Title	Limit theorems for linear spectrum statistics of orthogonal polynomial ensembles and their applications in random matrix theory
Author(s)	Pan, Guangming; Wang, Shaochen; Zhou, Wang
Citation	Pan, G., Wang, S., & Zhou, W. (2017). Limit theorems for linear spectrum statistics of orthogonal polynomial ensembles and their applications in random matrix theory. Journal of Mathematical Physics, 58(10), 103301-.
Date	2017
URL	http://hdl.handle.net/10220/44176
Rights	

JOURNAL OF MATHEMATICAL PHYSICS **58**, 103301 (2017)

Limit theorems for linear spectrum statistics of orthogonal polynomial ensembles and their applications in random matrix theory

Guangming Pan,[1,a)] Shaochen Wang,[2,b)] and Wang Zhou[3,c)]
[1]*Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371*
[2]*School of Mathematics, South China University of Technology, Guangzhou 510640, China*
[3]*Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546*

(Received 10 November 2016; accepted 25 September 2017; published online 9 October 2017)

In this paper, we consider the asymptotic behavior of $X_{f_n}^{(n)} := \sum_{i=1}^n f_n(x_i)$, where x_i, $i = 1, \ldots, n$ form orthogonal polynomial ensembles and f_n is a real-valued, bounded measurable function. Under the condition that $\mathrm{Var}X_{f_n}^{(n)} \to \infty$, the Berry-Esseen (BE) bound and Cramér type moderate deviation principle (MDP) for $X_{f_n}^{(n)}$ are obtained by using the method of cumulants. As two applications, we establish the BE bound and Cramér type MDP for linear spectrum statistics of Wigner matrix and sample covariance matrix in the complex cases. These results show that in the edge case (which means f_n has a particular form $f(x)I(x \geq \theta_n)$ where θ_n is close to the right edge of equilibrium measure and f is a smooth function), $X_{f_n}^{(n)}$ behaves like the eigenvalues counting function of the corresponding Wigner matrix and sample covariance matrix, respectively. *Published by AIP Publishing.* https://doi.org/10.1063/1.5006507

I. INTRODUCTION

Assume that the random points $x_1, \ldots, x_n$ ($x_i \in \mathbb{R}, i = 1, \ldots, n$) have a joint density function

$$\mathscr{P}(\mathrm{d}x_1, \ldots, \mathrm{d}x_n) = \frac{1}{Z_n} \prod_{j<k} (x_j - x_k)^2 \prod_{k=1}^n e^{-V(x_i)} \mathrm{d}\mu(x_1) \cdots \mathrm{d}\mu(x_n), \tag{1.1}$$

where Z_n is a normalization constant, μ is a measure, and V is some real-valued function such that $\int x^k e^{-V(x)} \mathrm{d}\mu(x)$ is finite for all positive integers k. We call $\{x_1, \ldots, x_n\}$ orthogonal polynomial ensembles with size $n \in \mathbb{N}$. During the past decades, many interesting models (such as non-colliding random walks, growth models, and last passage percolation) have been described by (1.1) with suitable choice of V and measure μ. One can see the work of Anderson *et al.*[1] and König[21] for a survey of this topic.

In random matrix theory (RMT), especially in statistical applications, there are two important class of random matrices which can be formulated by (1.1), i.e., the Gaussian Unitary Ensemble (GUE, with $V(x) = x^2, x \in \mathbb{R}$) and Laguerre Unitary Ensemble (LUE, with $V_n(x) = x - (1 - m/n)\log x, x \in \mathbb{R}_+$), where in both cases, μ is the Lebesgue measure. In RMT, we usually call $x_1, \ldots, x_n$ eigenvalues.

In this paper, we are interested in the asymptotic behavior of

$$X_{f_n}^{(n)} = \sum_{i=1}^n f_n(x_i) \tag{1.2}$$

[a)]E-mail: gmpan@ntu.edu.sg
[b)]Author to whom correspondence should be addressed: mascwang@scut.edu.cn
[c)]E-mail: stazw@nus.edu.sg

for some real-valued measurable function f_n. In the words of RMT, $X_{f_n}^{(n)}$ is known as linear spectrum statistics. There are extensive studies on this issue. One can see the literature reviews in Subsections III A and III B, respectively. A standard and effective approach to studying (1.1) is using orthogonal polynomials. Many important results can be obtained by this method, such as the distribution of largest eigenvalue and the spacing distribution. One can refer to the work of Deift[10] or König[21] for reference about this topic.

For general orthogonal polynomial ensembles, there are also many papers on the asymptotic behavior of $X_{f_n}^{(n)}$. For example, when $f_n(x) = I(a \le x \le b)$ for some $a < b \in \mathbb{R}$ ($X_{f_n}^{(n)}$ is known as the eigenvalues counting function), Costin and Lebowitz[8] and Ben Hough *et al.*[20] proved a central limit theorem (CLT) for $X_{f_n}^{(n)}$. After that, Döring and Eichelsbacher[15] proved the corresponding moderate deviation principle (MDP) for $X_{f_n}^{(n)}$. Chen, Gao, and Wang[7] also obtained a similar result independently. Besides, we would like to mention Breuer and Duits,[5] who obtained the local law of large numbers for $X_{f_n}^{(n)}$. Their approach is based on the asymptotic expansion of Fredholm determinants. As a byproduct, their results also give the estimates of cumulants of $X_{f_n}^{(n)}$. Using their method, in this paper, we establish the Berry-Esseen bound and the Cramér type MDP for $X_{f_n}^{(n)}$. The Berry-Esseen bound and Cramér MDP characterize the limiting distribution of $X_{f_n}^{(n)}$ and also provide the convergence rates in different scales. One can refer to the work of Dembo and Zeitouni[11] for more history and theories on large and moderate deviations.

For simplicity, in the rest of paper, we may write X_{f_n} instead of $X_{f_n}^{(n)}$. If $x_1, \ldots, x_n$ are the eigenvalues of a matrix W_n, we write X_{f_n} as $X_{f_n}(W_n)$, and if there is no confusion, we will compress the notation W_n. The uniform norm of a function f is defined as $\|f\|_\infty = \sup_x |f(x)|$. The constant C may change from line to line (even in one line), but C will not depend on n.

Our main result of this paper is the following theorem.

Theorem 1.1. *Let $(x_1, \ldots, x_n)$ be random with density (1.1) and f_n a real-valued measurable function with uniformly bounded $\|f_n\|_\infty$. Define $X_{f_n} = \sum_{i=1}^n f_n(x_i)$ and*

$$\xi_n := \frac{X_{f_n} - \mathbb{E}X_{f_n}}{\sqrt{\mathrm{Var}X_{f_n}}} \tag{1.3}$$

and its distribution function $F_{\xi_n}(x) := \mathbb{P}(\xi_n \le x)$ for all $x \in \mathbb{R}$. Assuming $\mathrm{Var}X_{f_n} \to \infty$, we have

(1). *(Berry-Esseen bound). There exists a universal constant $C > 0$ such that*

$$\sup_{x\in\mathbb{R}} |F_{\xi_n}(x) - \Phi(x)| \le \frac{C}{\sqrt{\mathrm{Var}X_{f_n}}},$$

where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution.

(2). *(Cramér type MDP). For any positive constant ρ, when $x \in [0, \rho(\mathrm{Var}X_{f_n})^{1/6})$, we have*

$$\frac{1 - F_{\xi_n}(x)}{1 - \Phi(x)} = 1 + O(1)\frac{1 + x^3}{\sqrt{\mathrm{Var}X_{f_n}}} \quad \text{and} \quad \frac{F_{\xi_n}(-x)}{\Phi(-x)} = 1 + O(1)\frac{1 + x^3}{\sqrt{\mathrm{Var}X_{f_n}}}. \tag{1.4}$$

As a corollary of Cramér type MDP in (2), we obtain the following CLT and MDP (cf. the proof of Theorem 1.1 of Döring and Eichelsbacher[14] and Fan *et al.*[17]).

