NONCOMMUTATIVE HARMONIC ANALYSIS
WITH APPLICATIONS TO PROBABILITY III
BANACH CENTER PUBLICATIONS, VOLUME 96
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 2012

SPECTRAL PROPERTIES OF LARGE RANDOM MATRICES WITH INDEPENDENT ENTRIES

P. DUECK, S. O'ROURKE, D. RENFREW and A. SOSHNIKOV

Department of Mathematics, University of California, Davis
One Shields Avenue, Davis, CA 95616-8633, USA

Abstract. We consider large Wigner random matrices and related ensembles of real symmetric and Hermitian random matrices. Our results are related to the local spectral properties of these ensembles.

1. Introduction. Wigner random matrices were introduced by E. Wigner in the 1950s ([41], see also [2, 1]). Let $\{X_{i,j}\}_{1\le i<j}$ be a family of independent, identically distributed, centered, real (or complex)-valued random variables independent from a family of $\{Y_j\}_{j\ge 1}$ independent, identically distributed, real-valued random variables. An $n\times n$ matrix W_n is defined as

$$W_n(i,j) = \overline{W_n}(j,i) =: w_{i,j} = \begin{cases} X_{i,j} & \text{if } \ i<j, \\ Y_i & \text{if } \ i=j. \end{cases} \tag{1}$$

We assume that $\mathbb{E}|X_{1,2}|^2 = \sigma^2 < \infty$. The matrix W_n is called a real symmetric (Hermitian in the complex case) Wigner random matrix. The Euclidean norm of any fixed column of W_n is proportional to $\sqrt{n}$. Therefore, it is natural to conjecture that typical eigenvalues of W_n are of order of $\sqrt{n}$. We define

$$M_n = \frac{1}{2\sigma\sqrt{n}}\, W_n. \tag{2}$$

The main result about the global distribution of the eigenvalues of M_n goes back to Wigner and is known as the Wigner Semicircle Law ([41, 2, 1]). To formulate this result, we first define the distribution function of the Wigner Semicircle Law

$$F(t) = \begin{cases} 1 & \text{if } \ t>1, \\ \frac{2}{\pi}\int_{-1}^{t}\sqrt{1-x^2}\,\mathrm{dx} & \text{if } \ -1\le t\le 1, \\ 0 & \text{if } \ -\infty<t<-1. \end{cases} \tag{3}$$

2010 *Mathematics Subject Classification*: 60B20.

Key words and phrases: Wigner random matrices, outliers in the spectrum.
The paper is in final form and no version of it will be published elsewhere.

DOI: 10.4064/bc96-0-7

Let us denote by $x_1 \le x_2 \le \ldots \le x_n$ the (ordered) eigenvalues of M_n defined in (2). We denote their empirical distribution function by F_n. In other words,

$$F_n(x) = \frac{1}{n}\,\#\{1 \le i \le n : \lambda_i \le x\}. \tag{4}$$

The Wigner Semicircle Law states that under the above conditions on the distribution of the matrix entries, the empirical distribution function $F_n(x)$ converges almost surely to $F(x)$ for all values of x. The immediate corollary of the Wigner Semicircle Law is

THEOREM 1. *Let $x_1 \le \ldots \le x_n$ denote the ordered eigenvalues of an $n \times n$ Wigner random matrix W_n defined in* (1)*. If $\frac{k}{n} \to \gamma \in (0,1)$, then $\frac{x_k}{2\sigma\sqrt{n}} \to F^{-1}(\gamma)$ as $n \to \infty$ a.s. where $F(t)$ is defined in* (3)*.*

The archetypal examples of Wigner random matrices are the Gaussian Unitary Ensemble (GUE) of Hermitian random matrices and the Gaussian Orthogonal Ensemble (GOE) of real symmetric random matrices. The GUE ensemble is defined as

$$A = \frac{1}{2}\,(B + B^*), \tag{5}$$

where the entries of B are i.i.d. complex Gaussian random variables, so that $\operatorname{Re} b_{j,k}$ and $\operatorname{Im} b_{j,k}$ are independent from each other and have $N(0, \sigma^2)$ distribution.

In a similar fashion, the GOE ensemble is defined as

$$A = \frac{1}{2}\,(B + B^t), \tag{6}$$

where the entries of B are i.i.d. $N(0, 2\sigma^2)$ random variables. Thus, A is a real symmetric random matrix with independent $N(0, (1 + \delta_{i,j})\sigma^2)$-distributed entries for $1 \le i \le j \le n$.

The joint distribution of the matrix entries in the GOE/GUE ensembles is given by the formula

$$\mathbb{P}(\,\mathrm{d}A) = C_n^{(\beta)} \exp\Big(-\frac{\beta}{4\sigma^2}\,\mathrm{Tr}(A^2)\Big)\,\mathrm{d}A, \tag{7}$$

where $\beta = 1$ for GOE, $\beta = 2$ for GUE, and $\mathrm{d}A$ is the Lebesgue measure on the space of $n \times n$ real-symmetric (Hermitian) matrices.

The other special value of β in (7), $\beta = 4$, corresponds to a so-called Gaussian Symplectic Ensemble (GSE) of $n \times n$ quaternion self-dual Hermitian matrices. We refer the reader to [25] for the details.

There are explicit formulas for the k-point correlation functions of eigenvalues in the Gaussian ensembles (see e.g. [25, 1]). In particular, the k-point correlation function in the GUE ensemble are determinantal and the k-point correlation functions in the GOE and GSE ensembles are Pfaffian. These formulas greatly simplify the analysis of the local spectral properties of Gaussian ensembles.

In Section 2, we will study the fluctuation of the k-th eigenvalue of a Wigner random matrix about the appropriate quantile of the Wigner Semicircle Law provided $k, n - k \to \infty$ as $n \to \infty$. The first result in this direction is due to J. Gustavsson ([19]) who studied the GUE case. Later, Gustavsson's results were extended to a sufficiently large class of Wigner Hermitian random matrices by T. Tao and V. Vu ([37]). In Section 2, we will discuss the extension of the Gustavsson, Tao–Vu results to the Wigner real symmetric random matrices as well as to the Wishart Ensemble of sample-covariance random matrices and Unitary Ensembles of Hermitian random matrices.

Section 3 is devoted to finite rank perturbations of Wigner random matrices

$$M_n = \frac{1}{\sqrt{n}} W_n + A_n.$$

Here W_n is a random Wigner Hermitian matrix and A_n is a deterministic, finite rank matrix. In [8, 9], M. Capitaine, C. Donati-Martin, and D. Féral studied the distribution of the largest eigenvalues of the deformed matrix provided the marginal distribution of the matrix entries of W_n is symmetric and satisfies the Poincaré inequality. We extend the results of [8] by lifting the assumption that the marginal distribution is symmetric. In particular, the third moment is not necessarily zero.

Finally, in Sections 4 and 5, we apply the resolvent technique to study recursive relations for local linear statistics in the bulk and at the edge of the spectrum of large random matrices.

2. Gaussian fluctuations of eigenvalues in Wigner random matrices. Let $x_1 \le \ldots \le x_n$, as above denote the ordered eigenvalues of an $n \times n$ Wigner random matrix $W_n = \{w_{ij}\}_{i,j=1}^n$. Without loss of generality we can assume that $\mathrm{Var}(w_{ij}) = \frac{1}{2}$ for $1 \le i < j \le n$, so $\sigma = 1/\sqrt{2}$. We wish to study eigenvalue number $k = k(n)$, x_k, as k and $n-k$ tend to infinity with n. Let $\frac{k}{n} \to \gamma \in (0,1)$ as $n \to \infty$. Theorem 1 states that x_k converges, with probability 1, to a particular value corresponding to the quantile determined by γ. Our goal is to study how, and on what order, x_k fluctuates about that value.

To study the fluctuations of x_k, we first consider the case when W_n is drawn from the Gaussian ensembles. The result can then be extended to a more general class of Wigner matrices by applying a university result by Tao and Vu called the Four Moment Theorem (see [37] and [38]).

The result below was first proven by Gustavsson [19] in the case when W_n is drawn from the GUE. Following Gustavsson's notation, we write $k(n) \sim n^\theta$ to mean that $k(n) = h(n)n^\theta$, where h is a function such that, for all $\epsilon > 0$,

$$\frac{h(n)}{n^\epsilon} \longrightarrow 0 \quad \text{and} \quad h(n)n^\epsilon \longrightarrow \infty,$$

as $n \to \infty$.

