CLT for large dimensional general Fisher matrices and its applications in high-dimensional data analysis

Shurong Zheng\textsuperscript{a}, Zhidong Bai\textsuperscript{a} and Jianfeng Yao\textsuperscript{b}

\textsuperscript{a}KLAS and School of Mathematics and Statistics, Northeast Normal University, China
\textsuperscript{b}Department of Statistics and Actuarial Science, The University of Hong Kong, China

Abstract

Random Fisher matrices arise naturally in multivariate statistical analysis and understanding the properties of its eigenvalues is of primary importance for many hypothesis testing problems like testing the equality between two covariance matrices, or testing the independence between sub-groups of a multivariate random vector. Most of the existing work on random Fisher matrices deals with a particular situation where the population covariance matrices are equal. In this paper, we consider general Fisher matrices with arbitrary population covariance matrices and develop their spectral properties when the dimensions are proportionally large compared to the sample size. The paper has two main contributions: first the limiting distribution of the eigenvalues of a general Fisher matrix is found and secondly, a central limit theorem is established for a wide class of functionals of these eigenvalues. Applications of the main results are also developed for testing hypotheses on high-dimensional covariance matrices.


Key words and phrases. large-dimensional covariance matrices; large-dimensional Fisher matrix; linear spectral statistics; central limit theorem; equality of covariance matrices

1 Introduction

For testing the equality of variances from two populations, a well-known statistic is the Fisher statistic defined as the ratio of two sample variances. Its multivariate counterpart is a random Fisher matrix defined by

\[ F := \mathbf{B}_1 \mathbf{B}_2^{-1} \] (1.1)

where \( \mathbf{B}_1 \) and \( \mathbf{B}_2 \) are \( p \)-dimensional sample covariance matrices from two independent samples, say \( \{\xi_k, 1 \leq k \leq n_1\} \) and \( \{\eta_\ell, 1 \leq \ell \leq n_2\} \) with population covariance matrices \( \Sigma_1 \) and \( \Sigma_2 \), respectively. Fisher matrices, especially their eigenvalues, arise in many hypothesis testing problems in multivariate analysis. Examples include the test of the equality...
hypothesis $\Sigma_1 = \Sigma_2$ where the likelihood ratio (LR) statistic can be written as a functional of the eigenvalues of a Fisher matrix, see Bai et al. [7]. In multivariate analysis of variance (MANOVA), the test on the equality of means is reduced to a statistic depending on a Fisher matrix which is a functional of the “between” sum of squares and the “within” sum of squares (Anderson [1, p. 346]). In multivariate linear regression, the likelihood ratio criterion for testing linear hypotheses about regression coefficients is expressed as a functional of the eigenvalues of a Fisher matrix (Anderson [1, p. 298]). To test the independence between sub-groups of a multivariate population, the LR statistic is a functional of a Fisher matrix defined by sub-matrices of sample covariance matrices (Anderson [1, p. 381]). Fisher matrices appear also in the canonical correlation analysis, see Yang and Pan [27] for a recent account.

This paper concerns the high-dimensional situation where the population dimension $p$ is large compared to the sample sizes $n_1$ and $n_2$. It is now well understood that classical procedures as those presented in Anderson [1] become impracticable or dramatically lose efficiency in high-dimensional data. For example, the deficiency of the Hotelling’s $T^2$ statistic has been reported in Dempster [12] and Bai and Saranadasa [2]. Regarding hypothesis testing on one population high-dimensional covariance matrix, many recent works appeared in the literature, see [14], [22], [23], [24], and [25], among others. About the equality of two population covariance matrices, please see [7], [10], [15] and [31] among others. For two-sample tests on covariance matrices, [10] and [15] use some ad-hoc distance measures to avoid the corresponding Fisher matrix, particularly when $p > n$ where the Fisher matrix is indeed undefined. However, all these works are not scale invariant and require some regularities of the population covariance matrix, e.g. [10] requires the sparsity condition, and [15] requires that the eigenvalues of the population covariance matrices are not dominated by a few of them. Therefore, tests defined by Fisher matrices are still important ones in multivariate analysis. Furthermore, [31] presented an example showing that tests defined by $F$ matrices has larger powers than the method of [15] in a few cases, after corrections using random matrix theory.

In the literature from random matrix theory and assuming that the dimension grows to infinity proportionally to sample sizes, the convergence of the eigenvalues of a Fisher matrix to a limiting distribution has been studied under the equality of two population covariance matrices by several authors, see e.g. [26], [3], [9], [17], [18], [19] and [28]. As an important contribution of the paper, we establish a limiting distribution for the eigenvalues of a general Fisher matrix with arbitrary population covariance matrices. Regarding central limit theorems for linear spectral statistics (or LSS), Chatterjee [11] establishes the existence of a Gaussian limit assuming that the populations are Gaussian. However, he
does not provide explicit formulae for the asymptotic mean and asymptotic covariances of
the Gaussian limit. A closely related work by Bai and Silverstein [5] establishes a CLT
for spectral statistics of a general sample covariance matrices of form $B_1 T_p$ where $B_1$ is a
sample covariance matrix and $T_p$ is a non-random Hermitian matrix. This CLT is later
refined in [16] where the original restriction on the values of the fourth moments of the
population components is removed. However, the CLT in [5] cannot cover spectral statistics
of a Fisher matrix through simply replacing $T_p$ by $B_2^{-1}$ for the reason that the centering
term of this CLT would thus become a random term without an explicit expression. To
overcome this difficulty, Zheng [29] establishes a CLT for LSS of a Fisher matrix which has a
non-random and explicit centering term. In particular, the components of the observations
$\{\xi_i\}$ and $\{\eta_j\}$ can have an arbitrary fourth moment. To our best knowledge, this is the
only CLT reported in the literature for LSS of a Fisher matrix. However, this CLT has a
severe limitation with the assumption that the population covariance matrices are equal
i.e. $\Sigma_1 = \Sigma_2$. Although the derivation of this CLT is complex and highly non trivial,
it has a small impact on the statistical problems mentioned above where the population
covariance matrices $\Sigma_i$ can be arbitrary and not necessarily equal. Specifically, for the test
of the equality hypothesis “$\Sigma_1 = \Sigma_2$”, this CLT enables us to find the distribution of the
LR statistic under the null hypothesis, but not under the alternative “$\Sigma_1 \neq \Sigma_2$”, that is,
the asymptotic test size of the test can be found by this CLT but not the power function.

The main contribution of the paper is the establishment of a central limit theorem for
linear spectral statistics $\{W_n\}$ of a general Fisher matrix where the population covariance
matrices $\Sigma_i$ are arbitrary. Under this scheme and as a preparatory step, we also establish
a limiting distribution for their eigenvalues and give an explicit equation satisfied by its
Stieltjes transform. Due to the fact that the population covariance matrices are arbitrary,
the proofs of these results have required several new techniques compared to the existing
literature on the central limit theory although the general scheme remains similar to the
one used in [5, 29]. A significantly different tool used here is another CLT reported in [30]
for random matrix of type $S^{-1}T_p$ where $S$ is a standard sample covariance matrix (with
i.i.d. standardised components) and $T_p$ a nonnegative definite and deterministic Hermitian
matrix. These two papers are related to each other but focus on different random matrices.

The paper is organized as follows. In Section 2 we first introduce the asymptotic scheme
and the technical assumptions used, and then establish the limiting spectral distribution
of the eigenvalues. Section 3 presents the CLT for linear spectral statistics of general
Fisher matrices which is the main result of the paper. Section 4 gives two algorithms
for approximation of the limiting spectral density, the asymptotic mean and covariance
functions involved in the new CLT. In Section 5, applications of the main results are
proposed for hypothesis testing and confidence intervals about high-dimensional covariance matrices. Technical lemmas and proofs are postponed to Appendix B.

2 Limiting spectral distribution of large dimensional general \( F \)-matrices

Following Bai and Silverstein [5] and Zheng [29], we will impose the following structure on the observation model. Assume that the samples can be expressed as

\[
\xi_k = Q_1 X_k, \quad 1 \leq k \leq n_1; \quad \eta_\ell = Q_2 Y_\ell, \quad 1 \leq \ell \leq n_2;
\]

where the observation matrices

\[
X := (X_{1}, \ldots, X_{n_1}) = (X_{jk} : 1 \leq j \leq p, 1 \leq k \leq n_1),
\]
\[
Y := (Y_{1}, \ldots, Y_{n_2}) = (Y_{j\ell} : 1 \leq j \leq p, 1 \leq \ell \leq n_2),
\]

are the upper-left corners, of size \( p \times n_1 \) and \( p \times n_2 \), of two independent arrays of independent random variables \( \{X_{jk}, j, k = 1, 2, \ldots\} \) and \( \{Y_{jk}, j, k = 1, 2, \ldots\} \), respectively, and \( Q_i \) is any square matrix such that \( Q_i Q_i^* = \Sigma_i, i = 1, 2 \) where \( * \) denotes the (complex) adjoint of a matrix (i.e. transpose and conjugate). The corresponding sample covariance matrices become

\[
B_1 = \frac{1}{n_1} \sum_{k=1}^{n_1} \xi_k \xi_k^* = Q_1 S_1 Q_1^*, \quad \text{with} \quad S_1 = \frac{1}{n_1} \sum_{k=1}^{n_1} X_k X_k^*, \quad (2.1)
\]
\[
B_2 = \frac{1}{n_2} \sum_{\ell=1}^{n_2} \eta_\ell \eta_\ell^* = Q_2 S_2 Q_2^*, \quad \text{with} \quad S_2 = \frac{1}{n_2} \sum_{\ell=1}^{n_2} Y_\ell Y_\ell^*. \quad (2.2)
\]

Because \( F = B_1 B_2^{-1} \) has the same eigenvalues as \( S_1 T_{p,1/2}^* S_2^{-1} T_{p,1/2} \) where \( T_{p,1/2} = Q_2^{-1} Q_1 \), we can define as well the Fisher matrix to be \( F := S_1 T_{p,1/2}^* S_2^{-1} T_{p,1/2} \). It is also noticed that obviously, the matrix \( S_2 \) should be invertible (almost surely) so that in our asymptotic analysis, we will assume \( n_2 > p \) for large \( p \) and \( n_2 \).

Throughout the paper, the empirical spectral distribution (or ESD) of a complex \( p \times p \) matrix \( A \) is the probability measure \( F_n^A(x) = p^{-1} \sum_{j=1}^{p} \delta_{\lambda_j} \) where \( \{\lambda_j\}_{j=1}^{p} \) are the eigenvalues of \( A \), and \( \delta_a \) denotes the Dirac mass at a point \( a \). We consider the following assumptions.

**Assumption [A]** The two double arrays \( \{X_{ki}, i, k = 1, 2, \ldots\} \) and \( \{Y_{ki}, i, k = 1, 2, \ldots\} \) consist of independent but not necessarily identically distributed random variables with mean 0 and variance 1.
Assumption [B1] For any fixed \( \eta > 0 \) and when \( n_1, n_2, p \to \infty \),
\[
\frac{1}{n_1 p} \sum_{j=1}^{p} \sum_{k=1}^{n_1} E \left[ |X_{jk}|^2 I_{|X_{jk}| \geq \eta \sqrt{n_1}} \right] \to 0, \quad \frac{1}{n_2 p} \sum_{j=1}^{p} \sum_{k=1}^{n_2} E \left[ |Y_{jk}|^2 I_{|Y_{jk}| \geq \eta \sqrt{n_2}} \right] \to 0.
\]

(2.3)

Assumption [B2] If the two arrays are either both real, we then set the indicator \( \kappa = 2 \); or both complex, we then set \( \kappa = 1 \), with homogeneous 4th moments: \( E|X_{jk}|^4 = 1 + \kappa + \beta_x + o(1) \), \( E|Y_{jk}|^4 = 1 + \kappa + \beta_y + o(1) \). Moreover, for any fixed \( \eta > 0 \) when \( n_1, n_2, p \to \infty \),
\[
\frac{1}{n_1 p} \sum_{j=1}^{p} \sum_{k=1}^{n_1} E \left[ |X_{jk}|^4 I_{|X_{jk}| \geq \eta \sqrt{n_1}} \right] \to 0, \quad \frac{1}{n_2 p} \sum_{j=1}^{p} \sum_{k=1}^{n_2} E \left[ |Y_{jk}|^4 I_{|Y_{jk}| \geq \eta \sqrt{n_2}} \right] \to 0.
\]

(2.4)

In addition, \( EX^2_{jk} = o(n_1^{-1}) \), \( EY^2_{jk} = o(n_2^{-1}) \) when both arrays \( \{X_{jk}\} \) and \( \{Y_{jk}\} \) are complex.

Assumption [C] The sample sizes \( n_1, n_2 \) and the dimension \( p \) grow to infinity in such a way that
\[
y_{n_1} := p/n_1 \to y_1 \in (0, +\infty), \quad y_{n_2} := p/n_2 \to y_2 \in (0, 1)
\]

(2.5)

where \( S_2 \) has the inverse matrix \( S_2^{-1} \).

Assumption [D] The matrices \( T_p = T_{p,1/2}^* T_{p,1/2} \) are non-random and nonnegative definite Hermitian matrices and the sequence \( \{T_p\} \) is bounded in spectral norm. Moreover, the ESD \( H_p \) of \( T_p \) tends to a proper probability measure \( H \) when \( p \to \infty \).

The assumptions (2.3) and (2.4) are standard Lindeberg type conditions which are necessary for the existence of the limiting spectral distribution for \( \mathbf{F} \), and for the CLT for LSS of \( \mathbf{F} \), respectively. Moreover, under these conditions, the variables \( \{X_{ik}\} \) and \( \{Y_{ik}\} \) can be truncated at size \( \eta_p \sqrt{p} \) \( (\eta_p \Downarrow 0) \) without altering asymptotic results.

The following notations are used throughout the paper:
\[
n = (n_1, n_2), \quad \mathbf{y} = (y_{n_1}, y_{n_2}), \quad \mathbf{y} = (y_1, y_2), \quad h^2 = y_1 + y_2 - y_1 y_2.
\]

In the sequel, the limiting results will be investigated under the regime (2.5) that will be simply referred as \( n \to \infty \). Some useful concepts are now recalled. The Stieltjes transform of a positive Borel measure \( G \) on the real line is defined by
\[
m_G(z) \equiv \int \frac{1}{\lambda - z} dG(\lambda), \quad z \in \mathbb{C}^+ = \{z : z \in \mathbb{C}, \Im(z) > 0\}.
\]

(2.6)
This transform has a natural extension to the lower-half plane by the formula

\[ m_G(z) = \overline{m_G}(\overline{z}), \quad \text{for} \quad z \in \mathbb{C}^- = \{z : z \in \mathbb{C}, \Im(z) < 0\}. \]

In addition to \( F \), we will also need several other matrices. Table 1 contains the notations used in the sequel for the characteristics of these matrices: ESD, LSD, the associated Stieltjes transforms and LSS.

Table 1: Notations for distributions, Stieltjes transforms (S.T.) and linear spectral statistics (LSS) of random matrices.

<table>
<thead>
<tr>
<th>Matrix</th>
<th>ESD / S.T.</th>
<th>LSD / S.T.</th>
<th>LSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F = S_1T_{p,1/2}^*S_2^{-1}T_{p,1/2} )</td>
<td>( U_n(x) / m_n(z) )</td>
<td>( U_y(x) / m_y(z) )</td>
<td>( U_n(f) = \int f(x)dU_n(x) )</td>
</tr>
<tr>
<td>( X^*T_{p,1/2}^*S_2^{-1}T_{p,1/2}X/n_1 )</td>
<td>( U_n(x) / m_n(z) )</td>
<td>( U_y(x) / m_y(z) )</td>
<td>( \int f(x)dU_y(x) )</td>
</tr>
</tbody>
</table>

where \( m_n(z) = \int \frac{1}{\lambda - z}dU_n(x) \), \( m_y(z) = \int \frac{1}{\lambda - z}dU_y(x) \), \( m_n(z) = \int \frac{1}{\lambda - z}dU_n(x) \), \( m_y(z) = \int \frac{1}{\lambda - z}dU_y(x) \). \( f \) is an analytic function, \( U_n(f) = p^{-1} \sum_{j=1}^p f(\lambda_j^F) \) with the eigenvalues \( \{\lambda_j^F\}_{j=1}^p \) of \( F \) and \( U_{y_n}(x) \) is obtained by substituting \( y_n = (y_{n_1}, y_{n_2}) \) for \( y = (y_1, y_2) \) in \( U_y(x) \).

The matrices \( F \) and \( X^*T_{p,1/2}^*S_2^{-1}T_{p,1/2}X/n_1 \) are companion matrices each other sharing same non null eigenvalues so that we have

\[ m_n(z) = \frac{1}{z}m_n(z) + y_{n_1}m_n(z), \quad \text{(2.7)} \]

\[ m_y(z) = \frac{1}{z}m_y(z) + y_1m_y(z). \quad \text{(2.8)} \]

Furthermore, when \( \Sigma_1 = \Sigma_2 \), i.e., \( T_p = I_p \), it is well-known that the LSD \( U_y \) of \( F \) and its Stieltjes transform \( m_y(z) \) can be found on Page 79 of Bai and Silverstein [6]. As a first result of the paper, we prove the existence of \( U_y \) and one of its characteristics for general Fisher matrix \( F \) where \( T_p \) is a Hermitian matrix.