Corollary 1. Under the setting of Theorem 1.1, we have

(1). *(CLT). ξ_n converges to standard normal distribution $N(0,1)$ in distribution as $n \to \infty$.*

(2). *(MDP). Let l_n be a sequence of positive numbers such that $l_n \to \infty$ and $l_n/(\mathrm{Var}X_{f_n})^{1/2} \to 0$ as $n \to \infty$. Then $\{\xi_n l_n^{-1}, n \ge 1\}$ satisfies MDP on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, with speed l_n^2 and good rate function $I(x) = x^2/2$, $x \in \mathbb{R}$. Particularly, for any $x \ge 0$,*

$$\lim_{n\to\infty} \frac{1}{l_n^2} \log \mathbb{P}\left(\left|\frac{X_{f_n} - \mathbb{E}X_{f_n}}{\sqrt{\mathrm{Var}X_{f_n}}}\right| \ge l_n x\right) = -\frac{x^2}{2}. \tag{1.5}$$

Remark 1. The assumption $\mathrm{Var}X_{f_n} \to \infty$ *in Theorem 1.1 is crucial, see the proof in Sec. II for more details. In RMT, for smooth* f_n *(here "smooth" can be weaken to "with only finitely many fixed discontinuous points"), we know that from the work of Bao, Pan and Zhou,[4] Kopel,[22] Lytova and Pastur,[23] Sosoe and Wong[30] and so on, there is no scaling factor before* $X_{f_n} - \mathbb{E}X_{f_n}$ *(where the variance is* $O(1)$*). This is mainly caused by the repulsion properties of the eigenvalues, cf. Diaconis.[12]*

When applying the above results to Wigner and sample covariance matrix, we choose $f_n(x) = f(x)I(x \geq \theta_n)$ *for some smooth function* f *and* θ_n *approaching the edge of spectrum with a suitable speed. We prove in this setting that* $\mathrm{Var}X_{f_n} \to \infty$*. Thus we believe that the approach used in this paper cannot be used to deal with the bulk case (which means that* θ_n *is fixed in the bulk of spectrum) in the work of Bao, Pan, and Zhou.[4] It is still an open problem to obtain similar results to Theorem 1.1 under the assumption that* $\mathrm{Var}X_{f_n}$ is $O(1)$.

As applications of Theorem 1.1, in Sec. III we consider the Berry-Esseen bound and Cramér type MDP for linear spectrum statistics of Wigner matrix (Subsection III A) and sample covariance matrix (Subsection III B) in the edge case. Our approach is based on the result of Theorem 1.1 and some comparison arguments. The main tasks are to verify the conditions of Theorem 1.1 and provide explicit expressions for the mean and variance of X_{f_n} in these two cases. Under the Gaussian assumption, by a Taylor expansion technique, the expressions for the mean and variance of X_{f_n} can be reduced to those of the mean and variance of the corresponding eigenvalues counting function which have been studied extensively. Then the results can be extended to general Wigner matrices and covariance matrices by some comparison methods developed by Erdös, Yau and Yin,[16] and Tao and Vu.[32,33]

The paper is organized as follows: In Sec. II, we prove the Berry-Esseen bound and Cramér MDP for orthogonal polynomial ensembles. Our proofs are based on the fine estimates of cumulants of X_{f_n}. Section III is devoted to studying the linear spectrum statistics for Wigner and sample covariance matrices in edge cases. Some technical proofs are postponed to Appendixes A and B.

II. BERRY-ESSEEN BOUND AND CRAMÉR MDP FOR ORTHOGONAL POLYNOMIAL ENSEMBLES

Let $\{x_1, \ldots, x_n\}$ be random points as defined in (1.1). Then $\{x_1, \ldots, x_n\}$ is said to be a determinantal point process with associated kernel K_n defined in the following. Let $p_j(x) = \gamma_j x^j +, \cdots$, $\gamma_j > 0, j \geq 0$, be the jth orthogonal polynomial with respect to the weight $e^{-nV(x)}$ and measure μ. That is

$$\int_{\mathbb{R}_+} p_i(x)p_j(x)e^{-nV_n(x)}\mathrm{d}\mu(x) = \delta_{ij},$$

for all nonnegative integers i, j. Define the correlation kernel,

$$K_n(x,y) = \sum_{j=0}^{n-1} p_j(x)p_j(y)e^{-\frac{n}{2}(V(x)+V(y))}, \quad x, y \in \mathbb{R}_+. \tag{2.1}$$

From the work of Breuer and Duits [Ref. 5, Eq.(2.13)], one has the following explicit formulas of expectation and variance of $X_{f_n}^{(n)}$:

$$\mathbb{E}X_{f_n}^{(n)} = \int_{\mathbb{R}} f_n(x)K_n(x,x)\mathrm{d}x \tag{2.2}$$

and

$$\mathrm{Var}X_{f_n}^{(n)} = \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} (f_n(x) - f_n(y))^2 K_n^2(x,y)\mathrm{d}x\mathrm{d}y. \tag{2.3}$$

In the following, (2.2) and (2.3) will be used frequently.

Assume that b_n is a sequence of positive constants so that $b_n^{-1}(X_{f_n} - \mathbb{E}X_{f_n})$ converges to a nontrivial distribution. We denote for any fixed $t \in \mathbb{R}$

$$\varphi[f_n, K_n] := \mathbb{E}e^{tb_n^{-1}(X_{f_n} - \mathbb{E}X_{f_n})} = e^{-tb_n^{-1}\mathbb{E}X_{f_n}}\mathbb{E}e^{tb_n^{-1}X_{f_n}}, \tag{2.4}$$

i.e., the Laplace transform of $b_n^{-1}(X_{f_n} - \mathbb{E}X_{f_n})$. Observe that if $|tb_n^{-1}f_n|$ is sufficiently small (so that $\|(e^{tb_n^{-1}f_n} - 1)\mathscr{K}\|_\infty < 1$, where $\mathscr{K}$ is the integral operator acting on L^2 with kernel function K_n, see the work of Anderson *et al.*[1] (Sec. III D) for a brief introduction to the Fredholm determinants theory), the expectation in (2.4) can be expressed in the Fredholm determinants,

$$\mathbb{E}e^{tb_n^{-1}X_{f_n}} = \det(1 + (e^{tb_n^{-1}f_n} - 1)\mathscr{K}). \tag{2.5}$$

Consequently,

$$\log \varphi[f_n, K_n] = \log\det(1 + (e^{tb_n^{-1}f_n} - 1)\mathscr{K}) - tb_n^{-1}\mathbb{E}X_{f_n}. \tag{2.6}$$

Before moving on, we give the definition of cumulants. Given any random variable ξ, we assume that its jth moment is finite. Then the jth cumulant of ξ is defined as

$$\mathscr{C}_j(\xi) = \frac{1}{i^j}\frac{\mathrm{d}^j}{\mathrm{d}t^j}\log\mathbb{E}(e^{it\xi})\bigg|_{t=0}, \tag{2.7}$$

where $i = \sqrt{-1}$. Moreover, if the Laplace transform of ξ exists and $\mathbb{E}(e^{t\xi}) \neq 0$ and finite for all t in some neighborhood of zero, we can equivalently define $\mathscr{C}_j(\xi)$ as

$$\mathscr{C}_j(\xi) = \frac{\mathrm{d}^j}{\mathrm{d}t^j}\log\mathbb{E}(e^{t\xi})\bigg|_{t=0}. \tag{2.8}$$

Note that cumulants and moments are in one to one correspondence. Thus if the limiting distribution is totally determined by moments, we can derive the asymptotic results by showing the jth cumulant convergence to the jth cumulant of limiting distribution instead of moments convergence. For example, the standard normal distribution has only one non-vanishing cumulant, the 2nd one and thus proving convergence in distribution to standard normal distribution is equivalent to proving that the 2nd cumulant converges to 1 and all other jth cumulants converge to 0.

The following lemma is a corollary of Lemmas 2.1 and 2.3 in the work of Saulis and Statulevicius.[27] It shows that proper bounds of cumulants can give suitable convergence rates of distributions. Its proof can be found in the work of Chen, Gao, and Wang.[7]

Lemma 2.1. Assume that $\xi = \xi(n)$ has mean zero and variance one. Suppose there exists $\Delta = \Delta(n) > 0$ such that $\Delta(n) \to \infty$ as $n \to \infty$ and

$$|\mathscr{C}_j(\xi)| \le \frac{j!}{\Delta^{j-2}}, \qquad j = 3, 4, \ldots. \tag{2.9}$$

Set $F_\xi(x) = \mathbb{P}(\xi \le x)$.

(1). *Then there exists a constant $C > 0$ such that*

$$\sup_{x\in\mathbb{R}}\left|F_\xi(x) - \Phi(x)\right| \le C\Delta^{-1},$$

where $\Phi(\cdot)$ is the standard normal distribution function.

(2). *For any constant ρ, we have*

$$\frac{1 - F_\xi(x)}{1 - \Phi(x)} = 1 + O(1)\frac{1 + x^3}{\Delta} \quad \text{and} \quad \frac{F_\xi(-x)}{\Phi(-x)} = 1 + O(1)\frac{1 + x^3}{\Delta}, \tag{2.10}$$

uniformly in $x \in [0, \rho\Delta^{1/3})$.

From the above lemma and Eq. (2.6), in order to derive the Berry-Esseen bound and Cramér MDP, it is sufficient to estimate the cumulants of X_{f_n}. The following lemma is slightly adapted from the work of Breuer and Duits (Ref. 5, Lemma 2.2). For convenience of readers, we sketch its proof.