THEOREM 2 (The bulk, [26]). *Let $x_1 < x_2 < \ldots < x_n$ be the ordered eigenvalues from a random matrix drawn from the GOE, GUE, or GSE. Consider $\{x_{k_i}\}_{i=1}^m$ such that $0 < k_i - k_{i+1} \sim n^{\theta_i}$, $0 < \theta_i \le 1$, and $\frac{k_i}{n} \to a_i \in (0,1)$ as $n \to \infty$. Define $s_i = s_i(k_i, n) = F^{-1}(k_i/n)$ and set*

$$X_i = \frac{x_{k_i} - s_i\sqrt{2n}}{\left(\frac{\log n}{2\beta(1-s_i^2)n}\right)^{1/2}}, \quad i = 1, \ldots, m,$$

where $\beta = 1, 2, 4$ corresponds to the GOE, GUE, or GSE. Then as $n \to \infty$,

$$\mathbb{P}[X_1 \le \xi_1, \ldots, X_m \le \xi_m] \longrightarrow \Phi_\Lambda(\xi_1, \ldots, \xi_m),$$

where Φ_Λ is the cdf[1] *for the m-dimensional normal distribution with covariance matrix* $\Lambda_{i,j} = 1 - \max\{\theta_k : i \le k < j < m\}$ *if* $i < j$ *and* $\Lambda_{i,i} = 1$.

THEOREM 3 (The edge, [26]). *Let* $x_1 < x_2 < \ldots < x_n$ *be the ordered eigenvalues from a random matrix drawn from the GOE, GUE, or GSE. Consider* $\{x_{n-k_i}\}_{i=1}^m$ *such that* $k_1 \sim n^\gamma$ *where* $0 < \gamma < 1$ *and* $0 < k_{i+1} - k_i \sim n^{\theta_i}$, $0 < \theta_i < \gamma$. *Set*

$$X_i = \frac{x_{n-k_i} - \sqrt{2n}\Big(1 - \big(\frac{3\pi k_i}{4\sqrt{2}n}\big)^{2/3}\Big)}{\Big(\big(\frac{1}{12\pi}\big)^{2/3} \frac{2\log k_i}{\beta n^{1/3} k_i^{2/3}}\Big)^{1/2}}, \quad i = 1, \ldots, m,$$

where $\beta = 1, 2, 4$ *corresponds to the GOE, GUE, or GSE. Then as* $n \to \infty$,

$$\mathbb{P}[X_1 \le \xi_1, \ldots, X_m \le \xi_m] \longrightarrow \Phi_\Lambda(\xi_1, \ldots, \xi_m),$$

where Φ_Λ is the cdf for the m-dimensional normal distribution with covariance matrix $\Lambda_{i,j} = 1 - \frac{1}{\gamma}\max\{\theta_k : i \le k < j < m\}$ *if* $i < j$ *and* $\Lambda_{i,i} = 1$.

REMARK 4. The GUE ($\beta = 2$) case in Theorems 2 and 3 was shown by Gustavsson in [19].

REMARK 5. In the case $m = 1$, Theorem 2 can be stated as follows. Set $t = t(k, n) = F^{-1}(k/n)$ where $k = k(n)$ is such that $k/n \to a \in (0, 1)$ as $n \to \infty$. If x_k denotes eigenvalue number k in the GOE, GUE, or GSE, then, as $n \to \infty$,

$$\frac{x_k - t\sqrt{2n}}{\big(\frac{\log n}{2\beta(1-t^2)n}\big)^{1/2}} \longrightarrow N(0,1),$$

in distribution where $\beta = 1, 2, 4$ corresponds to the GOE, GUE, or GSE.

REMARK 6. In the case $m = 1$, Theorem 3 can be stated as follows. Let k be such that $k \to \infty$ but $\frac{k}{n} \to 0$ as $n \to \infty$ and let x_{n-k} denote eigenvalue number $n - k$ in the GOE, GUE, or GSE. Then it holds that, as $n \to \infty$,

$$\frac{x_{n-k} - \sqrt{2n}\Big(1 - \big(\frac{3\pi k}{4\sqrt{2}n}\big)^{2/3}\Big)}{\Big(\big(\frac{1}{12\pi}\big)^{2/3} \frac{2\log k}{\beta n^{1/3} k^{2/3}}\Big)^{1/2}} \longrightarrow N(0,1),$$

in distribution where $\beta = 1, 2, 4$ corresponds to the GOE, GUE, or GSE.

REMARK 7. One can omit the assumption that $k_i/n \to a_i$ in Theorem 2 and the conclusion still holds. To see this, first consider the case $m = 1$. Let x_k denote a sequence of eigenvalues from the bulk with $k = k(n)$ (where k/n does not necessarily converge as $n \to \infty$). Since $k/n < 1$, there exists a subsequence, say $k' = k(n_l)$, such that $k'/n_l \to a$ as $l \to \infty$ for some $a \in (0, 1)$. By Theorem 2, the centered and scaled eigenvalues from the subsequence $x_{k'}$ converge to the standard normal distribution. It follows that every subsequence has a further subsequence which converges in distribution to the standard Gaussian distribution. Therefore, the entire sequence must converge in distribution to the standard Gaussian distribution.

[1] Cumulative distribution function

A similar argument allows one to omit the assumption that $k_i/n \to a_i$ in the case $m > 1$.

REMARK 8. It is also possible to extend Theorems 2 and 3 to other random matrix ensembles. In particular, for the complex Wishart distribution, the p non-negative eigenvalues $x_1, \ldots, x_p$ have probability density given by

$$P_p(x_1, \ldots, x_p) = C_{n,p} \prod_{1 \le i < j \le p} (x_i - x_j)^2 \prod_{i=1}^{p} x_i^{\alpha_p} e^{-x_i},$$

where $\alpha_p = n - p$ and $C_{n,p}$ is a normalizing constant. The eigenvalues of the complex Wishart distribution form a determinantal random point process and hence $P_p(x_1, \ldots, x_p)$ can be rewritten as

$$P_p(x_1, \ldots, x_p) = \frac{1}{p!} \det\big(S_p(x_i, x_j)\big)_{1 \le i,j \le p},$$

where

$$S_p(x, y) = \sum_{j=0}^{p-1} \phi_j^{(\alpha_p)}(x) \phi_j^{(\alpha_p)}(y),$$

with

$$\phi_j^{(\alpha_p)}(x) = \sqrt{\frac{j!}{(j + \alpha_p)!}}\, x^{\alpha_p/2} \exp(-x/2) L_j^{\alpha_p}(x),$$

and $L_j^{\alpha_p}$ are the generalized Laguerre polynomials.

One can then follow Gustavsson's proof for the GUE [19] in which Gustavsson uses the asymptotic expansion for the Hermite polynomials. For the complex Wishart case, the kernel $S_p(x, y)$ is given in terms of the Laguerre polynomials.

REMARK 9. Theorems 2 and 3 should also be extended to a more general class of unitary ensembles. That is, for a Hermitian $n \times n$ matrix H with probability distribution given by

$$\mathbb{P}(\,\mathrm{d}H) = C_n e^{-\operatorname{Tr} v(H)}\, \mathrm{d}H,$$

where

$$v(x) = \gamma_{2j} x^{2j} + \ldots + \gamma_0, \quad \gamma_{2j} > 0.$$

In such ensembles, the eigenvalues form a determinantal random point process where the kernel is given in terms of orthogonal polynomials with respect to the exponential weight $e^{-v(x)}$. The asymptotics of such orthogonal polynomials has been recently studied using a Riemann–Hilbert approach (see e.g. [12, 11]).

T. Tao and V. Vu extended Gustavsson's GUE results to a sufficiently large class of Hermitian Wigner matrices using the technique developed in [37] and [38] to prove the universality of the local distribution of the eigenvalues in Wigner matrices. The key ingredient of their approach is the Four Moment Theorem proved for Hermitian matrices (see Theorem 15 in [37] and Theorem 1.13 in [38]). The technical conditions imposed in [37, 38] on the distribution of matrix entries are the exponential decay of the marginal tail distribution

$$\mathbb{P}(|w_{i,j}| > t^C) \le \exp(-t), \tag{8}$$

for all $|t| > C_1$, and the requirement that the first four moments of the marginal distribution coincide with the Gaussian moments.

Extending the Four Moment Theorem to the real symmetric case, one obtains the following theorem.

Theorem 10 (Real Symmetric Wigner Matrices, [26]). *The conclusions of Theorems* 2 *and* 3 *also hold with* $\beta = 1$ *when* $x_1 \leq x_2 \leq \ldots \leq x_n$ *are the ordered eigenvalues of any other real symmetric Wigner matrix* $W_n = (w_{ij})_{1\leq i,j\leq n}$ *where* w_{ij} *has exponential decay, mean* 0 *and variance* $(1 + \delta_{ij})/2$ *for* $1 \leq i \leq j \leq n$ *and* $\mathbb{E}(w_{ij}^3) = 0$, $\mathbb{E}(w_{ij}^4) = 3/4$ *for* $1 \leq i < j \leq n$.

We now turn our attention to outlining the proof of Theorem 2. The first step, namely Theorem 1, immediately follows from the Wigner Semicircle Law and the fact that the almost sure convergence of the empirical distribution function $F_n(x)$ to the Wigner Semicircle distribution function $F(x)$ implies the almost sure convergence of the quantiles.