**Theorem 2.1** Under Assumptions [A], [B1], [C] and [D], The Fisher matrix

\[ F = S_1T_{p,1/2}^*S_2^{-1}T_{p,1/2} \]

has a non-random LSD \( U_y \). Moreover, \( U_y \) is characterized by the fact that the Stieltjes transform \( m_y(z) \) of its companion measure \( U_y \) is the unique solution to the equation

\[ \frac{h^2m_0(z)}{y_2(-1 + y_2 \int \frac{m_0(z)dH(t)}{t + m_0(z)})} + \frac{y_1}{y_2}m_0(z), \quad z \in \mathbb{C}^+, \quad \text{(2.9)} \]
where $m_0(z) = m_{y^2}(-m_y(z))$ and $m_{y^2}(z)$ satisfies the equation $z = -(m_{y^2}(z))^{-1} + y_2 \int (t + m_{y^2}(z))^{-1} dH(t)$.

The proof of this theorem is given in Appendix B.1.

Remark 2.1 If the condition [B1] in Theorem 2.1 is strengthened to [B2], following Bai and Yin [8], after truncation, one can show that with probability 1, $S_2$ is invertible. However, under the condition [B1], it may be singular with a small positive probability. In this case, Theorem 2.1 remains true if $S_2^{-1}$ is understood as the Moore-Penrose generalized inverse of $S_2$.

Remark 2.2 For a given $z \in \mathbb{C}^+$, the equation (2.9) has a unique solution $m_0$ such that $\Im(m_0) < 0$.

In fact by Silverstein [20], $m_y(z)$ is the unique solution to the equation

$$z = -\frac{1}{m_y(z)} + y_1 \int \frac{xdG_{y^2}(x)}{1 + xm_y(z)},$$

(2.10)

subject to the condition that $\Im(\frac{1-y_1}{z} + m_y(z))\Im(z) > 0$. Here, $G_{y^2}$ denotes the limiting spectral distribution of the random matrix $T^*_{p,1/2}S_2^{-1}T_{p,1/2}$.

In the sequel, for brevity, the notations $m_y(z)$ and $m_y^*(z)$ will be simplified to $m(z)$ and $m$, or even to $m$ and $m$, respectively, if no confusion would be possible. We will use the notation $G_{y^2_n}$ that is obtained by substituting $y_{n^2} = p/n^2$ for $y_2$ in $G_{y^2}$.

3 CLT for LSS of large dimensional general Fisher matrices

We consider a centered version of the LSS

$$p[U_n(f) - U_{y_n}(f)] = \sum_{j=1}^{p} f(\lambda_j^F) - p \int f(x)dU_{y_n}(x)$$

(3.1)

with the eigenvalues $\{\lambda_j^F\}_{j=1}^{p}$ of $F$. Note that (3.1) is translation-invariant in $f$, we may assume $f(0) = 0$. Due to the exact separation theorem (see Bai and Silverstein [4]), with probability one, for large enough $n_j$ and $p$, the possible point mass at the origin of $U_n$ will coincide exactly with that of $U_{y_n}$. By Theorem 1.1 of Silverstein and Choi [21], $U_y$ has the
continuous density \( u_y \) except \( x = 0 \). Therefore, we can restrict the integral (3.1) to their continuous components on \((0, \infty)\), i.e.

\[
p [U_n(f) - U_{y_n}(f)] = \sum_{j=1}^p f(\lambda^F_j) I(\lambda^F_j > 0) - p \int f(x) u_{y_n}(x) \, dx
\]  

(3.2)

where \( u_{y_n}(x) \) is the density of \( U_{y_n}(x) \).

Regarding the central limit theory on linear spectral statistics of random matrices, it has been well-known ([5, 16, 29]) that the mean and covariance parameters of the limiting Gaussian distribution depend on the values of the fourth moments of \( \{X_{jk}\} \) and \( \{Y_{jk}\} \). When these moments match the Gaussian case, i.e. \( \beta_x = 0 \) or \( \beta_y = 0 \) in Assumption [B2], the limiting parameters have a simpler expression. Otherwise, they have a more involved expression that depends on other limiting functionals of sample covariance matrices. More specifically, if \( \beta_x \neq 0 \), we will need the existence of the following limits

\[
\frac{1}{p} \sum_{i=1}^p E \left[ e_i' T^*_{p,1/2} S_2^{-\frac{1}{2}} D_1^{-\frac{1}{2}} S_2^{-\frac{3}{2}} T_{p,1/2} e_i \right] \cdot e_i' T^*_{p,1/2} S_2^{-\frac{3}{2}} D_1^{-1} \left( m(z) S_2^{-1/2} T_p S_2^{-1/2} + I_p \right)^{-1} S_2^{-\frac{1}{2}} T_{p,1/2} e_i \left| S_2 \right] \rightarrow h_{m1}(z),
\]

(3.3)

\[
\frac{1}{n_1 p} \sum_{j=1}^{n_1} \sum_{i=1}^p e_i' T^*_{p,1/2} S_2^{-\frac{3}{2}} \left[ E_j D_{1j}^{-1}(z_1) \right] S_2^{-\frac{1}{2}} T_{p,1/2} e_i \rightarrow h_{v1}(z_1, z_2),
\]

(3.4)

where \( E_j \) denotes the conditional expectation given \( X_1, \ldots, X_j \) and \( S_2 \), and if \( \beta_y \neq 0 \), we will need the existence of the limits

\[
\frac{1}{p} \sum_{i=1}^p E e_i' (m(z) T_p + S_2)^{-1} e_i \cdot e_i' (m(z) T_p + S_2)^{-1} T_p (m(z) T_p + m(z)m_0(z)I_p)^{-1} e_i \\
\rightarrow h_{m2}(m^{-1}(z)),
\]

(3.5)

\[
\frac{1}{n_2 p} \sum_{j=1}^{n_2} \sum_{i=1}^p e_i' E_j (m(z_1) T_p + S_{2,j})^{-1} e_i \cdot e_i' E_j (m(z_2) T_p + S_{2,j})^{-1} e_i \\
\rightarrow h_{v2}(m^{-1}(z_1), m^{-1}(z_2))
\]

(3.6)

where \( E_j \) denotes the conditional expectation given \( Y_1, \ldots, Y_j \). Here

\[
S_{2,j} = S_2 - \eta_j \eta_j^*, \quad D_1(z) = \sum_{i=1}^{n_1} \gamma_i \gamma_i^* - z I_p, \quad D_{1j}(z) = D_1(z) - \gamma_j \gamma_j^*,
\]

8
and $\mathbf{e}_i$ denotes the $i$-th vector of the canonical basis of $\mathbb{C}^p$ where $\eta_j = n_2^{-1/2}Y_j$ and $\gamma_j = n_1^{-1/2}S^{-1/2}_{2j}T_{p,1/2}X_j$.

The following CLT is the main result of the paper.

**Theorem 3.1** Under the Assumptions [A], [B2], [C] and [D], assume that the limits (3.3)-(3.4) exist whenever $\beta_\gamma \neq 0$, and the limits (3.5)-(3.6) exist whenever $\beta_\gamma \neq 0$. Let $f_1, \ldots, f_s$ be s functions analytic in an open domain of the complex plane that enclosed the support interval $[c_1, c_2]$ of the continuous component of the LSD $U_y$. Then, as $n \to \infty$, the random vector $$\{p[U_n(f_j) - U_n(f_j)], 1 \leq j \leq s\},$$ converges weakly to a Gaussian vector $(X_{f_1}, \ldots, X_{f_s})$ with the mean function

$$EX_f = \frac{\kappa - 1}{4\pi i} \oint_C f(z) d\log \left( \frac{h^2}{y_1} - \frac{y_1}{y_2} \cdot \frac{1 - y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t)}{1 - y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t)} \right),$$

$$- \frac{\kappa - 1}{4\pi i} \oint_C f(z) d\log \left( 1 - y_2 \int \frac{m_0(z)dH(t)}{(t + m_0(z))^2} \right),$$

$$+ \frac{\kappa - 1}{4\pi i} \oint_C f(z) d\log \left( 1 - y_2 \int \frac{m_0(z)dH(t)}{(t + m_0(z))^2} \right),$$

$$+ \frac{\kappa - 1}{4\pi i} \oint_C f(z) d\log \left( 1 - y_2 \int \frac{m_0(z)dH(t)}{(t + m_0(z))^2} \right),$$

$$+ \frac{\kappa - 1}{4\pi i} \oint_C f(z) d\log \left( 1 - y_2 \int \frac{m_0(z)dH(t)}{(t + m_0(z))^2} \right),$$

and the covariance function

$$\text{Cov}(X_{f_1}, X_{f_2}) =$$

$$- \frac{\beta_\gamma y_1}{4\pi^2} \oint_{C_1} \oint_{C_2} f_i(z_1)f_j(z_2) \frac{\partial^2 \left[ z_1z_2m(z_1)m(z_2)h_{\nu_1}(z_1, z_2) \right]}{\partial z_1 \partial z_2} dz_1 dz_2$$

$$- \frac{\kappa}{4\pi^2} \oint_{C_1} \oint_{C_2} \frac{f_i(z_1)f_j(z_2)}{(m_0(z_1) - m_0(z_2))^2} dm_0(z_1) dm_0(z_2)$$

$$- \frac{\beta_\gamma y_2}{4\pi^2} \oint_{C_1} \oint_{C_2} f_i(z_1)f_j(z_2) \frac{\partial^2 \left[ m(z_1)m_0(z_1)m(z_2)m_0(z_2)h_{\nu_2}(z_1, z_2) \right]}{\partial z_1 \partial z_2} dz_1 dz_2,$$

(3.8)

where the contours $C$, $C_1$ and $C_2$ all enclose the support of $U_y$, and $C_1$ and $C_2$ are disjoint.

Similar to CLT’s developed in [5, 29], all the limiting parameters depend on contour integrals using the associated Stieltjes transforms. Some specific examples of calculations of such integrals can be found in these references.
When \( T_p \)'s are diagonal, we find explicit expressions for the limiting functions \( h_{mj}(z) \) and \( h_{vj}(z_1, z_2), \ j = 1, 2 \). This in turn simplifies the expressions of limiting mean and covariance functions in the CLT. Two propositions about \( h_{mj}(z) \) and \( h_{vj}(z_1, z_2) \) are given in Appendix A. In this case, the limiting mean and covariance function in Theorem 3.1 are simplified as follows.

**Theorem 3.2** Under the conditions of Theorem 3.1 and assuming that the matrices \( T_p \)'s are diagonal, we obtain that as \( n \to \infty \), the random vector

\[
\{ p[U_n(f_j) - U_y(f_j)] \ , \ \ 1 \leq j \leq s \},
\]

converges weakly to a Gaussian vector \((X_{f_1}, \ldots, X_{f_s})\) with the mean function

\[
EX_f = \frac{\kappa - 1}{4\pi i} \int_C f(z) \ d\log \left( \frac{h^2}{y_2} - \frac{y_1}{y_2} \cdot \left( 1 - y_2 \int \frac{m_0(z) \ dm H(t)}{(t + m_0(z))^2} \right)^2 \right) + \kappa - 1 \int_C f(z) \ d\log \left( 1 - y_2 \int \frac{m_0^2(z) \ dm H(t)}{(t + m_0(z))^2} \right) + \frac{\beta z y_1}{2\pi i} \int_C f(z) \int \left( \frac{t^2 \ dm H(t)}{(t + m_0(z))^3} \right) dm_0(z) + \frac{\beta y}{4\pi i} \int_C f(z) \left( 1 - y_2 \int \frac{m_0^2(z) \ dm H(t)}{(t + m_0(z))^2} \right) d\log \left( 1 - y_2 \int \frac{m_0^2(z) \ dm H(t)}{(t + m_0(z))^2} \right)
\]

and the covariance function

\[
\text{Cov}(X_{f_i}, X_{f_j}) = -\frac{\kappa}{4\pi^2} \int_{C_1} \int_{C_2} \frac{f_i(z_1) f_j(z_2)}{(m_0(z_1) - m_0(z_2))^2} dm_0(z_1) dm_0(z_2) - \frac{\beta z y_1 + \beta y y_2}{4\pi^2} \cdot \int_{C_1} \int_{C_2} f_i(z_1) f_j(z_2) \left( \int \frac{t^2 \ dm H(t)}{(t + m_0(z_1))^2 (t + m_0(z_2))^2} \right) dm_0(z_1) dm_0(z_2)
\]

where the contours \( C, C_1 \) and \( C_2 \) all enclose the support of \( U_y \), and \( C_1 \) and \( C_2 \) are disjoint.

**Remark 3.1** The contours in Theorem 3.1 and Proposition A.1-A.2 are taken in the \( z \) space. In this case, the contours can be arbitrary provided that they enclose the support of the LSD \( U_y \). Since the integrands are functions of \( m_0 \), thus the integrals can be taken in the \( m_0 \) space using the change of variable \( z \mapsto m_0(z) \).

**Remark 3.2** When \( T_p \) is an identity matrix, (A.3) and (A.4) are the same as (3.6) and (3.7) in Zheng [29]. That is, Theorem 3.2 in Zheng [29] is a special case of Theorem A.1 in this paper when \( T_p = I_p \).
4 Evaluation of the asymptotic parameters $EX_f$, Cov $(X_{f_i}, X_{f_j})$ and the limiting density $u_y(x)$

Notice that practical applications of Theorem 3.1 or Proposition A.1 require the knowledge of the limiting spectral density $u_y(x)$, the asymptotic mean $EX_f$ and covariance function $Cov(X_{f_i}, X_{f_j})$. In particular, the last two functions depend on some non trivial contour integrals. In the simple case where $T_p = I_p$ and for simple functions like $f(x) = x^j$ or $f(x) = \log(x)$, analytical results can be found exactly, see [29]. However, this is a very particular case and for general population matrices or more complex functions $f$, such exact results are not available. In this section, we introduce some numerical procedures to approximate these asymptotic parameters while deliberately place ourselves in the context of practical applications with real data sets. In such a situation, the sample sizes and dimension of data $(n_1, n_2, p)$ are given and the empirical spectral distribution $H_p$ of $T_p = T_p, 1/2 T_p^* 1/2$ is known. In this section, we denote the eigenvalues of $T_p$ simply by $\{\lambda_j^0\}_{j=1}^p$ so that $H_p(t) = p^{-1} \sum_{j=1}^p I(\lambda_j^0 \leq t)$. However in such a concrete application situation, the LSD $H$ is never known and we need an estimate of $H$. A very reasonable and widely used estimate of $H$ is indeed just $H_p$.

Notice that the parameters $u_y(x)$, $EX_f$ and Cov($X_{f_i}, X_{f_j}$) all depend on the Stieltjes transform $m_0(z)$. By (2.9), we have

$$z = \frac{h^2 m_0(z)}{y_2 (-1 + y_2 \int \frac{m_0(z)dH(t)}{t + m_0(z)})} + \frac{y_1}{y_2} m_0(z).$$

That is,

$$m_0(z) = z \left( \frac{h^2}{y_2 (-1 + y_2 \int \frac{m_0(z)dH(t)}{t + m_0(z)})} + \frac{y_1}{y_2} \right)^{-1}$$

which leads to numerical approximations of the Stieltjes transform $m_0(z)$. Choose $s_0(z)$ be the initial values of $m_0(z)$ and iterate for $k \geq 0$ the above mapping

$$s_{k+1}(z) = z \left( \frac{h^2}{y_2 (-1 + y_2 \int \frac{s_k(z)dH(t)}{t + s_k(z)})} + \frac{y_1}{y_2} \right)^{-1}$$

(4.1)

until convergence and let $s_k(z)$ be the final values. Define the approximate value for $m_0(z)$ be $s_k(z)$. Next, the limiting spectral density $u_y(x)$ can be approximated as follows.
Remark 4.1 By Theorem B.10 of Bai and Silverstein [6], we have

\[ m(z) = \frac{1}{m_0(z)} - y_2 \int \frac{dH(t)}{t + m_0(z)} \approx \frac{1}{m_0(z)} - y_2 p^{-1} \sum_{j=1}^{p} \frac{1}{\lambda_j + m_0(z)} \]  
(4.2)

and

\[ u_y(x) = \frac{1}{\pi y_1} \lim_{\varepsilon \to 0^+} \Im(m(x + \varepsilon i)) \]  
(4.3)

where \( \{\lambda_j^p\}_{j=1}^{p} \) are the eigenvalues of \( T_p \).

The following remark will give a simplified form of the asymptotic mean function \( E X_f \) and asymptotic covariance function \( \text{Cov}(X_{f_t}, X_{f_j}) \).

