Convergence rates of spectral distributions of large sample covariance matrices

Z. D. BAI [*]
Dept. of Stat. & Appl. Prob.
Natl. U. of Singapore
matbaizd@nus.edu.sg

Baiqi MIAO [†]
Dept. of Stat. & Fin.
U. of Sci. & Tech. of China
bqmiao@ustc.edu.cn

Jianfeng YAO
SAMOS-MATISSE
Université Paris I
yao@univ-paris1.fr

March 17, 2000

Abstract

In this paper, we improve known results on the convergence rates of spectral distributions of large dimensional sample covariance matrices of size $p \times n$. Depending on the limiting value y of the ratio p/n and by using the tool of Stieltjes transforms, we first prove that the expected spectral distribution converges to the limiting Marčenko-Pastur distribution at a rate of $O(n^{-\frac{1}{2}})$ for $y \notin \{0, 1\}$, and of $O(n^{-\frac{1}{4}})$ for $y = 1$, under the assumption that the entries have a finite 8-th order moment. Furthermore, the rates for both the convergence in probability and the almost sure convergence are investigated.

AMS 1991 subject classification: 60F99.
Keywords and phrases: Convergence rate, random matrix, spectral distribution, Manc̆henko-Pastur distribution

[*]Supported in part by the NUS...**TO BE CLARIFIED**
[†]Supported in part by the National Foundation of Natural Science of China.

1 Introduction

The spectral analysis of large dimensional random matrices has been actively developed in the last decades since the initial contributions of Wigner (1955, 1958), see the review by Bai (1999) and the references therein. Various limiting distributions were discovered including the Wigner semicircular law (Wigner, 1955), the Marčenko-Pastur law (Marčenko and Pastur, 1967) and the circular law (Bai,1997).

Let A be an $n \times n$ symmetric matrix, and $\lambda_1 \leq \cdots \leq \lambda_n$ be the eigenvalues of A. The spectral distribution F^A of A is defined as

$$F^A(x) = \frac{1}{n} \times \text{number of elements in } \{k : \lambda_k \leq x\} .$$

Let $\mathbf{X}_p = (x_{ij})_{p\times n}$ be a $p \times n$ observation matrix whose entries are mutually independent and have a common mean zero and variance 1. The entries of $\mathbf{X}_p$ may depend on n but we suppress the index n for simplicity. In this paper, we consider the sample covariance matrix $\mathbf{S}_p = n^{-1}\mathbf{X}_p\mathbf{X}_p'$. Assume that the ratio p/y of sizes tends to a positive limit y as $n \to \infty$. Under suitable moment conditions on the entries x_{ij}'s, it is known that the empirical spectral distribution (ESD) $F_p := F^{\mathbf{S}_p}$ converges to the following Marčenko-Pastur distribution F_y with index y with density

$$F_y'(x) = \begin{cases} \frac{1}{2\pi xy}\sqrt{(x-a)(b-x)}\ , & \text{if } a < x < b \\ \\ 0\ , & \text{otherwise,} \end{cases}$$

where $a = (1-\sqrt{y})^2$, $b = (1+\sqrt{y})^2$.

An important question arose here is the problem of the convergence rates. However, no significant progress had been made before the introduction of a novel and powerful tool, namely the Stieltjes transforms, by Bai (1993a,1993b). Using this methodology, Bai (1993b) proved that the expected ESD, $\mathbb{E}F_p$ converges to F_y at a rate of $O(n^{-1/4})$ and $O(n^{-5/48})$ according to $y \neq 1$ or $y = 1$, respectively. In a further work by Bai *et al.* (1997), these rates are also established for the convergence in probability of the ESD F_p itself.

In this work, we prove the following theorems which give a significant improvement of these rates.

The following conditions will be used.

(C.1) $\quad \mathbb{E}x_{ij} = 0, \quad \mathbb{E}x_{ij}^2 = 1, \ \ 1 \leq i \leq p, \ 1 \leq j \leq n,$

(C.2) $\quad \sup_{i,j,n} \mathbb{E}|x_{ij}|^8 < \infty$,

(C.3) $\quad \sum_{ij} \mathbb{E}x_{ij}^8 I_{(|x_{ij}|\geq\varepsilon\sqrt{n})} = o(n^2)$, for any $\varepsilon > 0$.

(C.2') $\quad \sup_{i,j,n} \mathbb{E}|x_{ij}|^k < \infty$, for all integer $k \geq 1$.

Throughout the text, we use the notation $Z_n = O_p(a_n)$ if the sequence $(a_n^{-1}Z_n)$ is tight, and $Z_n = o_p(a_n)$ when $a_n^{-1}Z_n$ tends to 0 in probability. Let be $\|f\| = \sup_x |f(x)|$.

Theorem 1.1 *Assume that the conditions* ***C.1-2-3*** *are satisfied. Then*

$$\|\mathbb{E}F_p - F_y\| = \begin{cases} O(n^{-\frac{1}{2}}), & \text{if} \quad 0 < y < 1 \\ \\ O(n^{-\frac{1}{4}}), & \text{if} \quad y = 1. \end{cases}$$

Theorem 1.2 *Assume that the conditions* ***C.1-2-3*** *are satisfied. Then*

$$\|F_p - F_y\| = \begin{cases} O_p(n^{-\frac{2}{5}}), & \text{if} \quad 0 < y < 1 \\ O_p(n^{-\frac{2}{9}}), & \text{if} \quad y = 1. \end{cases}$$

Theorem 1.3 *Assume that the conditions* ***C.1-2'-3*** *are satisfied. Then, for all $\eta > 0$ and almost surely,*

$$\|F_p - F_y\| = \begin{cases} o(n^{-\frac{2}{5}+\eta}), & \text{if} \quad y \neq 1, \\ o(n^{-\frac{2}{9}+\eta}), & \text{if} \quad y = 1. \end{cases}$$

It is worth noticing that the convergence rates given above for the case $0 < y < 1$ also apply to the case $y > 1$, since the last case can be reduced to the first case by interchanging the roles of row and column sizes p and n.

The proofs of these main results will be given in Section 4. To simplify their presentation, we first establish several intermediate results in Section 3 after the introduction of some necessary notations and preliminary consequences in Section 2.

2 Definitions and easy consequences

Throughout the paper, the transpose of a possibly complex matrix $\mathbf{A}$ is denoted by $\mathbf{A}^{\mathrm{T}}$, and its conjugate by $\overline{\mathbf{A}}$. For each fixed $p,\ n$ and $k = 1, \ldots, p$, let us denote by $\mathbf{x}_k = (x_{k1}, \ldots, x_{kn})^T$ the k-th row of $\mathbf{X}_p$ arranged as a column vector, $\mathbf{X}_p(k)$ be the $(p-1) \times n$ sub-matrix obtained from $\mathbf{X}_p$ by deleting its k-th row. Let us define

$$\begin{aligned} \alpha_k &:= \frac{1}{n}\mathbf{X}_p(k)\mathbf{x}_k, & \mathbf{S}_k &:= \frac{1}{n}\mathbf{X}_p(k)\mathbf{X}_p^T(k) & \mathbf{B}_k &:= \frac{1}{n}\mathbf{X}_p^T(k)\mathbf{D}_k\mathbf{X}_p(k), \\ \mathbf{B} &:= \frac{1}{n}\mathbf{X}_p^T\mathbf{D}\mathbf{X}_p & \mathbf{D}_k &:= (\mathbf{S}_k - z\mathbf{I}_{p-1})^{-1}, & \mathbf{D} &:= (\mathbf{S} - z\mathbf{I}_p)^{-1}, \\ \mathbf{\Gamma}_k &:= \mathbf{D}_k\overline{\mathbf{D}}_k, & \mathbf{\Lambda}_k &:= \mathbf{D}_k\mathbf{S}_k\overline{\mathbf{D}}_k. \end{aligned} \tag{2.1}$$

Here $\mathbf{I}_m$ is the m-dimensional identity matrix and z a complex number with a positive imaginary part.

Following Bai (1993b), the Stieltjes transform of the spectral distribution F_p of the sample covariance matrix $\mathbf{S}_p$ is defined for $z = u + iv$ with $v > 0$, by

$$m_p(z) = \int_{-\infty}^{\infty} \frac{1}{x - z} dF_p(x),$$

and it is well-known that

$$m_p(z) = \frac{1}{p} tr(\mathbf{S} - z\mathbf{I}_p)^{-1} .$$

Similarly, the Stieltjes transform of the spectral distribution $F_p^{(k)}$ of the sub-matrix $\mathbf{S}_k$ satisfies

$$m_p^{(k)}(z) = \int_{-\infty}^{\infty} \frac{1}{x - z} dF_p^{(k)}(x) = \frac{1}{p-1} tr(\mathbf{S}_k - z\mathbf{I}_{p-1})^{-1} .$$

Lastly, the Stieltjes transform of the (limiting) Marčenko-Pastur distribution F_y is

$$\begin{aligned} m(z) &:= \int_{-\infty}^{\infty} \frac{1}{x-z} dF_y(x) \\ &= \begin{cases} -\dfrac{y+z-1-\sqrt{(1-y-z)^2-4y}}{2yz}, & 0 < y < 1, \\[2ex] -\dfrac{z-\sqrt{z^2-4z}}{2z}, & y = 1. \end{cases} \end{aligned} \tag{2.2}$$

Here the square root $\sqrt{z}$ is the one with a positive imaginary part. Bai (1993b) also provided the following bounds for $m(z)$ which will play a key role in next derivations :

$$m(z) \leq \begin{cases} \dfrac{1+3\sqrt{y}}{\sqrt{y}(1-y)}, & 0 < y < 1, \\[2ex] \dfrac{2}{\sqrt{v}}, & y = 1. \end{cases} \tag{2.3}$$

Lemma 2.1 *Let $\mathbf{x} = (x_1, \ldots, x_n)^T$ and $\mathbf{y} = (y_1, ..., y_n)^T$ be independent real random vectors with independent elements. Suppose that for all $1 \leq j \leq n$, $\mathbb{E}x_j = Ey_j = 0$, $\mathbb{E}|x_j|^2 = \mathbb{E}|y_j|^2 = 1$, $\mathbb{E}|x_j|^4 \leq L < \infty$, and that $\mathbf{A}$ is an $n \times n$ complex symmetric matrix. Let $\nu_k = \max_{j \leq n}(\ \mathbb{E}|x_j|^k,\ \mathbb{E}|y_j|^k)$. Then*

(i). $\mathbb{E}|\mathbf{x}^T\mathbf{A}\mathbf{y}|^2 = tr(\mathbf{A}\bar{\mathbf{A}})$ *;*

(ii). $\mathbb{E}|\mathbf{x}^T\mathbf{A}\mathbf{x}|^2 \leq (L-1)tr(\mathbf{A}\bar{\mathbf{A}}) + |tr\mathbf{A}|^2$ *;*

(iii). $\mathbb{E}|\mathbf{x}^T\mathbf{A}\mathbf{x} - tr\mathbf{A}|^2 \leq (L-1)(tr\mathbf{A}\bar{\mathbf{A}})$ *;*

(iv). $\mathbb{E}|\mathbf{x}^T\mathbf{A}\mathbf{x} - tr\mathbf{A}|^{2k} \leq d_k\left[\nu_{4k} tr(\mathbf{A}\bar{\mathbf{A}})^k + (Ltr(\mathbf{A}\bar{\mathbf{A}}))^k\right]$ *for $k \geq 2$ and some positive constant d_k depending on k only.*

Lemma 2.1 can be proved in an elementary way and is stated in Bai *et al.* (1997).