To prove Theorem 2, we remark that $\{x_k < t\} = \{\#([t, \infty)) < n - k\}$, where $\#(I)$ denotes the number of the eigenvalues in the interval I. Thus, one is interested in studying the asymptotic distribution of the counting random variables $\#(I)$ in the limit $n \to \infty$. We will outline the proof of Theorem 2 for the GOE in the case when $m = 1$ (see Remark 5). In the proof of the GUE case of Theorem 2, Gustavsson relies on the fact that the GUE defines a determinantal random point process. Gustavsson utilizes a theorem due to Costin, Lebowitz, and Soshnikov ([10, 22, 34]).

Theorem 11 (Costin-Lebowitz, Soshnikov). *If* $\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n)) \to \infty$ *as* $n \to \infty$, *then*

$$\frac{\#_{\mathrm{GUE}_n}(I_n) - \mathbb{E}[\#_{\mathrm{GUE}_n}(I_n)]}{\sqrt{\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n))}} \longrightarrow N(0,1),$$

in distribution as $n \to \infty$.

Remark 12. We stated the theorem here in terms of the GUE, but the result is actually more general and holds for any sequence of determinantal random point fields.

Our goal is to prove a version of Theorem 11 for the GOE and the GSE. The difficulty here is that there is no general Central Limit Theorem for counting random variables for Pfaffian random point processes. To do this, we utilize the fact that Gustavsson already proved the GUE case of Theorems 2 and 3 in [19] and we use the result due to P. Forrester and E. Rains (see [16]) that relates the eigenvalues of the different ensembles.

Theorem 13 (Forrester-Rains). *The following relations hold between matrix ensembles*:

$$\mathrm{GUE}_n = \mathrm{even}(\mathrm{GOE}_n \cup \mathrm{GOE}_{n+1})$$
$$\mathrm{GSE}_n = \mathrm{even}(\mathrm{GOE}_{2n+1}) \cdot \frac{1}{\sqrt{2}}\,.$$

Remark 14. The result by Forrester and Rains in [16] is actually much more general. Here we only consider two specific cases.

Remark 15. The multiplication by $1/\sqrt{2}$ denotes scaling the $(2n + 1) \times (2n + 1)$ GOE matrix by a factor of $1/\sqrt{2}$.

REMARK 16. The first statement can be interpreted in the following way. Take two independent matrices from the GOE: one of size $n \times n$ and one of size $(n+1) \times (n+1)$. Superimpose the eigenvalues on the real line to form a random point process with $2n+1$ particles. Then the new random point process formed by taking the n even particles has the same distribution as the eigenvalues of an $n \times n$ matrix from the GUE.

From Theorems 11 and 13 we are able to show that if $\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n)) \to \infty$ as $n \to \infty$, then

$$\frac{\#_{\mathrm{GOE}_n}(I_n) - \mathbb{E}[\#_{\mathrm{GOE}_n}(I_n)]}{\sqrt{2\,\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n))}} \longrightarrow N(0,1),$$

in distribution as $n \to \infty$.

Set

$$I_n = \left[t\sqrt{2n} + \xi\Big(\frac{\log n}{2(1-t^2)n}\Big)^{1/2}, \infty\right).$$

Then the proof in the GOE case is completed by computing $\mathbb{E}[\#_{\mathrm{GOE}_n}(I_n)]$ and $\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n))$ and noting that

$$\begin{aligned}
\mathbb{P}\left[\frac{x_k - t\sqrt{2n}}{\big(\frac{\log n}{2(1-t^2)n}\big)^{1/2}} \le \xi\right] &= \mathbb{P}\left[x_k \le t\sqrt{2n} + \xi\Big(\frac{\log n}{2(1-t^2)n}\Big)^{1/2}\right] \\
&= \mathbb{P}[\#_{GOE_n}(I_n) \le n-k] \\
&= \mathbb{P}\left[\frac{\#_{\mathrm{GOE}_n}(I_n) - \mathbb{E}[\#_{\mathrm{GOE}_n}(I_n)]}{\sqrt{2\,\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n))}} \le \frac{n-k-\mathbb{E}[\#_{\mathrm{GOE}_n}(I_n)]}{\sqrt{2\,\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n))}}\right] \\
&= \mathbb{P}\left[\frac{\#_{\mathrm{GOE}_n}(I_n) - \mathbb{E}[\#_{\mathrm{GOE}_n}(I_n)]}{\sqrt{2\,\mathrm{Var}(\#_{\mathrm{GUE}_n}(I_n))}} \le \xi + \epsilon(n)\right],
\end{aligned}$$

$x_1 < \ldots < x_n$ are the ordered eigenvalues of a GOE matrix and $\epsilon(n) \to 0$ as $n \to \infty$. We refer the reader to [26] for the details.

3. Deformed Wigner matrices. In this section, we study deformed Wigner matrices given by

$$M_n = \frac{1}{\sqrt{n}} W_n + A_n = X_n + A_n,$$

where W_n is a random Wigner Hermitian matrix satisfying some technical assumptions on the marginal distribution of matrix entries and A_n is a deterministic, finite rank Hermitian matrix.

Perturbations of classical matrix models have been studied in several different contexts. In [4], J. Baik, G. Ben Arous and S. Péché studied perturbations of Wishart matrices, called spiked population models. They consider Y_N, a $p \times N$ complex matrix whose columns are i.i.d., centered, Gaussian with covariance matrix Σ, and study the asymptotic spectrum of $S_N = \frac{1}{N} Y_N^* Y_N$. The size of Y_N grows taken to infinity in such a way that $N, p \to \infty$, $p/N \to c \ge 1$. In the classical case (known as the Wishart model) $\Sigma = I$, and the limiting behavior of the spectral measure is the Marchenko–Pastur law. We recall that the Marchenko–Pastur distribution is supported on the interval $[a,b]$ where $a = (1-c^{-1/2})^2$, $b = (1+c^{-1/2})^2$, and its density equals $\frac{c}{2\pi x}\sqrt{(b-x)(x-a)}$. The largest eigenvalue converges to the edge of the support of this distribution, with fluctuations given by the Tracy–Widom distribution ([23]).

In the perturbed model, all but finitely many of the eigenvalues of Σ are equal to one. Once an eigenvalue of Σ is large enough, a phase transition occurs and the largest eigenvalue of S_N leaves the support of the Marchenko–Pastur law. These results are extended to the case when the matrix entries are not necessarily Gaussian in [5]. J. Baik and J. Silverstein show the limiting distribution of the eigenvalues converge to the same universal limit as in the Gaussian case. Additionally, the fluctuations of the largest eigenvalues are shown to be universal in the sense that they do not depend on the distribution of the entries of Y_N.

The additive analog of the spiked population model are deformed Wigner matrices. As before, we shall denote a Wigner Hermitian matrix by W_n. We assume that the n^2 random variables $(W_n)_{ii}$, $\sqrt{2}\,\mathrm{Re}((W_n)_{ij})_{i<j}$, $\sqrt{2}\,\mathrm{Im}((W_n)_{ij})_{i<j}$ are independent and identically distributed with distribution μ. This distribution has zero expectation and variance σ^2.

Deformed Wigner matrices were first studied in [17]. Z. Füredi and J. Komlós consider real symmetric random matrices where the entries have the same non-zero mean. This can be viewed as adding a rank one perturbation to a real symmetric Wigner matrix with zero mean on the entries. They specifically consider $W_n + C$ where W_n is a real symmetric matrix with independent, identically distributed entries of mean zero, and C is a matrix with each entry equal to c. In this model the entries are not rescaled, so the largest eigenvalue is $O(n)$ and the second largest eigenvalue, given by the edge of the semi-circle, is $O(\sqrt{n})$. The fluctuations of the largest eigenvalue are Gaussian and only depend on the second moment of the entries of the random matrix.

The more difficult case when the constant matrix is scaled so that the largest eigenvalue is the same order as the edge of the semi-circle. This case is considered in [14]. In this paper, Féral and Péché show the existence of a phase transition. When the eigenvalue of the scaled constant matrix is larger than σ the fluctuations of the largest eigenvalue are Gaussian and only depend on the variance of the entries. When the eigenvalue is less than σ the fluctuations are given by the Tracy–Widom distribution and in the case when the eigenvalue equals σ the fluctuations are a generalized Tracy–Widom distribution.

Recently, more general perturbations have been considered. S. Péché [28] considered perturbations to GUE matrices of the form $\frac{1}{\sqrt{n}} W_n + A_n$, where W_n is a GUE matrix and A_n is any finite rank Hermitian perturbation. Due to the unitary invariance of the GUE, the spectrum of the deformed matrix depends only on the spectrum of A_n. Her results are extended to the general Wigner case by M. Capitaine, C. Donati-Martin and D. Féral in [8]. Both papers show that when the largest eigenvalue of A_n is sufficiently large, the largest eigenvalue of M_n leaves the support of the semi-circle and converges to the same limit, independent of the distribution of the matrix entries. In contrast to the Wishart case, the fluctuations of the largest eigenvalues are shown to depend on both the distribution of matrix entries and the form of the perturbation. In [8], the fluctuations of the largest eigenvalue are given by a convolution of the matrix entries with a Gaussian.

