

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

Qianqian Jiang*, Jiaxin Qiu†, Zeng Li‡

(This version: December 26, 2023)

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. Simulation experiments demonstrate the accuracy of theoretical results.

Contents

1	Introduction	2
2	Main Results	4
2.1	Preliminaries and Notations	4
2.2	Limiting spectral distribution and Extreme eigenvalues	5
2.3	CLT for LSS	7
3	Numerical experiments	9
3.1	Limiting spectral distribution	9
3.2	CLT for LSS	10
4	Proof of Theorem 2.5	12
4.1	Substitution principle	14
4.2	Some important lemmas	15
4.3	CLT for LSS of the centralized sample covariance B_p^0	16

*Southern University of Science and Technology, jiangqq@sustech.edu.cn

†The University of Hong Kong, qiujx@connect.hku.hk

‡Southern University of Science and Technology, liz9@sustech.edu.cn

4.3.1	Step 1: Truncation	16
4.3.2	Step 2: Finite dimensional convergence of $M_p^{(1)}(z)$ in distribution	17
4.3.3	Step 3: Tightness of $M_p^{(1)}(z)$	23
4.3.4	Step 4: Convergence of $M_p^{(2)}(z)$	23
A Some technical lemmas		25
B Proofs		26
B.1	Proof of Theorem 2.3	26
B.2	Proof of Proposition 2.4	27
B.3	Proof of Lemma 4.2	29
B.4	Proof of Lemma 4.3	30
B.5	Proof of Lemma 4.4	33
B.6	Proof of Lemma 4.5	37
B.7	Proof of Lemma 4.9	44
B.8	Proof of Lemma 4.10	47
B.9	Proof of Lemma 4.11	48
B.10	Proof of Lemma 4.12	49
B.11	Proof of Lemma 4.13	52
B.12	Proof of Lemma A.4	54
B.13	Proof of Corollary 2.6	56
B.14	Tightness of $M_p^{(1)}(z)$	62
C Simulation of CLT for $M_p(z)$		64

I 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. [Cao et al. \[2018\]](#) extended the maximum test framework by [Cai et al. \[2014\]](#) 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 statistics, like Hotelling's T^2 -statistic, rely on the sample covariance matrix. For bacterial species correlation, [Faust et al. \[2012\]](#) 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, [Wu et al. \[2011\]](#) 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 [Marčenko and Pastur \[1967\]](#) and [Jonsson \[1982\]](#) 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 [Yin and Krishnaiah \[1983\]](#) and [Silverstein and Bai \[1995\]](#) extended these findings to the sample covariance matrix $n^{-1}\mathbf{X}\mathbf{\Sigma}\mathbf{X}'$ for data with a linear dependence structure. [Zhang \[2007\]](#) 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. [Johnstone \[2001\]](#) 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 [[El Karoui, 2007](#)], Kendall rank correlation coefficient matrices [[Bao, 2019](#)], among others. Considerable attention has also been given to the study of linear functionals of eigenvalues. [Bai and Silverstein \[2004\]](#) 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 [[Gao et al., 2017](#)], and separable product matrices [[Bai et al., 2019](#)]. 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. Jiang [2004] demonstrated that the LSD of sample correlation matrix for i.i.d data is the well-known Marčenko-Pastur law. Gao et al. [2017] 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 2.3 is established following the strategy in Jiang [2004], and we further investigate the extreme eigenvalues in Proposition 2.4. The proof strategy of CLT for LSS in Theorem 2.5 follows the methodologies outlined in Bai and Silverstein [2004] for the sample covariance matrix and Gao et al. [2017] 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 4.2 and Section 4.3), where we develop three crucial technique lemmas (see Lemmas 4.3 - 4.5) essential for the proof. Finally, it is noteworthy that the mean and variance-covariance in Theorem 2.5 differ from those in Bai and Silverstein [2004], and additional terms are present in both the mean and variance-covariance.

The paper is organized as follows. Section 2.2 investigates the LSD and extreme eigenvalues of the sample covariance matrix for HCD. Section 2.3 establishes our main CLT for LSS of the sample covariance matrix for HCD. Section 3 reports numerical studies. Technical proofs and lemmas are relegated to Appendix.

Before moving forward, let us introduce some notations that will be used throughout this paper. We adopt the convention of using regular letters for scalars and using bold-face 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 $\text{tr}(\mathbf{A})$, 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 $\mathbb{P}(|X_n/a_n| \geq C) \leq \varepsilon$ holds for all $n \geq N$; and we write $X_n = o_p(a_n)$ if $\lim_{n \rightarrow \infty} \mathbb{P}(|X_n/a_n| \geq \varepsilon) = 0$ holds for any $\varepsilon > 0$; and we write $X_n \xrightarrow{a.s.} a$ ($X_n \xrightarrow{i.p.} a$, resp.) if X_n converges almost surely (in probability, resp.) to a . We denote by C and K are constants, which may be different from line to line.

2 Main Results

2.1 Preliminaries and Notations

Let $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$ denote the $n \times p$ observed data matrix, where each \mathbf{x}_i represents compositions that lie in the $(p - 1)$ -dimensional simplex $\mathcal{S}^{p-1} = \{(y_1, \dots, y_p) : \sum_{j=1}^p y_j = 1, y_j \geq 0\}$. 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

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

The unbiased sample covariance matrix of \mathbf{X}_n is defined by $\mathbf{S}_{n,N} = \frac{1}{N}(\mathbf{C}_n \mathbf{X}_n)'(\mathbf{C}_n \mathbf{X}_n)$, where $\mathbf{C}_n = \mathbf{I}_n - (1/n)\mathbf{1}_n \mathbf{1}_n'$, $\mathbf{1}_n$ is a n -dimensional vector of all ones, and $N = n - 1$ is the adjusted sample size.

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 $\lambda_1, \dots, \lambda_p$, 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 $p \rightarrow \infty$, 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

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

Many classes of statistics related to the eigenvalues of the sample covariance matrix $\mathbf{B}_{p,N}$ 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* of $\mathbf{B}_{p,N}$ given by

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

where $\lambda_i, i = 1, \dots, p$, 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 2.3), the behavior of extreme eigenvalues (see, Proposition 2.4), and the CLT for LSS (see, Theorem 2.5).

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

$$\mathbf{X}_n = \begin{pmatrix} \frac{1}{\sum_{j=1}^p w_{1j}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sum_{j=1}^p w_{nj}} \end{pmatrix}_{n \times n} \begin{pmatrix} w_{11} & \cdots & w_{1p} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{np} \end{pmatrix}_{n \times p} := \mathbf{\Lambda}_n \mathbf{W}_n.$$

Here we assume \mathbf{W}_n has i.i.d. components w_{ij} satisfying $\mathbb{E}(w_{ij}) = \mu > 0$, $\text{Var}(w_{ij}) = \sigma^2$. Recall that the Pearson sample correlation matrix for \mathbf{W}_n expressed as

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

where $\|\mathbf{w}_j\|_2 = \left\{ (1/n) \sum_{i=1}^n (w_{ij} - \bar{w}_j)^2 \right\}^{1/2}$, $\bar{w}_j = (1/n) \sum_{i=1}^n w_{ij}$, $j = 1, \dots, p$, and

$$\tilde{\mathbf{X}}_n = \begin{pmatrix} w_{11} & \cdots & w_{1p} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{np} \end{pmatrix}_{n \times p} \begin{pmatrix} \|\mathbf{w}_1\|_2^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|\mathbf{w}_p\|_2^{-1} \end{pmatrix}_{p \times p} := \mathbf{W}_n \tilde{\mathbf{\Lambda}}_p.$$

It can be seen that the normalizing matrix $\mathbf{\Lambda}_n$ of $\mathbf{S}_{n,N}$ is very similar to $\tilde{\mathbf{\Lambda}}_p$ of \mathbf{R}_n . The former uses $(\sum_{j=1}^p w_{ij})^{-1}$ for normalization, while the latter utilizes $\|\mathbf{w}_j\|_2^{-1}$. 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 linear functionals of eigenvalues 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 $\mathbb{E}w_{11} = \mu > 0$, $\mathbb{E}(w_{11} - \mu)^2 = \sigma^2$ and $\mathbb{E}|w_{11} - \mu|^4 < \infty$.

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

Theorem 2.3. Under Assumptions 2.1 and 2.2, with probability one, the ESD of $\mathbf{B}_{p,N}$ converges weakly to a deterministic probability distribution with a density function

$$f(x) = \begin{cases} \frac{\mu^2}{2\pi c \sigma^2 x} \sqrt{(b-x)(x-a)}, & \text{if } x \in [a, b], \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

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

The proof of Theorem 2.3 is postponed to Appendix.

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 $z \in \mathbb{C}^+ = \{z : \Im(z) > 0\}$, by Theorem 2.3 the Stieltjes transform $m(z) = m_{F^c}(z)$ is the unique solution of $m = \frac{1}{\sigma^2/\mu^2(1-c-zm)-z}$ in the set $\{m \in \mathbb{C} : \frac{1-c}{z} + \underline{m}(z) \in \mathbb{C}^+\}$. Define $\underline{m}(z)$ to be 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 $\{\underline{m} \in \mathbb{C} : \frac{1-c}{z} + \underline{m}(z) \in \mathbb{C}^+\}$ of the equation

$$z = -\frac{1}{\underline{m}(z)} + \frac{c\sigma^2/\mu^2}{1 + \sigma^2/\mu^2 \underline{m}(z)}, \quad z \in \mathbb{C}^+. \quad (2)$$

Proposition 2.4. Under Assumptions 2.1 and 2.2, we have

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

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 $\ell > 0$, $\eta_1 > \frac{\sigma^2}{\mu^2}(1 + \sqrt{c})^2$ and $0 < \eta_2 < \frac{\sigma^2}{\mu^2}(1 - \sqrt{c})^2 \cdot \mathbb{I}_{\{0 < c < 1\}}$, we have

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

The proof of Proposition 2.4 is postponed to Appendix.

Remark 1. The LSD has support $\left[\frac{\sigma^2}{\mu^2}(1 - \sqrt{c})^2, \frac{\sigma^2}{\mu^2}(1 + \sqrt{c})^2 \right]$, where it has a density function. The results of extreme eigenvalues find application in locating eigenvalues of the population covariance matrix and in proving the CLT for LSS. Proposition 2.4 shows that with probability 1, there are no eigenvalues of $\mathbf{B}_{p,N}$ outside the support of LSD under Assumptions 2.1-2.2. These lemmas are crucial for applying the Cauchy integral formula (see, equation (6)) and proving tightness.

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 LSD of $\mathbf{B}_{p,N}$, i.e. $\int f(x) dF^c(x)$. In this section, we explore second order fluctuation of $\frac{1}{p} \sum_{i=1}^p f(\lambda_i)$ describing how such LSS converges to its first order limit. Define

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

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 2.1 – 2.2 and the analyticity of f , the rate $\int f(x) d\{F^{\mathbf{B}_{p,N}}(x) - F^{c_N}(x)\}$, approaching zero is essentially $1/n$ and $G_{p,N}(f)$ convergence weakly to a Gaussian variable. Before presenting the main result, we first recall some notation. Let $\underline{m}(z)$ be the Stieltjes transform of the LSD $F^c(x)$ and $\underline{m}(z)$ be the Stieltjes transform of the companion LSD $F^c(x)$. Furthermore, we define $\underline{m}'(z)$ as the first derivative of $\underline{m}(z)$ with respect to z throughout the rest of this paper. The main result is stated in the following theorem.

Theorem 2.5. *Under Assumptions 2.1 and 2.2, let f_1, f_2, \dots, f_k be functions on \mathbb{R} and analytic on an open interval containing*

$$\left[\frac{\sigma^2}{\mu^2} (1 - \sqrt{c})^2, \frac{\sigma^2}{\mu^2} (1 + \sqrt{c})^2 \right]. \quad (3)$$

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

$$\begin{aligned} \mathbb{E}X_f &= \frac{1}{2\pi i} \oint_{\mathcal{C}} c \frac{\sigma^4}{\mu^4} f(z) \underline{m}^3(z) \left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right\}^{-3} \left[1 - c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right\}^{-2} \right]^{-2} \\ &\quad - \frac{1}{2\pi i} \oint_{\mathcal{C}} f(z) \underline{m}(z) \left[1 - c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right\}^{-2} \right]^{-1} \\ &\quad \times z \underline{m}(z) \left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right\}^{-1} \times \left\{ h_1 \underline{m}(z) + \frac{\sigma^2}{\mu^2} \underline{m}(z) + \frac{\sigma^2}{\mu^2} \frac{1}{z} \right\} dz \\ &\quad - \frac{1}{2\pi i} \oint_{\mathcal{C}} c f(z) z^2 \underline{m}^3(z) \left[1 - c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right\}^{-2} \right]^{-1} \\ &\quad \times \left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right\}^{-1} \left\{ (\alpha_1 + \alpha_2) \underline{m}^2(z) + 2 \frac{\sigma^4}{\mu^4} \underline{m}'(z) \right\} dz, \end{aligned} \quad (4)$$

and covariance function

$$\begin{aligned} \text{Cov}(X_f, X_g) &= -\frac{1}{2\pi^2} \oint_{\mathcal{C}_1} \oint_{\mathcal{C}_2} \frac{f(z_1)g(z_2)}{\{\underline{m}(z_1) - \underline{m}(z_2)\}^2} d\underline{m}(z_1) d\underline{m}(z_2) \\ &\quad - \frac{c(\alpha_1 + \alpha_2)}{4\pi^2} \oint_{\mathcal{C}_1} \oint_{\mathcal{C}_2} \frac{f(z_1)g(z_2)}{\left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z_1) \right\}^2 \left\{ 1 + \frac{\sigma^2}{\mu^2} \underline{m}(z_2) \right\}^2} d\underline{m}(z_1) d\underline{m}(z_2), \end{aligned} \quad (5)$$

where

$$\alpha_1 = \lim_{p \rightarrow \infty} \left[\mathbb{E} \left(\frac{w_{11}}{w_1} - 1 \right)^4 - 3 \mathbb{E} \left(\frac{w_{11}}{w_1} - 1 \right)^2 \left(\frac{w_{12}}{w_1} - 1 \right)^2 \right],$$

$$\alpha_2 = \lim_{p \rightarrow \infty} p \left[\mathbb{E} \left(\frac{w_{11}}{\bar{w}_1} - 1 \right)^2 \left(\frac{w_{12}}{\bar{w}_1} - 1 \right)^2 - \left\{ \mathbb{E} \left(\frac{w_{11}}{\bar{w}_1} - 1 \right)^2 \right\}^2 \right],$$

$$h_1 = \lim_{p \rightarrow \infty} p \left[\mathbb{E} \left(\frac{w_{11}}{\bar{w}_1} - 1 \right)^2 - \frac{\sigma^2}{\mu^2} \right],$$

and $\bar{w}_1 = \sum_{j=1}^p w_{1j}/p$. The contours $\mathcal{C}, \mathcal{C}_1, \mathcal{C}_2$ in (4) and (5) are closed and taken in the positive direction in the complex plane, each enclosing the support of $F^c(x)$, i.e., $[\frac{\sigma^2}{\mu^2}(1 - \sqrt{c}), \frac{\sigma^2}{\mu^2}(1 + \sqrt{c})^2]$.

Remark 2. The emergence of parameters h_1 and α_2 in our limiting mean and covariance functions may appear unconventional, but it stems from the unique aspects of our analysis. This phenomenon arises from the non-negligible influence of terms h_1/p and α_2/p in the approximation of $\mathbb{E}(\frac{w_{11}}{\bar{w}_1} - 1)^2$ and $\mathbb{E}(\frac{w_{11}}{\bar{w}_1} - 1)^2(\frac{w_{12}}{\bar{w}_1} - 1)^2$, driven by the multiplication by p in the CLT (refer to Lemma 4.3 and Lemma 4.5). Furthermore, our results introduce parameters α_1 and α_2 in place of conventional parameters like $\mathbb{E}|w_{11}|^4 - 3$ and $\mathbb{E}|w_{11}|^4 - 1$ in the limiting mean and covariance functions of the sample correlation matrix in Gao et al. [2017]. Remarkably, our findings also bring forth a novel parameter, h_1 , in the mean function, setting our results apart from conventional approaches.

Applying Theorem 2.5 to three polynomial functions, we obtain the following corollary. The proof of Theorem 2.5 is postponed to Section 4, and detailed calculations in these applications are postponed to Appendix.

Corollary 2.6. *Under conditions and notations in Theorem 2.5, let $f_i = x^i$ for $i = 1, 2, 3$, we have*

$$G_p(f_1) = \text{tr}(\mathbf{B}_{p, \mathcal{N}}) - p \frac{\sigma^2}{\mu^2} \xrightarrow{d} \mathcal{N}(\mu_1, V_1),$$

$$G_p(f_2) = \text{tr}(\mathbf{B}_{p, \mathcal{N}}^2) - p(1 + c_{\mathcal{N}}) \left(\frac{\sigma^2}{\mu^2} \right)^2 \xrightarrow{d} \mathcal{N}(\mu_2, V_2),$$

$$G_p(f_3) = \text{tr}(\mathbf{B}_{p, \mathcal{N}}^3) - p(1 + 3c_{\mathcal{N}} + c_{\mathcal{N}}^2) \left(\frac{\sigma^2}{\mu^2} \right)^3 \xrightarrow{d} \mathcal{N}(\mu_3, V_3),$$

where $c_{\mathcal{N}} = \frac{p}{\mathcal{N}}$, and

$$\mu_1 = h_1, \quad \mu_2 = (1 + c) \left(\frac{\sigma^2}{\mu^2} \right)^2 + 2(1 + c) \frac{\sigma^2}{\mu^2} h_1 + c(\alpha_1 + \alpha_2),$$

$$\mu_3 = (2 + 6c + 3c^2) \left(\frac{\sigma^2}{\mu^2} \right)^3 + 3(1 + 3c + c^2) \left(\frac{\sigma^2}{\mu^2} \right)^2 h_1 + 3c(1 + c) \frac{\sigma^2}{\mu^2} (\alpha_1 + \alpha_2),$$

$$V_1 = 2c \left(\frac{\sigma^2}{\mu^2} \right)^2 + c(\alpha_1 + \alpha_2),$$

$$V_2 = 4c(2 + c)(1 + 2c) \left(\frac{\sigma^2}{\mu^2} \right)^4 + 4c(1 + c)^2 \left(\frac{\sigma^2}{\mu^2} \right)^2 (\alpha_1 + \alpha_2),$$

$$V_3 = 6c(1 + 6c + 3c^2)(3 + 6c + c^2) \left(\frac{\sigma^2}{\mu^2} \right)^6 + 9c(1 + 3c + c^2)^2 \left(\frac{\sigma^2}{\mu^2} \right)^4 (\alpha_1 + \alpha_2).$$

3 Numerical experiments

3.1 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 2.3. Compositional data $\{x_{ij}\}_{1 \leq i \leq n, 1 \leq j \leq p}$ 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 $TN(0,1;0,10)$, 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 1 – 2. Figures 1 – 2 reveal that all histograms align with their theoretical limits, affirming the accuracy of our theoretical results.

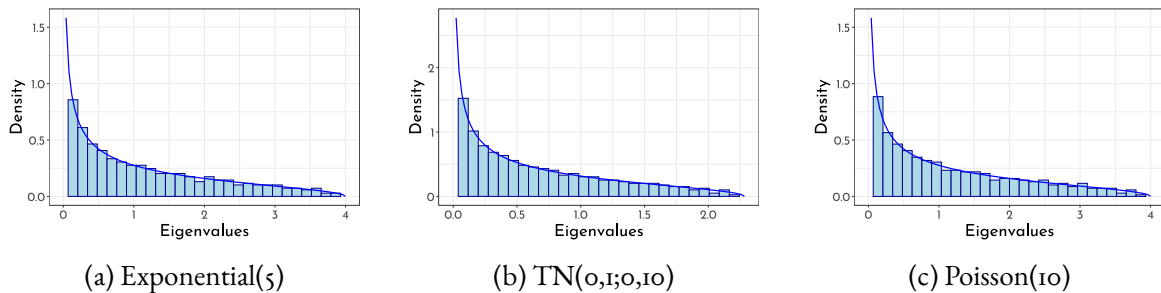


Figure 1: Histograms of sample eigenvalues of $\mathbf{B}_{p,N}$ with $(p, n) = (500,500)$. The curves are density functions of their corresponding limiting spectral distribution.

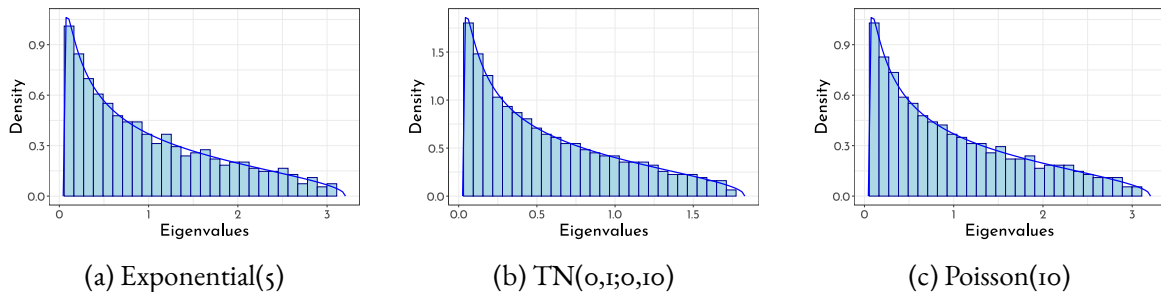


Figure 2: Histograms of sample eigenvalues of $\mathbf{B}_{p,N}$ with $(p, n) = (500,800)$. The curves are density functions of their corresponding limiting spectral distribution.

3.2 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 2.6.

In the following, we present the numerical simulation of CLT for LSS. First, we compare the empirical mean and variance of $G_{p,N}(x^r) = \text{tr}(\mathbf{B}_{p,N}^r) - p \int x^r dF^{cN}(x)$, $r = 1,2,3$, with their corresponding theoretical limits in Corollary 2.6. 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 $\{G_{p,N}(x^r)\}$, $r = 1,2,3$, 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 1 – 2 report the empirical results for Exp(5) population and $\chi^2(1)$ population, respectively. As shown in Tables 1 – 2, 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, $\bar{G}_{p,N}(x^r) = (G_{p,N}(x^r) - \mu_r)/\sqrt{V_r}$, $r = 1,2,3$, where μ_r and V_r are defined in Corollary 2.6, and compare them with the standard normal density. Figures 3 and 4 depict the histograms of $\bar{G}_{p,N}(x^r)$ 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 3 – 4 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 $w_{ij} \sim \text{Exp}(5)$.

		$G_{p,N}(x)$		$G_{p,N}(x^2)$		$G_{p,N}(x^3)$		
	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	

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

		$G_{p,N}(x)$		$G_{p,N}(x^2)$		$G_{p,N}(x^3)$		
	p/n	n	mean	var	mean	var	mean	var
Emp	3/4	100	-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

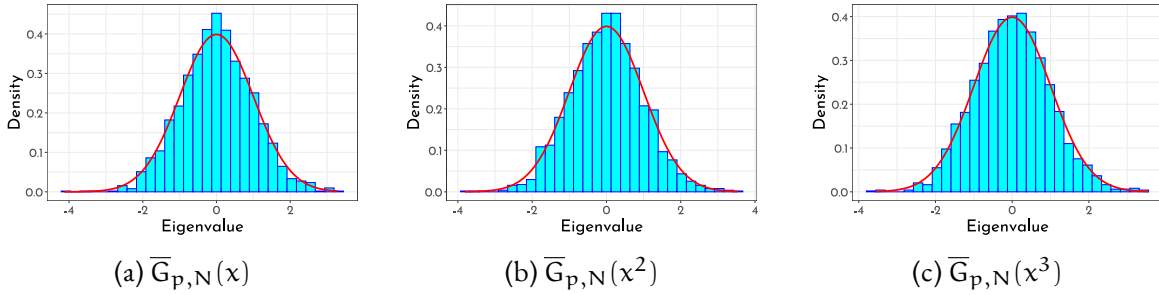


Figure 3: Histograms of normalized LSS $\bar{G}_{p,N}(x^r) = (G_{p,N}(x^r) - \mu_r)/\sqrt{V_r}$, $r = 1,2,3$, with $w_{ij} \sim \text{Exp}(5)$ and $p = n = 400$. The curves are density functions of the standard normal distribution.

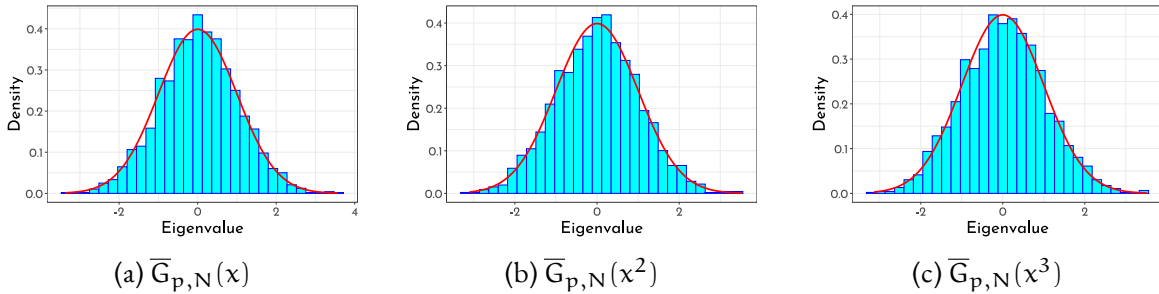


Figure 4: Histograms of normalized LSS $\bar{G}_{p,N}(x^r) = (G_{p,N}(x^r) - \mu_r)/\sqrt{V_r}$, $r = 1,2,3$, with $w_{ij} \sim \chi^2(1)$ and $p = 300$, $n = 400$. The curves are density functions of the standard normal distribution.

4 Proof of Theorem 2.5

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 4.1, where

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

and $\mathbf{Y}_n = (y_{ij})_{n \times p}$, $y_{ij} = \frac{w_{ij}}{\bar{w}_i} - 1$ and $\bar{w}_i = \frac{1}{p} \sum_{j=1}^p w_{ij}$. 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 the centralized sample covariance \mathbf{B}_p^0 share the same CLT (see, Section 4.1). The general strategy of the main proof of Theorem 2.5 is explained in the following and four major steps of the general strategy are presented in Section 4.3.

