The rate of convergence of spectra of sample covariance matrices

F. Götze
Faculty of Mathematics
University of Bielefeld
Germany

A. Tikhomirov[1]
Faculty of Mathematics and Mechanics
Sankt-Peterburg State University
S.-Peterburg, Russia

December 24, 2007

Abstract

It is shown that the Kolmogorov distance between the spectral distribution function of a random covariance matrix $\frac{1}{p}XX^T$, where X is a $n \times p$ matrix with independent entries and the distribution function of the Marchenko-Pastur law is of order $O(n^{-1/2})$. The bounds hold *uniformly* for any p, including $\frac{p}{n}$ equal or close to 1.

1 Introduction

Let $X_{ij}, 1 \le i \le p, 1 \le j \le n$, be independent random variables with $\mathbf{E}\,X_{ij} = 0$ and $\mathbf{E}\,X_{ij}^2 = 1$ and $\mathbf{X}_p = \Big(X_{ij}\Big)_{\{1\le i\le p, 1\le j\le n\}}$. Denote by $\lambda_1 \le \ldots \le \lambda_p$ the eigenvalues of the symmetric matrix

$$\mathbf{W} := \mathbf{W}_p := \frac{1}{n}\mathbf{X}_p\mathbf{X}_p^T$$

and define its empirical distribution by

$$F_p(x) = \frac{1}{p}\sum_{k=1}^{p} I_{\{\lambda_k \le x\}},$$

where $I_{\{B\}}$ denotes the indicator of an event B. We shall investigate the rate of convergence of the expected spectral distribution $\mathbf{E}\,F_p(x)$ as well as $F_p(x)$ to the Marchenko-Pastur distribution function $F_y(x)$ with density

$$f_y(x) = \frac{1}{2xy\pi}\sqrt{(b-x)(x-a)}I_{\{[a,b]\}}(x) + I_{\{[1,\infty)\}}(y)(1-y^{-1})\delta(x),$$

[1]Partially supported by RFBF grant N 07-01-00583-a, by RF grant of the leading scientific schools NSh-4222.2006.1. Partially supported by CRC 701 “Spectral Structures and Topological Methods in Mathematics”, Bielefeld

where $y \in (0, \infty)$ and $a = (1 - \sqrt{y})^2$, $b = (1 + \sqrt{y})^2$. Here we denote by $\delta(x)$ the Dirac delta-function and by $I_{\{[a,b]\}}(x)$ the indicator function of the interval $[a, b]$. As in Marchenko and Pastur [9] and Pastur [11] assume that X_{ij}, $i, j \geq 1$, are independent identically distributed random variables such that

$$\mathbf{E}\, X_{ij} = 0, \qquad \mathbf{E}\, X_{ij}^2 = 1 \quad \text{and} \qquad \mathbf{E}\, |X_{ij}|^4 < \infty, \qquad \text{for all } i, j.$$

Then $\mathbf{E}\, F_p \to F_y$ and $F_p \to F_y$ in probability, where
$y = \lim_{n\to\infty} y_p := \lim_{n\to\infty}(\frac{p}{n}) \in (0, \infty)$.

Let $y := y_p := p/n$. We introduce the following distance between the distributions $\mathbf{E}\, F_p(x)$ and $F_y(x)$

$$\Delta_p := \sup_x |\mathbf{E}\, F_p(x) - F_y(x)|$$

as well as another distance between the distributions $F_p(x)$ and $F_y(x)$

$$\Delta_p^* := \mathbf{E} \sup_x |F_p(x) - F_y(x)|.$$

We shall use the notation $\xi_n = O_P(a_n)$ if, for any $\varepsilon > 0$, there exists an $L > 0$ such that $\Pr\{|\xi_n| \geq L a_n\} \leq \varepsilon$. Note that, for any $L > 0$,

$$\Pr\{\sup_x |F_p(x) - F_y(x)| \geq L\} \leq \frac{\Delta_p^*}{L}.$$

Hence bounds for Δ_p^* provide bounds for the rate of convergence in probability of the quantity $\sup_x |F_p(x) - F_y(x)|$ to zero. Using our techniques it is straightforward though technical to prove that the rate of almost sure convergence is at least $O(n^{-1/2+\epsilon})$, for any $\epsilon > 0$. In view of the length of the proofs for the results stated above we refrain from including those details in this paper as well.

Bai [1] proved that $\Delta_p = O(n^{-\frac{1}{4}})$, assuming $\mathbf{E}\, X_{ij} = 0$, $\mathbf{E}\, X_{ij}^2 = 1$, $\sup_n \sup_{i,j} \mathbf{E}\, X_{ij}^4 \mathbf{I}_{\{|X_{ij}|>M\}} \to 0$, as $M \to \infty$, and

$$y \in (\theta, \Theta) \text{ such that } 0 < \theta < \Theta < 1 \text{ or } 1 < \theta < \Theta < \infty.$$

If y is close to 1 the limit density and the Stieltjes transform of the limit density have a singularity. In this case the investigation of the rate of convergence is more difficult. Bai [1] has shown that, if $0 < \theta \leq y_p \leq \Theta < \infty$, $\Delta_p = O(n^{-\frac{5}{48}})$. Recently Bai et al. [2] have shown for y_p equal to 1 or asymptotically near 1 that $\Delta_p = O(n^{-\frac{1}{8}})$ (see also [3]). It is clear that the case $y_p \approx 1$ requires different techniques. Results of the authors [4] show that for Gaussian r.v. X_{ij} actually the rate $\Delta_p = O(n^{-1})$ is the correct rate of approximation including the case $y = 1$.

By C (with an index or without it) we shall denote generic absolute constants, whereas $C(\,\cdot\,,\,\cdot\,)$ will denote positive constants depending on arguments. Introduce the notation, for $k \geq 1$,

$$M_k := M_k^{(n)} := \sup_{1\leq j,k\leq n} \mathbf{E}\, |X_{jk}|^k.$$

Our main results are the following

Theorem 1.1. *Let $1 \geq y > \theta > 0$, for some positive constant θ. Assume that $\mathbf{E}\,X_{jk} = 0$, $\mathbf{E}\,|X_{jk}|^2 = 1$, and*

$$M_4 := \sup_{1\leq j,k\leq n} \mathbf{E}\,|X_{jk}|^4 < \infty. \tag{1.1}$$

Then there exists a positive constant $C(\theta) > 0$ depending on θ such that

$$\Delta_p \leq C(\theta)\, M_4^{\frac{1}{2}} n^{-1/2}.$$

Theorem 1.2. *Let $1 \geq y > \theta > 0$, for some positive constant θ. Assume that X_{ij} $\mathbf{E}\,X_{jk} = 0$, $\mathbf{E}\,|X_{jk}|^2 = 1$, and*

$$M_{12} := \sup_{1\leq j,k\leq n} \mathbf{E}\,|X_{jk}|^{12} < \infty.$$

Then there exists a positive constant $C(\theta) > 0$ depending on θ such that

$$\Delta_p^* = \mathbf{E}\,\sup_x |F_p(x) - G(x)| \leq C(\theta) M_{12}^{\frac{1}{6}}\, n^{-1/2}.$$