In [8], M. Capitaine, C. Donati-Martin and D. Féral assume the marginal distribution $\mu(\mathrm{d}x)$ of the entries of W_n is symmetric and satisfies the Poincaré inequality: there exists a positive constant C such that for any differentiable function $f : \mathbb{R} \to \mathbb{C}$ such that $\int |f|^2(x)\,\mathrm{d}\mu(x) < \infty$, $\int |f'|^2(x)\,\mathrm{d}\mu(x) < \infty$ one has

$$\mathrm{Var}(f) \leq C \int |f'|^2(x)\,\mathrm{d}\mu(x), \tag{9}$$

where $\mathrm{Var}(f) = \int |f - \mathbb{E}(f)|^2\,\mathrm{d}\mu$.

The Poincaré inequality assumption implies that all moments are finite and the tail distribution decays exponentially (see e.g. [1]). The odd moments of symmetric distributions are 0; in particular the third moment vanishes. The assumption that the third moment vanishes is quite important in the above mentioned results, as it removes the lowest order error term.

The deterministic matrix, A_n, is Hermitian and similar to a diagonal matrix with finitely many non-zero eigenvalues. The non-zero eigenvalues of A_n are denoted by $\theta_1 > \ldots > \theta_J$. The multiplicity of θ_j is denoted by k_j for $j = 1, \ldots, J$. The value of J and each k_j do not depend on n.

Hermitian matrices induce a measure on the real line, called the empirical spectral distribution (ESD), given by its eigenvalues. Given X_n, a Hermitian matrix, with eigenvalues $\lambda_1 \leq \ldots \leq \lambda_n$, the ESD is defined as $\mu_{X_n} = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i}$. We recall that for a rescaled Wigner Hermitian matrix $X_n = \frac{1}{\sqrt{n}} W_n$, the Wigner semicircle law states that the ESD converges a.s. to the semicircle, whose density is given by $\frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\,\mathbf{1}_{[-2\sigma,2\sigma]}$. Furthermore, if the fourth moment is finite the largest eigenvalue of $\frac{1}{\sqrt{n}} W_n$ converges to 2σ a.s. [2] with the fluctuations given by the Tracy–Widom distribution, assuming moment conditions on the distribution are met ([39, 40, 32, 38]).

In the deformed Wigner model, the semi-circle still holds on the global level, but the location of largest eigenvalue undergoes a phase transition when the largest eigenvalue of A_n is sufficiently large. The first result of [8] gives the location of the largest eigenvalues of M_N. Let k be the number of eigenvalues, counting repetitions, of A_n that are greater than σ. Label these eigenvalues θ_j^+, for $j = 1, \ldots, k$. Then the k largest eigenvalues of M_n converge almost surely to $\rho_j^+ = \theta_j^+ + \frac{\sigma^2}{\theta_j^+}$. The $(k+1)$-th largest eigenvalue converges to 2σ a.s. An equivalent statement is true for all eigenvalues of A_n that are less than $-\sigma$, labeled θ_j^-. This implies that all the other eigenvalues lie in the support of the semicircle. To be precise, let

$$K = \{\rho_j^-\}_j \cup [-2\sigma, 2\sigma] \cup \{\rho_j^+\}_j \quad \text{and} \quad K^\epsilon = K + \Big[-\frac{\epsilon}{2}, \frac{\epsilon}{2}\Big], \tag{10}$$

then for n large $\mathrm{Spect}(M_n) \subset K^\epsilon$ almost surely.

The results of [8] can be extended to the case of non-symmetric marginal distribution:

Theorem 17 ([30, 29]). *Let M_n be a sequence of deformed Wigner matrices with distribution of the entries that satisfies the Poincaré inequality, J_{σ^+} be the number of j's such that $\theta_j > \sigma$, and J_{σ^-} be the number of j's such that $\theta_j < -\sigma$. Then*

1. *For all $j = 1, \ldots, J_{\sigma^+}$ and $i = 1, \ldots, k_j$, $\lambda_{k_1+\ldots+k_{j-1}+i} \to \rho_j$ almost surely.*
2. *$\lambda_{k_1+\ldots+k_{J_{\sigma^+}}+1} \to 2\sigma$ almost surely.*
3. *$\lambda_{k_1+\ldots+k_{J-J_{\sigma^-}}} \to -2\sigma$ almost surely.*
4. *For all $j = J - J_{\sigma^-} + 1, \ldots, J$ and $i = 1, \ldots, k_j$, $\lambda_{k_1+\ldots+k_{j-1}+i} \to \rho_j$ almost surely.*

The convergence in probability has been established in [29]. The almost sure convergence has been proved in [30]. The next results concerns the distribution of the outliers, i.e. the eigenvalues of M_N corresponding to $\theta_j > \sigma$.

THEOREM 18 ([30]). *Let $1 \le j \le J_{\sigma+}$ (so that the eigenvalue θ_j of A_n is such that $\theta_j > \sigma$). Then the sequence of random vectors*

$$\big(\sqrt{n}\,(\lambda_{k_1+\ldots+k_{j-1}+i} - \rho_j),\ i = 1, \ldots, k_j\big)$$

is bounded in probability. In addition, the following bound holds with probability 1

$$\lambda_{k_1+\ldots+k_{j-1}+i} - \rho_j = O\Big(\frac{\log(n)}{\sqrt{n}}\Big),\ i = 1, \ldots, k_j. \tag{11}$$

One can study the limiting distribution of $\big(c_{\theta_j}\sqrt{n}\,(\lambda_{k_1+\ldots+k_{j-1}+i} - \rho_j),\ i = 1, \ldots, k_j\big)$ in some special cases. For example, the following result holds.

THEOREM 19 ([8, 30]). *Suppose that the orthonormal eigenvectors of A_n corresponding to θ_j, $1 \le j \le J_{\sigma+}$, depend on a finite number K_j of canonical basis vectors of $\mathbb{C}^n$ (without loss of generality we can assume those canonical vectors to be $e_1, \ldots, e_{K_j}$), and their coordinates are independent of n. Let*

$$c_{\theta_j} = \frac{\theta_j^2}{\theta_j^2 - \sigma^2}\,. \tag{12}$$

Then the k_j-dimensional vector

$$\big(c_{\theta_j}\sqrt{n}\,(\lambda_{k_1+\ldots+k_{j-1+i}} - \rho_j),\ i = 1, \ldots, k_j\big)$$

converges in distribution to the distribution of the ordered eigenvalues of the $k_j \times k_j$ random matrix V_j defined as

$$V_j = U_j^*(W_j + H_j)U_j, \tag{13}$$

where W_j is a Wigner random matrix of size K_j with the same marginal distribution of the matrix entries as W_N,
H_j is a centered Hermitian Gaussian matrix of size K_j, independent of W_j, with independent entries H_{st}, $1 \le s \le t \le K_j$, with the variance of the entries given by

$$\mathbb{E}(H_{ss}^2) = \frac{1}{\beta}\Big(\frac{m_4 - 3\sigma^2}{\theta_j^2}\Big) + \frac{2}{\beta}\,\frac{\sigma^4}{\theta_j^2 - \sigma^2}\,, \qquad s = 1, \ldots, K_j, \tag{14}$$

$$\mathbb{E}(|H_{st}|^2) = \frac{\sigma^4}{\theta_j^2 - \sigma^2}\,, \qquad 1 \le s < t \le K_j, \tag{15}$$

and U_j is a $K_j \times k_j$ such that the (K_j-dimensional) columns of U_j are written from the first K_j coordinates of the orthonormal eigenvectors corresponding to θ_j.

When the eigenvectors of A_n are delocalized, the limiting distribution of

$$\big(c_{\theta_j}\sqrt{n}\,(\lambda_{k_1+\ldots+k_{j-1+i}} - \rho_j),\ i = 1, \ldots, k_j\big)$$

does not depend on μ (provided certain technical conditions are satisfied). Here we present the simplest case.

THEOREM 20 ([8]). *Let M_n be a sequence of deformed Wigner matrices with distribution μ of the entries that satisfies the Poincaré inequality, but is not necessarily symmetric. Let the deformation, A_n be of the form $A_n = \mathrm{diag}(\theta, 0, \ldots, 0)$ with $\theta > \sigma$. Then*

$$\sqrt{n}(\lambda_1 - \rho_\theta) \to \Big(1 - \frac{\sigma^2}{\theta^2}\Big)\{\mu * \mathcal{N}(0, v_\theta)\},$$

where convergence is in distribution, $\mu * \mathcal{N}(0, v_\theta)$ denotes the mixture of μ and $\mathcal{N}(0, v_\theta)$, and $v_\theta = \frac{1}{2}\big(\frac{m_4 - 3\sigma^4}{\theta^2}\big) + \frac{\sigma^4}{\theta^2 - \sigma^2}$, with $m_4 = \int x^4 \, \mathrm{d}\mu(x)$.

