On Cramér-von Mises statistic for the spectral distribution of random matrices

Zhigang Bao∗
HKUST
maizbao@ust.hk

Yukun He†
University of Zürich
yukun.he@math.uzh.ch

Let $F_N$ and $F$ be the empirical and limiting spectral distributions of an $N \times N$ Wigner matrix. The Cramér-von Mises (CvM) statistic is a classical goodness-of-fit statistic that characterizes the distance between $F_N$ and $F$ in $\ell^2$-norm. In this paper, we consider a mesoscopic approximation of the CvM statistic for Wigner matrices, and derive its limiting distribution. In the appendix, we also give the limiting distribution of the CvM statistic (without approximation) for the toy model CUE.

1. Introduction and main result

1.1. Background. Let $H = (H_{ij})_{N,N}$ be a Wigner matrix, i.e., $H_{ij} = H_{ji}^*$ and $H_{ij}$’s are independent (up to symmetry). Further, we assume

(1) $\mathbb{E}H_{ij} = 0$ for all $i, j$.
(2) $\mathbb{E}|H_{ij}|^2 = \frac{1}{4N}$ for $i \neq j$, and $\mathbb{E}|H_{ii}|^2 = \frac{\sigma^2}{4N}$ for all $i$.
(3) $\mathbb{E}|H_{ij}|^k = O(N^{-k/2})$ for all $i, j$ and all fixed $k$.

We distinguish the real symmetric case ($\beta = 1$), where $H_{ij} \in \mathbb{R}$ for all $i, j$; the complex Hermitian case ($\beta = 2$), where $\mathbb{E}H_{ij}^2 = 0$ for $i \neq j$. We further assume the fourth moments of $H_{ij}$ are homogeneous, i.e. $\mathbb{E}|H_{ij}|^4 = m_4/N^2$ for all $i \neq j$, and denote $c_4 = m_4 - (4 - \beta)/16$.

Let $
abla_1 \geq \nabla_2 \geq \ldots \geq \nabla_N$
be the ordered eigenvalues of $H$. Denote the empirical spectral distribution of $H$ by

$$F_N(x) := \frac{1}{N} \sum_{i=1}^{N} 1(\lambda_i \leq x).$$

Arguably, the most fundamental result in Random matrix Theory (RMT) is Wigner’s semicircle law [35]. It states that almost surely $F_N(x)$ converges weakly to the semicircle law $F(x)$ with the density function given by

$$\rho_{sc}(x) = \frac{2}{\pi} \sqrt{1 - x^2}_+. $$

Based on this fundamental weak convergence result, some natural questions can be asked further. For instance, one can ask how to characterize the distance between $F_N$ and $F$. In addition, if a specific distance is chosen, we can take one more step to ask: What is the limiting behaviour of this distance between $F_N$ and $F$? The first question has a rich answer, considering that there are many widely used statistical distances for distributions in the literature. However, it could be very challenging to answer the second question for some specific distances, especially when one aims at some fine result like the limiting distributions of these distances between $F_N$ and $F$.

In applied probability and statistics theory, two widely used distances (or statistics) between the empirical distribution and the limiting one, are the Cramér-von Mises (CvM)

∗Supported by Hong Kong RGC grant ECS 26301517, GRF 16300618, and GRF 16301519.
†Supported by NCCR Swissmap, SNF grant No. 20020_1726.
statistic and the Kolmogorov-Smirnov (KS) statistic, which are $\ell^2$ and $\ell^\infty$ distances, respectively. More specifically, the CvM statistic for the spectral distribution of $H$ is defined as

$$ A_N := \int (F_N(t) - F(t))^2 dF(t), \quad (1.1) $$

and the well-known KS statistic is defined by

$$ K_N := \sup_t |F_N(t) - F(t)|. \quad (1.2) $$

We remark here that most of the literature on these two statistics are about the empirical distribution of the i.i.d. samples of a random variable. Here we consider the same distances for the empirical distribution of the highly correlated eigenvalues of random matrices. In the statistics literature, both $A_N$ and $K_N$ belong to the category of the goodness-of-fit statistics, which can be used to test how well the empirical distribution fits a given distribution. Specifically, both two statistics are fundamental for the following hypothesis testing problem: test if or not a random variable $X$ follows a given distribution $(F)$, based on the empirical distribution $(F_N)$ of its i.i.d. samples, see [2] for instance. Here we use $F_N$ and $F$ to denote the empirical and limiting distribution for the i.i.d. samples as well, with certain abuse of notations. For the i.i.d. setting, identifying the limiting distributions of $A_N$ and $K_N$ and their variants had been the primary task for this topic, due to their importance for the hypothesis testing problems. Nowadays, in the classical i.i.d. setting, under the null hypothesis, it is well-known that the limiting distributions of both $A_N$ and $K_N$ follow from the classical Donsker’s theorem. Specifically, since $\sqrt{N}(F_N(t) - F(t))$ converges to the Brownian bridge, $A_N$ and $K_N$ converge to the corresponding functionals of Brownian bridge, after appropriate normalization. In the classical i.i.d. case, the weight $dF(t)$ in (1.1) is chosen to make the statistic nonparametric. That means, after a simple change of variable in (1.1), no matter the distribution of the i.i.d. sample, it is always the same as the case of i.i.d. uniformly distributed samples. But in the literature, there are indeed many other choices of the weights. We refer to the reference [2, 33]. In general, for the i.i.d. case, the distribution of the CvM type statistics with general weight is given by an infinite sum of independent weighted $\chi^2$ random variables. The characteristic function of this distribution can be written as a Fredholm determinant, see [2, 33] for instance.

To the best of our knowledge, in the context of RMT, the limiting distribution of $A_N$ and $K_N$ for the empirical spectral distribution have never been obtained. Nevertheless, many recent results in RMT are related to this topic in one way or another. For Wigner matrices, as a consequence of the rigidity of the eigenvalues, a large deviation result for $K_N$ has been established in [14]. It states that $K_N$ is bounded by $(\log N)^{O(\log \log N)}/N$ with high probability. We also refer to [19, 12, 13, 1] for some related developments. More precise bound of $K_N$ is available for unitarily invariant ensembles. For instance, in the recent work [7], it is proved that the bound $\log N/N$ holds for $K_N$ with high probability, for a class of unitarily invariant ensembles. However, to identify the limiting distribution of $K_N$ precisely is still far beyond the current studies. It is worth mentioning that $K_N$ can also be written as

$$ K_N = \max_t \left| \frac{1}{N\pi} \Im \log \det \left( H - (t + i0^+)I_N \right) - F(t) \right|, $$

with the convention $\Im \log(x + i\omega) = \frac{\omega}{2}$. In contrast to the maximum of the random field for the imaginary part of log characteristic polynomial stated above, more research has been devoted to understanding the maximum of the random field for the real part of log characteristic polynomial of random matrices [3, 6, 16, 31, 15, 18, 27, 8, 25]. But none of
these works are on generally distributed Wigner matrices, and also no rigorous result on the limiting distribution (fluctuation) is obtained, although conjectures have been made in [15, 16, 18] for CUE and GUE. The related study for $\mathcal{A}_N$ is even less. In [32] (see (1.1) therein), Rains discussed a quadratic goodness-of-fit statistic for CUE, which is constructed in the same spirit as $\mathcal{A}_N$. But only the expectation of Rains’ statistic was derived in [32]. In the Appendix, we will show that CUE is actually a toy model for the CvM statistic. We can derive the limiting distribution of a CvM statistic explicitly in this case.

From the application perspective, unlike the i.i.d. case where the distribution of a random variable is the central object for the hypothesis testing, we can use the goodness-of-fit statistics of the spectral distribution to test the goodness-of-fit for general features of the matrix ensembles rather than the limiting spectral distribution itself. For instance, for the sample covariance matrix, one often uses spectral statistics to test the structure of the population covariance matrix; for a random graph, one can use spectral statistics of the adjacency matrix to test the graph parameters. Although the limiting distribution of the CvM statistic for Wigner type or covariance type matrices is not available so far, it has already been used in some statistics works, such as [30, 34], where the applications are based on numerical simulation of the CvM statistic.

For Wigner matrices, it is known that $F_N(t)$ behaves asymptotically as a log-correlated Gaussian field in the finite-dimensional sense, on both macroscopic and mesoscopic scale, see [23, 4, 28, 20, 29, 5] for instance. However, unlike the i.i.d. case, this log-correlated Gaussian field asymptotic for random matrix is not precise enough to tell the fluctuation of $A_N$ and $K_N$. It can be viewed as a common difficulty for $A_N$ and $K_N$, whose distributions both rely on a rather delicate understanding of the limiting behaviour of the field $F_N(t)$, $t \in \mathbb{R}$. On the other hand, an elementary calculation leads to the alternative representation

$$A_N = \frac{1}{N} \sum_{i=1}^{N} \left( F(\lambda_i) - \frac{N - i + 1}{2N} \right)^2 + \frac{1}{12N^2}, \quad (1.3)$$

which together with the spectral rigidity in [14] allows us to write

$$A_N = \frac{1}{N} \sum_{i=1}^{N} \rho_{sc}(\gamma_i)^2(\lambda_i - \gamma_i)^2 + \frac{1}{12N^2} + O(N^{-3+\epsilon}) \quad (1.4)$$

with high probability, where $\rho_{sc}$ represents the density function of $F$ and $\gamma_i$’s represent the quantiles of $F$, i.e., $F(\gamma_i) = (N - i + 1)/N$. Hence, one can also regard $A_N$ as certain quadratic eigenvalue statistic.

