equation and the identity $\underline{m}_p^0 = (-z + \frac{\lambda c_n}{1+\lambda\underline{m}_p^0})^{-1}$, we get

$$\mathbb{E}\underline{m}_p(z) - \underline{m}_p^0(z) = -\underline{m}_p^0(z)A_p(z)\left[1 - \frac{c_n\lambda^2\underline{m}_p^0(z)\mathbb{E}\underline{m}_p(z)}{\{1+\lambda\mathbb{E}\underline{m}_p(z)\}\{1+\lambda\underline{m}_p^0(z)\}}\right]^{-1}. \qquad (22)$$

Note that $\mathbb{E}\underline{m}_p(z) \to \underline{m}(z)$, $\underline{m}_p^0(z) \to \underline{m}(z)$. It suffices to derive the limit of $nA_p(z)$, which is provided in the following lemma.

Lemma 5.8. *Under Assumptions* **??** *and* **??**, *as $n \to \infty$, we have*

$$nA_p(z) \to -\frac{\underline{m}(z)\{z(\lambda+h_1)m(z)+\lambda\}}{1+\lambda\underline{m}(z)} - \frac{cz^2\underline{m}^2(z)\{(\alpha_1+\alpha_2)m^2(z)+2\lambda^2 m'(z)\}}{1+\lambda\underline{m}(z)} + \frac{c\lambda^2\underline{m}^2(z)}{\{1+\lambda\underline{m}(z)\}[\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)]}.$$

The proof of Lemma **??** is postponed to supplementary material. By (**??**) and Lemma **??**, we have $M_p^{(2)}(z) \to$ (**??**) as $n \to \infty$. Combining two parts above yields Lemma **??**.

References

Zhidong Bai and Jack W. Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probability*, 32(1A):553–605, 2004. doi: 10.1214/aop/1078415845.

Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer, 2nd edition, 2010.

Zhidong Bai and Yongquan Yin. Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. *The Annals of Probability*, 21(3):1275–1294, 1993. doi: 10.1214/aop/1176989118.

Zhidong Bai, Baiqi Miao, and Guangming Pan. On asymptotics of eigenvectors of large sample covariance matrix. *The Annals of Probability*, 35(4):1532 – 1572, 2007. doi: 10.1214/009117906000001079.

Zhidong Bai, Huiqin Li, and Guangming Pan. Central limit theorem for linear spectral statistics of large dimensional separable sample covariance matrices. *Bernoulli*, 25(3):1838–1869, 2019. doi: 10.3150/18-bej1038.

Zhigang Bao. Tracy-widom limit for kendall's tau. *The Annals of Statistics*, 47(6):3504–3532, 2019. doi: 10.1214/18-aos1786.

Florent Benaych-Georges and Raj Rao Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111:120–135, 2012. doi: 10.1016/j.jmva.2012.04.019.

Patrick Billingsley. *Convergence of probability measures*. New York: Wiley, 1968.

Patrick Billingsley. *Probability and Measure*. New York: Wiley, 1995.

Charles Bordenave, Pietro Caputo, and Djalil Chafaï. Circular law theorem for random markov matrices. *Probability Theory and Related Fields*, 152(3):751–779, 2012. doi: 10.1007/s00440-010-0336-1.

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

Tony Cai, Weidong Liu, and Yin Xia. Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(2):349 – 372, 2014. doi: 10.1111/rssb.12034.

Yuanpei Cao, Wei Lin, and Hongzhe Li. Two-sample tests of high-dimensional means for compositional data. *Biometrika*, 105(1):115 – 132, 2018. doi: 10.1093/biomet/asx060.

Noureddine El Karoui. Tracy–widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *The Annals of Probability*, 35(2):470 – 490, 2007. doi: 10.1214/009117906000000917.

Karoline Faust, Jarupon Fah Sathirapongsasuti, Jacques Izard, N. Segata, Dirk Gevers, Jeroen Raes, and Curtis Huttenhower. Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology*, 8(7):e1002606, 2012. doi: 10.1371/journal.pcbi.1002606.

Jiti Gao, Xiao Han, Guangming Pan, and Yanrong Yang. High dimensional correlation matrices: The central limit theorem and its applications. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3):677–693, 2017. doi: 10.1111/rssb.12189.

Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge University Press, 2012.

Tiefeng Jiang. The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā: The Indian Journal of Statistics*, 66(1):35–48, 2004. doi: 10.2307/25053330.

Iain M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *The Annals of Statistics*, 29(2):295 – 327, 2001. doi: 10.1214/aos/1009210544.

Dag Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *Journal of Multivariate Analysis*, 12(1):1–38, 1982. doi: 10.1016/0047-259x(82)90080-x.

V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457–483, 1967. doi: 10.1070/sm1967v001n04abeh001994.

Jamal Najim and Jianfeng Yao. Gaussian fluctuations for linear spectral statistics of large random covariance matrices. *The Annals of Applied Probability*, 26(3), 2016. doi: 10.1214/15-aap1135.

Jack W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis*, 55(2):331–339, 1995. doi: 10.1006/jmva.1995.1083.

Jack W. Silverstein and Zhidong Bai. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2):175–192, 1995. doi: 10.1006/jmva.1995.1051.

Gary D. Wu, Jun Chen, Christian Hoffmann, Kyle Bittinger, Ying-Yu Chen, Sue A. Keilbaugh, Meenakshi Bewtra, Dan Knights, William A. Walters, Rob Knight, Rohini Sinha, Erin Gilroy, Kernika Gupta, Robert N. Baldassano, Lisa C. Nessel, Hongzhe Li, Frederic D. Bushman, and James D. Lewis. Linking long-term dietary patterns with gut microbial enterotypes. *Science*, 334 (6052):105 – 108, 2011. doi: 10.1126/science.1208344.

Han Xiao and Wang Zhou. Almost sure limit of the smallest eigenvalue of some sample correlation matrices. *Journal of Theoretical Probability*, 23(1):1–20, 2010. doi: 10.1007/s10959-009-0270-2.

Yongquan Yin and Paruchuri R. Krishnaiah. A limit theorem for the eigenvalues of product of two random matrices. *Journal of Multivariate Analysis*, 13(4):489–507, 1983. doi: 10.1016/0047-259x(83)90035-0.

Lixin Zhang. *Spectral analysis of large dimentional random matrices*. Phd thesis, National University of Singapore, 2007.

Shurong Zheng, Zhidong Bai, and Jianfeng Yao. Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *The Annals of Statistics*, 43(2):546–591, 2015. doi: 10.1214/14-AOS1292.

Supplementary Material for "On eigenvalues of sample covariance matrices based on high-dimensional compositional data"

S1 Auxiliary lemmas

This section introduces several auxiliary lemmas used in the technical proofs of our theoretical results. Lemmas **??** – **??** are from existing literature, while Lemma **??** is our original contribution, and its proof is provided in Section **??**.

Lemma S1.1 (Weyl's inequality, Corollary 7.3.5 of **?**). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $p \times n$ *matrices and let* $r = \min\{p, n\}$. *Let* $s_1(\mathbf{A}) \geq \cdots \geq s_r(\mathbf{A})$ *and* $s_1(\mathbf{B}) \geq \cdots \geq s_r(\mathbf{B})$ *be the nonincreasingly ordered singular values of* $\mathbf{A}$ *and* $\mathbf{B}$, *respectively. Then*

$$\max_{1\leq i\leq r} |s_i(\mathbf{A}) - s_i(\mathbf{B})| \leq \|\mathbf{A} - \mathbf{B}\|,$$

where $\|\mathbf{A} - \mathbf{B}\|$ *denotes the spectral norm of* $\mathbf{A} - \mathbf{B}$.

Lemma S1.2 (Burkholder's inequality, **?**). *Let* $\{X_k\}$ *be a complex martingale difference sequence with respect to the increasing* σ*-field* $\{\mathcal{F}_k\}$, *and let* $\mathbb{E}_k$ *denote conditional expectation with respect to* $\mathcal{F}_k$. *Then, for* $q > 1$,

$$\mathbb{E}\left|\sum X_k\right|^q \leq K_q \left\{\mathbb{E}\left(\sum \mathbb{E}_{k-1}|X_k|^2\right)^{q/2} + \mathbb{E}\sum |X_k|^q\right\}.$$

Lemma S1.3 (Uniform law of large numbers, Lemma 2 of **?**). *Let* $\{X_{ij}, i, j = 1, 2, \ldots\}$ *be a double array of i.i.d. random variables and let* $\alpha > 1/2$, $\beta \geq 0$ *and* $M > 0$ *be constants. Then as* $n \to \infty$,

$$\max_{j\leq Mn^\beta}\left|n^{-\alpha}\sum_{i=1}^{n}(X_{ij} - c)\right| \xrightarrow{a.s.} c,$$

if and only if the following hold:

$$\mathbb{E}|X_{11}|^{(1+\beta)/\alpha} < \infty, \qquad c = \begin{cases} \mathbb{E}X_{11}, & \text{if } \alpha \leq 1, \\ \text{any number}, & \text{if } \alpha > 1. \end{cases}$$

Lemma S1.4 (Martingale CLT, Theorem 35.12 of **?**). *Suppose for each* n, $\{Y_{n1}, Y_{n2}, \ldots, Y_{nr_n}\}$ *is a real martingale difference sequence with respect to the increasing* σ*-field* $\{\mathcal{F}_{nj}\}$ *having second moments. If as* $n \to \infty$,

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^2 | \mathcal{F}_{n,j-1}) \xrightarrow{P} \sigma^2,$$

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

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^2 I_{\{|Y_{nj}|\geq\varepsilon\}}) \to 0,$$

then $\sum_{j=1}^{r_n} Y_{nj} \xrightarrow{D} \mathcal{N}(0, \sigma^2)$.

Lemma S1.5. *Suppose that* $\mathbf{x}_p = \frac{1}{\sqrt{p}}(1, 1, \ldots, 1)'$ *is a p-dimensional normalized all-one vector, then for the truncated random variable satisfying* (**??**), *we have* $\mathbb{E}|\mathbf{x}_p'\mathbf{D}^{-1}(z)\mathbf{x}_p + 1/z|^2 \to 0$.

S2 Proofs

S2.1 Proof of Equation (??)

Recall that $\mathbf{Y}_n = p\mathbf{C}_n\mathbf{\Lambda}_n\mathbf{W}_n/\sqrt{N}$, $\check{\mathbf{Y}}_n = \mathbf{C}_n\mathbf{W}_n/(\sqrt{N}\mu)$, and $\mathbf{B}_{p,N} = \mathbf{Y}_n'\mathbf{Y}_n$. Let $\check{\mathbf{B}}_{p,N} = \check{\mathbf{Y}}_n'\check{\mathbf{Y}}_n$. For any positive constant ε small enough such that

$$\eta_1 - \varepsilon > \lambda(1+\sqrt{c})^2, \qquad \eta_2 + 2\varepsilon < \lambda(1-\sqrt{c})^2 I_{\{0<c<1\}}, \tag{S2.1}$$

we have

$$\begin{aligned}
&\mathbb{P}\Big(\lambda_{\max}(\mathbf{B}_{p,N}) \geq \eta_1\Big)\\
&= \mathbb{P}\Big(\lambda_{\max}(\mathbf{B}_{p,N}) \geq \eta_1, \lambda_{\max}(\check{\mathbf{B}}_{p,N}) \geq \eta_1 - \varepsilon\Big) + \mathbb{P}\Big(\lambda_{\max}(\mathbf{B}_{p,N}) \geq \eta_1, \lambda_{\max}(\check{\mathbf{B}}_{p,N}) < \eta_1 - \varepsilon\Big)\\
&\leq \mathbb{P}\Big(\lambda_{\max}(\check{\mathbf{B}}_{p,N}) \geq \eta_1 - \varepsilon\Big) + \mathbb{P}\Big(|\lambda_{\max}(\mathbf{B}_{p,N}) - \lambda_{\max}(\check{\mathbf{B}}_{p,N})| \geq \varepsilon\Big)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{P}\Big(\lambda_{\min}(\mathbf{B}_{p,N}) \leq \eta_2\Big)\\
&= \mathbb{P}\Big(\lambda_{\min}(\mathbf{B}_{p,N}) \leq \eta_2, \lambda_{\min}(\check{\mathbf{B}}_{p,N}) \leq \eta_2 + \varepsilon\Big) + \mathbb{P}\Big(\lambda_{\min}(\mathbf{B}_{p,N}) \leq \eta_2, \lambda_{\min}(\check{\mathbf{B}}_{p,N}) > \eta_2 + \varepsilon\Big)\\
&\leq \mathbb{P}\Big(\lambda_{\min}(\check{\mathbf{B}}_{p,N}) \leq \eta_2 + \varepsilon\Big) + \mathbb{P}\Big(|\lambda_{\min}(\mathbf{B}_{p,N}) - \lambda_{\min}(\check{\mathbf{B}}_{p,N})| \geq \varepsilon\Big).
\end{aligned}$$

To prove Equation (**??**), it suffices to give the following three estimations:

$$\mathbb{P}\Big(\max_{1\leq i\leq p} |\lambda_i(\mathbf{B}_{p,N}) - \lambda_i(\check{\mathbf{B}}_{p,N})| \geq \varepsilon\Big) = o(n^{-\ell}), \tag{S2.2}$$

$$\mathbb{P}\Big(\lambda_{\max}(\check{\mathbf{B}}_{p,N}) \geq \eta_1 - \varepsilon\Big) = o(n^{-\ell}), \tag{S2.3}$$

$$\mathbb{P}\Big(\lambda_{\min}(\check{\mathbf{B}}_{p,N}) \leq \eta_2 + \varepsilon\Big) = o(n^{-\ell}). \tag{S2.4}$$

The proof of these estimates are provided as follows:
Proof of (**??**): By Lemma **??**, we have

$$\max_{1\leq i\leq p} |\lambda_i(\mathbf{B}_{p,N}) - \lambda_i(\check{\mathbf{B}}_{p,N})| \leq \|\mathbf{B}_{p,N} - \check{\mathbf{B}}_{p,N}\| \leq \|\mathbf{Y}_n - \check{\mathbf{Y}}_n\|^2 + 2\|\mathbf{Y}_n - \check{\mathbf{Y}}_n\|\|\check{\mathbf{Y}}_n\|.$$

Note that $\|\mathbf{Y}_n - \check{\mathbf{Y}}_n\| \leq \|p\mu\mathbf{\Lambda}_n - \mathbf{I}_n\|\|\mathbf{W}_n/(\mu\sqrt{N})\|$. We get from Theorem 2.9 in **?** that $\|\mathbf{W}_n/(\mu\sqrt{N})\|$ is bounded almost surely. In view of the above inequalities and (**??**) (will be proved below), it suffices to show that, for any $\ell > 0$ and $\varepsilon > 0$, $\mathbb{P}(\|p\mu\mathbf{\Lambda}_n - \mathbf{I}_n\| \geq \varepsilon) = o(n^{-\ell})$, which is guaranteed by

$$\mathbb{P}\left(\max_{1\leq i\leq n}\left|\frac{\sum_{j=1}^p w_{ij}/p}{\mu} - 1\right| \geq \varepsilon\right) = o(n^{-\ell}). \tag{S2.5}$$

This inequality follows from Equation (B.115) in **?**, and thus we complete the proof of (**??**).
Proof of (**??**): Let $\mathbf{B}^\circ_{p,N} = (\mathbf{Y}^\circ_n)'\mathbf{Y}^\circ_n$, where $\mathbf{Y}^\circ_n = \frac{\mathbf{W}_n - \mathbb{E}\mathbf{W}_n}{\sqrt{N}\mu}$. From **?**, we have

$$\mathbb{P}\Big(\lambda_{\max}(\mathbf{B}^\circ_{p,N}) \geq \eta_1 - \varepsilon\Big) = o(n^{-\ell}), \tag{S2.6}$$

$$\mathbb{P}\Big(\lambda_{\min}(\mathbf{B}^\circ_{p,N}) \leq \eta_2 + \varepsilon\Big) = o(n^{-\ell}). \tag{S2.7}$$

By the identity $\check{\mathbf{Y}}_n = \mathbf{C}_n \mathbf{Y}_n^\circ$, we have

$$\check{\mathbf{B}}_{p,N} = \mathbf{B}_{p,N}^\circ - \frac{1}{n}(\mathbf{Y}_n^\circ)'\mathbf{1}_n\mathbf{1}_n'\mathbf{Y}_n^\circ. \tag{S2.8}$$

This, together with Cauchy interlacing theorem, implies that

$$\lambda_1(\mathbf{B}_{p,N}^\circ) \geq \lambda_1(\check{\mathbf{B}}_{p,N}) \geq \lambda_2(\mathbf{B}_{p,N}^\circ) \geq \lambda_2(\check{\mathbf{B}}_{p,N}) \geq \cdots \geq \lambda_p(\mathbf{B}_{p,N}^\circ) \geq \lambda_p(\check{\mathbf{B}}_{p,N}). \tag{S2.9}$$

For the largest eigenvalue, we have

$$\lambda_{\max}(\mathbf{B}_{p,N}^\circ) = \lambda_1(\mathbf{B}_{p,N}^\circ) \geq \lambda_1(\check{\mathbf{B}}_{p,N}) = \lambda_{\max}(\check{\mathbf{B}}_{p,N}),$$

which, together with (**??**), implies (**??**).

Proof of (**??**): When $p \geq n$, the smallest eigenvalue of $\check{\mathbf{B}}_{p,N}$ is its $(n-1)$-th largest eigenvalue. By using (**??**), we have

$$\lambda_{\min}(\check{\mathbf{B}}_{p,N}) = \lambda_{n-1}(\check{\mathbf{B}}_{p,N}) \geq \lambda_n(\mathbf{B}_{p,N}^\circ) = \lambda_{\min}(\mathbf{B}_{p,N}^\circ),$$

which, together with (**??**), implies (**??**). When $p < n$, the smallest eigenvalue of $\check{\mathbf{B}}_{p,N}$ is its p-th largest eigenvalue, and all eigenvalues of $\check{\mathbf{B}}_{p,N}$ and $\mathbf{B}_{p,N}^\circ$ are interlaced each other as in (**??**). From (**??**), we have

$$\operatorname{tr}(\mathbf{B}_{p,N}^\circ) = \operatorname{tr}(\check{\mathbf{B}}_{p,N}) + \frac{n}{N}\sum_{j=1}^{p}\Delta_{j,n}^2, \qquad \Delta_{j,n} = \frac{\sum_{i=1}^n w_{ij}/n}{\mu} - 1.$$

Hence, there exists some constant C such that $\lambda_{\min}(\check{\mathbf{B}}_{p,N}) = \lambda_{\min}(\mathbf{B}_{p,N}^\circ) - \frac{C}{p}\frac{n}{N}\sum_{j=1}^p \Delta_{j,n}^2$, and thus

$$\begin{aligned}
&\Pr\Big(\lambda_{\min}(\check{\mathbf{B}}_{p,N}) \leq \eta_2 + \varepsilon\Big) \\
=&\Pr\left(\lambda_{\min}(\check{\mathbf{B}}_{p,N}) \leq \eta_2 + \varepsilon,\ \frac{C}{p}\frac{n}{N}\sum_{j=1}^{p}\Delta_{j,n}^2 \leq \varepsilon\right) + \Pr\left(\lambda_{\min}(\check{\mathbf{B}}_{p,N}) \leq \eta_2 + \varepsilon,\ \frac{C}{p}\frac{n}{N}\sum_{j=1}^{p}\Delta_{j,n}^2 \geq \varepsilon\right) \\
\leq&\Pr\Big(\lambda_{\min}(\mathbf{B}_{p,N}^\circ) \leq \eta_2 + 2\varepsilon\Big) + \Pr\Big(\max_{1\leq j\leq p}|\Delta_{j,n}|^2 \geq \varepsilon/C\Big).
\end{aligned} \tag{S2.10}$$

From (**??**) and **?**, the first term in (**??**) is of order $o(n^{-\ell})$ for any $\ell > 0$. Similar to (**??**), for any $\ell > 0$ and $\varepsilon > 0$, we have $\mathbb{P}(\max_{1\leq j\leq p}|\Delta_{j,n}|^2 \geq \varepsilon) = o(n^{-\ell})$. Therefore, we conclude that (**??**) still holds true when $p < n$.

S2.2 Proof of Lemma ??

