Max-Planck-Institut für Mathematik Bonn

Large gaps of CUE and GUE

by

Renjie Feng
Dongyi Wei

Max-Planck-Institut für Mathematik
Preprint Series 2022 (72)

Date of submission: November 10, 2022

Large gaps of CUE and GUE

by

Renjie Feng
Dongyi Wei

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

School of Mathematical Sciences
Peking University
Beijing 100871
China

MPIM 22-72

LARGE GAPS OF CUE AND GUE

RENJIE FENG AND DONGYI WEI

Abstract. In this article, we will study the largest gaps of the classical random matrices of CUE and GUE. The main result is that the rescaling largest gaps will converge to a Poisson point process, and the limiting densities are given by the Gumbel distributions.

1. Introduction

In random matrix theory, the typical spacings between eigenvalues of classical random matrices have been well understood for a long time [1, 4]. But there are only few results known for the extremal spacings. The rescaling limits of the smallest gaps of CUE and GUE (where the point processes of eigenvalues are both determinintal point processes) were proved by Vinson and he also suggest the decay order of the largest gap [11]. Later on, Soshnikov studied the smallest gaps for the general determinantal point processes with translation invariant kernels [9], and proved that the point processes of the smallest gaps after rescaling are asymptotic to the Poisson distributions. In [2], Ben Arous-Bourgade adapted Soshnikov's technique and reproved the smallest gaps for CUE and GUE, and they further proved the decay order of the largest gap for these two ensembles and confirmed Vinson's prediction. The proofs in [2, 9, 11] highly depend on the determinantal structures of the point processes. For the point processes without determinantal structures, in [5], we developed a new technique based on the Selberg integral to prove the smallest gaps for the circular log-gas β-ensemble for any positive integer β. As special cases, our result implies the limiting distributions of the smallest gaps of the classical random matrices of COE, CUE and CSE. The same technique can be applied to GOE which has Pfaffian structures [6].

In this paper, we will study the rescaling limits of the largest gaps of CUE and GUE. We will prove that the point processes of the rescaling largest gaps in both cases are asymptotic to the Poisson point processes with some explicitly given intensities. As a direct consequence, we can derive the laws of the rescaling limits of the k-th largest gap, which are given by the Gumbel distributions. Our results are further proved to be universal [8].

To state our results, let's first consider CUE. Let u_n be a Haar-distributed unitary matrix $U(n)$ over $\mathbb{C}^n$. Suppose u_n has eigenvalues $e^{i\theta_k}$'s with ordered eigenangles $0 < \theta_1 < \cdots < \theta_n < 2\pi$. Let $m_1 > m_2 > \cdots$ be the largest gaps between successive eigenangles of u_n i.e., m_k $(1 \le k \le n)$ is the decreasing rearrangement of $\theta_{k+1} - \theta_k$ $(1 \le k \le n)$ with $\theta_{k+n} = \theta_k + 2\pi$. Ben Arous-Bourgade showed that

Date: November 13, 2022.

for any $p > 0$ and $l_n = n^{o(1)}$, one has [2]

$$\frac{nm_{l_n}}{\sqrt{32\ln n}} \xrightarrow{L^p} 1.$$

In this article, we will give the rescaling limit law for m_k as follows.

Theorem 1. *Let's denote m_k as the k-th largest gap of CUE, and*

$$\tau_k = (2\ln n)^{\frac{1}{2}}(nm_k - (32\ln n)^{\frac{1}{2}})/4 - (3/8)\ln(2\ln n),$$

then $\{\tau_k\}$ will converge to a Poisson point process as $n \to +\infty$ where for any bounded interval $I \subset \mathbb{R}$, the limiting density is given by the Gumbel distribution,

$$\lim_{n\to+\infty} \mathbb{P}(\tau_k \in I) = \int_I \frac{e^{k(c_1-x)}}{(k-1)!} e^{-e^{c_1-x}} dx.$$

Here, $c_1 = c_0 + \ln(\pi/2)$ where $c_0 = \frac{1}{12}\ln 2 + 3\zeta'(-1)$ is the constant in the expansion (6) and $\zeta(x)$ is the Riemann zeta function. In particular, the limiting density for the largest gap τ_1 is,

$$e^{c_1-x}e^{-e^{c_1-x}}.$$

Let's sketch the main ideas to prove Theorem 1. First, by the uniform asymptotic expansion (6) of the gap probability for a given arc of the circle to be free of eigenvalues, we can find the correct rescaling formula for the largest gap m_k and our crucial observation is the rescaling limit (11) in Lemma 2. Then the rest main task is to prove that the point process of the rescaling largest gaps is asymptotic to the Poisson point process as $n \to +\infty$, and hence a Gumbel distribution will be derived. In order to do this, we will first prove Lemma 1 as a criterion for a sequence of (decreasing) point processes on the real line converging to the Poisson point processes. Lemma 1 implies that Theorem 1 is proved by the upper bound (17) and the lower bound (18). The upper bound can be proved by the negatively associated property of the determinantal point processes. The lower bound is the most essential part of the whole proof, which is based on the asymptotic splitting formula (24) for the gap probabilities in Lemma 5. The proof of Lemma 5 is further based on Lemma 6 and Lemma 7 for the eigenvalue estimates of some symmetric operators.

For GUE, the joint density of the eigenvalues is

$$\frac{1}{Z_n} e^{-n\sum_{i=1}^{n}\lambda_i^2/2} \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^2 \tag{1}$$

with respect to the Lebesgue product measure on the simplex $\lambda_1 < \cdots < \lambda_n$. And the empirical spectral distribution converges in probability to the semicircle law [1]

$$\rho_{sc}(x) = \sqrt{(4-x^2)_+}/(2\pi),$$

where we denote $f_+ := \max(f, 0)$.

For largest gaps of GUE, the result is completely different inside the bulk and on the edge of the semicircle law. On the edge, the largest gap is of order $n^{-2/3}$ which is indicated by the Tracy-Widom law [1]; while inside the bulk, the largest gap is of order $\sqrt{\log n}/n$ [2, 11]. To be more precise, given $I = [a, b]$ which is a compact subinterval of $(-2, 2)$, let $m_1^* > m_2^* > \cdots$ be the largest gaps of type $\lambda_{i+1} - \lambda_i$

with $\lambda_{i+1}, \lambda_i \in I$, then Ben Arous-Bourgade [2] showed that for any $p > 0$ and $l_n = n^{o(1)}$,

$$\left(\inf_I \sqrt{4-x^2}\right) \frac{nm^*_{l_n}}{\sqrt{32 \ln n}} \xrightarrow{L^p} 1.$$

Regarding the GUE case, we have

Theorem 2. *Given $I = [a, b]$ which is a compact subinterval of $(-2, 2)$, let m^*_k be the k-th largest gap of GUE falling in I, we denote $S(I) = \inf_I \sqrt{4-x^2}$ and*

$$\tau^*_k = (2 \ln n)^{\frac{1}{2}} (nS(I)m^*_k - (32 \ln n)^{\frac{1}{2}})/4 + (5/8) \ln(2 \ln n),$$

then $\{\tau^*_k\}$ will converge to a Poisson point process as $n \to +\infty$ where for any bounded interval $I_1 \subset \mathbb{R}$, the limiting density is given by the Gumbel distribution,

$$\lim_{n\to+\infty} \mathbb{P}(\tau^*_k \in I_1) = \int_{I_1} \frac{e^{k(c_2-x)}}{(k-1)!} e^{-e^{c_2-x}} dx.$$

Here, the constant $c_2 = c_0 + M_0(I)$ depending on I, where $c_0 = \frac{1}{12} \ln 2 + 3\zeta'(-1)$ and $M_0(I) = (3/2) \ln(4-a^2) - \ln(4|a|)$ if $a + b < 0$, $M_0(I) = (3/2) \ln(4-b^2) - \ln(4|b|)$ if $a + b > 0$ and $M_0(I) = (3/2) \ln(4-a^2) - \ln(2|a|)$ if $a + b = 0$. In particular, the limiting density for the largest gap τ^*_1 is,

$$e^{c_2-x} e^{-e^{c_2-x}}.$$

Note that the constant $M(I)$ depends on whether or not I is a symmetric subinterval about origin, this is because the semicircle law is symmetric about the origin.

The starting point to prove Theorem 2 is the observation (38) in Lemma 8, which is another rescaling limit regarding the gap probability for CUE. Such rescaling limit about CUE will play an important role in the proof of the largest gaps of GUE. We still need to prove that the point process of the rescaling largest gaps tends to the Poisson point process as in Theorem 1, where we need to prove the upper bound (45) and lower bound (46). And another key ingredient to prove the GUE case is the comparisons regarding the kernels and the Fredholm determinants between CUE and GUE in the proofs of Lemma 12 and Lemma 14.

2. A criterion for the Poisson convergence

We first prove the following general criterion for a sequence of (decreasing) point processes on the real line converging to Poisson point processes.

Lemma 1. *Let $\chi^{(n)} = \sum_{k=1}^{k_n} \delta_{\tau^{(n)}_k}$ be a sequence of point processes on $\mathbb{R}$ such that the sequence $\tau^{(n)}_k$ $(1 \le k \le k_n)$ is decreasing for every fixed n, $f \in C^2(\mathbb{R})$ satisfies $f(x) > 0,\ f'(x) < 0,\ f''(x) > 0$ for $x \in \mathbb{R}$ and $\lim_{x\to+\infty} f'(x) = 0$. Assume that for every positive integer k and $x_1, \cdots, x_k \in \mathbb{R}$, we have*

$$\lim_{n\to+\infty} \mathbb{E} \sum_{i_1,\cdots,i_k \ \text{all distinct}} \prod_{j=1}^{k} (\tau^{(n)}_{i_j} - x_j)_+ = \prod_{j=1}^{k} f(x_j). \tag{2}$$

Then for $A = (x, +\infty)$ or $A = [x, +\infty)$, we have the convergence

$$\chi^{(n)}(A) \xrightarrow{law} \chi(A), \tag{3}$$

where $\chi(A)$ is a Poisson random variable with mean $-f'(x)$. Furthermore, for any bounded interval $I\subset\mathbb{R}$, we have the limiting distribution,

$$\lim_{n\to+\infty}\mathbb{P}(\tau_k^{(n)}\in I)=\int_I\frac{f''(x)(-f'(x))^{k-1}}{(k-1)!}e^{f'(x)}dx.$$

Here, we denote $\tau_k^{(n)}=-\infty$ for $k>k_n$.

Proof. For $a<b,\ x\in\mathbb{R}$, we simply have

$$(b-a)\chi_{\{x\geq b\}}\leq(x-a)_+-(x-b)_+\leq(b-a)\chi_{\{x>a\}},$$

then for $a_1<a_{-1}$, we have

$$(a_{-1}-a_1)^k\prod_{j=1}^k\chi_{\{\tau_{i_j}^{(n)}\geq a_{-1}\}}\leq\prod_{j=1}^k((\tau_{i_j}^{(n)}-a_1)_+-(\tau_{i_j}^{(n)}-a_{-1})_+)$$
$$=\sum_{\varepsilon_1,\cdots,\varepsilon_k\in\{\pm1\}}\prod_{j=1}^k\varepsilon_j(\tau_{i_j}^{(n)}-a_{\varepsilon_j})_+\leq(a_{-1}-a_1)^k\prod_{j=1}^k\chi_{\{\tau_{i_j}^{(n)}>a_1\}}.$$

We denote

$$\rho^{(n,k)}=\sum_{i_1,\cdots,i_k\text{ all distinct}}\delta_{\tau_{i_1}^{(n)},\cdots,\tau_{i_k}^{(n)}},$$

then we have

$$\rho^{(n,k)}(A^k)=\frac{(\chi^{(n)}(A))!}{(\chi^{(n)}(A)-k)!}$$

for every interval $A\subset\mathbb{R}$. By taking summation over distinct points, we have

$$(a_{-1}-a_1)^k\rho^{(n,k)}([a_{-1},+\infty)^k)$$
$$\leq\sum_{i_1,\cdots,i_k\text{ all distinct}}\sum_{\varepsilon_1,\cdots,\varepsilon_k\in\{\pm1\}}\prod_{j=1}^k\varepsilon_j(\tau_{i_j}^{(n)}-a_{\varepsilon_j})_+$$
$$\leq(a_{-1}-a_1)^k\rho^{(n,k)}((a_1,+\infty)^k).$$

Using (2), taking expectation and the limit, we have

$$(a_{-1}-a_1)^k\limsup_{n\to+\infty}\mathbb{E}\rho^{(n,k)}([a_{-1},+\infty)^k)$$
$$\leq\sum_{\varepsilon_1,\cdots,\varepsilon_k\in\{\pm1\}}\lim_{n\to+\infty}\mathbb{E}\sum_{i_1,\cdots,i_k\text{ all distinct}}\prod_{j=1}^k\varepsilon_j(\tau_{i_j}^{(n)}-a_{\varepsilon_j})_+$$
$$=\sum_{\varepsilon_1,\cdots,\varepsilon_k\in\{\pm1\}}\prod_{j=1}^k\big(\varepsilon_jf(a_{\varepsilon_j})\big)=\big(f(a_1)-f(a_{-1})\big)^k$$
$$\leq(a_{-1}-a_1)^k\liminf_{n\to+\infty}\mathbb{E}\rho^{(n,k)}((a_1,+\infty)^k).$$

For every $x\in\mathbb{R}$ and $\delta>0$, taking $(a_1,a_{-1})=(x,x+\delta)$ and $(a_1,a_{-1})=(x-\delta,x)$, we will have

$$((f(x)-f(x+\delta))/\delta)^k\leq\liminf_{n\to+\infty}\mathbb{E}\rho^{(n,k)}((x,+\infty)^k)$$
$$\leq\limsup_{n\to+\infty}\rho^{(n,k)}([x,+\infty)^k)\leq((f(x-\delta)-f(x))/\delta)^k.$$

Letting $\delta \to 0+$ and using $\rho^{(n,k)}((x,+\infty)^k) \le \rho^{(n,k)}([x,+\infty)^k)$, we have the following convergence of the factorial moments,

$$\lim_{n\to+\infty} \mathbb{E}\frac{(\chi^{(n)}(A))!}{(\chi^{(n)}(A)-k)!} = \lim_{n\to+\infty} \mathbb{E}\rho^{(n,k)}(A^k) = (-f'(x))^k,$$

where $A=(x,+\infty)$ or $A=[x,+\infty)$, which implies the convergence of (3).

Now for every $k \ge 0,\ k\in\mathbb{Z}$, we have

$$\lim_{n\to+\infty} \mathbb{P}(\chi^{(n)}(A)=k) = \mathbb{P}(\chi(A)=k) = (-f'(x))^k\, e^{f'(x)}/k!.$$

Therefore, for $A=(x,+\infty)$ or $A=[x,+\infty)$, we have

$$\lim_{n\to+\infty} \mathbb{P}(\tau_k^{(n)}\in A) = \lim_{n\to+\infty} \mathbb{P}(\chi^{(n)}(A)\ge k) = \mathbb{P}(\chi(A)\ge k) = \varphi_k\left(-f'(x)\right), \tag{4}$$

where

$$\varphi_k\left(\lambda\right) = 1-\sum_{j=0}^{k-1}\frac{\lambda^j}{j!}e^{-\lambda},$$

thus

$$\varphi_k(0)=0,\ \varphi_k'\left(\lambda\right) = -\sum_{j=1}^{k-1}\frac{\lambda^{j-1}}{(j-1)!}e^{-\lambda}+\sum_{j=0}^{k-1}\frac{\lambda^j}{j!}e^{-\lambda} = \frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}$$

and

$$\varphi_k\left(\lambda\right) = \int_0^\lambda \varphi_k'\left(s\right)ds = \int_0^\lambda \frac{s^{k-1}}{(k-1)!}e^{-s}ds.$$

Changing variables $s=-f'(x)$, we have

$$\varphi_k\left(-f'(a)\right) = \int_0^{-f'(a)}\frac{s^{k-1}}{(k-1)!}e^{-s}ds = \int_a^{+\infty}\frac{f''(x)(-f'(x))^{k-1}}{(k-1)!}e^{f'(x)}dx \tag{5}$$

for every $a\in\mathbb{R}$. Now for any bounded interval $I\subset\mathbb{R}$, we can write $I=(a,b)$ or $I=(a,b]$ or $I=[a,b)$ or $I=[a,b]$ where $a<b$, thus $I=A_1\setminus A_2$ with $A_1=(a,+\infty)$ or $A_1=[a,+\infty)$ and $A_2=(b,+\infty)$ or $A_2=[b,+\infty)$, and by (4) and (5) we have

$$\begin{aligned}\lim_{n\to+\infty} \mathbb{P}(\tau_k^{(n)}\in I) &= \lim_{n\to+\infty} \mathbb{P}(\tau_k^{(n)}\in A_1) - \lim_{n\to+\infty} \mathbb{P}(\tau_k^{(n)}\in A_2)\\ &= \varphi_k\left(-f'(a)\right)-\varphi_k\left(-f'(b)\right) = \int_a^b \frac{f''(x)(-f'(x))^{k-1}}{(k-1)!}e^{f'(x)}dx.\end{aligned}$$

This completes the proof. □

3. The CUE case

3.1. A rescaling limit. For CUE, the gap probability of having no eigenvalue in an arc of size 2α is equal to the Toeplitz determinant

$$D_n(\alpha) = \det_{1\le j,k\le n}\left(\frac{1}{2\pi}\int_\alpha^{2\pi-\alpha} e^{i(j-k)\theta}d\theta\right).$$

All the asymptotics we need are direct consequences of the precise analysis of $D_n(\alpha)$ given by Deift et al. [3]. More precisely they proved that for some sufficiently large s_0 and any $\varepsilon>0$, uniformly in $s_0/n<\alpha<\pi-\varepsilon$, one has

$$\ln D_n(\alpha) = n^2\ln\cos\frac{\alpha}{2} - \frac{1}{4}\ln\left(n\sin\frac{\alpha}{2}\right)+c_0+O\left(\frac{1}{n\sin(\alpha/2)}\right), \tag{6}$$

here $c_0 = \frac{1}{12}\ln 2 + 3\zeta'(-1)$ where $\zeta(x)$ is the Riemann zeta function.

