DOI: 10.1214/EJP.v16-963

Electronic Journal of Probability

Vol. 16 (2011), Paper no. 88, pages 2439–2451.

Journal URL
http://www.math.washington.edu/~ejpecp/

A Note on Rate of Convergence in Probability to Semicircular Law[*]

Zhidong Bai[†] Jiang Hu[‡] Guangming Pan[§] Wang Zhou[¶]

Abstract

In the present paper, we prove that under the assumption of the finite sixth moment for elements of a Wigner matrix, the convergence rate of its empirical spectral distribution to the Wigner semicircular law in probability is $O(n^{-1/2})$ when the dimension n tends to infinity..

Key words: convergence rate, Wigner matrix, Semicircular Law, spectral distribution.

AMS 2010 Subject Classification: Primary 60F15; Secondary: 62H99.

Submitted to EJP on March 18, 2010, final version accepted November 17, 2011.

[*]Z. D. Bai was partially supported by CNSF 10871036. J. Hu was partially supported by the Fundamental Research Funds for the Central Universities 10ssxt149. G.M. Pan was partially supported by a grant M58110052 at the Nanyang Technological University. W. Zhou was partially supported by grant R-155-000-095-112 at the National University of Singapore

[†]KLASMOE and School of Mathematics & Statistics, Northeast Normal University, Changchun, 130024, P.R.C. and Department of Statistics and Applied Probability, National University of Singapore. E-mail: baizd@nenu.edu.cn

[‡]KLASMOE and School of Mathematics & Statistics, Northeast Normal University, Changchun, 130024, P.R.C. E-mail: huj156@nenu.edu.cn

[§]Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University. E-mail: gmpan@ntu.edu.sg

[¶]Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546. E-mail: stazw@nus.edu.sg

1 Introduction and the result.

A Wigner matrix $\mathbf{W}_n = n^{-1/2}\left(x_{ij}\right)_{i,j=1}^{n}$ is defined to be a Hermitian random matrix whose entries on and above the diagonal are independent zero-mean random variables. It is an important model for depicting heavy-nuclei atoms, which began with the seminal work of Wigner in 1955 ([16]). Details in this area can be found in [13].

There are various mathematical tools in the study of Wigner matrices in the past half century (see [1]). One of the most popular instruments is the limit theory of empirical spectral distribution (ESD). Here, for any $n \times n$ matrix $\mathbf{A}$ with real eigenvalues, the ESD of $\mathbf{A}$ is defined by

$$F^{\mathbf{A}}(x) = \frac{1}{n}\sum_{i=1}^{n} I(\lambda_i^{\mathbf{A}} \le x),$$

where $\lambda_i^{\mathbf{A}}$ denotes the i-th smallest eigenvalue of $\mathbf{A}$ and $I(B)$ denotes the indicator function of an event B. It is proved that, under assumptions of for all i, j, $\mathbb{E}|x_{ij}|^2 = \sigma^2$, the ESD $F^{\mathbf{W}_n}(x)$ converges almost surely to a non-random distribution $F(x)$ which has the destiny function

$$f(x) = \frac{1}{2\pi\sigma}\sqrt{4\sigma^2 - x^2}, \quad x \in [-2\sigma, 2\sigma]. \tag{1}$$

This is also known as the Wigner semicircular law (see [16], [6]).

The rate of convergence is important in establishing the central limit theorem for linear spectral statistics of Wigner matrices ([7, 6]). There are some partial results in this area. In [2], Bai proved that under the assumption of $\sup_n \sup_{i,j} \mathbb{E}x_{ij}^4 < \infty$, the rate of

$$\Delta_n = \|\mathbb{E}F^{\mathbf{W}_n} - F\| := \sup_x |\mathbb{E}F^{\mathbf{W}_n}(x) - F(x)|$$

tending to 0 is $O(n^{-1/4})$. Bai et al. in [4] obtained that the rate established in [2] was still valid in probability for

$$\Delta_p = \|F^{\mathbf{W}_n} - F\| := \sup_x |F^{\mathbf{W}_n}(x) - F(x)|$$

Under a stronger condition that $\sup_n \sup_{i,j} \mathbb{E}x_{ij}^8 < \infty$, Bai et al. in [5] showed that $\Delta_n = O(n^{-1/2})$ and $\Delta_p = O_p(n^{-2/5})$ (Bai and Silverstein improved this condition up to $\sup_n \sup_{i,j} \mathbb{E}x_{ij}^6 < \infty$ in their book [6]). Here and in the sequel, the notation $R_n = O_p(r_n)$ means for any $\varepsilon > 0$, there exists an $C > 0$ such that $\sup_n \mathbb{P}(|R_n| \ge C r_n) < \varepsilon$. Later, Götze et al. in [10] derived $\Delta_n = O(n^{-1/2})$ as well assuming fourth moment, and $\Delta_p = O_p(n^{-1/2})$ at the cost of the twelfth moment of the matrix entries. There are some other results with some special assumptions on the matrix entries. If the entries of $\mathbf{W}_n$ have a normal distribution, then the optimal order $\Delta_n = O(n^{-1})$ was shown in [11]. When the distribution of the entries satisfies a Poincare inequality or a uniformly subexponential decay, the order of Δ_p can be improved to $O_p(n^{-2/3}\log^2 n)$ and $O_p(n^{-1}\log^C n)$ with some constant C respectively. For which one can refer to [9, 8].