In this paper, we take a first step to understand the limiting distribution of the CvM type statistics for Wigner matrices. Instead of $A_N$, we turn to study a more accessible mesoscopic approximation of $A_N$, see Definition 1.1. Our aim is to derive the limiting distribution for this approximated $A_N$. Numerical study strongly suggests that the fluctuation of this approximation shall be the same as the fluctuation of the original $A_N$. The mesoscopic approximation can be regarded as a regularization of $F_N(t)$ on mesoscopic scale, which resembles the classical idea in RMT of using the imaginary part of Stieltjes transform on mesoscopic scale to regularize the eigenvalue density. Similar ideas of regularizing the entire field $F_N(t)$ can also be found in [26, 17], for instance.

To study the mesoscopic approximation of $A_N$, we first use the Fourier-Chebyshev expansion to factorize it into linear statistics of Chebychev polynomials. The main step of the proof lies at analyzing the covariance of squares of linear statistics of Chebychev polynomials with $N$-dependent degrees (see Proposition 3.1 below). To this end, we explicitly
compute the four-point correlation functions for mesoscopic linear statistics of Green functions, by the cumulant expansion formula. We present a simple argument that analyzes the four-point correlation functions both in the bulk and near the edge of the spectrum; the result is precise in the sense that it reveals the cancellations coming from the orthogonality of Chebyshev polynomials, for all degrees $k \ll N^{1/3}$. We remark here that the fact that the Chebyshev polynomials form an orthogonal basis for the covariance structure of the linear eigenvalue statistics for random Hermitian matrix was first noticed in [23], for general $\beta$ ensembles. It was recently rigorously shown in [17] that the field of log characteristic polynomial of GUE converges to log correlated Gaussian fields, on macroscopic and mesoscopic scales. Especially, on macroscopic scale, the limiting log correlated Gaussian field is given by a random Fourier-Chebyshev series, and the convergence is understood in the sense of distributions in a suitable Sobolev space. Fourier-Chebyshev expansion of the test function plays a significant role in [17] as well.

1.2. Mesoscopic approximation of CvM statistics and main result. In this paper, we will turn to study a mesoscopic approximation of the CvM statistic for Wigner matrix. The statistic is constructed via a Poisson convolution. Specifically, we set the Poisson type kernel on $[-1, 1]$

$$P^\pm_\omega(x, y) = \frac{1 - r^2_\omega}{1 - 2r_\omega(xy \pm \sqrt{(1 - x^2)(1 - y^2)}) + r^2_\omega}, \quad r_\omega = 1 - \omega.$$ 

Then we define the transformation for any function $f \in L^1((-1, 1))$,

$$f_\omega(x) = \int_{-1}^1 (P^+_\omega(x, y) + P^-_\omega(x, y)) \frac{f(y)}{\sqrt{1 - y^2}} dy. \quad (1.5)$$

Observe that $f_\omega(\cos \theta)$ is the Poisson transform of $f(\cos \theta)$ on $(-\pi, \pi)$. Hence, $f_\omega(x)$ is an approximation of $f(x)$ on scale $\omega$. Especially, for any $x \in (-1, 1)$ which is a continuity point of $f$, we have $f_\omega(x) \to f(x)$ if $\omega \downarrow 0$. In addition, observe that the domain of the integral in the definition (1.1) is $t \in (-1, 1)$. For any $t \in (-1, 1)$. We can also write

$$F_N(t) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(\bar{x}_i \leq t),$$

where $\bar{x} = -1 \vee x \wedge 1$ for any $x \in \mathbb{R}$. In order to raise our mesoscopic approximation of CvM statistic, we approximate the indicator function $\mathbf{1}(\cdot \leq t)$ by smooth approximation in the sense of (1.5). More specifically, we denote by

$$\chi^\omega_\omega(x) = \int_{-1}^t (P^+_\omega(x, y) + P^-_\omega(x, y)) \frac{1}{\sqrt{1 - y^2}} dy, \quad x \in (-1, 1), \quad t \in (-1, 1). \quad (1.6)$$

Then we define the approximation of $F_N(t)$ and $F(t)$ on scale $\omega$ as the following

$$F_{N, \omega}(t) := \frac{1}{N} \sum_{i=1}^N \chi^\omega_i(\bar{\lambda}_i), \quad F_\omega(t) := \int_{-1}^1 \chi^\omega(x) dF(x). \quad (1.7)$$

With these approximations, we are ready to define our mesoscopic approximation of $A_N$.

**Definition 1.1** (MCvM statistic). We call the statistic

$$A_{N, \omega} := \int (F_{N, \omega}(t) - F_\omega(t))^2 dF(t)$$

the mesoscopic approximation of the CvM statistic on scale $\omega$, which will be abbreviated as MCvM statistic (on scale $\omega$).
In this paper, we will investigate the limiting distribution of $A_{N,\omega}$ when $\omega = N^{-\alpha}$ with some constant $\alpha > 0$. Let $(Z_i)$ be a sequence of independent Standard Gaussian random variables. Our main theorem is as following.

**Theorem 1.2.** Let $H$ be a Wigner matrix, and $A_{N,\omega}$ be as in Definition 1.1. Let $\omega = N^{-\alpha}$. For any fixed $\alpha \in (0, \frac{1}{2})$, we have

$$N^2 A_{N,\omega} - \frac{\alpha \log N}{\beta \pi^2} - b_\beta \xrightarrow{d} \frac{1}{\beta \pi^2} \sum_{k=1}^{\infty} \left( \frac{1}{k} (Z_k^2 - 1) - \frac{1}{\sqrt{k(k+2)}} Z_k Z_{k+2} \right)$$

$$+ \frac{2 - \beta}{\sqrt{\beta \pi^2}} \sum_{k=1}^{\infty} \left( \frac{k + 2}{4k^{3/2}(k+1)} Z_{2k} - \frac{1}{4k^{3/2}(k+1)} Z_{2k+2} \right) + a_\beta,$$  (1.8)

where

$$a_\beta := \frac{1}{8\pi^2}((4\sigma^2 - 16c_4 - 5)(\beta^{-1} + 8c_4)^{1/2} - (7 - 4\beta)\beta^{-1/2})Z_2 + \frac{2\sqrt{2}}{\sqrt{\beta \pi^2}} (c_4 - \frac{\sigma^2 + \beta - 3}{32})Z_4$$

$$- \frac{2}{\sqrt{3\beta \pi^2}} c_4 Z_6 + \frac{1}{\pi} \left( \frac{\sigma^2 + \beta - 3}{4} - \frac{1}{2\beta} \right) (Z_2^2 - 1) - \frac{1}{2\sqrt{3\pi^2}} ((2(\beta(\sigma^2 + \beta - 3) + 4)^{1/2} - 2) Z_3 Z_3$$

$$- \frac{4}{\pi^2} c_4 (Z_2^2 - 1) - \frac{1}{2\sqrt{2\beta \pi^2}} ((1 + 8c_4 \beta)^{1/2} - 1) Z_4 Z_4,$$

and

$$b_\beta := -\frac{\log 2 - 1/2}{\beta \pi^2} + (2 - \beta) \left( \frac{1}{48} - \frac{1}{8\pi^2} \right) + \frac{1}{16\pi^2} (\sigma^2 + \beta - 3)(2\sigma^2 + \beta - 3) + \frac{19 - 2\beta - 3\sigma^2}{3\pi^2} c_4 + \frac{8}{\pi^2} c_4^2.$$

**Remark 1.3.** Note that in the Gaussian case where $\sigma^2 = 3 - \beta$ and $c_4 = 0$, Theorem 1.2 simplifies to

$$N^2 A_{N,\omega} - \frac{\alpha \log N}{\beta \pi^2} + \frac{\log 2 - 1/2}{\beta \pi^2} - (2 - \beta) \left( \frac{1}{48} - \frac{1}{8\pi^2} \right)$$

$$\xrightarrow{d} \frac{1}{\beta \pi^2} \sum_{k=1}^{\infty} \left( \frac{1}{k} (Z_k^2 - 1) - \frac{1}{\sqrt{k(k+2)}} Z_k Z_{k+2} \right)$$

$$+ \frac{2 - \beta}{\sqrt{\beta \pi^2}} \sum_{k=1}^{\infty} \left( \frac{k + 2}{4k^{3/2}(k+1)} Z_{2k} - \frac{1}{4k^{3/2}(k+1)} Z_{2k+2} \right) - \frac{1}{2\beta \pi^2} (Z_2^2 - 1).$$  (1.9)

**Remark 1.4.** Although $\frac{\alpha \log N}{\beta \pi^2} + b_\beta$ shall be far away from the expectation of the original CvM statistic $A_N$, we believe that the fluctuation in the RHS shall be the same as the fluctuation of $A_N$. We refer to Figure 1 for a simulation result in the case of GUE. We notice that the centered histogram of the original $A_N$ (red curve) matches perfectly the plot of the density of the random variables on the RHS of (1.8) in case of GUE (blue curve).

