Universality for a global property of the eigenvectors of Wigner matrices

Bao, Zhigang; Pan, Guangming; Zhou, Wang

2014

Bao, Z., Pan, G., & Zhou, W. (2014). Universality for a global property of the eigenvectors of Wigner matrices. *Journal of Mathematical Physics*, 55(2), 023303-.

<https://hdl.handle.net/10356/103709>

<https://doi.org/10.1063/1.4864735>

© 2014 AIP Publishing LLC. This paper was published in *Journal of Mathematical Physics* and is made available as an electronic reprint (preprint) with permission of AIP Publishing LLC. The paper can be found at the following official DOI:

<http://dx.doi.org/10.1063/1.4864735>. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.

Downloaded on 26 Apr 2023 21:39:19 SGT

Universality for a global property of the eigenvectors of Wigner matrices

Zhigang Bao, Guangming Pan, and Wang Zhou

Citation: *Journal of Mathematical Physics* **55**, 023303 (2014); doi: 10.1063/1.4864735

View online: <http://dx.doi.org/10.1063/1.4864735>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jmp/55/2?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Local deformed semicircle law and complete delocalization for Wigner matrices with random potential](#)
J. Math. Phys. **54**, 103504 (2013); 10.1063/1.4823718

[Ensemble of transition states for two-state protein folding from the eigenvectors of rate matrices](#)
J. Chem. Phys. **121**, 9186 (2004); 10.1063/1.1802674

[An algorithm for eigenvalues and eigenvectors of quaternion matrices in quaternionic quantum mechanics](#)
J. Math. Phys. **45**, 3334 (2004); 10.1063/1.1769106

[Duals of colored quantum universal enveloping algebras and colored universal T-matrices](#)
J. Math. Phys. **39**, 1199 (1998); 10.1063/1.532378

[Construction of a complete set of states in relativistic scattering theory](#)
J. Math. Phys. **38**, 115 (1997); 10.1063/1.531845

AIP | Journal of
Applied Physics

Journal of Applied Physics is pleased to
announce **André Anders** as its new Editor-in-Chief

Universality for a global property of the eigenvectors of Wigner matrices

Zhigang Bao,^{1,a)} Guangming Pan,^{2,b)} and Wang Zhou^{3,c)}

¹Department of Mathematics, Zhejiang University, People's Republic of China

²Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

³Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546, Singapore

(Received 13 November 2013; accepted 25 January 2014; published online 12 February 2014)

Let M_n be an $n \times n$ real (resp. complex) Wigner matrix and $U_n \Lambda_n U_n^*$ be its spectral decomposition. Set $(y_1, y_2, \dots, y_n)^T = U_n^* \mathbf{x}$, where $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is a real (resp. complex) unit vector. Under the assumption that the elements of M_n have 4 matching moments with those of GOE (resp. GUE), we show that the process $X_n(t) = \sqrt{\frac{\beta n}{2}} \sum_{i=1}^{\lfloor nt \rfloor} (|y_i|^2 - \frac{1}{n})$ converges weakly to the Brownian bridge for any $\mathbf{x}$ satisfying $\|\mathbf{x}\|_\infty \rightarrow 0$ as $n \rightarrow \infty$, where $\beta = 1$ for the real case and $\beta = 2$ for the complex case. Such a result indicates that the orthogonal (resp. unitary) matrices with columns being the eigenvectors of Wigner matrices are asymptotically Haar distributed on the orthogonal (resp. unitary) group from a certain perspective. © 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4864735>]

I. INTRODUCTION

Let $M_n = \frac{1}{\sqrt{n}}(v_{ij})_{n,n}$ be an $n \times n$ real or complex Wigner matrix whose definition is stated below.

Definition 1.1 (Real Wigner matrix). We call M_n a real Wigner matrix if it is a symmetric random matrix such that $\{v_{ij}, 1 \leq i \leq j \leq n\}$ is a collection of independent real random variables. Here $v_{ij}, 1 \leq i < j \leq n$ are i.i.d with common mean 0 and variance 1, and $v_{ii}, 1 \leq i \leq n$ are i.i.d with common mean 0 and finite variance.

Definition 1.2 (Complex Wigner matrix). We call M_n a complex Wigner matrix if it is an Hermitian random matrix such that $\{v_{ij}, 1 \leq i \leq j \leq n\}$ is a collection of independent random variables. Here $v_{ij}, 1 \leq i < j \leq n$ are i.i.d complex random variables with common mean 0 and variance 1. And $v_{ii}, 1 \leq i \leq n$ are i.i.d real random variables with common mean 0 and finite variance.

Example 1.1 (GOE and GUE). A real Wigner matrix M_n is called GOE if $v_{ij} \sim N(0, 1), i \neq j$, and $v_{ii} \sim N(0, 2)$. And a complex Wigner matrix M_n is called GUE if $v_{ij} \sim N(0, 1/2) + \sqrt{-1}N(0, 1/2), i < j$ and $v_{ii} \sim N(0, 1)$. Here $N(0, 1/2) + \sqrt{-1}N(0, 1/2)$ stands for the standard complex normal variable whose real and imaginary parts are independent.

Now we denote the ordered eigenvalues of M_n by $\lambda_1 \leq \dots \leq \lambda_n$ and the corresponding normalized eigenvectors by $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$. Set $\Lambda_n = \text{diag}(\lambda_1, \dots, \lambda_n)$ and write the spectral decomposition of M_n as

$$M_n = U_n \Lambda_n U_n^*,$$

a)E-mail: zhigangbao@zju.edu.cn

b)E-mail: gmpan@ntu.edu.sg

c)E-mail: stazw@nus.edu.sg

where

$$U_n = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n).$$

Conventionally, we require the coefficients of the eigenvectors to be real in the real case. However, the choices of the normalized eigenvectors are not unique owing to the following two reasons.

(r1): If there is an $i \in \{1, 2, \dots, n\}$ such that λ_i is not simple, one can arbitrarily choose an orthogonal basis of the eigenspace corresponding to λ_i .

(r2): If every eigenvalue is simple, we can still rotate the eigenvector by multiplying a sign -1 in the real case or any phase $e^{\sqrt{-1}\theta}$ ($\theta \in \mathbb{R}$) in the complex case.

Note that when the matrix elements are continuously distributed, **(r1)** will cause no ambiguity since the eigenvalues are simple with probability one in this case. Moreover, even for the discontinuous case, it will be shown that whichever eigenvectors for the multiple eigenvalues have been taken will not affect the results of this paper (see Sec. II B for details). In addition, **(r2)** will actually cause no trouble since the concerned quantities in this paper only depend on the projections

$$\mathbf{u}_i \mathbf{u}_i^*, \quad i = 1, \dots, n,$$

which are uniquely defined as long as the eigenvalues are simple. However, in order to eliminate the ambiguities in some issues in the sequel, we will adopt the following viewpoint to fix the definition of the eigenvector $\mathbf{u}_i$ when λ_i is simple.

A viewpoint: In the complex case, one can replace $\mathbf{u}_i$ by $e^{\sqrt{-1}\theta_i} \mathbf{u}_i$, where θ_i is uniformly distributed on $[0, 2\pi)$. Moreover, $\theta_i, i = 1, \dots, n$ are i.i.d and independent of the matrix M_n . In the real case, one can replace $\mathbf{u}_i$ by $b_i \mathbf{u}_i$, where $b_i, i = 1, \dots, n$ are i.i.d ± 1 Bernoulli variables which are independent of M_n .

Under the above viewpoint, it is well known that U_n is Haar distributed on the orthogonal group $O(n)$ (resp. unitary group $U(n)$) when M_n is GOE (resp. GUE). Then for more general Wigner matrices, it is natural to ask whether U_n is “asymptotically” Haar distributed in some sense. In other words, we are concerned with the universality problem of the eigenvectors. More specifically, we want to study whether the universality property holds for some statistic constructed from U_n . The word *universality* means that the limiting behavior of the statistic concerned depends only on the symmetry class of the random matrix but not on the distribution details of its entries. In the past few decades, a lot of work has been devoted to the study of the universality properties of various statistics constructed from the eigenvalues. By contrast, the work on the universality properties of eigenvector statistics is much less. For recent progresses on this aspect, we refer to the delocalization or localization property of the eigenvectors (see Refs. 7, 8, 3, 5, 6, and 13 for instance) and the universality for the local statistics of the eigenvector coefficients (see Refs. 11 and 17).

In this paper, we will establish a universality result for a global property of the eigenvectors. Below we give the definition of the main object of our paper. Let $\mathbf{x} = (x_1, \dots, x_n)^T$ be a definite unit vector. That is to say, $\mathbf{x} \in S^{n-1}$ in the real case, and $\mathbf{x} \in S^{2n-1}$ in the complex case. Here

$$S^{n-1} := \{\mathbf{r} \in \mathbb{R}^n : \|\mathbf{r}\| = 1\}, \quad S^{2n-1} := \{\mathbf{z} \in \mathbb{C}^n : \|\mathbf{z}\| = 1\}.$$

Now we set the vector

$$\mathbf{y} = (y_1, \dots, y_n)^T = U_n^* \mathbf{x}.$$

We are concerned with the process $X_n(t) \in D[0, 1]$ constructed from the vector $\mathbf{y}$ as

$$X_n(t) = \sqrt{\frac{\beta n}{2}} \sum_{i=1}^{\lfloor nt \rfloor} (|y_i|^2 - \frac{1}{n}), \quad (1.1)$$

where $\beta = 1$ for the real case and $\beta = 2$ for the complex case. Hereafter, the notation $\lfloor x \rfloor$ stands for integer part of x . In this paper, we will discuss the limit of the process (1.1) in the weak sense. Such a

problem was raised by Silverstein in Ref. 14 and was shown to be closely related to the universality problem on U_n .

Provided that U_n is Haar distributed on the orthogonal group $O(n)$ (resp. unitary group $U(n)$), it is well known that for any real (resp. complex) unit vector $\mathbf{x}$ one has $\mathbf{y}$ is uniformly distributed on S^{n-1} (resp. S^{2n-1}). Then in the real case, one has

$$\mathbf{y} \stackrel{d}{=} \frac{\mathbf{g}_{\mathbb{R}}}{\|\mathbf{g}_{\mathbb{R}}\|},$$

where $\mathbf{g}_{\mathbb{R}} := (g_1, g_2, \dots, g_n)^T$ is a Gaussian vector with i.i.d. $N(0, 1)$ coefficients. Similarly, in the complex case one has

$$\mathbf{y} \stackrel{d}{=} \frac{\mathbf{g}_{\mathbb{C}}}{\|\mathbf{g}_{\mathbb{C}}\|},$$

where $\mathbf{g}_{\mathbb{C}} := (\eta_1 + \sqrt{-1}\zeta_1, \dots, \eta_n + \sqrt{-1}\zeta_n)^T$ is a Gaussian vector with i.i.d. $N(0, 1/2) + \sqrt{-1}N(0, 1/2)$ coefficients. Then by using the classical results on weak convergence, it is straightforward to check that in the Gaussian case,

$$X_n(t) \stackrel{d}{\Rightarrow} W^\circ(t), \quad t \in [0, 1], \quad (1.2)$$

where $W^\circ(t)$ is the standard Brownian bridge.

Conversely, (1.2) reflects the fact that $\mathbf{y}$ is uniformly distributed on sphere in a global sense. Analogously, if (1.2) is valid for general Wigner matrices with a large class of $\mathbf{x}$, we can regard that U_n in general case is “asymptotically” Haar distributed from such a certain perspective. Such a measure of closeness between the distribution of U_n and the Haar distribution was originally raised by Silverstein in Ref. 14 for the sample covariance matrices. However, Silverstein only succeeded in proving the result for the unit vectors $\mathbf{x} = (\pm 1/\sqrt{n}, \dots, \pm 1/\sqrt{n})^T$ under the additional assumption that the matrix elements are symmetrically distributed. As explained above, in the Gaussian case, $\mathbf{x}$ can be arbitrary. Therefore, it is crucial to verify (1.2) for more general $\mathbf{x}$ rather than those in Ref. 14. In this paper, we will prove (1.2) for a large class of Wigner matrices under the slight restriction of $\|\mathbf{x}\|_\infty \rightarrow 0$ which will be shown to be necessary in the general case (see Remark 3.5 below). Here $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$ is the maximum norm of $\mathbf{x}$. In addition, we do not need the symmetrical distribution condition imposed on the matrix elements.

To state our main result, we need an *ad hoc* terminology.

Definition 1.3 (Matching to the k th moments). We say that two Wigner matrices $M_n = 1/\sqrt{n}(v_{ij})$ and $\tilde{M}_n = 1/\sqrt{n}(\tilde{v}_{ij})$ match to the k th moments if for all $1 \leq i, j \leq n$,

$$\mathbb{E}(\Re(v_{ij})^l \Im(v_{ij})^m) = \mathbb{E}(\Re(\tilde{v}_{ij})^l \Im(\tilde{v}_{ij})^m), \quad 0 \leq l, m \leq l+m \leq k.$$

In the sequel, we will specify $k = 4$. That means we require the elements of two Wigner matrices involved have the same first four moments. Moreover, throughout the paper, we will need the following additional condition on the matrix elements.