The general strategy of the proof follows the method established in Bai and Silverstein [2004] and Gao et al. [2017], 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 Jiang [2004] to establish the LSD of $\mathbf{B}_{p,N}$ in Theorem 2.3. Then, we develop Proposition 2.4 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 (6) and proving tightness. Given that compositional data $x_{ij} = w_{ij} / \sum_{j=1}^p w_{ij}$ 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 [Zheng et al., 2015] 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 4.1). 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 $\mathbf{r}_i = \frac{1}{\sqrt{n}} (\frac{w_{i1}}{\bar{w}_i} - 1, \dots, \frac{w_{ip}}{\bar{w}_i} - 1)'$, $i = 1, 2, \dots, n$, 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

$$\int f(x) dG(x) = -\frac{1}{2\pi i} \oint_{\mathcal{C}} f(z) m_G(z) dz \quad (6)$$

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, $G(x) = p(\mathbb{F}^{\mathbf{B}_p^0}(x) - \mathbb{F}^{c_n}(x))$. Therefore, the problem of finding the limiting distribution reduces to the study of $M_p(z)$ defined as follows:

$$\begin{aligned} M_p(z) &= p [\mathbf{m}_p(z) - \mathbf{m}_p^0(z)] = n [\underline{\mathbf{m}}_p(z) - \underline{\mathbf{m}}_p^0(z)], \\ \mathbf{m}_p(z) &= \mathbf{m}_{\mathbb{F}^{\mathbf{B}_p^0}}(z) = \frac{1}{p} \text{tr} [(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1}], \quad \mathbf{m}_p^0(z) = \mathbf{m}_{\mathbb{F}^{c_n}}(z), \\ \underline{\mathbf{m}}_p(z) &= \mathbf{m}_{\mathbb{F}^{\mathbf{B}_p^0}}(z) = \frac{1}{p} \text{tr} [(\mathbf{B}_p^0 - z\mathbf{I}_n)^{-1}], \quad \underline{\mathbf{m}}_p^0(z) = \underline{\mathbf{m}}_{\mathbb{F}^{c_n}}(z), \\ \mathbf{B}_p^0 &= p^2 \underline{\mathbf{S}}_n^0 = \frac{p^2}{n} (\mathbf{X}_n - \mathbb{E}\mathbf{X}_n)(\mathbf{X}_n - \mathbb{E}\mathbf{X}_n)'. \end{aligned}$$

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 2.3, Proposition 2.4). Then the contour \mathcal{C} can be appropriately chosen. Moreover, as in Bai and Silverstein [2004], by Proposition 2.4, 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 $\frac{\sigma^2}{\mu^2}(1 + \sqrt{c})^2$. Let x_l be any negative number if the left endpoint of (3) is zero. Otherwise we choose $x_l \in (0, \frac{\sigma^2}{\mu^2}(1 - \sqrt{c})^2)$. 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 $\alpha \in (0, 1)$, $\varepsilon_n \geq n^{-\alpha}$. 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 $\mathcal{C}_r = \{x_r + iv : v \in [n^{-1}\varepsilon, v_0]\}$. 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 4.1. *Under Assumption 2.1 and 2.2, then $\widehat{M}_p(\cdot)$ converges weakly to a two-dimensional Gaussian process $M(\cdot)$ for $z \in \mathcal{C}$, with means*

$$\begin{aligned} \mathbb{E}M(z) = & -\underline{m}(z) \left[1 - c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right)^{-2} \right]^{-1} \\ & \times \left[-z \underline{m}(z) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right)^{-1} \times \left(h_1 \underline{m}(z) + \frac{\sigma^2}{\mu^2} \underline{m}(z) + \frac{\sigma^2}{\mu^2} \frac{1}{z} \right) \right. \\ & - cz^2 \underline{m}^2(z) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right)^{-1} \left(\alpha_1 \underline{m}^2(z) + \alpha_2 \underline{m}^2(z) + 2 \frac{\sigma^4}{\mu^4} \underline{m}'(z) \right) \\ & \left. + c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right)^{-3} \left(1 - c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right)^{-2} \right)^{-1} \right], \quad (7) \end{aligned}$$

and covariance function

$$\begin{aligned} \text{Cov}(M(z_1), M(z_2)) = & 2 \left[\frac{\underline{m}'(z_1) \underline{m}'(z_2)}{(\underline{m}(z_1) - \underline{m}(z_2))^2} - \frac{1}{(z_1 - z_2)^2} \right] \\ & + c(\alpha_1 + \alpha_2) \times \frac{\underline{m}'(z_1) \underline{m}'(z_2)}{(1 + \sigma^2/\mu^2 \underline{m}(z_1))^2 (1 + \sigma^2/\mu^2 \underline{m}(z_2))^2}. \quad (8) \end{aligned}$$

Now we explain how Theorem 2.5 follows from the above proposition. As in Bai and Silverstein [2004], with probability 1, $\left| \int f(z)(M_p(z) - \widehat{M}_p(z)) dz \right| \rightarrow 0$ as $n \rightarrow \infty$. Combining this observation with equation (6), Theorem 2.5 follows from Proposition 4.1. To prove Proposition 4.1, 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

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

The random part contributes to the covariance function and the deterministic part contributes to the mean function. By Theorem 8.1 in Billingsley [1968], the proof of Proposition 4.1 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 (8).

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 (7) 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 4.1 and 4.2 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 .

4.1 Substitution principle

By the Cauchy integral formula, we have

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

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

$$\begin{aligned} m_N^0(z) &\equiv m_{Fc_N}(z) = \frac{1}{\sigma^2/\mu^2(1 - c_N - c_N z m_N^0) - z}, \\ \underline{m}_N^0(z) &\equiv \underline{m}_{Fc_N}(z) = -\frac{1 - c_N}{z} + c_N m_N^0(z), \\ z &= -\frac{1}{\underline{m}_N^0(z)} + c_N \frac{\sigma^2/\mu^2}{1 + \sigma^2/\mu^2 \underline{m}_N^0(z)} \end{aligned}$$

with $c_N = \frac{p}{N}$. To obtain the asymptotic distribution of $G_{p,N}(f)$, it is necessary to find the asymptotic distribution of $\text{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - pm_N^0(z)$. To achieve this, we derive the following Lemma 4.2 whose proof is postponed to Appendix.

Lemma 4.2. *Under conditions and notations in Theorem 2.5, as $n \rightarrow \infty$,*

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

By Lemma 4.2, the asymptotic distribution of $G_{p,N}(f)$ is identical to that of $G_p^0(f)$, i.e.,

$$G_{p,N}(f) = \sum_{i=1}^p f(\lambda_i(\mathbf{B}_{p,N})) - p \int f(x) dF^{c_N}(x) \implies \mathcal{N}(m(f), v(f)),$$

$$G_p^0(f) = \sum_{i=1}^p f(\lambda_i(\mathbf{B}_p^0)) - p \int f(x) dF^{c_n}(x) \implies \mathcal{N}(m(f), v(f)),$$

a Gaussian distribution whose parameters $m(f)$ and $v(f)$ depend only on the LSD $F^c(x)$ and f , where

$$G_p^0(f) = -\frac{1}{2\pi i} \oint_{\mathcal{C}} f(z) [\text{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - pm_n^0(z)] dz$$

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

4.2 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 4.3 is crafted to estimate essential parameters, facilitating the derivation of estimates of any order. Concerning v_2 and v_{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 $B_p^c(\epsilon)$ decays polynomially to 0 and leverage Taylor expansion on the event $B_p(\epsilon) = \{\omega : |\bar{w}_i - \mu| \leq \epsilon, \bar{w}_i = \sum_{j=1}^p w_{ij}/p\}$ 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 4.4 establishes the former step, essential for converting them into the corresponding traces, while Lemma 4.5 establishes the latter step, enabling the application of CLT for martingale differences. Both Lemma 4.4 and Lemma 4.5 heavily hinge on the estimation of parameters v_2, v_4 , and v_{12} in Lemma 4.3. The proof of Lemmas 4.3 – 4.5 are postponed to Appendix.

Lemma 4.3. *Suppose that $\mathbf{w} = (w_1, \dots, w_p)'$ has i.i.d. entries with $\mathbb{E}w_1 = \mu$, $\mathbb{E}(w_1 - \mu)^2 = \sigma^2$, and $\mathbb{E}|w_1 - \mu|^4 < \infty$, let $\bar{w} = \frac{1}{p} \sum_{j=1}^p w_j$, then there exists a constant $K > 0$, such that for any $0 < \epsilon < 1/2$ and $p > 0$,*

$$v_2 := \mathbb{E}\left(\frac{w_1}{\bar{w}} - 1\right)^2 = \mathbb{E}\left(\frac{w_1}{\mu} - 1\right)^2 + \frac{1}{p}h_1 + o(p^{-1}),$$

$$v_{12} := \mathbb{E}\left(\frac{w_1}{\bar{w}} - 1\right)^2 \left(\frac{w_2}{\bar{w}} - 1\right)^2 = \left[\mathbb{E}\left(\frac{w_1}{\mu} - 1\right)^2\right]^2 + \frac{1}{p}h_2 + o(p^{-1}),$$

$$v_4 := \mathbb{E}\left(\frac{w_1}{\bar{w}} - 1\right)^4 = \mathbb{E}\left(\frac{w_1}{\mu} - 1\right)^4 + o(1),$$

where

$$h_1 = -2\frac{\mathbb{E}w_1^3}{\mu^3} + 3\left(\frac{\sigma^2}{\mu^2}\right)^2 + 5\frac{\sigma^2}{\mu^2} + 2, h_2 = -8\frac{\sigma^2}{\mu^2}\frac{\mathbb{E}w_1^3}{\mu^3} + 10\left(\frac{\sigma^2}{\mu^2}\right)^3 + 22\left(\frac{\sigma^2}{\mu^2}\right)^2 + 8\frac{\sigma^2}{\mu^2}.$$

Lemma 4.4. Suppose that $\mathbf{w} = (w_1, \dots, w_p)'$ 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 $q \geq 2$, we have there is a positive constant K_q depending on q such that

$$\mathbb{E} \left| \mathbf{r}' \mathbf{A} \mathbf{r} - \frac{1}{n} \nu_2 \text{tr} \mathbf{A} \right|^q \leq K_q \left[n^{-q} \left((\mathbb{E} |w_1|^4 \text{tr}(\mathbf{A} \mathbf{A}'))^{q/2} + \mathbb{E} |w_1|^{2q} \text{tr}(\mathbf{A} \mathbf{A}')^{q/2} \right) + n^q \mathbb{P}(\mathbf{B}_p^c(\epsilon)) \|\mathbf{A}\|^q + n^{-q} \|\mathbf{A}^q\| h_1^q \right],$$

where $\mathbf{r} = \frac{1}{\sqrt{n}}(w_1/\bar{w} - 1, \dots, w_p/\bar{w} - 1)'$, h_1 is defined in Lemma 4.3, $\mathbf{B}_p(\epsilon) = \{\omega : |\bar{w} - \omega| \leq \epsilon, \bar{w} = \sum_{j=1}^p w_j/p\}$, and

$$\mathbb{P}(\mathbf{B}_p^c(\epsilon)) \leq K \epsilon^{-kq_1} (\sigma^{kq_1} p^{-kq_1/2} + p^{-kq_1+1} \mathbb{E} |w_1|^{kq_1}), \quad (9)$$

in which $\epsilon, k, q_1 > 0$ are constants. Furthermore, if $\|\mathbf{A}\| \leq K$ and $|w_j - \mu| < \delta_n \sqrt{n}$ for all $j = 1, \dots, p$, then, for any $q \geq 2$,

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

Lemma 4.5. Suppose that $\mathbf{w} = (w_1, \dots, w_p)'$ 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 $\|\mathbf{A}\| \leq K$ and $\|\mathbf{B}\| \leq K$, then

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

4.3 CLT for LSS of the centralized sample covariance \mathbf{B}_p^0

4.3.1 Step 1: Truncation

We begin the proof of Proposition 4.1 with the replacement of the entries of \mathbf{W}_n with truncated variables. We can choose a positive sequence of $\{\delta_n\}$ such that

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

Let $\hat{\mathbf{B}}_p^0 = \frac{p^2}{n} (\hat{\mathbf{X}}_n - \mathbb{E} \hat{\mathbf{X}}_n)' (\hat{\mathbf{X}}_n - \mathbb{E} \hat{\mathbf{X}}_n)$, where $\hat{\mathbf{W}}_n$ is $n \times p$ matrix having $\hat{w}_{ij} = w_{ij} I_{\{|w_{ij} - \mu| < \delta_n \sqrt{n}\}}$. We then have

$$\begin{aligned} \mathbb{P}(\mathbf{B}_p^0 \neq \hat{\mathbf{B}}_p^0) &\leq \mathbb{P} \left(\bigcup_{i \leq n, j \leq p} (|w_{ij} - \mu| \geq \delta_n \sqrt{n}) \right) \leq np \cdot \mathbb{P}(|w_{11} - \mu| \geq \delta_n \sqrt{n}) \\ &\leq K \delta_n^{-4} \int_{\{|w_{11} - \mu| \geq \delta_n \sqrt{n}\}} |w_{11}|^4 = o(1). \end{aligned}$$

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

$$\int f_j(x) dG_p^0(x) = \int f_j(x) d\hat{G}_p^0(x) + o_p(1).$$

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

$$|w_{ij} - \mu| < \delta_n \sqrt{n}, \quad \mathbb{E} w_{ij} = \mu > 0, \quad \mathbb{E} |w_{ij} - \mu|^2 = \sigma^2, \quad \mathbb{E} |w_{ij} - \mu|^4 < \infty. \quad (10)$$

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

Lemma 4.6. *Under conditions and notations in Theorem 2.5, as $p \rightarrow \infty$, for any set of r points $\{z_1, z_2, \dots, z_r\} \cup \mathcal{C}$, the random vector $(M_p^{(1)}(z_1), \dots, M_p^{(1)}(z_r))$ converges weakly to a r -dimensional centered Gaussian distribution with covariance function in (8).*

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 α_j , 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 decompose the random part $M_p^{(1)}(z)$ as a sum of martingale difference, which is given in (18). Then, we apply the martingale CLT (Lemma A.3) to obtain the asymptotic distribution of $M_p^{(1)}(z)$. Details of these two steps are provided in the following two parts.

Part I: 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

$$\begin{aligned} \mathbf{r}_j &= \frac{1}{\sqrt{n}} \left(\frac{w_{j1}}{\bar{w}_j} - 1, \dots, \frac{w_{jp}}{\bar{w}_j} - 1 \right)', \quad \mathbf{D}(z) = \mathbf{B}_p^0 - z\mathbf{I}_p, \quad \mathbf{D}_j(z) = \mathbf{D}(z) - \mathbf{r}_j \mathbf{r}_j', \\ \beta_j(z) &= \frac{1}{1 + \mathbf{r}_j' \mathbf{D}_j^{-1}(z) \mathbf{r}_j}, \quad \bar{\beta}_j(z) = \frac{1}{1 + \frac{1}{n} v_2 \text{tr} \mathbf{D}_j^{-1}(z)}, \quad b_p(z) = \frac{1}{1 + \frac{1}{n} v_2 \mathbb{E} \text{tr} \mathbf{D}_1^{-1}(z)}, \end{aligned}$$

$\varepsilon_j(z) = \mathbf{r}_j' \mathbf{D}_j^{-1}(z) \mathbf{r}_j - \frac{1}{n} v_2 \text{tr} \mathbf{D}_j^{-1}(z)$ and $\delta_j(z) = \mathbf{r}_j' \mathbf{D}_j^{-2}(z) \mathbf{r}_j - \frac{1}{n} v_2 \text{tr} \mathbf{D}_j^{-2}(z) = \frac{d}{dz} \varepsilon_j(z)$. By Lemma 4.4, we have, for any $r \geq 2$,

$$\mathbb{E} |\varepsilon_j(z)|^r \leq \frac{K}{v^{2r}} n^{-1} \delta_n^{2r-4} \quad \text{and} \quad \mathbb{E} |\delta_j(z)|^r \leq \frac{K}{v^{2r}} n^{-1} \delta_n^{2r-4}. \quad (11)$$

It is easy to see that

$$\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z) = -\mathbf{D}_j^{-1}(z) \mathbf{r}_j \mathbf{r}_j' \mathbf{D}_j^{-1}(z) \beta_j(z), \quad (12)$$

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 $|\beta_j(z)|$, $|\bar{\beta}_j(z)|$ and $|b_p(z)|$ are bounded by $\frac{|z|}{v}$. We also get that for any j ,

$$\mathbb{E}(\mathbf{r}_j \mathbf{r}_j') = \frac{1}{n} v_2 \left(-\frac{1}{p-1} \mathbf{1}_p \mathbf{1}_p' + \frac{1}{p-1} \mathbf{I}_p + \mathbf{I}_p \right). \quad (13)$$

Let $\mathbb{E}_0(\cdot)$ denote expectation and $\mathbb{E}_j(\cdot)$ denote conditional expectation with respect to the σ -field generated by $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_j$, where $j = 1, 2, \dots, n$. Next, we write $M_p^{(1)}(z)$ as a sum of martingale difference sequences (MDS), and then utilize the CLT of MDS (Lemma A.3) to derive the asymptotic distribution of $M_p^{(1)}(z)$, which can be written as

$$p[\mathbf{m}_p(z) - \mathbb{E} M_p(z)] = \sum_{j=1}^n [\text{tr}(\mathbb{E}_j - \mathbb{E}_{j-1}) \mathbf{D}^{-1}(z)] = -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_j(z) \mathbf{r}_j' \mathbf{D}_j^{-2}(z) \mathbf{r}_j. \quad (14)$$

Write $\beta_j(z) = \bar{\beta}_j(z) - \beta_j(z)\bar{\beta}_j(z)\varepsilon_j(z) = \bar{\beta}_j(z) - \bar{\beta}_j^2(z)\varepsilon_j(z) + \bar{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z)$. From this and the definition of $\delta_j(z)$, (14) has the following expression

$$\begin{aligned} & (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z)\mathbf{r}'_j\mathbf{D}_j^{-2}(z)\mathbf{r}_j \\ &= (\mathbb{E}_j - \mathbb{E}_{j-1})\left[(\bar{\beta}_j(z) - \bar{\beta}_j^2(z)\varepsilon_j(z) + \bar{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z))(\delta_j(z) + \frac{1}{n}\mathbf{v}_2\text{tr}\mathbf{D}_j^{-2}(z))\right] \\ &= -Y_j(z) + \mathbb{E}_{j-1}Y_j(z) - (\mathbb{E}_j - \mathbb{E}_{j-1})\left[\bar{\beta}_j^2(z)(\varepsilon_j(z)\delta_j(z) - \beta_j(z)\mathbf{r}'_j\mathbf{D}_j^{-2}(z)\mathbf{r}_j\varepsilon_j^2(z))\right], \end{aligned} \quad (15)$$

where the second equality uses the fact that $(\mathbb{E}_j - \mathbb{E}_{j-1})\bar{\beta}_j(z)\text{tr}\mathbf{D}_j^{-2}(z) = 0$, and

$$Y_j(z) = -\mathbb{E}_j\left(\bar{\beta}_j(z)\delta_j(z) - \bar{\beta}_j^2(z)\varepsilon_j(z)\frac{1}{n}\mathbf{v}_2\text{tr}\mathbf{D}_j^{-2}(z)\right) = -\mathbb{E}_j\frac{d}{dz}(\bar{\beta}_j(z)\varepsilon_j(z)).$$

By (ii), we have

$$\mathbb{E}\left|\sum_{j=1}^n(\mathbb{E}_j - \mathbb{E}_{j-1})\bar{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z)\right|^2 \leq 4\sum_{j=1}^n\mathbb{E}|\bar{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z)|^2 = o(1), \quad (16)$$

here we leverage the the martingale difference property of $(\mathbb{E}_j - \mathbb{E}_{j-1})\bar{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z)$. Thus, $\sum_{j=1}^n(\mathbb{E}_j - \mathbb{E}_{j-1})\bar{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z)$ converges to zero in probability. By the same argument, we have

$$\sum_{j=1}^n(\mathbb{E}_j - \mathbb{E}_{j-1})\bar{\beta}_j^2(z)\beta_j(z)\mathbf{r}'_j\mathbf{D}_j^{-2}(z)\mathbf{r}_j\varepsilon_j^2(z) \xrightarrow{i.p.} 0. \quad (17)$$

Then, equations (14) – (17) imply that

$$M_p^{(1)}(z) = \sum_{j=1}^n\{Y_j(z) - \mathbb{E}_{j-1}Y_j(z)\} + o_p(1), \quad (18)$$

where $\{Y_j(z) - \mathbb{E}_{j-1}Y_j(z), j = 1, \dots, n\}$ is a sequence of martingale difference.

Part 2: Application of martingales CLT to (18). 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:

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

where $\mathcal{J}(z_i) \neq 0$, $\{\alpha_i : i = 1, 2, \dots, r\}$ are constants. We apply the martingale CLT (Lemma A.3) to this martingale difference decomposition of $\sum_{i=1}^r\alpha_i M_p^{(1)}(z_i)$. To this end, we need to check two conditions:

Condition 4.7.

$$\sum_{j=1}^n\mathbb{E}\left(\left|\sum_{i=1}^r\alpha_i(Y_j(z_i) - \mathbb{E}_{j-1}Y_j(z_i))\right|^2 \mathbf{I}_{\{|\sum_{i=1}^r\alpha_i(Y_j(z_i) + \mathbb{E}_{j-1}Y_j(z_i))| \geq \varepsilon\}}\right) \rightarrow 0. \quad (19)$$

Condition 4.8.

$$\sum_{j=1}^n \mathbb{E}_{j-1} [(Y_j(z_1) - \mathbb{E}_{j-1} Y_j(z_1))(Y_j(z_2) - \mathbb{E}_{j-1} Y_j(z_2))] \quad (20)$$

converges in probability to a constant.

First, we verify Condition 4.7. By Lemma 4.4, we obtain

$$\mathbb{E}|Y_j(z)|^4 \leq K\mathbb{E}|\varepsilon_j(z)|^4 = o\left(\frac{1}{p}\right). \quad (21)$$

Furthermore, by Jensen's inequality and (21),

$$\mathbb{E}|\mathbb{E}_{j-1} Y_j(z)|^4 \leq \mathbb{E}(\mathbb{E}_{j-1}|Y_j(z)|^4) = \mathbb{E}|Y_j(z)|^4 = o\left(\frac{1}{p}\right). \quad (22)$$

It follows from (21) and (22) that

$$\text{the left hand side of (19)} \leq K \left(\frac{1}{\varepsilon^2} \sum_{j=1}^n \mathbb{E} \left| \sum_{i=1}^r \alpha_i Y_j(z_i) \right|^4 + \frac{1}{\varepsilon^2} \sum_{j=1}^n \mathbb{E} \left| \sum_{i=1}^r \alpha_i \mathbb{E}_{j-1} Y_j(z_i) \right|^4 \right) \rightarrow 0.$$

Then, we verify Condition 4.8. Since

$$\begin{aligned} (20) &= \sum_{j=1}^n \mathbb{E}_{j-1} [Y_j(z_1)Y_j(z_2)] - \sum_{j=1}^n [\mathbb{E}_{j-1} Y_j(z_1)][\mathbb{E}_{j-1} Y_j(z_2)] \\ &= \frac{\partial^2}{\partial z_1 \partial z_2} \left(\sum_{j=1}^n \mathbb{E}_{j-1} [\mathbb{E}_j(\bar{\beta}_j(z_1)\varepsilon_j(z_1))\mathbb{E}_j(\bar{\beta}_j(z_2)\varepsilon_j(z_2))] \right) \\ &\quad - \frac{\partial^2}{\partial z_1 \partial z_2} \left(\sum_{j=1}^n [\mathbb{E}_{j-1} \bar{\beta}_j(z_1)\varepsilon_j(z_1)][\mathbb{E}_{j-1} \bar{\beta}_j(z_2)\varepsilon_j(z_2)] \right), \end{aligned}$$

it is enough to consider the limits of

$$\sum_{j=1}^n \mathbb{E}_{j-1} [\mathbb{E}_j(\bar{\beta}_j(z_1)\varepsilon_j(z_1))\mathbb{E}_j(\bar{\beta}_j(z_2)\varepsilon_j(z_2))] \quad (23)$$

and

$$\sum_{j=1}^n [\mathbb{E}_{j-1} \bar{\beta}_j(z_1)\varepsilon_j(z_1)][\mathbb{E}_{j-1} \bar{\beta}_j(z_2)\varepsilon_j(z_2)]. \quad (24)$$

The limit of (24) is provided in the following lemma.

Lemma 4.9. *Under conditions and notations in Theorem 2.5, then*

$$\sum_{j=1}^n [\mathbb{E}_{j-1} \bar{\beta}_j(z_1)\varepsilon_j(z_1)][\mathbb{E}_{j-1} \bar{\beta}_j(z_2)\varepsilon_j(z_2)] \xrightarrow{i.P.} 0.$$

The proof of Lemma 4.9 is postponed to Appendix. By Lemma 4.9, the remaining work is to consider the limit of (23). Since the following inequalities hold:

$$|\text{tr}(\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z))\mathbf{A}| \leq \frac{\|\mathbf{A}\|}{\mathfrak{J}(z)}, \quad (25)$$

$$\mathbb{E} \left| \frac{1}{n} \nu_2 \text{tr} \mathbf{D}^{-1}(z) - \mathbb{E} \frac{1}{n} \nu_2 \text{tr} \mathbf{D}^{-1}(z) \right|^q \leq C_q n^{-q/2} \nu_0^{-q}, \quad (26)$$

$$\mathbb{E} |\bar{\beta}_j(z_i) - b_n(z_i)|^2 \leq \frac{K|z_i|^4}{n\nu_0^6}, \quad (27)$$

it is enough to prove that

$$b_p(z_1)b_p(z_2) \sum_{j=1}^n \mathbb{E}_{j-1} [\mathbb{E}_j(\varepsilon_j(z_1))\mathbb{E}_j(\varepsilon_j(z_2))] \quad (28)$$

converges to a constant in probability, which further gives the limit of (23). By Lemma 4.5, we have

$$\begin{aligned} (28) &= b_p(z_1)b_p(z_2) \sum_{j=1}^n \left[\sum_{i=1}^p \frac{1}{n^2} (\nu_4 - 3\nu_{12}) \mathbb{E}_j(\mathbf{D}_j^{-1}(z_1))_{ii} \mathbb{E}_j(\mathbf{D}_j^{-1}(z_2))_{ii} \right. \\ &\quad + \frac{1}{n^2} \nu_{12} \left(\text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_1) \mathbb{E}_j(\mathbf{D}_j^{-1}(z_2))'] + \text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_1) \mathbb{E}_j \mathbf{D}_j^{-1}(z_2)] \right) \\ &\quad \left. + \frac{1}{n^2} (\nu_{12} - \nu_2^2) \text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_1)] \text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_2)] \right] + o(1) =: I_1 + I_2 + I_3 + I_4 + o(1), \end{aligned}$$

where

$$\begin{aligned} I_1 &= \frac{1}{n^2} (\nu_4 - 3\nu_{12}) b_p(z_1)b_p(z_2) \sum_{j=1}^n \sum_{i=1}^p \mathbb{E}_j(\mathbf{D}_j^{-1}(z_1))_{ii} \mathbb{E}_j(\mathbf{D}_j^{-1}(z_2))_{ii}, \\ I_2 &= \frac{1}{n^2} \nu_{12} b_p(z_1)b_p(z_2) \sum_{j=1}^n \text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_1) \mathbb{E}_j(\mathbf{D}_j^{-1}(z_2))'], \\ I_3 &= \frac{1}{n^2} \nu_{12} b_p(z_1)b_p(z_2) \sum_{j=1}^n \text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_1) \mathbb{E}_j \mathbf{D}_j^{-1}(z_2)], \\ I_4 &= \frac{1}{n^2} (\nu_{12} - \nu_2^2) b_p(z_1)b_p(z_2) \sum_{j=1}^n \text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_1)] \text{tr} [\mathbb{E}_j \mathbf{D}_j^{-1}(z_2)]. \end{aligned}$$

In the following Steps (i)-(iii), we derive as $p \rightarrow \infty$,

$$\frac{\partial^2}{\partial z_1 \partial z_2} I_1 \xrightarrow{i.p.} c \alpha_1 \frac{\underline{\mathbf{m}}'(z) \underline{\mathbf{m}}'(z_2)}{\left(1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z_1)\right)^2 \left(1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z_2)\right)^2}, \quad (29)$$

$$\frac{\partial^2}{\partial z_1 \partial z_2} I_2, \frac{\partial^2}{\partial z_1 \partial z_2} I_3 \xrightarrow{i.p.} \frac{\partial}{\partial z_2} \left(\frac{\partial \alpha(z_1, z_2) / \partial z_1}{1 - \alpha(z_1, z_2)} \right) = \frac{\underline{\mathbf{m}}'(z_1) \underline{\mathbf{m}}'(z_2)}{(\underline{\mathbf{m}}(z_1) - \underline{\mathbf{m}}(z_2))^2} - \frac{1}{(z_1 - z_2)^2}, \quad (30)$$

$$\frac{\partial^2}{\partial z_1 \partial z_2} I_4 \xrightarrow{i.p.} c \left(h_2 - 2 \frac{\sigma^2}{\mu^2} h_1 \right) \frac{\underline{\mathbf{m}}'(z_1) \underline{\mathbf{m}}'(z_2)}{\left(1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z_1)\right)^2 \left(1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z_2)\right)^2}. \quad (31)$$

