Electronic Journal of Statistics
Vol. 15 (2021) 135–210
ISSN: 1935-7524
https://doi.org/10.1214/20-EJS1783

Two-sample tests for high-dimensional covariance matrices using both difference and ratio

Tingting Zou

Northeast Normal University, Changchun, China
e-mail: zoutt260@nenu.edu.cn

Ruitao Lin[†]

The University of Texas MD Anderson Cancer Center, Houston, Texas, U.S.A.
e-mail: rlin@mdanderson.org

Shurong Zheng[*]

Northeast Normal University and KLAS, Changchun, China
e-mail: zhengsr@nenu.edu.cn

Guo-Liang Tian[†]

Southern University of Science and Technology, Shenzhen, China
e-mail: tiangl@sustech.edu.cn

Abstract: By borrowing strengths from the difference and ratio between two sample covariance matrices, we propose three tests for testing the equality of two high-dimensional population covariance matrices. One test is shown to be powerful against dense alternatives, and the other two tests are suitable for general cases, including dense and sparse alternatives, or the mixture of the two. Based on random matrix theory, we investigate the asymptotical properties of these three tests under the null hypothesis as the sample size and the dimension tend to infinity proportionally. Limiting behaviors of the new tests are also studied under various local alternatives. Extensive simulation studies demonstrate that the proposed methods outperform or perform equally well compared with the existing tests.

MSC2020 subject classifications: Primary 62H15, 62H10; secondary 60E05.

Keywords and phrases: Asymptotic normality, high-dimensional covariance matrices, power enhancement, random matrix theory.

Received April 2019.

[*]Corresponding author: Shurong Zheng (e-mail: zhengsr@nenu.edu.cn). Shurong Zheng is supported by NSFC 12071066 and 11690012.

[†]Ruitao Lin's research was partially supported by NIH/NCI grant P30 CA 016672. Guo-Liang Tian's research was fully supported by a grant from the National Natural Science Foundation of China (No. 11771199).

Contents

1 Introduction . . . 136
2 Main results . . . 137
  2.1 Joint limiting null distribution of T_{d} and T_{r} . . . 138
  2.2 Three test procedures . . . 141
3 Power analysis . . . 143
  3.1 Power analysis in the dense alternative set Π_1 . . . 144
  3.2 Power analysis in the dense alternative set Π_2 . . . 145
  3.3 Power analysis in the alternative set Π_3 . . . 147
4 Simulation studies . . . 147
Appendix . . . 150
  A.1 Review of Conditions (C1), (C2) (or (C2*)) and (C3) of [5] . . . 151
  A.2 Verification of the considered simulation scenarios . . . 154
  A.3 Proofs of some theorems, propositions and lemmas . . . 157
    A.3.1 Preparatory works for proving Theorems 2.1, 3.1 and 3.2 . . . 157
    A.3.2 Proofs of the theorems and propositions . . . 173
    A.3.3 Technical lemmas . . . 185
Acknowledgements . . . 209
References . . . 209

1. Introduction

Testing whether two populations share the same covariance matrix is of fundamental interest in statistical inference, as many statistical procedures, such as the linear discriminant analysis, require the task of the equality test of covariance matrices. Let $\mathbf{x}_{k1}, \mathbf{x}_{k2}, \dots, \mathbf{x}_{kN_k}$ be an independent and identically distributed sample from the kth p-dimensional population with mean vector $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$, where N_k is the sample size, $k = 1, 2$. Let $n_k = N_k - 1$ and $y_{n_k} = p/n_k$. The sample covariance matrix of $\boldsymbol{\Sigma}_k$ is given by $\mathbf{S}_k = n_k^{-1} \sum_{i=1}^{N_k} (\mathbf{x}_{ki} - \bar{\mathbf{x}}_k)(\mathbf{x}_{ki} - \bar{\mathbf{x}}_k)^{\mathrm{T}}$, where $\bar{\mathbf{x}}_k = N_k^{-1} \sum_{i=1}^{N_k} \mathbf{x}_{ki}$ is the sample mean of the kth population, $k = 1, 2$. We are interested in testing the following hypothesis

$$H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 \quad \text{versus} \quad H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2. \tag{1.1}$$

In the conventional low-dimensional setting where p is relatively small compared with the sample sizes, such a testing problem has been well studied. See, for example, [16], [8], [13], [7], [12] and [1]. In the high-dimensional setting where p is large relative to N_1 and N_2, several different approaches have been proposed to address the failure of classical methods, see [4, 14, 15, 10, 5, 19, 21], and among others. In particular, many of the existing testing procedures are developed based on $\mathrm{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$, the square of the Frobenius norm of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. For example, under the Gaussian population assumption, [14] developed an unbiased estimator by correcting the bias of the sample statistic $T_{\mathrm{d}} = \mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2$ in estimating $\mathrm{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$. [11] considered a U-statistic for $\mathrm{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2$ and showed that their proposed test can yield desirable performance in various situations, especially when the two population matrices have dense differences. [21] also studied the limiting behavior of T_{d}, under general populations including both Gaussian and non-Gaussian distributions. On the other hand, known as a relative measure, the ratio measure, as given by $\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}-\mathbf{I}_p)^2$ with $\mathbf{I}_p$ denoting the identity matrix, can also characterize the relative discrepancy between $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$. When $\mathbf{S}_2$ is invertible, a natural estimator of $\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}-\mathbf{I}_p)^2$ is $T_{\mathrm{r}} = \mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}-\mathbf{I}_p)^2$. Unfortunately, such a test statistic has been seldom studied. This may be due to the reason that the statistic T_{r} involves the inverse of $\mathbf{S}_2$, which requires the dimension p does not exceed the sample size N_2; and when p is close to N_2, the behavior of T_{r} is complicated. In fact, the relative performance between the tests T_{d} and T_{r} varies case by case: as will be shown in the simulation studies as well as through the theoretical power analyses, any one of T_{d} and T_{r} cannot dominate the other in terms of statistical power under all scenarios.

This paper makes several contributions to the literature. (1) Firstly, we investigate the relative performances of T_{d} and T_{r} through theoretical power analyses in the high-dimensional setting. In particular, the statistic T_{r} is not well studied in literature. To adaptively borrow strengths from T_{d} and T_{r}, we develop a new testing procedure through establishing the joint limiting null distribution of T_{d} and T_{r} using random matrix theory. The proposed method is able to perform nearly the same as the better one of T_{d} and T_{r} under all scenarios. (2) Since T_{d} and T_{r} only target on dense alternatives where there are many small differences between $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, they may be not powerful for sparse alternatives where there are only few but large differences between $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$. Our second contribution is to propose another two testing procedures, by combining T_{d}, T_{r}, and a maximum norm statistic, to maintain high power for testing the equality of two high-dimensional covariance matrices under both dense and sparse alternatives. To the best of our knowledge, there is no existing approach to two-sample testing problems with high-dimensional covariance matrices using the combination of three statistics. (3) The theoretical properties of the proposed tests based on different combination procedures are extensively investigated in the paper, which sheds lights on the pros and cons of different procedures for combining multiple statistics.

The rest of this paper is organized as follows. Section 2 establishes the joint limiting null distribution of T_{d} and T_{r}, and introduces three novel procedures for testing the equality of two covariance matrices. Section 3 investigates the power functions of the proposed tests under representative alternatives. Section 4 presents various simulation results to demonstrate the performance of the proposed methods. All technical details are presented in the Appendix.

2. Main results

Before presenting the main results, we first introduce some basic notations. The empirical spectral distribution (*ESD*) is defined as

$$F^{\mathbf{A}} = p^{-1}\sum_{j=1}^{p}\delta_{\lambda_j},$$

where $\mathbf{A}$ is a $p \times p$ non-negative definite matrix, $\{\lambda_j, 1 \le j \le p\}$ are the eigenvalues of $\mathbf{A}$ and δ_a denotes the Dirac mass at point a. Define the operator as

$$\mathbf{A} \circ \mathbf{B} = (a_{ij} b_{ij})_{i=1,\ldots,m, j=1,\ldots,n}$$

with $m \times n$ matrices $\mathbf{A} = (a_{ij})_{i=1,\ldots,m,j=1,\ldots,n}$ and $\mathbf{B} = (b_{ij})_{i=1,\ldots,m,j=1,\ldots,n}$. The sample covariance matrices are given by

$$\mathbf{S}_1 = n_1^{-1} \sum_{i=1}^{N_1} (\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)^{\mathrm{T}}, \ \mathbf{S}_2 = n_2^{-1} \sum_{i=1}^{N_2} (\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)^{\mathrm{T}},$$

where $\bar{\mathbf{x}}_1 = N_1^{-1} \sum_{i=1}^{N_1} \mathbf{x}_{1i}$ and $\bar{\mathbf{x}}_2 = N_2^{-1} \sum_{i=1}^{N_2} \mathbf{x}_{2i}$ are the sample means. Let the aforementioned two statistics be

$$T_{\mathrm{d}} = \mathrm{tr}(\mathbf{S}_1 - \mathbf{S}_2)^2, \quad T_{\mathrm{r}} = \mathrm{tr}(\mathbf{S}_1 \mathbf{S}_2^{-1} - \mathbf{I}_p)^2,$$

where $\mathbf{I}_p$ is the $p \times p$ identity matrix.

Next, we impose the following two assumptions, which are commonly used in random matrix theory, for studying the limiting behaviors of the considered statistics throughout the paper.

- **Assumption A**. The random vector $\mathbf{x}_{ki}$ satisfies the independent component structure $\mathbf{x}_{ki} = \boldsymbol{\mu}_k + \boldsymbol{\Sigma}_k^{1/2} \mathbf{w}_{ki}$, where $\mathbf{w}_{ki} = (w_{k1i}, \cdots, w_{kpi})^{\mathrm{T}}$, the elements $\{w_{kli}, k = 1, 2; l = 1, \cdots, p; i = 1, \cdots, N_k\}$ are independent and identically distributed with $\mathrm{E}w_{kli} = 0, \mathrm{E}w_{kli}^2 = 1$ and $\beta_k = \mathrm{E}w_{kli}^4 - 3$. Moreover, $\boldsymbol{\Sigma}_k > 0$ has a bounded spectral norm. The *ESD* of $\boldsymbol{\Sigma}_k$ converges weakly to a limit spectral distribution (*LSD*) L_k for $k = 1, 2$.
- **Assumption B**. The dimension p and n_1, n_2 grow to infinity under the convergence regime

  $$y_{n_1} = p/n_1 \to y_1 \in (0, +\infty), \quad y_{n_2} = p/n_2 \to y_2 \in (0, 1),$$

  where $n_1 = N_1 - 1$ and $n_2 = N_2 - 1$.

In Assumption B, we require the data dimension p to be less than n_2 so that the sample covariance matrix of the second population is invertible.

2.1. *Joint limiting null distribution of T_{d} and T_{r}*

The following theorem establishes the joint limiting null distribution of the two statistics T_{d} and T_{r}. Let $L(x)$ be the *LSD* of $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$ under the null hypothesis in (1.1).

Theorem 2.1. *Under Assumptions A-B and the null hypothesis H_0 in (1.1), let $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$, then*

$$\begin{pmatrix} T_{\mathrm{d}} - \mu_0 - \mu_{10} \\ T_{\mathrm{r}} - \mu_{20} \end{pmatrix} \xrightarrow{d} N \left(\mathbf{0}_2, \begin{pmatrix} \sigma_{110} & \sigma_{120} \\ \sigma_{210} & \sigma_{220} \end{pmatrix} \right),$$

where $\mathbf{0}_2 = (0, 0)^{\mathrm{T}}$,

$$\begin{aligned}
\mu_0 &= n_1^{-1}\mathrm{tr}^2\mathbf{S}_1 + n_2^{-1}\mathrm{tr}^2\mathbf{S}_2,\ m_{10} = \int x dL(x),\ m_{20} = \int x^2 dL(x),\\
\mu_{10} &= (y_{n_1} + y_{n_2})p^{-1}\mathrm{tr}\mathbf{\Sigma}^2 + (\beta_1 y_{n_1} + \beta_2 y_{n_2})p^{-1}\mathrm{tr}(\mathbf{\Sigma}\circ\mathbf{\Sigma}),\\
\mu_{20} &= p\Big[-\frac{2}{1-y_{n_2}} + \frac{y_{n_1}}{(1-y_{n_2})^2} + \frac{1}{(1-y_{n_2})^3}\Big] + p\\
&\quad -\frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{2y_{n_1}y_{n_2} + y_{n_1}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} - \frac{2\beta_2 y_{n_2}}{1-y_{n_2}}\\
&\quad +\frac{2\beta_2 y_{n_1}y_{n_2} + \beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1-y_{n_2})^2} + \frac{2\beta_2 y_{n_2}}{(1-y_{n_2})^3},\\
\sigma_{110} &= 4(y_1+y_2)^2 m_{20}^2,\\
\sigma_{120} &= \sigma_{210} = \Big[\frac{8y_2(y_1+y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3}\Big]m_{10}^2,\\
\sigma_{220} &= \frac{8y_1^3 + 16y_1^2y_2}{(1-y_2)^5} + \frac{4y_1^2 + 40y_1^2y_2 + 64y_1y_2^2}{(1-y_2)^6}\\
&\quad +\frac{8y_1y_2^4 + 56y_1y_2^2 + 48y_2^3 + 8y_1y_2}{(1-y_2)^7} + \frac{8y_2^5 + 24y_2^3 + 4y_2^2}{(1-y_2)^8}\\
&\quad +4(\beta_1y_1 + \beta_2y_2)\Big[\frac{(y_1+y_2)^2}{(1-y_2)^4} + \frac{2y_2(y_1+y_2)}{(1-y_2)^5} + \frac{y_2^2}{(1-y_2)^6}\Big].
\end{aligned}$$

Remark 2.1. *From Hölder's inequality,* $m_{20} \geq m_{10}^2$ *and* $\beta_k \geq -2$ *for* $k = 1, 2$. *Thus,*

$$\begin{aligned}
&\sigma_{110}\sigma_{220} - \sigma_{120}\sigma_{210}\\
\geq\ & 16\Big[\frac{(4y_2^3 + 2y_1y_2)(y_1+y_2)^4 + 4y_1^2y_2^2(y_1+y_2)^2 + y_1^4y_2}{(1-y_2)^5}\\
&+\frac{(2y_2^3 + 4y_2^2)(y_1+y_2)^4 + 4y_1y_2^2(y_1+y_2)^3 + 2y_1^2y_2^4}{(1-y_2)^6}\\
&+\frac{(4y_1y_2^3 + 4y_1y_2^2)(y_1+y_2)^2 + y_1^3y_2(y_1+y_2) + 3y_1^3y_2^2}{(1-y_2)^6}\\
&+\frac{14y_2^3(y_1+y_2)^3 + (8y_2^5 + 4y_1y_2^2)(y_1+y_2)^2}{(1-y_2)^7}\\
&+\frac{(y_2^5 + 2y_1y_2^3)(y_1+y_2) + y_2^7 + 3y_1y_2^5 + y_2^5 + 2y_1y_2^3}{(1-y_2)^7}\\
&+\frac{(4y_2^7 + 4y_2^3)(y_1+y_2)^2 + y_2^4(y_1+y_2) + 2y_1y_2^4}{(1-y_2)^8}\Big]m_{10}^4 > 0.
\end{aligned}$$

Therefore, the covariance matrix $\begin{pmatrix} \sigma_{110} & \sigma_{120} \\ \sigma_{210} & \sigma_{220} \end{pmatrix}$ *is positive definite.*

Theorem 2.1 shows that the joint limiting distribution of T_{d} and T_{r} is a bivariate normal distribution under the null hypothesis. Its proof is provided in Appendix. The marginal limiting distribution of $T_{\rm d}$ is consistent with that given in the Proposition 1 of [21]. Due to $\sigma_{120} > 0$, $T_{\rm d}$ and $T_{\rm r}$ are asymptotically positively-correlated.

Because μ_{10}, σ_{110}, σ_{120} and σ_{210} involve the functionals of the unknown population covariance matrix $\mathbf{\Sigma}$, we have to estimate these quantities. Let

$$\mathbf{S} = (n_1+n_2)^{-1}(n_1\mathbf{S}_1 + n_2\mathbf{S}_2),\ \hat{h}_{10} = p^{-1}\mathrm{tr}(\mathbf{S}\circ\mathbf{S}),$$
$$\hat{m}_{10} = p^{-1}\mathrm{tr}\mathbf{S},\ \hat{m}_{20} = p^{-1}[\mathrm{tr}\mathbf{S}^2 - (n_1+n_2)^{-1}\mathrm{tr}^2\mathbf{S}].$$

Theorem 2.2. *Under the conditions of Theorem 2.1, the weakly consistent estimators of μ_{10}, σ_{110}, σ_{120}, σ_{210} and σ_{220} are as follows:*

$$\begin{aligned}
\hat{\mu}_{10} &= (y_{n_1}+y_{n_2})\hat{m}_{20} + (\beta_1 y_{n_1} + \beta_2 y_{n_2})\hat{h}_{10},\\
\hat{\sigma}_{110} &= 4(y_{n_1}+y_{n_2})^2\hat{m}_{20}^2,\\
\hat{\sigma}_{120} &= \hat{\sigma}_{210} = \Big[\frac{8y_{n_2}(y_{n_1}+y_{n_2})^2 + 4y_{n_1}^2 + 4y_{n_2}^2}{(1-y_{n_2})^2} + \frac{8y_{n_1}y_{n_2} + 8y_{n_2}^3}{(1-y_{n_2})^3}\Big]\hat{m}_{10}^2,\\
\hat{\sigma}_{220} &= \frac{8y_{n_1}^3 + 16y_{n_1}^2 y_{n_2}}{(1-y_{n_2})^5} + \frac{4y_{n_1}^2 + 40y_{n_1}^2 y_{n_2} + 64y_{n_1}y_{n_2}^2}{(1-y_{n_2})^6}\\
&\quad + \frac{8y_{n_1}y_{n_2}^4 + 56y_{n_1}y_{n_2}^2 + 48y_{n_2}^3 + 8y_{n_1}y_{n_2}}{(1-y_{n_2})^7} + \frac{8y_{n_2}^5 + 24y_{n_2}^3 + 4y_{n_2}^2}{(1-y_{n_2})^8}\\
&\quad + 4(\beta_1 y_{n_1} + \beta_2 y_{n_2})\Big[\frac{(y_{n_1}+y_{n_2})^2}{(1-y_{n_2})^4} + \frac{2y_{n_2}(y_{n_1}+y_{n_2})}{(1-y_{n_2})^5} + \frac{y_{n_2}^2}{(1-y_{n_2})^6}\Big]\\
&\quad + 8y_{n_1}\Big[\frac{2(y_{n_1}+y_{n_2}-y_{n_1}y_{n_2})(y_{n_1}+3y_{n_2}-y_{n_1}y_{n_2}-y_{n_2}^2)}{(1-y_{n_2})^6}\\
&\qquad + \frac{(\beta_1 y_{n_1}+\beta_2 y_{n_2})(y_{n_1}+2y_{n_2}-y_{n_1}y_{n_2}-y_{n_2}^2)}{(1-y_{n_2})^4}\Big]v_{10p}\\
&\quad + 8y_{n_1}\Big[\frac{(y_{n_1}+y_{n_2}-1)^2}{(1-y_{n_2})^2} + \frac{y_{n_1}}{(1-y_{n_2})^3}\Big]v_{20p}\\
&\quad + \frac{16y_{n_1}(y_{n_1}+y_{n_2}-1)}{(1-y_{n_2})}v_{30p} + 8y_{n_1}v_{40p}\\
&\quad + 4y_{n_1}^2\Big[\frac{2(y_{n_1}+y_{n_2}-y_{n_1}y_{n_2})}{(1-y_{n_2})^4} + \frac{\beta_1 y_{n_1}+\beta_2 y_{n_2}}{(1-y_{n_2})^2}\Big]v_{10p}^2\\
&\quad + \frac{16y_{n_1}^2(y_{n_1}+y_{n_2}-1)}{(1-y_{n_2})}v_{10p}v_{20p} + 16y_{n_1}^2 v_{10p}v_{30p}\\
&\quad + 4y_{n_1}^2 v_{20p}^2 + 8y_{n_1}^3 v_{10p}^2 v_{20p},
\end{aligned}$$

where

$$\begin{aligned}
v_{10p} &= \frac{1}{p}\Big[\frac{y_{n_2}}{(1-y_{n_2})^2} + \frac{\beta_2 y_{n_2}}{(1-y_{n_2})}\Big],\\
v_{20p} &= \frac{1}{p}\Big[\frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} - \frac{\beta_2(y_{n_2}^2 - 3y_{n_2})}{(1-y_{n_2})^3}\Big],
\end{aligned}$$

$$v_{30p} = \frac{1}{p}\Big[\frac{y_{n_2}^3+9y_{n_2}^2+6y_{n_2}}{(1-y_{n_2})^6}+\frac{6\beta_2 y_{n_2}}{(1-y_{n_2})^5}\Big],$$
$$v_{40p} = \frac{1}{p}\Big[\frac{y_{n_2}^4+18y_{n_2}^3+35y_{n_2}^2+10y_{n_2}}{(1-y_{n_2})^8}+\frac{10\beta_2(y_{n_2}^2+y_{n_2})}{(1-y_{n_2})^7}\Big].$$

Based on the Slutsky's theorem, we further have that the random vector

$$\left(\frac{T_{\mathrm{d}}-\mu_0-\hat{\mu}_{10}}{\sqrt{\hat{\sigma}_{110}}}, \frac{T_{\mathrm{r}}-\mu_{20}}{\sqrt{\hat{\sigma}_{220}}}\right)^{\mathrm{T}}$$

converges in distribution to $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$ with $\rho = \sigma_{120}/\sqrt{\sigma_{110}\sigma_{220}}$.

Remark 2.2. *Instead of simply replacing y_1 and y_2 in σ_{220} respectively with y_{n_1} and y_{n_2}, the expression of $\hat{\sigma}_{220}$ has additional terms involving v_{i0p}, $i = 1,2,3,4$. This is because even if the order of v_{i0p} is $o(1)$, the value of v_{i0p} is much greater than 0 when p is small. For more details, one can refer to the proof of Theorem 2.2 in Appendix.*

2.2. Three test procedures

For testing the hypothesis (1.1), three test procedures are proposed as follows:

Test 1: Let the first statistic be

$$T_{\mathrm{dr}} = \max\{|T_{\mathrm{d}}-\mu_0-\hat{\mu}_{10}|/\sqrt{\hat{\sigma}_{110}}, |T_{\mathrm{r}}-\mu_{20}|/\sqrt{\hat{\sigma}_{220}}\}. \tag{2.1}$$

That is, T_{dr} is constructed by the maximum absolute value of the standardized statistics T_{d} and T_{r}. For a given test level α, the rejection region is

$$\{\mathbf{x}_{11}, ...\mathbf{x}_{1N_1}, \mathbf{x}_{21}, ..., \mathbf{x}_{2N_2} : T_{\mathrm{dr}} > t_\alpha\}, \tag{2.2}$$

where the critical value t_α is obtained by

$$\alpha = 1 - \int_{-t_\alpha}^{t_\alpha}\int_{-t_\alpha}^{t_\alpha} f(x_{\mathrm{d}}, x_{\mathrm{r}})dx_{\mathrm{d}}dx_{\mathrm{r}}, \tag{2.3}$$

with $f(x_{\mathrm{d}}, x_{\mathrm{r}})$ being the density of $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \hat{\rho} \\ \hat{\rho} & 1 \end{pmatrix}\right)$ and $\hat{\rho} = \hat{\sigma}_{120}/\sqrt{\hat{\sigma}_{110}\hat{\sigma}_{220}}$.

Test 2: T_{d} and T_{r} are powerful to measure the dense differences between $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, so is T_{dr}. To enhance the power of the test procedure (2.2) for sparse alternatives, we use a theoretical result of [5]. Under Conditions (C1), (C2) (or (C2*)) and (C3) of [5] (see Appendix) and the null hypothesis H_0, for any $t \in \mathbb{R}$,

$$P(T_{\mathrm{x}} - 4\log p + \log\log p \le t) \to \exp\Big(-\frac{1}{\sqrt{8\pi}}\exp\Big(-\frac{t}{2}\Big)\Big), \tag{2.4}$$

as $N_1, N_2, p \to \infty$, where

$$T_{\mathrm{x}} = \max_{1\le l_1\le l_2\le p}\frac{(s_{1l_1l_2}-s_{2l_1l_2})^2}{\hat{\theta}_{1l_1l_2}/n_1+\hat{\theta}_{2l_1l_2}/n_2}, \tag{2.5}$$

with

$$\hat{\theta}_{kl_1l_2} = n_k^{-1} \sum_{i=1}^{N_k} \{(x_{kl_1i} - \bar{x}_{kl_1})(x_{kl_2i} - \bar{x}_{kl_2}) - s_{kl_1l_2}\}^2,$$

$\mathbf{S}_1 = (s_{1l_1l_2})_{l_1,l_2=1,\dots,p}$ and $\mathbf{S}_2 = (s_{2l_1l_2})_{l_1,l_2=1,\dots,p}$. Borrowing the idea of [6] and [21], let the second test statistic be

$$T_{\mathrm{drx}_1} = T_{\mathrm{dr}} + p^2 I(T_{\mathrm{x}} > s(N_1, N_2, p)),$$

where $I(\cdot)$ is the indicator function, and $s(N_1, N_2, p)$ is a pre-specified threshold depending on the sample sizes N_1, N_2 and the dimension p. Specifically, with a carefully selected threshold $s(N_1, N_2, p)$, the second term in T_{drx_1} converges to zero under the null hypothesis, and it takes effect as long as T_{x} detects a strong signal. As a result, the first term T_{dr} plays a dominant role, and the second term serves as a power enhancer for screening sparse disturbances between the two covariance matrices. Therefore, such a weighted statistic is able to adaptively combine the information from T_{d}, T_{r} and T_{x}. The corresponding rejection region is

$$\{\mathbf{x}_{11}, \dots \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2} : T_{\mathrm{drx}_1} > t_\alpha\}, \tag{2.6}$$

with t_α obtained from (2.3).

Test 3: To incorporate the information of T_{dr} and T_{x}, we propose another statistic T_{drx_2} as follow,

$$T_{\mathrm{drx}_2} = \max\{T_{\mathrm{dr}}, c_\alpha T_{\mathrm{x}}\},$$

where $c_\alpha = t_{\alpha/2}/q_{\alpha/2}$, $t_{\alpha/2}$ satisfies the equation (2.3) with α replaced by $\alpha/2$, and

$$q_{\alpha/2} = -\log(8\pi) - 2\log\log(1-\alpha/2)^{-1} + 4\log p - \log\log p.$$

The rejection region is

$$\{\mathbf{x}_{11}, \dots \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2} : T_{\mathrm{drx}_2} > t_{\alpha/2}\}. \tag{2.7}$$

The statistic T_{drx_2} has a similar form as T_{dr}, but incorporates the additional contribution from the statistic T_{x}.

Theorem 2.3. *We have the following results:*

- *Under the conditions of Theorem 2.1, the test based on* T_{dr} *has an asymptotic significance level* α*;*
- *Under the conditions of Theorem 2.1 and Conditions (C1), (C2) (or (C2*)) and (C3) of [5], if the threshold* $s(N_1, N_2, p)$ *satisfies* $s(N_1, N_2, p) - 4\log p \geq 0$*, the test based on* T_{drx_1} *has an asymptotic significance level* α*;*
- *Under the conditions of Theorem 2.1 and Conditions (C1), (C2) (or (C2*)) and (C3) of [5], the size of the test based on* T_{drx_2} *is asymptotically equal to or less than the significance level* α*.*

Theorem 2.3 states that all of the three proposed methods can maintain the nominal test level asymptotically. According to Theorem 2.3, in this paper, we take

$$s(N_1, N_2, p) = \{[\log\log(N_1/2 + N_2/2) - 1]^2 + 4\}(\log p - \log\log p/4) + q \quad (2.8)$$

with

$$\exp\{-(8\pi)^{-1/2}\exp(-q/2)\} = 0.99.$$

It is easy to show that such a specification satisfies $s(N_1, N_2, p) - 4\log p \geq 0$, thus the result of T_{drx_1} in Theorem 2.3 holds.

Remark 2.3. *Last, several remarks are in place for comparing T_{drx_1} and T_{drx_2}. Both of the two tests borrow information from the maximum norm statistic T_{x} for power enhancement under the sparse alternatives. But the role of T_{x} functions in different ways: T_{drx_1} directly adds the contribution from T_{x} to the main term T_{dr} without changing the critical value; T_{drx_2} integrates the information from T_{drx_1} and T_{x} using a similar idea as the Tippett's combination test [18, 9], and it is essentially a multiple-testing procedure that requires multiplicity adjustments. Since $t_{\alpha/2} > t_\alpha$, when the main signals come from T_{dr}, it can be shown that the power of T_{drx_1} is greater than that of T_{drx_2}. On the other hand, to guarantee the asymptotic unbiasedness of T_{drx_1}, the contribution of T_{x} to T_{drx_1} is heavily penalized. As a result, T_{drx_2} sometimes is more powerful than T_{drx_1} under sparse alternatives. These results are all confirmed in the simulation studies.*

3. Power analysis

From Theorems 2.1-2.2, the rejection regions of the tests based on the statistics T_{d} and T_{r} are defined as

$$\{\mathbf{x}_{11}, ...\mathbf{x}_{1N_1}, \mathbf{x}_{21}, ..., \mathbf{x}_{2N_2} : |T_{\mathrm{d}} - \mu_0 - \hat{\mu}_{10}|/\sqrt{\hat{\sigma}_{110}} > z_{1-\alpha/2}\} \quad (3.1)$$

and

$$\{\mathbf{x}_{11}, ...\mathbf{x}_{1N_1}, \mathbf{x}_{21}, ..., \mathbf{x}_{2N_2} : |T_{\mathrm{r}} - \mu_{20}|/\sqrt{\hat{\sigma}_{220}} > z_{1-\alpha/2}\}, \quad (3.2)$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of $N(0,1)$.

In this section, we study the asymptotic powers of T_{d}, T_{r}, T_{dr}, T_{drx_1} and T_{drx_2} under the following alternative sets:

$$\begin{aligned}
\Pi_1 &= \Big\{(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) : \mathbf{\Sigma}_2 = \tau_p \mathbf{\Sigma}_1, \tau_p = p/(p+a_1), a_1 > 0 \text{ is a constant}\Big\}, \\
\Pi_2 &= \Big\{(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) : \mathbf{\Sigma}_2 = \mathbf{\Sigma}_1 + \frac{a_2}{p}\mathbf{J}_p, \mathbf{\Sigma}_1 = \mathbf{I}_p, a_2 > 0 \text{ is a constant}\Big\}, \\
\Pi_3 &= \Big\{(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) : \max_{1 \leq l_1 \leq l_2 \leq p} \frac{(\sigma_{1l_1l_2} - \sigma_{2l_1l_2})^2}{\theta_{1l_1l_2}/n_1 + \theta_{2l_1l_2}/n_2} \geq 16\log p\Big\},
\end{aligned}$$

where $\mathbf{J}_p$ is a $p \times p$ matrix with all elements being 1, and $\theta_{kl_1l_2} = \mathrm{Var}((x_{kl_11} - \mu_{kl_1})(x_{kl_21} - \mu_{kl_2}))$ for $k = 1, 2$. The first two sets correspond to local dense alternatives, and the last set includes sparse alternatives as a special case.