We shall prove the same result for the following class of sparse matrices. Let ε_{jk}, $j = 1, \ldots, n$, $k = 1, \ldots, p$, denote Bernoulli random variables which are independent in aggregate and independent of (X_{jk}) with $p_n := \Pr\{\varepsilon_{jk} = 1\}$. Consider the matrix $\mathbf{X}^{(\varepsilon)} = \frac{1}{\sqrt{np_n}}(\varepsilon_{jk}X_{jk})$. Let $\lambda_1^{(\varepsilon)}, \ldots, \lambda_p^{(\varepsilon)}$ denote the (complex) eigenvalues of the matrix $\mathbf{X}^{(\varepsilon)}$ and denote by $F_p^{(\varepsilon)}(x)$ the empirical spectral distribution function of the matrix $\mathbf{X}^{(\varepsilon)}$, i. e.

$$F_p^{(\varepsilon)}(x) := \frac{1}{p}\sum_{j=1}^{p} I_{\{\lambda_j^{(\varepsilon)} \leq x,\}}. \tag{1.2}$$

Theorem 1.3. *Let X_{jk} be independent random variables with*

$$\mathbf{E}\,X_{jk} = 0, \qquad \mathbf{E}\,|X_{jk}|^2 = 1, \qquad \text{and} \quad \mathbf{E}\,|X_{jk}|^4.$$

Assume that $np_n \to \infty$ as $n \to \infty$ Then

$$\Delta_n^{(\varepsilon)} := \sup_x |\mathbf{E}\,F_p^{(\varepsilon)}(x) - F_p(x)| \leq C M_4^{1/2}(np_n)^{-\frac{1}{2}}. \tag{1.3}$$

.

We have developed a new approach to the investigation of convergence of spectra of sample covariance matrices based on the so-called Hadamar matrices. Note that our approach allows us to obtain a bound of the rate of convergence to the Marchenko-Pastur distribution uniformly in $1 \geq y \geq \theta$ (including $y = 1$). In this paper we give the proof of Theorem 1.1 only. To prove Theorem 1.2 and 1.3 it is enough to repeat the proof of Theorem 1.2 and Corollary 1.3 in [5] with inessential changes.

2 Inequalities for the distance between distributions via Stieltjes transforms.

We define the Stieltjes transform $s(z)$ of a random variables ξ with the distribution function $F(x)$ (the Stieltjes transform $s(z)$ of distribution function $F(x)$)

$$s(z) := \mathbf{E}\,\frac{1}{\xi - z} = \int_{-\infty}^{\infty} \frac{1}{x-z} d\,F(x), \quad z = u + iv, \quad v > 0.$$

Lemma 2.1. *Let F and G be a distribution functions such that*

$$\int_{-\infty}^{\infty} |F(x) - G(x)|\,dx < \infty. \tag{2.1}$$

Denote their Stieltjes transforms by $s(z)$ and $t(z)$ respectively. Assume that the distribution $G(x)$ has support contained in the bounded interval $I = [a, b]$. Assume that there exists a positive constant c_g such that

$$\sup_x \frac{d}{dx} G(x) \le c_g. \tag{2.2}$$

Denote their Stieltjes transforms by $s(z)$ and $t(z)$ respectively. Let $v > 0$. Then there exist some constants $C_1(c_g)$, $C_2(c_g)$, $C_3(c_g)$ depending only on c_g, such that

$$\Delta(F, G) := \sup_x |F(x) - G(x)| \tag{2.3}$$

$$\le C_1 \sup_{x \in I} |\mathrm{Im}\Big(\int_{-\infty}^{x} (s(z) - s_y(z))\,du\Big)| + C_2\, v, \tag{2.4}$$

where $z = u + iv$.

A proof of Lemma 2.1 in Götze, Tikhomirov [5], .

Corollary 2.2. *The following inequality holds, for any $0 < v < V$,*

$$\Delta(F, G) \le C_1 \int_{-\infty}^{\infty} |(s(u + iV) - t(u + iV))\,|du + C_2\, v \tag{2.5}$$

$$+ C_1 \sup_{x \in I} \left|\mathrm{Re}\, \left\{\int_{v}^{V} (s(x + iu) - t(x + iu))du\right\}\right|. \tag{2.6}$$

3 The main Lemma

Let $\xi \ge 0$ be a positive random variables with distribution function $F(x)$. Let $\varkappa$ be a Rademacher random variable with value ± 1 with porbability $1/2$. Consider a random variable $\widetilde{\xi} := \varkappa\xi$ and denote its distribution function by $\widetilde{F}(x)$. For any x, we have

$$\widetilde{F}(x) = \frac{1}{2}(1 + \mathrm{sgn} x\, F(x^2)) \tag{3.1}$$

This equality implies that

$$\widetilde{p}(x) := \frac{d}{dx}\widetilde{F}(x) = |x|p(x), \tag{3.2}$$

where

$$p(x) = \frac{d}{dx}F(x). \tag{3.3}$$

For the Marchenko–Pastur distribution with parameter $y \in (0,1]$, we have

$$\widetilde{p}_y(x) = |x|p_y(x) = \frac{1}{2\pi y|x|}\sqrt{(x^2-a)(b-x^2)}. \tag{3.4}$$

It is straighforward to check that, for $y \in (0,1]$,

$$\sup_x \widetilde{p}_y(x) \le \frac{1}{\pi\sqrt{y}(1+\sqrt{y})}. \tag{3.5}$$

Note also that the distribution $\widetilde{F}_y(x)$ has a support which is contained in the union of the intervals $[-(1+\sqrt{y}),-(1-\sqrt{y})] \cup [(1-\sqrt{y}),(1+\sqrt{y})]$.

Introduce the following matrix

$$\mathbf{H} := \begin{pmatrix} \mathbf{O} & \mathbf{X} \\ \mathbf{X}^* & \mathbf{O} \end{pmatrix}, \tag{3.6}$$

where $\mathbf{O}$ is the matrix with zero entries only. Consider the resolvent matrix

$$\mathbf{R}(z) = (\mathbf{H} - z\mathbf{I})^{-1}, \tag{3.7}$$

where $\mathbf{I}$ denotes the identity matrix of order $n+p$.

