Bernoulli **16**(4), 2010, 1086–1113
DOI: 10.3150/10-BEJ250

Functional CLT for sample covariance matrices

ZHIDONG BAI[1], XIAOYING WANG[2] and WANG ZHOU[3]

[1] *KLASMOE and School of Mathematics and Statistics, Northeast Normal University, Changchun, 130024, P.R. China and Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546. E-mail: baizd@nenu.edu.cn*
[2] *School of Mathematics and Physics North China Electric Power University, Beijing, 102206, P.R. China. E-mail: wangxy022@gmail.com*
[3] *Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546. E-mail: stazw@nus.edu.sg*

Using Bernstein polynomial approximations, we prove 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 $[(1-\sqrt{y})^2, (1+\sqrt{y})^2]$, the support of the Marčenko–Pastur law. We also derive the explicit expressions for asymptotic mean and covariance functions.

Keywords: Bernstein polynomial; central limit theorem; sample covariance matrices; Stieltjes transform

1. Introduction and main result

Let $X_n = (x_{ij})_{p\times n}$, $1 \le i \le p$, $1 \le j \le n$, be an observation matrix and $x_j = (x_{1j}, \ldots, x_{pj})^{\mathrm{t}}$ be the jth column of X_n. The sample covariance matrix is then

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

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 unbiased estimator of the population covariance matrix and, more importantly, many statistics in multivariate statistical analysis (e.g., principle component analysis, factor analysis and multivariate regression analysis) can be expressed as functionals of the empirical spectral distributions of sample covariance matrices. The *empirical spectral distribution* (ESD) of a symmetric (or Hermitian, in the complex case) $p \times p$ matrix A is defined as

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

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

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

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

since F^{B_n} and F^{S_n} have the same liming properties, according to Theorem 11.43 in [8]. We refer to [3] for a review of 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 proven 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

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

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

$$s_F(z) \triangleq \int \frac{1}{x-z}\,\mathrm{d}F(x), \qquad \Im z \neq 0.$$

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

$$s = \frac{1}{1 - y - z - yzs} \tag{1.1}$$

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

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

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

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

Instead of studying the convergence rate directly, Bai and Silverstein [7] 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 $G_n(x) = p[F^{B_n}(x) - F_{y_n}(x)]$. Bai and Silverstein proved the central limit theorem (CLT) for the LSS,

$$G_n(f) \triangleq \int_{-\infty}^{\infty} f(x)\,\mathrm{d}G_n(x), \qquad f \in \mathcal{A}(\mathcal{D}).$$

Their result is very useful for testing large-dimensional hypotheses. However, the analytic assumption on f seems inflexible in practical applications because in many cases of application, the kernel functions f can only be defined on the real line, instead of on the complex plane. On the other hand, it is proved in [8] that the CLT of LSS does not hold for indicator functions. Therefore, it is natural to ask what the weakest continuity condition is that should be imposed on the kernel functions so that the CLT of the LSS holds. For the CLT for other types of matrices, one can refer to [1].

In this paper, we consider the CLT for

$$G_n(f) \triangleq \int_{-\infty}^{\infty} f(x)\,\mathrm{d}G_n(x), \qquad f \in C^4(\mathcal{U}),$$

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

Denote by $\underline{s}(z)$ the Stieltjes transform of $\underline{F}_y(x) = (1-y)\mathbb{I}_{(0,\infty)}(x) + yF_y(x)$ and set $k(z) = \underline{s}(z)/(\underline{s}(z)+1)$, where, for $x \in \mathbb{R}$, $\underline{s}(x) = \lim_{z\to x+\mathrm{i}0} \underline{s}(z)$.

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* $\mathbb{E}x_{11} = 0$, $\mathbb{E}|x_{11}|^2 = 1$, $\mathbb{E}|x_{11}|^8 < \infty$ *and if* x_{ij} *are complex variables*, $\mathbb{E}x_{11}^2 = 0$;

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

The LSS $G_n = \{G_n(f)\colon f \in C^4(\mathcal{U})\}$ *then converges weakly in finite dimensions to a Gaussian process* $G = \{G(f)\colon f \in C^4(\mathcal{U})\}$ *with mean function*

$$\mathbb{E}G(f) = \frac{\kappa_1}{2\pi}\int_a^b f'(x)\arg\bigl(1 - yk^2(x)\bigr)\,\mathrm{d}x - \frac{\kappa_2}{\pi}\int_a^b f(x)\Im\left(\frac{yk^3(x)}{1-yk^2(x)}\right)\mathrm{d}x \tag{1.3}$$

and covariance function

$$\begin{aligned} c(f,g) &\triangleq \mathbb{E}[\{G(f) - \mathbb{E}G(f)\}\{G(g) - \mathbb{E}G(g)\}] \\ &= \frac{\kappa_1+1}{2\pi^2}\int_a^b\int_a^b f'(x_1)g'(x_2)\ln\left|\frac{\overline{\underline{s}(x_1)} - \underline{s}(x_2)}{\underline{s}(x_1) - \underline{s}(x_2)}\right|\mathrm{d}x_1\,\mathrm{d}x_2 &(1.4)\\ &\quad - \frac{\kappa_2 y}{2\pi^2}\int_a^b\int_a^b f'(x_1)g'(x_2)\Re[k(x_1)k(x_2) - \overline{k(x_1)}k(x_2)]\,\mathrm{d}x_1\,\mathrm{d}x_2, &(1.5)\end{aligned}$$

where the parameter $\kappa_1 = |\mathbb{E}x_{11}^2|^2$ *takes the value* 1 *if* x_{ij} *are real*, 0 *otherwise*, *and* $\kappa_2 = \mathbb{E}|x_{11}|^4 - \kappa_1 - 2$.

Remark 1.2. In the definition of $G_n(f)$, $\theta = \int f(x)\,\mathrm{d}F(x)$ can be regarded as a population parameter. The linear spectral statistic $\hat{\theta} = \int f(x)\,\mathrm{d}F_n(x)$ is then an estimator of θ. We remind the reader that the center $\theta = \int f(x)\,\mathrm{d}F(x)$, rather than $E\int f(x)\,\mathrm{d}F_n(x)$, 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 a statistical test of the ideal hypothesis. However, in this test, one cannot 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. The problem is then reduced to the analytic case. The truncation and renormalization steps are 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], then 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 that $\tilde{f}(y) \in C^4[0, 1]$. A Taylor expansion gives

$$\tilde{f}\left(\frac{k}{m}\right) = \tilde{f}(y) + \left(\frac{k}{m} - y\right)\tilde{f}'(y) + \frac{1}{2}\left(\frac{k}{m} - y\right)^2 \tilde{f}''(y) + \frac{1}{3!}\left(\frac{k}{m} - y\right)^3 \tilde{f}^{(3)}(y) + \frac{1}{4!}\left(\frac{k}{m} - y\right)^4 \tilde{f}^{(4)}(\xi_y),$$

where ξ_y is a number between k/m and y. Hence,

$$\tilde{f}_m(y) - \tilde{f}(y) = \frac{y(1-y)\tilde{f}''(y)}{2m} + \mathrm{O}\left(\frac{1}{m^2}\right). \tag{2.1}$$

For the function $f \in C^4(\mathcal{U})$, there exist $0 < a_\mathrm{l} < a < b < b_\mathrm{r}$ such that $[a_\mathrm{l}, b_\mathrm{r}] \subset \mathcal{U}$. If we let $\epsilon \in (0, 1/2)$ and perform a linear transformation $y = Lx + c$, where $L = (1 - 2\epsilon)/(b_\mathrm{r} - a_\mathrm{l})$ and $c = ((a_\mathrm{l} + b_\mathrm{r})\epsilon - a_\mathrm{l})/(b_\mathrm{r} - a_\mathrm{l})$, then $y \in [\epsilon, 1-\epsilon]$ if $x \in [a_\mathrm{l}, b_\mathrm{r}]$. Define $\tilde{f}(y) \triangleq f((y-c)/L) = f(x)$, $y \in [\epsilon, 1-\epsilon]$ and

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

From (2.1), we have

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

Since $\tilde{h}(y) \triangleq y(1-y)\tilde{f}''(y)$ has a second order derivative, we can once again use Bernstein polynomial approximation to get

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

So, with $h_m(x) = \tilde{h}_m(y)$,

$$f(x) = f_m(x) - \frac{1}{2m} h_m(x) + \mathrm{O}\left(\frac{1}{m^2}\right).$$

Therefore, $G_n(f)$ can be split into three parts:

$$\begin{aligned} G_n(f) &= p \int_{-\infty}^{\infty} f(x)[F^{B_n} - F_{y_n}](\mathrm{d}x) \\ &= p \int f_m(x)[F^{B_n} - F_{y_n}](\mathrm{d}x) - \frac{p}{2m} \int h_m(x)[F^{B_n} - F_{y_n}](\mathrm{d}x) \\ &\quad + p \int \left(f(x) - f_m(x) + \frac{1}{2m} h_m(x) \right) [F^{B_n} - F_{y_n}](\mathrm{d}x) \\ &= \Delta_1 + \Delta_2 + \Delta_3. \end{aligned}$$

For Δ_3, under the conditions in Theorem 1.1, by Lemma A.1 in the Appendix,

$$\|F^{B_n} - F_{y_n}\| = \mathrm{O}_p(n^{-2/5}),$$

where $a = \mathrm{O}_p(b)$ means that $\lim_{x\to\infty} \lim_{n\to\infty} P(|a/b| \geq x) = 0$.

Taking $m^2 = [n^{3/5+\epsilon_0}]$ for some $\epsilon_0 > 0$ and using integration by parts, we have that

$$\begin{aligned} \Delta_3 &= -p \int \left(f(x) - f_m(x) + \frac{1}{2m} h_m(x) \right)' \bigl(F_n(x) - F(x)\bigr)\,\mathrm{d}x \\ &= \mathrm{O}_p(n^{-\epsilon_0}) \end{aligned}$$

since $(f(x) - f_m(x) + \frac{1}{2m} h_m(x))' = \mathrm{O}(m^{-2})$. From now on, we choose $\epsilon_0 = 1/20$, so $m = [n^{13/40}]$.

Note that $f_m(x)$ and $h_m(x)$ are both analytic. Based on Conditions 4.1 and 4.2 in Section 4 and a martingale CLT ([9], Theorem 35.12), replacing f_m by h_m, we obtain

$$\Delta_2 = \frac{\mathrm{O}(\Delta_1)}{m} = \mathrm{o}_p(1).$$

It suffices to consider $\Delta_1 = G_n(f_m)$. Clearly, the two polynomials $f_m(x)$ and $\tilde{f}_m(y)$, defined only on the real line, can be extended to $[a_{\mathrm{l}}, b_{\mathrm{r}}] \times [-\xi, \xi]$ and $[\epsilon, 1-\epsilon] \times [-L\xi, L\xi]$, respectively.