3.1. Power analysis in the dense alternative set Π_1

The following theorem gives the joint limiting distribution of T_{d} and T_{r} when $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_1$.

Theorem 3.1. *Under Assumptions A-B and $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_1$, we have*

$$\begin{pmatrix} T_{\mathrm{d}} - \mu_0 - \mu_{11} \\ T_{\mathrm{r}} - \mu_{21} \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}_2, \begin{pmatrix} \sigma_{111} & \sigma_{121} \\ \sigma_{211} & \sigma_{220} \end{pmatrix}\right),$$

where

$$\begin{aligned}
m_{11} &= \int x dL_1,\ m_{21} = \int x^2 dL_1, \\
\mu_{11} &= (y_{n_1} + y_{n_2}\tau_p^2)p^{-1}\mathrm{tr}\mathbf{\Sigma}_1^2 + (\beta_1 y_{n_1} + \beta_2 y_{n_2}\tau_p^2)p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1 \circ \mathbf{\Sigma}_1), \\
\mu_{21} &= p\Big[-\frac{2}{(1-y_{n_2})\tau_p} + \frac{y_{n_1}}{(1-y_{n_2})^2\tau_p^2} + \frac{1}{(1-y_{n_2})^3\tau_p^2}\Big] + p \\
&\quad -\frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} - \frac{2\beta_2 y_{n_2}}{(1-y_{n_2})} \\
&\quad +\frac{2\beta_2 y_{n_1}y_{n_2} + \beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1-y_{n_2})^2} + \frac{2\beta_2 y_{n_2}}{(1-y_{n_2})^3}, \\
\sigma_{111} &= 4(y_1+y_2)^2 m_{21}^2, \\
\sigma_{121} &= \sigma_{211} = \Big[\frac{8y_2(y_1+y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3}\Big] m_{11}^2.
\end{aligned}$$

Proposition 3.1. *Under the conditions of Theorem 3.1, we have*

(I) *For the test based on T_{d},*

$$P(|T_{\mathrm{d}} - \mu_0 - \hat{\mu}_{10}|/\sqrt{\hat{\sigma}_{110}} > z_{1-\alpha/2}) \to \alpha;$$

(II) *For the test based on T_{r},*

$$\begin{aligned}
&P(|T_{\mathrm{r}} - \mu_{20}|/\sqrt{\hat{\sigma}_{220}} > z_{1-\alpha/2}) \\
\to\ & 1 - \Big[\Phi(z_{1-\alpha/2} - \Delta_1) - \Phi(-z_{1-\alpha/2} - \Delta_1)\Big] > \alpha,
\end{aligned}$$

where $\Phi(\cdot)$ is the distribution function of $N(0,1)$ and

$$\Delta_1 = \frac{2a_1}{\sqrt{\sigma_{220}}}\Big[\frac{y_1+y_2}{(1-y_2)^2} + \frac{y_2}{(1-y_2)^3}\Big];$$

(III) *For the test based on T_{dr},*

$$P(T_{\mathrm{dr}} > t_\alpha) \to 1 - \int_{-t'_\alpha - \Delta_1}^{t'_\alpha - \Delta_1} \int_{-t'_\alpha}^{t'_\alpha} f'(x_{\mathrm{d}}, x_{\mathrm{r}}) dx_{\mathrm{d}} dx_{\mathrm{r}} > \alpha,$$

where $f'(x_{\mathrm{d}}, x_{\mathrm{r}})$ *is the density of* $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix}\right)$ *with*

$$\rho_1 = \sigma_{121}/\sqrt{\sigma_{111}\sigma_{220}}, \quad \alpha = 1 - \int_{-t'_\alpha}^{t'_\alpha} \int_{-t'_\alpha}^{t'_\alpha} f'(x_{\mathrm{d}}, x_{\mathrm{r}}) dx_{\mathrm{d}} dx_{\mathrm{r}}; \tag{3.3}$$

(IV) For the test based on T_{drx_1}*,*

$$\lim_{p\to\infty} P(T_{\mathrm{drx}_1} > t_\alpha) \geq \lim_{p\to\infty} P(T_{\mathrm{dr}} > t_\alpha) > \alpha;$$

(V) For the test based on T_{drx_2}*,*

$$\lim_{p\to\infty} P(T_{\mathrm{drx}_2} > t_{\alpha/2}) \geq 1 - \int_{-t'_{\alpha/2}-\Delta_1}^{t'_{\alpha/2}-\Delta_1} \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} f'(x_{\mathrm{d}}, x_{\mathrm{r}}) dx_{\mathrm{d}} dx_{\mathrm{r}} > \alpha/2,$$

where $t'_{\alpha/2}$ *satisfies the equation (3.3) with replacing* α *by* $\alpha/2$*.*

Proposition 3.1 shows that the test based on T_{d} suffers from low power if $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_1$. Since the asymptotical power functions of the tests based on T_{r}, T_{dr}, T_{drx_1} and T_{drx_2} are increasing functions of Δ_1, these tests will enjoy high powers if Δ_1 is large enough.

3.2. Power analysis in the dense alternative set $\mathbf{\Pi_2}$

The following theorem gives the joint limiting distribution of T_{d} and T_{r} when $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_2$.

Theorem 3.2. *Under Assumptions A-B and* $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_2$*, if* $\beta_2 = 0$*, then*

$$\begin{pmatrix} T_{\mathrm{d}} - \mu_0 - \mu_{12} \\ T_{\mathrm{r}} - \mu_{22} \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}_2, \begin{pmatrix} \sigma_{112} & \sigma_{122} \\ \sigma_{212} & \sigma_{222} \end{pmatrix}\right),$$

where

$$\begin{aligned}
\mu_{12} &= a_2^2 + y_{n_1} + y_{n_2} + \beta_1 y_{n_1}, \\
\mu_{22} &= p\Big[-\frac{2}{(1-y_{n_2})} + \frac{y_{n_1}}{(1-y_{n_2})^2} + \frac{1}{(1-y_{n_2})^3}\Big] + p \\
&\quad + \frac{a_2^2 - 2a_2(1+a_2)(y_{n_1}+y_{n_2})}{(1-y_{n_2})^2(1+a_2)^2} - \frac{2a_2 y_{n_2}}{(1-y_{n_2})^3(1+a_2)} \\
&\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} + \frac{\beta_1 y_{n_1}}{(1-y_{n_2})^2}, \\
\sigma_{112} &= 4(y_1+y_2)^2, \\
\sigma_{122} &= \sigma_{212} = \frac{8y_2(y_1+y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3}, \\
\sigma_{222} &= \frac{8y_1^3 + 16y_1^2y_2}{(1-y_2)^5} + \frac{4y_1^2 + 40y_1^2y_2 + 64y_1y_2^2}{(1-y_2)^6}
\end{aligned}$$

$$+\frac{8y_1y_2^4+56y_1y_2^2+48y_2^3+8y_1y_2}{(1-y_2)^7}+\frac{8y_2^5+24y_2^3+4y_2^2}{(1-y_2)^8}$$
$$+4\beta_1y_1\Big[\frac{(y_1+y_2)^2}{(1-y_2)^4}+\frac{2y_2(y_1+y_2)}{(1-y_2)^5}+\frac{y_2^2}{(1-y_2)^6}\Big].$$

Proposition 3.2. *Under the conditions of Theorem 3.2, we have*

(I) For the test based on T_{d},

$$\begin{aligned}&P(|T_{\mathrm{d}}-\mu_0-\hat{\mu}_{10}|/\sqrt{\hat{\sigma}_{110}}>z_{1-\alpha/2})\\ \to\quad &1-\Big[\Phi(z_{1-\alpha/2}-\Delta_2)-\Phi(-z_{1-\alpha/2}-\Delta_2)\Big]>\alpha,\end{aligned}$$

where $\Delta_2=a_2^2/[2(y_1+y_2)]$;

(II) For the test based on T_{r},

$$\begin{aligned}&P(|T_{\mathrm{r}}-\mu_{20}|/\sqrt{\hat{\sigma}_{220}}>z_{1-\alpha/2})\\ \to\quad &1-\Big[\Phi(z_{1-\alpha/2}-\Delta_3)-\Phi(-z_{1-\alpha/2}-\Delta_3)\Big]\geq\alpha,\end{aligned}$$

where

$$\Delta_3=\frac{1}{\sqrt{\sigma_{222}}}\Big[\frac{a_2^2-2a_2(1+a_2)(y_1+y_2)}{(1-y_2)^2(1+a_2)^2}-\frac{2a_2y_2}{(1-y_2)^3(1+a_2)}\Big];$$

(III) For the test based on T_{dr},

$$P(T_{\mathrm{dr}}>t_\alpha)\to1-\int_{-t_\alpha^*-\Delta_3}^{t_\alpha^*-\Delta_3}\int_{-t_\alpha^*-\Delta_2}^{t_\alpha^*-\Delta_2}f^*(x_{\mathrm{d}},x_{\mathrm{r}})dx_{\mathrm{d}}dx_{\mathrm{r}}>\alpha,$$

where $f^*(x_{\mathrm{d}},x_{\mathrm{r}})$ *is the density of* $N\left(\mathbf{0}_2,\begin{pmatrix}1&\rho_2\\ \rho_2&1\end{pmatrix}\right)$ *with*

$$\rho_2=\sigma_{122}/\sqrt{\sigma_{112}\sigma_{222}},\quad\alpha=1-\int_{-t_\alpha^*}^{t_\alpha^*}\int_{-t_\alpha^*}^{t_\alpha^*}f^*(x_{\mathrm{d}},x_{\mathrm{r}})dx_{\mathrm{d}}dx_{\mathrm{r}};\tag{3.4}$$

(IV) For the test based on T_{drx_1},

$$\lim_{p\to\infty}P(T_{\mathrm{drx}_1}>t_\alpha)\geq\lim_{p\to\infty}P(T_{\mathrm{dr}}>t_\alpha)>\alpha;$$

(V) For the test based on T_{drx_2},

$$\lim_{p\to\infty}P(T_{\mathrm{drx}_2}>t_{\alpha/2})\geq1-\int_{-t_{\alpha/2}^*-\Delta_3}^{t_{\alpha/2}^*-\Delta_3}\int_{-t_{\alpha/2}^*-\Delta_2}^{t_{\alpha/2}^*-\Delta_2}f^*(x_{\mathrm{d}},x_{\mathrm{r}})dx_{\mathrm{d}}dx_{\mathrm{r}}>\alpha/2,$$

where $t_{\alpha/2}^*$ *satisfies the equation (3.4) with parameter* $\alpha/2$ *instead of* α.

Proposition 3.2 shows that the tests based on T_{d}, T_{r}, T_{dr} and T_{drx_1} are all asymptotically unbiased if $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) \in \Pi_2$ and $\beta_2 = 0$. Apparently, Δ_2 is an increasing function of a_2, the tests based on T_{d}, T_{dr}, T_{drx_1} and T_{drx_2} will enjoy high powers when a_2 is large enough. However, with the increase of a_2, Δ_3 will converge to

$$\frac{1}{\sqrt{\sigma_{222}}}\Big[\frac{1-2(y_1+y_2)}{(1-y_2)^2} - \frac{2y_2}{(1-y_2)^3}\Big].$$

Thus, the test based on T_{r} will suffer from low power in the case of small Δ_3. For example, when $y_1 = 0.5 - (2y_2 - y_2^2)/(1-y_2)$ and $a_2 \to \infty$, it follows that the power of the test based on T_{r} converges to α.

3.3. Power analysis in the alternative set $\boldsymbol{\Pi_3}$

The following proposition gives some results on the power functions of the tests based on T_{drx_1} and T_{drx_2} under the alternative set Π_3. Let

$$\boldsymbol{\Sigma}_d = \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2, \qquad \boldsymbol{\Sigma}_w = \frac{n_1}{n_1+n_2}\boldsymbol{\Sigma}_1 + \frac{n_2}{n_1+n_2}\boldsymbol{\Sigma}_2.$$

Proposition 3.3. *When* $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) \in \Pi_3$, *we have*

(I) Under Assumptions A-B and Conditions (C2) or (C2) in [5], if* $p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_d$, $p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_w$ *and* $p^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_w^2)$ *converge and*

$$\max_{1\le l_1\le l_2\le p} \frac{(\sigma_{1l_1l_2} - \sigma_{2l_1l_2})^2}{\theta_{1l_1l_2}/n_1 + \theta_{2l_1l_2}/n_2} \ge 2s(N_1, N_2, p) + 8\log p,$$
$$P(T_{\mathrm{drx}_1} > t_\alpha) \to 1;$$

(II) Under Conditions (C2) or (C2) in [5],*

$$P(T_{\mathrm{drx}_2} > t_{\alpha/2}) \to 1.$$

It can be seen from Proposition 3.3 that the power functions of the tests based on T_{drx_1} and T_{drx_2} will tend to 1 if some entries of $\boldsymbol{\Sigma}_d$ are large enough.

4. Simulation studies

We perform extensive simulation studies to examine the finite-sample performance of the proposed new tests. The observations are drawn from $\mathbf{x}_{ki} = \boldsymbol{\Sigma}_k^{1/2}\mathbf{w}_{ki}$, where $\{w_{kli}, k = 1,2; l = 1,\cdots,p; i = 1,\cdots,N_k\}$ are independent and identically distributed (*i.i.d.*) from the Gaussian population $N(0,1)$ or the Gamma population $Gamma(4,2) - 2$. We evaluate our proposed tests under four different scenarios for $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$.

- **Scenario 1**.
  The population covariance matrix is $\boldsymbol{\Sigma}_k = \boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^{\mathrm{T}}$, where $\boldsymbol{\Gamma}_k = \mathbf{I}_p + \theta_k(u_{l_1l_2})_{l_1,l_2=1}^p$ for $k = 1,2$, and $\{u_{l_1l_2}, l_1, l_2 = 1,\cdots,p\}$ are *i.i.d.* from Unif

$(-(2p)^{-2/3}, (2p)^{-2/3})$. We evaluate the empirical size with $\theta_1 = \theta_2 = 0$, and the empirical power with $(\theta_1, \theta_2) = (0, 1)$, which leads to a dense alternative. This data-generating model was used in [21].

- **Scenario 2**.
We take $\mathbf{\Sigma}_k = (\rho_k^{|l_1-l_2|})_{l_1,l_2=1}^p$, where $|\rho_k| < 1$ for $k = 1, 2$. When $\rho_1 = \rho_2$, $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$. When $\rho_1 \neq \rho_2$, $\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2$ has relatively dense signals. We evaluate the empirical size with $\rho_1 = \rho_2 = 0.5$ and the empirical power with $(\rho_1, \rho_2) = (0.5, 0.55)$.
- **Scenario 3**.
Denote $\mathbf{C} = \big((0.1^{|l_1-l_2|} + 0.2^{|l_1-l_2|})/2\big)_{l_1,l_2=1}^p$. Let $\mathbf{U}$ be a $p \times p$ symmetric matrix having only four nonzero entries from Unif$(0.4, 0.6)$ in its upper triangle. The locations of these four nonzero entries are selected randomly from the upper triangle of $\mathbf{U}$. For $k = 1, 2$, we take the population covariance matrix as
$$\mathbf{\Sigma}_k = \mathbf{C} + \delta_0 \mathbf{I}_p + \theta_k \mathbf{U},$$
where $\delta_0 = |\min\{\lambda_{\min}(\mathbf{C}+\mathbf{U}), \lambda_{\min}(\mathbf{C})\}| + 0.05$. As $\theta_1 = \theta_2 = 0$, $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$. As $(\theta_1, \theta_2) = (0, 1)$, $\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2$ is extremely sparse. We evaluate the empirical size with $\theta_1 = \theta_2 = 0$ and the empirical power with $(\theta_1, \theta_2) = (0, 1)$. This scenario was considered in [5].
- **Scenario 4**.
Denote $\mathbf{C} = \big((0.5 - |l_1 - l_2|/10) I(|l_1 - l_2| \leq 4)\big)_{l_1,l_2=1}^p + 0.5\mathbf{I}_p$. Let $\mathbf{U}$ be a $p \times p$ symmetric matrix having $\lceil p^2/4 \rceil$ entries from Unif$(0, 0.04)$ in its upper triangle and $\mathbf{E}$ be a $p \times p$ symmetric matrix having only one entry $\log p/10$ in its upper triangle, where $\lceil x \rceil$ denotes the smallest integer not less than x. The locations of the nonzero entries in both $\mathbf{U}$ and $\mathbf{E}$ are selected randomly from their upper triangle. For $k = 1, 2$, the population covariance matrix is taken as
$$\mathbf{\Sigma}_k = \mathbf{C} + \delta_0 \mathbf{I}_p + \theta_k (\mathbf{U} + \mathbf{E}),$$
where $\delta_0 = |\min\{\lambda_{\min}(\mathbf{C} + \mathbf{U} + \mathbf{E}), \lambda_{\min}(\mathbf{C})\}| + 0.05$. When $\theta_1 = \theta_2 = 0$, $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$. When $(\theta_1, \theta_2) = (0, 1)$, $\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2$ has a mixture of dense and sparse signals. We evaluate the empirical size with $\theta_1 = \theta_2 = 0$ and the empirical power with $(\theta_1, \theta_2) = (0, 1)$.

The above four scenarios all satisfy the conditions listed in Theorem 2.3 with the justifications provided in Appendix. We compare our new tests with two existing tests LC [11] and CLX [5]. The nominal significant level for all the tests is set at $\alpha = 0.05$. Based on 10,000 replications under each scenario, the empirical sizes for the Gaussian and Gamma populations are reported in Tables 1-4, and the empirical powers for each of the considered methods are exhibited Figures 1-2. For reference, we also provide detailed numerical values for power comparison in Tables A.1-A.4 of Appendix.

It is observed that all the tests can maintain the nominal level for both Gaussian and Gamma populations. For power comparisons, relative performance of the seven tests depend on the different scenarios. In Scenarios 1-2, the test CLX

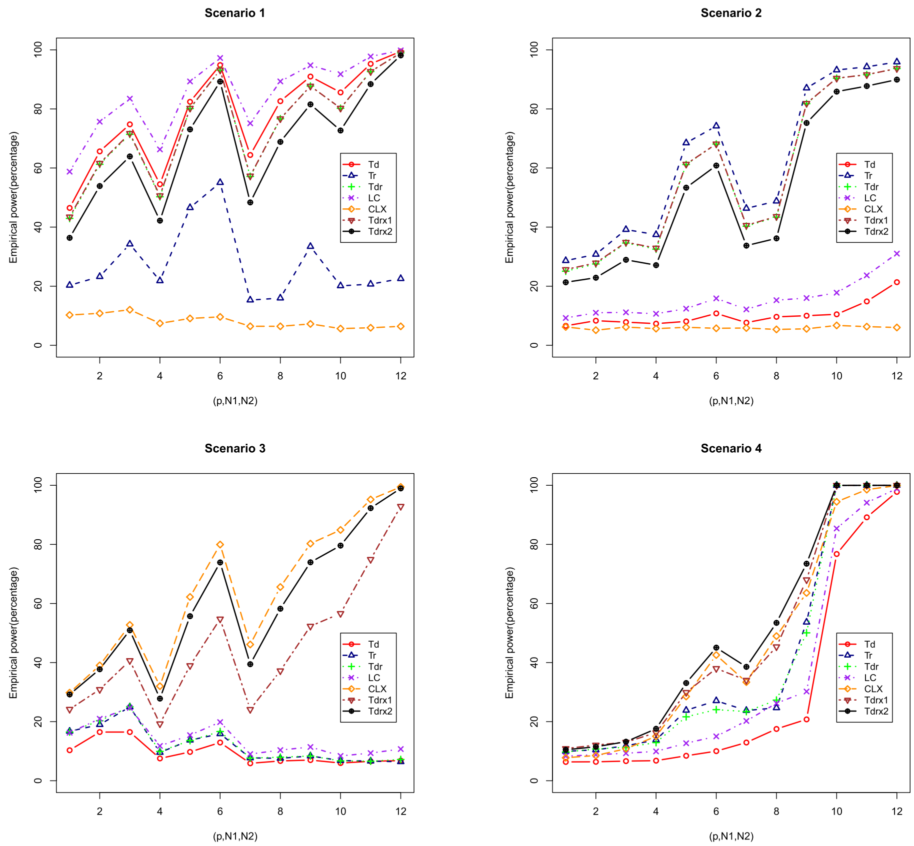

FIG 1. *Comparison of empirical powers (in percentages) of the proposed tests with existing tests under the considered four scenarios when the observations are sampled from a Gaussian population.* (p, N_1, N_2) *is varied from (40,80,120) to (320,480,480), and x-axis takes values correspond to the first column in Tables 1-4.*

fails to detect the dense but small disturbances and yields low empirical testing powers. The test based on T_{d} performs better than the test based on T_{r} in Scenario 1; while the test based on T_{r} is more powerful than the test based on T_{d} in Scenario 2. Since the proposed three tests based on T_{dr}, T_{drx_1} and T_{drx_2} all have the advantages of the tests based on T_{d} and T_{r}, these three tests can maintain competitive testing powers in Scenarios 1-2. Specifically, consistent with Remark 2.3, T_{drx_1} delivers higher powers than T_{drx_2} in these two scenarios. Scenario 3 examines the performance of the considered tests under a sparse alternative, as expected, the tests based on T_{d}, T_{r} and T_{dr} fail to detect the sparse signals and have low powers. By contrast, the tests based on T_{drx_1} and T_{drx_2} are adaptive and thus yield high powers. Scenario 4 is a hybrid alternative that has a mixture of dense and sparse signals, in this case, the proposed tests based on T_{drx_1} and T_{drx_2} have better performance than the other methods in

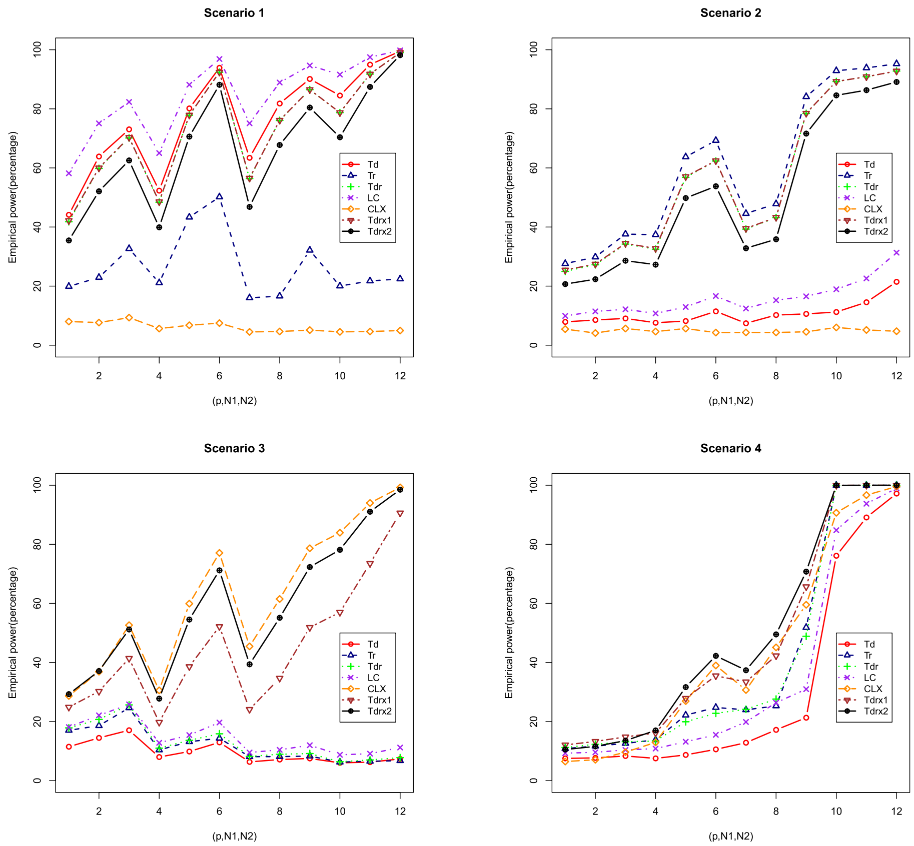

FIG 2. *Comparison of empirical powers (in percentages) of the proposed tests with existing tests under the considered four scenarios when the observations are sampled from a Gamma population.* (p, N_1, N_2) *is varied from (40,80,120) to (320,480,480), and x-axis takes values correspond to the first column in Tables 1-4.*

terms of empirical powers. Since sparse differences exist in last two scenarios, the test based on T_{drx_2} is more powerful than that using T_{drx_1}. In summary, the simulation studies show that the proposed three tests can effectively borrow strengths from each of the individual statistics in signal detection, and the nominal testing level is well maintained under all methods.

Appendix

The appendix includes three sections: Section A.1 gives Conditions (C1), (C2) (or (C2*)) and (C3) of [5], Section A.2 justifies the conditions of the four scenarios considered in the simulation studies, and Section A.3 provides the proofs of some theorems and propositions.

TABLE 1
Comparison of empirical sizes (in percentages) of the proposed tests with the tests LC and CLX under Scenario 1 when the observations are from the Gaussian and Gamma population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	5.42	6.06	6.11	5.84	5.03	6.54	5.45
2	40	(120,120)	5.09	6.09	6.23	5.53	4.85	6.69	6.01
3	40	(120,160)	5.33	5.52	5.51	5.49	4.88	5.99	5.24
4	80	(120,160)	5.13	6.35	6.05	5.32	4.81	6.55	5.74
5	80	(160,240)	4.94	5.42	5.29	5.42	5.21	5.63	5.2
6	80	(240,240)	5.01	5.52	5.59	5.31	4.69	5.77	5.21
7	160	(160,240)	5.26	6.07	5.98	5.15	4.08	6.26	5.29
8	160	(240,240)	5.28	6.48	6.27	5	4.65	6.62	5.51
9	160	(240,320)	5.11	5.38	5.3	4.94	4.48	5.48	5.11
10	320	(240,480)	5.12	5.43	5.46	5.28	4.38	5.6	4.9
11	320	(320,480)	4.61	5.37	5.33	4.76	4.71	5.48	4.97
12	320	(480,480)	4.79	5.59	5.5	4.72	4.62	5.59	4.92
			Gamma population						
1	40	(80,120)	6.64	6.65	7.76	6.26	3.98	8.1	6.37
2	40	(120,120)	6.43	7.01	7.71	6.41	3.42	7.93	6.1
3	40	(120,160)	6.61	6.85	7.46	6.33	3.53	7.69	5.78
4	80	(120,160)	5.64	6.51	6.63	5.72	3.36	6.9	5.44
5	80	(160,240)	5.8	6.15	6.45	5.49	3.39	6.64	5.24
6	80	(240,240)	5.76	6.35	6.31	5.41	3.62	6.46	5
7	160	(160,240)	5.56	7.09	6.86	5.6	3.2	7.04	5.28
8	160	(240,240)	4.83	6.84	6.2	4.77	3.35	6.4	4.95
9	160	(240,320)	5.5	5.59	5.63	5.08	3.34	5.76	4.43
10	320	(240,480)	5.22	5.9	5.83	5.1	3.42	5.95	4.69
11	320	(320,480)	5.31	5.75	5.84	5.38	3.17	5.88	4.57
12	320	(480,480)	5.01	5.54	5.6	4.95	3.45	5.67	4.57

A.1. Review of Conditions (C1), (C2) (or (C2)) and (C3) of [5]*

Before giving the Conditions (C1), (C2) (or (C2*)) and (C3) of [5], we introduce some basic notations. For two sequences of real numbers $\{a_n\}$ and $\{b_n\}$, write $a_n \asymp b_n$ if there exist constants $C > c > 0$ such that $c|b_n| \le |a_n| \le C|b_n|$ for all sufficiently large n. Let $N = \max\{N_1, N_2\}$.

Let $\mathbf{R}_k = (\rho_{kl_1l_2})$ be the correlation matrix of the kth population for $k = 1, 2$.

(C1). Assume that there is a positive constant α_0 and a subset $\Upsilon \subset \{1, 2, \cdots, p\}$ with $\mathrm{Card}(\Upsilon) = o(p)$ such that $\max_{1 \le l_2 \le p, l_2 \notin \Upsilon} s_{l_2}(\alpha_0) = o(p^\gamma)$ for all $\gamma > 0$, where s_{l_2}, $l_2 = 1, \ldots, p$, is the cardinality of the set of indices l_1 ($l_1 \ne l_2$) that are highly correlated with variable l_2 in population 1 or 2, as given by

$$s_{l_2} = s_{l_2}(\alpha_0) := \mathrm{card}\{l_1 : |\rho_{1l_1l_2}| \ge (\log p)^{-1-\alpha_0} \text{ or } |\rho_{2l_1l_2}| \ge (\log p)^{-1-\alpha_0}\}.$$

TABLE 2

Comparison of empirical sizes (in percentages) of the proposed tests with the tests LC and CLX under Scenario 2 when the observations are from the Gaussian and Gamma population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	5.75	6.11	6.53	6.59	5	7.08	6.23
2	40	(120,120)	6.12	6.23	6.74	6.48	4.45	7.2	5.92
3	40	(120,160)	5.85	6.13	6.39	6.5	4.63	6.8	5.87
4	80	(120,160)	5.31	6.01	6.1	6	4.83	6.55	5.77
5	80	(160,240)	5.37	5.53	5.57	5.65	4.7	5.93	5.1
6	80	(240,240)	5.68	5.46	5.85	5.98	4.71	6.14	5.26
7	160	(160,240)	5.36	6.25	5.99	5.65	4.66	6.26	5.52
8	160	(240,240)	5.17	5.56	5.44	5.13	4.64	5.73	5.29
9	160	(240,320)	4.84	5.35	5.3	4.77	4.46	5.56	4.65
10	320	(240,480)	4.85	5.69	5.48	5.08	4.48	5.57	5.13
11	320	(320,480)	4.99	5.86	5.45	5.13	4.76	5.61	4.89
12	320	(480,480)	5.09	5.72	5.41	5.32	4.75	5.55	5.33
			Gamma population						
1	40	(80,120)	6.65	6.89	7.79	6.91	4.09	8.19	6.49
2	40	(120,120)	6.25	7.2	7.8	6.67	3.27	8.12	6.12
3	40	(120,160)	6.73	6.83	7.86	7.14	4.3	8.05	6.56
4	80	(120,160)	6.18	7.01	7.14	6.51	3.62	7.43	5.83
5	80	(160,240)	5.91	5.75	6.09	5.97	3.93	6.36	5.02
6	80	(240,240)	5.92	5.85	6.31	6.23	3.51	6.54	5.1
7	160	(160,240)	5.42	6.42	6.63	5.46	3.2	6.79	5.24
8	160	(240,240)	5.64	6.97	6.57	5.63	3.68	6.71	5.27
9	160	(240,320)	5.84	5.79	6.06	5.62	3.15	6.26	4.76
10	320	(240,480)	5.11	5.39	5.08	5.52	3.63	5.14	4.06
11	320	(320,480)	5.23	5.53	5.72	5.35	3.7	5.8	4.64
12	320	(480,480)	5.43	5.79	5.98	5.51	3.74	6.06	4.98

In addition, there is a sequence of numbers $\Lambda_{p,r}$ such that $\mathrm{card}(\Lambda(r)) \leq \Lambda_{p,r} = o(p)$ for some constant $0 < r < 1$, where

$$\Lambda(r) = \{1 \leq l_1 \leq p : |\rho_{1l_1l_2}| \geq r \text{ or } |\rho_{2l_1l_2}| \geq r \text{ for some } l_2 \neq l_1\},$$

that is, $\Lambda(r)$ is the set of highly correlated variable incidences.

