On the Rate of Convergence to the Marchenko–Pastur Distribution

F. Götze  
Faculty of Mathematics  
University of Bielefeld  
Germany

A. Tikhomirov  
Department of Mathematics  
Komi Research Center of Ural Branch of RAS,  
Syktyvkar state University  
Syktyvkar, Russia

Abstract
Let $X = (X_{jk})$ denote $n \times p$ random matrix with entries $X_{jk}$, which are independent for $1 \leq j \leq n, 1 \leq k \leq p$. We consider the rate of convergence of empirical spectral distribution function of the matrix $W = \frac{1}{p}XX^*$ to the Marchenko–Pastur law. We assume that $EX_{jk} = 0, EX^2_{jk} = 1$ and that the distributions of the matrix elements $X_{jk}$ have a uniformly sub exponential decay in the sense that there exists a constant $\kappa > 0$ such that for any $1 \leq j \leq n, 1 \leq k \leq p$ and any $t \geq 1$ we have
\[
Pr\{|X_{jk}| > t\} \leq \kappa^{-1} \exp\{-t^{\kappa}\}.
\]
By means of a recursion argument it is shown that the Kolmogorov distance between the empirical spectral distribution of the sample covariance matrix $W$ and the Marchenko–Pastur distribution is of order $O(n^{-1} \log^b n)$ for some positive constant $b > 0$ with high probability.

1 Introduction

For any $n, p \geq 1$, consider a family of independent random variables $\{X_{jk}, 1 \leq j \leq n, 1 \leq k \leq p\}$, defined on some probability space $(\Omega, \mathfrak{M}, \Pr)$. Let $X = (X_{jk})$ be a matrix of order $n \times p$ and let $W = \frac{1}{p}XX^*$. Denote by $\{s_1^2, \ldots, s_n^2\}$ the eigenvalues of the matrix $W$ and introduce the associated spectral distribution function
\[
F_n(x) = \frac{1}{n} \text{card}\{j \leq n : s_j^2 \leq x\}, \quad x \in \mathbb{R}.
\]
Averaging over the random values $X_{ij}(\omega)$, define the expected (non-random) empirical distribution functions $F_n(x) = \mathbb{E} \mathcal{F}_n(x)$. We assume that $p = p(n)$ and $\lim_{n \to \infty} \frac{n}{p} = y \in (0, \infty)$. Without loss of generality we shall assume that $y \in (0, 1]$. Let $G_y(x)$ denote the Marchenko–Pastur distribution function with density $g_y(x) = G'_y(x) = \frac{1}{2\pi x} \sqrt{(x-a)(b-x)}I_{[a,b]}(x)$, where $I_{[a,b]}(x)$ denotes the indicator–function of the interval $[a, b]$, $a = (1-\sqrt{y})^2$, $b = (1+\sqrt{y})^2$. We shall study the rate of convergence $F_n(x)$ to the Marchenko–Pastur law assuming that

$$\Pr\{|X_{jk}| > t\} \leq \nu^{-1} \exp\{-t^\nu\}, \quad \nu > 0$$

for some $\nu > 0$ and any $t \geq 1$. The rate of convergence to the Marchenko–Pastur law has been studied by several authors. In particular, we proved in [11] that the Kolmogorov distance between $\mathcal{F}_n(x)$ and the distribution function $G_y(x)$, $\Delta_n := \sup_x |\mathcal{F}_n(x) - G_y(x)|$ is of order $O_P(n^{-\frac{1}{2}})$. Bai et al. showed in [1] that $\Delta_n := \sup_x |\mathcal{F}_n(x) - G_y(x)| = O(n^{-\frac{1}{2}})$. For the Laguerre Unitary Ensemble Götze and Tikhomirov proved in [5] that $\Delta_n = O(n^{-1})$. Let $y = \frac{a}{p} \in (0, 1]$ in what follows. For any positive constants $\alpha > 0$ and $\nu > 0$ define the quantities

$$l_{n,\alpha} := \log n (\log \log n)\alpha \quad \text{and} \quad \beta_n := (l_{n,\alpha})^{\frac{1}{\nu}} + \frac{1}{\nu}.$$  

The main result of this paper is the following.

**Theorem 1.1.** Let $\mathbb{E}X_{jk} = 0$, $\mathbb{E}X^2_{jk} = 1$ and assume that there exists a constant $\nu > 0$ such that for any $1 \leq j \leq n$ and $1 \leq k \leq p$ and any $t \geq 1$, condition (1.1) holds. Then for any $\alpha > 0$ there exist positive constants $C$ and $c$, depending on $\nu$, $\alpha$ and $y$ such that

$$\Pr\{|\mathcal{F}_n(x) - G_y(x)| > n^{-1/2} \beta_n^2\} \leq C \exp\{-c l_{n,\alpha}\}.$$  

We apply the result of Theorem 1.1 to the investigation of eigenvectors of the matrix $W$. Let $u_j = (u_{j1}, \ldots, u_{jn})^T$ denote the eigenvectors of the matrix $W$ corresponding to the eigenvalues $s_j^2$, $j = 1, \ldots, n$. We prove the following result.

**Theorem 1.2.** Assuming the conditions of Theorem 1.1 for any $\alpha > 0$ there exist constants $C$, $c$, depending on $\nu$, $\alpha$ and $y$ such that

$$\Pr\{\max_{1 \leq j, k \leq n} |u_{jk}|^2 > \frac{\beta_n^4}{n}\} \leq C \exp\{-c l_{n,\alpha}\}$$

and

$$\Pr\{\max_{1 \leq k \leq n} \frac{1}{k} \sum_{\nu=1}^k |u_{j\nu}|^2 - \frac{k}{n} > \frac{\beta_n^2}{\sqrt{n}}\} \leq C \exp\{-c l_{n,\alpha}\}.$$  

We use a relatively short recursion argument based on the approach developed in [6] and [7] and ideas similar to those used in Erdős, Yau and Yin [9], Lemma 3.4.
2 Estimation of Kolmogorov distances via Stieltjes Transforms

To bound $\Delta^*_n$ we shall use an approach developed in Götze and Tikhomirov [6] and [11]. We modify a bound for the Kolmogorov distance between distribution functions based on their Stieltjes transforms obtained in [5], Lemma 2.1. Let $\tilde{G}_y(x)$ denote the distribution function defined by the equality

$$\tilde{G}_y(x) = \frac{1 + \text{sign}(x)G_y(x^2)}{2}, \quad (2.1)$$

Recall that $G_y(x)$ is the Marchenko–Pastur distribution function with parameter $y \in (0, 1]$. The distribution function $\tilde{G}_y(x)$ has a density $\tilde{G}'_y(x) = \frac{1}{2\pi|x|}\sqrt{(x^2 - a^2)(b^2 - x^2)}I\{a \leq |x| \leq b\}$. \quad (2.2)

For $y = 1$ the distribution function $\tilde{G}_y(x)$ is the distribution function of the semi-circular law. Given $\sqrt{y} \geq \varepsilon > 0$ introduce the interval $J_\varepsilon = [1 - \sqrt{y} + \varepsilon, 1 + \sqrt{y} - \varepsilon]$ and $J'_\varepsilon = [1 - \sqrt{y} + \frac{1}{2}\varepsilon, 1 + \sqrt{y} - \frac{1}{2}\varepsilon]$. For any $x$ such that $|x| \in [a, b]$, define $\gamma = \gamma(x) := \min\{|x| - 1 + \sqrt{y}, 1 + \sqrt{y} - |x|\}$. Note that $0 \leq \gamma \leq \sqrt{y}$. For any $x : |x| \in J_\varepsilon$, we have $\gamma \geq \varepsilon$, respectively, for any $x : |x| \in J'_\varepsilon$, we have $\gamma \geq \frac{1}{2}\varepsilon$. For a distribution function $F$ denote by $S_F(z)$ its Stieltjes transform,

$$S_F(z) = \int_{-\infty}^{\infty} \frac{1}{x - z}dF(x).$$

Proposition 2.1. Let $v > 0$ and $H > 0$ and $\varepsilon > 0$ be positive numbers such that

$$\tau = \frac{1}{\pi} \int_{|u| \leq H} \frac{1}{u^2 + 1}du = \frac{3}{4}, \quad (2.3)$$

and

$$2vH \leq \varepsilon^\frac{3}{4}. \quad (2.4)$$

If $\tilde{G}_y$ denotes the distribution function of the symmetrized Marchenko–Pastur law (as in (2.1)), and $F$ denotes any distribution function, there exists some absolute constants $C_1, C_2, C_3$ depending on $y$ only such that

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

$$\leq 2 \sup_{x : |x| \in J_\varepsilon} \left| \text{Im} \int_{-\infty}^{x} (S_F(u + \frac{v}{\sqrt{\gamma}}) - S_{\tilde{G}_y}(u + \frac{v}{\sqrt{\gamma}}))du \right| + C_1v + C_2\varepsilon^\frac{3}{4} \quad (2.5)$$

with $C_1 = \begin{cases} \frac{2H^2\sqrt{3}}{\pi}\sqrt{y(1-\sqrt{y})} & \text{if } 0 < y < 1, \\ \frac{H^2}{\pi} & \text{if } y = 1 \end{cases}$ and $C_2 = \begin{cases} \frac{4}{\pi}\sqrt{y(1-\sqrt{y})} & \text{if } 0 < y < 1, \\ \frac{1}{\pi} & \text{if } y = 1. \end{cases}$
Remark 2.2.

\[ H = \tan \frac{3\pi}{8} = 1 + \sqrt{2}. \quad (2.6) \]

Proof. Without loss of generality we may assume that \(0 < y < 1\). The case \(y = 1\) is considered in \([7]\). The proof of Proposition 2.1 is an adaption of the proof of Proposition 4.1 in \([7]\). We provide it here for completeness.

By Lemma 9.2 in the Appendix, we have

\[ \sup_{x} |F(x) - \tilde{G}_y(x)| \leq \sup_{|x| \in J'} |F(x) - \tilde{G}_y(x)| + \frac{2}{\pi \sqrt{y(1 - y)}} \varepsilon^3. \quad (2.7) \]

Let \(x \in J'_\varepsilon\). Recall that \(\gamma = \min\{|x| - 1 + \sqrt{y}, 1 + \sqrt{y} - |x|\}\). Then, according to condition (2.4) we have \(x + \frac{\nu H}{\sqrt{y}} \in J'_\varepsilon\). Denote by \(v' = \frac{v}{\sqrt{y}}\). For any \(x \in J'_\varepsilon\), we have

\[
\left| \frac{1}{\pi} \operatorname{Im} \left( \int_{-\infty}^{x} (S_F(u + iv') - S_{\tilde{G}_y}(u + iv')) du \right) \right| \\
\geq \frac{1}{\pi} \operatorname{Im} \left( \int_{-\infty}^{x} (S_F(u + iv') - S_{\tilde{G}_y}(u + iv')) du \right) \\
= \frac{1}{\pi} \int_{-\infty}^{x} \left[ \int_{-\infty}^{\infty} \frac{v' d(F(t) - \tilde{G}_y(t))}{(t - u)^2 + v'^2} \right] du \\
= \frac{1}{\pi} \int_{-\infty}^{x} \left[ \int_{-\infty}^{\infty} \frac{2v'(t - u)(F(t) - \tilde{G}_y(t)) dt}{((t - u)^2 + v'^2)^2} \right] \\
= \frac{1}{\pi} \int_{-\infty}^{x} \frac{(F(x) - \tilde{G}_y(x)) \left[ \int_{-\infty}^{x} \frac{2v'(t - u) du}{((t - u)^2 + v'^2)^2} \right]}{t^2 + 1}.
\]

Since \(F\) is non-decreasing, we obtain

\[
\frac{1}{\pi} \int_{|t| \leq H} \frac{(F(x - v't) - \tilde{G}_y(x - v't)) dy}{t^2 + 1} \\
\geq \tau (F(x - v'H) - \tilde{G}_y(x - v'H)) - \frac{1}{\pi} \int_{|t| \leq H} \left| \tilde{G}_y(x - v't) - \tilde{G}_y(x - v'H) \right| dt \\
\geq \tau (F(x - v'H) - \tilde{G}_y(x - v'H)) - \frac{1}{v'\pi} \int_{|t| \leq v'H} \left| \tilde{G}_y(x - t) - \tilde{G}_y(x - v'H) \right| dt.
\]

Moreover, by inequality (2.7), we have

\[
\left| \frac{1}{\pi} \int_{|t| > H} \frac{(F(x - v't) - \tilde{G}_y(x - v't)) dy}{t^2 + 1} \right| \leq (1 - \tau) \Delta(F, \tilde{G}_y).
\]

(2.9)