Condition 1.4. We assume that the matrix elements v_{ij} 's have a uniform subexponential decay. That is,

$$\mathbb{P}(|v_{ij}| \geq t) \leq C^{-1} \exp(-t^C)$$

with some positive constant C independent of i, j .

Now we can state our main result.

Theorem 1.5. Assume that M_n is a real (resp. complex) Wigner matrix matching GOE (resp. GUE) to the 4th moments. Moreover, we assume that M_n satisfies Condition 1.4. For any definite

real (resp. complex) unit vector $\mathbf{x}$ satisfying $\|\mathbf{x}\|_\infty \rightarrow 0$ as n tends to infinity, we have

$$X_n(t) = \sqrt{\frac{\beta n}{2}} \sum_{i=1}^{\lfloor nt \rfloor} (|y_i|^2 - \frac{1}{n}) \implies W^\circ(t).$$

Here $\beta = 1, 2$ in the real case and complex case, respectively.

Hereafter, when we refer to “limit” and “accumulation point” of a random sequence, they are always in the sense of weak convergence. Moreover, for simplicity, when there is no confusion we may omit the time parameter t from $X_n(t)$ and $W^\circ(t)$.

The main proof strategy will benefit from the discussions in Ref. 14. Specifically, a criteria for the weak convergence of a random sequence on $D[0, 1]$ with its limit supported on $C[0, 1]$ was provided in Ref. 14 (see Theorem 3.1 therein). Such a criteria can be regarded as a slight modification of the classical “*finitedimensionalconvergence + tightness*” issue. The discussions in Ref. 14 and the recent result of Bai and Pan¹ can help us to confirm that the unique possible $C[0, 1]$ -supported accumulation point of $(X_n)_{n \geq 1}$ is W° . What remains therefore is to show that $(X_n)_{n \geq 1}$ is tight and can only have $C[0, 1]$ -supported accumulation point. However, it has been shown in Ref. 14 that the proof of the tightness of the sequence $(X_n)_{n \geq 1}$ is an obstacle for this kind of problem. In order to show the tightness, Silverstein imposed the additional symmetrical distribution condition on the matrix elements and restricted the discussions on the special case of $\mathbf{x} = (\pm 1/\sqrt{n}, \dots, \pm 1/\sqrt{n})^T$ in Ref. 14 for the sample covariance matrices. Actually, Silverstein’s proof strategy can also be used on the Wigner matrices under similar restrictions as those imposed in Ref. 14.

In order to remove the restrictions mentioned above, we will use a totally different method. A main new input is the so-called *isotropic local semicircle law* proposed by Knowles and Yin in Ref. 12 quite recently. Crudely speaking, we can verify the tightness of $(X_n)_{n \geq 1}$ through providing some good upper bounds on the fourth moments of the increments of $X_n(t)$. Such bounds will turn out to be easily obtained for the Gaussian case owing to the explicit distribution information of $\mathbf{y}$. For more general Wigner matrices, we will use the idea of comparing the general case with the Gaussian case. Such a comparison method relies on the celebrated Lindeberg strategy. More specifically, one need to replace the matrix elements by those of the “reference” matrix one pair (or one unit in the diagonal case) each time and then evaluate the change of the quantity concerned induced by the replacement on each step. Then by a telescoping argument, one can get the difference of the quantities concerned of two Wigner matrices. Such an approach was successfully employed in the literature of the Random Matrix Theory recently. One can see Refs. 15, 9, and 10 for instance. Particularly, one can refer to Refs. 11, 17, and 2 for the applications of the comparison strategy on some problems about the eigenvectors of the Wigner matrices.

More precisely, to provide the upper bounds on the fourth moments of the increments of $X_n(t)$, we will mainly pursue the idea of the Green function comparison approach raised by Erdős, Yau, and Yin in Ref. 9. To this end, first, we will approximate the increment of the process by a quantity expressed in terms of the Green function. Then we will perform a replacement issue on the Green functions to achieve the purpose of comparison. It will be clear that the isotropic local semicircle law provided in Ref. 12 will serve as a main technical tool in pursuing the Green function comparison approach on our problem.

Our paper will be organized as follows. In Sec. II, we will present some necessary preliminaries. And in Sec. III, we will provide a criteria of the weak convergence of $X_n(t)$ which contains two statements. It will be shown that the first statement can be implied by a recent result of Bai and Pan,¹ thus we will just sketch the proof of this statement at the end of Sec. III. The second statement is mainly about the tightness of the sequence $(X_n)_{n \geq 1}$, which will be handled in Sec. IV.

Throughout the paper, the notations C , C_1 , C' , and K will be used to denote some n -independent positive constants whose values may defer from line to line. The notation $\|\cdot\|_{op}$ stands for the operator norm of a matrix. Moreover, we will use $M_n(\mathbb{C})$ to represent the space of $n \times n$ matrices with entries in $\mathbb{C}$. We will say an event $\mathbf{E}$ occurs with overwhelming probability if and only if

$$\mathbb{P}(\mathbf{E}) \geq 1 - n^{-K}$$

for any given positive number K when n is sufficiently large. In addition, we will conventionally use $\mathbf{1}_E$ to denote the indicator function of the event E .

II. PRELIMINARIES

In this section, we will state some basic notions and recent results, especially Knowles and Yin's isotropic semicircle law and isotropic delocalization property which will be frequently used in the proof of Theorem 1.5. In addition, relying on these known results, we will show that the ambiguity caused by $(\mathbf{r}\mathbf{1})$ can be cleared up.

A. Basic notions and known results

The so-called empirical spectral distribution (ESD) of M_n is defined by

$$F_n(x) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\lambda_i \leq x\}}.$$

It is well known that $F_n(x)$ almost surely converges weakly to Wigner's semicircle law $F_{sc}(x)$ whose density function is given by

$$\rho_{sc}(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{\{|x| \leq 2\}}. \quad (2.1)$$

The Stieltjes transform of a probability measure μ can be defined for all complex number $z = E + \sqrt{-1}\eta \in \mathbb{C} \setminus \mathbb{R}$ as

$$m^\mu(z) = \int \frac{1}{x - z} \mu(dx).$$

Here E and η are the real and imaginary parts of z , respectively. Thus by definition, we have

$$m^{F_n}(z) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda_i - z} = \frac{1}{n} \text{tr}(M_n - zI_n)^{-1} \quad (2.2)$$

and

$$m^{F_{sc}}(z) = \int_{-2}^2 \frac{1}{x - z} \rho_{sc}(x) dx.$$

In order to simplify the notation, we will briefly write $m^{F_n}(z)$ and $m^{F_{sc}}(z)$ as $m_n(z)$ and $m_{sc}(z)$, respectively. We remind here the well known fact that

$$\sup_{z \in \mathbb{C} \setminus \mathbb{R}} |m_{sc}(z)| = O(1).$$

The Green function $G_n(z)$ of M_n is defined by

$$G_n(z) := (G_{ij}(z))_{n,n} = (M_n - zI_n)^{-1},$$

which is also called the resolvent of M_n . Then it follows from (2.2) that

$$m_n(z) = \frac{1}{n} \text{tr} G_n(z) = \frac{1}{n} \sum_{i=1}^n G_{ii}(z).$$

Now we set

$$\mathbf{S} := \{E + \sqrt{-1}\eta : |E| \leq 5, n^{-1}(\log n)^{C \log \log n} < \eta \leq 10\} \quad (2.3)$$

for some positive constant C . It was shown by Erdős, Yau, and Yin in Ref. 10 that when $z \in \mathbf{S}$, $m_n(z)$ can be well approximated by $m_{sc}(z)$ with overwhelming probability. Moreover, it was proved in

Ref. 10 that the diagonal entries of the Green function $G_{ii}(z)$'s are close to $m_{sc}(z)$ and the off-diagonal entries $G_{jk}(z)$'s ($j \neq k$) are small in the sense that for some positive constant C ,

$$\max_i |G_{ii}(z) - m_{sc}(z)| + \max_{j \neq k} |G_{jk}(z)| \leq (\log n)^{C \log \log n} \left(\sqrt{\frac{\Im m_{sc}(z)}{n\eta}} + \frac{1}{n\eta} \right) \quad (2.4)$$

holds uniformly in $z \in \mathbf{S}$ with overwhelming probability. See Theorem 2.1 of Ref. 10 for details.

Now we denote the standard basis of $\mathbb{R}^n$ by $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ conventionally, i.e., $\mathbf{e}_i$ is the $n \times 1$ vector with only the i th component being 1 and the others being 0. Then we can write

$$G_{ij}(z) = \mathbf{e}_i^* G_n(z) \mathbf{e}_j.$$

Recently, Knowles and Yin generalized the estimation (2.4) to the quantities

$$G_{\mathbf{v}\mathbf{w}} := \mathbf{v}^* G_n(z) \mathbf{w}$$

for any definite unit vectors $\mathbf{v}, \mathbf{w}$ in Ref. 12 and provided the so-called isotropic local semicircle law. Meanwhile, for any unit vector $\mathbf{v}$ they also provided in Ref. 12 the uniform upper bounds for the quantities

$$|\langle \mathbf{u}_i, \mathbf{v} \rangle|, i = 1, \dots, n.$$

And they named the control on the quantities above as the isotropic delocalization of the eigenvectors, which can be viewed as a generalization of the delocalization property for eigenvectors raised in Ref. 7. Both the isotropic local semicircle law and isotropic delocalization property will be crucial to our analysis in the sequel. We remark here the assumptions imposed in Ref. 12 are weaker than those made in our paper. We refer to Ref. 12 for details and will not mention this fact again in the sequel. For convenience, we will reformulate their results as the following lemma under our assumptions.

Lemma 2.1 (Knowles and Yin (Ref. 12)). Under the assumptions of Theorem 1.5, we have the following two statements.

(1): (Isotropic local semicircle law). For $z \in \mathbf{S}$, there exists some positive constant C such that

$$|\langle \mathbf{v}, G_n(z) \mathbf{w} \rangle - m_{sc}(z) \langle \mathbf{v}, \mathbf{w} \rangle| \leq (\log n)^{C \log \log n} \left(\sqrt{\frac{\Im m_{sc}(z)}{n\eta}} + \frac{1}{n\eta} \right) \quad (2.5)$$

with overwhelming probability for all deterministic and normalized vectors $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$.

(2): (Isotropic delocalization). For any deterministic and normalized vector $\mathbf{v} \in \mathbb{C}^n$, we have

$$\sup_i |\langle \mathbf{u}_i, \mathbf{v} \rangle|^2 \leq \frac{(\log n)^{C \log \log n}}{n} \quad (2.6)$$

for some positive constant C with overwhelming probability.

Remark 2.2. We remind here that the validity of (2.6) does not depend on the ambiguity of the choices of the eigenvectors caused by $(\mathbf{r}1)$ and $(\mathbf{r}2)$. For details, see the proof of Theorem 2.5 of Ref. 12.

Remark 2.3. Actually, given any deterministic normalized vectors $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$, (2.5) holds simultaneously for all $z \in \mathbf{S}$ with overwhelming probability. That is to say, we can strengthen (2.5) to be the statement that the event

$$\bigcap_{z \in \mathbf{S}} \left\{ |\langle \mathbf{v}, G_n(z) \mathbf{w} \rangle - m_{sc}(z) \langle \mathbf{v}, \mathbf{w} \rangle| \leq (\log n)^{C \log \log n} \left(\sqrt{\frac{\Im m_{sc}(z)}{n\eta}} + \frac{1}{n\eta} \right) \right\}$$

holds with overwhelming probability. To see this, we need the following elementary bound:

$$|m'_{sc}(z)|, \|G'(z)\|_{op} \leq \eta^{-2}. \quad (2.7)$$

Now we choose an ε -net of $\mathbf{S}$ with $\varepsilon = n^{-K}$, where $K > 0$ is some sufficiently large constant. Then it is apparent that (2.5) holds simultaneously on the ε -net with overwhelming probability. By using (2.7) and the elementary mean value theorem, we can get that (2.5) holds simultaneously for all $z \in \mathbf{S}$ with overwhelming probability by slightly adjusting the constant C in (2.5).