The proof of this lemma is quite similar to Sections 5.3.1, 5.3.2, and 5.5 of **?**, it is then omitted. For readers' convenience, we present the outline of the proof for this Lemma. In this situation, $\mathbf{B}_p^0 = \frac{1}{n}\mathbf{Y}_n'\mathbf{Y}_n = \sum_{i=1}^n \mathbf{r}_i\mathbf{r}_i'$, $\mathbf{r}_i = \frac{1}{\sqrt{n}}(\frac{w_{i1}}{\overline{w}_i} - 1, \ldots, \frac{w_{ip}}{\overline{w}_i} - 1)' = \frac{1}{\sqrt{n}}(y_{i1}, \ldots, y_{ip})'$. As for moments of y_{ij}, by (**??**), for any $q > 0$, we have

$$\mathbb{E}y_{ij}^q = \mathbb{E}\left[\left(\frac{w_{ij}}{\overline{w}_i} - 1\right)^q I_{B_p(\varepsilon)}\right] + \mathbb{E}\left[\left(\frac{w_{ij}}{\overline{w}_i} - 1\right)^q I_{B_p^c(\varepsilon)}\right] \leq K\mathbb{E}\Big(w_{ij} - \mu\Big)^q.$$

Therefore, in the following proof, the requirement of truncation of y_{ij} reduces to truncation of w_{ij}. First, we get that

$$\begin{aligned}
&\operatorname{tr}(\mathbf{B}_{p,N}-z\mathbf{I}_p)^{-1}-pm_N^0(z)\\
&=\operatorname{tr}\{\mathbf{A}^{-1}(z)\}-pm_n^0(z)+p\{m_n^0(z)-m_N^0(z)\}+\operatorname{tr}\{\mathbf{A}^{-2}(z)\boldsymbol{\Delta}\}\\
&\quad+\operatorname{tr}\Big[\mathbf{A}^{-1}(z)\{\boldsymbol{\Delta}\mathbf{A}^{-1}(z)\}^2\Big]+\operatorname{tr}\Big[\{\mathbf{A}(z)-\boldsymbol{\Delta}\}^{-1}\{\boldsymbol{\Delta}\mathbf{A}^{-1}(z)\}^3\Big],
\end{aligned}$$

where $\mathbf{A}(z)=\mathbf{B}_p^0-z\mathbf{I}_p$ and $\boldsymbol{\Delta}=\mathbf{B}_p^0-\mathbf{B}_{p,N}$. Moreover, after truncation and normalization, for every $z\in\mathbb{C}^+=\{z:\Im z>0\}$,

$$p\{m_n^0(z)-m_N^0(z)\}=\{1+z\underline{m}(z)\}\frac{\underline{m}(z)+z\underline{m}'(z)}{z\underline{m}(z)}+o_P(1),\quad \operatorname{tr}\{\mathbf{A}^{-2}(z)\boldsymbol{\Delta}\}=o_P(1), \tag{S2.11}$$

$$\operatorname{tr}\{\mathbf{A}^{-2}(z)\boldsymbol{\Delta}\mathbf{A}^{-1}(z)\boldsymbol{\Delta}\}=\{\underline{m}(z)+z\underline{m}'(z)\}\{1+z\underline{m}(z)\}+o_P(1), \tag{S2.12}$$

$$\operatorname{tr}\Big[\{\boldsymbol{\Delta}\mathbf{A}^{-1}(z)\}^3\{\mathbf{A}(z)-\boldsymbol{\Delta}\}^{-1}\Big]=\frac{\{1+z\underline{m}(z)\}^2\{\underline{m}(z)+z\underline{m}'(z)\}}{-z\underline{m}(z)}+o_P(1). \tag{S2.13}$$

Note that, we also need to check the tightness of $\operatorname{tr}(\mathbf{B}_{p,N}-z\mathbf{I}_p)^{-1}-pm_N^0(z)$. Since

$$\begin{aligned}
&\operatorname{tr}(\mathbf{B}_{p,N}-z\mathbf{I}_p)^{-1}-pm_N^0(z)\\
&=\operatorname{tr}(\mathbf{B}_{p,N}-z\mathbf{I}_p)^{-1}-\operatorname{tr}\{\mathbf{A}^{-1}(z)\}+\operatorname{tr}\{\mathbf{A}^{-1}(z)\}-pm_n^0(z)+p\{m_n^0(z)-m_N^0(z)\},
\end{aligned}$$

and the tightness of $\operatorname{tr}\{\mathbf{A}^{-1}(z)\}-pm_n^0(z)$ is proved in Step 2 of Section **??**, it suffices to prove tightness of $\operatorname{tr}(\mathbf{B}_{p,N}-z\mathbf{I}_p)^{-1}-\operatorname{tr}\{\mathbf{A}^{-1}(z)\}$. It can be obtained from similar arguments in Section 5.3.2 of **?** and we omit the details. Finally, the proof is completed.

S2.3 Proof of Lemma ??

Note that, by Taylor expansion, there exist $C_1>0$ such that, for any $-1/2\le x\le 1/2$,

$$\frac{1}{(1+x)^2}=1-2x+3x^2+a(x),\quad |a(x)|\le C_1x^3.$$

Hence, there exist $C_1>0$ such that, for any $0<\varepsilon<1/2$, on the event $B_p(\varepsilon)=\{\omega:|\overline{w}-\mu|\le\varepsilon,\overline{w}=\sum_{j=1}^p w_j/p\}$,

$$\frac{1}{\overline{w}^2}=\frac{1}{\mu^2(\frac{\overline{w}-\mu}{\mu}+1)^2}=\frac{1}{\mu^2}\bigg[1-\frac{2(\overline{w}-\mu)}{\mu}+\frac{3(\overline{w}-\mu)^2}{\mu^2}+a\Big(\frac{\overline{w}-\mu}{\mu}\Big)\bigg],$$

where $|a((\overline{w}-\mu)/\mu)|<C_1\varepsilon^3$. Hence, we have

$$\frac{w_1^2}{\overline{w}^2}I_{B_p(\varepsilon)}=\left\{\frac{w_1^2}{\mu^2}-\frac{2w_1^2(\overline{w}-\mu)}{\mu^3}+\frac{3w_1^2(\overline{w}-\mu)^2}{\mu^4}\right\}I_{B_p(\varepsilon)}+\frac{w_1^2}{\mu^2}a\Big(\frac{\overline{w}-\mu}{\mu}\Big)I_{B_p(\varepsilon)}. \tag{S2.14}$$

This, together with the fact $I_{B_p(\varepsilon)}=1-I_{B_p^c(\varepsilon)}$, implies that

$$\begin{aligned}
\frac{w_1^2}{\overline{w}^2}-\frac{w_1^2}{\mu^2}=&-\frac{2w_1^2(\overline{w}-\mu)}{\mu^3}+\frac{3w_1^2(\overline{w}-\mu)^2}{\mu^4}\\
&-\left\{\frac{w_1^2}{\mu^2}-\frac{2w_1^2(\overline{w}-\mu)}{\mu^3}+\frac{3w_1^2(\overline{w}-\mu)^2}{\mu^4}\right\}I_{B_p^c(\varepsilon)}+a+\frac{w_1^2}{\overline{w}^2}I_{B_p^c(\varepsilon)},
\end{aligned} \tag{S2.15}$$

where $|a| \le C_1 \frac{w_1^2}{\mu^2}\varepsilon^3$. Taking expectation for (**??**) yields that

$$\mathbb{E}\frac{w_1^2}{\overline{w}^2} - \mathbb{E}\frac{w_1^2}{\mu^2} = -\frac{2\mathbb{E}w_1^2(\overline{w}-\mu)}{\mu^3} + \frac{3\mathbb{E}w_1^2(\overline{w}-\mu)^2}{\mu^4} - \mathbb{E}b + \mathbb{E}a + \mathbb{E}c. \tag{S2.16}$$

where $b := \{w_1^2/\mu^2 - 2w_1^2(\overline{w}-\mu)/\mu^3 + 3w_1^2(\overline{w}-\mu)^2/\mu^4\} I_{B_p^c(\varepsilon)}$ and $c := (w_1^2/\overline{w}^2) I_{B_p^c(\varepsilon)}$. Note that

$$|\mathbb{E}c| \le p^2 \mathbb{P}(B_p^c(\varepsilon)). \tag{S2.17}$$

Next, we bound $\mathbb{E}b$. In save of notation, we denote by

$$y_1 = \frac{w_1^2}{\mu^2}, \qquad y_2 = \frac{2w_1^2(\overline{w}-\mu)}{\mu^3}, \qquad y_3 = \frac{3w_1^2(\overline{w}-\mu)^2}{\mu^4}.$$

It is obvious that

$$\left|\mathbb{E}\Big(y_1 I_{B_p^c(\varepsilon)}\Big)\right| \le C_2 \left|\mathbb{E}w_1^2 I_{B_p^c(\varepsilon)}\right| \le C_2 \mathbb{P}^{1/2}(B_p^c(\varepsilon)). \tag{S2.18}$$

Note that

$$\frac{\mathbb{E}w_1^2(\overline{w}-\mu)}{\mu^3} = \frac{1}{p}\frac{\mathbb{E}w_1^2(w_1-\mu)}{\mu^3}, \tag{S2.19}$$

$$\frac{\mathbb{E}w_1^2(\overline{w}-\mu)^2}{\mu^4} = \frac{1}{p^2}\frac{\mathbb{E}w_1^2(w_1-\mu)^2}{\mu^4} + \frac{p-1}{p^2}\frac{\mathbb{E}w_1^2\mathbb{E}(w_1-\mu)^2}{\mu^4}. \tag{S2.20}$$

By (**??**), we get

$$\begin{aligned}
\left|\mathbb{E}\Big(y_2 I_{B_p^c(\varepsilon)}\Big)\right| &= \frac{2}{p\mu^3}\left|\mathbb{E}w_1^2(w_1-\mu) I_{B_p^c(\varepsilon)}\right| \\
&\le \frac{C_3}{p}\Big(\left|\mathbb{E}w_1^2 I_{B_p^c(\varepsilon)}\right| + \left|\mathbb{E}w_1^3 I_{B_p^c(\varepsilon)}\right|\Big) \\
&\le \frac{C_3}{p}\Big\{\mathbb{P}^{1/2}(B_p^c(\varepsilon)) + \left|(\mathbb{E}w_1^4)^{1/2}\Big(\mathbb{E}w_1^2 I_{B_p^c(\varepsilon)}\Big)^{1/2}\right|\Big\} \\
&\le \frac{C_3}{p}\Big\{\mathbb{P}^{1/2}(B_p^c(\varepsilon)) + \mathbb{P}^{1/4}(B_p^c(\varepsilon))\Big\}.
\end{aligned} \tag{S2.21}$$

By (**??**), we get

$$\begin{aligned}
\left|\mathbb{E}\Big(y_3 I_{B_p^c(\varepsilon)}\Big)\right| &\le \frac{1}{p^2\mu^4}\left|\mathbb{E}w_1^2(w_1-\mu)^2 I_{B_p^c(\varepsilon)}\right| + \frac{p-1}{p^2\mu^4}\left|\Big\{\mathbb{E}w_1^2 I_{B_p^c(\varepsilon)}\Big\}\Big\{\mathbb{E}(w_1-\mu)^2 I_{B_p^c(\varepsilon)}\Big\}\right| \\
&\le \frac{C_4}{p^2} + \frac{C_4}{p}\Big\{\mathbb{P}^{1/2}(B_p^c(\varepsilon)) \cdot \mathbb{P}^{1/2}(B_p^c(\varepsilon))\Big\} \\
&\le C_4\Big\{\frac{1}{p^2} + \frac{1}{p}\mathbb{P}(B_p^c(\varepsilon))\Big\}.
\end{aligned} \tag{S2.22}$$

By (**??**) – (**??**), we have

$$|\mathbb{E}b| \le C_5\Big\{\mathbb{P}^{1/2}(B_p^c(\varepsilon)) + \frac{1}{p}\mathbb{P}^{1/4}(B_p^c(\varepsilon)) + \frac{1}{p^2}\Big\}. \tag{S2.23}$$

By using (**??**) and $\mathbb{E}\left|w_1-\mu\right|^4<\infty$ and $|\frac{w_j-\mu}{\sigma}|\leq\sqrt{n}\delta_n$, we have $\mathbb{E}a-\mathbb{E}b+\mathbb{E}c=o(p^{-1})$. From (**??**), (**??**) and (**??**), we get

$$\mathbb{E}\frac{w_1^2}{\overline{w}^2}-\mathbb{E}\frac{w_1^2}{\mu^2}=-\frac{2\mathbb{E}w_1^2(\overline{w}-\mu)}{\mu^3}+\frac{3\mathbb{E}w_1^2(\overline{w}-\mu)^2}{\mu^4}+o(p^{-1}). \tag{S2.24}$$

Plugging (**??**) – (**??**) into (**??**), we get

$$\mathbb{E}\frac{w_1^2}{\overline{w}^2}-\mathbb{E}\frac{w_1^2}{\mu^2}=\frac{h_1}{p}+o(p^{-1}),$$

which implies the first equation in Lemma **??**.

Similar to the previous calculation, we obtain

$$\begin{aligned}&\left(\frac{w_1}{\overline{w}}-1\right)^2\left(\frac{w_2}{\overline{w}}-1\right)^2 I_{B_p(\varepsilon)}\\ =&\left(\frac{w_1}{\mu}-1\right)^2\left(\frac{w_2}{\mu}-1\right)^2 I_{B_p(\varepsilon)}+\left(\frac{\overline{w}}{\mu}-1\right)f_1(w_1,w_2,\mu)I_{B_p(\varepsilon)}+\left(\frac{\overline{w}}{\mu}-1\right)^2 f_2(w_1,w_2,\mu)I_{B_p(\varepsilon)}+o(p^{-1}),\end{aligned}$$

where

$$\begin{aligned}f_1(w_1,w_2,\mu)&=-\frac{2w_1}{\mu}\left(\frac{w_1}{\mu}-1\right)\left(\frac{w_2}{\mu}-1\right)^2-\frac{2w_2}{\mu}\left(\frac{w_2}{\mu}-1\right)\left(\frac{w_1}{\mu}-1\right)^2,\\ f_2(w_1,w_2,\mu)&=4\frac{w_1}{\mu}\frac{w_2}{\mu}\left(\frac{w_1}{\mu}-1\right)\left(\frac{w_2}{\mu}-1\right)\\ &\quad+\left(\frac{w_1}{\mu}-1\right)^2\left(\frac{w_2}{\mu}-1\right)^2\left\{\frac{w_1^2/\mu^2}{(w_1/\mu-1)^2}+\frac{2w_1/\mu}{w_1/\mu-1}\right\}\\ &\quad+\left(\frac{w_1}{\mu}-1\right)^2\left(\frac{w_2}{\mu}-1\right)^2\left\{\frac{w_2^2/\mu^2}{(w_2/\mu-1)^2}+\frac{2w_2/\mu}{w_2/\mu-1}\right\}.\end{aligned} \tag{S2.25}$$

Similar to (**??**) – (**??**), we get

$$\begin{aligned}&\mathbb{E}\left(\frac{w_1}{\overline{w}}-1\right)^2\left(\frac{w_2}{\overline{w}}-1\right)^2\\ =&\,\mathbb{E}\left(\frac{w_1}{\mu}-1\right)^2\left(\frac{w_2}{\mu}-1\right)^2+\mathbb{E}\left(\frac{\overline{w}}{\mu}-1\right)f_1(w_1,w_2,\mu)+\mathbb{E}\left(\frac{\overline{w}}{\mu}-1\right)^2 f_2(w_1,w_2,\mu)+o(p^{-1})\\ =:&\,T_1+T_2+T_3+o(p^{-1}).\end{aligned} \tag{S2.26}$$

Similar to (**??**) and (**??**), we obtain

$$T_2=\frac{1}{p}\mathbb{E}\left(\frac{w_1}{\mu}+\frac{w_2}{\mu}-2\right)f_1(w_1,w_2,\mu) \tag{S2.27}$$

and

$$T_3=\frac{1}{p^2}\sum_{i=1}^{2}\mathbb{E}\left(\frac{w_i}{\mu}-1\right)^2 f_2(w_1,w_2,\mu)+\frac{p-2}{p^2}\mathbb{E}\left(\frac{w_1}{\mu}-1\right)^2\mathbb{E}f_2(w_1,w_2,\mu). \tag{S2.28}$$

Thus, by (**??**) – (**??**), we get

$$\mathbb{E}\left(\frac{w_1}{\overline{w}}-1\right)^2\left(\frac{w_2}{\overline{w}}-1\right)^2=\mathbb{E}\left(\frac{w_1}{\mu}-1\right)^2\left(\frac{w_2}{\mu}-1\right)^2+\frac{1}{p}\mathbb{E}\left(\frac{w_1}{\mu}+\frac{w_2}{\mu}-2\right)f_1(w_1,w_2,\mu)$$

$$
\begin{aligned}
&+\frac{\lambda}{p}\mathbb{E}f_2(w_1,w_2,\mu)+o(p^{-1})\\
&=\mathbb{E}\Big(\frac{w_1}{\mu}-1\Big)^2\Big(\frac{w_2}{\mu}-1\Big)^2+\frac{1}{p}h_2+o(p^{-1}).
\end{aligned} \tag{S2.29}
$$

which is the second equation in Lemma **??**.

Similarly, we get

$$
\mathbb{E}\Big(\frac{w_1}{\overline{w}}-1\Big)^4=\mathbb{E}\Big(\frac{w_1}{\mu}-1\Big)^4+o(1),
$$

which is the third equation in Lemma **??**.

S2.4 Proof of Lemma ??

First, we prove the estimation of $\mathbb{P}(B_p^c(\varepsilon))$. By Markov's inequality and Burkhölder inequality, we get

$$
\begin{aligned}
\mathbb{P}(B_p^c(\varepsilon))&=\mathbb{P}(|\overline{w}-\mu|\geq\varepsilon)\\
&\leq\varepsilon^{-kq_1}\mathbb{E}\Big|\frac{1}{p}\sum_{j=1}^{p}(w_j-\mu)\Big|^{kq_1}\\
&\leq K_{kq_1}\varepsilon^{-kq_1}p^{-kq_1}\Big\{\mathbb{E}\Big(\sum_{j=1}^{p}\mathbb{E}_{j-1}|w_j-\mu|^2\Big)^{kq_1/2}+\mathbb{E}\sum_{j=1}^{p}|w_j-\mu|^{kq_1}\Big\}\\
&=K_{kq_1}\varepsilon^{-kq_1}p^{-kq_1}\Big\{\Big(\sum_{j=1}^{p}\mathbb{E}|w_j-\mu|^2\Big)^{kq_1/2}+\mathbb{E}\sum_{j=1}^{p}|w_j-\mu|^{kq_1}\Big\}\\
&=K_{kq_1}\varepsilon^{-kq_1}p^{-kq_1}\Big\{(p\sigma^2)^{kq_1/2}+p\mathbb{E}|w_1-\mu|^{kq_1}\Big\}\\
&=K_{kq_1}\sigma^{kq_1}\varepsilon^{-kq_1}\Big(p^{-kq_1/2}+p^{-kq_1+1}\mathbb{E}\Big|\frac{w_j-\mu}{\sigma}\Big|^{kq_1}\Big).
\end{aligned} \tag{S2.30}
$$

Next, we prove the estimation of the q-th moment of $\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\nu_2}{n}\mathrm{tr}\mathbf{A}$. For any $q\geq 2$,

$$
\mathbb{E}\Big|\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big|^q\leq K_q\Big(\mathbb{E}\Big|\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\Big|^q+\mathbb{E}\Big|\frac{\lambda}{n}\mathrm{tr}\mathbf{A}-\frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big|^q\Big).
$$

By Lemma **??**, we have

$$
\mathbb{E}\Big|\frac{\lambda}{n}\mathrm{tr}\mathbf{A}-\frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big|^q\leq K_qp^{-q}\|\mathbf{A}\|^qh_1^q. \tag{S2.31}
$$

Now, we consider $\mathbb{E}\Big|\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\Big|^q$. There exists a positive constant K_q such that

$$
\mathbb{E}\Big|\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\Big|^q\leq K_q\Big\{\mathbb{E}\Big|\Big(\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\Big)I_{B_p^c(\varepsilon)}\Big|^q+\mathbb{E}\Big|\Big(\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\Big)I_{B_p(\varepsilon)}\Big|^q\Big\}. \tag{S2.32}
$$

<u>Estimating $\mathbb{E}|(\mathbf{r}'\mathbf{A}\mathbf{r}-n^{-1}\lambda\mathrm{tr}\mathbf{A})I_{B_p^c(\varepsilon)}|^q$</u>: Since $\Big|\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\Big|\leq\|\mathbf{r}\|\|\mathbf{A}\|+C\|\mathbf{A}\|\leq C\cdot p\|\mathbf{A}\|$, we have

$$
\mathbb{E}\Big|\Big(\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\Big)I_{B_p^c(\varepsilon)}\Big|^q\leq Cp^q\|\mathbf{A}\|^q\mathbb{P}(B_p^c(\varepsilon)). \tag{S2.33}
$$

Estimating $\mathbb{E}|(\mathbf{r}'\mathbf{A}\mathbf{r}-n^{-1}\lambda\mathrm{tr}\mathbf{A})I_{B_p(\varepsilon)}|^q$: Write

$$\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}=\frac{\sigma^2}{n\overline{w}^2}\left\{\frac{(\mathbf{w}-\overline{w}\mathbf{1}_p)'\mathbf{A}(\mathbf{w}-\overline{w}\mathbf{1}_p)}{\sigma^2}-\mathrm{tr}\mathbf{A}\right\}+\frac{\sigma^2/\overline{w}^2-\lambda}{n}\mathrm{tr}\mathbf{A}=:v_1+v_2, \tag{S2.34}$$

where $\mathbf{1}_p=(1,1,\dots,1)'\in\mathbb{R}^p$. For $0<\varepsilon<1/2$, there exists a positive constant K_q such that

$$\mathbb{E}\left|\left(\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\right)I_{B_p(\varepsilon)}\right|^q\le K_q\left(\mathbb{E}|v_1I_{B_p(\varepsilon)}|^q+\mathbb{E}|v_2I_{B_p(\varepsilon)}|^q\right).$$