(2.10)
Let \( \Delta_\varepsilon(F, \tilde{G}_y) = \sup_{x \in \mathbb{I}_\varepsilon} |F(x) - \tilde{G}_y(x)| \) and let \( x_n \in J_\varepsilon \) such that \( F(x_n) - \tilde{G}_y(x_n) \to \Delta_\varepsilon(F, \tilde{G}) \). Then \( x'_n = x_n + v'a \in \mathbb{J}_\varepsilon \). We have

\[
\sup_{x \in \mathbb{J}_\varepsilon} \left| \text{Im} \int_{-\infty}^{x} (S_F(u + iv') - S_{\tilde{G}_y}(u + iv'))du \right| \geq \tau(F(x_n) - \tilde{G}_y(x_n)) - \frac{1}{\pi v'} \sup_{x \in \mathbb{J}_\varepsilon} \sqrt{\gamma} \int_{|t| \leq 2v'H} |\tilde{G}_y(x + t) - \tilde{G}_y(x)|dt - (1 - \tau)\Delta(F, \tilde{G}_y). \tag{2.11}
\]

Furthermore, assume for definiteness that \( t \geq 0 \). Using Lemma 9.1 in the Appendix, we get

\[
|\tilde{G}_y(x + t) - \tilde{G}_y(x)| \leq |t| \sup_{u \in [x, x + t]} \tilde{G}'_y(u) \leq \frac{2|t|\sqrt{\gamma + t}}{\pi \sqrt{1 - \sqrt{y}}} \leq \frac{2|t|\sqrt{\gamma + \varepsilon}}{\pi \sqrt{y(1 - \sqrt{y})}} \tag{2.12}
\]

for \( |t| \leq 2v'H \leq \varepsilon \). This implies after integration

\[
\frac{1}{\pi v'} \sup_{x \in \mathbb{J}_\varepsilon} \sqrt{\gamma} \int_{|t| \leq 2v'H} |\tilde{G}_y(x + t) - \tilde{G}_y(x)|dt \leq \frac{2H^2\sqrt{\gamma}}{\pi^2 \sqrt{y(1 - \sqrt{y})}} \sup_{x \in \mathbb{J}_\varepsilon} \frac{\sqrt{\gamma + \varepsilon}}{\sqrt{\gamma}} \leq \frac{2H^2\sqrt{3\varepsilon}}{\pi^2 \sqrt{y(1 - \sqrt{y})}} \tag{2.13}
\]

We use here that for \( |x| \in J'_\varepsilon \) the inequality \( \gamma \geq \frac{1}{2} \varepsilon \) holds. Inequalities (2.11) and (2.13) together imply

\[
\sup_{x \in \mathbb{J}_\varepsilon} \left| \text{Im} \int_{-\infty}^{x} (S_F(u + iv') - S_{\tilde{G}_y}(u + iv'))du \right| \geq (2\tau - 1)\Delta(F, \tilde{G}_y) - \frac{1}{2} C_1 v - (1 - \tau)C_2 \varepsilon^2, \tag{2.14}
\]

where \( C_1 = \frac{2H^2\sqrt{3}}{\pi^2 \sqrt{y(1 - \sqrt{y})}} \) and \( C_2 = \frac{2}{\pi \sqrt{y(1 - \sqrt{y})}} \). Similar arguments may be used to prove this inequality in case that there is a sequence \( x_n \in J_\varepsilon \) such that \( F(x_n) - G(x_n) \to -\Delta_\varepsilon(F, G) \). In view of (2.14) and \( 2\alpha - 1 = 1/2 \) this completes the proof. \( \Box \)

**Corollary 2.1.** Under the conditions of Proposition 2.1, for any \( V > v \), the following inequality holds

\[
\sup_{x \in \mathbb{J}_\varepsilon} \left| \int_{-\infty}^{x} (\text{Im}(S_F(u + iv') - S_{\tilde{G}_y}(u + iv'))du \right| \leq \int_{-\infty}^{\infty} |S_F(u + iv) - S_{\tilde{G}_y}(u + iV)|du \leq \sup_{x \in \mathbb{J}_\varepsilon} \int_{V}^{\infty} |S_F(x + iv) - S_{\tilde{G}_y}(x + iv)|du \right|. \tag{2.15}
\]
Proof. Let $x : |x| \in \mathcal{J}'_\varepsilon$ be fixed. Let $\gamma = \gamma(x) = \min\{ |x| - 1 + \sqrt{y}, 1 + \sqrt{y} - |x| \}$. Set $z = u + iv'$ with $v' = \frac{u}{\sqrt{\gamma}}$, $v' \leq V$. Since the functions of $S_F(z)$ and $S_{\tilde{G}_y}(z)$ are analytic in the upper half-plane, it is enough to use Cauchy’s theorem. We can write

$$\int_{-\infty}^{x} \text{Im}(S_F(z) - S_{\tilde{G}_y}(z))du = \lim_{L \to \infty} \int_{-L}^{x} (S_F(u + iv') - S_{\tilde{G}_y}(u + iv'))du,$$

(2.16)

for $x \in \mathcal{J}'_\varepsilon$. Since $v' = \frac{u}{\sqrt{\gamma}} \leq \frac{\varepsilon}{2L}$, without loss of generality we may assume that $v' \leq 2$. By Cauchy’s integral formula, we have

$$\int_{-L}^{x} (S_F(z) - S_{\tilde{G}_y}(z))du = \int_{-L}^{x} (S_F(u + iV) - S_{\tilde{G}_y}(u + iV))du$$

$$+ \int_{v'}^{V} (S_F(-L + iu) - S_{\tilde{G}_y}(-L + iu))du$$

$$- \int_{v'}^{V} (S_F(x + iu) - S_{\tilde{G}_y}(x + iu))du. \quad (2.17)$$

Denote by $\xi$ (resp. $\eta$) a random variable with distribution function $F(x)$ (resp. $\tilde{G}_y(x)$). Then we have

$$|S_F(-L + iv')| = \left| \frac{1}{\xi + L - iv'} \right| \leq v'^{-1} \Pr\{|\xi| > L/2\} + \frac{2}{L}. \quad (2.18)$$

Similarly,

$$|S_{\tilde{G}_y}(-L + iv')| \leq v'^{-1} \Pr\{|\eta| > L/2\} + \frac{2}{L}. \quad (2.19)$$

These inequalities imply that

$$\left| \int_{v'}^{V} (S_F(-L + iu) - S_{\tilde{G}_y}(-L + iu))du \right| \to 0 \quad \text{as} \quad L \to \infty, \quad (2.20)$$

which completes the proof. \qed

Combining the results of Proposition 2.1 and Corollary 2.1 we get

**Corollary 2.2.** Under the conditions of Proposition 2.1 the following inequality holds

$$\Delta(F, \tilde{G}_y) \leq 2 \int_{-\infty}^{\infty} |S_F(u + iV) - S_{\tilde{G}_y}(u + iV)|du + C_1v + C_2\varepsilon^{\frac{3}{2}}$$

$$+ 2 \sup_{u \in \mathcal{J}'_\varepsilon} \int_{v'}^{V} |S_F(x + iu) - S_{\tilde{G}_y}(x + iu)|du, \quad (2.21)$$

where $v' = \frac{v}{\sqrt{\gamma}}$ with $\gamma = \min\{ |x| - 1 + \sqrt{y}, 1 + \sqrt{y} - |x| \}$. 

We shall apply Corollary 2.2 to bound the Kolmogorov distance between the empirical spectral distribution $F_n$ and the Marchenko–Pastur distribution $G_y$. We denote the Stieltjes transform of $F_n(x)$ by $m_n(z)$ and the Stieltjes transform of the Marchenko–Pastur law by $s_y(z)$. We shall use a “symmetrization” of the spectral sample covariance matrix as in [6]. Introduce the $(p + n) \times (p + n)$ matrix
\[
V = \frac{1}{\sqrt{p}} \begin{bmatrix} O & X \\ X^* & O \end{bmatrix},
\]
where $O$ denotes a matrix with zero entries. Note that the eigenvalues of the matrix $V$ are $\pm s_1, \ldots, \pm s_n$, and $0$ with multiplicity $p - n$. Let $R = R(z)$ denote the resolvent matrix of $V$ defined by the equality
\[
R = (V - zI_{n+p})^{-1},
\]
for all $z = u + iv$ with $v \neq 0$. Here and in what follows $I_k$ denotes the identity matrix of order $k$. Sometimes we shall omit the sub index in the notation of the identity matrix.

It is well-known that the Stieltjes transform of the Marchenko–Pastur distribution satisfies the equation
\[
y z s_y^2(z) + (y - 1 + z) s_y(z) + 1 = 0
\]
(see, for example, equality (3.9) in [5]). If we consider the Stieltjes transforms $S_y(z)$ of the “symmetrized” Marchenko-Pastur distribution $\tilde{G}_y(x)$ (see formula (2.1)), then it is straightforward to check that $S_y(z) = z s_y(z^2)$ and
\[
y S_y^2(z) + \left(\frac{y - 1}{z} + z\right) S_y(z) + 1 = 0.
\]
(see Section 3 in [6]). Furthermore, for the Stieltjes transform $\tilde{m}_n(z)$ of the “symmetrized” empirical spectral distribution function
\[
\tilde{F}_n(x) = \frac{1 + \text{sign } x F_n(x^2)}{2},
\]
we have
\[
\tilde{m}_n(z) = \frac{1}{n} \sum_{j=1}^n R_{jj} = \frac{1}{n} \sum_{j=n+1}^{n+p} R_{jj} + \frac{1 - y}{yz}.
\]
(see, for instance, Section 3 in [6]). Note that the definition of the symmetrized distribution (2.24) yields
\[
\sup_x |F_n(x) - G_y(x)| = 2 \sup_x |\tilde{F}_n(x) - \tilde{G}_y(x)|.
\]
(2.25)

In what follows we shall consider these symmetrized quantities only and shall omit the symbol "\text{\texttilde}" in the notation of the distribution functions and their Stieltjes transforms. Let $T_j = \{1, \ldots, n\} \setminus \{j\}$. For $j = 1, \ldots, n$, introduce the matrices $V^{(j)}$, obtained from
The rate of convergence to the Marchenko–Pastur distribution

$V$ by deleting the $j$-th row and $j$-th column, and define the corresponding resolvent matrix $R^{(j)}$ by the equality $R^{(j)} = (V^{(j)} - zI_{n+p-1})^{-1}$. Let $m_n^{(j)}(z) = \frac{1}{n-1} \sum_{l \in \mathbb{T}_j} R^{(j)}_{ll}$.

We shall use the representation, for $j = 1, \ldots, n$,

$$R_{jj} = \frac{1}{-z - \frac{1}{p} \sum_{k,l=1}^{p} X_{jk} X_{jl} R^{(j)}_{kk+n,l+n}}$$

(see, for example, Section 3 in [6]). We may rewrite it as follows

$$R_{jj} = -\frac{1}{z + ym_n(z) + \frac{y-1}{z}} + \frac{1}{z + ym_n(z) + \frac{y-1}{z}} \varepsilon_j R_{jj},$$

where $\varepsilon_j = \varepsilon_{j1} + \varepsilon_{j2} + \varepsilon_{j3}$ and

$$\varepsilon_{j1} := \frac{1}{p} \sum_{k=1}^{p} (X_{jk}^2 - 1) P^{(j)}_{k+n,k+n}, \quad \varepsilon_{j2} := \frac{1}{p} \sum_{1 \leq k \neq l \leq p} X_{jk} X_{jl} P^{(j)}_{k+n,l+n},$$

$$\varepsilon_{j3} := \frac{1}{p} \left( \sum_{l=1}^{p} R^{(j)}_{l+n,l+n} - \sum_{l=1}^{p} R^{(j)}_{l+n,l+n} \right).$$

This relation immediately implies the following equations

$$R_{jj} = -\frac{1}{z + ym_n(z) + \frac{y-1}{z}} - \sum_{\nu=1}^{2} \frac{\varepsilon_{j\nu}}{(z + ym_n(z) + \frac{y-1}{z})^2} +$$

$$\sum_{\nu=1}^{2} \frac{1}{(z + ym_n(z) + \frac{y-1}{z})^2} \varepsilon_{j\nu} \varepsilon_j R_{jj} + \frac{1}{z + ym_n(z) + \frac{y-1}{z}} \varepsilon_{j3} R_{jj},$$

and

$$m_n(z) = -\frac{1}{z + ym_n(z) + \frac{y-1}{z}} + \frac{1}{(z + ym_n(z) + \frac{y-1}{z})} \frac{1}{n} \sum_{j=1}^{n} \varepsilon_j R_{jj}$$

$$= -\frac{1}{z + ym_n(z) + \frac{y-1}{z}} - \frac{1}{(z + ym_n(z) + \frac{y-1}{z})^2} \frac{1}{n} \sum_{\nu=1}^{2} \sum_{j=1}^{n} \varepsilon_{j\nu}$$

$$+ \frac{1}{(z + ym_n(z) + \frac{y-1}{z})^2} \frac{1}{n} \sum_{\nu=1}^{2} \sum_{j=1}^{n} \varepsilon_{j\nu} \varepsilon_j R_{jj} + \frac{1}{z + ym_n(z) + \frac{y-1}{z}} \frac{1}{n} \sum_{j=1}^{n} \varepsilon_{j3} R_{jj}.$$  \hfill (2.26)

3 Large deviations I

In the following Lemmas we bound $\varepsilon_{j\nu}$, for $\nu = 1, 2, 3$ and $j = 1, \ldots, n$. 