Remark 4.2 In Theorem A.1, assuming that \( T_p \)'s are diagonal, then by (A.3)-(A.4) and (1.7) in Bai and Silverstein [5], the mean and covariance functions have the alternative expressions

\[ E X_f = -\frac{\kappa - 1}{4\pi i} \int_C f'(z) \log \left( \frac{h^2}{y_2} - \frac{y_1}{y_2} \cdot \frac{1 - y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t)}{1 - y_2 \int \frac{m_0(z)}{(t + m_0(z))^2} dH(t)} \right) \, dz \]

\[ -\frac{\kappa - 1}{4\pi i} \int_C f'(z) \log \left( 1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2} \right) \, dz \]

\[ -\frac{\beta_x y_1}{2\pi i} \int_C f(z) \frac{1}{(t + m_0(z))^2} \int \frac{t^2 dH(t)}{dH(t)} \left( -1 + y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t) \right)^2 \, dz \]

\[ + \frac{\beta_y y_2}{2\pi i} \int_C f(z) \frac{m_0(z)}{1 - y_2 \int \frac{m_0(z)}{(t + m_0(z))^2} dH(t)} \left( h^2 - \frac{y_1}{y_2} - \frac{m_0^2(z)}{\int (t + m_0(z))^2 \, dH(t)} \right) \, dz \]

(4.4)

and

\[ \text{Cov}(X_{f_t}, X_{f_j}) = -\frac{\kappa}{2\pi^2} \int \int f'_t(x) f'_j(y) \log \left| \frac{m(x) - m(y)}{m(x) - m(y)} \right| \, dx \, dy \]

\[ -\frac{\beta_x y_1 + \beta_y y_2}{4\pi^2} \int \int f'_t(z_1) f'_j(z_2) \int \frac{t^2 dH(t)}{(t + m_0(z_1))(t + m_0(z_2))} \, d z_1 \, d z_2. \]

(4.5)

Combining Remarks 4.1 and 4.2, we now describe the general procedure to approximate the limiting spectral density \( u_y(x) \), the asymptotic mean and covariance functions.
Algorithm 1: approximating the limiting spectral density $u_y(x)$

Cut the support set $[c_1, c_2]$ of the LSD of the Fisher matrix $F$ into a mesh set as

$$\mathcal{A} = \left\{ z_j = x_j + \varepsilon i, x_j = c_1 + \frac{(c_2 - c_1) j}{K}, \ j = 0, \ldots, K \right\} ,$$

where $\varepsilon$ is a small step size, e.g. $10^{-3}$. By (4.1), we obtain $m_0(z_j)$ with $z_j \in \mathcal{A}$. By (4.2), we obtain $m(z_j)$ with $z_j \in \mathcal{A}$. Then by (4.3), let

$$u_y(x_j) \approx \frac{1}{\pi y_1^2} \Im(m(z_j)) \tag{4.6}$$

which is an approximation of the density $u_y(x_j)$. Then we have

$$\int f(x)u_y(x)dx \approx \frac{c_2 - c_1}{K} \sum_{l=1}^{K} f(x_l)u_y(x_l). \tag{4.7}$$

Algorithm 2: approximating the asymptotic mean function (4.4) and covariance function (4.5)

**Step 1.** Choose two disjoint contours $C_1$ and $C_2$ both enclosing the support $[c_1, c_2]$ of $u_y$ as depicted on Figure 1 where $\varepsilon$ and $\zeta$ are small numbers, e.g. $\varepsilon = \zeta = 10^{-3}$.

**Step 2.** Let $K_1, K_2$ be large integers, e.g. $10^3$. Then $C_1$ and $C_2$ are cut into a grid set as

$$\mathcal{A}_1 = \left\{ z_k = c_1 - \varepsilon + \left( \zeta - \frac{2\zeta k}{K_1} \right) i, \ k = 0, \ldots, K_1 \right\} ,$$

$$z_{K_1+j} = c_1 - \varepsilon + \frac{(c_2 - c_1 + 2\varepsilon) j}{K_2} - \zeta i, \ j = 0, \ldots, K_2$$

$$z_{K_1+K_2+k} = c_2 + \varepsilon + \left( -\zeta + \frac{2\zeta k}{K_1} \right) i, \ k = 0, \ldots, K_1$$

$$z_{2K_1+K_2+j} = c_2 + \varepsilon - \frac{(c_2 - c_1 + 2\varepsilon) j}{K_2} + \zeta i, \ j = 0, \ldots, K_2 \right\} ,$$

$$\mathcal{A}_2 = \left\{ z_k = c_1 - \frac{\varepsilon}{2} + \left( \frac{\zeta}{2} - \frac{\zeta k}{K_1} \right) i, k = 0, \ldots, K_1 \right\} ,$$

$$z_{K_1+j} = c_1 - \frac{\varepsilon}{2} + \frac{(c_2 - c_1 + \varepsilon) j}{K_2} - \frac{\zeta}{2} i, j = 0, \ldots, K_2$$

$$z_{K_1+K_2+k} = c_2 + \frac{\varepsilon}{2} + \left( -\frac{\zeta}{2} + \frac{\zeta k}{K_1} \right) i, k = 0, \ldots, K_1$$

$$z_{2K_1+K_2+j} = c_2 + \frac{\varepsilon}{2} - \frac{(c_2 - c_1 + 2\varepsilon) j}{K_2} + \frac{\zeta}{2} i, j = 0, \ldots, K_2 \right\} ,$$
Figure 1: Contours used to approximate the asymptotic mean and covariance.

\[ z_{2K_1+K_2+j} = c_2 + \frac{\varepsilon}{2} - \frac{(c_2 - c_1 + \varepsilon)j}{K_2} + \frac{\zeta}{2}i, \ j = 0, \ldots, K_2 \}

**Step 3.** By (4.1), we obtain \( m_0(z_j) \). By (4.2), we obtain \( m(z_j) \). Then the mean and covariance functions are approximated by

\[ \text{EX}_f \approx \frac{\kappa - 1}{4\pi} \sum_{j=0}^{2K_1+2K_2-1} \Im \left[ f'(z_j) \log \left( \frac{h^2}{y_2} - \frac{y_1}{y_2} \cdot \frac{1 - y_2 \int \frac{m_0(z_j)}{t+m_0(z_j)} dH(t)}{1 - y_2 \int \frac{m_0(z_j)}{t+m_0(z_j)} dH(t)} \right) (z_{j+1} - z_j) \right] \\
- \frac{\kappa - 1}{4\pi} \sum_{j=0}^{2K_1+2K_2-1} \Im \left[ f'(z_j) \log \left( 1 - y_2 \int \frac{m_0^2(z_j) dH(t)}{(t + m_0(z_j))^2} \right) (z_{j+1} - z_j) \right] \\
- \frac{\beta x y_1}{2\pi} \sum_{j=0}^{2K_1+2K_2-1} \Im \left[ f(z_j) \int \frac{t^2}{(t+m_0(z_j))^5} dH(t) \cdot \left( 1 + y_2 \int \frac{m_0(z_j)}{t+m_0(z_j)} dH(t) \right)^2 (z_{j+1} - z_j) \right] \\
\]
and
\[
\begin{align*}
\text{Cov}(X_{f_i}, X_{f_j}) \approx & -\frac{\kappa}{2\pi^2} \sum_{k,l=0}^{K_1} I_{(k\neq l)} \Re \left( f'_i(z_k) f'_j(z_l) \log \left| \frac{m(x) - m(y)}{m(z_k) - m(z_l)} \right| \right) (c_2 - c_1)^2 / K^2_1 \\
- & \frac{\beta_2 y_2 + \beta_2 y_2}{4\pi^2} \sum_{j,k=0}^{2K_1+2K_2-1} \Re \left( f'_i(z_{1k}) f'_j(z_{2l}) \int t^2 dH(t)(z_{1k+1} - z_{1k})(z_{2l+1} - z_{2l}) \right)
\end{align*}
\]
\tag{4.9}

where \( z_k \in A_1, z_{1k} \in A_1 \) and \( z_{2l} \in A_2 \) and \( I(\cdot) = 1 \) is an indicator function.

### 5 Applications to high-dimensional statistical analysis

In this section, we introduce two applications of the theory developed in the paper to two high-dimensional statistical problems.

#### 5.1 Power function for testing the equality of two high-dimensional covariance matrices

First we consider the two-sample test of the hypothesis that two high-dimensional covariance matrices are equal, i.e.

\[
H_0 : \Sigma_1 = \Sigma_2 \quad \text{v.s.} \quad H_1 : \Sigma_1 \neq \Sigma_2.
\tag{5.1}
\]

By Zheng, Bai and Yao [31], the likelihood ratio test statistic for (5.1) is

\[
T_n = \sum_{i=1}^{p} \log(y_{N_1} + y_{N_2} \lambda_i) - \sum_{i=1}^{p} \frac{y_{N_2}}{y_{N_1} + y_{N_2}} \log \lambda_i - \log(y_{N_1} + y_{N_2})
\]

where \( \lambda_i \)’s are eigenvalues of a Fisher matrix \( AB^{-1} \) with

\[
A = \frac{1}{N_1} \sum_{k=1}^{n_1} \Sigma_1^{1/2}(X_k - \bar{X})(X_k - \bar{X})^T \Sigma_1^{1/2}, \quad B = \frac{1}{N_2} \sum_{k=1}^{n_2} \Sigma_2^{1/2}(Y_k - \bar{Y})(Y_k - \bar{Y})^T \Sigma_2^{1/2}.
\]
where \( N_i = n_i - 1 \) and \( y_{N_i} = \frac{p}{N_i} \) for \( i = 1, 2 \). Under \( H_0 \) and as \( n \to \infty \), we have

\[
\hat{T}_n = v(f)^{-\frac{1}{2}} \left[ T_n - p \cdot F_{y_{N_1},y_{N_2}}(f) - m(f) \right] \xrightarrow{H_0} N(0, 1) \tag{5.2}
\]

where \( f(x) = \log(y_{N_1} + y_{N_2}x) - \frac{y_{N_2}}{y_{N_1} + y_{N_2}} \log x, F_{y_{N_1},y_{N_2}}(f), m(f) \) and \( v(f) \) are

\[
F_{y_{N_1},y_{N_2}}(f) = \frac{h^2}{y_{N_1} y_{N_2}} \log\left( \frac{y_{N_1} y_{N_2}}{h^2} \right) - \frac{y_{N_2}(1 - y_{N_1})}{y_{N_1} + y_{N_2}} \log(1 - y_{N_1}) + \frac{y_{N_1}(1 - y_{N_2})}{y_{N_2}(y_{N_1} + y_{N_2})} \log(1 - y_{N_2})
\]

\[
m(f) = \frac{1}{2} \left[ \log \left( \frac{h^2}{y_{N_1} + y_{N_2}} \right) - \frac{y_{N_2} \log(1 - y_{N_1})}{y_{N_1} + y_{N_2}} - \frac{y_{N_1} \log(1 - y_{N_2})}{y_{N_1} + y_{N_2}} + \frac{y_{N_1} y_{N_2} (\beta_x y_{N_2} + \beta_y y_{N_1})}{(y_{N_1} + y_{N_2})^2} \right]
\]

\[
v(f) = -\frac{2y_{N_1}^2}{(y_{N_1} + y_{N_2})^2} \log(1 - y_{N_2}) - \frac{2y_{N_2}^2}{(y_{N_1} + y_{N_2})^2} \log(1 - y_{N_1}) + 2 \log \frac{h^2}{y_{N_1} + y_{N_2}}.
\]

Then the critical region at the test size \( \alpha = 0.05 \) is

\[
\{ T_n > 1.96v(f)^{\frac{1}{2}} + p \cdot F_{y_{N_1},y_{N_2}}(f) + m(f), \quad \text{or} \quad T_n < -1.96v(f)^{\frac{1}{2}} + p \cdot F_{y_{N_1},y_{N_2}}(f) + m(f) \}
\]

By Theorem 3.1 in this paper, under \( H_1 \) we have

\[
v^{1}(f)^{-\frac{1}{2}} \left[ T_n - p \cdot F^{1}_{y_{N_1},y_{N_2}}(f) - m^{1}(f) \right] \xrightarrow{H_1} N(0, 1),
\]

where \( m^{1}(f) \) and \( v^{1}(f) \) can be approximated by (4.8) and (4.9), and \( F^{1}_{y_{N_1},y_{N_2}}(f) \) by

\[
F^{1}_{y_{N_1},y_{N_2}}(f) = \int_{c_1}^{\infty} f(x) u_y(x) dx \approx \frac{c_2 - c_1}{10^3} \sum_{j=1}^{10^3} f(x_j) u_y(x_j), \quad x_j = c_1 + \frac{(c_2 - c_1)j}{10^3}
\]

and \( u_y(x_j) \) is computed by (4.6). Since

\[
T_n \geq 1.64v(f)^{\frac{1}{2}} + p \cdot F_{y_{N_1},y_{N_2}}(f) + m(f)
\]

\[
\Leftrightarrow v^{1}(f)^{-\frac{1}{2}} \left[ T_n - p \cdot F^{1}_{y_{N_1},y_{N_2}}(f) - m^{1}(f) \right]
\]

\[
\geq v^{1}(f)^{-\frac{1}{2}} \left[ 1.64v(f)^{\frac{1}{2}} + p \cdot F_{y_{N_1},y_{N_2}}(f) + m(f) - p \cdot F^{1}_{y_{N_1},y_{N_2}}(f) - m^{1}(f) \right],
\]

then the power function of the test is

\[
1 - \Phi \left( v^{1}(f)^{-\frac{1}{2}} \left[ 1.64v(f)^{\frac{1}{2}} + p \cdot F_{y_{N_1},y_{N_2}}(f) + m(f) - p \cdot F^{1}_{y_{N_1},y_{N_2}}(f) - m^{1}(f) \right] \right) \tag{5.3}
\]

where \( \Phi(\cdot) \) is the standardized normal distribution function.

To show the feasibility of the proposed numerical methods (4.7)-(4.8)-(4.9), we plot the power function (5.3) for the testing problem (5.1) in the Gaussian case where \( p = 300, n_1 = 600, n_2 = 6000, \Sigma_1 \) is the identity matrix and \( \Sigma_2 = \text{diag}(1, \ldots, 1, c, \ldots, c) \) with the parameter \( c = 1.00, 1.05, 1.10, 1.15, 1.20, 1.25, 1.30, 1.40, 1.50 \) and the number of \( c \) in \( \Sigma_2 \) being \( p/2 \). The numerical power function (5.3) for the testing problem (5.1) is plotted in Figure 2. Moreover, we compare the true values with numerical values of \( F_{y_{N_1},y_{N_2}}(f), m(f), Cov(f) \) in Table 2 in the case of \( \Sigma_1 = \Sigma_2 \). These values are very close each other and the proposed numerical methods perform well.
5.2 Confidence interval of $\theta$ in $T_p(\theta)$

As the second application, we consider $T_p = T_p(\theta)$, that is, $T_p$ is determined by the parameter $\theta$ which takes values in an interval $[a, b]$. We are interested in the confidence interval for the parameter $\theta$. Then using the fact

$$v^\theta(f)^{-\frac{1}{2}} \left[T_n - p \cdot F_{y_{N_1},y_{N_2}}^\theta(f) - m^\theta(f)\right] \xrightarrow{H_0} N(0, 1),$$

we will give a method to determine the confidence interval of the parameter $\theta$.

First cut $[a, b]$ as $A_3 = \{\theta_j = a + \frac{(b-a)j}{N}, \; j = 0, \ldots, N\}$ where $N$ is a large integer, e.g. $10^3$. Giving $\theta_j$, that is, $T_p = T_p(\theta_j)$ and using Algorithms 1-2, we obtain $m^\theta_j(f) = EX_f$, $v^\theta_j(f) = \text{Cov}(X_f, X_f)$ and $F_{y_{N_1},y_{N_2}}^\theta_j(f), \; j = 0, \ldots, N$. Then the confidence interval of $\theta$ is $[\theta_L, \theta_U]$ where

$$\theta_L = \min \left\{ \theta_j : v^\theta_j(f)^{-\frac{1}{2}} \left[T_n - p \cdot F_{y_{N_1},y_{N_2}}^\theta_j(f) - m^\theta_j(f)\right] \leq 1.64 \right\},$$

and

$$\theta_U = \max \left\{ \theta_j : v^\theta_j(f)^{-\frac{1}{2}} \left[T_n - p \cdot F_{y_{N_1},y_{N_2}}^\theta_j(f) - m^\theta_j(f)\right] \leq 1.64 \right\}.$$
6 Concluding remarks

In this paper, we have considered a general Fisher matrix $F$ where the (high-dimensional) population covariance matrices $\Sigma_1$ and $\Sigma_2$ can be arbitrary and not necessarily equal. First the limiting distribution of its eigenvalues has been found. Next and more importantly, we establish a CLT for its linear spectral statistics. This CLT is unavoidable in many two-sample statistical problems with high-dimensional data. Besides, this CLT extends and covers the CLT of Zheng [29] which is related to standard Fisher matrices.

An important and unsolved issue on the developed theory is about the evaluation of the limiting mean and covariance function in the CLT. These functions have a very complex structure depending on non-trivial contour integrals. In the special case where the matrices $\Sigma_2^{-1}\Sigma_1$ are diagonal, we have proposed some simplifications though the obtained results are still complex. In Section 4, we have devised some numerical procedures to approximate numerically these asymptotic parameters. The advantage of these procedures is that they depend on the observed data only. However, the accuracy of these procedure is currently unknown. A precise analysis of these procedures or finding other more accurate procedures for the approximation are certainly a valuable and challenging question in future research.

Acknowledgement. We are grateful to the Editors and the Referee for their valuable comments that have led to important improvements of the paper. Shurong Zheng was partially supported by NSFC grants NSFC-11171058 and NECT-11-0616. Zhidong Bai was partially supported by NSFC grant NSFC-11171057, PCSIRT and Fundamental Research Funds for the Central Universities. Jianfeng Yao was partially supported by GRF grant 17305814.