B. Elimination of the ambiguity

With the aid of the isotropic delocalization property, we can explain in this subsection that the ambiguity caused by $(\mathbf{r}\mathbf{1})$ does not influence the limit property of $X_n(t)$. Now let $\gamma_i := \gamma_{i,n} \in [-2, 2]$ be the classical location of λ_i in the sense that

$$\int_{-2}^{\gamma_i} \rho_{sc}(x) dx = \frac{i}{n}.$$

It is easy to check that

$$|\gamma_i - \gamma_{i+1}| \leq C[\min(i, n-i+1)]^{-1/3} n^{-2/3}, \quad i = 1, \dots, n-1 \quad (2.8)$$

for some positive constant C . By using the rigidity property of the eigenvalues proved by Erdős, Yau, and Yin in Ref. 10, we see that with overwhelming probability, the event

$$\bigcap_{i=1}^n \{|\lambda_i - \gamma_i| \leq (\log n)^{C_1 \log \log n} [\min(i, n-i+1)]^{-1/3} n^{-2/3}\} \quad (2.9)$$

holds for some positive constant C_1 when n is sufficiently large. Now we assume that there is an n_0 such that

$$\lambda_{n_0} < \lambda_{n_0+1} = \dots = \lambda_{n_0+k} < \lambda_{n_0+k+1}.$$

Observe that although the eigenvectors $\mathbf{u}_{n_0+1}, \dots, \mathbf{u}_{n_0+k}$ can be chosen in many different ways, the choice of the projection matrix

$$(\mathbf{u}_{n_0+1}, \dots, \mathbf{u}_{n_0+k})(\mathbf{u}_{n_0+1}, \dots, \mathbf{u}_{n_0+k})^*$$

is unique, which yields that the quantity

$$\sum_{i=n_0+1}^{n_0+k} (|y_i|^2 - \frac{1}{n}) = \mathbf{x}^*(\mathbf{u}_{n_0+1}, \dots, \mathbf{u}_{n_0+k})(\mathbf{u}_{n_0+1}, \dots, \mathbf{u}_{n_0+k})^* \mathbf{x} - \frac{k}{n}$$

is uniquely defined. This implies that the definition of $X_n(t)$ does not depend on the choice of the eigenvectors as long as $\lambda_{[nt]}$ is a simple eigenvalue. Now if $\lambda_{[nt]}$ is not simple, we can assume that $n_0 + 1 \leq [nt] \leq n_0 + k$ without loss of generality. Following from (2.8) and (2.9), it is not difficult to see that with overwhelming probability, there is no eigenvalue with multiplicity larger than $(\log n)^{2C_1 \log \log n}$. For $n_0 + 1 \leq [nt] \leq n_0 + k$, we can write

$$X_n(t) = X_n(n_0/n) + \sqrt{\frac{\beta n}{2}} \sum_{i=n_0+1}^{[nt]} (|y_i|^2 - \frac{1}{n}). \quad (2.10)$$

Then by the fact that $k \leq (\log n)^{2C_1 \log \log n}$ with overwhelming probability and the isotropic delocalization property (2.6), we see that the second term on the right hand side of (2.10) (not well-defined term) can be discarded in probability. Moreover, since both the upper bound of the multiplicity of eigenvalue and the isotropic delocalization property hold uniformly in $i \in \{1, \dots, n\}$, the above discussion also holds uniformly in $t \in [0, 1]$. Hence, we can get the conclusion that the limit behavior of $X_n(t)$ does not depend on the ambiguity caused by $(\mathbf{r}\mathbf{1})$.

III. UNIQUENESS OF THE $C[0, 1]$ -SUPPORTED ACCUMULATION POINT

In this section, we will provide some known results on the process concerned and mainly show that W° is the unique $C[0, 1]$ -supported accumulation point of $(X_n)_{n \geq 1}$. Analogous to the discussions in Ref. 2, to prove Theorem 1.5, it suffices to verify the following two lemmas.

Lemma 3.1. Under the assumptions of Theorem 1.5, the sequence $(X_n)_{n \geq 1}$ has W° as its unique possible accumulation point supported on $C[0, 1]$.

Lemma 3.2. Under the assumptions of Theorem 1.5, the sequence $(X_n)_{n \geq 1}$ is tight and can only have $C[0, 1]$ -supported accumulation points.

The remaining part of this section will be devoted to the proof of Lemma 3.1. To this end, we will need Theorem 1.2 of Ref. 1. For the convenience of the reader, we rewrite it here. Let

$$W_n(g) = \sqrt{\beta n}(\mathbf{x}^* g(M_n) \mathbf{x} - \frac{1}{n} \text{tr} g(M_n)),$$

where $g(x)$ is a function analytic on a region in the complex plane including the interval $[-2, 2]$. Then we have the following theorem provided by Bai and Pan in Ref. 1.

Theorem 3.3 (Bai and Pan (Ref. 1)). *Let M_n be a real or complex Wigner matrix satisfying $\mathbb{E}|v_{12}|^4 < \infty$. Suppose that $g_1, \dots, g_k$ are analytic on an open interval including $[-2, 2]$, and that*

$$\|\mathbf{x}\|_\infty \rightarrow 0.$$

(1) If M_n is real, i.e., $\beta = 1$, and $\mathbb{E}v_{12}^3 = 0$, then $W_n(g_1), \dots, W_n(g_k)$ converges weakly to a Gaussian vector W_f with mean zero and covariance function

$$\text{Cov}(W_{g_1}, W_{g_2}) = 2 \left(\int g_1(x)g_2(x)dF_{sc}(x) - \int g_1(x)dF_{sc}(x) \int g_2(x)dF_{sc}(x) \right). \quad (3.1)$$

(2) If M_n is complex, i.e., $\beta = 2$, and $\mathbb{E}v_{12}^2 = 0$, while $\mathbb{E}v_{12}^2 \bar{v}_{12} = 0$, then (1) remains true.

Remark 3.4. We remind here that $W_n(g)$ in the complex case is different from $X_n(g)$ in Ref. 1 in scaling.

Remark 3.5. It has been shown that the condition $\|\mathbf{x}\|_\infty \rightarrow 0$ is necessary for Theorem 3.3. See Remark 1.4 of Ref. 1.

Now we commence the proof of Lemma 3.1.

Proof of Lemma 3.1. Changing the variable t by $F_{sc}(u)$, one see that it is equivalent to verify the sequence $(X_n(F_{sc}(u)))_{n \geq 1}$ has $W^\circ(F_{sc}(u))$ as its unique possible accumulation point supported on $C[-2, 2]$. We claim that it suffices to show the following two statements.

(a): *We have*

$$\left\{ \int_{-2}^2 u^r X_n(F_{sc}(u)) du \right\}_{r=0}^\infty \Longrightarrow \left\{ \int_{-2}^2 u^r W^\circ(F_{sc}(u)) du \right\}_{r=0}^\infty.$$

(b): *The distribution of a process $X(u)$ supported on $C[-2, 2]$ is uniquely determined by the distribution of*

$$\left\{ \int_{-2}^2 u^r X(u) du \right\}_{r=0}^\infty.$$

Below we sketch the proof of Lemma 3.1 providing (a) and (b) at first. Supposing that one convergent subsequence $\{X_{n'}(F_{sc}(u))\}$ converges weakly to some $C[-2, 2]$ -supported process $X(u)$, by

Theorem 5.1 of Ref. 4, one has

$$\left\{ \int_{-2}^2 u^r X_{n'}(F_{sc}(u)) du \right\}_{r=0}^{\infty} \Longrightarrow \left\{ \int_{-2}^2 u^r X(u) du \right\}_{r=0}^{\infty}.$$

Meanwhile, by (b) we also know that if $X(u)$ is $C[-2, 2]$ -supported, its distribution is uniquely determined by the distribution of

$$\left\{ \int_{-2}^2 u^r X(u) du \right\}_{r=0}^{\infty}.$$

Thus we have $\{X_{n'}(F_{sc}(u))\}$ converges weakly to $W^\circ(F_{sc}(u))$ as $n' \rightarrow \infty$. That means each convergent subsequence of $\{X_n(F_{sc}(u))\}$ converges weakly to $W^\circ(F_{sc}(u))$. Consequently, we get Lemma 3.1 by (a) and (b).

It remains to verify (a) and (b). The proof of (b) is nearly the same as the counterpart in the proof of Theorem 3.1 of Ref. 14. Thus here we will not reproduce the details. To verify (a) for $X_n(F_{sc}(u))$, we will work on its slight modification $X_n(F_n(u))$ instead. Note that by the rigidity property proved by Erdős, Yau, and Yin in Ref. 10, we can see that for some positive constant C there exists

$$\sup_{|u| \leq 5} |F_n(u) - F_{sc}(u)| \leq \frac{(\log n)^{C \log \log n}}{n} \quad (3.2)$$

with overwhelming probability (see Theorem 2.2 of Ref. 10). Consequently, we have for some positive constant C' ,

$$\sup_{|u| \leq 5} |X_n(F_{sc}(u)) - X_n(F_n(u))| \leq \sqrt{n} (\log n)^{C \log \log n} \max_i (|y_i|^2 + \frac{1}{n}) \leq \frac{(\log n)^{C' \log \log n}}{\sqrt{n}}$$

with overwhelming probability. Above we have used the isotropic delocalization property (2.6). Therefore, it suffices to study the limit behavior of

$$\left\{ \int_{|u| \leq 5} u^r X_n(F_n(u)) du \right\}_{r=0}^{\infty}.$$

Moreover, by using the rigidity property in Ref. 10 again, we observe that all the eigenvalues of M_n are in the interval $[-5, 5]$ with overwhelming probability. Combining this observation with the fact that

$$X_n(0) = X_n(1) = 0,$$

we have

$$\left\{ \int_{|u| \leq 5} u^r X_n(F_n(u)) du \right\}_{r=0}^{\infty} = \left\{ \int_{-\infty}^{+\infty} u^r X_n(F_n(u)) du \right\}_{r=0}^{\infty}$$

with overwhelming probability. Relying on the discussion above one can transfer the problem to show that

$$\left\{ \int_{-\infty}^{\infty} u^r X_n(F_n(u)) du \right\}_{r=0}^{\infty} \Longrightarrow \left\{ \int_{-2}^2 u^r W^\circ(F_{sc}(u)) du \right\}_{r=0}^{\infty}.$$

Moreover, by integration by parts, it suffices to verify

$$\left\{ \int_{-\infty}^{\infty} u^r dX_n(F_n(u)) \right\}_{r=0}^{\infty} \Longrightarrow \left\{ \int_{-2}^2 u^r dW^\circ(F_{sc}(u)) \right\}_{r=0}^{\infty}. \quad (3.3)$$

Now we draw attention to the fact that

$$\int_{-\infty}^{\infty} u^r dX_n(F_n(u)) = \sqrt{\frac{\beta n}{2}} \left(\mathbf{x}^* M_n^r \mathbf{x} - \frac{1}{n} \text{tr} M_n^r \right).$$

Thus we arrive at the stage to use Theorem 3.3. Observe that the assumptions in Theorem 3.3 only depend on the first four moments of the matrix elements. It is obvious that GOE and GUE satisfy the

moment assumptions in Theorem 3.3. Moreover, by the discussions above and (1.2) for the Gaussian case, (3.3) is valid for GOE and GUE obviously. Hence, (3.3) also holds for general Wigner matrices under the assumptions of Theorem 1.5. So we complete the proof. $\square$

IV. TIGHTNESS OF $(X_n)_{n \geq 1}$

In this section, we will prove Lemma 3.2. At first, we show that the process sequence $(X_n)_{n \geq 1}$ can only have $C[0, 1]$ -supported accumulation points. To this end, it suffices to check that the maximal jump of the process $X_n(t)$ converges to zero in probability. This can be seen directly from the isotropic delocalization property (2.6). Hence, the remaining part of this section will be devoted to showing the tightness of the process sequence $(X_n)_{n \geq 1}$, which is the main part of our proof. To this end, we begin with the modulus of continuity of the process X_n as

$$w_{X_n}(\delta) = w(X_n, \delta) := \sup_{|t_1 - t_2| < \delta} |X_n(t_1) - X_n(t_2)|, \quad 0 < \delta \leq 1.$$

By Theorem 8.2 of the Billingsley's book,⁴ to prove the tightness of $(X_n)_{n \geq 1}$, it suffices to show the following two statements.

(I): For each positive η , there exists an a such that

$$\mathbb{P}(|X_n(0)| > a) \leq \eta, \quad n \geq 1$$

and

(II): For each positive ε and η , there exists a δ , with $0 < \delta < 1$, and an integer n_0 such that

$$\mathbb{P}(w_{X_n}(\delta) \geq \varepsilon) \leq \eta, \quad n \geq n_0.$$

At first, we notice that (I) is obvious in our case since

$$\mathbb{P}(X_n(0) = 0) = 1, \quad n \geq 1.$$

Therefore, it remains to show (II) in the sequel. We will rely on the following lemma.

Lemma 4.1. Assume that M_n satisfies the assumptions imposed in Theorem 1.5. Let ϵ be some sufficiently small but fixed positive constant. If for any $t_1, t_2 \in [0, 1]$ satisfying $|t_2 - t_1| \geq n^{-1/2 - \epsilon}$ there exists

$$\mathbb{E} (X_n(t_2) - X_n(t_1))^4 \leq C|t_2 - t_1|^\alpha \quad (4.1)$$

with some positive constants C and $\alpha > 1$ which are both independent of t_1, t_2 , then (II) holds.