We denote

$$F_n(x) = \frac{8x + 3\ln(2\ln n)}{2n(2\ln n)^{\frac{1}{2}}} + \frac{(32\ln n)^{\frac{1}{2}}}{n}, \tag{7}$$

then we have

$$m_k = F_n(\tau_k),$$

where m_k and τ_k are as defined in Theorem 1.

From the definition of $F_n(x)$, we have

$$\tau_k - x = (F_n(\tau_k) - F_n(x))(n/4)(2\ln n)^{\frac{1}{2}} = (m_k - F_n(x))(n/4)(2\ln n)^{\frac{1}{2}}, \tag{8}$$

and for every fixed x, we have

$$\lim_{n\to+\infty} \frac{nF_n(x)}{(32\ln n)^{\frac{1}{2}}} = 1, \ \lim_{n\to+\infty} nF_n(x) = +\infty, \ \lim_{n\to+\infty} n^{\gamma} F_n(x) = 0, \ \forall\, \gamma < 1. \tag{9}$$

For every fixed $\alpha \in (0, \pi)$, by (6) we have

$$\lim_{n\to+\infty} (n/4)(2\ln n)^{\frac{1}{2}} D_n(\alpha) = 0. \tag{10}$$

Another important consequence of (6) is the following rescaling limit

Lemma 2.

$$\lim_{n\to+\infty} n(2\ln n)^{\frac{1}{2}} D_n(F_n(x)/2) = e^{c_0 - x}. \tag{11}$$

Proof. Let $\alpha_n = F_n(x)/2$, then by (9) we have $\alpha_n \to 0$, $n\alpha_n \to +\infty$ as $n \to +\infty$, thus $s_0/n < \alpha_n < \pi - \varepsilon$ for n sufficiently large, and

$$\lim_{n\to+\infty} \frac{1}{n\sin(\alpha_n/2)} = \lim_{n\to+\infty} \frac{2}{n\alpha_n} \lim_{n\to+\infty} \frac{\alpha_n/2}{\sin(\alpha_n/2)} = 0. \tag{12}$$

Thus, by (6) we have

$$\lim_{n\to+\infty} \left(\ln D_n(\alpha_n) - n^2 \ln\cos\frac{\alpha_n}{2} + \frac{1}{4}\ln\left(n\sin\frac{\alpha_n}{2}\right) - c_0 \right) = 0. \tag{13}$$

By (9) we have

$$\begin{aligned}\lim_{n\to+\infty} \frac{(2\ln n)^{\frac{1}{2}}}{n\sin(\alpha_n/2)} &= \lim_{n\to+\infty} \frac{(2\ln n)^{\frac{1}{2}}}{n\alpha_n/2} \lim_{n\to+\infty} \frac{\alpha_n/2}{\sin(\alpha_n/2)} = \lim_{n\to+\infty} \frac{(2\ln n)^{\frac{1}{2}}}{n\alpha_n/2} \\ &= \lim_{n\to+\infty} \frac{(2\ln n)^{\frac{1}{2}}}{nF_n(x)/4} = \lim_{n\to+\infty} \frac{(32\ln n)^{\frac{1}{2}}}{nF_n(x)} = 1,\end{aligned}$$

and thus we have

$$\lim_{n\to+\infty} \left(\frac{1}{8}\ln(2\ln n) - \frac{1}{4}\ln\left(n\sin\frac{\alpha_n}{2}\right) \right) = 0. \tag{14}$$

By (9) and the Taylor expansion $\ln\cos y = -y^2/2 + O(y^4)$ as $y \to 0$, we have

$$n^2 \ln\cos\frac{\alpha_n}{2} + \frac{n^2\alpha_n^2}{8} = n^2 O(\alpha_n^4) = n^2 O(F_n^4(x)) = O\left((n^{1/2}F_n(x))^4\right) \to 0,$$

and

$$\frac{n^2\alpha_n^2}{8} = \frac{n^2F_n^2(x)}{32} = \frac{32\ln n}{32} + \frac{8x + 3\ln(2\ln n)}{(2\ln n)^{\frac{1}{2}}} \frac{(32\ln n)^{\frac{1}{2}}}{32} + \frac{(8x + 3\ln(2\ln n))^2}{32\cdot 4\cdot (2\ln n)}$$

$$= \ln n + \frac{8x + 3\ln(2\ln n)}{8} + o(1)$$

as $n \to +\infty$, which implies

$$\lim_{n\to+\infty} \left(n^2 \ln\cos\frac{\alpha_n}{2} + \ln n + x + \frac{3\ln(2\ln n)}{8} \right) = 0. \tag{15}$$

By (13)(14)(15), we have

$$\lim_{n\to+\infty} \left(\ln D_n(\alpha_n) + \ln n + x + \frac{\ln(2\ln n)}{2} - c_0 \right) = 0,$$

and thus we have

$$\lim_{n\to+\infty} \ln\left(n(2\ln n)^{\frac{1}{2}} D_n(\alpha_n) \right) = c_0 - x.$$

As $\alpha_n = F_n(x)/2$, we finally have

$$\lim_{n\to+\infty} n(2\ln n)^{\frac{1}{2}} D_n(F_n(x)/2) = e^{c_0 - x},$$

which completes the proof of (11). □

3.2. **The strategy to prove Theorem 1.** Now we take $c_1 = c_0 + \ln(\pi/2)$, $f(x) = e^{c_1 - x} = (2\pi)e^{c_0 - x}/4$, then we have $-f'(x) = f''(x) = e^{c_1 - x}$. Thanks to Lemma 1, for every positive integer k, $x_1, \cdots, x_k \in \mathbb{R}$, and τ_j is as defined in Theorem 1, if we can prove the following convergence

$$\lim_{n\to+\infty} \mathbb{E} \sum_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k} (\tau_{i_j} - x_j)_+ = (2\pi)^k \prod_{j=1}^{k} \left(e^{c_0 - x_j}/4 \right), \tag{16}$$

then Theorem 1 will be proved.

We need to introduce some notations. For a set $A \subset \mathbb{R}$, we denote $A(\text{mod } 2\pi) := \{x + 2\pi k | x \in A, k \in \mathbb{Z}\} \cap [0, 2\pi)$. Then $I(x,a) := [x, x+a](\text{mod } 2\pi)$ is an arc of size a for $a \in (0, 2\pi)$. For $0 < \theta_1 < \cdots < \theta_n < 2\pi$ and $\theta_{k+n} = \theta_k + 2\pi$, denote $J_k(a) := \{x \in [0, 2\pi) | I(x,a) \subset (\theta_k, \theta_{k+1})(\text{mod } 2\pi)\}$ for $a \in (0, 2\pi)$, $1 \le k \le n$, then we have $J_k(a) = (\theta_k, \theta_{k+1} - a)(\text{mod } 2\pi)$ for $\theta_{k+1} - \theta_k > a$ and $J_k(a) = \emptyset$ for $\theta_{k+1} - \theta_k \le a$, thus $J_k(a)$ is an arc of size $(\theta_{k+1} - \theta_k - a)_+$, moreover, $J_k(a) \subset (\theta_k, \theta_{k+1})(\text{mod } 2\pi)$ and $J_k(a) \cap J_l(b) = \emptyset$ for $k \ne l$. Now let the set

$$\Sigma_k(a_1, \cdots, a_k) := \bigcup_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k} J_{i_j}(a_j) \subset [0, 2\pi)^k,$$

then this is in fact a disjoint union and

$$|\Sigma_k(a_1, \cdots, a_k)| = \sum_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k} (\theta_{i_j+1} - \theta_{i_j} - a_j)_+$$

$$= \sum_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k} (m_{i_j} - a_j)_+,$$

here, we denote $|X|$ as the k-dimensional Lebesgue measure of a set $X \subset \mathbb{R}^k$. By (9), for every fixed $x_1, \cdots, x_k \in \mathbb{R}$, there exists $N_0 > 0$ such that $0 < 2s_0/n <$

$F_n(x_j) < 1 < 2\pi$ for $n > N_0,\ 1 \le j \le k$. Now we always assume $n > N_0$. By (8), we have

$$\begin{aligned}&\sum_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k}(\tau_{i_j} - x_j)_+\\ =&(n/4)^k(2\ln n)^{\frac{k}{2}} \sum_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k}(m_{i_j} - F_n(x_j))_+\\ =&(n/4)^k(2\ln n)^{\frac{k}{2}}|\Sigma_k(F_n(x_1),\cdots,F_n(x_k))|.\end{aligned}$$

For fixed $x_1,\cdots,x_k \in \mathbb{R}$ and $y_1,\cdots,y_k \in [0,2\pi)$, let's denote

$$\begin{aligned}&\phi_{k,n}(y_1,\cdots,y_k) := (n/4)^k(2\ln n)^{\frac{k}{2}}\times\\ &\mathbb{P}\Big((y_1,\cdots,y_k) \in \Sigma_k(F_n(x_1),\cdots,F_n(x_k))\Big),\end{aligned}$$

then we can rewrite

$$\begin{aligned}&\mathbb{E}\sum_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k}(\tau_{i_j} - x_j)_+\\ =&\mathbb{E}(n/4)^k(2\ln n)^{\frac{k}{2}}|\Sigma_k(F_n(x_1),\cdots,F_n(x_k))|\\ =&\int_{[0,2\pi)^k}\phi_{k,n}(y_1,\cdots,y_k)dy_1\cdots dy_k.\end{aligned}$$

Hence, (16) will be the direct consequence of the following two inequalities and the dominated convergence theorem: we will prove the upper bound

$$\limsup_{n\to+\infty}\ \sup_{y_1,\cdots,y_k\in[0,2\pi)} \phi_{k,n}(y_1,\cdots,y_k) \le \prod_{j=1}^{k}\left(e^{c_0-x_j}/4\right); \tag{17}$$

and if all y_k's are distinct, then we will prove the lower bound

$$\liminf_{n\to+\infty} \phi_{k,n}(y_1,\cdots,y_k) \ge \prod_{j=1}^{k}\left(e^{c_0-x_j}/4\right). \tag{18}$$

3.3. **The proof of Theorem 1.** Let's prove Theorem 1.

3.3.1. *An equivalent condition.* We first need the following equivalent condition for a point $(y_1,\cdots,y_k)$ in the set $\Sigma_k(a_1,\cdots,a_k)$.

Lemma 3. $(y_1,\cdots,y_k) \in \Sigma_k(a_1,\cdots,a_k)$ *is equivalent to the following conditions: (i)* $I(y_l,a_l) \cap I(y_j,a_j) = \emptyset$ *for* $1 \le l < j \le k$, *and (ii)* $\theta_l \notin I(y_j,a_j)$, *for* $1 \le j \le k,\ 1 \le l \le n$, *and (iii)* $\{\theta_1,\cdots,\theta_n\} \cap (y_p,y_q) \ne \emptyset,\ \{\theta_1,\cdots,\theta_n\} \setminus [y_p,y_q] \ne \emptyset$ *for every* $p,q \in \{1,\cdots,k\}$ *such that* $y_p < y_q$.

Proof. If $(y_1,\cdots,y_k) \in \Sigma_k(a_1,\cdots,a_k)$, then we can find $i_1,\cdots,i_k \in \{1,\cdots,n\}$ all distinct such that $y_j \in J_{i_j}(a_j)$, thus $I(y_j,a_j) \subset (\theta_{i_j},\theta_{i_j+1})(\text{mod } 2\pi)$, and $I(y_l,a_l)\cap I(y_j,a_j) \subseteq (\theta_{i_l},\theta_{i_l+1})(\text{mod } 2\pi) \cap (\theta_{i_j},\theta_{i_j+1})(\text{mod } 2\pi) = \emptyset$ for $1 \le l < j \le k$, since $i_l \ne i_j$, which gives (i).

Since $0 < \theta_1 < \cdots < \theta_n < 2\pi$, we have $\theta_l \notin (\theta_j,\theta_{j+1})(\text{mod } 2\pi)$ for $1 \le j,l \le n$. Thus for $1 \le j \le k,\ 1 \le l \le n$, we have $\theta_l \notin (\theta_{i_j},\theta_{i_j+1})(\text{mod } 2\pi)$ and $I(y_j,a_j) \subset (\theta_{i_j},\theta_{i_j+1})(\text{mod } 2\pi)$, which implies (ii) $\theta_l \notin I(y_j,a_j)$.

For every $p, q \in \{1, \cdots, k\}$, such that $y_p < y_q$, we have $i_p \neq i_q$. If $i_p, i_q \neq n$ then we have $y_p \in I(y_p, a_p) \subset (\theta_{i_p}, \theta_{i_p+1}) (\text{mod } 2\pi) = (\theta_{i_p}, \theta_{i_p+1})$, and similarly $y_q \in (\theta_{i_q}, \theta_{i_q+1})$. Therefore, $\theta_{i_p} < y_p < y_q < \theta_{i_q+1}$ and $i_p < i_q + 1$, since $i_p, i_q \in \mathbb{Z}$, we have $i_p \leq i_q$, since $i_p \neq i_q$, we have $i_p < i_q$ and $i_p + 1 \leq i_q$. Thus $0 < \theta_{i_p} < y_p < \theta_{i_p+1} \leq \theta_{i_q} < y_q$ and $\theta_{i_p+1} \in (y_p, y_q)$, $\theta_{i_p} \notin [y_p, y_q]$, which implies (iii).

If $i_p \neq i_q = n$, then we have $y_p \in (\theta_{i_p}, \theta_{i_p+1})$ and $y_q \in (\theta_{i_q}, \theta_{i_q+1}) (\text{mod } 2\pi) = (\theta_n, 2\pi) \cup [0, \theta_1)$. Thus $\theta_1 \leq \theta_{i_p} < y_p < y_q$, which implies $y_q \notin [0, \theta_1)$ and $y_q \in (\theta_n, 2\pi)$. Now we have $i_p < n$, $0 < \theta_{i_p} < y_p < \theta_{i_p+1} \leq \theta_n$, and $\theta_{i_p+1} \in (y_p, y_q)$, $\theta_{i_p} \notin [y_p, y_q]$, which implies (iii).

If $i_p = n \neq i_q$, then we have $y_p \in (\theta_n, 2\pi) \cup [0, \theta_1)$ and $y_q \in (\theta_{i_q}, \theta_{i_q+1}),\ i_q < n$. Thus $y_p < y_q < \theta_{i_q+1} \leq \theta_n$, which implies $y_p \notin (\theta_n, 2\pi)$ and $y_p \in [0, \theta_1)$. Now we have $y_p < \theta_1 \leq \theta_{i_q} < y_q < \theta_{i_q+1} < \pi$ and $\theta_1 \in (y_p, y_q)$, $\theta_{i_q+1} \notin [y_p, y_q]$, which also implies (iii). Now we finish the proof of the first part.

Conversely if (i)(ii)(iii) are true, by (ii) there exists a unique $i_j \in \{1, \cdots, n\}$ such that $I(y_j, a_j) \subset (\theta_{i_j}, \theta_{i_j+1}) (\text{mod } 2\pi)$, by (i) we know that all y_k's are distinct.

If $i_p = i_q$ for some $p, q \in \{1, \cdots, k\}$ with $p \neq q$, we can assume $y_p < y_q$. If $i_p = i_q < n$, then we have $y_p \in (\theta_{i_p}, \theta_{i_p+1})$ and $y_q \in (\theta_{i_q}, \theta_{i_q+1}) = (\theta_{i_p}, \theta_{i_p+1})$, thus $\theta_{i_p} < y_p < y_q < \theta_{i_p+1}$, and $\{\theta_1, \cdots, \theta_n\} \cap (y_p, y_q) = \emptyset$, which contradicts (iii).

If $i_p = i_q = n$, then we have $y_p \in (\theta_{i_p}, \theta_{i_p+1}) (\text{mod } 2\pi) = (\theta_n, 2\pi) \cup [0, \theta_1)$ and $y_q \in (\theta_{i_q}, \theta_{i_q+1}) (\text{mod } 2\pi) = (\theta_n, 2\pi) \cup [0, \theta_1)$. Thus, if $y_q < \theta_1$, then $y_p < y_q < \theta_1$ and $\{\theta_1, \cdots, \theta_n\} \cap (y_p, y_q) = \emptyset$; if $y_p > \theta_n$, then $\theta_n < y_p < y_q$ and $\{\theta_1, \cdots, \theta_n\} \cap (y_p, y_q) = \emptyset$; if $y_p \leq \theta_n,\ y_q \geq \theta_1$, then $y_p \in [0, \theta_1)$,$y_q \in (\theta_n, 2\pi)$, and $\{\theta_1, \cdots, \theta_n\} \setminus [y_p, y_q] = \emptyset$. All the 3 cases contradict (iii).

Therefore, we must have $i_p \neq i_q$ for every $p, q \in \{1, \cdots, k\}$, $p \neq q$, i.e., $i_1, \cdots, i_k \in \{1, \cdots, n\}$ are all distinct, and $I(y_j, a_j) \subset (\theta_{i_j}, \theta_{i_j+1}) (\text{mod } 2\pi)$, $y_j \in J_{i_j}(a_j)$, which implies $(y_1, \cdots, y_k) \in \Sigma_k(a_1, \cdots, a_k)$. This completes the proof. □

3.3.2. *Upper bound.* The proof of the upper bound is based on the following negative correlation of the vacuum events for the determinantal point processes.