**Remark 1.5.** We also remark here that our proof of the main theorem will rely on the local semicircle law on scale $\omega$. We believe that a finer argument shall allow one to push the scale $\omega$ to $\frac{1}{\sqrt{N}}$, i.e., $\alpha = 1$, which is the optimal scale of local law. However, within the framework of the current proof strategy, getting the limiting distribution of the original $A_N$ requires one to go even below the scale $\frac{1}{\sqrt{N}}$. For the toy model CUE, one can get the limiting distribution of the CvM statistic explicitly. We state the details in Appendix. The argument for CUE also inspires the definition of the MCvM for Wigner matrix.
Figure 1. Blue curve represents the limiting distribution in the RHS of (1.9) with \( \beta = 2 \). For the simulation purpose, we truncate the series of (1.9) at \( k = 300 \). Green curve is a smooth approximation of the histogram of the original CvM statistic \( A_N \) for GUE with dimension \( N = 300 \). Both curves are plotted based on 6000 repetitions of simulation study. The red curve is a shift of the Green one to mean 0. We notice that the blue and red curves match perfectly.

1.3. Organization. The paper is organized as follows. In Section 2, we state some preliminaries. Section 3 will be devoted to the proof of the main result, Theorem 1.2, based on Propositions 3.1 and 3.2, whose proofs will be stated in Section 4. In Section 5, we prove our main technical result, Proposition 4.2, which is used in the proofs in Section 4. In Section 6 we provide some further discussion on the CvM statistics for generally distributed random matrices and in Appendix we derive the limiting distribution of a CvM type statistic for the toy model CUE.

2. Preliminaries

Throughout the paper, for an \( N \times N \) matrix \( A \), we write \( \mathbb{A} := \frac{1}{N} \text{Tr} A \), and we abbreviate \( A_{ij}^m := (A_{ij})^m \). For \( u, v \in \mathbb{C}^N \), we abbreviate

\[
A_{uv} := \langle uA, v \rangle, \quad A_{iu} := \langle uA, e_i \rangle \quad \text{and} \quad A_{i}u := \langle e_iA, u \rangle,
\]

where \( e_i \) is the standard \( i \)-th basis vector of \( \mathbb{R}^N \). We denote \( \langle X \rangle := X - EX \) for any random variable \( X \) with finite expectation.

2.1. Green function and the local semicircle law. For \( z \in \mathbb{C} \setminus \mathbb{R} \), we denote the Green function of \( H \) and the Stieltjes transform of its empirical eigenvalue distribution by

\[
G(z) := (H - z)^{-1}, \quad G(z) = \frac{1}{N} \text{Tr} G(z) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i - z}.
\]

Correspondingly, we denote the Stieltjes transform of the semicircle law by

\[
m(z) = \int \frac{1}{x - z} \rho_{sc}(x) dx = 2(-z + \sqrt{z^2 - 1}). \tag{2.1}
\]
For $n \in \mathbb{N}_+$, we use the shorthand notation $[1, n] := \{1, 2, \ldots, n\}$. We also adopt the notion of stochastic domination introduced in [11]. It provides a convenient way of making precise statements of the form “$X^{(N)}$ is bounded by $Y^{(N)}$ up to small powers of $N$ with high probability”.

**Definition 2.1 (Stochastic domination).** Let 

$$X = (X^{(N)}(u) : n \in \mathbb{N}, u \in U(n)), \quad Y = (Y^{(N)}(u) : n \in \mathbb{N}, u \in U(n)),$$

be two families of random variables, where $Y$ are nonnegative, and $U^{(N)}$ is a possibly $n$-dependent parameter set. We say that $X$ is stochastically dominated by $Y$, uniformly in $u$, if for all small $\varepsilon$ and large $D$, we have

$$\sup_{u \in U^{(N)}} \mathbb{P}\left(\left|X^{(N)}(u)\right| > N^\varepsilon Y^{(N)}(u)\right) \leq N^{-D},$$

for large enough $N \geq N_0(\varepsilon, D)$. In addition, we use the notation $X = O(Y)$ if $X$ is stochastically dominated by $Y$, uniformly in $u$. Throughout this paper, the stochastic domination will always be uniform in all parameters (mostly are matrix indices and the spectral parameter $z$) that are not explicitly fixed.

Fix $\tau > 0$, let us define the spectral domains

$$\mathcal{S} := \{E + i\eta : |E| \leq 10, 0 < \eta \leq 10\}$$

and

$$\mathcal{S}^\circ := \mathcal{S}^\circ(\tau) := \{E + i\eta \in \mathcal{S} : |E| \geq 1 + N^{-2/3+\tau}\}.$$

We also define the distance to spectral edge by

$$\kappa \equiv \kappa_E := |E^2 - 1|.$$

We have the following isotropic local semicircle law for Wigner matrices from [24].

**Theorem 2.2 (Local semicircle law).** Let $u, v \in \mathbb{C}^N$ be deterministic with $\|u\|_2 = \|v\|_2 = 1$. For Green function, we have the following estimates

$$|G_{uv}(z) - m(z)\langle u, v \rangle| \lesssim \sqrt{\frac{\text{Im} m(z)}{N\eta}} + \frac{1}{N\eta} \quad \text{and} \quad |G(z) - m(z)| \lesssim \frac{1}{N\eta} \quad (2.2)$$

uniformly for $z = E + i\eta \in \mathcal{S}$. Moreover, outside the bulk of the spectral, we have the stronger estimates

$$|G_{uv}(z) - m(z)\langle u, v \rangle| \lesssim \frac{1}{\sqrt{N(\kappa + \eta)^{1/4}}} \quad \text{and} \quad |G(z) - m(z)| \lesssim \frac{1}{N(\eta + \kappa)} \quad (2.3)$$

uniformly for $z = E + i\eta \in \mathcal{S}^\circ$.

One can easily deduce from the above theorem that

$$|\lambda_i - \gamma_i| \lesssim N^{-\frac{2}{3}} (i \wedge (N - i + 1))^{\frac{1}{3}} \quad (2.4)$$

uniformly for all $i = 1, \ldots, N$, and

$$\left|\{i : \lambda_i \in \mathcal{I}\} - N \int_{\mathcal{I}} \rho_{sc}(x)dx\right| \lesssim 1 \quad (2.5)$$

uniformly for any interval $\mathcal{I} \subset \mathbb{R}$.

If $h$ is a real-valued random variable with finite moments of all order, we denote by $C_n(h)$ the $n$th cumulant of $h$, i.e.

$$C_n(h) := \langle (-i)^n \cdot (\partial_x^n \log \mathbb{E}e^{i\lambda x})\rangle_{\lambda = 0}.$$

We state the cumulant expansion formula, whose proof is given in e.g. [22, Appendix A].
Lemma 2.3 (Cumulant expansion). Let $f : \mathbb{R} \to \mathbb{C}$ be a smooth function, and denote by $f^{(n)}$ its $n$th derivative. Then, for every fixed $\ell \in \mathbb{N}$, we have

$$\mathbb{E}[h \cdot f(h)] = \sum_{n=0}^{\ell} \frac{1}{n!} C_{n+1}(h) \mathbb{E}[f^{(n)}(h)] + R_{\ell+1},$$

(2.6)

assuming that all expectations in (2.6) exist, where $R_{\ell+1}$ is a remainder term (depending on $f$ and $h$), such that for any $t > 0$,

$$R_{\ell+1} = O(1) \cdot \left( \mathbb{E} \sup_{|x| \leq |h|} |f^{(\ell+1)}(x)|^2 \cdot \mathbb{E} |h|^{2\ell+4} \cdot \mathbb{E} \sup_{|x| \leq t} |f^{(\ell+1)}(x)| \right)^{1/2} + O(1) \cdot \mathbb{E} |h|^{\ell+2} \cdot \mathbb{E} \sup_{|x| \leq t} |f^{(\ell+1)}(x)|.$$  

(2.7)

The following result gives bounds on the cumulants of the entries of $H$, whose proof follows by the homogeneity of the cumulants.

Lemma 2.4. For every $n \in \mathbb{N}$ we have

$$C_n(H_{ij}) = O_n(N^{-n/2})$$

uniformly for all $i, j$.

We conclude with the magical Ward identity.

Lemma 2.5 (Ward identity). We have

$$\sum_j |G_{ij}(z)|^2 = \frac{\text{Im} G_{ii}(z)}{\eta}$$

for all $z = E + i \eta$ with $\eta \neq 0$.

2.2. Chebyshev’s Polynomial. In the sequel, we will consider the Fourier-Chebyshev expansion of a function $f : [-1, 1] \to \mathbb{R}$ which admits

$$f(x) = \frac{1}{2} c_0 + \sum_{k=1}^{\infty} c_k T_k(x) =: \sum_k' c_k T_k(x), \quad a.e.$$  

(2.8)

where we used the notation $\sum_k'$ to denote the sum from $k = 0$ to $\infty$ with the first summand ($k = 0$) halved and $T_k$’s are the Chebyshev polynomials of the first kind, i.e.,

$$T_k(\cos \theta) = \cos(k \theta).$$

In the whole $\mathbb{C}$, one can also write the polynomials as

$$T_k(z) = \frac{1}{2} \left( (z - \sqrt{z^2 - 1})^k + (z + \sqrt{z^2 - 1})^k \right), \quad z \in \mathbb{C}$$  

(2.9)

where we make the convention that the square-root takes the branch-cut at the negative real axis. Note from (2.1) that

$$T_k(z) = \frac{1}{2} \left( \left( \frac{m(z)}{2} \right)^k + \left( \frac{m(z)}{2} \right)^{-k} \right).$$

(2.10)

Let us denote

$$t_k(A) := \frac{1}{N} \text{Tr} T_k(A) - \int T_k(x) \rho_{sc}(x) dx$$

for a matrix $A \in \mathbb{C}^{N \times N}$. We shall use the following results from [4].
Theorem 2.6 (Corollary 6.1 of [4]). For any fixed \( k \in \mathbb{N} \), the random vector \((N_{t_1}(H), \ldots, N_{t_k}(H))\) converges weakly to Gaussian vector \((g_1, \ldots, g_k)\) with independent components and the means and variances are given by

\[
\mathbb{E}g_k = \frac{2 - \beta}{4} (1 + (-1)^k) + \frac{1}{2} (\sigma^2 + \beta - 3) \delta_{k2} + 8c_4 \delta_{k4},
\]

\[
\text{Var}(g_k) = \frac{1}{4} ((3 - \beta)k + (\sigma^2 + \beta - 3) \delta_{k1} + 32c_4 \delta_{k2}),
\]

where the parameters \( \beta, \sigma^2 \) and \( c_4 \) are defined in Section 1.1.