Proof. Note that by definition, we need to show that for any positive ϵ and η , there exists a $\delta \in (0, 1)$ and a sufficiently large n_0 such that for $n \geq n_0$,

$$\mathbb{P} \left(\sup_{|t_2 - t_1| \leq \delta} |X_n(t_2) - X_n(t_1)| \geq \varepsilon \right) \leq \eta.$$

By the discussions in Ref. 4 (see (8.12) therein), it suffices to show that for $n \geq n_0$ and $0 \leq t_1 \leq 1$,

$$\frac{1}{\delta} \mathbb{P} \left(\sup_{t_1 \leq t_2 \leq t_1 + \delta} |X_n(t_2) - X_n(t_1)| \geq \varepsilon/3 \right) \leq \eta.$$

To this end, we set

$$m = m(n) := \lfloor n^{1/2 + \epsilon/2} \rfloor.$$

Observe that

$$\begin{aligned} 0 &\leq \sup_{t_1 \leq t_2 \leq t_1 + \delta} |X_n(t_2) - X_n(t_1)| - \max_{0 \leq j \leq m} |X_n(t_1 + \frac{j}{m}\delta) - X_n(t_1)| \\ &\leq C\sqrt{n} \max_{1 \leq j \leq m} \sum_{i=\lfloor n(t_1 + \frac{j-1}{m}\delta) \rfloor}^{\lfloor n(t_1 + \frac{j}{m}\delta) \rfloor} (|y_i|^2 + \frac{1}{n}) \leq C\sqrt{n} \frac{n\delta}{m} \max_i (|y_i|^2 + \frac{1}{n}) \leq C\delta n^{-\epsilon/4} \end{aligned}$$

with overwhelming probability for sufficiently large n . Here in the last inequality above we employed the isotropic delocalization property (2.6) again. Consequently, it suffices to verify that for $m = \lfloor n^{1/2+\epsilon/2} \rfloor$,

$$\frac{1}{\delta} \mathbb{P} \left(\max_{0 \leq j \leq m} |X_n(t_1 + \frac{j}{m}\delta) - X_n(t_1)| \geq \epsilon/4 \right) \leq \eta/2. \quad (4.2)$$

By Theorem 12.2 and the proof of Theorem 12.3 of Ref. 4, in order to obtain (4.2), it suffices to show that for any $t_1, t_2 \in [0, 1]$ such that $|t_2 - t_1| \geq m^{-1}\delta$, one has

$$\mathbb{E} (X_n(t_2) - X_n(t_1))^4 \leq C|t_2 - t_1|^\alpha$$

with some positive constants C and $\alpha > 1$. When n is sufficiently large, by the definition of m , it suffices to have (4.1) when $t_2 - t_1 \geq n^{-1/2-\epsilon}$. Therefore, we complete the proof. $\square$

Below we will verify the condition (4.1) of Lemma 4.1 for $t_1, t_2 \in [0, 1]$ such that $t_2 - t_1 \geq n^{-1/2-\epsilon}$. At first, we construct a modified process as

$$Y_n(F_{sc}^{-1}(t)) = \sqrt{\frac{\beta n}{2}} \sum_{i=1}^{nF_n(F_{sc}^{-1}(t))} (|y_i|^2 - \frac{1}{n}).$$

Here we specify $F_{sc}^{-1}(0) = -2$ and $F_{sc}^{-1}(1) = 2$. Note that $Y_n(s), s \in [-2, 2]$ is just $X_n(F_n(s))$ restricted on $[-2, 2]$. In addition,

$$X_n(t) = X_n(F_{sc}(F_{sc}^{-1}(t))), \quad t \in [0, 1].$$

Then by using (3.2) and (2.6) again, one obtains that with overwhelming probability,

$$\begin{aligned} \sup_{t \in [0, 1]} |Y_n(F_{sc}^{-1}(t)) - X_n(t)| &\leq \sqrt{n}(\log n)^{C \log \log n} \max_i (|y_i|^2 + \frac{1}{n}) \\ &\leq \frac{(\log n)^{C' \log \log n}}{\sqrt{n}}. \end{aligned} \quad (4.3)$$

Moreover, we also have the definite bound

$$\sup_{t \in [0, 1]} |X_n(t)|, |Y_n(F_{sc}^{-1}(t))| \leq \sqrt{n} \sum_{i=1}^n (|y_i|^2 + \frac{1}{n}) = 2\sqrt{n}. \quad (4.4)$$

Then by combining (4.3) and (4.4), for $t_2 - t_1 \geq n^{-1/2-\epsilon}$ we have

$$\mathbb{E}((Y_n(F_{sc}^{-1}(t_2)) - Y_n(F_{sc}^{-1}(t_1))) - (X_n(t_2) - X_n(t_1)))^4 \leq C(t_2 - t_1)^2. \quad (4.5)$$

Therefore, it suffices to show that for any $t_1, t_2 \in [0, 1]$ satisfying $t_2 - t_1 \geq n^{-1/2-\epsilon}$, one has

$$\mathbb{E}(Y_n(F_{sc}^{-1}(t_2)) - Y_n(F_{sc}^{-1}(t_1)))^4 \leq C(t_2 - t_1)^\alpha \quad (4.6)$$

with some positive constants C and $\alpha > 1$. Now we set

$$s_1 = F_{sc}^{-1}(t_1), \quad s_2 = F_{sc}^{-1}(t_2).$$

Using the explicit formula of the semicircle law (2.1), it is elementary to see that when $t_2 - t_1 \geq n^{-1/2-\epsilon}$, there exists $s_2 - s_1 \geq Cn^{-1/2-\epsilon}$ for some positive constant C . Actually, one can get that

$$C(t_2 - t_1) \leq s_2 - s_1 \leq C'(t_2 - t_1)^{2/3} \quad (4.7)$$

holds uniformly in s_1, s_2 with some positive constants C and C' . Hence, what remains is to verify that when

$$s_2 - s_1 \geq Cn^{-1/2-\epsilon},$$

one has

$$\mathbb{E}(Y_n(s_2) - Y_n(s_1))^4 \leq (s_2 - s_1)^2 \leq C'(t_2 - t_1)^{4/3}. \quad (4.8)$$

Then (4.5) together with (4.8) implies (4.6) with $\alpha = 4/3$. Note that, by definition we have

$$Y_n(s) = \sqrt{\frac{\beta n}{2}} \sum_{i=1}^{nF_n(s)} (|y_i|^2 - \frac{1}{n}) = \sqrt{\frac{\beta n}{2}} \sum_{i=1}^n (|y_i|^2 - \frac{1}{n}) \mathbf{1}_{\{\lambda_i \leq s\}}, \quad s \in [-2, 2],$$

which implies

$$Y_n(s_2) - Y_n(s_1) = \sqrt{\frac{\beta n}{2}} \sum_{i=1}^n (|y_i|^2 - \frac{1}{n}) \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}}, \quad s_1, s_2 \in [-2, 2].$$

In the sequel, we will show (4.8) by a Green function comparison strategy. To this end, at first, we will approximate the indicator functions $\mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}}$, $i = 1, \dots, n$ by smooth functions expressed in terms of the Green function with the help of the following lemma.

Lemma 4.2. Under the assumptions of Theorem 1.5, for $\eta = n^{-1/2-\epsilon}(s_2 - s_1)^{1/2}$ with some sufficiently small but fixed positive constant ϵ , when $s_2 - s_1 \geq Cn^{-1/2-\epsilon}$ with some positive constant C , we have

$$\mathbb{E} \left(\sqrt{n} \sum_{i=1}^n (|y_i|^2 - \frac{1}{n}) \left(\mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} - \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + \sqrt{-1}\eta)} dE \right) \right)^4 \leq C'(s_2 - s_1)^2$$

with some positive constant C' .

Proof. By the isotropic delocalization property (2.6) and the definite bound $|y_i|^2 \leq 1$ we see that it suffices to show for some sufficiently small constant $\epsilon > 0$, there exists

$$n^{-2+\epsilon} \mathbb{E} \left(\sum_{i=1}^n \left| \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} - \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + \sqrt{-1}\eta)} dE \right| \right)^4 \leq C'(s_2 - s_1)^2. \quad (4.9)$$

Now we choose

$$\theta := n^{-1/2-\epsilon/2}(s_2 - s_1)^{1/2} \gg \eta \geq Cn^{-\frac{3}{4}-\frac{3\epsilon}{2}}. \quad (4.10)$$

Observe that both η and θ are much less than $s_2 - s_1$. Now we split the real line into $\mathbb{R} = \mathbf{L}_1 \cup \mathbf{L}_2$, where

$$\mathbf{L}_1 = (-\infty, s_1 - \theta) \cup (s_1 + \theta, s_2 - \theta) \cup (s_2 + \theta, \infty), \quad \mathbf{L}_2 = \mathbb{R} \setminus \mathbf{L}_1.$$

We will show that when $\lambda_i \in \mathbf{L}_1$, there is

$$\left| \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} - \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + \sqrt{-1}\eta)} dE \right| \leq C\eta \left(\frac{1}{|\lambda_i - s_1|} + \frac{1}{|\lambda_i - s_2|} \right). \quad (4.11)$$

To see (4.11), we use the following elementary fact:

$$\begin{aligned} \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + \sqrt{-1}\eta)} dE &= \frac{1}{\pi} \int_{s_1}^{s_2} \frac{\eta}{(\lambda_i - E)^2 + \eta^2} dE \\ &= \frac{1}{\pi} \left(\arctan \frac{s_2 - \lambda_i}{\eta} - \arctan \frac{s_1 - \lambda_i}{\eta} \right). \end{aligned}$$

Note that when $\lambda_i \in \mathbf{L}_1$, one has

$$|\lambda_i - s_2|, |\lambda_i - s_1| \geq \theta \gg \eta.$$

By the basic asymptotic properties of $\arctan(x)$, one has for $\lambda_i \in \mathbf{L}_1$,

$$\left| \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} - \frac{1}{\pi} \left(\arctan \frac{s_2 - \lambda_i}{\eta} - \arctan \frac{s_1 - \lambda_i}{\eta} \right) \right| \leq C\eta \left(\frac{1}{|\lambda_i - s_1|} + \frac{1}{|\lambda_i - s_2|} \right)$$

with some positive constant C . Let $N_n(I)$ be the number of the eigenvalues falling into the region $I \in \mathbb{R}$. Then we have

$$\begin{aligned} & \sum_{i=1}^n \left| \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} - \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + i\eta)} \right| \\ & \leq C \left(\eta \sum_{i: \lambda_i \in \mathbf{L}_1} \left(\frac{1}{|\lambda_i - s_1|} + \frac{1}{|\lambda_i - s_2|} \right) + N_n(\mathbf{L}_2) \right). \end{aligned} \quad (4.12)$$

Now we employ the so-called local semicircle law (for instance, see Theorem 1.8 of Ref. 16) in the sense that for any interval $I \in \mathbb{R}$ with its length $|I| \geq n^{-1+c}$ for any sufficiently small but fixed constant $c > 0$,

$$N_n(I) = O(n|I|) \quad (4.13)$$

with overwhelming probability. We decompose the real line as

$$\mathbb{R} = (-\infty, -5) \cup \left(\bigcup_{k=1}^{K_n} I_k \right) \cup (5, +\infty),$$

where $K_n = O(n^{1-c})$ and

$$I_k = [-5 + (k-1)n^{-1+c}, -5 + kn^{-1+c}].$$

Here we can choose K_n appropriately such that

$$-5 + (K_n - 1)n^{-1+c} < 5, \quad -5 + K_n n^{-1+c} \geq 5.$$

We will show that

$$\sum_{i: \lambda_i \in \mathbf{L}_1} \left(\frac{1}{|\lambda_i - s_1|} + \frac{1}{|\lambda_i - s_2|} \right) \leq Cn \log n \quad (4.14)$$

with overwhelming probability. At first, by the rigidity property in Ref. 10 we see that all the eigenvalues of M_n are in $[-5, 5]$ with overwhelming probability. Thus it suffices to show with overwhelming probability,

$$\begin{aligned} \sum_{i: \lambda_i \in \mathbf{L}_1 \cap [-5, 5]} \left(\frac{1}{|\lambda_i - s_1|} + \frac{1}{|\lambda_i - s_2|} \right) & \leq \sum_{k=1}^{K_n} \sum_{i: \lambda_i \in \mathbf{L}_1 \cap I_k} \left(\frac{1}{|\lambda_i - s_1|} + \frac{1}{|\lambda_i - s_2|} \right) \\ & \leq C \sum_{k=1}^{K_n} \sum_{i: \lambda_i \in \mathbf{L}_1 \cap I_k} \frac{1}{\theta + (k-1)n^{-1+c}} \\ & \leq C \sum_{k=1}^{K_n} \frac{n|I_k|}{\theta + (k-1)n^{-1+c}} \\ & \leq Cn \log n. \end{aligned}$$