Since $\tilde{f} \in C^4[0,1]$, there exists a constant M such that $|\tilde{f}(y)| < M \ \forall y \in [\epsilon, 1-\epsilon]$. Noting that for $(u, v) \in [\epsilon, 1-\epsilon] \times [-L\xi, L\xi]$,

$$
\begin{aligned}
|u + \mathrm{i}v| + |1 - (u + \mathrm{i}v)| &= \sqrt{u^2 + v^2} + \sqrt{(1-u)^2 + v^2} \\
&\le u\left[1 + \frac{v^2}{2u^2}\right] + (1-u)\left[1 + \frac{v^2}{2(1-u)^2}\right] \le 1 + \frac{v^2}{\epsilon},
\end{aligned}
$$

we have, for $y = Lx + c = u + \mathrm{i}v$,

$$
|\tilde{f}_m(y)| = \left|\sum_{k=0}^{m} \binom{m}{k} y^k (1-y)^{m-k} \tilde{f}\left(\frac{k}{m}\right)\right| \le M\left(1 + \frac{v^2}{\epsilon}\right)^m .
$$

If we take $|\xi| \le L/\sqrt{m}$, then $|\tilde{f}_m(y)| \le M(1 + L^2/(m\epsilon))^m \to M e^{L^2/\epsilon}$ as $m \to \infty$. Therefore, $\tilde{f}_m(y)$ is bounded when $y \in [\epsilon, 1-\epsilon] \times [-L/\sqrt{m}, L/\sqrt{m}]$. In other words, $f_m(x)$ is bounded when $x \in [a_{\mathrm{l}}, b_{\mathrm{r}}] \times [-1/\sqrt{m}, 1/\sqrt{m}]$.

Let $v = 1/\sqrt{m} = n^{-13/80}$ and γ_m be the contour formed by the boundary of the rectangle with vertices $(a_{\mathrm{l}} \pm \mathrm{i}v)$ and $(b_{\mathrm{r}} \pm \mathrm{i}v)$. Similarly, one can show that $h_m(x)$, $f_m'(x)$ and $h_m'(x)$ are bounded on γ_m.

3. Simplification by truncation and normalization

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

By condition (a) in Theorem 1.1, for any $\delta > 0$,

$$
\delta^{-8} \mathbb{E}|x_{11}|^8 \mathbb{I}_{\{|x_{11}| \ge \sqrt{n}\delta\}} \to 0,
$$

which implies the existence of a sequence $\delta_n \downarrow 0$ such that

$$
\delta_n^{-8} \mathbb{E}|x_{11}|^8 \mathbb{I}_{\{|x_{11}| \ge \sqrt{n}\delta_n\}} \to 0
$$

as $n \to \infty$. Let $\hat{x}_{ij} = x_{ij} \mathbb{I}_{\{|x_{ij}| \le \sqrt{n}\delta_n\}}$ and $\tilde{x}_{ij} = (\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij})/\sigma_n$, where $\sigma_n^2 = \mathbb{E}|\hat{x}_{ij} - \mathbb{E}\hat{x}_{ij}|^2$. We then have $\mathbb{E}\tilde{x}_{ij} = 0$ and $\sigma_n^2 \to 1$ as $n \to \infty$. We use $\hat{X}_n$ and $\tilde{X}_n$ to denote the analogs 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 analogs of B_n, and let $\hat{G}_n$ and $\tilde{G}_n$ be analogs of G_n. We then have

$$
\begin{aligned}
P(G_n \ne \hat{G}_n) &\le P(B_n \ne \hat{B}_n) \le npP\big(|x_{11}| \ge \sqrt{n}\delta_n\big) \\
&\le pn^{-3}\delta_n^{-8} \mathbb{E}|x_{11}|^8 \mathbb{I}_{\{|x_{11}| \ge \sqrt{n}\delta_n\}} = \mathrm{o}(n^{-2}).
\end{aligned}
\tag{3.1}
$$

From Yin, Bai and Krishnaiah [19], we know that $\lambda_{\max}^{\hat{B}_n}$ and $\lambda_{\max}^{\tilde{B}_n}$ are a.s. bounded by $b = (1+\sqrt{y})^2$. Let λ_j^A denote the jth largest eigenvalue of matrix A. Since

$$|\sigma_n^2 - 1| \le 2\mathbb{E}|x_{11}|^2 \mathbb{I}_{\{|x_{11}|\ge\sqrt{n}\delta_n\}} \le 2\big(\sqrt{n}\delta_n\big)^{-6}\mathbb{E}|x_{11}|^8 \mathbb{I}_{\{|x_{11}|\ge\sqrt{n}\delta_n\}} = \mathrm{o}(\delta_n^2 n^{-3})$$

and

$$|\mathbb{E}\hat{x}_{11}|^2 \le \mathbb{E}|x_{11}|^2 \mathbb{I}_{\{|x_{11}|\ge\sqrt{n}\delta_n\}} \le \mathrm{o}(\delta_n^2 n^{-3}),$$

we have

$$\begin{aligned}
&\left|\int f(x)\,\mathrm{d}\hat{G}_n(x) - \int f(x)\,\mathrm{d}\tilde{G}_n(x)\right| \\
&\quad\le K\sum_{j=1}^{p} |\lambda_j^{\hat{B}_n} - \lambda_j^{\tilde{B}_n}| \\
&\quad\le K\big(\mathrm{tr}(\hat{X}_n - \tilde{X}_n)(\hat{X}_n - \tilde{X}_n)^*\big)^{1/2} \\
&\quad\le 2(1-\sigma_n^{-1})^2\,\mathrm{tr}\,\hat{B}_n + 2\sigma_n^{-2}\,\mathrm{tr}\,\mathbb{E}\hat{X}_n\mathbb{E}\hat{X}_n^* \\
&\quad\le \frac{2(1-\sigma_n^2)^2}{\sigma_n^2(1+\sigma_n)^2} p\lambda_{\max}^{\hat{B}_n} + 2\sigma_n^{-2} np|\mathbb{E}\hat{x}_{11}|^2 = \mathrm{o}(\delta_n^2 n^{-1}).
\end{aligned} \tag{3.2}$$

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

$$\int f(x)\,\mathrm{d}G_n(x) = \int f(x)\,\mathrm{d}\tilde{G}_n(x) + \mathrm{o}_p(1).$$

Therefore, we only need to find the limiting distribution of $\int f(x)\,\mathrm{d}\tilde{G}_n(x)$ 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 $\mathbb{E}\tilde{x}_{11}^2 = \mathrm{o}(n^{-2})$ for complex variables. For brevity, in the sequel, we shall suppress the superscript on the variables and still use x_{ij} to denote the truncated and renormalized variable $\tilde{x}_{ij}$. Note that in this paper, we use K as a generic positive constant which is independent of n and which may differ from one line to the next.

4. Convergence of $\Delta - \mathbb{E}\Delta$

If we let $\underline{B}_n = n^{-1}X_n^*X_n$, then $F^{\underline{B}_n}(x) = (1-y_n)\mathbb{I}_{(0,\infty)}(x) + y_n F^{B_n}(x)$. Correspondingly, we define $\underline{F}_{y_n}(x) = (1-y_n)\mathbb{I}_{(0,\infty)}(x) + y_n F_{y_n}(x)$. Let $s_n(z)$ and $s_n^0(z)$ be the Stieltjes transforms of F^{B_n} and F_{y_n}, respectively; let $\underline{s}_n(z)$ and $\underline{s}_n^0(z)$ be the Stieltjes transforms of $F^{\underline{B}_n}$ and $\underline{F}_{y_n}$, respectively. By Cauchy's theorem, we then have

$$\Delta_1 = \frac{1}{2\pi \mathrm{i}}\int\oint_{\gamma_m} \frac{f_m(z)}{z-x} p[F^{B_n} - F_{y_n}](\mathrm{d}x)\,\mathrm{d}z = -\frac{1}{2\pi\mathrm{i}}\oint_{\gamma_m} f_m(z)p[s_n(z) - s_n^0(z)]\,\mathrm{d}z.$$

It is easy to verify that

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

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

$$r_j = \bigl(1/\sqrt{n}\bigr)x_j, \qquad D(z) = B_n - zI_p, \qquad D_j(z) = D(z) - r_j r_j^*,$$

$$\beta_j(z) = \frac{1}{1 + r_j^* D_j^{-1}(z) r_j}, \qquad \bar{\beta}_j(z) = \frac{1}{1 + (1/n)\operatorname{tr} D_j^{-1}(z)},$$

$$b_n(z) = \frac{1}{1 + (1/n)\mathbb{E}\operatorname{tr} D_j^{-1}(z)}, \qquad \varepsilon_j(z) = r_j^* D_j^{-1}(z) r_j - \frac{1}{n}\operatorname{tr} D_j^{-1}(z),$$

$$\delta_j(z) = r_j^* D_j^{-1}(z) r_j - \frac{1}{n}\operatorname{tr}\mathbb{E} D_j^{-1}(z)$$

and equalities

$$D^{-1}(z) - D_j^{-1}(z) = -\beta_j(z) D_j^{-1}(z) r_j r_j^* D_j^{-1}(z), \tag{4.1}$$

$$\beta_j(z) - \bar{\beta}_j(z) = -\beta_j(z)\bar{\beta}_j(z)\varepsilon_j(z) = -\bar{\beta}_j^2(z)\varepsilon_j(z) + \beta_j(z)\bar{\beta}_j^2(z)\varepsilon_j^2(z), \tag{4.2}$$

$$\beta_j(z) - b_n(z) = -\beta_j(z) b_n(z)\delta_j(z) = -b_n^2(z)\delta_j(z) + \beta_j(z) b_n^2(z)\delta_j^2(z). \tag{4.3}$$

Note that by (3.4) of Bai and Silverstein [6], the quantities $\beta_j(z)$, $\bar{\beta}_j(z)$ and $b_n(z)$ are bounded in absolute value by $|z|/v$.