Lemma 2.2 *Let G_1 and G_2 be probability distribution functions and $z = u + iv$, $v > 0$. Then for each positive integer m,*

$$\left| \int_{-\infty}^{\infty} \frac{1}{|x-z|^m} d(G_1(x) - G_2(x)) \right| \leq \frac{2}{v^m} \|G_1 - G_2\|.$$

Proof. Let be $G^* := G_1 - G_2$. We have, by integration by parts,

$$\begin{aligned}
&\left|\int_{-\infty}^{\infty} \frac{1}{|x-z|^m} dG^*\right| \\
&= \left|-\int_{-\infty}^{\infty} G^*(x) d\left[\frac{1}{|x-z|^m}\right]\right| \\
&= \left|-\int_{-\infty}^{Re(z)} G^*(x) d\left[\frac{1}{|x-z|^m}\right] + \int_{Re(z)}^{\infty} G^*(x) d\left[-\frac{1}{|x-z|^m}\right]\right| \\
&\leq \|G^*\| \left\{\int_{-\infty}^{Re(z)} d\left[\frac{1}{|x-z|^m}\right] + \int_{Re(z)}^{\infty} d\left[-\frac{1}{|x-z|^m}\right]\right\} \\
&= \|G^*\| \left\{\frac{1}{|x-z|^m}\Big|_{-\infty}^{Re(z)} + \left(-\frac{1}{|x-z|^m}\Big|_{Re(z)}^{\infty}\right)\right\} = \|G^*\| \frac{2}{v^m}\ . \qquad \blacksquare
\end{aligned}$$

We will need the following auxiliary variables.

$$\begin{aligned}
\varepsilon_k &= -\frac{1}{n}\sum_{j=1}^{n}(x_{kj}^2 - 1) + \frac{1}{n}(\mathbf{x}_k{}'\mathbf{B}_k\mathbf{x}_k - \mathbb{E}tr\mathbf{B}), \\
\varepsilon_k^* &= -\frac{1}{n}\sum_{j=1}^{n}(x_{kj}^2 - 1) + \frac{1}{n}(\mathbf{x}_k{}'\mathbf{B}_k\mathbf{x}_k - tr\mathbf{B}_k), \\
\widetilde{\varepsilon}_k &= \frac{1}{n}(tr\mathbf{B}_k - \mathbb{E}tr\mathbf{B}_k) = \frac{z}{n}(tr\mathbf{D}_k - \mathbb{E}tr\mathbf{D}_k), \\
\pi_k &= \frac{1}{n}\mathbb{E}(tr\mathbf{B}_k - tr\mathbf{B}) = \frac{z}{n}\mathbb{E}(tr\mathbf{D}_k - tr\mathbf{D}) - \frac{1}{n}, \\
\beta_k &= -\frac{1}{n}\sum_{j=1}^{n}(x_{kj}^2 - 1) + z - 1 + \frac{1}{n}\mathbf{x}_k{}'\mathbf{B}_k\mathbf{x}_k, \\
\beta_k^* &= z - 1 + \frac{1}{n}tr\mathbf{B}_k, \\
\beta &= z - 1 + \frac{1}{n}tr\mathbf{B}
\end{aligned}$$

We summarize below some inequalities which will be used in the derivations. Let $\Delta = \|\ \mathbb{E}F_p - F\|$ and $M := \sup_{i,j,n}\ \mathbb{E}|x_{ij}|^4$. For fixed (n,p) and $1 \leq k \leq p$, we define the σ-algebra $\mathcal{F}^{(k)} = \sigma(\mathbf{x}_i\ :\ i = 1,\ldots,p\ ;\ i \neq k)$ and $\mathcal{F}_k = \sigma(\mathbf{x}_i\ :\ i = 1,\ldots,p\ ;\ i > k)$. Notice that $\mathcal{F}_k \subseteq \mathcal{F}^{(k)}$.

(i). (Lemma 3.3 of Bai (1993a)) :

$$|(p-1)F_p^{(k)}(x) - pF_p(x)| \leq 1. \tag{2.4}$$

(ii). ((3.11) of Bai (1993a)) :

$$|tr\mathbf{D} - tr\mathbf{D}_k| = \left|\int_{-\infty}^{\infty} \frac{d[pF_p(x) - (p-1)F_p^{(k)}(x)]}{x-z}\right| \leq v^{-1}. \tag{2.5}$$

(iii). ((4.7) of Bai (1993a)) :

$$m_p(z) = -\frac{1}{p}\sum_{k=1}^{p}\frac{1}{\beta_k}. \tag{2.6}$$

(iv). (Lemma 2.2 of Bai *et al.* (1997)) :

$$\mathbb{E}|m_p(z) - \mathbb{E}(m_p(z))|^2 \le p^{-1}v^{-2}. \tag{2.7}$$

(v). (from $|\beta_k^*| \ge Im(\beta_k^*) = v(1 + n^{-1}tr\mathbf{\Lambda}_k)$) :

$$|\beta_k^*|^{-1}(1 + n^{-1}tr\mathbf{\Lambda}_k) \le v^{-1}. \tag{2.8}$$

(vi).

$$|\beta_k| \ge Im(\beta_k) = v(1 + \frac{1}{n}\alpha_k^T \mathbf{D}_k \overline{\mathbf{D}_k}\alpha_k). \tag{2.9}$$

(vii).

$$|1 + \frac{1}{n}\alpha_k^T {\mathbf{D}_k}^2 \alpha_k| \le 1 + \frac{1}{n}\alpha_k^T \mathbf{D}_k \overline{\mathbf{D}_k}\alpha_k. \tag{2.10}$$

Let $\lambda_{kj},\ j = 1, 2, \ldots, p-1$, be the eigenvalues of $\mathbf{S}_k$ which can be decomposed in a diagonal form on a basis of orthonormal and real eigenvectors. Let $\mathbf{L}$ be a complex matrix having the product form $\mathbf{L} = \mathbf{M}^{\ell}\mathbf{N}^{\ell'}$ for some integers $\ell,\ \ell'$ and factors $\mathbf{M},\ \mathbf{N}$ equal to one of the matrices $\{\mathbf{D}_k, \overline{\mathbf{D}}_k, \mathbf{B}_k, \overline{\mathbf{B}}_k\}$. An important feature that we will frequently use in the sequel is that such a matrix $\mathbf{L}$ can be decomposed into a diagonal form *on the same basis of the eigenvectors of* $\mathbf{S}_k$. Moreover, the eigenvalues of $\mathbf{L}$ can be straightforwardly expressed in term of the λ_{kj}'s. In particular, we have the following

Lemma 2.3 *Assume that $|z| \le T$ where $T \ge 1$. Then for all integers $\ell \ge 1$*

$$tr(\mathbf{\Gamma}_k)^{\ell} \le \left(\frac{1}{v^2}\right)^{\ell-1} tr\mathbf{\Gamma}_k\ , \tag{2.11}$$

$$tr(\mathbf{\Lambda}_k)^{\ell} \le \left(\frac{T}{v^2}\right)^{\ell-1} tr\mathbf{\Lambda}_k\ . \tag{2.12}$$

Proof. (i) The inequality (2.11) follows from

$$tr(\mathbf{\Gamma}_k)^{\ell} = \sum_{j=1}^{p-1}\frac{1}{|\lambda_{kj} - z|^{2\ell}} \le v^{-2(\ell-1)}\sum_{j=1}^{p-1}\frac{1}{|\lambda_{kj} - z|^2} = v^{-2(\ell-1)}tr\mathbf{\Gamma}_k\ .$$

(ii) For the inequality (2.12), we have

$$tr(\mathbf{\Lambda}_k)^{\ell} = \sum_{j=1}^{p-1}\frac{\lambda_{kj}^{\ell}}{|\lambda_{kj} - z|^{2\ell}}.$$

The conclusion follows from that The function $\varphi(\lambda) := \lambda^{-1}|\lambda - z|^2$ defined on $(0, \infty)$ is convex and has an unique minimum of value φ^* satisfying

$$\varphi^* = 2(\sqrt{u^2 + v^2}) - u = 2\frac{v^2}{|z| + u} \ge \frac{v^2}{T}. \qquad \blacksquare$$

Lemma 2.4 *For the Marčenko-Pastur distribution F_y, we have*

$$\int_a^b \frac{1}{|x-z|^2} dF_y(x) \leq \begin{cases} \frac{1}{(1-y)\sqrt{y}} v^{-1}, & 0 < y < 1, \\ |z|^{-1} v^{-1/2}, & y = 1. \end{cases} \tag{2.13}$$

Proof. For $0 < y < 1$, we have by elementary calculus that the density function $F_y{}'(x)$ has an unique maximum of value $(\pi(1-y)\sqrt{y})^{-1}$. Thus

$$\begin{aligned} \int_a^b \frac{1}{|x-z|^2} dF_y(x) &\leq \frac{1}{\pi(1-y)\sqrt{y}} \int_a^b \frac{1}{|x-z|^2} dx \\ &\leq \frac{1}{(1-y)\sqrt{y}} v^{-1}. \end{aligned}$$

When $y = 1$, $a = 0$ and $b = 4$. We find that

$$\begin{aligned} &\int_a^b \frac{1}{|x-z|^2} dF_y(x) \\ &\leq \frac{1}{\pi} \int_0^4 \frac{dx}{\sqrt{x}[(x-u)^2 + v^2]} \leq \frac{1}{\pi} \int_0^\infty \frac{dx}{\sqrt{x}[(x-u)^2+v^2]} \leq |z|^{-1} v^{-1/2}. \qquad \blacksquare \end{aligned}$$

Lemma 2.5 *For the Marčenko-Pastur distribution F_y, we have for any $0 < v < 4\sqrt{y}$,*

$$\sup_x \int_{|u|\leq v} |F_y(x+u) - F_y(x)|\, du \leq \frac{14\sqrt{2(1+y)}}{3\pi y} \frac{1}{\sqrt{v} + (1-\sqrt{y})} v^2.$$

Proof. It is enough to consider the part $0 \leq u \leq v$ only in the integral since the remaining part for $-v \leq u \leq 0$ can be handled in a similar way. Set $x = a + \lambda$ with $\lambda \geq 0$ and $\Phi(\lambda) := \int_0^v [F_y(x+u) - F_y(x)] du$. Then

$$\begin{aligned} \Phi(\lambda) &= \int_0^v du \int_x^{x+u} F_y{}'(t)\, dt = \int_{a+\lambda}^{a+\lambda+v} \frac{a+\lambda+v-t}{2\pi y t} \sqrt{(t-a)(b-t)} dt \\ &= \int_\lambda^{\lambda+v} \frac{\lambda+v-u}{2\pi y(u+a)} \sqrt{u(4\sqrt{y}-u)}\, du. \end{aligned} \tag{2.14}$$

Let $\phi(u) := (u+a)^{-1}\sqrt{u(4\sqrt{y}-u)}$.