In this note we prove that the twelfth moment condition in [10] could be reduced to the sixth moment assumption while still getting $\Delta_p = O_p(n^{-1/2})$. Our main result of this paper is as follows.

Theorem 1.1. *Assume that*

- $\mathbb{E}x_{ij} = 0$, *for all* $1 \le i \le j \le n$,
- $\mathbb{E}|x_{ii}^2| = \sigma^2 > 0, \mathbb{E}|x_{ij}|^2 = 1$, *for all* $1 \le i < j \le n$,
- $\sup_n \sup_{1\le i<j\le n} \mathbb{E}|x_{ii}^3|, \mathbb{E}|x_{ij}|^6 < \infty$.

Then we have

$$\Delta_p := \|F^{\mathbf{W}_n} - F\| = O_p(n^{-1/2}). \tag{2}$$

Remark 1.2. *It is not clear what the exact rate and the optimal conditions are in Theorem 1.1.*

The rest of this paper is organized as follows. The main tool of proving the theorem is introduced in Section 2. Theorem 1.1 is proved in Section 3 and some technical lemmas are given in Section 4. Throughout this paper, constants appearing in inequalities are represented by C which are nonrandom and may take different values from one appearance to another.

2 The main tool

For any function of bounded variation H on the real line, its Stieltjes transform is defined by

$$s_H(z) = \int \frac{1}{\lambda - z} dH(\lambda), \quad z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C}^+ : \Im z > 0\}.$$

Our main tool to prove the theorem is a Berry-Esseen type inequality which is proved in [2].

Lemma 2.1. *(Bai inequality) Let F be a distribution function and let G be a function of bounded variation satisfying $\int |F(x) - G(x)|dx < \infty$. Denote their Stieltjes transforms by $s_F(z)$ and $s_G(z)$ respectively, where $z = u + iv \in \mathbb{C}^+$. Then*

$$\begin{aligned}\|F - G\| \le \frac{1}{\pi(1-\zeta)(2\rho-1)} &\left(\int_{-A}^{A} |s_F(z) - s_G(z)|du \right.\\ &+ 2\pi v^{-1} \int_{|x|>B} |F(x) - G(x)|dx \\ &\left. + v^{-1} \sup_x \int_{|u|\le 2v\epsilon} |G(x+u) - G(x)|du \right),\end{aligned}$$

where the constants $A > B > 0$, ζ and ϵ are restricted by $\rho = \frac{1}{\pi}\int_{|u|\le\epsilon} \frac{1}{u^2+1} du > \frac{1}{2}$, and $\zeta = \frac{4B}{\pi(A-B)(2\rho-1)} \in (0,1)$.

Here we should notice that we can use the same methods in [10] to prove our theorem. However, Götze-Tikhomirov inequality (see Corollary 2.3 in [10]) involves the supremum of $|s_n(z) - \mathbb{E}s_n(z)|$ over $\Im z$ in some interval. This makes the proof rather complicated. Therefore in this paper, we use Bai inequality instead of Götze-Tikhomirov inequality which could make the presentation simpler.

3 The proof of Theorem 1.1.

We will firstly introduce a new technique which can handle the moment conditions efficiently. That is given in Lemma 3.2. Then, by using this lemma and dividing the expression of $\mathbb{E}|s_n - \mathbb{E}s_n|^2$, we prove our theorem step by step.

Before proving the theorem, we introduce some notation. Denote $\mathbf{I}_n$ be the identity matrix of size n and $\mathbf{a}_i$ be the ith column of $\mathbf{W}_n$ with x_{ii} removed. Define $\mathbf{D}(z) = n^{-1/2}\mathbf{W}_n - z\mathbf{I}_n$, $\mathbf{D}_i(z) = \mathbf{D}(z) - n^{-1}\mathbf{a}_i\mathbf{a}_i^*$ and $s_n = s_n(z) = s_{F^{\mathbf{W}_n}}(z)$. Moreover write

$$\beta_i = \left(n^{-1/2}x_{ii} - z - n^{-1}\mathbf{a}_i^*\mathbf{D}_i^{-1}\mathbf{a}_i\right)^{-1}, \quad \gamma_i = \mathbf{a}_i^*\mathbf{D}_i^{-1}\mathbf{a}_i - tr\mathbf{D}_i^{-1}$$
$$\varepsilon_i = n^{-1/2}x_{ii} - n^{-1}\mathbf{a}_i^*\mathbf{D}_i^{-1}\mathbf{a}_i + \mathbb{E}s_n(z), \quad \hat{\gamma}_i = \mathbf{a}_i^*\mathbf{D}_i^{-2}\mathbf{a}_i - tr\mathbf{D}_i^{-2}$$
$$\xi_i = tr\mathbf{D}^{-1} - tr\mathbf{D}_i^{-1}, \quad a_n = (z + \mathbb{E}s_n(z))^{-1}.$$

Throughout this section, we denote $z = u + iv$, $u \in [-16, 16]$ and $1 \geq v \geq v_0 = C_0 n^{-1/2}$ with an appropriate constant C_0. Let $s = s(z) = s_F(z)$, we know that (see (3.2) in [2])

$$s(z) = -\frac{1}{2}\left(z - \sqrt{z^2 - 4}\right) \text{ for all } z \in \mathbb{C}^+.$$