Denote the σ-field generated by $r_1, \ldots, r_j$ by $\mathcal{F}_j = \sigma(r_1, \ldots, r_j)$, and let conditional expectations $\mathbb{E}_j(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_j)$ and $\mathbb{E}_0(\cdot) = \mathbb{E}(\cdot)$. Using the equality

$$D^{-1}(z) - D_j^{-1}(z) = -\beta_j(z) D_j^{-1}(z) r_j r_j^* D_j^{-1}(z), \tag{4.4}$$

we have the following well-known martingale decomposition:

$$\begin{aligned}
p[s_n(z) - \mathbb{E}s_n(z)] &= \operatorname{tr}\bigl(D^{-1}(z) - \mathbb{E}D^{-1}(z)\bigr) = \sum_{j=1}^n \operatorname{tr}\bigl(\mathbb{E}_j D^{-1}(z) - \mathbb{E}_{j-1} D^{-1}(z)\bigr) \\
&= \sum_{j=1}^n \operatorname{tr}(\mathbb{E}_j - \mathbb{E}_{j-1})\bigl(D^{-1}(z) - D_j^{-1}(z)\bigr) \\
&= -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_j(z) r_j^* D_j^{-2}(z) r_j \\
&= -\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1}) \frac{\mathrm{d}\log\beta_j(z)}{\mathrm{d}z}.
\end{aligned}$$

Integrating by parts, we obtain

$$\Delta_1 - \mathbb{E}\Delta_1 = \frac{1}{2\pi i}\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\oint_{\gamma_m} f_m'(z)\log\frac{\overline{\beta}_j(z)}{\beta_j(z)}\,dz$$
$$= \frac{1}{2\pi i}\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\oint_{\gamma_m} f_m'(z)\log\bigl(1+\varepsilon_j(z)\overline{\beta}_j(z)\bigr)\,dz.$$

Let $R_j(z) = \log(1+\varepsilon_j(z)\overline{\beta}_j(z)) - \varepsilon_j(z)\overline{\beta}_j(z)$ and write

$$\Delta_1 - \mathbb{E}\Delta_1 = \frac{1}{2\pi i}\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\oint_{\gamma_m} f_m'(z)\bigl(\varepsilon_j(z)\overline{\beta}_j(z) + R_j(z)\bigr)\,dz$$
$$= \frac{1}{2\pi i}\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\int_{\gamma_{mh}} f_m'(z)[\varepsilon_j(z)\overline{\beta}_j(z) + R_j(z)]\,dz \tag{4.5}$$
$$+ \frac{1}{2\pi i}\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1})\int_{\gamma_{mv}} f_m'(z)[\varepsilon_j(z)\overline{\beta}_j(z) + R_j(z)]\,dz, \tag{4.6}$$

where here, and in the sequel, γ_{mh} denotes the union of the two horizontal parts of γ_m, and γ_{mv} the union of the two vertical parts.

We first prove (4.6) $\to 0$ in probability. Let $A_n = \{a - \epsilon_1 \le \lambda^{B_n} \le b + \epsilon_1\}$ for any $0 < \epsilon_1 < a - a_1$ and $A_{nj} = \{a - \epsilon_1 \le \lambda^{B_{nj}} \le b + \epsilon_1\}$, where $B_{nj} = B_n - r_j r_j^*$ and λ^B denotes all eigenvalues of matrix B. By the interlacing theorem (see [15], page 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 [7], when $y \in (0,1)$, for any $l \ge 0$,

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

We have $P(A_n^c) = o(n^{-l})$ for any $l \ge 0$.

By continuity of $s(z)$, for large n, there exist positive constants M_l and M_u such that for all $z \in \gamma_{mv}$, $M_l \le |y_n s(z)| \le M_u$. Letting $C_{nj} = \{|\overline{\beta}_j(z)|^{-1}\mathbb{I}_{A_{nj}} > \epsilon_2\}$, where $0 < \epsilon_2 < M_l/2$ and $C_n = \bigcap_{j=1}^n C_{nj}$, we have

$$P(C_n^c) = P\left(\bigcup_{j=1}^{n} C_{nj}^c\right) \le \sum_{j=1}^{n} P(C_{nj}^c) = \sum_{j=1}^{n} P\{|\overline{\beta}_j(z)|^{-1}\mathbb{I}_{A_{nj}} \le \epsilon_2\}$$
$$\le \sum_{j=1}^{n} P\left\{\left|\frac{1}{n}\operatorname{tr} D_j^{-1}(z) - y_n s(z)\right|\mathbb{I}_{A_{nj}} \ge \epsilon_2\right\} + \sum_{j=1}^{n} P(A_{nj}^c)$$

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

$$\leq \frac{1}{\epsilon_2^4}\sum_{j=1}^n \mathrm{O}(n^{-2/5})^4 + nP(A_n^c) \leq \mathrm{O}(n^{-2/5}),$$

where we have used Lemma A.1. Defining $Q_{nj} = A_{nj} \cap C_{nj}$ and $Q_n = \bigcap_{j=1}^n Q_{nj}$, it is easy to show that Q_{nj} is independent of r_j and $P(Q_n^c) \leq P(A_n^c) + P(C_n^c) \to 0$ as $n \to \infty$. (4.6) now becomes

$$\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_{mv}} f_m'(z)[\varepsilon_j(z)\overline{\beta}_j(z) + R_j(z)]\mathbb{I}_{Q_{nj}}\,\mathrm{d}z + \mathrm{o}_p(1).$$

From the Burkholder inequality, Lemma A.3 and the inequalities $|n^{-1}\operatorname{tr} D_j(z)D_j(\bar{z})|\mathbb{I}_{A_{nj}} \leq 1/(a-\epsilon_1-a_1)^2$ and $|\bar{\beta}_j(z)|\mathbb{I}_{Q_{nj}} \leq 1/\epsilon_2$, we have

$$\mathbb{E}\left|\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_{mv}} f_m'(z)[\varepsilon_j(z)\overline{\beta}_j(z)]\mathbb{I}_{Q_{nj}}\,\mathrm{d}z\right|^2$$

$$\leq K\|\gamma_{mv}\|^2 \sum_{j=1}^n \sup_{z\in\gamma_{mv}} \mathbb{E}|\varepsilon_j(z)\overline{\beta}_j(z)|^2\mathbb{I}_{Q_{nj}}$$

$$\leq Kn^{-13/40}\sum_{j=1}^n \sup_{z\in\gamma_{mv}} \mathbb{E}|\varepsilon_j(z)|^2\mathbb{I}_{A_{nj}} \leq Kn^{-13/40}.$$

By Lemma A.3, for $z \in \gamma_{mv}$, we have

$$\sum_{j=1}^n P\big(|\varepsilon_j(z)\overline{\beta}_j(z)|\mathbb{I}_{Q_{nj}} \geq 1/2\big) \leq K\sum_{j=1}^n \mathbb{E}|\varepsilon_j(z)\overline{\beta}_j(z)|^4\mathbb{I}_{Q_{nj}} \leq K/n.$$

From the inequality $|\log(1+x) - x| \leq Kx^2$ for $|x| < 1/2$, we get

$$\mathbb{E}\left|\sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_{mv}} f_m'(z)R_j(z)\mathbb{I}_{Q_{nj}\cap\{|\varepsilon_j(z)\overline{\beta}_j(z)|<1/2\}}\,\mathrm{d}z\right|^2$$

$$\leq K\|\gamma_{mv}\|^2 \sum_{j=1}^n \sup_{z\in\gamma_{mv}} \mathbb{E}|R_j(z)|^2\mathbb{I}_{Q_{nj}\cap\{|\varepsilon_j(z)\overline{\beta}_j(z)|<1/2\}} \tag{4.7}$$

$$\leq Kn^{-13/40}\sum_{j=1}^n \sup_{z\in\gamma_{mv}} \mathbb{E}|\varepsilon_j(z)|^4\mathbb{I}_{A_{nj}} \leq Kn^{-53/40}.$$

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 the following estimates:

$$\sum_{j=1}^{n} P\big(|\varepsilon_j(z)\overline{\beta}_j(z)| \geq 1/2\big) \leq K \sum_{j=1}^{n} \mathbb{E}|\varepsilon_j(z)\overline{\beta}_j(z)|^4$$

and

$$\begin{aligned}
&\mathbb{E}\left|\sum_{j=1}^{n}(\mathbb{E}_j - \mathbb{E}_{j-1}) \int_{\gamma_{mh}} f_m'(z) R_j(z) \mathbb{I}_{\{|\varepsilon_j(z)\overline{\beta}_j(z)|<1/2\}} \,\mathrm{d}z\right|^2 \\
&\quad \leq K \|\gamma_{mh}\|^2 \sum_{j=1}^{n} \sup_{z\in\gamma_{mh}} \mathbb{E}|R_j(z)|^2 \mathbb{I}_{\{|\varepsilon_j(z)\overline{\beta}_j(z)|<1/2\}} \qquad (4.8)\\
&\quad \leq K \sum_{j=1}^{n} \sup_{z\in\gamma_{mh}} \mathbb{E}|\varepsilon_j(z)\overline{\beta}_j(z)|^4.
\end{aligned}$$

Thus, we get

$$\begin{aligned}
(4.5) &= -\frac{1}{2\pi\mathrm{i}} \sum_{j=1}^{n} \mathbb{E}_j \int_{\gamma_{mh}} f_m'(z)[\varepsilon_j(z)\overline{\beta}_j(z)] \,\mathrm{d}z + \mathrm{o}_p(1) \\
&\triangleq -\frac{1}{2\pi\mathrm{i}} \sum_{j=1}^{n} Y_{nj} + \mathrm{o}_p(1),
\end{aligned}$$

where $\mathrm{o}_p(1)$ 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 $Y_{nj} \in \mathcal{F}_j$ and $\mathbb{E}_{j-1}Y_{nj} = 0$, $\{Y_{nj}, j = 1, \ldots, n\}$ is a martingale difference sequence and thus $\sum_{j=1}^{n} Y_{nj}$ is a sum of a martingale difference sequence. In order to apply a martingale CLT ([9], Theorem 35.12) 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 $f, g \in C^4(\mathcal{U})$ and f_m, g_m are their corresponding Bernstein polynomial approximations, respectively.