Lemma 4. *Let $\xi^{(n)}$ be the point process associated to the eigenvalues of Haar-distributed unitary matrix (resp., an element of the GUE). Let I_1 and I_2 be compact disjoint subsets of $[0, 2\pi)$ (resp., $\mathbb{R}$). Then*

$$\mathbb{P}(\xi^{(n)}(I_1 \cup I_2) = 0) \leq \mathbb{P}(\xi^{(n)}(I_1) = 0)\mathbb{P}(\xi^{(n)}(I_2) = 0). \tag{19}$$

By monotone convergence theorem, we have

$$\mathbb{P}(\xi^{(n)}(\cup_{j=1}^{+\infty} J_j) = 0) = \lim_{k \to +\infty} \mathbb{P}(\xi^{(n)}(\cup_{j=1}^{k} J_j) = 0),$$

thus (19) is also true if I_1 and I_2 are disjoint F_σ subsets (i.e. $I_k = \cup_{j=1}^{+\infty} I_{k,j}$ and $I_{k,j}$ are compact), especially the subsets in the form of $(a, b)(\text{mod } 2\pi)$ or $[a, b](\text{mod } 2\pi)$. By induction, for disjoint F_σ subsets $I_1, \cdots, I_k$, we also have

$$\mathbb{P}(\xi^{(n)}(\cup_{j=1}^{k} I_j) = 0) \leq \prod_{j=1}^{k} \mathbb{P}(\xi^{(n)}(I_j) = 0). \tag{20}$$

By definition of $D_n(\alpha)$, for $a \in (0, 2\pi),\ x \in \mathbb{R}$, we have

$$\mathbb{P}(\xi^{(n)}(I(x, a)) = 0) = D_n(a/2). \tag{21}$$

We consider the point process

$$\xi^{(n)} = \sum_{i=1}^{n} \delta_{\theta_i}.$$

For fixed $x_1, \cdots, x_k \in \mathbb{R},\ n > N_0$, let's denote

$$A_n := \Big\{(y_1, \cdots, y_k) \in [0, 2\pi)^k \big| I(y_i, F_n(x_i)) \cap I(y_j, F_n(x_j)) = \emptyset,\ \forall\ 1 \le i < j \le k\Big\}.$$

If $(y_1, \cdots, y_k) \in A_n$, then all y_k's are distinct, let

$$I_{k,n} = \cup_{j=1}^{k} I(y_j, F_n(x_j)),\ J_{k,n,j} = (z_j, z_{j+1})(\text{mod } 2\pi),\ 1 \le j \le k, \tag{22}$$

here, $z_j (1 \le j \le k)$ is the increasing rearrangement of $y_j (1 \le j \le k)$ and $z_{k+1} = z_1 + 2\pi$. Then $I_{k,n}$ is a disjoint union and by Lemma 3 we have

$$\begin{aligned} \phi_{k,n}(y_1, \cdots, y_k) &= (n/4)^k (2 \ln n)^{\frac{k}{2}} \times \\ &\mathbb{P}(\xi^{(n)}(I_{n,k}) = 0,\ \xi^{(n)}(J_{n,k,j}) > 0,\ \forall\ 1 \le j \le k). \end{aligned} \tag{23}$$

By (20) and (21) we have

$$\begin{aligned} \phi_{k,n}(y_1, \cdots, y_k) &\le (n/4)^k (2 \ln n)^{\frac{k}{2}} \mathbb{P}(\xi^{(n)}(I_{n,k}) = 0) \\ &\le (n/4)^k (2 \ln n)^{\frac{k}{2}} \prod_{j=1}^{k} \mathbb{P}(\xi^{(n)}(I(y_j, F_n(x_j))) = 0) \\ &= (n/4)^k (2 \ln n)^{\frac{k}{2}} \prod_{j=1}^{k} D_n(F_n(x_j)/2). \end{aligned}$$

For $(y_1, \cdots, y_k) \in [0, 2\pi)^k \setminus A_n$, by Lemma 3, we have

$$\phi_{k,n}(y_1, \cdots, y_k) = 0 \le (n/4)^k (2 \ln n)^{\frac{k}{2}} \prod_{j=1}^{k} D_n(F_n(x_j)/2).$$

Therefore, by (11) we always have

$$\begin{aligned} &\sup_{y_1, \cdots, y_k \in [0, 2\pi)} \phi_{k,n}(y_1, \cdots, y_k) \le (n/4)^k (2 \ln n)^{\frac{k}{2}} \prod_{j=1}^{k} D_n(F_n(x_j)/2) \\ &= \prod_{j=1}^{k} (n (2 \ln n)^{\frac{1}{2}} D_n(F_n(x_j)/2)/4) \to \prod_{j=1}^{k} \left(e^{c_0 - x_j}/4\right),\ \ n \to +\infty, \end{aligned}$$

which gives the upper bound (17).

3.3.3. *Lower bound.* Now we consider the lower bound.

If all y_k's are distinct, let z_j be the increasing rearrangement of y_j and $z_{k+1} = z_1 + 2\pi$ as above. By (9), there further exists $N_1 > N_0$ (depending only on $x_1, \cdots, x_k$ and $y_1, \cdots, y_k$) such that $0 < 2s_0/n < F_n(x_j) < \min\{z_{i+1} - z_i | 1 \le i \le k\}/2$ for $n > N_1,\ 1 \le j \le k$. Then we have $(y_1, \cdots, y_k) \in A_n$ for $n > N_1$, and we can still use the notation (22) and formula (23) in this case. The proof of the lower bound is based on the following asymptotic splitting property,

Lemma 5.

$$\lim_{n\to+\infty} \mathbb{P}(\xi^{(n)}(I_{n,k})=0)/\prod_{j=1}^{k}\mathbb{P}(\xi^{(n)}(I(y_j,F_n(x_j)))=0)=1. \tag{24}$$

For a nuclear operator T in the form of

$$Tf(x)=\int K(x,y)f(y)dy,$$

the Hilbert-Schmidt norm is given by

$$|T|_2^2=\int\int |K(x,y)|^2dxdy$$

and the trace is given by

$$\operatorname{Tr} T=\int K(x,x)dx.$$

For a bounded operator T on L^2-space, the operator norm is given by

$$\|T\|=\sup\Big\{\|Tf\|_{L^2}|\ \|f\|_{L^2}=1\Big\}.$$

Let's recall that the probability that a determinantal point process ξ with kernel K has no point in a measurable subset A is given by the Fredholm determinant [1]

$$\mathbb{P}(\xi(A)=0)=\det(\mathrm{Id}-K_A).$$

In the case of CUE, the point process of eigenvalues $\xi^{(n)}$ is a determinantal point process with kernel [1],

$$\begin{aligned} K_n(x,y):=K^{CUE(n)}(x,y)&=\frac{1}{2\pi}\frac{\sin(n(x-y)/2)}{\sin((x-y)/2)} \\ &=\frac{1}{2\pi}\sum_{k=0}^{n-1}e^{(k-(n-1)/2)i(x-y)}. \end{aligned} \tag{25}$$

Therefore, the probability that $\xi^{(n)}$ has no point in a measurable subset I is

$$\mathbb{P}(\xi^{(n)}(I)=0)=\det(\mathrm{Id}-\chi_I P_n\chi_I),$$

where P_n is the orthogonal projection from $L^2([0,2\pi))$ to the finite dimensional space $V_n:=\operatorname{span}\{e^{i(k-(n-1)/2)x}|0\le k<n, k\in\mathbb{Z}\}$ with kernel $K_n(x,y)$ in (25), and χ_I is the characteristic function supported on I. Assume $n>N_1$ and denote

$$A=\chi_{I_{n,k}}P_n\chi_{I_{n,k}},\ B=\sum_{j=1}^{k}B_j,\ B_j=\chi_{I(y_j,F_n(x_j))}P_n\chi_{I(y_j,F_n(x_j))}, \tag{26}$$

then we have

$$\mathbb{P}(\xi^{(n)}(I_{n,k})=0)=\det(\mathrm{Id}-A);$$

since the support $I(y_j,F_n(x_j))$'s are disjoint, we also have

$$\prod_{j=1}^{k}\mathbb{P}(\xi^{(n)}(I(y_j,F_n(x_j)))=0)=\prod_{j=1}^{k}\det(\mathrm{Id}-B_j)=\det(\mathrm{Id}-B).$$

Now (24) is equivalent to

$$\lim_{n\to+\infty}\det(\mathrm{Id}-A)/\det(\mathrm{Id}-B)=1. \tag{27}$$

By (6) we have

$$\det(\mathrm{Id}-B)=\prod_{j=1}^{k}\mathbb{P}(\xi^{(n)}(I(y_j,F_n(x_j)))=0)=\prod_{j=1}^{k}D_n(F_n(x_j)/2)>0,$$

and thus $\det(\mathrm{Id}-A)/\det(\mathrm{Id}-B)$ is well defined. Since P_n is a finite rank orthogonal projection operator, we know that $A,\ B$ are both finite rank symmetric operators. As $\langle Af,f\rangle=\langle P_n\chi_{I_{n,k}}f,\chi_{I_{n,k}}f\rangle=\|P_n\chi_{I_{n,k}}f\|_{L^2}^2$, we have

$$0\le\langle Af,f\rangle=\|P_n\chi_{I_{n,k}}f\|_{L^2}^2\le\|\chi_{I_{n,k}}f\|_{L^2}^2\le\|f\|_{L^2}^2,$$

here, we use the L^2 inner product

$$\langle f,g\rangle:=\int_0^{2\pi}f(x)\overline{g(x)}dx,\ \ \|f\|_{L^2}^2=\langle f,f\rangle.$$

Similarly, we have $0\le\langle B_jf,f\rangle\le\|\chi_{I(y_j,F_n(x_j))}f\|_{L^2}^2$, and if $n>N_1$, then $I(y_j,F_n(x_j))$'s are disjoint and

$$0\le\sum_{j=1}^{k}\langle B_jf,f\rangle=\langle Bf,f\rangle\le\sum_{j=1}^{k}\|\chi_{I(y_j,F_n(x_j))}f\|_{L^2}^2\le\|f\|_{L^2}^2.$$

Therefore, we can conclude that $A,\ B,\ \mathrm{Id}-A,\ \mathrm{Id}-B$ are all semi-positive definite. As $\det(\mathrm{Id}-B)>0$, then $\mathrm{Id}-B$ is further positive definite, so is its inverse $(\mathrm{Id}-B)^{-1}$. Such results are also true for the GUE case in §4.

We will need the following general comparison inequalities regarding the Fredholm determinants which will be used in both CUE and GUE cases.

Lemma 6. *Assume A,B are finite rank symmetric operators on a Hilbert space, $Id-B$ is positive definite and $Id-A$ is semi-positive definite, then we have*

$$1-|A-B|_2^2\|(Id-B)^{-1}\|^2\le\exp(\mathrm{Tr}(A-B)(Id-B)^{-1})\det(Id-A)/\det(Id-B)\le 1$$

and

$$|\mathrm{Tr}((A-B)(Id-B)^{-1})|\le|\mathrm{Tr}(A-B)|+|A-B|_2|B|_2\|(Id-B)^{-1}\|.$$

In the proof we need to use the following formulas [7]

- If A,B are finite rank operators, then $\det(\mathrm{Id}-A)\det(\mathrm{Id}-B)=\det((\mathrm{Id}-A)(\mathrm{Id}-B))$ and $|\mathrm{Tr}\,AB|\le|A|_2|B|_2$.
- If A is a finite rank operator, B is a bounded operator, then $\mathrm{Tr}\,AB=\mathrm{Tr}\,BA$ and $|AB|_2\le|A|_2\|B\|$.

If B is a finite rank symmetric operator and $\mathrm{Id}-B$ is positive definite, let $\{e_k\}$ be eigenfunctions forming a complete orthonormal basis with $Be_k=\lambda_k(B)e_k$, then $\lambda_k(B)\in\mathbb{R},\ \lambda_k(B)<1$. Now we have

$$\det(\mathrm{Id}-B)=\prod(1-\lambda_k(B)),\ \mathrm{Tr}\,B=\sum\lambda_k(B).$$

We can also define $(\mathrm{Id}-B)^p$ for every $p\in\mathbb{R}$ as

$$(\mathrm{Id}-B)^pf=\sum(1-\lambda_k(B))^p\langle f,e_k\rangle e_k=f+\sum((1-\lambda_k(B))^p-1)\langle f,e_k\rangle e_k.$$

Then $(\mathrm{Id}-B)^p$ is also positive definite, $(\mathrm{Id}-B)^p(\mathrm{Id}-B)^q=(\mathrm{Id}-B)^{p+q}$ and $\det(\mathrm{Id}-B)^p=(\det(\mathrm{Id}-B))^p$. Moreover, for $p<0$, we have $\|(\mathrm{Id}-B)^p\|=(1-\lambda_1(B))^p$ where $\lambda_1(B)$ is the largest eigenvalue of B.

Proof. Since $\mathrm{Id}-B$ is positive definite, so is its inverse $(\mathrm{Id}-B)^{-1}$ and $(\mathrm{Id}-B)^{-1}$ has a positive square root $(\mathrm{Id}-B)^{-1/2}$. Moreover, $\|(\mathrm{Id}-B)^{-1/2}\|^2 = \|(\mathrm{Id}-B)^{-1}\| = (1-\lambda_1(B))^{-1}$, where $\lambda_1(B)$ is the largest eigenvalue of B and $\lambda_1(B) < 1$. We also have $(\det(\mathrm{Id}-B)^{-1/2})^2 = \det(\mathrm{Id}-B)^{-1} = (\det(\mathrm{Id}-B))^{-1}$. Since $\mathrm{Id}-A$ is semi-positive definite, so is $A_1 := (\mathrm{Id}-B)^{-1/2}(\mathrm{Id}-A)(\mathrm{Id}-B)^{-1/2}$, and $\det A_1 = \det(\mathrm{Id}-A)/\det(\mathrm{Id}-B)$. Let $B_1 := (\mathrm{Id}-B)^{-1/2}(A-B)(\mathrm{Id}-B)^{-1/2}$, then we have $A_1+B_1 = \mathrm{Id}$, B_1 is a finite rank symmetric operator, $\mathrm{Tr}\, B_1 = \mathrm{Tr}(A-B)(\mathrm{Id}-B)^{-1}$, and its eigenvalues $\lambda_j(B_1)$ are real. Since $A_1 = \mathrm{Id}-B_1$ is semi-positive definite, we have $\lambda_j(B_1) \le 1$ and $\det A_1 = \det(\mathrm{Id}-B_1) = \prod_j(1-\lambda_j(B_1))$. Now we can rewrite

$$
\begin{aligned}
(28) \qquad & \exp(\mathrm{Tr}(A-B)(\mathrm{Id}-B)^{-1})\det(\mathrm{Id}-A)/\det(\mathrm{Id}-B) \\
&= \exp(\mathrm{Tr}\, B_1)\det A_1 \\
&= \exp(\sum_j \lambda_j(B_1))\prod_j(1-\lambda_j(B_1)) = \prod_j(e^{\lambda_j(B_1)}(1-\lambda_j(B_1))).
\end{aligned}
$$

Since $e^\lambda(1-\lambda) \le 1$ and $1+\lambda \le e^\lambda$, we have $(1+\lambda)_+ \le e^\lambda$ and thus $1 \ge e^\lambda(1-\lambda) \ge (1+\lambda)_+(1-\lambda) = (1-\lambda^2)_+$ for $\lambda \le 1$. Therefore, we have

$$
(29) \qquad 1 \ge \prod_j(e^{\lambda_j(B_1)}(1-\lambda_j(B_1))) \ge \prod_j(1-\lambda_j(B_1)^2)_+ \ge 1 - \sum_j \lambda_j(B_1)^2.
$$

Moreover, we have

$$
\begin{aligned}
(30) \qquad & \sum_j \lambda_j(B_1)^2 = |B_1|_2^2 = |(\mathrm{Id}-B)^{-1/2}(A-B)(\mathrm{Id}-B)^{-1/2}|_2^2 \\
&\le \|(\mathrm{Id}-B)^{-1/2}\|^2|A-B|_2^2\|(\mathrm{Id}-B)^{-1/2}\|^2 = \|(\mathrm{Id}-B)^{-1}\|^2|A-B|_2^2.
\end{aligned}
$$

Therefore, the first inequality follows if we combine (28)(29)(30). We also have

$$
\begin{aligned}
& |\mathrm{Tr}((A-B)(\mathrm{Id}-B)^{-1})| \\
=& |\mathrm{Tr}((A-B) + (A-B)B(\mathrm{Id}-B)^{-1})| \\
\le& |\mathrm{Tr}(A-B)| + |\mathrm{Tr}((A-B)B(\mathrm{Id}-B)^{-1})| \\
\le& |\mathrm{Tr}(A-B)| + |A-B|_2|B(\mathrm{Id}-B)^{-1}|_2 \\
\le& |\mathrm{Tr}(A-B)| + |A-B|_2|B|_2\|(\mathrm{Id}-B)^{-1}\|,
\end{aligned}
$$

which is the second inequality. This completes the proof. □

Thanks to Lemma 6 and the fact that $\lim_{n\to+\infty}(\ln n)^2 e^{-(\ln n)^{1/2}} = 0$, for every positive integer k, $x_1, \cdots, x_k \in \mathbb{R}$ and distinct $y_1, \cdots, y_k \in [0, 2\pi)$, if we can prove the following bound for $n > N_1$,

$$
(31) \qquad \mathrm{Tr}((A-B)(\mathrm{Id}-B)^{-1}) = 0,
$$

$$
(32) \qquad |A-B|_2^2 = O\left(\frac{\ln n}{n^2}\right), \ \|(\mathrm{Id}-B)^{-1}\| = O(n(\ln n)^{\frac{1}{2}} e^{-(\ln n)^{\frac{1}{2}}/2}),
$$

then (27) will be proved, and thus (24).