Step (i): Consider I_2 and I_3 . Let $\mathbf{D}_{ij}(z) = \mathbf{D}(z) - \mathbf{r}_i \mathbf{r}_i' - \mathbf{r}_j \mathbf{r}_j'$, $\mathbf{b}_1(z) = \frac{1}{1 + \frac{1}{n} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_{12}^{-1}(z)}$ and $\beta_{ij}(z) = \frac{1}{1 + \mathbf{r}_i' \mathbf{D}_{ij}^{-1}(z) \mathbf{r}_i}$. We have the equality $\mathbf{D}_j(z_1) + z_1 \mathbf{I}_p - \frac{n-1}{n} \nu_2 \mathbf{b}_1(z_1) \mathbf{I}_p = \sum_{i \neq j}^n \mathbf{r}_i \mathbf{r}_i' - \frac{n-1}{n} \nu_2 \mathbf{b}_1(z_1) \mathbf{I}_p$. Multiplying by $(z_1 \mathbf{I}_p - \frac{n-1}{n} \nu_2 \mathbf{b}_1(z_1) \mathbf{I}_p)^{-1}$ on the left-hand side and $\mathbf{D}_j^{-1}(z_1)$ on the right-hand side, 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

$$\begin{aligned} \mathbf{D}_j^{-1}(z_1) &= -\mathbf{Q}_p(z_1) + \sum_{i \neq j}^n \beta_{ij}(z_1) \mathbf{Q}_p(z_1) \mathbf{r}_i \mathbf{r}_i' \mathbf{D}_{ij}^{-1}(z_1) - \frac{n-1}{n} \nu_2 \mathbf{b}_1(z_1) \mathbf{Q}_p(z_1) \mathbf{D}_j^{-1}(z_1) \\ &= -\mathbf{Q}_p(z_1) + \mathbf{b}_1(z_1) \mathbf{A}(z_1) + \mathbf{B}(z_1) + \mathbf{C}(z_1), \end{aligned} \quad (32)$$

where $\mathbf{Q}_p(z_1) = (z_1 \mathbf{I}_p - \frac{n-1}{n} \nu_2 \mathbf{b}_1(z_1) \mathbf{I}_p)^{-1}$, $\mathbf{A}(z_1) = \sum_{i \neq j}^n \mathbf{Q}_p(z_1) (\mathbf{r}_i \mathbf{r}_i' - \frac{1}{n} \nu_2 \mathbf{I}_p) \mathbf{D}_{ij}^{-1}(z_1)$, $\mathbf{B}(z_1) = \sum_{i \neq j}^n (\beta_{ij}(z_1) - \mathbf{b}_1(z_1)) \mathbf{Q}_p(z_1) \mathbf{r}_i \mathbf{r}_i' \mathbf{D}_{ij}^{-1}(z_1)$, $\mathbf{C}(z_1) = \frac{1}{n} \nu_2 \mathbf{b}_1(z_1) \mathbf{Q}_p(z_1) \sum_{i \neq j}^n (\mathbf{D}_{ij}^{-1}(z_1) - \mathbf{D}_j^{-1}(z_1))$. For any real t ,

$$\left| 1 - \frac{t}{z(1 + n^{-1} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_{12}^{-1}(z))} \right|^{-1} \leq \frac{|z(1 + n^{-1} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_{12}^{-1}(z))|}{\Im(z(1 + n^{-1} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_{12}^{-1}(z)))} \leq \frac{|z|(1 + p/(n\nu_0))}{\nu_0}.$$

Thus,

$$\|\mathbf{Q}_p(z_1)\| \leq \frac{1 + p/(n\nu_0)}{\nu_0}. \quad (33)$$

For any random matrix \mathbf{M} , denote its nonrandom bound on the spectrum norm of \mathbf{M} by $\|\|\mathbf{M}\|\|$. From (27), Lemma 4.4, (33) and (25), we get, for any \mathbf{M} ,

$$\mathbb{E}|\text{tr} \mathbf{B}(z_1) \mathbf{M}| \leq K \|\|\mathbf{M}\|\| \frac{|z_1|^2 (1 + p/(n\nu_0))}{\nu_0^5} n^{1/2}, \quad (34)$$

$$|\text{tr} \mathbf{C}(z_1) \mathbf{M}| \leq \|\|\mathbf{M}\|\| \frac{|z_1| (1 + p/(n\nu_0))}{\nu_0^3}, \quad (35)$$

$$\mathbb{E}|\text{tr} \mathbf{A}(z_1) \mathbf{M}| \leq K \|\|\mathbf{M}\|\| \frac{1 + p/(n\nu_0)}{\nu_0^2} n^{1/2}. \quad (36)$$

Note that

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

therefore, by using (12), we can write

$$\text{tr} \mathbb{E}_j(\mathbf{A}(z_1)) \mathbf{D}_j^{-1}(z_2) = \mathbf{A}_1(z_1, z_2) + \mathbf{A}_2(z_1, z_2) + \mathbf{A}_3(z_1, z_2) + \mathbf{R}(z_1, z_2), \quad (37)$$

where

$$\begin{aligned}
A_1(z_1, z_2) &= - \sum_{i < j}^n \beta_{ij}(z_2) \mathbf{r}'_i \mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2) \mathbf{r}_i \mathbf{r}'_i \mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1) \mathbf{r}_i, \quad (38) \\
A_2(z_1, z_2) &= - \text{tr} \sum_{i < j}^n \mathbf{Q}_p(z_1) n^{-1} \nu_2 \mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) (\mathbf{D}_j^{-1}(z_2) - \mathbf{D}_{ij}^{-1}(z_2)), \\
A_3(z_1, z_2) &= \text{tr} \sum_{i < j}^n \mathbf{Q}_p(z_1) (\mathbf{r}_i \mathbf{r}'_i - n^{-1} \nu_2 \mathbf{I}_p) \mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2), \\
R(z_1, z_2) &= \text{tr} \sum_{i > j}^n \mathbf{Q}_p(z_1) \left(-\frac{1}{n(p-1)} \nu_2 \mathbf{1}_p \mathbf{1}'_p + \frac{1}{n(p-1)} \nu_2 \mathbf{I}_p \right) \mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_j^{-1}(z_2), \quad (39)
\end{aligned}$$

and $\mathbf{1}_p$ is a p -dimensional vector with all elements being 1. It is easy to see that $R(z_1, z_2) = O(1)$. We get from (25) and (33) that $|A_2(z_1, z_2)| \leq \frac{1+p/(n\nu_0)}{\nu_0^2}$. Similar to (36), we have $\mathbb{E}|A_3(z_1, z_2)| \leq \frac{1+p/(n\nu_0)}{\nu_0^3} n^{1/2}$. By similar calculation of Bai and Silverstein [2004], we get the following lemma and its proof is postponed to Appendix.

Lemma 4.10. *Under conditions and notations in Theorem 2.5, for any $1 \leq j \leq n$,*

$$\begin{aligned}
& \text{tr}(\mathbb{E}_j(\mathbf{D}_j^{-1}(z_1)) \mathbf{D}_j^{-1}(z_2)) \\
& \quad \times \left[1 - \frac{j-1}{n} \nu_2^2 \underline{\mathbf{m}}_p^0(z_1) \underline{\mathbf{m}}_p^0(z_2) \frac{c_n}{(1 + \frac{n-1}{n} \nu_2 \underline{\mathbf{m}}_p^0(z_1))(1 + \frac{n-1}{n} \nu_2 \underline{\mathbf{m}}_p^0(z_2))} \right] \\
&= \frac{n c_n}{z_1 z_2} \frac{1}{(1 + \frac{n-1}{n} \nu_2 \underline{\mathbf{m}}_p^0(z_1))(1 + \frac{n-1}{n} \nu_2 \underline{\mathbf{m}}_p^0(z_2))} + S(z_1, z_2),
\end{aligned}$$

where $\mathbb{E}|S(z_1, z_2)| \leq K n^{1/2}$.

By Lemma 4.10, I_3 can be written as

$$I_3 = \alpha_p(z_1, z_2) \frac{1}{n} \sum_{j=1}^n \frac{1}{1 - \frac{j-1}{n} \alpha_p(z_1, z_2)} + A_4(z_1, z_2), \quad (40)$$

where $\alpha_p(z_1, z_2) = \nu_2^2 \frac{c_n \underline{\mathbf{m}}_p^0(z_1) \underline{\mathbf{m}}_p^0(z_2)}{(1 + \frac{n-1}{n} \nu_2 \underline{\mathbf{m}}_p^0(z_1))(1 + \frac{n-1}{n} \nu_2 \underline{\mathbf{m}}_p^0(z_2))}$ and $\mathbb{E}|A_4(z_1, z_2)| \leq K n^{-1/2}$. By Lemma 4.3, the limit of $\alpha_p(z_1, z_2)$ is $\alpha(z_1, z_2) = \frac{\sigma^4}{\mu^4} \frac{c \underline{\mathbf{m}}(z_1) \underline{\mathbf{m}}(z_2)}{(1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z_1))(1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z_2))}$. Thus, by (40), the in probability (i.p.)

limit of $\frac{\partial^2}{\partial z_2 \partial z_1} I_3$ is in (30). Similarly, we get the i.p. limit of $\frac{\partial^2}{\partial z_2 \partial z_1} I_2$, which is also given by (30).

Step (ii): Consider I_1 . It is enough to find the limit of $\frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^p \mathbb{E}_j(\mathbf{D}_j^{-1}(z_1))_{ii} \mathbb{E}_j(\mathbf{D}_j^{-1}(z_2))_{ii}$. By similar calculation of Gao et al. [2017], we get the following lemma and its proof is postponed to Appendix.

Lemma 4.11. *Under conditions and notations in Theorem 2.5, for any $1 \leq j \leq n$,*

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

By (26), the formula (2.2) of Silverstein [1995], $\underline{m}_p(z) = -\frac{1}{zn} \sum_{j=1}^n \beta_j(z)$, and Lemma 4.4, we have

$$|\mathbf{b}_p(z) - \mathbb{E}\beta_1(z)| \leq \mathbf{K}n^{-1/2}, \quad \mathbb{E}\beta_1(z) = -z\mathbb{E}\underline{m}_p(z), \quad |\mathbf{b}_p(z) + z\underline{m}_p^0(z)| \leq \mathbf{K}n^{-1/2}. \quad (41)$$

Thus, by (41), Lemma 4.3 and Lemma 4.11, we have

$$I_1 \xrightarrow{i.p.} c\alpha_1 z_1 z_2 \underline{m}(z_1) \underline{m}(z_2) \underline{m}(z_1) \underline{m}(z_2) = c\alpha_1 \frac{\underline{m}(z_1) \underline{m}(z_2)}{\left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z_1)\right) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z_2)\right)},$$

where the equality above follows from $\underline{m}(z) = -z^{-1} \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z)\right)^{-1}$. Thus, the i.p. limit of $\frac{\partial^2}{\partial z_2 \partial z_1} I_1$ is in (29).

Step (iii): Consider I_4 . We have $\mathbb{E} \left| \frac{1}{p} \text{tr} \mathbb{E}_j \mathbf{D}_j^{-1}(z_1) \frac{1}{p} \text{tr} \mathbb{E}_j \mathbf{D}_j^{-1}(z_2) - \underline{m}_p^0(z_1) \underline{m}_p^0(z_2) \right| = o(1)$. By Lemma 4.3, we get

$$\lim_{p \rightarrow \infty} p(\nu_{12} - \nu_2^2) = h_2 - 2 \frac{\sigma^2}{\mu^2} h_1.$$

This, together with (41), implies that

$$I_4 \xrightarrow{i.p.} c \left(h_2 - 2 \frac{\sigma^2}{\mu^2} h_1 \right) z_1 \underline{m}(z_1) z_2 \underline{m}(z_2) \underline{m}(z_1) \underline{m}(z_2) = c \left(h_2 - 2 \frac{\sigma^2}{\mu^2} h_1 \right) \frac{\underline{m}(z_1) \underline{m}(z_2)}{\left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z_1)\right) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z_2)\right)}.$$

Then the i.p. limit of the second derivative $\frac{\partial^2}{\partial z_2 \partial z_1} I_4$ is in (31).

4.3.3 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 Billingsley [1968], 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 Bai and Silverstein [2004], and is postponed to Appendix.

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

Similar to Bai and Silverstein [2004], one can prove the inequality:

$$\mathbb{E} |\text{tr} \mathbf{D}_1^{-1}(z) \mathbf{M} - \mathbb{E} \text{tr} \mathbf{D}_1^{-1}(z) \mathbf{M}|^2 \leq \mathbf{K} \|\mathbf{M}\|^2. \quad (42)$$

We first present the following equations for later use, $M_p^{(2)}(z) = p(\mathbb{E} m_{\mathbb{F}_p^2 \mathbb{S}_n^0}(z) - m_{\mathbb{F}^c n}(z)) = n(\mathbb{E} \underline{m}_p(z) - \underline{m}_p^0(z))$, $\underline{m}(z) = -\frac{1-c}{z} + c\underline{m}(z)$. The next step is to find $\mathbb{E} \underline{m}_p(z)$. From the identity (2), which is the inverse of $\underline{m}(z)$, we define

$$R_p(z) := \frac{1}{\mathbb{E} \underline{m}_p(z)} + z - c_n \frac{\sigma^2/\mu^2}{1 + \sigma^2/\mu^2 \mathbb{E} \underline{m}_p(z)} = \frac{1}{\mathbb{E} \underline{m}_p(z)} \left(1 - c_n + z \mathbb{E} \underline{m}_p(z) + \frac{c_n}{1 + \sigma^2/\mu^2 \mathbb{E} \underline{m}_p(z)} \right),$$

thus,

$$\mathbb{E} \underline{m}_p(z) = \left(-z + c_n \frac{\sigma^2/\mu^2}{1 + \sigma^2/\mu^2 \mathbb{E} \underline{m}_p(z)} + A_p(z) / \mathbb{E} \underline{m}_p(z) \right)^{-1}, \quad (43)$$

where $\mathbf{A}_p(z) = \frac{c_n}{1 + \sigma^2/\mu^2 \mathbb{E} \underline{\mathbf{m}}_p(z)} + z c_n \mathbb{E} \mathbf{M}_p(z)$. Note that

$$\underline{\mathbf{m}}_p^0 = \left(-z + c_n \frac{\sigma^2/\mu^2}{1 + \sigma^2/\mu^2 \underline{\mathbf{m}}_p^0} \right)^{-1}. \quad (44)$$

From (43) and (44), we get

$$\mathbb{E} \underline{\mathbf{m}}_p(z) - \underline{\mathbf{m}}_p^0(z) = -\underline{\mathbf{m}}_p^0(z) \mathbf{A}_p(z) \left[1 - c_n \underline{\mathbf{m}}_p^0(z) \mathbb{E} \underline{\mathbf{m}}_p(z) \frac{(\sigma^2/\mu^2)^2}{(1 + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z)) (1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}_p^0(z))} \right]^{-1}. \quad (45)$$

Our next task is to investigate the limiting behavior of $n \mathbf{A}_p$. Let $\tilde{\mathbf{Q}}_p(z) = \mathbf{I}_p + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z) \mathbf{I}_p$, then

$$n \mathbf{A}_p = p \frac{1}{1 + \sigma^2/\mu^2 \mathbb{E} \underline{\mathbf{m}}_p(z)} + p z \mathbb{E} \mathbf{M}_p(z) = \mathbb{E}(\beta_1 \mathbf{P}_1(z)) + \mathbb{E}(\beta_1 \mathbf{P}_2(z)), \quad (46)$$

where

$$\begin{aligned} \mathbf{P}_1(z) &= \left[n \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 - \frac{\sigma^2}{\mu^2} \text{tr} \left(\tilde{\mathbf{Q}}_p^{-1}(z) \mathbb{E} \mathbf{D}_1^{-1}(z) \right) \right], \\ \mathbf{P}_2(z) &= \left[\frac{\sigma^2}{\mu^2} \text{tr}(\mathbf{Q}_p^{-1}(z) \mathbb{E} \mathbf{D}_1^{-1}(z)) - \frac{\sigma^2}{\mu^2} \text{tr}(\mathbf{Q}_p^{-1}(z) \mathbb{E} \mathbf{D}^{-1}(z)) \right]. \end{aligned}$$

Since $\beta_1 = \mathbf{b}_p - \mathbf{b}_p^2 \gamma_1 + \beta_1 \mathbf{b}_p^2 \gamma_1^2$, where $\gamma_1(z) = \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_1^{-1}(z)$, we have

$$\mathbb{E}(\beta_1 \mathbf{P}_1(z)) = \mathbf{b}_p(z) \mathbb{E} \mathbf{P}_1(z) - \mathbf{b}_p^2(z) \mathbb{E}(\gamma_1 \mathbf{P}_1(z)) + \mathbf{b}_p^2(z) \mathbb{E}(\beta_1 \gamma_1^2 \mathbf{P}_1(z)). \quad (47)$$

For $\mathbb{E} \mathbf{P}_1(z)$, by (13), we get

$$\mathbb{E} \mathbf{P}_1(z) = \frac{n}{1 + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z)} \left(\mathbb{E} \gamma_1(z) + \frac{1}{n} \left(\nu_2 - \frac{\sigma^2}{\mu^2} \right) \mathbb{E} \text{tr} \mathbf{D}_1^{-1}(z) \right). \quad (48)$$

The estimates for $\mathbb{E}(\gamma_1 \mathbf{P}_1(z))$, $\mathbb{E}(\beta_1 \gamma_1^2 \mathbf{P}_1(z))$, and $\mathbb{E}(\beta_1 \mathbf{P}_2(z))$ are provided in the following lemma.

Lemma 4.12. *Under conditions and notations in Theorem 2.5, we have*

$$\begin{aligned} \mathbb{E}(\gamma_1 \mathbf{P}_1(z)) &= n \mathbb{E} \left[\left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} \mathbf{D}_1^{-1}(z) \right) \times \left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \right] \right) \right] \\ &+ \frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\text{tr} \mathbf{D}_1^{-1}(z) \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \right] \right) - \frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\mathbf{1}'_p \mathbf{D}_1^{-1}(z) \mathbf{1}_p \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \right] \right) \\ &- \frac{\sigma^2/\mu^2}{1 + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z)} \text{tr}[\mathbb{E} \mathbf{D}_1^{-1}(z)] \mathbb{E} \gamma_1(z) + o(1), \end{aligned} \quad (49)$$

and

$$\begin{aligned} \mathbb{E}(\beta_1(z) \gamma_1^2(z) \mathbf{P}_1(z)) &= \mathbb{E} \left(n \beta_1(z) \gamma_1^2(z) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 \right) - \mathbb{E} \left(\beta_1(z) \gamma_1^2(z) \text{tr} \left[\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right] \right) \\ &+ \text{Cov} \left(\beta_1(z) \gamma_1^2(z), \text{tr} \left[\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right] \right) = O(\delta_n^2), \end{aligned} \quad (50)$$

and

$$\mathbb{E}(\beta_1(z) \mathbf{P}_2(z)) = \frac{p}{n(p-1)} \nu_2 \frac{\sigma^2}{\mu^2} \mathbf{b}_p^2(z) \mathbb{E} \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right] + O(n^{-1/2}). \quad (51)$$

The proof of Lemma 4.12 is postponed to Appendix. Therefore, from (46) – (51), we get

$$n\mathcal{A}_p = J_1 + J_2 + J_3 + o(1), \quad (52)$$

where

$$\begin{aligned} J_1 &= \frac{nb_p(z)}{1 + \frac{\sigma^2}{\mu^2}\mathbb{E}\underline{m}_p(z)} \left(\mathbb{E}\gamma_1(z) + \frac{1}{n}(\nu_2 - \frac{\sigma^2}{\mu^2})\mathbb{E}\text{tr}\mathbf{D}_1^{-1}(z) \right) \\ &\quad + \left(b_p^2(z) \frac{\sigma^2}{\mu^2} \frac{1}{1 + \frac{\sigma^2}{\mu^2}\mathbb{E}\underline{m}_p(z)} \right) \text{tr}[\mathbb{E}\mathbf{D}_1^{-1}(z)]\mathbb{E}\gamma_1(z) \\ &\quad - b_p^2(z) \left[\frac{1}{n(p-1)}\nu_2^2\mathbb{E}\left(\text{tr}\mathbf{D}_1^{-1}(z) \text{tr}[\mathbf{D}_1^{-1}(z)\tilde{\mathbf{Q}}_p^{-1}(z)] \right) \right. \\ &\quad \left. - \frac{1}{n(p-1)}\nu_2^2\mathbb{E}\left(\mathbf{1}'_p\mathbf{D}_1^{-1}(z)\mathbf{1}_p \text{tr}[\mathbf{D}_1^{-1}(z)\tilde{\mathbf{Q}}_p^{-1}(z)] \right) \right], \\ J_2 &= -nb_p^2(z)\mathbb{E}\left(\mathbf{r}'_1\mathbf{D}_1^{-1}(z)\mathbf{r}_1 - \frac{1}{n}\nu_2 \text{tr}\mathbf{D}_1^{-1}(z) \right) \\ &\quad \times \left(\mathbf{r}'_1\mathbf{D}_1^{-1}(z)\tilde{\mathbf{Q}}_p^{-1}(z)\mathbf{r}_1 - \frac{1}{n}\nu_2 \text{tr}[\mathbf{D}_1^{-1}(z)\tilde{\mathbf{Q}}_p^{-1}(z)] \right), \\ J_3 &= \frac{p}{n(p-1)}b_p^2(z)\frac{\sigma^2}{\mu^2}\nu_2\mathbb{E}\text{tr}[\mathbf{D}_1^{-1}(z)\tilde{\mathbf{Q}}_p^{-1}(z)\mathbf{D}_1^{-1}(z)]. \end{aligned}$$

The limits of J_1 , J_2 and J_3 are provided in the following lemma. The proof of Lemma 4.13 is postponed to Appendix.

Lemma 4.13. *Under conditions and notation in Theorem 2.5, as $n \rightarrow \infty$,*

$$\begin{aligned} J_1 &\xrightarrow{i.p.} \left(\frac{-z\underline{m}(z)}{1 + \frac{\sigma^2}{\mu^2}\underline{m}(z)} \right) \left(\frac{\sigma^2}{\mu^2}\underline{m}(z) + \frac{\sigma^2}{\mu^2}\frac{1}{z} + h_1\underline{m}(z) \right), \\ J_2 &\xrightarrow{i.p.} -\frac{c\alpha_1 z^2 \underline{m}^2(z) \underline{m}^2(z)}{1 + \frac{\sigma^2}{\mu^2}\underline{m}(z)} - \frac{2cz^2 m'(z) \underline{m}^2(z) |\mathbb{E}|w_1 - \mu|^2|^2}{1 + \frac{\sigma^2}{\mu^2}\underline{m}(z) \mu^4} - c(h_2 - 2\frac{\sigma^2}{\mu^2}h_1) \frac{z^2 \underline{m}^2(z) \underline{m}^2(z)}{1 + \frac{\sigma^2}{\mu^2}\underline{m}(z)}, \\ J_3 &\xrightarrow{i.p.} c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) (1 + \frac{\sigma^2}{\mu^2}\underline{m}(z))^{-3} \left[1 - c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) (1 + \frac{\sigma^2}{\mu^2}\underline{m}(z))^{-2} \right]^{-1}. \end{aligned}$$

From (45), (52), and Lemma 4.13, we get (7). The proof is completed.

A Some technical lemmas

Lemma A.1 (Weyl's inequality, Corollary 7.3.5 in Horn and Johnson [2012]). *Let \mathbf{A} and \mathbf{B} be two $p \times n$ matrices and let $r = \min\{p, n\}$. Let $\sigma_1(\mathbf{A}) \geq \dots \geq \sigma_r(\mathbf{A})$ and $\sigma_1(\mathbf{B}) \geq \dots \geq \sigma_r(\mathbf{B})$ be the nonincreasingly ordered singular values of \mathbf{A} and \mathbf{B} , respectively. Then*

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

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

Lemma A.2 (Burkholder [1973]). Let $\{X_k\}$ be a complex martingale difference sequence with respect to the increasing σ -field $\{\mathcal{F}_k\}$. Then for $q > 1$,

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

Lemma A.3 (Theorem 35.12 of Billingsley [1995]). Suppose for each n , $\{Y_{n1}, Y_{n2}, \dots, Y_{nr_n}\}$ is a real martingale difference sequence with respect to the increasing σ -field $\{\mathcal{F}_{nj}\}$ having second moments. If as $n \rightarrow \infty$,

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^2 | \mathcal{F}_{n,j-1}) \xrightarrow{i.p.} \sigma^2, \quad (53)$$

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

$$\sum_{j=1}^{r_n} \mathbb{E}(Y_{nj}^2 I_{\{|Y_{nj}| \geq \varepsilon\}}) \rightarrow 0, \quad (54)$$

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

Lemma A.4 is used in the proof of Lemma 4.9, and its proof is provided in Section B.12.

Lemma A.4. Suppose that $\mathbf{x}_p = \frac{1}{\sqrt{p}}(1, 1, \dots, 1)'$ is a p -dimensional normalized all-one vector, then for the truncated random variable satisfying (10), we have $\mathbb{E}|\mathbf{x}_p' \mathbf{D}^{-1}(z) \mathbf{x}_p + \frac{1}{z}|^2 \rightarrow 0$.

B Proofs

This section contains proofs of Theorem 2.3, Proposition 2.4, Lemmas 4.3 – 4.5, Lemmas 4.9 – 4.13, Lemma A.4, Corollary 2.6, tightness of $M_p^{(1)}(z)$.

B.1 Proof of Theorem 2.3

We write

$$\mathbf{X}_n = \begin{pmatrix} \frac{1}{\sum_{j=1}^p w_{1j}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sum_{j=1}^p w_{nj}} \end{pmatrix} \begin{pmatrix} w_{11} & \cdots & w_{1p} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{np} \end{pmatrix} =: \mathbf{\Lambda}_n \mathbf{W}_n,$$

and let $\widetilde{\mathbf{W}}_n = \frac{\mathbf{C}_n \mathbf{W}_n}{\sqrt{N\mu}}$. By Theorem 3.6 in Bai and Silverstein [2010], the LSD of $\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n$ has a density function given by (1). Let $L(\cdot, \cdot)$ be the Levy distance. To prove Theorem 2.3, we need to show that $L(\mathbb{F}^{\mathbf{B}_{p,N}}, \mathbb{F}^{\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n}) \xrightarrow{a.s.} 0$. Let $\mathbf{Y}_n = \frac{p \mathbf{C}_n \mathbf{X}_n}{\sqrt{N}} = \frac{p \mathbf{C}_n \mathbf{\Lambda}_n \mathbf{W}_n}{\sqrt{N}}$, then we can write $\mathbf{B}_{p,N} = \mathbf{Y}_n' \mathbf{Y}_n$. By Corollary A.42 in Bai and Silverstein [2010], we have

$$\begin{aligned} L^4(\mathbb{F}^{\mathbf{B}_{p,N}}, \mathbb{F}^{\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n}) &= L^4(\mathbb{F}^{\mathbf{Y}_n' \mathbf{Y}_n}, \mathbb{F}^{\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n}) \\ &\leq \frac{2}{p^2} \text{tr}[(\mathbf{Y}_n - \widetilde{\mathbf{W}}_n)'(\mathbf{Y}_n - \widetilde{\mathbf{W}}_n)] \cdot \text{tr}(\mathbf{Y}_n' \mathbf{Y}_n + \widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n). \end{aligned}$$

By using Von Neumann's trace inequality and the law of large numbers, we have

$$\begin{aligned}\frac{1}{p}\text{tr}(\mathbf{Y}'_n \mathbf{Y}_n) &= \frac{p}{N}\text{tr}(\mathbf{W}'_n \boldsymbol{\Lambda}'_n \mathbf{C}_n \boldsymbol{\Lambda}_n \mathbf{W}_n) \\ &\leq \frac{p\|\mathbf{C}_n\|}{N}\text{tr}(\mathbf{W}'_n \boldsymbol{\Lambda}'_n \boldsymbol{\Lambda}_n \mathbf{W}_n) = \frac{1}{N}\sum_{i=1}^n \frac{\sum_{j=1}^p w_{ij}^2/p}{(\sum_{j=1}^p w_{ij}/p)^2} = O_p(1), \\ \frac{1}{p}\text{tr}(\widetilde{\mathbf{W}}'_n \widetilde{\mathbf{W}}_n) &= \frac{1}{Np\mu^2}\text{tr}(\mathbf{W}'_n \mathbf{C}_n \mathbf{W}_n) \leq \frac{\|\mathbf{C}_n\|}{Np\mu^2}\text{tr}(\mathbf{W}'_n \mathbf{W}_n) = \frac{1}{Np\mu^2}\sum_{i=1}^n \sum_{j=1}^p w_{ij}^2 = O_p(1).\end{aligned}$$

Thus, to complete the proof, it is sufficient to prove that

$$\frac{1}{p}\text{tr}[(\mathbf{Y}_n - \widetilde{\mathbf{W}}_n)'(\mathbf{Y}_n - \widetilde{\mathbf{W}}_n)] = o_p(1).$$

Let \mathbf{v}_i be the i -th column of $\mathbf{Y}_n - \widetilde{\mathbf{W}}_n$, that is, $\mathbf{v}_i = \mathbf{C}_n \left(\frac{\mathbf{w}_i}{\sqrt{N}(\sum_{j=1}^p w_{ij}/p)} - \frac{\mathbf{w}_i}{\sqrt{N}\mu} \right)$. Then,

$$\begin{aligned}&\frac{1}{p}\text{tr}[(\mathbf{Y}_n - \widetilde{\mathbf{W}}_n)'(\mathbf{Y}_n - \widetilde{\mathbf{W}}_n)] \\ &= \frac{1}{p}\sum_{i=1}^n \|\mathbf{v}_i\|^2 \leq \frac{\|\mathbf{C}_n\|}{Np}\sum_{i=1}^n \sum_{j=1}^p \left(\frac{1}{\mu} - \frac{1}{\sum_{j=1}^p w_{ij}/p} \right)^2 w_{ij}^2 \\ &= \frac{1}{\mu^2} \left(\frac{1}{Np}\sum_{i=1}^n \sum_{j=1}^p w_{ij}^2 \right) - \frac{2}{N\mu}\sum_{i=1}^n \frac{\sum_{j=1}^p w_{ij}^2/p}{\sum_{j=1}^p w_{ij}/p} + \frac{1}{N}\sum_{i=1}^n \frac{\sum_{j=1}^p w_{ij}^2/p}{(\sum_{j=1}^p w_{ij}/p)^2} \\ &= o_p(1).\end{aligned}$$

This completes the proof.