Proof of Condition 4.1. By Lemmas A.5 and A.6, for any $z \in \gamma_{mh}$,

$$\mathbb{E}|\varepsilon_j(z)|^6 \le \frac{K}{n^6}[(\mathbb{E}|x_{11}|^4 \operatorname{tr} D_j^{-1}(z) D_j^{-1*}(z))^3 + \mathbb{E}|x_{11}|^{12} \operatorname{tr}(D_j^{-1}(z) D_j^{-1*}(z))^3]$$
$$\le \frac{K}{n^6 v^6}[n^3 + \delta_n^4 n^3] \le \frac{K}{n^3 v^6}.$$

Hence, we get

$$\begin{aligned}\sum_{j=1}^n \mathbb{E}|Y_{nj}|^4 &\le K \sum_{j=1}^n \int_{\gamma_{mh}} \mathbb{E}|\varepsilon_j(z)\overline{\beta}_j(z)|^4 \,\mathrm{d}z \\ &\le K \sum_{j=1}^n \int_{\gamma_{mh}} (\mathbb{E}|\overline{\beta}_j(z)|^{12})^{1/3} (\mathbb{E}|\varepsilon_j(z)|^6)^{2/3} \,\mathrm{d}z \\ &\le \frac{K}{nv^4} \to 0.\end{aligned}$$

□

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

$$\begin{aligned}&-\frac{1}{4\pi^2}\sum_{j=1}^n \mathbb{E}_{j-1}[Y_{nj}(f_m)\cdot Y_{nj}(g_m)] \\ &= \frac{-1}{4\pi^2}\sum_{j=1}^n \mathbb{E}_{j-1}\left[\int_{\gamma_{mh}} f_m'(z)\mathbb{E}_j(\varepsilon_j(z)\overline{\beta}_j(z))\,\mathrm{d}z \cdot \int_{\gamma_{mh}'} g_m'(z)\mathbb{E}_j(\varepsilon_j(z)\overline{\beta}_j(z))\,\mathrm{d}z\right] \\ &= \frac{-1}{4\pi^2}\iint_{\gamma_{mh}\times\gamma_{mh}'} f_m'(z_1)g_m'(z_2)\sum_{j=1}^n \mathbb{E}_{j-1}[\mathbb{E}_j(\varepsilon_j(z_1)\overline{\beta}_j(z_1))\mathbb{E}_j(\varepsilon_j(z_2)\overline{\beta}_j(z_2))]\,\mathrm{d}z_1\,\mathrm{d}z_2 \\ &\triangleq \frac{-1}{4\pi^2}\iint_{\gamma_{mh}\times\gamma_{mh}'} f_m'(z_1)g_m'(z_2)\Gamma_n(z_1,z_2)\,\mathrm{d}z_1\,\mathrm{d}z_2,\end{aligned}$$

where $\Gamma_n(z_1,z_2) = \sum_{j=1}^n \mathbb{E}_{j-1}[\mathbb{E}_j(\varepsilon_j(z_1)\overline{\beta}_j(z_1))\mathbb{E}_j(\varepsilon_j(z_2)\overline{\beta}_j(z_2))]$ and γ_m' is the contour formed by the rectangle with vertices $a_l' \pm \mathrm{i}/2\sqrt{m}$ and $b_r' \pm \mathrm{i}/2\sqrt{m}$. Here, $0 < a_l < a_l' < a < b < b_r' < b_r$, which means that the contour γ_m encloses the contour γ_m'. γ_{mh}' is the union of the horizontal parts of γ_m'.

First, we show that

$$\Gamma_n(z_1,z_2) - \Gamma(z_1,z_2) \xrightarrow{\text{Pr.}} 0 \quad \text{uniformly on } \gamma_{mh}\times\gamma_{mh}',$$

where

$$\Gamma(z_1,z_2) = \kappa_2 y k(z_1)k(z_2) - (\kappa_1+1)\ln\frac{\underline{s}(z_1)\underline{s}(z_2)(z_1-z_2)}{\underline{s}(z_1)-\underline{s}(z_2)}.$$

From Lemma A.6, for all $z \in \gamma_{mh} \cup \gamma'_{mh}$ and any $l \geq 2$,

$$\begin{aligned}\mathbb{E}|\overline{\beta}_j(z) - b_n(z)|^l &= \mathbb{E}\big|\overline{\beta}_j(z)b_n(z)n^{-1}\big(\operatorname{tr} D_j(z) - \mathbb{E}\operatorname{tr} D_j(z)\big)\big|^l \\ &\leq M\big(\mathbb{E}\big|n^{-1}\big(\operatorname{tr} D_j(z) - \mathbb{E}\operatorname{tr} D_j(z)\big)\big|^{2l}\big)^{1/2} \leq K\big(\sqrt{nv}\big)^l.\end{aligned} \tag{4.9}$$

This leads to

$$\mathbb{E}\left|\Gamma_n(z_1, z_2) - b_n(z_1)b_n(z_2)\sum_{j=1}^{n}\mathbb{E}_{j-1}(\mathbb{E}_j\varepsilon_j(z_1)\mathbb{E}_j\varepsilon_j(z_2))\right| \leq \frac{K}{\sqrt{n}v^3} = \mathrm{O}(n^{-1/80}).$$

Thus, we need to consider

$$b_n(z_1)b_n(z_2)\sum_{j=1}^{n}\mathbb{E}_{j-1}(\mathbb{E}_j\varepsilon_j(z_1)\mathbb{E}_j\varepsilon_j(z_2)). \tag{4.10}$$

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

$$\begin{aligned}&\mathbb{E}(x_1^* A x_1 - n\operatorname{tr} A)(x_1^* B x_1 - n \operatorname{tr} B) \\ &= (\mathbb{E}|x_{11}|^4 - |\mathbb{E}x_{11}^2|^2 - 2)\sum_{i=1}^{p} a_{ii}b_{ii} + |\mathbb{E}x_{11}^2|^2\sum_{i,j} a_{ij}b_{ij} + \sum_{i,j} a_{ij}b_{ji} \\ &= \kappa_2 \sum_{i=1}^{p} a_{ii}b_{ii} + \kappa_1 \operatorname{tr} AB^{\mathrm{T}} + \operatorname{tr} AB,\end{aligned} \tag{4.11}$$

from which (4.10) becomes

$$\begin{aligned}&(\kappa_1 + 1)b_n(z_1)b_n(z_2)\frac{1}{n^2}\sum_{j=1}^{n}\operatorname{tr}\mathbb{E}_j D_j^{-1}(z_1)\mathbb{E}_j D_j^{-1}(z_2) \\ &\quad + \kappa_2 b_n(z_1)b_n(z_2)\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{p}\mathbb{E}_j[D_j^{-1}(z_1)]_{ii}\mathbb{E}_j[D_j^{-1}(z_2)]_{ii} \\ &\triangleq \Gamma_{n1}(z_1, z_2) + \Gamma_{n2}(z_1, z_2).\end{aligned}$$

For $\Gamma_{n2}(z_1, z_2)$, by Lemmas A.6, A.7 and $-zs(z)(\underline{s}(z) + 1) = 1$, we get

$$\Gamma_{n2}(z_1, z_2) = \kappa_2 y_n k(z_1)k(z_2) + \mathrm{o}_p(1),$$

where $\mathrm{o}_p(1)$ denotes uniform convergence in probability on $\gamma_{mh} \times \gamma'_{mh}$.

It is easy to check that $k(\bar{z})=\overline{k(z)}$ since $s(\bar{z})=\overline{\underline{s}(z)}$. As $n\to\infty$, $a_{\mathrm{l}}\to a$ and $b_{\mathrm{r}}\to b$, we then get

$$\begin{aligned}
&-\frac{1}{4\pi^2}\iint_{\gamma_{mh}\times\gamma'_{mh}} f'_m(z_1)g'_m(z_2)\Gamma_{n2}(z_1,z_2)\,\mathrm{d}z_1\,\mathrm{d}z_2\\
&\quad=-\frac{\kappa_2 y_n}{4\pi^2}\iint_{\gamma_{mh}\times\gamma'_{mh}} f'_m(z_1)g'_m(z_2)k(z_1)k(z_2)\,\mathrm{d}z_1\,\mathrm{d}z_2+\mathrm{o}_p(1)\\
&\quad\to-\frac{\kappa_2 y}{2\pi^2}\int_a^b\int_a^b f'(x_1)g'(x_2)\Re[k(x_1)k(x_2)-\overline{k(x_1)}k(x_2)]\,\mathrm{d}x_1\,\mathrm{d}x_2,
\end{aligned}$$

which is (1.5) in Theorem 1.1.

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

$$b_n(z_1)b_n(z_2)\frac{1}{n^2}\sum_{j=1}^n \operatorname{tr}\mathbb{E}_j D_j^{-1}(z_1)\mathbb{E}_j D_j^{-1}(z_2). \tag{4.12}$$

Let $D_{ij}(z)=D(z)-r_jr_j^*-r_ir_i^*$, $\beta_{ij}(z)=(1+r_i^*D_{ij}^{-1}(z)r_i)^{-1}$, $b_{12}(z)=(1+\frac{1}{n}\mathbb{E}\operatorname{tr}D_{12}^{-1}(z))^{-1}$ and $t(z)=(z-\frac{n-1}{n}b_{12}(z))^{-1}$. Write

$$D_j(z)+zI_p-\frac{n-1}{n}b_{12}(z)I_p=\sum_{i\neq j}^n r_ir_i^*-\frac{n-1}{n}b_{12}(z)I_p.$$

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

$$r_i^*D_j^{-1}(z)=\beta_{ij}(z)r_i^*D_{ij}^{-1}(z), \tag{4.13}$$

we obtain

$$\begin{aligned}
D_j^{-1}(z)&=-t(z)I_p+\sum_{i\neq j}^n t(z)\beta_{ij}(z)r_ir_i^*D_{ij}^{-1}(z)-\frac{n-1}{n}b_{12}(z)t(z)D_j^{-1}(z)\\
&=-t(z)I_p+b_{12}(z)A(z)+B(z)+C(z),
\end{aligned}\tag{4.14}$$

where

$$A(z)=\sum_{i\neq j}^n t(z)(r_ir_i^*-n^{-1}I_p)D_{ij}^{-1}(z),\qquad B(z)=\sum_{i\neq j}^n t(z)\bigl(\beta_{ij}(z)-b_{12}(z)\bigr)r_ir_i^*D_{ij}^{-1}(z)$$

and

$$C(z)=\frac{1}{n}t(z)b_{12}(z)\sum_{i\neq j}^n\bigl(D_{ij}^{-1}(z)-D_j^{-1}(z)\bigr).$$

It is easy to verify that for all $z \in \gamma_{mh} \cup \gamma'_{mh}$,

$$|t(z)| = \left| z + \frac{n-1}{n} \frac{1}{1 + n^{-1}\mathbb{E}\,\mathrm{tr}\, D_{12}^{-1}(z)} \right|^{-1} = \left| \frac{1 + n^{-1}\mathbb{E}\,\mathrm{tr}\, D_{12}^{-1}(z)}{z(1 + n^{-1}\mathbb{E}\,\mathrm{tr}\, D_{12}^{-1}(z)) + (n-1)/n} \right|$$
$$\le \frac{1}{|z|}\left[1 + \frac{1}{\Im z(1 + n^{-1}\mathbb{E}\,\mathrm{tr}\, D_{12}^{-1}(z))}\right] \le \frac{K}{v}$$

since $a_{\mathrm{l}} \le |z| \le b_{\mathrm{r}} + 1$. Thus, by Lemmas A.6, A.4 and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\mathbb{E}|\,\mathrm{tr}(B(z_1)\mathbb{E}_j D_j^{-1}(z_2))| &= \mathbb{E}\left|\sum_{i\ne j}^{n} t(z_1)\big(\beta_{ij}(z_1) - b_{12}(z_1)\big) r_i^* D_{ij}^{-1}(z_1)\mathbb{E}_j D_j^{-1}(z_2) r_i\right| \\ &\le \frac{Kn}{v}\mathbb{E}\big|\big(\beta_{ij}(z_1) - b_{12}(z_1)\big) r_i^* D_{ij}^{-1}(z_1)\mathbb{E}_j D_j^{-1}(z_2) r_i\big| \\ &\le \frac{Kn}{v}\frac{1}{\sqrt{n}v}\frac{1}{\sqrt{n}v^2} = \frac{K}{v^4}.\end{aligned} \tag{4.15}$$