Let $s_y(z)$ denote the Stieltjes transform of the Marchenko–Pastur distribution function with parameter y. Denote by $\widetilde{s}_y(z)$ the Stieltjes transform of the distribution function $\widetilde{F}_y(x)$. It is straighforward to check that

$$\widetilde{s}_y(z) = zs_y(z^2). \tag{3.8}$$

For the Stieltjes transform of the expected spectral distribution function of the sample covariance matrix $s_p(z)$ and its "symmetrization" $\widetilde{s}_p(z)$ we have,

$$\widetilde{s}_p(z) = zs_p(z^2). \tag{3.9}$$

From the equation for $s_y(z)$

$$s_y(z) = -\frac{1}{z+y-1+yzs_y(z)} \tag{3.10}$$

it follows that

$$\widetilde{s}_y(z) = -\frac{1}{z+y\widetilde{s}_y(z)+\frac{y-1}{z}}. \tag{3.11}$$

By inversion of the partitioned matrix formula (see [8], p. 18, Section 0.7.3) , we have

$$\mathbf{R}(z) = \begin{pmatrix} z(\mathbf{X}\mathbf{X}^* - z^2\mathbf{I}_n)^{-1} & \mathbf{X}(\mathbf{X}^*\mathbf{X} - z^2\mathbf{I}_p)^{-1} \\ (\mathbf{X}^*\mathbf{X} - z^2\mathbf{I}_p)^{-1}\mathbf{X}^* & (\mathbf{X}^*\mathbf{X} - z^2\mathbf{I}_p)^{-1} \end{pmatrix} \tag{3.12}$$

This equality implies that

$$\widetilde{s}_p(z) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}\, R_{jj}(z) = \frac{1}{n}\sum_{j=1}^{p} R_{j+n,j+n}(z) + \frac{y-1}{z} \tag{3.13}$$

and

$$\frac{1}{p}\sum_{j=1}^{p} R_{j+n,j+n}(z) = y\frac{1}{n}\sum_{j=1}^{n} R_{j,j}(z) + \frac{1-y}{z}. \tag{3.14}$$

Tfor the readers convenient we state here two Lemmas, which follow from Shur's complement formula (see, for example, [5]). Let $\mathbf{A} = \left(a_{kj}\right)$ denote a matrix of order n and $\mathbf{A}_k$ denote the principal sub-matrix of order $n-1$, i.e. $\mathbf{A}_k$ is obtained from $\mathbf{A}$ by deleting the k-th row and the k-th column. Let $\mathbf{A}^{-1} = \left(a^{jk}\right)$. Let $\mathbf{a}'_k$ denote the vector obtained from the k-th row of $\mathbf{A}$ by deleting the k-th entry and $\mathbf{b}_k$ the vector from the k-th column by deleting the k-th entry. Let $\mathbf{I}$ with subindex or without denote the identity matrix of corresponding size.

Lemma 3.1. *Assume that* $\mathbf{A}$ *and* $\mathbf{A}_k$ *are nonsingular. Then we have*

$$a^{kk} = \frac{1}{a_{kk} - \mathbf{a}'_k\mathbf{A}_k^{-1}\mathbf{b}_k}.$$

Lemma 3.2. *Let* $z = u + iv$, *and* $\mathbf{A}$ *be an* $n \times n$ *symmetric matrix. Then*

$$\begin{aligned} \operatorname{Tr}(\mathbf{A} - z\mathbf{I}_n)^{-1} - \operatorname{Tr}(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-1} &= \frac{1 + \mathbf{a}'_k(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-2}\mathbf{a}_k}{a_{kk} - z - \mathbf{a}'_k(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-1}\mathbf{a}_k} \\ &= (1 + \mathbf{a}'_k(\mathbf{A}_k - z\mathbf{I}_{n-1})^{-2}\mathbf{a}_k)\, a^{kk}. \end{aligned} \tag{3.15}$$

and

$$\left|\operatorname{Tr}\Big(\mathbf{A} - z\mathbf{I}_n\Big)^{-1} - \operatorname{Tr}\Big(\mathbf{A}_k - z\mathbf{I}_{n-1}\Big)^{-1}\right| \le v^{-1}.$$

Applying Lemma 3.1 with $\mathbf{A} = \mathbf{W}$ we may write, for $j = 1, \ldots, n$

$$\begin{aligned} R_{j,j} &= -\frac{1}{z + y\widetilde{s}_p(z) + \frac{y-1}{z} - \varepsilon_j} = -\frac{1}{z + y\widetilde{s}_p(z) + \frac{y-1}{z}} \\ &\quad + \frac{\varepsilon_j}{(z + y\widetilde{s}_p(z) + \frac{y-1}{z})(z + y\widetilde{s}_p(z) + \frac{y-1}{z} - \varepsilon_j)} \\ &= -\frac{1}{z + y\widetilde{s}_p(z) + \frac{y-1}{z}}\,(1 - \varepsilon_j R_{j,j})\,, \end{aligned} \tag{3.16}$$

where

$$\varepsilon_j = \varepsilon_j^{(1)} + \varepsilon_j^{(2)} + \varepsilon_j^{(3)} + \varepsilon_j^{(4)} \tag{3.17}$$

with

$$\varepsilon_j^{(1)} = \frac{1}{p}\sum_{1\le k\ne l\le p} X_{jk}X_{jl}^* R_{k+n,l+n}^{(j)}, \quad \varepsilon_j^{(2)} = \frac{1}{p}\sum_{k=1}^{p}(|X_{j,k}|^2-1)R_{k+n,k+n}^{(j)}$$

$$\varepsilon_j^{(3)} = \frac{1}{p}\sum_{k=1}^{p} R_{k+n,k+n}^{(j)} - \frac{1}{p}\sum_{k=1}^{p} R_{k+n,k+n}, \quad \varepsilon_j^{(4)} = \frac{1}{p}\sum_{k=1}^{p} R_{k+n,k+n} - \frac{1}{p}\mathbf{E}\left(\sum_{k=1}^{p} R_{k+n,k+n}\right).$$

This implies that

$$\widetilde{s}_p(z) = -\frac{1}{z + y\widetilde{s}_p(z) + \frac{y-1}{z}} + \delta_p(z), \tag{3.18}$$

where

$$\delta_p(z) = \frac{1}{n\left(z + y\widetilde{s}_p(z) + \frac{y-1}{z}\right)}\sum_{j=1}^{n}\varepsilon_j R_{jj}. \tag{3.19}$$

Throughout this paper we shall consider $z = u + iv$ with $a \le |u| \le b$ and $0 < v < C$.