B.2 Proof of Proposition 2.4

First, we prove the convergence of extreme eigenvalues. Recall that $\widetilde{\mathbf{W}}_n = \frac{\mathbf{C}_n \mathbf{W}_n}{\sqrt{N}\mu}$. Let $\lambda_{\max}(\widetilde{\mathbf{W}}'_n \widetilde{\mathbf{W}}_n)$ and $\lambda_{\min}(\widetilde{\mathbf{W}}'_n \widetilde{\mathbf{W}}_n)$ be the largest eigenvalue and the smallest non-zero of $\widetilde{\mathbf{W}}'_n \widetilde{\mathbf{W}}_n$. By the equation (2.7) in Jiang [2004] and Theorem 1.4 in Xiao and Zhou [2010], we have

$$\lim_{n \rightarrow \infty} \lambda_{\max}(\widetilde{\mathbf{W}}'_n \widetilde{\mathbf{W}}_n) = \frac{\sigma^2}{\mu^2}(1 + \sqrt{c})^2 \quad \text{and} \quad \lim_{n \rightarrow \infty} \lambda_{\min}(\widetilde{\mathbf{W}}'_n \widetilde{\mathbf{W}}_n) = \frac{\sigma^2}{\mu^2}(1 - \sqrt{c})^2. \quad (55)$$

By Lemma A.1, we have

$$\max_{1 \leq i \leq p} \left| \sqrt{\lambda_i(\mathbf{B}_{p,N})} - \sqrt{\lambda_i(\widetilde{\mathbf{W}}'_n \widetilde{\mathbf{W}}_n)} \right| \leq \left\| \mathbf{C}_n(p\boldsymbol{\Lambda}_n - \mathbf{I}) \frac{\mathbf{W}_n}{\mu\sqrt{N}} \right\| \leq \|p\mu\boldsymbol{\Lambda}_n - \mathbf{I}\| \cdot \left\| \frac{\mathbf{W}_n}{\mu\sqrt{N}} \right\|. \quad (56)$$

By Lemma 2 in Bai and Yin [1993], we have $\max_{1 \leq i \leq n} |\sum_{j=1}^p w_{ij}/(p\mu) - 1| \xrightarrow{a.s.} 0$, which implies that $\|p\mu\boldsymbol{\Lambda} - \mathbf{I}\| = \max_{1 \leq i \leq n} |\frac{\mu}{\sum_{j=1}^p w_{ij}/p} - 1| \xrightarrow{a.s.} 0$. From the Theorem 2.9 in Benaych-Georges and

Nadakuditi [2012], we conclude that $\|\mathbf{W}_n/(\mu\sqrt{N})\| < \infty$. Combining these facts with (55) and (56), we have

$$\sqrt{\lambda_{\max}(\mathbf{B}_{p,N})} - \sqrt{\lambda_{\max}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n)} \xrightarrow{\text{a.s.}} 0 \quad \text{and} \quad \sqrt{\lambda_{\min}(\mathbf{B}_{p,N})} - \sqrt{\lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n)} \xrightarrow{\text{a.s.}} 0,$$

which together with (55) implies that

$$\begin{aligned} & \sqrt{\lambda_{\max}(\mathbf{B}_{p,N})} + \sqrt{\lambda_{\max}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n)} \xrightarrow{\text{a.s.}} \frac{2\sigma}{\mu} |1 + \sqrt{c}|, \\ & \sqrt{\lambda_{\min}(\mathbf{B}_{p,N})} + \sqrt{\lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n)} \xrightarrow{\text{a.s.}} \frac{2\sigma}{\mu} |1 - \sqrt{c}|. \end{aligned}$$

Therefore,

$$\lambda_{\max}(\mathbf{B}_{p,N}) - \lambda_{\max}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) \xrightarrow{\text{a.s.}} 0 \quad \text{and} \quad \lambda_{\min}(\mathbf{B}_{p,N}) - \lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) \xrightarrow{\text{a.s.}} 0.$$

This, together with (55), completes the proof of Proposition 2.4.

Now, we show that, with probability one, no eigenvalues of $\mathbf{B}_{p,N}$ outside the support of LSD. For any positive constant ε small enough such that $\eta_1 - \varepsilon > \frac{\sigma^2}{\mu^2} (1 + \sqrt{c})^2$ and $\eta_2 + \varepsilon < \mathbf{1}_{(0,1)}(c) \cdot \frac{\sigma^2}{\mu^2} (1 - \sqrt{c})^2$, we have

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

and

$$\begin{aligned} & \mathbb{P}(\lambda_{\min}(\mathbf{B}_{p,N}) \leq \eta_2) \\ &= \mathbb{P}(\lambda_{\min}(\mathbf{B}_{p,N}) \leq \eta_1, \lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) \leq \eta_2 - \varepsilon) \\ & \quad + \mathbb{P}(\lambda_{\min}(\mathbf{B}_{p,N}) \leq \eta_1, \lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) > \eta_2 - \varepsilon) \\ &\leq \mathbb{P}(\lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) \leq \eta_2 - \varepsilon) + \mathbb{P}(|\lambda_{\min}(\mathbf{B}_{p,N}) - \lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n)| \geq \varepsilon). \end{aligned}$$

To prove this theorem, it suffices to give the following three estimations:

$$\mathbb{P}\left(\max_{1 \leq i \leq p} |\lambda_i(\mathbf{B}_{p,N}) - \lambda_i(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n)| \geq \varepsilon\right) = o(n^{-\ell}), \quad (57)$$

$$\mathbb{P}(\lambda_{\max}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) \geq \eta_1 - \varepsilon) = o(n^{-\ell}), \quad (58)$$

$$\mathbb{P}(\lambda_{\min}(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) \leq \eta_2 - \varepsilon) = o(n^{-\ell}). \quad (59)$$

First, we prove (57). In view of the inequalities (56) and (58) (will be proved below), it suffices to show that, for any $\ell > 0$ and $\varepsilon > 0$,

$$\mathbb{P}\left(\max_{1 \leq i \leq n} \left| \frac{\sum_{j=1}^p w_{ij}/p}{\mu} - 1 \right| \geq \varepsilon\right) = o(n^{-\ell}), \quad (60)$$

which follows from the equation (9) in [Chen and Pan \[2012\]](#). (Although [Chen and Pan \[2012\]](#)'s ultra-high dimensional setting ($p/n \rightarrow \infty$ as $p \wedge n \rightarrow \infty$) is different from our setting, their equation (9) still holds when p and n are of the same order.)

Second, we prove (58) and (59). Let $\mathbf{W}_n^\circ = \frac{1}{\mu\sqrt{N}}(\mathbf{W}_n - \mu\mathbf{1}_n\mathbf{1}'_n)$. From [Bai and Silverstein \[2004\]](#), we have

$$\mathbb{P}\left(\lambda_{\max}((\mathbf{W}_n^\circ)' \mathbf{W}_n^\circ) \geq \eta_1 - \varepsilon/2\right) = o(n^{-\ell}) \quad (61)$$

and

$$\mathbb{P}\left(\lambda_{\min}((\mathbf{W}_n^\circ)' \mathbf{W}_n^\circ) \leq \eta_2 - \varepsilon/2\right) = o(n^{-\ell}). \quad (62)$$

By using Lemma A.1, we have

$$\begin{aligned} & \max_{1 \leq i \leq p} \left| \lambda_i(\widetilde{\mathbf{W}}_n' \widetilde{\mathbf{W}}_n) - \lambda_i((\mathbf{W}_n^\circ)' \mathbf{W}_n^\circ) \right| \\ & \leq \|\widetilde{\mathbf{W}}_n - \mathbf{W}_n^\circ\| = \left\{ \frac{1}{N} \sum_{j=1}^p \left(\frac{\sum_{i=1}^n w_{ij}/n}{\mu} - 1 \right)^2 \right\}^{1/2} \\ & \leq \left\{ \frac{p}{N} \cdot \max_{1 \leq j \leq p} \left(\frac{\sum_{i=1}^n w_{ij}/n}{\mu} - 1 \right)^2 \right\}^{1/2}. \end{aligned} \quad (63)$$

Similarly to (60), for any $\ell > 0$ and $\varepsilon > 0$,

$$\mathbb{P}\left(\max_{1 \leq j \leq p} \left| \frac{\sum_{i=1}^n w_{ij}/n}{\mu} - 1 \right| \geq \varepsilon\right) = o(n^{-\ell}). \quad (64)$$

This, together with (61) ~ (63), completes the proof of (58) and (59).

B.3 Proof of Lemma 4.2

The proof of Lemma 4.2 is quite similar to Sections 5.3.1, 5.3.2, and 5.5 of [Zheng et al. \[2015\]](#), it is then omitted. For readers' convenience, we present the outline of the proof for Lemma 4.2. In this situation, $\mathbf{B}_p^0 = \frac{1}{n} \mathbf{Y}'_n \mathbf{Y}_n = \sum_{i=1}^n \mathbf{r}_i \mathbf{r}'_i$, $\mathbf{r}_i = \frac{1}{\sqrt{n}} \left(\frac{w_{i1}}{w_i} - 1, \dots, \frac{w_{ip}}{w_i} - 1 \right)' = \frac{1}{\sqrt{n}} (y_{i1}, \dots, y_{ip})'$, $i = 1, 2, \dots, n$ and $\mathbb{E} \mathbf{r}_i \mathbf{r}'_i = \frac{1}{n} \Sigma$, where Σ is the population covariance matrix of p mutiple of the compositional data (i.e., $p\mathbf{X}_n$) and $\Sigma = \nu_2 \left(-\frac{1}{p-1} \mathbf{1}_p \mathbf{1}'_p + \frac{1}{p-1} \mathbf{I}_p + \mathbf{I}_p \right)$. As for moments of y_{ij} , by (9), for any $q > 0$, we have

$$\mathbb{E} (y_{ij})^q = \mathbb{E} \left[\left(\frac{w_{ij}}{w_i} - 1 \right)^q I(\mathbf{B}_p(\varepsilon)) \right] + \mathbb{E} \left[\left(\frac{w_{ij}}{w_i} - 1 \right)^q I(\mathbf{B}_p^c(\varepsilon)) \right] \leq K \mathbb{E} (w_{ij} - \mu)^q.$$

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

$$\begin{aligned} \text{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - p m_N^0(z) &= \text{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - p m_n^0(z) + p(m_n^0(z) - m_N^0(z)) \\ &+ \text{tr} \mathbf{A}^{-2}(z) \Delta + \text{tr} \mathbf{A}^{-1}(z) (\Delta \mathbf{A}^{-1}(z))^2 + \text{tr}(\mathbf{A}(z) - \Delta)^{-1} (\Delta \mathbf{A}^{-1}(z))^3, \end{aligned} \quad (65)$$

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

$$p(m_n^0(z) - m_N^0(z)) \xrightarrow{a.s.} (1 + z \underline{m}(z)) \frac{\underline{m}(z) + z \underline{m}'(z)}{z \underline{m}(z)}, \quad \mathbb{E} |\text{tr}(\mathbf{A}^{-2}(z) \Delta)|^2 = o(1), \quad (66)$$

$$\text{tr}(\mathbf{A}^{-2}(z)\Delta\mathbf{A}^{-1}(z)\Delta) \stackrel{L_2}{\rightarrow} (\underline{\mathbf{m}}(z) + z\underline{\mathbf{m}}'(z))(1 + z\underline{\mathbf{m}}(z)), \quad (67)$$

$$\text{tr}(\mathbf{A}^{-1}(z)\Delta)^3(\mathbf{A}(z) - \Delta)^{-1} = (1 + z\underline{\mathbf{m}}(z)) \text{tr}((\mathbf{A}^{-1}(z)\Delta)^2(\mathbf{A}(z) - \Delta)^{-1}) + o_p(1). \quad (68)$$

Thus, by (65) – (68), Lemma 4.2 is obtained.

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

$$\begin{aligned} \text{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - \text{pm}_N^0(z) &= \text{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - \text{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} \\ &\quad + \text{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - \text{pm}_n^0(z) + \text{pm}_n^0(z) - \text{pm}_N^0(z), \end{aligned}$$

and the tightness of $\text{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - \text{pm}_n^0(z)$ is proved in Step 2 of Section 4.3, it suffices to prove tightness of $\text{tr}(\mathbf{B}_{p,N} - z\mathbf{I}_p)^{-1} - \text{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1}$. It can be obtained from similar arguments in Section 5.3.2 of Zheng et al. [2015] and we omit the details. Finally, the proof is completed.

B.4 Proof of Lemma 4.3

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

$$\frac{1}{(1+x)^2} = 1 - 2x + 3x^2 + a(x), |a(x)| \leq C_1 x^3.$$

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

$$\frac{1}{\bar{w}^2} = \frac{1}{\mu^2(\frac{\bar{w}-\mu}{\mu} + 1)^2} = \frac{1}{\mu^2} \left[1 - \frac{2(\bar{w}-\mu)}{\mu} + \frac{3(\bar{w}-\mu)^2}{\mu^2} + a\left(\frac{\bar{w}-\mu}{\mu}\right) \right],$$

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

$$\frac{w_1^2}{\bar{w}^2} I(B_p(\epsilon)) = \left[\frac{w_1^2}{\mu^2} - \frac{2w_1^2(\bar{w}-\mu)}{\mu^3} + \frac{3w_1^2(\bar{w}-\mu)^2}{\mu^4} \right] I(B_p(\epsilon)) + \frac{w_1^2}{\mu^2} a\left(\frac{\bar{w}-\mu}{\mu}\right) I(B_p(\epsilon)), \quad (69)$$

where $\left| \frac{w_1^2}{\mu^2} a\left(\frac{\bar{w}-\mu}{\mu}\right) I(B_p(\epsilon)) \right| < C_1 \frac{w_1^2}{\mu^2} \epsilon^3$. Since $I(B_p(\epsilon)) = 1 - I(B_p^c(\epsilon))$,

$$\begin{aligned} \text{the right hand side of (69)} &= \frac{w_1^2}{\mu^2} - \frac{2w_1^2(\bar{w}-\mu)}{\mu^3} + \frac{3w_1^2(\bar{w}-\mu)^2}{\mu^4} \\ &\quad - \left[\frac{w_1^2}{\mu^2} - \frac{2w_1^2(\bar{w}-\mu)}{\mu^3} + \frac{3w_1^2(\bar{w}-\mu)^2}{\mu^4} \right] I(B_p^c(\epsilon)) + a, \end{aligned}$$

where $|a| \leq C_1 \frac{w_1^2}{\mu^2} \epsilon^3$. Therefore,

$$\begin{aligned} \frac{w_1^2}{\bar{w}^2} I(B_p(\epsilon)) - \frac{w_1^2}{\mu^2} &= -\frac{2w_1^2(\bar{w}-\mu)}{\mu^3} + \frac{3w_1^2(\bar{w}-\mu)^2}{\mu^4} \\ &\quad - \left[\frac{w_1^2}{\mu^2} - \frac{2w_1^2(\bar{w}-\mu)}{\mu^3} + \frac{3w_1^2(\bar{w}-\mu)^2}{\mu^4} \right] I(B_p^c(\epsilon)) + a, \end{aligned} \quad (70)$$

where $|a| \leq C_1 \frac{w_1^2}{\mu^2} \epsilon^3$. Taking expectation for (70) yields that

$$\mathbb{E}\left(\frac{w_1^2}{\bar{w}^2} I(B_p(\epsilon))\right) - \mathbb{E}\frac{w_1^2}{\mu^2} = -\frac{2\mathbb{E}w_1^2(\bar{w}-\mu)}{\mu^3} + \frac{3\mathbb{E}w_1^2(\bar{w}-\mu)^2}{\mu^4}$$

$$\begin{aligned}
& -\mathbb{E}\left[\frac{w_1^2}{\mu^2} - \frac{2w_1^2(\bar{w} - \mu)}{\mu^3} + \frac{3w_1^2(\bar{w} - \mu)^2}{\mu^4}\right]I(B_p^c(\epsilon)) + \mathbb{E}a \\
& = -\frac{2\mathbb{E}w_1^2(\bar{w} - \mu)}{\mu^3} + \frac{3\mathbb{E}w_1^2(\bar{w} - \mu)^2}{\mu^4} - \mathbb{E}b + \mathbb{E}a,
\end{aligned}$$

where $|\mathbb{E}a| \leq C_1 \frac{w_1^2}{\mu^2} \epsilon^3$. Thus,

$$\mathbb{E}\frac{w_1^2}{\bar{w}^2} - \mathbb{E}\frac{w_1^2}{\mu^2} = -\frac{2\mathbb{E}w_1^2(\bar{w} - \mu)}{\mu^3} + \frac{3\mathbb{E}w_1^2(\bar{w} - \mu)^2}{\mu^4} - \mathbb{E}b + \mathbb{E}a + \mathbb{E}\left(\frac{w_1^2}{\bar{w}^2}I(B_p^c(\epsilon))\right). \quad (71)$$

Note that

$$|\mathbb{E}c| = \left|\mathbb{E}\left(\frac{w_1^2}{\bar{w}^2}I(B_p^c(\epsilon))\right)\right| \leq p^2\mathbb{P}(B_p^c(\epsilon)). \quad (72)$$

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

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

It is obvious that

$$\left|\mathbb{E}(y_1 I(B_p^c(\epsilon)))\right| \leq C_2 \left|\mathbb{E}w_1^2 I(B_p^c(\epsilon))\right| \leq C_2 \mathbb{P}(B_p^c(\epsilon))^{1/2}. \quad (73)$$

Note that

$$\mathbb{E}\frac{w_1^2(\bar{w} - \mu)}{\mu^3} = \frac{1}{p} \sum_{i=1}^p \frac{\mathbb{E}w_1^2(w_i - \mu)}{\mu^3} = \frac{1}{p} \frac{\mathbb{E}w_1^2(w_1 - \mu)}{\mu^3}, \quad (74)$$

and

$$\begin{aligned}
\mathbb{E}\frac{w_1^2(\bar{w} - \mu)^2}{\mu^4} &= \frac{1}{p^2} \sum_{i,j=1}^p \frac{\mathbb{E}w_1^2(w_i - \mu)(w_j - \mu)}{\mu^4} \\
&= \frac{1}{p^2} \sum_{i=1}^p \frac{\mathbb{E}w_1^2(w_i - \mu)^2}{\mu^4} \\
&= \frac{1}{p^2} \frac{\mathbb{E}w_1^2(w_1 - \mu)^2}{\mu^4} + \frac{p-1}{p^2} \frac{\mathbb{E}w_1^2\mathbb{E}(w_1 - \mu)^2}{\mu^4}.
\end{aligned} \quad (75)$$

By (74), we get

$$\begin{aligned}
\left|\mathbb{E}(y_2 I(B_p^c(\epsilon)))\right| &= \frac{2}{p\mu^3} \left|\mathbb{E}w_1^2(w_1 - \mu)I(B_p^c(\epsilon))\right| \\
&\leq \frac{C_3}{p} \left(\left|\mathbb{E}w_1^2 I(B_p^c(\epsilon))\right| + \left|\mathbb{E}w_1^3 I(B_p^c(\epsilon))\right|\right) \\
&\leq \frac{C_3}{p} \left(\mathbb{P}(B_p^c(\epsilon))^{1/2} + \left|(\mathbb{E}w_1^4)^{1/2}(\mathbb{E}w_1^2 I(B_p^c(\epsilon)))^{1/2}\right|\right) \\
&\leq \frac{C_3}{p} \left(\mathbb{P}(B_p^c(\epsilon))^{1/2} + \mathbb{P}(B_p^c(\epsilon))^{1/4}\right).
\end{aligned} \quad (76)$$

By (75), we get

$$\begin{aligned}
\left| \mathbb{E}(y_3 I(B_p^c(\epsilon))) \right| &\leq \frac{1}{p^2 \mu^4} \left| \mathbb{E} w_1^2 (w_1 - \mu)^2 I(B_p^c(\epsilon)) \right| + \frac{p-1}{p^2 \mu^4} \left| (\mathbb{E} w_1^2 I(B_p^c(\epsilon))) (\mathbb{E}(w_1 - \mu)^2 I(B_p^c(\epsilon))) \right| \\
&\leq \frac{C_4}{p^2} + \frac{C_4}{p} (\mathbb{P}(B_p^c(\epsilon))^{1/2} \cdot \mathbb{P}(B_p^c(\epsilon))^{1/2}) \\
&\leq C_4 \left(\frac{1}{p^2} + \frac{1}{p} \mathbb{P}(B_p^c(\epsilon)) \right).
\end{aligned} \tag{77}$$

By (76)–(73), we have

$$\begin{aligned}
|\mathbb{E}b| &= \left| \mathbb{E}(y_1 + y_2 + y_3) I(B_p^c(\epsilon)) \right| \\
&\leq \left| \mathbb{E}(y_1 I(B_p^c(\epsilon))) \right| + \left| \mathbb{E}(y_2 I(B_p^c(\epsilon))) \right| + \left| \mathbb{E}(y_3 I(B_p^c(\epsilon))) \right| \\
&\leq C_5 \left(\mathbb{P}(B_p^c(\epsilon))^{1/2} + \frac{1}{p} \mathbb{P}(B_p^c(\epsilon))^{1/4} + p^{-2} \right).
\end{aligned} \tag{78}$$

By using (9) and $\mathbb{E} |w_1 - \mu|^4 < \infty$ and $|\frac{w_1 - \mu}{\sigma}| \leq \sqrt{n} \delta_n$, we have $\mathbb{E}a - \mathbb{E}b + \mathbb{E}c = o(p^{-1})$.

Thus, by (71), (72) and (78), we get

$$\mathbb{E} \frac{w_1^2}{\bar{w}^2} - \mathbb{E} \frac{w_1^2}{\mu^2} = -\frac{2\mathbb{E}w_1^2(\bar{w} - \mu)}{\mu^3} + \frac{3\mathbb{E}w_1^2(\bar{w} - \mu)^2}{\mu^4} + o(p^{-1}). \tag{79}$$

Plugging (74)–(75) into (79), we get

$$\begin{aligned}
\mathbb{E} \left(\frac{w_1}{\bar{w}} \right)^2 - \mathbb{E} \left(\frac{w_1}{\mu} \right)^2 &= \frac{1}{p} \left[\frac{3\mathbb{E}w_1^2 \mathbb{E}(w_1 - \mu)^2}{\mu^4} - \frac{2\mathbb{E}w_1^2 (w_1 - \mu)}{\mu^3} \right] + o(p^{-1}) \\
&= \frac{1}{p} h_1 + o(p^{-1}),
\end{aligned} \tag{80}$$

which is the first equation in Lemma 4.3.

Similar to the previous calculation, we obtain

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

where

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

$$+ \left(\frac{w_1}{\mu} - 1\right)^2 \left(\frac{w_2}{\mu} - 1\right)^2 \left(\frac{w_2^2/\mu^2}{(w_2/\mu - 1)^2} + 2\frac{w_2/\mu}{w_2/\mu - 1}\right). \quad (81)$$

Similar to (69)–(79), we get

$$\begin{aligned} & \mathbb{E}\left(\frac{w_1}{\bar{w}} - 1\right)^2 \left(\frac{w_2}{\bar{w}} - 1\right)^2 \\ &= \mathbb{E}\left(\frac{w_1}{\mu} - 1\right)^2 \left(\frac{w_2}{\mu} - 1\right)^2 + \mathbb{E}\left(\frac{\bar{w}}{\mu} - 1\right) f_1(w_1, w_2, \mu) + \mathbb{E}\left(\frac{\bar{w}}{\mu} - 1\right)^2 f_2(w_1, w_2, \mu) + o(p^{-1}) \\ &=: T_1 + T_2 + T_3 + o(p^{-1}). \end{aligned} \quad (82)$$

Similar to (74) and (75), we obtain

$$T_2 = \frac{1}{p} \mathbb{E} \sum_{i=1}^p \left(\frac{w_i}{\mu} - 1\right) f_1(w_1, w_2, \mu) = \frac{1}{p} \mathbb{E} \left(\frac{w_1}{\mu} + \frac{w_2}{\mu} - 2\right) f_1(w_1, w_2, \mu) \quad (83)$$

and

$$\begin{aligned} T_3 &= \frac{1}{p^2} \mathbb{E} \sum_{i,j=1}^p \left(\frac{w_i}{\mu} - 1\right) \left(\frac{w_j}{\mu} - 1\right) f_2(w_1, w_2, \mu) \\ &= \frac{1}{p^2} \mathbb{E} \sum_{i=1}^p \left(\frac{w_i}{\mu} - 1\right)^2 f_2(w_1, w_2, \mu) \\ &= \frac{1}{p^2} \mathbb{E} \sum_{i=1}^2 \left(\frac{w_i}{\mu} - 1\right)^2 f_2(w_1, w_2, \mu) + \frac{p-2}{p^2} \mathbb{E} \left(\frac{w_1}{\mu} - 1\right)^2 \mathbb{E} f_2(w_1, w_2, \mu). \end{aligned} \quad (84)$$

Thus, by (82) – (84), we get

$$\begin{aligned} \mathbb{E}\left(\frac{w_1}{\bar{w}} - 1\right)^2 \left(\frac{w_2}{\bar{w}} - 1\right)^2 &= \mathbb{E}\left(\frac{w_1}{\mu} - 1\right)^2 \left(\frac{w_2}{\mu} - 1\right)^2 + \frac{1}{p} \mathbb{E} \left(\frac{w_1}{\mu} + \frac{w_2}{\mu} - 2\right) f_1(w_1, w_2, \mu) \\ &\quad + \frac{1}{p} \frac{\sigma^2}{\mu^2} \mathbb{E} f_2(w_1, w_2, \mu) + o(p^{-1}) \\ &= \mathbb{E}\left(\frac{w_1}{\mu} - 1\right)^2 \left(\frac{w_2}{\mu} - 1\right)^2 + \frac{1}{p} h_2 + o(p^{-1}). \end{aligned} \quad (85)$$

which is the second equation in Lemma 4.3.

Similarly, we get

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

which is the third equation in Lemma 4.3.