Lemma 2.2. Let $\mathscr{K}$ be a self-adjoint projection operator (i.e., $\mathscr{K}^2 = \mathscr{K}$ and $\mathscr{K}^* = \mathscr{K}$) on a separable Hilbert space. Then for any bounded operator h such that $h\mathscr{K}$ and $\mathscr{K}h$ are trace class operators, there exists a constant $A > 0$ such that

$$\left|\log\det(1 + (\mathrm{e}^{th} - 1)\mathscr{K}) - t\mathrm{Tr}h\mathscr{K} - \frac{t^2}{2}\mathrm{Tr}(h\mathscr{K}h\mathscr{K} - h^2\mathscr{K})\right| \le A\|h\|_\infty|t|^3\,\|[h,\mathscr{K}]\|_2^2, \tag{2.11}$$

for $|t| \le 1/(3\|h\|_\infty)$. Here $\|\cdot\|_2$ denotes the Hilbert-Schmidt norm, $\|\cdot\|_\infty$ denotes the operator norm, and $[h, \mathscr{K}] = h\mathscr{K} - \mathscr{K}h$ stands for the commutator of h and $\mathscr{K}$.

Moreover, for any $m \ge 3$,

$$|\mathscr{C}_m^{(n)}| \le \frac{1}{\sqrt{2\pi}} m^{3/2} e^m \|h\|_\infty^{m-2} \|[h, \mathscr{K}]\|_2^2\, m!, \tag{2.12}$$

where $\mathscr{C}_m^{(n)}$ is the m-th cumulant of $X_h = \sum_{i=1}^n h(x_i)$, $\{x_1, \ldots, x_n\}$ is defined as (1.1) and here $\mathscr{K}$ is its associated integral operator with kernel K_n defined in (2.1).

Proof. It follows from Lemma 2.2, Eq. (2.5) in the work of Breuer and Duits[5] that

$$\log\det(1 + (e^{th} - 1)\mathscr{K}) - t\mathrm{Tr}h\mathscr{K} = \sum_{m=2}^{\infty} t^m \sum_{j=2}^{m} \frac{(-1)^{j+1}}{j} \sum_{\substack{l_1+\cdots+l_j=m \\ l_i \ge 1}} \frac{\mathrm{Tr}h^{l_1}\mathscr{K}\cdots h^{l_j}\mathscr{K} - \mathrm{Tr}h^m\mathscr{K}}{l_1!\cdots l_j!},$$

where the last summation is over all positive integers $l_1, \ldots, l_j$ such that $l_1 + \cdots + l_j = m$. For $m \ge 3$, we use the following estimates [Breuer and Duits,[5] Lemma 2.2, Eq.(2.6)]:

$$\left|\mathrm{Tr}h^{l_1}\mathscr{K}\cdots h^{l_j}\mathscr{K} - \mathrm{Tr}h^m\mathscr{K}\right| \le jm^2\|h\|_\infty^{m-2}\|[h, \mathscr{K}]\|_2^2, \tag{2.13}$$

for any $l_1, \ldots, l_j$ and $m \ge 2$ such that $l_1+, \cdots + l_j = m$ and $l_i \ge 1$. Using

$$\sum_{j=1}^{m} \sum_{\substack{l_1+\cdots+l_j=m \\ l_i \ge 1}} \frac{m!}{l_1!\cdots l_j!} = (\underbrace{1 + 1 + \cdots + 1}_{m})^m = m^m$$

and (2.13), we have

$$\begin{aligned} \text{LHS of (2.11)} &\le \|[h, \mathscr{K}]\|_2^2 \sum_{m=3}^{\infty} |t|^m \|h\|_\infty^{m-2} \frac{m^{m+2}}{m!} \\ &\le \frac{1}{\sqrt{2\pi}} \|[h, \mathscr{K}]\|_2^2 |t|^3 \sum_{m=3}^{\infty} |t|^{m-3} \|h\|_\infty^{m-2} m^{3/2} e^m, \end{aligned} \tag{2.14}$$

where in the above last inequality we used the Stirling estimates $m! \ge \sqrt{2\pi} m^{m+1/2} e^{-m}$. Note that

$$\sum_{m=3}^{\infty} |t|^{m-3} \|h\|_\infty^{m-2} m^{3/2} e^m \le C\|h\|_\infty \sum_{m=3}^{\infty} (|t|(e + 0.1)\|h\|_\infty)^{m-3}$$

for some universal constant C. Now choosing $t\, (t \le 1/(3\|h\|_\infty))$, the assertion (2.11) follows immediately.

Next we observe that for the orthogonal polynomial ensembles defined in (1.1) with its correlation kernel (2.1) and associated integral operator $\mathscr{K}$, it is well known that $\mathscr{K}$ fulfills the assumptions of this lemma. Thus from the definition of cumulant and the estimate in (2.14), we have (2.12). □

Remark 2. Note that for the orthogonal polynomial ensembles defined in (1.1) with its correlation kernel (2.1) and associated integral operator $\mathscr{K}$, the coefficient of t^2 in the above lemma reduces to

$$-\frac{1}{2}\mathrm{Tr}(h\mathscr{K}h\mathscr{K} - h^2\mathscr{K}) = -\frac{1}{2}\mathrm{Tr}(h\mathscr{K} - \mathscr{K}h)(h\mathscr{K} - \mathscr{K}h) = -\frac{1}{4}\|[h, \mathscr{K}]\|_2^2 = -\frac{1}{2}\mathrm{Var}(X_h),$$

where we have used the facts $\mathrm{Tr}AB = \mathrm{Tr}BA$, $\mathscr{K}^2 = \mathscr{K}$ and h is a real-valued function. The deviation bound (2.11) is sufficient for the CLT if $\mathrm{Var}(X_h)$ tends to infinity. (2.12) can be used to get the Berry-Esseen bound and Cramér MDP. Although the Berry-Esseen bound implies the CLT, (2.11) is useful when the bounds for cumulants are not available.

Now we can start the proof.

Proof of Theorem 1.1. The proof of the Berry-Esseen bound and Cramér type MDP for orthogonal polynomial ensembles is based on the cumulant estimates given in Lemma 2.1. Note that ξ_n has mean zero and variance one, so by Lemma 2.2 and Remark 2,

$$|\mathscr{C}_m^{(n)}(X_{f_n})| \leq \frac{1}{\sqrt{2\pi}} m^{3/2} e^m \|f_n\|_\infty^{m-2} m! \, \|[f_n, \mathscr{K}]\|_2^2$$
$$= \sqrt{\frac{2}{\pi}} m^{3/2} e^m \|f_n\|_\infty^{m-2} m! \mathrm{Var} X_{f_n}. \tag{2.15}$$

It is easy to check that for any random variable η and constant c, $\mathscr{C}_m(\eta + c) = \mathscr{C}_m(\eta)$ and $\mathscr{C}_m(c\eta) = c^m \mathscr{C}_m(\eta)$ for all $m \geq 2$. Applying this fact to ξ_n and (2.15), we derive

$$\mathscr{C}_m^{(n)}(\xi_n) \leq \frac{1}{\sqrt{2\pi}} m^{3/2} e^m \|f_n\|_\infty^{m-2} \frac{m!}{[\mathrm{Var} X_{f_n}]^{(m-2)/2}}$$
$$\leq \frac{m!}{[l_m \mathrm{Var} X_{f_n}]^{(m-2)/2}},$$

for all $m \geq 3$, where

$$l_m = \left(\frac{\sqrt{2\pi}}{m^{3/2} e^m \|f_n\|_\infty^{m-2}} \right)^{2/(m-2)}.$$

Note that we have assumed that $\|f_n\|_\infty$ is uniformly bounded from above, and thus it is easy to check that l_m is bounded from below (in fact,

$$\log l_m = \frac{2}{m-2} \left\{ \log \sqrt{2\pi} - (m-2) \log \|f_n\|_\infty - m - \frac{3}{2} \log m \right\}$$

which tends to $-2 - 2\log \|f_n\|_\infty$ as $m \to \infty$. This shows that l_m is bounded away from zero for all $m \geq 3$). We denote this lower bound by C ($C > 0$), so we get that

$$\mathscr{C}_m^{(n)}(\xi_n) \leq \frac{m!}{[C \mathrm{Var} X_{f_n}]^{(m-2)/2}}. \tag{2.16}$$

Applying this estimate in Lemma 2.1 with $\Delta = \sqrt{C \mathrm{Var} X_{f_n}}$ completes the proof. □

III. APPLICATIONS IN RANDOM MATRIX THEORY

A. Berry-Esseen bound and Cramér type MDP for linear spectrum statistics of Wigner matrix in the edge case

The complex Wigner Ensemble is defined as a family of $n \times n$ random Hermitian matrices M_n of the form

$$W_n = \frac{1}{\sqrt{n}} M_n = \frac{1}{\sqrt{n}} \{\xi_{jk}\}_{j,k=1}^n,$$

in which $\xi_{ll} \in \mathbb{R}, 1 \leq l \leq n$, $\xi_{jk} = \bar{\xi}_{kj} \in \mathbb{C}, 1 \leq j < k \leq n$, and $\{\xi_{ll}, \xi_{jk}; 1 \leq i \leq n, 1 \leq j < k \leq n\}$ is a collection of independent variables such that $\mathbb{E}\xi_{ll} = \mathbb{E}\xi_{jk} = 0$, $\mathbb{E}|\xi_{jk}|^2 = 1$, and $\mathbb{E}\xi_{ll}^2 = \sigma^2 < \infty$. Moreover, if the entries are Gaussian distributed, i.e.,

$$\xi_{ll} \sim N(0,1)_{\mathbb{R}}, \quad 1 \leq l \leq n, \quad \xi_{jk} \sim N(0,1)_{\mathbb{C}}, \quad 1 \leq j < k \leq n,$$

where $N(0,1)_{\mathbb{R}}$ (respectively, $N(0,1)_{\mathbb{C}}$) represents the standard real (respectively, complex) Gaussian distribution, we say that W_n is Gaussian Unitary Ensemble (GUE).