The main result of this Section is

Lemma 3.3. *Let*

$$\operatorname{Im}\left\{y\delta_p(z) + z + \frac{y-1}{z}\right\} \ge 0.$$

Then

$$\left|z + \frac{y-1}{z} + ys_p(z)\right| \ge 1.$$

Proof. From representation (3.18) it follows that

$$\operatorname{Im}\left\{ys_p(z) + z + \frac{y-1}{z}\right\} = \frac{\operatorname{Im}\left\{ys_p(z) + z + \frac{y-1}{z}\right\}}{|ys_p(z) + z + \frac{y-1}{z}|^2} + \operatorname{Im}\{\delta_p(z) + z + \frac{y-1}{z}\}. \tag{3.20}$$

This equality concludes the proof. □

4 Bounds for $\delta_p(z)$

We start from the simple bound for the $\delta_p(z)$.

Lemma 4.1. *Under the conditions of Theorem 1.1 the following bound holds for* $1 \ge v \ge CM^{1/2}n^{-1/2}$

$$|\delta_p(z)| \le \frac{1}{|z + y\widetilde{s}_p(z) + \frac{y-1}{z}|^2}\frac{C}{nv^4}. \tag{4.1}$$

Proof. Note that

$$|\delta_p(z)| \le \frac{1}{|z + y\widetilde{s}_p(z) + \frac{y-1}{z}|^2}\Big(\frac{1}{n}\sum_{j=1}^{n}|\mathbf{E}\,\varepsilon_j| + \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,\varepsilon_j|^2|R_{j,j}|\Big). \tag{4.2}$$

Using inequalities (4.5), (4.6), (4.14), and (4.15) below and inequality $|R_{j,j}| \le 1/v$, we get

$$\begin{aligned}|\delta_p(z)| &\le \frac{1}{|z + y\widetilde{s}_p(z) + \frac{y-1}{z}|^2}\Big(\frac{1}{nv} + \frac{1}{nv}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_{(j)}|^2 \\ &\le \frac{1}{|z + y\widetilde{s}_p(z) + \frac{y-1}{z}|^2}\Big(\frac{1}{nv} + \frac{C}{nv^3}\Big)\end{aligned} \tag{4.3}$$

Thus the Lemma is proved. □

In this Section we give bounds for remainder term $\delta_p(z)$ in the equation (3.18).We first start with bounds assuming that there exist positive constants a_1, a_2 such that

$$a_1 \le \left|z + \frac{y-1}{z} + ys_p(z)\right| \le a_2. \tag{4.4}$$

Lemma 4.2. *There exists a positive absolute constant C such that, for $v \ge cn^{-1}$ with some other positive absolute constant c,*

$$\mathbf{E}\,|\varepsilon_j^{(1)}|^2 \le \frac{C(1 + |s_p(z)|)}{nv} \tag{4.5}$$

$$\mathbf{E}\,|\varepsilon_j^{(2)}|^2 \le \frac{C(1 + |s_p(z)|)}{nv} \tag{4.6}$$

and

$$\mathbf{E}\,|\varepsilon_j^{(1)}|^4 \le \frac{CM_4^2(1 + |\widetilde{s}_p(z)|)}{n^2v^2}. \tag{4.7}$$

Proof. Consider inequality (4.5). We have

$$\mathbf{E}\,|\varepsilon_j^{(1)}|^2 \le \frac{2}{p^2}\sum_{k,l=1}^{p}\mathbf{E}\,|R_{k,l}^{(j)}|^2 \le \frac{1}{p^2}\mathbf{E}\,\mathrm{Tr}\,\mathbf{R}^{(j)}(\mathbf{R}^{(j)})^* \le \frac{2}{p^2v}\mathbf{E}\,\mathrm{Im}\,\mathrm{Tr}\,\mathbf{R}^{(j)}. \tag{4.8}$$

Applying Lemma 3.2, we get

$$|\mathrm{Tr}\,\mathbf{R} - \mathrm{Tr}\,\mathbf{R}^{(j)}| \le 1/v. \tag{4.9}$$

Note that

$$\frac{1}{2n}\mathbf{E}\,\mathrm{Im}\,\mathrm{Tr}\,\mathbf{R}(z) \le (1+y)|\widetilde{s}_p(z)| + \left|\mathrm{Im}\left\{\frac{1-y}{z}\right\}\right|. \tag{4.10}$$

It is straighforward to check that

$$\left|\operatorname{Im}\left\{\frac{1-y}{z}\right\}\right| \le 1 \tag{4.11}$$

The last inequalities together conclude the proof of inequality (4.5). The proof of inequality (4.6) is similar. Furthermore,

$$\mathbf{E}\,|\varepsilon_j^{(1)}|^4 \le \frac{CM_4^2}{p^4}\mathbf{E}\left(\sum_{k,l=1}^{p}|R_{k,l}^{(j)}|^2\right)^2 \le \frac{CM_4^2}{p^2v^2}\mathbf{E}\left(\frac{1}{p}\operatorname{Im}\operatorname{Tr}\mathbf{R}^{(j)}\right)^2. \tag{4.12}$$

Similar to inequality (4.5) we get

$$\mathbf{E}\,|\varepsilon_j^{(1)}|^4 \le \frac{CM_4^2(1+|\widetilde{s}_p(z)|)^2}{p^2v^2} \tag{4.13}$$

Thus the Lemma is proved. □

Lemma 4.3. *For any $j = 1, \ldots, n$ the following inequality*

$$|\varepsilon_j^{(3)}| \le \frac{1}{nv} \tag{4.14}$$

holds.

Proof. The result follows immediately from Lemma 3.2 with $\mathbf{A} = \mathbf{H}$. □

Lemma 4.4. *The follwoing bound holds for all $v > 0$*

$$\mathbf{E}\,|\varepsilon_j^{(4)}|^2 \le \frac{4}{nv^2}. \tag{4.15}$$

There exist positive constants c and C depending on a_1 and a_2 such that for any $v \ge cn^{-\frac{1}{2}}$

$$\mathbf{E}\,|\varepsilon_j^{(4)}|^2 \le \frac{CM_4(1+|\widetilde{s}_p(z)|)}{n^2v^3} \tag{4.16}$$

and

$$\mathbf{E}\,|\varepsilon_j^{(4)}|^3 \le \frac{CM_4(1+|\widetilde{s}_p(z)|)}{n^{\frac{5}{2}}v^4} \tag{4.17}$$

and

$$\mathbf{E}\,|\varepsilon_j^{(4)}|^4 \le \frac{CM_4(1+|\widetilde{s}_p(z)|)}{n^3v^5}. \tag{4.18}$$

Proof. Note that

$$\varepsilon_j^{(4)} = \frac{1}{p}(\sum_{j=1}^{p} R_{j+n,j+n} - \mathbf{E}\sum_{j=1}^{p} R_{j+n,j+n}) = \frac{1}{p}(\operatorname{Tr}\mathbf{R}(z) - \mathbf{E}\operatorname{Tr}\mathbf{R}(z)) \tag{4.19}$$

