# On Resolvent Identities in Gaussian Ensembles at the Edge of the Spectrum


Alexander Soshnikov *


August 15, 2008


**Abstract**

We obtain the recursive identities for the joint moments of the traces of the powers of the resolvent for Gaussian ensembles of random matrices at the soft and hard edges of the spectrum. We also discuss the possible ways to extend these results to the non-Gaussian case.


## 1 Introduction

Consider the Gaussian Orthogonal Ensemble (GOE) of real symmetric $n \times n$ random matrices

$$A_n = \frac{1}{\sqrt{n}} \left(a_{ij}\right)_{i,j=1}^{n}, \tag{1}$$

where $\{a_{ij} = a_{ji}\}_{i \le j}$ are independent $N(0, 1 + \delta_{ij})$ random variables. GOE is the archetypal example of a Wigner real symmetric random matrix where the matrix entries $\{a_{ij} = a_{ji}\}_{i \le j}$ are assumed to be independent up from the diagonal, centralized, and to have the same variance (except, possibly, on the diagonal). It follows from the classical Wigner semi-circle law (([20], [21], [1]) that the empirical distribution function of the eigenvalues of $A_n$ converges as $n \to \infty$ to the limiting distribution with the probability density $\frac{1}{2\pi}\sqrt{4-x^2}$ supported on the interval $[-2, 2]$. Celebrated work by Tracy and Widom (see [17] for the GOE case) proved that the largest eigenvalues of $A_n$ deviate from the right edge of the spectrum on the order of $n^{-2/3}$. In particular, Tracy and Widom calculated the limiting distribution of the largest eigenvalue:

$$\lim_{n\to\infty} \mathbb{P}\left(\lambda_{max} \le 2 + x n^{-2/3}\right) = F_1(x) = \exp\left(-1/2 \int_x^{\infty} q(t) + (t-x) q^2(t) dt\right), \tag{2}$$


*Department of Mathematics, University of California at Davis, One Shields Ave., Davis, CA 95616, USA. E-mail address: soshniko@math.ucdavis.edu.


Research was supported in part by the NSF grant DMS-0707145.

where $q(x)$ is the solution of the Painléve II differential equation $q''(x) = xq(x) + 2q^3(x)$ with the asymptotics at infinity $q(x) \sim Ai(x)$ as $x \to +\infty$. Here $Ai(x)$ denotes the Airy function.

To consider the joint distribution of the largest eigenvalues at the edge of the spectrum, we rescale the eigenvalues as

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

where $\lambda_1^{(n)} \geq \lambda_2^{(n)} \ldots \geq \lambda_n^{(n)}$ are the ordered eigenvalues of $A_n$. It then follows from the results of [13], [18] that the random point configuration $\{\xi_j^{(n)}, \quad 1 \leq j \leq n\}$ converges in distribution on the cylinder sets to the random point process on the real line with the $k$-point correlation functions given by

$$\rho_k(x_1, \ldots, x_k) = \left(\det\left(K(x_i, x_j)\right)_{1 \leq i,j \leq k}\right)^{1/2}, \tag{4}$$

where $K(x, y)$ is a $2 \times 2$ matrix-valued kernel with the entries

$$K_{11}(x, y) = K_{Airy}(x, y) + \frac{1}{2} Ai(x) \left(1 - \int_y^{+\infty} Ai(z) dz\right), \tag{5}$$

$$K_{12}(x, y) = -\partial_y K_{Airy}(x, y) - \frac{1}{2} Ai(x) Ai(y), \tag{6}$$

$$K_{21}(x, y) = -\int_x^{+\infty} K_{Airy}(z, y) dz + \frac{1}{2}\left(\int_y^x Ai(z) dz + \int_x^{+\infty} Ai(z) dz \int_y^{+\infty} Ai(z) dz\right), \tag{7}$$

$$K_{22}(x, y) = K_{11}(y, x), \tag{8}$$

and the Airy kernel $K_{Airy}(x, y)$ is defined as

$$K_{Airy}(x, y) = \int_0^{+\infty} Ai(x + z) Ai(y + z) dz = \frac{Ai(x) Ai'(y) - Ai'(x) Ai(y)}{x - y}. \tag{9}$$

Therefore, the $k$-point correlation function of the limiting random point process is given by the square root of the determinant of the $2k \times 2k$ matrix defined in (4) and (5). One can also rewrite the $k$-point correlation function in the pfaffian form (see e.g. [9]) which shows that the limiting random point process belongs to the family of the pfaffian random point processes (see e.g. [14]). In particular, the right-most particle of this pfaffian random point process is given by the Tracy-Widom distribution (2). Moreover, it was shown in [12] that the asymptotic behavior of the largest eigenvalues is universal for Wigner real symmetric matrices with sub-Gaussian and symmetrically distributed entries.