For Wigner matrix W_n, we denote its ordered eigenvalues as $\lambda_1(W_n) \leq \lambda_2(W_n) \leq \cdots \leq \lambda_n(W_n)$. The famous semi-circle law says that the empirical spectral distribution (ESD) of W_n which is defined by

$$F^{W_n}(x) =: \frac{1}{n} \sum_{i=1}^n I(\lambda_i(W_n) \leq x)$$

converges weakly to the semi-circle law ρ_{sc}, where

$$\rho_{sc}(\mathrm{d}x) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathrm{d}x, \quad -2 \leq x \leq 2.$$

This means for any bounded and continuous function f,

$$\frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(W_n)) \to \int f(x)\rho_{sc}(\mathrm{d}x)$$

in probability as $n\to\infty$. The above convergence can be proved under weaker conditions. One can refer to the work of Bai and Silverstein[2] as one of the standard references in random matrix theory. Note that the above result describes the global limiting behavior of eigenvalues of Wigner ensembles, which can be seen as a universal result corresponding to the classical law of large numbers (LLN) for sums of independent random variables.

A natural question is to consider the fluctuation problem for $\sum_{i=1}^{n} f(\lambda_i(W_n))$. This question has been studied extensively in literature in different settings, for example one can refer to the work of Bai, Wang and Zhou,[3] Lytova and Pastur,[23] Pastur,[25] Shcherbina,[28] and Soshnikov.[29] One can see the work of Bao, Pan, and Zhou[4] for a review and more references. For convenience, we denote

$$X_f^{(n)}(W_n) = \sum_{i=1}^{n} f(\lambda_i(W_n)).$$

$X_f^{(n)}(W_n)$ is usually referred as the global linear spectrum statistics of Wigner matrices with test function f. A remarkable work on this topic is due to Lytova and Pastur.[23] Particularly for GUE, Lytova and Pastur showed that for any bounded test function f with bounded derivatives,

$$X_f^{(n)}(W_n) - \mathbb{E}X_f^{(n)}(W_n) \longrightarrow_d N(0, V_f), \tag{3.1}$$

where $\longrightarrow_d$ means convergence in distribution and

$$V_f = \frac{1}{4\pi^2}\int_{-2}^{2}\int_{-2}^{2}\left(\frac{f(\lambda)-f(\mu)}{\lambda-\mu}\right)^2 \frac{4-\lambda\mu}{\sqrt{4-\lambda^2}\sqrt{4-\mu^2}}\mathrm{d}\lambda\mathrm{d}\mu. \tag{3.2}$$

Moreover, under more smoothness assumptions on the test function f, they extended the above result to general Wigner matrices. See Kopel[22] (for GUE), Sosoe and Wong[30] (for Wigner matrix) for regularity conditions on f which make the CLT hold. It is remarkable that there is no normalizing constant for the difference in (3.1). This is mainly caused by the repulsion properties of the eigenvalues.[12]

In a recent paper by Bao, Pan, and Zhou,[4] they extended the above results to the case $f_u(x) = g(x)I(x\geq u)$, where g is a smooth function and $u\in(-2,2)$ is a fixed number. They obtained that

$$X_{f_u}^{(n)}(W_n) - \mathbb{E}X_{f_u}^{(n)}(W_n) \longrightarrow_d N(0, V_{f_u}), \tag{3.3}$$

where

$$V_{f_u} = \frac{1}{4\pi^2}\int_{u}^{2}\int_{u}^{2}\left(\frac{f(\lambda)-f(\mu)}{\lambda-\mu}\right)^2 \frac{4-\lambda\mu}{\sqrt{4-\lambda^2}\sqrt{4-\mu^2}}\mathrm{d}\lambda\mathrm{d}\mu. \tag{3.4}$$

They also established the weak convergence of the corresponding process by considering u as a parameter in the space $\mathbb{D}$. Their approach is based on the idea of Lytova and Pastur[23] and comparison theorems for linear eigenvalue statistics. Note that they require that $u\in(-2,2)$, i.e., u should be in the bulk region. When $f_u(x) = I(x\geq u)$ (in this case $X_{f_u}^{(n)}(W_n)$ is known as eigenvalues counting function) and u is close to the edge, Dallaporta and Vu[9] established a central limit theorem for $X_{f_u}^{(n)}(W_n)$ in the Wigner matrix case, Döring and Eichelsbacher[13,14] also obtained the same CLT and MDP results for the same quantity. In this subsection, we will continue to study the Berry-Esseen bound and Cramér type MDP for more general f_u.

For the Wigner matrix W_n, we assume the following classical condition throughout this subsection.

Condition $\mathbf{C_0}$: *We say that a complex Wigner matrix* W_n *obeys condition* $\mathbf{C_0}$ *if* $\{\xi_{ll}, \mathrm{Re}\xi_{jk}, \mathrm{Im}\xi_{jk}; 1\leq l\leq n, 1\leq j<k\leq n\}$ *is a collection of independent variables whose distributions are all supported on at least three points, and we have the exponential decay condition on the elements in the sense that*

$$\mathbb{P}(|\xi_{jk}| \geq t^C) \leq e^{-t}$$

holds for all $t \geq C'$ with some positive constants C, C' (independent of l,j,k,n).

If ξ_{ij} is Gaussian, condition $\mathbf{C_0}$ holds trivially. We say that two complex random variables ξ and ξ' match up to order k if

$$\mathbb{E}\left[\Re(\xi)^m \Im(\xi)^l\right] = \mathbb{E}\left[\Re(\xi')^m \Im(\xi')^l\right]$$

for all $m, l \geq 0$ such that $m+l \leq k$, where $\Re(\xi)$, $\Im(\xi)$ are the real and imaginary parts of ξ, respectively.

Our main result of this subsection is listed as follows.

Theorem 3.1 (Wigner matrix). *Let $(\theta_n)_{n\geq 1}$ be a sequence of numbers such that $-2 < \theta_n < 2$ and*

$$\theta_n \to 2^-, \qquad n(2-\theta_n)^{3/2} \to \infty, \tag{3.5}$$

as $n \to \infty$. Let f be differentiable and have bounded derivative in the neighborhood of 2. Moreover, suppose $f(2) \neq 0$ and

$$\frac{n^2(2-\theta_n)^5}{\log[n(2-\theta_n)^{3/2}]} \to 0 \tag{3.6}$$

as $n \to \infty$. Denote $f_n(x) = f(x)I(x \geq \theta_n)$, $X_{f_n}^{(n)}(W_n) = \sum_{i=1}^n f_n(\lambda_i(W_n))$, and

$$\xi_n(W_n) := \frac{X_{f_n}^{(n)}(W_n) - \frac{2}{3\pi} f(2) n(2-\theta_n)^{3/2}}{\sqrt{\frac{f^2(2)}{2\pi^2}\log[n(2-\theta_n)^{3/2}]}}. \tag{3.7}$$

Then the conclusions of Theorem 1.1 hold for $\xi_n(W_n)$, where W_n is drawn from GUE.

Moreover, if W_n is a Wigner matrix whose off-diagonal entries match up to order 4 and diagonal entries match up to order 2, to the corresponding entries of a GUE matrix, the conclusions of Theorem 1.1 also hold for $\xi_n(W_n)$ in this case.

First we give the explicit expressions for the mean and variance of $X_{f_n}^{(n)}$.

Lemma 3.1 (Pan, Wang, and Zhou[26]).

$$\mathbb{E}X_{f_n}^{(n)}(W_n) = \frac{2}{3\pi} f(2) n(2-\theta_n)^{3/2} + \max\left\{O(1)n(2-\theta_n)^{5/2}, O(1)\right\} \tag{3.8}$$

and

$$\mathrm{Var}X_{f_n}^{(n)}(W_n) = (1+o(1))\frac{f^2(2)}{2\pi^2}\log[n(2-\theta_n)]^{3/2}, \tag{3.9}$$

where W_n is drawn from GUE or W_n is Wigner matrices which satisfies the moment matching condition in Theorem 3.1.