References


A Two propositions about $h_{m_2}(z)$ and $h_{v_1,v_2}(z_1,z_2)$ when $T_p$’s are diagonal

Here we develop an important special example where the matrices $T_p$’s are diagonal. In this case, the explicit expressions for the limiting functions $h_{m_j}(z)$ and $h_{v_j}(z_1,z_2)$, $j = 1, 2$ are given.

**Proposition A.1** In addition to the assumptions of Theorem 3.1, assume that the matrices $T_p$’s are diagonal. Then, the limits (3.5) and (3.6) exist and equal to

$$h_{m_2}(-m^{-1}(z)) = -\frac{1}{m^3(z)} \int \frac{t}{(t + m_0(z))^3} dH(t), \quad (A.1)$$

$$h_{v_2}(-m^{-1}(z_1),-m^{-1}(z_2)) = \frac{1}{m(z_1)m(z_2)} \int \frac{1}{(t + m_0(z_1))(t + m_0(z_2))} dH(t). \quad (A.2)$$

Consequently, the same conclusions as in Theorem 3.1 hold where the last term of $E_X f_i$ in (3.7) is simplified to

$$\frac{\beta_y}{4\pi i} \int_C f(z) \left(1 - y_2 \int \frac{m_0^2(z)H(t)}{(t + m_0(z))^2} dH(t) \right) d\log \left(1 - y_2 \int \frac{m_0^2(z)H(t)}{(t + m_0(z))^2} \right) \quad (A.3)$$

and the last term of $\text{Cov}(X_{f_1}, X_{f_2})$ in (3.8) is simplified to

$$-\frac{\beta_y y_2}{4\pi^2} \int_{C_1} \int_{C_2} f_i(z_1)f_j(z_2) \left[ \int \frac{t^2 H(t)}{(t + m_0(z_1))^2(t + m_0(z_2))^2} dt \right] dm_0(z_1) dm_0(z_2) \quad (A.4)$$

where each of the contours $C, C_1$ and $C_2$ encloses the support of $U_y$, $C_1$ and $C_2$ are disjoint and $m_0(z) = m_{y_2}(-m(z))$.

**Proposition A.2** In addition to the assumptions of Theorem 3.1, assume that the matrices $T_p$’s are diagonal. Then (B.64) and (B.65) exist and equal to

$$h_{m_1}(z) = -\frac{m_0^2(z)}{E_m(z)} \int \frac{t^2 H(t)}{(t + m_0(z))^2} dt \quad (A.5)$$
and
\[ h_{v1}(z_1, z_2) = \frac{1}{z_1 z_2 m(z_1) m(z_2)} \int \frac{t^2}{(t + m_0(z_1)) (t + m_0(z_2))} dH(t). \] (A.6)

Consequently, the same conclusions as in Theorem 3.1 hold where
\[ -\beta x y_1 \frac{1}{2\pi i} \oint \oint f(z_1) f(z_2) \left[ \int \frac{t^2}{(t + m_0(z))} dH(t) \right] dm_0(z_1) dm_0(z_2). \] (A.7)

\[ -\beta x y_1 \frac{1}{4\pi^2} \oint \oint f(z_1) f(z_2) \left[ \int \frac{t^2}{(t + m_0(z))} dH(t) \right] dm_0(z_1) dm_0(z_2). \] (A.8)

B Proofs

B.1 Proof of Theorem 2.1

Let
\[ s_{n_2}(z) = \int_0^\infty \frac{1}{t - z} dG_{n_2}(t), \quad s_{y_2}(z) = \int_0^\infty \frac{1}{t - z} dG_{y_2}(t), \]
be the Stieltjes transforms of the ESD \( G_{n_2}(t) \) and LSD \( G_{y_2}(t) \) of random matrix \( T_{p,1/2} S^{-1} T_{p,1/2} \), respectively. Let
\[ m_{y_2}(z) = \int_0^\infty \frac{t}{1 - tz} dG_{y_2}(t), \] (B.1)
which is the Stieltjes transform of the image measure of \( G_{y_2} \) by the reciprocal transformation \( \lambda \mapsto 1/\lambda \) on \((0, \infty)\). It is easily checked that the Stieltjes transforms are
\[ m_{y_2}(z) = -\frac{1}{z} - \frac{1}{z^2} s_{y_2}(1/z). \] (B.2)

Similarly, consider the image measure and the associated Stieltjes transform
\[ m_{n_2}(z) = -\frac{1}{z} - \frac{1}{z^2} s_{n_2}(1/z). \] (B.3)

Let
\[ m_{y_2}(z) = -\frac{1 - y_2}{z} + y_2 m_{y_2}(z), \] (B.4)
then by Theorem 2.1 of Zheng, Bai and Yao [30], we have
\[ z = - \frac{1}{m_{y_2}(z)} + y_2 \int \frac{dH(t)}{t + m_{y_2}(z)}, \]  
(B.5)
where \( H(t) \) is the LSD of \( T_p \). In fact, we have
\[ m_{y_2}(z) = -1 - \frac{y_2}{z^2}s_{y_2}(1/z) \quad \text{or} \quad -1 - \frac{z}{y_2} = 1 + y_2zs_{y_2}(z). \]

By Silverstein and Choi [21], we have
\[ z = - \frac{1}{m(z)} + y_1 \int \frac{tdG_{y_2}(t)}{1 + tm(z)} = - \frac{1}{m(z)} + y_1m_{y_2}(-m(z)). \]  
(B.6)
So by (B.4) the above equation reduces to
\[ z = - \frac{h^2}{m(z)} + \frac{y_1}{y_2}m_{y_2}(-m(z)) \]  
(B.7)
where \( h^2 = y_1 + y_2 - y_1y_2 \). Write \( m_0(z) = m_{y_2}(-m(z)) = \frac{1-y_2}{m(z)} + y_2 \int \frac{tdG_{y_2}(t)}{1 + tm(z)} \). Replacing \( z \) by \(-m(z)\), Eq. (B.5) becomes
\[ -m(z) = - \frac{1}{m_0(z)} + y_2 \int \frac{dH(t)}{t + m_0(z)}. \]  
(B.8)
Therefore, Eq. (B.7) reduces to
\[ z = \frac{h^2m_0(z)}{y_2(-1 + y_2 \int \frac{m_0(z)dH(t)}{t + m_0(z)})} + \frac{y_1}{y_2}m_0(z). \]  
(B.9)
The proof of Theorem 2.1 is then completed.

**B.2 Uniqueness of the Solution to Equation (2.9)**

Rewrite (2.9) as
\[ \frac{h^2}{y_2z - y_1m_0(z)} = - \frac{1}{m_0(z)} + y_2 \int \frac{dH(t)}{t + m_0(z)}, \]  
(B.10)
and denote \( m_0(z) = m_r(z) + im_i(z) \), where \( m_r(z) \) and \( m_i(z) \) are real. Comparing the imaginary parts of both sides of (B.10) and then dividing \( m_i \), we obtain
\[ \frac{h^2(y_2v - y_1m_i)}{m_i |y_2z - y_1m_0|^2} = \frac{1}{|m_0|^2} - y_2 \int \frac{dH(t)}{|t + m_0|^2}. \]  
(B.11)
Noting the fact that $m_i < 0$, we have
\[
\frac{1}{|m_0|^2} - y_2 \int \frac{dH(t)}{|t + m_0|^2} > \frac{h^2y_1}{|y_2z - y_1m_0|^2}.
\] (B.12)

Now suppose that $m_1 \neq m_2$ are two roots to the equation (B.10) with negative imaginary parts. Making difference of (B.10) with $m_1$ and $m_2$, we obtain
\[
\frac{h^2y_1}{(y_2z - y_1m_1)(y_2z - y_1m_2)} = \frac{1}{m_1m_2} - y_2 \int \frac{dH(t)}{(t + m_1)(t + m_2)}. \tag{B.13}
\]

Applying Cauchy-Schwarz to (B.13), then we obtain
\[
\frac{1}{|m_1m_2|} - y_2 \int \frac{dH(t)}{|t + m_1|^2} \left( \frac{1}{|m_2|^2} - y_2 \int \frac{dH(t)}{|t + m_2|^2} \right)^{1/2}
\geq \left( \frac{1}{|m_1|^2} - y_2 \int \frac{dH(t)}{|t + m_1|^2} \right)^{1/2} \left( \frac{1}{|m_2|^2} - y_2 \int \frac{dH(t)}{|t + m_2|^2} \right)^{1/2}

> \frac{h^2y_1}{|(y_2z - y_1m_1)(y_2z - y_1m_2)|}.
\]

This contradicts the equation (B.13).

The proof of the uniqueness is done. \hfill ■

**B.3 Some useful identities**

**Lemma B.1** Let $m_0(z) = m_{y_2}(-m(z))$ where $m(z)$ is the solution of (B.6), then we have the following identities

\[
1 - y_1 \int \frac{m^2(z)x^2dG_y(x)}{(1 + xm(z))^2} = \frac{h^2}{y_2} - \frac{y_1}{y_2} \cdot \left( \frac{1 - \int \frac{y_2m_0}{t + m_0}dH(t)}{1 - \int \frac{y_2m_0}{(t + m_0)^2}dH(t)} \right)^2,
\] (B.14)

\[
\log \left( \frac{h^2}{y_2} - \frac{y_1}{y_2} \cdot \left( \frac{1 - \int \frac{y_2m_0}{t + m_0}dH(t)}{1 - \int \frac{y_2m_0}{(t + m_0)^2}dH(t)} \right)^2 \right)' = -2y_1 \int \frac{m^3(z)x^2dG_y(x)}{(1 + xm(z))^2} \left[ 1 - y_1 \int \frac{m^3(z)x^2dG_y(x)}{(1 + xm(z))^2} \right],
\] (B.15)

\[
\left( \frac{h^2}{y_2} - \frac{y_1}{y_2} \cdot \left( \frac{1 - \int \frac{y_2m_0}{t + m_0}dH(t)}{1 - \int \frac{y_2m_0}{(t + m_0)^2}dH(t)} \right)^2 \right)' = -2y_1 \int \frac{m^3(z)x^2dG_y(x)}{(1 + xm(z))^2},
\] (B.16)
Differentiating both sides of (B.7) with respect to $z$

$$\left[ \log \left( 1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2} \right) \right]' = \frac{2m'(z)y_2 \int \frac{tm_0^2 dH(t)}{(t + m_0)^2}}{\left( 1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2} \right)^2},$$  \hfill (B.17)

where $m_0(z) = \frac{1}{m(z)} \left( 1 - \frac{y_2}{m(z)} s y_2 \left( -\frac{1}{m(z)} \right) \right)$, $m'(z) = \frac{-m'm_0^2}{1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2}}$. \hfill (B.18)

$$1 - y_1 \int \frac{(m(z))^2 x^2 dG_{y_2}(x)}{(1 + x m(z))^2} = \frac{(m(z))^2}{m'(z)}, \hfill (B.19)$$

$$\left( 1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2} \right)' = 2m'(z) \frac{y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2}}{1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2}}, \hfill (B.20)$$

Proof. By (B.1), we have $m'(z) = \int_0^\infty x^2 dG_{y_2}(x) \frac{1}{(1 - x z)^2}$ where $'$ denotes the derivative with respect to $z$. So by (B.4) we have

$$\int \frac{x^2 dG_{y_2}(x)}{(1 + x m(z))^2} = m'(z)(-m(z)) = -\frac{1 - y_2}{y_2} \cdot \frac{1}{(m(z))^2} + \frac{1}{y_2} \cdot m'(z)(-m(z)).$$  \hfill (B.21)

where $m'(z)(-m(z)) = \frac{d}{dz}m(z)(\xi)|_{z=-m(z)}$ instead of $\frac{d}{dz}m(z)(-m(z))$. By (B.21), we have

$$1 - y_1 \int \frac{(m(z))^2 x^2 dG_{y_2}(x)}{(1 + x m(z))^2} = \frac{h^2}{y_2} \cdot \frac{y_1(m(z))^2 m'(z)(-m(z))}{y_2}.$$  \hfill (B.22)

Differentiating both sides of (B.5) and then replacing $z$ by $-m$, we obtain

$$1 = \left( \frac{1}{m_0^2} - y_2 \int \frac{dH(t)}{(t + m_0)^2} \right) m'(z)(-m).$$  \hfill (B.23)

This equation, together with (B.8), (B.22) and (B.23) imply that

$$1 - y_1 \int \frac{m_0^2(z) x^2 dG_{y_2}(x)}{(1 + x m(z))^2} = \frac{h^2}{y_2} \cdot \frac{y_1(m(z))^2 m'(z)(-m(z))}{y_2}.$$  \hfill (B.24)

Differentiating both sides of (B.7) with respect to $z$, we obtain

$$1 = \frac{h^2}{y_2(m(z))^2} m'(z) - \frac{y_1 m'(z)(-m(z))}{y_2}.$$  \hfill (B.7)

This implies that

$$m'(z) = \frac{y_2(m(z))^2}{h^2 - y_1(m(z))^2 m'(z)(-m(z))}.$$  \hfill (B.24)

25
or equivalently

\[ y_1(m(z))^2 m_{y_2}(\neg m(z)) = h^2 - \frac{y_2(m(z))^2}{m'(z)}. \] (B.25)

So by (B.22) and (B.25), we have

\[ 1 - y_1 \int \frac{(m(z))^2 x^2 dG_{y_2}(x)}{(1 + xm(z))^2} = \frac{(m(z))^2}{m'(z)}. \] (B.26)

Differentiating both sides of (B.8), we have

\[ m'_0 = \frac{-m'm^2_0}{1 - y_2 \int \frac{m^2(t)}{(t+m_0(z))^2} dH(t)}. \] (B.27)

So we have

\[ \left(1 - y_2 \int \frac{m^2_0(z) dH(t)}{(t+m_0(z))^2}\right)' = 2m'(z) \frac{y_2 \int \frac{m^3(t)}{(t+m_0(z))^3} dH(t)}{1 - y_2 \int \frac{m^2_0(z) dH(t)}{(t+m_0(z))^2}}. \]

So by (B.24) and (B.26), we obtain (B.15). By (B.27), we have the conclusion (B.17). By (B.2) and (B.4), we have

\[ m_0(z) = \frac{1}{m(z)} \left(1 - \frac{y_2}{m(z)} s_{y_2} \left(-\frac{1}{m(z)}\right)\right). \]

The proof of the lemma is completed. \(\square\)

In the sequel, for brevity, \(s_{y_2}(z)\) will be denoted as \(s(z)\) if no confusion would be possible.

### B.4 Proof of Theorem 3.1

#### B.4.1 Deriving CLT of general Fisher matrix

Following the same techniques of truncation and normalization given in Bai and Silverstein [5] (see lines -9 to -6 from the bottom of Page 559), the assumptions can be simplified as the following:

- \(|X_{jk}| < \eta_p \sqrt{p}, \ |Y_{jk}| < \eta_p \sqrt{p}\), for some \(\eta_p \to 0\), as \(p \to \infty\),
- \(EX_{jk} = 0, \ EY_{jk} = 0\) and \(E|X_{jk}|^2 = 1, \ E|Y_{jk}|^2 = 1\);
- \(E|X_{jk}|^4 = 1 + \kappa + \beta_x + o(1)\) and \(E|Y_{jk}|^4 = 1 + \kappa + \beta_y + o(1)\);
- For the complex case, \(EX_{jk}^2 = o(n_1^{-1})\) and \(EY_{jk}^2 = o(n_2^{-1})\).
We have

\[ n_1 \left[ m_n(z) - m_n(y_{1.2}, G_{n_2}) \right] = n_1 \left[ m_n(z) - m_{y_{1.2}}(z) \right] + n_1 \left[ m_{y_{1.2}}(z) - m_n(y_{1.2}, G_{n_2}) \right], \]

where \( m_n(y_{1.2}, G_{n_2}) \) and \( m_n(y_{1.2}) \) are the unique roots with imaginary parts having the same signs as that of \( z \) to the following equations, see Eq. (2.10),

\[ z = -\frac{1}{m_n(y_{1.2}, G_{n_2})} + y_n \cdot \int \frac{td_{G_{n_2}}(t)}{1 + tm_n(y_{1.2}, G_{n_2})} \text{ and } z = -\frac{1}{m_n(y_{1.2})} + y_n \cdot \int \frac{td_{G_{y_2}}(t)}{1 + tm_n(y_{1.2})}. \]

The proof follows two steps and we unify the real and complex cases with the indicator notation \( \kappa \).