Case $0 < y < 1$ **:** We have $a > 0$ and the derivative of $\log(\phi(u))^2$ is

$$\frac{1}{u} - \frac{1}{4\sqrt{y}-u} - \frac{2}{u+a} = \frac{2(2\sqrt{y}a - (1+y)u)}{u(4\sqrt{y}-u)(u+a)}$$

Let $\rho := (1+y)^{-1}(2a\sqrt{y})$. Thus $\phi(u)$ is decreasing when $u > \rho$ and increasing when $u < \rho$. Since

$$\frac{d\Phi(\lambda)}{d\lambda} = \frac{1}{2\pi y} \left(\int_\lambda^{\lambda+v} [\phi(u) - \phi(v)]\, du \right),$$

it follows that for $\lambda > \rho$, $\Phi(\lambda)$ is decreasing and then $\Phi(\lambda) \leq \Phi(\rho)$; and for $\lambda < \rho - v$, $\Phi(\lambda)$ is increasing and then $\Phi(\lambda) \leq \Phi(\rho - v)$. Hence $\Phi(\lambda)$ reaches its maximum only for some $\lambda \in (\max(\rho - v, 0), \rho)$. Now suppose that $\lambda \in (\rho - v, \rho)$, it follows from (2.14) that

$$\begin{aligned}
\Phi(\lambda) &\leq \frac{2y^{1/4}}{2\pi y} \int_{\lambda}^{\lambda+v} \frac{\lambda + v - u}{u + a} \sqrt{u} du \\
&= 2(\pi y^{3/4})^{-1} \Big\{ (\lambda + v + a) \Big[(\sqrt{\lambda + v} - \sqrt{\lambda}) \\
&\quad - \sqrt{a} \left(\arctan \sqrt{\frac{\lambda + v}{a}} - \arctan \sqrt{\frac{\lambda}{a}} \right) \Big] - \frac{1}{3} \left[(\lambda + v)^{3/2} - \lambda^{3/2} \right] \Big\} .
\end{aligned}$$

Notice that $-\sqrt{a} \arctan \frac{x}{\sqrt{a}}$ is convex , we get

$$\frac{1}{\sqrt{a}} \left(\arctan \sqrt{\frac{\lambda + v}{a}} - \arctan \sqrt{\frac{\lambda}{a}} \right) \geq \frac{a}{\lambda + v + a} \left(\sqrt{\lambda + v} - \sqrt{\lambda} \right) \ ,$$

and by setting $\lambda^* = \sqrt{\lambda + v} - \sqrt{\lambda}$, we have

$$\begin{aligned}
\Phi(\lambda) &\leq \frac{2}{\pi y^{3/4}} \{ (a + \lambda + v)(\lambda^* - \frac{a}{a + \lambda + v} \lambda^*) - (\lambda^*(\lambda + \sqrt{\lambda}\lambda^* + \frac{1}{3} \lambda^{*2}) \} \\
&= \frac{2}{\pi y^{3/4}} \left[\sqrt{\lambda} \lambda^{*2} + \frac{2}{3} \lambda^{*3} \right] .
\end{aligned} \tag{2.15}$$

Let $c^2 = \frac{1+y}{2\sqrt{y}}$. Since $\lambda + v \geq c^{-2} a$, we have

$$\begin{aligned}
\frac{\sqrt{\lambda}}{(\sqrt{\lambda + v} + \sqrt{\lambda})^2} &\leq \frac{c}{\sqrt{a} + \sqrt{v}} \ , \\
\frac{1}{(\sqrt{\lambda + v} + \sqrt{\lambda})^3} &\leq \frac{2c}{(\sqrt{a} + \sqrt{v}) v} \ .
\end{aligned}$$

Hence

$$\Phi(\lambda) \leq \frac{2}{\pi y^{3/4}} \cdot \frac{7c}{3(\sqrt{a} + \sqrt{v})} v^2 = \frac{7\sqrt{2(1 + y)}}{3\pi y} \frac{1}{\sqrt{v} + (1 - \sqrt{y})} v^2.$$

Case $y = 1$ **:** Here $a = 0$ and

$$\begin{aligned}
\Phi(\lambda) &= \int_{\lambda}^{\lambda+v} \frac{\lambda + v - u}{2\pi} \sqrt{\frac{4 - u}{u}} \, du \\
\frac{d\Phi(\lambda)}{d\lambda} &= \frac{1}{2\pi} \int_{\lambda}^{\lambda+v} \left[\sqrt{\frac{4 - u}{u}} - \sqrt{\frac{4 - \lambda}{\lambda}} \right] du \ .
\end{aligned}$$

But $(4 - u)/u$ is decreasing for $u > 0$, thus $\Phi(\lambda)$ is decreasing for $\lambda \geq 0$. Hence

$$\Phi(\lambda) \leq \Phi(0) = \int_0^v \frac{v - u}{2\pi} \sqrt{\frac{4 - u}{u}} du \leq \frac{2}{\pi} v^{3/2}.$$

Combining these two cases proves the lemma. ∎

3 Intermediate lemmas

In this section, we establish some more technical lemmas. Let $\nu_\ell = \sup_{i,j,n}\{E|x_{ij}|^\ell\}$.

Lemma 3.1 *For each $\ell > 1/2$ with $\nu_{4\ell} < \infty$, there exist positive constants c_ℓ independent of n and v, such that for all n, v satisfying $nv \geq T$, we have*

$$\mathbb{E}\left(|\varepsilon_k^*|^{2\ell} \middle| \mathcal{F}^{(k)} \right) \leq c_\ell n^{-\ell} \left(1 + \frac{1}{n} tr \mathbf{\Lambda}_k\right)^\ell \tag{3.1}$$

and

$$\mathbb{E}\left(\frac{(\varepsilon_k^*)^{2\ell}}{|\beta_k^*|^\ell} \middle| \mathcal{F}^{(k)} \right) \leq c_\ell n^{-\ell} v^{-\ell} \; . \tag{3.2}$$

Proof. We have

$$\begin{aligned}
\mathbb{E}\left(|\varepsilon_k^*|^{2\ell} |\mathcal{F}^{(k)} \right) &= \mathbb{E}\left(\left| -\frac{1}{n}\sum_{j=1}^n (x_{kj}^2 - 1) + \frac{1}{n}(\mathbf{x}_k' \mathbf{B}_k \mathbf{x}_k - tr\mathbf{B}_k) \right|^{2\ell} \middle| \mathcal{F}^{(k)} \right) \\
&\leq 2^{2\ell-1} n^{-2\ell} \left\{ \mathbb{E}\left| \sum_{j=1}^n (x_{kj}^2 - 1) \right|^{2\ell} + \mathbb{E}\left(|\mathbf{x}_k'\mathbf{B}_k\mathbf{x}_k - tr\mathbf{B}_k|^{2\ell} \middle| \mathcal{F}^{(k)} \right) \right\} \\
&:= A + B \; .
\end{aligned}$$

For the first term A, by the Burkholder inequality, we get

$$\begin{aligned}
&\mathbb{E}\left| \sum_{j=1}^n (x_{kj}^2 - 1) \right|^{2\ell} \\
&\leq c_\ell \, \mathbb{E}\left[\sum_{j=1}^n (x_{kj}^2 - 1)^2 \right]^\ell \leq c_\ell n^{\ell-1} \, \mathbb{E}\left[\sum_{j=1}^n (x_{kj}^2 - 1)^{2\ell} \right] \leq c_\ell \nu_{4\ell} n^\ell \; .
\end{aligned}$$

For the second term B, we first notice that

$$tr\left(\mathbf{B}_k \overline{\mathbf{B}}_k\right) = tr\mathbf{B}_k + \overline{z} tr\mathbf{\Lambda}_k \; ,$$

and

$$\frac{1}{n}|tr\mathbf{B}_k| = \left| y + \frac{z}{n} tr\mathbf{D}_k \right| \leq 1 + \frac{T}{nv} \leq 2 \; .$$

Hence

$$\frac{1}{n} tr\left(\mathbf{B}_k \overline{\mathbf{B}}_k\right) \leq 2 + \frac{T}{n} tr\mathbf{\Lambda}_k \leq T\left(1 + \frac{1}{n} tr\mathbf{\Lambda}_k\right) \; .$$

Therefore by Lemma 2.1,

$$\begin{aligned}
&E\left(|\mathbf{x}_k'\mathbf{B}_k\mathbf{x}_k - tr\mathbf{B}_k|^{2\ell} \middle| \mathcal{F}^{(k)} \right) \\
&\leq c_\ell(\nu_{4\ell} + M^\ell)(tr\mathbf{B}_k\overline{\mathbf{B}}_k)^\ell \leq c_\ell T^{-\ell} n^{-\ell} \left(1 + \frac{1}{n} tr\mathbf{\Lambda}_k\right)^\ell \; .
\end{aligned}$$

Combining the bounds for A and B proves the first conclusion. The second conclusion immediately follows by taking into account the inequality (2.8). ∎

Lemma 3.2 *If $n^{-1/2} \le v < 1$, then there are positive constants $C_1,\ C_2$ such that for large n and each $k \le n$,*

(i). $|\ \mathbb{E}tr(\mathbf{D}_k\overline{\mathbf{D}}_k)| \le C_1 p\dfrac{\Delta+v}{v^2}$.