By (26), we can write

$$
A - B = \sum_{i\ne j} \chi_{I(y_i, F_n(x_i))} P_n \chi_{I(y_j, F_n(x_j))} := \sum_{i\ne j} \chi_i P_n \chi_j,
$$

here, we denote $\chi_j = \chi_{I(y_j,F_n(x_j))}$. For $i \neq j$, we have $\mathrm{Tr}(\chi_i P_n \chi_j (\mathrm{Id}-B)^{-1}) = \mathrm{Tr}(P_n\chi_j(\mathrm{Id}-B)^{-1}\chi_i)$. Since $I(y_j,F_n(x_j))$'s are disjoint, we have $\chi_j B = B_j = B\chi_j$ and thus $(\mathrm{Id}-B)\chi_j(\mathrm{Id}-B)^{-1}\chi_i = \chi_j(\mathrm{Id}-B)(\mathrm{Id}-B)^{-1}\chi_i = \chi_j\chi_i = 0$. Since $(\mathrm{Id}-B)$ is invertible, we further have $\chi_j(\mathrm{Id}-B)^{-1}\chi_i = 0$, which implies $\mathrm{Tr}(\chi_i P_n \chi_j(\mathrm{Id}-B)^{-1}) = 0$. And thus (31) follows.

By definition of N_1 and z_j, for $x \in I(y_i, F_n(x_i)),\ y \in I(y_j, F_n(x_j)),\ i \neq j,\ n > N_1$, we have

$$\begin{aligned}&\min(|x-y|, 2\pi - |x-y|) \geq \min(|y_i - y_j|, 2\pi - |y_i - y_j|) - \max(F_n(x_i), F_n(x_j))\\ &\geq \min\{z_{i+1} - z_i | 1 \leq i \leq k\} - \min\{z_{i+1} - z_i | 1 \leq i \leq k\}/2\\ &= \min\{z_{i+1} - z_i | 1 \leq i \leq k\}/2 := a_0 \in (0, 2\pi),\end{aligned}$$

and

$$|K_n(x,y)| = \left|\frac{1}{2\pi}\frac{1-e^{in(x-y)}}{1-e^{i(x-y)}}\right| \leq \frac{1}{\pi}\frac{1}{|1-e^{i(x-y)}|} \leq \frac{1}{\pi}\frac{1}{|1-e^{ia_0}|} = O(1),$$

using this and (9) we have

$$\begin{aligned}|A-B|_2^2 = &\sum_{i\neq j}\int_{I(y_i,F_n(x_i))} dx \int_{I(y_j,F_n(x_j))} |K_n(x,y)|^2 dy\\ =&\sum_{i\neq j}\int_{I(y_i,F_n(x_i))} dx \int_{I(y_j,F_n(x_j))} O(1) dy = \sum_{i\neq j} F_n(x_i)F_n(x_j)O(1)\\ =&\sum_{i\neq j} O\left(\frac{\ln n}{n^2}\right)O(1) = k(k-1)O\left(\frac{\ln n}{n^2}\right) = O\left(\frac{\ln n}{n^2}\right),\end{aligned}$$

which is the first inequality in (32). It remains to estimate $\|(\mathrm{Id}-B)^{-1}\|$, we need the following eigenvalue esitmate.

Lemma 7. *Let B be a finite rank symmetric operator on a Hilbert space and $Id - B$ is positive definite, let $\lambda_1(B)$ be the largest eigenvalue of B, then we have $1 - \lambda_1(B) \geq \det(Id - B)e^{\mathrm{Tr}\,B-1}$.*

Proof. Let $\lambda_k(B)$ be the eigenvalues of B, then we have $\lambda_k(B) < 1$ and

$$\det(\mathrm{Id}-B)e^{\mathrm{Tr}\,B-1} = \prod_k (1-\lambda_k(B))e^{\sum_k \lambda_k(B) - 1}.$$

Using the fact that $0 < (1-\lambda)e^{\lambda} \leq 1$ for $\lambda < 1$ again, we have

$$\begin{aligned}\det(\mathrm{Id}-B)e^{\mathrm{Tr}\,B-1} &= e^{-1}\prod_k (1-\lambda_k(B))e^{\lambda_k(B)}\\ =&(1-\lambda_1(B))e^{\lambda_1(B)-1}\prod_{k\neq 1}(1-\lambda_k(B))e^{\lambda_k(B)}\\ \leq&(1-\lambda_1(B))e^{\lambda_1(B)-1} \leq 1 - \lambda_1(B).\end{aligned}$$

This completes the proof. □

Recall the definitions of B and B_j in (26), assume $0 \neq f \in L^2([0, 2\pi))$ such that $Bf = \lambda_1(B)f$ where $\lambda_1(B)$ is the largest eigenvalue of B, then we have

$$\lambda_1(B)f = Bf = \sum_{j=1}^k B_j f.$$

For $n > N_1,\ i \neq j$, by definition we have $I(y_i, F_n(x_i)) \cap I(y_j, F_n(x_j)) = \emptyset$ and then $\chi_{I(y_i,F_n(x_i))}B_j = 0$, thus we further have

$$\begin{aligned}\lambda_1(B)\chi_{I(y_i,F_n(x_i))}f &= \chi_{I(y_i,F_n(x_i))}Bf \\ &= \chi_{I(y_i,F_n(x_i))}B_i f = B_i f = B_i\chi_{I(y_i,F_n(x_i))}f,\end{aligned} \tag{33}$$

i.e., $\chi_{I(y_i,F_n(x_i))}f$ is an eigenfunction of B_i and its largest eigenvalue $\lambda_1(B_i) \geq \lambda_1(B)$.

If $\chi_{I(y_i,F_n(x_i))}f \neq 0$ for some $1 \leq i \leq k$, then by Lemma 7 we have $1-\lambda_1(B) \geq 1-\lambda_1(B_i) \geq \det(\mathrm{Id}-B_i)e^{\mathrm{Tr}\,B_i - 1}$. Notice that

$$\det(\mathrm{Id}-B_i) = \mathbb{P}(\xi^{(n)}(I(y_i, F_n(x_i))) = 0) = D_n(F_n(x_i)/2),$$
$$K_n(x,x) = n/(2\pi),$$

and

$$\mathrm{Tr}\,B_i = \int_{I(y_i,F_n(x_i))} K_n(x,x)dx = nF_n(x_i)/(2\pi),$$

thus we have

$$1-\lambda_1(B) \geq D_n(F_n(x_i)/2)e^{nF_n(x_i)/(2\pi)-1}.$$

By (9)(11) and $32 > \pi^2$, there exists a constant $N_2 > N_1$ such that $nF_n(x_i) > \pi(\ln n)^{\frac{1}{2}}$ and $n(4\ln n)^{\frac{1}{2}}D_n(F_n(x_i)/2) > e^{c_0-x_i}$ for $1 \leq i \leq k$. Thus, we further have

$$1-\lambda_1(B) \geq n^{-1}(4\ln n)^{-\frac{1}{2}}e^{c_0-x_i}e^{(\ln n)^{\frac{1}{2}}/2-1}.$$

If $\chi_{I(y_i,F_n(x_i))}f = 0$ for every $1 \leq i \leq k$, then we have $B_i f = 0$, and thus $\lambda_1(B)f = 0$, $\lambda_1(B) = 0,\ 1-\lambda_1(B) = 1$. In both cases for $n > N_2$ we always have

$$1-\lambda_1(B) \geq \min(1, n^{-1}(4\ln n)^{-\frac{1}{2}}e^{c_0-\max\{x_j|1\leq j\leq k\}}e^{(\ln n)^{\frac{1}{2}}/2-1}),$$

therefore,

$$\begin{aligned}\|(\mathrm{Id}-B)^{-1}\| = (1-\lambda_1(B))^{-1} &\leq 1 + n(4\ln n)^{\frac{1}{2}}e^{\max\{x_j|1\leq j\leq k\}-c_0}e^{1-(\ln n)^{\frac{1}{2}}/2} \\ &\leq 1 + O(n(\ln n)^{\frac{1}{2}}e^{-(\ln n)^{\frac{1}{2}}/2}) = O(n(\ln n)^{\frac{1}{2}}e^{-(\ln n)^{\frac{1}{2}}/2}),\end{aligned}$$

which finishes the second inequality in (32), and hence, we finish the proof of (24) in Lemma 5.

Now we can use (24) to prove the lower bound (18). For $n > N_1$, by (23) we have

$$\begin{aligned}\phi_{k,n}(y_1,\cdots,y_k) \geq\ & (n/4)^k(2\ln n)^{\frac{k}{2}}\mathbb{P}(\xi^{(n)}(I_{n,k}) = 0) \\ &- (n/4)^k(2\ln n)^{\frac{k}{2}}\sum_{j=1}^{k}\mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0).\end{aligned} \tag{34}$$

Now we claim that

$$(n/4)^k(2\ln n)^{\frac{k}{2}}\mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0) \to 0,\ \ n \to +\infty.$$

Let $x_0 = \min\{x_j|1 \leq j \leq k\}$, then we have $F(x_j) \geq F(x_0)$, $I_{n,k} \supseteq \cup_{j=1}^k I(y_j, F_n(x_0)) = \cup_{j=1}^k I(z_j, F_n(x_0))$. Therefore, we have $I_{n,k} \cup J_{n,k,j} \supseteq J_{n,k,j} \cup (\cup_{i\neq j} I(z_i, F_n(x_0)))$, and the right hand side is a disjoint union for $n > N_1$. If $k = 1$, then $J_{n,k,j} = (0, 2\pi)$ and $\mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0) = \mathbb{P}(\xi^{(n)}((0,2\pi)) = 0) = 0$. If $k > 1$, by (20) and (21) we have

$$0 \leq \mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0) = \mathbb{P}(\xi^{(n)}(I_{n,k} \cup J_{n,k,j}) = 0)$$

$$
\begin{aligned}
&\leq \mathbb{P}(\xi^{(n)}(J_{n,k,j} \cup (\cup_{i\neq j} I(z_i, F_n(x_0)))) = 0)\\
&\leq \mathbb{P}(\xi^{(n)}(J_{n,k,j}) = 0) \prod_{i\neq j} \mathbb{P}(\xi^{(n)}(I(z_i, F_n(x_0))) = 0)\\
&= D_n((z_{j+1} - z_j)/2)(D_n(F_n(x_0)/2))^{k-1}.
\end{aligned}
$$

Thus by (10) and (11), we have

$$
\begin{aligned}
&0 \leq \limsup_{n\to+\infty} (n/4)^k (2\ln n)^{\frac{k}{2}} \mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0)\\
&\leq \lim_{n\to+\infty} (n/4)(2\ln n)^{\frac{1}{2}} D_n((z_{j+1} - z_j)/2) \left(\lim_{n\to+\infty} (n/4)(2\ln n)^{\frac{1}{2}} D_n(F_n(x_0)/2) \right)^{k-1}\\
&= 0 \cdot \left(e^{c_0 - x_0}/4\right)^{k-1} = 0, \quad \forall\ 1 \leq j \leq k,
\end{aligned}
$$

which implies the claim. Therefore, combining the cases $k = 1$ and $k > 1$, using (11)(21)(24)(34), we have

$$
\begin{aligned}
&\liminf_{n\to+\infty} \phi_{k,n}(y_1, \cdots, y_k)\\
&\geq \liminf_{n\to+\infty} (n/4)^k (2\ln n)^{\frac{k}{2}} \mathbb{P}(\xi^{(n)}(I_{n,k}) = 0)\\
&= \liminf_{n\to+\infty} (n/4)^k (2\ln n)^{\frac{k}{2}} \prod_{j=1}^{k} \mathbb{P}(\xi^{(n)}(I(y_j, F_n(x_j))) = 0)\\
&= \liminf_{n\to+\infty} (n/4)^k (2\ln n)^{\frac{k}{2}} \prod_{j=1}^{k} D_n(F_n(x_j)/2)\\
&= \prod_{j=1}^{k} \lim_{n\to+\infty} (n(2\ln n)^{\frac{1}{2}} D_n(F_n(x_j)/2)/4) = \prod_{j=1}^{k} (e^{c_0 - x_j}/4),
\end{aligned}
$$

which is the lower bound (18). Therefore, we finish the proof of Theorem 1.

4. The GUE case

In this section, let's denote $\mathbb{P}^{CUE(n)}$ (or $\mathbb{P}^{GUE(n)}$) as the probability taken with respect to the Haar measure of $U(n)$ (or GUE), when we drop the superscript, the expectation $\mathbb{E}$ and the probability $\mathbb{P}$ are taken with respect to GUE.

4.1. **Another rescaling limit.** We first need another rescaling limit of $D_n(\alpha)$. Let's denote

$$
G_n(x) = \frac{8x - 5\ln(2\ln n)}{2n(2\ln n)^{\frac{1}{2}}} + \frac{(32\ln n)^{\frac{1}{2}}}{n}. \tag{35}
$$

Given a compact subinterval $I = [a, b]$ in $(-2, 2)$, let's denote $S(I) = \inf_I \sqrt{4 - x^2}$, then we have

$$
S(I)m_k^* = G_n(\tau_k^*),
$$

where m_k^* and τ_k^* are as defined in Theorem 2.

From the definition of $G_n(x)$ we have

$$
y - x = (G_n(y) - G_n(x))(n/4)(2\ln n)^{\frac{1}{2}}, \tag{36}
$$

and for every fixed x,

$$\lim_{n\to+\infty}\frac{nG_n(x)}{(32\ln n)^{\frac12}}=1,\ \lim_{n\to+\infty}nG_n(x)=+\infty,\ \lim_{n\to+\infty}n^{\gamma}G_n(x)=0,\ \forall\ \gamma<1. \tag{37}$$

Now we need the following rescaling limit which is similar to (11).

Lemma 8. *For fixed $x,z\in\mathbb{R}$, we have*

$$\lim_{n\to+\infty}n(2\ln n)^{-\frac12}D_n((1+z/\ln n)G_n(x)/2)=e^{c_0-x-2z}. \tag{38}$$

Proof. Let $\alpha_n=(1+z/\ln n)G_n(x)/2$, then by (37) we have $\alpha_n\to 0,\ n\alpha_n\to+\infty$ as $n\to+\infty$, thus $s_0/n<\alpha_n<\pi-\varepsilon$ for n sufficiently large. Therefore, (12) holds for such α_n, and we still have

$$\lim_{n\to+\infty}\left(\ln D_n(\alpha_n)-n^2\ln\cos\frac{\alpha_n}{2}+\frac14\ln\left(n\sin\frac{\alpha_n}{2}\right)-c_0\right)=0. \tag{39}$$

By (37) we have

$$\begin{aligned}\lim_{n\to+\infty}\frac{(2\ln n)^{\frac12}}{n\sin(\alpha_n/2)}&=\lim_{n\to+\infty}\frac{(2\ln n)^{\frac12}}{n\alpha_n/2}\\&=\lim_{n\to+\infty}\frac{(2\ln n)^{\frac12}}{n(1+z/\ln n)G_n(x)/4}=\lim_{n\to+\infty}\frac{(32\ln n)^{\frac12}}{nG_n(x)}=1,\end{aligned}$$

and thus we have

$$\lim_{n\to+\infty}\left(\frac18\ln(2\ln n)-\frac14\ln\left(n\sin\frac{\alpha_n}{2}\right)\right)=0. \tag{40}$$

By (37) and Taylor expansion of $\ln\cos y$ as $y\to 0$, we have

$$n^2\ln\cos\frac{\alpha_n}{2}+\frac{n^2\alpha_n^2}{8}=n^2O(\alpha_n^4)=n^2O(G_n^4(x))\to 0,$$

and

$$\begin{aligned}\frac{n^2G_n^2(x)}{32}&=\frac{32\ln n}{32}+\frac{8x-5\ln(2\ln n)}{(2\ln n)^{\frac12}}\frac{(32\ln n)^{\frac12}}{32}+\frac{(8x-5\ln(2\ln n))^2}{32\cdot 4\cdot(2\ln n)}\\&=\ln n+\frac{8x-5\ln(2\ln n)}{8}+o(1),\end{aligned}$$

and

$$\begin{aligned}&\frac{n^2\alpha_n^2}{8}-\frac{n^2G_n^2(x)}{32}=\frac{n^2(1+z/\ln n)^2G_n(x)^2}{32}-\frac{n^2G_n^2(x)}{32}\\&=\frac{(z/\ln n)(2+z/\ln n)n^2G_n(x)^2}{32}\\&=(z/\ln n)(2+z/\ln n)(\ln n+o(\ln n))\to 2z\end{aligned}$$

as $n\to+\infty$, which implies

$$\lim_{n\to+\infty}\left(n^2\ln\cos\frac{\alpha_n}{2}+\ln n+x-\frac{5\ln(2\ln n)}{8}+2z\right)=0. \tag{41}$$

By (39)(40)(41) we have

$$\lim_{n\to+\infty}\left(\ln D_n(\alpha_n)+\ln n+x+2z-\frac{\ln(2\ln n)}{2}-c_0\right)=0,$$

and thus we have

$$\lim_{n\to+\infty} \ln\left(n(2\ln n)^{-\frac{1}{2}} D_n(\alpha_n)\right) = c_0 - x - 2z.$$

As $\alpha_n = (1 + z/\ln n)G_n(x)/2$, the above limit is equivalent to

$$\lim_{n\to+\infty} n(2\ln n)^{-\frac{1}{2}} D_n((1 + z/\ln n)G_n(x)/2) = e^{c_0-x-2z},$$

this completes the proof of (38). □

4.2. **One integral lemma.** In this subsection, we will prove one integral Lemma 10 which will be used in the proof of Theorem 2. We first have the bound,

Lemma 9. *For every fixed* $x \in \mathbb{R}$ *and* $A > 1$, *there exists a constant* $N_3 > 0$ *depending only on* x, A *such that for* $n > N_3,\ w \in [1, A]$, *we have* $s_0/n < G_n(x)/2 < AG_n(x)/2 < \pi/2$ *and* $D_n(wG_n(x)/2) \le e^{1-(w-1)\ln n} D_n(G_n(x)/2)$.