**Step 1.** Consider the conditional distribution of

\[ n_1 \left[ m_n(z) - m_{y_{1.2}}(z) \right] \quad (B.28) \]

given \( S_2 = \{ \text{all } S_2 \} \). In the proof of Theorem 2.1, we have proved that \( G_{n_2} \) converges to \( G_{y_2} \). Using Lemma 1.1 of Bai and Silverstein (2004), we conclude that the conditional distribution of

\[ n_1 \left[ m_n(z) - m_{y_{1.2}}(z) \right] = p \left[ m_n(z) - m_{y_{1.2}}(z) \right] \]

given \( S_2 \) converges to a Gaussian process \( M_1(z) \) on the contour \( C \) enclosing the support \([a, b] \) of the LSD \( U_x \) of Fisher matrix. Moreover, its mean function equals

\[
E(M_1(z)|S_2) = (\kappa - 1) \frac{y_1 \int m(z)^3 x^2 [1 + xm(z)]^{-3} dG_{y_2}(x)}{[1 - y_1 \int m(z)x^2 (1 + xm(z))^{-2} dG_{y_2}(x)]^2} \nonumber \\
+ \beta_x \frac{y_1 z^2 m(z) \cdot h_{m1}(z)}{[1 - y_1 \int x^2 m(z) |z| dG_{y_2}(x)]}, \quad (B.29)
\]

where \( h_{m1}(z) \) is the limit of

\[
\frac{1}{p} \sum_{i=1}^{p} E \left[ e_i' \left( S_{2^{-\frac{1}{2}}} T_p^{-\frac{1}{2}} \right)^* D_1^{-1} \left( S_{2^{-\frac{1}{2}}} T_p^{-\frac{1}{2}} \right) e_i \right] \\
\times e_i' \left( S_{2^{-\frac{1}{2}}} T_p^{-\frac{1}{2}} \right)^* D_1^{-1} \left( m(z) \{ T_p \}^* S_{2^{-\frac{1}{2}}} T_p^{-\frac{1}{2}} + I_p \right) \left( S_{2^{-\frac{1}{2}}} T_p^{-\frac{1}{2}} \right) e_i \bigg| S_2 \bigg]
\]

which is obtained through replacing \( S_2^{-1/2} \) by \( S_2^{-\frac{1}{2}} T_p^{-\frac{1}{2}} \) in (6.40) of Zheng (2012). The conditional covariance function of the process \( M_1(z) \) equals

\[
\text{Cov}(M_1(z_1), M_1(z_2)|S_2) = \kappa \cdot \left( \frac{m'(z_1) \cdot m'(z_2)}{(m(z_1) - m(z_2))^2} - \frac{1}{(z_1 - z_2)^2} \right).
\]

27
\[-\frac{\beta_z y_1}{4\pi^2} \cdot \oint_{C_1} \oint_{C_2} f_i(z_1) f_j(z_2) \frac{\partial^2 [z_1 z_2 m(z_1) m(z_2) h_{v1}(z_1, z_2)]}{\partial z_1 \partial z_2} dz_1 dz_2\]

where \(h_{v1}(z_1, z_2)\) is obtained through replacing \(S_2^{-1/2}\) by \(S_2^{-1/2} T_p^2\) in (6.41) of Zheng (2012).

It is remarkable that these limiting functions are independent of the conditioning \(\mathcal{J}_2\), which shows that the limiting process \(M_1(z)\) is independent of the limit of the second part below.

Step 2. Now, we consider the limiting process of

\[n_1 [m_{\{y_{n1}, G_{n2}\}}(z) - m_{y_n}(z)] = p [m_{\{y_{n1}, G_{n2}\}}(z) - m_{y_n}(z)]. \quad (B.30)\]

By (B.1), we have

\[z = -\frac{1}{m_{y_n}} + y_1 \int \frac{t}{1 + t \cdot m_{y_n}} dG_{y_n^2}(t) = -\frac{1}{m_{y_n}} + y_1 \cdot m_{y_n^2} (-m_{y_n}(z)). \quad (B.31)\]

On the other hand, \(m_{\{y_{n1}, G_{n2}\}}\) is the solution to the equation

\[z = -\frac{1}{m_{\{y_{n1}, G_{n2}\}}} + y_1 \int \frac{t \cdot dG_{n_2}(t)}{1 + t \cdot m_{\{y_{n1}, G_{n2}\}}},\]

and

\[z = -\frac{1}{m_{\{y_{n1}, G_{n2}\}}} + y_1 \int \left\{ \frac{tdG_{n_2}(t)}{1 + t m_{\{y_{n1}, G_{n2}\}}} - \frac{tdG_{n_2}(t)}{1 + tm_{y_n}} \right\} + y_1 \int \frac{tdG_{n_2}(t)}{1 + tm_{y_n}}, \quad (B.32)\]

where

\[\int \frac{t}{1 + t \cdot m_{y_n}} dG_{y_n^2}(t) = m_{n_2} (-m_{y_n}(z)).\]

Taking the difference of (B.31) and (B.32) yields

\[0 = -\frac{1}{m_{\{y_{n1}, G_{n2}\}}} + \frac{1}{m_{y_n}} y_1 \int \left\{ \frac{tdG_{n_2}(t)}{1 + tm_{\{y_{n1}, G_{n2}\}}} - \frac{tdG_{n_2}(t)}{1 + tm_{y_n}} \right\}
+ y_1 \cdot \int \frac{t \cdot dG_{n_2}(t)}{1 + t \cdot m_{y_n}} - y_1 \cdot \int \frac{t}{1 + t \cdot m_{y_n}} dG_{y_n^2}(t).\]

That is,

\[0 = \frac{m_{\{y_{n1}, G_{n2}\}} - m_{y_n}}{m_{y_n} \cdot m_{\{y_{n1}, G_{n2}\}}} - y_1 \int \frac{(m_{\{y_{n1}, G_{n2}\}} - m_{y_n}) t^2 dG_{n_2}(t)}{(1 + tm_{\{y_{n1}, G_{n2}\}})(1 + tm_{y_n})}
+ y_1 \left\{ m_{n_2} (-m_{y_n}) - m_{y_n} (-m_{y_n}) \right\}.\]
Therefore, we obtain
\[
\begin{align*}
    n_1 \cdot [m_{\{yn_1,Gn_2\}}(z) - m_{yn}(z)] &= -y_{n_1} \cdot m_{\{yn_1,Gn_2\}} m_{yn} \cdot \frac{n_1 [m_{n_2} (-m_{yn}) - m_{yn_2} (-m_{yn})]}{1 - y_{n_1} \cdot \int \frac{m_{yn} m_{\{yn_1,Gn_2\}}^2 dGn_2(t)}{(1+tm_{yn})(1+tm_{\{yn_1,Gn_2\}})}} \\
    &= -m_{\{yn_1,Gn_2\}} m_{yn} \cdot \frac{p [m_{n_2} (-m_{yn}) - m_{yn_2} (-m_{yn})]}{1 - y_{n_1} \cdot \int \frac{m_{yn} m_{\{yn_1,Gn_2\}}^2 dGn_2(t)}{(1+tm_{yn})(1+tm_{\{yn_1,Gn_2\}})}}. 
\end{align*}
\]

(B.33)

We then consider the limiting process of
\[
p [m_{n_2} (-m_{yn}(z)) - m_{yn_2} (-m_{yn}(z))] = -\frac{p}{(m_{yn}(z))^2} \left[ s_{n_2} \left( -\frac{1}{m_{yn}(z)} \right) - s_{yn_2} \left( -\frac{1}{m_{yn}(z)} \right) \right]
\]
by (B.3). Noticing that for any \( z \in \mathbb{C}\backslash\mathbb{R} \), \( m_{yn}(z) \to m(z) \), the limiting distribution of
\[
\begin{align*}
    &-\frac{p}{(m_{yn}(z))^2} \left[ s_{n_2} \left( -\frac{1}{m(z)} \right) - s_{yn_2} \left( -\frac{1}{m(z)} \right) \right] \\
\end{align*}
\]
is the same as that of
\[
-\frac{p}{(m_{yn}(z))^2} \left[ s_{n_2} \left( -\frac{1}{m(z)} \right) - s_{yn_2} \left( -\frac{1}{m(z)} \right) \right].
\]

From now on, we use the notation \( g(z) = -1/m(z) \). By Theorem 2.2 of Zheng, Bai and Yao (2013), we conclude that
\[
-pg^2(z) \left[ s_{n_2} (g(z)) - s_{yn_2} (g(z)) \right],
\]
converges weakly to a Gaussian process \( M_2(\cdot) \) on \( z \in \mathcal{C} \) with the mean function
\[
E(M_2(z)) = (\kappa - 1) \cdot \frac{y_2 \int \frac{[1+y_2g(z)s(g(z))]^3 dH(t)}{[-m(z)-1-y_2g(z)s(g(z))]^2}}{\left(1 - y_2 \int \frac{[1+y_2g(z)s(g(z))]^2 dH(t)}{[-m(z)-1-y_2g(z)s(g(z))]^2} \right)^2} \tag{B.34}
\]
\[
+ \beta y_2 \frac{[1+y_2g(z)s(g(z))]^3 h_{m_2}(g(z))}{1 - y_2 \int \frac{[1+y_2g(z)s(g(z))]^3 dH(t)}{[-m(z)-1-y_2g(z)s(g(z))]^2}}, \tag{B.35}
\]
and covariance function \( \text{Cov}(M_2(z_1), M_2(z_2)) \) equating
\[
\kappa g^2(z_1)g^2(z_2) \left( \frac{\partial^2 [g(z_1)[1+y_2g(z_1)s(g(z_1))] - g(z_1)[1+y_2g(z_2)s(g(z_2))]}{\partial (-1/m(z_1)) \partial (-1/m(z_2))} - \frac{1}{g(z_1)-g(z_2)^2} \right) \\
+ \beta y_2 g^2(z_1)g^2(z_2) \frac{\partial^2 [1+y_2g(z_1)s(g(z_1))] \partial^2 [1+y_2g(z_2)s(g(z_2))]}{\partial (-1/m(z_1)) \partial (-1/m(z_2))}, \tag{B.36}
\]
29
for $z_1, z_2 \in \mathcal{C}$, where $H(t)$ is the LSD of $T_p$. Here we have used the fact that the limits $h_{m2}(z)$ and $h_{n2}(z_1, z_2)$ in (3.5)-(3.6) exist whenever $\beta_y \neq 0$. Since

$$1 - n_1 \cdot \int \frac{m_{y_n}(z) \cdot m_G(y_{n2})}{(1 + tm_{y_n}(z))} t^2 dG_{y_2}(t) \to 1 - n_1 \int \frac{t^2 m_2(z) dG_{y_2}(t)}{(1 + tm(z))^2},$$

almost surely, this limit equals $\frac{m^2}{m}$ by (B.26). Then by (B.33) we obtain that

$$n_1 \left[ m_G(y_{n2}) - m_{y_n}(z) \right],$$

converges weakly to a Gaussian process

$$M_3(z) = -m'(z)M_2(z),$$

with the mean function $E(M_3(z)) = -m'(z)EM_2(z)$ and covariance functions

$$\text{Cov}(M_3(z_1), M_3(z_2)) = m'(z_1)m'(z_2)\text{Cov}(M_2(z_1), M_2(z_2)).$$

Since the limit process $M_1(z)$ of

$$n_1 \cdot \left[ m_n(z) - m_G(y_{n2}) \right]$$

is independent of the ESD of $S_{n2}$, we know that

$$\left\{ n_1 \cdot \left[ m_n(z) - m_G(y_{n2}) \right], \quad n_1 \cdot \left[ m_G(y_{n2}) - m_{y_n}(z) \right] \right\}$$

converge to a two-dimensional Gaussian process $(M_1(z), M_3(z))$ with independent components. Consequently, $n_1 \cdot \left[ m_n(z) - m_{y_n}(z) \right]$ converges weakly to $M_4(z)$, a Gaussian process with mean function

$$E(M_4(z)) = (\kappa - 1) \cdot \frac{y_1 \int m^3(z)x^2[1 + xm(z)]^{-3}dG_{y_2}(x)}{[1 - y_1 \int m^2(z)x^2(1 + xm(z))^{-2}dG_{y_2}(x)]^2} \quad (B.37)$$

$$+ \beta_x \cdot \frac{y_1 z^2 m^3(z) \cdot h_{m1}(z)}{1 - y_1 \int \frac{x^2 m^2(z)}{1 + xm(z)} \cdot dG_{y_2}(x)} \quad (B.38)$$

$$- (\kappa - 1) m'(z) \cdot \frac{y_2 \int \frac{t[1 + yg(g(z)) s(g(z))]^3 dH(t)}{[-tm(z) - 1 - yg(g(z)) s(g(z))]^2}}{[1 - y_2 \int \frac{[1 + yg(g(z)) s(g(z))]^2 dH(t)}{[-tm(z) - 1 - yg(g(z)) s(g(z))]^2}]^2} \quad (B.39)$$

$$- \beta_y \cdot m'(z) \cdot \frac{y_2 [1 + yg(g(z)) s(g(z))]^3 h_{m2}(g(z))}{1 - y_2 \int \frac{[1 + yg(g(z)) s(g(z))]^3 dH(t)}{[-tm(z) - 1 - yg(g(z)) s(g(z))]^2}}, \quad (B.40)$$

and covariance function

$$\text{Cov}(M_4(z_1), M_4(z_2))$$

30
\[ B.4.2 \text{ Simplifying the mean expressions (B.37) to (B.40)} \text{ and the covariance expressions (B.42)-(B.43)} \]

Recall that \( m_0(z) = m_{yw}(-m(z)) \). By (B.15), we have

\[
(B.37) = \frac{(\kappa - 1) \cdot y_1 \int \frac{m_3(x) x^2}{1 + x m(z)} dG_{y_2}(x)}{\left[ 1 - y_1 \int \frac{m_2(z) x^2}{(1 + x m(z))^2} dG_{y_2}(x) \right]^2} = -\left( \frac{\kappa - 1}{2} \right) \frac{d \log \left( \frac{h^2 y_2 - \frac{m_0(z)}{(1+y_2 g(z)) dH(t)}}{1-y_2 \int \frac{m_0(z) dH(t)}{(1+y_2 g(z))^2 dH(t)}} \right)}{d z}. 
\]

By (B.14) we have

\[
(B.38) = \beta_x \left[ y_1 z^2 m_3(z) \right]^{-1} \cdot h_{m_1}(z) \left[ 1 - y_1 \int \frac{x^2 m_2(z)}{(1 + x m(z))^2} dG_{y_2}(x) \right] = \beta_x \left[ \frac{h^2 y_2 - \frac{m_0(z)}{(1+y_2 g(z)) dH(t)}}{1-y_2 \int \frac{m_0(z) dH(t)}{(1+y_2 g(z))^2 dH(t)}} \right]^2 
\]

By (B.17) and (B.18) we have

\[
(B.39) = -\left( \frac{\kappa - 1}{2} \right) \frac{d \log \left( \frac{1 - y_2 \int \frac{m_0(z) dH(t)}{(t+m_0(z))^2 dH(t)}}{1-y_2 \int \frac{m_0(z) dH(t)}{(1+y_2 g(z))^2 dH(t)}} \right)}{d z}. 
\]

We have

\[
(B.40) = -\beta_y m'(z) \frac{y_2 [1 + y_2 g(z) s(g(z))]^3 h_{m_2}(g(z))}{1 - y_2 \int \frac{[1+y_2 g(z) s(g(z))]^2 dH(t)}{[-t m(z) - 1 - y_2 g(z) s(g(z))]^2}} = -\beta_y m'(z) \frac{y_2 m_3(z) m_0(z) h_{m_2}(g(z))}{1 - y_2 \int \frac{m_0(z) dH(t)}{(t+m_0(z))^2 dH(t)}}. 
\]
By (B.18) we have

\[(B.42) \quad \frac{\partial [g(z_1)(1+y_2g(z_1))s(g(z_1))] \partial [g(z_2)(1+y_2g(z_2))s(g(z_2))]}{\partial (-1/m(z_1)) \partial (-1/m(z_2))} \]

By (B.18) we have

\[(B.42) \quad \kappa g'(z_1)g'(z_2) \frac{\partial [g(z_1)(1+y_2g(z_1))s(g(z_1))] \partial [g(z_2)(1+y_2g(z_2))s(g(z_2))]}{\partial (-1/m(z_1)) \partial (-1/m(z_2))} \]

By (B.18) we have

\[(B.42) \quad \frac{1}{(m_0(z_1) - m_0(z_2))^2} \frac{\partial m_0(z_1)}{\partial z_1} \frac{\partial m_0(z_2)}{\partial z_2}, \]

and

\[(B.43) \quad \beta y_2g'(z_1)g'(z_2) \frac{\partial^2 [m(z_1)m_0(z_1)m(z_2)m_0(z_2)h_{v_2}(g(z_1), g(z_2))]}{\partial (-1/m(z_1)) \partial (-1/m(z_2))} \]

So we obtain

\[(B.44) \quad -\frac{1}{2\pi i} \oint f_i(z) \cdot (B.37)dz \]

\[(B.44) \quad -\frac{1}{2\pi i} \oint f_i(z) \cdot (B.38)dz = -\frac{1}{2\pi i} \oint f_i(z) \cdot (B.39)dz \]

\[(B.44) \quad -\frac{1}{2\pi i} \oint f_i(z) \cdot (B.40)dz \]

for the mean function and

\[\frac{1}{4\pi^2} \oint f_i(z_1)f_j(z_2) \cdot (B.41)dz \]

\[= -\frac{\beta y_1}{4\pi^2} \oint f_i(z_1)f_j(z_2) \frac{\partial^2 [z_1z_2m(z_1)m(z_2)h_{v_1}(z_1, z_2)]}{\partial z_1 \partial z_2} dz_1dz_2 \]
Proof. Let
\[ \int \int \int \int f_i(z_1) f_j(z_2) \cdot \{m(z_1)m_0(z_2) - h_{v1}(z_1, z_2)\} dz_1 dz_2. \]  
(B.48)

\[ -\frac{1}{4\pi^2} \int \int f_i(z_1) f_j(z_2) \cdot (B.42) dz \]
\[ = -\frac{\kappa}{4\pi^2} \int \int f_i(z_1) f_j(z_2) \frac{dm_0(z_1)}{dm_0(z_2)} \]
\[ = -\frac{\kappa}{4\pi^2} \int \int f_i(z_1) f_j(z_2) \log(m_0(z_1) - m_0(z_2)) dz_1 dz_2. \]  
(B.49)

for the covariance function where \( h^2 = y_1 + y_2 - y_1 y_2 \). The respective sums lead to the mean and covariance functions of the theorem.