Finally, we set \( \bar{H} = -1 \vee H \wedge 1 \), and we have the following comparison result.

Lemma 2.7. For \( k \ll N^{1/3} \), we have

\[
t_k(H) - t_k(\bar{H}) \prec k^2 N^{-5/3}.
\]

Proof. By definition in (2.10), we have

\[
N(t_k(H) - t_k(\bar{H})) = \sum_{i=1}^{N} (T_k(\lambda_i) - T_k(\bar{\lambda}_i)) = \sum_{i=1}^{N} (T_k(\lambda_i) - T_k(\bar{\lambda}_i)) \mathbf{1}(|\lambda_i| > 1)
\]

\[
\prec k^2 \max_i |\lambda_i - 1| \{i : |\lambda_i| > 1\} \prec k^2 N^{-\frac{2}{3}}.
\]

where in the last two steps we used (2.4) and its consequence (2.5).

3. Proof of Theorem 1.2

For the remaining of this paper we set

\[
\zeta := 1/3 - \alpha > 0.
\]

Let \( f \) be a function satisfying (2.8). Recall the definition of \( f_\omega \) in (1.5). We have

\[
f_\omega(\cos \theta) = \frac{1}{2\pi} \int_0^\pi \frac{1 - r_\omega^2}{1 - 2r_\omega \cos(\theta - t) + r_\omega^2} f(\cos t) dt
\]

\[
= \frac{1}{\pi} \sum_k c_k \int_0^\pi \left( 1 + \sum_{\ell=1}^{\infty} r_\omega^\ell (\cos(\ell(\theta - t)) + \cos(\ell(\theta + t))) \right) \cos(kt) dt
\]

\[
= \frac{1}{\pi} \sum_k c_k \int_0^\pi \left( 1 + 2 \sum_{\ell=1}^{\infty} r_\omega^\ell \cos(\ell \theta) \cos(\ell t) \right) \cos(kt) dt
\]

\[
= \frac{1}{2} c_0 + \sum_{k=1}^{\infty} c_k r_\omega^k \cos(k\theta).
\]

Hence, for any \( x \in [-1, 1] \), we have

\[
f_\omega(x) = \frac{1}{2} c_0 + \sum_{k=1}^{\infty} c_k r_\omega^k T_k(x).
\]

For any \( x \in \mathbb{R} \), we set

\[
\bar{x} := -1 \vee x \wedge 1
\]

Set

\[
f_t(x) := 1(x \leq t).
\]
Observe that \( f_t(\bar{x}) = f_t(x) \) for any \( t \in (-1, 1) \). Note that \( \chi^\omega_n(\bar{x}) = (f_t)_n(\bar{x}) \), see (1.6) for the definition of \( \chi^\omega_n(x) \). It is then easy to compute

\[
\chi^\omega_n(\bar{x}) = \frac{1}{2} d_0^\omega + \sum_{k=1}^{\infty} d_k^\omega T_k(\bar{x}),
\]

\[
d_k^\omega := \frac{2}{\pi} \int_{-1}^{1} T_k(s) \frac{1}{\sqrt{1 - s^2}} ds = -\frac{2}{\pi k} \sin(k \cos^{-1} t).
\]

(3.1)

According to the definitions in (3.1) and (1.7), we can write

\[
F_{N, \omega}(t) = \frac{1}{2} d_0^\omega (\bar{x}) + \frac{1}{N} \sum_{k=1}^{\infty} d_k^\omega r_k T_k(\bar{H}), \quad F_\omega(t) = \frac{1}{2} d_0^\omega (\bar{x}) + \sum_{k=1}^{\infty} d_k^\omega t_k(\bar{H}) \int T_k(x) \rho_{ae}(x) dx.
\]

Therefore, from Definition 1.1, we have

\[
\mathcal{A}_{N, \omega} = \int (F_{N, \omega}(t) - F_\omega(t))^2 dF(t) = \sum_{j,k=1}^{\infty} r_j^{+k} t_j(H) t_k(H) \int d_j^\omega d_k^\omega dF(t),
\]

where we used the notation in (2.10).

According to the definition of \( d_j^\omega \) in (3.1), we have

\[
\int d_j^\omega d_k^\omega dF(t) = \frac{8}{\pi^3 jk} \int_{-1}^{1} \sin(j \cos^{-1} t) \sin(k \cos^{-1} t) \sqrt{1 - t^2} dt
\]

\[
= \frac{8}{\pi^3 jk} \int_{0}^{\pi} \sin(j \theta) \sin(k \theta) \sin^2 \theta d\theta
\]

\[
= \frac{2}{\pi^2 jk} \left( j(j + 1) - \frac{1}{2} j(j + 2) - \frac{1}{2} j(j + k) \right).
\]

Therefore, we have

\[
\mathcal{A}_{N, \omega} = \frac{2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \omega^2 (t_k(H))^2 - \frac{2}{\pi^2} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{k(j + 2)} t_k(H) t_{j+k+2}(H) - \frac{1}{\pi^2} \omega^2 (t_1(H))^2.
\]

(3.2)

According to the definition in (2.10), we have \( |t_j(H)| \leq 2 \). Further, recall \( r_\omega = 1 - \omega \) and \( \omega = N^{-\alpha} \). We can trivially truncate the sum in (3.2) to

\[
\mathcal{A}_{N, \omega} = \frac{2}{\pi^2} \sum_{k=1}^{n_\omega} \frac{1}{k^2} \omega^2 (t_k(H))^2 - \frac{2}{\pi^2} \sum_{k=1}^{n_\omega} \frac{1}{k(k + 2)} t_k(H) t_{k+2}(H)
\]

\[
- \frac{1}{\pi^2} \omega^2 (t_1(H))^2 + o(N^{-K}),
\]

(3.3)

for any large constant \( K \) when \( N \) is sufficiently large. Here

\[
n_\omega := \lfloor \omega^{-1} (\log N)^2 \rfloor.
\]

Now for \( k \in \mathbb{N}_+ \) and \( \alpha > 0 \), we define

\[
\gamma_{a,k}^{(1)} := \{ E + i\eta \in \mathbb{C} : E \in [-1, 1], \eta = \pm ak^{-1} \sqrt{|E^2 - 1| + k^{-2}} \},
\]

and

\[
\gamma_{a,k}^{(2)} := \{ E + i\eta \in \mathbb{C} : |E| \geq 1, \text{dist}(\gamma_{a,k}, \{-1, 1\}) = ak^{-2} \}.
\]

We then consider the contour

\[
\gamma_{a,k} := \gamma_{a,k}^{(1)} \cup \gamma_{a,k}^{(2)}
\]

(3.4)

whose figure is sketched below.
By (2.4), with high probability, we can write
\[ t_k(H) = \frac{i}{2\pi} \oint_{\gamma_{a,k}} T_k(z)(m_N(z) - m_{sc}(z))dz =: \frac{i}{2\pi} \oint_{\gamma_{a,k}} T_k(z)m_N^2(z)dz. \quad (3.5) \]

It is elementary to check that
\[ \sup_{z \in \gamma_{a,k}} |T_k(z)| \leq C_a \]
for all \( k \leq n_\omega \). Furthermore, by Theorem 2.2 and (3.5) one easily deduces that
\[ t_k(H) \prec \frac{k}{N}. \]

Together with Lemma 2.7, (3.3), and \( \omega N^{1/3} = N^\delta \), we arrive at
\[ A_{N,\omega} = \frac{2}{\pi^2} \sum_{k=1}^{n_\omega} \frac{1}{k^2} \sqrt{\omega} (t_k(H))^2 - \frac{2}{\pi^2} \sum_{k=1}^{n_\omega} \sqrt{\omega} \frac{1}{k(k+2)} t_k(H)t_{k+2}(H) \]
\[ - \frac{1}{\pi^2} r_{a,k}^2 (t_1(H))^2 + O_{\prec}(N^{-2\delta}). \]

For the rest of this paper, we shall only work on the case when \( H \) is real and symmetric \((\beta = 1)\); in the complex Hermitian case \((\beta = 2)\), one only needs to apply the complex analogue of Lemma 2.3 (see e.g. [21, Lemma 7.1]) and the proof works in the same way.

Now Theorem 1.2 follows easily from the following results, whose proofs are given in Section 4 below.