Then we have

$$\int_{-16}^{16} \frac{1}{|z + 2s(z)|} du \leq \int_{-16}^{16} \frac{1}{\sqrt{|z^2 - 4|}} du \leq \int_{-16}^{16} \frac{1}{\sqrt{|u^2 - 4|}} du < 10. \tag{3}$$

In addition, by Lemma 2.1 and Theorem 8.2 in [6], we have for some positive constant C,

$$\mathbb{E}\|F^{\mathbf{W}_n} - F\| \leq C \int_{-16}^{16} \mathbb{E}|s_n(z) - \mathbb{E}s_n(z)| du + O(n^{-1/2}). \tag{4}$$

Therefore, the rest of the proof is reduced to the lemma below.

Lemma 3.1. *Under the assumptions in Theorem 1.1, for any $1 > v \geq v_0 = C_0 n^{-1/2}$ with sufficiently large $C_0 > 0$, we have*

$$\mathbb{E}\left|s_n(z) - \mathbb{E}s_n(z)\right|^2 \leq \frac{C}{n|z + 2s(z)|^2}.$$

3.1 Known results and a preliminary lemma

Following the same truncation, centralization and rescaling steps in [6], in this section we may assume the random variables satisfy the conditions as follows

$$|x_{ij}| \leq n^{1/4}, \quad \mathbb{E}x_{ij} = 0, \quad \mathbb{E}|x_{ij}|^2 = 1 \text{ for all } i, j.$$

Bai in [2] derived the equation

$$s_n(z) = \frac{1}{n} tr\mathbf{D}^{-1} = \frac{1}{n}\sum_{i=1}^{n} \beta_i,$$

which together with the fact

$$\beta_i = -a_n + a_n \beta_i \varepsilon_i, \tag{5}$$

implies

$$s_n(z) = -a_n + \frac{a_n}{n} \sum_{i=1}^{n} \beta_i \varepsilon_i. \tag{6}$$

For each i we have

$$|\Im \beta_i^{-1}| = |\Im \left(z + n^{-1} \mathbf{a}_i^* \mathbf{D}_i^{-1} \mathbf{a}_i \right)| \geq v.$$

Thus we have

$$|\beta_i| \leq v^{-1}. \tag{7}$$

From the definition of ε_i it follows that

$$\varepsilon_i = n^{-1/2} x_{ii} - n^{-1} \gamma_i + n^{-1} \xi_i - (s_n - \mathbb{E} s_n), \tag{8}$$

and

$$s_n = -a_n + \frac{a_n}{n^{3/2}} \sum_{i=1}^{n} \beta_i x_{ii} + \frac{a_n}{n^2} \sum_{i=1}^{n} \beta_i \gamma_i + \frac{a_n}{n^2} \sum_{i=1}^{n} \beta_i \xi_i - a_n (s_n - \mathbb{E} s_n) s_n. \tag{9}$$

Then, we have the the following lemma.

Lemma 3.2. *Under the assumption in Theorem 1.1, we have for any $v > v_0$*

$$\mathbb{P}\left(|\beta_i| > 2\right) \leq \frac{C}{n^2 v^2}. \tag{10}$$

Proof. From integration by parts and Theorem 1.1 in [10], we have for $1 > v > v_0$,

$$\begin{aligned} |\mathbb{E} s_n(z) - s(z)| &= \left| \int_{-\infty}^{\infty} \frac{d(\mathbb{E} F^{\mathbf{W}_n}(x) - F(x))}{x - z} \right| \\ &= \left| \int_{-\infty}^{\infty} \frac{\mathbb{E} F^{\mathbf{W}_n}(x) - F(x)}{(x - z)^2} dx \right| \leq C, \end{aligned}$$

which together with the fact that $|s(z)| \leq 1$ (see (3.3) in [2]) implies

$$|\mathbb{E} s_n(z)| \leq C.$$

Then applying Lemma 4.2 and Lemma 4.3 we have for $1 > v > v_0$,

$$\mathbb{E}|s_n(z)| \leq C \quad \text{and} \quad \frac{1}{n} \mathbb{E}|tr \mathbf{D}_i^{-1}| \leq C.$$

By Lemma 4.1 we can check that

$$\begin{aligned}\mathbb{E}|\gamma_i|^4 &\le C\mathbb{E}\left(\left(tr\mathbf{D}_i^{-1}(\mathbf{D}_i^{-1})^*\right)^2 + n^{1/2}tr\left(\mathbf{D}_i^{-1}(\mathbf{D}_i^{-1})^*\right)^2\right)\\ &\le C\left(v^{-2}\mathbb{E}|tr\mathbf{D}_i^{-1}|^2 + n^{1/2}v^{-3}\mathbb{E}|tr\mathbf{D}_i^{-1}|\right)\\ &\le \frac{Cn^2}{v^2}. \end{aligned}\tag{11}$$