(C2). Assume that $\log p = o(N^{1/5})$ and $N_1 \asymp N_2$. There exist constants $\eta > 0$ and $K > 0$ satisfying the following moment conditions

$$\begin{aligned} \mathrm{E}\exp(\eta(\mathbf{x}_{1l1} - \mu_{1l})^2/\sigma_{1ll}) &\leq K, \\ \mathrm{E}\exp(\eta(\mathbf{x}_{2l1} - \mu_{2l})^2/\sigma_{2ll}) &\leq K, \text{ for } l = 1, \cdots, p. \end{aligned}$$

Additionally, for some constants $\tau_1 > 0$ and $\tau_2 > 0$,

$$\min_{1\leq l_1 \leq l_2 \leq p} \frac{\theta_{1l_1l_2}}{\sigma_{1l_1l_1}\sigma_{1l_2l_2}} \geq \tau_1 \quad \text{and} \quad \min_{1\leq l_1 \leq l_2 \leq p} \frac{\theta_{2l_1l_2}}{\sigma_{2l_1l_1}\sigma_{2l_2l_2}} \geq \tau_2. \quad \text{(A.1)}$$

TABLE 3
Comparison of empirical sizes (in percentages) of the proposed tests with the tests LC and CLX under Scenario 3 when the observations are from the Gaussian population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	4.96	5.86	5.95	5.32	5.39	6.4	5.89
2	40	(120,120)	5.44	6.18	6.18	5.46	4.83	6.57	5.6
3	40	(120,160)	5.52	6.03	6.17	5.8	4.42	6.56	5.58
4	80	(120,160)	5.29	6.23	6.09	5.3	4.86	6.44	5.77
5	80	(160,240)	5.3	5.76	5.55	4.96	4.79	5.84	4.95
6	80	(240,240)	5.31	5.51	5.5	5.29	5.24	5.89	5.27
7	160	(160,240)	4.71	6.17	5.94	5.12	4.9	6.18	5.74
8	160	(240,240)	4.87	6.02	6.17	4.9	4.41	6.5	5.43
9	160	(240,320)	5.11	5.67	5.65	4.84	4.53	5.89	4.89
10	320	(240,480)	5.03	5.73	5.57	5.26	4.48	5.74	4.9
11	320	(320,480)	4.99	5.45	5.62	5.49	4.36	5.73	5.01
12	320	(480,480)	5.05	5.58	5.48	4.93	4.84	5.59	5.31
			Gamma population						
1	40	(80,120)	6.49	6.47	7.54	6.21	3.65	7.75	5.92
2	40	(120,120)	6.67	7.68	7.97	6.33	3.47	8.22	6.13
3	40	(120,160)	6.3	6.98	7.22	6.32	3.9	7.57	5.88
4	80	(120,160)	5.45	6.7	6.77	5.73	3.25	7.01	5.42
5	80	(160,240)	5.5	5.89	5.76	5.25	3.58	5.98	4.92
6	80	(240,240)	5.63	5.78	6.01	5.32	3.61	6.16	5.19
7	160	(160,240)	5.48	6.71	6.48	5.4	3.54	6.67	5.26
8	160	(240,240)	5.25	6.7	6.76	5.37	3.02	6.87	5.3
9	160	(240,320)	5.33	5.65	5.51	5.43	3.49	5.67	4.51
10	320	(240,480)	4.92	5.37	5.28	5.51	3.67	5.35	4.31
11	320	(320,480)	5.36	5.61	5.72	5.26	3.26	5.82	4.37
12	320	(480,480)	5.65	5.96	6.05	5.57	3.73	6.15	4.84

(C2*). Suppose that condition (A.1) holds, and that $N_1 \asymp N_2$ and $p \leq c_1 N^{\gamma_0}$ for some γ_0, $c_1 > 0$. Furthermore, the following moment conditions hold

$$\begin{aligned} \mathrm{E}|(\mathbf{x}_{1l1} - \mu_{1l})/\sigma_{1ll}^{1/2}|^{4\gamma_0+4+\epsilon} &\leq K, \\ \mathrm{E}|(\mathbf{x}_{2l1} - \mu_{2l})/\sigma_{2ll}^{1/2}|^{4\gamma_0+4+\epsilon} &\leq K, \text{ for } l = 1, \cdots, p, \end{aligned}$$

for some constants $\epsilon > 0$ and $K > 0$.

(C3). For any $l_1, l_2, l_3, l_4 \in \{1, 2, \cdots, p\}$, and for some constants $\kappa_1, \kappa_2 \geq \frac{1}{3}$,

$$\begin{aligned} &\mathrm{E}(\mathbf{x}_{1l_11} - \mu_{1l_1})(\mathbf{x}_{1l_21} - \mu_{1l_2})(\mathbf{x}_{1l_31} - \mu_{1l_3})(\mathbf{x}_{1l_41} - \mu_{1l_4}) \\ &\qquad = \kappa_1(\sigma_{1l_1l_2}\sigma_{1l_3l_4} + \sigma_{1l_1l_3}\sigma_{1l_2l_4} + \sigma_{1l_1l_4}\sigma_{1l_2l_3}), \\ &\mathrm{E}(\mathbf{x}_{2l_11} - \mu_{2l_1})(\mathbf{x}_{2l_21} - \mu_{2l_2})(\mathbf{x}_{2l_31} - \mu_{2l_3})(\mathbf{x}_{2l_41} - \mu_{2l_4}) \\ &\qquad = \kappa_2(\sigma_{2l_1l_2}\sigma_{2l_3l_4} + \sigma_{2l_1l_3}\sigma_{2l_2l_4} + \sigma_{2l_1l_4}\sigma_{2l_2l_3}). \end{aligned}$$

Table 4
Comparison of empirical sizes (in percentages) of the proposed tests with the tests LC and CLX under Scenario 4 when the observations are from the Gaussian and Gamma population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	6.17	5.73	6.7	6.53	5.29	7.2	6.19
2	40	(120,120)	5.65	6.26	6.46	6.19	4.92	6.98	6.03
3	40	(120,160)	5.48	5.77	5.95	6.2	4.78	6.4	5.62
4	80	(120,160)	5.76	6.18	6.23	5.54	4.85	6.62	5.67
5	80	(160,240)	5.97	5.77	6.13	6	4.52	6.46	5.34
6	80	(240,240)	5.36	5.88	5.93	5.56	4.54	6.29	5.29
7	160	(160,240)	5.14	6.09	5.83	5.35	4.83	6.16	5.34
8	160	(240,240)	5.16	6.39	6.33	5.37	4.31	6.66	5.54
9	160	(240,320)	5.12	5.47	5.48	5.22	4.14	5.7	4.84
10	320	(240,480)	4.88	5.05	4.94	5.05	4.3	5.14	4.77
11	320	(320,480)	5.24	5.46	5.53	5.48	4.6	5.7	4.82
12	320	(480,480)	5.44	5.82	5.86	5.28	4.47	6.03	5.05
			Gamma population						
1	40	(80,120)	7.27	6.82	7.91	7.03	4.24	8.29	6.5
2	40	(120,120)	6.83	7.22	7.84	6.74	3.54	8.05	6.62
3	40	(120,160)	7.36	6.82	7.95	7.46	3.99	8.27	6.18
4	80	(120,160)	6.02	6.46	6.57	6.33	3.66	6.8	5.54
5	80	(160,240)	5.94	5.94	6.53	6.02	3.66	6.69	5.15
6	80	(240,240)	5.96	5.96	6.43	5.99	3.61	6.6	5.36
7	160	(160,240)	5.29	6.36	6.2	5.74	3.54	6.42	5.06
8	160	(240,240)	5.7	6.94	6.67	5.62	3.51	6.75	5.33
9	160	(240,320)	5.49	5.76	5.85	5.41	3.7	5.98	4.81
10	320	(240,480)	5.65	5.85	5.97	5.49	3.68	6.12	4.95
11	320	(320,480)	5.08	5.62	5.66	5.01	3.82	5.74	4.62
12	320	(480,480)	5.13	6.03	5.69	5.45	3.55	5.78	4.77

A.2. Verification of the considered simulation scenarios

In what follows, we verify our simulation settings in the simulation part satisfy the three conditions as shown in Appendix A.1. For easiness of presentation, some notations are given:

- Let $\mathbf{\Sigma}_{1k} = (\sigma_{1kl_1l_2})_{l_1,l_2=1,\dots,p}$ and $\mathbf{\Sigma}_{2k} = (\sigma_{2kl_1l_2})_{l_1,l_2=1,\dots,p}$ be the population covariance matrices in Scenarios $k = 1, 2, 3, 4$;
- Let $\mathbf{R}_{1k} = (\rho_{1kl_1l_2})_{l_1,l_2=1,\dots,p}$ and $\mathbf{R}_{2k} = (\rho_{2kl_1l_2})_{l_1,l_2=1,\dots,p}$ be the population correlation matrices in Scenarios $k = 1, 2, 3, 4$;
- Let $\mathbf{r}_{1kl}$ and $\mathbf{r}_{2kl}$ denote the lth column of the matrices $\mathbf{\Sigma}_{1k}^{1/2}$ and $\mathbf{\Sigma}_{2k}^{1/2}$, respectively;
- Let $\mathbf{w} = (w_1, \cdots, w_p)$ denote the vector with elements being *i.i.d.* from $N(0, 1)$ or $Gamma(4, 2) - 2$.

- In our simulation studies, under the null hypothesis $\mathbf{\Sigma}_{1k} = \mathbf{\Sigma}_{2k}$, the data are *i.i.d.* from $\mathbf{\Sigma}_{1k}^{1/2}\mathbf{w}$.

(1) All $\mathbf{\Sigma}_{1k}$'s in Scenarios 1-4 satisfy Condition (C1) under the null hypothesis H_0.

For a fixed constant $\alpha_0 > 0$, define

$$s_{kl_2} = s_{kl_2}(\alpha_0) := \text{card}\{l_1 : |\rho_{1kl_1l_2}| \geq (\log p)^{-1-\alpha_0} \text{ or } |\rho_{2kl_1l_2}| \geq (\log p)^{-1-\alpha_0}\}$$

and

$$s'_{kl_2} = s'_{kl_2}(\alpha_0) := \text{card}\{l_1 : |\sigma_{1kl_1l_2}| \geq (\log p)^{-1-\alpha_0} \text{ or } |\sigma_{2kl_1l_2}| \geq (\log p)^{-1-\alpha_0}\}.$$

For $0 < r < 1$, define the sets

$$\Lambda_k(r) = \{1 \leq l_1 \leq p : |\rho_{1kl_1l_2}| \geq r \text{ or } |\rho_{2kl_1l_2}| \geq r \text{ for some } l_2 \neq l_1\}$$

and

$$\Lambda'_k(r) = \{1 \leq l_1 \leq p : |\sigma_{1kl_1l_2}| \geq r \text{ or } |\sigma_{2kl_1l_2}| \geq r \text{ for some } l_2 \neq l_1\}.$$

Since all diagonal elements of $\mathbf{\Sigma}_{1k}$ are greater than or equal to 1, we have $\rho_{1kl_1l_2} \leq \sigma_{1kl_1l_2}$, which indicates that $s_{kl_2} \leq s'_{kl_2}$ and $\text{card}(\Lambda_k(r)) \leq \text{card}(\Lambda'_k(r))$. Therefore, we only need to show that s'_{kl_2} and $\text{card}(\Lambda'_k(r))$ satisfy Condition (C1) for $k = 1, 2, 3, 4$.

- In Scenario 1, when $l_1 \neq l_2$, $\sigma_{11l_1l_2} = 0$. For any fixed constant $\alpha_0 > 0$ and $0 < r < 1$, we have $s'_{1l_2} = 1$ and $\text{card}(\Lambda'_1(r)) = 0$. Thus, Condition (C1) is satisfied.
- In Scenario 2, when $l_1 \neq l_2$, $\sigma_{12l_1l_2} = 0.5^{|l_1-l_2|}$. When we take $r = 0.5$, we have $\text{card}(\Lambda'_2(r)) = 0$, indicating Condition (C1) is satisfied. For any $\gamma > 0$, any fixed constant $\alpha_0 > 0$ and sufficiently large p, we have $0.5^{p^{\gamma/2}} \leq (\log p)^{-1-\alpha_0}$, which indicates that $\max_{1\leq l_2 \leq p} s'_{2l_2} \leq 2p^{\gamma/2}+1 = o(p^\gamma)$, that is, s'_{2l_2} satisfies Condition (C1).
- In Scenario 3, when $l_1 \neq l_2$, $\sigma_{13l_1l_2} = (0.1^{|l_1-l_2|} + 0.2^{|l_1-l_2|})/2$. Mimicking the discussion in Scenario 2, we can show that s'_{3l_2} and $\text{card}(\Lambda'_3(r))$ satisfy Condition (C1).
- In Scenario 4, when $l_1 \neq l_2$, $\sigma_{14l_1l_2} = (0.5 - |l_1 - l_2|/10)I(|l_1 - l_2| \leq 4)$, where $I(\cdot)$ is the indicator function. When $r = 0.4$, we have $\text{card}(\Lambda'_4(r)) = 0$, which obviously satisfies Condition (C1). For any fixed constant $\alpha_0 > 0$, we have $\max_{1\leq l_2 \leq p} s'_{4l_2} \leq 9$, which also satisfies Condition (C1).

(2) The data of group 1 are *i.i.d.* from $\mathbf{\Sigma}_{1k}^{1/2}\mathbf{w}$, and those of group 2 are *i.i.d.* from $\mathbf{\Sigma}_{2k}^{1/2}\mathbf{w}$, $k = 1, 2, 3, 4$. We show that $\mathbf{\Sigma}_{1k}^{1/2}\mathbf{w}$ and $\mathbf{\Sigma}_{2k}^{1/2}\mathbf{w}$ satisfy Condition (C2*), regardless $\{w_1, \cdots, w_p\}$ are *i.i.d.* from $N(0, 1)$ or $Gamma(4, 2) - 2$.

When $\gamma_0 = 1$, $c_1 = y_1$ and $\epsilon = 2$ in Condition (C2*) for $k = 1, 2, 3, 4$, we have

$$\mathrm{E}|\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{w}/\sigma_{1kll}^{1/2}|^{4\gamma_0+4+\epsilon} = \mathrm{E}|\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{w}/\sigma_{1kll}^{1/2}|^{10} = \mathrm{E}(\mathbf{w}^{\mathrm{T}}\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{w})^5/\sigma_{1kll}^5.$$

From Lemma 2.7 in [2] and $\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{r}_{1kl} = \sigma_{1kll}$, we have

$$\begin{aligned}
& \mathrm{E}(\mathbf{w}^{\mathrm{T}}\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{w})^5 \\
\leq\ & 16\mathrm{E}|\mathbf{w}^{\mathrm{T}}\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{w} - \mathrm{tr}(\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}})|^5 + 16|\mathrm{tr}(\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}})|^5 \\
\leq\ & 16K_5((\mathrm{E}|w_1|^4\mathrm{tr}(\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}}))^{5/2} + \\
& \mathrm{E}|w_1|^{10}\mathrm{tr}(\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{r}_{1kl}\mathbf{r}_{1kl}^{\mathrm{T}})^{5/2}) + 16(\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{r}_{1kl})^5 \\
=\ & 16K_5((\mathrm{E}|w_1|^4)^{5/2} + \mathrm{E}|w_1|^{10})(\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{r}_{1kl})^5 + 16(\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{r}_{1kl})^5 \\
=\ & 16K_5((\mathrm{E}|w_1|^4)^{5/2} + \mathrm{E}|w_1|^{10})\sigma_{1kll}^5 + 16\sigma_{1kll}^5,
\end{aligned}$$

where K_5 is a constant. No matter w_1 is from $N(0,1)$ or $Gamma(4,2) - 2$, $\mathrm{E}|w_1|^4$ and $\mathrm{E}|w_1|^{10}$ are bounded. Therefore, it can be shown that

$$\mathrm{E}|\mathbf{r}_{1kl}^{\mathrm{T}}\mathbf{w}/\sigma_{1kll}^{1/2}|^{4\gamma_0+4+\epsilon} \leq K \text{ for all } l.$$

Similarly, for $k = 1,2,3,4$, we can also prove

$$\mathrm{E}|\mathbf{r}_{2kl}^{\mathrm{T}}\mathbf{w}/\sigma_{2kll}^{1/2}|^{4\gamma_0+4+\epsilon} \leq K \text{ for all } l.$$

Let $\theta_{1kl_1l_2} = \mathrm{Var}(\mathbf{r}_{1kl_1}^{\mathrm{T}}\mathbf{w}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{w})$, from the equation (A.76), we have

$$\begin{aligned}
& \mathrm{Var}(\mathbf{r}_{1kl_1}^{\mathrm{T}}\mathbf{w}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{w}) \\
=\ & \mathrm{E}(\mathbf{w}^{\mathrm{T}}\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{w} - \mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{r}_{1kl_1})^2 \\
=\ & \mathrm{tr}(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}) + \mathrm{tr}(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{r}_{1kl_2}\mathbf{r}_{1kl_1}^{\mathrm{T}}) + \beta_w\sum_{i=1}^{p}(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}})_{ii}^2 \\
=\ & \sigma_{1kl_1l_2}^2 + \sigma_{1kl_1l_1}\sigma_{1kl_2l_2} + \beta_w\sum_{i=1}^{p}(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}})_{ii}^2,
\end{aligned}$$

where $(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}})_{ii}$ denotes the ith diagonal element of matrix $\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}$ and $\beta_w = \mathrm{E}w_1^4 - 3 \geq 0$. Therefore, we have

$$\min_{1\leq l_1\leq l_2\leq p}\frac{\theta_{1kl_1l_2}}{\sigma_{1kl_1l_1}\sigma_{1kl_2l_2}} \geq 1.$$

Similarly, denoting $\theta_{2kl_1l_2} = \mathrm{Var}(\mathbf{r}_{2kl_1}^{\mathrm{T}}\mathbf{w}\mathbf{r}_{2kl_2}^{\mathrm{T}}\mathbf{w})$, we also have

$$\min_{1\leq l_1\leq l_2\leq p}\frac{\theta_{2kl_1l_2}}{\sigma_{2kl_1l_1}\sigma_{2kl_2l_2}} \geq 1.$$

(3) When $\{w_1,\cdots,w_p\}$ are *i.i.d.* from $N(0,1)$, we will prove that $\mathbf{\Sigma}_{1k}^{1/2}\mathbf{w}$ satisfies Condition (C3). When $\{w_1,\cdots,w_p\}$ are *i.i.d.* from $Gamma(4,2) - 2$, we show that $\mathbf{\Sigma}_{1k}^{1/2}\mathbf{w}$ does not satisfy Condition (C3).

For any $l_1, l_2, l_3, l_4 \in \{1, 2, \cdots, p\}$, from the equation (A.76), we have

$$\begin{aligned}
&\mathrm{E}(\mathbf{r}_{1kl_1}^{\mathrm{T}}\mathbf{w}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{w}\mathbf{r}_{1kl_3}^{\mathrm{T}}\mathbf{w}\mathbf{r}_{1kl_4}^{\mathrm{T}}\mathbf{w}) \\
= \;& \mathrm{E}(\mathbf{w}^{\mathrm{T}}\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{r}_{1kl_3}\mathbf{r}_{1kl_4}^{\mathrm{T}}\mathbf{w}) \\
= \;& \mathrm{tr}(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}})\mathrm{tr}(\mathbf{r}_{1kl_3}\mathbf{r}_{1kl_4}^{\mathrm{T}}) + \mathrm{tr}(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{r}_{1kl_3}\mathbf{r}_{1kl_4}^{\mathrm{T}}) \\
& + \mathrm{tr}(\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}}\mathbf{r}_{1kl_4}\mathbf{r}_{1kl_3}^{\mathrm{T}}) + \beta_w \sum_{i=1}^{p} (\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}})_{ii}(\mathbf{r}_{1kl_3}\mathbf{r}_{1kl_4}^{\mathrm{T}})_{ii} \\
= \;& \sigma_{1kl_1l_2}\sigma_{1kl_3l_4} + \sigma_{1kl_2l_3}\sigma_{1kl_1l_4} + \sigma_{1kl_1l_3}\sigma_{1kl_2l_4} \\
& + \beta_w \sum_{i=1}^{p} (\mathbf{r}_{1kl_1}\mathbf{r}_{1kl_2}^{\mathrm{T}})_{ii}(\mathbf{r}_{1kl_3}\mathbf{r}_{1kl_4}^{\mathrm{T}})_{ii}.
\end{aligned}$$

- When $\{w_1, \cdots, w_p\}$ are *i.i.d.* from $N(0, 1)$, we have $\beta_w = 0$. That is, Condition (C3) is satisfied;
- When $\{w_1, \cdots, w_p\}$ are *i.i.d.* from $Gamma(4, 2) - 2$, we have $\beta_w = 1.5$. That is, Condition (C3) is not satisfied.

However, by Proposition 1 in [5], under H_0 and Condition (C2) or (C2*), we have

$$\mathrm{P}(T_{\mathrm{x}} - 4\log p + \log\log p \geq c_\alpha) \leq -\log(1-\alpha) + o(1),$$

where c_α is the $1-\alpha$ quantile of the Type I extreme value distribution. It indicates that even without Conditions (C1) and (C3), the test based on T_{x} can effectively control the Type I error. This is the reason that when $\{w_1, \cdots, w_p\}$ are *i.i.d.* from $Gamma(4, 2) - 2$ with Condition (C3) being violated, the empirical test sizes are still satisfactory.

A.3. Proofs of some theorems, propositions and lemmas

This section is divided into three subsections. Subsection A.3.1 gives some preparations. Subsection A.3.2 contains the proofs of theorems and propositions. Lemmas used in this paper are placed in Subsection A.3.3.

A.3.1. Preparatory works for proving Theorems 2.1, 3.1 and 3.2

The preparatory works include two steps. The skeletons of the two steps are as follows:

Skeleton of Step 1. Define the linear combination T of $T_{\mathrm{d}} - \mu_0$ and T_{r} as

$$T = \omega_1(T_{\mathrm{d}} - \mu_0) + \omega_2 T_{\mathrm{r}},$$

where ω_1 and ω_2 are two arbitrary constants. In this step, we will obtain

$$T = T_A + T_B + \omega_1\mu_1 + \omega_2\mu_2 + o_p(1), \tag{A.2}$$

where T_A and T_B are given in (A.14) and (A.15) and μ_1 and μ_2 are given in (A.12) and (A.13).

Skeleton of Step 2. Let $\mathrm{E}_0(\cdot)$ denote expectation, $\mathrm{E}_j(\cdot)$ denote the conditional expectation with respect to the σ-field generated by $\mathbf{w}_{21}, \cdots, \mathbf{w}_{2j}$ and $\mathrm{E}_{2,i}(\cdot)$ denote the conditional expectation with respect to the σ-field generated by $\mathbf{w}_{21}, \cdots, \mathbf{w}_{2N_2}, \mathbf{w}_{11}, \cdots, \mathbf{w}_{1i}$. Then we have

$$T_A + T_B = \sum_{j=1}^{N_2} (\mathrm{E}_j - \mathrm{E}_{j-1}) T_A + \sum_{i=1}^{N_1} (\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1}) T_B,$$

that is, $T_A + T_B$ is a sum of the martingale difference sequence

$$\{(\mathrm{E}_j - \mathrm{E}_{j-1})T_A, (\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})T_B, j = 1, \cdots, N_2, i = 1, \cdots, N_1\}.$$

Therefore, based on the central limit theorem of martingale difference sequences, it suffices to consider

$$\sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})T_A)^2] + \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})T_B)^2].$$

Routine algebra shows that

$$\begin{aligned} & \sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})T_A)^2] + \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})T_B)^2] \\ = \;& \omega_1^2 \sigma_{11p} + 2\omega_1\omega_2(\sigma_{12p}^1 + \sigma_{12p}^2) + \omega_2^2(\sigma_{22p}^1 + \sigma_{22p}^2) + o_p(1), \end{aligned}$$

where σ_{11p}, σ_{12p}^1, σ_{12p}^2, σ_{22p}^1 and σ_{22p}^2 are given in (A.24), (A.25), (A.26), (A.27) and (A.28), respectively. Under the conditions of Theorems 2.1, 3.1 and 3.2, based on the expressions of σ_{11p}, σ_{12p}^1, σ_{12p}^2, σ_{22p}^1 and σ_{22p}^2, we will prove the Theorems 2.1, 3.1 and 3.2.

Step 1: Letting ω_1 and ω_2 be two arbitrary constants, define

$$\begin{aligned} T &= \omega_1(T_\mathrm{d} - \mu_0) + \omega_2 T_\mathrm{r} \\ &= \omega_1[\mathrm{tr}\mathbf{S}_1^2 - 2\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2) + \mathrm{tr}\mathbf{S}_2^2 - n_1^{-1}\mathrm{tr}^2\mathbf{S}_1 - n_2^{-1}\mathrm{tr}^2\mathbf{S}_2] \\ &\quad + \omega_2[\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - 2\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) + p] \\ &= \omega_1[\mathrm{tr}\mathbf{S}_1^2 - 2\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2) + \mathrm{tr}\mathbf{S}_2^2 - n_1^{-1}(\mathrm{tr}^2\mathbf{S}_1 - \mathrm{tr}^2\boldsymbol{\Sigma}_1) - n_1^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_1 \\ &\qquad + n_2^{-1}(\mathrm{tr}^2\mathbf{S}_2 - \mathrm{tr}^2\boldsymbol{\Sigma}_2) - n_2^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_2] \\ &\quad + \omega_2[\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - 2\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) + p]. \end{aligned}$$

Denote the noncentralized sample covariance matrices as

$$\mathbf{B}_k = N_k^{-1} \sum_{i=1}^{N_k} \boldsymbol{\Sigma}_k^{1/2} \mathbf{w}_{ki} \mathbf{w}_{ki}^\mathrm{T} \boldsymbol{\Sigma}_k^{1/2}, \quad k = 1, 2.$$

Letting $\boldsymbol{\gamma}_{ki} = (1/\sqrt{N_k})\mathbf{w}_{ki}$, we have

$$\mathbf{S}_k = \mathbf{B}_k - n_k^{-1} \sum_{i \neq j} \boldsymbol{\Sigma}_k^{1/2} \boldsymbol{\gamma}_{ki} \boldsymbol{\gamma}_{kj}^\mathrm{T} \boldsymbol{\Sigma}_k^{1/2}, \quad k = 1, 2. \tag{A.3}$$

Based on (A.3), we obtain $\mathrm{tr}\mathbf{S}_1 = \mathrm{tr}\mathbf{B}_1 - n_1^{-1}\sum_{i\neq j}\boldsymbol{\gamma}_{1j}^{\mathrm{T}}\boldsymbol{\Sigma}_1\boldsymbol{\gamma}_{1i}$. Under Assumptions A-B, from (A.76), we have $n_1^{-1}\sum_{i\neq j}\boldsymbol{\gamma}_{1j}^{\mathrm{T}}\boldsymbol{\Sigma}_1\boldsymbol{\gamma}_{1i} = o_p(1)$ and

$$\mathrm{E}(\mathrm{tr}\mathbf{B}_1 - \mathrm{tr}\boldsymbol{\Sigma}_1)^2 = N_1^{-1}[2\mathrm{tr}\boldsymbol{\Sigma}_1^2 + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1\circ\boldsymbol{\Sigma}_1)], \tag{A.4}$$

which yields that $n_1^{-1}(\mathrm{tr}\mathbf{S}_1 - \mathrm{tr}\boldsymbol{\Sigma}_1)^2 = o_p(1)$. Then we have

$$n_1^{-1}(\mathrm{tr}^2\mathbf{S}_1 - \mathrm{tr}^2\boldsymbol{\Sigma}_1) = 2n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1(\mathrm{tr}\mathbf{S}_1 - \mathrm{tr}\boldsymbol{\Sigma}_1) + o_p(1).$$

Similarly, we also have

$$n_2^{-1}(\mathrm{tr}^2\mathbf{S}_2 - \mathrm{tr}^2\boldsymbol{\Sigma}_2) = 2n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2(\mathrm{tr}\mathbf{S}_2 - \mathrm{tr}\boldsymbol{\Sigma}_2) + o_p(1).$$

Let

$$\begin{aligned} T^* \;=\;& \omega_1[\mathrm{tr}\mathbf{S}_1^2 - 2\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2) + \mathrm{tr}\mathbf{S}_2^2 - 2n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1(\mathrm{tr}\mathbf{S}_1 - \mathrm{tr}\boldsymbol{\Sigma}_1) \\ & -n_1^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_1 - 2n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2(\mathrm{tr}\mathbf{S}_2 - \mathrm{tr}\boldsymbol{\Sigma}_2) - n_2^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_2] \\ & +\omega_2[\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - 2\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) + p]. \end{aligned}$$

Rewrite

$$T^* = T^* - \mathrm{E}(T^*|\mathbf{S}_2) + \mathrm{E}(T^*|\mathbf{S}_2),$$

where

$$\begin{aligned} \mathrm{E}(T^*|\mathbf{S}_2) \;=\;& \omega_1[\mathrm{E}\mathrm{tr}\mathbf{S}_1^2 - 2\mathrm{tr}(\mathbf{S}_2\boldsymbol{\Sigma}_1) + \mathrm{tr}\mathbf{S}_2^2 - n_1^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_1 \\ & -2n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2(\mathrm{tr}\mathbf{S}_2 - \mathrm{tr}\boldsymbol{\Sigma}_2) - n_2^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_2] \\ & +\omega_2\big[\mathrm{E}\big(\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2|\mathbf{S}_2\big) - 2\mathrm{E}\big(\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})|\mathbf{S}_2\big) + p\big] \end{aligned}$$

and

$$\begin{aligned} T^* - \mathrm{E}(T^*|\mathbf{S}_2) \;=\;& \omega_1[\mathrm{tr}\mathbf{S}_1^2 - \mathrm{E}\mathrm{tr}\mathbf{S}_1^2 - 2(\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2) - \mathrm{tr}(\mathbf{S}_2\boldsymbol{\Sigma}_1)) \\ & -2n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1(\mathrm{tr}\mathbf{S}_1 - \mathrm{tr}\boldsymbol{\Sigma}_1)] \\ & +\omega_2\big[\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2 - \mathrm{E}\big(\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})^2|\mathbf{S}_2\big) \\ & -2\big(\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1}) - \mathrm{E}\big(\mathrm{tr}(\mathbf{S}_1\mathbf{S}_2^{-1})|\mathbf{S}_2\big)\big)\big]. \end{aligned}$$