B.5 Proofs of Proposition A.1 and A.2

B.5.1 Some Lemmas

Write \( \eta_j = \frac{1}{\sqrt{\pi}} v_j \) and \( \gamma_j = \frac{1}{\sqrt{\pi}} S_2^{-1/2} T_{p,1/2} x_j \), then

\[ D_1(z) = \sum_{j=2}^{n_1} \gamma_j \gamma_j^* - z I_p, \quad D_2(z) = \sum_{j=1}^{n_2} \eta_j \eta_j^* + z I_p. \]

Let \( S_{ij} = S_1 - \gamma_j^* \gamma_j, \quad S_{2j} = S_2 - \eta_j \eta_j^* \), \( D_{1j}(z) = D_1(z) - \gamma_j \gamma_j^*, \quad D_{2j}(z) = D_2(z) - \eta_j \eta_j^* \), \( D_{1jk}(z) = D_{1j}(z) - \gamma_k \gamma_k^* \), and \( D_{2jk}(z) = D_{2j}(z) - \eta_k \eta_k^* \). Moreover, Let \( \beta_{1j}(z) = \frac{1}{1+\gamma_j^* D_{1j}(z) \gamma_j}, \quad \beta_{2j}(z) = \frac{1}{1+\eta_j^* D_{2j}(z) \eta_j}, \quad \beta_{1jk}(z) = \frac{1}{1+\gamma_j^* D_{1jk}(z) \gamma_k}, \quad \beta_{2jk}(z) = \frac{1}{1+\eta_j^* D_{2jk}(z) \eta_k} \).

Lemma B.2 Under the conditions of Theorem 3.1, in \( h_{m1} \) and \( h_{v1} \), the matrix \( D_1^{-1}(z) \) can be replaced by \( -z^{-1} \{m(z)S_2^{-1/2} T_{p} S_2^{-1/2} + I_p \}^{-1} \).

Proof. Let
\[ K = -\frac{z m(z)}{n_1} S_2^{-1/2} T_{p} S_2^{-1/2} \]
and let \( e_i \) denote the \( i \)-th vector of the canonical basis of \( \mathbb{C}^p \). For any non-random \( M \) (\( p \times p \)) with bound operation norm (\( M \) will be \( I_p \) and \( m(z)T_{p,1/2} S_2^{-1/2} T_{p,1/2} + I_p \) in application), we have \( (K - z I_p)^{-1} = \sum_{j=2}^{n_1} d_j \) where \( d_j = d_j^1 + d_j^2 + d_j^3 \) and
\[ d_j^1 = (K - z I_p)^{-1} \gamma_j \gamma_j^* D_{1j}(z) (\beta_j + z m(z)) \]

\[ 33 \]
Furthermore, we have

\[ d_j^2 = -zm(z)(K - zI_p)^{-1}(\gamma_j \gamma_j^*) - \frac{1}{n_1} S_2^{-1/2} T_p S_2^{-1/2} D_{ij}^{-1} \]

\[ d_j^0 = \frac{1}{n_1 - 1} (K - zI_p)^{-1} K (D_{ij}^{-1} - D_1^{-1}). \]

Then, we have

\[
E \left| \sum_{j=2}^{n_t} e_i' T_{p,1/2}^* S_2^{-1/2} d_j^1 M S_2^{-1/2} T_{p,1/2} e_i \right|^2 \\
\leq E \left| \sum_{j=2}^{n_t} e_i' T_{p,1/2}^* S_2^{-1/2} (K - zI_p)^{-1} \gamma_j^* D_{ij}^{-1} (\beta_j + zm(z)) M S_2^{-1/2} T_{p,1/2} e_i \right|^2 = o(1) \ (B.51) \]

because, with \( S_{11} = \sum_{j=2}^{n_t} \gamma_j \gamma_j^* \),

\[
E \left| \sum_{j=2}^{n_t} e_i' T_{p,1/2}^* S_2^{-1/2} (K - zI_p)^{-1} \gamma_j \right|^2 \\
= E e_i' T_{p,1/2}^* S_2^{-1/2} (K - zI_p)^{-1} S_{11} (K - zI_p)^{-1} S_2^{-1/2} T_{p,1/2} e_i \\
\leq K v^{-2} \left| e_i'(T_{p,1/2}^* S_2^{-1/2} T_{p,1/2})^2 e_i \right| = O(1) \\
\]

and since \( \beta_j + zm(z) = o_p(1) \) and by the control convergence theorem, we have

\[
E \left| \sum_{j=2}^{n_t} \gamma_j D_{ij}^{-1} M S_2^{-1/2} T_{p,1/2} e_i (\beta_j + zm(z)) \right|^2 \\
= E \left| \sum_{j=2}^{n_t} e_i' T_{p,1/2}^* S_2^{-1/2} M^* D_{ij}^{-1} \gamma_j^* D_{ij}^{-1} M S_2^{-1/2} T_{p,1/2} e_i \times o_p(1) \right| \\
= v^{-2} (E e_i' T_{p,1/2}^* S_2^{-1/2} M^* S_{11} M S_2^{-1/2} T_{p,1/2} e_i) \times o(1) = o(1). \]

Furthermore, we have

\[
E \left| \sum_{j=2}^{n_t} e_i' T_{p,1/2}^* S_2^{-1/2} d_j^1 M S_2^{-1/2} T_{p,1/2} e_i \right|^2 \\
= \sum_{j=2}^{n_t} E \left| e_i' T_{p,1/2}^* S_2^{-1/2} (K - zI_p)^{-1} (\gamma_j \gamma_j^*) - n_1^{-1} S_2^{-1/2} T_p S_2^{-1/2} D_{ij}^{-1} M S_2^{-1/2} T_{p,1/2} e_i \right|^2 \\
\leq \frac{K}{n_1^2} \sum_{j=2}^{n_t} \left( E e_i' T_{p,1/2}^* S_2^{-1/2} (K - zI_p)^{-1} S_2^{-1/2} T_p S_2^{-1/2} (K^* - zI)^{-1} S_2^{-1/2} T_{p,1/2} e_i \cdot e_i' T_{p,1/2}^* S_2^{-1/2} M^* D_{ij}^{-1} S_2^{-1/2} T_p S_2^{-1/2} D_{ij}^{-1} M S_2^{-1/2} T_{p,1/2} e_i \right) = O(n_1^{-1}). \ (B.52) \]
and
\[
\left| \sum_{j_1 \neq j_2}^{n_1} E e_i' T_{p,1/2}^* S_2^{-1/2} d_{j_1}^2 M S_2^{-1/2} T_{p,1/2} e_i e_i' T_{p,1/2}^* S_2^{-1/2} M^* d_{j_2}^2 S_2^{-1/2} T_{p,1/2} e_i \right|
\]
\[
= \left| \sum_{j_1 \neq j_2}^{n_1} E e_i' T_{p,1/2}^* S_2^{-1/2} (K - z I_p)^{-1}(\gamma_{j_1} \gamma_{j_1}^* - \frac{1}{n_1} S_2^{-1/2} T_p S_2^{-1/2}) D_{1j_1,j_2}^{-1} \gamma_{j_2}
\gamma_{j_1}^* D_{1j_1,j_2}^{-1} (\gamma_{j_2} \gamma_{j_2}^* - \frac{1}{n_1} S_2^{-1/2} T_p S_2^{-1/2}) S_2^{-1/2} T_{p,1/2} e_i \right| \cdot |z m(z)|^2 = O(n_1^{-1}) \quad (B.53)
\]
because
\[
E \left| e_i' T_{p,1/2}^* S_2^{-1/2} (K - z I_p)^{-1}(\gamma_{j_1} \gamma_{j_1}^* - \frac{1}{n_1} S_2^{-1/2} T_p S_2^{-1/2}) D_{1j_1,j_2}^{-1} \gamma_{j_2} \gamma_{j_2}^* D_{1j_1,j_2}^{-1} (\gamma_{j_2} \gamma_{j_2}^* - \frac{1}{n_1} S_2^{-1/2} T_p S_2^{-1/2}) S_2^{-1/2} T_{p,1/2} e_i \right|^2 
\]
\[
\leq \frac{K}{n_1^2} E \left\| e_i' T_{p,1/2}^* S_2^{-1/2} (K - z I_p)^{-1} S_2^{-1/2} T_{p,1/2} \right\|^2 
\]
\[
\leq \frac{K}{n_1} E \left\| T_{p,1/2}^* S_2^{-1/2} D_{1j_1,j_2}^{-1} \gamma_{j_1} \gamma_{j_2}^* D_{1j_1,j_2}^{-1} (\gamma_{j_2} \gamma_{j_2}^* - \frac{1}{n_1} S_2^{-1/2} T_p S_2^{-1/2}) S_2^{-1/2} T_{p,1/2} \right\|^2 
\]
\[
\leq \frac{K}{n_1} \text{Etr} \left( T_{p,1/2}^* S_2^{-1/2} D_{1j_1,j_2}^{-1} \gamma_{j_1} \gamma_{j_2}^* D_{1j_1,j_2}^{-1} (\gamma_{j_2} \gamma_{j_2}^* - \frac{1}{n_1} S_2^{-1/2} T_p S_2^{-1/2}) S_2^{-1/2} T_{p,1/2} \right) = O(n_1^{-3}).
\]

The two estimates above imply that
\[
\sum_{j=2}^{n_1} e_i' T_{p,1/2}^* S_2^{-1/2} d_j^2 M S_2^{-1/2} T_{p,1/2} e_i = O_p(n_1^{-1/2}). \quad (B.54)
\]

We can similarly prove that
\[
E \left| \sum_{j=2}^{n_1} e_i' T_{p,1/2}^* S_2^{-1/2} d_j^2 M S_2^{-1/2} T_{p,1/2} e_i \right|^2 
\]
\[
\leq \frac{1}{n_1^2} E \left| \sum_{j=2}^{n_1} e_i' T_{p,1/2}^* S_2^{-1/2} (K - z I_p)^{-1} K D_{1j}^{-1} \gamma_{j} \gamma_{j}^* D_{1j}^{-1} M S_2^{-1/2} T_{p,1/2} e_i \right|^2 
\]
\[
= O(n_1^{-2}). \quad (B.55)
\]

Therefore, the matrix $D_{1j}^{-1}(z)$ can be replaced by
\[
(K - z I_p)^{-1} = -z^{-1}(m(z)^* S_2^{-1/2} T_p S_2^{-1/2} + I_p)^{-1}
\]

35
in \( h_{m_1} \) and \( h_{v_1} \).

The proof is completed.

Define \( m_1 \) and \( m_T \) as follows

\[
m_I(z) = \int \frac{1}{1 + y_2 m_1(z)} + zt \, dH(t), \quad m_T(z) = \int \frac{t}{1 + y_2 m_1(z)} + zt \, dH(t). \tag{B.56}
\]

**Lemma B.3** For any \( \Im(z) \neq 0 \), the equation (B.56) has a unique solution \( m_M(z) \) such that \( \Im(m_M(z)) \Im(z) < 0 \) where \( M = I \) or \( M = T \).

Proof. We only need to prove the lemma for \( M = I \). Without loss of generality, assume that \( \Im(z) = v > 0 \). At first, comparing the imaginary parts of both sides, we obtain

\[
m_I(z) = \int -tv + \frac{y_2 m_I(z)}{1 + y_2 m_I(z)} \, dH(t) < \int \frac{1}{1 + y_2 m_I(z)} \, dH(t),
\]

where \( m_I(z) < 0 \) is the imaginary part of \( m_I(z) \). Therefore, we have

\[
\int \frac{y_2}{|1 + y_2 m_I(z)|^2} \, dH(t) < 1. \tag{B.57}
\]

Suppose that the equation (B.56) has two solutions \( m_1 \neq m_2 \) with \( \Im(m_k) < 0 \), \( k = 1, 2 \). Then making difference of (B.56) and dividing both sides by \( m_1 - m_2 \), we obtain

\[
1 = \int \frac{y_2}{(1 + y_2 m_1)(1 + y_2 m_2)} \left( \frac{1}{1 + y_2 m_1} - \frac{1}{1 + y_2 m_2} \right) \, dH(t).
\]

Using Cauchy-Schwarz to the above equation and then using (B.57), we obtain a contradiction.

The proof of the Lemma is completed.

**Lemma B.4** Define \( m_{y_2}(z) = y_2 m_T(z) - \frac{1}{z} \). Then, we have

\[
\frac{1}{1 + y_2 m_T(z)} = -y_2 z m_T(z) + (1 - y_2) = -z m_{y_2}(z). \tag{B.58}
\]

Moreover, \( m_{y_2}(z) \) is the same as that defined in Theorem 2.1.

Proof. By the second equation of (B.56), we obtain

\[
m_T(z) = \int \frac{t}{t z + \frac{1}{1 + y_2 m_I(z)}} \, dH(t)
\]
\[
= \frac{1}{z} \left( 1 + y_2 m_1(-z) \right) - \frac{1}{z(1 + y_2 m_1(-z))} \int tz + \frac{1}{1 + y_2 m_1(-z)} dH(t) = \frac{1}{z} - \frac{1}{z} \frac{m_1(-z)}{1 + y_2 m_1(-z)}.
\]

Then we have \( \frac{1}{1 + y_2 m_1(-z)} = y_2 m_T(-z) + (1 - y_2) \). Thus we obtain
\[
\frac{1}{1 + y_2 m_1(z)} = -y_2 m_T(z) + (1 - y_2) = -m_{y_2}(z).
\]

This implies (B.58). Moreover, replacing (B.58) to (B.56), we obtain
\[
z = -\frac{1}{m_{y_2}(z)} + y_2 \int \frac{dH(t)}{t + m_{y_2}(z)}
\]
where the equation has the unique solution. Then \( m_{y_2}(z) \) is the same as that defined in Theorem 2.1.

The proof is completed.

Lemma B.5 Defining \( \bar{K}(z) = \frac{I_p}{1 + y_2 m_1(-z)} \), then we have
\[
(\bar{K}(z) + z T_p)^{-1} - (S_2 + z T_p)^{-1} = \sum_{j=1}^{n_2} \hat{d}_j(z) = \sum_{j=1}^{n_2} \check{d}_j(z), \tag{B.59}
\]
where \( \hat{d}_j(z) = \bar{d}_j(z) + \bar{d}_j(z) + \hat{d}_j(z) \), and by symmetry, \( \check{d}_j(z) = \check{d}_j(z) + \check{d}_j(z) + \check{d}_j(z) \)
\[
\hat{d}_j(z) = (\bar{K}(z) + z T_p)^{-1} \eta_j \eta_j^* D_{2j}^{-1}(z) \left( \beta_{2j}(z) - \frac{1}{1 + y_2 m_1(-z)} \right)
\]
\[
\bar{d}_j(z) = \frac{1}{1 + y_2 m_1(-z)} (\bar{K}(z) + z T_p)^{-1} (\eta_j \eta_j^* - \frac{1}{n_2} I_p) D_{2j}^{-1}(z)
\]
\[
\check{d}_j(z) = n_2^{-1} (\bar{K}(z) + z T_p)^{-1} \bar{K}(z) (D_{2j}^{-1}(z) - D_{2j}^{-1}(z))
\]

and
\[
\check{d}_j(z) = D_{2j}^{-1}(z) \eta_j \eta_j^* (\bar{K}(z) + z T_p)^{-1} \left( \beta_{2j}(z) - \frac{1}{1 + y_2 m_1(-z)} \right)
\]
\[
\bar{d}_j(z) = \frac{1}{1 + y_2 m_1(-z)} D_{2j}^{-1}(z) (\eta_j \eta_j^* - \frac{1}{n_2} I_p) (\bar{K}(z) + z T_p)^{-1}
\]
\[
\hat{d}_j(z) = n_2^{-1} (D_{2j}^{-1}(z) - D_{2j}^{-1}(z)) \bar{K}(z) (\bar{K}(z) + z T_p)^{-1}.
\]

Assuming that \( T_p \)'s are diagonal, we have
\[
\sum_{j=1}^{n_2} e_j^* T_{p,1/2} \bar{d}_{jk}(z) T_{p,1/2} e_i \to 0, \ a.s. \quad k = 1, 2, 3, \tag{B.60}
\]

37
for any random matrix $A$ with bounded non-random norm

$$
\sum_{j=1}^{n_2} e_i^T p_{1/2} A T_p p_{1/2} e_i \to 0, \ a.s. \ k = 1, 3, \quad (B.61)
$$

and for any non-random nonnegative definite matrix $B$ with bounded norm

$$
\sum_{j=1}^{n_2} e_i^T B T_p p_{1/2} (z) T_p, p_{1/2} e_i \to 0, \ a.s. \quad (B.62)
$$

Proof. We have $D_2^{-1}(z) = (zT_p + S_2)^{-1}$ and $\tilde{K}(z) = \frac{1}{1 + y_2 m_1(z)} I_p$. We write

$$
D_2(z) - (zT_p + \tilde{K}(z)) = \sum_{j=1}^{n_2} \eta_j \eta_j^* - \tilde{K}(z).
$$

Multiplying by $(zT_p + \tilde{K}(z))^{-1}$ on the left, $D_2^{-1}$ on the right and using

$$
\eta_j^* D_2^{-1}(z) = \eta_j^* D_2^{-1}(z) \beta_{2j},
$$

we obtain the first equation of $(B.59)$. And similarly we get the second equation of $(B.59)$ by multiplying $(zT_p + \tilde{K}(z))^{-1}$ on the right, and $D_2^{-1}$ on the left. At first, we derive a limit of $\frac{1}{p} \text{tr}(D_2^{-1}(z) M)$, where $M$ is a non-random $p \times p$ n.n.d. matrix of bounded operation norm and $= I_p$ or $T_p$ in accordance of application. For fixed $z = x + iy$ with $v \neq 0$ and any $n_2$, we show that $\frac{1}{p} \text{tr}(D_2^{-1}(M))$ is bounded. In fact, let $t_j$ and $x_j$ denote the eigenvalues and eigenvectors of $S_2^{-1/2} T_p S_2^{-1/2}$ and $u_j^2 = x_j^* S_2^{-1/2} M S_2^{-1/2} x_j$, we have

$$
\left| \frac{1}{p} \text{tr}(D_2^{-1} M) \right| = \left| \frac{1}{p} \sum_{j=1}^{p} \frac{|u_j|^2}{z t_j + 1} \right| \\
\leq \frac{1}{|z|} \frac{1}{p} \sum_{j=1}^{p} \frac{|u_j|^2}{|\Im(t_j + z)|} \leq \frac{|z|}{|v|} \text{tr}(S_2^{-1} M) \leq K < \infty,
$$

for some constant $K$.