Define

$$G_n(z) = (A_n - 2 - zn^{-2/3})^{-1} \tag{10}$$

for complex $z$ with non-zero imaginary part $\Im z \neq 0$. Here and throughout the paper, we will use $(A_n - z)^{-1}$ as the shorthand notation for the resolvent matrix $(A_n - zId)^{-1}$.

Let

$$g_{n,k}(z) = n^{-2k/3}\mathrm{Tr}G_n^k(z) = n^{-2k/3}\mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-k} = \sum_1^n (\xi_j^{(n)} - z)^{-k} \tag{11}$$

for positive integers $k = 1, 2, \ldots$. It can be shown that for $k \geq 2$, $g_{n,k}(z)$ is a "local" statistic of the largest eigenvalues in the GOE. Indeed, only eigenvalues from the $O(n^{-2/3})$-neighborhood of the right edge of the spectrum give non-vanishing contribution to $g_{n,k}(z)$ in the limit $n \to \infty$. For example, the joint contribution of the eigenvalues from $(-\infty, 2-\delta]$ can be trivially bounded in absolute value by $n^{1-2k/3}|\delta + zn^{-2/3}|^k = o(1)$ for large $n$ uniformly in $z$ with $\Re z$ bounded from below. More delicate estimates involving the asymptotics of the one-point correlation function, imply that the joint contribution of the eigenvalues from $(-\infty, 2 - n^{-2/3+\varepsilon})$ to $g_{n,k}(z)$ is still negligible for all $\varepsilon > 0$ and $k > 1$. Moreover, the one-point correlation function $\rho_1(x)$ of the limiting pfaffian random point process defined in (3)-(9) decays super-exponentially at $+\infty$ and grows proportionally to $|x|^{1/2}$ at $-\infty$. Consequently, if $\xi = \{\xi_j,\ j \in \mathbb{Z}\}$ is a random point configuration of the limiting pfaffian random process then

$$\mathbb{E}\sum_j |\xi_j - z|^{-k} = \int_{-\infty}^{+\infty} |x - z|^{-k}\rho_1(x)dx < \infty \tag{12}$$

for any integer $k \geq 2$. The integral at the r.h.s. of (12) diverges for $k = 1$ which emphasizes the fact that $g_{n,1}(z)$ is not a "local statistic" as the main contribution to $g_{n,1}(z) = n^{-2/3}\times$ $\mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-1}$ comes from the eigenvalues in the bulk. Moreover, it could be shown from the asymptotics of the GOE one-point correlation function that

$$\mathbb{E}\left(\mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-1}\right) = -n + O(n^{2/3}). \tag{13}$$

Eventhough the eigenvalues from the bulk of the spectrum give the main contribution to the mathematical expectation of $\mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-1}$, their joint contribution to the fluctuations of $\mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-1}$ around its mean is much smaller (namely, it can be shown to be of order of constant if one smoothes their contribution by a test function with the support inside $[-2+\delta, 2$
$delta]$.) On the other hand, the largest eigenvalues of $A_n$ give smaller (namely, of the order of $O(n^{2/3})$) contribution to the mean of $\mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-1}$, but they give the main contribution to the fluctuations of $\mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-1}$ around its mean. This suggests to consider

$$g_{n,1}^c(z) = n^{-2/3}\left(n + \mathrm{Tr}G_n(z)\right) = n^{-2/3}\left(n + \mathrm{Tr}(A_n - 2 - zn^{-2/3})^{-1}\right) \tag{14}$$

which is a "local" statistic in a sense that the main contribution to $g_{n,1}^c(z)$ comes from the largest eigenvalues (i.e. the eigenvalues that deviate from the right edge of the spectrum on the order of $O(n^{-2/3})$.)

In Theorem 1.1, we obtain the recursive relations on the joint moments of the local linear statistics $g^c_{n,1}(z)$ and $g_{n,k}(z), \quad k \geq 2$. Let $K$ be a multi-index, $K = (k_1, \dots, k_j), \quad j \geq 1$, with the components $k_l, \ 1 \leq l \leq j$, nonnegative integers. The number of components $j$ is not fixed. We will denote by $m_K$ the corresponding joint moment of $g^c_{n,1}(z)$ and $g_{n,k}(z), \quad k \geq 2$, namely:

$$m_K = \mathbb{E}\left( (g^c_{n,1}(z))^{k_1} \prod_{l=2}^{j} (g_{n,l}(z))^{k_l} \right). \tag{15}$$

Let $e_l$ denote the multi-index with the $l$-th component equal to 1 and the other components equal to zero.