Proof. Let $\alpha_n = G_n(x)/2$, then by (37) we have $\alpha_n \to 0,\ n\alpha_n \to +\infty$ as $n \to +\infty$, thus there exists a constant $N_{3,0} > 0$ such that $s_0/n < \alpha_n < w\alpha_n < A\alpha_n < \pi/2$ for $n > N_{3,0}$ and

$$\lim_{n\to+\infty} \sup_{w\in[1,A]} \frac{1}{n\sin(w\alpha_n/2)} = \lim_{n\to+\infty} \frac{1}{n\sin(\alpha_n/2)} = \lim_{n\to+\infty} \frac{2}{n\alpha_n} = 0.$$

By (6) there exists a constant $N_{3,1} > N_{3,0}$ such that

$$\left|\ln D_n(w\alpha_n) - n^2 \ln\cos\frac{w\alpha_n}{2} + \frac{1}{4}\ln\left(n\sin\frac{w\alpha_n}{2}\right) - c_0\right| < 1/2$$

for $n > N_{3,1},\ z \in [1, A]$, thus we have

$$\begin{aligned}&\ln(D_n(w\alpha_n)/D_n(\alpha_n)) = \ln D_n(w\alpha_n) - \ln D_n(\alpha_n)\\ \le& n^2 \ln\cos\frac{w\alpha_n}{2} - \frac{1}{4}\ln\left(n\sin\frac{w\alpha_n}{2}\right) - n^2\ln\cos\frac{\alpha_n}{2} + \frac{1}{4}\ln\left(n\sin\frac{\alpha_n}{2}\right) + 1.\end{aligned}$$

Let's denote $F(y) = \ln\cos(y/2)$, since $\sin\frac{w\alpha_n}{2} \ge \sin\frac{\alpha_n}{2}$, we further have

$$\ln\frac{D_n(w\alpha_n)}{D_n(\alpha_n)} \le n^2\ln\cos\frac{w\alpha_n}{2} - n^2\ln\cos\frac{\alpha_n}{2} + 1 = n^2(F(w\alpha_n) - F(\alpha_n)) + 1.$$

Since $F'(y) = -\tan(y/2)/2 < -y/4 \le -\alpha_n/4$ for $n > N_{3,1},\ y \in [\alpha_n, A\alpha_n] \subset (0, \pi)$, we have $F(w\alpha_n) - F(\alpha_n) \le -(w\alpha_n - \alpha_n)\alpha_n/4$ and thus

$$\ln\frac{D_n(w\alpha_n)}{D_n(\alpha_n)} \le -n^2(w\alpha_n - \alpha_n)\alpha_n/4 + 1 = -(w-1)\frac{n^2\alpha_n^2}{4} + 1$$

for $n > N_{3,1},\ w \in [1, A]$. By (37) we have

$$\frac{n^2\alpha_n^2}{4\ln n} = \frac{n^2G_n^2(x)}{16\ln n} \to 2$$

as $n \to +\infty$, and there exists a constant $N_3 > N_{3,1}$ such that $n^2\alpha_n^2 > 4\ln n$ for $n > N_3$, which implies

$$\ln(D_n(w\alpha_n)/D_n(\alpha_n)) \le -(w-1)\ln n + 1.$$

As $\alpha_n = G_n(x)/2$, for $n > N_3 > N_{3,0}$ and $w \in [1, A]$, we have

$$\begin{aligned}&D_n(wG_n(x)/2) = D_n(w\alpha_n) = \exp(\ln(D_n(w\alpha_n)/D_n(\alpha_n)))D_n(\alpha_n)\\ \le& e^{-(w-1)\ln n+1}D_n(\alpha_n) = e^{1-(w-1)\ln n}D_n(G_n(x)/2),\end{aligned}$$

this completes the proof. □

Using (38) and Lemma 9, we have the limit of the integral,

Lemma 10. *For $I=[a,b]\subset(-2,2)$, let $S(I)=\inf_I\sqrt{4-y^2}$, then we have*

$$\lim_{n\to+\infty} n(2\ln n)^{\frac12}\int_I D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)dy = M(I)e^{c_0-x},$$

where $M(I)=(4-a^2)/|a|$ if $a+b<0$, $M(I)=(4-b^2)/|b|$ if $a+b>0$, and $M(I)=2(4-a^2)/|a|$ if $a+b=0$.

Proof. **Case 1:** $a+b<0$. In this case we have $a<0$, $S(I)=\sqrt{4-a^2}$. Let $A=2/S(I)$, then we have $1\le\sqrt{4-y^2}/S(I)\le 2/S(I)=A$ for $y\in I$. Let N_3 be determined in Lemma 9 with $w=\sqrt{4-y^2}/S(I)\in[1,A]$, for $n>N_3$, we have

$$D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)\le e^{1-(\sqrt{4-y^2}/S(I)-1)\ln n}D_n(G_n(x)/2).$$

Let $b_0=(a+b)/2$, then we have $a<b_0<\min(b,0)$ and we can write $I=I_1\cup I_2$ such that $I_1=[a,b_0]$, $I_2=[b_0,b]$. Now we have for n large enough,

$$\begin{aligned}
&n(2\ln n)^{\frac12}\int_{I_2} D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)dy\\
\le&n(2\ln n)^{\frac12}\int_{I_2} e^{1-(\sqrt{4-y^2}/S(I)-1)\ln n}D_n(G_n(x)/2)dy\\
\le&n(2\ln n)^{\frac12}\int_{I_2} e^{1-(S(I_2)/S(I)-1)\ln n}D_n(G_n(x)/2)dy\\
=&n(2\ln n)^{\frac12}(b-b_0)e^{1-(S(I_2)/S(I)-1)\ln n}D_n(G_n(x)/2),
\end{aligned}$$

where $S(I_2)=\min(\sqrt{4-b_0^2},\sqrt{4-b^2})>S(I)>0$. By (38), we have

$$\begin{aligned}
&\lim_{n\to+\infty} n(2\ln n)^{\frac12}e^{1-(S(I_2)/S(I)-1)\ln n}D_n(G_n(x)/2)\\
=&\lim_{n\to+\infty}(2\ln n)e^{1-(S(I_2)/S(I)-1)\ln n}\lim_{n\to+\infty} n(2\ln n)^{-\frac12}D_n(G_n(x)/2)=0,
\end{aligned}$$

which implies

$$\lim_{n\to+\infty} n(2\ln n)^{\frac12}\int_{I_2} D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)dy=0. \tag{42}$$

As to the integration in I_1, we change variable $y=-\sqrt{4-z^2}$ to obtain

$$\begin{aligned}
&\int_{I_1} D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)dy=\int_{a_1}^{b_1} D_n(z/S(I)\cdot G_n(x)/2)\frac{z}{\sqrt{4-z^2}}dz\\
=&S(I)(\ln n)^{-1}\int_0^{(b_1/a_1-1)\ln n} D_n((1+z/\ln n)G_n(x)/2)\frac{(1+z/\ln n)a_1}{\sqrt{4-(1+z/\ln n)^2a_1^2}}dz,
\end{aligned}$$

here $a_1=\sqrt{4-a^2}=S(I)$, $b_1=\sqrt{4-b_0^2}>a_1$, thus we have

$$\begin{aligned}
&n(2\ln n)^{\frac12}\int_{I_1} D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)dy=n(2\ln n)^{-\frac12}D_n(G_n(x)/2)\times\\
&2S(I)\int_0^{(b_1/a_1-1)\ln n}\frac{D_n((1+z/\ln n)G_n(x)/2)}{D_n(G_n(x)/2)}\frac{(1+z/\ln n)a_1}{\sqrt{4-(1+z/\ln n)^2a_1^2}}dz.
\end{aligned}$$

Since $b_1/a_1 = b_1/S(I) \le 2/S(I) = A$, by Lemma 9 the integrand above has the uniform bound

$$\begin{aligned}
&\sup_{z\in[0,(b_1/a_1-1)\ln n]} e^z \frac{D_n((1+z/\ln n)G_n(x)/2)}{D_n(G_n(x)/2)} \frac{(1+z/\ln n)a_1}{\sqrt{4-(1+z/\ln n)^2a_1^2}} \\
=&\sup_{w\in[1,b_1/a_1]} e^{(w-1)\ln n} \frac{D_n(wG_n(x)/2)}{D_n(G_n(x)/2)} \frac{wa_1}{\sqrt{4-w^2a_1^2}} \\
\le&\sup_{w\in[1,b_1/a_1]} e^{(w-1)\ln n} e^{1-(w-1)\ln n} \frac{b_1}{\sqrt{4-b_1^2}} = \frac{eb_1}{\sqrt{4-b_1^2}}
\end{aligned}$$

for n large enough. By (38) (with $z=0$) we have

$$\begin{aligned}
&\lim_{n\to+\infty} \frac{D_n((1+z/\ln n)G_n(x)/2)}{D_n(G_n(x)/2)} \frac{(1+z/\ln n)a_1}{\sqrt{4-(1+z/\ln n)^2a_1^2}} \\
=&\lim_{n\to+\infty} \frac{D_n((1+z/\ln n)G_n(x)/2)}{D_n(G_n(x)/2)} \frac{a_1}{\sqrt{4-a_1^2}} = e^{-2z}\frac{a_1}{\sqrt{4-a_1^2}}.
\end{aligned}$$

Therefore, we can apply the dominated convergence theorem to get

$$\begin{aligned}
&\lim_{n\to+\infty} 2S(I)\int_0^{(b_1/a_1-1)\ln n} \frac{D_n((1+z/\ln n)G_n(x)/2)}{D_n(G_n(x)/2)} \frac{(1+z/\ln n)a_1}{\sqrt{4-(1+z/\ln n)^2a_1^2}}dz \\
=&2S(I)\int_0^{+\infty} e^{-2z}\frac{a_1}{\sqrt{4-a_1^2}}dz = \frac{S(I)a_1}{\sqrt{4-a_1^2}} = \frac{S(I)\sqrt{4-a^2}}{|a|},
\end{aligned}$$

and by (38) with $z=0$ again, we have

$$\lim_{n\to+\infty} n(2\ln n)^{-\frac{1}{2}} D_n(G_n(x)/2) = e^{c_0-x},$$

which implies

$$\lim_{n\to+\infty} n(2\ln n)^{\frac{1}{2}} \int_{I_1} D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)dy = e^{c_0-x}S(I)\frac{\sqrt{4-a^2}}{|a|}, \tag{43}$$

which finishes the proof by the fact that $S(I)\frac{\sqrt{4-a^2}}{|a|} = (4-a^2)/|a| = M(I)$.

Case 2: $a+b>0$. By symmetry, we can consider $-I=[-b,-a]$ and the result follows **Case 1**.

Case 3: $a+b=0$. We can write $I=I_1\cup I_2$ such that $I_1=[a,0]$, $I_2=[0,b]$, then we have $S(I)=S(I_1)=S(I_2)$, $M(I)=M(I_1)+M(I_2)$, and by the results of **Case 1**, **Case 2** we have

$$\begin{aligned}
&n(2\ln n)^{\frac{1}{2}} \int_I D_n(\sqrt{4-y^2}/S(I)\cdot G_n(x)/2)dy \\
=&\sum_{j=1}^{2} n(2\ln n)^{\frac{1}{2}} \int_{I_j} D_n(\sqrt{4-y^2}/S(I_j)\cdot G_n(x)/2)dy \\
\to& M(I_1)e^{c_0-x} + M(I_2)e^{c_0-x} = M(I)e^{c_0-x},\ \ n\to+\infty,
\end{aligned}$$

this completes the proof. ∎

4.3. **The strategy to prove Theorem 2.** The strategy to prove Theorem 2 is similar to that of Theorem 1, but we will still give all the detailed definitions and computations. Now we consider the point process of eigenvalues of GUE,

$$\xi^{(n)} = \sum_{i=1}^{n} \delta_{\lambda_i}.$$

By definition of $M_0(I)$ in Theorem 2 and $M(I)$ in Lemma 10, we have $M_0(I) = \ln(M(I)S(I)/4)$. Take $c_2 = c_0 + M_0(I)$, $f(x) = e^{c_2 - x} = M(I)S(I)e^{c_0 - x}/4$, then we have $-f'(x) = f''(x) = e^{c_2 - x}$. By Lemma 1, for every positive integer k and $x_1, \cdots, x_k \in \mathbb{R}$, for τ_j^* defined in Theorem 2, if we can prove the following convergence

$$\lim_{n\to+\infty} \mathbb{E} \sum_{i_1,\cdots,i_k \text{ all distinct}} \prod_{j=1}^{k} (\tau_{i_j}^* - x_j)_+ = (M(I)S(I))^k \prod_{j=1}^{k} \left(e^{c_0 - x_j}/4\right), \tag{44}$$

then Theorem 2 will be proved.

For $\lambda_1 < \cdots < \lambda_n$, denote $J_k(a) := \{x \in \mathbb{R} | [x, x+a] \subset (\lambda_k, \lambda_{k+1})\}$ for $a > 0$, $1 \le k < n$, then we have $J_k(a) = (\lambda_k, \lambda_{k+1} - a)$ for $\lambda_{k+1} - \lambda_k > a$ and $J_k(a) = \emptyset$ for $\lambda_{k+1} - \lambda_k \le a$, thus $J_k(a)$ is an interval of size $(\lambda_{k+1} - \lambda_k - a)_+$, and $J_k(a) \subset (\lambda_k, \lambda_{k+1})$ and $J_k(a) \cap J_l(b) = \emptyset$ for $k \ne l$. Now let $\Lambda(I) = \{i | \lambda_i, \lambda_{i+1} \in I\}$,

$$\Sigma_k(a_1, \cdots, a_k) := \bigcup_{i_1,\cdots,i_k \in \Lambda(I) \text{ all distinct}} \prod_{j=1}^{k} J_{i_j}(a_j) \subset (a, b)^k,$$

then the right hand side is a disjoint union and

$$\begin{aligned} |\Sigma_k(a_1, \cdots, a_k)| &= \sum_{i_1,\cdots,i_k \in \Lambda(I) \text{ all distinct}} \prod_{j=1}^{k} (\lambda_{i_j+1} - \lambda_{i_j} - a_j)_+ \\ &= \sum_{i_1,\cdots,i_k \in \Lambda(I) \text{ all distinct}} \prod_{j=1}^{k} (m_{i_j}^* - a_j)_+. \end{aligned}$$

Let $A = 2/S(I) > 1$, thanks to Lemma 9, for every fixed $x_1, \cdots, x_k \in \mathbb{R}$ there exists $N_3 > 0$ such that $0 < 2s_0/n < G_n(x_j) < AG_n(x_j) < \pi$ for $n > N_3$, $1 \le j \le k$. Now we always assume $n > N_3$. By (36) and the fact that $S(I)m_k^* = G_n(\tau_k^*)$, we have $\tau_k^* - x = (G_n(\tau_k^*) - G_n(x))(n/4)(2\ln n)^{\frac{1}{2}} = (S(I)m_k^* - G_n(x))(n/4)(2\ln n)^{\frac{1}{2}}$, and

$$\begin{aligned} &\sum_{i_1,\cdots,i_k \in \Lambda(I) \text{ all distinct}} \prod_{j=1}^{k} (\tau_{i_j}^* - x_j)_+ \\ =&(nS(I)/4)^k (2\ln n)^{\frac{k}{2}} \sum_{i_1,\cdots,i_k \in \Lambda(I) \text{ all distinct}} \prod_{j=1}^{k} (m_{i_j}^* - G_n(x_j)/S(I))_+ \\ =&(nS(I)/4)^k (2\ln n)^{\frac{k}{2}} |\Sigma_k(G_n(x_1)/S(I), \cdots, G_n(x_k)/S(I))|. \end{aligned}$$

For fixed $x_1, \cdots, x_k \in \mathbb{R}$ and $y_1, \cdots, y_k \in I$, let

$$\begin{aligned} &\phi_{k,n}(y_1, \cdots, y_k) = n^k (2\ln n)^{\frac{k}{2}} \times \\ &\mathbb{P}((y_1, \cdots, y_k) \in \Sigma_k(G_n(x_1)/S(I), \cdots, G_n(x_k)/S(I))), \end{aligned}$$

then

$$
\begin{aligned}
&\mathbb{E}\sum_{i_1,\cdots,i_k \text{ all distinct}}\prod_{j=1}^{k}(\tau_{i_j}^{*}-x_j)_+\\
=&\mathbb{E}(nS(I)/4)^k(2\ln n)^{\frac{k}{2}}|\Sigma_k(G_n(x_1)/S(I),\cdots,G_n(x_k)/S(I))|\\
=&(S(I)/4)^k\int_{I^k}\phi_{k,n}(y_1,\cdots,y_k)dy_1\cdots dy_k.
\end{aligned}
$$

Now we prove the following upper bound and lower bound separately

$$
\limsup_{n\to+\infty}\int_{I^k}\phi_{k,n}(y_1,\cdots,y_k)dy_1\cdots dy_k\le (M(I))^k\prod_{j=1}^{k}\left(e^{c_0-x_j}\right), \tag{45}
$$

$$
\liminf_{n\to+\infty}\int_{I^k}\phi_{k,n}(y_1,\cdots,y_k)dy_1\cdots dy_k\ge (M(I))^k\prod_{j=1}^{k}\left(e^{c_0-x_j}\right), \tag{46}
$$

in fact (45) and (46) imply (44), and thus Theorem 2 follows.

4.4. **The proof of Theorem 2.** We first need the following equivalent condition for a point in $\Sigma_k(a_1,\cdots,a_k)$, the proof is similar to that of Lemma 3 and we omit it here.