(2.27)
Lemma 3.1. Under the conditions of Theorem 1.1 we have, for any \( z = u + iv \) with \( u \in \mathbb{R}, v > 0 \), and for any \( j = 1, \ldots, n \),
\[
|\epsilon_{j3}| \leq \frac{y}{nv}.
\]

Proof. It is straightforward to check that
\[
\sum_{l=1}^{p} R_{l+n,l+n} = nm_n(z) - \frac{p-n}{z} = \sum_{l=1}^{p+n} R_{ll} - nm_n(z) \quad (3.1)
\]
and
\[
\sum_{l=1}^{p} R_{l+n,l+n}^{(j)} = \sum_{l=1, l\neq j}^{p+n} R_{ll}^{(j)} - (n-1)m_n^{(j)}(z) \quad (3.2)
\]
Furthermore,
\[
nm_n(z) = \frac{1}{2} \text{Tr} R + \frac{p-n}{2z}, \quad (n-1)m_n^{(j)}(z) = \frac{1}{2} \text{Tr} R^{(j)} + \frac{p-n+1}{2z}. \quad (3.3)
\]
This implies
\[
\sum_{l=1}^{p} R_{l+n,l+n} - \sum_{l=1}^{p} R_{l+n,l+n}^{(j)} = \frac{1}{2} \text{Tr} R - \frac{1}{2} \text{Tr} R^{(j)} - \frac{1}{2z}. \quad (3.4)
\]
The conclusion of Lemma 3.1 follows immediately from the inequality
\[
|\text{Tr} R - \text{Tr} R^{(j)}| \leq v^{-1} \quad \text{and} \quad \frac{1}{z} \leq v^{-1} \quad \text{(see Lemma 4.1 in [4]).}
\]
Remark 3.1. Equalities (3.3) imply that
\[
|m_n(z) - m_n^{(j)}(z)| \leq v^{-1}. \quad (3.5)
\]

Lemma 3.2. Assuming conditions of Theorem 1.1 for any \( \alpha > 0 \) there exist positive constants \( C \) and \( c \), depending on \( \alpha \) and \( \kappa \) only such that for any \( z = u + iv \) with \( u \in \mathbb{R}, v > 0 \), the following inequality holds
\[
\Pr\{|\epsilon_{j1}| > 2p^{-\frac{1}{2} + \frac{\alpha}{2} + \frac{1}{2} (p-1) \sum_{l=1}^{p} |R_{l+n,l+n}^{(j)}|^2}^{\frac{1}{2}}\} \leq C \exp\{-ct_{n,\alpha}\}
\]
Proof. The proof of this Lemma is similar to the proof of Lemma 2.3 in [7]. Introduce, for \( k = 1, \ldots, p \), \( \eta_k = X_{jk}^2 - 1 \), and define
\[
\xi_k = (\eta_k \mathbb{I}\{|X_{jk}| \leq l_{n,\alpha}^{\frac{1}{2}}\}) - E\eta_k \mathbb{I}\{|X_{jk}| \leq l_{n,\alpha}^{\frac{1}{2}}\}R_{k+n,k+n}^{(j)}.
\]
Note that \( E\xi_k = 0 \) and \( |\xi_k| \leq 2l_{n,\alpha}^{\frac{1}{2}} |R_{k+n,k+n}^{(j)}| \). For any \( j = 1, \ldots, n \), introduce the \( \sigma \)-algebra \( \mathfrak{A}^{(j)} \) generated by the random variables \( X_{lk} \) with \( 1 \leq l \neq j \leq n, 1 \leq k \leq p \).
Let $E_j$ and $Pr_j$ denote the conditional expectation and the conditional probability given $\mathcal{M}^{(j)}$. Note that the random variables $X_{jk}$ and $\sigma$-algebra $\mathcal{M}^{(j)}$ are independent.

Applying Lemma 9.3 in the Appendix with $\sigma^2 = 4p l_n^{\frac{1}{2}} \left( p^{-1} \sum_{k=1}^n |R_{k+n,k+n}^{(j)}|^2 \right)$ and $x := l_n^{\frac{1}{4}} \sigma$, we get

$$\Pr\{|\sum_{k=1}^n \xi_k| > x\} = E \Pr_j\{|\sum_{k=1}^n \xi_k| > x\} \leq E \exp\left\{-\frac{x^2}{\sigma^2}\right\} \leq C \exp\{-c l_n^{\frac{1}{2}}\}. \quad (3.6)$$

Furthermore, since $E_j \eta_j = 0$, we have

$$|E_j \eta_k \mathbb{I}\{|X_{jk}| \leq l_n^{\frac{1}{2}} \}| \leq E_j \frac{1}{|\eta_k^2|} \Pr_j \{|X_{jk}| > l_n^{\frac{1}{2}} \} \leq E_j \eta_k^2 \exp\{-\frac{c}{2} l_n^{\frac{1}{2}}\}, \quad (3.7)$$

for $k = 1, \ldots, p$. The last inequality implies that

$$|\sum_{k=1}^n E_j \eta_k \mathbb{I}\{|X_{jk}| \leq l_n^{\frac{1}{2}} \} R_{k+n,k+n}^{(j)}| \leq C \exp\{-\frac{c}{2} l_n^{\frac{1}{2}}\} \left( \frac{1}{n} \sum_{k \in T_j} |R_{k+n,k+n}^{(j)}|^2 \right)^{\frac{1}{2}}. \quad (3.8)$$

The inequalities (3.6) and (3.8) together conclude the proof of Lemma 3.2. Thus the Lemma is proved.

**Corollary 3.3.** Under the conditions of Theorem 1.1 there exist constants $c$ and $C$ depending on $\alpha$ and $z$ only such that for any $z = u + iv$ with $u \in \mathbb{R}$ and with $v > 0$,

$$\Pr\{|\varepsilon_{j1}| > 2\beta_n^{\frac{1}{2}} \sqrt{y(nv)^{-\frac{1}{2}}} \left( y |\text{Im} m_n(z)| + \frac{(1 - y)v}{|z|^2} + \frac{1}{pv} \right)^{\frac{1}{2}}\} \leq C \exp\{-c l_n^{\frac{1}{2}}\}. \quad (3.9)$$

**Proof.** Note that

$$\frac{1}{p} \sum_{k=1}^p |R_{k+n,k+n}^{(j)}|^2 \leq \frac{1}{p} \text{Tr} R^{(j)} R^{(j)*} \leq \frac{1}{pv} \text{Im} \text{Tr} R^{(j)} \leq \frac{2y}{v} |\text{Im} m_n^{(j)}(z)| + \frac{p - n + 1}{p} \text{Im} \frac{1}{z}, \quad (3.10)$$

where $|R^{(j)}|^2 = R^{(j)} R^{(j)*}$. Recall that $n = py$. The result follows now from Lemma 3.2 and inequality (3.5). \qed

**Lemma 3.4.** Under the conditions of Theorem 1.1, for any $j = 1, \ldots, n$ and for any $z = u + iv$ with $u \in \mathbb{R}$ and with $v > 0$, the following inequality holds

$$\Pr\{|\varepsilon_{j2}| > 4\beta_n^{\frac{1}{2}} \beta_p^{-\frac{1}{2}} \left( \frac{1}{p} \sum_{1 \leq k \neq l \leq p} |R_{k+n,l+n}^{(j)}|^2 \right)^{\frac{1}{2}}\} \leq C \exp\{-c l_n^{\frac{1}{2}}\}. \quad (3.11)$$

**Proof.** In order to bound $\varepsilon_{j2}$ we use Proposition 9.1 in Appendix with

$$\xi_k = (X_{jk} \mathbb{I}\{|X_{jk}| \leq l_n^{\frac{1}{2}} \} - E X_{jk} \mathbb{I}\{|X_{jk}| \leq l_n^{\frac{1}{2}} \})/2l_n^{\frac{1}{2}}, \quad (3.12)$$

and
for $k = 1, \ldots, p$. Note that the random variables $X_{jk}$, $k = 1, \ldots, p$ and the matrix $R^{(j)}$ are mutually independent for any fixed $j = 1, \ldots, n$. Moreover, $|\xi_k| \leq 1$. Put $Z := \sum_{1 \leq l \neq k \leq p} \xi_l R_{kl}^{(j)}$. Note that $R^{(j)} = R^{(j)^T}$ and $E_j |Z|^2 = 2 \sum_{1 \leq l \neq k \leq p} |R_{kl}^{(j)}|^2$.

Applying Proposition 9.1 with $\sigma = \sqrt{2(\sum_{1 \leq l \neq k \leq p} |R_{kl}^{(j)}|^2)^{\frac{1}{2}}}$ and $t := l_{n, \alpha}$, we get

$$E_{\Pr} \{ |Z| \geq p^{\frac{1}{2}} l_{n, \alpha} (p^{-1} \sum_{1 \leq l \neq k \leq p} |R_{lk}^{(j)}|^2)^{\frac{1}{2}} \} \leq C \exp \{ -cl_{n, \alpha} \}. \quad (3.13)$$

Furthermore,

$$\Pr \{ \exists j \in [1, \ldots, n] \text{ and } \exists k \in [1, \ldots, p] : |X_{jk}| > \frac{l_{n, \alpha}}{2} \} \leq C \exp \{ -cl_{n, \alpha} \} \quad (3.14)$$

and, for any $k = 1, \ldots, p$,

$$|\mathbb{E} X_{jk} I \{ |X_{jk}| \leq \frac{l_{n, \alpha}}{2} \}| \leq \frac{l_{n, \alpha}}{2} \Pr \{ \exists j \in [1, \ldots, n], k \in [1, \ldots, p] : |X_{jk}| > \frac{l_{n, \alpha}}{2} \} \leq C \exp \{ -cl_{n, \alpha} \}. \quad (3.15)$$

Introduce the random variables

$$\widehat{\xi}_k := X_{jk} I \{ |X_{jk}| \leq \frac{l_{n, \alpha}}{2} \}/2l_{n, \alpha} \quad \text{and} \quad \widehat{Z} := \sum_{l=1}^{p} \sum_{k=1}^{n} \widehat{\xi}_k R_{l+n, k+n}^{(j)}.$$

Note that

$$\Pr \left\{ \sum_{l,k=1}^{p} X_{jk} X_{jl} R_{k+n, l+n}^{(j)} \neq 4l_{n, \alpha} \frac{1}{2} \widehat{Z} \right\} \leq C \exp \{ -cl_{n, \alpha} \}. \quad (3.16)$$

Inequalities (3.13)–(3.16) together imply

$$\Pr \left\{ |\varepsilon_{j2}| > 4\beta_{n, \alpha}^{-1} p^{-\frac{1}{2}} \left( \frac{1}{p} \sum_{1 \leq k \neq l \leq p} |R_{k+n, l+n}^{(j)}|^2 \right)^{\frac{1}{2}} \right\} \leq C \exp \{ -cl_{n, \alpha} \}. \quad (3.17)$$

Thus, Lemma 3.4 is proved.

**Corollary 3.5.** Under the conditions of Theorem 1.1 there exist constants $c$ and $C$, depending on $\kappa$ and $\alpha$ such that for any $z = u + iv$ with $v > 0$

$$\Pr \left\{ |\varepsilon_{j2}| > 4\beta_{n}^{2} \sqrt{y(nv)}^{-\frac{1}{2}} \left( y \text{Im} \left( m_{n}(z) + \frac{1 - y}{|z|^{2}} + \frac{1}{pv} \right)^{\frac{1}{2}} \right) \right\} \leq C \exp \{ -cl_{n, \alpha} \}. \quad (3.18)$$

**Proof.** The result follows from Lemma 3.4 and inequalities (3.10) and (3.5). \qed

Collecting these results, recall the definition

$$\beta_{n} = l_{n, \alpha}^{\frac{1}{2} + \frac{1}{2}}. \quad (3.19)$$
Without loss of generality we may assume that $\beta_n \geq 1$ and $l_{n, \alpha} \geq 1$. Taking these relations into account and applying Lemma 3.1 and Corollaries 3.3 and 3.5 we may write, for $\nu = 1, 2, 3$

$$\Pr\left\{ |\varepsilon_{j, \nu}| > \frac{4\beta_n^2 \sqrt{y}}{\sqrt{nv}} \left( (y \operatorname{Im} m_n(z))^\frac{1}{2} + \frac{(1-y)v}{|z|} + \frac{1}{\sqrt{nv}} \right) \right\} \leq C \exp\{-cl_{n, \alpha}\}. \quad (3.20)$$

Denote by

$$\Omega_n(z, \theta) := \left\{ \omega \in \Omega : |\varepsilon_j| \leq \frac{8\beta_n^2 \sqrt{y} \theta}{\sqrt{nv}} \left( (y \operatorname{Im} m_n(z))^\frac{1}{2} + \frac{1}{\sqrt{nv}} + \frac{(1-y)v}{|z|} \right) \right\}. \quad (3.21)$$