**Theorem 1.1.** *Let $K$ be a non-zero multi-index, then the following equation holds:*

$$\begin{aligned} & m_K(z + O(n^{-2/3})) - m_{K+2e_1}(1 + O(n^{-2/3})) - m_{K+e_2}(1 + O(n^{-2/3})) \\ - \quad & 2\sum_{l\geq 1} l k_l m_{K-e_l+e_{l+2}}(1 + O(n^{-2/3})) = O(n^{-1/3}) m_{K+e_1}. \end{aligned} \tag{16}$$

*Also, the following "boundary" condition holds:*

$$z + O(n^{-2/3}) - m_{2e_1}(1 + O(n^{-2/3})) - m_{e_2}(1 + O(n^{-2/3})) = O(n^{-1/3}) m_{e_1}. \tag{17}$$

**Remark.** We will always assume in (16) that $k_l m_{K-e_l+e_{l+2}} = 0$ if $k_l = 0$.

Theorem 1.1 will be proved in the next section. Let us now consider the Gaussian Unitary Ensemble (GUE) of Hermitian $n \times n$ random matrices.

$$A_n = \frac{1}{\sqrt{n}} \left(a_{jk}\right)^n_{j,k=1}, \tag{18}$$

where $\{\Re a_{jk} = \Re a_{kj}\}_{j<k}$ and $\{\Im a_{jk} = -\Im a_{kj}\}_{j<k}$ are i.i.d. $N(0, 1/2)$ random variables, and $\{a_{ii}\}_{1\leq i\leq n}$ are i.i.d. $N(0, 1)$ random variables.

The global distribution of the eigenvalues of $A_n$ still satisfies the Wigner semi-circle law in the limit $n \to \infty$. The limiting local distribution of the largest eigenvalues of $A_n$ was calculated by Tracy and Widom in [16]. In particular,

$$\lim_{n\to\infty} \mathbb{P}\left(\lambda_{max} \leq 2 + xn^{-2/3}\right) = F_2(x) = \exp\left(-\int_x^\infty (t - x) q^2(t) dt\right), \tag{19}$$

where, as before, $q(x)$ is the solution of the Painléve II differential equation with the same asymptotics at infinity.

Consider the same rescaling at the right edge of the spectrum as in the GOE case, namely

$$\lambda_j^{(n)} = 2 + \xi_j^{(n)} n^{-2/3}, \quad j = 1, 2, \dots, \tag{20}$$

where $\lambda_1^{(n)} \geq \lambda_2^{(n)} \ldots \geq \lambda_n^{(n)}$ are the ordered eigenvalues of $A_n$. It then follows from the results of [16] that the random point configuration $\{\xi_j^{(n)}, \ 1 \leq j \leq n\}$ converges in distribution on the cylinder sets to the random point process on the real line with the $k-$point correlation functions given by

$$\rho_k(x_1, \ldots, x_k) = \det\left(K(x_i, x_j)\right)_{1 \leq i,j \leq k} \tag{21}$$

where $K(x,y) = K_{Airy}(x,y)$ is the Airy kernel defined in (9). The limiting random point process belongs to the class of determinantal random point processes (see [15], [6]).

Let us use the same notations $G_n(z), \ g_{n,k}(z), \ g_{n,1}^c(z)$, and $M_K$ in the GUE case as they were defined in (10), (11), (14), and (15) in the GOE case above. The following analogue of the Theorem 1.1 holds:

**Theorem 1.2.** *Let $K$ be a non-zero multi-index, then the following equation holds:*

$$\begin{aligned} & m_K(z + O(n^{-2/3})) - m_{K+2e_1}(1 + O(n^{-2/3})) \\ - & \sum_{l \geq 1} l k_l m_{K-e_l+e_{l+2}}(1 + O(n^{-2/3})) = O(n^{-1/3}) m_{K+e_1}. \end{aligned} \tag{22}$$

*Also, the following "boundary" condition holds:*

$$z + O(n^{-2/3}) - m_{2e_1}(1 + O(n^{-2/3})) = O(n^{-1/3}) m_{e_1}. \tag{23}$$

We now turn our attention to Wishart (a.k.a Laguerre) ensembles of random matrices. Again, we start with the real case. Let $A = A_{n,N} = \frac{1}{\sqrt{n}}(a_{ij})$ be a rectangular $n \times N$ matrix with $\{a_{ij}, \ 1 \leq i \leq n, \ 1 \leq j \leq N,\}$ real i.i.d. $N(0,1)$ random variables. Let us assume that $N \geq n$, and $N - n = \nu$ is fixed. Consider a nonnegative-definite random matrix

$$M_{n,N} = AA^t. \tag{24}$$

The ensemble of random matrices $M_{n,N}$ is known as the real Wishart distribution in the statistical literature or the Laguerre ensemble in the mathematical physics. The empirical distribution function of the eigenvalues of $M_{n,N}$ converges to the Marchenko-Pastur law as $n \to \infty$ ([8], [11]). The density of the Marchenko-Pastur law is given by

$$\rho_{MP}(x) = \begin{cases} \frac{1}{2\pi\sqrt{x}}\sqrt{4-x} & : \quad \text{if } \ 0 \leq x \leq 4, \\ 0 & : \quad \text{otherwise.} \end{cases} \tag{25}$$

