Dear Prof. Holger Rootzén,

RE: BEJ0901-010: “Functional CLT for Sample Covariance matrices”

Thank you very much for your email of August 13, 2009, and referees’ reports in the e-mail. I’d like to take this opportunity to thank you, one Associate Editor and one referee for all the constructive comments. The paper has now been revised again according to these suggestions. All the major changes are outlined and attached with this letter. Would you consider it for publication in Bernoulli? Thank you very much.

Yours sincerely,

Dr. Wang Zhou
Reply to Associate Editor

We have revised the paper according to your suggestions. Thank you very much!
Reply to Referee

1. **Referee states:** “A recent paper somewhat close to the paper under review is arXiv:0809.4698 by A. Lytova and L. Pastur, in which CLTs for linear eigenvalue statistics corresponding to the test-functions having five continuous derivatives were proved. It seems worthwhile to explore the connections between the paper and the paper under review.”

**Answer:** We agree with the referee. We have added one remark after Theorem 1.1, comparing their results and ours. Thank you!

All the other comments are taken into account when we revise the paper. Thank you very much!
FUNCTIONAL CLT FOR SAMPLE COVARIANCE MATRICES

ZHIDONG BAI, XIAOYING WANG, AND WANG ZHOU

ABSTRACT. Using Bernstein polynomial approximations, we prove that the central
limit theorem for linear spectral statistics of sample covariance matrices, indexed
by a set of functions with continuous fourth-order derivatives on an open interval
including $[(1 - \sqrt{\gamma})^2, (1 + \sqrt{\gamma})^2]$, the support of the MP law. We also derive the
explicit expressions for asymptotic mean and covariance functions.

1. INTRODUCTION AND MAIN RESULT

Let $X_n = (x_{ij})_{p \times n}$, $1 \leq i \leq p, 1 \leq j \leq n$, be an observation matrix and $x_j = (x_{1j}, \cdots, x_{pj})^t$ be the $j$-th column of $X_n$. Then the sample covariance matrix is

$$S_n = \frac{1}{n-1} \sum_{j=1}^{n} (x_j - \bar{x})(x_j - \bar{x})^*,$$

where $\bar{x} = n^{-1} \sum_{j=1}^{n} x_j$ and $A^*$ is the complex conjugate transpose of $A$. The sample
covariance matrix plays an important role in multivariate analysis since it is an un-
biased estimator of the population covariance matrix and, more importantly, many
statistics in multivariate statistical analysis, e.g., principle component analysis, fac-
tor analysis and multivariate regression analysis, can be expressed as functionals of
the empirical spectral distributions (ESD) of sample covariance matrices. The ESD
of a symmetric (or Hermitian in complex case) $p \times p$ matrix $A$ is defined as

$$F^A(x) = \frac{1}{p} \times \text{cardinal number of } \{ j : \lambda_j \leq x \},$$

where $\lambda_1, \cdots, \lambda_p$ are the eigenvalues of $A$.

Assuming that the magnitude of dimension $p$ is proportional to sample size $n$, we
will study a simplified version of sample covariance matrices

$$B_n = \frac{1}{n} \sum_{j=1}^{n} x_j x_j^* = \frac{1}{n} X_n X_n^*,$$

1991 Mathematics Subject Classification. Primary 15A52, 60F15; Secondary 62H99.
Key words and phrases. Bernstein polynomial, central limit theorem, Stieltjes transform, sample
covariance matrices.

The first two authors were partially supported by NSFC 10871036 and NUS grant R-155-000-
079-112. Wang Zhou was partially supported by grant R-155-000-083-112 at the National University
of Singapore.
since $F_{B_n}$ and $F_{S_n}$ have the same liming properties according to Theorem 11.43 in [7]. We refer to [2] for a review in this field.

The first success in finding the limiting spectral distribution (LSD) of sample covariance matrices is due to to Marčenko and Pastur [13]. Subsequent work was done in [11], [12], [16], [17] and [18], where it was proved that under suitable moment conditions on $x_{ij}$, with probability 1, the ESD $F_{B_n}$ converges to the Marčenko-Pastur (MP) law $F_y$ with density function

$$F_y'(x) = \frac{1}{2\pi xy} \sqrt{(x - a)(b - x)}, \ x \in [a, b],$$

with point mass $1 - 1/y$ at the origin if $y > 1$, where $a = (1 - \sqrt{y})^2$ and $b = (1 + \sqrt{y})^2$; the constant $y$ is the dimension to sample size ratio index. The commonly used method to study the convergence of $F_{B_n}$ is the Stieltjes transform, which is defined for any distribution function $F$ by

$$s_F(z) \triangleq \int \frac{1}{x - z} dF(x), \ \Im z \neq 0.$$

It is easy to see that $s_F(\bar{z}) = s_F(z)$, here $\bar{z}$ denotes the conjugate of complex number $z$. As is known, the Stieltjes transform of MP law $s(z) \triangleq s_F$, is the unique solution to the equation

$$(1.1) \quad s = \frac{1}{1 - y - z - yzs'},$$

for each $z \in \mathbb{C}^+ \triangleq \{z \in \mathbb{C} : \Im z > 0\}$ in the set $\{s \in \mathbb{C} : -(1 - y)z^{-1} + ys \in \mathbb{C}^+\}$. Explicitly,

$$(1.2) \quad s(z) = -\frac{1}{2} \left( \frac{1}{y} - \frac{1}{yz} \sqrt{z^2 - (1 + y)z + (1 - y)^2} - \frac{1 - y}{yz} \right).$$

Here and in the sequel, $\sqrt{z}$ denotes the square root of complex number $z$ with positive imaginary part.

Using a Berry-Esseen type inequality established in terms of Stieltjes transforms, Bai [1] was able to show that the convergence rate of $\mathbb{E}F_{B_n}$ to $F_{y_n}$ is $O(n^{-5/48})$ or $O(n^{-1/4})$ according to whether $y_n$ is close to 1 or not. In [3], Bai, Miao and Tsay improved these rates in the case of the convergence in probability. Later, Bai, Miao and Yao [4] proved that $F_{B_n}$ converges to $F_{y_n}$ at a rate of $O(n^{-2/5})$ in probability and $O(n^{-2/5+\eta})$ a.s. when $y_n = p/n$ is away from 1; when $y_n = p/n$ is close to 1, both rates are $O(n^{-1/8})$. The exact convergence rate still remains unknown for the ESD of sample covariance matrices.

Instead of studying the convergence rate directly, Bai and Silverstein [6] considered the limiting distribution of the linear spectral statistics (LSS) of the general form of sample covariance matrices, indexed by a set of functions analytic on an open region covering the support of the LSD. More precisely, let $\mathcal{D}$ denote any region including $[a, b]$, and $\mathcal{A}(\mathcal{D})$ be the set of analytic functions on $\mathcal{D}$. Write
G_n(x) = p[F_n(x) − F_y(x)]. They proved the central limit theorem (CLT) for the LSS

\[ G_n(f) \doteq \int_{-\infty}^{\infty} f(x) dG_n(x), \quad f \in \mathcal{A}(D). \]

Their result is very useful to test large dimensional hypotheses. However, the analytic assumption on \( f \) seems inflexible in practical applications, because in many cases of applications, the kernel functions \( f \) can only be defined on the real line instead of the complex plane. On the other hand, it is proved in [7] that the CLT of LSS doesn’t hold for indicator functions. Therefore, it is natural to ask what is the weakest continuity condition that should be imposed on the kernel functions so that the CLT of the LSS holds. For the CLT of other type matrices, one can refer to [10].

In this paper, we consider the CLT for

\[ G_n(f) \doteq \int_{-\infty}^{\infty} f(x) dG_n(x), \quad f \in C^4(\mathcal{U}), \]

where \( \mathcal{U} \) denotes any open interval including \([a, b]\), and \( C^4(\mathcal{U}) \) means the set of functions \( f : \mathcal{U} \to \mathbb{C} \) which have continuous fourth-order derivatives.

Denote \( s(z) \) as the Stieltjes transform of \( F_z(x) = (1 - y)F_{(0, \infty)}(x) + yF_y(x) \), and set \( k(z) = s(z)/(s(z) + 1) \), where for \( x \in \mathbb{R} \), \( s(x) = \lim_{y \to 0} s(y) \).

Our main result is as follows.

**Theorem 1.1.** Assume that

a) For each \( n \), \( X_n = (x_{ij})_{p \times n} \), where \( x_{ij} \) are independent identically distributed (i.i.d.) for all \( i, j \) with \( \mathbb{E}x_{11} = 0, \mathbb{E}|x_{11}|^2 = 1, \mathbb{E}|x_{11}|^4 < \infty \) and if \( x_{ij} \) are complex variables, \( \mathbb{E}x_{11}^2 = 0; \)

b) \( y_n = p/n \to y \in (0, \infty) \) and \( y \neq 1. \)

Then, the LSS \( G_n = \{G_n(f) : f \in C^4(\mathcal{U})\} \) converges weakly in finite dimensions to a Gaussian process \( G = \{G(f) : f \in C^4(\mathcal{U})\} \) with mean function

\[ G(f) = \frac{k_1}{2\pi} \int_a^b f'(x) \text{arg} \left(1 - yk^2(x)\right) dx - \frac{k_2}{\pi} \int_a^b f(x) \Re \left(\frac{yk^2(x)}{1 - yk^2(x)}\right) dx \]

and covariance function

\[ c(f, g) \doteq \mathbb{E}[(G(f) - \mathbb{E}G(f))(G(g) - \mathbb{E}G(g))] \]

\[ = \frac{k_1 + 1}{2\pi^2} \int_a^b \int_a^b f'(x_1)g'(x_2) \ln \left|\frac{s(x_1) - \overline{s(x_2)}}{\overline{s(x_1) - s(x_2)}}\right| dx_1 dx_2 \]

\[ - \frac{k_2 y}{2\pi^2} \int_a^b \int_a^b f'(x_1)g'(x_2) \Re \{k(x_1)k(x_2) - \overline{k(x_1)k(x_2)}\} dx_1 dx_2, \]

where the parameter \( k_1 = |\mathbb{E}x_{11}^2| \) takes value 1 if \( x_{ij} \) are real, 0 otherwise and \( k_2 = \mathbb{E}|x_{11}|^4 - k_1 - 2. \)
Remark 1.2. In the definition of $G_n(f)$, $\theta = \int f(x)dF(x)$ can be regarded as a population parameter. The linear spectral statistic $\hat{\theta} = \int f(x)dF_n(x)$ is then an estimator of $\theta$. We remind the reader that the center $\theta = \int f(x)dF(x)$, instead of $\mathbb{E}\int f(x)dF_n(x)$, has its strong statistical meaning in the application of Theorem 1.1.

Using the limiting distribution of $G_n(f) = n(\hat{\theta} - \theta)$, one may perform statistical test of the ideal hypothesis. However, in the above test, one can not apply the limiting distribution of $n(\hat{\theta} - \mathbb{E}\hat{\theta})$, which was studied in [14].