(ii). $\mathbb{E}|\varepsilon_k^*|^2 \le C_2\dfrac{1}{n}\left(1+|z|^2\dfrac{\Delta+v}{v^2}\right)$.

Proof. (i). Recall that $\Delta = \|\ \mathbb{E}F_p - F_y\|$. By Lemma 2.2,

$$\left|\int_{-\infty}^{\infty}\frac{1}{|x-z|^2}d(\ \mathbb{E}F_p(x)-F_y(x))\right| \le \frac{2\Delta}{v^2}\ .$$

Application of Lemmas 2.1 and (2.4) yields that

$$\begin{aligned}
|\ \mathbb{E}tr(\mathbf{D}_k\overline{\mathbf{D}}_k)| &= \left|(p-1)\int_{-\infty}^{\infty}\frac{1}{|x-z|^2}d[\ \mathbb{E}F_p^{(k)}(x)]\right| \\
&\le \left|\int_{-\infty}^{\infty}\frac{1}{|x-z|^2}d[(p-1)\ \mathbb{E}F_p^{(k)}(x)-p\ \mathbb{E}F_p(x)]\right| \\
&\quad +p\left|\int_{-\infty}^{\infty}\frac{1}{|x-z|^2}d[\ \mathbb{E}F_p(x)-F_y(x)]\right|+p\left|\int_{-\infty}^{\infty}\frac{1}{|x-z|^2}dF_y(x)\right| \\
&\le \frac{2}{v^2}+p\frac{2\Delta}{v^2}+p\left|\int_{-\infty}^{\infty}\frac{1}{|x-z|^2}dF_y(x)\right|\ .
\end{aligned}$$

By Lemma 2.4, the last term is bounded by C_3pv^{-1} or $C_3p(|z|\sqrt{v})^{-1}$ according to $0<y<1$ or $y=1$. Taking into account the condition $v\sqrt{n}\ge 1$, we have for large n, $pv\ge 2C_3$ for the first case and for the second one, since $\sqrt{v}\le v\le |z|$, $p\sqrt{v}\ge 2C_3$. The conclusion (i) follows in both cases.

(ii). The conclusion follows from (i), (3.1) and the fact

$$tr\mathbf{B}_k\overline{\mathbf{B}}_k = tr(\mathbf{I}_{p-1}+z\mathbf{D}_k)(I+\overline{z}\overline{\mathbf{D}}_k) \le 2(p+|z|^2 tr\mathbf{D}_k\overline{\mathbf{D}}_k)\ . \quad ∎$$

Let us define $v_y = v$ for $0<y<1$ and $v_y=\sqrt{v}$ for $y=1$.

Lemma 3.3 *Assume $|z|\le T$ with $T\ge 2$, and $\sqrt{n}v \ge 6\sqrt{2T(M+2)}$. Then for large n and a positive constants C_1 ,*

$$\sum_{k=1}^{p}\mathbb{E}(|\beta_k^*|^{-1}) \le C_1 n(\Delta+v_y)v^{-1}\ . \tag{3.3}$$

Proof. First notice that from the definition of ε_k^*, we have $(\beta_k^*)^{-1} = \beta_k^{-1}(1+\beta_k^{-1}\varepsilon_k^*)$. By (2.5),

$$|\beta_k^*-\beta| = \frac{1}{n}|-1+z(trD_k-trD)| \le \frac{1}{n}(1+\frac{|z|}{v}) \le \frac{2T}{nv}\ .$$

Taking account of (2.6) and (3.2), we obtain

$$\begin{aligned}
&\sum_{k=1}^{p} \mathbb{E}(|\beta_k^*|^{-1}) \\
\le\ & \sum_{k=1}^{p} \mathbb{E}\left|\frac{1}{|\beta_k^*|} - \frac{1}{|\beta|}\right| + \mathbb{E}\left|\sum_{k=1}^{p}(\frac{1}{\beta} - \frac{1}{\beta_k^*})\right| + \mathbb{E}\left|\sum_{k=1}^{p}(\frac{1}{\beta_k^*} - \frac{1}{\beta_k})\right| + \mathbb{E}\left|\sum_{k=1}^{p}\beta_k^{-1}\right| \\
\le\ & 2\sum_{k=1}^{p} \mathbb{E}\frac{|\beta_k^* - \beta|}{|\beta||\beta_k^*|} + \sum_{k=1}^{p} \mathbb{E}\frac{|\varepsilon_k^*|}{|\beta_k^*|^2} + \sum_{k=1}^{p} \mathbb{E}\frac{|\varepsilon_k^*|^2}{|\beta_k||\beta_k^*|^2} + p\ \mathbb{E}|m_p(z)| \\
\le\ & \frac{4T}{nv^2}\sum_{k=1}^{p} \mathbb{E}(|\beta_k^*|^{-1}) + \sum_{k=1}^{p} \mathbb{E}\frac{(\ \mathbb{E}(|\varepsilon_k^*|^2|\mathcal{F}^{(k)}))^{1/2}}{|\beta_k^*|^2} + \sum_{k=1}^{p} \mathbb{E}\frac{\mathbb{E}(|\varepsilon_k^*|^2|\mathcal{F}^{(k)})}{v|\beta_k^*|^2} + p\ \mathbb{E}|m_p(z)| \\
\le\ & (\frac{4T}{nv^2} + (2MT)^{1/2}n^{-1/2}v^{-1})\sum_{k=1}^{p} \mathbb{E}(|\beta_k^*|^{-1}) + \frac{2MT}{nv^2}\sum_{k=1}^{p} \mathbb{E}(|\beta_k^*|^{-1}) + p\ \mathbb{E}|m_p(z)| \\
\le\ & (2T(2+M)v^{-2}n^{-1} + (2MT)^{1/2}n^{-1/2}v^{-1})\sum_{k=1}^{p} \mathbb{E}(|\beta_k^*|^{-1}) + p\ \mathbb{E}|m_p(z) -\ \mathbb{E}(m_p(z))| \\
& +p|\ \mathbb{E}(m_p(z)) - m(z)| + p|m(z)| \\
\le\ & 2[2T(2+M)]^{1/2}v^{-1}n^{-1/2}\sum_{k=1}^{p} \mathbb{E}(|\beta_k^*|^{-1}) + \sqrt{p}v^{-1} + 2p\Delta v^{-1} + p|m(z)|.
\end{aligned}$$

Since $2[2T(2+M)]^{1/2}v^{-1}n^{-1/2} < 1/3$, we find

$$\sum_{k=1}^{p} \mathbb{E}(|\beta_k^*|^{-1}) \le \frac{3}{2}\left(\sqrt{p}v^{-1} + 2p\Delta v^{-1} + p|m(z)|\right)\ .$$

Notice that for large n, $\frac{1}{2}yn \le p \le \frac{3}{2}yn$. The conclusion follows by taking into account the bounds for $m(z)$ given in Eq. (2.3). ∎

Lemma 3.4 *Let* $z_k =\ \mathbb{E}(tr\mathbf{D}|\mathcal{F}_{k-1}) -\ \mathbb{E}(tr\mathbf{D}|\mathcal{F}_k)$. *Then* $tr\mathbf{D} -\ \mathbb{E}tr\mathbf{D} = \sum_{k=1}^{p} z_k$ *and* (z_k) *is a martingale difference with respect to* $(\mathcal{F}_k)$, $k = p, p-1, \ldots, 0$. *Moreover, we have the following formula for* z_k

$$z_k = \{\ \mathbb{E}(a_k|\mathcal{F}_{k-1}) -\ \mathbb{E}(a_k|\mathcal{F}_k)\} -\ \mathbb{E}(b_k|\mathcal{F}_{k-1})\ ,$$

with

$$a_k = \frac{\varepsilon_k^*(1 + \alpha_k^T\mathbf{D}_k{}^2\alpha_k)}{\beta_k^*\beta_k}\ , \quad b_k = \frac{\alpha_k^T\mathbf{D}_k{}^2\alpha_k - \frac{1}{n}tr[(\mathbf{I} + z\mathbf{D}_k)\mathbf{D}_k]}{\beta_k^*}\ . \tag{3.4}$$

Proof. Since $\mathbb{E}(tr\mathbf{D}_k|\mathcal{F}_{k-1}) =\ \mathbb{E}(tr\mathbf{D}_k|\mathcal{F}_k)$, we have

$$z_k =\ \mathbb{E}[(tr\mathbf{D} - tr\mathbf{D}_k)|\mathcal{F}_{k-1}] -\ \mathbb{E}[(tr\mathbf{D} - tr\mathbf{D}_k)|\mathcal{F}_k].$$

On the other hand,

$$
\begin{aligned}
tr\mathbf{D} - tr\mathbf{D}_k &= -\frac{1+\frac{1}{n}\alpha_k^T\mathbf{D}_k{}^2\alpha_k}{\beta_k} \\
&= -\frac{1+\frac{1}{n}tr[(\mathbf{I}+z\mathbf{D}_k)\mathbf{D}_k]}{\beta_k^*} + \frac{\varepsilon_k^*(1+\alpha_k^T\mathbf{D}_k{}^2\alpha_k)}{\beta_k^*\beta_k} - \frac{\alpha_k^T\mathbf{D}_k{}^2\alpha_k - \frac{1}{n}tr[(\mathbf{I}+z\mathbf{D}_k)\mathbf{D}_k]}{\beta_k^*} \\
&= -\frac{1+\frac{1}{n}tr[(\mathbf{I}+z\mathbf{D}_k)\mathbf{D}_k]}{\beta_k^*} + a_k - b_k\ .
\end{aligned}
$$

The conclusion follows from

$$
\mathbb{E}\left(\left. \frac{1+\frac{1}{n}tr[(\mathbf{I}+z\mathbf{D}_k)\mathbf{D}_k]}{\beta_k^*} \right| \mathcal{F}_{k-1} \right) = \mathbb{E}\left(\left. \frac{1+\frac{1}{n}tr[(\mathbf{I}+z\mathbf{D}_k)\mathbf{D}_k]}{\beta_k^*} \right| \mathcal{F}_{k} \right),
$$

and

$$
\mathbb{E}\left(\alpha_k^T\mathbf{D}_k{}^2\alpha_k \middle| \mathcal{F}^{(k)} \right) = \frac{1}{n}tr[(\mathbf{I}+z\mathbf{D}_k)\mathbf{D}_k]\ . \quad \blacksquare
$$

Lemma 3.5 *For each $\ell > 1/2$ with $\nu_{4\ell} < \infty$, there exist positive constants c_ℓ and L_0 independent of n and v, such that for all n, v satisfying $L_0 n^{-1/2} \le v < 1$,*

$$
\mathbb{E}|m_p(z) - Em_p(z)|^{2\ell} \le c_\ell n^{-2\ell} v^{-4\ell} (\Delta + v_y)^\ell\ .
$$