B.5 Proof of Lemma 4.4

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

$$\mathbb{P}(B_p^c(\epsilon)) = \mathbb{P}(|\bar{w} - \mu| \geq \epsilon)$$

$$\begin{aligned}
&\leq \epsilon^{-kq_1} \mathbb{E} \left| \sum_{j=1}^p (w_j - \mu) / p \right|^{kq_1} \\
&\leq K_{kq_1} \epsilon^{-kq_1} p^{-kq_1} \left[\mathbb{E} \left(\sum_{j=1}^p \mathbb{E}_{j-1} |w_j - \mu|^2 \right)^{kq_1/2} + \mathbb{E} \sum_{j=1}^p |w_j - \mu|^{kq_1} \right] \\
&= K_{kq_1} \epsilon^{-kq_1} p^{-kq_1} \left[\left(\sum_{j=1}^p \mathbb{E} |w_j - \mu|^2 \right)^{kq_1/2} + \mathbb{E} \sum_{j=1}^p |w_j - \mu|^{kq_1} \right] \\
&\leq K_{kq_1} \epsilon^{-kq_1} p^{-kq_1} \left[(p\sigma^2)^{kq_1/2} + p \mathbb{E} |w_j - \mu|^{kq_1} \right] \\
&= K_{kq_1} \sigma^{kq_1} \epsilon^{-kq_1} \left[p^{-kq_1/2} + p^{-kq_1+1} \mathbb{E} \left| \frac{w_j - \mu}{\sigma} \right|^{kq_1} \right]. \tag{86}
\end{aligned}$$

Next, we prove the estimation of the q -th moment of $\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n}v_2 \text{tr} \mathbf{A}$. For any $q \geq 2$,

$$\mathbb{E} \left| \mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n}v_2 \text{tr} \mathbf{A} \right|^q \leq K_q \left(\mathbb{E} \left| \mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right|^q + \mathbb{E} \left| \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} - \frac{1}{n} v_2 \text{tr} \mathbf{A} \right|^q \right). \tag{87}$$

By Lemma 4.3, we have

$$\mathbb{E} \left| \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} - \frac{1}{n} v_2 \text{tr} \mathbf{A} \right|^q \leq K_q p^{-q} \|\mathbf{A}\|^q h_1^q. \tag{88}$$

Next, consider $\mathbb{E} \left| \mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right|^q$. There exists a positive constant K_q such that

$$\begin{aligned}
&\mathbb{E} \left| \mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right|^q \\
&\leq K_q \left(\mathbb{E} \left| \left(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right) I(B_p(\epsilon)) \right|^q + \mathbb{E} \left| \left(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right) I(B_p^c(\epsilon)) \right|^q \right). \tag{89}
\end{aligned}$$

Step 1: Since $\left| \mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right| \leq \|\mathbf{r}\| \|\mathbf{A}\| + C \|\mathbf{A}\| \leq C \cdot p \|\mathbf{A}\|$, we have

$$\mathbb{E} \left| \left(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right) I(B_p^c(\epsilon)) \right|^q \leq C \cdot p^q \|\mathbf{A}\|^q \mathbb{P}(B_p^c(\epsilon)). \tag{90}$$

Step 2: Estimating $\mathbb{E} \left| \left(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right) I(B_p(\epsilon)) \right|^q$. Write

$$\begin{aligned}
\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} &= \frac{1}{n\bar{w}^2} (\mathbf{w} - \bar{w}\mathbf{1}_p)' \mathbf{A} (\mathbf{w} - \bar{w}\mathbf{1}_p) - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \\
&= \frac{\sigma^2}{n\bar{w}^2} \left[\frac{(\mathbf{w} - \bar{w}\mathbf{1}_p)' \mathbf{A} (\mathbf{w} - \bar{w}\mathbf{1}_p)}{\sigma} - \text{tr} \mathbf{A} \right] + \left[\frac{1}{n} \frac{\sigma^2}{\bar{w}^2} \text{tr} \mathbf{A} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right] \\
&=: v_1 + v_2, \tag{91}
\end{aligned}$$

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

$$\mathbb{E} \left| \left(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right) I(B_p(\epsilon)) \right|^q \leq K_q \left(\mathbb{E} |v_1 I(B_p(\epsilon))|^q + \mathbb{E} |v_2 I(B_p(\epsilon))|^q \right).$$

By (86), we also have

$$\mathbb{E}|\bar{w} - \mu|^q \leq K_q \sigma^q (p^{-q/2} + p^{-q+1} \mathbb{E} \left| \frac{w_j - \mu}{\sigma} \right|^q). \quad (92)$$

On the event $B_p(\epsilon)$, we have $-\epsilon \leq \bar{w} - \mu \leq \epsilon$, by this fact and (92), we get

$$\begin{aligned} \mathbb{E}|v_2 I(B_p(\epsilon))|^q &= \frac{1}{n} \frac{\sigma^2}{\bar{w}^2} \text{tr} \mathbf{A} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \\ &= \sigma^{2q} \left| \frac{\text{tr} \mathbf{A}}{n} \right|^q \left| \mathbb{E} \left(\frac{1}{\bar{w}^2} - \frac{1}{\mu} \right) I(B_p(\epsilon)) \right|^q \\ &= \sigma^{2q} \left| \frac{\text{tr} \mathbf{A}}{n} \right|^q \left| \mathbb{E} \left(\frac{(\bar{w} - \mu)(\bar{w} + \mu)}{\bar{w}^2 \mu^2} \right) I(B_p(\epsilon)) \right|^q \\ &= K_{\mu, \sigma, q} \left| \frac{\text{tr} \mathbf{A}}{n} \right|^q \mathbb{E} |\bar{w} - \mu|^q \\ &\leq K_{\mu, \sigma, q} \left| \frac{\text{tr} \mathbf{A}}{n} \right|^q \left[p^{-q/2} + p^{-q+1} \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^q \right] \\ &\leq K_{\mu, \sigma, q} \|\mathbf{A}\|^q \left[p^{-q/2} + p^{-q+1} \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^q \right]. \end{aligned} \quad (93)$$

Next, we consider v_1 . Note that $\frac{w - \bar{w} \mathbf{1}_p}{\sigma} = \frac{w - \mu \mathbf{1}_p}{\sigma} - \frac{\bar{w} \mathbf{1}_p - \mu \mathbf{1}_p}{\sigma}$, we get

$$\begin{aligned} \mathbb{E}|v_1 I(B_p(\epsilon))|^q &= \mathbb{E} \left| \frac{\sigma^2}{n \bar{w}^2} \left[\frac{(w - \bar{w} \mathbf{1}_p)'}{\sigma} \mathbf{A} \frac{(w - \bar{w} \mathbf{1}_p)}{\sigma} - \text{tr} \mathbf{A} \right] I(B_p(\epsilon)) \right|^q \\ &\leq K_{\mu, \sigma, q} n^{-q} \mathbb{E} \left| \left[\frac{(w - \bar{w} \mathbf{1}_p)'}{\sigma} \mathbf{A} \frac{(w - \bar{w} \mathbf{1}_p)}{\sigma} - \text{tr} \mathbf{A} \right] I(B_p(\epsilon)) \right|^q \\ &\leq K_{\mu, \sigma, q} n^{-q} \mathbb{E} \left| \left[\frac{(w - \mu \mathbf{1}_p)'}{\sigma} \mathbf{A} \frac{(w - \mu \mathbf{1}_p)}{\sigma} - \text{tr} \mathbf{A} \right] I(B_p(\epsilon)) \right|^q \\ &\quad + K_{\mu, \sigma, q} n^{-q} \mathbb{E} \left| \left[\frac{(\bar{w} \mathbf{1}_p - \mu \mathbf{1}_p)'}{\sigma} \mathbf{A} \frac{(w - \mu \mathbf{1}_p)}{\sigma} \right] I(B_p(\epsilon)) \right|^q \\ &\quad + K_{\mu, \sigma, q} n^{-q} \mathbb{E} \left| \left[\frac{(w - \mu \mathbf{1}_p)'}{\sigma} \mathbf{A} \frac{(\bar{w} \mathbf{1}_p - \mu \mathbf{1}_p)}{\sigma} \right] I(B_p(\epsilon)) \right|^q \\ &\quad + K_{\mu, \sigma, q} n^{-q} \mathbb{E} \left| \left[\frac{(\bar{w} \mathbf{1}_p - \mu \mathbf{1}_p)'}{\sigma} \mathbf{A} \frac{(\bar{w} \mathbf{1}_p - \mu \mathbf{1}_p)}{\sigma} \right] I(B_p(\epsilon)) \right|^q \\ &=: K_{\mu, \sigma, q} n^{-q} (V_{11} + V_{12} + V_{13} + V_{14}). \end{aligned} \quad (94)$$

By Lemma 2.2 in Bai and Silverstein [2004], we have

$$V_{11} \leq K_q \left[\left(\mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^4 \text{tr}(\mathbf{A} \mathbf{A}') \right)^{q/2} + \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^{2q} \text{tr}(\mathbf{A} \mathbf{A}')^{q/2} \right]. \quad (95)$$

In the following, we consider V_{12} , V_{13} and V_{14} . Note that $\frac{(\bar{w} - \mu) \mathbf{1}_p'}{\sigma} = \frac{1}{p} \mathbf{1}_p \mathbf{1}_p' \frac{(w - \mu \mathbf{1}_p)}{\sigma}$ and

$$\mathbb{E} \frac{1}{p} \text{tr}(\mathbf{1}_p \mathbf{1}_p' \mathbf{A}) = \mathbb{E} \frac{1}{p} \text{tr}(\mathbf{A} \mathbf{1}_p \mathbf{1}_p') = \mathbb{E} \frac{1}{p^2} \text{tr}(\mathbf{1}_p \mathbf{1}_p' \mathbf{A} \mathbf{1}_p \mathbf{1}_p') \leq \sqrt{p} \|\mathbf{A}\|,$$

we get

$$\begin{aligned}
V_{12} &= K_{\mu, \sigma, q} n^{-q} \mathbb{E} \left| \left[\frac{1}{p} \frac{(\mathbf{w} - \mu \mathbf{1}_p)'}{\sigma} \mathbf{1}_p \mathbf{1}_p' \mathbf{A} \frac{(\mathbf{w} - \mu \mathbf{1}_p)}{\sigma} \right] I(B_p(\epsilon)) \right|^q \\
&\leq K_q \left(\mathbb{E} \left| \frac{(\mathbf{w} - \mu \mathbf{1}_p)'}{\sigma} \left(\frac{1}{p} \mathbf{1}_p \mathbf{1}_p' \mathbf{A} \right) \frac{(\mathbf{w} - \mu \mathbf{1}_p)}{\sigma} - \text{tr} \left(\frac{1}{p} \mathbf{1}_p \mathbf{1}_p' \mathbf{A} \right) \right|^q + \mathbb{E} \left| \text{tr} \left(\frac{1}{p} \mathbf{1}_p \mathbf{1}_p' \mathbf{A} \right) \right|^q \right) \\
&\leq K_q \left(\mathbb{E} \left| \frac{(\mathbf{w} - \mu \mathbf{1}_p)'}{\sigma} \left(\frac{1}{p} \mathbf{1}_p \mathbf{1}_p' \mathbf{A} \right) \frac{(\mathbf{w} - \mu \mathbf{1}_p)}{\sigma} - \text{tr} \left(\frac{1}{p} \mathbf{1}_p \mathbf{1}_p' \mathbf{A} \right) \right|^q + p^{q/2} \|\mathbf{A}\|^q \right) \\
&\leq V_{11}.
\end{aligned} \tag{96}$$

Similarly, we get

$$V_{13} \leq V_{11}, \quad V_{14} \leq V_{11}. \tag{97}$$

By (94)–(97), we get

$$\mathbb{E} |v_1 I(B_p(\epsilon))|^q \leq K_q n^{-q} \left[\left(\mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^4 \text{tr}(\mathbf{A}\mathbf{A}') \right)^{q/2} + \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^{2q} \text{tr}(\mathbf{A}\mathbf{A}')^{q/2} \right]. \tag{98}$$

From (89)–(93) and (98), we have

$$\begin{aligned}
&\mathbb{E} \left| \mathbf{r}' \mathbf{A} \mathbf{r} - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \mathbf{A} \right|^q \\
&\leq K_q \left(n^{-q} \left[\left(\mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^4 \text{tr}(\mathbf{A}\mathbf{A}') \right)^{q/2} + \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^{2q} \text{tr}(\mathbf{A}\mathbf{A}')^{q/2} \right] \right. \\
&\quad \left. + \|\mathbf{A}\|^q \left[p^{-q/2} + p^{-q+1} \mathbb{E} \left| \frac{w_j - \mu}{\sigma} \right|^q \right] + p^q \|\mathbf{A}\|^q \mathbb{P}(B_p^c(\epsilon)) \right) \\
&\leq K_q \left(n^{-q} \left[\left(\mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^4 \text{tr}(\mathbf{A}\mathbf{A}') \right)^{q/2} + \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^{2q} \text{tr}(\mathbf{A}\mathbf{A}')^{q/2} \right] + n^q \|\mathbf{A}\|^q \mathbb{P}(B_p^c(\epsilon)) \right). \tag{99}
\end{aligned}$$

By (88) and (99), we get

$$\begin{aligned}
&\mathbb{E} \left| \mathbf{r}' \mathbf{A} \mathbf{r} - \frac{1}{n} v_2 \text{tr} \mathbf{A} \right|^q \\
&\leq K_q \left(n^{-q} \left[\left(\mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^4 \text{tr}(\mathbf{A}\mathbf{A}') \right)^{q/2} + \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^{2q} \text{tr}(\mathbf{A}\mathbf{A}')^{q/2} \right] \right. \\
&\quad \left. + n^q \|\mathbf{A}\|^q \mathbb{P}(B_p^c(\epsilon)) + n^{-q} \|\mathbf{A}\|^q h_1^q \right). \tag{100}
\end{aligned}$$

If $\mathbb{E} |w_1 - \mu|^4 < \infty$, $\|\mathbf{A}\| \leq K$ and $|\frac{w_1 - \mu}{\sigma}| \leq \sqrt{n} \delta_n$, then, for any $q \geq 2$,

$$n^{-q} \left[\left(\mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^4 \text{tr}(\mathbf{A}\mathbf{A}') \right)^{q/2} + \mathbb{E} \left| \frac{w_1 - \mu}{\sigma} \right|^{2q} \text{tr}(\mathbf{A}\mathbf{A}')^{q/2} \right] \leq K_q n^{-1} \delta_n^{2q-4} \tag{101}$$

and

$$n^{-q} \|\mathbf{A}\|^q h_1^q \leq K_q n^{-1} \delta_n^{2q-4}. \tag{102}$$

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

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

Hence, we have

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

which is derived by the following calculation,

$$\begin{aligned} \mathbb{P}(\mathcal{B}_p^c(\epsilon)) &\leq K_{kq_1} \sigma^{kq_1} \epsilon^{-kq_1} \left[p^{-kq_1/2} + p^{-kq_1+1} \mathbb{E} \left| \frac{w_j - \mu}{\sigma} \right|^{kq_1} \right] \\ &= K_{kq_1} \sigma^{kq_1} (P_1 + P_2), \end{aligned}$$

and

$$\begin{aligned} P_2 &= \epsilon^{-kq_1} p^{-kq_1+1} \mathbb{E} \left| \frac{w_j - \mu}{\sigma} \right|^{kq_1} \\ &\leq \epsilon^{-kq_1} p^{-kq_1+1} (n^{1/2} \delta_n)^{kq_1-4} \\ &\leq \epsilon^{-kq_1} n^{-kq_1/2-1} \\ &\leq P_1. \end{aligned}$$

Therefore,

$$\mathbb{P}(\mathcal{B}_p^c(\epsilon)) \leq 2K_{kq_1} \sigma^{kq_1} P_1. \quad (103)$$

Take $\epsilon = n^{-\alpha}$ ($0 < \alpha < 1/2$) and $kq_1 > \frac{3q}{1-2\alpha}$ into (103), we have

$$n^q \mathbb{P}(\mathcal{B}_p^c(\epsilon)) \leq K_q n^q n^{-(\frac{1}{2}-\alpha)kq_1} \leq K_q n^{-q/2} \leq K_q n^{-1} \delta_n^{2q-4}.$$

B.6 Proof of Lemma 4.5

It is obvious that

$$\begin{aligned} &\mathbb{E} \left(\mathbf{r}' \mathbf{A} \mathbf{r} - \frac{1}{n} \nu_2 \operatorname{tr} \mathbf{A} \right) \left(\mathbf{r}' \mathbf{B} \mathbf{r} - \frac{1}{n} \nu_2 \operatorname{tr} \mathbf{B} \right) \\ &= \mathbb{E} \left(\mathbf{r}' \mathbf{A} \mathbf{r} \mathbf{r}' \mathbf{B} \mathbf{r} \right) - \left(\frac{1}{n} \nu_2 \operatorname{tr} \mathbf{B} \right) \mathbb{E} \left(\mathbf{r}' \mathbf{A} \mathbf{r} - \frac{1}{n} \nu_2 \operatorname{tr} \mathbf{A} \right) \\ &\quad - \left(\frac{1}{n} \nu_2 \operatorname{tr} \mathbf{A} \right) \mathbb{E} \left(\mathbf{r}' \mathbf{B} \mathbf{r} - \frac{1}{n} \nu_2 \operatorname{tr} \mathbf{B} \right) - \frac{1}{n^2} \nu_2^2 \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B}. \end{aligned} \quad (104)$$

Step i: Consider $\mathbb{E} \left(\mathbf{r}' \mathbf{A} \mathbf{r} - \frac{1}{n} \nu_2 \operatorname{tr} \mathbf{A} \right)$.

Recall that $\mathbf{r} = \frac{1}{\sqrt{n}} \left(\frac{w_1}{w} - 1, \dots, \frac{w_p}{w} - 1 \right)'$. Note that $\{w_j, 1 \leq j \leq p\}$ are i.i.d., and

$$\sum_{j=1}^p \frac{w_j}{w} = p, \quad (105)$$

and

$$\sum_{j=1}^p \left(\frac{w_j}{w} - 1 \right)^2 + \sum_{j_1 \neq j_2} \left(\frac{w_{j_1}}{w} - 1 \right) \left(\frac{w_{j_2}}{w} - 1 \right) = 0.$$

Taking expectation on the above two identities yields that $\mathbb{E}\frac{w_j}{\bar{w}} = 1$, and, for $1 \leq j_1 \neq j_2 \leq p$,

$$\mathbb{E}\left(\frac{w_{j_1}}{\bar{w}} - 1\right)\left(\frac{w_{j_2}}{\bar{w}} - 1\right) = -\frac{\nu_2}{p-1},$$

where $\nu_2 = \text{Var}\left(\frac{w_j}{\bar{w}}\right)$. This implies that

$$\mathbb{E}(\mathbf{r}\mathbf{r}') = \frac{\nu_2}{n} \left(-\frac{1}{p-1} \mathbf{1}_p \mathbf{1}_p' + \frac{1}{p-1} \mathbf{I}_p + \mathbf{I}_p \right). \quad (106)$$

By (106), we have

$$\begin{aligned} \mathbb{E}\left(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{\nu_2}{n} \text{tr} \mathbf{A}\right) &= \mathbb{E} \text{tr}(\mathbf{A}\mathbf{r}\mathbf{r}') - \frac{\nu_2}{n} \text{tr} \mathbf{A} \\ &= \text{tr}(\mathbf{A}\mathbb{E}\mathbf{r}\mathbf{r}') - \frac{\nu_2}{n} \text{tr} \mathbf{A} \\ &= -\frac{\nu_2}{n(p-1)} \sum_{k \neq l} \mathbf{A}_{kl}, \end{aligned}$$

thus,

$$\left(\frac{1}{n} \nu_2 \text{tr} \mathbf{B}\right) \mathbb{E}\left(\mathbf{r}'\mathbf{A}\mathbf{r} - \frac{1}{n} \nu_2 \text{tr} \mathbf{A}\right) = -\frac{1}{n^2(p-1)} \nu_2^2 \text{tr} \mathbf{B} \sum_{k \neq l} \mathbf{A}_{kl}. \quad (107)$$

Similarly,

$$\left(\frac{1}{n} \nu_2 \text{tr} \mathbf{A}\right) \mathbb{E}\left(\mathbf{r}'\mathbf{B}\mathbf{r} - \frac{1}{n} \nu_2 \text{tr} \mathbf{B}\right) = -\frac{1}{n^2(p-1)} \nu_2^2 \text{tr} \mathbf{A} \sum_{k \neq l} \mathbf{B}_{kl}. \quad (108)$$

Step 2: Consider $\mathbb{E}(\mathbf{r}'\mathbf{A}\mathbf{r}\mathbf{r}'\mathbf{B}\mathbf{r})$.

Let $R_j = \frac{1}{\sqrt{n}}\left(\frac{w_j}{\bar{w}} - 1\right)$, $1 \leq j \leq p$, then we have $\sum_{j=1}^p R_j = 0$, and

$$\begin{aligned} \nu_4 &:= \mathbb{E}\left(\frac{w_1}{\bar{w}} - 1\right)^4 = n^2 \mathbb{E}R_1^4, \\ \nu_{12} &:= \mathbb{E}\left[\left(\frac{w_1}{\bar{w}} - 1\right)^2 \left(\frac{w_2}{\bar{w}} - 1\right)^2\right] = n^2 \mathbb{E}(R_1^2 R_2^2). \end{aligned}$$

It follows from (105) that

$$\mathbb{E}R_1^3 R_2 = \frac{1}{p-1} \mathbb{E}\left[R_1^3 \left(\sum_{j=1}^p R_j - R_1\right)\right] = -\frac{1}{p-1} \nu_4, \quad (109)$$

$$\begin{aligned} \mathbb{E}R_1^2 R_2 R_3 &= \frac{1}{p-2} \mathbb{E}\left[R_1^2 R_2 \left(\sum_{j=1}^p R_j - R_1 - R_2\right)\right] \\ &= -\frac{1}{p-2} \mathbb{E}(R_1^3 R_2) - \frac{1}{p-2} \mathbb{E}(R_1^2 R_2^2) \\ &= \frac{1}{(p-1)(p-2)} \nu_4 - \frac{1}{p-2} \nu_{12}, \end{aligned} \quad (110)$$

$$\begin{aligned}
\mathbb{E}R_1R_2R_3R_4 &= \frac{1}{p-3}\mathbb{E}\left[R_1R_2R_3\left(\sum_{j=1}^pR_j - R_1 - R_2 - R_3\right)\right] \\
&= -\frac{3}{p-3}\mathbb{E}(R_1^2R_2R_3) \\
&= -\frac{3}{(p-1)(p-2)(p-3)}\nu_4 + \frac{3}{(p-2)(p-3)}\nu_{12}.
\end{aligned} \tag{III}$$

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

$$\mathbb{E}(\mathbf{r}'\mathbf{A}\mathbf{r}\mathbf{r}'\mathbf{B}\mathbf{r}) = \mathbb{E}\left(\sum_{i,j}R_iA_{ij}R_j\sum_{k,l}R_kB_{kl}R_l\right) = \sum_{i,j,k,l}\mathbb{E}(R_iR_jR_kR_lA_{ij}B_{kl}). \tag{II2}$$

To calculate (II2), we split it into the following cases:

1

$$i = j = k = l, \sum_i (R_i^4) A_{ii} B_{ii};$$

2

$$i = j, k = l, i \neq k, \sum_{\substack{i,k \\ i \neq k}} (R_i^2 R_k^2) A_{ii} B_{kk};$$

3

$$i = j, k \neq l, \sum_{\substack{i,k,l \\ k \neq l}} (R_i^2 R_k R_l) A_{ii} B_{kl};$$

4

$$i \neq j, k = l, \sum_{\substack{i,j,k \\ i \neq j}} (R_i R_j R_k^2) A_{ij} B_{kk};$$

5

$$i \neq j, k \neq l, i = k, j = l, \sum_{\substack{i,j \\ i \neq j}} (R_i^2 R_j^2) A_{ij} B_{ij};$$

6

$$i \neq j, k \neq l, i = l, j = k, \sum_{\substack{i,j \\ i \neq j}} (R_i^2 R_j^2) A_{ij} B_{ji};$$

7

$$i \neq j, k \neq l, i = k, l \neq j, \sum_{\substack{i,j,l \\ i \neq j \neq l}} (R_i R_j R_i R_l) A_{ij} B_{il};$$

8

$$i \neq j, k \neq l, l = j, i \neq k, \sum_{\substack{i,j,k \\ i \neq j \neq k}} (R_i R_j R_k R_j) A_{ij} B_{kj};$$

9

$$i \neq j, k \neq l, k = j, i \neq l, \sum_{\substack{i,j,l \\ i \neq j \neq l}} (R_i R_j R_k R_l) A_{ij} B_{jl};$$

10

$$i \neq j, k \neq l, i = l, k \neq j, \sum_{\substack{i,j,k \\ i \neq j \neq k}} (R_i R_j R_k R_i) A_{ij} B_{ki};$$

11

$$i \neq j, k \neq l, l \neq j, i \neq k, \sum_{\substack{i,j,k,l \\ i \neq j \neq k \neq l}} (R_i R_j R_k R_l) A_{ij} B_{kl}.$$

For ease of presentation, we still keep v_4 in the expectations although we have obtained its value. The expectations of all cases are listed as follows.

Case 1:

$$\mathbb{E} \sum_i (R_i^4) A_{ii} B_{ii} = \mathbb{E} (R_i^4) \sum_i A_{ii} B_{ii} = \frac{1}{n^2} v_4 \sum_i A_{ii} B_{ii}.$$

Case 2:

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

Case 3: From (109) and (110), we obtain

$$\begin{aligned} \mathbb{E} \sum_{\substack{i,k,l \\ k \neq l}} (R_i^2 R_k R_l) A_{ii} B_{kl} &= \mathbb{E} R_1^2 R_2 R_3 \sum_{\substack{i,k,l \\ i \neq k \neq l}} A_{ii} B_{kl} + \mathbb{E} R_1^3 R_2 \sum_{\substack{i,l \\ l \neq i}} A_{ii} B_{il} + \mathbb{E} R_1^3 R_2 \sum_{\substack{i,k \\ k \neq i}} A_{ii} B_{ki} \\ &= \left(\frac{v_4}{n^2(p-1)(p-2)} - \frac{v_{12}}{n^2(p-2)} \right) \sum_{\substack{i,k,l \\ i \neq k \neq l}} A_{ii} B_{kl} \\ &\quad - \frac{v_4}{n^2(p-1)} \left(\sum_{\substack{i,l \\ l \neq i}} A_{ii} B_{il} + \sum_{\substack{i,k \\ k \neq i}} A_{ii} B_{ki} \right). \end{aligned} \quad (113)$$

Note that

$$\mathbb{E} \text{tr} \mathbf{A} \leq \mathbb{E} \left(p \sum_i A_{ii}^2 \right)^{1/2} \leq \mathbb{E} \left(p \text{tr}(\mathbf{A} \mathbf{A}') \right)^{1/2} \leq \mathbb{E} \left(p^2 \|\mathbf{A}\|^2 \right)^{1/2} = p \mathbb{E} \|\mathbf{A}\|, \quad (114)$$

and

$$\begin{aligned} \mathbb{E} \left| \sum_{i,k,l} A_{ii} B_{kl} \right| &= \mathbb{E} \left| \text{tr} \mathbf{A} \mathbf{1}'_p \mathbf{B} \mathbf{1}_p \right| \\ &\leq p \mathbb{E} (\|\mathbf{A}\| \cdot |\mathbf{1}'_p \mathbf{B} \mathbf{1}_p|) \end{aligned}$$

$$\begin{aligned}
&\leq p^2 \mathbb{E}(\|\mathbf{A}\| \cdot \|\mathbf{B}\|) \\
&\leq p^2 (\mathbb{E}\|\mathbf{A}\|^{q_1})^{1/q_1} (\mathbb{E}\|\mathbf{B}\|)^{1/q_2} \\
&= O(p^2),
\end{aligned} \tag{II5}$$

and, for $q_1, q_2 > 0$ and $1/q_1 + 1/q_2 = 1$,

$$\begin{aligned}
\mathbb{E} \left| \sum_{i,l} A_{ii} B_{il} \right| &= \mathbb{E} \left| \sum_i A_{ii} (\mathbf{1}_p^i)' \mathbf{B} \mathbf{1}_p \right| \\
&\leq \mathbb{E} \left(\left| \sum_i A_{ii}^2 \right| \right)^{1/2} \left(\left| \sum_i ((\mathbf{1}_p^i)' \mathbf{B} \mathbf{1}_p) ((\mathbf{1}_p^i)' \mathbf{B} \mathbf{1}_p) \right| \right)^{1/2} \\
&\leq p \mathbb{E} \|\mathbf{A}\| \cdot \|(\mathbf{1}_p' \mathbf{B} \mathbf{1}_p) (\mathbf{1}_p' \mathbf{B} \mathbf{1}_p)\|^{1/2} \\
&\leq p^{3/2} \mathbb{E} \|\mathbf{A}\| \cdot \|\mathbf{B}\| \\
&\leq p^{3/2} (\mathbb{E}\|\mathbf{A}\|^{q_1})^{1/q_1} (\mathbb{E}\|\mathbf{B}\|)^{1/q_2} \\
&= O(p^{3/2}),
\end{aligned} \tag{II6}$$

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

$$\mathbb{E} \sum_{i,l} A_{ii} B_{il} = \mathbb{E} \operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{B} \leq p^2 \mathbb{E}(\|\mathbf{A}\| \cdot \|\mathbf{B}\|) = O(p^2), \tag{II7}$$

and

$$\mathbb{E} \sum_i A_{ii} B_{ii} \leq p \mathbb{E} \|\operatorname{diag}(\mathbf{A}) \operatorname{diag}(\mathbf{B})\| \leq p \mathbb{E}(\|\mathbf{A}\| \cdot \|\mathbf{B}\|). \tag{II8}$$