Thus, from (8), Lemma 4.2 and Lemma 4.3 we have for $v > v_0$,

$$\mathbb{E}|\varepsilon_i|^4 \le \frac{C}{n^2v^2}.\tag{12}$$

In addition, by the proof of Lemma 5.1 and (5.59) in [10], we can see that

$$|a_n| \le 1 \text{ for all } v \ge v_0.\tag{13}$$

Therefore from the equation (5) we can obtain

$$\mathbb{P}\left(|\beta_i| > 2\right) \le \mathbb{P}\left(|a_n\varepsilon_i| > \frac{1}{2}\right) \le 2^4\mathbb{E}|\varepsilon_i|^4 \le \frac{C}{n^2v^2},$$

which complete the proof. □

3.2 The proof of Lemma 3.1

Notice that in this subsection, we will use the equality (5) and (8) frequently. From (9), we have

$$\begin{aligned}\mathbb{E}\left|s_n - \mathbb{E}s_n\right|^2 &= \mathbb{E}(\overline{s_n - \mathbb{E}s_n})(s_n - \mathbb{E}s_n)\\ =&\mathbb{E}(\overline{s_n - \mathbb{E}s_n})s_n = a_n(S_1 + S_2 + S_3 + S_4),\end{aligned}$$

where

$$S_1 = \frac{1}{n^{3/2}}\sum_{i=1}^n \mathbb{E}(\overline{s_n - \mathbb{E}s_n})x_{ii}\beta_i, \qquad S_2 = -\frac{1}{n^2}\sum_{i=1}^n \mathbb{E}(\overline{s_n - \mathbb{E}s_n})\gamma_i\beta_i,$$

$$S_3 = \frac{1}{n^2}\sum_{i=1}^n \mathbb{E}(\overline{s_n - \mathbb{E}s_n})\xi_i\beta_i, \qquad S_4 = -\mathbb{E}|s_n - \mathbb{E}s_n|^2 s_n.$$

We first consider S_1. From (5), we have

$$\begin{aligned}S_1 &= \frac{1}{n^{3/2}}\sum_{i=1}^n \mathbb{E}(\overline{s_n - \mathbb{E}s_n})x_{ii}\beta_i\\ &= \frac{a_n}{n^{3/2}}\sum_{i=1}^n \mathbb{E}\left(-(\overline{s_n - \mathbb{E}s_n})x_{ii} + a_n\mathbb{E}(\overline{s_n - \mathbb{E}s_n})x_{ii}\varepsilon_i - a_n(\overline{s_n - \mathbb{E}s_n})x_{ii}\beta_i\varepsilon_i^2\right)\\ &= \frac{a_n}{n^{3/2}}\sum_{i=1}^n \mathbb{E}\left(-S_{11} + S_{12} - S_{13}\right).\end{aligned}$$

By Lemma 4.2 we have

$$|\mathbb{E}S_{11}| = \left|\frac{1}{n}\mathbb{E}\xi_i x_{ii}\right| \le \frac{\mathbb{E}|x_{ii}|}{nv} = O\left(\frac{1}{nv}\right). \tag{14}$$

From (12), (13), Hölder's inequality and Lemma 4.3, we obtain

$$|\mathbb{E}S_{12}| \le \left(\mathbb{E}|s_n - \mathbb{E}s_n|^4 \mathbb{E}|\varepsilon_i|^4 (\mathbb{E}|x_{ii}|^2)^2\right)^{1/4} \le \frac{C}{n^{3/2}v^2}. \tag{15}$$

Next we consider the term S_{13}. Using (12), (13), Lemma 3.2, Lemma 4.3 and the fact $|\beta_i| \le v^{-1}$ we have

$$\begin{aligned}
|\mathbb{E}S_{13}| \le\, & 2\left(\mathbb{E}|s_n - \mathbb{E}s_n|^4 (\mathbb{E}|\varepsilon_i|^4)^2 \mathbb{E}|x_{ii}|^4\right)^{1/4} \\
& + v^{-1}\left(\mathbb{E}|s_n - \mathbb{E}s_n|^4 (\mathbb{E}|\varepsilon_i|^4)^2\right)^{1/4}\left(\mathbb{E}|x_{ii}|^2 I(|\beta_i| > 2)\right)^{1/2} \\
\le\, & \frac{C}{nv}.
\end{aligned} \tag{16}$$

Therefore combining inequalities (13)-(16) we obtain

$$|S_1| = O\left(\frac{1}{n}\right). \tag{17}$$

Furthermore, we have the following expression for S_2,

$$\begin{aligned}
S_2 &= -\frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}(\overline{s_n - \mathbb{E}s_n})\gamma_i\beta_i \\
&= \frac{a_n}{n^2}\sum_{i=1}^{n}\mathbb{E}(\overline{s_n - \mathbb{E}s_n})\gamma_i - \frac{a_n^2}{n^2}\sum_{i=1}^{n}\mathbb{E}(\overline{s_n - \mathbb{E}s_n})\gamma_i\varepsilon_i\beta_i \\
&= S_{21} + S_{22} + S_{23} + S_{24} + S_{25},
\end{aligned}$$

where

$$S_{21} = \frac{a_n}{n^2}\sum_{i=1}^{n}\mathbb{E}(\overline{s_n - n^{-1}tr\mathbf{D}_i})\gamma_i, \qquad S_{22} = -\frac{a_n}{n^{5/2}}\sum_{i=1}^{n}\mathbb{E}(\overline{s_n - \mathbb{E}s_n})x_{ii}\gamma_i\beta_i,$$

$$S_{23} = \frac{a_n}{n^3}\sum_{i=1}^{n}\mathbb{E}(\overline{s_n - \mathbb{E}s_n})\beta_i\gamma_i^2, \qquad S_{24} = -\frac{a_n}{n^3}\sum_{i=1}^{n}\mathbb{E}(\overline{s_n - \mathbb{E}s_n})\gamma_i\beta_i\xi_i,$$

$$S_{25} = \frac{a_n}{n^2}\sum_{i=1}^{n}\mathbb{E}|s_n - \mathbb{E}s_n|^2\gamma_i\beta_i.$$