The strategy of the proof is to use Bernstein polynomials to approximate functions in $C^4(\mathcal{U})$. This will be done in Section 2. Then the problem is reduced to the analytic case. The truncation and re-normalization step is in Section 3. The convergence of the empirical processes is proved in Section 4. We derive the mean function of the limiting process in Section 5.

2. Bernstein polynomial approximations

It is well-known that if $\tilde{f}(y)$ is a continuous function on the interval $[0, 1]$, the Bernstein polynomials

$$\tilde{f}_m(y) = \sum_{k=0}^{m} \binom{m}{k} y^k (1 - y)^{m-k} \tilde{f} \left( \frac{k}{m} \right)$$

converge to $\tilde{f}(y)$ uniformly on $[0, 1]$ as $m \to \infty$.

Suppose $\tilde{f}(y) \in C^4[0, 1]$. Taylor’s expansion gives

$$\tilde{f} \left( \frac{k}{m} \right) = \tilde{f}(y) + \left( \frac{k}{m} - y \right) \tilde{f}'(y) + \frac{1}{2} \left( \frac{k}{m} - y \right)^2 \tilde{f}''(y)$$

$$+ \frac{1}{3!} \left( \frac{k}{m} - y \right)^3 \tilde{f}'''(y) + \frac{1}{4!} \left( \frac{k}{m} - y \right)^4 \tilde{f}^{(4)}(\xi_y),$$

where $\xi_y$ is a number between $k/m$ and $y$. Hence

$$\tilde{f}_m(y) - \tilde{f}(y) = \frac{y(1-y)\tilde{f}'''(y)}{2m} + O\left( \frac{1}{m^2} \right).$$

For the function $f \in C^4(\mathcal{U})$, there exist $0 < a_i < a < b < b_i$ such that $[a_i, b_i] \subset \mathcal{U}$. Let $\epsilon \in (0, 1/2)$ and make a linear transformation $y = Lx + c$, where $L = (1 - 2\epsilon)/(b_i - a_i)$ and $c = ((a_i + b_i)\epsilon - a_i)/(b_i - a_i)$, then $y \in [\epsilon, 1 - \epsilon]$ if $x \in [a_i, b_i]$. Define $\tilde{f}(y) \doteq f((y - c)/L) = f(x), \ y \in [\epsilon, 1 - \epsilon]$, and

$$f_m(x) \doteq \tilde{f}_m(y) = \sum_{k=0}^{m} \binom{m}{k} y^k (1 - y)^{m-k} \tilde{f} \left( \frac{k}{m} \right).$$
From (2.1), we have
\[ f_m(x) - f(x) = \tilde{f}_m(y) - \tilde{f}(y) = \frac{y(1-y)\tilde{f}''(y)}{2m} + O\left(\frac{1}{m^2}\right). \]

Since \( \tilde{h}(y) \triangleq y(1-y)\tilde{f}''(y) \) has a second-order derivative, we can use Bernstein polynomial approximation once again to get
\[ \tilde{h}_m(y) - \tilde{h}(y) = \sum_{k=0}^{m} \binom{m}{k} y^k (1-y)^{m-k} \tilde{h}\left(\frac{k}{m}\right) - \tilde{h}(y) = O\left(\frac{1}{m}\right). \]

So, with \( h_m(x) = \tilde{h}_m(y) \),
\[ f(x) = f_m(x) - \frac{1}{2m} h_m(x) + O\left(\frac{1}{m^2}\right). \]

Therefore, \( G_n(f) \) can be split into three parts:
\[ G_n(f) = p \int_{-\infty}^{\infty} f(x)[F^{B_n} - F_{y_n}](dx) \]
\[ = p \int f_m(x)[F^{B_n} - F_{y_n}](dx) - \frac{p}{2m} \int h_m(x)[F^{B_n} - F_{y_n}](dx) \]
\[ + p \int O\left(\frac{1}{m^2}\right)[F^{B_n} - F_{y_n}](dx) \]
\[ = \Delta_1 + \Delta_2 + \Delta_3. \]

For \( \Delta_3 \), under the conditions in Theorem 1.1, by Lemma 6.1 in the Appendix,
\[ ||F^{B_n} - F_{y_n}|| = O_p(n^{-2/5}), \]
where \( a = O_p(b) \) means \( \lim_{n \to \infty} \lim_{m \to \infty} P(|a/b| \geq x) = 0. \)

Taking \( m^2 = [n^{3/5+\epsilon_0}] \), for some \( \epsilon_0 > 0 \), we have
\[ \Delta_3 = p \int O\left(\frac{1}{m^2}\right)[F^{B_n} - F_{y_n}](dx) = O\left(\frac{p}{m^2}\right) O_p(n^{-2/5}) = O_p(n^{-\epsilon_0}). \]

From now on we choose \( \epsilon_0 = 1/20 \), then \( m = [n^{13/40}] \).

Note that \( f_m(x) \) and \( h_m(x) \) are both analytic. Based on condition 4.1, 4.2 in Section 4 and martingale CLT (Theorem 35.12 in [8]), replacing \( f_m \) by \( h_m \), we obtain
\[ \Delta_2 = \frac{O(\Delta_1)}{m} = o_p(1). \]

It suffices to consider \( \Delta_1 = G_n(f_m) \). Clearly, the two polynomials \( f_m(x) \) and \( \tilde{f}_m(y) \)
only defined on the real line can be extended to \([a_l, b_r] \times [-\xi, \xi] \) and \([\epsilon, 1-\epsilon] \times [-L\xi, L\xi] \), respectively.
Since \( \tilde{f} \in C^4[0, 1] \), there is a constant \( M \), such that \( |\tilde{f}(y)| < M, \ \forall y \in [\varepsilon, 1 - \varepsilon] \).
Noting that for \((u, v) \in [\varepsilon, 1 - \varepsilon] \times [-L\xi, L\xi] \),
\[
|u + iv| + |1 - (u + iv)| = \sqrt{u^2 + v^2 + \sqrt{(1-u)^2 + v^2}}
\leq u\left[1 + \frac{v^2}{2u^2}\right] + (1 - u)\left[1 + \frac{v^2}{2(1-u)^2}\right] \leq 1 + \frac{v^2}{\varepsilon}
\]
we have, for \( y = Lx + c = u + iv \),
\[
|\tilde{f}_m(y)| = \left| \sum_{k=0}^{m} \left( \frac{m}{k}\right) x^{k}(1 - y)^{m-k} \tilde{f}\left( \frac{k}{m}\right) \right| \leq M \left( 1 + \frac{v^2}{\varepsilon} \right)^m.
\]

If we take \( |\xi| \leq L/\sqrt{m} \), then \( |\tilde{f}_m(y)| \leq M \left( 1 + L^2/(me) \right)^m \rightarrow Me^{L/\varepsilon} \), as \( m \rightarrow \infty \).
So \( \tilde{f}_m(y) \) is bounded when \( y \in [\varepsilon, 1 - \varepsilon] \times [-L/\sqrt{m}, L/\sqrt{m}] \). In other words \( f_m(x) \) is bounded when \( x \in [a_j, b_j] \times [-1/\sqrt{m}, 1/\sqrt{m}] \).
Denote \( v = 1/\sqrt{m} = n^{-1/80} \) and let \( \gamma_m \) be the contour formed by the boundary of the rectangle with vertices \((a_j \pm iv)\) and \((b_j \pm iv)\). Similarly, one can show that \( h_m(x) \), \( f'_m(x) \) and \( h'_m(x) \) are bounded on \( \gamma_m \).

3. Simplification by truncation and normalization

In this section, we will truncate the variables at a suitable level and re-normalize the truncated variables. As we will see, the truncation and re-normalization do not affect the weak limit of the spectral process.

By condition (a) in Theorem 1.1, for any \( \delta > 0 \),
\[
\delta^{-8}\mathbb{E}|x_{11}|^{8\delta}\mathbb{I}_{|x_{11}| \geq \sqrt{n}\delta} \rightarrow 0,
\]
which implies the existence of a sequence \( \delta_n \downarrow 0 \) such that
\[
\delta_n^{-8}\mathbb{E}|x_{11}|^{8\delta}\mathbb{I}_{|x_{11}| \geq \sqrt{n}\delta} \rightarrow 0,
\]
as \( n \rightarrow \infty \). Let \( \hat{x}_{ij} = x_{ij}^{\hat{B}_n}[x_{ij} \leq \sqrt{n}\delta] \) and \( \tilde{x}_{ij} = (\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij})/\sigma_n \), where \( \sigma_n^2 = \mathbb{E}|\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij}|^2 \).
Then \( \mathbb{E}\tilde{x}_{ij} = 0 \) and \( \sigma_n^2 \rightarrow 1 \) as \( n \rightarrow \infty \). We use \( \hat{X}_n \) and \( \tilde{X}_n \) to denote the analogues of \( X_n \) when the entries \( x_{ij} \) are replaced by \( \hat{x}_{ij} \) and \( \tilde{x}_{ij} \), respectively; let \( \hat{B}_n \) and \( \tilde{B}_n \) be analogues of \( B_n \); let \( \hat{G}_n \) and \( \tilde{G}_n \) be analogues of \( G_n \). Then,
\[
P(G_n \neq \hat{G}_n) \leq P(B_n \neq \tilde{B}_n) \leq np \mathbb{P}|x_{11}| \geq \sqrt{n}\delta_n \)
\[
\leq np^{-3}\delta_n^{-8}\mathbb{E}|x_{11}|^{8\delta}\mathbb{I}_{|x_{11}| \geq \sqrt{n}\delta} = o(n^{-2}).
\]

From Yin, Bai and Krishnaiah [19], we know that \( \lambda_j^{\hat{B}_n} \) and \( \lambda_j^{\tilde{B}_n} \) are a.s. bounded by \( b = (1 + \sqrt{\gamma})^2 \). Let \( \lambda_j^A \) denote the \( j \)th largest eigenvalue of matrix \( A \). Since
\[
|\sigma_n^2 - 1| \leq 2\mathbb{E}|x_{11}|^2\mathbb{I}_{|x_{11}| \geq \sqrt{n}\delta} \leq 2(\sqrt{n}\delta)^{-\gamma}\mathbb{E}|x_{11}|^{8\delta}\mathbb{I}_{|x_{11}| \geq \sqrt{n}\delta} = o(\delta_n^2n^{-3})
\]
and

\[ |\mathbb{E} \tilde{x}_{11}|^2 \leq \mathbb{E}|x_{11}|^2 1_{|x_{11}| \geq \sqrt{n}m_a} \leq o(\delta_n^2 n^{-3}) \]

we have

\[
(3.2) \quad \left| \int f(x) d\tilde{G}_n(x) - \int f(x) d\tilde{G}_n(x) \right| \leq K \sum_{j=1}^{\rho} |\lambda_j^{\tilde{n}} - \lambda_j^{B_n}| \\
\leq K(\text{tr}(\tilde{X}_n - \tilde{X}_n)(\tilde{X}_n - \tilde{X}_n)^*)^{1/2} \leq 2(1 - \sigma_n^{-1})^2 \text{tr} \hat{B}_n + 2\sigma_n^{-2} \text{tr} \tilde{X}_n \mathbb{E} \tilde{X}_n^* \\
\leq \frac{2(1 - \sigma_n^2)^2}{\sigma_n^2(1 + \sigma_n^2)} p \lambda_{\max}^{\tilde{n}} + 2\sigma_n^{-2} n \mathbb{E} |\mathbb{E} \tilde{x}_{11}|^2 = o(\delta_n^2 n^{-1}).
\]

From the above estimates in (3.1) and (3.2), we obtain

\[
\int f(x) d\tilde{G}_n(x) = \int f(x) d\tilde{G}_n(x) + o_n(1).
\]

Therefore, we only need to find the limiting distribution of \( \int f(x) d\tilde{G}_n(x) \) with the conditions that \( \mathbb{E} \tilde{x}_{11} = 0, \mathbb{E}|\tilde{x}_{11}|^2 = 1, \mathbb{E}|\tilde{x}_{11}|^8 < \infty \) and \( \mathbb{E} \tilde{x}_{11}^2 = o(n^{-2}) \) for complex variables. For brevity, in the sequel we shall suppress superscript on the variables and still use \( x_{ij} \) to denote the truncated and re-normalized variable \( \tilde{x}_{ij} \). Note that, in this paper, we use \( K \) as a generic positive constant independent of \( n \) and may be distinct on different occasions.