**Proposition 3.1.** For any \( k, l \in \mathbb{N} \), we have
\[
\text{Cov}\left((t_k(H))^2, (t_l(H))^2\right) = 4\left(\frac{1}{4N}(1 + (-1)^k) + \frac{8c_4}{N}\delta_{kl} + \frac{\sigma^2 - 2}{2N}\delta_{kl}\right)^2 \left(\frac{k}{2N^2} + \frac{8}{N}\delta_{kl} + \frac{\sigma^2 - 2}{4N^2}\delta_{kl}\right)\]
\[ + 2\left(\frac{k}{2N^2} + \frac{8}{N}\delta_{kl} + \frac{\sigma^2 - 2}{4N^2}\delta_{kl}\right)^2 \delta_{kl} + O_{\prec}(klN^{-4-\delta}) \]
\[ \text{(3.9)} \]

and
\[
\text{Cov}\left(t_k(H)t_{k+2}(H), t_l(H)t_{l+2}(H)\right) = \left(\frac{1}{4N}(1 + (-1)^{k+2}) + \frac{8c_4}{N}\delta_{kl}\right)^2 \left(\frac{k}{2N^2} + \frac{8}{N}\delta_{kl} + \frac{\sigma^2 - 2}{4N^2}\delta_{kl}\right)\delta_{kl} \]
\[ + \left(\frac{1}{4N}(1 + (-1)^k) + \frac{8c_4}{N}\delta_{kl} + \frac{2(\sigma^2 - 2)}{N}\delta_{kl}\right)^2 \delta_{kl} \]
\[ + \frac{k + 2}{N^2} \left(\frac{k}{2N^2} + \frac{8}{N}\delta_{kl} + \frac{\sigma^2 - 2}{4N^2}\delta_{kl}\right)\delta_{kl} + O_{\prec}(klN^{-4-\delta}), \]
\[ \text{(3.10)} \]

as well as
\[
\text{Cov}\left(t_k(H)t_k(H), t_l(H)t_{l+2}(H)\right) \prec klN^{-4-\delta}. \]
\[ \text{(3.11)} \]
Furthermore, we also need the following proposition on the estimate of the expectations.

**Proposition 3.2.** For any $k \in [1, n_\omega]$, we have

\[
E(t_k(H))^2 = \frac{k}{2N^2} + \frac{8}{N^2} c_4 \delta_{k2} + \frac{\sigma^2 - 2}{4N^2} \delta_{k1} + \left( \frac{1}{4N} (1 + (-1)^k) + \frac{8c_4}{N} \delta_{k4} + \frac{\sigma^2 - 2}{2N} \delta_{k2} \right)^2 + O_\omega(kN^{-2-\zeta})
\]

and

\[
E t_k(H) t_{k+2}(H) = \left( \frac{1}{4N} (1 + (-1)^k) + \frac{8c_4}{N} \delta_{k2} \right) \times \left( \frac{1}{4N} (1 + (-1)^k) + \frac{8c_4}{N} \delta_{k2} \right) + O_\omega(kN^{-2-\zeta}).
\]

With the discussion above, we can now prove Theorem 1.2.

**Proof of Theorem 1.2.** Fix an even integer $M \geq 4$. Recall the notation $\langle X \rangle := X - EX$. We have

\[
N^2 A_{N, \omega} = N^2 \left( \frac{2}{\pi^2} \sum_{k=1}^{M} \frac{1}{k^2} \langle t_k(H)^2 \rangle - \frac{2}{\pi^2} \sum_{k=1}^{M} \frac{1}{k(k+2)} r_\omega^{2k+2} \langle t_k(H) t_{k+2}(H) \rangle - \frac{1}{\pi^2} \langle t_1(H)^2 \rangle \right)
\]

\[+ N^2 \left( \frac{2}{\pi^2} \sum_{k=M+1}^{\infty} \frac{1}{k^2} \omega^{2k} \langle t_k(H)^2 \rangle - \frac{2}{\pi^2} \sum_{k=M+1}^{\infty} \frac{1}{k(k+2)} \omega^{2k+2} \langle t_k(H) t_{k+2}(H) \rangle \right)
\]

\[+ \text{EN}^2 \left( \frac{2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \omega^{2k} \langle t_k(H)^2 \rangle - \frac{2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k(k+2)} \omega^{2k+2} \langle t_k(H) t_{k+2}(H) \rangle \right)
\]

\[+ O_\omega(N^{-4\zeta}) = :X^{(1, M)} + X^{(2, M)} + X^{(3)} + O_\omega(N^{-4\zeta}). \tag{3.12}
\]

By Proposition 3.2 we have

\[
X^{(3)} = \frac{1}{\pi^2} \sum_{k=1}^{n_\omega} \frac{1}{k^2} t_k^2 + \frac{4}{\pi^2} c_4 r_\omega^4 + \frac{\sigma^2 - 2}{2\pi^2} r_\omega^4 + \frac{\sigma^2}{4\pi^2} r_\omega^4 + \frac{1}{\pi^2} \sum_{n=1}^{[n_\omega/2]} \frac{n}{4n^2} r_n^{4n} + \frac{8}{\pi^2} c_4^2 r_\omega^8
\]

\[+ \frac{(\sigma^2 - 2)^2}{8\pi^2} r_\omega^4 + \frac{1}{\pi^2} c_4 r_\omega^8 + \frac{\sigma^2 - 2}{4\pi^2} r_\omega^4 - \frac{1}{2\pi^2} \sum_{n=1}^{[n_\omega/2]} \frac{1}{4n(n+1)} r_n^{4n+2} - \frac{1}{\pi^2} c_4 r_\omega^8 - \frac{1}{3\pi^2} c_4 r_\omega^{10}
\]

\[= - \frac{\sigma^2 - 2}{16\pi^2} r_\omega^4 - \frac{\sigma^2 - 2}{\pi^2} c_4 - \frac{1}{2\pi^2} r_\omega^2 + O_\omega(N^{-2-\zeta})
\]

\[= - \frac{\log(2\omega)}{\pi^2} + \frac{4}{\pi^2} c_4 + \frac{\sigma^2 - 2}{2\pi^2} + \frac{1}{4\pi^2} c_4^2 + \frac{\sigma^2 - 2}{4\pi^2} + \frac{\sigma^2 - 2}{8\pi^2} + \frac{1}{\pi^2} c_4 + \frac{\sigma^2 - 2}{4\pi^2} - \frac{1}{8\pi^2} - \frac{1}{\pi^2} c_4
\]

\[= - \frac{1}{3\pi^2} c_4 - \sigma^2 - 2 c_4 = \frac{1}{\pi^2} b_1 + O((\log N)^{-1}) \tag{3.13}
\]
Let \((Z_i)_{i \in \mathbb{N}}\) be independent standard Gaussian random variables, and by Theorem 2.6 we see that
\[
X^{(1,M)} \xrightarrow{d} \frac{2}{\pi^2} \sum_{k=1}^{M} \frac{1}{k^2} \left( \frac{k}{2} + \frac{\sigma^2 - 2}{4} \delta_{k1} + 8c_4 \delta_{k2} \right) (Z_k^2 - 1) \\
+ \frac{4}{\pi^2} \sum_{k=1}^{M} \frac{1}{k^2} \left( \frac{1}{4} (1 + (-1)^k) + \frac{1}{2} (\sigma^2 - 2) \delta_{k2} + 8c_4 \delta_{k4} \right) \left( \frac{k}{2} + \frac{\sigma^2 - 2}{4} \delta_{k1} + 8c_4 \delta_{k2} \right)^{1/2} Z_k \\
- \frac{2}{\pi^2} \sum_{k=1}^{M} \frac{1}{k(k+2)} \left( \frac{k}{2} + \frac{\sigma^2 - 2}{4} \delta_{k1} + 8c_4 \delta_{k2} \right)^{1/2} \left( \frac{k+2}{2} \right)^{1/2} Z_k Z_{k+2} \\
- \frac{2}{\pi^2} \sum_{k=1}^{M} \frac{1}{k(k+2)} \left( \frac{1}{4} (1 + (-1)^k) + \frac{\sigma^2 - 2}{4} \delta_{k2} + 8c_4 \delta_{k4} \right) \left( \frac{k+2}{2} \right)^{1/2} Z_k Z_{k+2} \\
- \frac{2}{\pi^2} \sum_{k=1}^{M} \frac{1}{k(k+2)} \left( \frac{1}{4} (1 + (-1)^k) + 8c_4 \delta_{k2} \right) \left( \frac{k}{2} + \frac{\sigma^2 - 2}{4} \delta_{k1} + 8c_4 \delta_{k2} \right)^{1/2} Z_k \xrightarrow{d} \frac{1}{\pi^2} \left( \frac{1}{2} + \frac{\sigma^2 - 2}{4} \right) (Z_1^2 - 1) 
\]
be independent standard Gaussian random variables, and by Theorem 1.2. Let us denote
\[
Y := \frac{1}{\pi^2} \sum_{k=1}^{\infty} \left( \frac{1}{k} (Z_k^2 - 1) - \frac{1}{\sqrt{k(k+2)}} Z_k Z_{k+2} \right) + \frac{1}{\pi^2} \sum_{n=1}^{\infty} \left( \frac{n + 2}{4\pi/\sqrt{n+1}} Z_{2n} - \frac{1}{4\pi/\sqrt{n+1}} Z_{2n+2} \right) 
\]
and it is easy to see that
\[
\mathbb{E} \left[ \left| Y - Y^{(M)} \right|^2 \right] = \text{Var}(Y - Y^{(M)}) \leq CM^{-1} 
\]
for some constant \(C > 0\) independent of \(N, M\). By Proposition 3.1 we have
\[
\mathbb{E} \left[ \left| X^{(2,M)} \right|^2 \right] = \text{Var}(X^{(2,M)}) \leq CM^{-1} .
\]
Thus for any fixed \(t > 0\),
\[
\left| \lim_{N \to \infty} \mathbb{E} \left( \exp(it(X^{(1,M)} + X^{(2,M)})) \right) - \mathbb{E} \left( \exp(it(Y + a_1)) \right) \right| \\
\leq \lim_{N \to \infty} \mathbb{E} \left( \exp(it(X^{(1,M)} + X^{(2,M)})) \right) - \mathbb{E} \left( \exp(it(X^{(1,M)})) \right) \\
+ \left| \mathbb{E} \left( \exp(it(Y^{(M)} + a_1)) \right) - \mathbb{E} \left( \exp(it(Y + a_1)) \right) \right| \\
\leq t \left\| X^{(2,M)} \right\|_1 + t \left\| Y - Y^{(M)} \right\|_1 \leq 2t \sqrt{\frac{C}{M}} .
\]
Since \(X^{(1,M)} + X^{(2,M)}\) is independent of \(M\), we have
\[
X^{(1,M)} + X^{(2,M)} \xrightarrow{d} Y + a_1
\]
as $N \to \infty$. We conclude the proof of Theorem 1.2 by combining (3.12) – (3.14).

\[ \square \]

4. PROOF OF PROPOSITIONS 3.1 AND 3.2

In this section, we prove (3.9); the other statements in Proposition 3.1 and 3.2 follow in a similar fashion. We shall rewrite all quantities in terms of the Green function, and then proceed the proof using some estimates of the 4-points and 2-points correlation functions of the Green functions.