Let $\mathbf{E}_k$ denote the conditional expectation given $X_{lm},\ 1 \le l \le k;\ 1 \le m \le p$.

$$\mathbf{E}\,|\varepsilon_j^{(4)}|^2 = \frac{1}{p^2}\sum_{k=1}^{n}\mathbf{E}\,|\gamma_k|^2, \tag{4.20}$$

where

$$\gamma_k = \mathbf{E}_k(\mathrm{Tr}\,\mathbf{R}) - \mathbf{E}_{k-1}(\mathrm{Tr}\,\mathbf{R}). \tag{4.21}$$

Since $\mathbf{E}_k\mathrm{Tr}\,\mathbf{R}^{(k)} = \mathbf{E}_{k-1}\mathrm{Tr}\,\mathbf{R}^{(k)}$ we have

$$\gamma_k = \mathbf{E}_k\sigma_k - \mathbf{E}_{k-1}\sigma_k, \tag{4.22}$$

where

$$\sigma_k = (\mathrm{Tr}\,\mathbf{R} - \mathrm{Tr}\,\mathbf{R}^{(k)}). \tag{4.23}$$

According to Lemma 3.2, we may represent σ_k as follows

$$\sigma_k = \sigma_k^{(1)} + \sigma_k^{(2)} + \sigma_k^{(3)} + \sigma_k^{(4)}, \tag{4.24}$$

where

$$\begin{aligned}
\sigma_k^{(1)} &= \frac{1 + \frac{1}{p}\sum_{r=1}^{n}\sum_{s=1}^{p} X_{kr}\overline{X}_{ks}(\mathbf{R}^{(k)})^2_{rs}}{z + y\widetilde{s}_p(z) + \frac{y-1}{z}} \\
\sigma_k^{(2)} &= \frac{\varepsilon_k\sigma_k}{z + y\widetilde{s}_p(z) + \frac{y-1}{z}} \\
\sigma_k^{(3)} &= \frac{\frac{1}{p}\left(\sum_{r=1}^{n}\sum_{s=1}^{p} X_{kr}\overline{X}_{ks}(\mathbf{R}^{(k)})^2_{rs} - \mathrm{Tr}\,(\mathbf{R}^{(k)})^2\right)}{z + y\widetilde{s}_p(z) + \frac{y-1}{z}}.
\end{aligned}$$

Since

$$\mathbf{E}_k\sigma_k^{(1)} = \mathbf{E}_{k-1}\sigma_k^{(1)}, \tag{4.25}$$

we get

$$\mathbf{E}\,|\gamma_k|^2 \le 2(\mathbf{E}\,|\sigma_k^{(2)}|^2 + \mathbf{E}\,|\sigma_k^{(3)}|^2) \le C(\frac{1}{v^2}\mathbf{E}\,|\varepsilon_k|^2 + \mathbf{E}\,|\sigma_k^{(3)}|^2). \tag{4.26}$$

By definition of ε_k, we have

$$\mathbf{E}\,|\varepsilon_k|^2 \le 4\mathbf{E}\,|\varepsilon_k^{(1)}|^2 + 4\mathbf{E}\,|\varepsilon_k^{(2)}|^2 + 4\mathbf{E}\,|\varepsilon_k^{(3)}|^2 + 4\mathbf{E}\,|\varepsilon_k^{(4)}|^2. \tag{4.27}$$

According to Lemmas 4.2 – 4.4, we have

$$\mathbf{E}\,|\varepsilon_k|^2 \le \frac{C(1 + |\widetilde{s}_p(z)|)}{nv} + 4\mathbf{E}\,|\varepsilon_k^{(4)}|^2. \tag{4.28}$$

Furthermore,

$$\mathbf{E}\,|\sigma_k^{(3)}|^2 \le \frac{C}{n^2v^3}\mathrm{Im}\,\mathrm{Tr}\,\mathbf{R}^{(k)} \le \frac{C(1 + |\widetilde{s}_p(z)|)}{nv^3}. \tag{4.29}$$

Inequalities (4.26), (4.28) and (4.29) together imply that

$$\mathbf{E}\,|\gamma_k|^2 \le \frac{C(1+|\widetilde{s}_p(z)|)}{nv^3} + \frac{C}{v^2}\mathbf{E}\,|\varepsilon_k^{(4)}|^2 \tag{4.30}$$

From the inequalities (4.20) and (4.30) it follwos that

$$\mathbf{E}\,|\varepsilon_k^{(4)}|^2 \le \frac{C(1+|\widetilde{s}_p(z)|)}{n^2v^3} + \frac{C}{nv^2}\mathbf{E}\,|\varepsilon_k^{(4)}|^2. \tag{4.31}$$

For $v \ge cn^{-\frac{1}{2}}$ with some sufficiently small positive absolute constant c, we get

$$\mathbf{E}\,|\varepsilon_k^{(4)}|^2 \le \frac{C(1+|\widetilde{s}_p(z)|)}{n^2v^3}. \tag{4.32}$$

Thus the inequality (4.16) is proved. To prove inequality (4.18) we use the Burkholder inequality for martingales (see Hall and Heyde [7], p.24). We get

$$\mathbf{E}\,|\varepsilon_k^{(4)}|^4 \le \frac{n}{p^4}\sum_{l=1}^{n}\mathbf{E}\,|\gamma_l|^4. \tag{4.33}$$

Using that $|\gamma_l| \le \frac{2}{v}$, we get

$$\mathbf{E}\,|\gamma_l|^4 \le \frac{2}{v^2}\mathbf{E}\,|\gamma_l|^2 \le \frac{CM_4(1+|\widetilde{s}_p(z)|^4)}{nv^5}. \tag{4.34}$$

Inequalities (4.33) and (4.34) together imply that

$$\mathbf{E}\,|\varepsilon_k^{(4)}|^4 \le \frac{CM_4(1+|\widetilde{s}_p(z)|^4)}{n^3v^5}. \tag{4.35}$$

Thus the Lemma is proved. □

Lemma 4.5. *There exist some positive constants c and C such that, for any $1 \ge v \ge cn^{-\frac{1}{2}}$, the following inequality holds*

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|R_{k,k}|^2 \le C. \tag{4.36}$$

Proof. To prove this Lemma we repeat the proof of Lemma 5.4 in [5]. Let

$$U^2 = \frac{1}{n}\sum_{j=1}^{n+p}\mathbf{E}\,|R_{k,k}|^2. \tag{4.37}$$

By equality (3.16), we have

$$U^2 \le C(1+\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_j|^2|R_{j,j}|^2). \tag{4.38}$$

Applying Lemmas 4.2–4.4, we obtain

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_j^{(1)}|^2|R_{j,j}|^2 \le \frac{CM_4}{nv^2}\left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|R_{j,j}|^2\right)^{\frac{1}{2}}. \tag{4.39}$$