4. Convergence of \( \Delta - \mathbb{E} \Delta \)

Let \( B_n = n^{-1}X_n^*X_n \), then \( F^{B_n}(x) = (1 - y_n)\mathbb{E}_{(0,\infty)}(x) + y_n F^{B_n}(x) \). Correspondingly, we define \( \bar{F}_n(x) = (1 - y_n)\mathbb{E}_{(0,\infty)}(x) + y_n F_{y_n}(x) \). Let \( s_n(z) \) and \( s_n^0(z) \) be the Stieltjes transforms of \( F^{B_n} \) and \( F_{y_n} \); \( \bar{s}_n(z) \) and \( \bar{s}_n^0(z) \) be the Stieltjes transforms of \( F^{\bar{B}_n} \) and \( F_{y_n} \), respectively. Then, by Cauchy’s theorem, we have

\[
\Delta_1 = \frac{1}{2\pi i} \int \int \frac{f_m(z)}{z - x} p[F^{B_n} - F_{y_n}](dx)dz = -\frac{1}{2\pi i} \int \int f_m(z) p[s_n(z) - s_n^0(z)]dz.
\]

It is easy to verify that

\[
G_n(x) = p[F^{B_n}(x) - F_{y_n}(x)] = n[F^{\bar{B}_n}(x) - \bar{F}_n(x)].
\]

Hence, we only need to consider \( y \in (0, 1) \). We shall use the following notation:

\[
\begin{align*}
& r_j = (1/\sqrt{n})x_j, \quad D(z) = B_n - zI_p, \quad D_j(z) = D(z) - r_j r_j^*, \\
& \beta_j(z) = \frac{1}{1 + r_j D_j^{-1}(z)r_j}, \quad \bar{\beta}_j(z) = \frac{1}{1 + \frac{1}{n} \text{tr} D_j^{-1}(z)}, \\
& b_n(z) = \frac{1}{1 + \frac{1}{n} \text{tr} D_j^{-1}(z)}, \quad \epsilon_j(z) = r_j D_j^{-1}(z)r_j - \frac{1}{n} \text{tr} D_j^{-1}(z), \\
& \delta_j(z) = r_j D_j^{-1}(z)r_j - \frac{1}{n} \text{tr} D_j^{-1}(z),
\end{align*}
\]
and equalities
\begin{align*}
(4.1) \quad D^{-1}(z) - D_{j}^{-1}(z) &= -\beta_j(z)D_{j}^{-1}(z)r_{j}D_{j}^{-1}(z), \\
(4.2) \quad \beta_j(z) - \bar{\beta}_j(z) &= -\beta_j(z)\bar{\beta}_j(z)e_j(z) = -\bar{\beta}_j^2(z)e_j(z) + \beta_j(z)\bar{\beta}_j(z)e_j(z), \\
(4.3) \quad \beta_j(z) - b_n(z) &= -\beta_j(z)b_n(z)\delta_j(z) = -b_n^2(z)\delta_j(z) + \beta_j(z)b_n(z)\delta_j(z).
\end{align*}

Note that by (3.4) of Bai and Silverstein [5], the quantities \( \beta_j(z), \bar{\beta}_j(z) \) and \( b_n(z) \) are bounded in absolute value by \( |z|/v \).

Denote the \( \sigma \)-field generated by \( r_1, \ldots, r_j \) by \( F_j = \sigma(r_1, \ldots, r_j) \), conditional expectation \( \mathbb{E}_j(\cdot) = \mathbb{E}(\cdot|F_j) \) and \( \mathbb{E}_0(\cdot) = \mathbb{E}(\cdot) \). Using the equality
\begin{equation}
(4.4) \quad D^{-1}(z) - D_{j}^{-1}(z) = -\beta_j(z)D_{j}^{-1}(z)r_{j}D_{j}^{-1}(z),
\end{equation}
we have the following well-known martingale decomposition
\[ p[s_n(z) - \mathbb{E}s_n(z)] = \text{tr}(D^{-1}(z) - \mathbb{E}D^{-1}(z)) = \sum_{j=1}^{n} \text{tr}(\mathbb{E}_j D^{-1}(z) - \mathbb{E}_{j-1} D^{-1}(z)) \]
\[ = \sum_{j=1}^{n} \text{tr}(\mathbb{E}_j - \mathbb{E}_{j-1})(D^{-1}(z) - D_{j}^{-1}(z)) = - \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z)r_{j}D_{j}^{-2}(z)r_{j} \]
\[ = - \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) \frac{d \log \beta_j(z)}{dz}. \]

Integrating by part, we obtain
\[ \Delta_1 - \mathbb{E}\Delta_1 = \frac{1}{2\pi i} \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_m} f_m'(z) \log \frac{\bar{\beta}_j(z)}{\beta_j(z)} dz \]
\[ = \frac{1}{2\pi i} \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_m} f_m'(z) \log(1 + e_j(z)\bar{\beta}_j(z)) dz. \]

Let \( R_j(z) = \log(1 + e_j(z)\bar{\beta}_j(z)) - e_j(z)\bar{\beta}_j(z) \), and write
\[ \Delta_1 - \mathbb{E}\Delta_1 = \frac{1}{2\pi i} \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_m} f_m'(z)(e_j(z)\bar{\beta}_j(z) + R_j(z)) dz \]
\[ = \frac{1}{2\pi i} \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_m} f_m'(z)[e_j(z)\bar{\beta}_j(z) + R_j(z)] dz \]
\begin{equation}
(4.5) \quad + \frac{1}{2\pi i} \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_{mv}} f_m'(z)[e_j(z)\bar{\beta}_j(z) + R_j(z)] dz,
\end{equation}
where and in the sequel \( \gamma_{mh} \) denotes the union of the two horizontal parts of \( \gamma_m \) and \( \gamma_{mv} \) the union of the two vertical parts.
We first prove (4.6) → 0 in probability. Let $A_n = \{a - \epsilon_1 \leq \lambda_{\text{max}}^B \leq b + \epsilon_1\}$ for any $0 < \epsilon_1 < a - a_1$ and $A_{nj} = \{a - \epsilon_1 \leq \lambda_{\text{max}}^{B_{nj}} \leq b + \epsilon_1\}$, where $B_{nj} = B_n - r_j f_j$. Let $\lambda^B$ denotes all eigenvalues of matrix $B$. By the interlacing theorem (see [15] P.328), it follows that $A_n \subseteq A_{nj}$. Clearly, $\mathbb{I}_{A_{nj}}$ and $r_j$ are independent. By Yin, Bai and Krishnaiah [19] and Bai and Silverstein [6], when $y \in (0, 1)$, for any $l \geq 0,$

$$P(\lambda_{\text{max}}^{B_n} \geq b + \epsilon_1) = o(n^{-l})$$
and
$$P(\lambda_{\text{min}}^{B_n} \leq a - \epsilon_1) = o(n^{-l}).$$

We have $P(A_{nj}^c) = o(n^{-l})$, for any $l \geq 0$.

By continuity of $s(z)$, for large $n$, there exist positive constants $M_t$ and $M_u$ such that for all $z \in \gamma_{mv}, M_t \leq |y_n s(z)| \leq M_u$. Let $C_{nj} = \{|\tilde{\beta}_j(z)|^{-1} \mathbb{I}_{A_{nj}} > \epsilon_2\}$, here $0 < \epsilon_2 < M_t/2$, and $C_n = \cap^n_{j=1} C_{nj}$,

$$P(C_{n}^c) = P(\cup^n_{j=1} C_{nj}^c) \leq \sum^n_{j=1} P(C_{nj}^c) = \sum^n_{j=1} P(\{|\tilde{\beta}_j(z)|^{-1} \mathbb{I}_{A_{nj}} \leq \epsilon_2\})$$

$$\leq \sum^n_{j=1} P \left( \left\{ \frac{1}{n} \text{tr} D_j^{-1}(z) - y_n s(z) \right\} \mathbb{I}_{A_{nj}} \geq \epsilon_2 \right) + \sum^n_{j=1} P(A_{nj}^c)$$

$$\leq \frac{1}{\epsilon_2^4} \sum^n_{j=1} \mathbb{E} \left| \frac{1}{n} \text{tr} D_j^{-1}(z) - y_n s(z) \right|^4 \mathbb{I}_{A_{nj}} + n P(A_{nj}^c)$$

$$\leq \frac{1}{\epsilon_2^4} \sum^n_{j=1} O(n^{-\frac{4}{2}}) + n P(A_{nj}^c) \leq O(n^{-\frac{4}{2}}),$$

where we have used Lemma 6.1. Define $Q_{nj} = A_{nj} \cap C_{nj}$ and $Q_n = \cap^n_{j=1} Q_{nj}$, it is easy to find that $Q_{nj}$ is independent of $r_j$ and $P(Q_{nj}^c) \leq P(A_{nj}^c) + P(C_{nj}^c) \to 0$ as $n \to \infty$. Now (4.6) becomes

$$\sum^n_{j=1} (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_{mv}} f_m'(z) [\epsilon_j(z) \tilde{\beta}_j(z) + R_j(z)] \mathbb{I}_{Q_{nj}} dz + o_p(1).$$