Here we will use the method which we used to handle the bound of S_1. Firstly, we express S_{21} as follows

$$\begin{aligned}
S_{21} &= \frac{a_n}{n^3}\sum_{i=1}^{n}\mathbb{E}(\overline{(1 + n^{-1}\mathbf{a}_i^*\mathbf{D}_i^{-2}\mathbf{a}_i)\beta_i})\gamma_i \\
&= S_{211} + S_{212},
\end{aligned}$$

where

$$S_{211} = -\frac{|a_n|^2}{n^4}\sum_{i=1}^n \mathbb{E}(\overline{\hat{\gamma}_i})\gamma_i \qquad\qquad S_{212} = \frac{|a_n|^2}{n^3}\sum_{i=1}^n \mathbb{E}(\overline{(1+n^{-1}\mathbf{a}_i^*\mathbf{D}_i^{-2}\mathbf{a}_i)\beta_i\varepsilon_i})\gamma_i.$$

From Lemma 4.1 and Hölder's inequality we get

$$|S_{211}| \le \frac{C}{n^4}\sum_{i=1}^n \left(\mathbb{E}|\hat{\gamma}_i|^2\mathbb{E}|\gamma_i|^2\right)^{1/2} \le \frac{C}{n^2v^2}.$$

Applying Lemma 4.2, Hölder's inequality and (12), we obtain

$$S_{212} = \frac{|a_n|^2}{n^2}\left|\sum_{i=1}^n \mathbb{E}(\overline{s_n - n^{-1}tr\mathbf{D}_i^{-1}\varepsilon_i})\gamma_i\right| \le \frac{C}{n^3 v}\sum_{i=1}^n \left(\mathbb{E}|\varepsilon_i|^2\mathbb{E}|\gamma_i|^2\right)^{1/2} \le \frac{C}{n^2v^2}.$$

From the last two inequalities we obtain

$$|S_{21}| = O\left(\frac{1}{n^2v^2}\right). \tag{18}$$

For S_{22}, we use Lemma 4.2 to get

$$|S_{22}| \le \frac{C}{n^{7/2}}\sum_{i=1}^n \mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}||x_{ii}\gamma_i\beta_i| + \frac{C}{n^{7/2}v}\sum_{i=1}^n \mathbb{E}|x_{ii}\gamma_i\beta_i|.$$

Notice that x_{ii} and γ_i are independent. From Hölder's inequality and Lemma 3.2 we have

$$\mathbb{E}|x_{ii}\gamma_i\beta_i| \le C\mathbb{E}|x_{ii}\gamma_i| + \left(\mathbb{E}|x_{ii}\gamma_i|^2\mathbb{E}|\beta_i I(|\beta_i|>2)|^2\right)^{1/2} = O(n^{1/2}v^{-1/2}).$$

Similarly we can get

$$\mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}||x_{ii}\gamma_i\beta_i| \le \left(\mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}|^4\mathbb{E}|\gamma_i|^4\right)^{1/4} = O(n^{1/2}v^{-2}),$$

which implies

$$|S_{22}| \le \frac{C}{n^2v^2}. \tag{19}$$

Now consider S_{23}. Using Lemma 4.2 again we obtain

$$\begin{aligned}|S_{23}| &\le \frac{C}{n^3}\sum_{i=1}^n \mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}||\gamma_i^2\beta_i| + \frac{C}{n^3 v}\sum_{i=1}^n \mathbb{E}|\gamma_i^2\beta_i| \\ &\le \frac{C}{n^3}\sum_{i=1}^n \mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}||\gamma_i^2\beta_i| + \frac{C}{n^2v^2}.\end{aligned}$$

Applying Lemma 3.2 and Hölder's inequality we obtain

$$\begin{aligned}\mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}||\gamma_i^2\beta_i| &\le 2\mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}||\gamma_i^2| \\ &+ \left((\mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}|^2|\gamma_i|^4)\mathbb{E}|\beta_i I(|\beta_i|>2)|^2\right)^{1/2} \\ &\le \left(\mathbb{E}|tr\mathbf{D}_i^{-1} - \mathbb{E}tr\mathbf{D}_i^{-1}|^2|\gamma_i|^4\right)^{1/2}.\end{aligned}$$