Let $k,l \in [1,n_w]$, and W.L.O.G assume $k \geq l$. According to (3.5), we can write

\[
\text{Cov} \left( (t_k(H)t_k(H), t_l(H)t_l(H)) \right) = \frac{1}{16\pi} \oint_{\gamma_1,k} \oint_{\gamma_2,k} \oint_{\gamma_3,l} \oint_{\gamma_4,l} T_k(z_1)T_k(z_2)T_l(z_3)T_l(z_4) \\
\times \text{Cov} \left( m_N^\Delta(z_1)m_N^\Delta(z_2), m_N^\Delta(z_3)m_N^\Delta(z_4) \right) 4 \prod_{i=1}^4 dz_i + O_\prec(N^{-10}).
\]

(4.1)

Analogously, with high probability, we can write

\[
\mathbb{E} E_k(H)t_k(H) = -\frac{1}{4\pi^2} \oint_{\gamma_1,k} \oint_{\gamma_2,k} T_k(z_1)T_k(z_2)E(m_N^\Delta(z_1)m_N^\Delta(z_2))dz_1dz_2 + O_\prec(N^{-10}).
\]

In the following we shall abbreviate $\gamma_i := \gamma_{i,k}$ for $i = 1, 2$ and $\gamma_i := \gamma_{i,l}$ for $i = 3, 4$.

Let us define

\[
u_i = \sqrt{|E_i^2 - 1| + |\eta_i|}
\]

(4.2)

for $i = 1, \ldots, 4$. We set

\[
t_{i,j} := \frac{1}{u_i z_i - z_j} + \frac{1}{u_i |z_i - z_j|} + \frac{1}{u_j^2 |z_i - z_j|}
\]

for $i \neq j \in \{1, 2, 3, 4\}$. For the remaining part of the paper we denote our fundamental error by

\[
\mathcal{E} := \frac{kl(t_1,2t_{3,4} + t_1,3t_{2,4} + t_1,4t_{2,3})}{N^{4+\varsigma}}.
\]

(4.3)

The following is an elementary estimate.

**Lemma 4.1.** For $i \neq j \in \{1, \ldots, 4\}$, we have

\[
\oint_{\gamma_i} \oint_{\gamma_j} t_{i,j} dz_i dz_j < 1.
\]

(4.4)

For $z_i \in \gamma_i, z_j \in \gamma_j, i \neq j$, we define

\[
f(z_i, z_j) := \frac{(z_i - z_j)^2 - (\sqrt{z_i^2 - 1} - \sqrt{z_j^2 - 1})^2}{2N^2(z_i - z_j)^2 \sqrt{z_i^2 - 1} \sqrt{z_j^2 - 1}} - \frac{\sigma^2 - 2}{4N^2} m'(z_i)m'(z_j),
\]

and

\[
g(z_i) := -\frac{1}{4N \sqrt{z_i^2 - 1}} \left( m'(z_i) + 4c_4 m(z_i) \right)^4 + (\sigma^2 - 2)(m(z_i))^2).
\]

(4.5)

(4.6)

The following result is our main technical estimate, whose proof is postponed to Section 5.
Proposition 4.2. Let $z_i \in \gamma_i$ for $i = 1, ..., 4$, we have

\[
\mathbb{E}\left(m_N^\Delta(z_1)m_N^\Delta(z_2)\right) = f(z_1, z_2) + g(z_1)g(z_2) + O_\omega (kt_{1.2}N^{-\zeta - \zeta}).
\]

and

\[
\operatorname{Cov}\left(m_N^\Delta(z_1)m_N^\Delta(z_2), m_N^\Delta(z_3)m_N^\Delta(z_4)\right)
= f(z_1, z_3)f(z_2, z_4) + f(z_1, z_4)f(z_2, z_3) + g(z_1)g(z_3)f(z_2, z_4)
+ g(z_1)g(z_4)f(z_2, z_3) + g(z_2)g(z_3)f(z_1, z_4) + g(z_2)g(z_4)f(z_1, z_3) + O_\omega (E).
\]

Based on Proposition 4.2, we can now prove (3.9).

Proof of (3.9). By (3.6) and Lemma 4.1, we have

\[
\oint_{\gamma_1} \oint_{\gamma_2} \oint_{\gamma_3} \oint_{\gamma_4} T_k(z_1)T_k(z_2)T_l(z_3)T_l(z_4) \mathcal{E} \prod_{i=1}^4 dz_i < \frac{kl}{N^4}\gamma_5.
\]

This together with (4.1) and (4.8) leads to

\[
\operatorname{Cov}\left((t_k(H)t_k(H), t_l(H)t_l(H))\right) = \frac{1}{16\pi} \oint_{\gamma_1} \oint_{\gamma_2} \oint_{\gamma_3} \oint_{\gamma_4} T_k(z_1)T_k(z_2)T_l(z_3)T_l(z_4)
\times \left(f(z_1, z_3)f(z_2, z_4) + f(z_1, z_4)f(z_2, z_3) + g(z_1)g(z_3)f(z_2, z_4)
+ g(z_1)g(z_4)f(z_2, z_3) + g(z_2)g(z_3)f(z_1, z_4) + g(z_2)g(z_4)f(z_1, z_3)\right) \prod_{i=1}^4 dz_i
+ O_\omega(klN^{-(4+\zeta)}).
\]

Let us denote $q_i = m(z_i)/2$ for $i = 1, ..., 4$. To compute the RHS of (4.10), we first consider

\[
\oint_{\gamma_1} \oint_{\gamma_3} T_k(z_1)T_l(z_3)f(z_1, z_3)dz_1dz_3 = \oint_{|q_1| = \rho_1} \oint_{|q_3| = \rho_3} T_k(z_1)T_l(z_3)f(z_1, z_3)dz_1dz_3 \tag{4.11}
\]

for some $0 < \rho_1 < \rho_3 < 1$. Here we used the assumption $k \geq l$, and the fact that $T_k(z_1)T_l(z_3)f(z_1, z_3)$ is analytic for $z_1, z_3 \in \mathbb{C} \setminus [-1, 1]$. By

\[
z_i = \frac{-q_i - q_i^{-1}}{2}, \quad \sqrt{z_i^2 - 1} = \frac{q_i - q_i^{-1}}{2}, \quad \text{and} \quad q_i' = -\frac{q_i}{\sqrt{z_i^2 - 4}},
\]

we can write

\[
f(z_1, z_3) = \frac{q_1'q_3'}{2N^2q_1q_3} \left(1 - \left(\frac{1 + q_1q_3}{1 - q_1q_3}\right)^2\right) - \frac{32}{N^2} \varepsilon_4q_1q_3q_3' - \frac{\sigma^2 - 2}{N^2} q_1' q_3'.
\]

Thus

\[
T_k(z_1) = \frac{1}{2}(q_1^k + q_1^{-k}).
\]

Thus

\[
(4.11) = \oint_{|q_1| = \rho_1} \oint_{|q_3| = \rho_3} \frac{1}{4}(q_1^k + q_1^{-k})(q_3' + q_3'^{-l})
\times \left(\frac{1}{2N^2q_1q_3} \left(1 - \left(\frac{1 + q_1q_3}{1 - q_1q_3}\right)^2\right) - \frac{32}{N^2} \varepsilon_4q_1q_3 - \frac{\sigma^2 - 2}{N^2}\right) dq_1 dq_3
= :(A) + (B) + (C) \tag{4.12}
\]
By writing \( q_i = \rho_i e^{i\theta_i} \) and using \( \rho_1, \rho_3 \in (0,1) \), we have

\[
(A) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{4} \left( \rho_1^k e^{i k \theta_1} + \rho_1^{-k} e^{-i k \theta_1} \right) \left( \rho_3^l e^{i l \theta_3} + \rho_3^{-l} e^{-i l \theta_3} \right) \left( \frac{i^2}{2 N^2} \left( 1 - \frac{1 + \rho_1 \rho_3 e^{(\theta_1 + \theta_3)}}{1 - \rho_1 \rho_3 e^{(\theta_1 + \theta_3)}} \right)^2 \right) d\theta_1 d\theta_3 \]

\[
= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{4} \left( \rho_1^k e^{i k \theta_1} + \rho_1^{-k} e^{-i k \theta_1} \right) \left( \rho_3^l e^{i l \theta_3} + \rho_3^{-l} e^{-i l \theta_3} \right) \times \frac{-1}{2 N^2} (-4) \sum_{n=1}^{\infty} n(\rho_1 \rho_3)^n e^{in(\theta_1 + \theta_3)} d\theta_1 d\theta_3 = \frac{2 \pi^2 k}{2 N^2} \delta_{kl} .
\]