On the event $B_p(\varepsilon)$, we have $-\varepsilon\le\overline{w}-\mu\le\varepsilon$, and

$$\begin{aligned}\mathbb{E}|v_2I_{B_p(\varepsilon)}|^q&=\sigma^{2q}\left|\frac{\mathrm{tr}\mathbf{A}}{n}\right|^q\mathbb{E}\left|\left(\frac{1}{\overline{w}^2}-\frac{1}{\mu^2}\right)I_{B_p(\varepsilon)}\right|^q\\&\le\sigma^{2q}\|\mathbf{A}\|^q\mathbb{E}\left|\frac{(\overline{w}-\mu)(\overline{w}+\mu)}{\overline{w}^2\mu^2}I_{B_p(\varepsilon)}\right|^q\\&\le K\|\mathbf{A}\|^q\mathbb{E}|\overline{w}-\mu|^q\\&\le K\|\mathbf{A}\|^q\left(p^{-q/2}+p^{-q+1}\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^q\right),\end{aligned} \tag{S2.35}$$

where the last inequality follows from the same argument in the proof of (**??**). By using $\mathbf{w}-\overline{w}\mathbf{1}_p=\mathbf{w}-\mu\mathbf{1}_p-(\overline{w}\mathbf{1}_p-\mu\mathbf{1}_p)$, we get

$$\begin{aligned}\mathbb{E}|v_1I_{B_p(\varepsilon)}|^q&\le Kn^{-q}\mathbb{E}\left|\left\{\frac{(\mathbf{w}-\overline{w}\mathbf{1}_p)'}{\sigma}\mathbf{A}\frac{(\mathbf{w}-\overline{w}\mathbf{1}_p)}{\sigma}-\mathrm{tr}\mathbf{A}\right\}I_{B_p(\varepsilon)}\right|^q\\&\le Kn^{-q}\mathbb{E}\left|\left\{\frac{(\mathbf{w}-\mu\mathbf{1}_p)'}{\sigma}\mathbf{A}\frac{(\mathbf{w}-\mu\mathbf{1}_p)}{\sigma}-\mathrm{tr}\mathbf{A}\right\}I_{B_p(\varepsilon)}\right|^q\\&\quad+Kn^{-q}\mathbb{E}\left|\left\{\frac{(\overline{w}\mathbf{1}_p-\mu\mathbf{1}_p)'}{\sigma}\mathbf{A}\frac{(\mathbf{w}-\mu\mathbf{1}_p)}{\sigma}\right\}I_{B_p(\varepsilon)}\right|^q\\&\quad+Kn^{-q}\mathbb{E}\left|\left\{\frac{(\mathbf{w}-\mu\mathbf{1}_p)'}{\sigma}\mathbf{A}\frac{(\overline{w}\mathbf{1}_p-\mu\mathbf{1}_p)}{\sigma}\right\}I_{B_p(\varepsilon)}\right|^q\\&\quad+Kn^{-q}\mathbb{E}\left|\left\{\frac{(\overline{w}\mathbf{1}_p-\mu\mathbf{1}_p)'}{\sigma}\mathbf{A}\frac{(\overline{w}\mathbf{1}_p-\mu\mathbf{1}_p)}{\sigma}\right\}I_{B_p(\varepsilon)}\right|^q\\&=:Kn^{-q}(V_{11}+V_{12}+V_{13}+V_{14}).\end{aligned} \tag{S2.36}$$

By Lemma 2.2 in **?**, we have

$$V_{11}\le K_q\left[\left\{\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^4\mathrm{tr}(\mathbf{A}\mathbf{A}')\right\}^{q/2}+\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^{2q}\mathrm{tr}(\mathbf{A}\mathbf{A}')^{q/2}\right]. \tag{S2.37}$$

From $(\overline{w}-\mu)\mathbf{1}_p'=\frac{1}{p}\mathbf{1}_p\mathbf{1}_p'(\mathbf{w}-\mu\mathbf{1}_p)$ and $p^{-1}\mathbb{E}\mathrm{tr}(\mathbf{1}_p\mathbf{1}_p'\mathbf{A})=p^{-2}\mathbb{E}\mathrm{tr}(\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\mathbf{1}_p\mathbf{1}_p')\le\|\mathbf{A}\|$, we get

$$\begin{aligned}V_{12}&=\mathbb{E}\left|\left\{\frac{1}{p}\frac{(\mathbf{w}-\mu\mathbf{1}_p)'}{\sigma}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\frac{(\mathbf{w}-\mu\mathbf{1}_p)}{\sigma}\right\}I_{B_p(\varepsilon)}\right|^q\\&\le K_q\left\{\mathbb{E}\left|\frac{(\mathbf{w}-\mu\mathbf{1}_p)'}{\sigma}\left(\frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\right)\frac{(\mathbf{w}-\mu\mathbf{1}_p)}{\sigma}-\mathrm{tr}\left(\frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\right)\right|^q+\mathbb{E}\left|\mathrm{tr}\left(\frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\right)\right|^q\right\}\end{aligned}$$

$$
\begin{aligned}
&\le K_q\left\{\mathbb{E}\left|\frac{(\mathbf{w}-\mu\mathbf{1}_p)'}{\sigma}\left(\frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\right)\frac{(\mathbf{w}-\mu\mathbf{1}_p)}{\sigma}-\mathrm{tr}\left(\frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\right)\right|^q+\|\mathbf{A}\|^q\right\}\\
&\le K_q\left[\left\{\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^4\mathrm{tr}\left(\frac{1}{p^2}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\mathbf{A}'\mathbf{1}_p\mathbf{1}_p'\right)\right\}^{q/2}+\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^{2q}\mathrm{tr}\left(\frac{1}{p^2}\mathbf{1}_p\mathbf{1}_p'\mathbf{A}\mathbf{A}'\mathbf{1}_p\mathbf{1}_p'\right)^{q/2}\right]+K_q\|\mathbf{A}\|^q\\
&\le V_{11}.
\end{aligned}
\tag{S2.38}
$$

Similarly, we get

$$
V_{13}\le V_{11},\qquad V_{14}\le V_{11}. \tag{S2.39}
$$

By (**??**) – (**??**), we get

$$
\mathbb{E}|v_1 I_{B_p(\varepsilon)}|^q\le K_q n^{-q}\left[\left\{\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^4\mathrm{tr}(\mathbf{A}\mathbf{A}')\right\}^{q/2}+\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^{2q}\mathrm{tr}(\mathbf{A}\mathbf{A}')^{q/2}\right]. \tag{S2.40}
$$

From (**??**) – (**??**) and (**??**), we have

$$
\begin{aligned}
&\mathbb{E}\left|\mathbf{r}'\mathbf{A}\mathbf{r}-\frac{\lambda}{n}\mathrm{tr}\mathbf{A}\right|^q\\
&\le K_q\left\{n^{-q}\left[\left\{\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^4\mathrm{tr}(\mathbf{A}\mathbf{A}')\right\}^{q/2}+\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^{2q}\mathrm{tr}(\mathbf{A}\mathbf{A}')^{q/2}\right]\right.\\
&\qquad\left.+\|\mathbf{A}\|^q\left(p^{-q/2}+p^{-q+1}\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^q\right)+p^q\|\mathbf{A}\|^q\mathbb{P}(B_p^c(\varepsilon))\right\}\\
&\le K_q\left(n^{-q}\left[\left\{\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^4\mathrm{tr}(\mathbf{A}\mathbf{A}')\right\}^{q/2}+\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^{2q}\mathrm{tr}(\mathbf{A}\mathbf{A}')^{q/2}\right]+n^q\|\mathbf{A}\|^q\mathbb{P}(B_p^c(\varepsilon))\right).
\end{aligned}
$$

By this inequality, (**??**) and (**??**), we get

$$
\begin{aligned}
\mathbb{E}\left|\mathbf{r}'\mathbf{A}r-\frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\right|^q\le K_q\Bigg(&n^{-q}\left[\left\{\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^4\mathrm{tr}(\mathbf{A}\mathbf{A}')\right\}^{q/2}+\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^{2q}\mathrm{tr}(\mathbf{A}\mathbf{A}')^{q/2}\right]\\
&+n^q\|\mathbf{A}\|^q\mathbb{P}(B_p^c(\varepsilon))+n^{-q}\|\mathbf{A}\|^q h_1^q\Bigg).
\end{aligned}
$$

Finally, we prove the last inequality in Lemma **??**. If $\mathbb{E}\left|w_1-\mu\right|^4<\infty$, $\|\mathbf{A}\|\le K$ and $|\frac{w_j-\mu}{\sigma}|\le\sqrt{n}\delta_n$, then, for any $q\ge 2$,

$$
n^{-q}\left[\left\{\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^4\mathrm{tr}(\mathbf{A}\mathbf{A}')\right\}^{q/2}+\mathbb{E}\left|\frac{w_1-\mu}{\sigma}\right|^{2q}\mathrm{tr}(\mathbf{A}\mathbf{A}')^{q/2}\right]\le K_q n^{-1}\delta_n^{2q-4}
$$

and

$$
n^{-q}\|\mathbf{A}\|^q h_1^q\le K_q n^{-1}\delta_n^{2q-4}.
$$

Taking $\varepsilon=n^{-\alpha}$, $0<\alpha<1/2$, and $kq_1>\frac{3q}{1-2\alpha}$ yields that

$$
n^q\|\mathbf{A}\|^q\mathbb{P}(B_p^c(\varepsilon))\le K_q n^{-1}\delta_n^{2q-4}.
$$

From (**??**), we write

$$
\mathbb{P}(B_p^c(\varepsilon))\le K_{kq_1}\sigma^{kq_1}\varepsilon^{-kq_1}\left(p^{-kq_1/2}+p^{-kq_1+1}\mathbb{E}\left|\frac{w_j-\mu}{\sigma}\right|^{kq_1}\right)=:K_{kq_1}\sigma^{kq_1}(P_1+P_2).
$$

Since $P_2 \le \varepsilon^{-kq_1} p^{-kq_1+1}(n^{1/2}\delta_n)^{kq_1-4} \le \delta_n^{kq_1-4}\varepsilon^{-kq_1} n^{-kq_1/2-1} \le P_1$, we obtain that

$$\mathbb{P}(B_p^c(\varepsilon)) \le 2K_{kq_1}\sigma^{kq_1} P_1. \tag{S2.41}$$

Take $\varepsilon = n^{-\alpha}$ $(0 < \alpha < 1/2)$ and $kq_1 > \frac{3q}{1-2\alpha}$ into (**??**), we have

$$n^q \mathbb{P}(B_p^c(\varepsilon)) \le K_q n^q n^{-(\frac{1}{2}-\alpha)kq_1} \le K_q n^{-q/2} \le K_q n^{-1}\delta_n^{2q-4}.$$

Combining all these estimates, we obtain the last inequality in the Lemma.

S2.5 Proof of Lemma ??

First, we denote $R_j = \frac{1}{\sqrt{n}}\Big(\frac{w_j}{\overline{w}} - 1\Big)$ and derive some identities that will be used in the proof. It is obvious that $\sum_{j=1}^p R_j = 0$ and $\sum_{j=1}^p R_j^2 + \sum_{j_1 \ne j_2} R_{j_1} R_{j_2} = 0$. Since $\{w_j\}_{j=1}^p$ are i.i.d., taking expectation on the above two identities yields that, for any $1 \le i \ne j \le p$,

$$\mathbb{E}R_j = 0, \quad \mathbb{E}R_j^2 = \frac{\nu_2}{n}, \quad \mathbb{E}R_iR_j = -\frac{\nu_2}{n(p-1)}, \quad \mathbb{E}R_j^4 = \frac{\nu_4}{n^2}, \quad \mathbb{E}R_i^2R_j^2 = \frac{\nu_{12}}{n^2}. \tag{S2.42}$$

Recall that $\mathbf{r} = (R_1, \dots, R_p)'$. From (**??**), we have

$$\mathbb{E}(\mathbf{r}\mathbf{r}') = \frac{p\nu_2}{n(p-1)}\left(\mathbf{I}_p - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\right), \tag{S2.43}$$

It is obvious that

$$\begin{aligned}
&\mathbb{E}\Big(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big)\Big(\mathbf{r}'\mathbf{B}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{B}\Big) \\
= &\ \mathbb{E}\Big(\mathbf{r}'\mathbf{A}\mathbf{r}\mathbf{r}'\mathbf{B}\mathbf{r}\Big) - \frac{\nu_2\mathrm{tr}\mathbf{B}}{n}\mathbb{E}\Big(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big) - \frac{\nu_2\mathrm{tr}\mathbf{A}}{n}\mathbb{E}\Big(\mathbf{r}'\mathbf{B}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{B}\Big) - \frac{\nu_2^2}{n^2}\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B}.
\end{aligned} \tag{S2.44}$$

We first estimate each terms in the RHS of the identity above, and finally prove the equation given in Lemma **??**. The details are provided in the following three steps.
Step 1: <u>Estimate the first term.</u> It follows from (**??**) and the identity $\sum_{j=1}^p R_j = 0$ that

$$\mathbb{E}R_1^3R_2 = \frac{1}{p-1}\mathbb{E}\left[R_1^3\Big(\sum_{j=1}^p R_j - R_1\Big)\right] = -\frac{\nu_4}{n^2(p-1)}, \tag{S2.45}$$

$$\begin{aligned}
\mathbb{E}R_1^2R_2R_3 &= \frac{1}{p-2}\mathbb{E}\left[R_1^2R_2\Big(\sum_{j=1}^p R_j - R_1 - R_2\Big)\right] \\
&= -\frac{1}{p-2}\mathbb{E}(R_1^3R_2) - \frac{1}{p-2}\mathbb{E}(R_1^2R_2^2) \\
&\overset{(??)}{=} -\frac{\nu_4}{n^2(p-1)(p-2)} - \frac{\nu_{12}}{n^2(p-2)},
\end{aligned} \tag{S2.46}$$

$$\begin{aligned}
\mathbb{E}R_1R_2R_3R_4 &= \frac{1}{p-3}\mathbb{E}\left[R_1R_2R_3\Big(\sum_{j=1}^p R_j - R_1 - R_2 - R_3\Big)\right] \\
&= -\frac{3}{p-3}\mathbb{E}(R_1^2R_2R_3)
\end{aligned}$$

$$\stackrel{(??)}{=} -\frac{3\nu_4}{n^2(p-1)(p-2)(p-3)} + \frac{3\nu_{12}}{n^2(p-2)(p-3)}. \tag{S2.47}$$

To calculate $\mathbb{E}(\mathbf{r}'\mathbf{A}\mathbf{r}\mathbf{r}'\mathbf{B}\mathbf{r})$, we expand it as

$$\mathbb{E}(\mathbf{r}'\mathbf{A}\mathbf{r}\mathbf{r}'\mathbf{B}\mathbf{r}) = \mathbb{E}\left(\sum_{i,j} R_i A_{ij} R_j \sum_{k,\ell} R_k B_{k\ell} R_\ell\right) = \sum_{i,j,k,\ell} \mathbb{E}(R_i R_j R_k R_\ell A_{ij} B_{k\ell}). \tag{S2.48}$$

To calculate (**??**), we split it into the following 11 cases:

1. $i = j = k = \ell$, $\sum_i (R_i^4) A_{ii} B_{ii}$;
2. $i = j, k = \ell, i \neq k, \sum_{\substack{i,k \\ i\neq k}} (R_i^2 R_k^2) A_{ii} B_{kk}$;
3. $i = j, k \neq \ell, \sum_{\substack{i,k,\ell \\ k\neq \ell}} (R_i^2 R_k R_\ell) A_{ii} B_{k\ell}$;
4. $i \neq j, k = \ell, \sum_{\substack{i,j,k \\ i\neq j}} (R_i R_j R_k^2) A_{ij} B_{kk}$;
5. $i \neq j, k \neq \ell, i = k, j = \ell, \sum_{\substack{i,j \\ i\neq j}} (R_i^2 R_j^2) A_{ij} B_{ij}$;
6. $i \neq j, k \neq \ell, i = \ell, j = k, \sum_{\substack{i,j \\ i\neq j}} (R_i^2 R_j^2) A_{ij} B_{ji}$;
7. $i \neq j, k \neq \ell, i = k, \ell \neq j, \sum_{\substack{i,j,\ell \\ i\neq j\neq \ell}} (R_i^2 R_j R_\ell) A_{ij} B_{i\ell}$;
8. $i \neq j, k \neq \ell, \ell = j, i \neq k, \sum_{\substack{i,j,k \\ i\neq j\neq k}} (R_i R_j^2 R_k) A_{ij} B_{kj}$;
9. $i \neq j, k \neq \ell, k = j, i \neq \ell, \sum_{\substack{i,j,\ell \\ i\neq j\neq \ell}} (R_i R_j^2 R_\ell) A_{ij} B_{j\ell}$;
10. $i \neq j, k \neq \ell, i = \ell, k \neq j, \sum_{\substack{i,j,k \\ i\neq j\neq k}} (R_i^2 R_j R_k) A_{ij} B_{ki}$;
11. $i \neq j, k \neq \ell, \ell \neq j, i \neq k, \sum_{\substack{i,j,k,\ell \\ i\neq j\neq k\neq \ell}} (R_i R_j R_k R_\ell) A_{ij} B_{k\ell}$.

For ease of presentation, we still keep ν_4 in the expectations although we have obtained its value. The expectations of all cases are listed as follows.

Case 1: From (**??**), we have

$$\mathbb{E} \sum_i R_i^4 A_{ii} B_{ii} = \frac{\nu_4}{n^2} \sum_i A_{ii} B_{ii}.$$

Case 2: From (**??**), we have

$$\mathbb{E} \sum_{\substack{i,k \\ i\neq k}} R_i^2 R_k^2 A_{ii} B_{kk} = \frac{\nu_{12}}{n^2} \sum_{\substack{i,k \\ i\neq k}} A_{ii} B_{kk} = \frac{\nu_{12}}{n^2} \left(\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B} - \sum_i A_{ii} B_{ii}\right).$$

Case 3: Note that

$$\mathbb{E}\sum_{\substack{i,k,\ell\\k\neq\ell}} R_i^2R_kR_\ell A_{ii}B_{k\ell} = \mathbb{E}R_1^2R_2R_3\sum_{\substack{i,k,\ell\\i\neq k\neq\ell}} A_{ii}B_{k\ell} + \mathbb{E}R_1^3R_2\Bigg(\sum_{\substack{i,\ell\\\ell\neq i}} A_{ii}B_{i\ell} + \sum_{\substack{i,k\\k\neq i}} A_{ii}B_{ki}\Bigg). \tag{S2.49}$$

Now, we estimate the magnitude of the summation terms on the RHS of (**??**). Recall that for any $p\times p$ matrix, we have $\mathbb{E}\mathrm{tr}(\mathbf{M})\le p\mathbb{E}\|\mathbf{M}\|$ and $\max_{1\le i\le p} M_{ii}\le\|\mathbf{M}\|$. By using these facts and the Hölder's inequality, we obtain that

$$\mathbb{E}\Big|\sum_{i,k,\ell} A_{ii}B_{k\ell}\Big| = \mathbb{E}\Big|(\mathrm{tr}\mathbf{A})\mathbf{1}_p'\mathbf{B}\mathbf{1}_p\Big| \le p\mathbb{E}(\|\mathbf{A}\|\cdot|\mathbf{1}_p'\mathbf{B}\mathbf{1}_p|)\le p^2\mathbb{E}(\|\mathbf{A}\|\cdot\|\mathbf{B}\|) = O(p^2), \tag{S2.50}$$

$$\mathbb{E}\sum_{i,l} A_{ii}B_{\ell\ell} = \mathbb{E}\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B}\le p^2\mathbb{E}(\|\mathbf{A}\|\cdot\|\mathbf{B}\|) = O(p^2), \tag{S2.51}$$

$$\mathbb{E}\sum_{i} A_{ii}B_{ii} \le p\mathbb{E}\|\mathrm{diag}(\mathbf{A})\mathrm{diag}(\mathbf{B})\| \le p\mathbb{E}(\|\mathbf{A}\|\cdot\|\mathbf{B}\|) = O(p). \tag{S2.52}$$

Let $\mathbf{1}_p^i$ be the p-dimensional vector with all components being 0 except for the i-th component being 1, then we have

$$\mathbb{E}\Big|\sum_{i,\ell} A_{ii}B_{i\ell}\Big| = \mathbb{E}\Big|\sum_{i} A_{ii}(\mathbf{1}_p^i)'\mathbf{B}\mathbf{1}_p\Big| \le \mathbb{E}\Big(\sum_i A_{ii}^2\Big)^{1/2}\Big\{\sum_i(\mathbf{1}_p^i)'\mathbf{B}\mathbf{1}_p\mathbf{1}_p'\mathbf{B}\mathbf{1}_p^i\Big\}^{1/2} = O(p^{3/2}). \tag{S2.53}$$

From (**??**) – (**??**), we have

$$\mathbb{E}\sum_{\substack{i,k,\ell\\i\neq k\neq\ell}} A_{ii}B_{k\ell} = \mathbb{E}\Big(\sum_{i,k,\ell} A_{ii}B_{k\ell} - \sum_{i,\ell} A_{ii}B_{i\ell} - \sum_{i,k} A_{ii}B_{ki} - \sum_{i,\ell} A_{ii}B_{\ell\ell} + 2\sum_i A_{ii}B_{ii}\Big) = O(p^2). \tag{S2.54}$$

It follows from (**??**), (**??**), (**??**), (**??**), and (**??**), that

$$\mathbb{E}\sum_{\substack{i,k,\ell\\k\neq\ell}} (R_i^2R_kR_\ell)A_{ii}B_{k\ell} = -\frac{\nu_{12}}{n^2(p-2)}\sum_{\substack{i,k,\ell\\i\neq k\neq\ell}} A_{ii}B_{k\ell} + o(n^{-1}).$$

Case 4: Similarly to Case 3, one can conclude that

$$\mathbb{E}\sum_{\substack{i,j,k\\i\neq j}} R_iR_jR_k^2A_{ij}B_{kk} = -\frac{\nu_{12}}{n^2(p-2)}\sum_{\substack{i,j,k\\i\neq j\neq k}} A_{ij}B_{kk} + o(n^{-1}).$$

Case 5:

$$\mathbb{E}\sum_{\substack{i,j\\i\neq j}} R_i^2R_j^2A_{ij}B_{ij} = \mathbb{E}\left(R_1^2R_2^2\right)\sum_{\substack{i,j\\i\neq j}} A_{ij}B_{ij} = \frac{\nu_{12}}{n^2}\Big\{\mathrm{tr}(\mathbf{A}\mathbf{B}') - \sum_i A_{ii}B_{ii}\Big\}.$$

Case 6:

$$\mathbb{E}\sum_{\substack{i,j\\i\neq j}} R_i^2R_j^2A_{ij}B_{ji} = \mathbb{E}\left(R_1^2R_2^2\right)\sum_{\substack{i,j\\i\neq j}} A_{ij}B_{ji} = \frac{\nu_{12}}{n^2}\Big\{\mathrm{tr}(\mathbf{A}\mathbf{B}) - \sum_i A_{ii}B_{ii}\Big\}.$$