From Burkholder inequality, Lemma 6.3 and the inequalities $|n^{-1} \text{tr} D_j(z) D_j(\tilde{z})| \mathbb{I}_{A_{nj}} \leq 1/(a - \epsilon_1 - a_1)^2$ and $|\tilde{\beta}_j(z)| \mathbb{I}_{Q_{nj}} \leq 1/\epsilon_2$, we have

$$\mathbb{E} \left( \sum^n_{j=1} (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_{mv}} f_m'(z) [\epsilon_j(z) \tilde{\beta}_j(z)] \mathbb{I}_{Q_{nj}} dz \right)^2$$

$$\leq K \|y_{mv}\|^2 \sum^n_{j=1} \sup_{z \in \gamma_{mv}} \mathbb{E} |\epsilon_j(z)\tilde{\beta}_j(z)|^2 \mathbb{I}_{Q_{nj}}$$

$$\leq Kn^{-\frac{12}{13}} \sum^n_{j=1} \sup_{z \in \gamma_{mv}} \mathbb{E} |\epsilon_j(z)|^2 \mathbb{I}_{A_{nj}} \leq Kn^{-\frac{13}{13}}.$$
By Lemma 6.3, for $z \in \gamma_{mv}$, we have
\[
\sum_{j=1}^{n} P(\varepsilon_j(z) \beta_j(z) \in Q_{nj} \geq 1/2) \leq K \sum_{j=1}^{n} \mathbb{E}|\varepsilon_j(z) \beta_j(z)|^4 \mathbb{I}_{Q_{nj}} \leq K/n.
\]
From the inequality $|\log(1+x) - x| \leq Kx^2$ for $|x| < 1/2$, we get
\[
(4.7) \quad \mathbb{E} \left| \sum_{j=1}^{n} (E_j - E_{j-1}) \int_{\gamma_{mv}} f_{m}^{\prime}(z) R_j(z) \mathbb{E}|\varepsilon_j(z) \beta_j(z)|^2 \mathbb{I}_{Q_{nj} \cap \{|\varepsilon_j(z) \beta_j(z)| < 1/2\}} dz \right|^2
\leq K \|\gamma_{mv}\|^2 \sum_{j=1}^{n} \sup_{z \in \gamma_{mv}} \mathbb{E}|R_j(z)|^2 \mathbb{I}_{Q_{nj} \cap \{|\varepsilon_j(z) \beta_j(z)| < 1/2\}}
\leq Kn \sup_{j=1}^{n} \mathbb{E}|\varepsilon_j(z) \beta_j(z)|^4 \leq K n^{5/4}.
\]
Therefore, from the above estimates, we can conclude that (4.6) converges to 0 in probability. Similarly, for $z \in \gamma_{mh}$, we also have these two estimates
\[
\sum_{j=1}^{n} P(\varepsilon_j(z) \beta_j(z) \in Q_{nj} \geq 1/2) \leq K \sum_{j=1}^{n} \mathbb{E}|\varepsilon_j(z) \beta_j(z)|^4,
\]
and
\[
(4.8) \quad \mathbb{E} \left| \sum_{j=1}^{n} (E_j - E_{j-1}) \int_{\gamma_{mh}} f_{m}^{\prime}(z) R_j(z) \mathbb{E}|\varepsilon_j(z) \beta_j(z)|^2 \mathbb{I}_{Q_{nj} \cap \{|\varepsilon_j(z) \beta_j(z)| < 1/2\}} dz \right|^2
\leq K \|\gamma_{mh}\|^2 \sum_{j=1}^{n} \sup_{z \in \gamma_{mh}} \mathbb{E}|R_j(z)|^2 \mathbb{I}_{Q_{nj} \cap \{|\varepsilon_j(z) \beta_j(z)| < 1/2\}}
\leq K \sum_{j=1}^{n} \sup_{z \in \gamma_{mh}} \mathbb{E}|\varepsilon_j(z) \beta_j(z)|^4.
\]
Thus, we get
\[
\begin{align*}
(4.5) &= -\frac{1}{2\pi i} \sum_{j=1}^{n} E_j \int_{\gamma_{mh}} f_{m}^{\prime}(z) |\varepsilon_j(z) \beta_j(z)| dz + o_p(1) \\
&\pm -\frac{1}{2\pi i} \sum_{j=1}^{n} Y_{nj} + o_p(1),
\end{align*}
\]
where $o_p(1)$ follows from (4.7), (4.8) and condition 4.1 below. Therefore, our goal reduces to the convergence of $\sum_{j=1}^{n} Y_{nj}$.

Since $Y_{nj} \in F_j$ and $E_{j-1} Y_{nj} = 0, \{Y_{nj}, j = 1, ..., n\}$ is a martingale difference sequence, thus $\sum_{j=1}^{n} Y_{nj}$ is a sum of a martingale difference sequence. In order to
apply martingale CLT (Theorem 35.12 in [8]) to it, we need to check the following two conditions:

Condition 4.1. Lyapunov condition

\[ \sum_{j=1}^{n} \mathbb{E}|Y_{nj}|^4 \to 0. \]

Condition 4.2. Conditional covariance

\[-\frac{1}{4\pi^2} \sum_{j=1}^{n} \mathbb{E}_{j-1}[Y_{nj}(f_m) \cdot Y_{nj}(g_m)] \]

converges to a constant \(c(f, g)\) in probability, where \(f, g \in \mathcal{C}^4(\mathcal{U})\) and \(f_m, g_m\) are their corresponding Bernstein polynomial approximations, respectively.

**Proof of condition 4.1.** By Lemmas 6.5 and 6.6, for any \(z \in \gamma_{mh}\),

\[ \mathbb{E} |\varepsilon_j(z)|^6 \leq \frac{K}{n^6} \left[ (\mathbb{E}|x_{11}|^4 \text{tr}(D_j^{-1}(z)D_j^{-1}(z))^3 + \mathbb{E}|x_{11}|^2 \text{tr}(D_j^{-1}(z)D_j^{-1}(z))^3) \right] \]

\[ \leq \frac{K}{n^6} \left[ n^3 + \delta_4 n^3 \right] \leq \frac{K}{n^3 \sqrt{6}}. \]

Hence we get

\[ \sum_{j=1}^{n} \mathbb{E}|Y_{nj}|^4 \leq K \sum_{j=1}^{n} \int_{\gamma_{mh}} \mathbb{E}|\varepsilon_j(z)|^4 |\tilde{\beta}_j(z)|^4 |dz| \leq K \sum_{j=1}^{n} \int_{\gamma_{mh}} (\mathbb{E}|\tilde{\beta}_j(z)|^6)^{1/2} (\mathbb{E}|\varepsilon_j(z)|^6)^{1/2} |dz| \]

\[ \leq \frac{K}{n \sqrt{6}} \to 0. \]

**Proof of condition 4.2.** Note that in Cauchy’s theorem the integral formula is independent of the choice of the contour. Hence, we have

\[ -\frac{1}{4\pi^2} \sum_{j=1}^{n} \mathbb{E}_{j-1}[Y_{nj}(f_m) \cdot Y_{nj}(g_m)] \]

\[ = -\frac{1}{4\pi^2} \sum_{j=1}^{n} \mathbb{E}_{j-1} \left[ \int_{\gamma_{mh}} f_m'(z)\mathbb{E}(\varepsilon_j(z)|\tilde{\beta}_j(z)) dz \cdot \int_{\gamma_{mh}} g_m'(z)\mathbb{E}(\varepsilon_j(z)|\tilde{\beta}_j(z)) dz \right] \]

\[ = -\frac{1}{4\pi^2} \int_{\gamma_{mh} \times \gamma_{mh}'} \int_{\gamma_{mh} \times \gamma_{mh}'} f_m'(z_1)g_m'(z_2) \sum_{j=1}^{n} \mathbb{E}_{j-1} \left[ \mathbb{E}(\varepsilon_j(z_1)|\tilde{\beta}_j(z_1))\mathbb{E}(\varepsilon_j(z_2)|\tilde{\beta}_j(z_2)) \right] dz_1 dz_2 \]

\[ = -\frac{1}{4\pi^2} \int_{\gamma_{mh} \times \gamma_{mh}'} \int_{\gamma_{mh} \times \gamma_{mh}'} f_m'(z_1)g_m'(z_2) \Gamma_n(z_1, z_2) dz_1 dz_2, \]

where \(\Gamma_n(z_1, z_2) = \sum_{j=1}^{n} \mathbb{E}_{j-1} \left[ \mathbb{E}(\varepsilon_j(z_1)|\tilde{\beta}_j(z_1))\mathbb{E}(\varepsilon_j(z_2)|\tilde{\beta}_j(z_2)) \right]\) and \(\gamma_{m}'\) is the contour formed by the rectangle with vertex \(a'_i \pm i/2 \sqrt{m}\) and \(b'_i \pm i/2 \sqrt{m}\), here \(0 < a_i < a'_i < \)
where
\[\Gamma(z_1, z_2) = \kappa_2 k(z_1) k(z_2) - (\kappa_1 + 1) \ln \frac{s(z_1) s(z_2)(z_1 - z_2)}{s(z_1) - s(z_2)}.
\]

From Lemma 6.6, for all \(z \in \gamma_{mh} \cup \gamma'_{mh}\) and any \(l \geq 2\),
\[E|\tilde{b}_j(z) - b_n(z)|^l = E|\tilde{b}_j(z)b_n(z)n^{-1}(tr D_j(z) - tr D_j(z))|^l \leq M(E|n^{-1}(tr D_j(z) - tr D_j(z))|^2)^{1/2} \leq K(\sqrt{n}v)\]
(4.9)
This leads to
\[\left| E\Gamma_n(z_1, z_2) - b_n(z_1)b_n(z_2) \sum_{j=1}^n \tilde{b}_j(z_1) \tilde{b}_j(z_2) \right| \leq \frac{K}{\sqrt{nv}^3} = O(n^{-\frac{3}{2}}).
\]
Thus we need to consider
\[b_n(z_1)b_n(z_2) \sum_{j=1}^n \tilde{b}_j(z_1) \tilde{b}_j(z_2).
\]

Let \([A]_{ii}\) denote the \((i, i)\) entry of matrix \(A\). For any two \(p \times p\) non-random matrices \(A\) and \(B\), we have
\[E(x_i^* Ax_i - tr A)(x_i^* Bx_i - tr B) = (E|x_{i1}^4 - |E x_{i1}^2|^2 - 2) \sum_{j=1}^p a_j b_{ii} + |E x_{i1}^2|^2 \sum_{j=1}^p a_{ij} b_{ij} + \sum_{j=1}^p a_{ij} b_{ji}
\]
\[= \kappa_2 \sum_{j=1}^p a_{ij} b_{ii} + \kappa_1 tr A B^T + tr A B,
\]
from which (4.10) becomes
\[\frac{(\kappa_1 + 1)b_n(z_1)b_n(z_2) 1}{n^2} \sum_{j=1}^n \tilde{b}_j(z_1) \tilde{b}_j(z_2)
\]
\[+ \kappa_2 b_n(z_1)b_n(z_2) \frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^p \tilde{b}_j(z_1) \tilde{b}_j(z_2) \tilde{b}_j(z_2) \tilde{b}_j(z_1) i \Gamma_n(z_1, z_2) + \Gamma_n(z_1, z_2).
\]
For \(\Gamma_n(z_1, z_2)\), by Lemmas 6.6, 6.7 and \(-zs(z)(s(z) + 1) = 1\), one can get
\[\Gamma_n(z_1, z_2) = \kappa_2 y_n k(z_1) k(z_2) + o_p(1),
\]
here \(o_p(1)\) denotes uniform convergence in probability on \(\gamma_{mh} \times \gamma'_{mh}\).
It is easy to check that \( k(\bar{z}) = k(z) \) since \( s(\bar{z}) = s(z) \). Then as \( n \to \infty \), \( a_l \to a \) and \( b_r \to b \), we get

\[
-\frac{1}{4\pi^2} \int \int_{\gamma_{nb} \times \gamma_{mb}} f_m'(z_1)g_m'(z_2)\Gamma_{n2}(z_1, z_2)dz_1dz_2
\]

\[
= -\frac{k_2 y_n}{4\pi^2} \int \int_{\gamma_{nb} \times \gamma_{mb}} f_m'(z_1)g_m'(z_2)k(z_1)k(z_2)dz_1dz_2 + o_p(1)
\]

\[
\to -\frac{k_2 y_n}{2\pi^2} \int_a^b \int_a^b f'(x_1)g'(x_2) \Re[k(x_1)k(x_2) - k(x_1)k(x_2)]dx_1dx_2,
\]

which is (1.5) in Theorem 1.1.