Here in the third step we have used (4.13). Moreover, by (4.10) and (4.13) we also have

$$N_n(\mathbf{L}_2) \leq Cn\theta \quad (4.15)$$

with overwhelming probability for some positive constant C . Combining (4.12), (4.14), and (4.15) we have

$$\sum_{i=1}^n \left| \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} - \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + \sqrt{-1}\eta)} \right| \leq C(\eta n \log n + n\theta) \quad (4.16)$$

with overwhelming probability. Now by noticing that the left hand side of (4.16) is bounded by $2n$ definitely, we also have

$$\mathbb{E} \left(\sum_{i=1}^n \left| \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} - \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + \sqrt{-1}\eta)} \right| \right)^4 \leq C(\eta n \log n + n\theta)^4. \quad (4.17)$$

Then by (4.17) and the definitions of θ and η we can get (4.9). Thus we complete the proof. $\square$

Now we observe that

$$\begin{aligned} & \sum_{i=1}^n (|y_i|^2 - \frac{1}{n}) \frac{1}{\pi} \int_{s_1}^{s_2} \Im \frac{1}{\lambda_i - (E + \sqrt{-1}\eta)} dE \\ &= \frac{1}{\pi} \int_{s_1}^{s_2} \Im(\mathbf{x}^* G(z) \mathbf{x} - \frac{1}{n} \text{tr} G(z)) dE, \end{aligned} \quad (4.18)$$

where the r.h.s. is expressed in terms of the Green function. For simplicity, we will adopt the notation in Ref. 12 to write $A_{\mathbf{v}\mathbf{w}} = \mathbf{v}^* A \mathbf{w}$ for any matrix A . Particularly, $A_{\mathbf{v}\mathbf{e}}$ and $A_{\mathbf{e}\mathbf{v}}$ will be simply denoted by $A_{\mathbf{v}i}$ and $A_{i\mathbf{v}}$ in the sequel. Then with the aid of Lemma 4.2 and (4.18), it suffices to prove the following lemma.

Lemma 4.3. Let $z = E + \sqrt{-1}\eta$ with $\eta = n^{-1/2-\epsilon}(s_2 - s_1)^{1/2}$, where ϵ is some sufficiently small but fixed positive constant. And we assume that $s_1, s_2 \in [-2, 2]$ such that $s_2 - s_1 \geq n^{-1/2-\epsilon}$. Under the assumptions of Theorem 1.5, one has

$$\mathbb{E} \left(\sqrt{n} \int_{s_1}^{s_2} \Im(G_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} G(z)) dE \right)^4 \leq C(s_2 - s_1)^2 \quad (4.19)$$

for some positive constant C .

Obviously, Lemma 4.3 can be implied by the following two lemmas.

Lemma 4.4. When M_n is GOE or GUE and η, s_1, s_2 satisfy the assumptions in Lemma 4.3 (4.19) holds.

Lemma 4.5. Let $\tilde{M}_n$ be GOE (resp. GUE) and $\tilde{G}(z)$ be its Green function. Assume that M_n is a real (resp. complex) Wigner matrix satisfying the assumption in Theorem 1.5 and $G(z)$ is its Green function. Under the assumptions on η, s_1, s_2 in Lemma 4.3, one has

$$\begin{aligned} & \left| \mathbb{E} \left(\int_{s_1}^{s_2} \Im(G_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} G(z)) dE \right)^4 - \mathbb{E} \left(\int_{s_1}^{s_2} \Im(\tilde{G}_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} \tilde{G}(z)) dE \right)^4 \right| \\ & \leq Cn^{-2}(s_2 - s_1)^2 \end{aligned} \quad (4.20)$$

for some positive constant C .

Hence, it suffices to prove Lemmas 4.4 and 4.5 in the sequel.

Proof of Lemma 4.4. Below, we will focus on the real case for simplicity. The proof for the complex case is just analogous. We will show that (4.19) holds for GOE below. Note that by Lemma 4.2 we can go back to the original quantity to verify

$$\mathbb{E}(Y_n(s_2) - Y_n(s_1))^4 \leq C(s_2 - s_1)^2 \quad (4.21)$$

instead. According to (4.7), it suffices to check

$$\mathbb{E} \left(\sqrt{n} \sum_{i=1}^n (|y_i|^2 - \frac{1}{n}) \mathbf{1}_{\{s_1 < \lambda_i \leq s_2\}} \right)^4 \leq C(t_2 - t_1)^2.$$

Recall the fact that in the Gaussian case $\mathbf{y}$ is uniformly distributed on S^{n-1} . By using (3.2) and the isotropic delocalization property (2.6) again, we see that it suffices to verify

$$\mathbb{E} \left(\sqrt{n} \sum_{i=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor} (|y_i|^2 - \frac{1}{n}) \right)^4 \leq C(t_2 - t_1)^2. \quad (4.22)$$

Now we use the fact

$$\mathbf{y} \stackrel{d}{=} \frac{\mathbf{g}_{\mathbb{R}}}{\|\mathbf{g}_{\mathbb{R}}\|}.$$

Here $\mathbf{g}_{\mathbb{R}} = (g_1, \dots, g_n)^T$ is the $n \times 1$ random vector with i.i.d $N(0, 1)$ coefficients. It is elementary to check that

$$\mathbb{E}|y_i^2|^m = \mathbb{E} \frac{g_i^{2m}}{\|\mathbf{g}_{\mathbb{R}}\|^{2m}} = O(n^{-m}) \quad (4.23)$$

for any given integer $m \geq 0$. Moreover, we have the following lemma.

Lemma 4.6. Assume that $\mathbf{y} = (y_1, \dots, y_n)^T$ is uniformly distributed on S^{n-1} . For any i, j, k, l different from each other,

$$\mathbb{E}(y_i^2 - \frac{1}{n})(y_j^2 - \frac{1}{n})(y_k^2 - \frac{1}{n})^2 = O(n^{-5}), \quad (4.24)$$

$$\mathbb{E}(y_i^2 - \frac{1}{n})(y_j^2 - \frac{1}{n})(y_k^2 - \frac{1}{n})(y_l^2 - \frac{1}{n}) = O(n^{-6}). \quad (4.25)$$

Proof of Lemma 4.6. Note that we always have

$$\sum_{i=1}^n (y_i^2 - \frac{1}{n}) = 0.$$

Therefore, we can get

$$\begin{aligned} 0 &= \mathbb{E} \left(\sum_{i=1}^n (y_i^2 - \frac{1}{n}) \right) \left(\sum_{j=1}^n (y_j^2 - \frac{1}{n}) \right) \left(\sum_{k=1}^n (y_k^2 - \frac{1}{n})^2 \right) \\ &= \sum_{i,k} \mathbb{E} (y_i^2 - \frac{1}{n})^2 (y_k^2 - \frac{1}{n})^2 + 2 \sum_{i \neq k} \mathbb{E} (y_i^2 - \frac{1}{n}) (y_k^2 - \frac{1}{n})^3 \\ &\quad + \sum_{i,j,k \text{ are mutually distinct}} \mathbb{E} (y_i^2 - \frac{1}{n}) (y_j^2 - \frac{1}{n}) (y_k^2 - \frac{1}{n})^2 \\ &= \sum_{i,j,k \text{ are mutually distinct}} \mathbb{E} (y_i^2 - \frac{1}{n}) (y_j^2 - \frac{1}{n}) (y_k^2 - \frac{1}{n})^2 + O(n^{-2}), \end{aligned} \quad (4.26)$$

where the last step follows from (4.23) directly. Now observe that $\mathbf{y}$ is an exchangeable random vector. Thus by symmetry, we see that every term in the summation in (4.26) is the same, then (4.24)

follows. (4.25) can be verified analogously below. Note

$$\begin{aligned}
 0 &= \mathbb{E} \left(\sum_{i=1}^n \left(y_i^2 - \frac{1}{n} \right) \right) \left(\sum_{j=1}^n \left(y_j^2 - \frac{1}{n} \right) \right) \left(\sum_{k=1}^n \left(y_k^2 - \frac{1}{n} \right) \right) \left(\sum_{l=1}^n \left(y_l^2 - \frac{1}{n} \right) \right) \\
 &= \sum_i \mathbb{E} \left(y_i^2 - \frac{1}{n} \right)^4 + 4 \sum_{i \neq j} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right)^3 \left(y_j^2 - \frac{1}{n} \right) \\
 &\quad + 3 \sum_{i \neq j} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right)^2 \left(y_j^2 - \frac{1}{n} \right)^2 \\
 &\quad + 6 \sum_{\substack{i, j, k \text{ are mutually distinct}}} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right) \left(y_j^2 - \frac{1}{n} \right) \left(y_k^2 - \frac{1}{n} \right)^2 \\
 &\quad + \sum_{\substack{i, j, k, l \text{ are mutually distinct}}} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right) \left(y_j^2 - \frac{1}{n} \right) \left(y_k^2 - \frac{1}{n} \right) \left(y_l^2 - \frac{1}{n} \right) \\
 &= \sum_{\substack{i, j, k, l \text{ are mutually distinct}}} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right) \left(y_j^2 - \frac{1}{n} \right) \left(y_k^2 - \frac{1}{n} \right) \left(y_l^2 - \frac{1}{n} \right) + O(n^{-2}),
 \end{aligned}$$

where the last step follows from (4.23) and (4.24). Now again by symmetry, we can get (4.25). So we complete the proof. $\square$

Using (4.23), (4.24), and (4.25), we can finally get

$$\begin{aligned}
 &\mathbb{E} \left(\sum_{i=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor} \left(y_i^2 - \frac{1}{n} \right) \right)^4 \\
 &= \sum_{i=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right)^4 + 4 \sum_{i \neq j=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right)^3 \left(y_j^2 - \frac{1}{n} \right) \\
 &\quad + 3 \sum_{i \neq j=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right)^2 \left(y_j^2 - \frac{1}{n} \right)^2 \\
 &\quad + 6 \sum_{\substack{i, j, k=\lfloor nt_1 \rfloor \\ i, j, k \text{ are mutually distinct}}}^{\lfloor nt_2 \rfloor} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right) \left(y_j^2 - \frac{1}{n} \right) \left(y_k^2 - \frac{1}{n} \right)^2 \\
 &\quad + \sum_{\substack{i, j, k, l=\lfloor nt_1 \rfloor \\ i, j, k, l \text{ are mutually distinct}}}^{\lfloor nt_2 \rfloor} \mathbb{E} \left(y_i^2 - \frac{1}{n} \right) \left(y_j^2 - \frac{1}{n} \right) \left(y_k^2 - \frac{1}{n} \right) \left(y_l^2 - \frac{1}{n} \right) \\
 &\leq Cn^{-2}(t_2 - t_1)^2,
 \end{aligned}$$

which implies (4.22) for GOE. Moreover, by the discussions above (4.19) for the Gaussian case follows immediately. Hence, we complete the proof of Lemma 4.4. $\square$

Therefore, it remains to compare the general case with the Gaussian case. That is to say, we only need to show (4.20).