Let

$$v_0 := \frac{dy\beta_n^4}{n} \quad (3.22)$$

with sufficiently large positive constant $d$. Recall that $J_{\varepsilon} := [1 - \sqrt{y} + \varepsilon, 1 + \sqrt{y} - \varepsilon]$, for any $\frac{\sqrt{y}}{2} > \varepsilon > 0$. We introduce the region $\mathcal{D} := \{z = u + iv \in \mathbb{C} : |u| \in J_{\varepsilon}, v_0 < v \leq 4\sqrt{y}\}$ and a sequence $z_l = u_l + v_l$ in $\mathcal{D}$, defined recursively via $u_{l+1} - u_l = \frac{1}{n}$ and $v_{l+1} - v_l = \frac{2}{n} \varepsilon$. We introduce the events

$$\Omega'_n(z_l, \theta) := \cap_{j=1}^{n} \left\{ \omega \in \Omega : |\varepsilon_j| \leq \frac{8\beta_n^2 \sqrt{y} \theta}{\sqrt{nv_l}} \left( (y \operatorname{Im} m_n(z_l))^\frac{1}{2} + \frac{1}{\sqrt{nv_l}} + \frac{(1-y)v_l}{|z_l|} \right) \right\}. \quad (3.23)$$

Using a union bound, we obtain

$$\Pr\{\cap_{z \in \mathcal{D}} \Omega'_n(z_l, \theta)\} \geq 1 - C \exp\{-cl_{n, \alpha}\}. \quad (3.24)$$

Using the resolvent equality $R(z) - R(z') = -(z - z')R(z)R'(z)$, we get

$$|R_{k+n, l+n}^{(j)}(z) - R_{k+n, l+n}^{(j)}(z')| \leq \frac{|z - z'|}{v^2}. \quad (3.25)$$

This inequality and definition of $\varepsilon_j$ together imply

$$\Pr\{|\varepsilon_{j}(z) - \varepsilon_{j}(z')| \leq \frac{C(n + 1)v_0^2}{v_0^2}|z - z'|, \text{ for } z, z' \in \mathcal{D}\} \geq 1 - C \exp\{-cl_{n, \alpha}\}. \quad (3.26)$$

This immediately implies that, for $|z - z_l| \leq \frac{5}{n\varepsilon}$,

$$\Pr\{\cap_{z \in \mathcal{D}} \Omega_n(z, 2)\} \geq \Pr\{\cap_{z \in \mathcal{D}} \Omega'_n(z_l, 1)\} - C \exp\{-cl_{n, \alpha}\} \geq 1 - C \exp\{-cl_{n, \alpha}\} \quad (3.27)$$

with some constants $C$ and $c$, depending on $\alpha$ and $\varepsilon$ only. Let

$$\Omega_n := \cap_{z \in \mathcal{D}} \Omega_n(z, 2). \quad (3.28)$$

Put now

$$v'_0 := v_0'(z) = \frac{\sqrt{y}v_0}{\sqrt{\gamma}}, \quad (3.29)$$

where $\gamma := \min\{1 - \sqrt{y} - |u|, 1 + \sqrt{y} - |u|\}$. $z = u + iv$ and $v_0$ is given by (3.22). Note that $0 \leq \gamma \leq \sqrt{y}$, for $u \in [1 - \sqrt{y}, 1 + \sqrt{y}]$ and $v'_0 \geq v_0$. Denote $\mathcal{D}' := \{z \in \mathcal{D} : v \geq v'_0\}$. 


4 Estimation of $|m_n(z)|$

In this section we bound the probability that $\text{Im } m_n(z) \leq C$ for some numerical constant $C$ and for any $z \in \mathcal{D}'$. We shall derive auxiliary bounds for the difference between the Stieltjes transforms $m_n(z)$ of the empirical spectral measure of the matrix $X$ and the Stieltjes transform $S_y(z)$ of the symmetrized Marchenko–Pastur law. Introduce the additional notations

$$\delta_n := \frac{1}{n} \sum_{j=1}^{n} \varepsilon_j R_{jj}.$$ 

By Lemma 9.5, we have

$$|S_y(z)| \leq \frac{1}{\sqrt{y}} \text{ and } |z + \frac{y-1}{z} + yS_y(z)| \geq \sqrt{y}. \quad (4.1)$$

Introduce notations $g_n(z) := m_n(z) - S_y(z)$, $a_n(z) = z + \frac{y-1}{z} + ym_n(z)$, and $b_n(z) = a_n(z) + yS_y(z)$. Equality (2.23) implies that

$$1 - \frac{y}{(z + \frac{y-1}{z} + yS_y(z))a_n(z)} = 1 + \frac{yS_y(z)}{a_n(z)} = \frac{b_n(z)}{a_n(z)} = \frac{\delta_n}{a_n(z)}.$$

The representation (2.29) implies

$$g_n(z) = \frac{yg_n(z)}{(z + \frac{y-1}{z} + yS_y(z))a_n(z)} + \frac{\delta_n}{a_n(z)}.$$

\[\text{From here it follows by solving for } g_n(z) \text{ that } g_n(z) = \frac{\delta_n}{b_n(z)}. \quad (4.3)\]

Lemma 4.1. Let

$$|g_n(z)| \leq \frac{1}{2\sqrt{y}}. \quad (4.5)$$

Then $|a_n(z)| \geq \frac{\sqrt{y}}{2}$ and $\text{Im } m_n(z) \leq |m_n(z)| \leq \frac{3}{2\sqrt{y}}$.

Proof. This is an immediate consequence of inequalities (4.1) and of

$$|a_n(z)| \geq |z + \frac{y-1}{z} + yS_y(z)| - y|g_n(z)| \geq \frac{1}{2} \frac{\sqrt{y}}{2},$$

$$|m_n(z)| \leq |s(z)| + |g_n(z)| \leq \frac{1}{\sqrt{y}} + \frac{1}{2\sqrt{y}} = \frac{3}{2\sqrt{y}}.$$

\[\Box\]
The rate of convergence to the Marchenko–Pastur distribution

Lemma 4.2. Assume condition (4.5) for \( z = u + iv \) with \( v \geq v_0 \) and \( 1 + \sqrt{y} \geq |u| \geq 1 - \sqrt{y} \). Then for any \( \omega \in \Omega_n \), defined in (3.28), we obtain \( |R_{jj}| \leq \frac{4}{\sqrt{y}} \).

Proof. By definition of \( \Omega_n \) in (3.28), we have

\[
|\varepsilon_j| \leq \frac{16\beta_n^2 \sqrt{y}}{\sqrt{nv}} \left( \sqrt{y} \text{Im} m_n(z) + \frac{1}{\sqrt{pv}} + \frac{(1 - y)v}{|z|} \right).
\] (4.6)

Note that, for \( 1 + \sqrt{y} \geq |u| \geq 1 - \sqrt{y} \),

\[
\frac{(1 - y)v}{|z|^2} \leq \frac{(1 - y)^2 + v^2}{v^2 + (1 - \sqrt{y})^2} \leq \frac{(1 + \sqrt{y})^2}{2} + \frac{1}{2} \leq \frac{5}{2}.
\] (4.7)

Applying Lemmas 4.1, inequality (4.7) and definition (3.22), we get \( |\varepsilon_j| \leq A \) with some \( A > 0 \) which doesn’t depending on the parameter \( d \geq 1 \) in (3.22). We may choose the parameter \( d \) such that

\[
|\varepsilon_j| \leq \frac{1}{2},
\] (4.8)

for any \( \omega \in \Omega_n \), \( n \geq 2 \), \( v \geq v_0 \) and \( 1 + \sqrt{y} \geq |u| \geq 1 - \sqrt{y} \). Using the representation (2.27) and applying Lemma 4.1 we get \( |R_{jj}| \leq \frac{4}{\sqrt{y}}. \)

Lemma 4.3. Assume condition (4.5). Then, for any \( \omega \in \Omega_n \) and \( v \geq v_0 \),

\[
|g_n(z)| \leq \frac{1}{100 \sqrt{y}^2}.
\] (4.9)

Proof. Lemma 4.2, inequality (4.8), and the representation (4.4) together imply

\[
|\delta_n| \leq \frac{4}{n \sqrt{y}} \sum_{j=1}^{n} |\varepsilon_j| \leq \frac{64 \beta_n^2}{\sqrt{nv}} \left( \sqrt{y} \text{Im} \frac{1}{2} m_n(z) + \frac{1}{\sqrt{pv}} + \frac{(1 - y)v}{|z|} \right).
\] (4.10)

Note that

\[
|z + \frac{y - 1}{z} + y m_n(z) + y S_y(z)| \geq \text{Im} z + \text{Im} \left( \frac{y - 1}{z} + y \text{Im} m_n(z) + y \text{Im} S_y(z) \right)
\geq \text{Im} \left( z + \frac{y - 1}{z} + y S_y(z) \right) \geq \frac{1}{2} \text{Im} \left\{ (z + \frac{y - 1}{z})^2 - 4y \right\}.\] (4.11)

These relations and Lemma 9.7 together imply

\[
\frac{|\delta_n|}{|z + m_n(z) + s(z)|} \leq \frac{64 \sqrt{2} \beta_n^2}{n \sqrt{y^2}} + \frac{128 \beta_n^2 \sqrt{y}}{(nv)^{\frac{3}{2}}} + \frac{128 \beta_n^2 c_0(y)}{\sqrt{nv}}.
\] (4.12)
where
\[
c_0(y) = \begin{cases} 
0 & \text{if } y = 1, \\
\frac{\sqrt{1+y}}{1-\sqrt{y}} & \text{if } 0 < y < 1.
\end{cases}
\] (4.13)

For \( v\sqrt{y} \geq v_0 \), we get
\[
|g_n(z)| \leq \frac{128\beta_n^2(1 + c_0(y))}{\sqrt{n\nu_0}} + \frac{128\beta_n^2 \sqrt{y}}{(n\nu_0)^{\frac{1}{2}}}
\leq \frac{128(1 + c_0(y))}{\sqrt{dy}} + \frac{128}{yd^{\frac{1}{2}}} \leq \frac{1}{100} \sqrt{y}
\] (4.14)

by choosing the constant \( d \geq 1 \) in \( v_0 \) appropriately large. Thus the lemma is proved.

Lemma 4.4. Assume that condition (4.5) holds, for some \( z = u + iv \in D' \) and for any \( \omega \in \Omega_n \), (see (3.28) and the subsequent notions). Then (4.5) holds as well for \( z' = u + i\hat{v} \in D' \) with \( v \geq \hat{v} \geq v - n^{-8} \), for any \( \omega \in \Omega_n \).

Proof. First of all note that
\[
|m_n(z) - m_n(z')| = \frac{1}{n}(v - \hat{v})|\text{Tr} R(z)R(z')| \leq \frac{v - \hat{v}}{v\hat{v}} \leq \frac{C}{n^4} \leq \frac{1}{100} \sqrt{y}
\]
and \( |s(z) - s(z')| \leq \frac{|z - z'|}{v\hat{v}} \leq \frac{1}{100\sqrt{y}} \). By Lemma 4.3 we have \( |g_n(z)| \leq \frac{1}{100\sqrt{y}} \). All these inequalities together imply \( |g_n(z')| \leq \frac{3}{100\sqrt{y}} < \frac{1}{2\sqrt{y}} \). Thus, Lemma 4.4 is proved.

Proposition 4.1. Assuming the conditions of Theorem 1.1 there exist constants \( C > 0 \) and \( c > 0 \) depending on \( \kappa \) and \( \alpha \) only such that
\[
\Pr\{|m_n(z)| \leq \frac{3\sqrt{y}}{2} \text{ for any } z \in D'\} \leq C \exp\{-cn_{n,\alpha}\}.
\] (4.15)

Proof. First we note that \( |g_n(z)| \leq \frac{1}{2\sqrt{y}} \) a.s., for \( z = u + 4\sqrt{y}i \). By Lemma 4.4, \( |g_n(z')| \leq \frac{1}{2\sqrt{y}} \) for any \( \omega \in \Omega_n \) and \( z \in D' \). Applying Lemma 4.1 and a union bound, we get
\[
\Pr\{|m_n(z)| \leq \frac{3\sqrt{y}}{2} \text{ for any } z \in D'\} \leq C \exp\{-cn_{n,\alpha}\}.
\] (4.16)
Thus the proposition is proved.

5 Large deviations II

In this Section we shall obtain bounds for the large deviation probabilities of the sum of the \( \epsilon_j \). We start with the quantity
\[
\delta_{n1} = \frac{1}{np} \sum_{j=1}^{n} \sum_{k=1}^{p} (X_{jk}^2 - 1) R_{k+n,k+n}^{(j)}.
\] (5.1)

We prove the following lemma.
Lemma 5.1. Under conditions of Theorem 1.1 there exist constants $C$ and $c$ depending on $\kappa$ and $\alpha$ such that
\[
\Pr\{|\delta_{n1}| \leq 2\sqrt{yn^{-{1}v^{-{1}3\beta_n^2/2}}(3y/2 + 1 + \sqrt{y} + \frac{1}{\sqrt{pv}})}\} \leq C \exp\{-cn,\alpha\},
\]
for any $z \in \mathcal{D}'$.