The proof of Theorem 20 was given in [8] for symmetric μ. Once Theorem 17 is extended to the non-symmetric case, the same arguments as in [8] immediately extend the result of Theorem 20 to the non-symmetric case as well. The full proofs of Theorems 17, 18, and 19 will appear in [30]. Below, we sketch the main ideas of the approach in [29] that proves convergence in probability in Theorem 17 by extending the technique of [8, 9] to the case of non-symmetric distribution μ. The approach employed in [30] relies on the ideas developed in [6, 7].

In order to find the asymptotic spectrum of the M_n we follow the techniques of [8] and study the Stieltjes transform of the expectation of the ESD M_n. Given a probability measure, μ, on $\mathbb{R}$ its Stieltjes transform is given by

$$g(z) = \int \frac{\mathrm{d}\mu(x)}{z - x},$$

for $z \in \mathbb{C} \setminus \mathbb{R}$. Of particular interest to us is the Stieltjes transform of the ESD of a matrix and the Stieltjes transform of the semi-circle distribution.

The Stieltjes transform of the expectation of the empirical spectral distribution of a matrix M_n is

$$g_n(z) = \mathbb{E}\big(\mathrm{Tr}_n(G_n(z))\big),$$

where $\mathbb{E}$ denotes expectation, Tr_n denotes normalized trace, and $G_n(z) = (zI_n - M_n)^{-1}$ is the resolvent of M_n. We take advantage of $g_n(z)$ being the trace of the a resolvent by using resolvent identities and estimates. The Stieltjes transform of the semi-circle distribution can be characterized as the solution to

$$\sigma^2 g_\sigma^2(z) - z g_\sigma(z) + 1 = 0, \tag{16}$$

that decays to zero as $|z| \to \infty$.

Our goal is to show that g_n satisfies the same algebraic equation with a small error term. This will allow us to show $g_n(z)$ approaches $g_\sigma(z)$. We will then study the contribution of the order $1/n$ term to get the location of the large eigenvalues.

We begin with the resolvent identity

$$0 = -I - GA - GX + zG,$$

and then take normalized trace and expectation to get:

$$0 = -1 - \mathbb{E}[\mathrm{Tr}_n(GA)] - \frac{1}{n}\sum_{i,j} \mathbb{E}[G_{ij}X_{ji}] + z\mathbb{E}[\mathrm{Tr}_n(G)]. \tag{17}$$

The following cumulant expansion [24] is used to separate the $\mathbb{E}[G_{ij}X_{ji}]$ term. Given ξ, a real-valued random variable with $p+2$ finite moments, and ϕ a function from $\mathbb{C} \to \mathbb{R}$

with $p+1$ continuous and bounded derivatives, then

$$\mathbb{E}(\xi\phi(\xi)) = \sum_{a=0}^{p} \frac{\kappa_{a+1}}{a!}\,\mathbb{E}(\phi^{(a)}(\xi)) + \epsilon \tag{18}$$

where κ_a are the cumulants of ξ, $|\epsilon| \le C\sup_t\big|\phi^{(p+1)}(t)\big|\mathbb{E}(|\xi|^{p+2})$, C depends only on p.

After expanding and estimating the error terms we have:

$$\left|\sigma^2 g_n^2(z) - z g_n(z) + 1 + \frac{1}{n}\,\mathbb{E}(\mathrm{Tr}(GA)) + \frac{\kappa_4}{2n}\,\mathbb{E}\left[\Big(\frac{1}{n}\sum_i G_{ii}^2\Big)^2\right]\right| \le \frac{P_5(|\operatorname{Im} z|^{-1})}{n^{3/2}}. \tag{19}$$

Here and throughout the paper, P_k denotes a polynomial of degree k with coefficients that do not depend on n. The leading coefficient of P_k is always positive. Note that the third cumulant terms give the contribution of order $O(n^{-3/2})$ at the right hand side of (19). If we assume the third moments vanish then the error term is $O(n^{-2})$. In (19) the leading order terms satisfy equation (16), this allows us to show that $|g_n(z) - g_\sigma(z)|$ is $O(n^{-1})$. Then the resolvent identity and cumulant expansion are applied to the $O(n^{-1})$ terms to determine their leading order term. This gives

$$\mathbb{E}[\mathrm{Tr}(G_n(z)A_n)] = \sum_{j=1}^{J} \frac{k_j\theta_j}{z - \sigma^2 g_\sigma(z) - \theta_j} + \frac{P_{10}(|\operatorname{Im} z|^{-1})}{n^{1/2}}, \tag{20}$$

and

$$\mathbb{E}\left[\Big(\frac{1}{N}\sum_i G_{ii}^2\Big)^2\right] = g_\sigma^4 + \frac{P_7(|\operatorname{Im} z|^{-1})}{n^{1/2}}, \tag{21}$$

$$\left|\sigma^2 g_n^2(z) - z g_n(z) + \frac{1}{n}\sum_{j=1}^{J} \frac{\theta_j}{z - \sigma^2 g_\sigma(z) - \theta_j} + \frac{\kappa_4}{2n}\, g_\sigma^4(z)\right| \le \frac{P_{17}(|\operatorname{Im} z|^{-1})}{n^{3/2}}. \tag{22}$$

This equation gives:

$$g_n(z) = g_\sigma(z) + \frac{1}{n}\, g_\sigma(z)^{-1}\left(\sum_{j=1}^{J} \frac{k_j\theta_j}{z - \sigma^2 g_\sigma(z) + \theta_j} + \frac{\kappa_4}{2}\, g_\sigma^4(z)\right)\int \frac{\mathrm{d}\mu_{sc}(x)}{(z-x)^2} + \frac{P_{17}(|\operatorname{Im} z|^{-1})}{n^{3/2}}. \tag{23}$$

The support of the ESD of the expectation of M_n is given by the singularities of its Stieltjes transform. Equation (23) thus gives that the support is $[-2\sigma, 2\sigma]$ and $\{\rho_1, \ldots, \rho_J\}$. The $[-2\sigma, 2\sigma]$ part comes from the order 1 terms and gives the semicircle. The $\{\rho_1, \ldots, \rho_J\}$ comes from the order n^{-1} term and gives the extremal eigenvalues. To get the convergence in probability of the eigenvalues let $F = \{t \in \mathbb{R} : d(t, K) \ge \epsilon\}$, where K is defined in (10). Then it follows from (23) that

$$\mathbb{P}\big(|Z_n| \ge n^{-1-\epsilon}\big) \le \frac{\mathbb{E}[|Z_n|^2]}{n^{-2-\epsilon}} = \frac{O(n^{-3})}{n^{-2-\epsilon}} = O(n^{-1+\epsilon}).$$

So $\mathbb{P}[\mathrm{Tr}_n(1_F(M_n)) \ge O(n^{-1-\epsilon})] \to 0$ and along any subsequence that grows faster than $n^{1+\epsilon}$ the probability of an eigenvalue being outside of K is summable so by Borel–Cantelli theorem there are almost surely no eigenvalues outside of K.

The final step is to show that the number of eigenvalues of A_n at θ_i is equal to the number of eigenvalues of M_n in a small neighborhood of ρ_{θ_i}. To do this, we introduce a continuous family of matrices that interpolate between A_n and M_n. Using Weyl's eigenvalue inequalities, it is shown that multiplicity of eigenvalues is preserved.

The proof of Theorem 18 relies on the results about fluctuations of resolvent entries of standard Wigner matrices (see Theorems 1.1, 1.3, and 1.5 in [30]). Consider a fixed eigenvalue θ_m of A_n such that $\theta_m > \sigma$ and denote by $v^{(1)}, \dots, v^{(k_m)}$ those of the eigenvectors of A_n that correspond to the eigenvalue θ_m. Let $\Theta^{(m)}$ be an $m \times m$ matrix

$$\Theta_{ij}^{(m)} = \sqrt{n}\,\big(\langle v^{(i)}, R_n(\rho_m) v^{(j)}\rangle - g_\sigma(\rho_m)\big) = \sqrt{n}\,\Big(\langle v^{(i)}, R_n(\rho_m) v^{(j)}\rangle - \frac{1}{\theta_m}\Big),$$

where $R_n(z) = (z - X_n)^{-1}$ is the resolvent of a standard Wigner matrix $X_n = \frac{1}{\sqrt{n}} W_n$. Below we formulate Lemma 4.3 from [30] which plays the central role in the proof.

Lemma 21. *Let $y_1 \geq \ldots \geq y_{k_m}$ be the ordered eigenvalues of the matrix $\Theta^{(m)}$. Then, almost surely,*

$$\sqrt{n}\,\big(\lambda_{k_1+\ldots+k_{j-1}+i} - \rho_m\big) = -\frac{1}{g'_\sigma(\rho_m)}\, y_i + O\Big(\frac{\log^2(n)}{\sqrt{n}}\Big), \quad i = 1, \dots, k_j. \tag{24}$$

The limiting distribution of $\sqrt{n}\,(\langle u, R_n(z) v\rangle - g_\sigma(z)\langle u, v\rangle)$ is studied in Theorems 1.1, 1.3, and 1.5 in [30].