Proof of Lemma 4.5. To simplify the discussions in the sequel, we truncate the matrix elements of $\sqrt{n}M_n = (v_{ij})_{n,n}$ and $\sqrt{n}\tilde{M}_n = (\tilde{v}_{ij})_{n,n}$ at n^ϵ , where $\epsilon > 0$ is the small constant chosen in Lemma

4.3. We denote the truncated variables as

$$v_{ij}^{(t)} := v_{ij} \mathbf{1}_{\{|v_{ij}| \leq n^\epsilon\}}, \quad \tilde{v}_{ij}^{(t)} := \tilde{v}_{ij} \mathbf{1}_{\{|\tilde{v}_{ij}| \leq n^\epsilon\}}$$

and set

$$M_n^{(t)} := \frac{1}{\sqrt{n}}(v_{ij}^{(t)})_{n,n}, \quad \tilde{M}_n^{(t)} := \frac{1}{\sqrt{n}}(\tilde{v}_{ij}^{(t)})_{n,n}.$$

Now we define the event

$$\mathbf{E}(\epsilon) := \{\max_{i,j} |v_{ij}| \leq n^\epsilon\}$$

and let $\mathbf{E}^c(\epsilon)$ be its complement. By Condition 1.4, it is elementary to see that $\mathbf{E}(\epsilon)$ holds with overwhelming probability. Moreover, for any well-defined functional $\mathcal{H} : \text{Mat}_n(\mathbb{C}) \rightarrow \mathbb{C}$, obviously we have

$$\mathcal{H}(M_n) \mathbf{1}_{\mathbf{E}(\epsilon)} = \mathcal{H}(M_n^{(t)}) \mathbf{1}_{\mathbf{E}(\epsilon)},$$

since M_n equals $M_n^{(t)}$ when $\mathbf{E}(\epsilon)$ occurs. Then we obtain that

$$\begin{aligned} \mathbb{E} \mathcal{H}(M_n) &= \mathbb{E} \mathcal{H}(M_n^{(t)}) \mathbf{1}_{\mathbf{E}(\epsilon)} + \mathbb{E} \mathcal{H}(M_n) \mathbf{1}_{\mathbf{E}^c(\epsilon)} \\ &= \mathbb{E} \mathcal{H}(M_n^{(t)}) + (\mathbb{E} \mathcal{H}(M_n) \mathbf{1}_{\mathbf{E}^c(\epsilon)} - \mathbb{E} \mathcal{H}(M_n^{(t)}) \mathbf{1}_{\mathbf{E}^c(\epsilon)}), \end{aligned}$$

which directly implies that

$$|\mathbb{E} \mathcal{H}(M_n) - \mathbb{E} \mathcal{H}(M_n^{(t)})| \leq |\mathbb{E} \mathcal{H}(M_n) \mathbf{1}_{\mathbf{E}^c(\epsilon)}| + |\mathbb{E} \mathcal{H}(M_n^{(t)}) \mathbf{1}_{\mathbf{E}^c(\epsilon)}|. \quad (4.27)$$

Now we choose $\mathcal{H} := \mathcal{H}_{\mathbf{x}, \eta, s_1, s_2}$ as

$$\mathcal{H}(M_n) = \left(\int_{s_1}^{s_2} \Im(G_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} G(z)) dE \right)^4.$$

Observe that we always have the elementary bound

$$\|(A - z)^{-1}\|_{op} \leq \eta^{-1}$$

for any Hermitian matrix $A \in \text{Mat}_n(\mathbb{C})$. Hence, it is easy to check that

$$|\mathcal{H}(M_n)|, |\mathcal{H}(M_n^{(t)})| \leq 16\eta^{-4}(s_1 - s_2)^4$$

holds definitely. In addition, by Condition 1.4, we see that

$$\mathbb{E} \mathbf{1}_{\mathbf{E}^c(\epsilon)} = \mathbb{P}(\mathbf{E}^c(\epsilon)) \leq n^{-K}$$

for any positive constant K when n is sufficiently large. Consequently, by (4.27) one has

$$|\mathbb{E} \mathcal{H}(M_n) - \mathbb{E} \mathcal{H}(M_n^{(t)})| \leq n^{-K}$$

for any positive constant K when n is sufficiently large. Apparently, analogous conclusion holds for $\tilde{M}_n$ and $\tilde{M}_n^{(t)}$. That means we can replace M_n and $\tilde{M}_n$ by their truncated counterparts $M_n^{(t)}$ and $\tilde{M}_n^{(t)}$ in (4.20), and then prove (4.20) with the truncated matrices instead.

Note that the truncation may change the first four moments of the original elements by tiny amounts. Actually, such minor changes are smaller than n^{-K} for any positive constant K when n is sufficiently large. More specifically, by Condition 1.4, we can easily get that for all $1 \leq i, j \leq n$,

$$\mathbb{E}(\Re(v_{ij}^{(t)})^l \Im(v_{ij}^{(t)})^m) = \mathbb{E}(\Re(\tilde{v}_{ij}^{(t)})^l \Im(\tilde{v}_{ij}^{(t)})^m) + O(n^{-K}), \quad 0 \leq l, m \leq l+m \leq 4 \quad (4.28)$$

for any large constant $K > 0$ when n is sufficiently large. It will be clear that such a slight modification on the moments matching condition does not affect our comparison procedure. Moreover, note that all the results needed from the references such as Refs. 10 and 12 hold with overwhelming probability for the original matrices, while the truncated matrices are equal to the original ones with overwhelming probability, thus the results from these references are still valid for the truncated

matrices. For simplicity, we will recycle the notation $M_n, \tilde{M}_n, v_{ij}, \tilde{v}_{ij}$ to denote $M_n^{(t)}, \tilde{M}_n^{(t)}, v_{ij}^{(t)}, \tilde{v}_{ij}^{(t)}$ in the sequel. Therefore, below we will make the additional assumption

$$\max_{i,j} |v_{ij}| \leq n^\epsilon, \quad \max_{i,j} |\tilde{v}_{ij}| \leq n^\epsilon.$$

The main idea to show (4.20) is a Green function comparison strategy based on the discussions in Ref. 12. To pursue this approach, we need to introduce some notation. At first we assign a bijective ordering map ϕ on the index set of the matrix elements,

$$\phi : \{(i, j) : 1 \leq i \leq j \leq n\} \rightarrow \{1, \dots, \frac{n(n+1)}{2}\}.$$

For $1 \leq \gamma \leq n(n+1)/2$, we define the matrix M_n^γ to be the Wigner matrix with its (i, j) -th element being $v_{ij}/\sqrt{n}$ if $\phi(i, j) \leq \gamma$ or $\tilde{v}_{ij}/\sqrt{n}$ otherwise. Correspondingly, we denote the Green function of M_n^γ by $G^\gamma(z)$. Our aim is to estimate the one step difference

$$\left| \mathbb{E} \left(\int_{s_1}^{s_2} \Im(G_{\mathbf{xx}}^\gamma(z) - \frac{1}{n} \text{tr} G^\gamma(z)) dE \right)^4 - \mathbb{E} \left(\int_{s_1}^{s_2} \Im(G_{\mathbf{xx}}^{\gamma-1}(z) - \frac{1}{n} \text{tr} G^{\gamma-1}(z)) dE \right)^4 \right|,$$

and at the end we will use a telescoping argument to sum up all these one step differences to obtain (4.20).

Without loss of generality, we assume that $\gamma = \phi(a, b)$. Apparently, M_n^γ and $M_n^{\gamma-1}$ only differ in (a, b) and (b, a) -th elements. Let

$$E^{ab} = \mathbf{e}_a \mathbf{e}_b^*.$$

Then we can write

$$M_n^\gamma = Q + n^{-1/2} V, \quad M_n^{\gamma-1} = Q + n^{-1/2} \tilde{V},$$

where

$$V = (1 - \delta_{ab}/2)(v_{ab} E^{ab} + v_{ba} E^{ba}),$$

$$\tilde{V} = (1 - \delta_{ab}/2)(\tilde{v}_{ab} E^{ab} + \tilde{v}_{ba} E^{ba}),$$

and Q is a random matrix independent of v_{ab} and $\tilde{v}_{ab}$. In addition, we set

$$R(z) := (Q - z)^{-1}.$$

For simplicity, we further rewrite $G^{\gamma-1}(z)$ and $G^\gamma(z)$ as $S(z)$ and $T(z)$, respectively. And when there is no confusion, we will omit the variable z from the above notation. Moreover, we will mainly focus on the real case for simplicity in the sequel. The discussion for the complex case is just analogous. Using the resolvent expansion formula, one can get

$$S = R + \sum_{k=1}^4 n^{-k/2} (-RV)^k R + n^{-5/2} (-RV)^5 S. \quad (4.29)$$

With the aid of (4.29) we can then write

$$\begin{aligned}
 & \mathbb{E} \left(\int_{s_1}^{s_2} \Im(S_{\mathbf{x}\mathbf{x}}(z) - n^{-1} \text{tr} S(z)) dE \right)^4 \\
 &= \mathbb{E} \left(\int_{s_1}^{s_2} \Im \left((R_{\mathbf{x}\mathbf{x}}(z) - n^{-1} \text{tr} R(z)) \right. \right. \\
 &\quad \left. \left. + \sum_{k=1}^4 n^{-k/2} ([(-R(z)V)^k R(z)]_{\mathbf{x}\mathbf{x}} - n^{-1} \text{tr} (-R(z)V)^k R(z)) \right. \right. \\
 &\quad \left. \left. + n^{-5/2} ([(-R(z)V)^5 S(z)]_{\mathbf{x}\mathbf{x}} - n^{-1} \text{tr} (-R(z)V)^5 S(z)) \right) dE \right)^4 \\
 &=: \mathbb{E} (\Im(\mathcal{F}_0 + \sum_{k=1}^4 n^{-k/2} \mathcal{F}_k + n^{-5/2} \mathcal{F}_5))^4.
 \end{aligned}$$

Here $\mathcal{F}_i := \mathcal{F}_i(a, b)$, $i = 0 \dots, 5$ whose definitions are given by

$$\begin{aligned}
 \mathcal{F}_0 &:= \int_{s_1}^{s_2} (R_{\mathbf{x}\mathbf{x}}(z) - n^{-1} \text{tr} R(z)) dE, \\
 \mathcal{F}_k &:= \int_{s_1}^{s_2} ([(-R(z)V)^k R(z)]_{\mathbf{x}\mathbf{x}} - n^{-1} \text{tr} (-R(z)V)^k R(z)) dE, \quad k = 1, \dots, 4, \\
 \mathcal{F}_5 &:= \int_{s_1}^{s_2} ([(-R(z)V)^5 S(z)]_{\mathbf{x}\mathbf{x}} - n^{-1} \text{tr} (-R(z)V)^5 S(z)) dE.
 \end{aligned}$$

Observe that in the real case, every $[(RV)^k R]_{\mathbf{x}\mathbf{x}}$ can be written as a summation of the terms of the form

$$(v_{ab})^k \mathbf{q}_{k,a,b}(R, \mathbf{x}),$$

where $\mathbf{q}_{k,a,b}(R, \mathbf{x})$ is some product of the factors $R_{\mathbf{x}\mathbf{a}}, R_{\mathbf{a}\mathbf{x}}, R_{\mathbf{x}\mathbf{b}}, R_{\mathbf{b}\mathbf{x}}, R_{\mathbf{a}\mathbf{a}}, R_{\mathbf{a}\mathbf{b}}, R_{\mathbf{b}\mathbf{a}},$ and $R_{\mathbf{b}\mathbf{b}}$. We remind here in the complex case, $(v_{ab})^k$ should be replaced by $(v_{ab})^{k_1} (v_{ba})^{k_2}$ with $k_1 + k_2 = k$. Moreover, the total number of the factors $R_{\mathbf{x}\mathbf{a}}, R_{\mathbf{a}\mathbf{x}}, R_{\mathbf{x}\mathbf{b}}, R_{\mathbf{b}\mathbf{x}}$ in every $\mathbf{q}_{k,a,b}(R, \mathbf{x})$ is 2. For example,

$$[(RV)^2 R]_{\mathbf{x}\mathbf{x}} = (v_{ab})^2 (R_{\mathbf{x}\mathbf{a}} R_{\mathbf{b}\mathbf{b}} R_{\mathbf{a}\mathbf{x}} + R_{\mathbf{x}\mathbf{b}} R_{\mathbf{a}\mathbf{b}} R_{\mathbf{a}\mathbf{x}} + R_{\mathbf{x}\mathbf{a}} R_{\mathbf{b}\mathbf{a}} R_{\mathbf{b}\mathbf{x}} + R_{\mathbf{x}\mathbf{b}} R_{\mathbf{a}\mathbf{a}} R_{\mathbf{b}\mathbf{x}}).$$

In the following Lemma 4.7, we will give some crude bounds for the quantities $\mathcal{F}_k$. These bounds will be used to provide a crude bound on

$$\mathbb{E} \left(\int_{s_1}^{s_2} \Im(G_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} G(z)) dE \right)^4 \quad (4.30)$$

via a comparison procedure. Then the crude bound on (4.30) will imply an improved bound on $\mathcal{F}_0$ in turn. Such an improved bound on $\mathcal{F}_0$ combined with another round of comparison can help us to obtain an improved bound for (4.30). In other words, our main route in the sequel is to use a “bootstrap” strategy to get a crude bound of (4.30) at first and then use the crude bound to get the improved bound in Lemma 4.3. To simplify the presentation, we will use the notation in Ref. 12 to set the parameter

$$\Psi(z) := \sqrt{\frac{\Im m_{sc}(z)}{n\eta}} + \frac{1}{n\eta}.$$

Lemma 4.7. Under the assumptions in Lemma 4.3, one has with overwhelming probability,

$$|\mathcal{F}_0(z)| \leq n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \Psi(z)(s_2 - s_1), \quad (4.31)$$

and for $1 \leq k \leq 5$,

$$|\mathcal{F}_k(z)| \leq n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \left(\left(\frac{\Im S_{\mathbf{x}\mathbf{x}}(z)}{n\eta} + \Psi^2(z) \right) + (|x_a|^2 + |x_b|^2) \right) (s_2 - s_1). \quad (4.32)$$

Proof of Lemma 4.7. First of all, by truncation, we see that the elements are bounded by n^ϵ . Moreover, by assumption one has

$$2n^{-1/2-\epsilon} \geq \eta \gg n^{-3/4-2\epsilon} \quad (4.33)$$

for some sufficiently small $\epsilon > 0$. Thus we always have $z = E + \sqrt{-1}\eta \in \mathbf{S}$ which is defined in (2.3). Now we come to verify (4.31). By definition,

$$\mathcal{F}_0 = \int_{s_1}^{s_2} (R_{\mathbf{x}\mathbf{x}}(z) - n^{-1} \text{tr} R(z)) dE.$$

Observe that

$$n^\epsilon \gg (\log n)^{\log \log n}$$

for sufficiently large n . Using the isotropic local semicircle law (2.5), one has with overwhelming probability,

$$|S_{\mathbf{x}\mathbf{x}}(z) - m_{sc}(z)| \leq n^{O(\epsilon)} \Psi(z), \quad |n^{-1} \text{tr} S(z) - m_{sc}(z)| \leq n^{O(\epsilon)} \Psi(z). \quad (4.34)$$

In addition, by (3.28) and (3.29) of Ref. 12 and the discussions above them we know

$$|S_{\mathbf{x}a}(z)|, |R_{\mathbf{x}a}(z)| \leq n^{O(\epsilon)} \left(\sqrt{\frac{\Im S_{\mathbf{x}\mathbf{x}}(z)}{n\eta}} + \Psi(z) + |x_a| \right) \quad (4.35)$$

and

$$|S_{ij}(z)|, |R_{ij}(z)| = O(1) \quad (4.36)$$

with overwhelming probability. Moreover, analogously, the bound in (4.35) also holds for $S_{a\mathbf{x}}(z)$, $R_{a\mathbf{x}}(z)$ with overwhelming probability. Then by (4.29), (4.34)–(4.36), one can get that

$$|R_{\mathbf{x}\mathbf{x}}(z) - m_{sc}(z)| \leq n^{O(\epsilon)} \Psi(z), \quad |n^{-1} \text{tr} R(z) - m_{sc}(z)| \leq n^{O(\epsilon)} \Psi(z).$$

Hence, (4.31) follows immediately.