Our goal is to study the distribution of the smallest eigenvalues of $M_{n,N}$ in the limit $n \to \infty, \ N - n = \nu$. It can be shown (see e.g. [4], [5], [3]) that the smallest eigenvalue of $M_{n,N}$ are of the order of $n^{-2}$. Moreover, if we consider the rescaling at the hard edge of the spectrum

$$\lambda_i^{(n,N)} = \frac{\xi_i^{(n,N)}}{4n^2}, \quad 1 \leq i \leq n, \tag{26}$$

one can show that the random point configuration $\{\xi_i^{(n)},\ 1 \le i \le n\}$ converges in distribution on the cylinder sets to the pfaffian random point process on $(0, +\infty)$. The $k$-point correlation functions of the limiting process are of the same form as in (4), where $K(x,y)$ is again a $2 \times 2$ matrix-valued kernel. The formulas for the entries of $K(x,y)$ are similar to (5) with the important difference being that the Airy kernel $K_{Airy}(x,y)$ is replaced by the Bessel kernel

$$K_{Bessel}(x,y) = \frac{J_\nu(\sqrt{x})\sqrt{y}J_{\nu+1}(\sqrt{y}) - J_\nu(\sqrt{y})\sqrt{x}J_{\nu+1}(\sqrt{x})}{2(x-y)}, \tag{27}$$

where $J_\nu(x)$ is the usual Bessel function of index $\nu$.

Define

$$G_{n,N}(t) = \left(M_{n,N} + \frac{t^2}{n^2}\right)^{-1}, \tag{28}$$

where $t$ is a real number. Then

$$g_k(t) = g_{n,N,k}(t) = n^{-2k} Tr G_{n,N}^k(t) = n^{-2k} Tr \left(M_{n,N} + \frac{t^2}{n^2}\right)^{-k} = \sum_1^n (\xi_i^{(n,N)} + t^2)^{-k} \tag{29}$$

is a "local" statistics for any positive integer $k = 1, 2, 3, \ldots$. Indeed, one can show that the eigenvalues from the bulk of the spectrum give vanishing contribution to $g_k(t)$. In particular, if $\xi = \{\xi_i,\ \ i \in \mathbb{N}\}$ is a random point configuration of the limiting pfaffian process, then

$$\mathbb{E} \sum_j (\xi_j + t^2)^{-k} = \int_0^{+\infty} \frac{1}{(x+t^2)^k} \rho_1(x) dx < \infty. \tag{30}$$

We are interested to study the joint moments of the linear statistics $g_k(t),\ k \ge 1$. Let, as before, $K = (k_1, \ldots, k_j),\ j \ge 1$, denote a multi-index, and $m_K$ stand for the corresponding joint moment

$$m_K = \mathbb{E} \prod_{l=1}^j (g_l(t))^{k_l}. \tag{31}$$

The following theorem holds.

**Theorem 1.3.** *Let $K$ be a non-zero multi-index, then the following equation holds:*

$$(\nu - t^{-2} + \frac{1}{n}) m_{K+e_1} + m_{K+e_2} + m_{K+2e_1} + 2\sum_{l=1}^j l k_l m_{K-e_l+e_{l+2}} - \frac{2}{t^2} \sum_{l=1}^j l k_l m_{K-e_l+e_{l+1}}$$
$$= \frac{1}{n}\frac{1}{t^2} m_K. \tag{32}$$

*Also, the following "boundary" condition holds:*

$$(\nu - t^{-2} + \frac{1}{n}) m_{e_1} + m_{e_2} + m_{2e_1} = t^{-2}. \tag{33}$$

We recall that $\nu = N - n \geq 0$ is the difference between the dimensions of the rectangular matrix $A_{n,N}$.