Therefore, for any subsequence $\{n'_2\}$ of $\{n_2\}$, there is a convergent subsequence $\{n''_2\}$ such that $\frac{1}{p} \text{tr}(D_2^{-1} T_p)$ converges to some limit $m_{y_2}(-z)$ when $n_2$ runs to infinity along $\{n'_2\}$. Thus, to show that $\frac{1}{p} \text{tr}(D_2^{-1}(M))$ converges to $m_M(-z)$, we only need to derive an equation that $m_M(-z)$ satisfies and that the equation has a unique solution with $\Im(m_M(-z)) \Im(-z) < 0$. Without loss of generality, we assume that $\frac{1}{p} \text{tr}(D_2^{-1} M) \rightarrow m_M(-z)$ as $n_2 \rightarrow \infty$ subject to $p/n_2 \rightarrow y_2 \in (0, 1)$. Therefore, we have $\beta_{2j}(z) \rightarrow \frac{1}{1 + y_2 m_1(-z)}$. 

38
Similar to the proofs of (B.51) – (B.55), one can prove that for any $i = 1, 2, 3$,
\[
\frac{1}{p} \sum_{j=1}^{n_2} \text{tr}(\tilde{d}_{ij}^2 M) = o_{a.s.}(1).
\]
Therefore, from (B.59), one concludes that
\[
m_1(z) = \int \frac{1}{1 + y_2 m_1(-z)} \cdot \frac{m_1(y_2 m_1(-z))}{y_2} \, dt, \quad m_T(z) = \int \frac{t}{1 + y_2 m_1(-z)} \cdot \frac{m_1(y_2 m_1(-z))}{y_2} \, dt. \quad (B.63)
\]
By Lemma B.3, the equation above has a unique solution of $m_1(z)$ subject to
\[
\Im(m_1(z)) \Im(z) < 0.
\]
Hence, we have proved that \( \frac{1}{p} \text{tr}(D_2^{-1} M) \to m_1(z) \) with probability 1.

Furthermore, from (B.59), similar to the proofs of (B.51) – (B.55), one can prove that for $k = 1, 2, 3$,
\[
\sum_{j=1}^{n_2} e_j^T T_{p,1/2}^* d_{ij}^k(z) T_{p,1/2} e_i \to 0, \quad a.s.
\]
and for any random matrix $A$ with bounded non-random norm $k = 1, 3$,
\[
\sum_{j=1}^{n_2} e_j^T T_{p,1/2}^* A T_{p} d_{ij}^k(z) T_{p,1/2} e_i \to 0, \quad a.s.
\]
and for any non-random nnd matrix $B$ with bounded norm
\[
\sum_{j=1}^{n_2} e_j^T T_{p,1/2}^* B T_{p} d_{ij}^k(z) T_{p,1/2} e_i \to 0, \quad a.s.
\]
The proof is completed. 

\section*{B.5.2 Proof of Proposition A.1}

By Lemma B.4-B.5, we have
\[
-m(z) \frac{1}{p} \sum_{j=1}^{p} e_j^T (m(z) T_p + S_{2k})^{-1} e_j \cdot e_j^T (m(z) T_p + S_{2k})^{-1} T_p \left( T_p + m_{y_2}(-m(z)) I_p \right)^{-1} e_j
\]
\[
= -m^3(z) \int \frac{t}{(t + m_{y_2}(-m(z)))^3} \, dt + o_{a.s.}(1)
\]
and
\[
\frac{1}{p} \sum_{j=1}^{p} e_j^T (m(z_1) T_p + S_{2i})^{-1} e_j \cdot e_j^T (m(z_2) T_p + S_{2i})^{-1} e_j
\]
\[
\begin{align*}
\int_0^t \int_0^t \frac{1}{(t + m_{y_1}(z_1)) (t + m_{y_2}(-m(z_2)))} dH(t) + o_{a.s.}(1).
\end{align*}
\]
That is, we have proved that
\[
\begin{align*}
h_{m_2}(g(z)) &= h_{m_2}(-m^{-1}(z)) = -\frac{1}{m^3(z)} \int_0^t \frac{t}{(t + m_0(z))^3} dH(t) \\
h_{v_2}(g(z_1), g(z_2)) &= h_{v_2}(-m^{-1}(z_1), m^{-1}(z_2)) \\
&= \frac{1}{m(z_1)m(z_2)} \int_0^t \frac{1}{(t + m_0(z_1))(t + m_0(z_2))} dH(t).
\end{align*}
\]
Then we obtain
\[
\begin{align*}
(B.40) &= -\beta_y m'(z) \frac{y_2 [1 + y_2 g(z) s(g(z))]^3 h_{m_2}(g(z))}{1 - y_2 \int \frac{[1 + y_2 g(z) s(g(z))]^2 dH(t)}{[-m(z) - 1 - y_2 g(z) s(g(z))]^2}} \\
&= \frac{\beta_y m'(z) y_2 [1 + y_2 g(z) s(g(z))]^3}{1 - y_2 \int \frac{[1 + y_2 g(z) s(g(z))]^2 dH(t)}{[-m(z) - 1 - y_2 g(z) s(g(z))]^2}} \frac{1}{m^3(z)} \int_0^t \frac{t}{(t + m_0(z))^3} dH(t) \\
&= -\beta_y \cdot m'(z) \frac{y_2 \int \frac{m_0(z) dH(t)}{(t + m_0(z))^3}}{1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2}} \\
&= \frac{\beta_y}{2} \left(1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2} \right) \left[\log \left(1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2}\right)\right]'
\end{align*}
\]
and
\[
\begin{align*}
(B.43) &= \beta_y y_2 \frac{\partial^2 [(1 + y_2 g(z_1) s(g(z_1)))(1 + y_2 g(z_2) s(g(z_2))) h_{v_2}(g(z_1), g(z_2))]}{\partial z_1 \partial z_2} \\
&= \beta_y y_2 \frac{\partial^2 \int \frac{m_0(z_1)m_0(z_2)}{(t + m_0(z_1))(t + m_0(z_2))} dH(t)}{\partial z_1 \partial z_2} \\
&= \beta_y y_2 \int \frac{t^2 dH(t)}{(t + m_0(z_1))^2(t + m_0(z_2))^2} \frac{\partial^2 m_0(z_1)m_0(z_2)}{\partial z_1 \partial z_2}.
\end{align*}
\]
We have
\[
\begin{align*}
(B.47) &= -\frac{1}{2\pi i} \oint f(z) \cdot (B.40) dz \\
&= \frac{\beta_y}{4\pi i} \oint f(z) \left(1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2}\right) d\log \left(1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2}\right).
\end{align*}
\]
and
\[
\begin{align*}
(B.50) &= -\frac{1}{4\pi^2} \oint f_1(z_1) f_2(z_2) \cdot (B.43) dz_1 dz_2 \\
&= -\frac{\beta_y y_2}{4\pi^2} \oint f_1(z_1) f_2(z_2) \left[\int \frac{t^2 dH(t)}{(t + m_0(z_1))^2(t + m_0(z_2))^2}\right] dm_0(z_1) dm_0(z_2).
\end{align*}
\]
B.5.3 Proof of Proposition A.2

$h_{m1}$ and $h_{v1}$ are the limits of

$$\frac{1}{p} \sum_{i=1}^{p} E \left[ e'_i T_{p,1/2}^{*} S_2^{-\frac{3}{2}} D_1^{-\frac{1}{2}} S_2^{-\frac{3}{2}} T_{p,1/2} e_i \right]$$

and

$$\frac{1}{n_1 p} \sum_{j=1}^{n_1} \sum_{i=1}^{p} e'_i T_{p,1/2}^{*} S_2^{-\frac{3}{2}} [E_j D_{ij}^{-1}(z_1)] S_2^{-\frac{3}{2}} T_{p,1/2} e_i$$

By Lemmas B.2 and B.5, we have

$$h_{v1}(z_1, z_2) = \lim_{z_1 \rightarrow z_2} \frac{1}{z_1 z_2} \frac{1}{p} \sum_{i=1}^{p} e'_i T_{p,1/2}^{*} (m(z) T_p + S_2)^{-1} T_{p,1/2} e_i \cdot$$

$$e'_i T_{p,1/2}^{*} (m(z) T_p + S_2)^{-1} T_{p,1/2} e_i$$

$$= \lim_{z_1 \rightarrow z_2} \frac{1}{z_1 z_2} \frac{1}{p} \sum_{i=1}^{p} \left[ \frac{1}{z_2} (e'_i T_{p,1/2}^{*} (m(z) T_p + S_2)^{-1} T_{p,1/2} e_i)^2 \right.$$

$$- \frac{m(z)}{z_2^2} e'_i T_{p,1/2}^{*} (m(z) T_p + S_2)^{-1} T_{p,1/2} e_i \cdot$$

$$\left. e'_i T_{p,1/2}^{*} (m(z) T_p + S_2)^{-1} T_{p} (m(z) T_p + S_2)^{-1} T_{p,1/2} e_i \right].$$

In order to obtain the limit $h_{m1}(z)$, we first prove

$$\sum_{j_1 \neq j_2} e'_i T_{p,1/2}^{*} \frac{\partial^2}{\partial z_1^2} (m(z)) T_p d_{j1}^2 (m(z)) T_{p,1/2} e_i = o_p(1).$$

Substituting the expressions of $\frac{\partial^2}{\partial j_1}$ and $\frac{\partial^2}{\partial j_2}$ into the left hand side (LHS) of (B.67), we have

$$\text{LHS of (B.67)} = \sum_{j_1 \neq j_2} m(z)^2 m_0^2(z) e'_i T_{p,1/2}^{*} D_{j21}^{-1}(\eta_{j1} \eta_{j1} - \frac{1}{n_2} I)(m(z) T_p + \tilde{K})^{-2} T_{p} \times$$
\[
(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I_p) D_{2j2}^{-1} T_{p,1/2} e_i
\]

Then, (B.67) is equivalent to

\[
\sum_{j_1 \neq j_2} e_i^* T_{p,1/2} D_{2j1}^{-1}(\eta_j, \eta_{j1}) - \frac{1}{n_2} I_p (m(z) T_p + \tilde{K})^{-2} T_p (\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I_p) D_{2j2}^{-1} T_{p,1/2} e_i = o_p(1).
\]

(B.68)

Define \( \alpha = T_{p,1/2} e_1 \), \( B = (m(z) T_p + \tilde{K})^{-2} T_p \), \( D_{2j1j2} = D_{j1} - \eta_{j1} \eta_{j2}^* \) and \( \beta_{2j1j2} = \frac{1}{1 + \eta_{j1}^* D_{2j1j2} \eta_{j1}} \).

Noting that \( \beta_{2j1j2} = m(z) m_0(z) + O_p(n_2^{-1}) \), we have for any given \( j_1 \)

\[
\alpha^* D_{2j1j2}^{-1} \eta_{j2} \eta_{j2}^* D_{2j1j2}^{-1} (\eta_j, \eta_{j1}) - \frac{1}{n_2} I B(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) D_{2j1j2}^{-1} \eta_{j1} \eta_{j1}^* D_{2j1j2}^{-1} \alpha \beta_{2j1j2} \beta_{2j1j2} = O_p(n_2^{-5/2}),
\]

(B.69)

which follows from the easily shown facts that \( \alpha^* \eta_{j1} = O_p(n_2^{-1/2}) \), \( \eta_{j1}^* D_{2j1j2}^{-1} \alpha = O_p(n_2^{-1}) \), and \( \eta_{j1}^* B \eta_{j2} = O_p(n_2^{-1/2}) \). From (B.69), we obtain that

\[
\sum_{j_1 \neq j_2} \alpha^*(D_{2j1j2}^{-1} - D_{2j1j2}^{-1})(\eta_{j1} \eta_{j1}^* - \frac{1}{n_2} I B(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I)(D_{2j1j2}^{-1} - D_{2j1j2}^{-1}) \alpha = O_p(n^{-1/2}).
\]

(B.70)

Furthermore, we have

\[
\alpha^* D_{2j1j2}^{-1} (\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) B(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) D_{2j1j2}^{-1} \eta_{j1} \eta_{j1}^* D_{2j1j2}^{-1} \alpha \beta_{2j1j2}
\]

\[
= \alpha^* D_{2j1j2}^{-1} (\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) B(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) D_{2j1j2}^{-1} \eta_{j1} \eta_{j1}^* D_{2j1j2}^{-1} \alpha m(z) m_0(z) + O_p(n_2^{-5/2})
\]

\[
= O_p(n_2^{-2})
\]

and by similarly defined \( D_{2j1j2j3} = D_{2j1j2} - \eta_{j3} \eta_{j3}^* \),

\[
\mathbb{E} \left| \sum_{j_1 \neq j_2} \alpha^* D_{2j1j2}^{-1} (\eta_{j1} \eta_{j1}^* - \frac{1}{n_2} I) B(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) D_{2j1j2}^{-1} \eta_{j1} \eta_{j1}^* D_{2j1j2}^{-1} \alpha \right|^2
\]

\[
= \sum_{j_1 \neq j_2} \mathbb{E} \alpha^* D_{2j1j2}^{-1} (\eta_{j1} \eta_{j1}^* - \frac{1}{n_2} I) B(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) D_{2j1j2}^{-1} \eta_{j1} \eta_{j1}^* D_{2j1j2}^{-1} \alpha \times \alpha^* D_{2j3j4}^{-1} \eta_{j3} \eta_{j3}^* D_{2j3j4}^{-1} (\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) B(\eta_{j3} \eta_{j3}^* - \frac{1}{n_2} I) D_{2j3j4}^{-1} \eta_{j3} \eta_{j3}^* D_{2j3j4}^{-1} \alpha
\]

\[
= \sum_{j_1 \neq j_2 \neq j_3 \neq j_4 \text{ distinct}} \mathbb{E} \alpha^* D_{2j1j2}^{-1} (\eta_{j1} \eta_{j1}^* - \frac{1}{n_2} I) B(\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) D_{2j1j2}^{-1} \eta_{j1} \eta_{j1}^* D_{2j1j2}^{-1} \alpha \times \alpha^* D_{2j3j4}^{-1} \eta_{j3} \eta_{j3}^* D_{2j3j4}^{-1} (\eta_{j2} \eta_{j2}^* - \frac{1}{n_2} I) B(\eta_{j3} \eta_{j3}^* - \frac{1}{n_2} I) D_{2j3j4}^{-1} \eta_{j3} \eta_{j3}^* D_{2j3j4}^{-1} \alpha
\]