Proof. In the proof of this lemma, c_ℓ and $c_{\ell,0}$ will be used to denote universal positive constants which may depend on the moments up to order ℓ of underlying variables and may represent different values at different appearance, even in one expression. Recall that we have

$$
m_p(z) - \mathbb{E}m_p(z) = \frac{1}{p}[tr\mathbf{D} - \mathbb{E}tr\mathbf{D}] = \sum_{k=1}^{p} z_k\ ,
$$

where the (z_k) are defined in Lemma 3.4. We have

$$
\begin{aligned}
\mathbb{E}\left(|z_k|^{2\ell} \middle| \mathcal{F}_k \right) &= \mathbb{E}\left\{ \left| [\,\mathbb{E}(a_k|\mathcal{F}_{k-1}) - \mathbb{E}(a_k|\mathcal{F}_k)] - \mathbb{E}(b_k|\mathcal{F}_{k-1}) \right|^{2\ell} \middle| \mathcal{F}_k \right\} \\
&\le 2^{2\ell-1}\,\mathbb{E}\left\{ [\,\mathbb{E}(a_k|\mathcal{F}_{k-1}) - \mathbb{E}(a_k|\mathcal{F}_k)]^{2\ell} + [\,\mathbb{E}(b_k|\mathcal{F}_{k-1})]^{2\ell} \middle| \mathcal{F}_k \right\} \\
&\le 2^{2\ell-1}\,\mathbb{E}\left\{ [\,\mathbb{E}(a_k|\mathcal{F}_{k-1})]^{2\ell} + [\,\mathbb{E}(b_k|\mathcal{F}_{k-1})]^{2\ell} \middle| \mathcal{F}_k \right\} \\
&\le 2^{2\ell-1}\left\{ \mathbb{E}\left((a_k)^{2\ell} \middle| \mathcal{F}_k \right) + \mathbb{E}\left((b_k)^{2\ell} \middle| \mathcal{F}_k \right) \right\}\ .
\end{aligned}
$$

Note that by (2.9) and (2.10), $|a_k| \le v^{-1}\,|\varepsilon_k^*/\beta_k^*|$. Hence by Lemma 3.1

$$
\mathbb{E}\left(|a_k|^{2\ell} \middle| \mathcal{F}^{(k)} \right) \le \frac{1}{v^{2\ell}}\,\mathbb{E}\left(\left| \frac{\varepsilon_k^*}{\beta_k^*} \right|^{2\ell} \middle| \mathcal{F}^{(k)} \right) \le c_{\ell,0} n^{-\ell} v^{-3\ell} |\beta_k^*|^{-\ell}\ .
$$

On the other hand, by Lemma 2.1 and assuming $\ell \ge 1$,

$$
\mathbb{E}\left(|b_k|^{2\ell} \middle| \mathcal{F}^{(k)} \right) \le c_{\ell,0}(n\beta_k^*)^{-2\ell}(\nu_{4\ell} + M^\ell)\,\left[tr(\mathbf{I}+z\mathbf{D}_k)(\mathbf{I}+\overline{z}\overline{\mathbf{D}}_k)\mathbf{D}_k\overline{\mathbf{D}}_k\right]^\ell\ .
$$

Since from (2.8) and (2.12), it holds that

$$|\beta_k^*|^{-1} tr(\mathbf{I} + z\mathbf{D}_k)(\mathbf{I} + \overline{z}\overline{\mathbf{D}}_k)\mathbf{D}_k\overline{\mathbf{D}}_k \le |\beta_k^*|^{-1} tr\mathbf{\Lambda}_k^2 \le nTv^{-3} ,$$

we obtain

$$\mathbb{E}\left\{ |b_k|^{2\ell} \Big| \mathcal{F}_k \right\} \le c_{\ell,0} n^{-\ell} v^{-3\ell}\, \mathbb{E}\left[|\beta_k^*|^{-\ell} \Big| \mathcal{F}_k \right] .$$

Therefore for all $\ell \ge 1$,

$$\begin{aligned} \mathbb{E}\left(|z_k|^{2\ell} \Big| \mathcal{F}_k \right) &\le c_{\ell,0} n^{-\ell} v^{-3\ell}\, \mathbb{E}\left[|\beta_k^*|^{-\ell} \Big| \mathcal{F}_k \right] \\ &\le c_{\ell,0} n^{-\ell} v^{-4\ell+1}\, \mathbb{E}\left[|\beta_k^*|^{-1} \big| \mathcal{F}_k \right] \end{aligned} \tag{3.5}$$

Applying Lemma 3.3 gives for $\ell \ge 1$

$$\sum_{k=1}^{n} \mathbb{E}|z_k|^{2\ell} \le c_{\ell,0} n^{-\ell+1} (\Delta + v_y) v^{-4\ell}. \tag{3.6}$$

Case $\ell = 1$ *:* Since that $\{z_k\}$ is a martingale difference sequence, the above inequality yields

$$\mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^2 = n^{-2} \sum_{k=1}^{p} \mathbb{E}|z_k|^2 \le c_{1,0} n^{-2} (\Delta + v_y) v^{-4}. \tag{3.7}$$

The lemma is proved in this case.

Case $\frac{1}{2} < \ell < 1$ *:* By applying the Burkholder inequality for martingales (see Burkholder (1973)) and using the the concavity of the function x^ℓ, we find

$$\begin{aligned} &\mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^{2\ell} \\ &\le c_\ell p^{-2\ell}\ \mathbb{E}\left(\sum_{k=1}^{p} |z_k|^2 \right)^\ell \le c_\ell n^{-2\ell} \left[\mathbb{E}(\sum_{k=1}^{p} |z_k|^2) \right]^\ell \le c_\ell n^{-2\ell} \left[(\Delta + v_y) v^{-4} \right]^\ell , \end{aligned}$$

where the last step follows from the previous case $\ell = 1$. The lemma is then proved in this case.

Case $\ell > 1$ *:*

We proceed by induction in this general case. First, by another Burkholder inequality for martingales, we have

$$\begin{aligned} \mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^{2\ell} &\le c_\ell p^{-2\ell} \left\{ \sum_{k=1}^{p} \mathbb{E}|z_k|^{2\ell} + \mathbb{E}\left(\sum_{k=1}^{p} \mathbb{E}(|z_k|^2|\mathcal{F}_k) \right)^\ell \right\} \\ &\hat{=} I_1 + I_2. \end{aligned} \tag{3.8}$$

By (3.6)

$$I_1 \le c_{\ell,0} (\Delta + v_y) n^{-3\ell+1} v^{-4\ell}. \tag{3.9}$$

The lemma has been already proved for $\frac{1}{2} < \ell \le 1$. Suppose that the lemma is true for $\ell \le 2^t$. Now, we consider the case where $2^t < \ell \le 2^{t+1}$. Application of (3.5) with $\ell = 1$ gives

$$\sum_{k=1}^{n} \mathbb{E}\left(|z_k|^2 \big| \mathcal{F}_k \right) \le c_{1,0} n^{-1} v^{-3} \sum_{k=1}^{p} \mathbb{E}\left(|\beta_k^*|^{-1} \big| \mathcal{F}_k \right) .$$

Hence,

$$I_2 \le c_{\ell,0}(nv)^{-3\ell}\,\mathbb{E}\left(\sum_{k=1}^{p}\,\mathbb{E}(|\beta_k^*|^{-1}|\mathcal{F}_k)\right)^{\ell} \le c_{\ell,0}n^{-2\ell-1}v^{-3\ell}\sum_{k=1}^{p}\,\mathbb{E}|\beta_k^*|^{-\ell}. \tag{3.10}$$

Notice that if $L_0 > \sqrt{2}$ then $nv^2 > 2$ and that

$$\left||\beta|^{-1} - |\beta_k^*|^{-1}\right| \le |\beta^{-1} - (\beta_k^*)^{-1}| = \frac{|tr\mathbf{D} - tr\mathbf{D}_k|}{p|\beta||\beta_k^*|} \le \frac{1}{pv^2}\min(|\beta|^{-1},\, |\beta_k^*|^{-1})$$

(this comes from (2.5) and $|\beta\beta_k^*|^{-1} \le v^{-1}\min(|\beta|^{-1},\, |\beta_k^*|^{-1})$). This yields

$$|\beta_k^*|^{-1} \le |\beta|^{-1} + p^{-1}v^{-2}|\beta_k^*|^{-1} \le 2|\beta|^{-1}$$

and

$$\begin{aligned}
|p\beta^{-1}| &\le \left|\sum_{k=1}^{p}(\beta_k^*)^{-1}\right| + \sum_{k=1}^{p}|(\beta_k^*)^{-1} - \beta^{-1}| \le \left|\sum_{k=1}^{p}(\beta_k^*)^{-1}\right| + v^{-2}|\beta|^{-1} \\
&\le 2\left|\sum_{k=1}^{p}(\beta_k^*)^{-1}\right| \le 2\left|\sum_{k=1}^{p}((\beta_k^*)^{-1} - \beta_k^{-1})\right| + 2\left|\sum_{k=1}^{p}\beta_k^{-1}\right| \\
&\le 2\sum_{k=1}^{p}\frac{|\varepsilon_k^*|^2}{|\beta_k||\beta_k^*|^2} + 2p|m_p(z)|.
\end{aligned}$$

Therefore, by applying Lemma 3.1 and if we choose $L_0 > (2c_{\ell,0})^{1/\ell}$ so that $c_{\ell,0}n^{-\ell}v^{-2\ell} < 1/2$, we have

$$\begin{aligned}
\sum_{k=1}^{p}\,\mathbb{E}|\beta_k^*|^{-\ell} &\le c_{\ell,0}\left(v^{-\ell}\sum_{k=1}^{p}\,\mathbb{E}\frac{|\varepsilon_k^*|^{2\ell}}{|\beta_k^*|^{2\ell}} + p\,\mathbb{E}|m_p(z)|^{\ell}\right) \\
&\le c_{\ell,0}\left(n^{-\ell}v^{-2\ell}\sum_{k=1}^{p}\,\mathbb{E}|\beta_k^*|^{-\ell} + p\,\mathbb{E}|m_p(z)|^{\ell}\right) \\
&\le 2c_{\ell,0}p\,\mathbb{E}|m_p(z)|^{\ell}.
\end{aligned}$$