We finish the Introduction by the discussion of the complex Wishart ensemble. Let $A = A_{n,N} = \frac{1}{\sqrt{n}}(a_{ij})$ be a rectangular $n \times N$ matrix with $\{\Re a_{ij}, \Im a_{ij},\ 1 \leq i \leq n,\ 1 \leq j \leq N,\}$ i.i.d. $N(0, 1/2)$ random variables. As before, we us assume that $N \geq n$, and $N - n = \nu$ is fixed. Consider now a nonnegative-definite random matrix

$$M_{n,N} = AA^*. \tag{34}$$

The ensemble of random matrices $M_{n,N}$ is known as the complex Wishart/Laguerre ensemble of random matrices. Consider the rescaling of the eigenvalues at the hard edge of the spectrum

$$\lambda_i^{(n,N)} = \frac{\xi_i^{(n,N)}}{4n^2}, \quad 1 \leq i \leq n, \tag{35}$$

It follows from the results of [4] that the random point configuration $\{\xi_i^{(n)},\ i \geq 1\}$ converges in distribution on the cylinder sets to the determinantal random point process on $(0, +\infty)$ with the correlation kernel given by the Bessel kernel (27) in the limit $n \to \infty$, $N = n + \nu$. Define $G_{n,N}$, $g_k(t)$, $k \geq 1$, and $m_K$ in the same way as in (28), (29), and (31). The following theorem holds.

**Theorem 1.4.** *Let $K$ be a non-zero multi-index, then the following equation holds:*

$$(\nu + \frac{1}{n})m_{K+e_1} + m_{K+2e_1} + \sum_{l=1}^{j} lk_l m_{K-e_l+e_{l+2}} - \frac{1}{t^2}\sum_{l=1}^{j} lk_l m_{K-e_l+e_{l+1}}$$
$$= \frac{1}{n}\frac{1}{t^2}m_K. \tag{36}$$

*Also, the following "boundary" condition holds:*

$$(\nu + \frac{1}{n})m_{e_1} + m_{2e_1} = t^{-2}. \tag{37}$$

We recall that $\nu = N - n \geq 0$ is the difference between the dimensions of the rectangular matrix $A_{n,N}$. Theorems 1.3 and 1.4 will be proved in Section 3.

# 2 Proof of Theorems 1.1 and 1.2

Let us start with the proof of Theorem 1.1. Our first goal is to establish (17). To this end, we consider $n^{1/3}m_{e_1} = \mathbb{E}\left(n^{-1/3}\left(n + TrG_n(z)\right)\right)$, where, as before, $G_n(z) = (A_n - 2 - zn^{-2/3})^{-1}$. By using the resolvent identity

$$G_n(z) = (A_n - 2 - zn^{-2/3})^{-1} = -(2 + zn^{-2/3})^{-1}Id + (2 + zn^{-2/3})^{-1}A_n(A_n - 2 - zn^{-2/3})^{-1} =$$
$$-(2 + zn^{-2/3})^{-1}Id + (2 + zn^{-2/3})^{-1}A_nG_n, \tag{38}$$

we arrive at

$$n^{1/3}m_{e_1} = n^{2/3} - (2+zn^{-2/3})^{-1}n^{2/3} + (2+zn^{-2/3})^{-1}n^{-1/3}\mathbb{E}\sum_{ij} A_{ij}G_{ji}. \tag{39}$$

Here $A_{ij} = \frac{a_{ij}}{\sqrt{n}}$ denote the matrix entries of $A_n$, and $G_{ij}$ denote the matrix entries of $G_n(z)$. To calculate $\mathbb{E}A_{ij}G_{ji}$, we recall that random variables $A_{ij}$, $1 \le i \le j \le n$, are independent. Therefore, we can first fix all matrix entries (up from the diagonal) except $A_{ij}$ and integrate with respect to $A_{ij}$. Applying the Gaussian decoupling formula

$$\mathbb{E}\eta f(\eta) = \sigma^2 \mathbb{E}f'(\eta), \quad \eta \sim N(0,\sigma^2), \tag{40}$$

with $\eta = A_{ij}$ and $f(\eta) = G_{ij}$, and taking into account that $Var(A_{ij}) = \frac{1+\delta_{ij}}{n}$, and

$$\frac{\partial G_{kl}}{\partial A_{ij}} = \begin{cases} -G_{ki}G_{jl} - G_{kj}G_{il} & : \quad i \ne j \\ -G_{ki}G_{jl} & : \quad i = j, \end{cases} \tag{41}$$

we arrive at

$$n^{1/3}m_{e_1} = n^{2/3} - (2+zn^{-2/3})^{-1}n^{2/3} - (2+zn^{-2/3})^{-1}n^{-4/3}\mathbb{E}\sum_{ij}\left(G_{ji}G_{ji} + G_{ii}G_{jj}\right). \tag{42}$$

The term $n^{-4/3}\mathbb{E}\sum_{ij} G_{ji}G_{ji}$ is equal to

$$n^{-4/3}\mathbb{E}\sum_{ij} G_{ji}G_{ji} = \mathbb{E}n^{-4/3}Tr(G_n^2(z)) = m_{e_2}. \tag{43}$$