Similarly, we can show that

\[
(B) = \frac{32 \pi^2}{N^2} e^{i \delta_{kl} \delta_{l2}} , \quad \text{ and } \quad (C) = \frac{(\sigma^2 - 2) \pi^2}{N^2} \delta_{kl} ,
\]

which implies

\[
\oint_{\gamma_1} \oint_{\gamma_2} T_k(z_1) T_l(z_3) f(z_1, z_3) dz_1 dz_3 = \frac{2 \pi^2 k}{2 N^2} \delta_{kl} + \frac{32 \pi^2}{N^2} e^{i \delta_{kl} \delta_{l2}} + \frac{(\sigma^2 - 2) \pi^2}{N^2} \delta_{11} \delta_{kl} .
\]

Similarly, we have

\[
g(z_1) = \frac{q_1 q_1'}{N(1 - q_1^2)} + \frac{16 c_4 q_1^3 q_1'}{N} + \frac{(\sigma^2 - 2) q_1 q_1'}{N} ,
\]

thus

\[
\oint_{\gamma_1} T_k(z_1) g(z_1) dz_1 = \oint_{|q_1| = q_1} \frac{1}{2} (q_1^k + q_1^{-k}) \left( \frac{q_1}{N(1 - q_1^2)} + \frac{16 c_4 q_1^3}{N} + \frac{(\sigma^2 - 2) q_1}{N} \right) dq_1
\]

\[
= \frac{\pi}{2 N} \left( 1 + (-1)^k \right) + \frac{16 c_4}{N} \delta_{kl} + \frac{\pi (\sigma^2 - 2) i}{N} \delta_{kl} , \quad (4.13)
\]

where in the second step we used the change of variable \( q_1 = \rho_1 e^{i\theta} \). Plugging the above into (4.10) we have (3.9) as desired.

\[\square\]

5. PROOF OF PROPOSITION 4.2

We shall only prove (4.8). The proof of (4.7) follows in a similar fashion. Let us write \( z_i = E_i + i\eta \) for \( i = 1, 2, ..., 4 \). For simplicity, we denote by

\[\mathcal{G} = G(z_1), \quad \mathcal{S} = G(z_2), \quad \mathcal{T} = G(z_3), \quad \mathcal{U} = G(z_4).\]

In sections 5.1 to 5.3 we shall prove the following centered estimate.

**Proposition 5.1.** Let \( z_i \in \gamma_i \) for \( i = 1, ..., 4 \), we have

\[
\text{Cov} \left( (\mathcal{G}, \mathcal{S}), (\mathcal{T}, \mathcal{U}) \right) = f(z_1, z_3) f(z_2, z_4) + f(z_1, z_4) f(z_2, z_3) + O_{\prec} (\mathcal{E}) , \quad (5.1)
\]

where \( f(\cdot, \cdot) \) is defined as in (4.5).
5.1. The first step. To study the LHS of (5.1), it suffices to estimate the 4-point and
2-point correlation functions of the Green functions
\[ E(G)\langle S \rangle \langle T \rangle \langle U \rangle, \quad E(G)\langle S \rangle. \] (5.2)

By the resolvent identity
\[ zG = HG - I, \] (5.3)
we have
\[ z_1 E(G)\langle S \rangle \langle T \rangle \langle U \rangle = \frac{1}{N} \sum_{i,j} EH_{ij} G_{ji} \langle S \rangle \langle T \rangle \langle U \rangle. \]

We compute the RHS of the above using Lemma 2.3 with \( \ell = 3 \), \( h = H_{ij} \), \( f \equiv f_{ij} = G_{ji} \langle S \rangle \langle T \rangle \langle U \rangle \), and get
\[ z_1 E(G)\langle S \rangle \langle T \rangle \langle U \rangle = \sum_{n=1}^{3} \frac{1}{n!} \sum_{i,j} C_{n+1}(H_{ij}) E \partial^0_{ij} (G_{ji} \langle S \rangle \langle T \rangle \langle U \rangle) + \mathcal{R} = \sum_{n=1}^{3} L_n + \mathcal{R}, \]
where we abbreviate \( \partial_{ij} := \frac{\partial}{\partial H_{ij}} \), and \( \mathcal{R} \) is the remainder term satisfying
\[ \mathcal{R} < \frac{1}{N} \sum_{i,j} \left( \frac{\mathbb{E} \sup_{|x| \leq t} |\partial^4_{ij} f(H^{ij} + x\Delta^{ij})|^2 \cdot \mathbb{E} |H^{ij}_{ij}|^3 1_{|H_{ij}| > t}}{1 + \delta_{ij}} \right)^{1/2} \]
\[ + \frac{1}{N} \sum_{i,j} \mathbb{E} |H_{ij}|^5 \cdot \mathbb{E} \sup_{|x| \leq t} |\partial^1_{ij} f(H^{ij} + x\Delta^{ij})| \] (5.4)
for any \( t > 0 \). Here we define \( \Delta^{ij} \in \mathbb{C}^{N \times N} \) such that \( \Delta^{ij}_{xy} = \delta_{x,y} (1 + \delta_{xy})^{-1} \), and \( H^{ij} := H - H_{ij} \Delta^{ij} \). Note that
\[ L_1 = \frac{1}{N} \sum_{i,j} \frac{1 + \delta_{ij}}{4N} E(\partial_{ij} G_{ij} \langle S \rangle \langle T \rangle \langle U \rangle) + \frac{1}{N} \sum_{i,j} \frac{1 + \delta_{ij}}{4N} E G_{ij} \partial_{ij} \langle S \rangle \langle T \rangle \langle U \rangle + K = : (A) + (B) + K, \]
where
\[ K := \frac{\sigma^2 - 2}{4N^2} \sum_{i} E \partial_{ii} (G_{ii} \langle S \rangle \langle T \rangle \langle U \rangle). \]

By the differential rule
\[ \partial_{ij} G_{xy} = -(G_{xi} G_{yx} + G_{xj} G_{iy})(1 + \delta_{ij})^{-1}, \]
we have
\[ (A) = -\frac{1}{4N^2} E(G_{ii} G_{jj} + G_{ij}^2) \langle S \rangle \langle T \rangle \langle U \rangle \]
\[ = -\frac{1}{4} \left( 2E(G)\langle S \rangle \langle T \rangle \langle U \rangle + E(G)^2 \langle S \rangle \langle T \rangle \langle U \rangle - E(G)^2 \langle S \rangle \langle T \rangle \langle U \rangle + \frac{1}{N} E(G^2) \langle S \rangle \langle T \rangle \langle U \rangle \right) \]
and
\[ (B) = -\frac{1}{4} \left( \frac{2}{N^2} EGS^2 \langle T \rangle \langle U \rangle + \frac{2}{N^2} EGT^2 \langle S \rangle \langle U \rangle + \frac{2}{N^2} EGU^2 \langle S \rangle \langle T \rangle \right). \]
Thus we arrive at
\[
E(GS(T)U) = \frac{1}{4s_1} \left( E(G^2S(T)U) - E(G2E(S)T)U + \frac{1}{N} E(G^2) (S)T(U) \right.
+ \frac{2}{N^2} \frac{E(GS^2(T)U)}{L_2} + \frac{2}{N^2} \frac{E(GT^2(S)U)}{L_3} + \frac{2}{N^2} \frac{E(GU^2(T))}{L_4} - 4L_2 - 4L_3 - 4K - 4R \right),
\]
where
\[
s_i := z_i + \frac{1}{2} E(G(z_i))
\]
for \(i = 1, \ldots, 4\). On the RHS of (5.5), the first three terms and \(L_i/s_1, R_i/s_1\) are the error terms, while other terms contain the leading contributions. The analysis of the error terms in (5.5) is broken down into the estimates in the following section.

5.2. The estimates. We begin with some preliminary estimates on Green functions. Recall \(u_i, i = 1, \ldots, 4\) defined as in (4.2).

**Lemma 5.2.** Let \(u, v \in \mathbb{C}^N\) be deterministic satisfying \(\|u\|_2 = \|v\|_2 = 1\). We have
\[
\frac{1}{|s_1|} \leq \frac{1}{u_1} \leq k, \quad |G(z_1) - m(z_1)| \leq \frac{k}{N u_1}, \quad \|G_{uv}(z_1) - m(z_1)\|_1 < \frac{\sqrt{k}}{N u_1}
\]
and
\[
(G^2)_{uv}(z_1) \leq \frac{1}{u_1^{3/2}}, \quad G^2(z_1) \leq \frac{1}{u_1}, \quad G^2(z_1) \leq \frac{\sqrt{N}}{u_1^{3/2}}.
\]
uniformly for \(z_1 \in \gamma_1\).