From the above inequality and (3.10), we get

$$\begin{aligned}
I_2 &\le c_\ell n^{-2\ell}v^{-3\ell}\,\mathbb{E}|m_p(z)|^{\ell} \\
&\le c_\ell n^{-2\ell}v^{-3\ell}\left[\,\mathbb{E}|m_p(z) - \,\mathbb{E}m_p(z)|^{\ell} + |\,\mathbb{E}m_p(z) - m(z)|^{\ell} + |m(z)|^{\ell}\right] \\
&\le c_\ell n^{-2\ell}v^{-3\ell}\left[\,\mathbb{E}|m_p(z) - \,\mathbb{E}m_p(z)|^{\ell} + (\Delta + v_y)^{\ell}v^{-\ell}\right]\;.
\end{aligned} \tag{3.11}$$

It can be readily checked that the ratio of the upper bound for I_1, Eq. (3.9), over the second term from the last inequality is bounded by a constant (for both cases $0 < y < 1$ and $y = 1$), namely

$$\frac{(\Delta + v_y)n^{-3\ell+1}v^{-4\ell}}{n^{-2\ell}v^{-3\ell}(\Delta + v_y)^{\ell}v^{-\ell}} = \left[\frac{1}{n(\Delta + v_y)}\right]^{\ell-1} \le 1\;,$$

because $nv^2 \geq 1$ and $v \leq 1$. Therefore by (3.8) and (3.11), it follows that

$$\begin{aligned}
&\mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^{2\ell} \\
&\leq c_\ell n^{-2\ell} v^{-3\ell}\, \mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^\ell + c_\ell n^{-2\ell}(\Delta + v_y)^\ell v^{-4\ell}.
\end{aligned} \tag{3.12}$$

Finally by the induction hypothesis, we find

$$\begin{aligned}
&\mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^{2\ell} \\
&\leq c_\ell n^{-2\ell} v^{-3\ell} \left[n^{-\ell} v^{-2\ell} (\Delta + v_y)^{\ell/2} \right] + c_\ell n^{-2\ell}(\Delta + v_y)^\ell v^{-4\ell} \\
&= \left[\left(\frac{1}{n^2 v^2 (\Delta + v_y)} \right)^{\ell/2} + 1 \right] c_\ell n^{-2\ell}(\Delta + v_y)^\ell v^{-4\ell} \\
&\leq 2c_\ell n^{-2\ell}(\Delta + v_y)^\ell v^{-4\ell}\ .
\end{aligned}$$

The proof of Lemma 3.5 is complete. ∎

Remark 3.1. Application of Lemma 3.5 to the case $\ell = 1$ gives that there is some constant $c_1 > 0$ such that

$$\mathbb{E}|tr\mathbf{D} - \mathbb{E}tr\mathbf{D}|^2 \leq c_1(\Delta + v_y)v^{-4}. \tag{3.13}$$

It is also worth noticing that if we substitute $\mathbf{D}$ for any $\mathbf{D}_k$ with $k \leq n$, Lemma 3.5 as well as the above consequence (3.13) are still valid, with slightly different constants c_ℓ's.

4 Proofs

Suppose that G is a function of bounded variation. The Stieltjes transform g of G is defined as

$$g(z) = \int_{-\infty}^{\infty} \frac{1}{x - z}\, dG(x),$$

where $z = u + iv$ and $v > 0$. Our main tool is the following proposition (Bai (1993a)).

Proposition 4.1 *Let G be a distribution function and H be a function of bounded variation satisfying $\int |G(x) - H(x)|\, dx < \infty$. Denote their Stieltjes transforms by $g(z)$ and $h(z)$, respectively. Then*

$$\begin{aligned}
\|G - H\| \leq\ & \frac{1}{\pi(1-\kappa)(2\gamma - 1)} \left[\int_{-A}^{A} |g(z) - h(z)|\, du + \frac{2\pi}{v} \int_{|x|>B} |G(x) - H(x)|\, dx \right. \\
& \left. + \frac{1}{v} \sup_x \int_{|y| \leq 2va} |H(x+y) - H(x)|\, dy \right],
\end{aligned}$$

where the constants $A > B$, γ and a are restricted by

$$\gamma = \frac{1}{\pi} \int_{|u| \leq a} \frac{1}{u^2 + 1} du > \frac{1}{2}, \ \ \textit{and} \ \ \kappa = \frac{4B}{\pi(A - B)(2\gamma - 1)} \in (0, 1).$$

Denote the Stieltjes transform of F_p and F_y by $m_p(z)$ and $m(z)$, respectively. By Proposition 4.1 with $A = 25$, $B = 5$ and Lemma 2.5, we have, for some constant $c > 0$,

$$\| \mathbb{E}F_p - F\| \leq c \left[\int_{-A}^{A} | \mathbb{E}m_p(z) - m(z)| \, du + \frac{1}{v} \int_{|x|>5} | \mathbb{E}F_p(x) - e(x)| \, dx + v_y \right], \tag{4.1}$$

where $e(x) = 1$ for $x > 0$ and $e(x) = 0$ otherwise.

In the sequel, for brevity, c will be an universal constant which is not related to the estimation of the order. Since it is already proved in Bai (1993a) that $\Delta = \| \mathbb{E}F_p - F_y\| = O(n^{-1/4})$, Δ will be treated as of order $n^{-1/4}$.

4.1 Proof of Theorem 1.1

We will estimate the first two terms on the right hand side of (4.1) with various choices of v, subject to $v \simeq n^{-1/2}$. We begin with the the second term. Let λ_p be the largest eigenvalue of $\mathbf{S}_p$ and recall that $b = (1 + \sqrt{y})^2$. Yin, Bai and Krishnaiah [1988] proved that under the conditions **C.1-2-3**, one can find two sequences (η_p) and (m_p) satisfying $\eta_p \to 0$ and $m_p^{-1} \log n \to 0$ such that

$$\mathbb{E}(\lambda_p)^m \leq (b + \eta_p)^{m_p} \, . \tag{4.2}$$

Notice that

$$1 - F_p(x) \leq I_{\{\lambda_p \geq x\}}, \text{ for } x \geq 0 \, . \tag{4.3}$$

Take $B = 5$, we get for all $t > 0$

$$\begin{aligned} & \int_B^{\infty} \mathbb{E}|F_p(x) - F_y(x)| \, dx \\ \leq \quad & \int_B^{\infty} P(\lambda_p \geq x) \, dx \leq \int_B^{\infty} \left(\frac{b + \eta_p}{B} \right)^{m_p} dx = o(n^{-t}) \, . \end{aligned}$$

Thus the second term of the equation (3.1) can be neglected. Therefore what remains is to estimate the order of the first term of (4.1).

By Eq. (3.14) of Bai [1993b],

$$m_p(z) = \int_0^{\infty} \frac{1}{x - z} \, dF_p(x) = \frac{1}{p} tr \mathbf{D} = -\frac{1}{p} \sum_{k=1}^{p} \frac{1}{\beta_k}.$$

Let us define δ_p such that

$$m_p(z) = -\frac{1}{z + y - 1 + yz\mathbb{E}m_p(z)} + \delta_p = -\frac{1}{\mathbb{E}\beta} + \delta_p \, .$$

Since

$$\frac{1}{\beta_k} = \frac{1}{\mathbb{E}\beta} \left(1 - \frac{\varepsilon_k}{\beta_k} \right) ,$$

it is easy to see that

$$\delta_p = \frac{1}{p} \sum_{k=1}^{p} \frac{1}{\mathbb{E}\beta} \frac{\varepsilon_k}{\beta_k} = \frac{1}{(\mathbb{E}\beta)^2} \left(\frac{1}{p} \sum_{k=1}^{p} \varepsilon_k - \frac{1}{p} \sum_{k=1}^{p} \frac{\varepsilon_k^2}{\beta_k} \right) .$$

Now

$$
\begin{aligned}
&| \mathbb{E}\delta_p| \\
&\leq \frac{1}{p| \mathbb{E}\beta|^2}\sum_{k=1}^{p}(| \mathbb{E}\varepsilon_k| + | \mathbb{E}\frac{\varepsilon_k^2}{\beta_k}|) \\
&= \frac{1}{p| \mathbb{E}\beta|^2}\sum_{k=1}^{p}\left[| \mathbb{E}(\varepsilon_k^* + \tilde{\varepsilon}_k) + \pi_k| + \left|\frac{1}{\mathbb{E}\beta}\mathbb{E}\varepsilon_k^2 - \frac{1}{(\mathbb{E}\beta)^2}\mathbb{E}\varepsilon_k^3 + \frac{1}{(\mathbb{E}\beta)^2}\mathbb{E}(\frac{\varepsilon_k^4}{\beta_k})\right|\right] \\
&\leq \frac{1}{p| \mathbb{E}\beta|^2}\left[\sum_{k=1}^{p}| \mathbb{E}(\varepsilon_k^* + \tilde{\varepsilon}_k) + \pi_k| + \sum_{k=1}^{p}\left|\frac{1}{\mathbb{E}\beta}\mathbb{E}\varepsilon_k^2\right| + \sum_{k=1}^{p}\left|\frac{1}{(\mathbb{E}\beta)^2}\mathbb{E}\varepsilon_k^3\right|\right. \\
&\quad \left. + \sum_{k=1}^{p}\left|\frac{1}{(\mathbb{E}\beta)^2}\mathbb{E}(\frac{\varepsilon_k^4}{\beta_k})\right|\right] \\
&= | \mathbb{E}\beta|^{-2}[I_0 + I_1 + I_2 + I_3].
\end{aligned}
$$

We will estimate each of I_i's to obtain a bound on $| \mathbb{E}\delta_p|$ (*cf.* (4.4)). Since that $\mathbb{E}(\varepsilon_k^* + \tilde{\varepsilon}_k) = 0$, by (2.5), we have

$$
I_0 = \frac{1}{p}\sum_{k=1}^{p}|\pi_k| = \frac{1}{pn}\sum_{k=1}^{p}| \mathbb{E}tr\mathbf{D}_k - \mathbb{E}tr\mathbf{D}| \leq 1/(nv) \leq C_v v.
$$