4. Multivariate resolvent identities at the edge of the spectrum. Consider the Gaussian Orthogonal Ensemble (GOE), that is random matrices $A_n = \frac{1}{\sqrt{n}}\,(a_{ij})_{ij=1}^n$ where $a_{ij} = \mathcal{N}(0, 1+\delta_{ij})$, $i \leq j$, are independent Gaussian random variables with mean zero.

The complex analog is the Gaussian Unitary Ensemble (GUE). In this case $A = \frac{1}{\sqrt{n}}\,(a_{ij})_{ij=1}^n$ is a Hermitian matrix with $a_{ij} = x_{ij} + i y_{ij}$. Here the upper triangular entries x_{ij} and y_{ij}, $i < j$, are independent Gaussian random variables with mean zero, $\mathcal{N}(0, \frac{1}{2})$, while the diagonal entries x_{ii} are $\mathcal{N}(0,1)$.

Consider the resolvent matrix $G(z) = (A - 2 - zn^{-2/3})^{-1}$, $\operatorname{Im} z > 0$. We will use the shorthand $(A - z)^{-1} = (A - z \cdot I)^{-1}$. To consider the joint distribution of the largest eigenvalues at the edge of the spectrum, we rescale the eigenvalues as

$$\lambda_j^{(n)} = 2 + \xi_j^{(n)} n^{-2/3}, \quad j = 1, 2, \dots, n, \tag{25}$$

where $\lambda_1^{(n)} \geq \lambda_2^{(n)} \ldots \geq \lambda_n^{(n)}$ are the ordered eigenvalues of A_n. Let

$$g_{n,L}(z) = n^{-2L/3} \operatorname{Tr} G_n^L(z) = n^{-2L/3} \operatorname{Tr}(A_n - 2 - zn^{-2/3})^{-L} = \sum_1^n (\xi_j^{(n)} - z)^{-L}, \tag{26}$$

for positive integers $L = 1, 2, \dots$.

It can be shown ([36]) that for $L \geq 2$, $g_{n,L}(z)$ is a "local" statistic of the largest eigenvalues in the GOE in a sense that only the eigenvalues from a $O(n^{-2/3})$-neighborhood of the right edge of the spectrum give non-vanishing contribution to $g_{n,L}(z)$ in the limit $n \to \infty$. For $L = 1$, the linear statistic $n^{-2/3} \operatorname{Tr} G_n(z)$ is not local in the above sense since the main contribution comes from the eigenvalues in the bulk of the spectrum. However, the centralized statistic $g_{n,1}^c(z) = n^{-2/3}(n + \operatorname{Tr} G_n(z))$ is again a local one. In [36] it was shown that the joint moments of $g_{n,L}(z)$, $L > 1$, and $g_{n,1}^c(z)$ satisfy certain recursive identities in the limit $n \to \infty$. Similar results were obtained for the GUE, as well as for the Wishart real and complex random matrices at the hard edge of the spectrum. One expects these identities do not depend on the marginal distribution of matrix entries.

Let

$$m_L(z_1, \dots, z_L) = \mathbb{E} \prod_{k=1}^{L} n^{-2/3}(n + \operatorname{Tr} G(z_k)), \quad \operatorname{Im} z_k > 0. \tag{27}$$

Clearly, $g_{n,L}(z) = m_L(z, \dots, z)$. One can extend the recursive identities from [36] to $m_L(z_1, \dots, z_L)$, $L \geq 1$.

THEOREM 22. *Let m_L be defined as above for the GOE. For $L \geq 2$ we have*

$$\begin{aligned} &z_1 \overline{m}_{L-1} - \frac{\partial m_L}{\partial z_1} - m_{L+1}\big|_{z_{L+1}=z_1} \\ &- 2\sum_{k=2}^{L}\Big[\frac{1}{z_k - z_1}\frac{\partial \overline{m}_{L-1}}{\partial z_k} - \frac{1}{(z_k - z_1)^2}\,\overline{m}_{L-1} + \frac{1}{(z_k - z_1)^2}\,\overline{m}_{L-1}\big|_{z_k \to z_1}\Big] = \mathcal{O}(n^{-1/3}), \end{aligned} \tag{28}$$

where $\overline{m}_{L-1} = m_{L-1}(z_2, \dots, z_L)$. For $L = 1$ we have

$$z_1 - m_2\big|_{z_2=z_1} - \frac{\partial m_1}{\partial z_1} = \mathcal{O}(n^{-1/3}). \tag{29}$$

THEOREM 23. *Let m_L be defined as above for the GUE. For $L \geq 2$ we have*

$$\begin{aligned} &z_1 \overline{m}_{L-1} - m_{L+1}\big|_{z_{L+1}=z_1} \\ &- \sum_{k=2}^{L}\Big[\frac{1}{z_k - z_1}\frac{\partial \overline{m}_{L-1}}{\partial z_k} - \frac{1}{(z_k - z_1)^2}\,\overline{m}_{L-1} + \frac{1}{(z_k - z_1)^2}\,\overline{m}_{L-1}\big|_{z_k \to z_1}\Big] = \mathcal{O}(n^{-1/3}), \end{aligned} \tag{30}$$

where $\overline{m}_{L-1} = m_{L-1}(z_2, \dots, z_L)$. For $L = 1$ we have

$$z_1 - m_2\big|_{z_2=z_1} = \mathcal{O}(n^{-1/3}). \tag{31}$$

For an explanation of the appearance of $(z_k - z_1)^{-1}$ see (34) below. When all variables are set equal, one obtains equations that agree with [36].

Proof of Theorem 22 (GOE). Let $g(z) = g^c_{n,1}(z) = n^{-2/3}(n + \operatorname{Tr} G(z))$, and begin with

$$n^{1/3}(2 + z_1 n^{-2/3}) m_L(z_1, \dots, z_L) = n^{1/3}(2 + z_1 n^{-2/3}) \mathbb{E} \prod_{k=1}^{L} g(z_k).$$

We rewrite the first factor using the resolvent identity

$$(A - z)^{-1} = (B - z)^{-1} - (A - z)^{-1}(A - B)(B - z)^{-1}, \tag{32}$$

obtaining

$$\begin{aligned} n^{1/3} g(z_1) &= n^{1/3}\big(n^{-2/3}(n + \operatorname{Tr} G(z_1))\big) \\ &= n^{2/3} - n^{2/3}(2 + z_1 n^{-2/3})^{-1} + (2 + z_1 n^{-2/3})^{-1} n^{-1/3} \operatorname{Tr} AG(z_1). \end{aligned}$$

This substitution gives us

$$(n^{2/3} + z_1) m_{L-1}(z_2, \dots, z_L) + n^{-1/3} \mathbb{E} \operatorname{Tr}\big(AG(z_1)\big) \prod_{k=2}^{L} g(z_k).$$

To deal with the second term, we use the special case of (18) for mean zero Gaussian random variables ξ,

$$\mathbb{E}\xi f(\xi) = \mathrm{Var}(\xi)\,\mathbb{E}f'(\xi) \quad (\mathbb{E}\xi = 0). \tag{33}$$

We have

$$n^{-1/3}\sum_{ij}\mathbb{E}A_{ij}G_{ji}(z_1)\prod_{k=2}^{L}g(z_k) = n^{-4/3}\sum_{ij}\mathbb{E}\frac{\partial}{\partial A_{ij}}\Big[G_{ji}(z_1)\prod_{k=2}^{L}g(z_k)\Big],$$

obtaining

$$-n^{-4/3}\sum_{ij}\mathbb{E}\big[G_{ji}(z_1)G_{ji}(z_1)+G_{jj}(z_1)G_{ii}(z_1)\big]\prod_{k=2}^{L}g(z_k)$$
$$+n^{-4/3}\sum_{ij}\mathbb{E}G_{ji}(z_1)\sum_{k=2}^{L}-2n^{-2/3}\big(G^2(z_k)\big)_{ij}\prod_{r\neq k}g(z_r).$$

We rewrite

$$-n^{-4/3}\sum_{ij}\mathbb{E}\big[G_{jj}(z_1)G_{ii}(z_1)\big]\prod_{k=2}^{L}g(z_k) = -E\big[n^{-2/3}(n+\mathrm{Tr}\,G(z_1))\big]^2\prod_{l=2}^{L}g(z_l)$$
$$+2n^{1/3}\mathbb{E}n^{-2/3}(n+\mathrm{Tr}\,G(z_1))\prod_{l=2}^{L}g(z_l) - n^{2/3}\mathbb{E}\prod_{l=2}^{L}g(z_l).$$