From Lemma 2.10 of Bai and Silverstein [6], for any $n \times n$ matrix A,

$$\big|\mathrm{tr}\big(D^{-1}(z) - D_j^{-1}(z)\big)A\big| \le \frac{\|A\|}{v}, \tag{4.16}$$

which, combined with Lemma A.6, gives

$$\begin{aligned}&\mathbb{E}|\,\mathrm{tr}(C(z_1)\mathbb{E}_j D_j^{-1}(z_2))| \\ &\quad = \mathbb{E}\left|\frac{1}{n} t(z_1) b_{12}(z_1) \sum_{i\ne j}^{n} \mathrm{tr}\big(\big(D_{ij}^{-1}(z_1) - D_j^{-1}(z_1)\big)\mathbb{E}_j D_j^{-1}(z_2)\big)\right| \\ &\quad \le \frac{K}{v}(\mathbb{E}|b_{12}(z_1)|^2)^{1/2}\big(\mathbb{E}\big|\mathrm{tr}\big(D_{ij}^{-1}(z_1) - D_j^{-1}(z_1)\big)\mathbb{E}_j D_j^{-1}(z_2)\big|^2\big)^{1/2} \\ &\quad \le \frac{K}{v}\frac{1}{v^2} = \frac{K}{v^3}.\end{aligned} \tag{4.17}$$

From the above estimates (4.15) and (4.17), we arrive at

$$\begin{aligned}&\mathrm{tr}\,\mathbb{E}_j D_j^{-1}(z_1)\mathbb{E}_j D_j^{-1}(z_2) \\ &\quad = -t(z_1)\,\mathrm{tr}\,\mathbb{E}_j D_j^{-1}(z_2) + b_{12}(z_1)\,\mathrm{tr}\,\mathbb{E}_j A(z_1)\mathbb{E}_j D_j^{-1}(z_2) + \frac{K}{v^4}.\end{aligned} \tag{4.18}$$

Using the identity

$$D_j^{-1}(z_2) - D_{ij}^{-1}(z_2) = -\beta_{ij}(z_2) D_{ij}^{-1}(z_2) r_i r_i^* D_{ij}^{-1}(z_2),$$

we can write

$$\operatorname{tr} \mathbb{E}_j(A(z_1)) D_j^{-1}(z_2) = A_1(z_1, z_2) + A_2(z_1, z_2) + A_3(z_1, z_2), \tag{4.19}$$

where

$$\begin{aligned} A_1(z_1, z_2) &= -\operatorname{tr} \sum_{i<j} t(z_1) r_i r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1)) \big(D_j^{-1}(z_2) - D_{ij}^{-1}(z_2)\big) \\ &= -\sum_{i<j} t(z_1) \beta_{ij}(z_2) r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1)) D_{ij}^{-1}(z_2) r_i r_i^* D_{ij}^{-1}(z_2) r_i, \\ A_2(z_1, z_2) &= -\operatorname{tr} \sum_{i<j}^{n} t(z_1) \frac{1}{n} \mathbb{E}_j(D_{ij}^{-1}(z_1)) \big(D_j^{-1}(z_2) - D_{ij}^{-1}(z_2)\big) \end{aligned}$$

and

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

From (4.16), we get

$$\begin{aligned} |A_2(z_1, z_2)| &= \left| \frac{1}{n} \sum_{i<j} t(z_1) \operatorname{tr}\big(D_j^{-1}(z_2) - D_{ij}^{-1}(z_2)\big) \mathbb{E}_j D_{ij}^{-1}(z_1) \right| \\ &\le \frac{j-1}{n} \frac{1}{v} \frac{K}{v^2} \le \frac{K}{v^3} \end{aligned} \tag{4.20}$$

and by Lemma A.3, we have

$$\begin{aligned} \mathbb{E}|A_3(z_1, z_2)| &\le \frac{K(j-1)}{v} \mathbb{E}\left| \operatorname{tr}\left(r_i r_i^* - \frac{1}{n} I_p \right) \mathbb{E}_j(D_{ij}^{-1}(z_1)) D_{ij}^{-1}(z_2) \right| \\ &\le \frac{Kn}{v} \frac{1}{\sqrt{n} v^2} = \frac{K\sqrt{n}}{v^3}. \end{aligned} \tag{4.21}$$

For $A_1(z_1, z_2)$, by Lemmas A.4 and A.5,

$$\begin{aligned} &\mathbb{E}\bigg| r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1)) D_{ij}^{-1}(z_2) r_i r_i^* D_{ij}^{-1}(z_2) r_i \\ &\qquad - \frac{1}{n^2} \operatorname{tr}[\mathbb{E}_j(D_{ij}^{-1}(z_1)) D_{ij}^{-1}(z_2)] \operatorname{tr} D_{ij}^{-1}(z_2) \bigg| \\ &\quad \le \mathbb{E}\left| \left[r_i^* \mathbb{E}_j(D_{ij}^{-1}(z_1)) D_{ij}^{-1}(z_2) r_i - \frac{1}{n} \operatorname{tr}(\mathbb{E}_j(D_{ij}^{-1}(z_1)) D_{ij}^{-1}(z_2)) \right] r_i^* D_{ij}^{-1}(z_2) r_i \right| \\ &\qquad + \mathbb{E}\left| \frac{1}{n} \operatorname{tr}(\mathbb{E}_j(D_{ij}^{-1}(z_1)) D_{ij}^{-1}(z_2)) \left[r_i^* D_{ij}^{-1}(z_2) r_i - \frac{1}{n} \operatorname{tr} D_{ij}^{-1}(z_2) \right] \right| \le \frac{K}{\sqrt{n} v^3}. \end{aligned}$$

Let $\varphi_j(z_1, z_2) = \operatorname{tr}(\mathbb{E}_j(D_j^{-1}(z_1))D_j^{-1}(z_2))$. Using the identity (4.16), we have

$$|\operatorname{tr}(\mathbb{E}_j(D_{ij}^{-1}(z_1))D_{ij}^{-1}(z_2))\operatorname{tr} D_{ij}^{-1}(z_2) - \varphi_j(z_1, z_2)\operatorname{tr} D_j^{-1}(z_2)| \leq Knv^{-3}.$$

Thus, in conjunction with Lemma A.6, we can get

$$\mathbb{E}\left|A_1(z_1, z_2) + \frac{j-1}{n^2}t(z_1)b_{12}(z_2)\varphi_j(z_1, z_2)\operatorname{tr} D_j^{-1}(z_2)\right| \leq \frac{K}{\sqrt{n}v^3}. \tag{4.22}$$

Therefore, from (4.14)–(4.22), it follows that

$$\begin{aligned}&\varphi_j(z_1, z_2)\left[1 + \frac{j-1}{n^2}t(z_1)b_{12}(z_1)b_{12}(z_2)\operatorname{tr} D_j^{-1}(z_2)\right]\\&\quad = -\operatorname{tr}(t(z_1)\operatorname{tr} D_j^{-1}(z_2)) + A_4(z_1, z_2),\end{aligned}$$

where $\mathbb{E}|A_4(z_1, z_2)| \leq K\sqrt{n}/v^3$.

Using Lemma A.6, the expression for $D_j^{-1}(z_2)$ in (4.14) and the estimate

$$\begin{aligned}\mathbb{E}|\operatorname{tr} A(z)| &= \mathbb{E}\left|\operatorname{tr}\sum_{i\neq j}^{n} t(z)(r_i r_i^* - n^{-1}I_p)D_{ij}^{-1}(z)\right|\\&\leq \frac{Kn}{v}\mathbb{E}|r_i D_{ij}^{-1}(z)r_i^* - n^{-1}\operatorname{tr} D_{ij}^{-1}(z)| \leq \frac{K\sqrt{n}}{v^2},\end{aligned}$$

we find that

$$\begin{aligned}&\varphi_j(z_1, z_2)\left[1 - \frac{(j-1)p}{n^2}t(z_1)b_{12}(z_1)t(z_2)b_{12}(z_2)\right]\\&\quad = -pt(z_1)t(z_2) + A_5(z_1, z_2),\end{aligned}$$

where

$$\mathbb{E}|A_5(z_1, z_2)| \leq \frac{K\sqrt{n}}{v^3}.$$

By Lemma A.6, we can write

$$\begin{aligned}&\varphi_j(z_1, z_2)\left[1 - \frac{(j-1)p}{n^2}\frac{\underline{s}_n^0(z_1)\underline{s}_n^0(z_2)}{(\underline{s}_n^0(z_1)+1)(\underline{s}_n^0(z_2)+1)}\right]\\&\quad = \frac{p}{z_1 z_2}\frac{1}{(\underline{s}_n^0(z_1)+1)(\underline{s}_n^0(z_2)+1)} + A_6(z_1, z_2),\end{aligned}$$

where $\mathbb{E}|A_6(z_1, z_2)| \leq K\sqrt{n}/v^3$.