Thus, by (II5)–(II8), we have

$$\mathbb{E} \sum_{\substack{i,j,k,l \\ i \neq k \neq l}} A_{ii} B_{kl} = \mathbb{E} \left(\sum_{i,k,l} A_{ii} B_{kl} - \sum_{i,l} A_{ii} B_{il} - \sum_{i,l} A_{ii} B_{li} + \sum_i A_{ii} B_{ii} \right) = O(p^2). \tag{II9}$$

It follows from (II3)–(II9) that

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

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

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

Case 5:

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

Case 6:

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

Case 7: By (ii0), we have

$$\begin{aligned} \mathbb{E} \sum_{\substack{i,j,l \\ i \neq j \neq l}} (R_i R_j R_l) A_{ij} B_{il} &= \mathbb{E} (R_1^2 R_2 R_3) \sum_{\substack{i,j,l \\ i \neq j \neq l}} A_{ij} B_{il} \\ &= \left(\frac{1}{n^2(p-1)(p-2)} \nu_4 - \frac{1}{n^2(p-2)} \nu_{12} \right) \sum_{\substack{i,j,l \\ i \neq j \neq l}} A_{ij} B_{il}. \end{aligned} \quad (120)$$

Note that, $\mathbb{E} \mathbf{1}'_p \mathbf{A}' \mathbf{B} \mathbf{1}_p \leq p \mathbb{E}(\|\mathbf{A}\| \cdot \|\mathbf{B}\|)$, $\mathbb{E} \text{tr}(\mathbf{A}' \mathbf{B}) \leq p \mathbb{E}(\|\mathbf{A}\| \cdot \|\mathbf{B}\|)$, and by (ii6) and (ii8), we have

$$\mathbb{E} \sum_{i \neq l} A_{ii} B_{il} = \mathbb{E} \sum_{i,l} A_{ii} B_{il} - \mathbb{E} \sum_i A_{ii} B_{ii} = O(p^{3/2}), \quad (121)$$

thus,

$$\mathbb{E} \sum_{\substack{i,j,l \\ i \neq j \neq l}} A_{ij} B_{il} = \mathbb{E} \left(\mathbf{1}'_p \mathbf{A}' \mathbf{B} \mathbf{1}_p - \text{tr}(\mathbf{A}' \mathbf{B}) - 2 \sum_{i \neq l} A_{ii} B_{il} \right) = O(p^{3/2}). \quad (122)$$

It follows from (120)–(122) that

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

Case 8: Similarly to Case 7, we have

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

Case 9: Similarly to Case 7, we have

$$\mathbb{E} \sum_{\substack{i,j,l \\ i \neq j \neq l}} (R_i R_j R_j R_l) A_{ij} B_{jl} = \left(\frac{\nu_4}{n^2(p-1)(p-2)} - \frac{\nu_{12}}{n^2(p-2)} \right) \sum_{\substack{i,j,l \\ i \neq j \neq l}} A_{ij} B_{jl} = o(n^{-1}).$$

Case 10: Similarly to Case 7, we have

$$\mathbb{E} \sum_{\substack{i \neq j, k \\ i \neq j \neq k}} (R_i R_j R_k R_i) A_{ij} B_{ki} = \left(\frac{\nu_4}{n^2(p-1)(p-2)} - \frac{\nu_{12}}{n^2(p-2)} \right) \sum_{\substack{i,j,k \\ i \neq j \neq k}} A_{ij} B_{ki} = o(n^{-1}).$$

Case II: By (iii), we have

$$\begin{aligned}
& \mathbb{E} \sum_{\substack{i,j,k,l \\ i \neq j \neq k \neq l}} (R_i R_j R_k R_l) A_{ij} B_{kl} \\
&= \mathbb{E} (R_1 R_2 R_3 R_4) \sum_{\substack{i,j,k,l \\ i \neq j \neq k \neq l}} A_{ij} B_{kl} \\
&= \left(-\frac{3\nu_4}{n^2(p-1)(p-2)(p-3)} + \frac{3\nu_{12}}{n^2(p-2)(p-3)} \right) \sum_{\substack{i,j,k,l \\ i \neq j \neq k \neq l}} A_{ij} B_{kl}.
\end{aligned}$$

Note that, by (i22) we have

$$\begin{aligned}
\mathbb{E} \sum_{\substack{i,j,k,l \\ i \neq j \neq k \neq l}} A_{ij} B_{kl} &= \mathbb{E} (\mathbf{1}'_p \mathbf{A} \mathbf{1}_p - \text{tr} \mathbf{A}) (\mathbf{1}'_p \mathbf{B} \mathbf{1}_p - \text{tr} \mathbf{B}) - 2 \mathbb{E} \sum_{\substack{i,k,l \\ i \neq k \neq l}} A_{ik} B_{il} \\
&\leq K p^2 \mathbb{E} (\|\mathbf{A}\| \cdot \|\mathbf{B}\|) - 2 \mathbb{E} \sum_{\substack{i,k,l \\ i \neq k \neq l}} A_{ik} B_{il} = O(p^2).
\end{aligned}$$

Thus,

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

Combining Cases I – II gives us

$$\begin{aligned}
\mathbb{E} (\mathbf{r}' \mathbf{A} \mathbf{r} \mathbf{r}' \mathbf{B} \mathbf{r}) &= \frac{\nu_4 - 3\nu_{12}}{n^2} \sum_{i=1}^p A_{ii} B_{ii} + \frac{\nu_{12}}{n^2} (\text{tr} \mathbf{A} \text{tr} \mathbf{B} + \text{tr}(\mathbf{A} \mathbf{B}') + \text{tr}(\mathbf{A} \mathbf{B})) \\
&\quad - \frac{1}{n^2(p-2)} \nu_{12} \left(\sum_{\substack{i,k,l \\ i \neq k, k \neq l, i \neq l}} A_{ii} B_{kl} + \sum_{\substack{i,j,k \\ i \neq j, j \neq k, i \neq k}} A_{ij} B_{kk} \right) + o(n^{-1}).
\end{aligned}$$

Step 3: By (i07)–(i08), we have

$$\begin{aligned}
& - \left(\frac{\nu_2}{n} \text{tr} \mathbf{B} \right) \mathbb{E} (\mathbf{r}' \mathbf{A} \mathbf{r} - \frac{\nu_2}{n} \text{tr} \mathbf{A}) - \left(\frac{\nu_2}{n} \text{tr} \mathbf{A} \right) \mathbb{E} (\mathbf{r}' \mathbf{B} \mathbf{r} - \frac{\nu_2}{n} \text{tr} \mathbf{B}) \\
&= \frac{\nu_2^2}{n^2(p-1)} \text{tr} \mathbf{B} \sum_{k \neq l} \mathbf{A}_{kl} + \frac{\nu_2^2}{n^2(p-1)} \text{tr} \mathbf{A} \sum_{k \neq l} \mathbf{B}_{kl} \\
&= \frac{\nu_2^2}{n^2(p-1)} \\
&\quad \times \left(\sum_{\substack{k,l \\ k \neq l}} B_{kk} A_{kl} + \sum_{\substack{k,l \\ k \neq l}} A_{kk} B_{kl} + \sum_{\substack{k,l \\ k \neq l}} B_{ll} A_{kl} + \sum_{\substack{k,l \\ k \neq l}} A_{ll} B_{kl} + \sum_{\substack{i,k,l \\ i \neq k, k \neq l, i \neq l}} B_{ii} A_{kl} + \sum_{\substack{i,k,l \\ i \neq k, k \neq l, i \neq l}} A_{ii} B_{kl} \right) \\
&= \frac{\nu_2^2}{n^2(p-1)} \left(\sum_{\substack{i,k,l \\ i \neq k, k \neq l, i \neq l}} B_{ii} A_{kl} + \sum_{\substack{i,k,l \\ i \neq k, k \neq l, i \neq l}} A_{ii} B_{kl} \right) + o(n^{-1}).
\end{aligned}$$

Following the above calculations, we have

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

B.7 Proof of Lemma 4.9

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

$$|\operatorname{tr}(\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z)) \mathbf{A}| \leq \frac{\|\mathbf{A}\|}{\mathfrak{J}(z)}. \quad (123)$$

By Lemma A.2 and equation (123), similarly to (4.3) in Bai and Silverstein [1998],

$$\mathbb{E} \left| \frac{\nu_2}{n} \operatorname{tr} \mathbf{D}^{-1}(z) - \mathbb{E} \frac{\nu_2}{n} \operatorname{tr} \mathbf{D}^{-1}(z) \right|^q \leq C_q n^{-q/2} \nu_0^{-q}, \quad (124)$$

which implies that

$$\mathbb{E} |\bar{\beta}_j(z) - \mathbf{b}_p(z)|^2 \leq \frac{K|z|^4}{n\nu_0^6}. \quad (125)$$

Now, we prove that the difference between

$$\sum_{j=1}^n [\mathbb{E}_{j-1} \bar{\beta}_j(z_1) \varepsilon_j(z_1)] [\mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2)]$$

and

$$\mathbf{b}_p(z_1)\mathbf{b}_p(z_2) \sum_{j=1}^n [\mathbb{E}_{j-1} \varepsilon_j(z_1)] [\mathbb{E}_{j-1} (\varepsilon_j(z_2))]$$

converges to zero. We write

$$\begin{aligned} & \mathbb{E} \left| [\mathbb{E}_{j-1} \bar{\beta}_j(z_1) \varepsilon_j(z_1)] [\mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2)] \right. \\ & \quad \left. - \mathbf{b}_p(z_1)\mathbf{b}_p(z_2) [\mathbb{E}_{j-1} \varepsilon_j(z_1)] [\mathbb{E}_{j-1} (\varepsilon_j(z_2))] \right| \\ & \leq \mathbb{E} \left| [\mathbb{E}_{j-1} (\bar{\beta}_j(z_1) \varepsilon_j(z_1) - \mathbf{b}_p(z_1) \varepsilon_j(z_1))] [\mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2)] \right| \\ & \quad + \mathbb{E} \left| [\mathbb{E}_{j-1} \mathbf{b}_p(z_1) \varepsilon_j(z_1)] [\mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2) - \mathbb{E}_{j-1} (\mathbf{b}_p(z_2) \varepsilon_j(z_2))] \right| \\ & =: I_1 + I_2. \end{aligned}$$

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

$$\begin{aligned} I_1 &= \mathbb{E} \left| [\mathbb{E}_{j-1} (\bar{\beta}_j(z_1) - \mathbf{b}_p(z_1)) \varepsilon_j(z_1)] [\mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2)] \right| \\ &\leq \left(\mathbb{E} \left| \mathbb{E}_{j-1} (\bar{\beta}_j(z_1) - \mathbf{b}_p(z_1)) \varepsilon_j(z_1) \right|^{q_1} \right)^{1/q_1} \left(\mathbb{E} \left| \mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2) \right|^{q_2} \right)^{1/q_2} \\ &\leq \left(\mathbb{E} \left| (\bar{\beta}_j(z_1) - \mathbf{b}_p(z_1)) \varepsilon_j(z_1) \right|^{q_1} \right)^{1/q_1} \left(\mathbb{E} \left| \bar{\beta}_j(z_2) \varepsilon_j(z_2) \right|^{q_2} \right)^{1/q_2} \\ &=: I_{11} \times I_{12}. \end{aligned} \tag{126}$$

From Lemma 4.4 and (125), for any $1 < q_1 < 2, q_2 > 2$, we have

$$\begin{aligned} (I_{11})^{q_1} &= \mathbb{E} \left| [(\bar{\beta}_j(z_1) - \mathbf{b}_p(z_1)) \varepsilon_j(z_1)] \right|^{q_1} \\ &\leq \left(\mathbb{E} \left| \bar{\beta}_j(z_1) - \mathbf{b}_p(z_1) \right|^{q_1 \frac{2}{q_1}} \right)^{q_1/2} \left(\mathbb{E} \left| \varepsilon_j(z_1) \right|^{q_1 \frac{2}{2-q_1}} \right)^{\frac{2-q_1}{2}} \\ &\leq \mathbf{K} n^{-\frac{q_1}{2}} n^{-\frac{2-q_1}{2} q_1} \delta_n^{4q_1-4} = n^{-1} \delta_n^{4q_1-4}, \end{aligned} \tag{127}$$

and

$$I_{12} = \mathbb{E} \left| \bar{\beta}_j(z_2) \varepsilon_j(z_2) \right|^{q_2} \leq \mathbf{K} \frac{|z|^{q_2}}{v^{q_2}} n^{-1} \delta_n^{2q_2-4}.$$

Thus,

$$I_1 \leq \mathbf{K} n^{-1} \delta_n^{\frac{4q_1-4}{q_1} + \frac{2q_2-4}{q_2}} = \mathbf{K} n^{-1} \delta_n^2. \tag{128}$$

Similarly, we can obtain $I_2 \leq \mathbf{K} n^{-1} \delta_n^2$. Thus,

$$\begin{aligned} & \mathbb{E} \left| [\mathbb{E}_{j-1} \bar{\beta}_j(z_1) \varepsilon_j(z_1)] [\mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2)] \right. \\ & \quad \left. - \mathbf{b}_p(z_1)\mathbf{b}_p(z_2) [\mathbb{E}_{j-1} \varepsilon_j(z_1)] [\mathbb{E}_{j-1} (\varepsilon_j(z_2))] \right| \\ & \leq \mathbf{K} n^{-1} \delta_n^2. \end{aligned} \tag{129}$$

From (129), we get

$$\begin{aligned} & \sum_{j=1}^n [\mathbb{E}_{j-1} \bar{\beta}_j(z_1) \varepsilon_j(z_1)] [\mathbb{E}_{j-1} \bar{\beta}_j(z_2) \varepsilon_j(z_2)] \\ & - \mathbf{b}_p(z_1) \mathbf{b}_p(z_2) \sum_{j=1}^n [\mathbb{E}_{j-1} \varepsilon_j(z_1)] [\mathbb{E}_{j-1} (\varepsilon_j(z_2))] \xrightarrow{i.p.} 0. \end{aligned} \quad (I30)$$

Next, we prove that

$$\mathbf{b}_p(z_1) \mathbf{b}_p(z_2) \sum_{j=1}^n [\mathbb{E}_{j-1} \varepsilon_j(z_1)] [\mathbb{E}_{j-1} (\varepsilon_j(z_2))] \xrightarrow{i.p.} 0. \quad (I31)$$

By (106), we get

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

Thus,

$$\begin{aligned} & \sum_{j=1}^n [\mathbb{E}_{j-1} \varepsilon_j(z_1)] [\mathbb{E}_{j-1} \varepsilon_j(z_2)] \\ &= \frac{\nu_2^2}{n^2(p-1)^2} \sum_{j=1}^n \mathbb{E}_{j-1} \left(\text{tr} \mathbf{D}_j^{-1}(z_1) - \mathbf{1}_p' \mathbf{D}_j^{-1}(z_1) \mathbf{1}_p \right) \\ & \quad \times \mathbb{E}_{j-1} \left(\text{tr} \mathbf{D}_j^{-1}(z_2) - \mathbf{1}_p' \mathbf{D}_j^{-1}(z_2) \mathbf{1}_p \right). \end{aligned} \quad (I32)$$

By Lemma 2.3 in Bai and Silverstein [2004], we have

$$\frac{1}{p} \text{tr} [\mathbf{D}_j^{-1}(z)] \xrightarrow{L^2} m(z), \text{ as } p \rightarrow \infty, \quad (I33)$$

and by Lemma A.4,

$$\frac{1}{p} \mathbf{1}_p' \mathbf{D}_j^{-1}(z) \mathbf{1}_p \xrightarrow{L^2} -\frac{1}{z}, \text{ as } p \rightarrow \infty. \quad (I34)$$

By (124), we have

$$|\mathbf{b}_p(z) - \mathbb{E} \beta_1(z)| \leq K n^{-1/2}. \quad (I35)$$

From the formula (2.2) of Silverstein [1995], $\underline{\mathbf{m}}_p(z) = -\frac{1}{zn} \sum_{j=1}^n \beta_j(z)$, we have

$$\mathbb{E}\beta_1(z) = -z\mathbb{E}\underline{\mathbf{m}}_p(z). \quad (I36)$$

From (I35), (I36), and Lemma 4.4, we have

$$|\mathbf{b}_p(z) + z\underline{\mathbf{m}}_p^0(z)| \leq \mathbf{K}n^{-1/2}. \quad (I37)$$

From (I32) – (I37), we get

$$\mathbf{b}_p(z_1)\mathbf{b}_p(z_2) \sum_{j=1}^n [\mathbb{E}_{j-1} \varepsilon_j(z_1)] [\mathbb{E}_{j-1} (\varepsilon_j(z_2))] \xrightarrow{i.p.} 0.$$

This, together with (I30), completes the proof of Lemma 4.9.

B.8 Proof of Lemma 4.10

Using Lemma 4.4 and equation (I25), we have, for $i < j$,

$$\begin{aligned} & \mathbb{E} \left| \beta_{ij}(z_2) \mathbf{r}'_i \mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \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 \right. \\ & \quad \left. - n^{-2} \nu_2^2 \mathbf{b}_1(z_2) \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2)) \operatorname{tr}(\mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1)) \right| \\ \leq & \mathbb{E} \left| \beta_{ij}(z_2) \mathbf{r}'_i \mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \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 \right. \\ & \quad \left. - n^{-2} \nu_2^2 \beta_{ij}(z_2) \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2)) \operatorname{tr}(\mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1)) \right| \\ & \quad + \mathbb{E} \left| n^{-2} \nu_2^2 (\beta_{ij}(z_2) - \mathbf{b}_1(z_2)) \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2)) \operatorname{tr}(\mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1)) \right| \\ \leq & \mathbb{E} \left| \beta_{ij}(z_2) \left[\mathbf{r}'_i \mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2) \mathbf{r}_i - \frac{1}{n} \nu_2 \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2)) \right] \mathbf{r}'_i \mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1) \mathbf{r}_i \right. \\ & \quad \left. + \frac{1}{n} \nu_2 \beta_{ij}(z_2) \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2)) \left[\mathbf{r}'_i \mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1) \mathbf{r}_i - \frac{1}{n} \nu_2 \operatorname{tr}(\mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1)) \right] \right| \\ & \quad + \mathbb{E} \left| n^{-2} \nu_2^2 (\beta_{ij}(z_2) - \mathbf{b}_1(z_2)) \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2)) \operatorname{tr}(\mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1)) \right| \\ \leq & \mathbf{K}n^{-1/2}. \end{aligned} \quad (I38)$$

By (I23), we have

$$\begin{aligned} & \left| \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_{ij}^{-1}(z_1)) \mathbf{D}_{ij}^{-1}(z_2)) \operatorname{tr}(\mathbf{D}_{ij}^{-1}(z_2) \mathbf{Q}_p(z_1)) \right. \\ & \quad \left. - \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_j^{-1}(z_1)) \mathbf{D}_j^{-1}(z_2)) \operatorname{tr}(\mathbf{D}_j^{-1}(z_2) \mathbf{Q}_p(z_1)) \right| \leq \mathbf{K}n. \end{aligned} \quad (I39)$$

It follows from (I38) and (I39) that

$$\mathbb{E} \left| \mathbf{A}_1(z_1, z_2) + \frac{j-1}{n^2} \nu_2^2 \mathbf{b}_1(z_2) \operatorname{tr}(\mathbb{E}_j(\mathbf{D}_j^{-1}(z_1)) \mathbf{D}_j^{-1}(z_2)) \operatorname{tr}(\mathbf{D}_j^{-1}(z_2) \mathbf{Q}_p(z_1)) \right| \leq \mathbf{K}n^{1/2}, \quad (I40)$$

where $\mathbf{A}_1(z_1, z_2)$ is defined in (38). Note that

$$\operatorname{tr}(\mathbb{E}_j(\mathbf{D}_j^{-1}(z_1)) \mathbf{D}_j^{-1}(z_2)) = \operatorname{tr}(\mathbb{E}_j[-\mathbf{Q}_p(z_1) + \mathbf{b}_1(z) \mathbf{A}(z_1)] \mathbf{D}_j^{-1}(z_2)) + O(n^{1/2})$$

$$\begin{aligned}
&= -\operatorname{tr}(\mathbf{Q}_p(z_1)\mathbf{D}_j^{-1}(z_2)) + b_1(z_1)\operatorname{tr}\mathbb{E}_j(\mathbf{A}(z_1))\mathbf{D}_j^{-1}(z_2) + O(n^{1/2}) \\
&= -\operatorname{tr}(\mathbf{Q}_p(z_1)\mathbf{D}_j^{-1}(z_2)) + b_1(z_1)\mathbf{A}_1(z_1, z_2) + O(n^{1/2}).
\end{aligned}$$

Therefore, from (32)–(37) and (138)–(140), we can write

$$\begin{aligned}
&\operatorname{tr}(\mathbb{E}_j(\mathbf{D}_j^{-1}(z_1))\mathbf{D}_j^{-1}(z_2)) \left[1 + \frac{j-1}{n^2} v_2^2 b_1(z_1) b_1(z_2) \operatorname{tr}(\mathbf{D}_j^{-1}(z_2) \mathbf{Q}_p(z_1)) \right] \\
&= -\operatorname{tr}(\mathbf{Q}_p(z_1)\mathbf{D}_j^{-1}(z_2)) + \mathbf{A}_4(z_1, z_2),
\end{aligned}$$

where $\mathbb{E}|\mathbf{A}_4(z_1, z_2)| \leq \mathbf{K}n^{1/2}$. Using the expression for $\mathbf{D}_j^{-1}(z_2)$ in (32), (25), and (34)–(39), we have

$$\begin{aligned}
&\operatorname{tr}(\mathbb{E}_j(\mathbf{D}_j^{-1}(z_1))\mathbf{D}_j^{-1}(z_2)) \left[1 - \frac{j-1}{n^2} v_2^2 b_1(z_1) b_1(z_2) \operatorname{tr}(\mathbf{Q}_p(z_2) \mathbf{Q}_p(z_1)) \right] \\
&= \operatorname{tr}(\mathbf{Q}_p(z_2) \mathbf{Q}_p(z_1)) + \mathbf{A}_5(z_1, z_2), \tag{141}
\end{aligned}$$

where $\mathbb{E}|\mathbf{A}_5(z_1, z_2)| \leq \mathbf{K}n^{1/2}$. By (137) and (141), we get

$$\begin{aligned}
&\operatorname{tr}(\mathbb{E}_j(\mathbf{D}_j^{-1}(z_1))\mathbf{D}_j^{-1}(z_2)) \left\{ 1 - \frac{j-1}{n^2} v_2^2 \underline{\mathbf{m}}_p^0(z_1) \underline{\mathbf{m}}_p^0(z_2) \right. \\
&\quad \left. \times \operatorname{tr} \left[\left(\mathbf{I}_p + \frac{n-1}{n} v_2 \underline{\mathbf{m}}_p^0(z_1) \right)^{-1} \left(\mathbf{I}_p + \frac{n-1}{n} v_2 \underline{\mathbf{m}}_p^0(z_2) \right)^{-1} \right] \right\} \\
&= \frac{1}{z_1 z_2} \operatorname{tr} \left[\left(\mathbf{I}_p + \frac{n-1}{n} v_2 \underline{\mathbf{m}}_p^0(z_1) \right)^{-1} \left(\mathbf{I}_p + \frac{n-1}{n} v_2 \underline{\mathbf{m}}_p^0(z_2) \right)^{-1} \right] + \mathbf{A}_6(z_1, z_2), \tag{142}
\end{aligned}$$

where $\mathbb{E}|\mathbf{A}_6(z_1, z_2)| \leq \mathbf{K}n^{1/2}$. Thus, we complete the proof of in Lemma 4.10.

B.9 Proof of Lemma 4.11

By Burkholder's inequality, the inequality $|\beta_{ij}(z)| \leq \frac{|z|}{v_0}$, and Lemma 4.4, we get

$$\begin{aligned}
&\mathbb{E} |(\mathbf{1}_p^i)'(\mathbf{D}_1^{-1}(z_1) - \mathbb{E}\mathbf{D}_1^{-1}(z_1))\mathbf{1}_p^i|^2 \\
&= \mathbb{E} \left| \sum_{j=1}^n \left(\mathbb{E}_j(\mathbf{1}_p^i)' \mathbf{D}_1^{-1}(z_1) \mathbf{1}_p^i - \mathbb{E}_{j-1}(\mathbf{1}_p^i)' \mathbf{D}_1^{-1}(z_1) \mathbf{1}_p^i \right) \right|^2 \\
&= \mathbb{E} \left| \sum_{j=1}^n \left[\mathbb{E}_j(\mathbf{1}_p^i)' \left(\mathbf{D}_1^{-1}(z_1) - \mathbf{D}_{1j}^{-1}(z_1) \right) \mathbf{1}_p^i - \mathbb{E}_{j-1}(\mathbf{1}_p^i)' \left(\mathbf{D}_1^{-1}(z_1) - \mathbf{D}_{1j}^{-1}(z_1) \right) \mathbf{1}_p^i \right] \right|^2 \\
&= \mathbb{E} \left| \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})(\mathbf{1}_p^i)' \left(\mathbf{D}_1^{-1}(z_1) - \mathbf{D}_{1j}^{-1}(z_1) \right) \mathbf{1}_p^i \right|^2 \\
&\leq \mathbf{K} \sum_{j=1}^n \mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_{1j}(z_1) \mathbf{r}_j' \mathbf{D}_{1j}^{-1} \mathbf{1}_p^i (\mathbf{1}_p^i)' \mathbf{D}_{1j}^{-1} \mathbf{r}_j \right|^2 \\
&\leq \mathbf{K} \sum_{j=1}^n \mathbb{E} |\beta_{1j}(z_1) \mathbf{r}_j' \mathbf{D}_{1j}^{-1} \mathbf{1}_p^i (\mathbf{1}_p^i)' \mathbf{D}_{1j}^{-1} \mathbf{r}_j|^2 \\
&\leq \mathbf{K}n^{-1}, \tag{143}
\end{aligned}$$

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

$$\begin{aligned} & \mathbb{E} \left| \frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^p \mathbb{E}_j(\mathbf{D}_j^{-1}(z_1) - \mathbb{E} \mathbf{D}_j^{-1}(z_1))_{ii} \mathbb{E}_j(\mathbf{D}_j^{-1}(z_2))_{ii} \right| \\ & \leq \frac{nK}{n^2 \nu_0} \sum_{i=1}^p \mathbb{E} \left| (\mathbf{1}_p^i)' (\mathbf{D}_1^{-1}(z_1) - \mathbb{E} \mathbf{D}_1^{-1}(z_1)) \mathbf{1}_p^i \right| \leq Kn^{-1/2}, \end{aligned}$$

and thus

$$\frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^p \mathbb{E}_j(\mathbf{D}_j^{-1}(z_1) - \mathbb{E} \mathbf{D}_j^{-1}(z_1))_{ii} \mathbb{E}_j(\mathbf{D}_j^{-1}(z_2))_{ii} = O_P(n^{-1/2}). \quad (144)$$

With (144), it remains to find the limit of

$$\frac{1}{p} \sum_{i=1}^p \mathbb{E}(\mathbf{D}_j^{-1}(z_1))_{ii} \mathbb{E}(\mathbf{D}_j^{-1}(z_2))_{ii}. \quad (145)$$

It is easy to see that the sum of expectations in (145) is exactly the same for any j . Moreover, we have

$$\frac{1}{p} \sum_{i=1}^p \mathbb{E}(\mathbf{D}_j^{-1}(z_1))_{ii} \mathbb{E}(\mathbf{D}_j^{-1}(z_2))_{ii} \stackrel{i.p.}{\rightarrow} m(z_1)m(z_2).$$

This completes the proof of Lemma 4.11.