Case 7: By (**??**), we have

$$\mathbb{E}\sum_{\substack{i,j,\ell\\ i\neq j\neq \ell}} R_i^2 R_j R_\ell A_{ij} B_{i\ell} = O(n^{-3}) \times \sum_{\substack{i,j,\ell\\ i\neq j\neq \ell}} A_{ij} B_{i\ell}. \tag{S2.55}$$

Note that, $\mathbf{1}_p'\mathbf{A}'\mathbf{B}\mathbf{1}_p \leq p\|\mathbf{A}\|\cdot\|\mathbf{B}\| = O(p)$, $\mathrm{tr}(\mathbf{A}'\mathbf{B}) \leq p\|\mathbf{A}\|\cdot\|\mathbf{B}\| = O(p)$, and by (**??**) and (**??**), we have

$$\mathbb{E}\sum_{i\neq \ell} A_{ii} B_{i\ell} = \mathbb{E}\sum_{i,\ell} A_{ii} B_{i\ell} - \mathbb{E}\sum_{i} A_{ii} B_{ii} = O(p^{3/2}), \tag{S2.56}$$

thus,

$$\mathbb{E}\sum_{\substack{i,j,\ell\\ i\neq j\neq \ell}} A_{ij} B_{i\ell} = \mathbb{E}\left\{\mathbf{1}_p'\mathbf{A}'\mathbf{B}\mathbf{1}_p - \mathrm{tr}(\mathbf{A}'\mathbf{B}) - \sum_{i\neq \ell} A_{ii} B_{i\ell} - \sum_{i\neq j} A_{ij} B_{ii}\right\} = O(p^{3/2}). \tag{S2.57}$$

It follows from (**??**) and (**??**) that

$$\mathbb{E}\sum_{\substack{i,j,\ell\\ i\neq j\neq \ell}} R_i^2 R_j R_\ell A_{ij} B_{i\ell} = o(n^{-1}).$$

Case 8: Similarly to Case 7, we have

$$\mathbb{E}\sum_{\substack{i,j,k\\ i\neq j\neq k}} R_i R_j^2 R_k A_{ij} B_{kj} = o(n^{-1}).$$

Case 9: Similarly to Case 7, we have

$$\mathbb{E}\sum_{\substack{i,j,\ell\\ i\neq j\neq \ell}} R_i R_j^2 R_\ell A_{ij} B_{j\ell} = o(n^{-1}).$$

Case 10: Similarly to Case 7, we have

$$\mathbb{E}\sum_{\substack{i,j,k\\ i\neq j\neq k}} (R_i^2 R_j R_k) A_{ij} B_{ki} = o(n^{-1}).$$

Case 11: By (**??**), we have

$$\mathbb{E}\sum_{\substack{i,j,k,\ell\\ i\neq j\neq k\neq \ell}} R_i R_j R_k R_\ell A_{ij} B_{k\ell} = O(n^{-4}) \times \sum_{\substack{i,j,k,\ell\\ i\neq j\neq k\neq \ell}} A_{ij} B_{k\ell}.$$

Note that, by (**??**) we have

$$\mathbb{E}\sum_{\substack{i,j,k,\ell\\ i\neq j\neq k\neq \ell}} A_{ij} B_{k\ell} = \mathbb{E}(\mathbf{1}_p'\mathbf{A}\mathbf{1}_p - \mathrm{tr}\mathbf{A})(\mathbf{1}_p'\mathbf{B}\mathbf{1}_p - \mathrm{tr}\mathbf{B}) - \mathbb{E}\sum_{\substack{i,j,\ell\\ i\neq j\neq \ell}} A_{ij} B_{i\ell} - \mathbb{E}\sum_{\substack{i,j,k\\ i\neq j\neq k}} A_{ij} B_{ki} - \mathbb{E}\sum_{\substack{i,j\\ i\neq j}} A_{ij} B_{ij}$$

$$= O(p^2).$$

Thus,

$$\mathbb{E}\sum_{\substack{i,j,k,\ell\\ i\neq j\neq k\neq \ell}} R_iR_jR_kR_\ell A_{ij}B_{k\ell} = o(n^{-1}).$$

Combining (**??**) and Cases 1 – 11 gives us

$$\begin{aligned}\mathbb{E}(\mathbf{r}'\mathbf{A}\mathbf{r}\mathbf{r}'\mathbf{B}\mathbf{r}) =& \frac{\nu_4-3\nu_{12}}{n^2}\sum_{i=1}^{p}A_{ii}B_{ii} + \frac{\nu_{12}}{n^2}\Big\{\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B} + \mathrm{tr}(\mathbf{A}\mathbf{B}') + \mathrm{tr}(\mathbf{A}\mathbf{B})\Big\}\\ &- \frac{\nu_{12}}{n^2(p-2)}\Bigg(\sum_{\substack{i,k,\ell\\ i\neq k\neq \ell}}A_{ii}B_{k\ell} + \sum_{\substack{i,j,k\\ i\neq j\neq k}}A_{ij}B_{kk}\Bigg) + o(n^{-1}).\end{aligned} \tag{S2.58}$$

Step 2: Estimate the second and third terms. From (**??**) and (**??**), we have

$$\begin{aligned}\frac{\nu_2\mathrm{tr}\mathbf{B}}{n}\mathbb{E}\Big(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big) = \frac{\nu_2\mathrm{tr}\mathbf{B}}{n}\Big\{\mathbb{E}\mathrm{tr}(\mathbf{A}\mathbf{r}\mathbf{r}') - \frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big\} &= -\frac{\nu_2^2\mathrm{tr}\mathbf{B}}{n^2(p-1)}\sum_{k\neq\ell}A_{k\ell}\\ &= -\frac{\nu_2^2}{n^2(p-1)}\Bigg(\sum_{\substack{k,\ell\\ k\neq\ell}}B_{kk}A_{k\ell} + \sum_{\substack{k,\ell\\ k\neq\ell}}B_{\ell\ell}A_{k\ell} + \sum_{\substack{i,k,\ell\\ i\neq k\neq\ell}}B_{ii}A_{k\ell}\Bigg)\\ &= -\frac{\nu_2^2}{n^2(p-1)}\sum_{\substack{i,k,\ell\\ i\neq k\neq\ell}}B_{ii}A_{k\ell} + o(n^{-1}).\end{aligned} \tag{S2.59}$$

Similarly, we obtain

$$\frac{\nu_2\mathrm{tr}\mathbf{A}}{n}\mathbb{E}\Big(\mathbf{r}'\mathbf{B}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{B}\Big) = -\frac{\nu_2^2}{n^2(p-1)}\sum_{\substack{i,k,\ell\\ i\neq k\neq\ell}}A_{ii}B_{k\ell} + o(n^{-1}). \tag{S2.60}$$

Step 3: From (**??**) and (**??**) – (**??**), we have

$$\begin{aligned}&\mathbb{E}\Big(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{A}\Big)\Big(\mathbf{r}'\mathbf{B}\mathbf{r} - \frac{\nu_2}{n}\mathrm{tr}\mathbf{B}\Big)\\ =&\frac{\nu_4-3\nu_{12}}{n^2}\sum_{i=1}^{p}A_{ii}B_{ii} + \frac{\nu_{12}}{n^2}\Big\{\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B} + \mathrm{tr}(\mathbf{A}\mathbf{B}') + \mathrm{tr}(\mathbf{A}\mathbf{B})\Big\}\\ &- \frac{\nu_{12}}{n^2(p-2)}\Bigg(\sum_{\substack{i,k,\ell\\ i\neq k\neq \ell}}A_{ii}B_{k\ell} + \sum_{\substack{i,j,k\\ i\neq j\neq k}}A_{ij}B_{kk}\Bigg) + \frac{\nu_2^2}{n^2(p-1)}\Bigg(\sum_{\substack{i,k,\ell\\ i\neq k\neq\ell}}B_{ii}A_{k\ell} + \sum_{\substack{i,k,\ell\\ i\neq k\neq\ell}}A_{ii}B_{k\ell}\Bigg)\\ &- \frac{\nu_2^2}{n^2}\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B} + o(n^{-1})\\ =&\frac{\nu_4-3\nu_{12}}{n^2}\sum_{i=1}^{p}A_{ii}B_{ii} + \frac{\nu_{12}}{n^2}\Big\{\mathrm{tr}(\mathbf{A}\mathbf{B}') + \mathrm{tr}(\mathbf{A}\mathbf{B})\Big\} + \frac{\nu_{12}-\nu_2^2}{n^2}\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B}\end{aligned}$$

$$
+\frac{\nu_2^2-\nu_{12}}{n^2(p-2)}\left(\sum_{\substack{i,k,\ell\\ i\neq k\neq \ell}} A_{ii}B_{k\ell}+\sum_{\substack{i,j,k\\ i\neq j\neq k}} A_{ij}B_{kk}\right)+o(n^{-1})
$$

$$
=\frac{\nu_4-3\nu_{12}}{n^2}\sum_{i=1}^{p}A_{ii}B_{ii}+\frac{\nu_{12}}{n^2}\Big\{\mathrm{tr}(\mathbf{A}\mathbf{B}')+\mathrm{tr}(\mathbf{A}\mathbf{B})\Big\}+\frac{\nu_{12}-\nu_2^2}{n^2}\mathrm{tr}\mathbf{A}\mathrm{tr}\mathbf{B}+o(n^{-1}),
$$

where in the last "=" we use Lemma **??** and Equation (**??**).

S2.6 Proof of Lemma ??

S2.6.1 Limit of $\partial^2\mathcal{Y}_1(z_1,z_2)/(\partial z_1\partial z_2)$

Since $\mathbb{E}\left|\overline{\beta}_j(z)-b_p(z)\right|^2\leq K|z|^4/(nv_0^6)$, it is enough to prove that

$$
\frac{\partial^2}{\partial z_1\partial z_2}\left[b_p(z_1)b_p(z_2)\sum_{j=1}^{n}\mathbb{E}_{j-1}\Big\{\mathbb{E}_j\varepsilon_j(z_1)\mathbb{E}_j\varepsilon_j(z_2)\Big\}\right]\xrightarrow{P}(\textbf{??}). \tag{S2.61}
$$

By Lemma **??**, we have

$$
\text{LHS of } (\textbf{??})=\frac{\partial^2}{\partial z_1\partial z_2}\Big\{\mathcal{Y}_{11}(z_1,z_2)+\mathcal{Y}_{12}(z_1,z_2)+\mathcal{Y}_{13}(z_1,z_2)\Big\}+o_P(1), \tag{S2.62}
$$

where

$$
\mathcal{Y}_{11}(z_1,z_2)=\frac{\nu_4-3\nu_{12}}{n^2}b_p(z_1)b_p(z_2)\sum_{j=1}^{n}\sum_{i=1}^{p}\mathbb{E}_j[\mathbf{D}_j^{-1}(z_1)]_{ii}\mathbb{E}_j[\mathbf{D}_j^{-1}(z_2)]_{ii},
$$

$$
\mathcal{Y}_{12}(z_1,z_2)=\frac{2\nu_{12}}{n^2}b_p(z_1)b_p(z_2)\sum_{j=1}^{n}\mathrm{tr}\Big\{\mathbb{E}_j\mathbf{D}_j^{-1}(z_1)\mathbb{E}_j\mathbf{D}_j^{-1}(z_2)\Big\},
$$

$$
\mathcal{Y}_{13}(z_1,z_2)=\frac{\nu_{12}-\nu_2^2}{n^2}b_p(z_1)b_p(z_2)\sum_{j=1}^{n}\mathrm{tr}\Big\{\mathbb{E}_j\mathbf{D}_j^{-1}(z_1)\Big\}\mathrm{tr}\Big\{\mathbb{E}_j\mathbf{D}_j^{-1}(z_2)\Big\}.
$$

We claim that the following statements hold true as $p\to\infty$ (to be proven later):

$$
\frac{\partial^2\mathcal{Y}_{11}(z_1,z_2)}{\partial z_1\partial z_2}\xrightarrow{P}\frac{c\alpha_1\underline{m}'(z_1)\underline{m}'(z_2)}{\{1+\lambda\underline{m}(z_1)\}^2\{1+\lambda\underline{m}(z_2)\}^2}, \tag{S2.63}
$$

$$
\frac{\partial^2\mathcal{Y}_{12}(z_1,z_2)}{\partial z_1\partial z_2}\xrightarrow{P}\frac{2\underline{m}'(z_1)\underline{m}'(z_2)}{\{\underline{m}(z_1)-\underline{m}(z_2)\}^2}-\frac{2}{(z_1-z_2)^2}, \tag{S2.64}
$$

$$
\frac{\partial^2\mathcal{Y}_{13}(z_1,z_2)}{\partial z_1\partial z_2}\xrightarrow{P}\frac{c\alpha_2\underline{m}'(z_1)\underline{m}'(z_2)}{\{1+\lambda\underline{m}(z_1)\}^2\{1+\lambda\underline{m}(z_2)\}^2}. \tag{S2.65}
$$

By (**??**) – (**??**), we obtain the limit of $\frac{\partial^2}{\partial z_1\partial z_2}\mathcal{Y}_1(z_1,z_2)$.

Now, we provide the proofs of (**??**) – (**??**) as follows:
<u>Proof of (**??**):</u> It is enough to find the limit of

$$
\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{p}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_1)\Big]_{ii}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_2)\Big]_{ii}.
$$

By similar calculation of **?**, we get the following lemma and its proof is postponed to Section **??**.

Lemma S2.1. *Under Assumptions* **??** *and* **??**, *for any* $1 \le j \le n$, *we have*

$$\frac{1}{p}\sum_{i=1}^{p}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_1)\Big]_{ii}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_2)\Big]_{ii} \xrightarrow{P} m(z_1)m(z_2).$$

By $\mathbb{E}|\frac{\nu_2}{n}\mathrm{tr}\mathbf{D}^{-1}(z)-\frac{\nu_2}{n}\mathbb{E}\mathrm{tr}\mathbf{D}^{-1}(z)|^q \le K_q n^{-q/2}v_0^{-q}$, the formula (2.2) of **?**, $\underline{m}_p(z) = -\frac{1}{zn}\sum_{j=1}^{n}\beta_j(z)$, and Lemma **??**, we have

$$|b_p(z)-\mathbb{E}\beta_1(z)| \le \frac{K}{\sqrt{n}},\ \mathbb{E}\beta_1(z) = -z\mathbb{E}\underline{m}_p(z),\ |b_p(z)+z\underline{m}_p^0(z)| \le \frac{K}{\sqrt{n}}. \tag{S2.66}$$

Thus, by (**??**), Lemmas **??** and **??**, we have

$$\mathcal{Y}_{11}(z_1,z_2) \xrightarrow{P} c\alpha_1 z_1 z_2 m(z_1)m(z_2)\underline{m}(z_1)\underline{m}(z_2) = \frac{c\alpha_1\underline{m}(z_1)\underline{m}(z_2)}{\{1+\lambda\underline{m}(z_1)\}\{1+\lambda\underline{m}(z_2)\}},$$

where the equality above follows from $m(z) = -z^{-1}\{1+\lambda\underline{m}(z)\}^{-1}$. Thus, the i.p. limit of $\frac{\partial^2}{\partial z_1\partial z_2}\mathcal{Y}_{11}(z_1,z_2)$ is in (**??**).
Proof of (**??**): By similar calculation of **?**, we get the following lemma and its proof is postponed to Section **??**.

Lemma S2.2. *Under Assumptions* **??** *and* **??**, *for any* $1 \le j \le n$, *we have*

$$\begin{aligned}&\mathrm{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_j^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big]\Bigg[1-\frac{\frac{j-1}{n}c_n\nu_2^2\underline{m}_p^0(z_1)\underline{m}_p^0(z_2)}{\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_1)\}\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_2)\}}\Bigg]\\ &= \frac{nc_n}{z_1z_2}\frac{1}{\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_1)\}\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_2)\}} + O_P(n^{1/2}).\end{aligned}$$

By using (**??**) and this lemma, $\mathcal{Y}_{12}(z_1,z_2)$ can be written as

$$\mathcal{Y}_{12}(z_1,z_2) = \frac{a_p(z_1,z_2)\nu_{12}}{n\nu_2^2}\sum_{j=1}^{n}\frac{2}{1-\frac{j-1}{n}a_p(z_1,z_2)} + O_P(n^{-1/2}),$$

where $a_p(z_1,z_2) = \frac{\nu_2^2 c_n\underline{m}_p^0(z_1)\underline{m}_p^0(z_2)}{\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_1)\}\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_2)\}}$. By Lemma **??**, the limit of $a_p(z_1,z_2)$ is $a(z_1,z_2) = \frac{c\lambda^2\underline{m}(z_1)\underline{m}(z_2)}{\{1+\lambda\underline{m}(z_1)\}\{1+\lambda\underline{m}(z_2)\}}$, and thus the in probability (i.p.) limit of $\frac{\partial^2}{\partial z_1\partial z_2}\mathcal{Y}_{12}(z_1,z_2)$ is (**??**).
Proof of (**??**): We have $\mathbb{E}\left|\frac{1}{p}\mathrm{tr}\mathbb{E}_j\mathbf{D}_j^{-1}(z_1)\frac{1}{p}\mathrm{tr}\mathbb{E}_j\mathbf{D}_j^{-1}(z_2) - m_p^0(z_1)m_p^0(z_2)\right| = o(1)$. By Lemma **??**, we get $\lim_{p\to\infty}p(\nu_{12}-\nu_2^2) = \alpha_2$. This, together with (**??**), implies that

$$\mathcal{Y}_{13}(z_1,z_2) \xrightarrow{P} \frac{c\alpha_2\underline{m}(z_1)\underline{m}(z_2)}{\{1+\lambda\underline{m}(z_1)\}\{1+\lambda\underline{m}(z_2)\}}.$$

Thus, the i.p. limit of $\frac{\partial^2}{\partial z_1\partial z_2}\mathcal{Y}_{13}(z_1,z_2)$ is (**??**).

S2.6.2 Limit of $\partial^2\mathcal{Y}_2(z_1,z_2)/(\partial z_1\partial z_2)$

For any $p\times p$ matrix $\mathbf{A}$, we have

$$\left|\mathrm{tr}\Big\{\mathbf{D}^{-1}(z)-\mathbf{D}_j^{-1}(z)\Big\}\mathbf{A}\right|\le\frac{\|\mathbf{A}\|}{\mathfrak{Im}(z)}. \tag{S2.67}$$

By Lemma **??** and (**??**), we have

$$\mathbb{E}\left|\frac{\nu_2}{n}\mathrm{tr}\mathbf{D}^{-1}(z)-\frac{\nu_2}{n}\mathbb{E}\mathrm{tr}\mathbf{D}^{-1}(z)\right|^q\le\frac{K_q}{n^{q/2}v_0^q}, \tag{S2.68}$$

which implies that

$$\mathbb{E}\left|\overline{\beta}_j(z)-b_p(z)\right|^2\le\frac{K|z|^4}{nv_0^6}. \tag{S2.69}$$

The above inequalities will be used in the following proof.