To deal with the term $n^{-4/3}\mathbb{E}\sum_{ij} G_{ii}G_{jj}$, we rewrite it as

$$n^{-4/3}\mathbb{E}\sum_{ij} G_{ii}G_{jj} = n^{-4/3}\mathbb{E}\left((TrG_n(z))^2\right) = n^{-4/3}\mathbb{E}\left((-n+n+TrG_n(z))^2\right) \tag{44}$$
$$= \quad n^{2/3} - 2n^{-1/3}\mathbb{E}(n+TrG_n(z)) + n^{-4/3}\mathbb{E}\left((n+TrG_n(z))^2\right) = n^{2/3} - 2n^{1/3}m_{e_1} + m_{2e_1}.$$

As a result, we obtain

$$n^{1/3}m_{e_1} = n^{2/3} - (2+zn^{-2/3})^{-1}n^{2/3} - (2+zn^{-2/3})^{-1}m_{e_2} - (2+zn^{-2/3})^{-1}n^{2/3}$$
$$+2(2+zn^{-2/3})^{-1}n^{1/3}m_{e_1} - (2+zn^{-2/3})^{-1}m_{2e_1}, \tag{45}$$

which is equivalent to

$$n^{2/3}(1-(1+zn^{-2/3}/2)^{-1}) - m_{2e_1}(2+zn^{-2/3})^{-1} - m_{e_2}(2+zn^{-2/3})^{-1}$$
$$= m_{e_1}n^{1/3}(1-(1+zn^{-2/3}/2)^{-1}). \tag{46}$$

After trivial arithmetical simplifications, this leads to (17). The formula (16) can be proven along the same lines if one starts with $n^{1/3}m_{K+e_1}$. One can say that the formula (17) gives us the boundary term in the recursive system of linear equations satisfied by $\{m_K\}$ since it corresponds to $K=0$. Turning our attention to (16), we write

$$n^{1/3}m_{K+e_1} = n^{1/3}\mathbb{E}\left( (g^c_{n,1}(z))^{k_1+1}\prod_{l=2}^{j}(g_{n,k}(z))^{k_l}\right) =$$
$$\mathbb{E}\left[ n^{-1/3}(n+TrG_n)\left(n^{-2/3}(n+TrG_n)\right)^{k_1}\prod_{l\geq 2}\left(n^{-2l/3}TrG^l\right)^{k_l}\right], \tag{47}$$

we then rewrite, as before, the first term $n^{-1/3}(n+TrG_n)$ as

$$n^{-1/3}(n+TrG_n) = n^{2/3} - (2+zn^{-2/3})^{-1}n^{2/3} + (2+zn^{-2/3})^{-1}n^{-1/3}\sum_{ij}A_{ij}G_{ji}. \tag{48}$$

This leads to

$$n^{1/3}m_{K+e_1} = n^{2/3}m_K - (2+zn^{-2/3})^{-1}n^{2/3}m_K +$$
$$(2+zn^{-2/3})^{-1}n^{-1/3}\mathbb{E}\left[ (\sum_{ij}A_{ij}G_{ji})\left(n^{-2/3}(n+TrG_n)\right)^{k_1}\prod_{l\geq 2}\left(n^{-2l/3}TrG^l\right)^{k_l}\right]. \tag{49}$$

As in the case $K=0$ considered above, we fix all matrix entries (up from the diagonal) except $A_{ij}$, and apply (40) with $\eta = A_{ij}$ and $f(\eta) = G_{ji}\left(n^{-2/3}(n+TrG_n)\right)^{k_1}\prod_{l\geq 2}\left(n^{-2l/3}TrG^l\right)^{k_l}$. Taking into account (41) and the equation

$$\frac{\partial Tr(G^l)}{\partial A_{ij}} = -2l(G^{l+1})_{ij}, \tag{50}$$

one then obtains (16) after some simple algebraic calculations. Theorem 1.1 is proven.

The proof of Theorem 1.2 is quite similar. The only alteration required in the GUE case is that one needs to replace (41) with

$$\frac{\partial G_{kl}}{\partial Re(A_{ij})} = \begin{cases} -G_{ki}G_{jl} - G_{kj}G_{il} & : \quad i\neq j \\ -G_{ki}G_{jl} & : \quad i=j, \end{cases} \tag{51}$$

and

$$\frac{\partial G_{kl}}{\partial Im(A_{ij})} = -i(G_{ki}G_{jl} - G_{kj}G_{il}) \ \text{ for } i\neq j. \tag{52}$$

The remaining calculations are very similar and are left to the reader.