Lemma 11. *For $(y_1,\cdots,y_k)\in(a,b)^k$, the condition $(y_1,\cdots,y_k)\in\Sigma_k(a_1,\cdots,a_k)$ is equivalent to the following conditions: (i) $[y_l,y_l+a_l]\cap[y_j,y_j+a_j]=\emptyset$ for $1\le l<j\le k$, and (ii) $\lambda_l\notin[y_l,y_l+a_l]$, for $1\le j\le k$, $1\le l\le n$, and (iii) $\{\lambda_1,\cdots,\lambda_n\}\cap[y_p,y_q]\neq\emptyset$, for every $p,q\in\{0,\cdots,k+1\}$, such that $y_p<y_q$, here we denote $y_0=a, y_{k+1}=b$.*

4.4.1. *Upper bound.* Now for fixed $x_1,\cdots,x_k\in\mathbb{R}$, as n large enough, let

$$
\begin{aligned}
A_n:=&\{(y_1,\cdots,y_k)\in(a,b)^k|[y_i,y_i+G_n(x_i)/S(I)]\\
&\cap[y_j,y_j+G_n(x_j)/S(I)]=\emptyset,\forall\ 1\le i<j\le k\},
\end{aligned} \tag{47}
$$

then for $(y_1,\cdots,y_k)\in(a,b)^k\setminus A_n$, by Lemma 11 we have $\phi_{k,n}(y_1,\cdots,y_k)=0$. If $(y_1,\cdots,y_k)\in A_n$, then all y_k's are distinct, let $y_0=a,\ y_{k+1}=b$, and

$$
I_{k,n}=\cup_{j=1}^{k}[y_j,y_j+G_n(x_j)/S(I)],\ J_{k,n,j}=[z_j,z_{j+1}],\ 0\le j\le k, \tag{48}
$$

here z_j $(0\le j\le k+1)$ is the increasing rearrangement of y_j $(0\le j\le k+1)$, then $I_{k,n}$ is a disjoint union and by Lemma 11 we have

$$
\begin{aligned}
\phi_{k,n}(y_1,\cdots,y_k)=&n^k(2\ln n)^{\frac{k}{2}}\times\\
&\mathbb{P}(\xi^{(n)}(I_{n,k})=0,\ \xi^{(n)}(J_{n,k,j})>0,\ \forall\ 0\le j\le k).
\end{aligned} \tag{49}
$$

By Lemma 4 and (20) we have,

$$
\begin{aligned}
&\phi_{k,n}(y_1,\cdots,y_k)\le n^k(2\ln n)^{\frac{k}{2}}\mathbb{P}(\xi^{(n)}(I_{n,k})=0)\\
&\le n^k(2\ln n)^{\frac{k}{2}}\prod_{j=1}^{k}\mathbb{P}(\xi^{(n)}([y_j,y_j+G_n(x_j)/S(I)])=0).
\end{aligned}
$$

and this inequality is clearly true for $(y_1,\cdots,y_k)\notin A_n$. Therefore, we have

$$
\int_{I^k}\phi_{k,n}(y_1,\cdots,y_k)dy_1\cdots dy_k
$$

$$\le \int_{I^k} n^k (2\ln n)^{\frac{k}{2}} \prod_{j=1}^{k} \mathbb{P}(\xi^{(n)}([y_j, y_j + G_n(x_j)/S(I)]) = 0) dy_1 \cdots dy_k$$

$$= \prod_{j=1}^{k} \left[n(2\ln n)^{\frac{1}{2}} \int_I \mathbb{P}(\xi^{(n)}([y_j, y_j + G_n(x_j)/S(I)]) = 0) dy_j \right].$$

Thus, (45) follows if we can prove the following inequality

$$\limsup_{n\to+\infty} n(2\ln n)^{\frac{1}{2}} \int_I \mathbb{P}(\xi^{(n)}([y, y + G_n(x)/S(I)]) = 0) dy \le M(I) e^{c_0 - x}, \tag{50}$$

and by Lemma 10, we only need to prove

$$\limsup_{n\to+\infty} n(2\ln n)^{\frac{1}{2}} \sup_{y\in I} \Big(\mathbb{P}(\xi^{(n)}([y, y + G_n(x)/S(I)]) = 0) - D_n(\sqrt{4-y^2}/S(I) \cdot G_n(x)/2) \Big) \le 0. \tag{51}$$

Let $\{h_n\}$ be the Hermite polynomials, which are the successive monic orthogonal polynomials with respect to the Gaussian weight $e^{-x^2/2}dx$. Following [1], we introduce the functions

$$\psi_k(x) = \frac{e^{-x^2/4}}{\sqrt{\sqrt{2\pi}k!}} h_k(x).$$

Then the set of points $\{\lambda_1, \cdots, \lambda_n\}$ with respect to the joint density (1) is a determinantal point process with the kernel given by [1]

$$K^{GUE(n)}(x, y) = \sqrt{n} \frac{\psi_n(x\sqrt{n})\psi_{n-1}(y\sqrt{n}) - \psi_{n-1}(x\sqrt{n})\psi_n(y\sqrt{n})}{x - y}. \tag{52}$$

The probability that $\xi^{(n)}$ has no point in a measurable subset J is

$$\mathbb{P}(\xi^{(n)}(J) = 0) := \mathbb{P}^{GUE(n)}(\lambda_i \notin J, 1 \le i \le n) = \det(\mathrm{Id} - \chi_J P_{GUE(n)} \chi_J),$$

where $P_{GUE(n)}$ is the orthogonal projection from $L^2(\mathbb{R})$ to $W_n := \mathrm{span}\{x^k e^{-nx^2/4} |0 \le k < n, k \in \mathbb{Z}\}$ with kernel $K^{GUE(n)}(x, y)$.

We will need the following inequality regarding the difference of the gap probabilities between CUE and GUE,

Lemma 12. *Let $\varepsilon_0 \in (0,1)$, $C_0 > c_* > 0$, $\rho_{sc}(x) = \sqrt{(4-x^2)_+}/(2\pi)$. Then uniformly for $x \in (-2+\varepsilon_0, 2-\varepsilon_0)$, $c_*(\ln n)^{\frac{1}{2}}/n < \delta_n < \min(C_0(\ln n)^{\frac{1}{2}}/n, 1/2)$,*

$$\mathbb{P}^{GUE(n)}(\lambda_i \notin [x, x + \delta_n/\rho_{sc}(x)], 1 \le i \le n) - \mathbb{P}^{CUE(n)}(\theta_i \notin [0, 2\pi\delta_n], 1 \le i \le n) \le O((n\ln n)^{-1}).$$

Proof. Let A, B be integral operators with respective kernels

$$A(u, v) = -\frac{1}{n\rho_{sc}(x)} K^{GUE(n)}_{(0,n\delta_n)} \left(x + \frac{u}{n\rho_{sc}(x)}, x + \frac{v}{n\rho_{sc}(x)} \right)$$

and

$$B(u, v) = -\frac{2\pi}{n} K^{CUE(n)}_{(0,n\delta_n)} \left(\frac{2\pi}{n} u, \frac{2\pi}{n} v \right).$$

From the proof of Lemma 3.5 in [2], we know that

$$|A-B|_2=O((\ln n)^{3/2}/n),\ |A|_2^2=O((\ln n)^{2/3}),\ |B|_2^2=O((\ln n)^{2/3}), \tag{53}$$

$$\operatorname{Tr} A=-n\delta_n+O((\ln n)^{3/2}/n),\ \operatorname{Tr} B=-n\delta_n+O((\ln n)^{3/2}/n). \tag{54}$$

We also have

$$\det(\mathrm{Id}+A)=\mathbb{P}^{GUE(n)}(\lambda_i\notin[x,x+\delta_n/\rho_{sc}(x)],1\le i\le n) \tag{55}$$

and

$$\det(\mathrm{Id}+B)=\mathbb{P}^{CUE(n)}(\theta_i\notin[0,2\pi\delta_n],1\le i\le n)=D_n(\pi\delta_n). \tag{56}$$

Since $D_n(\alpha)$ is a continuous function for $\alpha\in[0,\pi]$, $D_n(0)=1$ and $D_n(\pi)=0$, for $n\ge 2$ there exists $\alpha_n\in(0,\pi)$ such that $D_n(\alpha_n)=(n\ln n)^{-1}$. Now we discuss the case $\pi\delta_n\le\alpha_n$ and the case $\pi\delta_n\ge\alpha_n$ separately.

If $\pi\delta_n\le\alpha_n$, recall the general comparison inequalities in Lemma 6, we have

$$\exp(\operatorname{Tr}(B-A)(\mathrm{Id}+B)^{-1})\det(\mathrm{Id}+A)/\det(\mathrm{Id}+B)\le 1, \tag{57}$$

and

$$|\operatorname{Tr}((B-A)(\mathrm{Id}+B)^{-1})|\le|\operatorname{Tr}(A-B)|+|A-B|_2|B|_2\|(\mathrm{Id}+B)^{-1}\|. \tag{58}$$

By Lemma 7 we have

$$\|(\mathrm{Id}+B)^{-1}\|=(1-\lambda_1(-B))^{-1}\le e^{1+\operatorname{Tr}B}(\det(\mathrm{Id}+B))^{-1}. \tag{59}$$

Since $D_n(\alpha)$ is decreasing and $\pi\delta_n\le\alpha_n$, by (56) we have

$$\det(\mathrm{Id}+B)=D_n(\pi\delta_n)\ge D_n(\alpha_n)=(n\ln n)^{-1}. \tag{60}$$

By (53)(54)(58)(59) and the fact that $c_*(\ln n)^{\frac12}/n<\delta_n$, we have

$$\begin{aligned}&|\operatorname{Tr}((B-A)(\mathrm{Id}+B)^{-1})|\\ \le&O((\ln n)^{3/2}/n)+O((\ln n)^{3/2+1/3}/n)e^{1+\operatorname{Tr}B}(\det(\mathrm{Id}+B))^{-1}\end{aligned} \tag{61}$$

and we also have

$$\det(\mathrm{Id}+B)\le e^{\operatorname{Tr}B}=e^{-n\delta_n+O((\ln n)^{3/2}/n)}=e^{O(1)-c_*(\ln n)^{1/2}}=O((\ln n)^{-3}). \tag{62}$$

By (60)(61)(62), we have

$$\begin{aligned}&|\operatorname{Tr}((B-A)(\mathrm{Id}+B)^{-1})|\\ \le&O((\ln n)^{3/2}/n)+O((\ln n)^{3/2+1/3-3}/n)(\det(\mathrm{Id}+B))^{-1}\\ \le&O((\ln n)^2/n)+O((\ln n)^{-7/6}/n)(n\ln n)=O(1),\end{aligned} \tag{63}$$

and thus we have

$$|\exp(-\operatorname{Tr}(B-A)(\mathrm{Id}+B)^{-1})-1|=O(|\operatorname{Tr}((B-A)(\mathrm{Id}+B)^{-1})|),$$

and we further have (using (57)(62)(63))

$$\begin{aligned}&\det(\mathrm{Id}+A)-\det(\mathrm{Id}+B)\\ \le&\exp(-\operatorname{Tr}(B-A)(\mathrm{Id}+B)^{-1})\det(\mathrm{Id}+B)-\det(\mathrm{Id}+B)\\ \le&O(|\operatorname{Tr}(B-A)(\mathrm{Id}+B)^{-1}|)\det(\mathrm{Id}+B)\\ \le&O((\ln n)^{3/2}/n)\det(\mathrm{Id}+B)+O((\ln n)^{3/2+1/3-3}/n)\\ \le&O((\ln n)^2/n)O((\ln n)^{-3})+O((\ln n)^{-1}/n)=O((n\ln n)^{-1}).\end{aligned}$$

Now the result follows from the identities (55) and (56).

If $\pi\delta_n \geq \alpha_n$, then we have (taking $\delta'_n = \alpha_n/\pi \leq \delta_n$)

$$\begin{aligned}
&\mathbb{P}^{GUE(n)}(\lambda_i \notin [x, x+\delta_n/\rho_{sc}(x)], 1 \leq i \leq n)\\
\leq&\mathbb{P}^{GUE(n)}(\lambda_i \notin [x, x+\delta'_n/\rho_{sc}(x)], 1 \leq i \leq n)\\
\leq&\mathbb{P}^{CUE(n)}(\theta_i \notin [0, 2\pi\delta'_n], 1 \leq i \leq n) + O((n\ln n)^{-1}) = O((n\ln n)^{-1}),
\end{aligned}$$

and the result is also true, here we used the fact that

$$\mathbb{P}^{CUE(n)}(\theta_i \notin [0, 2\pi\delta'_n], 1 \leq i \leq n) = D_n(\pi\delta'_n) = D_n(\alpha_n) = (n\ln n)^{-1}.$$

This completes the proof. □

Now we prove (51). For $y \in I,\ x \in \mathbb{R}$, take $\delta_n = [\sqrt{4-y^2}/S(I)] \cdot [G_n(x)/(2\pi)]$, then we have $\delta_n/\rho_{sc}(y) = 2\pi\delta_n/\sqrt{4-y^2} = G_n(x)/S(I)$. By (37), there exists a constant $N_4 > 0$ depending only on x such that $4(\ln n)^{\frac{1}{2}}/n < G_n(x) < 8(\ln n)^{\frac{1}{2}}/n < \pi S(I)/2$ for $n > N_4$. Then we have $(2/\pi)(\ln n)^{\frac{1}{2}}/n < G_n(x)/(2\pi) \leq [\sqrt{4-y^2}/S(I)]\cdot [G_n(x)/(2\pi)] = \delta_n \leq [2/S(I)] \cdot [G_n(x)/(2\pi)] < (\pi S(I))^{-1} \cdot 8(\ln n)^{\frac{1}{2}}/n < 1/2$ for $y \in I,\ n > N_4$, thus by Lemma 12 we deduce that

$$\begin{aligned}
&\mathbb{P}(\xi^{(n)}([y, y+G_n(x)/S(I)]) = 0) - D_n(\sqrt{4-y^2}/S(I) \cdot G_n(x)/2)\\
=&\mathbb{P}(\xi^{(n)}([y, y+\delta_n/\rho_{sc}(y)]) = 0) - D_n(\pi\delta_n)\\
=&\mathbb{P}^{GUE(n)}(\lambda_i \notin [y, y+\delta_n/\rho_{sc}(y)], 1 \leq i \leq n)\\
&- \mathbb{P}^{CUE(n)}(\theta_i \notin [0, 2\pi\delta_n], 1 \leq i \leq n) \leq O((n\ln n)^{-1}),
\end{aligned}$$

and the estimate is uniform for $y \in I,\ n > N_4$. Thus we have

$$\begin{aligned}
&n(2\ln n)^{\frac{1}{2}} \sup_{y\in I}\Big(\mathbb{P}(\xi^{(n)}([y, y+G_n(x)/S(I)]) = 0) - D_n(\sqrt{4-y^2}/S(I) \cdot G_n(x)/2)\Big)\\
&\leq n(2\ln n)^{\frac{1}{2}} O((n\ln n)^{-1}) = O((\ln n)^{-1/2}) \to 0, \quad n \to +\infty,
\end{aligned}$$

and thus (51) is true, so is (50) and hence the upper bound (45).

4.4.2. *Lower bound.* For the lower bound (46), we discuss the 3 cases separately.

Case 1: $a+b < 0$. Let $b_0 = (a+b)/2 < 0,\ I_1 = (a, b_0) \subset I,\ a_* = \sqrt{4-a^2} = S(I)$, $b_* = \sqrt{4-b_0^2} > a_*$. We change variables $y_j = -\sqrt{4-v_j^2},\ 0 < v_j = (1+u_j/\ln n)a_*$ to obtain

$$\begin{aligned}
&\int_{I^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k \geq \int_{I_1^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k\\
=&\int_{(a_*, b_*)^k} \phi_{k,n}\left(-\sqrt{4-v_1^2}, \cdots, -\sqrt{4-v_k^2}\right) \prod_{j=1}^{k} \frac{v_j}{\sqrt{4-v_1^2}} dv_1 \cdots dv_k\\
=&a_*^k (\ln n)^{-k} \int_{(0,(b_*/a_*-1)\ln n)^k} \phi_{k,n}\Big(-\sqrt{4-(1+u_1/\ln n)^2 a_*^2}, \cdots,\\
&-\sqrt{4-(1+u_k/\ln n)^2 a_*^2}\Big) \prod_{j=1}^{k} \frac{(1+u_j/\ln n)a_*}{\sqrt{4-(1+u_j/\ln n)^2 a_*^2}} du_1 \cdots du_k.
\end{aligned}$$

Denote $l_n = (b_*/a_* - 1)\ln n$ and

$$\gamma_n(u) = -\sqrt{4-(1+u/\ln n)^2 S(I)^2}, \quad \beta_n(u) = \frac{(1+u/\ln n)S(I)}{\sqrt{4-(1+u/\ln n)^2 S(I)^2}}, \tag{64}$$

then γ_n maps $(0, l_n)$ to $I_1 \subset (a,b)$ and

$$\int_{I^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k \geq S(I)^k (\ln n)^{-k} \int_{(0,l_n)^k} \phi_{k,n}\big(\gamma_n(u_1), \cdots, \gamma_n(u_k)\big) \prod_{j=1}^{k} \beta_n(u_j) du_1 \cdots du_k. \tag{65}$$

Case 2: $a+b > 0$. Let $b_0 = (a+b)/2 > 0$, $I_1 = (b_0, b) \subset I$, $a_* = \sqrt{4-b^2} = S(I)$, $b_* = \sqrt{4-b_0^2} > a_*$, $l_n = (b_*/a_* - 1)\ln n$ and

$$\gamma_n(u) = \sqrt{4 - (1-u/\ln n)^2 S(I)^2}, \quad \beta_n(u) = \frac{(1-u/\ln n)S(I)}{\sqrt{4-(1-u/\ln n)^2 S(I)^2}}. \tag{66}$$

Similar to **Case 1** we have $\gamma_n : (-l_n, 0) \to I_1 \subset (a,b)$ and

$$\int_{I^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k \geq \int_{I_1^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k = S(I)^k (\ln n)^{-k} \int_{(-l_n,0)^k} \phi_{k,n}\big(\gamma_n(u_1), \cdots, \gamma_n(u_k)\big) \prod_{j=1}^{k} \beta_n(u_j) du_1 \cdots du_k. \tag{67}$$

Case 3: $a + b = 0$. Let $a_0 = a/2 < 0$, $b_0 = b/2 = -a_0 > 0$, $I_1 = (a, a_0) \cup (b_0, b) \subset I$, $a_* = \sqrt{4-a^2} = \sqrt{4-b^2} = S(I)$, $b_* = \sqrt{4-a_0^2} = \sqrt{4-b_0^2} > a_*$, $l_n = (b_*/a_* - 1)\ln n$ and functions $\gamma_n(u), \beta_n(u)$ be defined as (64) for $u > 0$ and as (66) for $u < 0$. Similar to **Case 1** we have $\gamma_n : (-l_n, l_n) \setminus \{0\} \to I_1 \subset (a,b)$ and

$$\int_{I^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k \geq \int_{I_1^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k = S(I)^k (\ln n)^{-k} \int_{(-l_n,l_n)^k} \phi_{k,n}\big(\gamma_n(u_1), \cdots, \gamma_n(u_k)\big) \prod_{j=1}^{k} \beta_n(u_j) du_1 \cdots du_k. \tag{68}$$

Now the lower bound (46) is the consequence of the following

Lemma 13. *For fixed $I = [a,b] \subset (-2,2)$, $k \in \mathbb{Z}$, $k > 0$, $x_1, \cdots, x_k \in \mathbb{R}$, let $\gamma_n(u)$ be defined as (64) for $u > 0$ and as (66) for $u < 0$. Assume that (i) $a+b < 0$, $u_1, \cdots, u_k \in (0, +\infty)$ all distinct, or (ii) $a+b > 0$, $u_1, \cdots, u_k \in (-\infty, 0)$ all distinct, or (iii) $a+b = 0$, $u_1, \cdots, u_k \in \mathbb{R} \setminus \{0\}$ and $|u_j|$'s are all distinct, then we have*

$$\liminf_{n \to +\infty} (\ln n)^{-k} \phi_{k,n}\big(\gamma_n(u_1), \cdots, \gamma_n(u_k)\big) \geq 2^k e^{\sum_{j=1}^{k}(c_0 - x_j - 2|u_j|)}.$$

Lemma 13 will imply the lower bound (46) as follows.