Denoting $\mathcal{Y}_{2,j}(z_1,z_2):=\mathbb{E}_{j-1}\Big\{\overline{\beta}_j(z_1)\varepsilon_j(z_1)\Big\}\mathbb{E}_{j-1}\Big\{\overline{\beta}_j(z_2)\varepsilon_j(z_2)\Big\}$. To prove the second part of the lemma, it is enough to show that, for each $1\le j\le n$,

$$\mathcal{Y}_{2,j}(z_1,z_2)-b_p(z_1)b_p(z_2)\Big\{\mathbb{E}_{j-1}\varepsilon_j(z_1)\Big\}\Big\{\mathbb{E}_{j-1}\varepsilon_j(z_2)\Big\}=o_P(n^{-1}), \tag{S2.70}$$

and

$$b_p(z_1)b_p(z_2)\{\mathbb{E}_{j-1}\varepsilon_j(z_1)\}\{\mathbb{E}_{j-1}\varepsilon_j(z_2)\}=o_P(n^{-1}). \tag{S2.71}$$

Proof of (**??**): We write

$$\begin{aligned}\text{LHS of } (\textbf{??})&=\mathbb{E}_{j-1}\Big[\Big\{\overline{\beta}_j(z_1)-b_p(z_1)\Big\}\varepsilon_j(z_1)\Big]\mathbb{E}_{j-1}\Big\{\overline{\beta}_j(z_2)\varepsilon_j(z_2)\Big\}\\&\quad+\mathbb{E}_{j-1}\Big\{b_p(z_1)\varepsilon_j(z_1)\Big\}\mathbb{E}_{j-1}\Big[\Big\{\overline{\beta}_j(z_2)-b_p(z_2)\Big\}\varepsilon_j(z_2)\Big]\\&=:\Delta\mathcal{Y}_{2,j}^{(1)}(z_1,z_2)+\Delta\mathcal{Y}_{2,j}^{(2)}(z_1,z_2).\end{aligned}$$

Note that, for any $q_1,q_2\ge1$ with $1/q_1+1/q_2=1$, we have

$$\begin{aligned}\mathbb{E}\left|\Delta\mathcal{Y}_{2,j}^{(1)}(z_1,z_2)\right|&\le\left(\mathbb{E}\left|\mathbb{E}_{j-1}\{\overline{\beta}_j(z_1)-b_p(z_1)\}\varepsilon_j(z_1)\right|^{q_1}\right)^{1/q_1}\left(\mathbb{E}\left|\mathbb{E}_{j-1}\overline{\beta}_j(z_2)\varepsilon_j(z_2)\right|^{q_2}\right)^{1/q_2}\\&\le\left(\mathbb{E}\left|\{\overline{\beta}_j(z_1)-b_p(z_1)\}\varepsilon_j(z_1)\right|^{q_1}\right)^{1/q_1}\left(\mathbb{E}\left|\overline{\beta}_j(z_2)\varepsilon_j(z_2)\right|^{q_2}\right)^{1/q_2}\end{aligned}$$

From Lemma **??** and (**??**), for any $1<q_1<2$ and $q_2>2$, we have

$$\begin{aligned}\mathbb{E}\left|\Big\{\overline{\beta}_j(z_1)-b_p(z_1)\Big\}\varepsilon_j(z_1)\right|^{q_1}&\le\left(\mathbb{E}\left|\overline{\beta}_j(z_1)-b_p(z_1)\right|^{q_1\frac{2}{q_1}}\right)^{q_1/2}\left(\mathbb{E}\left|\varepsilon_j(z_1)\right|^{q_1\frac{2}{2-q_1}}\right)^{\frac{2-q_1}{2}}\\&\le Kn^{-\frac{q_1}{2}}n^{-\frac{2-q_1}{2}}\delta_n^{4q_1-4}\\&=n^{-1}\delta_n^{4q_1-4},\end{aligned}$$

and

$$\mathbb{E}\left|\overline{\beta}_j(z_2)\varepsilon_j(z_2)\right|^{q_2}\le K\frac{|z|^{q_2}}{v_0^{q_2}}n^{-1}\delta_n^{2q_2-4}.$$

Thus,

$$\mathbb{E}\Big|\Delta\mathcal{Y}^{(1)}_{2,j}(z_1,z_2)\Big| \le Kn^{-1}\delta_n^{\frac{4q_1-4}{q_1}+\frac{2q_2-4}{q_2}} = o(n^{-1}).$$

Similarly, we can obtain $\mathbb{E}\Big|\Delta\mathcal{Y}^{(2)}_{2,j}(z_1,z_2)\Big| = o(n^{-1})$. Thus, we complete the proof of (**??**).
Proof of (**??**): By (**??**), we get

$$\begin{aligned}
\mathbb{E}_{j-1}\varepsilon_j(z) &= \mathbb{E}_{j-1}\left\{\mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{r}_j - \frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_j^{-1}(z)\right\} \\
&= \mathrm{tr}\Big\{\mathbb{E}_{j-1}\mathbf{D}_j^{-1}(z)\mathbb{E}\mathbf{r}_j\mathbf{r}_j'\Big\} - \frac{\nu_2}{n}\mathbb{E}_{j-1}\mathrm{tr}\mathbf{D}_j^{-1}(z) \\
&= \frac{p\nu_2}{n(p-1)}\mathrm{tr}\Big\{\mathbb{E}_{j-1}\mathbf{D}_j^{-1}(z)\Big(\mathbf{I}_p - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\Big)\Big\} - \frac{\nu_2}{n}\mathbb{E}_{j-1}\mathrm{tr}\mathbf{D}_j^{-1}(z) \\
&= \frac{\nu_2}{n(p-1)}\mathbb{E}_{j-1}\Big\{\mathrm{tr}\mathbf{D}_j^{-1}(z) - \mathbf{1}_p'\mathbf{D}_j^{-1}(z)\mathbf{1}_p\Big\}.
\end{aligned}$$

By Lemma 2.3 in **?** and our Lemma **??**, we have

$$\frac{1}{p}\mathrm{tr}\mathbf{D}_j^{-1}(z) \xrightarrow{P} m(z), \qquad \frac{1}{p}\mathbf{1}_p'\mathbf{D}_j^{-1}(z)\mathbf{1}_p \xrightarrow{P} -\frac{1}{z}.$$

By (**??**) and the identity $\underline{m}_p(z) = -\frac{1}{zn}\sum_{j=1}^n \beta_j(z)$, we have

$$|b_p(z) + \mathbb{E}\underline{m}_p(z)| \le Kn^{-1/2}.$$

This, together with Lemma **??**, yields that

$$|b_p(z) + z\underline{m}_p^0(z)| \le Kn^{-1/2}. \tag{S2.72}$$

Equation (**??**) follows from the above estimates.

S2.7 Proof of Lemma ??

By Lemma **??**, the inequality $|\beta_{ij}(z)| \le \frac{|z|}{v_0}$, and Lemma **??**, we get

$$\begin{aligned}
&\mathbb{E}\Big|(\mathbf{1}_p^i)'\Big\{\mathbf{D}_1^{-1}(z_1) - \mathbb{E}\mathbf{D}_1^{-1}(z_1)\Big\}\mathbf{1}_p^i\Big|^2 \\
&= \mathbb{E}\Big|\sum_{j=1}^n(\mathbb{E}_j - \mathbb{E}_{j-1})(\mathbf{1}_p^i)'\Big\{\mathbf{D}_1^{-1}(z_1) - \mathbf{D}_{1j}^{-1}(z_1)\Big\}\mathbf{1}_p^i\Big|^2 \\
&\le K\sum_{j=1}^n\mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{1j}(z_1)\mathbf{r}_j'\mathbf{D}_{1j}^{-1}\mathbf{1}_p^i(\mathbf{1}_p^i)'\mathbf{D}_{1j}^{-1}\mathbf{r}_j\Big|^2 \\
&\le K\sum_{j=1}^n\mathbb{E}\Big|\beta_{1j}(z_1)\mathbf{r}_j'\mathbf{D}_{1j}^{-1}\mathbf{1}_p^i(\mathbf{1}_p^i)'\mathbf{D}_{1j}^{-1}\mathbf{r}_j\Big|^2 \\
&\le Kn^{-1},
\end{aligned} \tag{S2.73}$$

where $\mathbf{1}_p^i$ is the p-dimensional vector with all components being 0 except for the i-th component being 1. Hence, we have

$$\mathbb{E}\Big|\frac{1}{n^2}\sum_{j=1}^n\sum_{i=1}^p\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_1) - \mathbb{E}\mathbf{D}_j^{-1}(z_1)\Big]_{ii}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_2)\Big]_{ii}\Big|$$

$$\leq \frac{nK}{n^2 v_0}\sum_{i=1}^{p}\mathbb{E}\Big|(\mathbf{1}_p^i)'\{\mathbf{D}_1^{-1}(z_1)-\mathbb{E}\mathbf{D}_1^{-1}(z_1)\}\mathbf{1}_p^i\Big| \leq Kn^{-1/2},$$

and thus

$$\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{p}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_1)-\mathbb{E}\mathbf{D}_j^{-1}(z_1)\Big]_{ii}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_2)\Big]_{ii} = O_P(n^{-1/2}).$$

Similarly, we have

$$\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{p}\mathbb{E}_j\Big[\mathbf{D}_j^{-1}(z_2)-\mathbb{E}\mathbf{D}_j^{-1}(z_2)\Big]_{ii}\mathbb{E}\Big[\mathbf{D}_j^{-1}(z_1)\Big]_{ii} = O_P(n^{-1/2}).$$

With the above two inequalities, it remains to find the limit of

$$\frac{1}{p}\sum_{i=1}^{p}\mathbb{E}\Big[\mathbf{D}_j^{-1}(z_1)\Big]_{ii}\mathbb{E}\Big[\mathbf{D}_j^{-1}(z_2)\Big]_{ii}. \tag{S2.74}$$

It is easy to see that the sum of expectations in (**??**) is exactly the same for any j. Moreover, we have

$$\frac{1}{p}\sum_{i=1}^{p}\mathbb{E}\Big[\mathbf{D}_j^{-1}(z_1)\Big]_{ii}\mathbb{E}\Big[\mathbf{D}_j^{-1}(z_2)\Big]_{ii} \xrightarrow{P} m(z_1)m(z_2).$$

This completes the proof of Lemma **??**.

S2.8 Proof of Lemma ??

Let

$$\mathbf{D}_{ij}(z)=\mathbf{D}(z)-\mathbf{r}_i\mathbf{r}_i'-\mathbf{r}_j\mathbf{r}_j', \quad b_1(z)=\frac{1}{1+n^{-1}\nu_2\mathbb{E}\mathrm{tr}\mathbf{D}_{12}^{-1}(z)}, \quad \beta_{ij}(z)=\frac{1}{1+\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z)\mathbf{r}_i}.$$

We have the equality

$$\mathbf{D}_j(z_1)+z_1\mathbf{I}_p-\frac{n-1}{n}\nu_2 b_1(z_1)\mathbf{I}_p=\sum_{i\neq j}^{n}\mathbf{r}_i\mathbf{r}_i'-\frac{n-1}{n}\nu_2 b_1(z_1)\mathbf{I}_p.$$

Multiplying by $\mathbf{Q}_p(z_1):=\{z_1\mathbf{I}_p-\frac{n-1}{n}\nu_2 b_1(z_1)\mathbf{I}_p\}^{-1}$ on the LHS and $\mathbf{D}_j^{-1}(z_1)$ on the RHS, and using $\mathbf{r}_i'\mathbf{D}_j^{-1}(z_1)=\beta_{ij}(z_1)\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z_1)$, we get

$$\mathbf{D}_j^{-1}(z_1)=-\mathbf{Q}_p(z_1)+b_1(z_1)\mathbf{A}(z_1)+\mathbf{B}(z_1)+\mathbf{C}(z_1), \tag{S2.75}$$

where

$$\mathbf{A}(z_1)=\sum_{i\neq j}^{n}\mathbf{Q}_p(z_1)\Big(\mathbf{r}_i\mathbf{r}_i'-\frac{\nu_2}{n}\mathbf{I}_p\Big)\mathbf{D}_{ij}^{-1}(z_1),$$

$$\mathbf{B}(z_1)=\sum_{i\neq j}^{n}\Big\{\beta_{ij}(z_1)-b_1(z_1)\Big\}\mathbf{Q}_p(z_1)\mathbf{r}_i\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z_1),$$

$$\mathbf{C}(z_1) = \frac{\nu_2}{n} b_1(z_1)\mathbf{Q}_p(z_1) \sum_{i \neq j}^{n} \Big\{ \mathbf{D}_{ij}^{-1}(z_1) - \mathbf{D}_j^{-1}(z_1) \Big\}.$$

For any $t \in \mathbb{R}$, $|1 - tb_1(z)/z|^{-1} \leq \frac{|z/b_1(z)|}{\mathrm{Im}\{z/b_1(z)\}} \leq \frac{|z|\{1+p/(nv_0)\}}{v_0}$. Thus,

$$\|\mathbf{Q}_p(z_1)\| \leq \frac{1 + p/(nv_0)}{v_0}. \tag{S2.76}$$

For any random matrix $\mathbf{M}$, denote its nonrandom bound on the spectrum norm of $\mathbf{M}$ by $|||\mathbf{M}|||$. Since the same argument in (**??**) holds for $\mathbf{D}_{12}^{-1}$, and from (**??**), Lemma **??**, we get

$$\mathbb{E}\Big|\mathrm{tr}\mathbf{B}(z_1)\mathbf{M}\Big| \leq K|||\mathbf{M}|||\frac{|z_1|^2\{1 + p/(nv_0)\}}{v_0^5} n^{1/2}. \tag{S2.77}$$

From (**??**), we have

$$\Big|\mathrm{tr}\mathbf{C}(z_1)\mathbf{M}\Big| \leq |||\mathbf{M}|||\frac{|z_1|\{1 + p/(nv_0)\}}{v_0^3}. \tag{S2.78}$$

From (**??**) and Lemma **??**, we get, for $\mathbf{M}$ nonrandom,

$$\mathbb{E}\Big|\mathrm{tr}\mathbf{A}(z_1)\mathbf{M}\Big| \leq K\|\mathbf{M}\|\frac{1 + p/(nv_0)}{v_0^2} n^{1/2}. \tag{S2.79}$$

Note that

$$\begin{aligned}
&\mathrm{tr}\Big[\mathbb{E}_j\Big\{\mathbf{A}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big] \\
&= \mathrm{tr}\sum_{i<j}^{n} \mathbf{Q}_p(z_1)\Big(\mathbf{r}_i\mathbf{r}_i' - n^{-1}\nu_2\mathbf{I}_p\Big)\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2) \\
&\quad + \mathrm{tr}\sum_{i<j}^{n} \mathbf{Q}_p(z_1)\Big(\mathbf{r}_i\mathbf{r}_i' - n^{-1}\nu_2\mathbf{I}_p\Big)\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\{\mathbf{D}_j^{-1}(z_2) - \mathbf{D}_{ij}^{-1}(z_2)\} \\
&\quad + \mathrm{tr}\,\mathbb{E}_j\Bigg\{\sum_{i>j}^{n} \mathbf{Q}_p(z_1)\Big(\mathbf{r}_i\mathbf{r}_i' - n^{-1}\nu_2\mathbf{I}_p\Big)\mathbf{D}_{ij}^{-1}(z_1)\Bigg\}\mathbf{D}_j^{-1}(z_2),
\end{aligned}$$

thus, by using a decomposition similar to (**??**), we can write

$$\mathrm{tr}\Big[\mathbb{E}_j\Big\{\mathbf{A}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big] = A_1(z_1,z_2) + A_2(z_1,z_2) + A_3(z_1,z_2) + R(z_1,z_2), \tag{S2.80}$$

where

$$A_1(z_1,z_2) = -\sum_{i<j}^{n} \beta_{ij}(z_2)\mathbf{r}_i'\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\mathbf{r}_i\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\mathbf{r}_i, \tag{S2.81}$$

$$A_2(z_1,z_2) = -\frac{\nu_2}{n}\mathrm{tr}\sum_{i<j}^{n} \mathbf{Q}_p(z_1)\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\Big\{\mathbf{D}_j^{-1}(z_2) - \mathbf{D}_{ij}^{-1}(z_2)\Big\},$$

$$A_3(z_1,z_2) = \operatorname{tr}\sum_{i<j}^{n} \mathbf{Q}_p(z_1)\Big(\mathbf{r}_i\mathbf{r}_i' - \frac{\nu_2}{n}\mathbf{I}_p\Big)\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2),$$

$$R(z_1,z_2) = \operatorname{tr}\mathbb{E}_j\Big\{\sum_{i>j}^{n} \mathbf{Q}_p(z_1)\Big(\mathbf{r}_i\mathbf{r}_i' - n^{-1}\nu_2\mathbf{I}_p\Big)\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2), \tag{S2.82}$$

and $\mathbf{1}_p$ is a p-dimensional vector with all elements being 1. It is easy to see that $R(z_1,z_2) = O_P(1)$. We get from (**??**) and (**??**) that $|A_2(z_1,z_2)| \le \frac{1+p/(nv_0)}{v_0^2}$. Similar to (**??**), we have $\mathbb{E}|A_3(z_1,z_2)| \le \frac{1+p/(nv_0)}{v_0^3}n^{1/2}$. It remains to derive the limit of $A_1(z_1,z_2)$. By using Lemma **??** and similar argument in (**??**), we have, for $i<j$,

$$\begin{aligned}
&\mathbb{E}\Big|\beta_{ij}(z_2)\mathbf{r}_i'\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\mathbf{r}_i\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\mathbf{r}_i \\
&\qquad - \frac{\nu_2^2}{n^2}b_1(z_2)\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\Big]\operatorname{tr}\Big\{\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big| \\
\le\ &\mathbb{E}\Big|\beta_{ij}(z_2)\mathbf{r}_i'\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\mathbf{r}_i\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\mathbf{r}_i \\
&\qquad - \frac{\nu_2^2}{n^2}\beta_{ij}(z_2)\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\Big]\operatorname{tr}\Big\{\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big| \\
&\quad + \mathbb{E}\Big|\frac{\nu_2^2}{n^2}\Big\{\beta_{ij}(z_2) - b_1(z_2)\Big\}\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\Big]\operatorname{tr}\Big\{\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big| \\
\le\ &\mathbb{E}\Big|\beta_{ij}(z_2)\Big(\mathbf{r}_i'\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\mathbf{r}_i - \frac{\nu_2}{n}\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\Big]\Big)\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\mathbf{r}_i \\
&\quad + \frac{\nu_2}{n}\beta_{ij}(z_2)\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\Big]\Big[\mathbf{r}_i'\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\mathbf{r}_i - \frac{\nu_2}{n}\operatorname{tr}\Big\{\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big]\Big| \\
&\quad + \mathbb{E}\Big|\frac{\nu_2^2}{n^2}\Big\{\beta_{ij}(z_2) - b_1(z_2)\Big\}\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\Big]\operatorname{tr}\Big\{\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big| \\
\le\ &Kn^{-1/2}.
\end{aligned} \tag{S2.83}$$

By (**??**), we have

$$\begin{aligned}
&\mathbb{E}\Big|\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_{ij}^{-1}(z_1)\Big\}\mathbf{D}_{ij}^{-1}(z_2)\Big]\operatorname{tr}\Big\{\mathbf{D}_{ij}^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\} \\
&\quad -\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_j^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big]\operatorname{tr}\Big\{\mathbf{D}_j^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big| \le Kn.
\end{aligned} \tag{S2.84}$$

It follows from (**??**) and (**??**) that

$$\mathbb{E}\Big|A_1(z_1,z_2) + \frac{j-1}{n^2}\nu_2^2 b_1(z_2)\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_j^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big]\operatorname{tr}\Big\{\mathbf{D}_j^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big| \le Kn^{1/2}. \tag{S2.85}$$

By using (**??**) – (**??**), we have

$$\begin{aligned}
\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_j^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big] &= \operatorname{tr}\Big[\mathbb{E}_j\Big\{-\mathbf{Q}_p(z_1) + b_1(z)\mathbf{A}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big] + O_P(n^{1/2}) \\
&= -\operatorname{tr}\Big\{\mathbf{Q}_p(z_1)\mathbf{D}_j^{-1}(z_2)\Big\} + b_1(z_1)\operatorname{tr}\mathbb{E}_j\Big\{\mathbf{A}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2) + O_P(n^{1/2}) \\
&= -\operatorname{tr}\Big\{\mathbf{Q}_p(z_1)\mathbf{D}_j^{-1}(z_2)\Big\} + b_1(z_1)A_1(z_1,z_2) + O_P(n^{1/2}).
\end{aligned}$$

This, together with (**??**), implies that

$$\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_j^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big]\Big[1 + \frac{j-1}{n^2}\nu_2^2 b_1(z_1)b_1(z_2)\operatorname{tr}\Big\{\mathbf{D}_j^{-1}(z_2)\mathbf{Q}_p(z_1)\Big\}\Big]$$

$$= -\operatorname{tr}\Big\{\mathbf{Q}_p(z_1)\mathbf{D}_j^{-1}(z_2)\Big\} + O_P(n^{1/2}).$$

By using (**??**) – (**??**) and (**??**), we have

$$\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_j^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big]\Big[1 - \frac{j-1}{n^2}\nu_2^2 b_1(z_1)b_1(z_2)\operatorname{tr}\Big\{\mathbf{Q}_p(z_2)\mathbf{Q}_p(z_1)\Big\}\Big]$$
$$= \operatorname{tr}\Big\{\mathbf{Q}_p(z_2)\mathbf{Q}_p(z_1)\Big\} + O_P(n^{1/2}). \tag{S2.86}$$

From (**??**), we have $|b_1(z) - b_p(z)| = O(n^{-1})$. This, together with (**??**) and (**??**), implies that

$$\operatorname{tr}\Big[\mathbb{E}_j\Big\{\mathbf{D}_j^{-1}(z_1)\Big\}\mathbf{D}_j^{-1}(z_2)\Big]\Bigg[1 - \frac{\frac{j-1}{n}c_n\nu_2^2\underline{m}_p^0(z_1)\underline{m}_p^0(z_2)}{\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_1)\}\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_2)\}}\Bigg]$$
$$= \frac{nc_n}{z_1z_2}\frac{1}{\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_1)\}\{1+\frac{n-1}{n}\nu_2\underline{m}_p^0(z_2)\}} + O_P(n^{1/2}).$$

This completes the proof of Lemma **??**.

S2.9 Proof of Lemma ??

Let $\widetilde{\mathbf{Q}}_p(z) = \mathbf{I}_p + \lambda\mathbb{E}\underline{m}_p(z)\mathbf{I}_p$, then

$$nA_p = \frac{p}{1+\lambda\mathbb{E}\underline{m}_p(z)} + pz\mathbb{E}m_p(z) = \mathbb{E}\{\beta_1(z)P_1(z)\} + \mathbb{E}\{\beta_1(z)P_2(z)\}, \tag{S2.87}$$

where

$$P_1(z) = n\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1 - \lambda\operatorname{tr}\Big\{\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbb{E}\mathbf{D}_1^{-1}(z)\Big\},$$
$$P_2(z) = \lambda\operatorname{tr}\Big\{\mathbf{Q}_p^{-1}(z)\mathbb{E}\mathbf{D}_1^{-1}(z)\Big\} - \lambda\operatorname{tr}\Big\{\mathbf{Q}_p^{-1}(z)\mathbb{E}\mathbf{D}^{-1}(z)\Big\}.$$

Since $\beta_1 = b_p - b_p^2\gamma_1 + \beta_1 b_p^2\gamma_1^2$, where $\gamma_1 \equiv \gamma_1(z) := \mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{r}_1 - n^{-1}\nu_2\mathbb{E}\operatorname{tr}\mathbf{D}_1^{-1}(z)$, we have

$$\mathbb{E}\{\beta_1(z)P_1(z)\} = b_p(z)\mathbb{E}P_1(z) - b_p^2(z)\mathbb{E}\{\gamma_1(z)P_1(z)\} + b_p^2(z)\mathbb{E}\{\beta_1(z)\gamma_1^2(z)P_1(z)\}. \tag{S2.88}$$

The estimates for $\mathbb{E}P_1(z)$, $\mathbb{E}\{\gamma_1(z)P_1(z)\}$, $\mathbb{E}\{\beta_1(z)\gamma_1^2(z)P_1(z)\}$, and $\mathbb{E}\{\beta_1(z)P_2(z)\}$ are provided in the following lemma, and its proof is postponed to Section **??**.