## 3 Proof of Theorems 1.3 and 1.4

The proofs will be similar to the ones given in the previous section. Let us start with the proof of Theorem 1.3. Our first goal is to establish (33). To this end, we consider $n^{-1} \times m_{e_1} = \mathbb{E}\left(n^{-3} Tr G_{n,N}(t)\right)$, where $G_{n,N}(t)$ was defined in (28). By using the resolvent identity

$$G_{n,N} = (M_{n,N} + \frac{t^2}{n^2})^{-1} = \frac{n^2}{t^2} Id - \frac{n^2}{t^2} AA^t G_{n,N} \tag{53}$$

we arrive at

$$n^{-1} m_{e_1} = \frac{1}{t^2} - \frac{1}{n}\frac{1}{t^2}\mathbb{E}\sum_{1\le i,j\le n}\sum_{1\le p\le N} A_{ip}A_{jp}G_{ji}. \tag{54}$$

Here $A_{ip} = \frac{a_{ip}}{\sqrt{n}}$ denote the matrix entries of $A_n$, and $G_{ji}$ denote the matrix entries of $G_{n,N}(z)$. To calculate $\mathbb{E}A_{ip}A_{jp}G_{ji}$, we again use the Gaussian decoupling formula (40) and the equation

$$\frac{\partial G_{kl}}{\partial A_{ip}} = -G_{ki}(A^t G)_{pl} - (GA)_{kp} G_{il}. \tag{55}$$

Therefore,

$$\begin{aligned} n^{-1} m_{e_1} &= \frac{1}{t^2} + \frac{1}{n^2}\frac{1}{t^2}\mathbb{E}\sum_{1\le i,j\le n}\sum_{1\le p\le N} A_{jp}\left(G_{ji}(GA)_{ip} + G_{ii}(GA)_{jp}\right) - \frac{1}{n^2}\frac{1}{t^2}\mathbb{E}\sum_{1\le i\le n}\sum_{1\le p\le N} G_{ii} \\ &= \frac{1}{t^2} + \frac{1}{n^2}\frac{1}{t^2}\mathbb{E}\left[Tr(GAA^tG) + Tr(GAA^t)TrG\right] - \frac{N}{n^2}\frac{1}{t^2}\mathbb{E}TrG. \end{aligned} \tag{56}$$

Using the identity $GAA^t = Id - \frac{t^2}{n^2}G$, the last formula can rewritten as

$$n^{-1} m_{e_1} = \frac{1}{t^2} + \frac{1}{t^2} m_{e_1} - m_{e_2} + \frac{n}{t^2} m_{e_1} - m_{2e_1} - \frac{n+\nu}{t^2} m_{e_1}, \tag{57}$$

which implies (33). The formula (32) can be proven along the same lines if one starts with $n^{-1} m_{K+e_1}$. Let us write

$$\begin{aligned} n^{-1} m_{K+e_1} = n^{-1}\mathbb{E}\left((g_1(t))\prod_{l=1}^{j}(g_l(t))^{k_l}\right) = \\ \mathbb{E}\left[n^{-3} Tr G_{n,N})\prod_{l\ge 1}\left(n^{-2l} Tr G_{n,N}^l\right)^{k_l}\right], \end{aligned} \tag{58}$$

Using the resolvent identity, we can rewrite the first term in the product as

$$G_{n,N} = \frac{n^2}{t^2} Id - \frac{n^2}{t^2} AA^t G_{n,N}.$$

After integration by parts and a few lines of careful calculations, we obtain (32). Theorem 1.4 can be proven along similar lines.

## 4 Non-Gaussian Case.

The generalization of the Gaussian decoupling formula (40) to the non-Gaussian case can be found, for example, in [7]:

$$\mathbb{E}[\xi f(\xi)] = \sum_{k=0}^{p} \frac{c_{k+1}}{k!} \mathbb{E}\left[\frac{d^k f}{dx^k}(\xi)\right] + \varepsilon, \tag{59}$$

where $\xi$ is a real random variable such that $\mathbb{E}\left(|\xi|^{p+2}\right) < \infty$, $c_l,\ l \geq 1$, are the cumulants of the random variable $\xi$, complex-valued function $f(x)$ has first $p+1$ continuous and bounded derivatives, and the error term satisfies the upper bound $|\varepsilon| \leq B_k sup_x |\frac{d^k f}{dx^k}(x)| \mathbb{E}\left(|\xi|^{p+2}\right)$ with the constant $B_k$ depending only on $k$.