For \( \Gamma_{n1}(z_1, z_2) \), we will find the limit of

\[
(4.12) \quad b_n(z_1)b_n(z_2) \frac{1}{n^2} \sum_{i,j=1}^n trE_jD_{ij}^{-1}(z_1)E_jD_{ij}^{-1}(z_2).
\]

Let \( D_{ij}(z) = D(z) - r_j r_j^* - r_i r_i^* \), \( \beta_{ij}(z) = (1 + r_j^* D_{ij}^{-1}(z) r_i)^{-1} \), \( b_{12}(z) = (1 + \frac{1}{n} trD_{12}^{-1}(z))^{-1} \) and \( t(z) = (z - \frac{a_l - b_{12}(z)}{n})^{-1} \). Write

\[
D_{ij}(z) + zI_p - \frac{n - 1}{n} b_{12}(z)I_p = \sum_{i \neq j} r_i r_i^* - \frac{n - 1}{n} b_{12}(z)I_p.
\]

Multiplying by \( t(z)I_p \) on the left, \( D_{ij}^{-1}(z) \) on the right and combining with the identity

\[
(4.13) \quad r_i^* D_{ij}^{-1}(z) = \beta_{ij}(z) r_i^* D_{ij}^{-1}(z),
\]

we obtain

\[
D_{ij}^{-1}(z) = -t(z)I_p + \sum_{i \neq j} t(z)\beta_{ij}(z) r_i r_i^* D_{ij}^{-1}(z) - \frac{n - 1}{n} b_{12}(z) t(z) D_{ij}^{-1}(z)
\]

\[
\to -t(z)I_p + b_{12}(z) A(z) + B(z) + C(z),
\]

where

\[
A(z) = \sum_{i \neq j} t(z)(r_i r_i^* - n^{-1} I_p) D_{ij}^{-1}(z), \quad B(z) = \sum_{i \neq j} t(z)(\beta_{ij}(z) - b_{12}(z)) r_i r_i^* D_{ij}^{-1}(z)
\]

and

\[
C(z) = \frac{1}{n} t(z) b_{12}(z) \sum_{i \neq j} (D_{ij}^{-1}(z) - D_{ij}^{-1}(z)).
\]
It is easy to verify that for all $z \in \gamma_{nh} \cup \gamma'_{nh}$
\[
|t(z)| = \left| z + \frac{n - 1}{n} \frac{1}{1 + n^{-1} \operatorname{E} \operatorname{tr} D_{ij}^{-1}(z)} \right|^{-1} = \left| \frac{1 + n^{-1} \operatorname{E} \operatorname{tr} D_{ij}^{-1}(z)}{z(1 + n^{-1} \operatorname{E} \operatorname{tr} D_{ij}^{-1}(z)) + \frac{n-1}{n}} \right|
\]
\[
\leq \frac{1}{|z|} \left[ 1 + \frac{1}{n} \frac{1}{\operatorname{E} z(1 + n^{-1} \operatorname{E} \operatorname{tr} D_{ij}^{-1}(z))} \right] \leq \frac{K}{v}
\]
since $a_l \leq |z| \leq b_r + 1$. Thus, by Lemmas 6.6, 6.4 and Cauchy-Schwartz inequality, we have
\[
\mathbb{E} |\operatorname{tr}(B(z_1) \mathbb{E} D_j^{-1}(z_2))| = \mathbb{E} \left| \sum_{i,j}^n t(z_1)(\beta_{ij}(z_1) - b_{ij}(z_1)) r_i^* D_{ij}^{-1}(z_1) \mathbb{E} D_j^{-1}(z_2) r_i \right|
\]
\[
\leq \frac{Kn}{v} \mathbb{E} (|\beta_{ij}(z_1) - b_{ij}(z_1)| r_i^* D_{ij}^{-1}(z_1) \mathbb{E} D_j^{-1}(z_2) r_i)
\]
\[
\leq \frac{Kn}{v} \frac{1}{\sqrt{nv}} \frac{1}{\sqrt{nv}} = \frac{K}{v^2}.
\]
From Lemma 2.10 of Bai and Silverstein [5] for any $n \times n$ matrix $A$,
\[
|\operatorname{tr}(D_j^{-1}(z) - D_j^{-1}(z)) A| \leq \frac{|A|}{v},
\]
which combined with Lemma 6.6 gives
\[
\mathbb{E} |\operatorname{tr}(C(z_1) \mathbb{E} D_j^{-1}(z_2))|
\]
\[
= \mathbb{E} \left| \frac{1}{n} t(z_1) b_{ij}(z_1) \sum_{i,j}^n \operatorname{tr}((D_{ij}^{-1}(z_1) - D_j^{-1}(z_1)) \mathbb{E} D_j^{-1}(z_2)) \right|
\]
\[
\leq \frac{K}{v} (|b_{ij}(z_1)|^2 \frac{1}{2} (\mathbb{E} |\operatorname{tr}(D_{ij}^{-1}(z_1) - D_j^{-1}(z_1)) \mathbb{E} D_j^{-1}(z_2)|^2)^\frac{1}{2}
\]
\[
\leq \frac{K}{v^2} \frac{1}{v^2} = \frac{K}{v^3}.
\]
From the above estimates (4.15) and (4.17), we arrive at
\[
\operatorname{tr} \mathbb{E} D_j^{-1}(z_1) \mathbb{E} D_j^{-1}(z_2)
\]
\[
= -t(z_1) \operatorname{tr} \mathbb{E} D_j^{-1}(z_2) + b_{ij}(z_1) \operatorname{tr} \mathbb{E} A(z_1) \mathbb{E} D_j^{-1}(z_2) + \frac{K}{v^4}.
\]
Using the identity
\[
D_j^{-1}(z_2) - D_j^{-1}(z_2) = -\beta_{ij}(z_2) D_{ij}^{-1}(z_2) r_i r_i^* D_{ij}^{-1}(z_2),
\]
we can write
\[
\operatorname{tr} \mathbb{E} (A(z_1)) D_j^{-1}(z_2) = A_1(z_1, z_2) + A_2(z_1, z_2) + A_3(z_1, z_2),
\]
where

\[ A_1(z_1, z_2) = -\text{tr} \sum_{i<j} t(z_1) r_i r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1))(D_{ij}^{-1}(z_2)) = -\sum_{i<j} t(z_1) \beta_{ij}(z_2) r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) r_i, \]

\[ A_2(z_1, z_2) = -\text{tr} \sum_{i<j} t(z_1) \frac{1}{n} \mathbb{E}_j(D_{ij}^{-1}(z_1))(D_{ij}^{-1}(z_2) - D_{ij}^{-1}(z_2)) \]

and

\[ A_3(z_1, z_2) = -\text{tr} \sum_{i<j} t(z_1) \left( r_i r_i^* - \frac{1}{n} I_p \right) \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2). \]

From (4.16), we get

\[ |A_2(z_1, z_2)| = \left| \frac{1}{n} \sum_{i<j} t(z_1) \text{tr}(D_{ij}^{-1}(z_2) - D_{ij}^{-1}(z_2)) \mathbb{E}_jD_{ij}^{-1}(z_1) \right| \]

\[ \leq \frac{j - 1}{n} \frac{K}{\nu^2} \leq K, \]

and by Lemma 6.3, we have

\[ \mathbb{E}|A_3(z_1, z_2)| \leq \frac{K(j - 1)}{v} \left| \text{tr} \left( r_i r_i^* - \frac{1}{n} I_p \right) \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) \right| \]

\[ \leq \frac{K}{v} \frac{1}{\sqrt{nv^2}} = \frac{K \sqrt{n}}{v^3}. \]

For \( A_1(z_1, z_2) \), by Lemmas 6.4 and 6.5,

\[ \mathbb{E} \left| r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) r_i r_i^* D_{ij}^{-1}(z_2) r_i \right. \]

\[ - \frac{1}{n^2} \text{tr} \left[ \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) \text{tr} D_{ij}^{-1}(z_2) \right] \]

\[ \leq \mathbb{E} \left| r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) r_i \right. \]

\[ \left. - \frac{1}{n} \text{tr} \left( \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) \right) r_i D_{ij}^{-1}(z_2) r_i \right| \]

\[ + \mathbb{E} \left| \frac{1}{n} \text{tr} \left( \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) \right) \right. \]

\[ \left. r_i^* D_{ij}^{-1}(z_2) r_i - \frac{1}{n} \text{tr} D_{ij}^{-1}(z_2) \right| \]

\[ \leq \frac{K}{\sqrt{nv^2}}. \]

Let \( \psi_j(z_1, z_2) = \text{tr}(\mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2)) \). Using the identity (4.16), we have

\[ \left| \text{tr} \left( \mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2) \right) \text{tr} D_{ij}^{-1}(z_2) - \psi_j(z_1, z_2) \text{tr} D_{ij}^{-1}(z_2) \right| \leq Knv^{-3}. \]

Thus, in conjunction with Lemma 6.6, we can get

\[ \mathbb{E} \left| A_1(z_1, z_2) + \frac{j - 1}{n^2} t(z_1) b_{12}(z_2) \psi_j(z_1, z_2) \text{tr} D_{ij}^{-1}(z_2) \right| \leq \frac{K}{\sqrt{nv^3}}. \]
Therefore, from (4.14)-(4.22), it follows that
\[
\phi_j(z_1, z_2) \left[ 1 + \frac{j - 1}{n^2} t(z_1) b_{12}(z_1) b_{12}(z_2) \text{tr} D_j^{-1}(z_2) \right] \\
= -\text{tr}(t(z_1) \text{tr} D_j^{-1}(z_2)) + A_4(z_1, z_2),
\]
where \( \mathbb{E}|A_4(z_1, z_2)| \leq K \sqrt{n}/v^3. \)