Let

$$a_n(z_1, z_2) = \frac{y_n \underline{s}_n^0(z_1)\underline{s}_n^0(z_2)}{(\underline{s}_n^0(z_1)+1)(\underline{s}_n^0(z_2)+1)}.$$

(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

$$\mathbb{E}|A_7(z_1,z_2)|\le \frac{K}{\sqrt{n}v^3}.$$

Since

$$a_n(z_1,z_2)\to a(z_1,z_2)=\frac{y\underline{s}(z_1)\underline{s}(z_2)}{(\underline{s}(z_1)+1)(\underline{s}(z_2)+1)}$$

as $n\to\infty$, we arrive at

$$(4.12)\xrightarrow{\text{Pr.}} a(z_1,z_2)\int_0^1\frac{1}{1-ta(z_1,z_2)}\,\mathrm{d}t=-\ln\bigl(1-a(z_1,z_2)\bigr)=-\ln\frac{l(z_1,z_2)}{\underline{s}(z_1)-\underline{s}(z_2)},$$

where $l(z_1,z_2)=\underline{s}(z_1)\underline{s}(z_2)(z_1-z_2)$, which implies that

$$\begin{aligned}\Gamma_{n1}(z_1,z_2)&=(\kappa_1+1)b_n(z_1)b_n(z_2)\frac{1}{n^2}\sum_{j=1}^{n}\operatorname{tr}\mathbb{E}_jD_j^{-1}(z_1)\mathbb{E}_jD_j^{-1}(z_2)\\&=-(\kappa_1+1)\ln(l(z_1,z_2))+(\kappa_1+1)\ln\bigl(\underline{s}(z_1)-\underline{s}(z_2)\bigr)+\mathrm{o}_p(1).\end{aligned}$$

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 $\Gamma_{n1}(z_1,z_2)$ is

$$\begin{aligned}&-\frac{1}{4\pi^2}\iint_{\gamma_{mh}\times\gamma_{mh}'}f_m'(z_1)g_m'(z_2)(\kappa_1+1)\ln(l(z_1,z_2))\,\mathrm{d}z_1\,\mathrm{d}z_2\\&=-\frac{\kappa_1+1}{4\pi^2}\oint\oint_{\gamma_m\times\gamma_m'}f_m'(z_1)g_m'(z_2)\ln(l(z_1,z_2))\,\mathrm{d}z_1\,\mathrm{d}z_2+\mathrm{O}(v)\\&=\mathrm{o}(1).\end{aligned}$$

For the second term of $\Gamma_{n1}(z_1,z_2)$, since $s(\bar z)=\overline{\underline{s}(z)}$, as $n\to\infty$, $a_{\mathrm{l}}\to a$ and $b_{\mathrm{r}}\to b$, we get

$$\begin{aligned}&-\frac{\kappa_1+1}{4\pi^2}\oint\oint_{\gamma_m\times\gamma_m'}f_m'(z_1)g_m'(z_2)\ln\bigl(\underline{s}(z_1)-\underline{s}(z_2)\bigr)\,\mathrm{d}z_1\,\mathrm{d}z_2+\mathrm{o}_p(1)\\&\longrightarrow\frac{\kappa_1+1}{2\pi^2}\int_a^b\int_a^b f'(x_1)g'(x_2)\ln\left|\frac{\overline{\underline{s}(x_1)}-\underline{s}(x_2)}{\underline{s}(x_1)-\underline{s}(x_2)}\right|\,\mathrm{d}x_1\,\mathrm{d}x_2,\end{aligned}$$

which is (1.4) in Theorem 1.1. □

5. Mean function

In this section, we will find the limit of

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

We shall first consider $M_n(z) = p[\mathbb{E}s_n(z) - s_n^0(z)] = n[\mathbb{E}\underline{s}_n(z) - \underline{s}_n^0(z)]$.

Since $D(z) + zI = \sum_{j=1}^n r_j r_j^*$, multiplying by $D^{-1}(z)$ on the right-hand side and using (4.13), we find that

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

Taking trace, dividing by n on both sides and combining with the identity $z\underline{s}_n(z) = -1 + y_n + y_n z s_n(z)$ leads to

$$\underline{s}_n(z) = -\frac{1}{nz} \sum_{j=1}^n \frac{1}{1 + r_j^* D_j^{-1}(z) r_j} = -\frac{1}{nz} \sum_{j=1}^n \beta_j(z). \tag{5.1}$$

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

$$\begin{aligned}\frac{I_p}{z(\mathbb{E}\underline{s}_n(z) + 1)} - D^{-1}(z) &= -\frac{1}{z(\mathbb{E}\underline{s}_n(z) + 1)} \left[\sum_{j=1}^n r_j r_j^* + z\mathbb{E}\underline{s}_n(z)\right] D^{-1}(z) \\ &= -\frac{1}{z(\mathbb{E}\underline{s}_n(z) + 1)} \sum_{j=1}^n \left[\beta_j(z) r_j r_j^* D_j^{-1}(z) - \mathbb{E}(\beta_j(z)) \frac{1}{n} D^{-1}(z)\right].\end{aligned}$$

Taking trace, dividing by p and taking expectation, we find that

$$\begin{aligned}\omega_n(z) &= -\frac{1}{z(\mathbb{E}\underline{s}_n(z) + 1)} - \mathbb{E}s_n(z) \\ &= -\frac{1}{pz(\mathbb{E}\underline{s}_n(z) + 1)} \sum_{j=1}^n \mathbb{E}(\beta_j(z) d_j(z)) \\ &\triangleq -\frac{1}{pz(\mathbb{E}\underline{s}_n(z) + 1)} J_n(z),\end{aligned} \tag{5.2}$$

where

$$d_j(z) = r_j^* D_j^{-1}(z) r_j - \frac{1}{n} \operatorname{tr} \mathbb{E} D^{-1}(z).$$

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

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

where

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

which implies that

$$\mathbb{E}\underline{s}_n(z) = \left(-z + \frac{y_n}{\mathbb{E}\underline{s}_n(z)+1} - R_n(z)\right)^{-1}. \tag{5.3}$$

For $s_n^0(z)$, since $s_n^0(z) = (1 - y_n - y_n z s_n^0(z) - z)^{-1}$ and $\underline{s}_n^0(z) = -(1-y_n)z^{-1} + y_n s_n^0(z)$, we have

$$\underline{s}_n^0(z) = \left(-z + \frac{y_n}{\underline{s}_n^0(z)+1}\right)^{-1}. \tag{5.4}$$

By (5.3) and (5.4), we get

$$\begin{aligned}
\mathbb{E}\underline{s}_n(z) - \underline{s}_n^0(z) &= \left(-z + \frac{y_n}{\mathbb{E}\underline{s}_n(z)+1} - R_n(z)\right)^{-1} - \left(-z + \frac{y_n}{\underline{s}_n^0(z)+1}\right)^{-1} \\
&= \mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)\left(\frac{y_n}{\underline{s}_n^0(z)+1} - \frac{y_n}{\mathbb{E}\underline{s}_n(z)+1} + R_n(z)\right) \\
&= \frac{y_n \mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)}{(\underline{s}_n^0(z)+1)(\mathbb{E}\underline{s}_n(z)+1)}\big(\mathbb{E}\underline{s}_n(z) - \underline{s}_n^0(z)\big) + \mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)R_n(z),
\end{aligned}$$

which, combined with (5.2), leads to

$$\begin{aligned}
&n\big(\mathbb{E}\underline{s}_n(z) - \underline{s}_n^0(z)\big)\left(1 - \frac{y_n \mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)}{(\underline{s}_n^0(z)+1)(\mathbb{E}\underline{s}_n(z)+1)}\right) \\
&\quad = n\mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)R_n(z) \\
&\quad = n\mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)\frac{y_n z}{\mathbb{E}\underline{s}_n(z)}\omega_n(z) \\
&\quad = -\frac{\underline{s}_n^0(z)}{\mathbb{E}\underline{s}_n(z)+1} J_n(z).
\end{aligned} \tag{5.5}$$

Thus, in order to find the limit of $M_n(z) = n[\mathbb{E}\underline{s}_n(z) - \underline{s}_n^0(z)]$, it suffices to find the limit of $J_n(z)$. Let $\bar{d}_j(z) = r_j^* D_j^{-1}(z) r_j - \frac{1}{n}\operatorname{tr} D^{-1}(z)$ and $\bar{J}_n(z) = \sum_{j=1}^n \mathbb{E}(\beta_j(z)\bar{d}_j(z))$. By (4.3), we have

$$\begin{aligned}J_n(z) &= \bar{J}_n(z) + \sum_{j=1}^{n} \mathbb{E}\left[\beta_j(z)\left(\frac{1}{n}\operatorname{tr} D^{-1}(z) - \frac{1}{n}\operatorname{tr}\mathbb{E}D^{-1}(z)\right)\right]\\ &= \bar{J}_n(z) + \sum_{j=1}^{n} \mathbb{E}\left[\big(\beta_j(z) - b_n(z)\big)\left(\frac{1}{n}\operatorname{tr} D^{-1}(z) - \frac{1}{n}\operatorname{tr}\mathbb{E}D^{-1}(z)\right)\right]\\ &= \bar{J}_n(z) - T_1(z) + T_2(z),\end{aligned}$$

where, from (4.16),

$$\begin{aligned}T_1(z) &= \sum_{j=1}^{n} \mathbb{E}\left[b_n^2(z)\delta_j(z)\left(\frac{1}{n}\operatorname{tr} D^{-1}(z) - \frac{1}{n}\operatorname{tr}\mathbb{E}D^{-1}(z)\right)\right]\\ &= \sum_{j=1}^{n} \mathbb{E}\left[b_n^2(z)\delta_j(z)\frac{1}{n}\Big(\operatorname{tr}\big(D^{-1}(z) - D_j^{-1}(z)\big) - \operatorname{tr}\mathbb{E}\big(D^{-1}(z) - D_j^{-1}(z)\big)\Big)\right]\\ &\le \sum_{j=1}^{n} \frac{K}{\sqrt{n}v}\cdot\frac{K}{nv} = \frac{K}{\sqrt{n}v^2}.\end{aligned}$$

It follows from Bai and Silverstein [6], (4.3) that for $l \ge 2$,

$$\mathbb{E}\left|\frac{1}{n}\operatorname{tr} D^{-1}(z) - \frac{1}{n}\operatorname{tr}\mathbb{E}D^{-1}(z)\right|^l \le \frac{K_l}{(\sqrt{n}v)^l}. \tag{5.6}$$

Hence,

$$T_2(z) = \sum_{j=1}^{n} \mathbb{E}\left[\beta_j(z)b_n^2(z)\delta_j^2(z)\left(\frac{1}{n}\operatorname{tr} D^{-1}(z) - \frac{1}{n}\operatorname{tr}\mathbb{E}D^{-1}(z)\right)\right] \le \frac{K}{\sqrt{n}v^3}.$$

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

$$J_n(z) = \bar{J}_n(z) + \bar{\epsilon}_n,$$

where here, and in the sequel, $\bar{\epsilon}_n = \mathrm{O}((\sqrt{n}v^3)^{-1})$.