Proof of Theorem 3.1. For the GUE case, the above lemma implies $\mathrm{Var}X_{f_n}^{(n)}(W_n) \to \infty$. So the result of Theorem 3.1 follows.

The proofs of Berry-Esseen bound and Cramér type MDP for general Wigner matrix are left in Appendix A. □

Remark 3. We would like to point out the difference between Ref. 26 and the results in this subsection. In Ref. 26, the authors obtained the law of large number and CLT for $X_{f_n}^{(n)}(W_n)$ under the same conditions as Theorem 3.1. The Berry-Esseen bound gives the bound on the difference of two cumulative distribution functions (which implies CLT). The approach used in Ref. 26 is a comparison with eigenvalues counting function, while in this paper, we use the estimate of cumulants.

Remark 4. Condition (3.5) is required in the Gaussian fluctuation of eigenvalue counting function of GUE in the edge case, see the work of Gustavsson[19] for details. When $f \equiv 1$, our result recovers the CLT for eigenvalue counting function, i.e.,

$$\frac{N_{I_n}(W_n) - \frac{2}{3\pi} n(2-\theta_n)^{3/2}}{\sqrt{\frac{1}{2\pi^2}\log[n(2-\theta_n)^{3/2}]}} \longrightarrow_d N(0,1)$$

under the condition (3.6), where $N_{I_n}(W_n)$ *means the number of eigenvalues that fall into the interval* $I_n = [\theta_n, +\infty)$. *This condition can be dropped if we only use the formal expression* $\mathbb{E}X_{f_n}$ *and* $\mathrm{Var}X_{f_n}$ *for centering and scaling. Similar condition (3.16) occurs in the case of sample covariance matrix. In fact, if one checks carefully the computation of mean* $\mathbb{E}X_{f_n}$ *[cf. Eq. (B6) in Appendix B for LUE], the term*

$$\max\left\{O(1)n(2-\theta_n)^{5/2}, O(1)\right\}$$

cannot be dropped since we do not know if this is an ignored term compared to variance.

B. Berry-Esseen bound and Cramér type MDP for linear spectrum statistics of sample covariance matrix in edge case

A parallel result is also established for sample covariance matrices (SCM). SCM are Hermitian or real symmetric semidefinite matrices $S_{m,n}$ such that $S_{m,n} = X^*X/n$ where X is an $m \times n$ random complex or real matrix (with $m \geq n$) whose entries are iid with mean 0 and variance 1. We only consider here the situation where $m/n \to \gamma \in [1, +\infty)$ as $n \to \infty$. If the entries are Gaussian, then the covariance matrix becomes the so-called Laguerre Unitary Ensemble (LUE) if it is complex and Laguerre Orthogonal Ensemble (LOE) if it is real.

Let $\lambda_1(S_{m,n}) \leq \lambda_2(S_{m,n}) \leq \cdots \leq \lambda_n(S_{m,n})$ be the real, ordered eigenvalues of $n \times n$ matrix $S_{m,n}$. The famous Marchenko-Pastur law[24] says that as $n \to \infty$ and $m/n \to \gamma \geq 1$, almost surely

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(S_{m,n})} \to \mu_\gamma, \tag{3.10}$$

where μ_γ is known as the MP law with parameter γ and supported on $[\alpha, \beta]$, whose density is

$$\mu_\gamma(\mathrm{d}x) = \frac{1}{2\pi x}\sqrt{(x-\alpha)(\beta-x)}\mathrm{d}x, \qquad \alpha \leq x \leq \beta, \tag{3.11}$$

where $\alpha = (1-\sqrt{\gamma})^2$, $\beta = (1+\sqrt{\gamma})^2$. Consequently, for any bounded continuous function f,

$$X_f^{(n)}(S_{m,n}) := \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(S_{m,n})) \to \int \varphi \mathrm{d}\mu_\gamma$$

almost surely as $n \to \infty$.

At the fluctuation level, Guionnet[18] proved that, if $S_{m,n}$ is from LUE (or LOE) and φ a polynomial function, the random variable $X_f^{(n)}(S_{m,n}) - \mathbb{E}X_f^{(n)}(S_{m,n})$ converges in distribution to a Gaussian random variable with mean 0 and variance

$$V^\beta_{\mathrm{Laguerre}}[\varphi] = \frac{1}{2\beta\pi^2}\int_{-\alpha}^{\beta}\int_{-\alpha}^{\beta}\left(\frac{\varphi(x)-\varphi(y)}{x-y}\right)^2 \frac{4\gamma-(x-\delta)(y-\delta)}{\sqrt{4\gamma-(x-\delta)^2}\sqrt{4\gamma-(y-\delta)^2}}\mathrm{d}x\mathrm{d}y,$$

where $\beta = 1$ if the matrix is from the LOE and $\beta = 2$ if it is from the LUE, and $\delta = (\alpha+\beta)/2 = 1 + \gamma$. Cabanal-Duvillard[6] derived this result using different techniques. Recently, Lytova, and Pastur[23] proved that this result is true for any continuous test function φ with a bounded derivative.

The main purpose of this section is to study the asymptotic behavior of $X_{f_n}^{(n)}(S_{m,n})$ with $f_n(x) = f(x)I(x \geq \theta_n)$ where θ_n is close to the edge β. Denote

$$X_{f_n}^{(n)}(S_{m,n}) := \sum_{i=1}^{n} f[\lambda_i(S_{m,n})]I(\lambda_i(S_{m,n}) \geq \theta_n), \tag{3.12}$$

i.e., the sum of eigenvalues that are bigger than θ_n, where f is real-valued measurable function. Let $\{\theta_n, n \geq 1\}$ be a sequence of constants such that

$$\beta_{m,n} - \theta_n \to 0^+ \qquad \text{and} \qquad n(\beta_{m,n}-\theta_n)^{3/2} \to \infty \tag{3.13}$$

as $n \to \infty$, where

$$\alpha_{n,m} = \left(\sqrt{\frac{m}{n}} - 1\right)^2 \qquad \text{and} \qquad \beta_{n,m} = \left(\sqrt{\frac{m}{n}} + 1\right)^2. \tag{3.14}$$

For convenience, we set

$$\mu_{m,n} = \frac{\sqrt{\beta_{m,n} - \alpha_{m,n}}}{3\pi\beta_{m,n}} n(\beta_{m,n} - \theta_n)^{3/2} \quad \text{and} \quad \sigma_{m,n} = \sqrt{\frac{1}{2\pi^2} \log\left[n(\beta_{m,n} - \theta_n)^{3/2}\right]}. \tag{3.15}$$

Throughout this subsection, we also assume that the elements of X (recall that $S_{m,n} = X^*X/n$) satisfy condition $\mathbf{C_0}$ defined in Subsection III A. Our main result of this subsection is the following theorem.

Theorem 3.2 (Sample covariance matrix). *Assume f is bounded, differentiable, and has bounded derivative in a small neighborhood of* β, $f(\beta_{m,n}) \neq 0$ *and*

$$\frac{n^2(\beta_{m,n} - \theta_n)^5}{\log[n(\beta_{m,n} - \theta_n)^{3/2}]} \to 0 \tag{3.16}$$

as $n \to \infty$, *where* $\theta_n, \beta_{m,n}$ *are defined in (3.13) and (3.14), respectively. Denote*

$$\xi_n(S_{m,n}) := \frac{X_{f_n}^{(n)}(S_{m,n}) - f(\beta_{m,n})\mu_{m,n}}{f(\beta_{m,n})\sigma_{m,n}}. \tag{3.17}$$

Then the conclusions of Theorem 1.1 hold for $\xi_n(S_{m,n})$, *where* $S_{m,n}$ *is drawn from LUE.*

Moreover, if $S_{m,n}$ *is a sample covariance matrix whose off-diagonal entries match up to order 4 and diagonal entries match up to order 2, to the corresponding entries of an LUE matrix, the conclusion of Theorem 1.1 (1) also hold for* $\xi_n(S_{m,n})$.

Proof. The proof here is similar to the Wigner case. We only need to check that $\mathrm{Var}X_{f_n}^{(n)}(S_{m,n}) \to \infty$ and provide explicit expressions for $\mathbb{E}X_{f_n}^{(n)}(S_{m,n})$ and $\mathrm{Var}X_{f_n}^{(n)}(S_{m,n})$. This is furnished by Lemma 3.3 and Lemma 3.4.