Using Lemma 6.6, the expression for \( D_j^{-1}(z_2) \) in (4.14) and the following estimate
\[
\mathbb{E}|\text{tr} A(z)| = \mathbb{E} \left| \text{tr} \sum_{i \neq j} t(z)(r_i r_i^* - n^{-1} I_p) D_i^{-1}(z) \right| \\
\leq \frac{Kn}{v} \mathbb{E}|r_i D_i^{-1}(z) r_i^* | - n^{-1} \text{tr} D_i^{-1}(z) | \leq \frac{K \sqrt{n}}{v^2},
\]
we find
\[
\phi_j(z_1, z_2) \left[ 1 + \frac{j - 1}{n^2} t(z_1) b_{12}(z_1) t(z_2) b_{12}(z_2) \right] \\
= -pt(z_1) t(z_2) + A_5(z_1, z_2),
\]
where
\[
\mathbb{E}|A_5(z_1, z_2)| \leq K \frac{\sqrt{n}}{v^3}.
\]

By Lemma 6.6, we can write
\[
\phi_j(z_1, z_2) \left[ 1 - \frac{(j - 1)p}{n^2} t(z_1) b_{12}(z_1) t(z_2) b_{12}(z_2) \right] \\
= \frac{1}{z_1 z_2} \left( \frac{\sum_{m} \sum_{n}^0(z_1) \sum_{m}^0(z_2)}{(\sum_{m}^0(z_1) + 1)(\sum_{m}^0(z_2) + 1)} + A_6(z_1, z_2),
\]
where \( \mathbb{E}|A_6(z_1, z_2)| \leq K \sqrt{n}/v^3. \)

Let
\[
a_n(z_1, z_2) = \frac{y_n \sum_{m}^0(z_1) \sum_{m}^0(z_2)}{(s_n^0(z_1) + 1)(s_n^0(z_2) + 1)},
\]
(4.12) can be written as
\[
a_n(z_1, z_2) \frac{1}{n} \sum_{j=1}^{n} \left( 1 - \frac{j - 1}{n} a_n(z_1, z_2) \right)^{-1} + A_7(z_1, z_2),
\]
where
\[
\mathbb{E}|A_7(z_1, z_2)| \leq \frac{K}{\sqrt{n}v^3}.
\]
Since

\[ a_n(z_1, z_2) \to a(z_1, z_2) = \frac{yg(z_1)g(z_2)}{(g(z_1) + 1)(g(z_2) + 1)}, \]

as \( n \to \infty \), we arrive at

\[ (4.12) \xrightarrow{pr.} a(z_1, z_2) \int_0^1 \frac{1}{1 - ta(z_1, z_2)} dt = -\ln(1 - a(z_1, z_2)) \]

\[ = -\ln \frac{l(z_1, z_2)}{g(z_1) - g(z_2)}, \]

where \( l(z_1, z_2) = g(z_1)g(z_2)(z_1 - z_2) \), which implies

\[ \Gamma_{m_1}(z_1, z_2) = (\kappa_1 + 1)b_n(z_1)b_n(z_2)\frac{1}{n^2} \sum_{j=1}^n \text{tr} E_j D_j^{-1}(z_1)E_j D_j^{-1}(z_2) \]

\[ = - (\kappa_1 + 1) \ln(l(z_1, z_2)) + (\kappa_1 + 1) \ln(g(z_1) - g(z_2)) + o_\rho(1). \]

Thus, adding the vertical parts of both contours and using the fact that \( f_m''(z) \) and \( g_m''(z) \) are analytic functions, the integral of the first term of \( \Gamma_{n_1}(z_1, z_2) \) is

\[ -\frac{1}{4\pi^2} \int_{\gamma_{m_1} \times \gamma_{m_2}} f_m''(z_1)g_m''(z_2)(\kappa_1 + 1) \ln(l(z_1, z_2)) dz_1 dz_2 \]

\[ = -\frac{\kappa_1 + 1}{4\pi^2} \int_{\gamma_{m_1} \times \gamma_{m_2}} f_m''(z_1)g_m''(z_2) \ln(l(z_1, z_2)) dz_1 dz_2 + O(v) \]

\[ = o(1). \]

For the second term of \( \Gamma_{n_1}(z_1, z_2) \), since \( s(z) = \overline{g(z)} \), as \( n \to \infty \), \( a_i \to a \) and \( b_r \to b \), we get

\[ -\frac{\kappa_1 + 1}{4\pi^2} \int_{\gamma_{m_1} \times \gamma_{m_2}} f_m''(z_1)g_m''(z_2) \ln(g(z_1) - g(z_2)) dz_1 dz_2 + o_\rho(1) \]

\[ \to -\frac{\kappa_1 + 1}{2\pi^2} \int_a^b \int_a^b f'(x_1)g'(x_2) \ln \left| \frac{s(x_1) - s(x_2)}{s(x_1) - s(x_2)} \right| dx_1 dx_2, \]

which is (1.4) in Theorem 1.1.

5. Mean function

In this section, we will find the limit of

\[ \mathbb{E} G_n(f_m) = -\frac{1}{2\pi i} \oint_{\gamma_m} f_m(z) p[\mathbb{E} s_n(z) - s_n^0(z)] dz. \]

We shall first consider \( M_n(z) = p[\mathbb{E} s_n(z) - s_n^0(z)] = n[\mathbb{E} s_n(z) - s_n^0(z)]. \)
Since $D(z) + zI = \sum_{j=1}^{n} r_j r_j^*$, multiplying by $D^{-1}(z)$ on the right side and using (4.13), we find

$$I + zD^{-1}(z) = \sum_{j=1}^{n} r_j r_j^* D^{-1}(z) = \sum_{j=1}^{n} \frac{r_j r_j^* D^{-1}(z)}{1 + r_j^* D^{-1}(z) r_j}.$$  

Taking trace, dividing by $n$ on both sides and combining with the identity $z s_n(z) = -1 + y_n + y_n z s_n(z)$ lead to

$$s_n(z) = -\frac{1}{nz} \sum_{j=1}^{n} \frac{1}{1 + r_j^* D^{-1}(z) r_j} = -\frac{1}{nz} \sum_{j=1}^{n} \beta_j(z).$$ (5.1)

Then, using (4.13) once again and $A^{-1} - B^{-1} = -A^{-1}(A - B)B^{-1}$, we get

$$\frac{I_p}{z(\mathbb{E}s_n(z) + 1)} - D^{-1}(z) = -\frac{1}{z(\mathbb{E}s_n(z) + 1)} \left[ \sum_{j=1}^{n} r_j r_j^* z \mathbb{E}s_n(z) \right] D^{-1}(z)$$

$$= -\frac{1}{z(\mathbb{E}s_n(z) + 1)} \sum_{j=1}^{n} \left[ \beta_j(z) r_j r_j^* D^{-1}(z) - \mathbb{E}(\beta_j(z)) \frac{1}{n} D^{-1}(z) \right].$$

Taking trace, dividing by $p$ and taking expectation, we derive

$$\omega_n(z) = -\frac{1}{z(\mathbb{E}s_n(z) + 1)} - \mathbb{E}s_n(z)$$

$$= -\frac{1}{pz(\mathbb{E}s_n(z) + 1)} \sum_{j=1}^{n} \mathbb{E}(\beta_j(z)) d_j(z)$$

$$= -\frac{1}{pz(\mathbb{E}s_n(z) + 1)} J_n(z),$$ (5.2)

where

$$d_j(z) = r_j^* D^{-1}(z) r_j - \frac{1}{n} \text{tr} D^{-1}(z).$$

On the other hand, by the identity $\mathbb{E}s_n(z) = -(1 - y_n) z^{-1} + y_n \mathbb{E}s_n(z)$, we have

$$\omega_n(z) = \frac{\mathbb{E}s_n(z)}{y_n z} \left( -z - \frac{1}{\mathbb{E}s_n(z)} + \frac{y_n}{\mathbb{E}s_n(z) + 1} \right) = \frac{\mathbb{E}s_n(z)}{y_n z} R_n(z),$$

where

$$R_n(z) = -z - \frac{1}{\mathbb{E}s_n(z)} + \frac{y_n}{\mathbb{E}s_n(z) + 1},$$

which implies that

$$\mathbb{E}s_n(z) = \left( -z + \frac{y_n}{\mathbb{E}s_n(z) + 1} - R_n(z) \right)^{-1}.$$ (5.3)
For $s_n^0(z)$, since $s_n^0(z) = (1 - y_n - y_n z_n^0(z) - z)^{-1}$ and $s_n^0(z) = -(1 - y_n)z^{-1} + y_n s_n^0(z)$, we have

\begin{equation}
(5.4) \quad s_n^0(z) = \left( -z + \frac{y_n}{s_n^0(z) + 1} \right)^{-1}.
\end{equation}

By (5.3) and (5.4), we get

\[
\begin{align*}
\mathbb{E}s_n(z) - s_n^0(z) &= \left( -z + \frac{y_n}{\mathbb{E}s_n(z) + 1} - R_n(z) \right)^{-1} - \left( -z + \frac{y_n}{s_n^0(z) + 1} \right)^{-1} \\
&= \mathbb{E}s_n(z) s_n^0(z) \left( \frac{y_n}{s_n^0(z) + 1} - \frac{y_n}{\mathbb{E}s_n(z) + 1} + R_n(z) \right) \\
&= \frac{y_n \mathbb{E}s_n(z) s_n^0(z)}{(s_n^0(z) + 1)(\mathbb{E}s_n(z) + 1)} \left( \mathbb{E}s_n(z) - s_n^0(z) \right) + \mathbb{E}s_n(z) s_n^0(z) R_n(z),
\end{align*}
\]

which combined with (5.2) leads to

\begin{equation}
(5.5) \quad n(\mathbb{E}s_n(z) - s_n^0(z)) \left( 1 - \frac{y_n \mathbb{E}s_n(z) s_n^0(z)}{(s_n^0(z) + 1)(\mathbb{E}s_n(z) + 1)} \right) = n\mathbb{E}s_n(z) s_n^0(z) R_n(z)
\end{equation}