Furthermore,

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_j^{(3)}|^2|R_{j,j}|^2 \le \frac{C}{n^2v^4}. \tag{4.40}$$

To bound $\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_j^{(4)}|^2|R_{j,j}|^2$ we use that $\varepsilon_j^{(4)}$ does not depend on j. We write

$$\begin{aligned}\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_j^{(4)}|^2|R_{j,j}|^2 &= \mathbf{E}\,|\varepsilon_1^{(4)}|^2\left(\frac{1}{n}\sum_{j=1}^{n}|R_{j,j}|^2\right)\\ &\le \frac{C}{v}\mathbf{E}\,|\varepsilon_1^{(4)}|^2\frac{1}{n}\mathrm{Im}\,\mathrm{Tr}\,\mathbf{R}(z)\\ &\le \frac{C|\widetilde{s}_p(z)|}{v}\mathbf{E}\,|\varepsilon_1^{(4)}|^2+\frac{C}{v}\mathbf{E}\,|\varepsilon_1^{(4)}|^2|\frac{1}{n}(\mathrm{Tr}\,\mathbf{R}(z)-\mathbf{E}\,\mathrm{Tr}\,\mathbf{R}(z)|\\ &\le \frac{C(1+|\widetilde{s}_p(z)|)}{v}\mathbf{E}\,|\varepsilon_1^{(4)}|^2+\frac{C}{v}\mathbf{E}\,|\varepsilon_1^{(4)}|^3 \end{aligned} \tag{4.41}$$

Inequalities (4.16), (4.18), and (4.41) together imply

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\,|\varepsilon_j^{(4)}|^2|R_{j,j}|^2 \le \frac{CM_4(1+|\widetilde{s}_p(z)|)}{n^2v^4}+\frac{CM_4(1+|\widetilde{s}_p(z)|)}{\sqrt{n^5v^{10}}}. \tag{4.42}$$

Let

$$T := \frac{1}{n}\sum_{j=1}^{n+p}\mathbf{E}\,|\varepsilon_j^{(2)}|^2|R_{j,j}|^2. \tag{4.43}$$

From inequalities (4.38), (4.39), (4.40), and (4.42) it follows that, for $v \ge cn^{-\frac{1}{2}}$,

$$U^2 \le C + \delta U + T. \tag{4.44}$$

Solving this equation with respect to U, we get

$$U^2 \le C + T. \tag{4.45}$$

To bound T we start from the obvious inequality

$$T \le \frac{1}{v^2}\frac{1}{n}\sum_{j=1}^{n+p}\mathbf{E}\,|\varepsilon_j^{(2)}|^2 \le \frac{C}{nv^2}\frac{1}{n}\sum_{j=1}^{n+p}\left(\frac{1}{n}\sum\nolimits^{(j)}\mathbf{E}\,|R_{k,k}^{(j)}|^2\right), \tag{4.46}$$

where $\sum^{(j)}$ denotes the sum over all $k = 1, \dots, n + p$ except $k = j$. Introduce now some integer number $m = m(n)$ depending on n such that
$mv^{-1} \le a_1/4$. Without loss of generality we may assume that $m \le n/2$. Since $|\widetilde{s}_{p-l}(z) - \widetilde{s}_{p-l-1}(z)| \le \frac{1}{n-l}$ we get

$$a_1/2 \le \min_{1\le l\le m} |\widetilde{s}_{p-l}(z) + z + \frac{y-1}{z}| \le \max_{1\le l\le m} |y\widetilde{s}_{p-l}(z) + z + \frac{y-1}{z}| \le \frac{3}{2}a_2.$$

Let $\mathbf{j}^{(r)} = (j_1, \dots, j_r)$ with $1 \le j_1 \ne j_2 \dots \ne j_r \le n$, $r = 1, \dots, m$. Denote by $\mathbf{H}^{(\mathbf{j}^{(r)})}$ the matrix which is obtained from $\mathbf{H}$ by deleting the j_1th, …, j_rth rows and columns, and let

$$\mathbf{R}^{(\mathbf{j}^{(r)})} = \Big(\frac{1}{\sqrt{n-r}}\mathbf{H}^{(\mathbf{j}^{(r)})} - z\mathbf{I}_{n+p-r}\Big)^{-1}.$$

Arguing similar as in inequality (4.46) we get that uniformly for $r = 1, \dots, m-1$, and for $v \ge C_1(a_1, a_2)n^{-\frac{1}{2}}M^{\frac{1}{2}}$

$$\frac{1}{n}\sum_{k=1,\,k\notin\mathbf{j}^{(r)}}^{n} \mathbf{E}\,|R_{k,k}^{(\mathbf{j}^{(r)})}|^2 \le \frac{C_0(a_1,a_2)M}{nv^2}\Big(\frac{1}{n}\sum_{k=1,\,k\notin\mathbf{j}^{(r)}}^{n}\Big(\frac{1}{n}\sum_{j=1,j\notin\mathbf{j}^{(r+1)}}^{n}\mathbf{E}\,|R_{j,j}^{(\mathbf{j}^{(r+1)})}|^2\Big)\Big) + C_0(a_1,a_2). \tag{4.47}$$

Note that the constants $C_0(a_1, a_2)$ and $C_1(a_1, a_2)$ do not depend on $l = 1, \dots, m$.

Applying inequality (4.47) recursively we get for $1 \ge v \ge C_1(a_1, a_2)n^{-1/2}M^{\frac{1}{2}}$,

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|R_{k,k}|^2 \le C_0(a_1,a_2)\sum_{r=0}^{m-1}\Big(\frac{C_0(a_1,a_2)M}{nv^2}\Big)^r + \Big(\frac{C_0(a_1,a_2)M}{nv^2}\Big)^m\Big(\frac{1}{n}\sum_{k=1,\,k\notin\mathbf{j}^{(m-1)}}^{n}\Big(\frac{1}{n}\sum_{j=1,j\notin\mathbf{j}^{(m)}}^{n}\mathbf{E}\,|R_{j,j}^{\mathbf{j}^{(m)}}|^2\Big)\Big) \tag{4.48}$$

Without loss of generality we may assume that

$$\frac{C_0(a_1,a_2)M}{nv^2} \le \frac{1}{2}.$$

Similar to inequality (4.8) we get that

$$\frac{1}{n}\sum_{j=1,\,j\notin\mathbf{j}^{(m)}}^{n}\mathbf{E}\,|R_{\mathbf{j}^{(m)}}(j,j)|^2 \le \mathbf{E}\,\mathrm{Tr}\,|R_{\mathbf{j}^{(m)}}|^2 \le \frac{C_0(a_1,a_2)}{v}. \tag{4.49}$$