The proof of Berry-Esseen bound for general sample covariance matrix is same as for Wigner matrix. We omit the parallel discussions here for convenience. See Appendix A for an illustration of the Wigner matrix case. □

In the remaining part of this section, we will compute the mean and variance of $X_{f_n}(S_{m,n})$. Denote the number of eigenvalues that fall into the interval I by $N_I(S_{m,n})$. We first recall the result about $N_I(S_{m,n})$ proved by Su.[31]

Lemma 3.2. Let $S_{m,n}$ *be an LUE matrix.*

- *The expected number of eigenvalues of* $S_{m,n}$ *in the interval* $I_n = [\theta_n, +\infty)$, *when* $\beta_{m,n} - \theta_n \to 0^+$ *and* $n(\beta_{m,n} - \theta_n)^{3/2} \geq C$ *for some* $C > 0$, *is given by*

$$\mathbb{E}[N_{I_n}(S_{m,n})] = \mu_{m,n}(1 + o(1)). \tag{3.18}$$

- *Assume that* θ_n *satisfies* $\beta_{m,n} - \theta_n \to 0^+$ *and* $n(\beta_{m,n} - \theta_n)^{3/2} \geq C$ *for some* $C > 0$. *Then the variance of the number of eigenvalues of* $S_{m,n}$ *in* $I_n = [\theta_n, +\infty)$ *satisfies*

$$\mathrm{Var}(N_{I_n}(S_{m,n})) = \sigma_{m,n}^2(1 + o(1)). \tag{3.19}$$

- *Moreover, assume that* θ_n *satisfies* $n(\beta_{m,n} - \theta_n)^{3/2} \to \infty$ *when* $n \to \infty$. *Then, as n goes to infinity,*

$$\frac{N_{I_n}(S_{m,n}) - \mu_{m,n}}{\sigma_{m,n}} \longrightarrow_d N(0, 1). \tag{3.20}$$

Next, we focus on the mean and variance of $X_{f_n}^{(n)}$. The proofs here are similar to that of Pan, Wang, and Zhou.[26]

Lemma 3.3.

$$\mathbb{E}X_{f_n}^{(n)} = f(\beta_{m,n})\mu_{m,n} + \max\left\{O(1)n(\beta_{m,n} - \theta_n)^{5/2}, O(1)\right\}. \tag{3.21}$$

Proof. By (2.2), for any fixed $\delta > 0$, we can write

$$\begin{aligned}\mathbb{E}X_{f_n}^{(n)} &= \int_{\theta_n}^{\beta_{m,n}} f(u)K_n(u,u)\mathrm{d}u + \int_{\beta_{m,n}}^{\beta_{m,n}+\delta} f(u)K_n(u,u)\mathrm{d}u \\ &\quad + \int_{\beta_{m,n}+\delta}^{\infty} f(u)K_n(u,u)\mathrm{d}u \\ &:= m_1 + e_1 + e_2. \end{aligned} \tag{3.22}$$

Since $K_n(x,x)$ is nonnegative, we have for some $\xi_n \in (\theta_n, \beta_{m,n})$ by applying the mean value theorem in the integral form

$$m_1 = f(\xi_n)\int_{\theta_n}^{\beta_{m,n}} K_n(u,u)\mathrm{d}u = \{f(\beta_{m,n}) + O(1)(\xi_n - \beta_{m,n})\}\int_{\theta_n}^{\beta_{m,n}} K_n(u,u)\mathrm{d}u,$$

where in the second equality, we have used the Taylor expansion of $f(\xi_n)$ at point $\beta_{m,n}$, and $O(1)$ is a constant independent of n.

By the expectation of eigenvalue counting function of LUE in the interval $(\theta_n,\beta_{m,n})$ [see (B1) in Appendix B], we derive that

$$m_1 = f(\beta_{m,n})\frac{\sqrt{\beta_{m,n}-\alpha_{m,n}}}{3\pi\beta_{m,n}}n(\beta_{m,n}-\theta_n)^{3/2} + \max\left\{O(1)n(\beta_{m,n}-\theta_n)^{5/2}, O(1)\right\}. \tag{3.23}$$

For the remainder terms in (3.22), by the fact that the Airy function and its derivatives decay to zero with an exponentially speed outside $(\theta_n,\beta_{m,n})$, we can easily conclude that $e_1 = O(1)$ and $e_2 = o(1)$ (see Appendix B for details). Combining these estimates together, we establish the desired mean estimation (3.21). □

Lemma 3.4.

$$\mathrm{Var}X_{f_n}^{(n)} = (1+o(1))\frac{f^2(\beta_{m,n})}{2\pi^2}\log\left[n(\beta_{m,n}-\theta_n)^{3/2}\right]. \tag{3.24}$$

Proof. By (2.3),

$$\mathrm{Var}X_{f_n}^{(n)} = \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}}(f_n(x)-f_n(y))^2K_n^2(x,y)\mathrm{d}x\mathrm{d}y. \tag{3.25}$$

Consider the region Ω_1 defined as

$$\Omega_1 = \left\{(x,y)\middle|(x,y)\in[\theta_n,\theta_n+\frac{\beta_{m,n}-\theta_n}{r(n)}]\times[\theta_n-\frac{\beta_{m,n}-\theta_n}{r(n)},\theta_n-\frac{1}{n(\beta_{m,n}-\theta_n)^{1/2}}]\right\},$$

where

$$\frac{1}{r(n)} = \max\left\{\sqrt{\beta_{m,n}-\theta_n},\quad \frac{1}{\log[n(\beta_{m,n}-\theta_n)^{3/2}]}\right\}.$$

Note that from Su [Ref. 31, p. 1318, Eq.(3.68)],

$$\begin{aligned}\mathrm{Var}(N_{I_n}(S_{m,n})) &= \iint_{\Omega_1} K_n^2(x,y)\mathrm{d}x\mathrm{d}y + O(\log(r(n))) \\ &= \frac{1}{2\pi^2}\log[n(\beta_{m,n}-\theta_n)^{3/2}] + O(\log(r(n))) \\ &= (1+o(1))\frac{1}{2\pi^2}\log[n(\beta_{m,n}-\theta_n)^{3/2}]. \end{aligned} \tag{3.26}$$

Note that if $(x,y)\in\Omega_1$, $f_n(y)\equiv 0$; if $(y,x)\in\Omega_1$, $f_n(x)\equiv 0$. Denote $\Omega = \Omega_1 \cup \Omega_2$, where $\Omega_2 := \{(x,y)|(y,x)\in\Omega_1\}$. By definition, $\Omega_i, i = 1,2$ are disjoint, thus

$$\begin{aligned}\mathrm{Var}X_{f_n}^{(n)} &= \frac{1}{2}\left(\iint_{\Omega} + \iint_{\mathbb{R}^2\setminus\Omega}\right)(f_n(x)-f_n(y))^2K_n^2(x,y)\mathrm{d}x\mathrm{d}y \\ &= \frac{1}{2}\iint_{\Omega}(f_n(x)-f_n(y))^2K_n^2(x,y)\mathrm{d}x\mathrm{d}y + O(1)\iint_{\mathbb{R}^2\setminus\Omega}K_n^2(x,y)\mathrm{d}x\mathrm{d}y \\ &= \iint_{\Omega_1}(f_n(x)-f_n(y))^2K_n^2(x,y)\mathrm{d}x\mathrm{d}y + O(\log(r(n))), \end{aligned} \tag{3.27}$$

where we have used the fact that K_n is symmetric in the last equality. Observe that

$$f_n(x) = f(\beta_{m,n}) + (f(x) - f(\beta_{m,n}))I(x \ge \theta_n)$$
$$= f(\beta_{m,n}) + O(1)(x - \beta_{m,n})I(x \ge \theta_n) = f(\beta_{m,n}) + O(\beta_{m,n} - \theta_n).$$

Combining this and (3.26) gives that

$$\frac{1}{2}\iint_{\Omega_1} (f_n(x) - f_n(y))^2 K_n^2(x,y)\mathrm{d}x\mathrm{d}y = (1 + o(1))\frac{f^2(\beta_{m,n})}{2\pi^2}\log[n(\beta_{m,n} - \theta_n)^{3/2}].$$

Substituting the above into (3.27) gives (3.24). □

ACKNOWLEDGMENTS

We would like to thank a referee for very careful reading and useful suggestions which improved the presentation and quality of the paper.

G. M. Pan was partially supported by MOE Tier 2 Grant No. 2014-T2-2-060 and by a MOE Tier 1 Grant No. RG25/14 at the Nanyang Technological University, Singapore; S. C. Wang was supported by the Project Funded by China Postdoctoral Science Foundation (Grant No. 2015M580713), the Fundamental Research Funds for the Central Universities (Grant No. 2017BQ108), and Tianyuan Foundation of Mathematics of NSFC (Grant No. 11626119); W. Zhou was partially supported by a Ministry of Education Grant No. MOE2015-T2-2-039 (Grant No. R-155-000-171-112) at the National University of Singapore.

APPENDIX A: PROOF OF THEOREM 3.1 FOR WIGNER MATRICES

Denote $\Phi(\cdot)$ as the cumulative distribution function of standard normal distribution throughout this section.