From (A.3) and (A.78), we have

$$\mathrm{E}\mathrm{tr}\mathbf{S}_1^2 = n_1^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_1 + \mathrm{tr}\boldsymbol{\Sigma}_1^2 + n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1^2 + \beta_1 n_1^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\circ\boldsymbol{\Sigma}_1) + o(1) \tag{A.5}$$

and

$$\mathrm{E}\mathrm{tr}\mathbf{S}_2^2 = n_2^{-1}\mathrm{tr}^2\boldsymbol{\Sigma}_2 + \mathrm{tr}\boldsymbol{\Sigma}_2^2 + n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2^2 + \beta_2 n_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_2\circ\boldsymbol{\Sigma}_2) + o(1). \tag{A.6}$$

When $\boldsymbol{\Sigma}_1$ is invertible, we denote $\boldsymbol{\Gamma} = \boldsymbol{\Sigma}_1^{-1/2}\boldsymbol{\Sigma}_2^{1/2}$ and $\widetilde{\mathbf{S}}_2 = \boldsymbol{\Sigma}_1^{-1/2}\mathbf{S}_2\boldsymbol{\Sigma}_1^{-1/2}$. Under Assumptions A-B, if $\mathbf{T}_p = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}$ and the ESD H_p of $\mathbf{T}_p$ satisfies the Assumptions c-d-f in Subsection A.3.3, we obtain that $\|\widetilde{\mathbf{S}}_2^{-1}\|$ is almost everywhere bounded. Therefore, from (A.5) and (A.6), we have

$$\begin{aligned}
\mathrm{E}(T^*|\mathbf{S}_2) &= \omega_1[\mathrm{tr}\mathbf{S}_2^2 - \mathrm{Etr}\mathbf{S}_2^2 - 2(\mathrm{tr}(\mathbf{S}_2\mathbf{\Sigma}_1) - \mathrm{tr}(\mathbf{\Sigma}_2\mathbf{\Sigma}_1)) \\
&\quad -2n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2(\mathrm{tr}\mathbf{S}_2 - \mathrm{tr}\mathbf{\Sigma}_2) + \mathrm{Etr}\mathbf{S}_1^2 + \mathrm{Etr}\mathbf{S}_2^2 \\
&\quad -2\mathrm{tr}(\mathbf{\Sigma}_2\mathbf{\Sigma}_1) - n_1^{-1}\mathrm{tr}^2\mathbf{\Sigma}_1 - n_2^{-1}\mathrm{tr}^2\mathbf{\Sigma}_2] \\
&\quad +\omega_2[n_1^{-1}\mathrm{tr}^2(\mathbf{S}_2^{-1}\mathbf{\Sigma}_1) + \mathrm{tr}(\mathbf{S}_2^{-1}\mathbf{\Sigma}_1)^2 \\
&\quad -2\mathrm{tr}(\mathbf{S}_2^{-1}\mathbf{\Sigma}_1) + n_1^{-1}\mathrm{tr}(\mathbf{S}_2^{-1}\mathbf{\Sigma}_1)^2 \\
&\quad +\beta_1 n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{S}_2^{-1}\mathbf{\Sigma}_1^{1/2} \circ \mathbf{\Sigma}_1^{1/2}\mathbf{S}_2^{-1}\mathbf{\Sigma}_1^{1/2}) + p] + o_p(1) \\
&= \omega_1[\mathrm{tr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1)^2 - \mathrm{Etr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1)^2 - 2(\mathrm{tr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1^2) - \mathrm{tr}(\mathbf{\Sigma}_2\mathbf{\Sigma}_1)) \\
&\quad -2n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2(\mathrm{tr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1) - \mathrm{tr}\mathbf{\Sigma}_2) + \mathrm{tr}(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2)^2 + n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1^2 \\
&\quad +n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2^2 + \beta_1 n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1 \circ \mathbf{\Sigma}_1) + \beta_2 n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2 \circ \mathbf{\Sigma}_2)] \\
&\quad +\omega_2[n_1^{-1}\mathrm{tr}^2\widetilde{\mathbf{S}}_2^{-1} - n_1^{-1}(\mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1})^2 + \mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 - \mathrm{Etr}(\widetilde{\mathbf{S}}_2^{-1})^2 \\
&\quad -2(\mathrm{tr}\widetilde{\mathbf{S}}_2^{-1} - \mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1}) + n_1^{-1}(\mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1})^2 + \mathrm{Etr}(\widetilde{\mathbf{S}}_2^{-1})^2 \\
&\quad -2\mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1} + n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 + \beta_1 n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1} \circ \widetilde{\mathbf{S}}_2^{-1}) + p] + o_p(1).
\end{aligned}$$

Based on the proof of Theorem 2.1 in [20], we know that the limiting distribution of $\mathrm{tr}\widetilde{\mathbf{S}}_2^{-1} - \mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1}$ is normal and

$$\mathrm{Etr}(\widetilde{\mathbf{S}}_2^{-1})^i = pc_i + \xi_i + o(1), \quad i = 1, 2, \tag{A.7}$$

where

$$c_i = \int f_i(x) dF^{y_{n_2}, H_p}(x) \tag{A.8}$$

and

$$\begin{aligned}
\xi_i &= -\frac{1}{2\pi\mathbf{i}} \oint_{\mathcal{C}} f_i(z) \frac{y_2 \int \underline{m}_{y_2}^3(z) t^2 (1 + t\underline{m}_{y_2}(z))^{-3} dH(t)}{[1 - y_2 \int \underline{m}_{y_2}^2(z) t^2 (1 + t\underline{m}_{y_2}(z))^{-2} dH(t)]^2} dz \\
&\quad -\frac{\beta_2}{2\pi\mathbf{i}} \oint_{\mathcal{C}} f_i(z) \frac{y_2 \int \underline{m}_{y_2}^3(z) t^2 (1 + t\underline{m}_{y_2}(z))^{-3} dH(t)}{1 - y_2 \int \underline{m}_{y_2}^2(z) t^2 (1 + t\underline{m}_{y_2}(z))^{-2} dH(t)} dz,
\end{aligned} \tag{A.9}$$

with $f_1(x) = 1/x$, $f_2(x) = 1/x^2$, $F^{y_{n_2}, H_p}$ and $\underline{m}_{y_2}(z)$ being the corresponding Marčenko-Pastur distribution of index (y_{n_2}, H_p) and Stieltjes transform of the companion *LSD* $\underline{F}^{y_2, H}$, H being the *LSD* of H_p, and $\mathcal{C}$ being closed contours in the complex plan enclosing the support of the *LSD* $F^{y_2, H}$. Thus, from (A.7), we obtain

$$n_1^{-1}\mathrm{tr}^2\widetilde{\mathbf{S}}_2^{-1} - n_1^{-1}(\mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1})^2 = 2y_{n_1}c_1(\mathrm{tr}\widetilde{\mathbf{S}}_2^{-1} - \mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1}) + o_p(1).$$

Therefore,

$$\begin{aligned}
\mathrm{E}(T^*|\mathbf{S}_2) &= \omega_1[\mathrm{tr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1)^2 - \mathrm{Etr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1)^2 - 2(\mathrm{tr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1^2) - \mathrm{tr}(\mathbf{\Sigma}_2\mathbf{\Sigma}_1)) \\
&\quad -2n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2(\mathrm{tr}(\widetilde{\mathbf{S}}_2\mathbf{\Sigma}_1) - \mathrm{tr}\mathbf{\Sigma}_2) + \mathrm{tr}(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2)^2 + n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1^2 \\
&\quad +n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2^2 + \beta_1 n_1^{-1}\mathrm{tr}(\mathbf{\Sigma}_1 \circ \mathbf{\Sigma}_1) + \beta_2 n_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_2 \circ \mathbf{\Sigma}_2)]
\end{aligned}$$

$$
\begin{aligned}
&+\omega_2[2(y_{n_1}c_1-1)(\mathrm{tr}\widetilde{\mathbf{S}}_2^{-1}-\mathrm{Etr}\widetilde{\mathbf{S}}_2^{-1})+\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1})^2-\mathrm{Etr}(\widetilde{\mathbf{S}}_2^{-1})^2\\
&\quad+p(y_{n_1}c_1^2-2c_1+c_2)+2y_{n_1}c_1\xi_1-2\xi_1+\xi_2\\
&\quad+n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1})^2+\beta_1 n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1}\circ\widetilde{\mathbf{S}}_2^{-1})+p]+o_p(1).
\end{aligned}
$$

Letting $\widetilde{\mathbf{S}}_1=N_1^{-1}\sum_{i=1}^{N_1}(\mathbf{w}_{1i}-\bar{\mathbf{w}}_1)(\mathbf{w}_{1i}-\bar{\mathbf{w}}_1)^{\mathrm{T}}$, we have

$$
\begin{aligned}
T^*-\mathrm{E}(T^*|\mathbf{S}_2) \;=\;& \omega_1[\mathrm{tr}(\widetilde{\mathbf{S}}_1\boldsymbol{\Sigma}_1)^2-\mathrm{Etr}(\widetilde{\mathbf{S}}_1\boldsymbol{\Sigma}_1)^2\\
&\quad-2(\mathrm{tr}(\widetilde{\mathbf{S}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{S}}_2\boldsymbol{\Sigma}_1)-\mathrm{tr}(\widetilde{\mathbf{S}}_2\boldsymbol{\Sigma}_1^2))\\
&\quad-2n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1(\mathrm{tr}(\widetilde{\mathbf{S}}_1\boldsymbol{\Sigma}_1)-\mathrm{tr}\boldsymbol{\Sigma}_1)]\\
&+\omega_2[\mathrm{tr}(\widetilde{\mathbf{S}}_1\widetilde{\mathbf{S}}_2^{-1})^2-\mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{S}}_1\widetilde{\mathbf{S}}_2^{-1})^2|\widetilde{\mathbf{S}}_2\big)\\
&\quad-2\big(\mathrm{tr}(\widetilde{\mathbf{S}}_1\widetilde{\mathbf{S}}_2^{-1})-\mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{S}}_1\widetilde{\mathbf{S}}_2^{-1})|\widetilde{\mathbf{S}}_2\big)\big)].
\end{aligned}
$$

Denote $\widetilde{\mathbf{B}}_1=N_1^{-1}\sum_{i=1}^{N_1}\mathbf{w}_{1i}\mathbf{w}_{1i}^{\mathrm{T}}$ and $\widetilde{\mathbf{B}}_2=N_2^{-1}\sum_{i=1}^{N_2}\boldsymbol{\Gamma}\mathbf{w}_{2i}\mathbf{w}_{2i}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}$. Using the proof of Theorem 2.1 in [20] again, we have

$$
\begin{aligned}
\mathrm{E}(T^*|\mathbf{S}_2) \;=\;& \omega_1[\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)^2-\mathrm{Etr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)^2-2(\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2)-\mathrm{tr}(\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1))\\
&\quad-2n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2(\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)-\mathrm{tr}\boldsymbol{\Sigma}_2)+\mathrm{tr}(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2)^2+n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1^2\\
&\quad+n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2^2+\beta_1 n_1^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\circ\boldsymbol{\Sigma}_1)+\beta_2 n_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_2\circ\boldsymbol{\Sigma}_2)]\\
&+\omega_2[2(y_{n_1}c_1-1)(\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1}-\mathrm{Etr}\widetilde{\mathbf{B}}_2^{-1})+\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2-\mathrm{Etr}(\widetilde{\mathbf{B}}_2^{-1})^2\\
&\quad+p(y_{n_1}c_1^2-2c_1+c_2)+2y_{n_1}c_1\xi_1-2\xi_1+\xi_2\\
&\quad+n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1})^2+\beta_1 n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1}\circ\widetilde{\mathbf{S}}_2^{-1})+p]+o_p(1)
\end{aligned}
$$

and

$$
\begin{aligned}
T^*-\mathrm{E}(T^*|\mathbf{S}_2) \;=\;& \omega_1[\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2-\mathrm{Etr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2\\
&\quad-2(\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{S}}_2\boldsymbol{\Sigma}_1)-\mathrm{tr}(\widetilde{\mathbf{S}}_2\boldsymbol{\Sigma}_1^2))\\
&\quad-2n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1(\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)-\mathrm{tr}\boldsymbol{\Sigma}_1)]\\
&+\omega_2[\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})^2-\mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})^2|\widetilde{\mathbf{S}}_2\big)\\
&\quad-2\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})-\mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})|\widetilde{\mathbf{S}}_2\big)\big)]+o_p(1).
\end{aligned}
$$

According to the discussion following Assumptions a-b-c-d-f in Subsection A.3.3, we know that $\|\widetilde{\mathbf{B}}_2\|$ and $\|\widetilde{\mathbf{B}}_2^{-1}\|$ are almost everywhere bounded. Since

$$
\widetilde{\mathbf{S}}_2=\frac{N_2}{n_2}\big(\widetilde{\mathbf{B}}_2-N_2\boldsymbol{\Gamma}\bar{\boldsymbol{\gamma}}_2\bar{\boldsymbol{\gamma}}_2^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\big) \tag{A.10}
$$

and

$$
\widetilde{\mathbf{S}}_2^{-1}=\frac{n_2}{N_2}\Big(\widetilde{\mathbf{B}}_2^{-1}+\frac{1}{1-N_2\bar{\boldsymbol{\gamma}}_2^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Gamma}\bar{\boldsymbol{\gamma}}_2}N_2\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Gamma}\bar{\boldsymbol{\gamma}}_2\bar{\boldsymbol{\gamma}}_2^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\widetilde{\mathbf{B}}_2^{-1}\Big), \tag{A.11}
$$

based on (A.76), (A.78), Lemma 5.2 and Lemma 5.3 of [20], we have

$$
\begin{aligned}
\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{S}}_2\boldsymbol{\Sigma}_1)-\mathrm{tr}(\widetilde{\mathbf{S}}_2\boldsymbol{\Sigma}_1^2)&=\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)-\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2)+o_p(1),\\
\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})-\mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})|\widetilde{\mathbf{S}}_2\big)&=\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})-\mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})|\widetilde{\mathbf{B}}_2\big)+o_p(1)
\end{aligned}
$$

and

$$\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})^2 - \mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{S}}_2^{-1})^2|\widetilde{\mathbf{S}}_2\big) = \mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2 - \mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2|\widetilde{\mathbf{B}}_2\big) + o_p(1).$$

Therefore, we get

$$\begin{aligned} T^* - \mathrm{E}(T^*|\mathbf{S}_2) \quad = \quad & \omega_1[\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2 - \mathrm{Etr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2 \\ & \quad -2(\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) - \mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2)) \\ & \quad -2n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1(\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1) - \mathrm{tr}\boldsymbol{\Sigma}_1)] \\ & +\omega_2[\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1}) - \mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})|\widetilde{\mathbf{B}}_2\big) \\ & \quad -2\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2 - \mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2|\widetilde{\mathbf{B}}_2\big)\big)] + o_p(1). \end{aligned}$$

Denote

$$\begin{aligned} \mu_1 \quad = \quad & \mathrm{tr}(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)^2 + n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1^2 + n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2^2 \qquad \text{(A.12)} \\ & +\beta_1 n_1^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1) + \beta_2 n_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_2 \circ \boldsymbol{\Sigma}_2), \\ \mu_2 \quad = \quad & p(y_{n_1}c_1^2 - 2c_1 + c_2) + p + 2y_{n_1}c_1\xi_1 - 2\xi_1 + \xi_2 \qquad \text{(A.13)} \\ & +n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1})^2 + \beta_1 n_1^{-1}\mathrm{tr}(\widetilde{\mathbf{S}}_2^{-1} \circ \widetilde{\mathbf{S}}_2^{-1}), \\ T_A \quad = \quad & \omega_1[\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)^2 - \mathrm{Etr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)^2 - 2(\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2) - \mathrm{tr}(\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1)) \qquad \text{(A.14)} \\ & \quad -2n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2(\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) - \mathrm{tr}\boldsymbol{\Sigma}_2)] \\ & +\omega_2[2(y_{n_1}c_1 - 1)(\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1} - \mathrm{Etr}\widetilde{\mathbf{B}}_2^{-1}) + \mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 - \mathrm{Etr}(\widetilde{\mathbf{B}}_2^{-1})^2] \end{aligned}$$

and

$$\begin{aligned} T_B \quad = \quad & \omega_1[\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2 - \mathrm{Etr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2 - 2(\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) - \mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2)) \\ & \quad -2n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1(\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1) - \mathrm{tr}\boldsymbol{\Sigma}_1)] \qquad \text{(A.15)} \\ & +\omega_2[\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2 - \mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2|\widetilde{\mathbf{B}}_2\big) \\ & \quad -2\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1}) - \mathrm{E}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})|\widetilde{\mathbf{B}}_2\big)\big)]. \end{aligned}$$

Therefore, we have

$$T = T_A + T_B + \omega_1\mu_1 + \omega_2\mu_2 + o_p(1).$$

Based on Slutsky's Theorem, we only need to derive the limiting distribution of $T_A + T_B$.

Step 2: We have

$$T_A + T_B = \sum_{j=1}^{N_2}(\mathrm{E}_j - \mathrm{E}_{j-1})T_A + \sum_{i=1}^{N_1}(\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})T_B,$$

where

$$\begin{aligned}
&\sum_{j=1}^{N_2}(\mathrm{E}_j-\mathrm{E}_{j-1})T_A\\
=\;& \omega_1\big[\sum_{j=1}^{N_2}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)^2-2\sum_{j=1}^{N_2}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1^2)\\
&-2n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2\sum_{j=1}^{N_2}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)\big]\\
&+\omega_2\big[2(y_{n_1}c_1-1)\sum_{j=1}^{N_2}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1}+\sum_{j=1}^{N_2}(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2\big]
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{i=1}^{N_1}(\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})T_B\\
=\;& \omega_1\big[\sum_{i=1}^{N_1}(\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\mathbf{\Sigma}_1)^2-2\sum_{i=1}^{N_1}(\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)\\
&-2n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1\sum_{i=1}^{N_1}(\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\mathbf{\Sigma}_1)\big]\\
&+\omega_2\big[\sum_{i=1}^{N_1}(\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2-2\sum_{i=1}^{N_1}(\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})\big].
\end{aligned}$$

Based on the central limit theorem of martingale difference sequences, it suffices to consider

$$\sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})T_A)^2]+\sum_{i=1}^{N_1}\mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})T_B)^2].$$

From (A.71)-(A.74) in Lemma A.1, (A.88)-(A.91) in Lemma A.4 and (2.5) in the Theorem 2.1 of [20], we obtain

$$\begin{aligned}
\sigma_A \;=\;& \sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})T_A)^2]\\
=\;& \omega_1^2\big[4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)^2\eta_{33}+4\eta_{44}+\eta_{55}\\
&+8(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\eta_{34}-4(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\eta_{35}-4\eta_{45}\big]\\
&+2\omega_1\omega_2\big[-4(y_{n_1}c_1-1)(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\eta_{13}-4(y_{n_1}c_1-1)\eta_{14}\\
&+2(y_{n_1}c_1-1)\eta_{15}-2(n_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2)\eta_{23}-2\eta_{24}+\eta_{25}\big]\\
&+\omega_2^2\big[4(y_{n_1}c_1-1)^2\eta_{11}+\eta_{22}+4(y_{n_1}c_1-1)\eta_{12}\big],
\end{aligned}$$

when $i,j\in\{1,2\}$ and $p\to\infty$,

$$\begin{aligned}\eta_{ij} &\to -\frac{1}{2\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}\frac{f_i(z_1)f_j(z_2)}{(\underline{m}_{y_2}(z_1)-\underline{m}_{y_2}(z_2))^2}d\underline{m}_{y_2}(z_1)d\underline{m}_{y_2}(z_2) \\ &\quad -\frac{y_2\beta_2}{4\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2}f_i(z_1)f_j(z_2)\Big[\int\frac{t^2dH(t)}{(\underline{m}_{y_2}(z_1)t+1)^2(\underline{m}_{y_2}(z_2)t+1)^2}\Big] \\ &\qquad \times d\underline{m}_{y_2}(z_1)d\underline{m}_{y_2}(z_2),\end{aligned}\tag{A.16}$$

where $f_1(x)=1/x$ and $f_2(x)=1/x^2$, $\mathcal{C}_1$ and $\mathcal{C}_2$ are closed contours in the complex plan enclosing the support of the LSD $F^{y_2,H}$, and $\mathcal{C}_1$ and $\mathcal{C}_2$ are nonoverlapping,

$$\begin{aligned}
\eta_{33} &= \sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1))^2]=\frac{1}{N_2}[2\mathrm{tr}\mathbf{\Sigma}_2^2+\beta_2\mathrm{tr}(\mathbf{\Sigma}_2\circ\mathbf{\Sigma}_2)],\\
\eta_{44} &= \sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1^2))^2]\\
&= \frac{1}{N_2}[2\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)^2+\beta_2\mathrm{tr}(\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2}\circ\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2})],\\
\eta_{55} &= \sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)^2)^2]\\
&= \frac{4}{N_2^3}\mathrm{tr}^2\mathbf{\Sigma}_2[2\mathrm{tr}\mathbf{\Sigma}_2^2+\beta_2\mathrm{tr}(\mathbf{\Sigma}_2\circ\mathbf{\Sigma}_2)]+\frac{8}{N_2^2}\mathrm{tr}\mathbf{\Sigma}_2[2\mathrm{tr}\mathbf{\Sigma}_2^3+\beta_2\mathrm{tr}(\mathbf{\Sigma}_2\circ\mathbf{\Sigma}_2^2)]\\
&\quad +\frac{4}{N_2}[2\mathrm{tr}\mathbf{\Sigma}_2^4+\beta_2\mathrm{tr}(\mathbf{\Sigma}_2^2\circ\mathbf{\Sigma}_2^2)]+\frac{4}{N_2^2}\mathrm{tr}^2\mathbf{\Sigma}_2^2+o_p(1),\\
\eta_{13} &= \sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1})((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1))]\\
&= -\frac{2}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}\mathbf{\Sigma}_1-\frac{\beta_2}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_2)+o_p(1),\\
\eta_{14} &= \sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1})((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1^2))]\\
&= -\frac{2}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}\mathbf{\Sigma}_1^2-\frac{\beta_2}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2})\\
&\quad +o_p(1),\\
\eta_{15} &= \sum_{j=1}^{N_2}\mathrm{E}_{j-1}[((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1})((\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)^2)]\\
&= -\frac{4}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)-\frac{2\beta_2}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_2^2)\\
&\quad -\frac{2\beta_2}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}\mathbf{\Sigma}_2N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_2)+o_p(1),
\end{aligned}$$

$$
\begin{aligned}
\eta_{23} &= \sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2)((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1))] \\
&= -\frac{4}{(1-y_{N_2})^2} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^2\boldsymbol{\Sigma}_2^{-1}) - \frac{4}{(1-y_{N_2})^3} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1 \\
&\quad -\frac{2\beta_2}{(1-y_{N_2})^2} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2) \\
&\quad -\frac{2\beta_2}{(1-y_{N_2})^3} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2) + o_p(1), \\
\eta_{24} &= \sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2)((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2))] \\
&= -\frac{4}{(1-y_{N_2})^2} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^3\boldsymbol{\Sigma}_2^{-1}) - \frac{4}{(1-y_{N_2})^3} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1^2 \\
&\quad -\frac{2\beta_2}{(1-y_{N_2})^2} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2}) \\
&\quad -\frac{2\beta_2}{(1-y_{N_2})^3} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2}) \\
&\quad + o_p(1), \\
\eta_{25} &= \sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2)((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)^2)] \\
&= -\frac{8}{(1-y_{N_2})^2} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2) \\
&\quad -\frac{8}{(1-y_{N_2})^3} N_2^{-2}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2) \\
&\quad +\frac{4}{(1-y_{N_2})^2} (N_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1)^2 \\
&\quad -\frac{4\beta_2}{(1-y_{N_2})^2} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2^2) \\
&\quad -\frac{4\beta_2}{(1-y_{N_2})^3} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2^2) \\
&\quad -\frac{4\beta_2}{(1-y_{N_2})^2} N_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2 N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2) \\
&\quad -\frac{4\beta_2}{(1-y_{N_2})^3} N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2 N_2^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2) \\
&\quad + o_p(1), \\
\eta_{34} &= \sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1))((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^2))] \\
&= \frac{1}{N_2}[2\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^2) + \beta_2\mathrm{tr}(\boldsymbol{\Sigma}_2\circ\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2})] + o_p(1),
\end{aligned}
$$

$$\begin{aligned}
\eta_{35} &= \sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1))((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)^2)] \\
&= \frac{2}{N_2^2}\mathrm{tr}\mathbf{\Sigma}_2[2\mathrm{tr}\mathbf{\Sigma}_2^2 + \beta_2\mathrm{tr}(\mathbf{\Sigma}_2 \circ \mathbf{\Sigma}_2)] + \frac{2}{N_2}[2\mathrm{tr}\mathbf{\Sigma}_2^3 + \beta_2\mathrm{tr}(\mathbf{\Sigma}_2 \circ \mathbf{\Sigma}_2^2)] + o_p(1), \\
\eta_{45} &= \sum_{j=1}^{N_2} \mathrm{E}_{j-1}[((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1^2))((\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)^2)] \\
&= \frac{2}{N_2^2}\mathrm{tr}\mathbf{\Sigma}_2[2\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^2) + \beta_2\mathrm{tr}(\mathbf{\Sigma}_2 \circ \mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2})] \\
&\quad + \frac{2}{N_2}[2\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^3) + \beta_2\mathrm{tr}(\mathbf{\Sigma}_2^2 \circ \mathbf{\Sigma}_2^{1/2}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{1/2})] + o_p(1).
\end{aligned}$$

From (A.81), (A.82), (A.83) and (A.84) in Lemma A.2, we have

$$\begin{aligned}
p^{-1}\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1} &= d_{1p} + o_p(1), \qquad p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 = d_{2p} + o_p(1), \\
p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^3 &= d_{3p} + o_p(1), \qquad p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^4 = d_{4p} + o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \widetilde{\mathbf{B}}_2^{-1}) &= l_{1p} + o_p(1), \\
p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ (\widetilde{\mathbf{B}}_2^{-1})^2) &= l_{2p} + o_p(1), \\
p^{-1}\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ (\widetilde{\mathbf{B}}_2^{-1})^2) &= l_{3p} + o_p(1),
\end{aligned}$$

where

$$d_{1p} = \frac{1}{(1-y_{N_2})}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}), \tag{A.17}$$

$$d_{2p} = \frac{1}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})^2 + \frac{y_{N_2}}{(1-y_{N_2})^3}(p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}))^2, \tag{A.18}$$

$$\begin{aligned}
d_{3p} = {} & \frac{1}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})^3 \\
& + \frac{3y_{N_2}}{(1-y_{N_2})^4}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})^2 \\
& + \frac{2y_{N_2}^2}{(1-y_{N_2})^5}(p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}))^3,
\end{aligned} \tag{A.19}$$

$$\begin{aligned}
d_{4p} = {} & \frac{1}{(1-y_{N_2})^4}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})^4 \\
& + \frac{4y_{N_2}}{(1-y_{N_2})^5}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})^3 \\
& + \frac{2y_{N_2}}{(1-y_{N_2})^5}(p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})^2)^2
\end{aligned} \tag{A.20}$$

$$+\frac{10y_{N_2}^2}{(1-y_{N_2})^6}(p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}))^2p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})^2$$
$$+\frac{5y_{N_2}^3}{(1-y_{N_2})^7}(p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}))^4$$

and

$$l_{1p}=\frac{1}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}), \tag{A.21}$$
$$l_{2p}=\frac{1}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}) \tag{A.22}$$
$$+\frac{y_{N_2}}{(1-y_{N_2})^4}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}),$$
$$l_{3p}=\frac{1}{(1-y_{N_2})^4}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}) \tag{A.23}$$
$$+\frac{2y_{N_2}}{(1-y_{N_2})^5}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2})$$
$$+\frac{y_{N_2}^2}{(1-y_{N_2})^6}(p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}))^2p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}).$$