The inequalities (4.48) and (4.49) together imply that

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|R(k,k)|^2 \le 2C_0(a_1,a_2) + \frac{1}{2^m}\frac{C}{v}. \tag{4.50}$$

Choosing $m = [C\log n]$ such that $2^{-m} \le Cv$ concludes the proof. $\square$

Lemma 4.6. *Assume that condition (4.4) holds. Then there exist positive constants $C_3(a_1,a_2)$ and $C_4(a_1,a_2)$ such that for $v \ge C_3(a_1,a_2)n^{-1/2}M^{1/2}$ the following inequality holds*

$$|\delta_p(z)| \le \frac{C_4(a_1,a_2)M}{nv}.$$

Proof. The equalities (4.5) and (4.6) imply that

$$|\delta_p(z)| \le \frac{C}{|z+y\widetilde{s}_p(z)+\frac{y-1}{z}|^2}\Big(\frac{1}{p}\sum_{k=1}^{n+p}|\mathbf{E}\,\varepsilon_k| + \frac{1}{p}\sum_{k=1}^{n+p}\mathbf{E}\,|\varepsilon_k|^2|R(j,j)|\Big). \tag{4.51}$$

According to Lemma 4.3 and inequality (4.4) we get

$$\frac{C}{|z+ys_n(z)+\frac{y-1}{z}|^2}\Big(\frac{1}{n}\sum_{k=1}^{n}|\mathbf{E}\,\varepsilon_k|\Big) \le \frac{C}{nva_1^2} \le \frac{C(a_1,a_2)}{nv}. \tag{4.52}$$

Using the representation (3.17), we obtain

$$\frac{C}{|z+ys_n(z)+\frac{y-1}{z}|^2}\Big(\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k|^2|R(j,j)|\Big) \le C(a_1,a_2)\sum_{\nu=1}^{4}\Big(\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k^{(\nu)}|^2|R(j,j)|\Big). \tag{4.53}$$

Similar to inequality (4.48) and by Lemma 3.3 we arrive at

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k^{((1)}|^2|R(k,k| \le \Big(\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k^{(1)}|^4\Big)^{1/2}\Big(\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|R(k,k)|^2\Big)^{1/2} \tag{4.54}$$

$$\le \frac{C(a_1,a_2)M^{\frac{1}{2}}}{nv}. \tag{4.55}$$

By Lemma 4.3, $|\varepsilon_k^{(3)}| \le (nv)^{-1}$ we have

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k^{(3)}|^2|R_{k,k}| \le \frac{1}{n^2v^3} \le \frac{C(a_1,a_2)}{nv}. \tag{4.56}$$

Finally, note that

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k^{(2)}|^2|R(k,k| \le \frac{1}{nv}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k^{(2)}|^2 \le \frac{C(a_1,a_2)M}{nv}\Big(\frac{1}{n}\sum_{j=1,j\ne k}\mathbf{E}\,|R_(j,j)^{(k)}|^2\Big).$$

Applying Lemma 4.5 to the matrix $\mathbf{H}^{(k)}$ we get

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\,|\varepsilon_k^{(2)}|^2|R(k,k| \le \frac{C(a_1,a_2)M}{nv}. \tag{4.57}$$

The inequalities (4.51)–(4.57) together imply that for $1 \ge v \ge C_1(a_1,a_2)n^{-1/2}M^{\frac{1}{2}}$

$$|\delta_n(z)| \le \frac{C(a_1,a_2)M}{nv},$$

which proves Lemma 4.6. ∎

Lemma 4.7. *Assuming the conditions of Theorem 1.1, there exists an absolute positive constant C such that for any $1 \geq v \geq CM^{1/2}n^{-1/2}$ and $u \in [a,b]$, the following inequality holds*

$$\operatorname{Im}\Big\{z + y\widetilde{s}_p(z) + \frac{y-1}{z}\Big\} > 0, \quad z = u + iv. \tag{4.58}$$

Proof of Lemma 4.7. Assume that for $r_n(z) := z + y\delta_p(z) + \frac{y-1}{z}$ the following equality holds

$$\operatorname{Im}\big\{r_n(z)\big\} = 0. \tag{4.59}$$

Denote be $t(z) := y\widetilde{s}_p(z) + \frac{y-1}{z} + z$. Since

$$t(z) = -\frac{y}{t(z)} + r_n(z)$$

this immediately implies that

$$\operatorname{Im} t(z) = -\operatorname{Im}\Big\{\frac{y}{t(z)}\Big\}.$$

Since $\operatorname{Im}\{t(z)\} \geq \operatorname{Im} z = v > 0$ this implies that

$$|t(z)| = \sqrt{y}.$$

Hence condition (4.4) holds with $a_1 = a_2 = \sqrt{y}$ and we have

$$|\delta_p(z)| \leq \frac{CM}{nv}.$$

Then for any $v \geq 2n^{-\frac{1}{2}}\sqrt{CM}$,

$$|\delta_n(z)| \leq \frac{1}{4}v < v,$$

holds. But condition (4.59) implies that

$$|\delta_p(z)| \geq v,$$

which is a contradiction. Hence we conclude that $\operatorname{Im}\{z + y\delta_p(z) + \frac{y-1}{z}\} \neq 0$ in the region $v \geq 2n^{-\frac{1}{2}}\sqrt{CM}$. From Lemma 4.1 it follows for example that, for $v = 1$, $\operatorname{Im}\{r_n(z)\} > 0$. Since the function $\operatorname{Im}\{r_n(z)\}$ is continuous in the region $v \geq C_1 n^{-\frac{1}{2}}\sqrt{M}$ we get that $\operatorname{Im}\{r_n(z)\} > 0$ for $v \geq C_1 n^{-\frac{1}{2}}\sqrt{M}$. This proves Lemma 4.7. □

Proof of Theorem 1.1. Recall that $1 \geq y \geq \theta > 0$. Let $v_0 = \max\{\gamma_0\Delta_p, 2n^{-\frac{1}{2}}C_1M^{\frac{1}{2}}\}$ with a γ_0 such that $1 > \gamma_0 > 0$ to be chosen later. By Lemma 4.7 for any $1 \geq v \geq v_0$ we have

$$\operatorname{Im}\{z + y\delta_p(z) + \frac{y-1}{z}\} > 0.$$

Note that the constant C_1 does not depend on γ_0. In addition we have

$$|\widetilde{s}_p(z)-\widetilde{s}_y(z)| = \Big|\int_{-\infty}^{\infty}\frac{1}{x-z}d\Big(\mathbf{E}\,\widetilde{F}_p(x)-\widetilde{F}_y(x)\Big)\Big| \tag{4.60}$$

$$= \Big|\int_{-\infty}^{\infty}\frac{\mathbf{E}\,\widetilde{F}_p(x)-\widetilde{F}_y(x)}{(x-z)^2}dx\Big| \le \frac{\Delta_p}{v} \le \frac{1}{\gamma_0}. \tag{4.61}$$