Proof. For any $j = 1, \ldots, n$ and any $k = 1, \ldots, p$, we introduce the truncated and centered random variables
\[
\xi_{jk} = \hat{X}_{jk}^2 - E\hat{X}_{jk}^2,
\]
where $\hat{X}_{jk} = X_{jk}1\{|X_{jk}| \leq l_{n,\alpha}^1\}$. It is straightforward to check that
\[
0 \leq 1 - E\hat{X}_{jk}^2 \leq C \exp\{-cn,\alpha\}.
\]
Introduce as well the quantities
\[
\tilde{\delta}_{n1} = \frac{1}{np} \sum_{j=1}^{n} \sum_{k=1}^{p} (\hat{X}_{jk}^2 - 1)R_{k+n,k+n}^{(j)}, \quad \tilde{\delta}_{n1} = \frac{1}{np} \sum_{j=1}^{n} \sum_{k=1}^{p} \xi_{jk}R_{k+n,k+n}^{(j)}
\]
By assumption (1.1),
\[
\Pr\{\delta_{n1} \neq \tilde{\delta}_{n1}\} \leq C \exp\{-cn,\alpha\}. \tag{5.6}
\]
Let
\[
\zeta_j := \frac{1}{p} \sum_{k=1}^{p} \xi_{jk}R_{k+n,k+n}^{(j)}.
\]
Then
\[
\tilde{\delta}_{n1} = \frac{1}{n} \sum_{j=1}^{n} \zeta_j. \tag{5.8}
\]
Let $\mathcal{R}_j$ denote $\sigma$–algebra denoted by $X_{lk}$ with $1 \leq l \leq j$ and $1 \leq k \leq p$, for $j = 1, \ldots, n$. Let $\mathcal{R}_0$ denote the trivial $\sigma$-algebra. Note that the sequence $\tilde{\delta}_{n1}$ is a martingale with respect to the $\sigma$–algebras $\mathcal{R}_j$. In fact,
\[
E(\zeta_j | \mathcal{R}_{j-1}) = E\left(E(\zeta_j | \mathcal{R}_j) | \mathcal{R}_{j-1}\right) = 0. \tag{5.9}
\]
In order to use large deviation bounds for $\tilde{\delta}_{n1}$ we replace the differences $\zeta_j$ by truncated random variables.

Since $\zeta_j$ is a sum of independent bounded random variables with mean zero (conditioning on $\mathcal{R}_j$), and applying inequality (9.12) with $\sigma^2 = 4p\frac{\sigma^2}{l_{n,\alpha}^1} \left(p^{-1} \sum_{k=1}^{p} |R_{k+n,k+n}^{(j)}|^2 \right)$ and $x := \frac{1}{2} l_{n,\alpha}^1 \sigma$, we get
\[
\Pr_j\left\{|\zeta_j| > 2p^{-\frac{1}{2}}l_{n,\alpha}^{\frac{1}{2}} \left(\frac{1}{p} \sum_{k=1}^{p} |R_{k+n,k+n}^{(j)}|^2 \right)^{\frac{1}{2}}\right\} \leq C \exp\{-cn,\alpha\}. \tag{5.10}
\]
Applying now inequalities (3.10) and (3.5), we obtain
\[
\Pr \left\{ |\zeta_j| > 2p^{-\frac{1}{2}}v^{-\frac{1}{2}}l_{n,\alpha}^{\frac{3}{2}+\frac{1}{2}} \left( \sqrt{\gamma} \text{Im} \left( \sqrt{\gamma} m_n(z) + \frac{1}{z} \frac{(1-y)v}{\sqrt{\gamma}} + \frac{1}{\sqrt{pv}} \right) \right) \right\} \leq C \exp \{-cl_{n,\alpha}\}. \tag{5.11}
\]
Note that \( \frac{(1-y)v}{\sqrt{\gamma} z^2} \leq \frac{5}{2} \) for any \( z \in \mathcal{D}' \) (see (4.7)). Denote by \( t_{nv}^2 = \frac{3y}{2} + 1 + \sqrt{\gamma} + \frac{1}{\sqrt{pv}} \).

Applying Proposition 4.1, we get
\[
\Pr_j \left\{ |\zeta_j| > 2p^{-\frac{1}{2}}v^{-\frac{1}{2}}l_{n,\alpha}^{\frac{3}{2}+\frac{1}{2}} t_{nv} \right\} \leq C \exp \{-cl_{n,\alpha}\}. \tag{5.12}
\]

Introduce \( \hat{\zeta}_j := \zeta_j \mathbb{1} \{ |\zeta_j| \leq 2p^{-\frac{1}{2}}v^{-\frac{1}{2}}l_{n,\alpha}^{\frac{3}{2}+\frac{1}{2}} t_{nv} \} \).
\[
\Pr \left\{ \sum_{j=1}^n \zeta_j \neq \sum_{j=1}^n \hat{\zeta}_j \right\} \leq C \exp \{-cl_{n,\alpha}\}. \tag{5.14}
\]

Furthermore, introduce the conditionally centered random variables
\[
\tilde{\zeta}_j = \hat{\zeta}_j - \mathbb{E} \{ \hat{\zeta}_j | \mathcal{R}_{j-1} \}. \tag{5.15}
\]

Using the Cauchy-Schwarz inequality and boundedness of the random variables \( \xi_{jk} R_{k+n,k+n}^{(j)} \) it follows that
\[
|\mathbb{E} \{ \hat{\zeta}_j | \mathcal{R}_{j-1} \}| \leq \mathbb{E}^{\frac{1}{2}} \{ |\zeta_j|^2 | \mathcal{R}_{j-1} \} \mathbb{E}^{\frac{1}{2}} \{ |\zeta_j| > 2p^{-\frac{1}{2}}v^{-\frac{1}{2}}l_{n,\alpha}^{\frac{3}{2}+\frac{1}{2}} t_{nv} | \mathcal{R}_{j-1} \} \leq C \exp \{-cl_{n,\alpha}\} \mathbb{E} \{ \xi_{jk}^2 | \mathcal{R}_{j-1} \} \leq C v^{-1} \left( \frac{1}{p} \sum_{k=1}^p \mathbb{E} \mathbb{X}_{jk}^2 \right)^{\frac{1}{2}} \exp \{-cl_{n,\alpha}\} \leq C \exp \{-cl_{n,\alpha}\}, \tag{5.16}
\]
for \( v \geq v_0 \) with a constant \( C \) which is independent of \( d \geq 1 \).

We shall use now Lemma 9.4 in the Appendix (an inequality by Bentkus for martingales) to bound \( \tilde{\delta}_{n1} \). By the definition of \( \tilde{\zeta}_j \), we may choose \( b_j = 2l_{n,\alpha}^{\frac{3}{2}+\frac{1}{2}} p^{-\frac{1}{2}} v^{-\frac{1}{2}} t_{nv} \), \( \sigma^2 = 4gy_{l_{n,\alpha}^{\frac{3}{2}+\frac{1}{2}}} v^{-1} l_{n,\alpha}^{\frac{3}{2}+\frac{1}{2}} \) and \( x = l_{n,\alpha}^{\frac{1}{2}} \sigma \). Using inequality (9.4), we obtain
\[
\Pr \{ \tilde{\delta}_{n1} > 2\sqrt{yn}^{-1} v^{-\frac{1}{2}} t_{nv} l_{n,\alpha}^{\frac{3}{2}+1} \} \leq C \exp \{-cl_{n,\alpha}\}. \tag{5.17}
\]
Inequalities (5.14)–(5.17) together conclude the proof of Lemma 5.1.

Let
\[
\delta_{n2} := \frac{1}{n^2} \sum_{j=1}^n \sum_{1 \leq \ell \neq k \leq p} X_{jk} X_{jk} R_{k+n,k+n}^{(j)}. \tag{5.18}
\]
Lemma 5.2. Under the conditions of Theorem 1.1 there exist constants $C$ and $c$ depending on $\kappa$ and $\alpha$ only such that

$$\Pr\left\{|\delta_{n2}| > \frac{2\sqrt{y}}{n\sqrt{v}} \beta_n^2 \sqrt{l_{n,\alpha}} \left(\frac{3}{2} + 1 + \sqrt{y} + \frac{1}{\sqrt{pv}}\right)\right\} \leq C \exp\{-cl_{n,\alpha}\} \quad (5.19)$$

Proof. The proof of this Lemma is similar to the proof of Lemma 5.1. We introduce the random variables

$$\eta_j = \frac{1}{n} \sum_{1 \leq t \neq k \leq p} X_{jk} X_{jl} R^{(j)}_{l+n,k+n} \quad (5.20)$$

and note that the sequence

$$M_j = \sum_{t=1}^{j} \eta_t, \quad j = 1, \ldots, n, \quad (5.21)$$

is a martingale with respect to the $\sigma$–algebras $\mathcal{R}_j$. In order to apply the martingale large deviation bound (9.4) we replace $\eta_j$ by truncated random variables. Note that $\eta_j = \varepsilon_j^2$. By Corollary 3.5 we have

$$\Pr\{|\eta_j| > \frac{4\sqrt{y}\beta_n^2 (nv)^{-\frac{1}{2}} (\sqrt{y} \Im \frac{1}{2} m_n(z) + \sqrt{(1-y)v} + \frac{1}{\sqrt{pv}})}\} \leq C \exp\{-cl_{n,\alpha}\}. \quad (5.22)$$

Applying Lemma 4.1 we get

$$\Pr\{|\eta_j| > 4\sqrt{y}\beta_n^2 (nv)^{-\frac{1}{2}} t_{nv}\} \leq C \exp\{-cl_{n,\alpha}\}. \quad (5.23)$$

Introduce now the random variables

$$\tilde{\eta}_j = \eta_j I\{|\eta_j| \leq 4\sqrt{y}\beta_n^2 (nv)^{-\frac{1}{2}} t_{nv}\} \quad (5.24)$$

and

$$\tilde{\eta}_j = \tilde{\eta}_j - E\tilde{\eta}_j. \quad (5.25)$$

Furthermore, we introduce the random variables

$$\theta_j := \tilde{\eta}_j - E\{\tilde{\eta}_j|\mathcal{R}_{j-1}\}. \quad (5.26)$$

We consider the martingale

$$\tilde{M}_j = \sum_{t=1}^{j} \theta_t, \quad j = 1, \ldots, n. \quad (5.27)$$

Applying Lemma 9.4 with $\sigma^2 = 16\sqrt{y}\beta_n^4 v^{-1} t_{nv}^2$ and $x = t_{n,\alpha}^\frac{1}{2}$, we get

$$\Pr\{|\delta_{n2}| > \frac{4\sqrt{y}\beta_n^2 \sqrt{l_{n,\alpha}}}{n\sqrt{v}} t_{nv}\} \leq C \exp\{-cl_{n,\alpha}\} \quad (5.28)$$

Thus the Lemma is proved. \qed
Finally we have to bound
\[ \delta_{n3} := \frac{1}{n^2} \sum_{j=1}^{n} \sum_{k=1}^{p} (R_{k+n,k+n} - R_{k+n,k+n}^{(j)}) R_{jj}. \] (5.29)

**Lemma 5.3.** There exists a positive constant \( C \) such that
\[ |\delta_{n3}| \leq \frac{1}{n v} \text{Im} \, m_{n}(z). \] (5.30)

**Proof.** It is easy to check that
\[ \sum_{k=1}^{p} (R_{k+n,k+n} - R_{k+n,k+n}^{(j)}) = \frac{1}{2} (\text{Tr} \, R - \text{Tr} \, R^{(j)}) + \frac{1}{2z} \] (5.31)

By formula (5.4) in [4], we have
\[ (\text{Tr} \, R - \text{Tr} \, R^{(j)}) R_{jj} = (1 + \frac{1}{p} \sum_{l,k=1}^{n} X_{jl} X_{jk} (R^{(j)})^2_{l+n,k+n}) R_{jj}^2 = -\frac{d}{dz} R_{jj}. \] (5.32)

¿From here it follows that
\[ \frac{1}{n^2} \sum_{j=1}^{n} (\text{Tr} \, R - \text{Tr} \, R^{(j)}) R_{jj} = -\frac{1}{n} \frac{d}{dz} m_{n}(z). \] (5.33)

Note that
\[ m_{n}(z) = \frac{z}{n} \sum_{k=1}^{n} \frac{1}{s_k^2 - z^2} \] (5.34)

and
\[ \frac{d}{dz} m_{n}(z) = \frac{m_{n}(z)}{z} - \frac{2z^2}{n} \sum_{k=1}^{n} \frac{1}{(s_k^2 - z^2)^2}. \] (5.35)

This implies that
\[ \delta_{n3} = \frac{z^2}{n^2} \sum_{k=1}^{n} \frac{1}{(s_k^2 - z^2)^2}. \] (5.36)

Finally, we note that
\[ \text{Im} \, m_{n}(z) = \frac{1}{n} \sum_{k=1}^{n} \frac{v(s_k^2 + |z|^2)}{|s_k^2 - z^2|^2}. \] (5.37)

The last relation implies
\[ \left| \frac{z^2}{n^2} \sum_{k=1}^{n} \frac{1}{(s_k^2 - z^2)^2} \right| \leq \frac{1}{nv} \text{Im} \, m_{n}(z). \] (5.38)