Lemma S2.3. *Under Assumptions* **??** *and* **??**, *we have*

$$\mathbb{E}P_1(z) = \frac{n}{1+\lambda\mathbb{E}\underline{m}_p(z)}\Big\{\mathbb{E}\gamma_1(z) + \frac{\nu_2-\lambda}{n}\mathbb{E}\operatorname{tr}\mathbf{D}_1^{-1}(z)\Big\}, \tag{S2.89}$$

$$\mathbb{E}\{\gamma_1(z)P_1(z)\}$$
$$= n\mathbb{E}\Bigg(\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{r}_1 - \frac{\nu_2}{n}\operatorname{tr}\mathbf{D}_1^{-1}(z)\Big\}\Big[\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1 - \frac{\nu_2}{n}\operatorname{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\Bigg)$$
$$+ \frac{\nu_2^2}{n(p-1)}\mathbb{E}\Big[\Big\{\operatorname{tr}\mathbf{D}_1^{-1}(z) - \mathbf{1}_p'\mathbf{D}_1^{-1}(z)\mathbf{1}_p\Big\}\operatorname{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]$$
$$- \frac{\lambda}{1+\lambda\mathbb{E}\underline{m}_p(z)}\mathbb{E}\operatorname{tr}\{\mathbf{D}_1^{-1}(z)\}\mathbb{E}\gamma_1(z) + o(1), \tag{S2.90}$$

$$
\begin{aligned}
&\mathbb{E}\Big\{\beta_1(z)\gamma_1^2(z)P_1(z)\Big\} \\
&= \mathbb{E}\Big\{n\beta_1(z)\gamma_1^2(z)\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1\Big\} - \mathbb{E}\Big[\beta_1(z)\gamma_1^2(z)\mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big] \\
&\quad + \mathrm{Cov}\Big(\beta_1(z)\gamma_1^2(z),\ \mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big) \\
&= O(\delta_n^2), \qquad \text{(S2.91)}
\end{aligned}
$$

$$
\mathbb{E}\{\beta_1(z)P_2(z)\} = \frac{p\nu_2\lambda b_p^2(z)}{n(p-1)}\mathbb{E}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\} + O(n^{-1/2}). \qquad \text{(S2.92)}
$$

From this lemma and (**??**), (**??**), we get

$$
nA_p = J_1 + J_2 + J_3 + o(1), \qquad \text{(S2.93)}
$$

where

$$
\begin{aligned}
J_1 &= \frac{nb_p(z)}{1+\lambda\mathbb{E}\underline{m}_p(z)}\Big\{\mathbb{E}\gamma_1(z) + \frac{\nu_2-\lambda}{n}\mathbb{E}\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\} + \frac{b_p^2(z)\lambda\mathrm{tr}\{\mathbb{E}\mathbf{D}_1^{-1}(z)\}\mathbb{E}\gamma_1(z)}{1+\lambda\mathbb{E}\underline{m}_p(z)} \\
&\quad - \frac{b_p^2(z)\nu_2^2}{n(p-1)}\mathbb{E}\Big[\Big\{\mathrm{tr}\mathbf{D}_1^{-1}(z) - \mathbf{1}_p'\mathbf{D}_1^{-1}(z)\mathbf{1}_p\Big\}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big], \\
J_2 &= -nb_p^2(z)\mathbb{E}\Big(\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{r}_1 - \frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}\Big[\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1 - \frac{\nu_2}{n}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\Big), \\
J_3 &= \frac{pb_p^2(z)\lambda\nu_2}{n(p-1)}\mathbb{E}\,\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}.
\end{aligned}
$$

The limits of J_1, J_2 and J_3 are provided in the following lemma, whose proof is postponed to Section **??**.

Lemma S2.4. *Under Assumptions* **??** *and* **??**, *as* $n \to \infty$,

$$
\begin{aligned}
J_1 &\to -\frac{\underline{m}(z)\{z(\lambda+h_1)m(z)+\lambda\}}{1+\lambda\underline{m}(z)}, \\
J_2 &\to -\frac{cz^2\underline{m}^2(z)\{(\alpha_1+\alpha_2)m^2(z)+2\lambda^2 m'(z)\}}{1+\lambda\underline{m}(z)}, \\
J_3 &\to \frac{c\lambda^2\underline{m}^2(z)}{\{1+\lambda\underline{m}(z)\}[\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)]}.
\end{aligned}
$$

By this Lemma and (**??**), we get the limit of nA_p.

S2.10 Proof of Lemma ??

Proof of (**??**): this equation follows from the definition of $\gamma_1(z)$.
Proof of (**??**): For $\mathbb{E}\{\gamma_1(z)P_1(z)\}$, we have

$$
\begin{aligned}
&\mathbb{E}\{\gamma_1(z)P_1(z)\} \\
&= n\mathbb{E}\Big(\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{r}_1 - \frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_1^{-1}(z) + \frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_1^{-1}(z) - \frac{\nu_2}{n}\mathbb{E}\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}
\end{aligned}
$$

$$
\begin{aligned}
&\times\Big[\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1-\frac{\nu_2}{n}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}+\frac{\nu_2}{n}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\Big)\\
&-\frac{\lambda}{1+\lambda\mathbb{E}\underline{m}_p(z)}\mathbb{E}\mathrm{tr}\{\mathbf{D}_1^{-1}(z)\}\mathbb{E}\gamma_1(z)\\
=\;& n\mathbb{E}\Big(\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{r}_1-\frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}\Big[\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1-\frac{\nu_2}{n}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\Big)\\
&+\frac{\nu_2^2}{n(p-1)}\mathbb{E}\Big[\Big\{\mathrm{tr}\mathbf{D}_1^{-1}(z)-\mathbf{1}_p'\mathbf{D}_1^{-1}(z)\mathbf{1}_p\Big\}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\\
&+\frac{p\nu_2^2}{n(p-1)}\mathrm{Cov}\Big(\mathrm{tr}\mathbf{D}_1^{-1}(z),\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big)\\
&-\frac{\nu_2^2}{n(p-1)}\mathrm{Cov}\Big(\mathrm{tr}\mathbf{D}_1^{-1}(z),\;\mathbf{1}_p'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{1}_p\Big)\\
&-\frac{\lambda}{1+\lambda\mathbb{E}\underline{m}_p(z)}\mathbb{E}\mathrm{tr}\{\mathbf{D}_1^{-1}(z)\}\mathbb{E}\gamma_1(z)\\
=\;& n\mathbb{E}\Big(\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{r}_1-\frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}\Big[\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1-\frac{\nu_2}{n}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\Big)\\
&+\frac{\nu_2^2}{n(p-1)}\mathbb{E}\Big[\Big\{\mathrm{tr}\mathbf{D}_1^{-1}(z)-\mathbf{1}_p'\mathbf{D}_1^{-1}(z)\mathbf{1}_p\Big\}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\\
&-\frac{\lambda}{1+\lambda\mathbb{E}\underline{m}_p(z)}\mathbb{E}\mathrm{tr}\{\mathbf{D}_1^{-1}(z)\}\mathbb{E}\gamma_1(z)+O(n^{-1}),
\end{aligned}
$$

which is the second equation in Lemma **??**. Below are some interpretations of the above equalities:

1. The second equality uses the following derivation: By (**??**), we get

$$
\begin{aligned}
& n\mathbb{E}\Big[\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{r}_1-\frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}\cdot\frac{\nu_2}{n}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\\
&+n\mathbb{E}\Big[\Big\{\frac{\nu_2}{n}\mathrm{tr}\mathbf{D}_1^{-1}(z)-\frac{\nu_2}{n}\mathbb{E}\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}\cdot\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1\Big]\\
=\;&\nu_2\mathrm{tr}\Big(\mathbb{E}(\mathbf{r}_1\mathbf{r}_1')\mathbb{E}\Big[\mathbf{D}_1^{-1}(z)\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\Big)-\frac{\nu_2^2}{n}\mathbb{E}\Big[\mathrm{tr}\mathbf{D}_1^{-1}(z)\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\\
&+\nu_2\mathrm{tr}\Big[\mathbb{E}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}\mathbb{E}(\mathbf{r}_1\mathbf{r}_1')\Big]-\nu_2\mathbb{E}\Big\{\mathrm{tr}\mathbf{D}_1^{-1}(z)\Big\}\mathrm{tr}\Big[\mathbb{E}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\mathbb{E}(\mathbf{r}_1\mathbf{r}_1')\Big]\\
=\;&\frac{\nu_2^2}{n(p-1)}\mathbb{E}\Big[\Big\{\mathrm{tr}\mathbf{D}_1^{-1}(z)-\mathbf{1}_p'\mathbf{D}_1^{-1}(z)\mathbf{1}_p\Big\}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big]\\
&+\frac{p\nu_2^2}{n(p-1)}\mathrm{Cov}\Big(\mathrm{tr}\mathbf{D}_1^{-1}(z),\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big)\\
&-\frac{\nu_2^2}{n(p-1)}\mathrm{Cov}\Big(\mathrm{tr}\mathbf{D}_1^{-1}(z),\;\mathbf{1}_p'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{1}_p\Big).
\end{aligned}
$$

2. The last equality is due to

$$
\frac{1}{n}\mathrm{Cov}\Big(\mathrm{tr}\mathbf{D}_1^{-1}(z),\;\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big)=O(n^{-1}),\tag{S2.94}
$$

$$
-\frac{1}{n(p-1)}\mathrm{Cov}\Big(\mathrm{tr}\mathbf{D}_1^{-1}(z),\;\mathbf{1}_p'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{1}_p\Big)=O(n^{-1}).\tag{S2.95}
$$

The equation (**??**) follows from the inequality

$$\mathbb{E}\Big|\mathrm{tr}\mathbf{D}_1^{-1}(z)\mathbf{M}-\mathbb{E}\mathrm{tr}\mathbf{D}_1^{-1}(z)\mathbf{M}\Big|^2\le K\|\mathbf{M}\|^2, \tag{S2.96}$$

where $\mathbf{M}$ is any deterministic $p\times p$ matrix. The proof of (**??**) is similar to Equation (4.7) of **?**. The equation (**??**) follows from (**??**) and Lemma **??**.

Proof of (**??**): For $\mathbb{E}\{\beta_1(z)\gamma_1^2(z)P_1(z)\}$, we have

$$\begin{aligned}
&\mathbb{E}\Big\{\beta_1(z)\gamma_1^2(z)P_1(z)\Big\}\\
&=\mathbb{E}\Big\{n\beta_1(z)\gamma_1^2(z)\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1\Big\}-\mathbb{E}\Big\{\beta_1(z)\gamma_1^2(z)\Big\}\mathbb{E}\Big[\lambda\mathrm{tr}\Big\{\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big]\\
&=\mathbb{E}\Big\{n\beta_1(z)\gamma_1^2(z)\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1\Big\}-\mathbb{E}\Big[\beta_1(z)\gamma_1^2(z)\mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big]\\
&\qquad+\mathrm{Cov}\Big(\beta_1(z)\gamma_1^2(z),\ \mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big).
\end{aligned}$$

From Lemma **??** and equation (**??**), we have

$$\begin{aligned}
&\mathbb{E}\Big\{n\beta_1(z)\gamma_1^2(z)\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1\Big\}-\mathbb{E}\Big[\beta_1(z)\gamma_1^2(z)\mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big]\\
&\le n\Big\{\mathbb{E}\Big|\gamma_1^2(z)\beta_1(z)\Big|^2\Big\}^{1/2}\Bigg[\mathbb{E}\Big|\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1-\frac{\lambda}{n}\mathrm{tr}\Big\{\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big|^2\Bigg]^{1/2}\\
&\le Kn(n^{-1}\delta_n^4)^{1/2}n^{-1/2}=K\delta_n^2,
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{Cov}\Big(\beta_1(z)\gamma_1^2(z),\ \mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big)\\
&\le\Big\{\mathbb{E}|\beta_1(z)|^4\Big\}^{1/4}\Big\{\mathbb{E}|\gamma_1^2(z)|^4\Big\}^{1/4}\Big(\mathbb{E}\Big|\mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}-\mathbb{E}\mathrm{tr}\Big\{\lambda\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big|^2\Big)^{1/2}\\
&\le Kn^{-1/4}\delta_n^3.
\end{aligned}$$

These estimates yield the third equation in the lemma.

Proof of (**??**): From (**??**), $\mathbf{D}_1^{-1}(z)-\mathbf{D}^{-1}(z)=\beta_1(z)\mathbf{D}_1^{-1}(z)\mathbf{r}_1\mathbf{r}_1'\mathbf{D}_1^{-1}(z)$, $\beta_1(z)=b_p(z)-b_p(z)\gamma_1(z)\beta_1(z)$ and $\mathbb{E}\beta_1(z)=b_p(z)+o(n^{-1/2})$, we have

$$\begin{aligned}
&\mathbb{E}\{\beta_1(z)P_2(z)\}\\
&=\lambda\mathbb{E}\{\beta_1(z)\}\mathrm{tr}\Big[\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbb{E}\Big\{\mathbf{D}_1^{-1}(z)-\mathbf{D}^{-1}(z)\Big\}\Big]\\
&=\lambda\mathbb{E}\{\beta_1(z)\}\mathbb{E}\Big[\Big\{b_p(z)-b_p(z)\beta_1(z)\gamma_1(z)\Big\}\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\mathbf{r}_1\Big]\\
&=\lambda b_p(z)\mathbb{E}\{\beta_1(z)\}\Big[\mathbb{E}\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\mathbf{r}_1\Big\}-\mathbb{E}\Big\{\beta_1(z)\gamma_1(z)\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\mathbf{r}_1\Big\}\Big]\\
&=\lambda b_p^2(z)\mathbb{E}\Big\{\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\mathbf{r}_1\Big\}+O(n^{-1/2})\\
&=\lambda b_p^2(z)\mathrm{tr}\Big[\mathbb{E}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\mathbb{E}(\mathbf{r}_1\mathbf{r}_1')\Big]+O(n^{-1/2})\\
&=\frac{\lambda b_p^2(z)p\nu_2}{n(p-1)}\mathrm{tr}\Big[\mathbb{E}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\}\Big(\mathbf{I}_p-\frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\Big)\Big]+O(n^{-1/2})
\end{aligned}$$

$$= \frac{\lambda b_p^2(z) p \nu_2}{n(p-1)} \mathbb{E}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\Big\} + O(n^{-1/2}),$$

which is the last equation. Below are some interpretations of the above equalities:

1. The fourth equality follows from

$$\mathbb{E}\Big\{\beta_1(z)\gamma_1(z)\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\mathbf{r}_1\Big\} = O(n^{-1/2}),$$

   which can proved by using Lemma **??**.

2. The last equality follows from

$$\frac{\lambda b_p^2(z)\nu_2}{n(p-1)} \mathbb{E}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\mathbf{1}_p\mathbf{1}_p'\Big\} = O(n^{-1}).$$

   This can be proved by using Lemma **??** and Lemma **??**.

S2.11 Proof of Lemma ??

<u>Step 1: Consider J_1.</u> By Lemma 2.3 in **?**, we have

$$\frac{1}{p}\mathrm{tr}\mathbf{D}_j^{-1}(z) \xrightarrow{P} m(z).$$

By this estimate, Equation (**??**) and Lemma **??**, we get

$$\begin{aligned} n\mathbb{E}\gamma_1(z) &= \frac{p\nu_2}{p-1}\mathbb{E}\mathrm{tr}\Big\{\Big(\mathbf{I}_p - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\Big)\mathbf{D}_1^{-1}(z)\Big\} - \nu_2\mathbb{E}\mathrm{tr}\mathbf{D}_1^{-1}(z) \\ &= -\frac{\nu_2}{p-1}\mathbb{E}\mathrm{tr}\Big\{\mathbf{1}_p'\mathbf{D}_1^{-1}(z)\mathbf{1}_p\Big\} + \frac{\nu_2}{p-1}\mathbb{E}\mathrm{tr}\mathbf{D}_1^{-1}(z) \\ &\to \lambda\{m(z) + 1/z\}. \end{aligned}$$

By Lemma **??**, we have

$$\begin{aligned} &-\frac{b_p^2(z)\nu_2^2}{n(p-1)}\mathbb{E}\Big[\Big\{\mathrm{tr}\mathbf{D}_1^{-1}(z) - \mathbf{1}_p'\mathbf{D}_1^{-1}(z)\mathbf{1}_p\Big\}\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big] \\ \to\ &-\frac{cz^2\underline{m}^2(z)\lambda^2}{1+\lambda\underline{m}(z)}\Big\{m^2(z) + \frac{m(z)}{z}\Big\}. \end{aligned}$$

By Lemma **??**, Equation (**??**) and the above estimates, we have

$$J_1 \to -\frac{z\underline{m}(z)}{1+\lambda\underline{m}(z)}\Big\{(\lambda+h_1)m(z) + \frac{\lambda}{z}\Big\} = -\frac{\underline{m}(z)\{z(\lambda+h_1)m(z)+\lambda\}}{1+\lambda\underline{m}(z)}.$$

<u>Step 2: Consider J_2.</u> By Lemma **??**, we have

$$J_2 = J_{21} + 2J_{22} + J_{23} + o(1),$$

where

$$J_{21} = -\frac{(\nu_4 - 3\nu_{12})b_p^2(z)}{n}\sum_{i=1}^{p}\mathbb{E}\Big(\Big[\mathbf{D}_1^{-1}(z)\Big]_{ii}\Big[\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big]_{ii}\Big),$$

$$J_{22} = -\frac{\nu_{12} b_p^2(z)}{n} \mathbb{E}\mathrm{tr}\Big\{\mathbf{D}_1^{-2}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\},$$
$$J_{23} = -\frac{(\nu_{12} - \nu_2^2) b_p^2(z)}{n} \mathbb{E}\Big[\mathrm{tr}\mathbf{D}_1^{-1}(z)\,\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\}\Big].$$

Since $\frac{1}{p}\sum_{i=1}^{p}\mathbb{E}\Big(\Big[\mathbf{D}_1^{-1}(z)\Big]_{ii}\Big[\mathbf{D}_1^{-1}(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\Big]_{ii}\Big) \to \frac{m^2(z)}{1+\lambda\underline{m}(z)}$, we have

$$J_{21} \to \frac{-c\alpha_1 z^2 m^2(z)\underline{m}^2(z)}{1+\lambda\underline{m}(z)}.$$

Note that $\frac{1}{p}\mathbb{E}\mathrm{tr}\mathbf{D}_1^{-2}(z) \to m'(z)$, thus we get

$$J_{22} \to -\frac{c\lambda^2 z^2 m'(z)\underline{m}^2(z)}{1+\lambda\underline{m}(z)}.$$

By Lemma **??**, we get

$$J_{23} \to -\frac{c\alpha_2 z^2 m^2(z)\underline{m}^2(z)}{1+\lambda\underline{m}(z)}.$$

From these estimates, we have

$$J_2 \to -\frac{cz^2\underline{m}^2(z)\{(\alpha_1+\alpha_2)m^2(z) + 2\lambda^2 m'(z)\}}{1+\lambda\underline{m}(z)}.$$

Step 3: Consider J_3. To calculate the limit of J_3, we can expand $\mathbf{D}_1^{-1}(z)$ like (**??**) and find the limit of J_3 using the method similarly to **?**. The limit of J_3 is

$$\frac{c\lambda^2\underline{m}^2(z)}{\{1+\lambda\underline{m}(z)\}[\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)]}.$$

S2.12 Proof of Lemma ??