Then from (A.71)-(A.75) in Lemma A.1, we have

$$\begin{aligned}\sigma_B&=\sum_{i=1}^{N_1}\mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})T_B)^2]\\&=\omega_1^2\big[4(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)^2\theta_{11}+4\theta_{33}+\theta_{55}+8(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)\theta_{13}-4(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)\theta_{15}-4\theta_{35}\big]\\&\quad+2\omega_1\omega_2\big[4(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)\theta_{12}-2(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)\theta_{14}+4\theta_{23}-2\theta_{25}-2\theta_{34}+\theta_{45}\big]\\&\quad+\omega_2^2\big[4\theta_{22}+\theta_{44}-4\theta_{24}\big],\end{aligned}$$

where

$$\theta_{11}=\sum_{i=1}^{N_1}\mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\mathbf{\Sigma}_1))^2]=\frac{1}{N_1}[2\mathrm{tr}\mathbf{\Sigma}_1^2+\beta_1\mathrm{tr}(\mathbf{\Sigma}_1\circ\mathbf{\Sigma}_1)],$$
$$\begin{aligned}\theta_{22}&=\sum_{i=1}^{N_1}\mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1}))^2]\\&=\frac{1}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2+\beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\widetilde{\mathbf{B}}_2^{-1})]=y_{N_1}(2d_{2p}+\beta_1l_{1p})+o_p(1),\end{aligned}$$
$$\begin{aligned}\theta_{33}&=\sum_{i=1}^{N_1}\mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1))^2]\\&=\frac{1}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1^2)^2+\beta_1\mathrm{tr}(\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1\circ\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)]\\&=\frac{2}{N_1N_2}\mathrm{tr}^2(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)+\frac{2}{N_1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2)^2\end{aligned}$$

$$
\begin{aligned}
&\quad + \frac{\beta_1}{N_1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2} \circ \boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2}) + o_p(1),\\
\theta_{44} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2)^2]\\
&= \frac{4}{N_1^3}\mathrm{tr}^2(\widetilde{\mathbf{B}}_2^{-1})[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \widetilde{\mathbf{B}}_2^{-1})] + \frac{4}{N_1^2}\mathrm{tr}^2(\widetilde{\mathbf{B}}_2^{-1})^2\\
&\quad + \frac{8}{N_1^2}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^3 + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ (\widetilde{\mathbf{B}}_2^{-1})^2)\\
&\quad + \frac{4}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^4 + \beta_1\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ (\widetilde{\mathbf{B}}_2^{-1})^2)]\\
&= 4y_{N_1}^3 d_{1p}^2(2d_{2p} + \beta_1 l_{1p}) + 4y_{N_1}^2 d_{2p}^2 + 8y_{N_1}^2 d_{1p}(2d_{3p} + \beta_1 l_{2p})\\
&\quad + 4y_{N_1}(2d_{4p} + \beta_1 l_{3p}) + o_p(1),\\
\theta_{55} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2)^2]\\
&= \frac{4}{N_1^3}\mathrm{tr}^2\boldsymbol{\Sigma}_1[2\mathrm{tr}\boldsymbol{\Sigma}_1^2 + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1)] + \frac{8}{N_1^2}\mathrm{tr}\boldsymbol{\Sigma}_1[2\mathrm{tr}\boldsymbol{\Sigma}_1^3 + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1^2)]\\
&\quad + \frac{4}{N_1}[2\mathrm{tr}\boldsymbol{\Sigma}_1^4 + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1^2 \circ \boldsymbol{\Sigma}_1^2)] + \frac{4}{N_1^2}\mathrm{tr}^2\boldsymbol{\Sigma}_1^2 + o_p(1),\\
\theta_{12} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1))((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1}))]\\
&= \frac{1}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1)],\\
\theta_{13} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1))((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1))]\\
&= \frac{1}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^3) + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)]\\
&= \frac{1}{N_1}[2\mathrm{tr}(\boldsymbol{\Sigma}_1^2\boldsymbol{\Sigma}_2) + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2})] + o_p(1),\\
\theta_{14} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1))((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2)]\\
&= \frac{2}{N_1^2}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1} \circ \boldsymbol{\Sigma}_1)]\\
&\quad + \frac{2}{N_1}[2\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2 \circ \boldsymbol{\Sigma}_1)] + o_p(1),\\
\theta_{15} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1))((\mathrm{E}_{2,i} - \mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2)]\\
&= \frac{2}{N_1^2}\mathrm{tr}\boldsymbol{\Sigma}_1[2\mathrm{tr}\boldsymbol{\Sigma}_1^2 + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_2)] + \frac{2}{N_1}[2\mathrm{tr}\boldsymbol{\Sigma}_1^3 + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1 \circ \boldsymbol{\Sigma}_1^2)] + o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
\theta_{23} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1}))((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1))] \\
&= \frac{1}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)], \\
\theta_{24} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1}))((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2)] \\
&= \frac{2}{N_1^2}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2 + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\widetilde{\mathbf{B}}_2^{-1})] \\
&\quad + \frac{2}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^3 + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ(\widetilde{\mathbf{B}}_2^{-1})^2)] \\
&= 2y_{N_1}^2 d_{1p}(2d_{2p}+\beta_1 l_{1p}) + 2y_{N_1}(2d_{3p}+\beta_1 l_{2p}) + o_p(1), \\
\theta_{25} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1}))((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2)] \\
&= \frac{2}{N_1^2}\mathrm{tr}\boldsymbol{\Sigma}_1[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\boldsymbol{\Sigma}_1)] \\
&\quad + \frac{2}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1^2) + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\boldsymbol{\Sigma}_1^2)] + o_p(1), \\
\theta_{34} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1))((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2)] \\
&= \frac{2}{N_1^2}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)] \\
&\quad + \frac{2}{N_1}[2\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)] + o_p(1), \\
\theta_{35} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1))((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2)] \\
&= \frac{2}{N_1^2}\mathrm{tr}\boldsymbol{\Sigma}_1[2\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^3) + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1\circ\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)] \\
&\quad + \frac{2}{N_1}[2\mathrm{tr}(\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1^4) + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1^2\circ\boldsymbol{\Sigma}_1\widetilde{\mathbf{B}}_2\boldsymbol{\Sigma}_1)] + o_p(1) \\
&= \frac{2}{N_1^2}\mathrm{tr}\boldsymbol{\Sigma}_1[2\mathrm{tr}(\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^2) + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1\circ\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2})] \\
&\quad + \frac{2}{N_1}[2\mathrm{tr}(\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^3) + \beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1^2\circ\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2})] + o_p(1), \\
\theta_{45} &= \sum_{i=1}^{N_1} \mathrm{E}_{2,i-1}[((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\widetilde{\mathbf{B}}_2^{-1})^2)((\mathrm{E}_{2,i}-\mathrm{E}_{2,i-1})\mathrm{tr}(\widetilde{\mathbf{B}}_1\boldsymbol{\Sigma}_1)^2)] \\
&= \frac{4}{N_1^3}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})\mathrm{tr}\boldsymbol{\Sigma}_1[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1) + \beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\boldsymbol{\Sigma}_1)]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{4}{N_1^2}\mathrm{tr}\boldsymbol{\Sigma}_1[2\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\boldsymbol{\Sigma}_1)+\beta_1\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ\boldsymbol{\Sigma}_1)]\\
&+\frac{4}{N_1^2}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1^2)+\beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\boldsymbol{\Sigma}_1^2)]+\frac{4}{N_1^2}\mathrm{tr}^2(\widetilde{\mathbf{B}}_2^{-1}\boldsymbol{\Sigma}_1)\\
&+\frac{4}{N_1}[2\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\boldsymbol{\Sigma}_1^2)+\beta_1\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ\boldsymbol{\Sigma}_1^2)]\\
&+o_p(1).
\end{aligned}
$$

After sorting and calculation, we obtain

$$
\begin{aligned}
\sigma_{11p}&=4(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)^2\eta_{33}+4\eta_{44}+\eta_{55}+8(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)\eta_{34}\\
&\quad-4(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)\eta_{35}-4\eta_{45}+4(n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1)^2\theta_{11}+4\theta_{33}\\
&\quad+\theta_{55}+8(n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1)\theta_{13}-4(n_1^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1)\theta_{15}-4\theta_{35}\\
&=\frac{4}{N_1^2}\mathrm{tr}^2\boldsymbol{\Sigma}_1^2+\frac{4}{N_2^2}\mathrm{tr}^2\boldsymbol{\Sigma}_2^2+\frac{8}{N_1N_2}\mathrm{tr}^2(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)\\
&\quad+\frac{4}{N_1}[2\mathrm{tr}(\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1)\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)\\
&\qquad+\beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}(\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1)\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2})]\\
&\quad+\frac{4}{N_1}[2\mathrm{tr}(\boldsymbol{\Sigma}_1^3(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2))+\beta_1\mathrm{tr}(\boldsymbol{\Sigma}_1^2\circ\boldsymbol{\Sigma}_1^{1/2}(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_1^{1/2}]\\
&\quad+\frac{4}{N_2}[2\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_2)\\
&\qquad+\beta_2\mathrm{tr}(\boldsymbol{\Sigma}_2^{1/2}(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_2^{1/2}\circ\boldsymbol{\Sigma}_2^{1/2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{1/2})]\\
&\quad+\frac{4}{N_2}[2\mathrm{tr}((\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1)\boldsymbol{\Sigma}_2^3)+\beta_2\mathrm{tr}(\boldsymbol{\Sigma}_2^2\circ\boldsymbol{\Sigma}_2^{1/2}(\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1)\boldsymbol{\Sigma}_2^{1/2})]+o_p(1),
\end{aligned}
\tag{A.24}
$$

$$
\begin{aligned}
\sigma_{12p}^1&=-4(y_{n_1}c_1-1)(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)\eta_{13}-4(y_{n_1}c_1-1)\eta_{14}\\
&\quad+2(y_{n_1}c_1-1)\eta_{15}-2(n_2^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2)\eta_{23}-2\eta_{24}+\eta_{25}\\
&=2(y_{n_1}c_1-1)\Big[\frac{4y_{N_2}^2}{(1-y_{N_2})}p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2\\
&\qquad+\frac{4y_{N_2}}{(1-y_{N_2})}p^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2))\\
&\qquad+\frac{2\beta_2y_{N_2}}{(1-y_{N_2})}p^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1^{1/2}\circ\boldsymbol{\Sigma}_2^{1/2}(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_2^{1/2})\Big]\\
&\quad+\frac{8y_{N_2}^2}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2p^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1^2\boldsymbol{\Sigma}_2^{-1})\\
&\quad+\frac{8y_{N_2}^3}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_2p^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})\\
&\quad+\frac{4y_{N_2}^2}{(1-y_{N_2})^2}(p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1)^2+\frac{8y_{N_2}}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2))
\end{aligned}
\tag{A.25}
$$

$$
\begin{aligned}
&+\frac{8y_{N_2}^2}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2))\\
&+\frac{4\beta_2 y_{N_2}}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_2^{1/2}(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)\mathbf{\Sigma}_2^{1/2})\\
&+\frac{4\beta_2 y_{N_2}^2}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_2^{1/2}(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)\mathbf{\Sigma}_2^{1/2})\\
&+o_p(1)
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma_{12p}^2 &= 4(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)\theta_{12}-2(n_1^{-1}\mathrm{tr}\mathbf{\Sigma}_1)\theta_{14}+4\theta_{23}-2\theta_{25}-2\theta_{34}+\theta_{45}\\
&= 4(y_{N_1}d_{1p}-1)N_1^{-1}\big[2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\mathbf{\Sigma}_1^2)-2\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)\\
&\qquad +\beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\mathbf{\Sigma}_1^2)-\beta_1\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)\big]\\
&\quad +4N_1^{-1}\big[2\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\mathbf{\Sigma}_1^2)-2\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)+\beta_1\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ\mathbf{\Sigma}_1^2)\\
&\quad -\beta_1\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)\big]+4N_1^{-2}\mathrm{tr}^2(\widetilde{\mathbf{B}}_2^{-1}\mathbf{\Sigma}_1)+o_p(1).
\end{aligned}
$$

Based on (A.81) and (A.82) in Lemma A.2 and (A.86) and (A.87) in Lemma A.3, we have

$$
\begin{aligned}
N_2^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\mathbf{\Sigma}_1) &= \frac{1}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2\mathbf{\Sigma}_2^{-1})+o_p(1),\\
N_2^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\mathbf{\Sigma}_1^2) &= \frac{1}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^3\mathbf{\Sigma}_2^{-1})+o_p(1),\\
N_2^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1) &= \frac{1}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})\\
&\quad -\frac{1}{(1-y_{N_2})}(N_2^{-1}\mathrm{tr}\mathbf{\Sigma}_1)^2+o_p(1),\\
N_2^{-1}\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\mathbf{\Sigma}_1^2) &= \frac{1}{(1-y_{N_2})^2}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^3\mathbf{\Sigma}_2^{-1})\\
&\quad +\frac{1}{(1-y_{N_2})^3}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^3\mathbf{\Sigma}_2^{-1})\\
&\quad +o_p(1),\\
N_2^{-1}\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1) &= \frac{1}{(1-y_{N_2})^2}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})\\
&\quad +\frac{1}{(1-y_{N_2})^3}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})\\
&\quad -\frac{2}{(1-y_{N_2})^2}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2\mathbf{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}\mathbf{\Sigma}_1\\
&\quad -\frac{1}{(1-y_{N_2})^3}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})(N_2^{-1}\mathrm{tr}\mathbf{\Sigma}_1)^2+o_p(1)
\end{aligned}
$$

and

$$
\begin{aligned}
&N_2^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\mathbf{\Sigma}_1^2)\\
=&\frac{1}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^2)+o_p(1),\\
&N_2^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1}\circ\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)\\
=&\frac{1}{(1-y_{N_2})}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2})+o_p(1),\\
&N_2^{-1}\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ\mathbf{\Sigma}_1^2)\\
=&\frac{1}{(1-y_{N_2})^2}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^2)\\
&+\frac{1}{(1-y_{N_2})^3}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^2)+o_p(1),\\
&N_2^{-1}\mathrm{tr}((\widetilde{\mathbf{B}}_2^{-1})^2\circ\mathbf{\Sigma}_1\widetilde{\mathbf{B}}_2\mathbf{\Sigma}_1)\\
=&\frac{1}{(1-y_{N_2})^2}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2})\\
&+\frac{1}{(1-y_{N_2})^3}N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})N_2^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{1/2})+o_p(1).
\end{aligned}
$$

Therefore, we get

$$
\begin{aligned}
\sigma_{12p}^2 = 4(y_{N_1}d_{1p}-1)\Big[&\frac{2y_{N_1}y_{N_2}}{(1-y_{N_2})}(p^{-1}\mathrm{tr}\mathbf{\Sigma}_1)^2\\
&+\frac{2y_{N_1}}{(1-y_{N_2})}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})\\
&+\frac{\beta_1 y_{N_1}}{(1-y_{N_2})}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)\mathbf{\Sigma}_1^{1/2})\Big]\\
+4\Big[&\frac{4y_{N_1}y_{N_2}}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}\mathbf{\Sigma}_1\\
+&\frac{2y_{N_1}y_{N_2}^2}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})(p^{-1}\mathrm{tr}\mathbf{\Sigma}_1)^2\\
+&\frac{2y_{N_1}}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2))\\
+&\frac{2y_{N_1}y_{N_2}}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2))\\
+&\frac{\beta_1 y_{N_1}}{(1-y_{N_2})^2}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)\mathbf{\Sigma}_1^{1/2})\\
+&\frac{\beta_1 y_{N_1}y_{N_2}}{(1-y_{N_2})^3}p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1^{1/2}\circ\mathbf{\Sigma}_1^{1/2}(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)\mathbf{\Sigma}_1^{1/2})\Big]\\
+&\frac{4y_{N_1}^2}{(1-y_{N_2})^2}(p^{-1}\mathrm{tr}(\mathbf{\Sigma}_1^2\mathbf{\Sigma}_2^{-1}))^2+o_p(1).
\end{aligned}
\tag{A.26}
$$

Finally, we have

$$\sigma_{22p}^{1} = 4(y_{n_1}c_1 - 1)^2\eta_{11} + \eta_{22} + 4(y_{n_1}c_1 - 1)\eta_{12} + o_p(1) \tag{A.27}$$

and

$$\begin{aligned}\sigma_{22p}^{2} &= 4\theta_{22} + \theta_{44} - 4\theta_{24} \\ &= 4y_{N_1}(y_{N_1}d_{1p} - 1)^2(2d_{2p} + \beta_1 l_{1p}) \\ &\quad +8y_{N_1}(y_{N_1}d_{1p} - 1)(2d_{3p} + \beta_1 l_{2p}) \\ &\quad +4y_{N_1}(2d_{4p} + \beta_1 l_{3p}) + 4y_{N_1}^2 d_{2p}^2 + o_p(1).\end{aligned} \tag{A.28}$$

Next, we consider μ_2, according to (A.11) and (A.13), we have

$$\begin{aligned}\mu_2 &= p(y_{n_1}c_1^2 - 2c_1 + c_2) + p + 2y_{n_1}c_1\xi_1 - 2\xi_1 + \xi_2 \\ &\quad +y_{n_1}d_{2p} + \beta_1 y_{n_1}l_{1p} + o_p(1).\end{aligned} \tag{A.29}$$

A.3.2. Proofs of the theorems and propositions

The Proof of Theorem 2.1. Now we prove the Theorem 2.1, because all quantities are computed under H_0, we add 0 to the subscripts of these quantities, for example, we use μ_{10} instead of μ_1 in this subsection. Under H_0, assuming that $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$, from (A.12), we have

$$\mu_{10} = (y_{n_1} + y_{n_2})p^{-1}\mathrm{tr}\mathbf{\Sigma}^2 + (\beta_1 y_{n_1} + \beta_2 y_{n_2})p^{-1}\mathrm{tr}(\mathbf{\Sigma} \circ \mathbf{\Sigma}).$$

When $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$, we have $\mathbf{\Gamma} = \mathbf{\Sigma}_1^{-1/2}\mathbf{\Sigma}_2^{1/2} = \mathbf{I}_p$ and $\mathbf{T}_p = \mathbf{\Gamma}\mathbf{\Gamma}^{\mathrm{T}} = \mathbf{I}_p$, which satisfy the Assumptions c-d-f when we treat $\widetilde{\mathbf{B}}_2$ as $\mathbf{B}_n$ in Subsection A.3.3, H_p and H, the *ESD* and *LSD* of $\mathbf{T}_p$, are δ_1, the $F^{y_{n_2},H_p}$ has an explicit density function,

$$\frac{1}{2\pi x y_{n_2}}\sqrt{(b_{y_{n_2}} - x)(x - a_{y_{n_2}})}, \qquad a_{y_{n_2}} \le x \le b_{y_{n_2}}, \tag{A.30}$$

which is the seminal Marčenko-Pastur law with index y_{n_2} and support $[a_{y_{n_2}}, b_{y_{n_2}}]$, where $a_{y_{n_2}} = (1 - \sqrt{y_{n_2}})^2$, $b_{y_{n_2}} = (1 + \sqrt{y_{n_2}})^2$, from (A.8), we have

$$c_{10} = \frac{1}{(1 - y_{n_2})}, \qquad c_{20} = \frac{1}{(1 - y_{n_2})^3}. \tag{A.31}$$

Based on (A.9), (A.16) and the Proposition A.1. in the Appendix of [20], we get

$$\xi_{10} = \frac{y_{n_2}}{(1 - y_{n_2})^2} + \frac{\beta_2 y_{n_2}}{(1 - y_{n_2})}, \tag{A.32}$$

$$\xi_{20} = \frac{y_{n_2}^2 + 3y_{n_2}}{(1 - y_{n_2})^4} + \frac{\beta_2(-y_{n_2}^2 + 3y_{n_2})}{(1 - y_{n_2})^3} \tag{A.33}$$

and

$$\eta_{110} = \frac{2y_2}{(1-y_2)^4} + \frac{\beta_2 y_2}{(1-y_2)^2}, \tag{A.34}$$

$$\eta_{120} = \frac{4y_2(1+y_2)}{(1-y_2)^6} + \frac{2\beta_2 y_2}{(1-y_2)^4}, \tag{A.35}$$

$$\eta_{220} = \frac{4y_2(2y_2^2+5y_2+2)}{(1-y_2)^8} + \frac{4\beta_2 y_2}{(1-y_2)^6}. \tag{A.36}$$

From (A.17)-(A.23), we have

$$d_{10p} = \frac{1}{(1-y_{N_2})} \to \frac{1}{(1-y_2)} = d_{10}, \tag{A.37}$$

$$d_{20p} = \frac{1}{(1-y_{N_2})^3} \to \frac{1}{(1-y_2)^3} = d_{20}, \tag{A.38}$$

$$d_{30p} = \frac{1+y_{N_2}}{(1-y_{N_2})^5} \to \frac{1+y_2}{(1-y_2)^5} = d_{30}, \tag{A.39}$$

$$d_{40p} = \frac{y_{N_2}^2+3y_{N_2}+1}{(1-y_{N_2})^7} \to \frac{y_2^2+3y_2+1}{(1-y_2)^7} = d_{40} \tag{A.40}$$

and

$$l_{10p} = \frac{1}{(1-y_{N_2})^2} \to \frac{1}{(1-y_2)^2} = l_{10}, \tag{A.41}$$

$$l_{20p} = \frac{1}{(1-y_{N_2})^4} \to \frac{1}{(1-y_2)^4} = l_{20}, \tag{A.42}$$

$$l_{30p} = \frac{1}{(1-y_{N_2})^6} \to \frac{1}{(1-y_2)^6} = l_{30}. \tag{A.43}$$

Therefore, from (A.29), (A.31), (A.32), (A.33), (A.38) and (A.41), we have

$$\begin{aligned}
\mu_{20} &= p(y_{n_1}c_{10}^2 - 2c_{10} + c_{20}) + p + 2y_{n_1}c_{10}\xi_{10} - 2\xi_{10} + \xi_{20} \\
&\quad + y_{n_1}d_{20p} + \beta_1 y_{n_1} l_{10p} + o_p(1) \\
&= p\Big[-\frac{2}{(1-y_{n_2})} + \frac{y_{n_1}}{(1-y_{n_2})^2} + \frac{1}{(1-y_{n_2})^3}\Big] + p \\
&\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1}+2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2+3y_{n_2}}{(1-y_{n_2})^4} \\
&\quad + \frac{2\beta_2 y_{n_1}y_{n_2} + \beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1-y_{n_2})^2} + \frac{2\beta_2 y_{n_2}}{(1-y_{n_2})^3} + o_p(1).
\end{aligned}$$

Letting $m_{10} = \int x dL(x)$ and $m_{20} = \int x^2 dL(x)$, based on (A.24), (A.25), (A.26), (A.31) and (A.37), we get

$$\sigma_{110p} = 4[(y_{N_1}+y_{N_2})p^{-1}\mathrm{tr}\boldsymbol{\Sigma}^2]^2 + o_p(1) \to 4(y_1+y_2)^2 m_{20}^2 = \sigma_{110}$$

and

$$\begin{aligned}\sigma_{120p}^1 &= \Big[\frac{8(y_{N_1}+y_{N_2})y_{N_2}^2+4y_{N_2}^2}{(1-y_{N_2})^2}+\frac{8y_{N_2}^3}{(1-y_{N_2})^3}\Big](p^{-1}\mathrm{tr}\mathbf{\Sigma})^2+o_p(1),\\ \sigma_{120p}^2 &= \Big[\frac{8(y_{N_1}+y_{N_2})y_{N_1}y_{N_2}+4y_{N_1}^2}{(1-y_{N_2})^2}+\frac{8y_{N_1}y_{N_2}}{(1-y_{N_2})^3}\Big](p^{-1}\mathrm{tr}\mathbf{\Sigma})^2+o_p(1),\end{aligned}$$

which yields that

$$\begin{aligned}\sigma_{120p} &= \sigma_{120p}^1+\sigma_{120p}^2\\ &= \Big[\frac{8y_{N_2}(y_{N_1}+y_{N_2})^2+4y_{N_1}^2+4y_{N_2}^2}{(1-y_{N_2})^2}\\ &\quad+\frac{8y_{N_1}y_{N_2}+8y_{N_2}^3}{(1-y_{N_2})^3}\Big](p^{-1}\mathrm{tr}\mathbf{\Sigma})^2+o_p(1)\\ &\to \Big[\frac{8y_2(y_1+y_2)^2+4y_1^2+4y_2^2}{(1-y_2)^2}+\frac{8y_1y_2+8y_2^3}{(1-y_2)^3}\Big]m_{10}^2=\sigma_{120}.\end{aligned}$$

From (A.27), (A.28), (A.31), (A.34), (A.35), (A.36) and (A.37)-(A.43), we get

$$\begin{aligned}\sigma_{220p} &= \sigma_{220p}^1+\sigma_{220p}^2 \qquad\qquad (A.44)\\ &\to 4(y_1d_{10}-1)^2(\eta_{110}+2y_1d_{20}+\beta_1y_1l_{10})\\ &\quad+4(y_1d_{10}-1)(\eta_{120}+4y_1d_{30}+2\beta_1y_1l_{20})\\ &\quad+\eta_{220}+4y_1(2d_{40}+\beta_1l_{30})+4y_1^2d_{20}^2\\ &= \frac{8y_1^3+16y_1^2y_2}{(1-y_2)^5}+\frac{4y_1^2+40y_1^2y_2+64y_1y_2^2}{(1-y_2)^6}\\ &\quad+\frac{8y_1y_2^4+56y_1y_2^2+48y_2^3+8y_1y_2}{(1-y_2)^7}+\frac{8y_2^5+24y_2^3+4y_2^2}{(1-y_2)^8}\\ &\quad+4(\beta_1y_1+\beta_2y_2)\Big[\frac{(y_1+y_2)^2}{(1-y_2)^4}+\frac{2y_2(y_1+y_2)}{(1-y_2)^5}+\frac{y_2^2}{(1-y_2)^6}\Big]=\sigma_{220}.\end{aligned}$$

Therefore, under the conditions of Theorem 2.1, based on the central limit theorem of martingale difference sequences, we have

$$T_A+T_B\xrightarrow{d}N(0,\omega_1^2\sigma_{110}+2\omega_1\omega_2\sigma_{120}+\omega_2^2\sigma_{220}).$$

Thus, we complete the proof of Theorem 2.1.

The Proof of Theorem 2.2. In this subsection, we prove the Theorem 2.2. Under the Assumptions A-B and H_0, denoting $\mathbf{\Sigma}_1=\mathbf{\Sigma}_2=\mathbf{\Sigma}$ and $\mathbf{S}=(n_1+n_2)^{-1}(n_1\mathbf{S}_1+n_2\mathbf{S}_2)$, from (A.3), (A.4) and (A.76), we have

$$\begin{aligned}p^{-1}\mathrm{tr}\mathbf{S} &= p^{-1}\mathrm{tr}\mathbf{\Sigma}+o_p(1),\\ p^{-1}\mathrm{tr}\mathbf{S}^2 &= p^{-1}\mathrm{tr}\mathbf{\Sigma}^2+(n_1+n_2)^{-1}p^{-1}\mathrm{tr}^2\mathbf{\Sigma}+o_p(1),\\ p^{-1}\mathrm{tr}(\mathbf{S}\circ\mathbf{S}) &= p^{-1}\mathrm{tr}(\mathbf{\Sigma}\circ\mathbf{\Sigma})+o_p(1),\end{aligned}$$

which implies that

$$p^{-1}[\mathrm{tr}\mathbf{S}^2 - (n_1+n_2)^{-1}\mathrm{tr}^2\mathbf{S}] = p^{-1}\mathrm{tr}\mathbf{\Sigma}^2 + o_p(1).$$

According to the discussion before (A.17)-(A.20), we find that d_{10}, d_{20}, d_{30}, d_{40} are the limits of the estimators of $p^{-1}\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1}$, $p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2$, $p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^3$, $p^{-1}\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^4$, respectively. When $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$, that is, $\mathbf{\Gamma} = \mathbf{\Sigma}_1^{-1/2}\mathbf{\Sigma}_2^{1/2} = \mathbf{I}_p$, we know $\mathrm{tr}\widetilde{\mathbf{B}}_2^{-1}$, $\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^2$, $\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^3$, $\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^4$ are the linear spectral statistics of the noncentralized sample covariance matrix $\widetilde{\mathbf{B}}_2$ with the population covariance matrix $\mathbf{I}_p$. Therefore, based on the proof of Theorem 2.1 and the Proposition A.1. of [20], for $i = 1,2,3,4$, we have

$$p^{-1}\big(\mathrm{tr}(\widetilde{\mathbf{B}}_2^{-1})^i - \mathrm{Etr}(\widetilde{\mathbf{B}}_2^{-1})^i\big) = o_p(1) \tag{A.45}$$

and

$$p^{-1}\mathrm{Etr}(\widetilde{\mathbf{B}}_2^{-1})^i = d_{i0p} + v_{i0p} + o(1), \tag{A.46}$$

where

$$\begin{aligned}
v_{10p} &= \frac{1}{p}\Big[\frac{y_{n_2}}{(1-y_{n_2})^2} + \frac{\beta_2 y_{n_2}}{(1-y_{n_2})}\Big],\\
v_{20p} &= \frac{1}{p}\Big[\frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} - \frac{\beta_2(y_{n_2}^2 - 3y_{n_2})}{(1-y_{n_2})^3}\Big],\\
v_{30p} &= \frac{1}{p}\Big[\frac{y_{n_2}^3 + 9y_{n_2}^2 + 6y_{n_2}}{(1-y_{n_2})^6} + \frac{6\beta_2 y_{n_2}}{(1-y_{n_2})^5}\Big],\\
v_{40p} &= \frac{1}{p}\Big[\frac{y_{n_2}^4 + 18y_{n_2}^3 + 35y_{n_2}^2 + 10y_{n_2}}{(1-y_{n_2})^8} + \frac{10\beta_2(y_{n_2}^2 + y_{n_2})}{(1-y_{n_2})^7}\Big].
\end{aligned}$$

Therefore, when we substitute $d_{i0p}+v_{i0p}$ for d_{i0} in (A.44) for $i = 1,2,3,4$, we can get the the expression of $\hat{\sigma}_{220}$. Thus, we complete the proof of Theorem 2.2.

The Proof of Theorem 2.3. In this subsection, we prove the Theorem 2.3. First, under the assumptions of Theorem 2.1, based on Theorem 2.1 and Theorem 2.2, we have

$$\begin{aligned}
P(T_{\mathrm{dr}} > t_\alpha) &= P(\max\{|T_{\mathrm{d}} - \mu_0 - \hat{\mu}_{10}|/\sqrt{\hat{\sigma}_{110}}, |T_{\mathrm{r}} - \mu_{20}|/\sqrt{\hat{\sigma}_{220}}\} > t_\alpha)\\
&= 1 - P(|T_{\mathrm{d}} - \mu_0 - \hat{\mu}_{10}|/\sqrt{\hat{\sigma}_{110}} \le t_\alpha, |T_{\mathrm{r}} - \mu_{20}|/\sqrt{\hat{\sigma}_{220}} \le t_\alpha)\\
&\to \alpha,
\end{aligned}$$

which implies that the test based on T_{dr} is asymptotically with the test level α. Secondly, under the conditions of Theorem 2.1 and the Conditions (C1), (C2) (or (C2*) and (C3) of [5], we have

$$\begin{aligned}
P(T_{\mathrm{drx}_2} > t_{\alpha/2}) &= P(\max\{T_{\mathrm{dr}}, c_\alpha T_{\mathrm{x}}\} > t_{\alpha/2})\\
&\le P(T_{\mathrm{dr}} > t_{\alpha/2}) + P(T_{\mathrm{x}} > q_{\alpha/2})\\
&\to \alpha/2 + \alpha/2 = \alpha,
\end{aligned}$$

which implies that the test based on T_{drx_2} is asymptotically equal to or less than α. Finally, under the conditions of Theorem 2.1 and the Conditions (C1), (C2) (or (C2*) and (C3) of [5], when the threshold $s(N_1,N_2,p)-4\log p\geq 0$, from (2.4), we have

$$P(T_{\text{x}}-4\log p+\log\log p\leq s(N_1,N_2,p)-4\log p+\log\log p)\to 1.$$

Therefore, we have

$$\begin{aligned}
P(T_{\text{dr}}>t_\alpha) &\leq P(T_{\text{drx}_1}>t_\alpha)\\
&= P(T_{\text{dr}}+p^2I(T_{\text{x}}>s(N_1,N_2,p))>t_\alpha)\\
&\leq P(T_{\text{dr}}>t_\alpha)+P(p^2I(T_{\text{x}}>s(N_1,N_2,p))>0)\\
&= P(T_{\text{dr}}>t_\alpha)+1-P(p^2I(T_{\text{x}}>s(N_1,N_2,p))=0)\\
&= P(T_{\text{dr}}>t_\alpha)+1-P(T_{\text{x}}\leq s(N_1,N_2,p))\\
&\to \alpha+1-1=\alpha.
\end{aligned}$$

It follows that the test based on T_{drx_1} is asymptotically with the test level α. Thus, we complete the proof of Theorem 2.3.