The inequality (5.38) concludes the proof. Thus Lemma 5.3 is proved. \( \square \)
6 Stieltjes transforms

In this section we derive auxiliary bounds for the difference between the Stieltjes transforms \( m_n(z) \) of the empirical spectral measure of the matrix \( V \) and the Stieltjes transform \( S_y(z) \) of the symmetrized Marchenko–Pastur law. We introduce the additional notations

\[
\tilde{\delta}_n = \delta_n^3, \quad \overline{\delta}_n = \frac{1}{n} \sum_{\nu=1}^{n} \sum_{j=1}^{n} \varepsilon_{j\nu} \varepsilon_j R_{jj}.
\] (6.1)

Recall that \( S_y(z) \) satisfies the equation

\[
S_y(z) = -\frac{1}{z + yS_y(z) + \frac{y-1}{z}}.
\] (6.2)

Recall that \( g_n(z) := m_n(z) - S_y(z) \), \( a_n(z) = z + \frac{y-1}{z} + ym_n(z) \), and \( b_n(z) = z + \frac{y-1}{z} + ym_n(z) + yS_y(z) \). The representation (2.29) and the equality (6.2) together imply

\[
g_n(z) = -yg_n(z) S_y(z) + \frac{\delta_n}{z + ym_n(z) + \frac{y-1}{z}} + \frac{\overline{\delta}_n}{z + ym_n(z) + \frac{y-1}{z}}^2.
\] (6.3)

From equality (6.3) it follows that

\[
|g_n(z)| \leq \frac{|\delta_n| + |\overline{\delta}_n|}{|b_n(z, y)||a_n(z, y)|} + \frac{|\overline{\delta}_n|}{|b_n(z, y)|}.
\] (6.4)

For any \( z \in \mathcal{D} \) introduce the events

\[
\tilde{\Omega}_n(z) = \left\{ \omega \in \Omega : |\delta_n| \leq \frac{2\sqrt{y} \beta^2 \sqrt{\log n}}{n \sqrt{y}} \left( \frac{3\sqrt{y}}{2} + \frac{1 + \sqrt{y}}{y} + \frac{1}{\sqrt{nv}} \right) \right\},
\] (6.5)

\[
\tilde{\Omega}_n(z) = \left\{ \omega \in \Omega : |\tilde{\delta}_n| \leq \frac{\text{Im} m_n(z)}{nv} \right\},
\] (6.5)

\[
\overline{\Omega}_n(z) = \left\{ \omega \in \Omega : |\overline{\delta}_n| \leq \frac{16y \beta^4}{nv} \left( \text{Im} m_n(z) + \frac{1}{nv} + \frac{(1 - y)w}{|z|^2} \right) \frac{1}{n} \sum_{j=1}^{n} |R_{jj}| \right\}.
\] (6.6)

Put \( \tilde{\Omega}_n^*(z) = \tilde{\Omega}_n(z) \cap \tilde{\Omega}_n(z) \cap \overline{\Omega}_n(z) \). By Lemmas 5.1–5.3, we have

\[
\Pr\{ \tilde{\Omega}_n(z) \} \geq 1 - C \exp\{-c l_{n,a}\}.
\] (6.7)

By Lemma 5.3,

\[
\Pr\{ \overline{\Omega}_n(z) \} = 1.
\] (6.8)
The rate of convergence to the Marchenko–Pastur distribution

Note that
\[ |\varepsilon_{j,\nu}|^2 \leq \frac{1}{2} (|\varepsilon_{j,\nu}|^2 + |\varepsilon_{j,3}|^2). \]  
(6.9)

By inequality (3.20), we have, for \( \nu = 1, 2, \)
\[ \Pr \left\{ |\varepsilon_{j,\nu}|^2 > \frac{16y\beta_n^4}{nv} \left( y\text{Im}m_n(z) + \frac{1}{pv} + \frac{(1-y)v}{|z|^2} \right) \right\} \leq C \exp\{-cl_{n,\alpha}\}. \]  
(6.10)

Furthermore,
\[ \Pr \{|\varepsilon_{j,3}|^2 \leq \frac{1}{n^2v^2}\} = 1. \]  
(6.11)

Similar to equality (3.27) we may show that
\[ \Pr \left\{ \bigcap_{z \in \mathcal{D}} (\Omega^*_n(z) \cap \Omega_n(z)) \right\} \geq 1 - C \exp\{-c_{n,\alpha}\}. \]  
(6.12)

Let
\[ \Omega^*_n := \bigcap_{z \in \mathcal{D}} (\Omega^*_n(z) \cap \Omega_n(z)). \]  
(6.13)

In the what follows we shall assume that
\[ v_0 = \frac{dy\beta_n^4}{n} \]  
(6.14)

with \( d \geq 32. \) We now prove the first essential bound.

**Lemma 6.1.** Let \( z = u + iv \in \mathcal{D}. \) Assume that
\[ |g_n(z)| \leq \frac{1}{2\sqrt{y}}. \]  
(6.15)

Then for any \( \omega \in \Omega^*_n, \) the following bound holds
\[ |g_n(z)| \leq \frac{32\beta_n^4(1 + \sqrt{y} + c_0(y))}{nv} + \frac{32\beta_n^4}{nv |b_n(z)|} \]  
(6.16)

**Proof.** Inequalities (6.4), (6.5) and (6.12) imply that for \( \omega \in \Omega^*_n, \)
\[ |g_n(z)| \leq \frac{2\sqrt{y}\beta_n^4 \sqrt{n,\alpha}}{nv |b_n(z)| |a_n(z, y)|} \left( \sqrt{\frac{3\sqrt{y}}{2}} + \frac{1}{pv} + \frac{(1-y)v}{|z|^2} \right) + \frac{\text{Im}m_n(z)}{nv |b_n(z)|} \]  
(6.16)

By Lemma 4.1 \( |a_n(z)| \geq \frac{\sqrt{y}}{2}. \) In addition, \( |b_n(z)| \geq y\text{Im}m_n(z). \) By Lemma 9.7 we have \( |b_n(z)| \geq \frac{1}{2} y^\frac{1}{4} \sqrt{\gamma + v}. \) Moreover, for \( z \in \mathcal{D}', \) \( \frac{(1-y)v}{|z|^2} \leq 1 + \sqrt{y} \) and \( pv \geq d \geq 1 \).

All these relations together imply
\[ |g_n(z)| \leq \frac{12(1 + \sqrt{y})\beta_n^4 \sqrt{n,\alpha}}{nv} + \frac{1}{nv} + \frac{32\beta_n^4}{nv |b_n(z)|} + \frac{32\beta_n^4c_0(y)}{nv} \]  
(6.17)
The rate of convergence to the Marchenko–Pastur distribution

where \( c_0(y) \) is defined in (4.13). We use here as well the inequalities

\[
\frac{\sqrt{1 - y}}{|z|} \leq c_0(y),
\]

(6.18)

for \( z \in \mathcal{D}' \). Note that the last inequality holds for any \( z = u + iv \) with \( u \in \mathbb{R} \) and \( v \geq v_0 \).

Inequality (6.17) completes the proof of the Lemma. \( \square \)

Recall that \( v'_0 := v'_0(z) := \frac{v_0}{\sqrt{\gamma}} \), where \( \gamma = \min\{|u| - 1 + \sqrt{v}, 1 + \sqrt{y} - |u|\} \) and \( z = u + iv \), and \( \mathcal{D}' := \{z \in \mathcal{D} : v \geq v'_0\} \).

**Proposition 6.1.** There exists constants \( C, c \), such that

\[
\Pr \left\{ |g_n(z)| > \frac{32\beta_4^4(1 + \sqrt{y} + c_0(y))}{nv} + \frac{32\beta_4^4}{n^2v^2\sqrt{\gamma + v}} \right\}.
\]

(6.19)

**Proof.** Note that for \( v = 4\sqrt{y} \) we have

\[
|g_n(z)| \leq \frac{2}{v} \leq \frac{1}{2}\sqrt{y}.
\]

(6.20)

By Lemma 6.1 we obtain the inequality (6.19). By Lemma 4.4, this inequality holds for any \( \frac{4}{\sqrt{y}} \geq v \geq v_0 \). Thus proposition 6.1 is proved. \( \square \)

**7 Proof of Theorem 1.1**

Here we conclude the proof of Theorem 1.1.

We apply now the result of Corollary 2.2 to the empirical spectral distribution function \( F_n(x) \) of the random matrix \( X \). First we bound the integral over the line with \( V = 4\sqrt{y} \). Note that in this case we have \( \text{Im} m_n(z) \leq \frac{1}{\sqrt{y}} \leq \frac{1}{4\sqrt{y}} \leq \frac{3}{2\sqrt{y}} \) and \( |g_n(z)| \leq \frac{1}{2\sqrt{y}} \) for any \( \omega \in \Omega \). We may now apply results of the previous Lemmas on large deviations. This ensures the following bound for \( g_n(z) \) for all \( z = u + iV \) with \( u \in \mathbb{R} \).

\[
|g_n(z)| \leq \frac{12(1 + \sqrt{y})\beta^2_n}{n|a_n(z, y)b_n(z, y)|} + \frac{128\beta_4^4}{n|a_n(z, y)b_n(z, y)|} + \frac{\text{Im} m_n(z)}{n\sqrt{y}|b_n(z, y)|}.
\]

(7.1)

Note that for \( V = 4\sqrt{y} \),

\[
|a_n(z, y)b_n(z, y)| \geq \begin{cases} 16y & \text{for } |u| \leq \frac{16y}{64y + |z|^2 + 1}, \\ \frac{1}{4}|z|^2 & \text{for } |u| > 4. \end{cases}
\]

(7.2)
The rate of convergence to the Marchenko–Pastur distribution

We may rewrite the bound (7.1) as follows

$$|g_n(z)| \leq \frac{256y(1 + \sqrt{y})\beta^2_n\sqrt{n\alpha}}{n(64y|z|^2 + 1)} + \frac{C\text{Im} m_n(z)}{n} + \frac{512\beta^4_n}{n|z|^4} + \frac{2048\beta^4_n}{n(64y|z|^2 + 1)}. \quad (7.3)$$

Note that for any distribution function $F(x)$ we have

$$\int_{-\infty}^{\infty} \text{Im} S_F(u + iv) du = \pi \quad (7.4)$$

From here it follows that, for $V = 4\sqrt{y}$

$$\int_{|u| \geq n} |m_n(z) - S_y(z)| du \leq \frac{C}{n} \quad (7.5)$$

Denote $\overline{D}_n := \{z = u + 2i : |u| \leq n\}$ and

$$\overline{\Omega}_n := \left( \bigcap_{z \in \overline{D}_n} \{\omega \in \Omega : |g_n(z)| \leq \frac{C\beta^2_n}{n|z|^2 + 1}\} \right) \cap \Omega^*_n$$

Using a union bound, we may show that

$$\Pr\{\overline{\Omega}_n\} \geq 1 - C \exp\{-c\ln n,\alpha\}. \quad (7.6)$$

It is straightforward to check that for $\omega \in \overline{\Omega}_n$

$$\int_{-\infty}^{\infty} |m_n(z) - S_y(z)| du \leq \frac{C\beta^4_n}{n}. \quad (7.7)$$

We choose $\varepsilon = (2Hv_0)^{\frac{1}{2}}$ and $v_0 = \frac{dy\beta^2_n}{n}$. To conclude the proof we need to consider the “vertical” path integrals in $z = x + iv'$ with $x \in \Re^\varepsilon$, $v' = \frac{v_0}{\sqrt{\gamma}}$ and $\gamma = 2 - |x|$. It will be enough to consider one of these integrals only, the others being similar, namely

$$\int_{v'}^{4\sqrt{y}} \frac{1}{n^2v^2\sqrt{\gamma + v}} dv \leq \frac{1}{n^2v_0^2\sqrt{\gamma}} \leq \frac{1}{n^2v_0} \leq \frac{1}{ndy\beta^4_n} \quad (7.8)$$

and

$$\int_{v'}^{4\sqrt{y}} \frac{1}{nv} dv \leq \frac{1}{n} |\ln \frac{v_0}{4\sqrt{y}}| \leq \frac{C + \ln n}{n}. \quad (7.9)$$

Finally, we obtain for any $\omega \in \overline{\Omega}_n$

$$\Delta(F_n, G) = \sup_x |F_n(x) - G_y(x)| \leq \frac{C\beta^4_n\ln n}{n} \quad (7.10)$$

with some constant $C > 0$ depending on $\varepsilon$, $\alpha$ and $y$ only. Thus Theorem 1.1 is proved.
8 Proof of Theorem 1.2

Consider the singular value decomposition of the matrix $X$. Let $U$ and $H$ be unitary matrices of dimension $n \times n$ and $p \times p$ respectively. Let $S$ be a $n \times n$ diagonal matrix whose entries are the singular value of the matrix $X$. Let $O_{p \times q}$ denote the $p \times q$-matrix with zero entries. Introduce the matrix $\tilde{S} = \begin{bmatrix} S & O_{n \times (p-n)} \end{bmatrix}$. We have the following representation

$$X = U \tilde{S} H^*.$$  \hfill (8.1)

We may represent the matrix $H$ in the form

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix},$$  \hfill (8.2)

where $H_{11}$ is a $n \times n$ matrix and $H_{22}$ is a $(p-n) \times (p-n)$ matrix. We introduce the matrix

$$Z^* = \begin{bmatrix} \frac{1}{\sqrt{2}} U^* & \frac{1}{\sqrt{2}} H_{11}^* \\ \frac{1}{\sqrt{2}} H_{12}^* \end{bmatrix}.$$  \hfill (8.3)

It is straightforward to check that

$$Z^* V Z = \begin{bmatrix} S & O_{n \times n} & O_{n \times (p-n)} \\ O_{n \times n} & -S & O_{n \times (p-n)} \\ O_{(p-n) \times n} & O_{(p-n) \times n} & O_{(p-n) \times (p-n)} \end{bmatrix},$$  \hfill (8.4)

where $S$ denotes the diagonal matrix with entries $s_j$. The equality (8.4) implies that the rows $z_j$ of the matrix $Z$, for $j = 1, \ldots, n$, are the eigenvectors of the matrix $V$ corresponding to the eigenvalues $s_j$. Similarly, the rows $z_{j+n}$ of the matrix $Z$, for $j = 1, \ldots, n$, are the eigenvectors of the matrix $V$ corresponding to the eigenvalues $-s_j$ and the rows $z_{2n+l}$, for $l = 1, \ldots, p - n$, are the eigenvectors of the matrix $V$ corresponding to the eigenvalues 0.