This implies that for $z=u+iv$ such that $|u|\in[a,b]$, $1\ge v\ge v_0$, we have

$$|y\widetilde{s}_p(z)+z+\frac{y-1}{z}| \le \frac{1}{\gamma_0}+5. \tag{4.62}$$

From equality (3.18) it follows that

$$s_p(z) = -\frac{1}{2y}\Big(z+\frac{y-1}{z}-y\delta_p(z)-\sqrt{(z+\frac{y-1}{z}+y\delta_p(z))^2-4y}\Big). \tag{4.63}$$

Introduce the function

$$q(z) := -\frac{1}{2y}(z-\sqrt{z^2-4y}). \tag{4.64}$$

Equalities (4.63) and (4.64) together imply that for $v\ge v_0$

$$z+y\widetilde{s}_p(z)+\frac{y-1}{z} = q(\omega+y\delta_p(z)) \tag{4.65}$$

where $\omega := z+\frac{y-1}{z}$. Let $s(z)$ denote the Stieltjes transform of the semicircular law. Then $q(z)=\frac{1}{sqrty}s(z/\sqrt{y})$. This implies in particular that $|q(z)|\le 1/\sqrt{y}$. Since $\operatorname{Im}\{y\delta_p(z)+\omega\}>0$ the equality (4.65) immediately implies that

$$|z+y\widetilde{s}_p(z)+\frac{y-1}{z}| \ge 1/\sqrt{y}, \qquad \text{for} \qquad v\ge v_0 \tag{4.66}$$

From the inequalities (4.65) and (4.66) it follows that condition (4.4) holds with $a_1=1$, and $a_2=\frac{1}{\gamma_0}+5$. The relation (4.65) implies that

$$|\widetilde{s}_p(z)-\widetilde{s}_y(z)| \le \frac{1}{\sqrt{y}}\,|q(\omega)-q(\omega+y\delta_p(z)|\,. \tag{4.67}$$

After a simple calculation we get

$$|\widetilde{s}_p(z)-\widetilde{s}_y(z)| \le \frac{y|\delta_n(z)|}{|\sqrt{(\omega+y\delta_p(z))^2-4y}+\sqrt{\omega^2-4y}}. \tag{4.68}$$

By Lemma 4.6 we obtain for $1\ge v\ge v_0$,

$$|\delta_n(z)| \le \frac{1}{4}v, \tag{4.69}$$

and for $z = u + iv$ such that $u \in I$ we get

$$\min\{\sqrt{|\omega^2 - 4y|}, \sqrt{|(\omega + y\delta_n(z))^2 - 4y|}\} \geq C\sqrt{v}. \tag{4.70}$$

Inequalities (5.61)–(5.63) imply that for $z = u + iv$ such that $u \in I$ and $1 \geq v \geq v_0$

$$|\widetilde{s}_p(z) - \widetilde{s}_y(z)| \leq \frac{C|\delta_p(z)|}{\sqrt{v}}. \tag{4.71}$$

By Lemma 4.6 we have

$$|\delta_p(z)| \leq \frac{C(\gamma_0)M}{nv}. \tag{4.72}$$

From (5.64) and (5.65) it follows that

$$|\widetilde{s}_p(z) - \widetilde{s}(z)| \leq \frac{C(\gamma_0)M}{nv^{\frac{3}{2}}}.$$

Choosing in Corollary 2.3 $V = 1$ and using the inequality (4.29) we get after integrating in u and v

$$\Delta_n \leq C_1 M n^{-1} + C_2 v_0 + C_3(\gamma_0) M n^{-1} v_0^{-1}.$$

Since $v_0 \geq 2n^{-\frac{1}{2}}\sqrt{C_1(\gamma_0)M}$ we get

$$\Delta_n \leq C(\gamma_0) M^{\frac{1}{2}} n^{-\frac{1}{2}} + C_3 v_0$$

Recall that C_2 does not depend on γ_0. If $v_0 = 2n^{-\frac{1}{2}} C_1(\gamma_0) M^{\frac{1}{2}}$ then

$$\Delta_n \leq C(\gamma_0) M^{\frac{1}{2}} n^{-\frac{1}{2}}.$$

We choose $\gamma_0 = \frac{1}{2C_3}$. If $v_0 = \gamma_0 \Delta_n$ then

$$\Delta_n \leq C(\gamma_0) M^{\frac{1}{2}} (1 - C_3\gamma_0)^{-1} n^{-\frac{1}{2}} \leq 2C(\gamma_0) M^{\frac{1}{2}} n^{-\frac{1}{2}}.$$

This completes the proof of Theorem 1.1. □

Acknowledgment. The authors would like to thank Dmitry Timushev for careful reading of the manuscript.

References

[1] Bai, Z. D. *Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices.* Ann. Probab. 21 (1993), no. 2, 649–672.

[2] Bai, Z. D.; Miao, Baiqi; Yao, Jian-Feng. Convergence rates of spectral distributions of large sample covariance matrices. (English summary) SIAM J. Matrix Anal. Appl. 25 (2003), no. 1, 105–127 (electronic). 60F15 (60E10 62H99)

[3] Bai, Z. D. and Silverstein, J. *Spectral analysis of large dimensional random matrices.* Mathematics Monograph Series **2**, Sciences Press, Beijing 2006

[4] Götze, F.; Tikhomirov, A. N. *The rate of convergence for spectra of GUE and LUE matrix ensembles.* Cent. Eur. J. Math. **3**, no. 4, (2005), 666–704

[5] Götze, F. and Tikhomirov, A. N. *Rate of convergence to the semi-circular law.* Probab. Theory Relat. Fields **127** (2003), 228–276

[6] Götze, F.; Tikhomirov, A. N. Rate of convergence in probability to the Marchenko-Pastur law. Bernoulli **10**, no. 3, (2004), 503–548.

[7] P.Hall, C.C. Heyde *Martingale limit theory and its application* Academic Press, 111 Fifth Avenue, New York, NY 10003, (1980) pp. 308

[8] Horn, R., Johnson, Ch. *Matrix analysis.* Cambridge University Press, 1991, pp. 561

[9] Marchenko and V., Pastur, L. *The eigenvalue distribution in some ensembles of random matrices.* Math.USSR Sbornik, **1** (1967), 457-483

[10] Mehta, M. L. *Random Matrices.* 2nd ed., Academic Press, San Diego 1991

[11] Pastur, L.A.*Spectra of random self adjoint operators.* Russian mathematical Surveys **28**, 1, 1–67, 1973

[12] Petrov, V. V. *Sums of independent random variables.* Springer Verlag, Berlin 1975, 345 pp.