By Lemma **??**, we obtain, for any $2 \le q \in \mathbb{N}^+$,

$$\begin{aligned}
&\mathbb{E}|\mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\mathbf{r}_j|^q \\
&\le K_q\Big(\mathbb{E}\Big|\mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\mathbf{r}_j - \frac{\nu_2}{n}\mathrm{tr}\Big\{\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\Big\}\Big|^q \\
&\qquad + \Big|\frac{\nu_2}{n}\mathbb{E}\mathrm{tr}\Big\{\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\Big\}\Big|^q\Big) \\
&\le K_q\Big(n^{-2}\delta_n^{2q-4} + n^{-q}\Big) = O(n^{-2}),
\end{aligned} \tag{S2.97}$$

where $\frac{1}{n}\mathrm{tr}\Big\{\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\Big\} = \frac{1}{np}\mathbf{1}_p'\mathbf{D}_j^{-2}(z)\mathbf{1}_p = O(n^{-1})$. Write

$$\begin{aligned}
&\mathbf{x}_p'\mathbf{D}^{-1}(z)\mathbf{x}_p - \mathbf{x}_p'\mathbb{E}\mathbf{D}^{-1}(z)\mathbf{x}_p \\
&= \sum_{j=1}^{n}\mathbf{x}_p'\Big\{(\mathbb{E}_j - \mathbb{E}_{j-1})\mathbf{D}^{-1}(z)\Big\}\mathbf{x}_p \\
&= \sum_{j=1}^{n}\mathbf{x}_p'\Big[(\mathbb{E}_j - \mathbb{E}_{j-1})\Big\{\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z)\Big\}\Big]\mathbf{x}_p
\end{aligned}$$

$$= -\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z)\mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\mathbf{r}_j.$$

By Lemma **??**, (**??**), and $|\beta_j(z)| \le |z|/v_0$, we have

$$\begin{aligned}
&\mathbb{E}\left|\mathbf{x}_p'\mathbf{D}^{-1}(z)\mathbf{x}_p - \mathbf{x}_p'\mathbb{E}\mathbf{D}^{-1}(z)\mathbf{x}_p\right|^2 \\
&\le K\sum_{j=1}^{n}\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z)\mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\mathbf{r}_j|^2 \\
&\le K\sum_{j=1}^{n}\mathbb{E}|\beta_j(z)\mathbf{r}_j'\mathbf{D}_j^{-1}(z)\mathbf{x}_p\mathbf{x}_p'\mathbf{D}_j^{-1}(z)\mathbf{r}_j|^2 \\
&\le Kn^{-1}.
\end{aligned}$$

Thus, we have

$$\mathbb{E}|\mathbf{x}_p'\mathbf{D}^{-1}(z)\mathbf{x}_p - \mathbf{x}_p'\mathbb{E}\mathbf{D}^{-1}(z)\mathbf{x}_p|^2 \to 0. \tag{S2.98}$$

Recalling that $\widetilde{\mathbf{Q}}_p(z) = \mathbf{I}_p + \lambda\mathbb{E}\underline{m}_p(z)\mathbf{I}_p$. Using the identities $\underline{m}_p(z) = -\frac{1}{nz}\sum_{j=1}^{n}\beta_j(z)$ and $\mathbf{r}_j'\mathbf{D}^{-1}(z) = \beta_j(z)\mathbf{r}_j'\mathbf{D}_j^{-1}(z)$, we obtain

$$\begin{aligned}
&\left\{-z\widetilde{\mathbf{Q}}_p(z)\right\}^{-1} - \mathbf{D}^{-1}(z) \\
&= -\frac{1}{z}\widetilde{\mathbf{Q}}_p^{-1}(z)\left\{\mathbf{D}(z) + z\widetilde{\mathbf{Q}}_p(z)\right\}\mathbf{D}^{-1}(z) \\
&= -\frac{1}{z}\widetilde{\mathbf{Q}}_p^{-1}(z)\left\{\sum_{j=1}^{n}\mathbf{r}_j\mathbf{r}_j' + z\lambda\mathbb{E}\underline{m}_p(z)\mathbf{I}_p\right\}\mathbf{D}^{-1}(z) \\
&= -\frac{1}{z}\sum_{j=1}^{n}\beta_j(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_j\mathbf{r}_j'\mathbf{D}_j^{-1}(z) + \frac{\lambda\mathbb{E}\{\beta_1(z)\}}{z}\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}^{-1}(z).
\end{aligned}$$

Taking expectation of the above identity yields that

$$\begin{aligned}
&\left\{-z\widetilde{\mathbf{Q}}_p(z)\right\}^{-1} - \mathbb{E}\mathbf{D}^{-1}(z) \\
&= -\frac{n}{z}\mathbb{E}\left\{\beta_1(z)\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1\mathbf{r}_1'\mathbf{D}_j^{-1}(z)\right\} + \frac{\lambda\mathbb{E}\{\beta_1(z)\}}{z}\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbb{E}\mathbf{D}^{-1}(z).
\end{aligned}$$

Multiplying by $-\mathbf{x}_p'$ on the left and $\mathbf{x}_p$ on the right, we have

$$\mathbf{x}_p'\mathbb{E}\mathbf{D}^{-1}(z)\mathbf{x}_p + \frac{1}{z\{1 + \lambda\mathbb{E}\underline{m}_p(z)\}} = \varrho_1(z) + \varrho_2(z) + \varrho_3(z),$$

where

$$\begin{aligned}
\varrho_1(z) &:= \frac{n}{z}\mathbb{E}\{\beta_1(z)\varrho_{11}(z)\}, \\
\varrho_{11}(z) &:= \mathbf{x}_p'\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1\mathbf{r}_1'\mathbf{D}_1^{-1}(z)\mathbf{x}_p - \frac{\lambda}{n}\mathbf{x}_p'\widetilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)\mathbf{x}_p, \\
\varrho_2(z) &:= \frac{\lambda}{z}\mathbb{E}\left[\beta_1(z)\mathbf{x}_p'\widetilde{\mathbf{Q}}_p^{-1}(z)\left\{\mathbf{D}_1^{-1}(z) - \mathbf{D}^{-1}(z)\right\}\mathbf{x}_p\right],
\end{aligned}$$

$$\varrho_3(z) := \frac{\lambda}{z}\mathbb{E}\Big[\beta_1(z)\mathbf{x}_p'\widetilde{\mathbf{Q}}_p^{-1}(z)\Big\{\mathbf{D}^{-1}(z) - \mathbb{E}\mathbf{D}^{-1}(z)\Big\}\mathbf{x}_p\Big].$$

Recalling the notations defined above and the following equalities:

$$\varrho_1(z) = \frac{n}{z}\mathbb{E}\Big\{\overline{\beta}_1(z)\varrho_{11}(z)\Big\} - \frac{n}{z}\mathbb{E}\Big\{\beta_1(z)\overline{\beta}_1(z)\varepsilon_1(z)\varrho_{11}(z)\Big\}, \tag{S2.99}$$

$$\overline{\beta}_1(z) = b_p(z) - \frac{\nu_2}{n}b_p(z)\overline{\beta}_1(z)\mathrm{tr}\Big\{\mathbf{D}_1^{-1}(z) - \mathbb{E}\mathbf{D}_1^{-1}(z)\Big\}. \tag{S2.100}$$

From (**??**) – (**??**), Lemma **??** and Lemma **??**, it is easy to see that

$$\varrho_1(z) = \frac{n\mathbb{E}\varrho_{11}(z)}{z}\left\{\frac{1}{1+c\lambda m(z)} + o(1)\right\}. \tag{S2.101}$$

From (**??**) and Lemma **??**, we have

$$\mathbb{E}\varrho_{11}(z) = -\frac{1}{\lambda\underline{m}(z)+1}\frac{\lambda}{n}\Big\{\mathbf{x}_p'\mathbb{E}\mathbf{D}_1^{-1}(z)\mathbf{x}_p + o(1)\Big\}. \tag{S2.102}$$

Therefore, by (**??**) – (**??**), and Lemma **??** we have

$$\varrho_1(z) = \frac{\lambda\underline{m}(z)}{\lambda\underline{m}(z)+1}\mathbf{x}_p'\mathbb{E}\mathbf{D}_1^{-1}(z)\mathbf{x}_p + o(1).$$

Similarly to **?**, one may have $\varrho_2(z) = o(1)$ and $\varrho_3(z) = o(1)$. Hence, we obtain

$$\frac{\mathbf{x}_p'\mathbb{E}\mathbf{D}^{-1}(z)\mathbf{x}_p}{1+\lambda\underline{m}(z)} + \frac{1}{z\{1+\mathbb{E}\lambda\underline{m}(z)\}} = o(1),$$

which implies that

$$\mathbf{x}_p'\mathbb{E}\mathbf{D}^{-1}(z)\mathbf{x}_p \to -\frac{1}{z}.$$

This, together with (**??**), completes the proof of Lemma **??**.

S2.13 Proof of Corollary ??

First, we provide some expressions as follows. The proof of these equations are routine and thus omitted.

$$z = -\frac{1}{\underline{m}} + \frac{c\lambda}{1+\lambda\underline{m}}, \tag{S2.103}$$

$$\mathrm{d}z = \frac{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2}{\underline{m}^2(1+\lambda\underline{m})^2}\,\mathrm{d}\underline{m}, \tag{S2.104}$$

$$m = \frac{-\underline{m}}{-1-\lambda\underline{m}+c\lambda\underline{m}}, \tag{S2.105}$$

$$\frac{\mathrm{d}m}{\mathrm{d}z} = \frac{\underline{m}^2(1+\lambda\underline{m})^2}{\{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2\}(-1-\lambda\underline{m}+c\lambda\underline{m})^2}. \tag{S2.106}$$

These formula will be used in the following calculations.

S2.13.1 Calculation of expectation

The contours $\mathcal{C}$ is closed and taken in the positive direction in the complex plane, enclosing the support of $F^{c,H}$. Let $\tilde{C}$ be $\underline{m}(\mathcal{C})$.

For $f_1 = x$,

$$\begin{aligned}
\mathbb{E}X_x = &\frac{c\lambda^2}{2\pi i}\oint_{\mathcal{C}} \frac{z\underline{m}^3(z)\{1+\lambda\underline{m}(z)\}}{[\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)]^2}\,\mathrm{d}z \\
&- \frac{1}{2\pi i}\oint_{\mathcal{C}} \frac{z^2\underline{m}^2(z)\{h_1 m(z) + \lambda m(z) + \lambda/z\}\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z \\
&- \frac{c}{2\pi i}\oint_{\mathcal{C}} \frac{z^3\underline{m}^3(z)\{(\alpha_1+\alpha_2)m^2(z) + 2\lambda^2 m'(z)\}\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z, \\
=: &I_1(f_1) + I_2(f_1) + I_3(f_1).
\end{aligned}$$

For $I_1(f_1)$, by using (**??**) and (**??**), we get

$$\begin{aligned}
I_1(f_1) &= \frac{c\lambda^2}{2\pi i}\oint_{\mathcal{C}} \frac{z\underline{m}^3(z)\{1+\lambda\underline{m}(z)\}}{[\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)]^2}\,\mathrm{d}z \\
&= \frac{c\lambda^2}{2\pi i}\oint_{\tilde{\mathcal{C}}} \frac{\underline{m}^3(1+\lambda\underline{m})}{\{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2\}^2} \frac{(-1-\lambda\underline{m}+c\lambda\underline{m})\{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2\}}{\underline{m}^3(1+\lambda\underline{m})^3}\,\mathrm{d}\underline{m} \\
&= \frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}} \frac{c\lambda^2(1+\lambda\underline{m} - c\lambda\underline{m})}{(1+\lambda\underline{m})^2\{c\lambda^2\underline{m}^2 - (1+\lambda\underline{m})^2\}}\,\mathrm{d}\underline{m}.
\end{aligned}$$

The poles of $I_1(f_1)$ are $-\lambda^{-1}, -\frac{1}{(1\pm\sqrt{c})\lambda}$, we have by the residue theorem

$$I_1(f_1) = \lambda(1+c) - \frac{\lambda}{2}(1+\sqrt{c})^2 - \frac{\lambda}{2}(1-\sqrt{c})^2 = 0.$$

For the second integral $I_2(f_1)$, by using (**??**) – (**??**), we have,

$$\begin{aligned}
I_2(f_1) &= -\frac{1}{2\pi i}\oint_{\mathcal{C}} \frac{z^2\underline{m}^2(z)\{1+\lambda\underline{m}(z)\}\{h_1 m(z) + \lambda m(z) + \lambda/z\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z \\
&= -\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}} \frac{\underline{m}^2(1+\lambda\underline{m})}{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2}\frac{1}{c}\left\{(h_1+\lambda)\underline{m} + (h_1+\lambda-h_1 c)\frac{\underline{m}(1+\lambda\underline{m})}{-1-\lambda\underline{m}+c\lambda\underline{m}}\right\} \\
&\qquad\times \frac{(-1-\lambda\underline{m}+c\lambda\underline{m})^2\{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2\}}{\underline{m}^4(1+\lambda\underline{m})^4}\,\mathrm{d}\underline{m} \\
&= -\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}} \frac{(\lambda^2\underline{m} - h_1)(-1-\lambda\underline{m}+c\lambda\underline{m})}{\underline{m}(1+\lambda\underline{m})^3}\,\mathrm{d}\underline{m}.
\end{aligned}$$

The pole of $I_2(f_1)$ is $-\lambda^{-1}$, we have by the residue theorem

$$I_2(f_1) = h_1.$$

For $I_3(f_1)$, by using (**??**) – (**??**), we get

$$\begin{aligned}
I_3(f_1) &= -\frac{c}{2\pi i}\oint_{\mathcal{C}} \frac{z^3\underline{m}^3(z)\{(\alpha_1+\alpha_2)m^2(z) + 2\lambda^2 m'(z)\}\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z \\
&= -\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}}\left[(\alpha_1+\alpha_2)\left(\frac{-\underline{m}}{-1-\lambda\underline{m}+c\lambda\underline{m}}\right)^2 + \frac{2\lambda^2\underline{m}^2(1+\lambda\underline{m})^2}{\{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2\}(-1-\lambda\underline{m}+c\lambda\underline{m})^2}\right]
\end{aligned}$$

$$\times \frac{c\underline{m}^3(1+\lambda\underline{m})}{(1+\lambda\underline{m})^2-c\lambda^2\underline{m}^2}\times\frac{(-1-\lambda\underline{m}+c\lambda\underline{m})^3\{(1+\lambda\underline{m})^2-c\lambda^2\underline{m}^2\}}{\underline{m}^5(1+\lambda\underline{m})^5}\mathrm{d}\underline{m}$$
$$=-\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}}\left\{\frac{c(\alpha_1+\alpha_2)(-1-\lambda\underline{m}+c\lambda\underline{m})}{(1+\lambda\underline{m})^4}\mathrm{d}\underline{m}+\frac{2\lambda^2c(-1-\lambda\underline{m}+c\lambda\underline{m})}{(1+\lambda\underline{m})^2\{(1+\lambda\underline{m})^2-c\lambda^2\underline{m}^2\}}\right\}\mathrm{d}\underline{m}.$$

The poles of $I_3(f_1)$ are $-\frac{1}{\lambda}$, $-\frac{1}{(1\pm\sqrt{c})\lambda}$, we have by the residue theorem

$$I_3(f_1)=-2\lambda(1+c)+\lambda(1+\sqrt{c})^2+\lambda(1-\sqrt{c})^2=0.$$

Thus, we get

$$\mathbb{E}X_x=h_1.$$

For $f_2=x^2$, we have

$$\begin{aligned}
\mathbb{E}X_{x^2}=&\frac{c\lambda^2}{2\pi i}\oint_{\mathcal{C}}\frac{z^2\underline{m}^3(z)\{1+\lambda\underline{m}(z)\}}{[\{1+\lambda\underline{m}(z)\}^2-c\lambda^2\underline{m}^2(z)]^2}\mathrm{d}z\\
&-\frac{1}{2\pi i}\oint_{\mathcal{C}}\frac{z^3\underline{m}^2(z)\{h_1m(z)+\lambda m(z)+\lambda/z\}\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2-c\lambda^2\underline{m}^2(z)}\mathrm{d}z\\
&-\frac{c}{2\pi i}\oint_{\mathcal{C}}\frac{z^4\underline{m}^3(z)\{(\alpha_1+\alpha_2)m^2(z)+2\lambda^2m'(z)\}\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2-c\lambda^2\underline{m}^2(z)}\mathrm{d}z\\
=&\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}}\frac{c\lambda^2(-c\lambda\underline{m}+\lambda\underline{m}+1)^2}{\underline{m}(1+\lambda\underline{m})^3\{(1+\lambda\underline{m})^2-c\lambda^2\underline{m}^2\}}\mathrm{d}\underline{m}\\
&-\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}}\frac{(\lambda^2\underline{m}-h_1)(-1-\lambda\underline{m}+c\lambda\underline{m})^2}{\underline{m}^2(1+\lambda\underline{m})^4}\mathrm{d}\underline{m}\\
&-\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}}\left\{\frac{c(\alpha_1+\alpha_2)(-1-\lambda\underline{m}+c\lambda\underline{m})^2}{\underline{m}(1+\lambda\underline{m})^5}\mathrm{d}\underline{m}+\frac{2\lambda^2c(-1-\lambda\underline{m}+c\lambda\underline{m})^2}{\underline{m}(1+\lambda\underline{m})^3\{(1+\lambda\underline{m})^2-c\lambda^2\underline{m}^2\}}\right\}\mathrm{d}\underline{m}\\
=:&\,I_1(f_2)+I_2(f_2)+I_3(f_2).
\end{aligned}$$

For the first integral $I_1(f_2)$, the poles of are $-\lambda^{-1}$, $-\frac{1}{(1\pm\sqrt{c})\lambda}$, we have by the residue theorem,

$$I_1(f_2)=-c\lambda^2.$$

For the second integral $I_2(f_2)$, the poles is $-\lambda^{-1}$, we have by the residue theorem,

$$I_2(f_2)=\lambda(\lambda+2ch_1+2h_1).$$

For the third integral $I_3(f_2)$, the poles are $-\lambda^{-1}$, $-\frac{1}{(1\pm\sqrt{c})\lambda}$, we have by the residue theorem,

$$I_3(f_2)=c(\alpha_1+\alpha_2)+2\lambda^2c.$$

Thus,

$$\mathbb{E}X_{x^2}=(c+1)\lambda^2+2(c+1)\lambda h_1+c(\alpha_1+\alpha_2).$$

For $f_3=x^3$, we have

$$\mathbb{E}X_{x^3}=\frac{c\lambda^2}{2\pi i}\oint_{\mathcal{C}}\frac{z^3\underline{m}^3(z)\{1+\lambda\underline{m}(z)\}}{[\{1+\lambda\underline{m}(z)\}^2-c\lambda^2\underline{m}^2(z)]^2}\mathrm{d}z$$

$$
\begin{aligned}
&- \frac{1}{2\pi i}\oint_{\mathcal{C}} \frac{z^4\underline{m}^2(z)\{h_1 m(z)+\lambda m(z)+\lambda/z\}\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z \\
&- \frac{c}{2\pi i}\oint_{\mathcal{C}} \frac{z^5\underline{m}^3(z)\{(\alpha_1+\alpha_2)m^2(z)+2\lambda^2 m'(z)\}\{1+\lambda\underline{m}(z)\}}{\{1+\lambda\underline{m}(z)\}^2 - c\lambda^2\underline{m}^2(z)}\,\mathrm{d}z \\
= &\frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}} \frac{c\lambda^2(-1-\lambda\underline{m}+c\lambda\underline{m})^3}{\underline{m}^2(1+\lambda\underline{m})^4\{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2\}}\,\mathrm{d}\underline{m} \\
&- \frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}} \frac{(\lambda^2\underline{m}-h_1)(-1-\lambda\underline{m}+c\lambda\underline{m})^3}{\underline{m}^3(1+\lambda\underline{m})^5}\,\mathrm{d}\underline{m} \\
&- \frac{1}{2\pi i}\oint_{\tilde{\mathcal{C}}} \left\{\frac{c(\alpha_1+\alpha_2)(-1-\lambda\underline{m}+c\lambda\underline{m})^3}{\underline{m}^2(1+\lambda\underline{m})^6}\,\mathrm{d}\underline{m} + \frac{2\lambda^2 c(-1-\lambda\underline{m}+c\lambda\underline{m})^3}{\underline{m}^2(1+\lambda\underline{m})^4\{(1+\lambda\underline{m})^2 - c\lambda^2\underline{m}^2\}}\right\}\mathrm{d}\underline{m} \\
=: &\, I_1(f_3)+I_2(f_3)+I_3(f_3).
\end{aligned}
$$

For the first integral $I_1(f_3)$, the poles are $-\lambda^{-1}$, $-\frac{1}{(1\pm\sqrt{c})\lambda}$, we have by the residue theorem

$$I_1(f_3) = -3c(1+c)\lambda^3.$$

For the second integral $I_2(f_3)$, the poles is $-\lambda^{-1}$, we have by the residue theorem

$$I_2(f_3) = \lambda^2\{(2+3c)\lambda + 3(1+3c+c^2)h_1\}.$$

For the third integral $I_3(f_3)$, the poles are $-\lambda^{-1}$, $-\frac{1}{(1\pm\sqrt{c})\lambda}$, we have by the residue theorem

$$I_3(f_3) = 3c(1+c)\lambda(\alpha_1+\alpha_2) + 6c(1+c)\lambda^3.$$

Thus,

$$\mathbb{E}X_{x^3} = (3c^2+6c+2)\lambda^3 + 3(c^2+3c+1)\lambda^2 h_1 + 3c(c+1)\lambda(\alpha_1+\alpha_2).$$

S2.13.2 Calculation of variance

We claim that

$$
\begin{aligned}
&\operatorname{Cov}(X_{x^{r_1}}, X_{x^{r_2}}) \\
&= 2(\lambda c)^{r_1+r_2}\sum_{k_1=0}^{r_1-1}\sum_{k_2=0}^{r_2}\binom{r_1}{k_1}\binom{r_2}{k_2}\left(\frac{1-c}{c}\right)^{k_1+k_2}\sum_{\ell=1}^{r_1-k_1}\ell\binom{2r_1-1-k_1-\ell}{r_1-1}\binom{2r_2-1-k_2+\ell}{r_2-1} &\text{(S2.107)}\\
&\quad + \frac{c}{\lambda^2}(\alpha_1+\alpha_2)(\lambda c)^{r_1+r_2}\sum_{k_1=0}^{r_1}\sum_{k_2=0}^{r_2}\binom{r_1}{k_1}\binom{r_2}{k_2}\left(\frac{1-c}{c}\right)^{k_1+k_2}\binom{2r_1-k_1}{r_1-1}\binom{2r_2-k_2}{r_2-1}, &\text{(S2.108)}
\end{aligned}
$$

where $r_1, r_2 \in \mathbb{N}^+$. By using this result, we obtain the variances in Corollary **??**. It suffices to prove the above equation. The contours $\mathcal{C}_1, \mathcal{C}_2$ are closed and taken in the positive direction in the complex plane, each enclosing the support of $F^{c,H}$. Let $\tilde{\mathcal{C}}_i$ be $\underline{m}(\mathcal{C}_i)$ for $i=1,2$. From Theorem **??**, we have

$$
\begin{aligned}
\operatorname{Cov}(X_{x^{r_1}}, X_{x^{r_2}}) = &-\frac{1}{2\pi^2}\oint_{\tilde{\mathcal{C}}_1}\oint_{\tilde{\mathcal{C}}_2}\frac{z_1^{r_1}z_2^{r_2}}{(\underline{m}_1-\underline{m}_2)^2}\,\mathrm{d}\underline{m}_1\,\mathrm{d}\underline{m}_2 \\
&-\frac{c\,(\alpha_1+\alpha_2)}{4\pi^2}\oint_{\tilde{\mathcal{C}}_1}\oint_{\tilde{\mathcal{C}}_2}\frac{z_1^{r_1}z_2^{r_2}}{(1+\lambda\underline{m}_1)^2(1+\lambda\underline{m}_2)^2}\,\mathrm{d}\underline{m}_1\,\mathrm{d}\underline{m}_2,
\end{aligned}
$$

$$=: \mathsf{Cov}_{r_1,r_2,1} + \mathsf{Cov}_{r_1,r_2,2}.$$

The proof of "$\mathsf{Cov}_{r_1,r_2,1}$ = (**??**)" is exactly analogous with **?**, it is then omitted. Now, we prove that "$\mathsf{Cov}_{r_1,r_2,2}$ = (**??**)". Note that

$$\oint_{\tilde{\mathcal{C}}_1}\oint_{\tilde{\mathcal{C}}_2} \frac{z_1^{r_1} z_2^{r_2}}{(1+\lambda\underline{m}_1)^2(1+\lambda\underline{m}_2)^2}\,\mathrm{d}\underline{m}_1\,\mathrm{d}\underline{m}_2 = \oint_{\tilde{\mathcal{C}}_1} \frac{z_1^{r_1}}{(1+\lambda\underline{m}_1)^2}\,\mathrm{d}\underline{m}_1 \times \oint_{\tilde{\mathcal{C}}_2} \frac{z_2^{r_2}}{(1+\lambda\underline{m}_2)^2}\,\mathrm{d}\underline{m}_2.$$