We note the following representation for the diagonal entries of the resolvent matrix $R$:

$$R_{jj} = \sum_{k=1}^{n+p} \frac{1}{\lambda_k - z} |Z_{kj}|^2.$$  \hfill (8.5)

Let $\lambda_1, \ldots, \lambda_{n+p}$ denote the eigenvalues of the matrix $V$ ordered in such way that

$$\lambda_j = \begin{cases} -s_j, & \text{if } 1 \leq j \leq n \\ s_j, & \text{if } n+1 \leq j \leq 2n \\ 0, & \text{if } 2n \leq j \leq n+p. \end{cases}$$  \hfill (8.6)

Consider the distribution function $F_{n_j}(x)$ of the following weighted empirical probability distribution on the eigenvalues $\lambda_1, \ldots, \lambda_{n+p}$

$$F_{n_j}(x) = \sum_{k=1}^{n+p} |Z_{kj}|^2 \mathbb{1}\{\lambda_k \leq x\}.$$  \hfill (8.7)
Then we have
\[ R_{jj} = R_{jj}(z) = \int_{-\infty}^{\infty} \frac{1}{x - z} dF_{nj}(x). \] (8.8)
which means that \( R_{jj} \) is the Stieltjes transform of the distribution \( F_{nj}(x) \). Note that, for any \( \lambda > 0 \)
\[ \max_{1 \leq k \leq n+p} |Z_{kj}|^2 \leq \sup_x (F_{nj}(x + \lambda) - F_{nj}(x)) =: Q_{nj}(\lambda). \] (8.9)
On the other hand, it is easy to check that
\[ Q_{nj}(\lambda) \leq 2 \sup_u \lambda \text{Im} R_{jj}(u + i\lambda). \] (8.10)
By the relations (3.23) and (3.27), for any \( v \geq v_0 \) with \( v_0 = \frac{d\beta}{d} \) with a sufficiently large constant \( d \), we have
\[ \Pr\left\{ \frac{|\varepsilon_j|}{|b_n(z, y)|} > \frac{1}{2} \right\} \leq C \exp\{-cl_{n,\alpha}\}. \] (8.11)
Furthermore, the representation (2.27) and inequality (8.11) together imply, for \( v \geq v_0 \)
\[ \text{Im} R_{jj} \leq |R_{jj}| \leq C. \] (8.12)
This implies that
\[ \Pr\left\{ \max_{1 \leq k \leq n+p} |Z_{kj}|^2 > \frac{C\beta^4}{n} \right\} \leq C \exp\{-cl_{n,\alpha}\}. \] (8.13)
By definition of \( H \), we obtain
\[ \Pr\left\{ \max_{1 \leq j,k \leq n} |u_{kj}|^2 > \frac{C\beta^4}{n} \right\} \leq C \exp\{-cl_{n,\alpha}\}. \] (8.14)
and
\[ \Pr\left\{ \max_{1 \leq j,k \leq p} |v_{kj}|^2 > \frac{C\beta^4}{n} \right\} \leq C \exp\{-cl_{n,\alpha}\}. \] (8.15)
By a union bound, the inequality (1.4) follows. To prove inequality (1.5), we consider the quantity
\[ r_j := R_{jj} - S_y(z), \quad j = 1, \ldots, n. \] (8.16)
Using equalities (2.27) and (6.2), we get
\[ r_j = -\frac{S_y(z)g_n(z)}{b_n(z)} + \frac{\varepsilon_j}{b_n(z)} R_{jj}. \] (8.17)
By inequalities (6.19) and (3.27), we have
\[ |r_j| \leq \frac{C\beta^2}{\sqrt{nv}} + \frac{C\beta^4}{nv\sqrt{\gamma + v}}. \] (8.18)
This implies that
\[ \sup_{x \in J} \int_{V'} |r_j(x + iv)| dv \leq \frac{C\beta^2_n}{\sqrt{n}} + \frac{C\beta^4_n}{n\sqrt{v'}}. \] (8.19)

Similar to (7.7) we get
\[ \int_{-\infty}^{\infty} |r_j(x + iV)| dx \leq \frac{C\beta^2_n}{\sqrt{n}}. \] (8.20)

Applying Corollary 2.2, we finally obtain
\[ \Pr\{ \sup_{x} |F_{nj}(x) - G_y(x)| \leq \frac{C\beta^2_n}{\sqrt{n}} \} \geq 1 - C\exp\{-cl_{n,\alpha}\}. \] (8.21)

In view of
\[ \Pr\{ \sup_{x} |F_n(x) - G_y(x)| \leq \frac{C\beta^4_n\ln n}{n} \} \geq 1 - C\exp\{-cl_{n,\alpha}\}, \] (8.22)
we get
\[ \Pr\{ \sup_{x} |F_{nj}(x) - G_y(x)| \leq \frac{C\beta^2_n}{\sqrt{n}} \} \geq 1 - C\exp\{-cl_{n,\alpha}\}. \] (8.23)

The last two inequalities together imply that
\[ \Pr\{ \sup_{x} |F_{nj}(x) - F_n(x)(x)| \leq \frac{C\beta^2_n}{\sqrt{n}} \} \geq 1 - C\exp\{-cl_{n,\alpha}\}. \] (8.24)

Note that \( F_n(x) \) is the distribution function of a random variable which is uniformly distributed on the set \( \{\pm s_1, \ldots, \pm s_n\} \) and
\[ \sup_{x} |F_{nj}(x) - F_n(x)| = \max_{k} \left| \sum_{l=1}^{k} |u_{lj}|^2 - \frac{k}{n} \right|. \] (8.25)

Thus Theorem 1.2 is proved.

9 Appendix

Lemma 9.1. Let \( 0 < y < 1 \). Let \( x : |x| \in [1 - \sqrt{y}, 1 + \sqrt{y}] \) and let \( \gamma := \gamma(x) = \min\{X|x - 1 + \sqrt{y}, 1 + \sqrt{y} - |x|\} \). Then, for \( 0 < y < 1 \),
\[ |G'_y(x)| \leq \frac{3\gamma}{\pi \sqrt{y(1 - \sqrt{y})}}. \] (9.1)

Proof. By equality (2.2), we have
\[ G'_y(x) = \frac{\sqrt{(-1 + \sqrt{y})^2 - x^2}((1 + \sqrt{y}) - x^2)}{2\pi y|x|} \mathbb{1}\{1 - \sqrt{y} \leq |x| \leq 1 + \sqrt{y}\}. \] (9.2)
Assume for definiteness that \( x = -1 + \sqrt{y} - \gamma \). Note that \( 0 \leq \gamma \leq \sqrt{y} < 1 \). It is straightforward to check that

\[
x - 1 + \sqrt{y} = -2 + 2\sqrt{y} - \gamma, \quad x + 1 - \sqrt{y} = -\gamma,
1 + \sqrt{y} - x = 2 + \gamma, \quad 1 + \sqrt{y} + x = 2\sqrt{y} - \gamma.
\] (9.3)

We may write

\[
G_y'(x) = \sqrt{\frac{2\gamma(1 - \sqrt{y} + \frac{1}{2}\gamma)(2 + \gamma)(2\sqrt{y} + \gamma)}{2\pi y(1 - \sqrt{y} + \gamma)}} \leq \frac{3\sqrt{\gamma}}{\pi \sqrt{y}(1 - \sqrt{y})}.
\] (9.4)

Similarly we consider the cases \( x = -1 - \sqrt{y} + \gamma \), \( x = 1 - \sqrt{y} + \gamma \), and \( x = 1 + \sqrt{y} - \gamma \). Thus Lemma 9.1 is proved.

**Lemma 9.2.** For any distribution function \( F \) and for any \( \frac{1}{2} \sqrt{y} > \varepsilon > 0 \) the following inequality holds

\[
\sup_x |F(x) - \tilde{G}_y(x)| \leq \sup_{x : |x| \in [1 - \sqrt{y} + \varepsilon, 1 + \sqrt{y} - \varepsilon]} |F(x) - G(x)| + \frac{2\varepsilon^{\frac{3}{2}}}{\pi \sqrt{y}(1 - \sqrt{y})}.
\] (9.5)

**Proof.** Recall that \( \mathbb{J}_\varepsilon = [1 - \sqrt{y} + \varepsilon, 1 + \sqrt{y} - \varepsilon] \). Note that

\[
\sup_x |F(x) - \tilde{G}_y(x)| = \sup_{x : |x| \in [1 - \sqrt{y}, 1 + \sqrt{y}]} |F(x) - \tilde{G}_y(x)|
\]

\[
= \max \left\{ \sup_{x : |x| \in [-1 - \sqrt{y}, -1 - \sqrt{y} + \varepsilon]} |F(x) - \tilde{G}_y(x)|, \sup_{x : |x| \in [-1 + \sqrt{y} - \varepsilon, -1 + \sqrt{y}]} |F(x) - \tilde{G}_y(x)|, \sup_{x : |x| \in [1 - \sqrt{y}, 1 - \sqrt{y} - \varepsilon]} |F(x) - \tilde{G}_y(x)|, \sup_{x : |x| \in [1 + \sqrt{y} - \varepsilon, 1 + \sqrt{y}]} |F(x) - \tilde{G}_y(x)| \right\}.
\] (9.6)

Furthermore, for \( x \in [-1 - \sqrt{y}, -1 - \sqrt{y} + \varepsilon] \), we have

\[
-\tilde{G}_y(-1 - \sqrt{y} + \varepsilon) \leq F(x) - \tilde{G}_y(x)
\]

\[
\leq F(-1 - \sqrt{y} + \varepsilon) - \tilde{G}_y(-1 - \sqrt{y} + \varepsilon) + \tilde{G}_y(-1 - \sqrt{y} + \varepsilon).
\] (9.7)

Inequality (9.7) implies that

\[
\sup_{x : |x| \in [-1 - \sqrt{y}, -1 - \sqrt{y} + \varepsilon]} |F(x) - \tilde{G}_y(x)| \leq \sup_{|x| \in \mathbb{J}_\varepsilon} |F(x) - \tilde{G}_y(x)| + \tilde{G}_y(-1 - \sqrt{y} + \varepsilon).
\]

Similarly we get

\[
\sup_{x : |x| \in [1 + \sqrt{y} - \varepsilon, 1 + \sqrt{y}]} |F(x) - \tilde{G}_y(x)| \leq \sup_{|x| \in \mathbb{J}_\varepsilon} |F(x) - \tilde{G}_y(x)| + 1 - \tilde{G}_y(1 + \sqrt{y} - \varepsilon).
\]
Furthermore, for $x \in [-1 + \sqrt{y} - \varepsilon, -1 + \sqrt{y}]$ we have

$$F(-1 + \sqrt{y} - \varepsilon) - \tilde{G}_y(-1 + \sqrt{y} - \varepsilon) - (\tilde{G}_y(-1 + \sqrt{y}) - \tilde{G}_y(-1 + \sqrt{y} - \varepsilon)) \leq F(x) - \tilde{G}_y(x) \leq F(1 - \sqrt{y} + \varepsilon) - \tilde{G}_y(1 - \sqrt{y} + \varepsilon) + (\tilde{G}_y(1 - \sqrt{y} + \varepsilon) - \tilde{G}_y(1 - \sqrt{y})). \quad (9.8)$$