B.10 Proof of Lemma 4.12

First, we prove the first equation in Lemma 4.12. For $\mathbb{E}(\gamma_1(z)P_1(z))$, we have

$$\begin{aligned} & \mathbb{E}(\gamma_1(z)P_1(z)) \\ & = \mathbb{E} \left[\gamma_1(z) \left(n\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 - \frac{\sigma^2}{\mu^2} \text{tr} \left(\tilde{\mathbf{Q}}_p^{-1}(z) \mathbb{E} \mathbf{D}_1^{-1}(z) \right) \right) \right] \\ & = n\mathbb{E} \left[\gamma_1(z) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 \right] - \frac{\sigma^2/\mu^2}{1 + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z)} \text{tr}[\mathbb{E} \mathbf{D}_1^{-1}(z)] \mathbb{E} \gamma_1(z) \\ & = n\mathbb{E} \left[\left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} \mathbf{D}_1^{-1}(z) + \frac{1}{n} \nu_2 \text{tr} \mathbf{D}_1^{-1}(z) - \frac{1}{n} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_1^{-1}(z) \right) \right. \\ & \quad \times \left. \left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \right] + \frac{1}{n} \nu_2 \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \right] \right) \right] \\ & \quad - \frac{\sigma^2/\mu^2}{1 + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z)} \text{tr}[\mathbb{E} \mathbf{D}_1^{-1}] \mathbb{E} \gamma_1(z) \\ & = n\mathbb{E} \left[\left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} \mathbf{D}_1^{-1}(z) \right) \left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \right] \right) \right] \\ & \quad + \frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\text{tr} \mathbf{D}_1^{-1}(z) \text{tr}[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)] \right) - \frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\mathbf{1}'_p \mathbf{D}_1^{-1}(z) \mathbf{1}_p \text{tr}[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)] \right) \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{n} \left(1 + \frac{1}{p-1}\right) \nu_2^2 \text{Cov} \left(\text{tr} \mathbf{D}_1^{-1}(z), \text{tr} [\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)] \right) \\
& - \frac{1}{n(p-1)} \nu_2^2 \left[\mathbb{E} \left(\text{tr} \mathbf{D}_1^{-1}(z) [\mathbf{1}'_p \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{1}_p] \right) - \mathbb{E} \left(\text{tr} \mathbf{D}_1^{-1}(z) \right) \mathbb{E} \left(\mathbf{1}'_p \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{1}_p \right) \right] \\
& - \frac{\sigma^2/\mu^2}{1 + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z)} \text{tr} [\mathbb{E} \mathbf{D}_1^{-1}(z)] \mathbb{E} \gamma_1(z) \\
= & n \mathbb{E} \left[\left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} \mathbf{D}_1^{-1}(z) \right) \left(\mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \nu_2 \text{tr} [\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)] \right) \right] \\
& + \frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\text{tr} \mathbf{D}_1^{-1}(z) \text{tr} [\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)] \right) - \frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\mathbf{1}'_p \mathbf{D}_1^{-1}(z) \mathbf{1}_p \text{tr} [\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)] \right) \\
& - \frac{\sigma^2/\mu^2}{1 + \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z)} \text{tr} [\mathbb{E} \mathbf{D}_1^{-1}(z)] \mathbb{E} \gamma_1(z) + o(1),
\end{aligned}$$

which is the first equation in Lemma 4.12. Below are some interpretations of the above equalities:

1. The fourth equality uses the following derivation: By (io6), i.e., $\mathbb{E}(\mathbf{r}\mathbf{r}') = \frac{1}{n} \nu_2 \left(-\frac{1}{p-1} \mathbf{1}_p \mathbf{1}'_p + \frac{1}{p-1} \mathbf{I}_p + \mathbf{I}_p \right)$, we get

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

2. The last equality is due to

$$\frac{1}{n} \text{Cov}(\text{tr} \mathbf{D}_1^{-1}(z), \text{tr}[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)]) = O(n^{-1}),$$

which follows from (42).

Next, we prove the second equation in Lemma 4.12. For $\mathbb{E}(\beta_1(z) \gamma_1^2(z) \mathbf{P}_1(z))$, we have

$$\begin{aligned} & \mathbb{E}(\beta_1(z) \gamma_1^2(z) \mathbf{P}_1(z)) \\ &= \mathbb{E}\left(n \beta_1(z) \gamma_1^2(z) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1\right) - \mathbb{E}(\beta_1(z) \gamma_1^2(z)) \mathbb{E}\left(\frac{\sigma^2}{\mu^2} \text{tr} \left[\tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right]\right) \\ &= \mathbb{E}\left(n \beta_1(z) \gamma_1^2(z) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1\right) - \mathbb{E}\left(\beta_1(z) \gamma_1^2(z) \text{tr} \left[\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right]\right) \\ & \quad + \text{Cov}\left(\beta_1(z) \gamma_1^2(z), \text{tr} \left[\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right]\right). \end{aligned} \tag{146}$$

From Lemma 4.4 and equation (42), we have

$$\begin{aligned} & \mathbb{E}\left(n \beta_1(z) \gamma_1^2(z) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1\right) - \mathbb{E}\left(\beta_1(z) \gamma_1^2(z) \text{tr} \left[\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right]\right) \\ & \leq n \left[\mathbb{E}|\gamma_1^2(z) \beta_1(z)|^2 \right]^{1/2} \left[\mathbb{E} \left| \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 - \frac{1}{n} \frac{\sigma^2}{\mu^2} \text{tr} \left(\tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right) \right|^2 \right]^{1/2} \\ & \leq \mathbf{K} n (n^{-1} \delta_n^4)^{1/2} n^{-1/2} = \mathbf{K} \delta_n^2, \end{aligned} \tag{147}$$

and

$$\begin{aligned} & \text{Cov}\left(\beta_1(z) \gamma_1^2(z), \text{tr} \left[\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right]\right) \\ & \leq (\mathbb{E}|\beta_1(z)|^4)^{1/4} (\mathbb{E}|\gamma_1^2(z)|^4)^{1/4} \\ & \quad \times \left(\mathbb{E} \left| \text{tr} \left(\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right) - \mathbb{E} \text{tr} \left(\frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right) \right|^2 \right)^{1/2} \\ & \leq \mathbf{K} n^{-1/4} \delta_n^3. \end{aligned} \tag{148}$$

From (146), (147), and (148), we complete the proof of equation (50).

Finally, we consider (51). From (106), $\mathbf{D}_1^{-1}(z) - \mathbf{D}^{-1}(z) = \beta_1(z) \mathbf{D}_1^{-1}(z) \mathbf{r}_1 \mathbf{r}'_1 \mathbf{D}_1^{-1}(z)$, $\beta_1(z) = \mathbf{b}_p(z) - \mathbf{b}_p(z) \gamma_1(z) \beta_1(z)$ and $\mathbb{E} \beta_1(z) = \mathbf{b}_p(z) + o(n^{-1/2})$, we have

$$\begin{aligned} & \mathbb{E}(\beta_1(z) \mathbf{P}_2(z)) \\ &= \frac{\sigma^2}{\mu^2} \mathbb{E}(\beta_1(z)) \text{tr} \left[\tilde{\mathbf{Q}}_p^{-1}(z) (\mathbb{E} \mathbf{D}_1^{-1}(z) - \mathbb{E} \mathbf{D}^{-1}(z)) \right] \\ &= \frac{\sigma^2}{\mu^2} \mathbb{E}(\beta_1(z)) \text{tr} \left[\tilde{\mathbf{Q}}_p^{-1}(z) \mathbb{E}(\beta_1(z) \mathbf{D}_1^{-1}(z) \mathbf{r}_1 \mathbf{r}'_1 \mathbf{D}_1^{-1}(z)) \right] \\ &= \frac{\sigma^2}{\mu^2} \mathbb{E}(\beta_1(z)) \mathbb{E} \left[(\mathbf{b}_p(z) - \mathbf{b}_p(z) \beta_1(z) \gamma_1(z)) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{r}_1 \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbf{b}_p(z) \frac{\sigma^2}{\mu^2} \mathbb{E}(\beta_1(z)) \left[\mathbb{E} \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{r}_1 \right. \\
&\quad \left. - \mathbb{E} \beta_1(z) \gamma_1(z) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{r}_1 \right] \\
&= \mathbf{b}_p^2(z) \frac{\sigma^2}{\mu^2} \left[\mathbb{E} \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{r}_1 \right] + O(n^{-1/2}) \\
&= \mathbf{b}_p^2(z) \frac{\sigma^2}{\mu^2} \text{tr} \left[\mathbb{E} \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbb{E} \mathbf{r}_1 \mathbf{r}'_1 \right] + O(n^{-1/2}) \\
&= \mathbf{b}_p^2(z) \frac{\sigma^2}{\mu^2} \text{tr} \left[\mathbb{E} \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right. \\
&\quad \left. \times \frac{1}{n} \nu_2 \left(-\frac{1}{p-1} \mathbf{1}_p \mathbf{1}'_p + \frac{1}{p-1} \mathbf{I}_p + \mathbf{I}_p \right) \right] + O(n^{-1/2}) \\
&= \frac{p}{n(p-1)} \nu_2 \frac{\sigma^2}{\mu^2} \mathbf{b}_p^2(z) \mathbb{E} \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right] \\
&\quad - \frac{1}{n(p-1)} \nu_2 \frac{\sigma^2}{\mu^2} \mathbf{b}_p^2(z) \mathbb{E} \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{1}_p \mathbf{1}'_p \right] + O(n^{-1/2}) \\
&= \frac{p}{n(p-1)} \nu_2 \frac{\sigma^2}{\mu^2} \mathbf{b}_p^2(z) \mathbb{E} \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \right] + O(n^{-1/2}),
\end{aligned}$$

which is the equation (51). Below are some interpretations of the above equalities:

1. The fifth equality follows from

$$\mathbb{E}(\beta_1(z) \gamma_1(z) \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{r}_1) \leq K n^{-1/2},$$

which can be proved by using Lemma 4.4.

2. The last equality follows from

$$\frac{1}{n(p-1)} \nu_2 \frac{\sigma^2}{\mu^2} \mathbf{b}_p^2(z) \mathbb{E} \text{tr} \left[\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{1}_p \mathbf{1}'_p \right] = O(n^{-1}).$$

This can be proved by using Lemma 4.3 and Lemma A.4.

B.II Proof of Lemma 4.13

Step 1: Consider J_1 . By (13), we get

$$\begin{aligned}
n \mathbb{E} \gamma_1(z) &= n \left(\frac{1}{n} \nu_2 \mathbb{E} \text{tr} \left[\left(-\frac{1}{p-1} \mathbf{1}_p \mathbf{1}'_p + \frac{1}{p-1} \mathbf{I}_p + \mathbf{I}_p \right) \mathbf{D}_1^{-1}(z) \right] - \frac{1}{n} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_1^{-1}(z) \right) \\
&= -\frac{1}{p-1} \nu_2 \mathbb{E} \text{tr} (\mathbf{1}'_p \mathbf{D}_1^{-1}(z) \mathbf{1}_p) + \frac{1}{p-1} \nu_2 \mathbb{E} \text{tr} \mathbf{D}_1^{-1}(z). \tag{149}
\end{aligned}$$

By Lemma A.4, we have

$$-\mathbf{b}_p^2(z) \left[\frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\text{tr} \mathbf{D}_1^{-1}(z) \text{tr} (\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)) \right) \right]$$

$$\begin{aligned}
& - \frac{1}{n(p-1)} \nu_2^2 \mathbb{E} \left(\mathbf{1}'_p \mathbf{D}_1^{-1}(z) \mathbf{1}_p \operatorname{tr}(\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z)) \right) \Big] \\
& \xrightarrow{\text{i.p.}} \frac{-cz^2 \underline{\mathbf{m}}^2(z) \sigma^4 / \mu^4}{1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z)} \left\{ \mathbf{m}^2(z) + \frac{\mathbf{m}(z)}{z} \right\}. \tag{I50}
\end{aligned}$$

By Lemma 2.3 in [Bai and Silverstein \[2004\]](#), we have

$$\frac{1}{p} \operatorname{tr} [\mathbf{D}_j^{-1}(z)] \xrightarrow{L^2} \mathbf{m}(z), \text{ as } p \rightarrow \infty. \tag{I51}$$

By Lemma 4.3, equations (149), (150), (151), and (41), we have

$$\mathbf{J}_1 \xrightarrow{\text{i.p.}} \left(\frac{-z \underline{\mathbf{m}}(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z)} \right) \left(\frac{\sigma^2}{\mu^2} \mathbf{m}(z) + \frac{\sigma^2}{\mu^2} \frac{1}{z} + h_1 \mathbf{m}(z) \right). \tag{I52}$$

Step 2: Consider \mathbf{J}_2 . By Lemma 4.5, we have

$$\mathbf{J}_2 = \mathbf{J}_{21} + \mathbf{J}_{22} + \mathbf{J}_{23} + \mathbf{J}_{24} + o(1), \tag{I53}$$

where

$$\begin{aligned}
\mathbf{J}_{21} &= -n b_p^2(z) \left[\frac{1}{n^2} (\nu_4 - 3\nu_{12}) \sum_{i=1}^p \mathbb{E} [(\mathbf{D}_1^{-1}(z))_{ii} (\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z))_{ii}] \right], \\
\mathbf{J}_{22} &= -n b_p^2(z) \left[\frac{1}{n^2} \nu_{12} \mathbb{E} \operatorname{tr}(\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z)) \right], \\
\mathbf{J}_{23} &= -n b_p^2(z) \left[\frac{1}{n^2} \nu_{12} \mathbb{E} \operatorname{tr}(\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z) (\mathbf{D}_1^{-1}(z))') \right], \\
\mathbf{J}_{24} &= -n b_p^2(z) \left[\frac{1}{n^2} (\nu_{12} - \nu_2^2) \mathbb{E} [\operatorname{tr} \mathbf{D}_1^{-1}(z) \operatorname{tr}(\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z))] \right].
\end{aligned}$$

Since $\frac{1}{p} \sum_{i=1}^p \mathbb{E}(\mathbf{D}_1^{-1}(z))_{ii} (\mathbf{D}_1^{-1}(z) \tilde{\mathbf{Q}}_p^{-1}(z))_{ii} \xrightarrow{\text{i.p.}} \frac{\mathbf{m}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z)}$,

$$\mathbf{J}_{21} \xrightarrow{\text{i.p.}} \frac{-cz^2 \mathbf{m}^2(z) \underline{\mathbf{m}}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z)} \left(\frac{\mathbb{E} |w_1 - \mu|^4}{\mu^4} - 3 \frac{|\mathbb{E} |w_1 - \mu|^2|^2}{\mu^4} \right). \tag{I54}$$

Note that

$$\frac{1}{p} \mathbb{E} \operatorname{tr} \mathbf{D}_1^{-2}(z) \xrightarrow{\text{i.p.}} \mathbf{m}'(z), \tag{I55}$$

thus we get

$$\mathbf{J}_{22} \xrightarrow{\text{i.p.}} \frac{-cz^2 \mathbf{m}'(z) \underline{\mathbf{m}}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z)} \frac{|\mathbb{E} |w_1 - \mu|^2|^2}{\mu^4}. \tag{I56}$$

Similarly, we have

$$\mathbf{J}_{23} \xrightarrow{\text{i.p.}} \frac{-cz^2 \mathbf{m}'(z) \underline{\mathbf{m}}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{\mathbf{m}}(z)} \frac{|\mathbb{E} |w_1 - \mu|^2|^2}{\mu^4}. \tag{I57}$$

By Lemma 4.3, we get

$$J_{24} \xrightarrow{i.p.} -c \left(h_2 - 2 \frac{\sigma^2}{\mu^2} h_1 \right) \frac{z^2 \underline{m}^2(z) \underline{m}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{m}(z)}. \quad (I58)$$

From (I53), (I54), (I56), (I57), and (I58), we have

$$\begin{aligned} J_2 \xrightarrow{i.p.} & - \frac{cz^2 \underline{m}^2(z) \underline{m}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{m}(z)} \left(\frac{\mathbb{E}|w_1 - \mu|^4}{\mu^4} - 3 \frac{|\mathbb{E}|w_1 - \mu|^2|^2}{\mu^4} \right) \\ & - \frac{2cz^2 \underline{m}'(z) \underline{m}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{m}(z)} \frac{|\mathbb{E}|w_1 - \mu|^2|^2}{\mu^4} - c \left(h_2 - 2 \frac{\sigma^2}{\mu^2} h_1 \right) \frac{z^2 \underline{m}^2(z) \underline{m}^2(z)}{1 + \frac{\sigma^2}{\mu^2} \underline{m}(z)}. \end{aligned} \quad (I59)$$

Step 3: Consider J_3 . To calculate the limit of J_3 , we can expand $\mathbf{D}_1^{-1}(z)$ like (32) and find the limit of J_3 using the method similarly to Bai and Silverstein [2004]. The limit of J_3 is

$$J_3 \xrightarrow{i.p.} c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right)^{-3} \left[1 - c \frac{\sigma^4}{\mu^4} \underline{m}^2(z) \left(1 + \frac{\sigma^2}{\mu^2} \underline{m}(z) \right)^{-2} \right]^{-1}. \quad (I60)$$

From (I52), (I59), and (I60), the proof is completed.

B.12 Proof of Lemma A.4

By Lemma 4.4, we obtain, for any $2 \leq r \in \mathbb{N}^+$,

$$\begin{aligned} & \mathbb{E} |\mathbf{r}'_j \mathbf{D}_j^{-1}(z) \mathbf{x}_p \mathbf{x}'_p \mathbf{D}_j^{-1}(z) \mathbf{r}_j|^r \\ & \leq K_r \left(\mathbb{E} \left| \mathbf{r}'_j \mathbf{D}_j^{-1}(z) \mathbf{x}_p \mathbf{x}'_p \mathbf{D}_j^{-1}(z) \mathbf{r}_j - \frac{1}{n} \nu_2 \operatorname{tr}[\mathbf{D}_j^{-1}(z) \mathbf{x}_p \mathbf{x}'_p \mathbf{D}_j^{-1}(z)] \right|^r \right. \\ & \quad \left. + \frac{1}{n} \nu_2 \mathbb{E} \left| \operatorname{tr}[\mathbf{D}_j^{-1}(z) \mathbf{x}_p \mathbf{x}'_p \mathbf{D}_j^{-1}(z)] \right|^r \right) \\ & \leq K_r (n^{-2} \delta_n^{2r-4} + n^{-r}), \end{aligned} \quad (I61)$$

where $\frac{1}{n} \operatorname{tr}[\mathbf{D}_j^{-1}(z) \mathbf{x}_p \mathbf{x}'_p \mathbf{D}_j^{-1}(z)] = \frac{1}{n} \mathbf{x}'_p \mathbf{D}_j^{-2}(z) \mathbf{x}_p = \frac{1}{n p} \mathbf{1}'_p \mathbf{D}_j^{-2}(z) \mathbf{1}_p = O(n^{-1})$.

Write

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

By Burkholder's inequality, (I61), and $|\beta_j(z)| \leq \frac{|z|}{v_0}$, we have

$$\mathbb{E} |\mathbf{x}'_p \mathbf{D}^{-1}(z) \mathbf{x}_p - \mathbf{x}'_p \mathbb{E} \mathbf{D}^{-1}(z) \mathbf{x}_p|^2$$

$$\begin{aligned}
&\leq K \sum_{j=1}^n \mathbb{E} |(\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_j(z) \mathbf{r}'_j \mathbf{D}_j^{-1}(z) \mathbf{x}_p \mathbf{x}'_p \mathbf{D}_j^{-1}(z) \mathbf{r}_j|^2 \\
&\leq K \sum_{j=1}^n \mathbb{E} |\beta_j(z) \mathbf{r}'_j \mathbf{D}_j^{-1}(z) \mathbf{x}_p \mathbf{x}'_p \mathbf{D}_j^{-1}(z) \mathbf{r}_j|^2 \\
&\leq K n^{-1}.
\end{aligned} \tag{162}$$

Thus, we have

$$\mathbb{E} |\mathbf{x}'_p \mathbf{D}^{-1}(z) \mathbf{x}_p - \mathbf{x}'_p \mathbb{E} \mathbf{D}^{-1}(z) \mathbf{x}_p|^2 \rightarrow 0. \tag{163}$$

Note that $\mathbf{D}(z) + z \tilde{\mathbf{Q}}_p(z) = \sum_{j=1}^n \mathbf{r}_j \mathbf{r}'_j + z \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z) \mathbf{I}_p$, where $\tilde{\mathbf{Q}}_p(z) = \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z) \mathbf{I}_p + \mathbf{I}_p$. Recalling $\underline{\mathbf{m}}_p(z) = -\frac{1}{nz} \sum_{j=1}^n \beta_j(z)$ and using the identity $\mathbf{r}'_j \mathbf{D}^{-1}(z) = \beta_j(z) \mathbf{r}'_j \mathbf{D}_j^{-1}(z)$, we obtain

$$\begin{aligned}
&(-z \tilde{\mathbf{Q}}_p(z))^{-1} - \mathbf{D}^{-1}(z) \\
&= - (z \tilde{\mathbf{Q}}_p(z))^{-1} (\mathbf{D}(z) + z \tilde{\mathbf{Q}}_p(z)) \mathbf{D}^{-1}(z) \\
&= -z^{-1} \tilde{\mathbf{Q}}_p^{-1}(z) \left(\sum_{j=1}^n \mathbf{r}_j \mathbf{r}'_j + z \frac{\sigma^2}{\mu^2} \mathbb{E} \underline{\mathbf{m}}_p(z) \mathbf{I}_p \right) \mathbf{D}^{-1}(z) \\
&= -z^{-1} \sum_{j=1}^n \beta_j(z) \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_j \mathbf{r}'_j \mathbf{D}_j^{-1}(z) - z^{-1} \tilde{\mathbf{Q}}_p^{-1}(z) \left(-\frac{\sigma^2}{\mu^2} \mathbb{E} \beta_1(z) \right) \mathbf{D}^{-1}(z).
\end{aligned}$$

Taking expectation of the above identity yields that

$$\begin{aligned}
&(-z \tilde{\mathbf{Q}}_p(z))^{-1} - \mathbb{E} \mathbf{D}^{-1}(z) \\
&= -z^{-1} n \mathbb{E} \beta_1(z) \left[\tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) - \frac{1}{n} \frac{\sigma^2}{\mu^2} \tilde{\mathbf{Q}}_p^{-1}(z) \mathbb{E} \mathbf{D}^{-1}(z) \right].
\end{aligned}$$

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

$$\begin{aligned}
&\mathbf{x}'_p \mathbb{E} \mathbf{D}^{-1}(z) \mathbf{x}_p - \mathbf{x}'_p (-z \tilde{\mathbf{Q}}_p(z))^{-1} \mathbf{x}_p \\
&= z^{-1} n \mathbb{E} \left\{ \beta_1(z) \left[\mathbf{x}'_p \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \mathbf{x}_p - \frac{1}{n} \frac{\sigma^2}{\mu^2} \mathbf{x}'_p \tilde{\mathbf{Q}}_p^{-1}(z) \mathbb{E} \mathbf{D}^{-1}(z) \mathbf{x}_p \right] \right\} \\
&=: \delta_1 + \delta_2 + \delta_3,
\end{aligned}$$

where

$$\begin{aligned}
\delta_1 &= \frac{n}{z} \mathbb{E} (\beta_1(z) \alpha_1(z)), \\
\alpha_1(z) &= \mathbf{x}'_p \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{r}_1 \mathbf{r}'_1 \mathbf{D}_1^{-1}(z) \mathbf{x}_p - \frac{1}{n} \frac{\sigma^2}{\mu^2} \mathbf{x}'_p \tilde{\mathbf{Q}}_p^{-1}(z) \mathbf{D}_1^{-1}(z) \mathbf{x}_p, \\
\delta_2 &= \frac{1}{z} \frac{\sigma^2}{\mu^2} \mathbb{E} \beta_1(z) \mathbf{x}'_p \tilde{\mathbf{Q}}_p^{-1}(z) (\mathbf{D}_1^{-1}(z) - \mathbf{D}^{-1}(z)) \mathbf{x}_p, \\
\delta_3 &= \frac{1}{z} \frac{\sigma^2}{\mu^2} \mathbb{E} \beta_1(z) \mathbf{x}'_p \tilde{\mathbf{Q}}_p^{-1}(z) (\mathbf{D}^{-1}(z) - \mathbb{E} \mathbf{D}^{-1}(z)) \mathbf{x}_p.
\end{aligned}$$

Recalling the notations defined above and the following equalities:

$$\begin{aligned}\delta_1 &= \frac{\mathfrak{n}}{z} \mathbb{E} \bar{\beta}_1(z) \alpha_1(z) - \frac{\mathfrak{n}}{z} \mathbb{E} [\beta_1(z) \bar{\beta}_1(z) \varepsilon_1(z) \alpha_1(z)], \\ \bar{\beta}_1(z) &= \mathfrak{b}_p(z) - \frac{\nu_2}{\mathfrak{n}} \mathfrak{b}_p(z) \bar{\beta}_1(z) \text{tr}(\mathbf{D}_1^{-1}(z) - \mathbb{E} \mathbf{D}_1^{-1}(z)), \\ \mathbb{E} \alpha_1(z) &= -(\nu_2 \mathbb{E} \underline{\mathfrak{m}}_p(z) + 1)^{-1} \frac{\nu_2}{\mathfrak{n}} [\mathbb{E} \mathbf{x}'_p \mathbf{D}_1^{-1}(z) \mathbf{x}_p + o(1)].\end{aligned}$$

From Lemma 4.3 and (106), it is easy to see that

$$\mathfrak{n} \mathbb{E} \beta_1(z) \alpha_1(z) = \left[\frac{1}{1 + c \nu_2 \mathbb{E} \underline{\mathfrak{m}}_p(z)} + o(1) \right] \mathfrak{n} \mathbb{E} \alpha_1(z).$$

Therefore, $\delta_1 = \frac{\nu_2 z \mathbb{E} \underline{\mathfrak{m}}_p(z)}{\nu_2 z \mathbb{E} \underline{\mathfrak{m}}_p(z) + z} \mathbf{x}'_p \mathbb{E}(\mathbf{D}_1^{-1}(z)) \mathbf{x}_p + o(1)$. Similarly to Bai et al. [2007], one may have $\delta_2 = o(1)$ and $\delta_3 = o(1)$. Hence, we obtain

$$\left(1 - \frac{\nu_2 z \mathbb{E} \underline{\mathfrak{m}}_p(z)}{\nu_2 z \mathbb{E} \underline{\mathfrak{m}}_p(z) + z} \right) \mathbf{x}'_p \mathbb{E}(\mathbf{D}_1^{-1}(z)) \mathbf{x}_p + \frac{1}{\sigma^2 / \mu^2 z \mathbb{E} \underline{\mathfrak{m}}_p(z) + z} \rightarrow 0,$$

which implies that

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

This, together with (163), completes the proof of Lemma A.4.