Now we start to verify (4.32). Note that for $1 \leq k \leq 5$, the total number of the factors $R_{\mathbf{x}a}$, $R_{\mathbf{x}b}$, $R_{a\mathbf{x}}$, $R_{b\mathbf{x}}$, $S_{\mathbf{x}a}$, $S_{\mathbf{x}b}$, $S_{a\mathbf{x}}$, and $S_{b\mathbf{x}}$ in each $[(-RV)^k R]_{\mathbf{x}\mathbf{x}}$ or $[(-RV)^k S]_{\mathbf{x}\mathbf{x}}$ is 2. Let p be the total number of the factors $R_{\mathbf{x}a}$, $R_{a\mathbf{x}}$, $S_{\mathbf{x}a}$, and $S_{a\mathbf{x}}$, and q be the total number of the factors $R_{\mathbf{x}b}$, $R_{b\mathbf{x}}$, $S_{\mathbf{x}b}$, and $S_{b\mathbf{x}}$.

Thus, one always has $p + q = 2$. Then by (4.35) and (4.36), one obtains

$$\begin{aligned}
 |\mathcal{F}_k| &\leq n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \sum_{p+q=2} \left(\sqrt{\frac{\Im S_{\mathbf{x}\mathbf{x}}(z)}{n\eta}} + \Psi(z) + |x_a| \right)^p \\
 &\quad \times \left(\sqrt{\frac{\Im S_{\mathbf{x}\mathbf{x}}(z)}{n\eta}} + \Psi(z) + |x_b| \right)^q (s_2 - s_1) \\
 &\leq n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \left(\sqrt{\frac{\Im S_{\mathbf{x}\mathbf{x}}(z)}{n\eta}} + \Psi(z) + |x_a| \right)^2 (s_2 - s_1) \\
 &\quad + n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \left(\sqrt{\frac{\Im S_{\mathbf{x}\mathbf{x}}(z)}{n\eta}} + \Psi(z) + |x_b| \right)^2 (s_2 - s_1) \\
 &\leq n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \left(\left(\frac{\Im S_{\mathbf{x}\mathbf{x}}(z)}{n\eta} + \Psi^2(z) \right) + (|x_a|^2 + |x_b|^2) \right) (s_2 - s_1)
 \end{aligned}$$

with overwhelming probability. Therefore, we conclude the proof of Lemma 4.7. $\square$

Following from Lemma 4.7 and the definite bound

$$|\mathcal{F}_0| \leq \eta^{-1}(s_2 - s_1),$$

one can see that

$$\mathbb{E}(\Im \mathcal{F}_0)^4 \leq n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \left(\left(\sqrt{\frac{\Im m_{sc}(z)}{n\eta}} + \frac{1}{n\eta} \right) \right)^4 (s_2 - s_1)^4 \leq n^{-1+O(\epsilon)}(s_2 - s_1)^2.$$

Here we used the definition of η in Lemma 4.3. However, relying on Lemma 4.7, we can provide a better bound on the 4th moment of $\Im \mathcal{F}_0$ by a “bootstrap” strategy. Precisely, we will show the following bound:

$$\mathbb{E}(\Im \mathcal{F}_0)^4 \leq n^{-3/2+O(\epsilon)}(s_2 - s_1)^2. \quad (4.37)$$

To verify (4.37), we need the following crude bound on (4.30) for any Wigner matrix M_n satisfying the assumptions of Theorem 1.5.

Lemma 4.8. Under the assumptions of Lemma 4.3, we have

$$\mathbb{E} \left(\int_{s_1}^{s_2} (\Im(G_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} G(z))) dE \right)^4 \leq n^{-3/2+O(\epsilon)}(s_2 - s_1)^2.$$

Before commencing the proof of Lemma (4.8), we explain here how it implies (4.37). Note that by definition,

$$\Im \mathcal{F}_0 = \int_{s_1}^{s_2} \Im(R_{\mathbf{x}\mathbf{x}}(z) - n^{-1} \text{tr} R(z)) dE.$$

By (4.29), it suffices to show

$$\mathbb{E} \left(\int_{s_1}^{s_2} \Im(S_{\mathbf{x}\mathbf{x}}(z) - n^{-1} \text{tr} S(z)) dE \right)^4 \leq n^{-3/2+O(\epsilon)}(s_2 - s_1)^2. \quad (4.38)$$

Note M_n^γ is also a Wigner matrix satisfying the assumptions in Theorem 1.5. Thus (4.38) follows immediately. So does (4.37). Now we prove Lemma 4.8 below.

Proof of Lemma 4.8. We will use the aforementioned strategy to estimate the one step difference

$$\left| \mathbb{E} \left(\int_{s_1}^{s_2} (\Im(G_{\mathbf{xx}}^\gamma(z) - \frac{1}{n} \text{tr} G^\gamma(z))) dE \right)^4 - \mathbb{E} \left(\int_{s_1}^{s_2} (\Im(G_{\mathbf{xx}}^{\gamma-1}(z) - \frac{1}{n} \text{tr} G^{\gamma-1}(z))) dE \right)^4 \right|$$

by inserting the crude bounds on $\mathcal{F}_0$ and $\mathcal{F}_k$ provided in Lemma 4.7 at first, and then using the telescoping argument to obtain the desired bound in Lemma 4.8. By the discussions above, we have

$$\begin{aligned} & \mathbb{E} \left(\Im \int_{s_1}^{s_2} (S_{\mathbf{xx}}(z) - n^{-1} \text{tr} S(z)) dE \right)^4 \\ &= \mathbb{E} (\Im(\mathcal{F}_0 + \sum_{k=1}^4 n^{-k/2} \mathcal{F}_k + n^{-5/2} \mathcal{F}_5))^4 \\ &= \mathcal{A}_{ab} + \mathbb{E} \sum_{k_1, \dots, k_4=0}^5 n^{-\sum_{i=1}^4 k_i/2} \mathbf{1}_{\{\sum_{i=1}^4 k_i \geq 5\}} \prod_{i=1}^4 \Im \mathcal{F}_{k_i}, \end{aligned} \quad (4.39)$$

where $\mathcal{A}_{ab}$ is a quantity only depending on the first four moments of v_{ab} and independent of v_{ab} itself. Analogously, one has

$$\begin{aligned} & \mathbb{E} \left(\Im \int_{s_1}^{s_2} (T_{\mathbf{xx}}(z) - n^{-1} \text{tr} T(z)) dE \right)^4 \\ &= \mathcal{A}_{ab} + \mathbb{E} \sum_{k_1, \dots, k_4=0}^5 n^{-\sum_{i=1}^4 k_i/2} \mathbf{1}_{\{\sum_{i=1}^4 k_i \geq 5\}} \prod_{i=1}^4 \Im \mathcal{F}_{k_i}(\gamma-1 \rightarrow \gamma) + O(n^{-K}) \end{aligned} \quad (4.40)$$

for some sufficiently large constant $K > 10$ (say), where $\mathcal{F}_k(\gamma-1 \rightarrow \gamma)$ stands for the quantity defined through replacing V and S by $\tilde{V}$ and T , respectively, in the definition of $\mathcal{F}_k$. We remind here the negligible term $O(n^{-K})$ in (4.40) is produced by the slight modification on the moment matching condition, see (4.28). Therefore, to get the one step difference one only needs to estimate the second terms of (4.39) and (4.40). We will only handle (4.39) below. The case of (4.40) is just analogous.

Note that it suffices to estimate the contribution of

$$\mathbb{E} n^{-\kappa/2} \sum_{k_1, \dots, k_4=0}^5 \mathbf{1}_{\{\sum_{i=1}^4 k_i = \kappa\}} \prod_{i=1}^4 \Im \mathcal{F}_{k_i}. \quad (4.41)$$

for $5 \leq \kappa \leq 20$. Now we denote the right hand sides of (4.31) and (4.32) by D_1 and D_2 , respectively. Observe that by the assumption on η , we have

$$D_1 \leq n^{-1/4+O(\epsilon)}(s_2 - s_1)^{3/4}$$

and

$$D_2 \leq n^{-1/2+O(\epsilon)}(s_2 - s_1)^{1/2} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)(s_2 - s_1) \quad (4.42)$$

with overwhelming probability. Relying on the definite bound

$$\|R(z)\|_{op}, \|S(z)\|_{op} \leq \eta^{-1}, \quad (4.43)$$

we can use the upper bound in (4.42) as a definite one when we take expectations towards the polynomials in D_2 . Now let m be the number of 0 in the collection $\{k_1, k_2, k_3, k_4\}$. Then we have

$$\begin{aligned}
 & |n^{-\kappa/2} \mathbb{E} \sum_{k_1, \dots, k_4=0}^5 \mathbf{1}_{\{\sum_{i=1}^4 k_i = \kappa\}} \prod_{i=1}^4 \Im \mathcal{F}_{k_i}| \\
 & \leq C n^{-\kappa/2} \sum_{m=0}^3 D_1^m \mathbb{E} D_2^{4-m} \\
 & \leq C n^{-\kappa/2} (s_2 - s_1)^2 \sum_{m=0}^3 (n^{-1/4+O(\epsilon)})^m (n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2))^{4-m} \\
 & \leq n^{-\kappa/2} (s_2 - s_1)^2 \sum_{m=0}^3 (n^{-m/4+O(\epsilon)}) (n^{-(4-m)/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)).
 \end{aligned}$$

Here we have used the fact that $|x_a| \leq 1$. Hence, it is not difficult to see

$$\begin{aligned}
 & \left| \sum_{a,b} n^{-\kappa/2} \mathbb{E} \sum_{k_1, \dots, k_4=0}^5 \mathbf{1}_{\{\sum_{i=1}^4 k_i = \kappa\}} \prod_{i=1}^4 \Im \mathcal{F}_{k_i} \right| \\
 & \leq \sum_{a,b} n^{-\kappa/2} (n^{-5/4+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)) (s_2 - s_1)^2 \\
 & \leq n^{-\kappa/2+1+O(\epsilon)} (s_2 - s_1)^2.
 \end{aligned} \tag{4.44}$$

Here we have used the fact that $\sum_a |x_a|^2 = 1$. Then by (4.39), (4.44), and the fact $\kappa \geq 5$, one has

$$\begin{aligned}
 & \left| \mathbb{E} \left(\int_{s_1}^{s_2} (\Im(G_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} G(z))) dE \right)^4 - \mathbb{E} \left(\int_{s_1}^{s_2} (\Im(\tilde{G}_{\mathbf{x}\mathbf{x}}(z) - \frac{1}{n} \text{tr} \tilde{G}(z))) dE \right)^4 \right| \\
 & \leq n^{-3/2+O(\epsilon)} (s_2 - s_1)^2.
 \end{aligned}$$

Using Lemma 4.3 for the Gaussian case which has been proved above, we can conclude the proof. $\square$

Now we proceed to the proof of Lemma 4.5. Again we will resort to the telescoping argument. But this time we will use (4.37) instead of the crude bound (4.31). We go back to (4.39) and provide a better bound for (4.41) below. Note that by (4.44), it suffices to consider the cases of $\kappa = 5$ and $\kappa = 6$.