Here and hereafter, the symbol C_v denotes a positive constant which may be made arbitrarily small by choosing $\sqrt{n}v$ large. From Lemma 3.2, Remark 3.1 and noticing that $v \leq v_y$, we have

$$
\begin{aligned}
I_1 &\leq \frac{1}{p| \mathbb{E}\beta|}\sum_{k=1}^{p}\mathbb{E}|\varepsilon_k|^2 = \frac{1}{p| \mathbb{E}\beta|}\sum_{k=1}^{p}(\mathbb{E}|\varepsilon_k^*|^2 + \mathbb{E}|\tilde{\varepsilon}_k|^2 + |\pi_k|^2) \\
&\leq \frac{c}{| \mathbb{E}\beta|}\left(\left[\frac{1}{n} + \frac{\Delta + v_y}{nv^2}\right] + \frac{\Delta + v_y}{n^2v^4} + \frac{1}{n^2v^2}\right) \leq \frac{c(\Delta + v_y)}{| \mathbb{E}\beta|nv^2} \leq \frac{C_v(\Delta + v_y)}{| \mathbb{E}\beta|}, \\
I_2 &= \frac{1}{p| \mathbb{E}\beta|^2}\sum_{k=1}^{p}| \mathbb{E}\varepsilon_k^3| \leq \frac{1}{p| \mathbb{E}\beta|^2}\sum_{k=1}^{p}(\mathbb{E}|\varepsilon_k|^2 + \mathbb{E}|\varepsilon_k|^4).
\end{aligned}
$$

Now

$$
\frac{1}{p}\sum_{k=1}^{p}\mathbb{E}|\varepsilon_k|^4 \leq \frac{27}{p}\sum_{k=1}^{p}(\mathbb{E}|\varepsilon_k^*|^4 + \mathbb{E}|\tilde{\varepsilon}_k|^4 + |\pi_k|^4) \triangleq c(I_{21} + I_{22} + I_{23}).
$$

Since

$$
tr\mathbf{B}_k\overline{\mathbf{B}}_k = tr(\mathbf{I}_{p-1} + z\mathbf{D}_k)(I + \overline{z}\overline{\mathbf{D}}_k) \leq 2(p + |z|^2 tr\mathbf{D}_k\overline{\mathbf{D}}_k),
$$

We have from the proof of Lemma 3.1,

$$
\begin{aligned}
\mathbb{E}|\varepsilon_k^*|^4 &\leq cn^{-2}\left\{1 + n^{-2}\mathbb{E}(tr\mathbf{B}_k\overline{\mathbf{B}}_k)^2\right\} \\
&\leq cn^{-2}\left\{1 + n^{-2}\mathbb{E}(tr\mathbf{D}_k\overline{\mathbf{D}}_k)^2\right\}.
\end{aligned}
$$

Now

$$\begin{aligned}
\mathbb{E}(tr(\mathbf{D}_k\overline{\mathbf{D}_k}))^2 &= v^{-2}\ \mathbb{E}(\mathrm{Im}(tr(\mathbf{D}_k)))^2 \\
&\leq 2v^{-2}[v^{-2} + \ \mathbb{E}(\mathrm{Im}(tr(\mathbf{D})))^2] \\
&= 2v^{-4} + 2p^2v^{-2}\ \mathbb{E}(\mathrm{Im}(m_p(z)))^2 \\
&\leq 2v^{-4} + 4p^2v^{-2}|\ \mathbb{E}m_p(z)|^2 + 4p^2v^{-2}\ \mathbb{E}|m_p(z) -\ \mathbb{E}m_p(z)|^2 \\
&\leq cp^2v^{-4}(\Delta + v_y)^2 + cv^{-6}(\Delta + v_y),
\end{aligned}$$

where the second inequality follows from (2.5) and the last step follows from Lemma 3.5 and $|\ \mathbb{E}m_p(z)| \leq |\ \mathbb{E}m_p(z) - m(z)| + |m(z)| \leq v^{-1}(2\Delta + \alpha_y v_y)$ with $\alpha_y := (1+3\sqrt{y})/[\sqrt{y}(1-y)]$ for $0 < y < 1$ and $\alpha_y := 2$ for $y = 1$. Thus

$$\begin{aligned}
I_{21} &\leq c\left\{n^{-2} + n^{-4}v^{-4} + n^{-2}v^{-4}(\Delta + v_y)^2 + n^{-4}v^{-6}(\Delta + v_y)\right\} \\
&\leq C_v[v_y^2 + \Delta^2]\ .
\end{aligned}$$

Also, considering $\mathbf{D}_k$ instead of $\mathbf{D}$ as in Lemma 3.5 and applying (2.4), one can show that for some L_0 such that for all $L_0 n^{-1/2} \leq v < 1$,

$$I_{22} \leq c(\Delta + v_y)^2 n^{-4} v^{-8} \leq C_v[v_y^2 + \Delta^2].$$

Since $|\pi_k| \leq |z|(nv)^{-1}$, we have $I_{23} \leq |z|^4(nv)^{-4}$, and hence,

$$\begin{aligned}
p^{-1}\sum_{k=1}^{p}\ \mathbb{E}|\varepsilon_k|^4 &\leq c(I_{21} + I_{22} + I_{23}) \\
&\leq c[C_v(\Delta^2 + v_y^2) + C_v(\Delta^2 + v_y^2) + (nv)^{-4}] \\
&\leq C_v(\Delta^2 + v_y^2).
\end{aligned}$$

Consequently, for some constant $C_v > 0$,

$$I_2 \leq \frac{c(\Delta + v_y)}{|\ \mathbb{E}\beta|^2 nv^2} + \frac{C_v}{|\ \mathbb{E}\beta|^2}(v_y^2 + \Delta^2) \leq \frac{C_v(\Delta + v_y)}{|\ \mathbb{E}\beta|^2},$$

and

$$I_3 \leq \frac{1}{pv|\ \mathbb{E}\beta|^2}\sum_{k=1}^{p}\ \mathbb{E}|\varepsilon_k|^4 \leq \frac{C_v}{v|\ \mathbb{E}\beta|^2}(\Delta^2 + v_y^2).$$

Summing up the above results, we obtain

$$\begin{aligned}
|\ \mathbb{E}\delta_p| &\leq \frac{1}{|\ \mathbb{E}\beta|^2}[I_0 + I_1 + I_2 + I_3] \\
&\leq \frac{C_v}{|\ \mathbb{E}\beta|^2}\left[v + \frac{\Delta + v_y}{|\ \mathbb{E}\beta|} + \frac{\Delta + v_y}{|\ \mathbb{E}\beta|^2} + \frac{\Delta^2 + v_y^2}{v|\ \mathbb{E}\beta|^2}\right]. \qquad (4.4)
\end{aligned}$$

On the other hand, by Lemma 2.2 and (4.1), we have

$$\frac{1}{|\ \mathbb{E}\beta|} = |-\ \mathbb{E}\delta_p +\ \mathbb{E}[m_p(z) - m(z)] + m(z)| \leq |\ \mathbb{E}\delta_p| + \frac{2\Delta + \alpha_y v_y}{v}. \qquad (4.5)$$

Note that the estimates (4.4) and (4.5) are valid for all $L_0 n^{-1/2} \le v < 1$. As proved in Bai (1993b (see Eq. (3.39)-(3.40) there), there is a constant c such that for every $v > 0$

$$\int_{-A}^{A} |\ \mathbb{E} m_p(z) - m(z)|\, du < cv$$

provided that $\sup_u |\ \mathbb{E}\delta_n| \le v$ (here and hereafter, $\sup_u$ refers to $\sup_{|u|\le A}$). Thus, if $\sup_u |\ \mathbb{E}\delta_p| \le v$, in view of (4.1), we can find a positive constant c_1 such that

$$\Delta < c_1 v_y. \tag{4.6}$$

Part (i) of Theorem 1.1 :

In this part, $0 < y < 1$ and $v_y = v$. Write $M_0 = (1 + 2c_1 + \alpha_y)$ and select $L > L_0$ such that when $Ln^{-1/2} \le v < 1$, we have

$$C_v^{-1} > M_0^2[1 + (1 + c_1)M_0 + (2 + c_1 + c_1^2)M_0^2]\ .$$

The proof will be complete once we have shown that for all large n and $Ln^{-1/2} \le v < 1$,

$$\sup_u |\ \mathbb{E}\delta_p| \le v. \tag{4.7}$$

It is proved in Bai (1993b) that (4.7) holds for all large n and $c_2 n^{-1/4} \le v < 1$, where $c_2 > 0$ is a constant, and hence $\Delta < c_1 v$. Applying these to (4.5), we have

$$|\ \mathbb{E}\beta|^{-1} \le v + 2\Delta/v + \alpha_y < M_0. \tag{4.8}$$

This means that for all large n and $c_2 n^{-1/4} \le v < 1$, both (4.7) and (4.8) hold. Now letting v decrease to $Ln^{-1/2}$, since $\sup_u |\ \mathbb{E}\delta_p|$ is continuous in v, one of the following cases must hold:

Case 1. $\sup_u |\ \mathbb{E}\delta_p| < v$ is true for all $Ln^{-1/2} \le v < 1$;

Case 2. There is a $v \in [Ln^{-1/2},\ c_2 n^{-1/4})$ such that $\sup_u |\ \mathbb{E}\delta_p| = v$ and $|\ \mathbb{E}\beta|^{-1} \le M_0$;

Case 3. There is a $v \in [Ln^{-1/2},\ c_2 n^{-1/4})$ such that $\sup_u |\ \mathbb{E}\delta_p| < v$ and $|\ \mathbb{E}\beta|^{-1} = M_0$.

The theorem then follows if Case 1 is true. Thus to complete the proof of the theorem, it suffices to show that Cases 2 and 3 are impossible. Note that in either Cases 2 or 3, we have $\Delta < c_1 v$ by (4.6).

If Case 3 happens, then there exist $v_0 \in [Ln^{-1/2},\ c_2 n^{-1/4})$ and u_0, such that $|\ \mathbb{E}\beta(z_0)|^{-1} = M_0$, where $z_0 = u_0 + iv_0$. Then, by (4.5), we have

$$|\ \mathbb{E}\beta(z_0)|^{-1} \le 2c_1 + \alpha_y + v_0 < 2c_1 + \alpha_y + 1 = M_0,$$

which leads to a contradiction to the equality assumption. If Case 2 happens, then there exist $v_0 \in [Ln^{-1/2},\ c_2 n^{-1/4})$ and u_0, such that $|\ \mathbb{E}\delta_p(z_0)| = v_0$, where $z_0 = u_0 + iv_0$. From (4.4) we have

$$|\ \mathbb{E}\delta_p(z_0)| \le v_0 C_v M_0^2[1 + c_1 M_0 + (1 + c_1^2)M_0^2] < v_0.$$

This is also a contradiction to the equality assumption. The proof of Theorem 1.1 is complete for the case $0 < y < 1$.