We use here that $\tilde{G}(-1 + \sqrt{y}) = \tilde{G}_y(1 - \sqrt{y})$. Inequality (9.8) implies

$$\sup_{x \in [-1 - \sqrt{y}, 1 - \sqrt{y} + \varepsilon]} |F(x) - \tilde{G}_y(x)| \leq \sup_{|x| \in J_\varepsilon} |F(x) - \tilde{G}_y(x)| + \tilde{G}_y(1 - \sqrt{y} + \varepsilon) - \tilde{G}_y(1 - \sqrt{y}).$$

Similarly we get

$$\sup_{x \in [1 - \sqrt{y}, 1 - \sqrt{y} - \varepsilon]} |F(x) - \tilde{G}_y(x)| \leq \sup_{|x| \in J'_\varepsilon} |F(x) - \tilde{G}_y(x)| + \tilde{G}_y(1 - \sqrt{y} + \varepsilon) - \tilde{G}_y(1 - \sqrt{y}).$$

We use here that for $\tilde{G}_y(x)$

$$\tilde{G}_y(1 - \sqrt{y} + \varepsilon) - \tilde{G}_y(1 - \sqrt{y}) = \tilde{G}_y(-1 + \sqrt{y}) - \tilde{G}_y(-1 + \sqrt{y} - \varepsilon). \quad (9.9)$$

These relations together imply

$$\sup_x |F(x) - \tilde{G}_y(x)| \leq \sup_{|x| \in J'_\varepsilon} |F(x) - \tilde{G}_y(x)| + \max\{\tilde{G}_y(-1 - \sqrt{y} + \varepsilon), \tilde{G}_y(-1 + \sqrt{y}) - \tilde{G}_y(-1 + \sqrt{y} - \varepsilon)\}. \quad (9.10)$$

We note as well that, by Lemma 9.1

$$\max\{\tilde{G}_y(-1 - \sqrt{y} + \varepsilon), \tilde{G}_y(-1 + \sqrt{y}) - \tilde{G}_y(-1 + \sqrt{y} - \varepsilon)\} \leq \frac{2\varepsilon^2}{\pi \sqrt{y(1 - \sqrt{y})}}. \quad (9.11)$$

\[\Box\]

**Lemma 9.3.** Let $\xi_1, \ldots, \xi_p$ be independent random variables such that $\mathbb{E}\xi_j = 0$ and $|\xi_j| \leq \sigma_j$. Then

$$\Pr\{|\sum_{j=1}^p \xi_j| > x\} \leq c(1 - \Phi(x/\sigma)) \leq \frac{c\sigma}{x} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}, \quad (9.12)$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{y^2}{2}\right\}dy$ and $\sigma^2 = \sigma_1^2 + \cdots + \sigma_p^2$. The last inequality in (9.12) holds for $x \geq \sigma$.

**Proof.** This inequality is proved, for instance in [3]. \[\Box\]
Lemma 9.4. Let \( \mathcal{A}_0 = \{\emptyset, \Omega\} \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_n \subset \mathcal{A} \) be a family of sub-\(\sigma\)-algebras of the measurable space \( \{\Omega, \mathcal{A}\} \) and let \( M_n = \xi_1 + \cdots + \xi_n \) be a martingale with bounded differences \( \xi_j = M_j - M_{j-1} \) such that

\[
\Pr\{|\xi_j| \leq b_j\} = 1, \quad \text{for} \quad j = 1, \ldots, n.
\]

Then, for \( x > \sqrt{b} \),

\[
\Pr\{|M_n| \geq x\} \leq c(1 - \Phi(\frac{x}{\sigma})) = \int_{-\infty}^{x} \varphi(t) dt, \quad \varphi(t) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} \tag{9.13}
\]

with some numerical constant \( c > 0 \) and \( \sigma^2 = b_1^2 + \cdots + b_n^2 \).

Proof. The result follows from Theorem 1.1 in \[3\]. \(\square\)

Proposition 9.1. Let \( \xi_1, \ldots, \xi_p \) be independent random variables such that \( |\xi_k| \leq 1 \). Let also \( a_{ik} \) be real numbers such that \( a_{ik} = a_{ik} \) and \( a_{kk} = 0 \). Let \( Z = \sum_{i,k=1}^{p} \xi_i \xi_k a_{ik} \). Let \( \sigma^2 = \sum_{i,k=1}^{p} |a_{ik}|^2 \). Then for every \( t > 0 \) there exists some positive constant \( c > 0 \) such that the following inequality holds

\[
\Pr\{|Z| \geq \frac{3}{2} \mathbb{E}^2 |Z|^2 + t\} \leq \exp\left\{-\frac{ct}{\sigma}\right\} \tag{9.14}
\]

Proof. The result follows from Theorem 3.1 in \[10\]. \(\square\)

Remark 9.2. The result of Proposition 9.1 holds for complex \( a_{ij} \). We may consider two quadratic forms with coefficients \( \text{Re} a_{ij} \) and \( \text{Im} a_{ij} \).

Lemma 9.5. For Stieltjes transform \( \tilde{S}_y(z) \) of symmetrized Marchenko–Pastur distribution the following inequalities hold

\[
|S_y(z)| \leq \frac{1}{\sqrt{y}}, \quad |z + \frac{y - 1}{z} + yS_y(z)| \geq \sqrt{y}. \tag{9.15}
\]

Proof. Let \( \tilde{S}_y(z) = \frac{-z - \frac{y - 1}{z} \sqrt{(z + \frac{y - 1}{z})^2 - 4y}}{2y} \). Note that \( S_y(z) \) and \( \tilde{S}_y(z) \) are the roots of equation

\[
yS_y(z)^2 + (z + \frac{y - 1}{z})S_y(z) + 1 = 0. \tag{9.16}
\]

From here it follows

\[
|S_y(z)||\tilde{S}_y(z)| = \frac{1}{y}. \tag{9.17}
\]

Similar to \[2\], Section 3, we note that

\[
\text{sign}\{\text{Re} z + \frac{y - 1}{z}\} = \text{sign}\{\text{Re} \sqrt{(z + \frac{y - 1}{z})^2 - 4y}\}. \tag{9.18}
\]
The rate of convergence to the Marchenko–Pastur distribution

(For more details, see [2], pp. 631–632.) This implies that

\[ |S_y(z)| \leq |\hat{S}_y(z)|. \]

(9.19)

Inequality (9.19) and equality (9.17) together imply the claim. Thus Lemma 9.3 is proved.

\[ \square \]

**Lemma 9.6.** For \( z = u + iv \) with \( 1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y} \) and \( v > 0 \) the following relation holds

\[ \text{Re}\{(z + \frac{y-1}{z})^2 - 4y\} \leq 0. \]

(9.20)

**Proof.** It is straightforward to check that

\[ A := \text{Re}\{(z + \frac{y-1}{z})^2 - 4y\} = u^2(1 - \frac{(1-y)^2}{|z|^2}) - 4y - v^2(1 + \frac{1-y}{|z|^2})^2. \]

(9.21)

We rewrite this equality as

\[ A = (u^2 - v^2)(1 + \frac{(1-y)^2}{|z|^4}) - 2(1 + y) \leq u^2 + \frac{(1-y)^2}{|u|^2} - 2(1 + y). \]

(9.22)

Write \( t = u^2 \) and consider the equation

\[ t^2 - 2(1 + y)t + (1 - y)^2 = 0. \]

(9.23)

Solving it, we find

\[ t_{1,2} = (1 + y) \pm \sqrt{4y} = (1 \pm \sqrt{y})^2. \]

(9.24)

This immediately implies that \( A \leq 0 \), for \( 1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y} \). Thus Lemma 9.6 is proved.

\[ \square \]

**Lemma 9.7.** For any \( z = u + iv \) with \( 1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y} \), the following inequality holds

\[ \text{Im}\sqrt{(z + \frac{y-1}{z})^2 - 4y} \geq \frac{1}{2} \sqrt{(z + \frac{y-1}{z})^2 - 4y} \geq \frac{y^2}{2} \sqrt{\gamma + v}. \]

(9.25)

**Proof.** By Lemma 9.6 for \( z = u + iv \) with \( 1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y} \), we get \( \text{Re}\{(z + \frac{y-1}{z})^2 - 4y\} \leq 0 \) and \( \frac{\pi}{2} \leq \text{arg}\{(z + \frac{y-1}{z})^2 - 4y\} \leq \frac{3\pi}{2} \). Therefore,

\[ \text{Im}\sqrt{(z + \frac{y-1}{z})^2 - 4y} \geq \frac{1}{2} \sqrt{(z + \frac{y-1}{z})^2 - 4y} \geq \frac{y^2}{2} \sqrt{\gamma + v}. \]

(9.26)

Furthermore, we have

\[ (z + \frac{y-1}{z})^2 - 4y = \frac{(z + 1 + \sqrt{y})(z - 1 + \sqrt{y})(z - 1 - \sqrt{y})(z + 1 - \sqrt{y})}{z^2}. \]

(9.27)
Let \( u = -1 - \sqrt{y} + \gamma \). Then, \( |z| \geq 1 \) and
\[
| (z + \frac{y - 1}{z})^2 - 4y | \geq \frac{1}{2|z|^2} (\gamma + v)|z - 1 + \sqrt{y}||z + 1 - \sqrt{y}||z - 1 - \sqrt{y} |
\]
\[
\geq \frac{1}{2|z|^2} (\gamma + v) | - 2 + \gamma + iv || - 2 \sqrt{y} + \gamma + iv || - 2 - 2 \sqrt{y} + \gamma + iv |.
\]
(9.28)

Note that
\[
\frac{| - 2 + \gamma + iv |}{|z|} = \sqrt{\frac{(2 - \gamma)^2 + v^2}{(1 + \sqrt{y} - \gamma)^2 + v^2}} \geq 1,
\]
(9.29)
and
\[
\frac{| - 2 - 2 \sqrt{y} + \gamma + iv |}{|z|} = \sqrt{\frac{(2 + 2 \sqrt{y} - \gamma)^2 + v^2}{(1 + \sqrt{y} - \gamma)^2 + v^2}} \geq 1.
\]
(9.30)

These inequalities together imply
\[
| (z + \frac{y - 1}{z})^2 - 4y | \geq \frac{\sqrt{y}}{2} (\gamma + v).
\]
(9.31)

For \( u = -1 + \sqrt{y} - \gamma \), we have \( |z| \geq 1 - \sqrt{y} + \gamma \) and
\[
| (z + \frac{y - 1}{z})^2 - 4y | \geq \frac{1}{2|z|^2} (\gamma + v) |2 \sqrt{y} - \gamma + iv || - 2 + 2 \sqrt{y} - \gamma + iv || - 2 - \gamma + iv |
\]
(9.32)

Note that
\[
\frac{(2(1 - \sqrt{y}) + \gamma)^2 + v^2}{(1 - \sqrt{y} + \gamma)^2 + v^2} \geq 1,
\]
(9.33)
and
\[
\frac{(2 + \gamma)^2 + v^2}{(1 - \sqrt{y} + \gamma)^2 + v^2} \geq 1.
\]
(9.34)

These inequalities imply
\[
| (z + \frac{y - 1}{z})^2 - 4y | \geq \frac{\sqrt{y}}{2} (\gamma + v).
\]
(9.35)

Similarly we consider \( u = 1 - \sqrt{y} + \gamma \) and \( u = 1 + \sqrt{y} - \gamma \). Thus Lemma 9.7 is proved. \( \square \)

References

[1] Bai, Z. D.; Miao, Baiqi; Yao, Jian-Feng Convergence rates of spectral distributions of large sample covariance matrices. SIAM J. Matrix Anal. Appl. 25 (2003), no. 1, 105-127
The rate of convergence to the Marchenko–Pastur distribution

[2] Bai, Z. D. Convergence rate of expected spectral distributions of large random matrices. I. Wigner matrices. Ann. Probab. 21 (1993), no. 2, 625-648

[3] Bentkus, V. Measure concentration for separately Lipschitz functions in product spaces. Israel Journal of Mathematics, 158 (2007), 1–17.

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

[5] 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.

[6] Götze, F.; Tikhomirov, A. N. The rate of convergence of spectra of sample covariance matrices. Teor. Veroyatn. Primen. 54 (2009), no. 1, 196–206. Translation in Theory Probab. Appl. 54 (2010), no. 1, 129–140.

[7] Götze, F.; Tikhomirov, A. N. The rate of convergence to the semi-circular law. arXiv:1109.0611.

[8] Erdös, L.; Yau; H.-T, Yin, J. Bulk universality for generalized Wigner matrices. arXiv:1001.3453.

[9] Erdös, L.; Yau, H.-T.; Yin, J. Rigidity of eigenvalues of generalized Wigner matrices. arXiv:1007.4652.

[10] Ledoux, M. On Talagrand’s deviation inequalities for product measures. ESAIM: Probability and Statistics 1, (1996), 65–87.

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

[12] Timushev, D. A.; Tikhomirov, A. N.; Kholopov, A. A. On the accuracy of the approximation of the GOE spectrum by the semi-circular law. Teor. Veroyatn. Primen. 52 (2007), no. 1, 180–185.

[13] Tikhomirov, A. N. On the rate of convergence of the expected spectral distribution function of a Wigner matrix to the semi-circular law. Siberian Adv. Math. 19, (2009), no. 3, 211–223.