It is conjectured that the distribution of the largest eigenvalues in Wigner random matrices is universal provided the fourth moment of the matrix entries is finite. Currently, we are unable to prove this conjecture. Instead, we speculate below on the possible approach to extend the results of Theorems 1.1-1.4 to the non-Gaussian case. Let us consider a real Wigner random matrix $A_n = \frac{1}{\sqrt{n}}(a_{ij})_{i,j=1}^n$, and assume that the entries $(a_{ij} = a_{ji})_{i<j}$ are i.i.d. centralized random variables with the unit variance and the finite fourth moment. In addition, we assume that the diagonal entries $a_{ii},\ 1 \leq i \leq n$, are i.i.d. centralized random variables, independent from the non-diagonal entries. We assume that the diagonal entries also have the finite fourth moment. Let us also assume for simplicity that $Var(a_{ii}) = 2$. In an attempt to extend the result of Theorem 1.1 to the non-Gaussian situation, we apply the generalized decoupling formula (59). To be specific, let us concentrate our attention on the "boundary" equation (17). Looking at (39), we apply (59) to $\mathbb{E}A_{ij}G_{ji}$. Since $c_1(a_{ij}) = 0,\ c_2(a_{ij}) = 1 + \delta_{ij},\ c_3(a_{ij}) = c_3$, for $i < j$, and $\mathbb{E}a_{ij}^4 \leq \infty$, one might wish to truncate (59) after the first three terms (i.e. $p = 2$) to obtain

$$n^{-1/3}\mathbb{E}\sum_{ij} A_{ij}G_{ji} = -n^{-4/3}\mathbb{E}\sum_{ij}(G_{ji}^2 + G_{ii}G_{jj}) + n^{-11/6}\mathbb{E}\sum_{ij}(3G_{ij}G_{ii}G_{jj} + G_{ij}^3) + \epsilon. \tag{60}$$

The first sum in (60) is the same as in the Gaussian case. The hope is to show that the second sum and the remainder term give negligible contributions in the limit $n \to \infty$. However, it is currently unclear to us how to efficiently bound the terms $n^{-11/6}\mathbb{E}\sum_{ij} G_{ij}G_{ii}G_{jj}$ and $n^{-11/6} \times \mathbb{E}\sum_{ij} G_{ij}^3$.

## References


[1] L. Arnold, *On Wigner's semicircle law for eigenvalues of random matrices.* J. Math. Anal. Appl. **20**, (1967), 262–268.

[2] G. Ben Arous, S. Péché, *Universality of local eigenvalue statistics for some sample covariance matrices.* Commun. Pure Appl. Math. **LVIII**, (2005), 1316–1357.

[3] A. Edelman, *The distribution and moments of the smallest eigenvalue of a random matrix of Wishart type,* Linear Algebra Appl. **159**, (1991), 55–80.

[4] P. Forrester, *The spectrum edge of random matrix ensembles,* Nuclear Phys. B **402**, (1993), 709–728.

[5] P. Forrester, *Exact results and universal asymptotics in the Laguerre random matrix ensemble,* J. Math. Phys. **35**, (1994), 2539–2553.

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

[7] A.M.Khorunzhy, B.A. Khoruzhenko, and L.A. Pastur, *Asymptotic properties of large random matrices with independent entries.* J. Math. Phys. **37**, (1996), 5033–5059.

[8] V. Marchenko, L. Pastur, *Distribution of eigenvalues of some sets of random matrices,* Math USSR-Sb. **1**, (1967), 457–486.

[9] M.L. Mehta, *Random Matrices, 3rd ed.* Elsevier/Academic Press, Amsterdam, 2004.

[10] S. Péché and A. Soshnikov, *Wigner random matrices with non symmetrically distributed entries.* J. Stat. Phys. **129** (2007), 857–883.

[11] J. Silverstein, Z. Bai, *On the empirical distribution of eigenvalues of a class of large-dimensional random matrices.* J. Multivariate Anal. **54** (1995), 175–192.

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

[13] A. Soshnikov, *A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices.* J. Stat. Phys. **108** (2002), 1033–1056.

[14] A. Soshnikov, *Janossy densities II. Pfaffian ensembles.* J. Stat. Phys. **113** (2003), 611–622.

[15] A. Soshnikov, *Determinantal random point fields.* Russ. Math. Surv. **55** (2000), 923–975.

[16] C. Tracy and H. Widom, *Level-spacing distribution and Airy kernel.* Commun. Math. Phys. **159** (1994), 151–174.

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

[18] C. Tracy and H. Widom, *Matrix kernels for the Gaussian orthogonal and symplectic ensembles.* Ann. Inst. Fourier **55** (2005), 2197–2207.

[19] C. Tracy and H. Widom, *Levl-spacing distribution and the Bessel kernel.* Commun. Math. Phys. **161** (1994), 289–309.

[20] E.Wigner, *Characteristic vectors of bordered matrices with infinite dimenisons.* Annals of Math. **62** (1955), 548–564.

[21] E.Wigner, *On the distribution of the roots of certain symmetric matrices.* Annals of Math. **68** (1958), 325–328.