By (**??**), we have

$$\begin{aligned}
&\oint_{\tilde{\mathcal{C}}_1} \frac{z_1^{r_1}}{(1+\lambda\underline{m}_1)^2}\,\mathrm{d}\underline{m}_1 \\
&= \oint_{\tilde{\mathcal{C}}_1} \frac{\left(-\frac{1}{\underline{m}_1} + \frac{c\lambda}{1+\lambda\underline{m}_1}\right)^{r_1}}{(1+\lambda\underline{m}_1)^2}\,\mathrm{d}\underline{m}_1 \\
&= (\lambda c)^{r_1} \oint_{\tilde{\mathcal{C}}_1} \left(\frac{1}{1+\lambda\underline{m}_1} + \frac{1-c}{c}\right)^{r_1} \{1-(1+\lambda\underline{m}_1)\}^{-r_1}(1+\lambda\underline{m}_1)^{-2}\,\mathrm{d}\underline{m}_1 \\
&= (\lambda c)^{r_1} \oint_{\tilde{\mathcal{C}}_1} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1}\left(\frac{1-c}{c}\right)^{k_1}(1+\lambda\underline{m}_1)^{k_1-r_1} \sum_{j=0}^{\infty}\binom{r_1+j-1}{j}(1+\lambda\underline{m}_1)^{j}(1+\lambda\underline{m}_1)^{-2}\,\mathrm{d}\underline{m}_1 \\
&= (\lambda c)^{r_1} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1}\left(\frac{1-c}{c}\right)^{k_1} \oint_{\tilde{\mathcal{C}}_1} \sum_{j=0}^{\infty}\binom{r_1+j-1}{j}(1+\lambda\underline{m}_1)^{k_1-r_1+j-2}\,\mathrm{d}\underline{m}_1,
\end{aligned}$$

by substitution $\underline{\widetilde{m}}_1 = \lambda\underline{m}_1$, we get

$$\oint_{\tilde{\mathcal{C}}_1} \frac{z_1^{r_1}\,\mathrm{d}\underline{m}_1}{(1+\lambda\underline{m}_1)^2} = \frac{(\lambda c)^{r_1}}{\lambda} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1}\left(\frac{1-c}{c}\right)^{k_1} \oint_{\tilde{\tilde{\mathcal{C}}}_1} \sum_{j=0}^{\infty}\binom{r_1+j-1}{j}(1+\underline{\widetilde{m}}_1)^{k_1-r_1+j-2}\,\mathrm{d}\underline{\widetilde{m}}_1,$$

where $\tilde{\tilde{\mathcal{C}}}_1$ is the $\underline{\widetilde{m}}_1$ contour. For this integral, the pole is -1, we have by residual theorem

$$\oint_{\tilde{\mathcal{C}}_1} \frac{z_1^{r_1}}{(1+\lambda\underline{m}_1)^2}\,\mathrm{d}\underline{m}_1 = \frac{2\pi i}{\lambda}(\lambda c)^{r_1} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1}\left(\frac{1-c}{c}\right)^{k_1}\binom{2r_1-k_1}{r_1-1}.$$

Similarly, we get

$$\oint_{\tilde{\mathcal{C}}_2} \frac{z_2^{r_2}}{(1+\lambda\underline{m}_2)^2}\,\mathrm{d}\underline{m}_2 = \frac{2\pi i}{\lambda}(\lambda c)^{r_2} \sum_{k_2=0}^{r_2} \binom{r_2}{k_2}\left(\frac{1-c}{c}\right)^{k_2}\binom{2r_2-k_2}{r_2-1}.$$

Using the two equations above, we derive (**??**).

S2.14 Tightness of $M_p^{(1)}(z)$

The tightness of $M_p^{(1)}(z)$ is similar to that provided in **?**. It is sufficient to prove the moment condition (12.51) of **?**, i.e.

$$\sup_{n;z_1,z_2\in\mathcal{C}_n} \frac{\mathbb{E}|M_p^{(1)}(z_1) - M_p^{(1)}(z_2)|^2}{|z_1-z_2|^2} \tag{S2.109}$$

is finite.

Before proceeding, we provide some results needed in the proof later. First, moments of $\|\mathbf{D}^{-1}(z)\|$, $\|\mathbf{D}_j^{-1}(z)\|$ and $\|\mathbf{D}_{ij}^{-1}(z)\|$ are bounded in p and $z \in \mathcal{C}_n$. It is easy to see that it is true for $z \in \mathcal{C}_u$ and for $z \in \mathcal{C}_\ell$ if $x_\ell < 0$. For $z \in \mathcal{C}_r$ or, if $x_\ell > 0$, $z \in \mathcal{C}_\ell$, we have from Proposition **??** that

$$\begin{aligned} \mathbb{E}\|\mathbf{D}_j^{-1}(z)\|^m &\leq K_1 + v^{-m} P(\|\mathbf{B}_{(j)}\| \geq \eta_r \text{ or } \lambda_{\min}(\mathbf{B}_{(j)}) \leq \eta_\ell) \\ &\leq K_1 + K_2 n^m \varepsilon^{-m} n^{-\ell} \leq K \end{aligned}$$

for large ℓ, where $\mathbf{B}_{(j)} = \mathbf{B}_p - \mathbf{r}_j \mathbf{r}_j'$. Here η_r is any number between $\lambda(1+\sqrt{c})^2$ and x_r; if $x_\ell > 0$, η_ℓ is any number between x_ℓ and $\lambda(1-\sqrt{c})^2$ and if $x_\ell < 0$, η_ℓ can be any negative number. So for any positive integer m,

$$\max\Big(\mathbb{E}\|\mathbf{D}^{-1}(z)\|^m, \mathbb{E}\|\mathbf{D}_j^{-1}(z)\|^m, \mathbb{E}\|\mathbf{D}_{ij}^{-1}(z)\|^m\Big) \leq K. \tag{S2.110}$$

By the argument above, we can extend Lemma **??** and get

$$\left|\mathbb{E}\Big(a(v) \prod_{\ell=1}^{q} \Big(\mathbf{r}_1' \mathbf{B}_{(\ell)}(v) \mathbf{r}_1 - n^{-1} \mathrm{tr} \mathbf{B}_{(\ell)}(v)\Big)\Big)\right| \leq K n^{-1} \delta_n^{2q-4}, \tag{S2.111}$$

where $\mathbf{B}_\ell(v)$ is independent of $\mathbf{r}_1$ and

$$\max\Big(|a(v)|, \|\mathbf{B}_{(\ell)}(v)\|\Big) \leq K\Big(1 + n^s I_{\{\|\mathbf{B}_p\| \geq \eta_r \text{ or } \lambda_{\min}(\widetilde{\mathbf{B}}) \leq \eta_\ell\}}\Big),$$

with $\widetilde{\mathbf{B}}$ being $\mathbf{B}_{(j)}$ or $\mathbf{B}_p$. By (**??**), we have

$$\mathbb{E}|\varepsilon_j(z)|^m \leq K_m n^{-1} \delta_n^{2m-4}. \tag{S2.112}$$

Let $\gamma_j(z) = \mathbf{r}_j' \mathbf{D}_j^{-1}(z) \mathbf{r}_j - n^{-1} \nu_2 \mathbb{E} \mathrm{tr} \mathbf{D}_j^{-1}(z)$. By Lemma **??**, (**??**) and Hölder's inequality, with similar derivation on page 580 of **?**, we have

$$\mathbb{E}|\gamma_j(z) - \varepsilon_j(z)|^m \leq \frac{K_m}{n^{m/2}}. \tag{S2.113}$$

It follows from (**??**) and (**??**) that

$$\mathbb{E}|\gamma_j(z)|^m \leq K_m n^{-1} \delta_n^{2m-4}, \quad m \geq 2. \tag{S2.114}$$

Next, we prove that $b_p(z)$ is bounded. With (**??**), we have for any $m \geq 1$,

$$\mathbb{E}|\beta_1(z)|^m \leq K_m. \tag{S2.115}$$

Since $b_p(z) = \beta_1(z) + \beta_1(z) b_p(z) \gamma_1(z)$, it is derived from (**??**) and (**??**) that $|b_p(z)| \leq K_1 + K_2 |b_p(z)| n^{-1/2}$. Hence, we have

$$|b_p(z)| \leq \frac{K_1}{1 - K_2 n^{-1/2}} \leq K. \tag{S2.116}$$

With (**??**) – (**??**) and the same approach on Page 581 – 583 of **?**, we can obtain that (**??**) is finite.

S2.15 Proof of Theorem ??

Proof. By direction calculations, we have

$$T = \frac{1}{p}\mathrm{tr}(\mathbf{B}_{p,N}^2) - \frac{2\nu_2}{p-1}\mathrm{tr}(\mathbf{B}_{p,N}) + \frac{p\nu_2^2}{p-1}.$$

Taking $r_1 = 1$ and $r_2 = 2$ in (**??**) – (**??**), we obtain that

$$V_{12} := \mathrm{Cov}\Big(\mathrm{tr}(\mathbf{B}_{p,N}),\ \mathrm{tr}(\mathbf{B}_{p,N}^2)\Big) = 2\lambda c(1+c)(2\lambda^2 + \alpha_1 + \alpha_2).$$

From Theorem **??** and Corollary **??**, we have the following joint CLT:

$$p\begin{pmatrix} \frac{1}{p}\mathrm{tr}(\mathbf{B}_{p,N}^2) - \lambda^2(1+c_N) - \frac{\mu_2}{p} \\ \frac{1}{p}\mathrm{tr}(\mathbf{B}_{p,N}) - \lambda - \frac{\mu_1}{p} \end{pmatrix} \xrightarrow{D} \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix} V_2 & V_{12} \\ V_{12} & V_1 \end{pmatrix}\right).$$

Defining a function $f(x,y) = x - \frac{2p\nu_2}{p-1}y + \frac{p\nu_2^2}{p-1}$, then we have

$$\begin{aligned}
&T = f\left(\frac{1}{p}\mathrm{tr}(\mathbf{B}_{p,N}^2), \frac{1}{p}\mathrm{tr}(\mathbf{B}_{p,N})\right),\\
&\nabla f\left(\lambda^2(1+c_N) + \frac{\mu_1}{p},\ \lambda + \frac{\mu_2}{p}\right) = \left(1,\ -\frac{2p\nu_2}{p-1}\right)',\\
&f\left(\lambda^2(1+c_N) + \frac{\mu_2}{p},\ \lambda + \frac{\mu_1}{p}\right) = \lambda^2 c_N + \frac{\mu_2}{p} - \frac{2\lambda h_1 + \lambda^2}{p-1} =: \mu_T.
\end{aligned}$$

By the Delta method, we obtain that $p(T - \mu_T)$ is asymptotically Gaussian with mean zero and variance

$$\sigma_T^2 = \lim_{p\to\infty}\left\{\left(1,\ -\frac{2p\nu_2}{p-1}\right)\begin{pmatrix} V_2 & V_{12} \\ V_{12} & V_1 \end{pmatrix}\begin{pmatrix} 1 \\ -\frac{2p\nu_2}{p-1} \end{pmatrix}\right\} = 4\lambda^2 V_1 - 4\lambda V_{12} + V_2.$$

This completes the proof of the theorem. □

S2.16 Proof of Theorem ??

Proof. Defining

$$\tilde{\mu}_T = \frac{p\lambda^2}{n_1 - 1} + \frac{\hat{\mu}_2}{p} - \frac{2\hat{\lambda}\hat{h}_1 + \hat{\lambda}^2}{p-1},$$

then we can write

$$p(\widetilde{T} - \hat{\mu}_T) = p(\widetilde{T} - \tilde{\mu}_T) + \frac{p^2}{n_1 - 1}(\lambda^2 - \hat{\lambda}^2). \tag{S2.117}$$

Denoting $\tilde{w}_{ij} := w_{ij}/\mu$ and $w_i^{(k)} \equiv w_i^{(k,p)} := \frac{1}{p}\sum_{j=1}^{p}\tilde{w}_{ij}^k$ for $k = 1, 2$. By the Lindeberg-Feller CLT, we have

$$\sqrt{p}\begin{pmatrix} w_i^{(2)} - \lambda - 1 \\ w_i^{(1)} - 1 \end{pmatrix} \xrightarrow{D} \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix} \mathbb{E}\tilde{w}_{11}^4 - (\lambda+1)^2 & \mathbb{E}\tilde{w}_{11}^3 - \lambda - 1 \\ \mathbb{E}\tilde{w}_{11}^3 - \lambda - 1 & \lambda \end{pmatrix}\right).$$

By Taylor's theorem, we have the following approximation:

$$\hat{\lambda} = \frac{1}{n_2}\sum_{i=1}^{n_2}\hat{\lambda}_i + O_P\left(\frac{1}{p^{3/2}}\right),$$

where

$$\hat{\lambda}_i := w_i^{(2)} - 2w_i^{(1)} + 1 - 2\left(w_i^{(2)} - w_i^{(1)}\right)\left(w_i^{(1)} - 1\right) + \left(3w_i^{(2)} - 2w_i^{(1)}\right)\left(w_i^{(1)} - 1\right)^2.$$

Now, we derive the asymptotic distribution of $\hat{\lambda}_i$. Defining a function $f(x,y) = x - 2y + 1 - 2(x-y)(y-1) + (3x-2y)(y-1)^2$, then we have

$$\hat{\lambda}_i = f\left(w_i^{(2)}, w_i^{(1)}\right), \quad f(\lambda+1,1) = \lambda, \quad \nabla f(\lambda+1,1) = (1, -2\lambda-2)'.$$

Note that

$$\begin{aligned}
&\{\nabla f(\lambda+1,1)\}'\begin{pmatrix}\mathbb{E}\tilde{w}_{11}^4 - (\lambda+1)^2 & \mathbb{E}\tilde{w}_{11}^3 - \lambda - 1\\ \mathbb{E}\tilde{w}_{11}^3 - \lambda - 1 & \lambda\end{pmatrix}\nabla f(\lambda+1,1)\\
&= \mathbb{E}\tilde{w}_{11}^4 - 4(\lambda+1)\mathbb{E}\tilde{w}_{11}^3 + (\lambda+1)^2(4\lambda+3)\\
&= \mathbb{E}(\tilde{w}_{11}-1)^4 - \lambda^2 + h_2 - 2\lambda h_1.
\end{aligned}$$

By the Delta method, we have

$$\sqrt{p}\left(\hat{\lambda}_i - \lambda\right) \xrightarrow{D} \mathcal{N}\left(0, \mathbb{E}(\tilde{w}_{11}-1)^4 - \lambda^2 + h_2 - 2\lambda h_1\right). \tag{S2.118}$$

Since

$$p(\hat{\lambda} - \lambda) = \sqrt{\frac{p}{n_2}}\frac{1}{\sqrt{n_2}}\sum_{i=1}^{n_2}\sqrt{p}\left(\hat{\lambda}_i - \lambda\right) + o_P(1), \tag{S2.119}$$

it is necessary to expand the expectation of $\hat{\lambda}_i$ up to the order $O(p^{-1})$. From direct calculations, we obtain that

$$\begin{aligned}
&\mathbb{E}\tilde{w}_i^{(1)} = 1, \quad \mathbb{E}\tilde{w}_i^{(2)} = \lambda + 1,\\
&\mathbb{E}\tilde{w}_i^{(1)}\left(\tilde{w}_i^{(1)} - 1\right) = \frac{\lambda}{p}, \quad \mathbb{E}\tilde{w}_i^{(2)}\left(\tilde{w}_i^{(1)} - 1\right) = \frac{\mathbb{E}\tilde{w}_{11}^3 - \lambda - 1}{p},\\
&\mathbb{E}\tilde{w}_i^{(1)}\left(\tilde{w}_i^{(1)} - 1\right)^2 = \frac{\lambda}{p} + O(p^{-2}), \quad \mathbb{E}\tilde{w}_i^{(2)}\left(\tilde{w}_i^{(1)} - 1\right)^2 = \frac{\lambda(\lambda+1)}{p} + O(p^{-2}),
\end{aligned}$$

and thus,

$$\mathbb{E}\hat{\lambda}_i = \lambda + \frac{-2\mathbb{E}\tilde{w}_{11}^3 + 3\lambda^2 + 5\lambda + 2}{p} + o_P(p^{-1}) = \lambda + \frac{h_1}{p} + o_P(p^{-1}). \tag{S2.120}$$

From (**??**), (**??**), (**??**) and the fact that $\{\hat{\lambda}_i\}_{i=1}^{n_2}$ are independent, we have

$$p(\hat{\lambda} - \lambda) \xrightarrow{D} \mathcal{N}\left(h_1, c_2\{\mathbb{E}(\tilde{w}_{11}-1)^4 - \lambda^2 + h_2 - 2\lambda h_1\}\right).$$

Using the Delta method again, we have

$$\frac{p^2}{n_1 - 1}(\hat{\lambda}^2 - \lambda^2) \xrightarrow{D} \mathcal{N}\left(2c_1\lambda h_1, 4\lambda^2 c_1^2 c_2\{\mathbb{E}(\tilde{w}_{11}-1)^4 - \lambda^2 + h_2 - 2\lambda h_1\}\right).$$

This, together with (**??**), Theorem **??**, and the fact that $\widetilde{T}$ and $\hat{\lambda}$ are independent, completes the proof of the theorem. $\square$

S3 Simulation of CLT for $M_p(z)$

In this section, we compare the empirical mean and covariance of $M_p(z) = \mathrm{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - pm_{F^{c_n}}(z)$ with their theoretical limits as stated in Proposition **??**. This proposition is a key step for the proof of our main result, Theorem **??**. Readers are referred to Section **??** for more details of $M_p(z)$. We consider two types of data distribution of w_{ij} as follows:

1. w_{ij} follows the exponential distribution with rate parameter 5;
2. w_{ij} follows the Chi-square distribution with degree of freedom 1.

Empirical values of $\mathbb{E}M_p(z)$ and $\mathrm{Cov}(M_p(z_1), M_p(z_2))$ are calculated for various combinations of (p,n) with $p/n = 3/4$ or $p/n = 1$. For each pair of (p,n), 2000 independent replications are used to obtain the empirical values. Table **??** reports the empirical mean of $M_p(z)$ with $z = \pm 3 + 2i$ for both Exp(5) population and $\chi^2(1)$ population. The empirical results of $\mathrm{Cov}(M_p(z_1), M_p(z_2))$ are reported in Tanle **??**. As shown in Tables **??** – **??**, the empirical values of $\mathbb{E}M_p(z)$ and $\mathrm{Cov}(M_p(z_1), M_p(z_2))$ closely match their respective theoretical limits under all scenarios.

Table S.5: Empirical mean of $M_p(z)$ with $z = \mp 3 + 2i$.

			Exp(5)		$\chi^2(1)$	
	p/n	n	-3+2i	3+2i	-3+2i	3+2i
Emp	3/4	100	0.0586+0.0857i	-0.0373-0.249i	0.1405+0.1628i	-0.55-0.2732i
		200	0.0582+0.0858i	-0.0311-0.2526i	0.1459+0.1697i	-0.5761-0.3089i
		300	0.0567+0.0844i	-0.0336-0.2566i	0.1465+0.1712i	-0.5705-0.3212i
		400	0.0596+0.0878i	-0.0352-0.2528i	0.1463+0.172i	-0.5631-0.3465i
Theo			**0.0587+0.0872i**	**-0.029-0.2529i**	**0.15+0.1768i**	**-0.5792-0.3764i**
Emp	5/4	100	0.0547+0.0766i	-0.1069-0.2671i	0.1366+0.1473i	-0.5458-0.1545i
		200	0.0572+0.0793i	-0.1109-0.2757i	0.1395+0.1518i	-0.5847-0.1787i
		300	0.0587+0.0808i	-0.1074-0.2752i	0.1382+0.1511i	-0.5747-0.1934i
		400	0.0559+0.0778i	-0.0949-0.2733i	0.1434+0.1553i	-0.5751-0.1933i
Theo			**0.0578+0.0804i**	**-0.0919-0.2764i**	**0.1432+0.1569i**	**-0.6025-0.2149i**

Table S.6: Empirical covariance between $M_p(z_1)$ and $M_p(z_2)$.

	p/n	n	Exp(5)		$\chi^2(1)$	
			(-3+2i,-1+1i)†	(3+2i,5+1i)	(-3+2i,-1+1i)	(3+2i,5+1i)
Emp	3/4	100	-0.0038+0.0147i	-0.04+0.0035i	0+0.0304i	0.089+0.014i
		200	-0.0041+0.0163i	-0.0418+0.0022i	0.0004+0.0326i	0.117+0.0284i
		300	-0.0043+0.0171i	-0.0446+0.0011i	0+0.0335i	0.1372+0.0294i
		400	-0.0043+0.0168i	-0.0465-0.0003i	0.0002+0.0356i	0.1273+0.036i
Theo			**-0.0044+0.0172i**	**-0.0448-0.0002i**	**0.0006+0.0363i**	**0.1491+0.0524i**
Emp	5/4	100	-0.0032+0.0197i	-0.0483+0.0765i	0.0025+0.0349i	0.0931-0.0373i
		200	-0.0032+0.0196i	-0.0545+0.0763i	0.0032+0.035i	0.0991-0.0406i
		300	-0.0036+0.0212i	-0.0566+0.0708i	0.0026+0.0336i	0.0955-0.0209i
		400	-0.0032+0.02i	-0.0594+0.0742i	0.0038+0.0374i	0.1138-0.0297i
Theo			**-0.0034+0.0206i**	**-0.0624+0.0743i**	**0.0035+0.0388i**	**0.1099-0.0323i**

†This row denotes different combinations of (z_1, z_2).