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

$$\begin{aligned}\bar{J}_n(z) &= \sum_{j=1}^{n} \mathbb{E}\big[\big(\beta_j(z) - \bar{\beta}_j(z)\big)\varepsilon_j(z)\big] + \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\big[\beta_j(z)\operatorname{tr}\big(D_j^{-1}(z) - D^{-1}(z)\big)\big]\\ &= -\sum_{j=1}^{n} \mathbb{E}(\bar{\beta}_j^2(z)\varepsilon_j^2(z)) + \sum_{j=1}^{n} \mathbb{E}(\bar{\beta}_j^2(z)\beta_j(z)\varepsilon_j^3(z)) + \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}(\beta_j^2(z)r_j^* D_j^{-2}(z)r_j)\\ &\triangleq \bar{J}_{n1}(z) + \bar{J}_{n2}(z) + \bar{J}_{n3}(z).\end{aligned}$$

From Lemmas A.3 and A.6, we find that

$$|\bar{J}_{n2}(z)| \le K \sum_{j=1}^{n} (\mathbb{E}|\varepsilon_j^6(z)|)^{1/2} \le \frac{K}{\sqrt{n}v^3}.$$

By Lemma A.6, $\bar{\beta}_j(z)$, $\beta_j(z)$ and $b_n(z)$ can be replaced by $-z\underline{s}(z)$, and so we get

$$\bar{J}_{n3}(z) = z^2\underline{s}^2(z)\frac{1}{n^2}\sum_{j=1}^{n} \mathbb{E}\,\mathrm{tr}\, D_j^{-2}(z) + \bar{\epsilon}_n \triangleq z^2\underline{s}^2(z)\psi_n(z) + \bar{\epsilon}_n.$$

By the identity of quadric form (4.11) and the fact, from Lemma A.7, that $\mathbb{E}[D_j^{-1}(z)]_{ii}$ can be replaced by $s(z) = -z^{-1}(\underline{s}(z)+1)^{-1}$, we have

$$\begin{aligned}
\bar{J}_{n1}(z) &= -z^2\underline{s}^2(z)\sum_{j=1}^{n}\mathbb{E}\varepsilon_j^2(z) + \bar{\epsilon}_n \\
&= -\frac{z^2\underline{s}^2(z)}{n^2}\sum_{j=1}^{n}\mathbb{E}\left[\sum_{i=1}^{p}\kappa_2[D_j^{-1}(z)]_{ii}^2 + \kappa_1\,\mathrm{tr}\, D_j^{-2}(z) + \mathrm{tr}\, D_j^{-2}(z)\right] + \bar{\epsilon}_n \qquad (5.7)\\
&= y_n\kappa_2 k^2(z) - z^2\underline{s}^2(z)(\kappa_1+1)\psi_n(z) + \bar{\epsilon}_n,
\end{aligned}$$

where κ_1, κ_2 and $k(z)$ were defined in Theorem 1.1. Our goal is now to find the limit of $\psi_n(z)$. Using the expansion of $D_j^{-1}(z)$ in (4.14), we get

$$\begin{aligned}
\psi_n(z) &= \frac{1}{n^2}\sum_{j=1}^{n}\frac{p}{(z+z\underline{s}(z))^2} + z^2\underline{s}^2(z)\frac{1}{n^2}\sum_{j=1}^{n}\mathbb{E} n\,\mathrm{tr}\, A^2(z) + \bar{\epsilon}_n \\
&= \frac{k^2(z)}{n^2}\sum_{j=1}^{n}\sum_{i,l\ne j}^{n}\mathbb{E}\,\mathrm{tr}\left[\left(r_ir_i^* - \frac{1}{n}I\right)D_{ij}^{-1}(z)D_{lj}^{-1}(z)\left(r_lr_l^* - \frac{1}{n}I\right)\right] \\
&\quad + \frac{1}{n^2}\sum_{j=1}^{n}\frac{p}{z^2(\underline{s}(z)+1)^2} + \bar{\epsilon}_n.
\end{aligned}$$

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

$$D_{ij}^{-1}(z) - D_{lij}^{-1}(z) = -\frac{D_{lij}^{-1}(z)r_lr_l^*D_{lij}^{-1}(z)}{1 + r_l^*D_{lij}^{-1}(z)r_l}.$$

Therefore, by (4.16), we conclude that the sum of cross terms is negligible and bounded by $K/(\sqrt{n}v^3)$. Thus, we find that

$$\begin{aligned}
&\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i,l\neq j}^{n}\mathbb{E}\,\mathrm{tr}\Big[\Big(r_ir_i^*-\frac{1}{n}I\Big)D_{ij}^{-1}(z)D_{lj}^{-1}(z)\Big(r_lr_l^*-\frac{1}{n}I\Big)\Big]\\
&\quad=\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i\neq j}^{n}\mathbb{E}\,\mathrm{tr}\Big[\Big(r_ir_i^*-\frac{1}{n}I\Big)D_{ij}^{-2}(z)\Big(r_ir_i^*-\frac{1}{n}I\Big)\Big]+\bar{\epsilon}_n\\
&\quad=\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i\neq j}^{n}\mathbb{E}\big[\big(r_i^*D_{ij}^{-2}(z)r_i\big)\big(r_i^*r_i\big)\big]+\bar{\epsilon}_n\\
&\quad=\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i\neq j}^{n}\frac{1}{n^2}\mathbb{E}\big[\mathrm{tr}\,D_{ij}^{-2}(z)\big(p+\mathrm{O}(1)\big)\big]+\bar{\epsilon}_n=y_n\psi_n(z)+\bar{\epsilon}_n.
\end{aligned}$$

From above, we get that

$$\psi_n(z)=\frac{y_n}{z^2(\underline{s}(z)+1)^2}+y_nk^2(z)\psi_n(z)+\bar{\epsilon}_n.$$

Combined with (5.7), we have

$$J_n(z)=\kappa_2y_nk^2(z)-\frac{\kappa_1y_nk^2(z)}{1-y_nk^2(z)}+\bar{\epsilon}_n.$$

Thus, from (5.5), it follows that

$$\begin{aligned}
M_n(z)&=n\mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)R_n(z)\Big/\left(1-\frac{y_n\mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)}{(\underline{s}_n^0(z)+1)(\mathbb{E}\underline{s}_n(z)+1)}\right)\\
&=-\frac{\underline{s}_n^0(z)}{\mathbb{E}\underline{s}_n(z)+1}J_n(z)\Big/\left(1-\frac{y_n\mathbb{E}\underline{s}_n(z)\underline{s}_n^0(z)}{(\underline{s}_n^0(z)+1)(\mathbb{E}\underline{s}_n(z)+1)}\right)\\
&=\frac{\kappa_1y_nk^3(z)}{(1-y_nk^2(z))^2}-\frac{\kappa_2y_nk^3(z)}{1-y_nk^2(z)}+\bar{\epsilon}_n\\
&\triangleq\widetilde{M}_1(z)+\widetilde{M}_2(z)+\bar{\epsilon}_n.
\end{aligned}$$

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

$$\begin{aligned}
&-\frac{1}{2\pi\mathrm{i}}\int_{\gamma_{mh}}f_m(z)\widetilde{M}_1(z)\,\mathrm{d}z\\
&\quad=-\frac{\kappa_1}{2\pi\mathrm{i}}\int_{\gamma_{mh}}f_m(z)\frac{y_nk^3(z)}{(1-y_nk^2(z))^2}\,\mathrm{d}z
\end{aligned}$$

$$= \frac{\kappa_1}{4\pi i} \int_{\gamma_{mh}} f_m(z) \frac{d}{dz} \ln\bigl(1 - y_n k^2(z)\bigr)\, dz = -\frac{\kappa_1}{4\pi i} \int_{\gamma_{mh}} f_m'(z) \ln\bigl(1 - y_n k^2(z)\bigr)\, dz$$

$$\longrightarrow \frac{\kappa_1}{2\pi} \int_a^b f'(x) \arg\bigl(1 - y k^2(x)\bigr)\, dx,$$

as $n \to \infty$, $a_l \to a$ and $b_r \to b$; similarly,

$$-\frac{1}{2\pi i} \int_{\gamma_{mh}} f_m(z) \widetilde{M}_2(z)\, dz = \frac{\kappa_2}{2\pi i} \int_{\gamma_{mh}} f_m(z) \frac{y_n k^3(z)}{1 - y_n k^2(z)}\, dz$$

$$\longrightarrow -\frac{\kappa_2}{\pi} \int_a^b f(x) \Im\left(\frac{y k^3(x)}{1 - y k^2(x)}\right) dx.$$

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

Appendix

Lemma A.1. *Under the conditions in Theorem 1.1, we have*

$$\|\mathbb{E}F_n - F\| = O(n^{-1/2}), \qquad \|F_n - F\| = O_p(n^{-2/5}),$$
$$\|F_n - F\| = O(n^{-2/5+\eta}) \qquad \text{a.s. for any } \eta > 0.$$

This follows from Theorems 1.1, 1.2 and 1.3 in [5].

Lemma A.2 [Burkholder (1973), [10]]. *Let X_k, $k = 1, 2, \ldots,$ be a complex martingale difference sequence with respect to the increasing σ-fields $\mathcal{F}_k$. Then, for $p > 1$,*

$$\mathbb{E}\Bigl|\sum X_k\Bigr|^p \le K_p \mathbb{E}\Bigl(\sum |X_k|^2\Bigr)^{p/2}.$$

In the reference [10], only real variables were considered. It is straightforward to extend to complex cases.

Lemma A.3. *For $x = (x_1, \ldots, x_n)^t$ with i.i.d. standardized real or complex entries such that $\mathbb{E}x_i = 0$ and $\mathbb{E}|x_i|^2 = 1$, and for C an $n \times n$ complex matrix, we have, for any $p \ge 2$,*

$$\mathbb{E}|x^* C x - \operatorname{tr} C|^p \le K_p[(\mathbb{E}|x_1|^4 \operatorname{tr} CC^*)^{p/2} + \mathbb{E}|x_1|^{2p} \operatorname{tr}(CC^*)^{p/2}].$$

This is Lemma 8.10 in [8].