The Proof of Theorem 3.1. In this subsection, we prove the Theorem 3.1. Similar with the proof of Theorem 2.1, because all quantities are computed under $(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2)\in\Pi_1$, we add 1 to the subscripts of these quantities. When $(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2)\in\Pi_1$, we have

$$\mathbf{\Sigma}_1-\mathbf{\Sigma}_2=a_1/(p+a_1)\mathbf{\Sigma}_1,\quad \mathbf{\Sigma}_2^{-1}=\tau_p^{-1}\mathbf{\Sigma}_1^{-1}. \tag{A.47}$$

Then we have

$$\text{tr}(\mathbf{\Sigma}_1-\mathbf{\Sigma}_2)^2=a_1^2/(p+a_1)^2\text{tr}\mathbf{\Sigma}_1^2=o(1),$$

from (A.12), we obtain

$$\begin{aligned}
\mu_{11} &= n_1^{-1}\text{tr}\mathbf{\Sigma}_1^2+n_2^{-1}\text{tr}\mathbf{\Sigma}_2^2+\beta_1n_1^{-1}\text{tr}(\mathbf{\Sigma}_1\circ\mathbf{\Sigma}_1)+\beta_2n_2^{-1}\text{tr}(\mathbf{\Sigma}_2\circ\mathbf{\Sigma}_2)+o(1)\\
&= (y_{n_1}+y_{n_2}\tau_p^2)p^{-1}\text{tr}\mathbf{\Sigma}_1^2+(\beta_1y_{n_1}+\beta_2y_{n_2}\tau_p^2)p^{-1}\text{tr}(\mathbf{\Sigma}_1\circ\mathbf{\Sigma}_1)+o(1).
\end{aligned}$$

When $\mathbf{\Sigma}_2=\tau_p\mathbf{\Sigma}_1$, we have $\mathbf{\Gamma}_1=\mathbf{\Sigma}_1^{-1/2}\mathbf{\Sigma}_2^{1/2}=\tau_p^{1/2}\mathbf{I}_p$ and $\mathbf{T}_{1p}=\mathbf{\Gamma}_1\mathbf{\Gamma}_1^{\text{T}}=\tau_p\mathbf{I}_p$, which also satisfy the Assumptions c-d-f when we treat $\widetilde{\mathbf{B}}_2$ as $\mathbf{B}_n$ in section A.3.3, H_{1p} and H_1, the *ESD* and *LSD* of $\mathbf{T}_{1p}$, are $\tau_p\delta_1$ and δ_1, respectively. Therefore, for $i,j\in\{1,2\}$, from (A.8), (A.9) and (A.16), we get

$$c_{i1}=\tau_p^{-i}c_{i0},\qquad \xi_{i1}=\xi_{i0},\qquad \eta_{ij1}=\eta_{ij0}. \tag{A.48}$$

From (A.47) and (A.17)-(A.23), we have

$$d_{11p} = \frac{1}{(1-y_{N_2})\tau_p} \to \frac{1}{(1-y_2)} = d_{10}, \quad (A.49)$$

$$d_{21p} = \frac{1}{(1-y_{N_2})^3\tau_p^2} \to \frac{1}{(1-y_2)^3} = d_{20}, \quad (A.50)$$

$$d_{31p} = \frac{1+y_{N_2}}{(1-y_{N_2})^5\tau_p^3} \to \frac{1+y_2}{(1-y_2)^5} = d_{30}, \quad (A.51)$$

$$d_{41p} = \frac{y_{N_2}^2+3y_{N_2}+1}{(1-y_{N_2})^7\tau_p^4} \to \frac{y_2^2+3y_2+1}{(1-y_2)^7} = d_{40} \quad (A.52)$$

and

$$l_{11p} = \frac{1}{(1-y_{N_2})^2\tau_p^2} \to \frac{1}{(1-y_2)^2} = l_{10}, \quad (A.53)$$

$$l_{21p} = \frac{1}{(1-y_{N_2})^4\tau_p^3} \to \frac{1}{(1-y_2)^4} = l_{20}, \quad (A.54)$$

$$l_{31p} = \frac{1}{(1-y_{N_2})^6\tau_p^4} \to \frac{1}{(1-y_2)^6} = l_{30}. \quad (A.55)$$

Therefore, from (A.29), (A.31), (A.32), (A.33), (A.48), (A.50) and (A.53), we have

$$\begin{aligned}\mu_{21} &= p(y_{n_1}c_{11}^2 - 2c_{11} + c_{21}) + p + 2y_{n_1}c_{11}\xi_{11} - 2\xi_{11} + \xi_{21} \\ &\quad + y_{n_1}d_{21p} + \beta_1 y_{n_1} l_{11p} + o_p(1) \\ &= p\Big[-\frac{2}{(1-y_{n_2})\tau_p} + \frac{y_{n_1}}{(1-y_{n_2})^2\tau_p^2} + \frac{1}{(1-y_{n_2})^3\tau_p^2}\Big] + p \\ &\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1}+2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2+3y_{n_2}}{(1-y_{n_2})^4} - \frac{2\beta_2 y_{n_2}}{(1-y_{n_2})} \\ &\quad + \frac{2\beta_2 y_{n_1}y_{n_2} + \beta_1 y_{n_1} + \beta_2 y_{n_2}}{(1-y_{n_2})^2} + \frac{2\beta_2 y_{n_2}}{(1-y_{n_2})^3} + o_p(1).\end{aligned}$$

Letting $m_{11} = \int x dL_1(x)$ and $m_{21} = \int x^2 dL_1(x)$, based on (A.24), (A.25), (A.26), (A.47), (A.48), (A.31) and (A.49), we get

$$\sigma_{111p} = 4[(y_{N_1} + y_{N_2}\tau_p^2)p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1^2]^2 + o_p(1) \to 4(y_1+y_2)^2 m_{21}^2 = \sigma_{111}$$

and

$$\begin{aligned}\sigma_{121p}^1 &= \Big[\frac{8(y_{N_1}+y_{N_2}\tau_p)y_{N_2}^2 + 4y_{N_2}^2}{(1-y_{N_2})^2} \\ &\quad + \frac{8y_{N_2}^2 - 8y_{N_2}^2\tau_p}{(1-y_{N_2})^2} + \frac{8y_{N_2}^3}{(1-y_{N_2})^3}\Big](p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1)^2 + o_p(1), \\ \sigma_{121p}^2 &= \Big[\frac{8(y_{N_1}+y_{N_2}\tau_p)y_{N_1}y_{N_2} + 8y_{N_1}y_{N_2} - 8y_{N_1}y_{N_2}\tau_p}{(1-y_{N_2})^2\tau_p} \\ &\quad + \frac{4y_{N_1}^2}{(1-y_{N_2})^2\tau_p^2} + \frac{8y_{N_1}y_{N_2}}{(1-y_{N_2})^3\tau_p}\Big](p^{-1}\mathrm{tr}\boldsymbol{\Sigma}_1)^2 + o_p(1),\end{aligned}$$

which yields that

$$\begin{aligned}
\sigma_{121p} &= \sigma^1_{121p} + \sigma^2_{121p} \\
&\to \Big[\frac{8y_2(y_1+y_2)^2 + 4y_1^2 + 4y_2^2}{(1-y_2)^2} + \frac{8y_1y_2 + 8y_2^3}{(1-y_2)^3}\Big] m_{11}^2 = \sigma_{121}.
\end{aligned}$$

As for σ_{221}, from (A.27), (A.28), (A.31), (A.34), (A.35), (A.36), (A.48) and (A.49)-(A.55), we get

$$\begin{aligned}
\sigma_{221p} &= \sigma^1_{221p} + \sigma^2_{221p} \\
&\to 4(y_1d_{10} - 1)^2(\eta_{110} + 2y_1d_{20} + \beta_1y_1l_{10}) \\
&\quad +4(y_1d_{10} - 1)(\eta_{120} + 4y_1d_{30} + 2\beta_1y_1l_{20}) \\
&\quad +\eta_{220} + 4y_1(2d_{40} + \beta_1l_{30}) + 4y_1^2d_{20}^2 \\
&= \frac{8y_1^3 + 16y_1^2y_2}{(1-y_2)^5} + \frac{4y_1^2 + 40y_1^2y_2 + 64y_1y_2^2}{(1-y_2)^6} \\
&\quad + \frac{8y_1y_2^4 + 56y_1y_2^2 + 48y_2^3 + 8y_1y_2}{(1-y_2)^7} + \frac{8y_2^5 + 24y_2^3 + 4y_2^2}{(1-y_2)^8} \\
&\quad +4(\beta_1y_1 + \beta_2y_2)\Big[\frac{(y_1+y_2)^2}{(1-y_2)^4} + \frac{2y_2(y_1+y_2)}{(1-y_2)^5} + \frac{y_2^2}{(1-y_2)^6}\Big] = \sigma_{220}.
\end{aligned}$$

Therefore, under the conditions of Theorem 3.1, based on the central limit theorem of martingale difference sequences, we have

$$T_A + T_B \xrightarrow{d} N(0, \omega_1^2\sigma_{111} + 2\omega_1\omega_2\sigma_{121} + \omega_2^2\sigma_{220}).$$

Thus, we complete the proof of Theorem 3.1.

The proof of Proposition 3.1. Under Assumptions A-B, from (A.3) and (A.76), we have

$$\begin{aligned}
p^{-1}\mathrm{tr}\mathbf{S} &= p^{-1}\mathrm{tr}\mathbf{\Sigma}_w + o_p(1), && \text{(A.56)} \\
p^{-1}\mathrm{tr}\mathbf{S}^2 &= p^{-1}\mathrm{tr}\mathbf{\Sigma}_w^2 + \frac{n_1}{p(n_1+n_2)^2}\mathrm{tr}^2\mathbf{\Sigma}_1 && \text{(A.57)} \\
&\quad + \frac{n_2}{p(n_1+n_2)^2}\mathrm{tr}^2\mathbf{\Sigma}_2 + o_p(1), \\
p^{-1}\mathrm{tr}(\mathbf{S}\circ\mathbf{S}) &= p^{-1}\mathrm{tr}(\mathbf{\Sigma}_w\circ\mathbf{\Sigma}_w) + o_p(1), && \text{(A.58)}
\end{aligned}$$

where $\mathbf{\Sigma}_w = n_1/(n_1+n_2)\mathbf{\Sigma}_1 + n_2/(n_1+n_2)\mathbf{\Sigma}_2$. When $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_1$, we obtain

$$\begin{aligned}
\hat\mu_{10} &= \mu_{11} + o_p(1), \qquad \hat\sigma_{110} = \sigma_{111} + o_p(1), \\
\hat\sigma_{120} &= \sigma_{121} + o_p(1), \qquad \hat\sigma_{220} = \sigma_{220} + o(1)
\end{aligned}$$

and

$$\mu_{21} - \mu_{20} = 2a_1\Big[\frac{y_{n_1} + y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_2}}{(1-y_{n_2})^3}\Big] + o(1).$$

Therefore, based on Theorem 3.1 and Slutsky's theorem, we have

$$\begin{aligned}
& P\Big(\frac{|T_{\rm d}-\mu_0-\hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}} > z_{1-\alpha/2}\Big) \\
= \quad & P\Big(\frac{\sqrt{\sigma_{111}}}{\sqrt{\hat{\sigma}_{110}}}\cdot\frac{|T_{\rm d}-\mu_0-\mu_{11}+\mu_{11}-\hat{\mu}_{10}|}{\sqrt{\sigma_{111}}} > z_{1-\alpha/2}\Big) \\
\to \quad & 1-\Big[\Phi(z_{1-\alpha/2})-\Phi(-z_{1-\alpha/2})\Big] = \alpha
\end{aligned}$$

and

$$\begin{aligned}
& P\Big(\frac{|T_{\rm r}-\mu_{20}|}{\sqrt{\hat{\sigma}_{220}}} > z_{1-\alpha/2}\Big) \\
= \quad & P\Big(\frac{\sqrt{\sigma_{220}}}{\sqrt{\hat{\sigma}_{220}}}\cdot\frac{|T_{\rm r}-\mu_{21}+\mu_{21}-\mu_{20}|}{\sqrt{\sigma_{220}}} > z_{1-\alpha/2}) \\
\to \quad & 1-\Big[\Phi(z_{1-\alpha/2}-\Delta_1)-\Phi(-z_{1-\alpha/2}-\Delta_1)\Big] > \alpha,
\end{aligned}$$

where $\Phi(\cdot)$ is the distribution function of $N(0,1)$ and

$$\Delta_1 = \frac{2a_1}{\sqrt{\sigma_{220}}}\Big[\frac{y_1+y_2}{(1-y_2)^2}+\frac{y_2}{(1-y_2)^3}\Big].$$

Under the conditions of Theorem 3.1, based on the definition of t_α, we have

$$\begin{aligned}
P(T_{\rm dr} > t_\alpha) \quad = \quad & 1-P\Big(\frac{|T_{\rm d}-\mu_0-\hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}}\le t_\alpha, \frac{|T_{\rm r}-\mu_{20}|}{\sqrt{\hat{\sigma}_{220}}}\le t_\alpha\Big) \\
= \quad & 1-P\Big(\frac{\sqrt{\sigma_{111}}}{\sqrt{\hat{\sigma}_{110}}}\cdot\frac{|T_{\rm d}-\mu_0-\mu_{11}+\mu_{11}-\hat{\mu}_{10}|}{\sqrt{\sigma_{111}}}\le t_\alpha, \\
& \qquad \frac{\sqrt{\sigma_{220}}}{\sqrt{\hat{\sigma}_{220}}}\cdot\frac{|T_{\rm r}-\mu_{21}+\mu_{21}-\mu_{20}|}{\sqrt{\sigma_{220}}}\le t_\alpha\Big) \\
\to \quad & 1-\int_{-t'_\alpha-\Delta_1}^{t'_\alpha-\Delta_1}\int_{-t'_\alpha}^{t'_\alpha} f'(x_{\rm d},x_{\rm r})dx_{\rm d}dx_{\rm r} > \alpha,
\end{aligned}$$

where $f'(x_{\rm d},x_{\rm r})$ is the density of $N\left(\mathbf{0}_2, \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix}\right)$ with $\rho_1 = \sigma_{121}/\sqrt{\sigma_{111}\sigma_{220}}$ and t'_α satisifies

$$\alpha = 1-\int_{-t'_\alpha}^{t'_\alpha}\int_{-t'_\alpha}^{t'_\alpha} f'(x_{\rm d},x_{\rm r})dx_{\rm d}dx_{\rm r}.$$

According to the definition of $T_{{\rm drx}_1}$, we know $T_{{\rm drx}_1}\ge T_{\rm dr}$, then we have

$$P(T_{{\rm drx}_1} > t_\alpha) \ge P(T_{\rm dr} > t_\alpha),$$

which yeilds that

$$\lim_{p\to\infty} P(T_{{\rm drx}_1} > t_\alpha) \ge \lim_{p\to\infty} P(T_{\rm dr} > t_\alpha) > \alpha.$$

Based on the definition of T_{drx_2}, we have

$$\begin{aligned} & P(\max\{T_{\mathrm{dr}}, c_\alpha T_{\mathrm{x}}\} > t_{\alpha/2}) \\ = \; & 1 - P(T_{\mathrm{dr}} \le t_{\alpha/2}, T_{\mathrm{x}} \le q_{\alpha/2}) \\ \ge \; & 1 - P(T_{\mathrm{dr}} \le t_{\alpha/2}) = P(T_{\mathrm{dr}} > t_{\alpha/2}), \end{aligned}$$

then we obtain

$$\lim_{p\to\infty} P(T_{\mathrm{drx}_2} > t_{\alpha/2}) \ge 1 - \int_{-t'_{\alpha/2}-\Delta_1}^{t'_{\alpha/2}-\Delta_1} \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} f'(x_{\mathrm{d}}, x_{\mathrm{r}}) dx_{\mathrm{d}} dx_{\mathrm{r}} > \alpha/2,$$

where $t_{1\alpha/2}$ satisfies the following equation

$$\alpha/2 = 1 - \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} \int_{-t'_{\alpha/2}}^{t'_{\alpha/2}} f'(x_{\mathrm{d}}, x_{\mathrm{r}}) dx_{\mathrm{d}} dx_{\mathrm{r}}.$$

The Proof of Theorem 3.2. In this subsection, we prove the Theorem 3.2. Similar with the proof of Theorem 2.1, because all quantities are computed under $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_2$, we add 2 to the subscripts of these quantities. When $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_2$, we have

$$\mathbf{\Sigma}_1 = \mathbf{I}_p, \quad \mathbf{\Sigma}_2 - \mathbf{\Sigma}_1 = a_2/p\mathbf{J}_p, \quad \mathbf{\Sigma}_2^{-1} = \mathbf{I}_p - a_2/[p(1+a_2)]\mathbf{J}_p. \tag{A.59}$$

Then we have $\mathrm{tr}(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2)^2 = a_2^2$, when $\beta_2 = 0$, from (A.12), we obtain

$$\mu_{12} = a_2^2 + y_{n_1} + y_{n_2} + \beta_1 y_{n_1} + o(1).$$

When $\mathbf{\Sigma}_1 = \mathbf{I}_p$ and $\mathbf{\Sigma}_2 = \mathbf{\Sigma}_1 + a_2/p\mathbf{J}_p$, we have $\mathbf{\Gamma}_2 = \mathbf{\Sigma}_1^{-1/2}\mathbf{\Sigma}_2^{1/2} = \mathbf{I}_p + (\sqrt{1+a_2} - 1)/p\mathbf{J}_p$ and $\mathbf{T}_{2p} = \mathbf{\Gamma}_2\mathbf{\Gamma}_2^{\mathrm{T}} = \mathbf{\Sigma}_2$, which satisfy the Assumptions c-d when we view $\widetilde{\mathbf{B}}_2$ as $\mathbf{B}_n$ in section A.3.3. The eigenvalues of $\mathbf{T}_{2p}$ is $1 + a_2, 1, \cdots, 1$, which is coincident with the spiked model considered in [17], based on the Theorem 2 in [17], for $i = 1, 2$, we have

$$\begin{aligned} c_{i2} \;=\; & -\frac{1}{2\pi \mathbf{i} p} \oint_{\mathcal{C}_1} f_i\Big(-\frac{1}{z} + \frac{y_{n_2}}{1+z}\Big)\Big(\frac{1}{y_{n_2} z} - \frac{(1+a_2)^2 z}{(1+(1+a_2)z)^2}\Big) dz \\ & -\frac{1}{2\pi \mathbf{i} p} \oint_{\mathcal{C}_1} f_i'\Big(-\frac{1}{z} + \frac{y_{n_2}}{1+z}\Big) \frac{a_2}{(1+(1+a_2)z)(1+z)} \Big(\frac{1}{z} - \frac{y_{n_2} z}{(1+z)^2}\Big) dz \\ & + \Big(1 - \frac{1}{p}\Big) c_{i0} + \frac{1}{p} f_i(\phi_p(a_2)) I(a_2 > \sqrt{y_{n_2}}) + o\Big(\frac{1}{p}\Big), \end{aligned}$$

where $f_1(x) = 1/x$, $f_2(x) = 1/x^2$, $\phi_p(a) = 1 + a_2 + y_{n_2}(1+a_2)/a_2$ and $\mathcal{C}_1$ is a contour counterclockwise, when restricted to the real axes, encloses the interval $[-1/(1-\sqrt{y_{n_2}}), -1/(1+\sqrt{y_{n_2}})]$. After calculation, we get

$$c_{12} = \frac{1}{1 - y_{n_2}} - \frac{a_2}{p(1-y_{n_2})(1+a_2)} + o\Big(\frac{1}{p}\Big) \tag{A.60}$$

and

$$c_{22} = \frac{1}{(1-y_{n_2})^3} - \frac{(1+y_{n_2})a_2^2 + 2a_2}{p(1-y_{n_2})^3(1+a_2)^2} + o\Big(\frac{1}{p}\Big). \tag{A.61}$$

Since H_{2p} and H_2, the *ESD* and *LSD* of $\mathbf{T}_{2p}$, are $(1-p^{-1})\delta_1 + p^{-1}\delta_{1+a_2}$ and δ_1 respectively, when $\beta_2 = 0$, from (A.9) and (A.16), we have

$$\xi_{12} = \frac{y_{n_2}}{(1-y_{n_2})^2}, \quad \xi_{22} = \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} \tag{A.62}$$

and

$$\eta_{112} = \frac{2y_2}{(1-y_2)^4}, \quad \eta_{122} = \frac{4y_2(1+y_2)}{(1-y_2)^6}, \quad \eta_{222} = \frac{4y_2(2y_2^2+5y_2+2)}{(1-y_2)^8}. \tag{A.63}$$

From (A.59) and (A.17)-(A.23), we have

$$d_{12p} = \frac{1}{(1-y_{N_2})} + o(1) \to \frac{1}{(1-y_2)} = d_{10}, \tag{A.64}$$

$$d_{22p} = \frac{1}{(1-y_{N_2})^3} + o(1) \to \frac{1}{(1-y_2)^3} = d_{20}, \tag{A.65}$$

$$d_{32p} = \frac{1+y_{N_2}}{(1-y_{N_2})^5} + o(1) \to \frac{1+y_2}{(1-y_2)^5} = d_{30}, \tag{A.66}$$

$$d_{42p} = \frac{y_{N_2}^2 + 3y_{N_2} + 1}{(1-y_{N_2})^7} + o(1) \to \frac{y_2^2+3y_2+1}{(1-y_2)^7} = d_{40} \tag{A.67}$$

and

$$l_{12p} = \frac{1}{(1-y_{N_2})^2} + o(1) \to \frac{1}{(1-y_2)^2} = l_{10}, \tag{A.68}$$

$$l_{22p} = \frac{1}{(1-y_{N_2})^4} + o(1) \to \frac{1}{(1-y_2)^4} = l_{20}, \tag{A.69}$$

$$l_{32p} = \frac{1}{(1-y_{N_2})^6} + o(1) \to \frac{1}{(1-y_2)^6} = l_{30}. \tag{A.70}$$

Therefore, from (A.29), (A.60), (A.61), (A.62), (A.65) and (A.68), we have

$$\begin{aligned}
\mu_{22} &= p(y_{n_1}c_{12}^2 - 2c_{12} + c_{22}) + p + 2y_{n_1}c_{12}\xi_{12} - 2\xi_{12} + \xi_{22} \\
&\quad + y_{n_1}d_{22p} + \beta_1 y_{n_1} l_{12p} + o_p(1) \\
&= p\Big[-\frac{2}{(1-y_{n_2})} + \frac{y_{n_1}}{(1-y_{n_2})^2} + \frac{1}{(1-y_{n_2})^3}\Big] + p \\
&\quad + \frac{a_2^2 - 2a_2(1+a_2)(y_{n_1}+y_{n_2})}{(1-y_{n_2})^2(1+a_2)^2} - \frac{2a_2y_{n_2}}{(1-y_{n_2})^3(1+a_2)} \\
&\quad - \frac{2y_{n_2}}{(1-y_{n_2})^2} + \frac{y_{n_1} + 2y_{n_1}y_{n_2}}{(1-y_{n_2})^3} + \frac{y_{n_2}^2 + 3y_{n_2}}{(1-y_{n_2})^4} + \frac{\beta_1 y_{n_1}}{(1-y_{n_2})^2} + o_p(1).
\end{aligned}$$

Based on (A.24), (A.25), (A.26), (A.59), (A.60) and (A.64), we get

$$\sigma_{112p} = 4(y_{N_1} + y_{N_2})^2 + o_p(1) \to 4[(y_1+y_2)]^2 = \sigma_{112}$$

and

$$\begin{aligned}\sigma_{122p}^1 &= \Big[\frac{8(y_{N_1}+y_{N_2})y_{N_2}^2+4y_{N_2}^2}{(1-y_{N_2})^2}+\frac{8y_{N_2}^3}{(1-y_{N_2})^3}\Big]+o_p(1),\\ \sigma_{122p}^2 &= \Big[\frac{8(y_{N_1}+y_{N_2})y_{N_1}y_{N_2}+4y_{N_1}^2}{(1-y_{N_2})^2}+\frac{8y_{N_1}y_{N_2}}{(1-y_{N_2})^3}\Big]+o_p(1),\end{aligned}$$

which yields that

$$\begin{aligned}\sigma_{122p} &= \sigma_{122p}^1+\sigma_{122p}^2\\ &= \Big[\frac{8y_{N_2}(y_{N_1}+y_{N_2})^2+4y_{N_1}^2+4y_{N_2}^2}{(1-y_{N_2})^2}+\frac{8y_{N_1}y_{N_2}+8y_{N_2}^3}{(1-y_{N_2})^3}\Big]+o_p(1)\\ &\to \Big[\frac{8y_2(y_1+y_2)^2+4y_1^2+4y_2^2}{(1-y_2)^2}+\frac{8y_1y_2+8y_2^3}{(1-y_2)^3}\Big]=\sigma_{122}.\end{aligned}$$

From (A.27), (A.28), (A.60), (A.63) and (A.64)-(A.70), we get

$$\begin{aligned}\sigma_{222p} &= \sigma_{222p}^1+\sigma_{222p}^2\\ &\to 4(y_1d_{10}-1)^2(\eta_{112}+2y_1d_{20}+\beta_1y_1l_{10})\\ &\quad +4(y_1d_{10}-1)(\eta_{122}+4y_1d_{30}+2\beta_1y_1l_{20})\\ &\quad +\eta_{222}+4y_1(2d_{40}+\beta_1l_{30})+4y_1^2d_{20}^2\\ &= \frac{8y_1^3+16y_1^2y_2}{(1-y_2)^5}+\frac{4y_1^2+40y_1^2y_2+64y_1y_2^2}{(1-y_2)^6}\\ &\quad +\frac{8y_1y_2^4+56y_1y_2^2+48y_2^3+8y_1y_2}{(1-y_2)^7}+\frac{8y_2^5+24y_2^3+4y_2^2}{(1-y_2)^8}\\ &\quad +4\beta_1y_1\Big[\frac{(y_1+y_2)^2}{(1-y_2)^4}+\frac{2y_2(y_1+y_2)}{(1-y_2)^5}+\frac{y_2^2}{(1-y_2)^6}\Big]=\sigma_{222}.\end{aligned}$$

Therefore, under the conditions of Theorem 3.2, based on the central limit theorem of martingale difference sequences, we have

$$T_A+T_B\xrightarrow{d}N(0,\omega_1^2\sigma_{112}+2\omega_1\omega_2\sigma_{122}+\omega_2^2\sigma_{222}).$$

Thus, we complete the proof of Theorem 3.2.

The proof of Proposition 3.2. Under Assumptions A-B, when $(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2)\in\Pi_2$ and $\beta_2=0$, from (A.56), (A.57) and (A.58), we have

$$\hat\sigma_{110}=\sigma_{112}+o_p(1),\quad \hat\sigma_{120}=\sigma_{122}+o_p(1),\quad \hat\sigma_{220}=\sigma_{222}+o(1),$$
$$\mu_{12}-\hat\mu_{10}=a_2^2+o_p(1)$$

and

$$\mu_{22}-\mu_{20}=\frac{a_2^2-2a_2(1+a_2)(y_{n_1}+y_{n_2})}{(1-y_{n_2})^2(1+a_2)^2}-\frac{2a_2y_{n_2}}{(1-y_{n_2})^3(1+a_2)}.$$

Therefore, based on Theorem 3.2 and Slutsky's theorem, we have

$$\begin{aligned} & P\Big(\frac{|T_{\mathrm{d}}-\mu_0-\hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}}>z_{1-\alpha/2}\Big) \\ = \quad & P\Big(\frac{\sqrt{\sigma_{112}}}{\sqrt{\hat{\sigma}_{110}}}\cdot\frac{|T_{\mathrm{d}}-\mu_0-\mu_{12}+\mu_{12}-\hat{\mu}_{10}|}{\sqrt{\sigma_{112}}}>z_{1-\alpha/2}\Big) \\ \to \quad & 1-\Big[\Phi(z_{1-\alpha/2}-\Delta_2)-\Phi(-z_{1-\alpha/2}-\Delta_2)\Big]>\alpha \end{aligned}$$

and

$$\begin{aligned} & P\Big(\frac{|T_{\mathrm{r}}-\mu_{20}|}{\sqrt{\hat{\sigma}_{220}}}>z_{1-\alpha/2}\Big) \\ = \quad & P\Big(\frac{\sqrt{\sigma_{222}}}{\sqrt{\hat{\sigma}_{220}}}\cdot\frac{|T_{\mathrm{r}}-\mu_{22}+\mu_{22}-\mu_{20}|}{\sqrt{\sigma_{222}}}>z_{1-\alpha/2}\Big) \\ \to \quad & 1-\Big[\Phi(z_{1-\alpha/2}-\Delta_3)-\Phi(-z_{1-\alpha/2}-\Delta_3)\Big]\geq\alpha, \end{aligned}$$

where $\Delta_2=a_2^2/[2(y_1+y_2)]$ and

$$\Delta_3=\frac{1}{\sqrt{\sigma_{222}}}\Big[\frac{a_2^2-2a_2(1+a_2)(y_1+y_2)}{(1-y_2)^2(1+a_2)^2}-\frac{2a_2y_2}{(1-y_2)^3(1+a_2)}\Big].$$

Under the conditions of Theorem 3.2, according to the definition of t_α, we have

$$\begin{aligned} P(T_{\mathrm{dr}}>t_\alpha) \quad &= \quad 1-P\Big(\frac{|T_{\mathrm{d}}-\mu_0-\hat{\mu}_{10}|}{\sqrt{\hat{\sigma}_{110}}}\leq t_\alpha,\frac{|T_{\mathrm{r}}-\mu_{20}|}{\sqrt{\hat{\sigma}_{220}}}\leq t_\alpha\Big) \\ &= \quad 1-P\Big(\frac{\sqrt{\sigma_{112}}}{\sqrt{\hat{\sigma}_{110}}}\cdot\frac{|T_{\mathrm{d}}-\mu_0-\mu_{12}+\mu_{12}-\hat{\mu}_{10}|}{\sqrt{\sigma_{112}}}\leq t_\alpha, \\ &\qquad\qquad \frac{\sqrt{\sigma_{222}}}{\sqrt{\hat{\sigma}_{220}}}\cdot\frac{|T_{\mathrm{r}}-\mu_{22}+\mu_{22}-\mu_{20}|}{\sqrt{\sigma_{222}}}\leq t_\alpha\Big) \\ &\to \quad 1-\int_{-t_\alpha^*-\Delta_3}^{t_\alpha^*-\Delta_3}\int_{-t_\alpha^*-\Delta_2}^{t_\alpha^*-\Delta_2}f^*(x_{\mathrm{d}},x_{\mathrm{r}})dx_{\mathrm{d}}dx_{\mathrm{r}}>\alpha, \end{aligned}$$

where $f^*(x_{\mathrm{d}},x_{\mathrm{r}})$ is the density of $N\left(\mathbf{0}_2,\begin{pmatrix}1 & \rho_2\\ \rho_2 & 1\end{pmatrix}\right)$ with $\rho_2=\sigma_{122}/\sqrt{\sigma_{112}\sigma_{222}}$ and t_α^* satisifies

$$\alpha=1-\int_{-t_\alpha^*}^{t_\alpha^*}\int_{-t_\alpha^*}^{t_\alpha^*}f^*(x_{\mathrm{d}},x_{\mathrm{r}})dx_{\mathrm{d}}dx_{\mathrm{r}}.$$

Similar with the proof of Propostion 3.1, we can obtain (IV) and (V). Thus, we complete the proof of Proposition 3.2.

The proof of Proposition 3.3. First, we prove (I). Under the conditions given in the conclusion (I), from (A.56), (A.57) and (A.58), we get

$$\begin{aligned} \hat{I}_{10} \quad &= \quad p^{-1}\mathrm{tr}\mathbf{\Sigma}_w+o_p(1), \\ \hat{I}_{20} \quad &= \quad p^{-1}\mathrm{tr}\mathbf{\Sigma}_w^2+\frac{y_{n_1}^2y_{n_2}^2}{(y_{n_1}+y_{n_2})^3}(p^{-1}\mathrm{tr}\mathbf{\Sigma}_d)^2+o_p(1), \end{aligned}$$

where $\mathbf{\Sigma}_d = \mathbf{\Sigma}_1 - \mathbf{\Sigma}_2$, which implies the critical value t_α tends to a constant. Based on (32) and (33) in [5], we get

$$\begin{aligned} T_{\mathrm{x}} &\overset{a.s.}{\geq} 0.5 \max_{1 \leq l_1 \leq l_2 \leq p} \frac{(\sigma_{1l_1l_2} - \sigma_{2l_1l_2})^2}{\hat{\theta}_{1l_1l_2}/n_1 + \hat{\theta}_{2l_1l_2}/n_2} - 4\log p + 0.5 \log\log p \\ &\overset{a.s.}{\geq} 0.5 \max_{1 \leq l_1 \leq l_2 \leq p} \frac{(\sigma_{1l_1l_2} - \sigma_{2l_1l_2})^2}{\theta_{1l_1l_2}/n_1 + \theta_{2l_1l_2}/n_2} - 4\log p + 0.5 \log\log p. \end{aligned}$$

Therefore, when p is large enough, we have

$$\begin{aligned} & P(T_{\mathrm{drx}_1} > t_\alpha) = P(T_{\mathrm{dr}} + p^2 I(T_{\mathrm{x}} > s(N_1, N_2, p)) > t_\alpha) \\ \geq \;& P(p^2 I(T_{\mathrm{x}} > s(N_1, N_2, p))) = P(T_{\mathrm{x}} > s(N_1, N_2, p)). \end{aligned}$$

Hence, as $p \to \infty$, we have $P(T_{\mathrm{drx}_1} > t_\alpha) \to 1$.