Combining these equations and simplifying algebraically gives

$$\mathcal{O}(n^{-1/3}) = z_1 m_{L-1}(z_2,\dots,z_L) - \mathbb{E}n^{-4/3}\,\mathrm{Tr}\,G^2(z_1)\prod_{k=2}^{L}g(z_k)$$
$$-\mathbb{E}n^{-4/3}(n+\mathrm{Tr}\,G(z_1))^2\prod_{k=2}^{L}g(z_k) - 2\sum_{k=2}^{L}\mathbb{E}n^{-2}\,\mathrm{Tr}\big[G(z_1)G^2(z_k)\big]\prod_{r\neq k}g(z_r).$$

We may now rewrite these expressions in terms of the m_L (27). Using another resolvent identity

$$(B-z_1)^{-1}(B-z_2)^{-1} = \frac{1}{z_2-z_1}\,(B-z_2)^{-1} - \frac{1}{z_2-z_1}\,(B-z_1)^{-1},$$

we rewrite

$$G(z_1)G^2(z_k) = \frac{n^{2/3}}{z_k-z_1}\,G^2(z_k) - \frac{n^{4/3}}{(z_k-z_1)^2}\,G(z_k) + \frac{n^{4/3}}{(z_k-z_1)^2}\,G(z_1). \tag{34}$$

Simplifying gives the desired identity (22). The proof of (23) is very similar and is left to the reader. ■

5. Resolvent identities in the bulk for Gaussian and Wishart ensembles. In this section we consider local statistics in the bulk of the spectrum of Gaussian and Wishart ensembles. We consider the GOE and GUE as defined in the previous section. We are concerned with the joint moments of the collection of random variables

$$g_{n,l}(z) = n^{-l}\,\mathrm{Tr}\,G_n^l(z), \qquad l \geq 1,$$

where $G_n(z) = (A_n - \lambda_0 - n^{-1}z)^{-1}$, $\operatorname{Im} z > 0$, $-2 < \lambda_0 < 2$. Let K be a multi-index, $K = (k_1, k_2, \dots)$, with finitely many non-zero natural numbers $k_i \geq 0$. Let

$$m_{n,K}(z) = \mathbb{E} \prod_{l \geq 1} \big(n^{-l} \operatorname{Tr} G_n^l(z)\big)^{k_l}.$$

For clarity we may suppress the dependence on n and z.

Theorem 24 (Bulk GOE). *Let $A_n = \frac{1}{\sqrt{n}} (a_{ij})_{ij=1}^n$ be a GOE matrix. For non-zero multi-indices K we have*

$$\lambda_0 m_{K+e_1} = -m_K - m_{K+e_2} - m_{K+2e_1} - 2 \sum_{l \geq 1} l k_l m_{K-e_l+e_{l+2}} + \mathcal{O}(n^{-1}), \tag{35}$$

with boundary condition

$$\lambda_0 m_{e_1} = -1 - m_{e_2} - m_{2e_1} + \mathcal{O}(n^{-1}). \tag{36}$$

Theorem 25 (Bulk GUE). *Let $A_n = \frac{1}{\sqrt{n}} (a_{ij})_{ij=1}^n$ be a GUE matrix. For non-zero multi-indices K we have*

$$\lambda_0 m_{K+e_1} = -m_K - m_{K+2e_1} - \sum_{l \geq 1} l k_l m_{K-e_l+e_{l+2}} + \mathcal{O}(n^{-1}), \tag{37}$$

with boundary condition

$$\lambda_0 m_{e_1} = -1 - m_{2e_1} + \mathcal{O}(n^{-1}). \tag{38}$$

Now let us consider the real and complex Wishart (Laguerre) ensembles. Let $A_{n,N}$ be an $n \times N$ matrix with independent standard normal entries $a_{ij} = \mathcal{N}(0,1)$. Assume that $N \geq n$ and $N - n = \nu$ is fixed. Let $M_{n,N} = n^{-1} A A^t$. The limiting (Marchenko–Pastur) distribution of the eigenvalues of $M_{n,N}$ is supported on the interval $[0,4]$ and has density $\frac{1}{2\pi} \sqrt{(4-x)/x}$. Let λ_0 be in the bulk of the spectrum, i.e. $\lambda_0 \in (0,4)$. Similarly to the Wigner case, we define

$$m_{n,K}(z) = \mathbb{E} \prod_{l \geq 1} (n^{-l} \operatorname{Tr} G_n^l(z))^{k_l},$$

where $G_n(z) = (M_{n,N} - \lambda_0 - n^{-1}z)^{-1}$, $\operatorname{Im} z > 0$.

Theorem 26 (Bulk Real Wishart). *Let $M_{n,N}$ be a real Wishart matrix. For non-zero multi-indices K we have*

$$m_{K+e_1} = -\frac{1}{\lambda_0} m_K - m_{K+e_2} - m_{K+2e_1} - 2 \sum_{l \geq 1} l k_l m_{K-e_l+e_{l+2}} + \mathcal{O}(n^{-1}), \tag{39}$$

with boundary condition

$$m_{e_1} = -\frac{1}{\lambda_0} - m_{e_2} - m_{2e_1} + \mathcal{O}(n^{-1}). \tag{40}$$

Theorem 27 (Bulk Complex Wishart). *Let $M_{n,N}$ be a complex Wishart matrix. For non-zero multi-indices K we have*

$$m_{K+e_1} = -\frac{1}{\lambda_0} m_K - m_{K+2e_1} - \sum_{l \geq 1} l k_l m_{K-e_l+e_{l+2}} + \mathcal{O}(n^{-1}), \tag{41}$$

with boundary condition

$$m_{e_1} = -\frac{1}{\lambda_0} - m_{2e_1} + \mathcal{O}(n^{-1}). \tag{42}$$

Proof of Theorem 26 (*Bulk Real Wishart*). Here we consider the boundary term in the real Wishart case. We begin with $(\lambda_0 + n^{-1}z)m_{e_1} = (\lambda_0 + n^{-1}z)\mathbb{E}n^{-1}\operatorname{Tr} G(z)$, where

$$G(z) = (AA^t - \lambda_0 - zn^{-1})^{-1}.$$

The resolvent identity (32) gives

$$G(z) = -(\lambda_0 + zn^{-1})^{-1} + (\lambda_0 + zn^{-1})^{-1}AA^tG, \tag{43}$$

and therefore

$$(\lambda_0 + n^{-1}z)m_{e_1} = -1 + n^{-1}\sum_{ijp}\mathbb{E}A_{ip}A_{jp}G_{ji},$$

where $i, j = 1, \ldots, n$ and $p = 1, \ldots, N$. We use the Gaussian decoupling formula (33) with $\xi = A_{ip}$ and $f(\xi) = A_{jp}G_{ji}$,

$$\mathbb{E}A_{ip}A_{jp}G_{ji} = \operatorname{Var}(A_{ip})\mathbb{E}\Big(\frac{\partial A_{jp}}{\partial A_{ip}}G_{ji} + A_{jp}\frac{\partial G_{ji}}{\partial A_{ip}}\Big).$$

In this setting we have

$$\frac{\partial G_{kl}}{\partial A_{ip}} = -G_{ki}(A^tG)_{pl} - (GA)_{kp}G_{il}, \tag{44}$$

which gives

$$\begin{aligned}
&(\lambda_0 + n^{-1}z)m_{e_1}\\
&= -1 + n^{-2}\sum_{ijp}\mathbb{E}\delta_{ij}G_{ji} - n^{-2}\sum_{ijp}\mathbb{E}A_{jp}G_{ji}(A^tG)_{pi} - n^{-2}\sum_{ijp}\mathbb{E}A_{jp}(GA)_{jp}G_{ii}\\
&= -1 + n^{-1}\mathbb{E}\operatorname{Tr} G - n^{-2}\mathbb{E}\operatorname{Tr}(GAA^tG) - n^{-2}\mathbb{E}(\operatorname{Tr} GAA^t)(\operatorname{Tr} G).
\end{aligned}$$

Using the simple identity $GAA^t = I + (\lambda_0 + n^{-1}z)G$, we see that the right hand side equals

$$-1 - (\lambda_0 + n^{-1}z)n^{-2}\mathbb{E}\operatorname{Tr} G^2 - (\lambda_0 + n^{-1}z)\mathbb{E}(n^{-1}\operatorname{Tr} G)^2 + \mathcal{O}(n^{-1}).$$

This gives us the boundary condition (40).