B.13 Proof of Corollary 2.6

For ease of presentation, we denote $\lambda = \sigma^2 / \mu^2$ in this section. Note that

$$z = -\frac{1}{\underline{\mathfrak{m}}} + \frac{c\lambda}{1 + \lambda \underline{\mathfrak{m}}}, \quad (164)$$

$$\begin{aligned}dz &= \left[\underline{\mathfrak{m}}^{-2} - c\lambda^2 (1 + \lambda \underline{\mathfrak{m}})^{-2} \right] d\underline{\mathfrak{m}} \\ &= \left[1 - c\lambda^2 \underline{\mathfrak{m}}^2 (1 + \lambda \underline{\mathfrak{m}}^{-2}) \right] \underline{\mathfrak{m}}^{-2} d\underline{\mathfrak{m}},\end{aligned} \quad (165)$$

$$\underline{\mathfrak{m}} = \frac{1}{c} \left(\underline{\mathfrak{m}} + \frac{1-c}{z} \right), \quad (166)$$

$$\underline{\mathfrak{m}}' = \frac{1}{c} \left(\underline{\mathfrak{m}}' - \frac{1-c}{z^2} \right). \quad (167)$$

For $1 \leq r_1, r_2 \in \mathbb{N}^+$,

$$\begin{aligned}& \text{Cov}(X_{x^{r_1}}, X_{x^{r_2}}) \\ &= 2 \times (\lambda c)^{r_1+r_2} \sum_{k_1=0}^{r_1-1} \sum_{k_2=0}^{r_2} \binom{r_1}{k_1} \binom{r_2}{k_2} \left(\frac{1-c}{c} \right)^{k_1+k_2} \\ & \quad \times \sum_{\ell=1}^{r_1-k_1} \ell \binom{2r_1-1-(k_1+\ell)}{r_1-1} \binom{2r_2-1-k_2+\ell}{r_2-1} \\ & \quad + \frac{1}{\lambda^2} (\beta + h_2 - 2\lambda h_1) (\lambda c)^{r_1+r_2}\end{aligned} \quad (168)$$

$$\times \sum_{k_1=0}^{r_1} \sum_{k_2=0}^{r_2} \binom{r_1}{k_1} \binom{r_2}{k_2} \left(\frac{1-c}{c}\right)^{k_1+k_2} \binom{2r_1-k_1}{r_1-1} \binom{2r_2-k_2}{r_2-1}. \quad (169)$$

The proof of (168) is exactly analogous with Bai and Silverstein [2004], it is then omitted. Next, we prove (169). 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 $F^{c,H}$. Note that

$$\oint_{\mathcal{C}_1} \oint_{\mathcal{C}_2} \frac{z_1^{r_1} z_2^{r_2}}{(1+\lambda \underline{m}_1)^2 (1+\lambda \underline{m}_2)^2} d\underline{m}_1 d\underline{m}_2 = \oint_{\mathcal{C}_1} \frac{z_1^{r_1}}{(1+\lambda \underline{m}_1)^2} d\underline{m}_1 \times \oint_{\mathcal{C}_2} \frac{z_2^{r_2}}{(1+\lambda \underline{m}_2)^2} d\underline{m}_2. \quad (170)$$

By (164),

$$\begin{aligned} \oint_{\mathcal{C}_1} \frac{z_1^{r_1}}{(1+\lambda \underline{m}_1)^2} d\underline{m}_1 &= \oint_{\mathcal{C}_1} \frac{\left(-\frac{1}{\underline{m}_1} + \frac{c\lambda}{1+\lambda \underline{m}_1}\right)^{r_1}}{(1+\lambda \underline{m}_1)^2} d\underline{m}_1 \\ &= (\lambda c)^{r_1} \oint_{\mathcal{C}_1} \left(\frac{1}{1+\lambda \underline{m}_1} + \frac{1-c}{c}\right)^{r_1} (1 - (1+\lambda \underline{m}_1))^{-r_1} (1+\lambda \underline{m}_1)^{-2} d\underline{m}_1 \\ &= (\lambda c)^{r_1} \oint_{\mathcal{C}_1} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1} \left(\frac{1-c}{c}\right)^{k_1} (1+\lambda \underline{m}_1)^{k_1-r_1} \\ &\quad \times \sum_{j=0}^{\infty} \binom{r_1+j-1}{j} (1+\lambda \underline{m}_1)^j (1+\lambda \underline{m}_1)^{-2} d\underline{m}_1 \\ &= (\lambda c)^{r_1} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1} \left(\frac{1-c}{c}\right)^{k_1} \\ &\quad \times \oint_{\mathcal{C}_1} \sum_{j=0}^{\infty} \binom{r_1+j-1}{j} (1+\lambda \underline{m}_1)^{k_1-r_1+j-2} d\underline{m}_1, \end{aligned} \quad (171)$$

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

$$\begin{aligned} \oint_{\mathcal{C}_1} \frac{z_1^{r_1}}{(1+\lambda \underline{m}_1)^2} d\underline{m}_1 &= \frac{1}{\lambda} (\lambda c)^{r_1} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1} \left(\frac{1-c}{c}\right)^{k_1} \\ &\quad \times \oint_{\mathcal{C}_1} \sum_{j=0}^{\infty} \binom{r_1+j-1}{j} (1+\underline{m}_1)^{k_1-r_1+j-2} d\underline{m}_1. \end{aligned} \quad (172)$$

For integral (172), the pole is -1 , we have by residual theorem

$$\oint_{\mathcal{C}_1} \frac{z_1^{r_1}}{(1+\lambda \underline{m}_1)^2} d\underline{m}_1 = \frac{2\pi i}{\lambda} (\lambda c)^{r_1} \sum_{k_1=0}^{r_1} \binom{r_1}{k_1} \left(\frac{1-c}{c}\right)^{k_1} \binom{2r_1-k_1}{r_1-1}. \quad (173)$$

Similarly, we get

$$\oint_{\mathcal{C}_2} \frac{z_2^{r_2}}{(1+\lambda \underline{m}_2)^2} d\underline{m}_2 = \frac{2\pi i}{\lambda} (\lambda c)^{r_2} \sum_{k_2=0}^{r_2} \binom{r_2}{k_2} \left(\frac{1-c}{c}\right)^{k_2} \binom{2r_2-k_2}{r_2-1}. \quad (174)$$

By (173) and (174), (169) is derived.

For $f_1 = x$,

$$\begin{aligned}
\mathbb{E}X_x &= \frac{1}{2\pi i} \oint_e c\lambda^2 z \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} dz \\
&\quad - \frac{1}{2\pi i} \oint_e z \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times (h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z) dz \\
&\quad - \frac{1}{2\pi i} \oint_e cz^3 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-1} [\beta \underline{m}^2(z) + (h_2 - 2\lambda h_1) \underline{m}^2(z) + 2\lambda^2 \underline{m}'(z)] dz \\
&=: I_1(f_1) + I_2(f_1) + I_3(f_1).
\end{aligned} \tag{175}$$

For $I_1(f_1)$, we get

$$\begin{aligned}
I_1(f_1) &= \frac{1}{2\pi i} \oint_e c\lambda^2 z \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} dz \\
&= \frac{1}{2\pi i} \oint_e c\lambda^2 z \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times [\underline{m}^{-2} - c\lambda^2(1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= \frac{1}{2\pi i} \oint_e \frac{c\lambda^2(-c\lambda \underline{m} + \lambda \underline{m} + 1)}{(1 + \lambda \underline{m})^2 [c(\lambda \underline{m})^2 - (1 + \lambda \underline{m})^2]} d\underline{m}.
\end{aligned}$$

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

$$I_1(f_1) = \lambda(1 + c) - \frac{\lambda}{2}(1 + \sqrt{c})^2 - \frac{\lambda}{2}(1 - \sqrt{c})^2 = 0. \tag{176}$$

For the second integral $I_2(f_1)$, the poles are $-\lambda^{-1}$, we have, by the residue theorem,

$$\begin{aligned}
I_2(f_1) &= -\frac{1}{2\pi i} \oint_e z \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times (h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z) dz \\
&= -\frac{1}{2\pi i} \oint_e z \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times \left(h_1 \underline{m}(z) + \lambda \underline{m}(z) + \frac{\lambda}{z} + (h_1 + \lambda) \frac{1 - c}{cz} \right) \\
&\quad \times [\underline{m}^{-2} - c\lambda^2(1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= -\frac{1}{2\pi i} \oint_e \frac{(\lambda^2 \underline{m} - h_1)(-c\lambda \underline{m} + \lambda \underline{m} + 1)}{\underline{m}(1 + \lambda \underline{m})^3} d\underline{m} \\
&= h_1.
\end{aligned} \tag{177}$$

For $I_3(f_1)$, we get

$$\begin{aligned}
I_3(f_1) &= -\frac{1}{2\pi i} \oint_e cz z^2 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times [1 + \lambda \underline{m}(z)]^{-1} [\beta \underline{m}^2(z) + (h_2 - 2\lambda h_1) \underline{m}^2(z) + 2\lambda^2 \underline{m}'(z)] dz \\
&= -\frac{1}{2\pi i} \oint_e cz^3 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times [1 + \lambda \underline{m}(z)]^{-1} \left[(\beta + h_2 - 2\lambda h_1) \left(\frac{1}{c^2} \underline{m}^2 z^2 + \frac{2(1-c)}{c^2} \underline{m} z + \frac{(1-c)^2}{c^2} \right) \right. \\
&\quad \left. + 2\lambda^2 \left(\frac{1}{c} z^2 \underline{m}' - \frac{1-c}{c} \right) \right] \times [\underline{m}^{-2} - c\lambda^2 (1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= -\frac{1}{2\pi i} \oint_e \left[c(\lambda \underline{m} + 1)^4 (c\lambda^2 \underline{m}^2 - (\lambda \underline{m} + 1)^2) \right]^{-1} \\
&\quad \times \left[2c\lambda^2 (\lambda \underline{m} + 1)^2 ((c-1)[c\lambda^2 \underline{m}^2 - (\lambda \underline{m} + 1)^2] - (-c\lambda \underline{m} + \lambda \underline{m} + 1)^2) \right. \\
&\quad \left. + (c\lambda^2 \underline{m}^2 - (\lambda \underline{m} + 1)^2) (-2\lambda h_1 + \beta + h_2) \right. \\
&\quad \left. \times (-c\lambda \underline{m} + \lambda \underline{m} + (c-1)(\lambda \underline{m} + 1) + 1)^2 \right] \\
&\quad \times (-c\lambda \underline{m} + \lambda \underline{m} + 1) d\underline{m}
\end{aligned}$$

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

$$I_3(f_1) = (1+c) + \lambda(1+\sqrt{c})^2 + \lambda(1-\sqrt{c})^2 = 0. \quad (178)$$

Thus, by (175), (176)–(178), we get

$$\mathbb{E}X_x = h_1. \quad (179)$$

For $f_2 = x^2$, we have

$$\begin{aligned}
\mathbb{E}X_{x^2} &= \frac{1}{2\pi i} \oint_e c\lambda^2 z^2 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} dz \\
&\quad - \frac{1}{2\pi i} \oint_e z^2 \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times (h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z) dz \\
&\quad - \frac{1}{2\pi i} \oint_e cz^4 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-1} [\beta \underline{m}^2(z) + (h_2 - 2\lambda h_1) \underline{m}^2(z) + 2\lambda^2 \underline{m}'(z)] dz \\
&=: I_1(f_2) + I_2(f_2) + I_3(f_2). \quad (180)
\end{aligned}$$

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

$$\begin{aligned}
I_1(f_2) &= \frac{1}{2\pi i} \oint_e c\lambda^2 z^2 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} dz
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi i} \oint_{\mathcal{C}} c\lambda^2 z^2 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times [\underline{m}^{-2} - c\lambda^2(1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= \frac{1}{2\pi i} \oint_{\mathcal{C}} \frac{-c\lambda^2(-c\lambda \underline{m} + \lambda \underline{m} + 1)^2}{\underline{m}(1 + \lambda \underline{m})^3 [c(\lambda \underline{m})^2 - (1 + \lambda \underline{m})^2]} d\underline{m} \\
&= \lambda^2(1 + 5c + c^2) - \frac{\lambda^2}{2}(1 + \sqrt{c})^4 - \frac{\lambda^2}{2}(1 - \sqrt{c})^4 \\
&= -c\lambda^2.
\end{aligned} \tag{i81}$$

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

$$\begin{aligned}
I_2(f_2) &= -\frac{1}{2\pi i} \oint_{\mathcal{C}} z^2 \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} [h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z] dz \\
&= -\frac{1}{2\pi i} \oint_{\mathcal{C}} z^2 \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times \left(h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z + (h_1 + \lambda) \frac{1-c}{cz} \right) \\
&\quad \times [\underline{m}^{-2} - c\lambda^2(1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= -\frac{1}{2\pi i} \oint_{\mathcal{C}} \frac{-(\lambda^2 \underline{m} - h_1)[-c\lambda \underline{m} + \lambda \underline{m} + 1]^2}{\underline{m}^2(1 + \lambda \underline{m})^4} d\underline{m} \\
&= \lambda(\lambda + 2ch_1 + 2h_1).
\end{aligned} \tag{i82}$$

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

$$\begin{aligned}
I_3(f_2) &= -\frac{1}{2\pi i} \oint_{\mathcal{C}} cz^2 z^2 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times [1 + \lambda \underline{m}(z)]^{-1} [\beta \underline{m}^2(z) + (h_2 - 2\lambda h_1) \underline{m}^2(z) + 2\lambda^2 \underline{m}'(z)] dz \\
&= -\frac{1}{2\pi i} \oint_{\mathcal{C}} cz^2 z^2 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times [1 + \lambda \underline{m}(z)]^{-1} \left[(\beta + h_2 - 2\lambda h_1) \left(\frac{1}{c^2} \underline{m}^2 z^2 + \frac{2(1-c)}{c^2} \underline{m} z + \frac{(1-c)^2}{c^2} \right) \right. \\
&\quad \left. + 2\lambda^2 \left(\frac{1}{c} z^2 \underline{m}' - \frac{1-c}{c} \right) \right] \times [\underline{m}^{-2} - c\lambda^2(1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= -2\lambda^2(1 + 5c + c^2) + c(\beta + h_2 - 2\lambda h_1) + \lambda^2(1 + \sqrt{c})^4 + \lambda^2(1 - \sqrt{c})^4 \\
&= 2c\lambda^2 + c(\beta + h_2 - 2\lambda h_1).
\end{aligned} \tag{i83}$$

Thus, by (i80), (i81)–(i83), we get

$$\begin{aligned}
\mathbb{E}X_{x^2} &= -c\lambda^2 + \lambda(\lambda + 2ch_1 + 2h_1) + 2c\lambda^2 + c(\beta + h_2 - 2\lambda h_1) \\
&= (c + 1)\lambda^2 + 2(c + 1)\lambda h_1 + c(\beta + h_2 - 2\lambda h_1).
\end{aligned}$$

For $f_3 = x^3$, we have

$$\begin{aligned}
\mathbb{E}X_{x^3} &= \frac{1}{2\pi i} \oint_{\mathcal{C}} c\lambda^2 z^3 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} dz \\
&\quad - \frac{1}{2\pi i} \oint_{\mathcal{C}} z^3 \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times [h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z] dz \\
&\quad - \frac{1}{2\pi i} \oint_{\mathcal{C}} cz^5 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-1} [\beta \underline{m}^2(z) + (h_2 - 2\lambda h_1) \underline{m}^2(z) + 2\lambda^2 \underline{m}'(z)] dz \\
&=: I_1(f_3) + I_2(f_3) + I_3(f_3).
\end{aligned} \tag{184}$$

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

$$\begin{aligned}
I_1(f_3) &= \frac{1}{2\pi i} \oint_{\mathcal{C}} c\lambda^2 z^3 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} dz \\
&= \frac{1}{2\pi i} \oint_{\mathcal{C}} c\lambda^2 z^3 \underline{m}^3(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times (1 + \lambda \underline{m}(z))^{-3} [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times [\underline{m}^{-2} - c\lambda^2(1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= \frac{1}{2\pi i} \oint_{\mathcal{C}} \frac{c\lambda^2(-c\lambda \underline{m} + \lambda \underline{m} + 1)^3}{\underline{m}^2(1 + \lambda \underline{m})^4 [c(\lambda \underline{m})^2 - (1 + \lambda \underline{m})^2]} d\underline{m} \\
&= \lambda^3(1 + 12c + 12c^2 + c^3) - \frac{\lambda^3}{2}(1 + \sqrt{c})^6 - \frac{\lambda^3}{2}(1 - \sqrt{c})^6 \\
&= -3c(c + 1)\lambda^3.
\end{aligned} \tag{185}$$

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

$$\begin{aligned}
I_2(f_3) &= -\frac{1}{2\pi i} \oint_{\mathcal{C}} z^3 \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times [h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z] dz \\
&= -\frac{1}{2\pi i} \oint_{\mathcal{C}} z^3 \underline{m}(z) [1 - c\lambda^2 \underline{m}^2(z)(1 + \lambda \underline{m}(z))^{-2}]^{-1} \\
&\quad \times z \underline{m}(z) (1 + \lambda \underline{m}(z))^{-1} \times \left(h_1 \underline{m}(z) + \lambda \underline{m}(z) + \lambda/z + (h_1 + \lambda) \frac{1 - c}{cz} \right) \\
&\quad \times [\underline{m}^{-2} - c\lambda^2(1 + \lambda \underline{m})^{-2}] d\underline{m} \\
&= -\frac{1}{2\pi i} \oint_{\mathcal{C}} \frac{[\lambda^2 \underline{m} - h_1][-c\lambda \underline{m} + \lambda \underline{m} + 1]^3}{\underline{m}^3(1 + \frac{\sigma^2}{\mu^3} \underline{m})^5} d\underline{m} \\
&= \lambda^2(3c\lambda + 2\lambda + 3c^2 h_1 + 9c h_1 + 3h_1).
\end{aligned} \tag{186}$$

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

$$\begin{aligned}
I_3(f_3) &= -\frac{1}{2\pi i} \oint_{\mathcal{C}} cz^3 z^2 \underline{\mathbf{m}}^3(z) [1 - c\lambda^2 \underline{\mathbf{m}}^2(z)(1 + \lambda \underline{\mathbf{m}}(z))^{-2}]^{-1} \\
&\quad \times [1 + \lambda \underline{\mathbf{m}}(z)]^{-1} [\beta \mathbf{m}^2(z) + (\mathbf{h}_2 - 2\lambda \mathbf{h}_1) \mathbf{m}^2(z) + 2\lambda^2 \mathbf{m}'(z)] dz \\
&= -\frac{1}{2\pi i} \oint_{\mathcal{C}} cz^3 z^2 \underline{\mathbf{m}}^3(z) [1 - c\lambda^2 \underline{\mathbf{m}}^2(z)(1 + \lambda \underline{\mathbf{m}}(z))^{-2}]^{-1} \\
&\quad \times [1 + \lambda \underline{\mathbf{m}}(z)]^{-1} \left[(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1) \left(\frac{1}{c^2} \underline{\mathbf{m}}^2 z^2 + \frac{2(1-c)}{c^2} \underline{\mathbf{m}} z + \frac{(1-c)^2}{c^2} \right) \right. \\
&\quad \left. + 2\lambda^2 \left(\frac{1}{c} z^2 \underline{\mathbf{m}}' - \frac{1-c}{c} \right) \right] \times \left[\underline{\mathbf{m}}^{-2} - c\lambda^2 (1 + \lambda \underline{\mathbf{m}})^{-2} \right] d\underline{\mathbf{m}} \\
&= -2\lambda^3 (1 + 12c + 12c^2 + c^3) + 3c(c+1)\lambda(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1) \\
&\quad + \lambda^3 (1 + \sqrt{c})^6 + \lambda^3 (1 - \sqrt{c})^6 \\
&= 3c(c+1)\lambda(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1) + 6c(c+1)\lambda^3. \tag{187}
\end{aligned}$$

Thus, by (184), (185)–(187), we get

$$\begin{aligned}
\mathbb{E}X_{x^3} &= -3c(c+1)\lambda^3 + \lambda^2(3c\lambda + 2\lambda + 3c^2\mathbf{h}_1 + 9c\mathbf{h}_1 + 3\mathbf{h}_1) \\
&\quad + 3c(c+1)\lambda(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1) + 6c(c+1)\lambda^3 \\
&= (3c^2 + 6c + 2)\lambda^3 + 3(c^2 + 3c + 1)\lambda^2\mathbf{h}_1 + 3c(c+1)\lambda(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1).
\end{aligned}$$

Thus,

$$\begin{aligned}
\mu_1 &:= \mathbb{E}X_{f_1} = \mathbf{h}_1, \\
\mu_2 &:= \mathbb{E}X_{f_2} = (1+c)\lambda^2 + 2(1+c)\lambda\mathbf{h}_1 + c(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1), \\
\mu_3 &:= \mathbb{E}X_{f_3} = (2 + 6c + 3c^2)\lambda^3 + 3(1 + 3c + c^2)\lambda^2\mathbf{h}_1 + 3c(1+c)\lambda(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1),
\end{aligned}$$

and

$$\begin{aligned}
V_{(1,1)} &= 2c\lambda^2 + c(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1), \\
V_{(2,2)} &= 4c(2+c)(1+2c)\lambda^4 + 4c(1+c)^2\lambda^2(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1), \\
V_{(3,3)} &= 6c(1+6c+3c^2)(3+6c+c^2)\lambda^6 + 9c(1+3c+c^2)^2\lambda^4(\beta + \mathbf{h}_2 - 2\lambda \mathbf{h}_1).
\end{aligned}$$

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

The tightness of $M_p^{(1)}(z)$ is similar to that provided in [Bai and Silverstein \[2004\]](#). It is sufficient to prove the moment condition (12.51) of [Billingsley \[1968\]](#), i.e.

$$\sup_{n; z_1, z_2 \in \mathcal{C}_n} \frac{\mathbb{E}|M_p^{(1)}(z_1) - M_p^{(1)}(z_2)|^2}{|z_1 - z_2|^2} \tag{188}$$

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 \mathbf{p} and $z \in \mathcal{C}_p$. 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 2.4](#) that

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

$$\leq K_1 + K_2 n^m \varepsilon^{-m} n^{-\ell} \leq K$$

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

$$\max \left(\mathbb{E} \|\mathbf{D}^{-1}(z)\|^m, \mathbb{E} \|\mathbf{D}_j^{-1}(z)\|^m, \mathbb{E} \|\mathbf{D}_{ij}^{-1}(z)\|^m \right) \leq K. \quad (189)$$

By the argument above, we can extend Lemma 4.4 and get

$$\left| \mathbb{E} \left(\alpha(v) \prod_{\ell=1}^q (\mathbf{r}'_1 \mathbf{B}_{(\ell)}(v) \mathbf{r}_1 - n^{-1} \text{tr} \mathbf{B}_{(\ell)}(v)) \right) \right| \leq K n^{-1} \delta_n^{2q-4}, \quad (190)$$

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

$$\max(|\alpha(v)|, \|\mathbf{B}_{(\ell)}(v)\|) \leq K(1 + n^s \mathbb{I}(\|\mathbf{B}_p\| \geq \eta_r \text{ or } \lambda_{\min}(\tilde{\mathbf{B}}) \leq \eta_\ell)),$$

with $\tilde{\mathbf{B}}$ being $\mathbf{B}_{(j)}$ or \mathbf{B}_p . By (190), we have

$$\mathbb{E} |\varepsilon_j(z)|^m \leq K_m n^{-1} \delta_n^{2m-4}. \quad (191)$$

Let $\gamma_j(z) = \mathbf{r}'_j \mathbf{D}_j^{-1}(z) \mathbf{r}_j - n^{-1} v_2 \mathbb{E} \text{tr} \mathbf{D}_j^{-1}(z)$. By Lemma A.2, (190) and Hölder's inequality, with similar derivation on page 580 of Bai and Silverstein [2004], we have

$$\mathbb{E} |\gamma_j(z) - \varepsilon_j(z)|^m \leq \frac{K_m}{n^{m/2}}. \quad (192)$$

It follows from (191) and (192) that

$$\mathbb{E} |\gamma_j|^m \leq K_m n^{-1} \delta_n^{2m-4}, \quad m \geq 2. \quad (193)$$

Next, we prove that $b_p(z)$ is bounded. With (190), we have for any $m \geq 1$,

$$\mathbb{E} |\beta_1(z)|^m \leq K_m. \quad (194)$$

Since $b_p(z) = \beta_1(z) + \beta_1(z) b_p(z) \gamma_1(z)$, it is derived from (193) and (194) that $|b_p(z)| \leq K_1 + K_2 |b_p(z)| n^{-1/2}$.

Then

$$|b_p(z)| \leq \frac{K_1}{1 - K_2 n^{-1/2}} \leq K. \quad (195)$$

With (190)–(195) and the same approach on page 581–583 of Bai and Silverstein [2004], we can obtain that (188) is finite.

C Simulation of CLT for $M_p(z)$

In this section, we compare the empirical mean and covariance of $M_p(z) = \text{tr}(\mathbf{B}_p^0 - z\mathbf{I}_p)^{-1} - p\mathbf{m}_{\text{Fc}n}(z)$ with their theoretical limits as stated in Proposition 4.1. This proposition is a key step for the proof of our main result, Theorem 2.5. Readers are referred to Section 4 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 $\text{Cov}(M_p(z_1), M_p(z_2))$ 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 3 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 $\text{Cov}(M_p(z_1), M_p(z_2))$ are reported in Tanle 4. As shown in Tables 3 – 4, the empirical values of $\mathbb{E}M_p(z)$ and $\text{Cov}(M_p(z_1), M_p(z_2))$ closely match their respective theoretical limits under all scenarios.

Table 3: 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	

References

- Z. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *The Annals of Probability*, 26(1), 1998. doi: [10.1214/aop/1022855421](https://doi.org/10.1214/aop/1022855421). (cited on page 44)
- Z. Bai and J. 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](https://doi.org/10.1214/aop/1078415845). (cited on pages 3, 4, 12, 13, 22, 23, 29, 35, 46, 53, 54, 57, 62, and 63)

Table 4: Empirical covariance between $M_p(z_1)$ and $M_p(z_2)$.

p/n	n	Exp(5)		$\chi^2(1)$		
		$(z_1, z_2) = (-3+2i, -1+i)$	$(3+2i, 5+i)$	$(-3+2i, -1+i)$	$(3+2i, 5+i)$	
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	

- Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer, 2nd edition, 2010. (cited on page 26)
- Z. Bai and Y. 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](https://doi.org/10.1214/aop/1176989118). (cited on page 27)
- Z. Bai, B. Miao, and G. Pan. On asymptotics of eigenvectors of large sample covariance matrix. *The Annals of Probability*, 35(4), 2007. doi: [10.1214/009117906000001079](https://doi.org/10.1214/009117906000001079). (cited on page 56)
- Z. Bai, H. Li, and G. Pan. Central limit theorem for linear spectral statistics of large dimensional separable sample covariance matrices. *Bernoulli*, 25(3), 2019. doi: [10.3150/18-bej1038](https://doi.org/10.3150/18-bej1038). (cited on page 3)
- Z. Bao. Tracy-widom limit for kendall’s tau. *The Annals of Statistics*, 47(6):3504–3532, 2019. doi: [10.1214/18-aos1786](https://doi.org/10.1214/18-aos1786). (cited on page 3)
- F. Benaych-Georges and R. R. 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](https://doi.org/10.1016/j.jmva.2012.04.019). (cited on page 27)
- P. Billingsley. *Convergence of probability measures*. New York: Wiley, 1968. (cited on pages 14, 23, and 62)
- P. Billingsley. *Probability and Measure*. New York: Wiley, 1995. (cited on page 26)
- D. L. Burkholder. Distribution function inequalities for martingales. *The Annals of Probability*, 1(1): 19–42, 1973. (cited on page 26)
- T. Cai, W. Liu, and Y. Xia. Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76, 2014. doi: [10.1111/rssb.12034](https://doi.org/10.1111/rssb.12034). (cited on page 2)

- Y. Cao, W. Lin, and H. Li. Two-sample tests of high-dimensional means for compositional data. *Biometrika*, 105:115–132, 2018. doi: [10.1093/biomet/asx060](https://doi.org/10.1093/biomet/asx060). (cited on page 2)
- B. Chen and G. Pan. Convergence of the largest eigenvalue of normalized sample covariance matrices when p and n both tend to infinity with their ratio converging to zero. *Bernoulli*, 18(4):1405–1420, 2012. doi: [10.3150/11-bej381](https://doi.org/10.3150/11-bej381). (cited on page 29)
- N. El Karoui. Tracy–widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *The Annals of Probability*, 35(2), 2007. doi: [10.1214/009117906000000917](https://doi.org/10.1214/009117906000000917). (cited on page 3)
- K. Faust, J. F. Sathirapongsasuti, J. Izard, N. Segata, D. Gevers, J. Raes, and C. Huttenhower. Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology*, 8(7):e1002606, 2012. doi: [10.1371/journal.pcbi.1002606](https://doi.org/10.1371/journal.pcbi.1002606). (cited on page 2)
- J. Gao, X. Han, G. Pan, and Y. 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](https://doi.org/10.1111/rssb.12189). (cited on pages 3, 4, 8, 12, and 22)
- R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 2012. (cited on page 25)
- T. Jiang. The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā: The Indian Journal of Statistics*, 66(1):35–48, 2004. (cited on pages 4, 12, and 27)
- I. M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *The Annals of Statistics*, 29(2), 2001. doi: [10.1214/aos/1009210544](https://doi.org/10.1214/aos/1009210544). (cited on page 3)
- D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *Journal of Multivariate Analysis*, 12(1):1–38, Mar. 1982. doi: [10.1016/0047-259x\(82\)90080-x](https://doi.org/10.1016/0047-259x(82)90080-x). (cited on page 3)
- 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](https://doi.org/10.1070/sm1967v001n04abeh001994). (cited on page 3)
- J. 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](https://doi.org/10.1006/jmva.1995.1083). (cited on pages 23 and 47)
- J. Silverstein and Z. 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](https://doi.org/10.1006/jmva.1995.1051). (cited on page 3)
- G. D. Wu, J. Chen, C. Hoffmann, K. Bittinger, Y.-Y. Chen, S. A. Keilbaugh, M. Bewtra, D. Knights, W. A. Walters, R. Knight, R. Sinha, E. Gilroy, K. Gupta, R. N. Baldassano, L. C. Nessel, H. Li, F. D. Bushman, and J. D. Lewis. Linking long-term dietary patterns with gut microbial enterotypes. *Science*, 334:105 – 108, 2011. doi: [10.1126/science.1208344](https://doi.org/10.1126/science.1208344). (cited on page 3)
- H. Xiao and W. 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](https://doi.org/10.1007/s10959-009-0270-2). (cited on page 27)

- Y. Yin and P. 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](https://doi.org/10.1016/0047-259x(83)90035-0). (cited on page 3)
- L. Zhang. Spectral analysis of large dimensional random matrices. 2007. (cited on page 3)
- S. Zheng, Z. Bai, and J. 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](https://doi.org/10.1214/14-AOS1292). (cited on pages 12, 29, and 30)