Next, we prove (II). According to the definition of T_{drx_2}, we get

$$\begin{aligned} P(T_{\mathrm{drx}_2} > t_{\alpha/2}) &= P(\max\{T_{\mathrm{dr}}, c_\alpha T_{\mathrm{x}}\} > t_{\alpha/2}) \\ &= 1 - P(T_{\mathrm{dr}} \leq t_{\alpha/2}, T_{\mathrm{x}} \leq q_{\alpha/2}) \\ &\geq 1 - P(T_{\mathrm{x}} \leq q_{\alpha/2}) = P(T_{\mathrm{x}} > q_{\alpha/2}), \end{aligned}$$

when $(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \in \Pi_3$, based on the Theorem 2 in [5], we have $P(T_{\mathrm{drx}_2} > t_{\alpha/2}) \to 1$. Thus, we complete the proof of Proposition 3.3.

A.3.3. Technical lemmas

We provide some useful lemmas in this subsection. In accordance with the notations in [3], in this subsection, the dimension and sample size are represented by n and N, respectively. The sample $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N$ is from a n-dimensional population $\mathbf{x}$. Let $\mathbf{x}_i = (x_{1i}, \cdots, x_{ni})^{\mathrm{T}}$ and $\mathbf{X}_n = (\mathbf{x}_1, \cdots, \mathbf{x}_N)$, define

$$\mathbf{B}_n = N^{-1} \mathbf{\Gamma} \mathbf{X}_n \mathbf{X}_n^{\mathrm{T}} \mathbf{\Gamma}^{\mathrm{T}},$$
$$m_n(z) = m_{F^{\mathbf{B}_n}}(z) = \int \frac{1}{\lambda - z} dF^{\mathbf{B}_n}(\lambda), \quad \Im(z) \neq 0,$$

where $\mathbf{X}_n^{\mathrm{T}}$ denotes the transpose of $\mathbf{X}_n$ and $\mathbf{\Gamma}$ is an invertible $n \times n$ matrix, $\Im(z)$ denotes the imaginary part of the complex number z, $m_n(z)$ denotes the Stieltjes transform of $F^{\mathbf{B}_n}$.

Let $\underline{\mathbf{B}}_n = N^{-1} \mathbf{X}_n^{\mathrm{T}} \mathbf{\Gamma}^{\mathrm{T}} \mathbf{\Gamma} \mathbf{X}_n$ (the spectra of which differs from that of $\mathbf{B}_n$ by $|n - N|$ zeros). Its Stieltjes transform is defined as follow

$$\underline{m}_n(z) = m_{F^{\underline{\mathbf{B}}_n}}(z) = -\frac{1 - c_n}{z} + c_n m_n(z),$$

where $c_n = n/N$.

- **Assumption a**. The elements $\{x_{ji}, j = 1, \cdots, n; i = 1, \cdots, N\}$ are *i.i.d.* with $\mathrm{E}x_{11} = 0$, $\mathrm{E}x_{11}^2 = 1$ and $\beta_x = \mathrm{E}x_{11}^4 - 3$.

- **Assumption b**. $N = N(n)$ with $c_n = n/N \to c > 0$ as $n \to \infty$.
- **Assumption c**. The spectral norm of $\mathbf{T}_n = \mathbf{\Gamma}\mathbf{\Gamma}^{\mathrm{T}}$ is bounded and the *ESD* H_n of $\mathbf{T}_n$ converges weakly to a *LSD* H as $n \to \infty$.
- **Assumption d**. The spectral norm of $\mathbf{T}_n^{-1}$, the inverse of $\mathbf{T}_n$, is bounded.
- **Assumption f**. $\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}$ is diagonal or $\beta_x = 0$.

From [20], under Assumptions a-b-c, we known the sample covariance matrix $\mathbf{B}_n$ has the *LSD* $F^{c,H}$, namely the Marčenko-Pastur distribution of index (c, H), which has support

$$[a, b] = \big[(1-\sqrt{c})^2 I_{(0<c<1)} \liminf_n \lambda_{\min}^{\mathbf{T}_n}, (1+\sqrt{c})^2 \limsup_n \lambda_{\max}^{\mathbf{T}_n}\big].$$

Moreover, the *LSD* $F^{c,H}$ has a Dirac mass $1 - 1/c$ at the origin when $c > 1$. Define $\underline{m}_c$ to be the Stieltjes transform of the companion *LSD*

$$\underline{F}^{c,H} = (1-c)\delta_0 + cF^{c,H},$$

where δ_0 is the point distribution at zero. Then $\underline{m}_c$ is the unique solution in

$$\frac{1-c}{z} + \underline{m}_c \in \mathbb{C}^+ = \{z : \Im z > 0\}$$

of the equation

$$z = -\frac{1}{\underline{m}_c(z)} + c\int \frac{tdH(t)}{1 + t\underline{m}_c(z)}, \qquad z \in \mathbb{C}^+ = \{z : \Im z > 0\}.$$

Let $\boldsymbol{\gamma}_j = (1/\sqrt{N})\mathbf{\Gamma}\mathbf{x}_j$, $\mathrm{E}_0(\cdot)$ denote expectation and $\mathrm{E}_j(\cdot)$ denote the conditional expectation with respect to the σ-field generated by $\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_j$.

Lemma A.1. *Let $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$ and $\mathbf{M}_4$ be symmetric and nonnegative definite matrices with bounded spectral norms, then under Assumptions a-b-c, we have*

$$\begin{aligned}
& \sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1)]^2\} \qquad\qquad (A.71)\\
= \;& \frac{1}{N}[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_1)^2 + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma})],\\
& \sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_3)^2]^2\} \qquad\qquad (A.72)\\
= \;& \frac{4}{N^3}\mathrm{tr}^2(\mathbf{T}_n\mathbf{M}_3)[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)^2 + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma})]\\
& + \frac{8}{N^2}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)^3 + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma})]\\
& + \frac{4}{N}[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)^4 + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{\Gamma})]\\
& + \frac{4}{N^2}\mathrm{tr}^2(\mathbf{T}_n\mathbf{M}_3)^2 + o_p(1),
\end{aligned}$$

$$\begin{aligned}
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1)][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_2)]\} \qquad \text{(A.73)}\\
= \ & \frac{1}{N}[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n\mathbf{M}_2) + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})],\\
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1)][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_3)^2]\} \qquad \text{(A.74)}\\
= \ & \frac{2}{N^2}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n\mathbf{M}_3) + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma}]\\
&+\frac{2}{N}[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3) + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{\Gamma})] + o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_3)^2][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_4)^2]\} \qquad \text{(A.75)}\\
=&\frac{4}{N^3}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)\mathrm{tr}(\mathbf{T}_n\mathbf{M}_4)[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_4) + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_4\mathbf{\Gamma})]\\
&+ \frac{4}{N^2}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_4)[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_4) + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_4\mathbf{\Gamma})]\\
&+ \frac{4}{N^2}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_4\mathbf{T}_n\mathbf{M}_4) + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_4\mathbf{T}_n\mathbf{M}_4\mathbf{\Gamma})]\\
&+ \frac{4}{N}[2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_4\mathbf{T}_n\mathbf{M}_4) + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{T}_n\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_4\mathbf{T}_n\mathbf{M}_4\mathbf{\Gamma})]\\
&+ \frac{4}{N^2}\mathrm{tr}^2(\mathbf{T}_n\mathbf{M}_3\mathbf{T}_n\mathbf{M}_4) + o_p(1).
\end{aligned}$$

The proof of Lemma A.1 is similar to that of Lemma 1 given in the Supplementary Material of [21]. Besides, it is worth noting that the identity holds [see (1.15) of [3]]

$$\begin{aligned}
&\mathrm{E}\big[(\mathbf{x}_1^{\mathrm{T}}\mathbf{A}\mathbf{x}_1 - \mathrm{tr}\mathbf{A})(\mathbf{x}_1^{\mathrm{T}}\mathbf{B}\mathbf{x}_1 - \mathrm{tr}\mathbf{B})\big]\\
&\qquad = \beta_x \sum_{i=1}^{n} a_{ii}b_{ii} + \mathrm{tr}\mathbf{A}\mathbf{B}^{\mathrm{T}} + \mathrm{tr}\mathbf{A}\mathbf{B}, \qquad \text{(A.76)}
\end{aligned}$$

for $n \times n$ matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$. From (A.76), we have

$$\begin{aligned}
\mathrm{Etr}(\mathbf{B}_n\mathbf{M}_1) &= \mathrm{tr}(\mathbf{T}_n\mathbf{M}_1), \qquad \text{(A.77)}\\
\mathrm{Etr}(\mathbf{B}_n\mathbf{M}_3)^2 &= \frac{1}{N}\mathrm{tr}^2(\mathbf{T}_n\mathbf{M}_3) + \mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)^2\\
&\quad +\frac{1}{N}[\mathrm{tr}(\mathbf{T}_n\mathbf{M}_3)^2 + \beta_x \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_3\mathbf{\Gamma})]. \qquad \text{(A.78)}
\end{aligned}$$

Before giving Lemma A.2, some preparatory work is needed. Let $\nu = \Im(z)$. For the following of this subsection we will assume $\nu > 0$. Without loss of generality, we assume $\|\mathbf{T}_n\| \le 1$ ($\|\cdot\|$ denotes the spectral norm of $\mathbf{T}_n$) for all n. Let $\mathbf{D}(z) = \mathbf{B}_n - z\mathbf{I}$, $\mathbf{D}_j(z) = \mathbf{D}(z) - \boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}$,

$$\begin{aligned}
\varepsilon_j(z) &= \boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j - \frac{1}{N}\mathrm{tr}\mathbf{T}_n\mathbf{D}_j^{-1}(z),\\
\delta_j(z) &= \boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-2}(z)\boldsymbol{\gamma}_j - \frac{1}{N}\mathrm{tr}\mathbf{T}_n\mathbf{D}_j^{-2}(z) = \frac{d}{dz}\varepsilon_j(z),\\
\beta_j(z) &= \frac{1}{1+\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j},\\
\bar{\beta}_j(z) &= \frac{1}{1+N^{-1}\mathrm{tr}\mathbf{T}_n\mathbf{D}_j^{-1}(z)},\\
b_n(z) &= \frac{1}{1+N^{-1}\mathrm{Etr}\mathbf{T}_n\mathbf{D}_1^{-1}(z)}.
\end{aligned}$$

All of the three latter quantities are bounded in absolute by $|z|/\nu$ [see (3.4) of [2]]. We have

$$\mathbf{D}^{-1}(z) - \mathbf{D}_j^{-1}(z) = -\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\beta_j(z). \tag{A.79}$$

For convenience, constants appearing in inequalities will be denoted by K and may be taken as different values from one expression to the next. Assume that matrix $\mathbf{A} = (a_{ij})$ is a $n \times n$ matrix with bounded spectral norm, from (A.76) we get

$$\mathrm{E}\big[(N^{-1}\mathbf{x}_1^{\mathrm{T}}\mathbf{A}\mathbf{x}_1 - N^{-1}\mathrm{tr}\mathbf{A})^2\big] = \frac{1}{N^2}\big[\beta_x\sum_{i=1}^{n} a_{ii}^2 + \mathrm{tr}\mathbf{A}\mathbf{A}^{\mathrm{T}} + \mathrm{tr}\mathbf{A}^2\big],$$

which implies that

$$N^{-1}\mathbf{x}_1^{\mathrm{T}}\mathbf{A}\mathbf{x}_1 = N^{-1}\mathrm{tr}\mathbf{A} + o_p(1). \tag{A.80}$$

Lemma A.2. *Let $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$ and $\mathbf{M}_4$ be $n\times n$ matrices with bounded spectral norms. Under Assumptions a-b-c-d, when $\mathbf{B}_n$ is invertible, we have*

$$N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1) = \frac{1}{1-c_n}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1) + o_p(1), \tag{A.81}$$

$$\begin{aligned}
&N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2) \qquad\qquad (\mathrm{A.82})\\
&\quad= \frac{1}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2)\\
&\qquad+\frac{1}{(1-c_n)^3}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2) + o_p(1),
\end{aligned}$$

$$
\begin{aligned}
& N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3) \\
= \quad & \frac{1}{(1-c_n)^3}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3) \\
& +\frac{1}{(1-c_n)^4}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3) \\
& +\frac{1}{(1-c_n)^4}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_3) \\
& +\frac{1}{(1-c_n)^4}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2) \\
& +\frac{2}{(1-c_n)^5}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)+o_p(1)
\end{aligned}
\tag{A.83}
$$

and

$$
\begin{aligned}
& N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4) \\
= \quad & \frac{1}{(1-c_n)^4}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{1}{(1-c_n)^5}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{1}{(1-c_n)^5}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_3\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{1}{(1-c_n)^5}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{1}{(1-c_n)^5}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_4)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3) \\
& +\frac{1}{(1-c_n)^5}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{1}{(1-c_n)^5}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_4)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3) \\
& +\frac{1}{(1-c_n)^6}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{1}{(1-c_n)^6}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_4)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_3) \\
& +\frac{2}{(1-c_n)^6}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{2}{(1-c_n)^6}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_4)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3) \\
& +\frac{2}{(1-c_n)^6}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_4) \\
& +\frac{2}{(1-c_n)^6}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_4)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2) \\
& +\frac{5}{(1-c_n)^7}N^{-4}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_3)\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_4)
\end{aligned}
\tag{A.84}
$$

$+o_p(1)$.

Proof. According to (4.13) of [3], we have

$$\mathbf{D}^{-1}(z) = -(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1} + b_n(z)\mathbf{A}(z) + \mathbf{B}(z) + \mathbf{C}(z), \tag{A.85}$$

where

$$\begin{aligned}
\mathbf{A}(z) &= \sum_{j=1}^{N}(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1}(\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}} - N^{-1}\mathbf{T}_n)\mathbf{D}_j^{-1}(z),\\
\mathbf{B}(z) &= \sum_{j=1}^{N}(\beta_j(z) - b_n(z))(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{C}(z) &= N^{-1}b_n(z)(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n\sum_{j=1}^{N}\left(\mathbf{D}_j^{-1}(z) - \mathbf{D}^{-1}(z)\right)\\
&= N^{-1}b_n(z)(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n\sum_{j=1}^{N}\beta_j(z)\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z).
\end{aligned}$$

Under Assumptions a-b-c-d, we know that $\|(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1}\|$ and $\|\mathbf{D}_1^{-1}(z)\|$ are bounded. Suppose $\mathbf{M}$ is an $n \times n$ nonrandom matrix and $\|\mathbf{M}\|$ is bounded, similar with (4.15) and (4.16) in [3], we have

$$N^{-1}\mathrm{tr}\mathbf{A}(z)\mathbf{M} = o_p(1),$$
$$N^{-1}\mathrm{tr}\mathbf{B}(z)\mathbf{M} = o_p(1)$$

and

$$N^{-1}\mathrm{tr}\mathbf{C}(z)\mathbf{M} = o_p(1),$$

therefore

$$N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1 = -N^{-1}\mathrm{tr}(z\mathbf{I} - b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1 + o_p(1).$$

Based on $b_n(z) = \mathrm{E}\beta_1(z) + o(1)$ [see (4.3) of [2]] and

$$\mathrm{E}\beta_1(z) = -z\mathrm{E}\underline{m}_n(z) = 1 - c_n - zc_n\mathrm{E}m_n(z),$$

when $z = 0$, we have $b_n(z) = 1 - c_n + o(1)$. Therefore, we get

$$N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1) = \frac{1}{1 - c_n}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1) + o_p(1),$$

which is (A.81). Next we will prove (A.82). From (A.79), (A.80) and (A.85), we have

$$\begin{aligned}
&N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\\
&\quad= -N^{-1}\mathrm{tr}(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\\
&\qquad+b_n(z)N^{-1}\mathrm{tr}\mathbf{A}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2+o_p(1)\\
&\quad= -N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\\
&\qquad+b_n(z)N^{-1}\sum_{j=1}^{N}\mathrm{tr}(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\\
&\qquad\quad\times\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1(\mathbf{D}^{-1}(z)-\mathbf{D}_j^{-1}(z))\mathbf{M}_2+o_p(1)\\
&\quad= -N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\\
&\qquad-b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_2(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j\\
&\qquad\quad\times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j+o_p(1)\\
&\quad= -N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\\
&\qquad-b_n^2(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n\\
&\qquad\quad\times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{T}_n+o_p(1).
\end{aligned}$$

Therefore, when $z=0$, we get

$$\begin{aligned}
&N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\\
&\quad=(1-c_n)^{-1}N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_1\\
&\qquad+(1-c_n)N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_n+o_p(1).
\end{aligned}$$

Let $\mathbf{M}_2=\mathbf{T}_n$, from (A.81), we have

$$N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_n=\frac{1}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1+o_p(1).$$

Hence, we have

$$\begin{aligned}
&N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2)\\
=\;&\frac{1}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2)\\
&+\frac{1}{(1-c_n)^3}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)+o_p(1),
\end{aligned}$$

which is (A.82). Next we will prove (A.83). Again from (A.79), (A.80) and (A.85), we get

$$
\begin{aligned}
& N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3 \\
= \; & -N^{-1}\mathrm{tr}(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3 \\
& +b_n(z)N^{-1}\mathrm{tr}\mathbf{A}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3+o_p(1) \\
= \; & -N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1 \\
& +b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j^2(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \qquad\qquad \times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j \\
& -b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \qquad\qquad \times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j \\
& -b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \qquad\qquad \times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j+o_p(1) \\
= \; & -N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1 \\
& +b_n^3(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_n \\
& \qquad\qquad \times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n \\
& -b_n^2(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{T}_n \\
& \qquad\qquad \times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n \\
& -b_n^2(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_n \\
& \qquad\qquad \times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_3(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n+o_p(1).
\end{aligned}
$$

Therefore, when $z=0$, we get

$$
\begin{aligned}
& N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3) \\
= \; & (1-c_n)^{-1}N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{T}_n^{-1}\mathbf{M}_1 \\
& -(1-c_n)^2N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_3 \\
& +(1-c_n)N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3 \\
& +(1-c_n)N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_3+o_p(1).
\end{aligned}
$$

Let $\mathbf{M}_3=\mathbf{T}_n$, from (A.81) and (A.82), we have

$$
\begin{aligned}
& N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_n) \\
= \; & \frac{1}{(1-c_n)^4}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2 \\
& +\frac{2}{(1-c_n)^5}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_2+o_p(1).
\end{aligned}
$$

Hence, we can obtain (A.83). Finally, we will prove (A.84). From (A.79), (A.80) and (A.85), we get

$$
\begin{aligned}
& N^{-1}\mathrm{tr}\big(\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4\big) \\
= \; & -N^{-1}\mathrm{tr}\big((z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4\big) \\
& +b_n(z)N^{-1}\mathrm{tr}\big(\mathbf{A}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4\big)+o_p(1) \\
= \; & -N^{-1}\mathrm{tr}\big(\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\big) \\
& -b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j^3(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \quad\times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_3\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j \\
& +b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j^2(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \quad\times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_3\mathbf{D}_j^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j \\
& +b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j^2(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_4 \\
& \quad\times(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\mathbf{M}_3\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& -b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\mathbf{M}_3\mathbf{D}_j^{-1}(z)\mathbf{M}_4 \\
& \quad\times(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& +b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j^2(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \quad\times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_3\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j \\
& -b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \quad\times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_3\mathbf{D}_j^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j \\
& -b_n(z)N^{-1}\sum_{j=1}^{N}\beta_j(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_1\mathbf{D}_j^{-1}(z)\mathbf{M}_2\mathbf{D}_j^{-1}(z)\mathbf{M}_3\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j \\
& \quad\times\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\boldsymbol{\gamma}_j+o_p(1) \\
= \; & -N^{-1}\mathrm{tr}\big(\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{M}_1\big) \\
& -b_n^4(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_n \\
& \quad\times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n \\
& +b_n^3(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_n \\
& \quad\times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n
\end{aligned}
$$

$$
\begin{aligned}
&+b_n^3(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\\
&\quad\times\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n\\
&-b_n^2(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n\\
&\quad\times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{T}_n\\
&+b_n^3(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\\
&\quad\times\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n\\
&-b_n^2(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n\\
&\quad\times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{T}_n\\
&-b_n^2(z)N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_1\mathbf{D}^{-1}(z)\mathbf{M}_2\mathbf{D}^{-1}(z)\mathbf{M}_3\mathbf{D}^{-1}(z)\mathbf{T}_n\\
&\quad\times N^{-1}\mathrm{tr}\mathbf{D}^{-1}(z)\mathbf{M}_4(z\mathbf{I}-b_n(z)\mathbf{T}_n)^{-1}\mathbf{T}_n+o_p(1).
\end{aligned}
$$

Therefore, when $z=0$, we get

$$
\begin{aligned}
&N^{-1}\mathrm{tr}\big(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4\big)\\
=\;&(1-c_n)^{-1}N^{-1}\mathrm{tr}\big(\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4\mathbf{T}_n^{-1}\mathbf{M}_1\big)\\
&+(1-c_n)^3N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_n\\
&\quad\times N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_4\\
&-(1-c_n)^2N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_n\\
&\quad\times N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4\\
&-(1-c_n)^2N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_n\\
&\quad\times N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_4\\
&+(1-c_n)N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4\\
&-(1-c_n)^2N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_n\\
&\quad\times N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_4\\
&+(1-c_n)N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_4\\
&+(1-c_n)N^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_nN^{-1}\mathrm{tr}\mathbf{B}_n^{-1}\mathbf{M}_4.
\end{aligned}
$$

Let $\mathbf{M}_4=\mathbf{T}_n$, from (A.81), (A.82) and (A.83), we have

$$
\begin{aligned}
&N^{-1}\mathrm{tr}\big(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{T}_n\big)\\
=\;&\frac{1}{(1-c_n)^5}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3\\
&+\frac{1}{(1-c_n)^6}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_2N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_3\\
&+\frac{2}{(1-c_n)^6}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n^{-1}\mathbf{M}_3\\
&+\frac{2}{(1-c_n)^6}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_3N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\\
&+\frac{5}{(1-c_n)^7}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_1N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_2N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}\mathbf{M}_3.
\end{aligned}
$$

Hence, we can obtain (A.84). Thus, we complete the proof of Lemma A.2. ☐

Lemma A.3. *Let N' be an integer satisfying $c_n' = n/N' \to c' > 0$ and $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ be $n \times n$ matrices with bounded spectral norm. Under Assumptions a-b-c-d, when $\mathbf{B}_n$ is invertible, we have*

$$\begin{aligned}\frac{1}{N'}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n\mathbf{M}_2) &= \frac{1}{N'(1-c_n)}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n\mathbf{M}_2) \\ &\quad -\frac{1}{N'N(1-c_n)}\mathrm{tr}\mathbf{M}_1\mathrm{tr}\mathbf{M}_2 + o_p(1)\end{aligned} \tag{A.86}$$

and

$$\begin{aligned}&\frac{1}{N'}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n\mathbf{M}_3) \\ =\ & \frac{1}{N'(1-c_n)^2}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_3) \\ &+\frac{1}{N'N(1-c_n)^3}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_3) \\ &-\frac{1}{N'N(1-c_n)^2}\mathrm{tr}(\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2)\mathrm{tr}\mathbf{M}_3 \\ &-\frac{1}{N'N(1-c_n)^2}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{M}_3)\mathrm{tr}\mathbf{M}_2 \\ &-\frac{1}{N'N^2(1-c_n)^3}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)\mathrm{tr}\mathbf{M}_2\mathrm{tr}\mathbf{M}_3 + o_p(1).\end{aligned} \tag{A.87}$$

Proof. Rewrite

$$\mathbf{B}_n = N^{-1}\mathbf{\Gamma}\mathbf{X}_n\mathbf{X}_n^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}} = \sum_{j=1}^{N}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}},$$

letting $\mathbf{B}_{nj} = \mathbf{B}_n - \boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}$ and $\eta_j = (1+\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j)^{-1}$, from (A.79), (A.80) and (A.81), we have

$$\begin{aligned}&\frac{1}{N'}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n\mathbf{M}_2) = \frac{1}{N'}\sum_{j=1}^{N}\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_2\mathbf{B}_n^{-1}\mathbf{M}_1\boldsymbol{\gamma}_j \\ =\ & \frac{1}{N'}\sum_{j=1}^{N}\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_2(\mathbf{B}_{nj}^{-1} - \eta_j\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1})\mathbf{M}_1\boldsymbol{\gamma}_j \\ =\ & \frac{1}{N'}\sum_{j=1}^{N}\Big[\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_2\mathbf{B}_{nj}^{-1}\mathbf{M}_1\boldsymbol{\gamma}_j - \eta_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_2\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1}\mathbf{M}_1\boldsymbol{\gamma}_j\Big] \\ =\ & \frac{1}{N'}\Big[\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{T}_n\mathbf{M}_2) - \frac{\eta_1}{N}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{T}_n)\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{T}_n\mathbf{M}_2)\Big] + o_p(1) \\ =\ & \frac{1}{N'(1-c_n)}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n\mathbf{M}_2) - \frac{1}{N'N(1-c_n)}\mathrm{tr}\mathbf{M}_1\mathrm{tr}\mathbf{M}_2 + o_p(1).\end{aligned}$$

Similarly, from (A.79), (A.80), (A.81) and (A.82), we have

$$\begin{aligned}
&\frac{1}{N'}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{B}_n\mathbf{M}_3)\\
=\;& \frac{1}{N'}\sum_{j=1}^{N}\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_3\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\boldsymbol{\gamma}_j\\
=\;& \frac{1}{N'}\sum_{j=1}^{N}\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_3(\mathbf{B}_{nj}^{-1}-\eta_j\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1})\mathbf{M}_1\\
&\quad\times(\mathbf{B}_{nj}^{-1}-\eta_j\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1})\mathbf{M}_2\boldsymbol{\gamma}_j\\
=\;& \frac{1}{N'}\sum_{j=1}^{N}\big(\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_3\mathbf{B}_{nj}^{-1}\mathbf{M}_1\mathbf{B}_{nj}^{-1}\mathbf{M}_2\boldsymbol{\gamma}_j\\
&\quad-\eta_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_3\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1}\mathbf{M}_1\mathbf{B}_{nj}^{-1}\mathbf{M}_2\boldsymbol{\gamma}_j\\
&\quad-\eta_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_3\mathbf{B}_{nj}^{-1}\mathbf{M}_1\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1}\mathbf{M}_2\boldsymbol{\gamma}_j\\
&\quad+\eta_j^2\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_3\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1}\mathbf{M}_1\mathbf{B}_{nj}^{-1}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{B}_{nj}^{-1}\mathbf{M}_2\boldsymbol{\gamma}_j\big)\\
=\;& \frac{N}{N'}\Big[N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_3)\\
&\quad-\eta_1N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{T}_n\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{T}_n)\\
&\quad-\eta_1N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_n\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{T}_n)\\
&\quad+\eta_1^2N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{T}_n\mathbf{M}_3)N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_1\mathbf{B}_n^{-1}\mathbf{T}_n)\\
&\quad\times N^{-1}\mathrm{tr}(\mathbf{B}_n^{-1}\mathbf{M}_2\mathbf{T}_n)\Big]+o_p(1)\\
=\;& \frac{1}{N'(1-c_n)^2}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_3)\\
&+\frac{1}{N'N(1-c_n)^3}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_3)\\
&-\frac{1}{N'N(1-c_n)^2}\mathrm{tr}(\mathbf{M}_1\mathbf{T}_n^{-1}\mathbf{M}_2)\mathrm{tr}\mathbf{M}_3\\
&-\frac{1}{N'N(1-c_n)^2}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1\mathbf{M}_3)\mathrm{tr}\mathbf{M}_2\\
&-\frac{1}{N'N^2(1-c_n)^3}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)\mathrm{tr}\mathbf{M}_2\mathrm{tr}\mathbf{M}_3+o_p(1). \qquad \square
\end{aligned}$$

Lemma A.4. *Let $\mathbf{M}_1$, $\mathbf{M}_2$ be symmetric and nonnegative definite matrices with bounded spectral norms. Under Assumptions a-b-c-d, when $\mathbf{B}_n$ is invertible, we have*

$$\begin{aligned}
&\sum_{j=1}^{N}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-1})][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1)]\} \qquad \text{(A.88)}\\
=\;& -\frac{2}{(1-c_n)}N^{-1}\mathrm{tr}\mathbf{M}_1-\frac{\beta_x}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_1\boldsymbol{\Gamma})+o_p(1),
\end{aligned}$$

$$\begin{aligned}
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-2})][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1)]\} \qquad (A.89)\\
= \;& -\frac{4}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1)-\frac{4}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}\mathbf{M}_1\\
&-\frac{2\beta_x}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma})\\
&-\frac{2\beta_x}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma})+o_p(1),\\
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-1})][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_2)^2]\} \qquad (A.90)\\
= \;& -\frac{4}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2^2)-\frac{2\beta_x}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})\\
&-\frac{2\beta_x}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})+o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-2})][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_2)^2]\} \qquad (A.91)\\
= \;& -\frac{8}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2)-\frac{8}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2^2)\\
&+\frac{4}{(1-c_n)^2}\left(N^{-1}\mathrm{tr}\mathbf{M}_2\right)^2-\frac{4\beta_x}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})\\
&-\frac{4\beta_x}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})\\
&-\frac{4\beta_x}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})\\
&-\frac{4\beta_x}{(1-c_n)^3}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})+o_p(1).
\end{aligned}$$

Proof. Based on (A.79), we have

$$\begin{aligned}
&(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}\mathbf{D}^{-1}(z)\\
= \;& \mathrm{tr}\Big[(\mathrm{E}_j-\mathrm{E}_{j-1})\big(\mathbf{D}_j^{-1}(z)-\beta_j(z)\mathbf{D}_j^{-1}(z)\boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^T\mathbf{D}_j^{-1}(z)\big)\Big]\\
= \;& -(\mathrm{E}_j-\mathrm{E}_{j-1})\beta_j(z)\boldsymbol{\gamma}_j^T\mathbf{D}_j^{-2}(z)\boldsymbol{\gamma}_j.
\end{aligned}$$