Let W_n be drawn from the Wigner matrix which satisfies the conditions in Theorem 3.1. Suppose W_n' is a GUE. Then observe that

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}(\xi_n(W_n) \le x) - \Phi(x)\right| \le \sup_{x\in\mathbb{R}}\left|\mathbb{P}(\xi_n(W_n') \le x) - \Phi(x)\right| + \sup_{x\in\mathbb{R}}\left|\mathbb{P}(\xi_n(W_n) \le x) - \mathbb{P}(\xi_n(W_n') \le x)\right|. \tag{A1}$$

As in the work of Pan, Wang, and Zhou[26] and Chen, Gao, and Wang,[7] we can show that $|F_{\xi_n(W_n)}(x) - F_{\xi_n(W_n')}(x)| = O(n^{-c_0})$ holds uniformly for all $x \in \mathbb{R}$ and some sufficiently small $c_0 > 0$. We remark that for proving this, we need to use the strong localization results of the eigenvalues of Erdös, Yau, and Yin,[16] the four moment theorem[32,33] of Tao and Vu, and an approximation argument, see the proofs in Refs. 7 and 26 for more details or one can also refer to the work of Döring and Eichelsbacher[14] for a similar discussion. With this fact, the Berry-Esseen bound follows immediately from (A1), the Berry-Esseen bound in GUE case (see Theorems 3.1 and 1.1), and the fact that $\sqrt{\log[n(2-\theta_n)^{3/2}]}$ is a much smaller term compared to n^{c_0}.

For the Cramér type MDP for $\xi_n(W_n)$, we only prove the first assertion in (2.10) for Wigner matrix. The second one is similar and omitted here. We observe that

$$\frac{1 - F_{\xi_n(W_n)}(x)}{1 - \Phi(x)} = \frac{1 - F_{\xi_n(W_n')}(x)}{1 - \Phi(x)} + \frac{F_{\xi_n(W_n')}(x) - F_{\xi_n(W_n)}(x)}{1 - \Phi(x)}. \tag{A2}$$

By the Cramér type MDP in the GUE case (see Theorems 3.1 and 1.1), we know that

$$\frac{1 - F_{\xi_n(W_n')}(x)}{1 - \Phi(x)} = 1 + O(1)\frac{1 + x^3}{\sqrt{\log[n(2-\theta_n)^{3/2}]}} \tag{A3}$$

holds uniformly for $x \in [0, \rho(\log[n(2-\theta_n)^{3/2}])^{1/6}]$, where ρ is any fixed constant. By virtue of inequality

$$1 - \Phi(x) \ge \frac{1}{\sqrt{2\pi}}\frac{x}{1 + x^2}e^{-x^2/2}, \qquad x \ge 0,$$

we can easily derive that

$$\frac{|F_{\xi_n(W_n)}(x)-F_{\xi_n(W_n')}(x)|}{1-\Phi(x)}=\frac{1}{O(n^{c_0})(1-\Phi(x))}=o(1)\frac{1+x^3}{\sqrt{\log[n(2-\theta_n)^{3/2}]}}$$

holds uniformly for $x\in[0,\rho(\log[n(2-\theta_n)^{3/2}])^{1/6}]$. Substituting this into (A2) and combining (A3), we derive that

$$\frac{1-F_{\xi_n(W_n)}(x)}{1-\Phi(x)}=\frac{o(1)}{\sqrt{\log[n(2-\theta_n)^{3/2}]}}+1+O(1)\frac{1+x^3}{\sqrt{\log[n(2-\theta_n)^{3/2}]}}.$$

We finally get that

$$\frac{1-F_{\xi_n(W_n)}(x)}{1-\Phi(x)}=1+O(1)\frac{1+x^3}{\sqrt{\log[n(2-\theta_n)^{3/2}]}}$$

holds uniformly for $x\in[0,\rho(\log[n(2-\theta_n)^{3/2}])^{1/6})$, which ends the proof.

APPENDIX B: A VERIFICATION OF EQ. (3.22)

In this section, for any fixed $\delta>0$, we verify that

$$\int_{\theta_n}^{\beta_{m,n}}K_n(u,u)\mathrm{d}u=\frac{\sqrt{\beta_{m,n}-\alpha_{m,n}}}{3\pi\beta_{m,n}}n(\beta_{m,n}-\theta_n)^{3/2}+\max\left\{O(1)n(\beta_{m,n}-\theta_n)^{5/2},O(1)\right\} \tag{B1}$$

and

$$\int_{\beta_{m,n}}^{\beta_{m,n}+\delta}K_n(u,u)\mathrm{d}u=O(1). \tag{B2}$$

First observe that for any $y\in(\beta_{m,n}-\delta,\beta_{m,n}+\delta)$ (cf. Su,[31] Eq. (3.4), pp. 1306–1307]

$$\begin{aligned}K_n(y,y)=&\left(\frac{\Psi'(y)}{4\Psi(y)}-\frac{\gamma'(y)}{\gamma(y)}\right)(2Ai(\Psi(y))Ai'(\Phi(y)))\\&+\Psi'(y)\left((Ai'(\Psi(y)))^2-\Psi(y)(Ai(\Phi(y)))^2\right)\\&+O\left(\frac{1}{n^2}\right)\left(\gamma^2(y)\Psi^{-1/2}(y)Ai'(\Psi(y))^2+\gamma^{-2}(y)\Psi^{1/2}(y)Ai(\Psi(y))^2\right)\\:=&I_1(y)+I_2(y)+I_3(y)\end{aligned} \tag{B3}$$

with

$$\Psi(y)=\begin{cases}-\left(\frac{3\pi}{2}n\int_x^{\beta_{m,n}}\mu_{m,n}(v)\,\mathrm{d}v\right)^{2/3}, & y\le\beta_{m,n};\\ \left(\frac{3\pi}{2}n\int_{\beta_{m,n}}^{x}\mu_{m,n}(v)\,\mathrm{d}v\right)^{2/3}, & y>\beta_{m,n},\end{cases}\qquad \gamma(y)=\left(\frac{y-\beta_{m,n}}{y-\alpha_{m,n}}\right)^{1/4},$$

where

$$\mu_{m,n}=\frac{1}{2\pi x}\sqrt{(x-\alpha_{m,n})(\beta_{m,n}-x)},\quad \alpha_{m,n}\le x\le\beta_{m,n}$$

[in the appendix $\mu_{m,n}$ is different from (3.15)] and $Ai(y)$ is the Airy function which has the following asymptotic expansions:

$$\begin{aligned}Ai(-r)&=\pi^{-1/2}r^{-1/4}\left\{\cos\left(\frac{2}{3}r^{2/3}-\frac{\pi}{4}\right)+O\left(\frac{1}{r^{3/2}}\right)\right\},\quad r>0,\\Ai'(-r)&=\frac{r^{1/4}}{\sqrt{\pi}}\left\{\sin\left(\frac{2}{3}r^{3/2}-\frac{\pi}{4}\right)+O\left(\frac{1}{r^{3/2}}\right)\right\},\quad r>0,\\Ai(r)&=2^{-1}\pi^{-1/2}r^{-1/4}e^{-\frac{2}{3}r^{3/2}}\left\{1+O\left(\frac{1}{r^{3/2}}\right)\right\},\quad r>0,\\Ai'(r)&=2^{-1}\frac{r^{1/4}}{\sqrt{\pi}}e^{-\frac{2}{3}r^{3/2}}\left\{1+O\left(\frac{1}{r^{3/2}}\right)\right\},\quad r>0.\end{aligned} \tag{B4}$$

Now using these facts, one can check that (use $Ai(0)Ai'(0) = O(1)$, cf. Ref. 10)

$$|Ai(y)Ai'(y)| = O(1), \qquad \frac{\Psi'(y)}{4\Psi(y)} - \frac{\gamma'(y)}{\gamma(y)} = O(1)$$

hold for all $y \in \mathbb{R}$. Consequently, we have

$$\int_{\theta_n}^{\beta_{m,n}+\delta} I_1(y)\mathrm{d}y = O(1) \tag{B5}$$

and

$$\int_{\theta_n}^{\beta_{m,n}+\delta} I_3(y)\mathrm{d}y = o(1). \tag{B6}$$

Notice that

$$\begin{aligned}\int_{\theta_n}^{\beta_{m,n}+\delta} I_2(y)\mathrm{d}y &= \int_{\theta_n}^{\beta_{m,n}} I_2(y)\mathrm{d}y + \int_{\beta_{m,n}}^{\beta_{m,n}+\delta} I_2(y)\mathrm{d}y \\ &= -\left[\frac{2}{3}(y^2 Ai'(y)^2 - yAi(y)^2) - \frac{1}{3}Ai(y)Ai'(y)\right]\Big|_{\Psi(\theta_n)}^{\Psi(\beta_{m,n})} \\ &\quad - \left[\frac{2}{3}(y^2 Ai'(y)^2 - yAi(y)^2) - \frac{1}{3}Ai(y)Ai'(y)\right]\Big|_{\Psi(\beta_{m,n})}^{\Psi(\beta_{m,n}+\delta)}.\end{aligned}$$