Thus, in order to find the limit of $M_n(z) = n[\mathbb{E}s_n(z) - s_n^0(z)]$, it suffices to find the limit of $J_n(z)$. Let $\tilde{a}_j(z) = r_j D_j^{-1}(z) r_j - \frac{1}{n} \text{tr} D_j^{-1}(z)$ and $J_n(z) = \sum_{j=1}^n \mathbb{E}(\beta_j(z) \tilde{a}_j(z))$. By (4.3), we have

\[
J_n(z) = \tilde{J}_n(z) + \sum_{j=1}^n \mathbb{E} \left[ \beta_j(z) \left( \frac{1}{n} \text{tr} D_j^{-1}(z) - \frac{1}{n} \text{tr} \mathbb{E} D_j^{-1}(z) \right) \right]
\]

\[
= \tilde{J}_n(z) + \sum_{j=1}^n \mathbb{E} \left[ (\beta_j(z) - b_n(z)) \left( \frac{1}{n} \text{tr} D_j^{-1}(z) - \frac{1}{n} \text{tr} \mathbb{E} D_j^{-1}(z) \right) \right]
\]

\[
= \tilde{J}_n(z) - T_1(z) + T_2(z),
\]

where from (4.16)

\[
T_1(z) = \sum_{j=1}^n \mathbb{E} \left[ b_n^2(z) \delta_j(z) \left( \frac{1}{n} \text{tr} D_j^{-1}(z) - \frac{1}{n} \text{tr} \mathbb{E} D_j^{-1}(z) \right) \right]
\]

\[
= \sum_{j=1}^n \mathbb{E} \left[ b_n^2(z) \delta_j(z) \frac{1}{n} \left( \text{tr}(D_j^{-1}(z) - D_j^{-1}(z)) - \text{tr}(D_j^{-1}(z) - D_j^{-1}(z)) \right) \right]
\]

\[
\leq \sum_{j=1}^n \frac{K}{\sqrt{nv}} \cdot \frac{K}{nv} = \frac{K}{\sqrt{nv^2}}.
\]
It follows from Bai and Silverstein [5] (4.3) that for $l \geq 2$

\[
E \left| \frac{1}{n} tr D^{-1}(z) - \frac{1}{n} tr \mathbb{E} D^{-1}(z) \right|^l \leq \frac{K}{(\sqrt{n}v)^l}.
\]

Hence,

\[
T_2(z) = \sum_{j=1}^{n} \mathbb{E} \left[ \beta_j(z)b_n^2(z)\hat{\delta}_j^2(z) \left( \frac{1}{n} tr D^{-1}(z) - \frac{1}{n} tr \mathbb{E} D^{-1}(z) \right) \right] \leq \frac{K}{\sqrt{nv^3}}.
\]

From the above estimates on $T_1$ and $T_2$, we conclude that

\[
J_n(z) = \bar{J}_n(z) + \varepsilon_n,
\]

here and in the sequel, $\varepsilon_n = O(\sqrt{nv^3})$.

Now we only need to consider the limit of $\bar{J}_n(z)$. By (4.2), we write

\[
\bar{J}_n(z) = \sum_{j=1}^{n} \mathbb{E}[(\beta_j(z) - \bar{\beta}_j(z))e_j(z)] + \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}[\beta_j(z)tr(D_j^{-1}(z) - D^{-1}(z))]
\]

\[
= - \sum_{j=1}^{n} \mathbb{E}(\bar{\beta}_j^2(z)e_j^2(z)) + \sum_{j=1}^{n} \mathbb{E}(\bar{\beta}_j^2(z)\beta_j(z)e_j^2(z)) + \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}(\bar{\beta}_j^2(z)r_j D_j^{-2}(z)r_j)
\]

\[
\triangleq \bar{J}_{n1}(z) + \bar{J}_{n2}(z) + \bar{J}_{n3}(z).
\]

From Lemmas 6.3 and 6.6, we find

\[
|\bar{J}_{n2}(z)| \leq K \sum_{j=1}^{n} \left( \mathbb{E}e_j^2(z) \right)^{1/2} \leq \frac{K}{\sqrt{nv^3}}.
\]

By Lemma 6.6, $\bar{\beta}_j(z), \beta_j(z)$ and $b_n(z)$ can be replaced by $-\zeta s(z)$, thus we get

\[
\bar{J}_{n3}(z) = \zeta^2 s^2(z) \frac{1}{n^2} \sum_{j=1}^{n} \mathbb{E}tr D_j^{-2}(z) + \varepsilon_n \triangleq \zeta^2 s^2(z) \psi_n(z) + \varepsilon_n.
\]

By the identity of quadric form (4.11) and the fact from Lemma 6.7 that $\mathbb{E}[D_j^{-1}(z)]_{ii}$ can be replaced by $s(z) = -z^{-1}(s(z) + 1)^{-1}$, we have

\[
\bar{J}_{n1}(z) = -\zeta^2 s^2(z) \sum_{j=1}^{n} \mathbb{E}e_j^2(z) + \varepsilon_n
\]

\[
= -\zeta^2 s^2(z) \sum_{j=1}^{n} \mathbb{E} \left[ \sum_{i=1}^{p} k_i [D_j^{-1}(z)]_{ii}^2 + k_1 tr D_j^{-2}(z) + tr D_j^{-2}(z) \right] + \varepsilon_n
\]

\[
= y_n k_2 \zeta^2 s^2(z) - \zeta^2 s^2(z)(k_1 + 1) \psi_n(z) + \varepsilon_n,
\]

\[
E \left| \frac{1}{n} tr D^{-1}(z) - \frac{1}{n} tr \mathbb{E} D^{-1}(z) \right|^l \leq \frac{K}{(\sqrt{n}v)^l}.
\]
where $\kappa_1, \kappa_2$ and $k(z)$ were defined in Theorem 1.1. Now our goal is to find the limit of $\psi_n(z)$. Using the expansion of $D_{j}^{-1}(z)$ in (4.14), we get

$$
\psi_n(z) = \frac{1}{n^2} \sum_{j=1}^{n} \left( \frac{p}{z + z^2(z)} \right)^2 + \frac{z^2(z)}{n^2} \sum_{j=1}^{n} \mathbb{E} \text{tr} A^2(z) + \epsilon_n
$$

$$
= \frac{k^2(z)}{n^2} \sum_{j=1}^{n} \sum_{i \neq j}^{n} \mathbb{E} \text{tr} \left[ \left( r_i r_j^* - \frac{1}{n} I \right) D_{ij}^{-1}(z) D_{ij}^{-1}(z) \left( r_i r_j^* - \frac{1}{n} I \right) \right]
$$

$$
+ \frac{1}{n^2} \sum_{j=1}^{n} \frac{p}{z^2(z) + 1}^2 + \epsilon_n.
$$

Note that the cross terms will be 0 if either $D_{ij}^{-1}(z)$ or $D_{lj}^{-1}(z)$ is replaced by $D_{ij}^{-1}(z)$, where $D_{lj}(z) = D_{ij}(z) - r_i r_j^* = D_{lj}^{-1}(z) - r_i r_j^*$ and

$$
D_{ij}^{-1}(z) - D_{lj}^{-1}(z) = -\frac{D_{ij}^{-1}(z) r_j^* D_{lj}^{-1}(z)}{1 + r_i^* D_{lj}^{-1}(z) r_i}.
$$

Therefore, by (4.16), we conclude that the sum of cross terms is negligible and bounded by $K(\sqrt{n}v)$, thus we find

$$
\frac{1}{n^2} \sum_{j=1}^{n} \sum_{i \neq j}^{n} \mathbb{E} \text{tr} \left[ \left( r_i r_j^* - \frac{1}{n} I \right) D_{ij}^{-1}(z) D_{ij}^{-1}(z) \left( r_i r_j^* - \frac{1}{n} I \right) \right]
$$

$$
= \frac{1}{n^2} \sum_{j=1}^{n} \sum_{i \neq j}^{n} \mathbb{E} \text{tr} \left[ \left( r_i r_j^* - \frac{1}{n} I \right) D_{ij}^{-1}(z) \left( r_i r_j^* - \frac{1}{n} I \right) \right] + \epsilon_n
$$

$$
= \frac{1}{n^2} \sum_{j=1}^{n} \sum_{i \neq j}^{n} \mathbb{E} \left[ (r_i^* D_{ij}^{-1}(z) r_j)(r_j^* r_i) \right] + \epsilon_n
$$

$$
= \frac{1}{n^2} \sum_{j=1}^{n} \sum_{i \neq j}^{n} \frac{1}{n^2} \mathbb{E} \left[ \text{tr} D_{ij}^{-1}(z)(p + O(1)) \right] + \epsilon_n = y_n \psi_n(z) + \epsilon_n.
$$

From above, we get

$$
\psi_n(z) = \frac{y_n}{z^2(z) + 1} + y_n k^2(z) \psi_n(z) + \epsilon_n.
$$

Combined with (5.7), we have

$$
J_n(z) = \kappa_2 y_n k^2(z) - \frac{\kappa_1 y_n k^2(z)}{1 - y_n k^2(z)} + \epsilon_n.
$$
Thus, from (5.5), it follows that

\[
M_n(z) = \frac{n\mathbb{E}_z(z)\mathcal{S}_n(z)R_n(z)}{1 - \frac{\mathcal{S}_n(z)\mathcal{S}_n(z)}{(\mathcal{S}_n(z)+1)(\mathcal{S}_n(z)+1)}} = \frac{-\mathcal{S}_n(z)I_n(z)}{1 - \frac{\mathcal{S}_n(z)\mathcal{S}_n(z)}{(\mathcal{S}_n(z)+1)(\mathcal{S}_n(z)+1)}}
\]

\[
= \frac{\kappa_1y_nk^3(z)}{(1 - y_nk^2(z))^2} - \frac{\kappa_2y_nk^3(z)}{1 - y_nk^2(z)} + \tilde{e}_n
\]

\[
\triangleq \tilde{M}_1(z) + \tilde{M}_2(z) + \tilde{e}_n.
\]

Therefore, we can calculate the mean function in the following two parts.

\[
-\frac{1}{2\pi i} \int_{\gamma_{ab}} f_m(z)\tilde{M}_1(z)dz = \frac{\kappa_1}{2\pi i} \int_{\gamma_{ab}} f_m(z)\frac{y_nk^3(z)}{(1 - y_nk^2(z))^2}dz
\]

\[
\quad = \frac{\kappa_1}{4\pi i} \int_{\gamma_{ab}} f_m(z)\frac{d}{dz}\ln(1 - y_nk^2(z))dz = -\frac{\kappa_1}{4\pi i} \int_{\gamma_{ab}} f'_m(z)\ln(1 - y_nk^2(z))dz
\]

\[
\quad \rightarrow \frac{\kappa_1}{2\pi} \int_a^b f'(x)\text{arg}\left(1 - yk^2(x)\right)dx,
\]

as \( n \to \infty, a_l \to a \) and \( b_r \to b \). Similarly,

\[
-\frac{1}{2\pi i} \int_{\gamma_{ab}} f_m(z)\tilde{M}_2(z)dz = \frac{\kappa_2}{2\pi i} \int_{\gamma_{ab}} f_m(z)\frac{y_nk^3(z)}{1 - y_nk^2(z)}dz
\]

\[
\quad \rightarrow -\frac{\kappa_2}{\pi} \int_a^b f(x)\Im\left(\frac{yk^3(x)}{1 - yk^2(x)}\right)dx.
\]

Hence, summing the two terms, we obtain the mean function of the limiting distribution in (1.3).