Since

$$\begin{aligned}
\beta_j(z) &= \overline{\beta}_j(z)-\beta_j(z)\overline{\beta}_j(z)\varepsilon_j(z)\\
&= \overline{\beta}_j(z)-\overline{\beta}_j^2(z)\varepsilon_j(z)+\overline{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z),
\end{aligned}$$

we obtain

$$\begin{aligned}
& -(\mathrm{E}_j - \mathrm{E}_{j-1})\beta_j(z)\boldsymbol{\gamma}_j^T\mathbf{D}_j^{-2}(z)\boldsymbol{\gamma}_j \\
= & -(\mathrm{E}_j - \mathrm{E}_{j-1})\Big(\overline{\beta}_j(z)\delta_j(z) - \overline{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z) \\
& \qquad -\overline{\beta}_j^2(z)\varepsilon_j(z)N^{-1}\mathrm{tr}(\mathbf{D}_j^{-2}(z)\mathbf{T}_n) + \overline{\beta}_j^2(z)\beta_j(z)\varepsilon_j^2(z)\boldsymbol{\gamma}_j^T\mathbf{D}_j^{-2}(z)\boldsymbol{\gamma}_j\Big) \\
= & -\mathrm{E}_j\Big(\overline{\beta}_j(z)\delta_j(z) - \overline{\beta}_j^2(z)\varepsilon_j(z)N^{-1}\mathrm{tr}(\mathbf{D}_j^{-2}(z)\mathbf{T}_n)\Big) \\
& +(\mathrm{E}_j - \mathrm{E}_{j-1})\overline{\beta}_j^2(z)\Big(\varepsilon_j(z)\delta_j(z) - \beta_j(z)\boldsymbol{\gamma}_j^T\mathbf{D}_j^{-2}(z)\boldsymbol{\gamma}_j\varepsilon_j^2(z)\Big).
\end{aligned}$$

Using (2.3) of [3], we have

$$\begin{aligned}
\mathrm{E}\Big|\sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\overline{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z)\Big|^2 &= \sum_{j=1}^N \mathrm{E}|(\mathrm{E}_j - \mathrm{E}_{j-1})\overline{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z)|^2 \\
&\le 4\sum_{j=1}^N \mathrm{E}|\overline{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z)|^2 = o(1).
\end{aligned}$$

This implies that

$$\sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\overline{\beta}_j^2(z)\varepsilon_j(z)\delta_j(z) \xrightarrow{i.p.} 0.$$

By the same argument, we have

$$\sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\beta_j(z)\boldsymbol{\gamma}_j^T\mathbf{D}_j^{-2}(z)\boldsymbol{\gamma}_j\varepsilon_j^2(z) \xrightarrow{i.p.} 0.$$

Therefore

$$\begin{aligned}
& \mathrm{tr}\big[\mathbf{D}^{-1}(z) - \mathrm{E}\mathbf{D}^{-1}(z)\big] = \sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}\mathbf{D}^{-1}(z) \\
= & -\sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\beta_j(z)\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{D}_j^{-2}(z)\boldsymbol{\gamma}_j \\
= & -\sum_{j=1}^N \mathrm{E}_j\big(\overline{\beta}_j(z)\delta_j(z) - \overline{\beta}_j^2(z)\varepsilon_j(z)N^{-1}\mathrm{tr}(\mathbf{D}_j^{-2}(z)\mathbf{T}_n)\big) + o_p(1).
\end{aligned}$$

Also because of

$$-\mathrm{E}_j\frac{d}{dz}\overline{\beta}_j(z)\varepsilon_j(z) = -\mathrm{E}_j\big(\overline{\beta}_j(z)\delta_j(z) - \overline{\beta}_j^2(z)\varepsilon_j(z)N^{-1}\mathrm{tr}(\mathbf{D}_j^{-2}(z)\mathbf{T}_n)\big),$$
$$(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1) = \boldsymbol{\gamma}_j^T\mathbf{M}_1\boldsymbol{\gamma}_j - N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_1),$$

based on the central limit theorem of martingale difference sequences, we consider the following sum

$$\begin{aligned}
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\big[-\overline{\beta}_j(z)\varepsilon_j(z)(\boldsymbol{\gamma}_j^{\mathrm{T}}\mathbf{M}_1\boldsymbol{\gamma}_j - N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_1))\big] \\
= \quad & -\frac{1}{N^2}\sum_{j=1}^{N} \mathrm{E}_{j-1}\Big[\overline{\beta}_j(z)\big(\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\boldsymbol{\Gamma}\mathbf{x}_j \\
& \qquad -\mathrm{tr}(\mathbf{T}_n\mathbf{D}_j^{-1}(z))\big)(\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_1\boldsymbol{\Gamma}\mathbf{x}_j - \mathrm{tr}(\mathbf{T}_n\mathbf{M}_1))\Big] \\
= \quad & -\frac{1}{N^2}\sum_{j=1}^{N} \mathrm{E}_{j-1}\Big[\overline{\beta}_j(z)\big(2\mathrm{tr}(\mathbf{D}_j^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) \\
& \qquad +\beta_x\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_1\boldsymbol{\Gamma})\big)\Big] \\
= \quad & -\frac{1}{N^2}\sum_{j=1}^{N} \mathrm{E}_{j-1}\Big[b_n(z)\big(2\mathrm{tr}(\mathbf{D}_j^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) \\
& \qquad +\beta_x\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_1\boldsymbol{\Gamma})\Big] + o_p(1) \\
= \quad & -\frac{1}{N^2}\sum_{j=1}^{N} \mathrm{E}_{j-1}\Big[b_n(z)\big(2\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) \\
& \qquad +\beta_x\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_1\boldsymbol{\Gamma})\big)\Big] + o_p(1),
\end{aligned}$$

where the last two equalities follow from (2.3) of [3] and

$$\mathrm{E}|\overline{\beta}_j(z) - b_n(z)|^2 \le K\frac{|z|^4}{\nu^6}N^{-1}.$$

Based on the dominated convergence theorem, we only need to calculate the limits of first derivative and second derivative of $b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n)$ and $b_n(z)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_1\boldsymbol{\Gamma})$. Since

$$b_n'(z) = -b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n) + o_p(1),$$

by Lemma A.2, when $z \to 0$, we have

$$\begin{aligned}
& \frac{d}{dz}[b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n)] \\
= \quad & -b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) \\
& +b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) + o_p(1) \\
= \quad & \frac{1}{(1-c_n)}N^{-1}\mathrm{tr}\mathbf{M}_1 + o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{d^2}{dz^2}\Big[b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n)\Big] \\
= \;& 2b_n^3(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\big)^2 N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) \\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) \\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) \\
& +2b_n(z)\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n\mathbf{M}_1\mathbf{T}_n) + o_p(1) \\
= \;& \frac{2}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1) + \frac{2}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}\mathbf{M}_1 + o_p(1).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\frac{d}{dz}[b_n(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma})] \\
= \;& -b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
& +b_n(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-2}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) + o_p(1) \\
= \;& \frac{1}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) + o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{d^2}{dz^2}\Big[b_n(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma})\Big] \\
= \;& 2b_n^3(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\big)^2 N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-2}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
& +2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-3}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
= \;& \frac{2}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
& +\frac{2}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) + o_p(1).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{j=1}^{N}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-1})][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1)]\} \\
=& -\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big[\frac{2}{(1-c_n)}N^{-1}\mathrm{tr}\mathbf{M}_1 + \frac{\beta_x}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma})\Big] + o_p(1) \\
=& -\frac{2}{(1-c_n)}N^{-1}\mathrm{tr}\mathbf{M}_1 - \frac{\beta_x}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) + o_p(1), \\
&\sum_{j=1}^{N}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-2})][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_1)]\}
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{N}\sum_{j=1}^{N} \mathrm{E}_{j-1}\Big[\frac{4}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1) \\
&\quad + \frac{4}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}\mathbf{M}_1 \\
&\quad + \frac{2\beta_x}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
&\quad + \frac{2\beta_x}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma})\Big] + o_p(1) \\
&= -\frac{4}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_1) - \frac{4}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}\mathbf{M}_1 \\
&\quad - \frac{2\beta_x}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) \\
&\quad - \frac{2\beta_x}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_1\mathbf{\Gamma}) + o_p(1).
\end{aligned}
$$

For the martingale difference of $\mathrm{tr}[(\mathbf{B}_n\mathbf{M}_2)^2]$, we have

$$
\begin{aligned}
&(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}[(\mathbf{B}_n\mathbf{M}_2)^2] \\
= \;& \frac{1}{N^2}\Big[2\big(\mathbf{x}_j^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2(\sum_{i=1}^{j-1}\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}})\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_j - \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2(\sum_{i=1}^{j-1}\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}})\mathbf{M}_2\mathbf{\Gamma})\big) \\
&+2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)\big(\mathbf{x}_j^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_j - \mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)\big) \\
&+2(N-j)\big(\mathbf{x}_j^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_j - \mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)^2\big)\Big] + o_p(1).
\end{aligned}
$$

Still based on the central limit theorem of martingale difference sequences, we consider the following sum

$$
\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_2)^2\big(-\overline{\beta}_j(z)\varepsilon_j(z)\big)\Big] \tag{A.92}
$$

$$
\begin{aligned}
= \;& -\frac{1}{N^3}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big[2\overline{\beta}_j(z)\big(2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2(\sum_{i=1}^{j-1}\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}})\mathbf{M}_2\mathbf{T}_n\mathbf{D}_j^{-1}(z)) \\
&\qquad +\beta_x\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2(\sum_{i=1}^{j-1}\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}})\mathbf{M}_2\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{\Gamma})\big) \\
&\qquad +2\overline{\beta}_j(z)\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)\big(2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{D}_j^{-1}(z)) \\
&\qquad +\beta_x\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{\Gamma})\big) \\
&\qquad +2\overline{\beta}_j(z)(N-j)\big(2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{D}_j^{-1}(z)) \\
&\qquad +\beta_x\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{\Gamma})\big)\Big] \\
= \;& -\frac{1}{N^3}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big[2\overline{\beta}_j(z)\big(2\sum_{i=1}^{j-1}\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{D}_j^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_i
\end{aligned}
$$

$$
\begin{aligned}
&+\beta_x \sum_{i=1}^{j-1} \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^T\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{\Gamma})\big) \\
&+2\overline{\beta}_j(z)\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)\big(2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{D}_j^{-1}(z)) \\
&+\beta_x\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{\Gamma})\big) \\
&+2\overline{\beta}_j(z)(N-j)\big(2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{D}_j^{-1}(z)) \\
&+\beta_x\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{\Gamma})\big)\Big].
\end{aligned}
$$

Let

$$
\begin{aligned}
\mathbf{D}_{ij}(z) &= \mathbf{D}(z) - \boldsymbol{\gamma}_i\boldsymbol{\gamma}_i^{\mathrm{T}} - \boldsymbol{\gamma}_j\boldsymbol{\gamma}_j^{\mathrm{T}}, \\
\beta_{ij}(z) &= \frac{1}{1+\boldsymbol{\gamma}_i^{\mathrm{T}}\mathbf{D}_{ij}^{-1}(z)\boldsymbol{\gamma}_i}, \\
b_1(z) &= \frac{1}{1+N^{-1}\mathrm{E}\mathrm{tr}\mathbf{T}_n\mathbf{D}_{12}^{-1}(z)}.
\end{aligned}
$$

Writing

$$
\begin{aligned}
&\sum_{i=1}^{j-1} \mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{D}_j^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_i \\
= &\sum_{i=1}^{j-1} \mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\big(\mathbf{D}_{ji}^{-1}(z) - \beta_{ji}(z)\mathbf{D}_{ji}^{-1}(z)\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i^{\mathrm{T}}\mathbf{D}_{ji}^{-1}(z)\big)\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_i \\
= &\sum_{i=1}^{j-1} \Big(\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{D}_{ji}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_i \\
&-\beta_{ji}(z)N^{-1}\big(\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{D}_{ji}^{-1}(z)\mathbf{\Gamma}\mathbf{x}_i\big)^2\Big),
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{i=1}^{j-1} \mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^T\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_j^{-1}(z)\mathbf{\Gamma}) \\
= &\sum_{i=1}^{j-1} \big(\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_{ji}^{-1}(z)\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}) \\
&-\beta_{ji}(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_{ji}^{-1}(z)\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}_{ji}^{-1}(z)\mathbf{\Gamma}\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})\big),
\end{aligned}
$$

we get

$$
\begin{aligned}
(\mathrm{A.92}) = &-\frac{1}{N^2}\sum_{j=1}^{N} \mathrm{E}_{j-1}\Big[2b_n(z)\big(2(j-1)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n) \\
&-2(j-1)b_n(z)(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n))^2
\end{aligned}
$$

$$
\begin{aligned}
&+2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&+2(N-j)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)\\
&+\beta_x b_n(z)\big(2(j-1)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma})\\
&+2\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\boldsymbol{\Gamma})\\
&+2(N-j)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma})\big)\Big]+o_p(1)\\
= \quad & -\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big[2b_n(z)\big(2N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&-2(j-1)N^{-1}b_n(z)(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n))^2\\
&+2N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)\\
&+\beta_x b_n(z)\big(2N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma})\\
&+2N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\boldsymbol{\Gamma})\big)\Big]+o_p(1).
\end{aligned}
$$

Mimicking the discussion above, we need to calculate the limits of the first derivative and second derivative of $b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)$, $b_n^2(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)^2$ and $b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)$; $b_n(z)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma})$ and $b_n(z)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\boldsymbol{\Gamma})$. By Lemma A.2, when $z\to 0$, we have

$$
\begin{aligned}
&\frac{d}{dz}\big[b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big]\\
= \quad & -b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&+b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)+o_p(1)\\
= \quad & \frac{1}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2^2)+o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{d}{dz}\Big[b_n^2(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)^2\Big]\\
= \quad & -2b_n^3(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)^2\\
&+2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)+o_p(1)\\
= \quad & \frac{2}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{M}_2)+o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{d}{dz}\Big[b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\Big]\\
= \quad & -b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&+b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)+o_p(1)\\
= \quad & \frac{1}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{M}_2)+o_p(1),
\end{aligned}
$$

$$
\frac{d}{dz}\Big[b_n(z)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma})\Big]
$$

$$
\begin{aligned}
&= -b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})\\
&\quad +b_n(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})+o_p(1)\\
&= \frac{1}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})+o_p(1)
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{d}{dz}\Big[b_n(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})\Big]\\
&= -b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})\\
&\quad +b_n(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-2}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})+o_p(1)\\
&= \frac{1}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})+o_p(1).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\sum_{j=1}^{N}\mathrm{E}_{j-1}\{[(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-1})][(\mathrm{E}_j-\mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_2)^2]\}\\
&= -\frac{4}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2^2)-\frac{2\beta_x}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma})\\
&\quad -\frac{2\beta_x}{(1-c_n)}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})+o_p(1).
\end{aligned}
$$

Next, the second derivative of those terms are given as follows

$$
\begin{aligned}
&\frac{d^2}{dz^2}\big[b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big]\\
&= 2b_n^3(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\big)^2N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&\quad -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&\quad -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&\quad +2b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&= \frac{2}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2)\\
&\quad +\frac{2}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2^2)+o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{d^2}{dz^2}\big[b_n^2(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)^2\big]\\
&= 6b_n^4(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\big)^2\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)^2\\
&\quad -4b_n^3(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)^2\\
&\quad -8b_n^3(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&\quad \times N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
&\quad +2b_n^2(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\big)^2
\end{aligned}
$$

TABLE A.1

Comparison of empirical powers (in percentages) of the proposed tests with the tests LC and CLX under Scenario 1 when the observations are from the Gaussian and Gamma population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	46.49	20.29	42.94	58.77	10.25	43.45	36.36
2	40	(120,120)	65.63	23.24	61.25	75.68	10.82	61.61	53.9
3	40	(120,160)	74.78	34.25	71.4	83.46	12.03	71.71	63.92
4	80	(120,160)	54.51	21.82	50.26	66.3	7.43	50.64	42.19
5	80	(160,240)	82.41	46.57	80.12	89.3	9.09	80.24	73.08
6	80	(240,240)	94.85	55.07	93.14	97.24	9.61	93.19	89.26
7	160	(160,240)	64.41	15.3	57.14	75.14	6.43	57.31	48.36
8	160	(240,240)	82.63	15.97	76.66	89.31	6.42	76.69	68.85
9	160	(240,320)	90.96	33.43	87.73	94.77	7.23	87.73	81.55
10	320	(240,480)	85.59	20.12	80.2	91.77	5.65	80.26	72.7
11	320	(320,480)	95.26	20.68	92.64	97.75	5.93	92.66	88.41
12	320	(480,480)	99.45	22.55	98.96	99.81	6.41	98.96	98.13
			Gamma population						
1	40	(80,120)	44.14	19.89	41.78	58.19	8.01	42.18	35.46
2	40	(120,120)	63.87	22.94	59.84	75.09	7.66	60.01	52.1
3	40	(120,160)	73.06	32.68	70.14	82.33	9.34	70.33	62.54
4	80	(120,160)	52.32	21.11	48.4	65.04	5.63	48.56	39.91
5	80	(160,240)	80.15	43.33	77.84	88.17	6.75	77.89	70.61
6	80	(240,240)	93.9	50.18	92.39	96.87	7.48	92.41	88.14
7	160	(160,240)	63.43	16.03	56.4	75.11	4.5	56.54	46.82
8	160	(240,240)	81.81	16.62	76.07	88.91	4.64	76.13	67.77
9	160	(240,320)	90.12	32.14	86.5	94.67	5.11	86.52	80.43
10	320	(240,480)	84.51	20.03	78.71	91.6	4.52	78.73	70.39
11	320	(320,480)	94.97	21.75	91.77	97.49	4.63	91.78	87.43
12	320	(480,480)	99.51	22.43	99.03	99.79	4.96	99.03	98.18

$$
\begin{aligned}
& +4b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
= \; & \frac{2}{(1-c_n)^2}(N^{-1}\mathrm{tr}\mathbf{M}_2)^2 + \frac{4}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2)\\
& \frac{4}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}\mathbf{M}_2N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2) + o_p(1),\\
& \frac{d^2}{dz^2}\left[b_n(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\right]\\
= \; & 2b_n^3(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\big)^2N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-1}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)\\
& +2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n\mathbf{M}_2\mathbf{T}_n)
\end{aligned}
$$

Table A.2
Comparison of empirical powers (in percentages) of the proposed tests with the tests LC and CLX under Scenario 2 when the observations are from the Gaussian and Gamma population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	6.59	28.64	25.07	9.24	6.2	25.64	21.32
2	40	(120,120)	8.32	30.77	27.52	11.01	5.11	27.87	22.87
3	40	(120,160)	7.83	39.17	34.6	11.14	6.16	34.87	28.93
4	80	(120,160)	7.32	37.44	32.54	10.67	5.63	32.91	27.12
5	80	(160,240)	8.08	68.49	61.18	12.4	6.08	61.3	53.34
6	80	(240,240)	10.79	74.19	68.1	15.87	5.75	68.16	60.83
7	160	(160,240)	7.66	46.33	40.3	12.15	5.84	40.6	33.73
8	160	(240,240)	9.63	48.79	43.37	15.26	5.37	43.6	36.19
9	160	(240,320)	10.01	87.06	81.78	15.99	5.56	81.88	75.25
10	320	(240,480)	10.48	93.17	90.38	17.8	6.72	90.39	85.86
11	320	(320,480)	14.85	94.25	91.58	23.66	6.31	91.6	87.74
12	320	(480,480)	21.36	95.88	93.67	31.03	6.02	93.69	89.92
			Gamma population						
1	40	(80,120)	7.89	27.61	24.98	9.91	5.44	25.44	20.69
2	40	(120,120)	8.54	29.84	27.22	11.44	4.11	27.45	22.34
3	40	(120,160)	9.06	37.6	34.22	12.15	5.65	34.46	28.61
4	80	(120,160)	7.6	37.4	32.49	10.76	4.62	32.75	27.27
5	80	(160,240)	8.18	63.73	56.82	12.94	5.64	57.07	49.81
6	80	(240,240)	11.46	69.29	62.34	16.65	4.28	62.44	53.8
7	160	(160,240)	7.42	44.57	39.29	12.42	4.32	39.52	32.82
8	160	(240,240)	10.21	47.85	43.17	15.24	4.31	43.27	35.86
9	160	(240,320)	10.59	84.11	78.44	16.54	4.53	78.49	71.62
10	320	(240,480)	11.23	92.9	89.25	18.93	6.01	89.27	84.58
11	320	(320,480)	14.54	93.86	90.87	22.6	5.16	90.88	86.33
12	320	(480,480)	21.46	95.25	92.78	31.33	4.73	92.8	89.1

$$
\begin{aligned}
&= \frac{2}{(1-c_n)^2} N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2) + \frac{2}{(1-c_n)^3} N^{-1}\mathrm{tr}\mathbf{T}_n^{-1} N^{-1}\mathrm{tr}\mathbf{M}_2 + o_p(1), \\
&\quad \frac{d^2}{dz^2}\left[b_n(z)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma})\right] \\
&= 2b_n^3(z)\big(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\big)^2 N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma}) \\
&\quad -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma}) \\
&\quad -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-2}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma}) \\
&\quad +2b_n(z)N^{-1}\mathrm{tr}(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{D}^{-3}(z)\boldsymbol{\Gamma}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma}) \\
&= \frac{2}{(1-c_n)^2} N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma}) \\
&\quad \frac{2}{(1-c_n)^3} N^{-1}\mathrm{tr}\mathbf{T}_n^{-1} N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\boldsymbol{\Gamma}) + o_p(1)
\end{aligned}
$$

TABLE A.3
Comparison of empirical powers (in percentages) of the proposed tests with the tests LC and CLX under Scenario 3 when the observations are from the Gaussian and Gamma population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	10.33	16.73	16.5	16.26	29.81	24.2	29.21
2	40	(120,120)	13.86	19	20.31	20.98	39.05	30.9	37.74
3	40	(120,160)	16.48	24.98	25.04	24.66	52.75	40.68	50.93
4	80	(120,160)	7.55	9.57	9.93	11.78	32.01	19.29	27.8
5	80	(160,240)	9.75	13.74	13.59	15.43	62.21	38.98	55.7
6	80	(240,240)	12.94	15.89	16.66	19.85	79.95	54.76	73.89
7	160	(160,240)	5.9	7.73	7.69	9.03	46.12	24.15	39.45
8	160	(240,240)	6.7	7.6	7.97	10.38	65.57	37.22	58.22
9	160	(240,320)	7	8.38	8.83	11.43	80.26	52.32	73.94
10	320	(240,480)	5.99	6.83	6.73	8.4	84.89	56.62	79.6
11	320	(320,480)	6.58	6.64	6.62	9.36	95.2	74.93	92.28
12	320	(480,480)	6.82	6.43	7.2	10.73	99.45	92.89	98.96
			Gamma population						
1	40	(80,120)	11.49	17.04	17.86	18.13	28.68	24.88	29.29
2	40	(120,120)	14.48	18.61	20.72	22.12	36.89	30.22	37.12
3	40	(120,160)	17.07	24.65	25.64	25.83	52.63	41.4	51.19
4	80	(120,160)	8	10.4	11	12.82	30.57	19.77	27.77
5	80	(160,240)	9.87	13.19	13.65	15.44	59.92	38.61	54.54
6	80	(240,240)	12.96	14.35	15.86	19.69	77.09	52.16	71.18
7	160	(160,240)	6.36	8.13	8.21	9.54	45.46	24.12	39.39
8	160	(240,240)	7.15	8.14	8.95	10.44	61.49	34.69	55.13
9	160	(240,320)	7.52	8.32	9.18	12	78.66	51.86	72.31
10	320	(240,480)	6.04	6.28	6.4	8.75	83.94	56.99	78.14
11	320	(320,480)	6.27	6.57	7.01	9.13	93.99	73.52	91.04
12	320	(480,480)	7.36	6.77	7.79	11.22	99.27	90.62	98.51

and

$$\begin{aligned}
& \frac{d^2}{dz^2}\left[b_n(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})\right] \\
= \; & 2b_n^3(z)\left(N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)\right)^2 N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}) \\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-3}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-1}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}) \\
& -2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{D}^{-2}(z)\mathbf{T}_n)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-2}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}) \\
& +2b_n^2(z)N^{-1}\mathrm{tr}(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{D}^{-3}(z)\mathbf{\Gamma}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}) \\
= \; & \frac{2}{(1-c_n)^2}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma}) \\
& +\frac{2}{(1-c_n)^3}N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})+o_p(1).
\end{aligned}$$

TABLE A.4
Comparison of empirical powers (in percentages) of the proposed tests with the tests LC and CLX under Scenario 4 when the observations are from the Gaussian and Gamma population.

No.	p	(N_1, N_2)	T_{d}	T_{r}	T_{dr}	LC	CLX	T_{drx_1}	T_{drx_2}
			Gaussian population						
1	40	(80,120)	6.37	9.95	9.74	8.32	7.79	10.89	10.48
2	40	(120,120)	6.41	10.48	10.74	8.8	8.54	12.07	11.5
3	40	(120,160)	6.65	11.82	11.33	9.31	10.83	13.04	13.21
4	80	(120,160)	6.8	13.82	12.91	9.94	15.04	16.27	17.58
5	80	(160,240)	8.45	23.84	21.64	12.7	28.51	29.9	33.08
6	80	(240,240)	10.02	27.07	24.07	15.02	42.55	38	45.05
7	160	(160,240)	12.94	23.77	23.41	20.24	33.38	34.02	38.55
8	160	(240,240)	17.53	24.68	27.29	26.22	48.96	45.36	53.45
9	160	(240,320)	20.77	53.66	50.06	30.21	63.57	68.04	73.47
10	320	(240,480)	76.75	99.98	99.95	85.39	94.43	100	99.99
11	320	(320,480)	89.17	100	99.99	94.12	98.54	100	100
12	320	(480,480)	97.78	100	100	98.87	99.91	100	100
			Gamma population						
1	40	(80,120)	7.54	10.86	11.24	9.31	6.48	12.1	10.54
2	40	(120,120)	7.71	11.69	12.31	9.6	7.08	13.28	11.6
3	40	(120,160)	8.35	12.62	13.17	10.4	9.65	14.81	13.56
4	80	(120,160)	7.54	13.82	13.43	10.86	12.96	16.15	17.01
5	80	(160,240)	8.72	22.18	19.94	13.18	26.94	27.86	31.66
6	80	(240,240)	10.6	24.77	22.8	15.52	38.99	35.5	42.24
7	160	(160,240)	12.85	23.98	24.16	19.87	30.69	33.45	37.37
8	160	(240,240)	17.2	25.3	27.62	26.39	45.11	42.28	49.51
9	160	(240,320)	21.31	51.85	48.88	30.98	59.54	65.66	70.76
10	320	(240,480)	76.13	99.96	99.92	84.79	90.68	99.98	99.96
11	320	(320,480)	89.09	99.97	100	93.75	96.66	100	100
12	320	(480,480)	97.18	100	100	98.93	99.55	100	100

Therefore

$$
\begin{aligned}
&\sum_{j=1}^{N} \mathrm{E}_{j-1}\{[(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n^{-2})][(\mathrm{E}_j - \mathrm{E}_{j-1})\mathrm{tr}(\mathbf{B}_n\mathbf{M}_2)^2]\} \\
= \; & -\frac{8}{(1-c_n)^2} N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2) - \frac{8}{(1-c_n)^3} N^{-1}\mathrm{tr}\mathbf{T}_n^{-1} N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2^2) \\
& + \frac{4}{(1-c_n)^2}\left(N^{-1}\mathrm{tr}\mathbf{M}_2\right)^2 - \frac{4\beta_x}{(1-c_n)^2} N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
& - \frac{4\beta_x}{(1-c_n)^3} N^{-1}\mathrm{tr}\mathbf{T}_n^{-1} N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{T}_n\mathbf{M}_2\mathbf{\Gamma}) \\
& - \frac{4\beta_x}{(1-c_n)^2} N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2) N^{-1}\mathrm{tr}(\mathbf{T}_n^{-2} \circ \mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})
\end{aligned}
$$

$$-\frac{4\beta_x}{(1-c_n)^3}N^{-1}\mathrm{tr}(\mathbf{T}_n\mathbf{M}_2)N^{-1}\mathrm{tr}\mathbf{T}_n^{-1}N^{-1}\mathrm{tr}(\mathbf{T}_n^{-1}\circ\mathbf{\Gamma}^{\mathrm{T}}\mathbf{M}_2\mathbf{\Gamma})+o_p(1).$$

Thus, the proof of Lemma A.4 is completed. □

Acknowledgements

The authors thank the Editor, the Associate Editor and the referees for their constructive comments and suggestions that led to a significant improvement of this manuscript.

References

[1] Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley, New York. MR0091588
[2] Bai, Z. D. and Silverstein, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Annals of Probability*, **26** 316-345. MR1617051
[3] Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Annals of Probability*, **32** 553-605. MR2040792
[4] Bai, Z. D., Jiang, D. D., Yao, J. F. and Zheng, S. R. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Annals of Statistics*, **37** 3822-3840. MR2572444
[5] Cai, T. T., Liu, W. D. and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *Journal of the American Statistical Association*, **108** 265-277. MR3174618
[6] Fan, J. Q., Liao, Y. and Yao, J. W. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica*, **83** 1497-1541. MR3384226
[7] Gupta, A. K. and Tang, J. (1984). Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate gaussian models. *Biometrika*, **71** 555-559. MR0775401
[8] Gupta, D. S. and Giri, N. (1973). Properties of tests concerning covariance matrices of normal distributions. *Annals of Statistics*, **1** 1222-1224. MR0350969
[9] Heard, N. A. and Rubin-Delanchy, P. (2018). Choosing between methods for combining p-values. *Biometrika*, **105** 239-246. MR3768879
[10] Jiang, T. F. and Yang, F. (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Annals of Statistics*, **41** 2029-2074. MR3127857
[11] Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *Annals of Statistics*, **40** 908-940. MR2985938
[12] O'Brien, P. C. (1992). Robust procedures for testing equality of covariance matrices. *Biometrics*, **48** 819-827.

[13] Perlman, M. D. (1980). Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations. *Annals of Statistics*, **8** 247-263. MR0560727

[14] Schott, J. R. (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample sizes. *Computational Statistics & Data Analysis*, **51** 6535-6542. MR2408613

[15] Srivastava, M. S. and Yanagihara, H. (2010). Testing the equality of several covariance matrices with fewer observations than the dimension. *Journal of Multivariate Analysis*, **101** 1319-1329. MR2609494

[16] Sugiura, N. and Nagao, H. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *Annals of Mathematical Statistics*, **39** 1686-1692. MR0237033

[17] Wang, Q. W., Silverstein, J. W. and Yao, J. F. (2014). A note on the CLT of the LSS for sample covariance matrix from a spiked population model. *Journal of Multivariate Analysis*, **130** 194-207. MR3229532

[18] Westberg, M. (1985). Combining independent statistical tests. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **34** 287-296.

[19] Yang, Q. and Pan, G. M. (2017). Weighted statistic in detecting faint and sparse alternatives for high-dimensional covariance matrices. *Journal of the American Statistical Association*, **112** 188-200. MR3646565

[20] Zheng, S. R., Bai, Z. D. and Yao, J. F. (2015). Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Annals of Statistics*, **43** 546-591. MR3316190

[21] Zheng, S. R., Lin, R. T., Guo, J. H. and Yin, G. S. (2020). Testing homogeneity of high-dimensional covariance matrices. *Statistica Sinica*, **30** 35-53.