By the definition of Ψ, we know $\Psi(\beta_{m,n}) = 0$. Thus using the asymptotic expansions (B3) above, we get that

$$\begin{aligned}&\int_{\theta_n}^{\beta_{m,n}} I_2(y)\mathrm{d}y \\ &\quad = \frac{2}{3}\left(\Psi^2(t)(Ai(\Psi(t)))^2 - \Psi(t)(Ai'(\Psi'(t)))^2\right) - \frac{1}{3}Ai(\Psi(t))Ai'(\Psi(t))\Big|_{t=\Psi(\theta_n)} + O(1)\end{aligned}$$

and

$$\int_{\beta_{m,n}}^{\beta_{m,n}+\delta} I_2(y)\mathrm{d}y = O(1). \tag{B7}$$

Observe that (write $v = \beta_{m,n} - u$ in the second equality)

$$\begin{aligned}\Psi(\theta_n) &= -\left(\frac{3\pi}{2}n\int_{\theta_n}^{\beta_{m,n}} \frac{1}{2\pi v}\sqrt{(v-\alpha_{m,n})(\beta_{m,n}-v)}\,\mathrm{d}v\right)^{2/3} \\ &= -\left(\frac{3\pi}{2}n\int_0^{\beta_{m,n}-\theta_n} \frac{1}{2\pi(\beta_{m,n}-u)}\sqrt{u(\beta_{m,n}-\alpha_{m,n}-u)}\,\mathrm{d}u\right)^{2/3} \\ &= -\left(\frac{3\pi}{2}n\left\{\frac{\sqrt{\beta_{m,n}-\alpha_{m,n}}}{2\pi\beta_{m,n}}\int_0^{\beta_{m,n}-\theta_n} \sqrt{u} - O(1)u^{3/2})\,\mathrm{d}u + o(1)\right\}\right)^{2/3},\end{aligned}$$

where we have used the fact that

$$\sqrt{u(l-u)} = \sqrt{lu}(1-u/l)^{1/2} = \sqrt{lu} - \frac{u^{3/2}}{2\sqrt{l}} + o(u^{3/2})$$

for any $l > 0$ and small $u > 0$ in the last equality. This shows that $\Psi(\theta_n)$ tends to $-\infty$ as $n \to \infty$. Using the asymptotic expansions (B3) for Airy function above, we have

$$\begin{aligned}\int_{\theta_n}^{\beta_{m,n}} I_2(y)\mathrm{d}y &= \frac{1}{\pi}|\Psi(\theta_n)|^{3/2} + O(1) \\ &= \frac{\sqrt{\beta_{m,n}-\alpha_{m,n}}}{3\pi\beta_{m,n}}n\left\{(\beta_{m,n}-\theta_n)^{3/2} + O(1)(\beta_{m,n}-\theta_n)^{5/2}\right\} + O(1) \\ &= \frac{\sqrt{\beta_{m,n}-\alpha_{m,n}}}{3\pi\beta_{m,n}}n(\beta_{m,n}-\theta_n)^{3/2} + \max\left\{O(1)n(\beta_{m,n}-\theta_n)^{5/2}, O(1)\right\}.\end{aligned} \tag{B8}$$

Now substituting estimates (B5)–(B8) into (B3) completes the proof of (B1) and (B2).

[1] Anderson, G., Guionnet, A., and Zeitouni, O., *An Introduction to Random Matrices*, Volume 118 of Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, 2010).
[2] Bai, Z. D. and Silverstein, J., *Spectral Analysis of Large Dimensional Random Matrices* (Science Press, 2006).
[3] Bai, Z. D., Wang, X. Y., and Zhou, W., "CLT for linear spectral statitics of Wigner matrices," Electron. J. Probab. **14**(83), 2391–2417 (2009).
[4] Bao, Z. G., Pan, G. M., and Zhou, W., "CLT for partial linear eigenvalue statistics of Wigner matrices," J. Stat. Probab. **150**(1), 88–129 (2013).
[5] Breuer, J. and Duits, M., "The Nevai condition and a local law of large numbers for orthogonal polynomial ensembles," Adv. Math. **265**, 441–484 (2014).
[6] Cabanal-Duvillard, T., "Fluctuations de la loi empirique de grandes matrices aléatoires," Ann. Inst. Henri Poincaré Probab. Stat. **37**, 373–402 (2001).
[7] Chen, L., Gao, F. Q., and Wang, S. C., "Berry-Esseen bounds and Cramér type large deviations for eigenvalues of random matrices," Sci. China Math. **58**(9), 1959–1980 (2015).
[8] Costin, O. and Lebowitz, J. L., "Gaussian fluctuations in random matrices," Phys. Rev. Lett. **75**, 69–72 (1995).
[9] Dallaporta, S. and Vu, V., "A note on the central limit theorem for the eigenvalue counting function of Wigner matrices," Electron. Commun. Probab. **16**(30), 214–322 (2011).
[10] Deift, P., *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*, Courant Lecture Notes in Mathematics 18 (American Mathematical Society, Providence, RI, 2009).
[11] Dembo, A. and Zeitouni, O., *Large Deviations Techniques and Applications*, Volume 38 (Springer Science & Business Media, 2009).
[12] Diaconis, P., "Patterns in eigenvalues: The 70th Josiah Willard Gibbs lecture," Bull. Am. Math. Soc. **40**(2), 155–178 (2003).
[13] Döring, H. and Eichelsbacher, P., "Moderate deviations for the eigenvalue counting function of Wigner matrices," ALEA, Lat. Am. J. Probab. Math. Stat. **10**(1), 27–44 (2013).
[14] Döring, H. and Eichelsbacher, P., "Edge fluctuations of eigenvalues of Wigner matrices," in *High Dimensional Probability VI*, Volume 66 of Progress in Probability (Birkhäuser/Springer, Basel, 2013), pp. 261–275.
[15] Döring, H. and Eichelsbacher, P., "Moderate deviations via cumulants," J. Theor. Probab. **26**(2), 360–385 (2013).
[16] Erdös, L., Yau, H.-T., and Yin, J., "Rigidity of eigenvalues of generalized Wigner matrices," Adv. Math. **229**(3), 1435–1515 (2012).
[17] Fan, X. Q., Grama, I., and Liu, Q. S., "Cramér large deviation expansions for martingales under Bernstein's condition," Stochastic Processes Appl. **123**, 3919–3942 (2013).
[18] Guionnet, A., "Large deviations, upper bounds, and central limit theorems for non-commutative functionals of Gaussian large random matrices," Ann. Inst. Henri Poincaré Probab. Stat. **38**, 341–384 (2002).
[19] Gustavsson, J., "Gaussian fluctuations of eigenvalues in the GUE," Ann. Inst. H. Poincar. Probab. Statist. **41**(2), 151–178 (2005).
[20] Hough, J. B., Krishnapur, M., Peres, Y., and Virág, B., "Determinantal processes and independence," Probab. Surv. **3**, 206–229 (2006).
[21] König, W., "Orthogonal polynomial ensembles in probability theory," Probab. Surv. **2**, 385–447 (2005).
[22] Kopel, P., "Regularity conditions for convergence of linear statistics of GUE," e-print arXiv:1510.02988 (2015).
[23] Lytova, A. and Pastur, L., "Central limit theorem for linear eigenvalue statistics of random matrices with independent entries," Ann. Probab. **37**(5), 1778–1840 (2009).
[24] Marčenko, V. A. and Pastur, L. A., "The distribution of eigenvalues in certain set of random matrices," Math. Sb. **72**, 507–536 (1967).
[25] Pastur, L., "Limiting laws of linear eigenvalue statistics for Hermitian matrix models," J. Math. Phys. **47**(10), 103303 (2006).
[26] Pan, G. M., Wang, S. C., and Zhou, W., "Fluctuations of linear eigenvalues statistics for Wigner matrices: Edge case," J. Stat. Phys. **165**, 507–520 (2016).
[27] Saulis, L. and Statulevicius, V. A., *Limit Theorems for Large Deviations* (Springer, New York, 1991).
[28] Shcherbina, M., "Central limit theorem for linear eigenvalue statistcs of the Wigner and sample covariance random matrices," J. Math. Phys. Anal. Geom. **7**(2), 176–192 (2011).
[29] Soshnikov, A., "Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices," Zh. Mat. Fiz. Anal. Geom. **7**(2), 176–192 (2011).
[30] Sosoe, P. and Wong, P., "Regularity conditions in the CLT for linear eigenvalue statistics of Wigner matrices," Adv. Math. **249**, 37–87 (2013).
[31] Su, Z., "Gaussian fluctuations in complex sample covariance matrices," Electron. J. Probab. **11**, 1284–1320 (2006).
[32] Tao, T. and Vu, V., "Random matrices: Universality of local eigenvalue statistics," Acta Math. **206**(1), 127–204 (2011).
[33] Tao, T. and Vu, V., "Random matrices: Universality of local eigenvalue statistics up to the edge," Commun. Math. Phys. **298**(2), 549–572 (2010).