For the case $a+b < 0$, denote $I_0 = (0, +\infty)$, then we have $\int_{I_0} 2e^{-2|u|} du = 1$, $S(I) = \sqrt{4-a^2}$ and $S(I)^2/\sqrt{4-S(I)^2} = (4-a^2)/|a| = M(I)$.

Since $l_n \to +\infty$, $\beta_n(u_j) \to S(I)/\sqrt{4-S(I)^2}$ as $n \to +\infty$, by (65), Lemma 13 and Fatou's Lemma, we have

$$\liminf_{n \to +\infty} \int_{I^k} \phi_{k,n}(y_1, \cdots, y_k) dy_1 \cdots dy_k \geq S(I)^k \int_{I_0^k} \liminf_{n \to +\infty} \left[(\ln n)^{-k} \phi_{k,n}\big(\gamma_n(u_1), \cdots, \gamma_n(u_k)\big) \prod_{j=1}^{k} \beta_n(u_j) \right] du_1 \cdots du_k$$

$$\geq S(I)^k \int_{I_0^k} 2^k e^{\sum_{j=1}^k (c_0 - x_j - 2|u_j|)} \left(S(I)/\sqrt{4-S(I)^2}\right)^k du_1 \cdots du_k$$

$$= \left(S(I)^2/\sqrt{4-S(I)^2}\right)^k \prod_{j=1}^k \left(e^{c_0-x_j}\right) \int_{I_0^k} \prod_{j=1}^k \left(2e^{-2|u_j|}\right) du_1 \cdots du_k$$

$$= \left(S(I)^2/\sqrt{4-S(I)^2}\right)^k \prod_{j=1}^k \left(e^{c_0-x_j}\right) = (M(I))^k \prod_{j=1}^k \left(e^{c_0-x_j}\right).$$

For the cases when $a+b>0$ and $a+b=0$, the proof follows similarly. This completes the proof of the lower bound (46), and hence Theorem 2.

All of the rest effort is to prove Lemma 13. We first need a lower bound of $\mathbb{P}(\xi^{(n)}(J)=0)$ when J is a finite union of intervals.

Lemma 14. *Let* $\varepsilon_0 \in (0,1)$, $C_0>0$, $k \in \mathbb{Z}^+$, $I=[a,b]=[y_0,y_{k+1}] \subset (-2,2)$. *Assume* $y_1,\cdots,y_k \in I$, $a_1,\cdots,a_k \in (G_n(-C_0)/S(I), G_n(C_0)/S(I)) \cap (0, \varepsilon_0(2\ln n)^{-1})$, $|y_i - y_j| \geq \varepsilon_0(\ln n)^{-1}$ *for every* $0 \leq i < j \leq k+1$, $\sqrt{4-y_i^2}/S(I) \leq 1 + C_0(\ln n)^{-1}$ *for every* $1 \leq i \leq k$. *Then there exists a constant* $N_5>0$ *depending only on* ε_0, C_0, k, I *such that for* $n > N_5$ *we have*

$$\mathbb{P}(\xi^{(n)}(\cup_{j=1}^k [y_j, y_j+a_j])=0) \geq (1-(\ln n)^{-1}) \prod_{j=1}^k D_n\left(a_j\sqrt{4-y_j^2}/2\right).$$

Proof. We use $f=O(g)$ to denote $|f| \leq Cg$ for a constant C depending only on ε_0, C_0, k, I. As $|y_i-y_j| \geq \varepsilon_0(\ln n)^{-1} > \varepsilon_0(2\ln n)^{-1}$ for $i \neq j$, if $1 \leq j \leq k$, then $y_0 \leq y_j < y_j + a_j < y_j + \varepsilon_0(2\ln n)^{-1} < y_j + |y_{k+1}-y_j| = y_{k+1}$, and thus $[y_j, y_j+a_j] \subset [y_0, y_{k+1}] = I$. If $1 \leq i < j \leq k$, then by assumption $a_i, a_j \in (0, \varepsilon_0(2\ln n)^{-1}) \subset (0, |y_i-y_j|)$, and thus $[y_i, y_i+a_i] \cap [y_j, y_j+a_j] = \emptyset$. Therefore, we have $J := \cup_{j=1}^k [y_j, y_j+a_j]$ is a disjoint union and $J \subset I$. Let's denote

$$A = \chi_J P_{GUE(n)} \chi_J, \quad A_{i,j} = \chi_{[y_i, y_i+a_i]} P_{GUE(n)} \chi_{[y_j, y_j+a_j]},$$

then we have

$$A = \sum_{i=1}^k \sum_{j=1}^k A_{i,j}$$

and

$$\mathbb{P}(\xi^{(n)}(\cup_{j=1}^k [y_j, y_j+a_j])=0) = \mathbb{P}(\xi^{(n)}(J)=0) = \det(\mathrm{Id}-A). \tag{69}$$

Let B_j be the integral operator with kernel

$$B_j(u,v) = 2\pi\rho_{sc}(y_j) K^{CUE(n)}_{(y_j, y_j+a_j)}\left(2\pi\rho_{sc}(y_j)u, 2\pi\rho_{sc}(y_j)v\right),$$

where $K^{CUE(n)}(x,y)$ is the kernel defined in (25). Let's denote

$$B = \sum_{j=1}^k B_j.$$

As $0 < a_j\sqrt{4-y_j^2}/2 \leq a_j < \varepsilon_0(2\ln n)^{-1} < 1$, we have

$$\begin{aligned} \det(\mathrm{Id}-B_j) &= \mathbb{P}^{CUE(n)}(\theta_i \notin [0, 2\pi\rho_{sc}(y_j)a_j], 1 \leq i \leq n) \\ &= D_n(\pi\rho_{sc}(y_j)a_j) = D_n\left(a_j\sqrt{4-y_j^2}/2\right), \end{aligned} \tag{70}$$

and

$$\det(\mathrm{Id}-B)=\prod_{j=1}^{k}\det(\mathrm{Id}-B_j)=\prod_{j=1}^{k}D_n\left(a_j\sqrt{4-y_j^2}/2\right). \tag{71}$$

Now we need to compare the Fredholm determinants, the key point is to estimate $|A-B|_2,\ \mathrm{Tr}(A-B),\ \|(\mathrm{Id}-B)^{-1}\|$. Comparing the support of the kernels, we have

$$|A-B|_2^2=\sum_{j=1}^{k}|A_{j,j}-B_j|_2^2+\sum_{i\neq j}|A_{i,j}|_2^2,\ \mathrm{Tr}(A-B)=\sum_{j=1}^{k}\mathrm{Tr}(A_{j,j}-B_j). \tag{72}$$

For $x\in[y_i,y_i+a_i]\subset I,\ y\in[y_j,y_j+a_j]\subset I,\ i\neq j,\ (1\le i,j\le k)$, we have

$$|x-y|\ge|y_i-y_j|-\max(a_i,a_j)\ge\varepsilon_0(\ln n)^{-1}-\varepsilon_0(2\ln n)^{-1}=\varepsilon_0(2\ln n)^{-1}.$$

From the Plancherel-Rotach asymptotics for the Hermite polynomials (Theorem 8.22.9 in [10]) for any nonnegative integer j, $\psi_{n-j}(\sqrt{n}x)$ is $O(n^{-1/4})$, uniformly in $x\in I$. Consequently, if $|x-y|\ge\varepsilon_0(2\ln n)^{-1},\ x,y\in I$, from (52), we have

$$|K^{GUE(n)}(x,y)|=\sqrt{n}\frac{O(n^{-1/4})O(n^{-1/4})}{|x-y|}=\frac{O(1)}{|x-y|}\le\frac{O(1)}{\varepsilon_0(2\ln n)^{-1}}=O(\ln n).$$

Using this and (37), for $i\neq j$ we have (recall that $0<a_i<G_n(C_0)/S(I)$)

$$\begin{aligned}|A_{i,j}|_2^2&=\int_{[y_i,y_i+a_i]}dx\int_{[y_j,y_j+a_j]}|K^{GUE(n)}(x,y)|^2dy\\&=\int_{[y_i,y_i+a_i]}dx\int_{[y_j,y_j+a_j]}O((\ln n)^2)dy=a_ia_jO((\ln n)^2)\\&\le(G_n(C_0)/S(I))^2O((\ln n)^2)=O\left(\frac{\ln n}{n^2}\right)O((\ln n)^2)=O\left(\frac{(\ln n)^3}{n^2}\right).\end{aligned} \tag{73}$$

Since $a_j=O(\varepsilon_0(2\ln n)^{-1})=o(1),\ 0<S(I)\le\sqrt{4-y_j^2}=2\pi\rho_{sc}(y_j)\le 2$, and the kernel of $A_{j,j}$ is $A_{j,j}(u,v)=K^{GUE(n)}_{(y_j,y_j+a_j)}(u,v)$, by Lemma 3.4 in [2] we have

$$\frac{1}{n\rho_{sc}(x)}K^{GUE(n)}(x,y)-\frac{\sin(n\pi\rho_{sc}(x)(x-y))}{n\pi\rho_{sc}(x)(x-y)}=O\left(\frac{1}{n}\right)+O\left(a_j\right)+O\left(na_j^2\right),$$

$$\frac{2\pi}{n}K^{CUE(n)}(2\pi\rho_{sc}(y_j)x,2\pi\rho_{sc}(y_j)y)-\frac{\sin(n\pi\rho_{sc}(y_j)(x-y))}{n\pi\rho_{sc}(y_j)(x-y)}=O\left(\frac{a_j}{n}\right),$$

uniformly for $x,y\in[y_j,y_j+a_j]$. Thus the difference between the two kernels $A_{j,j}$ and B_j is $O(1+n^2a_j^2)$, integrating on a domain $[y_j,y_j+a_j]^2$ of area a_j^2, we have

$$|A_{j,j}-B_j|_2^2=O((1+n^2a_j^2)^2)a_j^2=O(a_j^2+n^4a_j^6);$$

and integrating on the diagonal $\{x=y\in[y_j,y_j+a_j]\}$ yields

$$|\mathrm{Tr}(A_{j,j}-B_j)|=O((1+n^2a_j^2))a_j=O((a_j^2+n^4a_j^6)^{1/2}).$$

Using $0<a_j<G_n(C_0)/S(I)$ and (37), we have

$$a_j^2\le(G_n(C_0)/S(I))^2=O\left(\frac{\ln n}{n^2}\right), \tag{74}$$

thus

$$|A_{j,j}-B_j|_2^2=O(a_j^2+n^4a_j^6)=O\left(\frac{\ln n}{n^2}+\frac{(\ln n)^3}{n^2}\right)=O\left(\frac{(\ln n)^3}{n^2}\right), \tag{75}$$

and

$$|\operatorname{Tr}(A_{j,j}-B_j)| = O((a_j^2+n^4a_j^6)^{1/2}) = O\left(\frac{(\ln n)^{3/2}}{n}\right). \tag{76}$$

Using (72)(73)(75)(76), we conclude that

$$|A-B|_2^2 = O\left(\frac{(\ln n)^3}{n^2}\right),\ |\operatorname{Tr}(A-B)| = O\left(\frac{(\ln n)^{3/2}}{n}\right). \tag{77}$$

Recall the formula (25), we have $K^{CUE(n)}(x,x) = \dfrac{n}{2\pi}$ and

$$|K^{CUE(n)}(x,y)| = O\left(\frac{n}{1+n|x-y|}\right),\ \ |x-y|\le 2.$$

Therefore, by definition of B_j, we have

$$B_j(u,u) = 2\pi\rho_{sc}(y_j)\frac{n}{2\pi} = n\rho_{sc}(y_j),\ u\in(y_j,y_j+a_j)$$

and

$$\operatorname{Tr} B_j = \int_{y_j}^{y_j+a_j} B_j(u,u)du = na_j\rho_{sc}(y_j) = na_i\sqrt{4-y_i^2}/(2\pi); \tag{78}$$

since $0 < 2\pi\rho_{sc}(y_j)a_j = \sqrt{4-y_j^2}a_j \le 2a_j < 2$ and $0 < S(I) \le \sqrt{4-y_j^2} = 2\pi\rho_{sc}(y_j)\le 2$, thus we have the off-diagonal estimate

$$|B_j(u,v)| = O\left(\frac{n}{1+n|u-v|}\right),\ u,v\in(y_j,y_j+a_j).$$

Therefore, we have

$$\begin{aligned}|B_j|_2^2 &= \int_{y_j}^{y_j+a_j}\int_{y_j}^{y_j+a_j}|B_j(u,v)|^2dudv\\ &= \int_{y_j}^{y_j+a_j}\int_{y_j}^{y_j+a_j} O\left(\frac{n^2}{(1+n|u-v|)^2}\right)dudv\\ &= \int_{y_j}^{y_j+a_j} O\left(\int_{\mathbb{R}}\frac{n^2}{(1+n|u-v|)^2}du\right)dv\\ &= \int_{y_j}^{y_j+a_j} O(n)\,dv = O(na_j) = O\left((\ln n)^{1/2}\right),\end{aligned}$$

here we used (74). Therefore, we have

$$|B|_2^2 = \sum_{j=1}^k |B_j|_2^2 = O\left((\ln n)^{1/2}\right). \tag{79}$$

Now we estimate $\|(\mathrm{Id}-B)^{-1}\|$. We have $\|(\mathrm{Id}-B)^{-1}\| = (1-\lambda_1(B))^{-1}$ where $\lambda_1(B)$ is the largest eigenvalue of B. Similar to the CUE case as in (33), we know that $\lambda_1(B)\le\lambda_1(B_i)$ for some $1\le i\le k$ or $\lambda_1(B)=0$. For every $1\le i\le k$, by Lemma 7, (70) and (78), we have

$$1-\lambda_1(B_i) \ge \det(\mathrm{Id}-B_i)e^{\operatorname{Tr}B_i-1} = D_n\left(a_i\sqrt{4-y_i^2}/2\right)e^{na_i\sqrt{4-y_i^2}/(2\pi)-1}.$$

By (37)(38) and $32 > \pi^2$, there exists a constant $N_{5,0} > 0$ such that

$$\pi(\ln n)^{\frac{1}{2}} < nG_n(-C_0),$$

and

$$n(\ln n)^{-\frac{1}{2}} D_n((1 + C_0/\ln n)G_n(C_0)/2) > e^{c_0 - 3C_0}$$

and $G_n(C_0) < 1,\ C_0 < \ln n$ for $n > N_{5,0}$. By assumption, $a_i < G_n(C_0)/S(I)$ and $\sqrt{4 - y_i^2}/S(I) \le 1 + C_0(\ln n)^{-1}$, we have $a_i\sqrt{4 - y_i^2} < (1 + C_0/\ln n)G_n(C_0)$. Since $a_i > G_n(-C_0)/S(I)$, $\sqrt{4 - y_i^2}/S(I) \ge 1$ for $y_i \in I$, we have $a_i\sqrt{4 - y_i^2} > G_n(-C_0)$. Thus if $n > N_{5,0},\ 1 \le i \le k$, we have

$$\begin{aligned} 1 - \lambda_1(B_i) &\ge D_n\left(a_i\sqrt{4 - y_i^2}/2\right) e^{na_i\sqrt{4-y_i^2}/(2\pi)-1} \\ &\ge D_n\left((1 + C_0/\ln n)G_n(C_0)/2\right) e^{nG_n(-C_0)/(2\pi)-1} \\ &\ge n^{-1}(\ln n)^{\frac{1}{2}} e^{c_0 - 3C_0} e^{(\ln n)^{\frac{1}{2}}/2-1}. \end{aligned}$$