Lemma A.4. *For any non-random $p \times p$ matrix A,*

$$\mathbb{E}|r_1^* A r_1|^2 \le K n^{-1} \|A\|^2.$$

Proof. For non-random $p \times p$ matrix A,

$$
\begin{aligned}
\mathbb{E}|r_1^* A r_1|^2 &= \frac{1}{n^2}\mathbb{E}\left|\sum_{l,k=1}^{p} \bar{x}_{l1} a_{lk} x_{k1}\right|^2 \\
&= \frac{1}{n^2}\mathbb{E}\left(\sum_{l\neq k}^{p} \bar{x}_{l1}^2 a_{lk}^2 x_{k1}^2 + \sum_{l\neq k}^{p} |x_{l1}|^2 |x_{k1}|^2 a_{lk} a_{kl} + \sum_{l=1}^{p} |x_{1l}|^4 a_{ll}^2\right) \\
&\leq \frac{K}{n^2}\mathbb{E}\left(\sum_{l,k=1}^{p} |a_{lk}|^2\right) = K n^{-2}\mathbb{E}\operatorname{tr}(A\bar{A}) \leq K n^{-1}\|A\|^2.
\end{aligned}
$$

□

Lemma A.5. *For non-random $p \times p$ matrices $A_k, k = 1, \ldots, s$,*

$$
\mathbb{E}\left|\prod_{k=1}^{s}\left(r_1^* A_l r_1 - \frac{1}{n}\operatorname{tr} A_l\right)\right| \leq K n^{-((s/2)\wedge 3)} \delta_n^{2(s-4)\vee 0} \prod_{k=1}^{s} \|A_l\|. \tag{A.1}
$$

Proof. Recalling the truncation steps $\mathbb{E}|x_{11}|^8 < \infty$ and Lemma A.3, we have, for all $l > 1$,

$$
\begin{aligned}
\mathbb{E}|r_1^* A_1 r_1 - n^{-1}\operatorname{tr} A_1|^l &\leq K\|A_1\|^l n^{-l}\left(n^{l/2} + \left(\sqrt{n}\delta_n\right)^{(2l-8)\vee 0} n\right) \\
&= K\|A_1\|^l n^{-((l/2)\wedge 3)} (\delta_n)^{2(l-4)\vee 0}.
\end{aligned} \tag{A.2}
$$

Then, (A.1) is the consequence of (A.2) and the Hölder inequality. □

Lemma A.6. *Under the conditions in Theorem 1.1, for any $l \geq 2, \mathbb{E}|\beta_j(z)|^l$, $\mathbb{E}|\bar{\beta}_j(z)|^l$ and $|b_n(z)|^l$ are uniformly bounded in γ_{mh}. Furthermore, $\beta_j(z)$, $\bar{\beta}_j(z)$ and $b_n(z)$ are uniformly convergent in probability to $-z\underline{s}(z)$ in γ_{mh}.*

Proof. By (4.2) and (4.3) in [6], we have, for any $l \leq 2$,

$$
\mathbb{E}|\operatorname{tr} D_j^{-1}(z) - \mathbb{E}\operatorname{tr} D_j^{-1}(z)|^l \leq K n^{l/2} v^{-l}, \tag{A.3}
$$

$$
\mathbb{E}|r_j D_j^{-1}(z) r_j - 1/n\mathbb{E}\operatorname{tr} D_j^{-1}(z)|^l \leq K n^{-l/2} v^{-l}. \tag{A.4}
$$

This lemma follows from Lemma A.3, (A.3), (A.4) and the following facts.

Fact 1. *Since $s_n^0(z) = -\frac{1}{2}(\frac{1}{y_n} - \frac{1}{y_n z}\sqrt{z^2 - (1+y_n)z + (1-y_n)^2} - \frac{1-y_n}{y_n z})$ and $\underline{s}_n^0(z) = -\frac{1-y_n}{z} + y_n s_n^0(z)$, we have*

$$
z\underline{s}_n^0(z) = -\tfrac{1}{2}\left(1 - y_n + z - \sqrt{z^2 - (1+y_n)z + (1-y_n)^2}\right).
$$

Thus, $z\underline{s}_n^0(z)$ is bounded in any bounded and closed complex region.

Fact 2.

$$
\begin{aligned}
|b_n(z) - \mathbb{E}\beta_j(z)| &\le \frac{1}{v^2}\mathbb{E}\left|r_j^* D_j^{-1}(z) r_j - \frac{1}{n}\mathbb{E}\operatorname{tr} D_j^{-1}(z)\right| \\
&\le \frac{1}{v^2}\left[\mathbb{E}|\varepsilon_j(z)| + \mathbb{E}\left|\frac{1}{n}\operatorname{tr} D_j^{-1}(z) - \frac{1}{n}\mathbb{E}\operatorname{tr} D_j^{-1}(z)\right|\right] \\
&\le \frac{1}{v^2}\left[\frac{K}{\sqrt{n}v} + \frac{K}{\sqrt{n}v}\right] = \frac{K}{\sqrt{n}v^3},
\end{aligned}
$$

where the last inequality follows from (5.6).

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

$$
z\mathbb{E}\underline{s}_n(z) = -\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\beta_j(z) = -\mathbb{E}\beta_j(z).
$$

Fact 4. *From Lemma* A.1, *we have*

$$
\begin{aligned}
|z\mathbb{E}\underline{s}_n(z) - z\underline{s}_n^0(z)| &\le z y_n \mathbb{E}|s_n(z) - s_n^0(z)| \\
&= z y_n \mathbb{E}\left|\int \frac{1}{x-z}(F^{B_n} - F^{y_n})(\mathrm{d}x)\right| \\
&\le \frac{K}{v}\|F^{B_n} - F^{y_n}\| \\
&= \frac{K}{v}\mathrm{O}_p(n^{-2/5}) = \mathrm{O}_p(n^{-2/5}v^{-1}).
\end{aligned}
$$

□

Lemma A.7. *Under the conditions in Theorem* 1.1, *as* $n \to \infty$,

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

uniformly in γ_{mh}, *where the maximum is taken over all* $1 \le i \le p$ *and* $1 \le j \le n$.

Proof. First, let e_j $(1 \le j \le n)$ be the p-vector whose jth element is 1, the rest being 0 and e_i', the transpose of e_i. Then,

$$
\begin{aligned}
\mathbb{E}|[D^{-1}(z)]_{ii} - [D_j^{-1}(z)]_{ii}| &= \mathbb{E}\left|e_i'\bigl(D^{-1}(z) - D_j^{-1}(z)\bigr)e_i\right| \\
&= \mathbb{E}|\beta_j(z) e_i' D_j^{-1}(z) r_j r_j^* D_j^{-1}(z) e_i| \\
&\le (\mathbb{E}|\beta_j(z)|^2)^{1/2}(\mathbb{E}|r_j^* D_j^{-1}(z) e_i e_i' D_j^{-1}(z) r_j|^2)^{1/2} \le \frac{K}{\sqrt{n}v^2}.
\end{aligned}
$$

Second, by martingale inequality, for any $\epsilon > 0$, we have

$$
\begin{aligned}
&P\Big(\max_{i,j} |\mathbb{E}_j[D^{-1}(z)]_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}| > \epsilon\Big)\\
&\quad \le \sum_{i=1}^{p} P\Big(\max_{j} |\mathbb{E}_j[D^{-1}(z)]_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}| > \epsilon\Big)\\
&\quad \le \sum_{i=1}^{p} \frac{1}{\epsilon^6}\mathbb{E}|[D^{-1}(z)]_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}|^6\\
&\quad = \frac{1}{\epsilon^6}\sum_{i=1}^{p}\mathbb{E}\left|\sum_{l=1}^{n}(\mathbb{E}_l - \mathbb{E}_{l-1})\beta_l(z)e_i' D_l^{-1}(z)r_l r_l^* D_l^{-1}(z)e_i\right|^6\\
&\quad \le K\sum_{i=1}^{p}\mathbb{E}\left(\sum_{l=1}^{n}|(\mathbb{E}_l - \mathbb{E}_{l-1})\beta_l(z)e_i' D_l^{-1}(z)r_l r_l^* D_l^{-1}(z)e_i|^2\right)^3.
\end{aligned}
$$

Let $Z_l(z) = e_i' D_l^{-1}(z)r_l r_l^* D_l^{-1}(z)e_i$. We have that

$$
|\mathbb{E}Z_l(z)| \le \frac{K}{nv^2} \quad \text{and} \quad \mathbb{E}|Z_l(z) - \mathbb{E}Z_l(z)|^2 \le \frac{K}{n^2v^4}.
$$

Thus, we obtain

$$
\begin{aligned}
&P\Big(\max_{i,j} |\mathbb{E}_j[D^{-1}(z)]_{ii} - \mathbb{E}[D^{-1}(z)]_{ii}| > \epsilon\Big)\\
&\quad \le K\sum_{i=1}^{p}\mathbb{E}\left(\sum_{l=1}^{n}\frac{K}{n^2v^4}\right)^3 = \frac{K}{n^2v^{12}}.
\end{aligned}
$$

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 γ_{mh}. The proof of Lemma A.7 is thus complete. □

Acknowledgements

The authors would like to thank one Associate Editor and one referee for many constructive comments. 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.

References

[1] Anderson, G.A. and Ofer, Z. (2006). A CLT for a band matrix model. *Probab. Theory Related Fields* **134** 283–338. MR2222385

[2] Bai, Z.D. (1993). Convergence rate of expected spectral distributions of large random matrices. Part II. Sample covariance matrices. *Ann. Probab.* **21** 649–672. MR1217560

[3] Bai, Z.D. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. *Statist. Sinica* **9** 611–677. MR1711663

[4] Bai, Z.D., Miao, B.Q. and Tsay, J. (1997). A note on the convergence rate of the spectral distributions of large random matrices. *Statist. Probab. Lett.* **34** 95–102. MR1457501

[5] Bai, Z.D., Miao, B.Q. and Yao, J.F. (2003). Convergence rates of spectral distributions of large sample covariance matrices. *SIAM J. Matrix Anal. Appl.* **25** 105–127. MR2002902

[6] Bai, Z.D. and Silverstein, J.W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices. *Ann. Probab.* **26** 316–345. MR1617051

[7] Bai, Z.D. and Silverstein, J.W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792

[8] Bai, Z.D. and Silverstein, J.W. (2006). *Spectral Analysis of Large Dimensional Random Matrices*. *Mathematics Monograph Series* **2**. Beijing: Science Press.

[9] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. New York: Wiley. MR1324786

[10] Burkholder, D.L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42. MR0365692

[11] Grenander, U. and Silverstein, J.W. (1977). Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.* **32** 499–519. MR0476178

[12] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38. MR0650926

[13] Marčenko, V.A. and Pastur, L.A. (1967). Distribution for some sets of random matrices. *Math. USSR-Sb.* **1** 457–483.

[14] Pastur, L.A. and Lytova, A. (2008). Central Limit Theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.* **37** 1778–1840. MR2561434

[15] Rao, C.R. and Rao, M.B. (2001). *Matrix Algebra and Its Applications to Statistics and Econometrics.* River Edge, NJ: World Scientific. MR1660868

[16] Silverstein, J.W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *J. Multivariate Anal.* **55** 331–339. MR1370408

[17] Wachter, K.W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Statist.* **6** 1–18. MR0467894

[18] Yin, Y.Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* **20** 50–68. MR0862241

[19] Yin, Y.Q., Bai, Z.D. and Krishnaiah, P.R. (1988). On limit of the largest eignevalue of the large dimensional sample covariance matrix. *Probab. Theory Related Fields* **78** 509–521. MR0950344

Received January 2009 and revised August 2009