Next we consider non-zero multi-indices K. Let $g_K = \prod_{l\geq 1}(n^{-l}\operatorname{Tr} G^l)^{k_l}$. Using (43) we have

$$(\lambda_0 + n^{-1}z)m_{K+e_1} = -m_K + n^{-1}\mathbb{E}(\operatorname{Tr} AA^tG)g_K = -m_K + n^{-1}\sum_{ijp}\mathbb{E}A_{ip}A_{jp}G_{ji}g_K.$$

Applying the Gaussian decoupling formula (33) with $\xi = A_{ip}$ and $f(\xi) = A_{jp}G_{ji}g_K$, we have $\mathbb{E}A_{ip}A_{jp}G_{ji}g_K = \operatorname{Var}(A_{ip})\mathbb{E}\frac{\partial}{\partial A_{ip}}\big(A_{jp}G_{ji}g_K\big)$. By using (44) the right hand side becomes

$$n^{-1}\mathbb{E}\Big[\delta_{ij}G_{ji}g_K - A_{jp}\big(G_{ji}(A^tG)_{pi} + (GA)_{jp}G_{ii}\big)g_K + A_{jp}G_{ji}\frac{\partial g_K}{\partial A_{ip}}\Big].$$

To compute the last term we use

$$\frac{\partial \operatorname{Tr} G^l}{\partial A_{ip}} = -2l(A^tG^{l+1})_{pi}. \tag{45}$$

Putting this together we have

$$
\begin{aligned}
&(\lambda_0 + n^{-1}z) m_{K+e_1} \\
&= -m_K + \mathbb{E}(n^{-1}\operatorname{Tr} G) g_K - \mathbb{E}(n^{-2}\operatorname{Tr} AA^t G^2) g_K - \mathbb{E}(n^{-1}\operatorname{Tr} AA^t G)(n^{-1}\operatorname{Tr} G) g_K \\
&= -2\,\mathbb{E}\sum_{l\geq 1} l k_l (n^{-l}\operatorname{Tr} G^l))^{k_l-1}\big(n^{-l-2}\operatorname{Tr} AA^t G^{l+2}\big)\prod_{r\neq l}(n^{-r} G^r)^{k_r} + \mathcal{O}(n^{-1}).
\end{aligned}
$$

By using $GAA^t = I + (\lambda_0 + n^{-1}z)G$, the right hand side becomes

$$
\begin{aligned}
&- m_K - (\lambda_0 + n^{-1}z) n^{-2}\mathbb{E}(\operatorname{Tr} G^2) g_K - (\lambda_0 + n^{-1}z) n^{-2}(\operatorname{Tr} G)^2 g_K \\
&- 2(\lambda_0 + n^{-1}z)\mathbb{E}\sum_{l\geq 1} l k_l (n^{-l}\operatorname{Tr} G^l)^{k_l-1}\big(n^{-l-2}\operatorname{Tr} G^{l+2}\big)\prod_{r\neq l}(n^{-r}\operatorname{Tr} G^r)^{k_r} + \mathcal{O}(n^{-1}).
\end{aligned}
$$

This gives us the identity (39). The proofs of (27), (24) and (25) are similar and left to the reader. ∎

References

[1] G. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Stud. Adv. Math. 118, Cambridge University Press, Cambridge, 2010.

[2] Z. D. Bai, *Methodologies in spectral analysis of large dimensional random matrices, a review*, Statist. Sinica 9 (1999), 611–677.

[3] Z. D. Bai, Y. Q. Yin, *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix*, Ann. Probab. 16 (1988), 1729–1741.

[4] J. Baik, G. Ben Arous, S. Péché, *Phase transition of the largest eigenvalue for non-null complex sample covariance matrices*, Ann. Probab. 33 (2005), 1643–1697.

[5] J. Baik, J. W. Silverstein, *Eigenvalues of large sample covariance matrices of spiked population models*, J. Multivariate Anal. 97 (2006), 1382–1408.

[6] F. Benaych-Georges, A. Guionnet, M. Maida, *Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices*, Electron. J. Probab. 16 (2011), 1621–1662.

[7] F. Benaych-Georges, A. Guionnet, M. Maida, *Large deviations of the extreme eigenvalues of random deformations of matrices*, `arXiv:1009.0135v2`.

[8] M. Capitaine, C. Donati-Martin, D. Féral, *The largest eigenvalue of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations*, Ann. Probab. 37 (2009), 1–47.

[9] M. Capitaine, C. Donati-Martin, D. Féral, *Central limit theorems for eigenvalues of deformations of Wigner matrices*, `arXiv:0903.4740 [math.PR]`.

[10] O. Costin, J. Lebowitz, *Gaussian fluctuations in random matrices*, Phys. Rev. Lett. 75 (1995), 69–72.

[11] P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*, Courant Lect. Notes Math. 3, New York University, Courant Inst. of Math. Sciences, New York; Amer. Math. Soc., Providence, RI, 1999.

[12] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides, X. Zhou, *Strong asymptotics of orthogonal polynomials with respect to exponential weights*, Comm. Pure Appl. Math. 52 (1999), 1491–1552.

[13] F. J. Dyson, *Statistical theory of the energy levels of complex systems. III*, J. Mathematical Phys. 3 (1962), 166–175.

[14] D. Féral, S. Péché, *The largest eigenvalue of rank one deformation of large Wigner matrices*, Comm. Math. Phys. 272 (2007), 185–228.

[15] S. Fisk, *A very short proof of Cauchy's interlace theorem*, Amer. Math. Monthly 112 (2005), 118.

[16] P. J. Forrester, E. M. Rains, *Interrelationships between orthogonal, unitary and symplectic matrix ensembles*, in: Random Matrix Models and their Applications, Math. Sci. Res. Inst. Publ. 40, Cambridge Univ. Press, Cambridge, 2001, 171–208.

[17] Z. Füredi, J. Komlós, *The eigenvalues of random symmetric matrices*, Combinatorica 1 (1981), 233–241.

[18] J. Gunson, *Proof of a conjecture of Dyson in the statistical theory of energy levels*, J. Mathematical Phys. 3 (1962), 752–753.

[19] J. Gustavsson, *Gaussian fluctuations of eigenvalues in the GUE*, Ann. Inst. H. Poincaré Probab. Statist. 41 (2005), 151–178.

[20] A. Gut, *Probability: A Graduate Course*, Springer Texts Statist., Springer, New York, 2005.

[21] U. Haagerup, S. Thorbjørnsen, *A new application of random matrices:* $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ *is not a group*, Ann. of Math. (2) 162 (2005), 711–775.

[22] J. B. Hough, M. Krishnapur, Y. Peres, B. Virág, *Determinantal processes and independence*, Probab. Surv. 3 (2006), 206–229.

[23] K. Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys. 209 (2000), 437–476.

[24] A. M. Khorunzhy, B. Khoruzhenko, L. Pastur, *Asymptotic properties of large random matrices with independent entries*, J. Math. Phys. 37 (1996), 5033–5060.

[25] M. L. Mehta, *Random Matrices*, 3rd ed., Pure Appl. Math. (Amst.) 142, Elsevier/ Academic Press, Amsterdam, 2004.

[26] S. O'Rourke, *Gaussian fluctuations of eigenvalues in Wigner random matrices*, J. Stat. Phys. 138 (2010), 1045–1066.

[27] L. A. Pastur, *On the spectrum of random matrices*, Teoret. Mat. Fiz. 10 (1972), 102–112 (in Russian).

[28] S. Péché, *The largest eigenvalue of small rank perturbations of Hermitian random matrices*, Probab. Theory Related Fields 134 (2006), 127–174.

[29] A. Pizzo, D. Renfrew, A. Soshnikov, *On finite rank deformations of Wigner matrices*, Ann. Inst. Henri Poincaré Probab. Stat., to appear; `arXiv:1103.3731 math.PR`.

[30] A. Pizzo, D. Renfrew, A. Soshnikov, *Fluctuations of matrix entries of regular functions of Wigner matrices*, J. Stat. Phys. 146 (2012), 550–591.

[31] M. Reed, B. Simon, *Methods of Modern Mathematical Physics, Vol. 1. Functional Analysis*, Academic Press, New York, 1980.

[32] A. B. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. 207 (1999), 697–733.

[33] A. B. Soshnikov, *Determinantal random point fields*, Uspekhi Mat. Nauk 55 (2000), no. 5, 107–160; English transl.: Russian Math. Surveys 55 (2000), 923–975.

[34] A. Soshnikov, *Gaussian fluctuation for the number of particles in Airy, Bessel, sine and other determinantal random point fields*, J. Statist. Phys. 100 (2000), 491–522.

[35] A. Soshnikov, *Gaussian limit for determinantal random point fields*, Ann. Probab. 30 (2002), 171–187.

[36] A. Soshnikov, *On resolvent identities in Gaussian ensembles at the edge of the spectrum*, in: New Trends in Mathematical Physics (Rio de Janeiro, 2006), Springer, Dordrecht, 2009, 615–627.

[37] T. Tao, V. Vu, *Random matrices: universality of local eigenvalue statistics*, Acta Math. 206 (2011), 127–204.

[38] T. Tao, V. Vu, *Random matrices: universality of local eigenvalue statistics up to the edge*, Comm. Math. Phys. 298 (2010), 549–572.

[39] C. A. Tracy, H. Widom, *Level spacing distributions and the Airy kernel*, Comm. Math. Phys. 159 (1994), 151–174.

[40] C. A. Tracy, H. Widom, *On orthogonal and symplectic matrix ensembles*, Comm. Math. Phys. 177 (1996), 727–754.

[41] E. Wigner, *On the distribution of the roots of certain symmetric matrices*, Ann. of Math. (2) 67 (1958), 325–327.