Now we always assume $n > N_{5,0}$, then similar to the CUE case, we have

$$\begin{aligned} \|(\mathrm{Id} - B)^{-1}\| &= (1 - \lambda_1(B))^{-1} \le \max_{1\le i\le k}(1 - \lambda_1(B_i))^{-1} + 1 \\ &\le n(\ln n)^{-\frac{1}{2}} e^{3C_0 - c_0 - (\ln n)^{\frac{1}{2}}/2+1} + 1 = O\left(n(\ln n)^{-\frac{1}{2}} e^{-(\ln n)^{\frac{1}{2}}/2}\right). \end{aligned} \tag{80}$$

By Lemma 6 and (77)(79)(80), we conclude that

$$\begin{aligned} b_2 :=& |\operatorname{Tr}((A - B)(\mathrm{Id} - B)^{-1})| \le |\operatorname{Tr}(A - B)| + |A - B|_2|B|_2\|(\mathrm{Id} - B)^{-1}\| \\ \le& O\left(\frac{(\ln n)^{3/2}}{n}\right) + O\left(\frac{(\ln n)^{3/2+1/4}}{n}\right) O\left(n(\ln n)^{-\frac{1}{2}} e^{-(\ln n)^{\frac{1}{2}}/2}\right) \\ =& O\left(\frac{(\ln n)^{3/2}}{n}\right) + O\left((\ln n)^{5/4} e^{-(\ln n)^{\frac{1}{2}}/2}\right) = O\left((\ln n)^{-2}\right), \end{aligned}$$

that

$$\begin{aligned} b_3 :=& |B - A|_2^2\|(\mathrm{Id} - B)^{-1}\|^2 \le O\left(\frac{(\ln n)^3}{n^2}\right) O\left(n^2(\ln n)^{-1} e^{-(\ln n)^{\frac{1}{2}}}\right) \\ =& O\left((\ln n)^2 e^{-(\ln n)^{\frac{1}{2}}}\right) = O\left((\ln n)^{-2}\right). \end{aligned}$$

Therefore, by Lemma 6 again, we have

$$\begin{aligned} 1 - b_3 =& 1 - |B - A|_2^2\|(\mathrm{Id} - B)^{-1}\|^2 \\ \le& \exp(\operatorname{Tr}(A - B)(\mathrm{Id} - B)^{-1}) \det(\mathrm{Id} - A)/\det(\mathrm{Id} - B) \\ \le& e^{b_2} \det(\mathrm{Id} - A)/\det(\mathrm{Id} - B). \end{aligned}$$

Thus there exists a constant $N_5 > N_{5,0}$ such that $b_2 < (2\ln n)^{-1} < 1,\ b_3 < (2\ln n)^{-1} < 1$ for $N > N_5$ and

$$\begin{aligned} \det(\mathrm{Id} - A)/\det(\mathrm{Id} - B) &\ge e^{-b_2}(1 - b_3) \ge (1 - b_2)(1 - b_3) \\ &\ge (1 - (2\ln n)^{-1})^2 \ge 1 - (\ln n)^{-1},\ \forall\, n > N_5. \end{aligned} \tag{81}$$

Now the result follows from (69)(71) and (81). □

Now we prove Lemma 13.

Proof. Let $u_0 = 0$ and

$$C_0 = \max_{1\le j\le k}(|x_j| + |u_j|),\ \varepsilon_1 = \min_{0\le i<j\le k}\big||u_i| - |u_j|\big|,\ \varepsilon_0 = \varepsilon_1 S(I)^2/(2+4\varepsilon_1). \tag{82}$$

Using $l_n = (b_*/a_* - 1)\ln n \to +\infty$ and (37), there exists a constant $N_{6,0} > 2$ such that $l_n > C_0$ and $0 < 4(\ln n)^{\frac{1}{2}}/n < G_n(-C_0) \le G_n(C_0) < 8(\ln n)^{\frac{1}{2}}/n$ for $n > N_{6,0}$. Let's denote

$$y_j = \gamma_n(u_j),\ a_j = G_n(x_j)/S(I),\quad \forall\ n > N_{6,0}. \tag{83}$$

Then we have $y_j \in (a, b)$ for $1 \le j \le k,\ n > N_{6,0}$ (See the range of γ_n in **Case 1-Case 3**). Now we need to check all assumptions in Lemma 14.

(a) Since $u_j \ne 0$, we have $C_0 > 0$. Since $|u_1|, \cdots, |u_k|$ are nonzero and all distinct in all the 3 cases, we have $\varepsilon_1 > 0$. Using this and $0 < S(I) \le 2$, we have

$$0 < \varepsilon_0 = \varepsilon_1 S(I)^2/(2+4\varepsilon_1) \le 4\varepsilon_1/(2+4\varepsilon_1) < 1.$$

(b) By (64)(66), we have $(\gamma_n(u))^2 = 4 - (1+|u|/\ln n)^2 S(I)^2$. Thus by (83), we have $y_j^2 = (\gamma_n(u_j))^2 = 4 - (1+|u_j|/\ln n)^2 S(I)^2$ and

$$\sqrt{4-y_j^2} = (1+|u_j|/\ln n)S(I),\ \sqrt{4-y_j^2}\,a_j = (1+|u_j|/\ln n)G_n(x_j). \tag{84}$$

For $1 \le i < j \le k,\ n > N_{6,0}$, we have $y_i, y_j \in (a,b) \subset (-2,2),\ |y_i + y_j| < 4$ and

$$\begin{aligned}
&4|y_i - y_j| \ge |y_i + y_j|\cdot|y_i - y_j| = \big|y_i^2 - y_j^2\big| = \big|(\gamma_n(u_i))^2 - (\gamma_n(u_j))^2\big|\\
=&\big|(1+|u_j|/\ln n)^2 - (1+|u_i|/\ln n)^2\big|\,S(I)^2\\
=&\big||u_j| - |u_i|\big|/\ln n\cdot(2+|u_j|/\ln n + |u_i|/\ln n)S(I)^2\\
\ge&\big||u_j| - |u_i|\big|/\ln n\cdot 2S(I)^2 \ge \varepsilon_1/\ln n\cdot 2S(I)^2,
\end{aligned}$$

thus we have

$$|y_i - y_j| \ge \varepsilon_1 S(I)^2/(2\ln n) = \varepsilon_0(1+2\varepsilon_1)(\ln n)^{-1} \ge \varepsilon_0(\ln n)^{-1}.$$

Actually, the similar arguments apply to the end points y_0 and y_{k+1} and we finally have $|y_i - y_j| \ge \varepsilon_0(\ln n)^{-1}$ for $0 \le i < j \le k+1$.

(c) For every $1 \le i \le k$ and $n > N_{6,0}$, we have $|u_i| \le C_0$, and by (84), we have

$$\sqrt{4-y_i^2}/S(I) = (1+|u_i|/\ln n)S(I)/S(I) = 1 + |u_i|/\ln n \le 1 + C_0/\ln n.$$

(d) For every $1 \le j \le k$ and $n > N_{6,0}$, since $a_j = G_n(x_j)/S(I),\ S(I) > 0,\ |x_j| < |x_j| + |u_j| \le C_0$ and G_n is increasing, we have $0 < 4(\ln n)^{\frac{1}{2}}/n < G_n(-C_0) < G_n(x_j) = a_j S(I) < G_n(C_0) < 8(\ln n)^{\frac{1}{2}}/n$ and thus $a_j \in (G_n(-C_0)/S(I), G_n(C_0)/S(I)) \cap (0, 8(\ln n)^{\frac{1}{2}}/(nS(I)))$. Since $\varepsilon_0 > 0,\ S(I) > 0$, there exists a constant $N_{6,1} > N_{6,0}$ such that $16(\ln n)^{\frac{3}{2}}/n < \varepsilon_0 S(I)$ for $n > N_{6,1}$. Thus $8(\ln n)^{\frac{1}{2}}/(nS(I)) < \varepsilon_0(2\ln n)^{-1}$ and we have $a_j \in (0, \varepsilon_0(2\ln n)^{-1})$ for $1 \le j \le k$ and $n > N_{6,1}$.

From the statements (a)-(d), we know that ε_0, C_0 defined in (82) and (83) satisfy all the assumptions in Lemma 14 for $n > N_{6,1}$. Thus $[y_i, y_i + G_n(x_i)/S(I)] \cap [y_j, y_j + G_n(x_j)/S(I)] = [y_i, y_i + a_i] \cap [y_j, y_j + a_j] = \emptyset$ for every $1 \le j \le k$ and $n > N_{6,1}$, then $(y_1, \cdots, y_n) \in A_n$ (recall (47)) for $n > N_{6,1}$ and we can use the notation (48) and formula (49) in this case. For $n > N_{6,1}$, by (49) we have

$$(\ln n)^{-k}\phi_{k,n}(y_1, \cdots, y_k) \ge (2n)^k(2\ln n)^{-\frac{k}{2}}\mathbb{P}(\xi^{(n)}(I_{n,k}) = 0) \tag{85}$$

$$- (2n)^k (2\ln n)^{-\frac{k}{2}} \sum_{j=0}^{k} \mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0).$$

As in the CUE case, we claim that

$$(2n)^k (2\ln n)^{-\frac{k}{2}} \mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0) \to 0.$$

Since $a_j = G_n(x_j)/S(I)$, by (48) we have $I_{n,k} = \cup_{j=1}^k [y_j, y_j + a_j]$. Let $d_0 := G_n(-C_0)/S(I)$, then we have $0 < d_0 < a_j < \varepsilon_0(2\ln n)^{-1}$ for $1 \le j \le k$ and $n > N_{6,1}$. Let $z'_j = (z_j + z_{j+1})/2$ for $0 \le j \le k$ where z_j $(0 \le j \le k+1)$ is the increasing rearrangement of y_j $(0 \le j \le k+1)$. Since $y_j \in I$ $(0 \le j \le k+1)$, we have $z_j \in I$ $(0 \le j \le k+1)$, $z'_j \in I$ $(0 \le j \le k)$ and

$$\begin{aligned}\min_{0\le i\le k+1} |z'_j - z_i| &= z_{j+1} - z'_j = z'_j - z_j = (z_{j+1} - z_j)/2 \\ &\ge \min_{0\le i<l\le k+1} |y_i - y_l|/2 \ge \varepsilon_0 (2\ln n)^{-1} > d_0\end{aligned}$$

for $0 \le j \le k$ and $n > N_{6,1}$. Thus $[z'_j, z'_j + d_0] \cap [z_i, z_i + d_0] = \emptyset$ and $[z'_j, z'_j + d_0] \subset [z_j, z_{j+1}] = J_{n,k,j}$ for $0 \le j \le k$, $0 \le i \le k$ and $n > N_{6,1}$. Since $d_0 < a_j$, we have $I_{n,k} \supseteq \cup_{j=1}^k [y_j, y_j + d_0] = \cup_{j=1}^k [z_j, z_j + d_0]$, and $I_{n,k} \cup J_{n,k,j} \supseteq [z'_j, z'_j + d_0] \cup \left(\cup_{i=1}^k [z_i, z_i + d_0]\right)$, and the right hand side is a disjoint union for $n > N_{6,1}$. By (20) we have

$$\begin{aligned} 0 \le & \mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0) = \mathbb{P}(\xi^{(n)}(I_{n,k} \cup J_{n,k,j}) = 0) \\ \le & \mathbb{P}([z'_j, z'_j + d_0] \cup \left(\cup_{i=1}^k [z_i, z_i + d_0]\right)) = 0) \\ \le & \mathbb{P}(\xi^{(n)}([z'_j, z'_j + d_0]) = 0) \prod_{i=1}^k \mathbb{P}(\xi^{(n)}([z_i, z_i + d_0]) = 0) \le p_{n,k}^{k+1}, \end{aligned}$$

where

$$p_{n,k} := \sup_{z\in I} \mathbb{P}(\xi^{(n)}([z, z + d_0]) = 0) = \sup_{z\in I} \mathbb{P}(\xi^{(n)}([z, z + G_n(-C_0)/S(I)]) = 0).$$

By (38) and (51), there exists a constant $N_{6,2} > N_{6,1}$ such that

$$\begin{aligned} n(2\ln n)^{\frac{1}{2}} \sup_{z\in I} \Big(& \mathbb{P}(\xi^{(n)}([z, z + G_n(-C_0)/S(I)]) = 0) \\ & - D_n(\sqrt{4 - z^2}/S(I) \cdot G_n(-C_0)/2) \Big) < 1, \end{aligned}$$

and

$$n(2\ln n)^{-\frac{1}{2}} D_n(G_n(-C_0)/2) < e^{c_0 + C_0 + 1}.$$

Then we further have

$$\begin{aligned} p_{n,k} &= \sup_{z\in I} \mathbb{P}(\xi^{(n)}([z, z + G_n(-C_0)/S(I)]) = 0) \\ &\le \sup_{z\in I} \Big(\mathbb{P}(\xi^{(n)}([z, z + G_n(-C_0)/S(I)]) = 0) - D_n(\sqrt{4 - z^2}/S(I) \cdot G_n(-C_0)/2) \Big) \\ &+ \sup_{z\in I} D_n(\sqrt{4 - z^2}/S(I) \cdot G_n(-C_0)/2) \\ &\le n^{-1}(2\ln n)^{-\frac{1}{2}} + D_n(G_n(-C_0)/2) \le n^{-1}(2\ln n)^{\frac{1}{2}} + n^{-1}(2\ln n)^{\frac{1}{2}} e^{c_0 + C_0 + 1}, \end{aligned}$$

where we used the fact that $D_n(\alpha)$ is decreasing. Thus, for every $0 \le j \le k$, we have,

$$\begin{aligned}
&\limsup_{n\to+\infty}(2n)^k(2\ln n)^{-\frac{k}{2}}\mathbb{P}(\xi^{(n)}(I_{n,k}) = \xi^{(n)}(J_{n,k,j}) = 0)\\
\le&\limsup_{n\to+\infty}(2n)^k(2\ln n)^{-\frac{k}{2}}p_{n,k}^{k+1}\\
\le&\limsup_{n\to+\infty}(2n)^k(2\ln n)^{-\frac{k}{2}}\left(n^{-1}(2\ln n)^{\frac{1}{2}} + n^{-1}(2\ln n)^{\frac{1}{2}}e^{c_0+C_0+1}\right)^{k+1}\\
\le&\limsup_{n\to+\infty}2^k n^{-1}(2\ln n)^{\frac{1}{2}}\left(1+e^{c_0+C_0+1}\right)^{k+1} = 0,
\end{aligned}$$

which completes the claim.

Now using (38)(84)(85) and Lemma 14, we have

$$\begin{aligned}
&\liminf_{n\to+\infty}(\ln n)^{-k}\phi_{k,n}(y_1,\cdots,y_k)\\
\ge&\liminf_{n\to+\infty}(2n)^k(2\ln n)^{-\frac{k}{2}}\mathbb{P}(\xi^{(n)}(I_{n,k}) = 0)\\
=&\liminf_{n\to+\infty}(2n)^k(2\ln n)^{-\frac{k}{2}}\mathbb{P}(\xi^{(n)}(\cup_{j=1}^k[y_j, y_j+a_j]) = 0)\\
\ge&\liminf_{n\to+\infty}(2n)^k(2\ln n)^{-\frac{k}{2}}(1-(\ln n)^{-1})\prod_{j=1}^k D_n\left(a_j\sqrt{4-y_j^2}/2\right)\\
=&\liminf_{n\to+\infty}(2n)^k(2\ln n)^{-\frac{k}{2}}\prod_{j=1}^k D_n((1+|u_j|/\ln n)G_n(x_j)/2))\\
=&2^k\prod_{j=1}^k\left(\lim_{n\to+\infty}n(2\ln n)^{-\frac{1}{2}}D_n((1+|u_j|/\ln n)G_n(x_j)/2)\right)\\
=&2^k\prod_{j=1}^k\left(e^{c_0-x_j-2|u_j|}\right) = 2^k e^{\sum_{j=1}^k(c_0-x_j-2|u_j|)}.
\end{aligned}$$

Now Lemma 13 follows from the definition of y_j in (83). ☐

References

[1] G. W. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices.* Cambridge Studies in Advanced Mathematics 118. Cambridge Univ. Press, Cambridge. (2010)

[2] G. Ben Arous, P. Bourgade, Extreme gaps between eigenvalues of random matrices. Ann. Prob. 41, 2648-2681 (2013).

[3] P. Deift, A. Its, I. Krasovsky, X. Zhou, The Widom-Dyson constant for the gap probability in random matrix theory. J. Comput. Appl. Math. 202 26-47 (2007).

[4] P. Deift and D. Gioev. *Random matrix theory: invariant ensembles and universality*, volume 18 of Courant Lecture Notes in Mathematics. Courant Institute of Mathematical Sciences, New York, 2009.

[5] R. Feng and D. Wei, Small gaps of circular β-ensemble, Annals of Probability, 2021, 49(2): 997-1032.

[6] R. Feng, G. Tian and D. Wei, Small gaps of GOE, GAFA. 29 (2019), no. 6, 1794-1827.

[7] I. Gohberg, S. Goldberg, N. Krupnik, Traces and Determinants of Linear Operators. Operator Theory: Advances and Applications 116. Birkhäuser, Basel. (2000).

[8] B. Landon, P. Lopatto, J. Marcinek, Comparison theorem for some extremal eigenvalue statistics. Annals of Probability, 48(6): 2894-2919.

[9] A. Soshnikov, Statistics of extreme spacing in determinantal random point processes. Mosc. Math. J. 5 705-719, 744.
[10] G. Szegö, *Orthogonal Polynomials*, 4th ed. American Mathematical Society Colloquium Publications XXIII. Amer. Math. Soc., Providence, RI. (1975).
[11] J. Vinson, Closest spacing of eigenvalues. Ph.D. thesis, Princeton University, 2001.

Max Planck Institute for Mathematics, Bonn, Germany, 53111.
E-mail address: `feng@mpim-bonn.mpg.de`

School of Mathematical Sciences, Peking University, Beijing, China, 100871.
E-mail address: `jnwdyi@pku.edu.cn`