42
Finally, noting that
\[
E|\alpha^*D^{-1}_{2j,j_2}(\eta_{j_1}, \eta_{j_2}^*) - \frac{1}{n_2} I]B(\eta_{j_2}, \eta_{j_2}^*) - \frac{1}{n_2} I]D^{-1}_{2j,j_2} \alpha|^2 = O_p(n^{-3})
\]
we obtain
\[
E \left| \sum_{j_1 \neq j_2} \alpha^*D^{-1}_{2j,j_2}(\eta_{j_1}, \eta_{j_2}^*) - \frac{1}{n_2} I]B(\eta_{j_2}, \eta_{j_2}^*) - \frac{1}{n_2} I]D^{-1}_{2j,j_2} \alpha \right|^2 = III_1 + 4III_2 + III_3,
\]
where
\[
III_1 = \sum_{j_1 \neq j_2} E|\alpha^*D^{-1}_{2j,j_2}(\eta_{j_1}, \eta_{j_2}^*) - \frac{1}{n_2} I]B(\eta_{j_2}, \eta_{j_2}^*) - \frac{1}{n_2} I]D^{-1}_{2j,j_2} \alpha|^2 = O(n^{-1}),
\]
\[
III_2 = \sum_{j_1 \neq j_2} E\alpha^*D^{-1}_{2j,j_2}(\eta_{j_1}, \eta_{j_2}^*) - \frac{1}{n_2} I]B(\eta_{j_2}, \eta_{j_2}^*) - \frac{1}{n_2} I]D^{-1}_{2j,j_2} \alpha
\]
\[
= \sum_{j_1 \neq j_2} E\alpha^*D^{-1}_{2j,j_2}(\eta_{j_1}, \eta_{j_2}^*) - \frac{1}{n_2} I]B(\eta_{j_2}, \eta_{j_2}^*) - \frac{1}{n_2} I]D^{-1}_{2j,j_2} \alpha + O(n^{-2})
\]
\[
= O(n^{-1/2}),
\]
and
\[
III_3 = \sum_{j_1 \neq j_2} E\alpha^*D^{-1}_{2j,j_2}(\eta_{j_1}, \eta_{j_2}^*) - \frac{1}{n_2} I]B(\eta_{j_2}, \eta_{j_2}^*) - \frac{1}{n_2} I]D^{-1}_{2j,j_2} \alpha
\]
\[
= O(n^{-1/2}).
\]
\[
\sum_{j_1,j_2,j_3,j_4 \text{ distinct}} E \left( \alpha^* D_{2j_1j_2}^{-1}(\eta_{j_1},\eta_{j_1}^*) - \frac{1}{n_2} I B(\eta_{j_2},\eta_{j_2}^*) - \frac{1}{n_2} I D_{2j_1j_2}^{-1} \alpha \right) \\
- \alpha^* D_{2j_1j_2j_3}^{-1}(\eta_{j_1},\eta_{j_1}^*) - \frac{1}{n_2} I B(\eta_{j_2},\eta_{j_2}^*) - \frac{1}{n_2} I D_{2j_1j_2j_3}^{-1} \alpha \right) \times \\
\left( \alpha^* D_{2j_1j_4}^{-1*}(\eta_{j_3},\eta_{j_3}^*) - \frac{1}{n_2} I B^*(\eta_{j_4},\eta_{j_4}^*) - \frac{1}{n_2} I D_{2j_1j_4}^{-1} \alpha \right) \\
- \alpha^* D_{2j_1j_3j_4}^{-1*}(\eta_{j_3},\eta_{j_3}^*) - \frac{1}{n_2} I B^*(\eta_{j_4},\eta_{j_4}^*) - \frac{1}{n_2} I D_{2j_1j_3j_4}^{-1} \alpha \right) \\
= m^2(z)m_0^2(z) \sum_{j_1,j_2,j_3,j_4 \text{ distinct}} E \left( \alpha^* D_{2j_1j_2j_3j_4}^{-1}(\eta_{j_1},\eta_{j_1}^*) - \frac{1}{n_2} I B(\eta_{j_2},\eta_{j_2}^*) - \frac{1}{n_2} I D_{2j_1j_2j_3j_4} \alpha \right) \times \\
+ \alpha^* D_{2j_1j_2}(\eta_{j_1},\eta_{j_1}^*) - \frac{1}{n_2} I B(\eta_{j_2},\eta_{j_2}^*) - \frac{1}{n_2} I D_{2j_1j_2j_3j_4} \eta_{j_3}^* D_{2j_1j_3j_4}^{-1} \alpha \right) \times \\
\left( \alpha^* D_{2j_1j_3j_4}^{-1*}(\eta_{j_3},\eta_{j_3}^*) - \frac{1}{n_2} I B^*(\eta_{j_4},\eta_{j_4}^*) - \frac{1}{n_2} I D_{2j_1j_3j_4}^{-1} \alpha \right) \times \\
+ \alpha^* D_{2j_1j_3j_4}^{-1*}(\eta_{j_3},\eta_{j_3}^*) - \frac{1}{n_2} I B^*(\eta_{j_4},\eta_{j_4}^*) - \frac{1}{n_2} I D_{2j_1j_3j_4}^{-1} \alpha \right) + O(n^{-1/2}) \\
= O(n^{-1/2}).
\]

In the derivation above, we have used the fact that once we change a $D^{-1}$ with one more subscript, the order of the error will increase $O(n^{-1/2})$. Finally, when all $D^{-1}$ factors have been changed to $D_{2j_1j_2j_3j_4}^{-1}$, the expectation is zero.

Then, we begin to derive the limit $h_{m_1}$. By (B.66), (B.60), (B.61) and Lemma B.4, we obtain

\[
h_{m_1}(z) = \lim_{p} \frac{1}{p} \sum_{i=1}^{p} \left[ \frac{1}{z^2} (e_i^* T_{p,1/2}^* (m(z)T_p + S_2)^{-1} T_{p,1/2} e_i)^2 \right.
\]

\[
\left. - \frac{m(z)}{z^2} e_i^* T_{p,1/2}^*(m(z)T_p + S_2)^{-1} T_{p,1/2} e_i \cdot e_i^* T_{p,1/2}^*(m(z)T_p + S_2)^{-1} T_{p,1/2} e_i \right]
\]

\[
= \lim_{p} \frac{1}{p} \sum_{i=1}^{p} \left[ \frac{1}{z^2} (e_i^* T_{p,1/2}^* (\tilde{K}(m) + mT_p)^{-1} T_{p,1/2} e_i)^2 \right.
\]

\[
\left. - \frac{m(z)}{z^2} e_i^* T_{p,1/2}^* (\tilde{K}(m) + mT_p)^{-1} T_{p,1/2} e_i \cdot e_i^* T_{p,1/2}^* D^{-1}(m)T_p (\tilde{K}(m) + mT_p)^{-1} T_{p,1/2} e_i \right]
\]

\[
- \frac{m(z)}{pz^2} \lim_{p} \sum_{i=1}^{p} e_i^* T_{p,1/2}^* D_2^{-1}(m(z))T_{p,1/2} e_i \sum_{j=1}^{n_2} e_i^* T_{p,1/2}^* D_2^{-1}(m(z))T_p D_j^2(m(z))T_{p,1/2} e_i
\]
\[
\begin{align*}
&= \lim_{p \to \infty} \frac{1}{p} \sum_{i=1}^{p} \left[ \frac{1}{z^2} \left( \mathbf{e}_i \mathbf{T}_{p,1/2}^* \{ \tilde{K}(m) + m \mathbf{T}_p \}^{-1} \mathbf{T}_{p,1/2} \mathbf{e}_i \right)^2 \
&\quad - \frac{m(z)}{z^2} \mathbf{e}_i \mathbf{T}_{p,1/2}^* \{ \tilde{K}(m) + m \mathbf{T}_p \}^{-1} \mathbf{T}_{p,1/2} \mathbf{e}_i \right] \\
&\quad \cdot \mathbf{e}_i \mathbf{T}_{p,1/2}^* \{ \tilde{K}(m) + m \mathbf{T}_p \}^{-1} \mathbf{T}_p \{ \tilde{K}(m) + m \mathbf{T}_p \}^{-1} \mathbf{T}_{p,1/2} \mathbf{e}_i \\
&\quad - \frac{m(z)}{pz^2} \lim_{p \to \infty} \sum_{i=1}^{p} \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{D}_2^{-1}(m(z)) \mathbf{T}_{p,1/2} \mathbf{e}_i \sum_{j_1,j_2=1}^{n_2} \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{d}_{j_1}^2(\mathbf{m}(z)) \mathbf{T}_p \mathbf{d}_{j_2}^2(\mathbf{m}(z)) \mathbf{T}_{p,1/2} \mathbf{e}_i \\
&= \frac{m_0(z)}{z^2 m_0^2(z)} \int \frac{t^2 dH(t)}{(t+m_0(z))^3} - \frac{m(z)}{pz^2} \lim_{p \to \infty} \sum_{i=1}^{p} \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{D}_2^{-1}(m(z)) \mathbf{T}_{p,1/2} \mathbf{e}_i \\
&\quad \cdot \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{D}_2^{-1}(m(z)) \mathbf{m}_0(1) \sum_{j=1}^{n_2} \eta_j \tilde{K}(m(z)) + m(z) \mathbf{T}_p \tilde{K}(m(z)) + m(z) \mathbf{T}_p \mathbf{T}^{-1} \mathbf{m}_0(1) \eta_j \\
&= \frac{m_0(z)}{z^2 m_0^2(z)} \int \frac{t^2 dH(t)}{(t+m_0(z))^3} - \frac{m(z)}{pz^2} \lim_{p \to \infty} \sum_{i=1}^{p} \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{D}_2^{-1}(m(z)) \mathbf{T}_{p,1/2} \mathbf{e}_i \\
&\quad \cdot \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{D}_2^{-1}(m(z)) \mathbf{m}_0(1) \sum_{j=1}^{n_2} \eta_j \tilde{K}(m(z)) + m(z) \mathbf{T}_p \tilde{K}(m(z)) + m(z) \mathbf{T}_p \mathbf{T}^{-1} \mathbf{m}_0(1) \eta_j \\
&= \frac{m_0(z)}{z^2 m_0^2(z)} \int \frac{t^2 dH(t)}{(t+m_0(z))^3} - \frac{m(z)}{pz^2} \lim_{p \to \infty} \sum_{i=1}^{p} \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{D}_2^{-1}(m(z)) \mathbf{T}_{p,1/2} \mathbf{e}_i \\
&\quad \cdot \mathbf{e}_i \mathbf{T}_{p,1/2}^* \mathbf{D}_2^{-1}(m(z)) \mathbf{m}_0(1) \sum_{j=1}^{n_2} \eta_j \tilde{K}(m(z)) + m(z) \mathbf{T}_p \tilde{K}(m(z)) + m(z) \mathbf{T}_p \mathbf{T}^{-1} \mathbf{m}_0(1) \eta_j.
\end{align*}
\]
where

\[
\sum_{j=1}^{n_2} e_j T_{p,1/2}^* D_2^{-1}(m(z)) \eta_j \eta_j^* D_2^{-1}(m(z)) T_{p,1/2} e_i = \\
\sum_{j=1}^{n_2} e_j T_{p,1/2}^* D_2^{-1}(m(z)) \eta_j \eta_j^* D_2^{-1}(m(z)) T_{p,1/2} e_i + \\
+ \sum_{j=1}^{n_2} e_j T_{1/2}^* D_2^{-1}(m(z)) \eta_j \eta_j^* D_2^{-1}(m(z)) T_{1/2} e_i \cdot \frac{\eta_j^* D_2^{-1}(m(z)) \eta_j}{1 + y_2 m_1(-m(z))} + o_p(1)
\]

and similarly we have

\[
\sum_{j=1}^{n_2} e_j T_{p,1/2}^* D_2^{-1}(m(z)) \eta_j \eta_j^* D_2^{-1}(m(z)) T_{p,1/2} e_i = \\
= \frac{1}{m(z)m_0(z)} \sum_{j=1}^{n_2} e_j T_{p,1/2}^* D_2^{-1}(m(z)) \eta_j \eta_j^* D_2^{-1}(m(z)) T_{p,1/2} e_i + o_p(1)
\]

Then we have

\[
h_{m_1} = \frac{m_0(z) \int \frac{t^2 dH(t)}{(t+m_0(z))^3}}{1 + y_2 \frac{t}{m(z)} \int \frac{t^2 dH(t)}{(t+m_0(z))^3}} = \frac{m_0^2(z) \int \frac{t^2 dH(t)}{(t+m_0(z))^3}}{1 - y_2 \frac{m_0^2(z) \int dH(t)}{(t+m_0(z))^2}}
\]

where

\[
1 + y_2 \frac{t}{m(z)} \int \frac{t^2 dH(t)}{(t+m_0(z))^2}
\]
Then by (B.72) and (B.75), we obtain
That is,
Therefore, by (B.73), we have
Moreover, we have
Because
then we have
Moreover, we have
Therefore, by (B.73), we have
That is,
Then by (B.72) and (B.75), we obtain

Thus we obtain

\[
\frac{\beta x y_1}{2\pi i} \cdot \oint_C f(z) \int \frac{t^2}{(t + m_0(z))^3} dH(t) \, dz
\]

then we have

\[
\frac{\beta x y_1}{2\pi i} \cdot \oint_C f(z) \left[ \int \frac{t^2}{(t + m_0(z))^3} dH(t) \right] \, dm_0(z).
\]

The proof is completed.

B.6 Proofs of Remark 4.2

We have

\[
\frac{\beta x y_1}{2\pi i} \oint_C f(z) \left[ \int \frac{t^2}{(t + m_0(z))^3} dH(t) \right] \, dm_0(z)
\]

\[
= -\frac{\beta x y_1}{2\pi i} \oint_C f(z) \left[ \int \frac{t^2}{(t + m_0(z))^3} dH(t) \cdot \left( -1 + y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t) \right)^2 \right] \, dz.
\]

Because

\[
m'_0(z) = \frac{-m'(z)m_0^2(z)}{1 - y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t)},
\]

then we have

\[
m'(z) = \frac{\left( -1 + y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t) \right)^2}{m_0^2(z) \left( \frac{h^2}{y^2} - \frac{y_1}{y_2} \frac{m_0^2(z)m_0^2(z)}{1 - y_2 \int \frac{m_0^2(z)}{t + m_0(z)} dH(t)} \right)}.
\]

Thus we obtain

\[
\left( 1 - y_2 \int \frac{m_0^2(z)}{t + m_0(z)} dH(t) \right)^2 \left[ \log \left( 1 - y_2 \int \frac{m_0^2(t)}{t + m_0(z)} \right) \right]'
\]

\[
= \left( 1 - y_2 \int \frac{m_0^2(z)}{t + m_0(z)} dH(t) \right)^2 \frac{2m'(z)y_2 \int \frac{m_0^2(dH(t))}{(t + m_0(z))^2}}{\left( 1 - y_2 \int \frac{m_0^2(t)}{t + m_0(z)} \right)^2}
\]

\[
= 2y_2m_0(z) \int \frac{tdH(t)}{(t + m_0(z))^3} \left( -1 + y_2 \int \frac{m_0(z)}{t + m_0(z)} dH(t) \right)^2
\]

\[
1 - y_2 \int \frac{m_0^2(dH(t))}{(t + m_0(z))^2} \left( \frac{h^2}{y^2} - \frac{y_1}{y_2} \frac{m_0^2(z)m_0^2(z)}{1 - y_2 \int \frac{m_0^2(z)}{t + m_0(z)} dH(t)} \right).
\]
Then we have

\[
\frac{\beta y}{4\pi i} \oint_C f(z) \left( 1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2} \right) d\log \left( 1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2} \right) \]

\[
= \frac{\beta y}{2\pi i} \oint_C f(z) \frac{y_2 m_0(z)}{1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2}} \left( -1 + y_2 \int \frac{m_0(z) dH(t)}{t + m_0(z)} \right)^2 d\log \left( 1 - y_2 \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2} \right) \]

\[
= \frac{\beta y}{2\pi i} \oint_C f(z) \frac{y_2 m_0(z)}{1 - y_2 \int \frac{m_0^2 dH(t)}{(t + m_0)^2}} \left( 2y_2 - \frac{y_1}{y_2} \int \frac{m_0^2(z) dH(t)}{(t + m_0(z))^2} \right) \]

So the proof of Remark 4.2 is completed. □
Table 2: True values and numerical values of three parameters $F_{y_{N_1} y_{N_2}}(f)$, $m(f)$, $v(f)$ where $a-3e = a \cdot 10^{-3}$.

<table>
<thead>
<tr>
<th></th>
<th>True values of $F_{y_{N_1} y_{N_2}}(f)$</th>
<th>$m(f)$</th>
<th>$v(f)$</th>
<th>Numerical values of $F_{y_{N_1} y_{N_2}}(f)$</th>
<th>$m(f)$</th>
<th>$v(f)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1 = 0.2, y_2 = 0.02$</td>
<td>9.775-3e</td>
<td>10.111-3e</td>
<td>0.376-3e</td>
<td>9.772-3e</td>
<td>10.151-3e</td>
<td>0.382-3e</td>
</tr>
<tr>
<td>$y_1 = 0.3, y_2 = 0.03$</td>
<td>15.269-3e</td>
<td>16.151-3e</td>
<td>0.918-3e</td>
<td>15.264-3e</td>
<td>16.232-3e</td>
<td>0.938-3e</td>
</tr>
<tr>
<td>$y_1 = 0.4, y_2 = 0.04$</td>
<td>21.282-3e</td>
<td>23.101-3e</td>
<td>1.790-3e</td>
<td>21.274-3e</td>
<td>23.254-3e</td>
<td>1.835-3e</td>
</tr>
<tr>
<td>$y_1 = 0.5, y_2 = 0.05$</td>
<td>27.998-3e</td>
<td>31.657-3e</td>
<td>3.121-3e</td>
<td>27.931-3e</td>
<td>31.562-3e</td>
<td>3.199-3e</td>
</tr>
<tr>
<td>$y_1 = 0.6, y_2 = 0.06$</td>
<td>35.449-3e</td>
<td>41.186-3e</td>
<td>5.050-3e</td>
<td>35.429-3e</td>
<td>41.730-3e</td>
<td>5.239-3e</td>
</tr>
<tr>
<td>$y_1 = 0.5, y_2 = 0.5$</td>
<td>170.085-3e</td>
<td>199.063-3e</td>
<td>110.863-3e</td>
<td>169.899-3e</td>
<td>202.7326-3e</td>
<td>117.783-3e</td>
</tr>
</tbody>
</table>