It follows from (11) that

$$
\begin{aligned}
&\mathbb{E}|tr\mathbf{D}_i^{-1}-\mathbb{E}tr\mathbf{D}_i^{-1}|^2|\gamma_i|^4\\
\leq&C\mathbb{E}|tr\mathbf{D}_i^{-1}-\mathbb{E}tr\mathbf{D}_i^{-1}|^2\left(v^{-2}|tr\mathbf{D}_i^{-1}|^2+n^{1/2}v^{-3}|tr\mathbf{D}_i^{-1}|\right)\\
\leq&Cv^{-2}\mathbb{E}|tr\mathbf{D}_i^{-1}-\mathbb{E}tr\mathbf{D}_i^{-1}|^4+\frac{Cn^2}{v^2}\mathbb{E}|tr\mathbf{D}_i^{-1}-\mathbb{E}tr\mathbf{D}_i^{-1}|^2\\
\leq&\frac{Cn^2}{v^2}\mathbb{E}|tr\mathbf{D}_i^{-1}-\mathbb{E}tr\mathbf{D}_i^{-1}|^2+\frac{C}{v^8}\\
\leq&\frac{Cn^2}{v^2}\mathbb{E}|tr\mathbf{D}^{-1}-\mathbb{E}tr\mathbf{D}^{-1}|^2+\frac{C}{v^8}.
\end{aligned}
$$

Then, we conclude that

$$
|S_{23}|\leq\frac{C}{nv}\left(\mathbb{E}\left|s_n-\mathbb{E}s_n\right|^2\right)^{1/2}+\frac{C}{n^2v^2}. \tag{20}
$$

From Lemma 3.2, Lemma 4.2, Lemma 4.3 and Hölder's inequality, it is easy to check that

$$
|S_{24}|\leq\frac{C}{n^2v}\sum_{i=1}^{n}\left(\mathbb{E}|s_n-\mathbb{E}s_n|^4\mathbb{E}|\gamma_i|^4\right)^{1/4}\left(\mathbb{E}|\beta_i|^2I(|\beta_i|>2)\right)^{1/2}\leq\frac{C}{n^2v^2}. \tag{21}
$$

For S_{25}, we use (5) to represent it in the form

$$
\begin{aligned}
S_{25}=&-\frac{a_n^2}{n^2}\sum_{i=1}^{n}\mathbb{E}|s_n-\mathbb{E}s_n|^2\gamma_i+\frac{a_n^2}{n^2}\sum_{i=1}^{n}\mathbb{E}|s_n-\mathbb{E}s_n|^2\gamma_i\varepsilon_i\beta_i\\
=&-S_{251}+S_{252}.
\end{aligned}
$$

Using Lemma 4.3 and Hölder's inequality we obtain

$$
\begin{aligned}
|S_{251}|\leq&\frac{C}{n^4}\sum_{i=1}^{n}\mathbb{E}|\xi_i-\mathbb{E}\xi_i|^2|\gamma_i|+\frac{C}{n^4}\sum_{i=1}^{n}\mathbb{E}|\xi_i-\mathbb{E}\xi_i||tr\mathbf{D}_i^{-1}-\mathbb{E}tr\mathbf{D}_i^{-1}||\gamma_i|\\
\leq&\frac{C}{n^{5/2}v^{5/2}}+\frac{C}{n^{5/2}v^3}=O\left(\frac{1}{n^2v^2}\right).
\end{aligned} \tag{22}
$$

Similarly we can obtain that

$$
\begin{aligned}
\mathbb{E}|s_n-\mathbb{E}s_n|^2\gamma_i\varepsilon_i\beta_i\leq&\left(\mathbb{E}(|s_n-\mathbb{E}s_n|^8|\gamma_i|^4)\mathbb{E}|\varepsilon_i|^4\right)^{1/4}\left(2+\mathbb{E}|\beta_i|^2I(|\beta_i|>2)\right)^{1/2}\\
\leq&\frac{C}{n^{1/2}v^{1/2}}\left(\mathbb{E}(|s_n-\mathbb{E}s_n|^8|\gamma_i|^4)\right)^{1/4}\\
\leq&\frac{C}{n^{1/2}v^{1/2}}\left(n^{-8}\mathbb{E}(|\xi_i-\mathbb{E}\xi_i|^8|\gamma_i|^4)+n^{-8}\mathbb{E}(|tr\mathbf{D}_i^{-1}-\mathbb{E}tr\mathbf{D}_i^{-1}|^8|\gamma_i|^4)\right)^{1/4}\\
\leq&\frac{C}{n^3v^4}.
\end{aligned}
$$

From the last inequality and (22) we obtain

$$
|S_{25}|\leq\frac{C}{n^2v^2}. \tag{23}
$$

Combining inequalities (18)-(21) and (23), we conclude that, for $v \geq v_0$

$$|S_2| \leq \frac{C}{nv}\left(\mathbb{E}\left|s_n - \mathbb{E}s_n\right|^2\right)^{1/2} + \frac{C}{n^2v^2}. \tag{24}$$

From Lemma 3.2, Lemma 4.3 and Hölder's inequality, we can check that

$$\begin{aligned}|S_3| \leq & \frac{C}{n^2v}\sum_{i=1}^{n}\left(\mathbb{E}|s_n - \mathbb{E}s_n|^2(2+\mathbb{E}|\beta_i|^2 I(|\beta_i|>2))\right)^{1/2} \\ \leq & \frac{C}{nv}\left(\mathbb{E}|s_n - \mathbb{E}s_n|^2\right)^{1/2}.\end{aligned} \tag{25}$$