At first, we deal with the case of $\kappa = 5$. To this end, we split the set $\{\{k_1, \dots, k_4\} : \sum_{i=1}^4 k_i = 5\}$ into the following four cases.

- (i): $\{k_1, k_2, k_3, k_4\} = \{0, 0, 0, 5\}$,
- (ii): $\{k_1, k_2, k_3, k_4\} = \{0, 0, 1, 4\}, \{0, 0, 2, 3\}$,
- (iii): $\{k_1, k_2, k_3, k_4\} = \{0, 1, 1, 3\}, \{0, 1, 2, 2\}$,
- (iv): $\{k_1, k_2, k_3, k_4\} = \{1, 1, 1, 2\}$.

Note that for case (i), by using (4.43), (4.37), and Lemma 4.7 we see that

$$\begin{aligned}
 & n^{-5/2} \sum_{a,b} |\mathbb{E} \mathfrak{F}_{k_1} \mathfrak{F}_{k_2} \mathfrak{F}_{k_3} \mathfrak{F}_{k_4}| \\
 & \leq n^{-5/2} \sum_{a,b} \prod_{i=1}^4 (\mathbb{E} (\mathfrak{F}_{k_i})^4)^{1/4} \\
 & \leq n^{-5/2} \sum_{a,b} n^{-9/8+O(\epsilon)} (n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)) (s_2 - s_1)^2 \\
 & = n^{-17/8+O(\epsilon)} (s_2 - s_1)^2.
 \end{aligned}$$

Now we handle the case (ii) as follows:

$$\begin{aligned}
 & n^{-5/2} \sum_{a,b} |\mathbb{E} \mathfrak{F}_{k_1} \mathfrak{F}_{k_2} \mathfrak{F}_{k_3} \mathfrak{F}_{k_4}| \\
 & \leq n^{-5/2} \sum_{a,b} n^{-3/4+O(\epsilon)} (n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2))^2 (s_2 - s_1)^2 \\
 & \leq n^{-5/2} \sum_{a,b} n^{-3/4+O(\epsilon)} (n^{-1+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)) (s_2 - s_1)^2 \\
 & = n^{-9/4+O(\epsilon)} (s_2 - s_1)^2.
 \end{aligned}$$

Therefore, in the two cases above, the desired bound can be obtained by our estimates in (4.32) and (4.37). Now we start to deal with the cases (iii) and (iv) whose estimates require more accurate bounding procedure on the products of $\mathfrak{F}_k$. The following discussion will rely on the observation that with overwhelming probability,

$$\begin{aligned}
 |\mathcal{F}_1| & \leq n^{O(\epsilon)} \sup_{E \in [s_1, s_2]} \left(\sqrt{\frac{\mathfrak{I} S_{\mathbf{x}\mathbf{x}}(z)}{n\eta}} + \Psi(z) + |x_a| \right) \left(\sqrt{\frac{\mathfrak{I} S_{\mathbf{x}\mathbf{x}}(z)}{n\eta}} + \Psi(z) + |x_b| \right) (s_2 - s_1) \\
 & \leq (n^{-1/2+O(\epsilon)} + n^{-1/4+O(\epsilon)}(|x_a| + |x_b|) + n^{O(\epsilon)}|x_a||x_b|) (s_2 - s_1)^{1/2}. \quad (4.45)
 \end{aligned}$$

Actually, (4.45) is a result of fact that $p = q = 1$ for $[-RV R]_{\mathbf{x}\mathbf{x}}$ (see page 24 for the definitions of p and q). Hence, when $\{k_1, k_2, k_3, k_4\} = \{0, 1, 1, 3\}$, one has

$$\begin{aligned}
 & n^{-5/2} \sum_{a,b} |\mathbb{E} \mathfrak{F}_0 (\mathfrak{F}_1)^2 \mathfrak{F}_3| \\
 & \leq n^{-5/2} \sum_{a,b} n^{-3/8+O(\epsilon)} (n^{-1/2+O(\epsilon)} + n^{-1/4+O(\epsilon)}(|x_a| + |x_b|) + n^{O(\epsilon)}|x_a||x_b|)^2 \\
 & \quad \times (n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)) (s_2 - s_1)^2 \\
 & \leq n^{-5/2} \sum_{a,b} n^{-3/8+O(\epsilon)} (n^{-1+O(\epsilon)} + n^{-1/2+O(\epsilon)}(|x_a|^2 + |x_b|^2) + n^{O(\epsilon)}|x_a|^2|x_b|^2) \\
 & \quad \times (n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)) (s_2 - s_1)^2 \\
 & \leq n^{-19/8+O(\epsilon)} (s_2 - s_1)^2.
 \end{aligned}$$

Now in the case of $\{k_1, k_2, k_3, k_4\} = \{0, 1, 2, 2\}$, we can perform the estimate as follows:

$$\begin{aligned}
 & n^{-5/2} \sum_{a,b} |\mathbb{E} \mathfrak{F}_0 \mathfrak{F}_1 (\mathfrak{F}_2)^2| \\
 & \leq n^{-5/2} \sum_{a,b} n^{-3/8+O(\epsilon)} \left(n^{-1/2+O(\epsilon)} + n^{-1/4+O(\epsilon)}(|x_a| + |x_b|) + n^{O(\epsilon)}|x_a||x_b| \right) \\
 & \quad \times \left(n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2) \right) (s_2 - s_1)^2 \\
 & \leq n^{-5/2} \sum_{a,b} n^{-3/8+O(\epsilon)} \left(n^{-1/2+O(\epsilon)} + n^{-1/4+O(\epsilon)}(|x_a| + |x_b|) + n^{O(\epsilon)}|x_a||x_b| \right) \\
 & \quad \times \left(n^{-1+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^4 + |x_b|^4) \right) (s_2 - s_1)^2 \\
 & \leq \left(n^{-5/2} \sum_{a,b} n^{-3/8+O(\epsilon)} |x_a||x_b| (n^{-1+O(\epsilon)} + |x_a|^4 + |x_b|^4) + n^{-17/8+O(\epsilon)} \right) (s_2 - s_1)^2.
 \end{aligned}$$

By the elementary fact $\sum_a |x_a| = O(\sqrt{n})$, we obtain

$$\sum_{a,b} |x_a||x_b| = O(n), \quad \sum_{a,b} |x_a|^5 |x_b| = O(\sqrt{n}).$$

Consequently, we have

$$n^{-5/2} \sum_{a,b} |\mathbb{E} \mathfrak{F}_0 \mathfrak{F}_1 (\mathfrak{F}_2)^2| \leq n^{-17/8+O(\epsilon)} (s_2 - s_1)^2.$$

The estimation towards case (iv) is similar. We do it as follows:

$$\begin{aligned}
 & n^{-5/2} \sum_{a,b} |\mathbb{E} (\mathfrak{F}_1)^3 \mathfrak{F}_2| \\
 & \leq n^{-5/2} \sum_{a,b} \left(n^{-1/2+O(\epsilon)} + n^{-1/4+O(\epsilon)}(|x_a| + |x_b|) + n^{O(\epsilon)}|x_a||x_b| \right)^3 \\
 & \quad \times \left(n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2) \right) (s_2 - s_1)^2 \\
 & \leq n^{-5/2} \sum_{a,b} \left(n^{-3/2+O(\epsilon)} + n^{-3/4+O(\epsilon)}(|x_a|^3 + |x_b|^3) + n^{O(\epsilon)}|x_a|^3 |x_b|^3 \right) \\
 & \quad \times \left(n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2) \right) (s_2 - s_1)^2 \\
 & \leq \left(n^{-5/2} \sum_{a,b} |x_a|^3 |x_b|^3 (n^{-1/2+O(\epsilon)} + n^{O(\epsilon)}(|x_a|^2 + |x_b|^2)) + n^{-9/4+O(\epsilon)} \right) (s_2 - s_1)^2 \\
 & \leq n^{-9/4+O(\epsilon)} (s_2 - s_1)^2.
 \end{aligned}$$

In summary, we have

$$|\mathbb{E} n^{-5/2} \mathbf{1}_{\{\sum_{i=1}^4 k_i=5\}} \prod_{i=1}^4 \mathfrak{F}_{k_i}| = n^{-17/8+O(\epsilon)} (s_2 - s_1)^2. \quad (4.46)$$

Now we come to deal with the case of $\kappa = 6$. When there is at least one k_i equal to 0, it will be easy to check that the contributions of such terms are negligible by using (4.32) and (4.37). We leave the details to the readers. Now we still have the case where there is no 0 in $\{k_1, k_2, k_3, k_4\}$. Since $\kappa = 6$, we have

$$\{k_1, k_2, k_3, k_4\} = \{1, 1, 2, 2\} \text{ or } \{1, 1, 1, 3\}.$$

The discussions for these two cases are quite similar to that of (iv) for $\kappa = 5$. Actually, we can get

$$|\mathbb{E} n^{-3} \mathbf{1}_{\{\sum_{i=1}^4 k_i=6\}} \prod_{i=1}^4 \Im \mathcal{F}_{k_i}| = n^{-5/2+O(\epsilon)} (s_2 - s_1)^2. \quad (4.47)$$

Then by (4.44), (4.46), and (4.47) we can complete the comparison procedure. Thus (4.20) follows. So we complete the proof. $\square$

ACKNOWLEDGMENTS

The authors are grateful to the anonymous referee for the careful reading of the paper and valuable comments.

Z. G. Bao was partially supported by NSFC Grant No. 11371317, ZJNSF Grant No. R6090034, and SRFDP Grant No. 20100101110001; G. M. Pan was partially supported by the Ministry of Education, Singapore, under Grant No. ARC 14/11; W. Zhou was partially supported by the Ministry of Education, Singapore, under Grant No. ARC 14/11, and by a Grant No. R-155-000-139-112 at the National University of Singapore.

- ¹ Z. D. Bai and G. M. Pan, "Limiting behavior of eigenvectors of large Wigner matrices," *J. Stat. Phys.* **146**, 519–549 (2012).
- ² F. Benaych-Georges, "A universality result for the global fluctuations of the eigenvectors of Wigner matrices," *Random Matrices: Theory Appl.* **01**, 1250011 (2012).
- ³ F. Benaych-Georges and S. Péché, "Localization and delocalization for heavy tailed band matrices," *Ann. Inst. Henri Poincaré Probab. Stat.* (to appear); e-print [arXiv:1210.7677](https://arxiv.org/abs/1210.7677).
- ⁴ P. Billingsley, *Convergence of Probability Measure* (Wiley, New York, 1968).
- ⁵ C. Bordenave and A. Guionnet, "Localization and delocalization of eigenvectors for heavy-tailed random matrices," *Probab. Theory Relat. Fields* **157**, 885–953 (2013).
- ⁶ L. Erdős and A. Knowles, "Quantum diffusion and eigenfunction delocalization in a random band matrix model," *Commun. Math. Phys.* **303**(2), 509–554 (2011).
- ⁷ L. Erdős, B. Schlein, and H.-T. Yau, "Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices," *Ann. Probab.* **37**(3), 815–852 (2009).
- ⁸ L. Erdős, B. Schlein, and H.-T. Yau, "Local semicircle law and complete delocalization for Wigner random matrices," *Commun. Math. Phys.* **287**, 641–655 (2009).
- ⁹ L. Erdős, H.-T. Yau, and J. Yin, "Bulk universality for generalized Wigner matrices," *Probab. Theory Relat. Fields* **154**, 341–407 (2012).
- ¹⁰ L. Erdős, H.-T. Yau, and J. Yin, "Rigidity of eigenvalues of generalized Wigner matrices," *Adv. Math.* **229**, 1435–1515 (2012).
- ¹¹ A. Knowles and J. Yin, "Eigenvector distribution of Wigner matrices. Probab.," *Theory Relat. Fields* **155**, 543–582 (2013).
- ¹² A. Knowles and J. Yin, "The isotropic semicircle law and deformation of Wigner matrices," *Commun. Pure Appl. Math.* **66**(11), 1663–1749 (2013).
- ¹³ J. Schenker, "Eigenvector localization for random band matrices with power law band width," *Commun. Math. Phys.* **290**(3), 1065–1097 (2009).
- ¹⁴ J. W. Silverman, "Weak convergence of random functions defined by the eigenvectors of sample covariance matrices," *Ann. Probab.* **18**(3), 1174–1194 (1990).
- ¹⁵ T. Tao and V. Vu, "Random matrices: Universality of local eigenvalue statistics," *Acta. Math.* **206**, 127–204 (2011).
- ¹⁶ T. Tao and V. Vu, "Random matrices: Universality of local eigenvalue statistics up to the edge," *Commun. Math. Phys.* **298**(2), 549–572 (2010).
- ¹⁷ T. Tao and V. Vu, "Random matrices: Universal properties of eigenvectors," *Random Matrices: Theory Appl.* **01**, 1150001 (2012).