### 6. Appendix

**Lemma 6.1.** *Under the conditions in Theorem 1.1, we have*

\[
\|\mathbb{E}_n - F\| = O(n^{1/2}), \quad \|F_n - F\| = O_p(n^{-2/5}),
\]

\[
\|F_n - F\| = O(n^{-2/(5+\eta)}) \text{ a.s., for any } \eta > 0.
\]

This follows from Theorems 1.1,1.2 and 1.3 in [4].

**Lemma 6.2.** [Burkholder (1973)] *Let \( X_k, k = 1, 2, \ldots, \) be a complex martingale difference sequence with respect to the increasing \( \sigma \)-fields \( \mathcal{F}_k \). Then, for \( p > 1 \),

\[
\mathbb{E}\left|\sum X_k\right|^p \leq K_p\mathbb{E}\left(\sum |X_k|^2\right)^{p/2}.
\]

In the reference [9], only real variables were considered. It is straightforward to extend to complex cases.
Lemma 6.3. For $x = (x_1, \ldots, x_n)^\top$ i.i.d. standardized real or complex entries with $\mathbb{E}x_i = 0$ and $\mathbb{E}|x_i|^2 = 1$, and $C$ is an $n \times n$ complex matrix, we have, for any $p \geq 2$

$$\mathbb{E}|x^*Cx - trC|^p \leq K_p \left( [\mathbb{E}|x_1|^4 trCC^*)^{p/2} + \mathbb{E}|x_1|^{2p} tr(CC^*)^{p/2} \right).$$

This is Lemma 8.10 in [7].

Lemma 6.4. For any nonrandom $p \times p$ matrix $A$, $\mathbb{E}|r_1^*Ar_1|^2 \leq Kn^{-1}\|A\|^2$.

Proof. For nonrandom $p \times p$ matrix $A$,

$$\mathbb{E}|r_1^*Ar_1|^2 = \frac{1}{n^2} \mathbb{E} \left\| \sum_{k=1}^p x_k a_{k,1} x_k \right\|^2$$

$$= \frac{1}{n^2} \mathbb{E} \left( \sum_{k=1}^p x_k^2 a_{k,1}^2 + \sum_{k \neq k} |x_k|^2 |a_{k,1}|^2 a_{k,1} a_{k,1} + \sum_{j=1}^p |x_1|^4 a_{j,1}^2 \right)$$

$$\leq K \frac{1}{n^2} \mathbb{E} \left( \sum_{k=1}^p |a_{k,1}|^2 \right) = Kn^{-2} \mathbb{E} tr(A^2) \leq Kn^{-1}\|A\|^2.$$

Lemma 6.5. For nonrandom $p \times p$ matrices $A_k, k = 1, \ldots, s$,

$$\mathbb{E} \left\| \prod_{k=1}^{s} (r_1^*A_1r_1 - \frac{1}{n} trA_1) \right\| \leq Kn^{-s} \delta_n^{2(s-4)\gamma_0} \prod_{k=1}^{s} \|A_k\|.$$  

Proof. Recalling the truncation steps $\mathbb{E}|x_1|^{8} < \infty$ and Lemma 6.3, we have, for all $l > 1$,

$$\mathbb{E} |r_1^*A_1r_1 - n^{-1} trA_1|^l \leq K\|A_1\|^l n^{-l}(\sqrt{n}\delta_n)^{(2l-4)\gamma_0}$$

Then, (6.1) is the consequence of (6.2) and H"older inequality.

Lemma 6.6. Under the conditions in Theorem 1.1, for any $l \geq 2$, $\mathbb{E}|\beta_j(z)|^l, \mathbb{E}|\tilde{\beta}_j(z)|^l, and |\hat{b}_n(z)|^l$ are uniformly bounded in $\gamma_{mh}$. Furthermore, $\beta_j(z), \tilde{\beta}_j(z)$ and $b_n(z)$ are uniformly convergent in probability to $-z\overline{\gamma}(z)$ in $\gamma_{mh}$.

Proof. By (4.2) and (4.3) in [5], we have, for any $l \leq 2$,

$$\mathbb{E}|trD_j^{-1}(z) - trD_j^{-1}(z)|^l \leq Kn^l, $$

$$\mathbb{E}|r_j D_j^{-1}(z) - 1/n trD_j^{-1}(z)|^l \leq Kn^l.$$
This lemma follows from Lemma 6.3, (6.3), (6.4) and the facts below.

Fact 1. Since \( s_n^0(z) = -\frac{1}{2} \left( \frac{1}{z_n} - \frac{1}{y_n^c} \sqrt{z^2 - (1 + y_n)z + (1 - y_n)^2 - \frac{1}{y_n^c}} \right) \) and \( s_n^0(z) = -\frac{1-y_n}{z} + y_n s_n^0(z) \), we have

\[
\begin{align*}
\varphi^0_n(z) = -\frac{1}{2} \left( 1 - y_n + z - \sqrt{z^2 - (1 + y_n)z + (1 - y_n)^2} \right).
\end{align*}
\]

Thus, \( \varphi^0_n(z) \) is bounded in any bounded and closed complex region.

Fact 2.

\[
\begin{align*}
|b_n(z) - \mathbb{E}\beta_j(z)| &\leq \frac{1}{v^2} \mathbb{E} \left| r_j D_j^{-1}(z) r_j - \frac{1}{n} \text{tr} D_j^{-1}(z) \right| \\
&\leq \frac{1}{v^2} \left[ \mathbb{E} |e_j(z)| + \mathbb{E} \left| \frac{1}{n} \text{tr} D_j^{-1}(z) - \frac{1}{n} \text{tr} D_j^{-1}(z) \right| \right] \\
&\leq \frac{1}{v^2} \left( \frac{K}{\sqrt{nv}} + \frac{K}{\sqrt{nv}} \right) = \frac{K}{\sqrt{nv}}.
\end{align*}
\]

where the last inequality follows from (5.6).

Fact 3. Taking expectation on (5.1), one can find

\[
\begin{align*}
\mathbb{E}\varphi_n(z) = -\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\beta_j(z) = -\mathbb{E}\beta_j(z).
\end{align*}
\]

Fact 4. From Lemma 6.1, we have

\[
\begin{align*}
|\varphi_n(z) - \varphi_n^0(z)| &\leq \gamma_n \mathbb{E}|s_n(z) - s_n^0(z)| = \gamma_n \mathbb{E} \left| \int \frac{1}{x-z} (F^{B_0} - F^{V_n})(dx) \right| \\
&\leq \frac{K}{v} \|F^{B_0} - F^{V_n}\| = \frac{K}{v} O_p(n^{-\frac{3}{2}}) = O_p(n^{-\frac{3}{2}} v^{-1}).
\end{align*}
\]

\( \square \)

**Lemma 6.7.** Under the conditions in Theorem 1.1, as \( n \to \infty \)

\[
\max_{i,j} |\mathbb{E}|D_j^{-1}(z)|_{ii} - s(z)| \to 0 \text{ in probability}
\]

uniformly in \( \gamma_{mb} \), where the maximum is taken over all \( 1 \leq i \leq p \) and \( 1 \leq j \leq n \).

**Proof.** Firstly, let \( e_j \) (\( 1 \leq j \leq n \)) be the \( p \)-vector whose \( j \)-th element is 1, the rest being 0 and \( e_i^t \) the transpose of \( e_i \), then

\[
\begin{align*}
\mathbb{E}|[D^{-1}(z)]_{ii} - [D_j^{-1}(z)]_{ii}| &= \mathbb{E}|e_i^t(D^{-1}(z) - D_j^{-1}(z))e_i| \\
&= \mathbb{E} |\beta_j(z)e_j^t D_j^{-1}(z) r_j^t D_j^{-1}(z) e_j| \\
&\leq (\mathbb{E} |\beta_j(z)|^2)^{1/2} (\mathbb{E} |r_j^t D_j^{-1}(z) e_j e_j^t D_j^{-1}(z) r_j|^2)^{1/2} \leq \frac{K}{\sqrt{nv}}.
\end{align*}
\]
Secondly, by martingale inequality, for any $\epsilon > 0$, we have
\[
P(\max_{i,j} |E_j[D^{-1}(z)]_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}| > \epsilon) \]
\[
\leq \sum_{i=1}^{p} P(\max_{j} |E_j[D^{-1}(z)]_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}| > \epsilon) \]
\[
\leq \sum_{i=1}^{p} \frac{1}{\epsilon^6} \mathbb{E}[(D^{-1}(z))_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}]^6 \]
\[
= \frac{1}{\epsilon^6} \sum_{i=1}^{p} \mathbb{E} \left( \sum_{l=1}^{n} (E_l - \mathbb{E}_{l-1}) \beta_l(z) e'_i D^{-1}_l(z) r_l r'_l D^{-1}_l(z) e_l \right)^6 \]
\[
\leq K \sum_{i=1}^{p} \mathbb{E} \left( \sum_{l=1}^{n} |(E_l - \mathbb{E}_{l-1}) \beta_l(z) e'_i D^{-1}_l(z) r_l r'_l D^{-1}_l(z) e_l|^2 \right)^3. \]

Let $Z_l(z) = e'_i D^{-1}_l(z) r_l r'_l D^{-1}_l(z) e_l$. We have that
\[
|\mathbb{E} Z_l(z)| \leq \frac{K}{n^2} \quad \text{and} \quad \mathbb{E} |Z_l(z) - \mathbb{E} Z_l(z)|^2 \leq \frac{K}{n^2 v^2}.
\]

Thus we obtain
\[
P(\max_{i,j} |E_j[D^{-1}(z)]_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}| > \epsilon) \]
\[
\leq K \sum_{i=1}^{p} \mathbb{E} \left( \sum_{l=1}^{n} \frac{K}{n^2 v^4} \right)^3 = \frac{K}{n^2 v^{12}}. \]

Finally,
\[
\mathbb{E}[D^{-1}]_{ii} = \frac{1}{p} \sum_{i=1}^{p} \mathbb{E}[D^{-1}]_{ii} = \mathbb{E} s_n(z). \]

In section 5 it is proved that $p(\mathbb{E} s_n(z) - s(z))$ converges to 0 uniformly on $\gamma_{mb}$. The proof of Lemma 6.6 is complete. \qed

**References**


KLASMOE and SCHOOL OF MATHEMATICS & STATISTICS, NORTHEAST NORMAL UNIVERSITY, CHANGCHUN, 130024, P.R.C., AND DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY, NATIONAL UNIVERSITY OF SINGAPORE

_E-mail address:_ baizd@nenu.edu.cn

KLASMOE and SCHOOL OF MATHEMATICS & STATISTICS, NORTHEAST NORMAL UNIVERSITY, CHANGCHUN, 130024, P.R.C., AND DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY, NATIONAL UNIVERSITY OF SINGAPORE

_E-mail address:_ Wangxy022@gmail.com

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY, NATIONAL UNIVERSITY OF SINGAPORE, SINGAPORE 117546

_E-mail address:_ stazw@nus.edu.sg