Therefore, it remians to get the bound of S_4. Now we recall the equality (9), then we have

$$S_4 = a_n\mathbb{E}|s_n - \mathbb{E}s_n|^2 - a_n(S_{41}+S_{42}+S_{43}+S_{44}),$$

and

$$S_4 = -\mathbb{E}s_n\mathbb{E}|s_n - \mathbb{E}s_n|^2 - \mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n). \tag{26}$$

Here

$$S_{41} = \frac{1}{n^{3/2}}\sum_{i=1}^{n}\mathbb{E}|s_n - \mathbb{E}s_n|^2 x_{ii}\beta_i, \quad S_{42} = -\frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}|s_n - \mathbb{E}s_n|^2\gamma_i\beta_i,$$

$$S_{43} = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}|s_n - \mathbb{E}s_n|^2\xi_i\beta_i,$$

$$S_{44} = -\mathbb{E}s_n\mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n) - \mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n)^2.$$

Comparing (26) with S_{44}, we obtain that

$$\begin{aligned}&(1+a_n\mathbb{E}s_n)\mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n) \\ =& -(a_n + \mathbb{E}s_n)\mathbb{E}|s_n - \mathbb{E}s_n|^2 \\ &+ a_n(S_{41}+S_{42}+S_{43} - \mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n)^2),\end{aligned}$$

which implies that

$$\begin{aligned}&-\mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n) \\ =& b_n a_n^{-1}(a_n + \mathbb{E}s_n)\mathbb{E}|s_n - \mathbb{E}s_n|^2 \\ &- b_n(S_{41}+S_{42}+S_{43} - \mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n)^2),\end{aligned}$$

where $b_n = (z + 2\mathbb{E}s_n(z))^{-1}$. Thus denoting $\delta_n = n^{-1}\sum_{i=1}^{n}\mathbb{E}\beta_i\varepsilon_i$, we conclude that

$$\begin{aligned}S_4 =& (-\mathbb{E}s_n + b_n a_n^{-1}(a_n + \mathbb{E}s_n))\mathbb{E}|s_n - \mathbb{E}s_n|^2 \\ &- b_n(S_{41}+S_{42}+S_{43} - \mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n)^2) \\ =& (a_n - \delta_n b_n\mathbb{E}s_n)\mathbb{E}|s_n - \mathbb{E}s_n|^2 \\ &- b_n(S_{41}+S_{42}+S_{43} - \mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n)^2) \\ =& (a_n + a_n\delta_n b_n)\mathbb{E}|s_n - \mathbb{E}s_n|^2 \\ &- b_n(\delta_n^2\mathbb{E}|s_n - \mathbb{E}s_n|^2 + S_{41}+S_{42}+S_{43} - \mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n)^2).\end{aligned}$$

It is obvious that S_{42} and S_{25} have the same bound. Using Lemma 3.2, Lemma 4.2 and Lemma 4.3 we get the following three inequalities

$$|\mathbb{E}|s_n - \mathbb{E}s_n|^2(s_n - \mathbb{E}s_n)^2| \le \mathbb{E}|s_n - \mathbb{E}s_n|^4 \le \frac{C}{n^4 v^6},$$

$$|S_{43}| \le \frac{1}{nv}\left(\mathbb{E}|s_n - \mathbb{E}s_n|^4\right)^{1/2} \le \frac{C}{n^3 v^4},$$

and

$$\begin{aligned}|\mathbb{E}|s_n - \mathbb{E}s_n|^2 x_{ii}\beta_i| &\le |\mathbb{E}|s_n - \mathbb{E}s_n|^2 x_{ii} a_n| + |\mathbb{E}|s_n - \mathbb{E}s_n|^2 x_{ii} a_n \varepsilon_i \beta_i| \\ &\le \frac{C}{n^2 v^3} + \left(\mathbb{E}|s_n - \mathbb{E}s_n|^{16}\mathbb{E}|x_{ii}|^8\right)^{1/8}\left(\mathbb{E}|\varepsilon_i|^4\right)^{1/4}\left(2 + \mathbb{E}|\beta_i|^2 I(|\beta_i| > 2)\right)^{1/2} \\ &= O\left(\frac{1}{n^2 v^3}\right).\end{aligned}$$

Furthermore, from the definition of δ_n and (12), we have

$$|\delta_n| = \left|n^{-1}\sum_{i=1}^{n}\left(\mathbb{E}n^{-1}\mathbf{D}_i^{-1} - \mathbb{E}s_n + \mathbb{E}\beta_i\varepsilon_i^2\right)\right| \le \frac{C}{nv}.$$

Therefore, we obtain

$$S_4 = a_n\mathbb{E}|s_n - \mathbb{E}s_n|^2 + O\left(\frac{|b_n|}{n^2 v^2}\right),$$

which combined with (17), (24) and (25) implies

$$|1 - a_n^2|\mathbb{E}|s_n - \mathbb{E}s_n|^2 \le \frac{C_1|a_n b_n|}{n} + \frac{C_2|a_n|}{\sqrt{n}}\left(\mathbb{E}|s_n - \mathbb{E}s_n|^2\right)^{1/2}.$$

Then, from (6.91) and (6.95) in [10] which are under the existing fourth moment assumption, for $1 > v > v_0$,

$$|1 - a_n^2| \ge |a_n(z + 2s(z))| \text{ and } |b_n| \le 2|z + 2s(z)|^{-1},$$

we obtain the following inequality

$$\mathbb{E}|s_n - \mathbb{E}s_n|^2 \le \frac{C_1}{n|z + 2s(z)|^2} + \frac{C_2}{\sqrt{n}|z + 2s(z)|}\left(\mathbb{E}|s_n - \mathbb{E}s_n|^2\right)^{1/2}.$$

Solving this inequality, we obtain

$$\mathbb{E}|s_n - \mathbb{E}s_n|^2 \le \frac{C}{n|z + 2s(z)|^2},$$

which complete the proof of the Lemma.