Part (ii) of Theorem 1.1 :

When $y = 1$, $F_p(x)$ and $F_y(x)$ satisfy the following conditions:

$$F_p(0) = F_y(0) \ , \qquad \int_0^\infty x dF_p(x) = \int_0^\infty x dF_y(x) = 1 \ .$$

Thus $\tilde{F}_p(x) = \int_0^x t dF_p(t)$ and $\tilde{F}_y(x) = \int_0^x t dF_y(t)$ are two distributions and $\tilde{F}_y(x)$ satisfies the Lipschitz condition, *i.e.* there exists a constant $L > 0$ for any x and y such that

$$|\tilde{F}_y(x') - \tilde{F}_y(x)| \le L|x' - x| \tag{4.9}$$

Therefore there is a constant c_1 such that

$$\frac{1}{v} \sup_x \int_{|u| \le 2\tau v} |\tilde{F}_y(x+u) - \tilde{F}_y(x)| \, du \le c_1 v$$

According to the definition of $\tilde{F}_p(x)$ and $\tilde{F}_y(x)$ it follows for any $\mu > 0$ and every $t > 0$ that

$$\begin{aligned} \int_{4+\mu}^\infty |\mathbb{E}\tilde{F}_p(x) - \tilde{F}_y(x)| \, dx &= o(n^{-t}) \\ \int_{4+\mu}^\infty \mathbb{E}|\tilde{F}_p(x) - \tilde{F}_y(x)| \, dx &= o(n^{-t}) \end{aligned}$$

Let $\tilde{m}_p(z)$ and $\tilde{m}(z)$ denote the Stieltjes transform of $\tilde{F}_p(x)$ and $\tilde{F}_y(x)$ respectively , then

$$\tilde{m}_p(z) = 1 + z m_p(z), \qquad \tilde{m}(z) = 1 + z m(z)$$

The proof of the Theorem 1.1, part (i) can be applied to the estimations of $\tilde{\Delta} = \|\tilde{F}_p(x) - \tilde{F}_y(x)\|$ and $\mathbb{E}|\tilde{m}_p(z) - \tilde{m}(z)|^k$. Therefore there is a constant $\tilde{c} > 0$, when $1/2 \ge v \ge \tilde{c} n^{-1/2}$ it is followed that

$$\sup_u \mathbb{E}|z\delta_p| \quad < \quad v, \tag{4.10}$$

$$\mathbb{E}|zm_p(z) - zm(z)| \quad = \quad \mathbb{E}|\tilde{m}_p(z) - \tilde{m}(z)| < v. \tag{4.11}$$

By (4.1) and Lemma 2.5, there is a constant c_2, such that

$$\begin{aligned} \Delta &\le \kappa \int_{|u| \le 25} |\mathbb{E}m_p(z) - m(z)| \, du + c_2\sqrt{v} \\ &= \kappa \int_{|u| \le 25} \frac{|\mathbb{E}zm_p(z) - zm(z)|}{|z|} \, du + c_2\sqrt{v} \\ &\le \kappa v \int_{|u| \le 25} \frac{du}{\sqrt{u^2 + v^2}} + c_2\sqrt{v} \le \kappa v \log \frac{c_3}{v} + c_2\sqrt{v}. \end{aligned}$$

Since $\kappa v \log \frac{c_3}{v} < \sqrt{v}$ when v is small enough, it is followed that

$$\Delta < (c_2 + 1)\sqrt{v}.$$

The proof of Theorem 1.1, part (ii) is complete. ∎.

4.2 Proof of Theorem 1.2

By Chebyshev inequality, it suffices to show that

$$\mathbb{E}||F_p - F_y|| = \begin{cases} O(n^{-\frac{2}{5}}), & \text{for } 0 < y < 1, \\ O(n^{-\frac{2}{9}}), & \text{for } y = 1 \end{cases}$$

Case $0 < y < 1$ **:** From (4.1), it follows that

$$\begin{aligned} &\mathbb{E}||F_p - F_y|| \\ &\leq c\left[\int_{-A}^{A} \mathbb{E}|m_p(z) - m(z)|du + \frac{1}{v}\int_{|x|>B} \mathbb{E}|F_p(x) - e(x)|dx + v\right] \\ &\leq c\left[\int_{-A}^{A} \mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|du + \int_{-A}^{A} |\mathbb{E}m_p(z) - m(z)|du\right. \\ &\quad \left. + \frac{1}{v}\int_{|x|>B} |\mathbb{E}F_p(x) - e(x)|dx + v\right]. \end{aligned}$$

In the above argument, we have used the fact that $\mathbb{E}|F_p(x) - e(x)| = |\mathbb{E}F_p(x) - e(x)|$ for all $|x| > B$.

As in the proof of Theorem 1.1, we have shown that the last three terms on the right hand of the above inequality are of order $O(v)$ for all $Ln^{-1/2} \leq v < 1$. Applying Cauchy-Schwarz inequality and Remark 3.1, and the result $\Delta = O(n^{-1/2})$ proved in Theorem 1.1, we conclude that

$$\begin{aligned} \int_{-A}^{A} \mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|du &\leq \int_{-A}^{A} (\mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^2)^{1/2}du \\ &\leq cn^{-1}v^{-3/2} \leq v, \end{aligned}$$

for some positive constant c and all $cn^{-2/5} \leq v < 1$. The proof of Theorem 1.2 in this case is complete.

Case $y = 1$ **:** Similarly we have for all $Ln^{-1/2} \leq v < 1$,

$$\mathbb{E}||F_p - F_y|| \leq c\left[\int_{-A}^{A} \mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|du + \sqrt{v}\right] .$$

Applying Cauchy-Schwarz inequality and Remark 3.1, and the result $\Delta = O(n^{-1/4})$ proved in Theorem 1.1, we conclude that

$$\begin{aligned} \int_{-A}^{A} \mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|du &\leq \int_{-A}^{A} (\mathbb{E}|m_p(z) - \mathbb{E}m_p(z)|^2)^{1/2}du \\ &\leq cn^{-1}v^{-2}v^{1/4} = cn^{-1}v^{-7/4} \leq v^{1/2}, \end{aligned}$$

for some positive constant c and all $cn^{-4/9} \leq v < 1$. The proof of Theorem 1.2 in this case is complete. ∎

4.3 Proof of Theorem 1.3

By Proposition 4.1 we have

$$\begin{aligned} \|F_p - F_y\| \quad \le \quad & c\left[\int_{-A}^{A} |m_p(z) - \mathbb{E}m_p(z)|du + \int_{-A}^{A} |\mathbb{E}m_p(z) - m(z)|du \right. \\ & \left. + \frac{1}{v}\int_{|x|\ge B} |F_p(x) - e(x)|dx + v_y\right]. \end{aligned} \tag{4.12}$$

Yin, Bai and Krishnaiah (1988) has proved that under the assumption of Theorem 1.3, with probability one, for all large n, S_p has no eigenvalues larger than B or less than $-B$ (recall that $B = 5$). Thus, with probability one, for all large n,

$$\int_{|x|\ge B} |F_p(x) - e(x)|\, dx = 0.$$

Moreover in the proof of Theorem 1.1, we have proved that the second term on the right hand of (4.12) has order $O(v_y)$ for all $Ln^{-1/2} \le v < 1$.

Case $0 < y < 1$ **:** Recall that in this case, $v_y = v$. To complete the proof of Theorem 1.3, set $v = \varepsilon n^{-2/5+\eta}$ with some $\varepsilon > 0$. We will show that

$$v^{-1}\int_{-A}^{A} |m_p(z) - \mathbb{E}m_p(z)|\, du \to 0 \qquad a.s. \tag{4.13}$$

Now, applying Lemma 3.5, we obtain for each $\xi > 0$,

$$\begin{aligned} & P\left(\int_{-A}^{A} |m_p(z) - \mathbb{E}m_p(z)|\, du \ge \xi v\right) \\ \le \quad & (v\xi)^{-2k}(2A)^{2k-1}\int_{-A}^{A} \mathbb{E}\,|m_p(z) - \mathbb{E}m_p(z)|^{2k}\, du \\ \le \quad & \xi^{-2k}(2A)^{2k}\left[c_k\left(n^{-2}v^{-5}\right)^k\right] \\ \le \quad & c_k'(\varepsilon\xi)^{-2k}n^{-5\eta k}. \end{aligned}$$

The right hand side of the above inequality is summable by choosing k such that $5\eta k > 1$. Thus, (4.13) is proved and the proof of Theorem 1.3 is complete in this case.

Case $y = 1$ **:** The proof in this case is similar with $v_y = \sqrt{v}$. By taking $v = \varepsilon n^{-4/9+\eta}$ with some $\varepsilon > 0$, we have

$$v^{-1/2}\int_{-A}^{A} |m_p(z) - \mathbb{E}m_p(z)|\, du \to 0 \qquad a.s. \tag{4.14}$$

∎

Acknowledgments. J. F. Yao thanks the National University of Singapore for the support during his visit to its Department of Statistics and Applied Probability.

References

[1] Bai, Z.D. and Yin, Y. Q. (1986) . Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang. *Probab. Th. Rel. Field*, **73**, 555-569.

[2] Bai, Z. D., Yin, Y. Q. and Krishnaiah, P.K. (1987). On the limiting empirical distribution function of the eigenvalues of a multivariate F-matrix. *The Probability Theory and Its Applications*, **32** , 490-500.

[3] Bai, Z. D. (1993a). Convergence rate of expected spectral distributions of large random matrices, Part I. Wigner matrices. *Ann. Probab.* **21**, 625-648.

[4] Bai, Z. D. (1993b). Convergence rate of expected spectral distributions of large random matrices, Part II. Sample covariance matrices. *Ann. Probab.* **21**, 6649-672.

[5] Bai, Z. D. (1998). Methodologies in spectral analysis of large dimensional random matrices. A review. *Statistica Sinica* **9**, 611-677

[6] Bai, Z. D., Miao, Baiqi and Tsay, Jhishen. (1997a). A note on the convergence rate of the spectral distribution of large random matrices, *Stat. & Probab. Let.* **34**,95-101.

[7] Bai, Z. D., Miao, Baiqi and Tsay, Jhishen. (1997b). Convergence rate of the spectral distribution of large Wigner matrices, (in preparation).

[8] Bai, Z. D., Miao, Baiqi and Tsay, Jhishen. (1998). Remarks on the convergence rate of the spectral distribution of Wigner matrices, *Ann. Appl. Probab.* (to appear)

[9] 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**, No.1, 316-345.

[10] Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mat. Sb.*, **72**, 507-536.

[11] Wigner, E. P. (1955). Characteristic vectors bordered matrices with infinite dimensions. *Ann. of Math.*, **62**, 548-564.

[12] Wigner, E. P. (1958). On the distributions of the roots of certain symmetric matrices. *Ann. Math.*, **67**, 325-327.

[13] Yin, Y. Q., Bai, Z.D. and Krishnaiah, P. R. (1988). On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probab. Theory and Related Fields*, **78**, 509-531.