4 Basic lemmas

In this section we list some results which are needed in the proof.

Lemma 4.1. *(Lemma B.26 of [6]) Let* $\mathbf{A}$ *be an* $n \times n$ *nonrandom matrix and* $\mathbf{X} = (x_1, \ldots, x_n)^*$ *be a random vector of independent entries. Assume that* $\mathbb{E}x_i = 0$, $\mathbb{E}|x_i|^2 = 1$, *and* $E|x_j|^l \le \nu_l$. *Then, for any* $p \ge 1$,

$$\mathbb{E}|\mathbf{X}^*\mathbf{A}\mathbf{X} - tr\mathbf{A}|^p \le C_p \left((\nu_4 tr(\mathbf{A}\mathbf{A}^*))^{p/2} + \nu_{2p} tr(\mathbf{A}\mathbf{A}^*)^{p/2} \right),$$

where C_p *is a constant depending on* p *only.*

Lemma 4.2. *(Lemma 2.6 of [14]). Let* $z \in \mathbb{C}^+$ *with* $v = \Im z$, $\mathbf{A}$ *and* $\mathbf{B}$ $n \times n$ *with* $\mathbf{B}$ *Hermitian,* $\tau \in \mathbb{R}$, *and* $\mathbf{q} \in \mathbb{C}^N$. *Then*

$$|tr((\mathbf{B} - z\mathbf{I})^{-1} - (\mathbf{B} + \tau\mathbf{q}\mathbf{q}^* - z\mathbf{I})^{-1})\mathbf{A}| \le \frac{\|\mathbf{A}\|}{v}.$$

Lemma 4.3. *(Lemma 8.7 of [6]) Under the assumption in Theorem 1.1, we have for any* $l \ge 1$

$$\mathbb{E}|s_n(z) - \mathbb{E}s_n(z)|^{2l} \le \frac{C}{n^{2l}v^{3l}}. \tag{27}$$

Acknowledgement: The authors would like to thank the referee for his/her helpful suggestions.

References

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*. Cambridge University Press, 2010. MR2760897

[2] Z. D. Bai. Convergence rate of expected spectral distributions of large random matrices. Part I. Wigner matrices. *The Annals of Probability*, 21(2):625–648, 1993. MR1217559

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

[4] Z. D. Bai, B. Q. Miao, and J. Tsay. A note on the convergence rate of the spectral distributions of large random matrices. *Statistics* & *Probability Letters*, 34:95–101, 1997. MR1457501

[5] Z. D. Bai, B. Q. Miao, and J. Tsay. Convergence rates of the spectral distributions of large Wigner matrices. *International Mathematical Journal*, 1:65–90, 2002. MR1825933

[6] Z. D. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices. Second Edition*. Springer Verlag, 2010. MR2567175

[7] Z. D. Bai, X. Y. Wang, and W. Zhou. CLT for linear spectral statistics of Wigner matrices. *Electronic Journal of Probability*, 14:2391–2417, 2009. MR2556016

[8] L. Erdős, H-T. Yau, and J. Yin. Rigidity of eigenvalues of generalized Wigner matrices. Preprint. arXiv:1007.4652v4

[9] S. G. Bobkov, F. Götze, and A. N. Tikhomirov. On concentration of empirical measures and convergence to the semi-circle law. *Journal of Theoretical Probability*, Apr. 2010.

[10] F. Götze and A. Tikhomirov. Rate of convergence to the semi-circular law. *Probability Theory and Related Fields*, 127(2):228–276, 2003. MR2013983

[11] F. Götze and A. Tikhomirov. The rate of convergence for spectra of GUE and LUE matrix ensembles. *Central European Journal of Mathematics*, 3(4):666–704, Dec. 2005. MR2171668

[12] F. Götze, A. N. Tikhomirov, and D. a. Timushev. Rate of convergence to the semi-circle law for the Deformed Gaussian Unitary Ensemble. *Central European Journal of Mathematics*, 5(2):305–334, June 2007. MR2300275

[13] M. L. Mehta. *Random matrices, Third Edition*. Academic Press, 2004. MR2129906

[14] J. W. Silverstein and Z. D. Bai. On the empirical distribution of eigenvalues of a class of largedimensional random matrices. *Journal of Multivariate Analysis*, 54(2):175–192, Aug. 1995. MR1345534

[15] A. Tikhomirov. On the rate of convergence of the expected spectral distribution function of a Wigner matrix to the semi-circular law. *Siberian Advances in Mathematics*, 19(3):211–223, 2009. MR2655022

[16] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62(3):548–564, 1955. MR0077805