**Proof.** The first three relations are simple consequences of Theorem 2.2 and the construction of our contour \(\gamma_1\). By Theorem 2.2 and the Helffer-Sjöstrand formula it can be shown (see [21, Lemma 4.4]) that
\[
(G^2)_{uv}(z_1) \leq |m'(z_1)| + \frac{\sqrt{\eta_1}}{|\eta_1|} \frac{1}{u_1^{3/2}}
\]
when \(|E_1| \leq 1 + 1/(2k^2)\). On the other hand, (2.3) and \(k \leq u_\omega\) imply
\[
(G^2)_{uv}(z_1) = \sum_i G_{ui} G_{iv} \leq \sum_i |(u, e_i)| \langle e_i, v \rangle + \sum_i \left( |(u, e_i)| + |\langle e_i, v \rangle| \right) \leq \sqrt{N} |E_1 - 1| + |\eta_1|^{3/4} + \frac{1}{\sqrt{|E_1 - 1| + |\eta_1|}} \leq \frac{1}{u_1}
\]
when \(|E_1| > 1 + 1/(2k^2)\), and \(e_i\) is the standard \(i\)-th basis vector of \(\mathbb{R}^N\). This prove the fourth relation. By Theorem 2.2 and the Helffer-Sjöstrand formula, we see that
\[
G^2(z_1) < m'(z_1) + \frac{1}{N |\eta_1|^2} \leq \frac{1}{u_1}
\]
when \(|E_1| \leq 1 + 1/(2k^2)\). Together with (5.6) we deduce the fifth relation. The proof of the last relation follows in a similar fashion. \(\square\)

**Lemma 5.3.** Let \(\mathcal{R}\) be as in (5.5). We have
\[
R/s_1 < \mathcal{E}.
\]
Proof. Recall that \( \Delta^{ij} \in \mathbb{C}^{N \times N} \) such that \( \Delta^{ij}_{y} = (\delta_{xy} + \delta_{yz} + \delta_{yx})(1 + \delta_{xy})^{-1} \), and \( H^{ij} := H - H_{ij} \Delta^{ij} \). Fix \( i, j \) and set \( S = H_{ij} \Delta^{ij} \), \( \hat{G} := (H^{ij} - z_1)^{-1} \). By repeating the proof of Theorem 2.2 one can show that

\[
|\hat{G}(z_1) - m(z_1)| < \frac{k}{Nu_1}, \quad \max_{i,j}|\hat{G}_{ij}(z_1) - \delta_{ij}m(z_1)| < \sqrt{\frac{k}{N u_1}} \tag{5.7}
\]

uniformly for \( z_1 \in \gamma_{1,k} \). We have the resolvent expansion

\[
G = \hat{G} - (\hat{G}S)(\hat{G}S)^2 + \cdots.
\]

Note that only two entries of \( S \) are nonzero, and they are stochastically dominated by \( N^{-1/2} \).

Thus by (5.7), (5.8), and the fact \( \hat{G} \) is independent of \( H_{ij} \), we have

\[
\max_{x,y \in [1,N]} \sup_{|H_{ij}| \leq N^{-1/2+\epsilon}} |G_{xy}| < 1 + N^{-1+2\epsilon} \max_{x,y \in [1,N]} \sup_{|H_{ij}| \leq N^{-1/2+\epsilon}} |G_{xy}|,
\]

which implies

\[
\max_{x,y \in [1,N]} \sup_{|H_{ij}| \leq N^{-1/2+\epsilon}} |G_{xy}| < 1.
\]

Inserting the above into (5.8) and applying (5.7), we have

\[
\max_{x,y \in [1,N]} \sup_{|H_{ij}| \leq N^{-1/2+\epsilon}} |G_{xy}| \leq \sqrt{\frac{k}{Nu_1}} \quad \text{and} \quad \max_{x \in [1,N]} \sup_{|H_{ij}| \leq N^{-1/2+\epsilon}} |G_{xx}| < 1, \tag{5.9}
\]

\[
\sup_{|H_{ij}| \leq N^{-1/2+\epsilon}} |\hat{G} - m(z_1)| \leq \frac{k}{Nu_1}, \tag{5.10}
\]

and

\[
\max_{x,y \in [1,N]} \sup_{|H_{ij}| \leq N^{-1/2+\epsilon}} |G_{xy}^2| < \frac{k}{u_1}. \tag{5.11}
\]

By setting \( t = N^{-1/2+\epsilon} \) in (5.4), we see that

\[
\mathcal{R}/s_1 < \mathcal{R}/u_1 < \frac{1}{N u_1} \sum_{i,j} \left( E \sup_{|x| \leq |H_{ij}|} \left| \partial_{ij}^1 f(H^{ij} + x \Delta^{ij}) \right|^2 \cdot E \sup_{|x| \leq |H_{ij}|} \left| \partial_{ij}^0 f(H^{ij} + x \Delta^{ij}) \right|^2 \right)^{1/2} + \frac{1}{N u_1} \sum_{i,j} E |H_{ij}|^4 \cdot E \sup_{|x| \leq N^{-1/2+\epsilon}} \left| \partial_{ij}^1 f(H^{ij} + x \Delta^{ij}) \right|, \tag{5.12}
\]

where recall \( f \equiv f_{ij} = G_{ij} \langle (S) (T) (U) \rangle \). Let us estimate the second term on RHS of (5.12).

By our assumption on \( H \), this term is bounded by

\[
O_{\prec}(N^{-3/2}u_1^{-1}) \cdot \max_{i,j \in [1,N]} E \sup_{|x| \leq N^{-1/2+\epsilon}} \left| \partial_{ij}^1 f(H^{ij} + x \Delta^{ij}) \right|.
\]

In terms of (5.9) – (5.11) the above is bounded by

\[
O_{\prec}(N^{-3/2}u_1^{-1}) \cdot \frac{kl^2}{N^{3}u_2u_3u_4} \ll \frac{kl}{N^{4+\epsilon}u_1u_2u_3u_4} \ll \mathcal{E}
\]

as desired. The first term on RHS of (5.12) can be easily bounded by \( O(N^{-D}) \) for any fixed \( D > 0 \), as we have \( \max_{i,j} |H_{ij}| \ll N^{-1/2} \).

\begin{lemma}
The first three terms on RHS of (5.5) are bounded by \( O_{\prec}(\mathcal{E}) \).
\end{lemma}
Proof. We consider the first term on RHS of (5.5). The proof for other two terms is similar. By \( z_4 U = H U - I \) and cumulant expansion, we have

\[
\frac{1}{4s_1} E(G)^2(T)(U) = \frac{1}{16s_1 s_4} \left(E(G)^2(T)(U)^2 - E(U)^2 E(G)^2(T)\right) \\
+ \frac{1}{N} E(G)^2(T)(U)^2 + \frac{4}{N^2} E U G^2(G)(T) + \frac{2}{N^2} E U S^2(G)^2(T) \\
+ \frac{2}{N^2} E U T^2(G)^2(S) - 4R^{(1)} \right). \tag{5.13}
\]

By Lemma 5.2, the first three terms on RHS of (5.13) are bounded by \( E \). Note that

\[
\frac{m(z_1) - m(z_4)}{z_1 - z_4} < 1 + \frac{|\sqrt{z_1^2 - 1}| + |\sqrt{z_4^2 - 1}|}{|z_1 - z_4|} < 1 + (u_1 + u_4) \left(\frac{k}{u_1} + \frac{l}{u_4}\right) < k + l
\]

uniformly for \( z_1 \in \gamma_1 \) and \( z_4 \in \gamma_4 \). Together with Lemma 5.2 and resolvent identity we have

\[
U G^2 = \frac{U - G}{z_4 - z_1} \times \frac{G^2}{z_4 - z_1} = \frac{m(z_4) - m(z_1)}{(z_4 - z_1)^2} + \frac{U - m(z_4) + m(z_1) - G}{(z_4 - z_1)^2} \\
\leq \frac{1}{|z_1 - z_4|} + \frac{1}{|z_1 - z_4|^2}\left(\frac{k}{N u_1} + \frac{l}{N u_4}\right) + \frac{1}{|z_1 - z_4|} < k + l
\]

uniformly for \( z_1 \in \gamma_1 \) and \( z_4 \in \gamma_4 \). Hence the fourth term on RHS of (5.13) is bounded by

\[
\frac{1}{u_1 u_4} \frac{1}{N^2} \frac{k^2}{|z_1 - z_4|} \frac{1}{N^{1/2+1/6}} < E.
\]

Similarly, the fifth and sixth terms on RHS of (5.13) are also bounded by \( O_{\gamma}(E) \). As in Lemma 5.3, we can apply Lemma 5.2 to show that

\[
\frac{R^{(1)}}{s_1 s_4} < \frac{1}{u_1 u_4} \cdot \frac{k^3}{N^4 u_1 u_2 u_3} < E.
\]

This completes the proof. \( \square \)

Lemma 5.5. We have

\[
L_{2,1}/s_1 < E. \tag{5.15}
\]

Proof. Let us examine the term

\[
L_{2,1}/s_1 := \frac{1}{2s_1 N^2} \sum_{i,j} C_3(H_{ij}) (1 + \delta_{ij})^{-2} E G_{ij} G_{jj}(S^2)_{ij} (T) (U) = \frac{1}{2s_1 N^2} \sum_{i,j} a_{ij} E G_{ij} G_{jj}(S^2)_{ij} (T) (U),
\]

where we write \( a_{ij} := N^3 C_3(H_{ij})(1 + \delta_{ij})^{-2} \) so that \( \max_{i,j} |a_{ij}| = O(1) \). By \( z_4 U = H U - I \) and cumulant expansion, we have

\[
L_{2,1}/s_1 = \frac{1}{8s_1 s_4 N^7/2} \sum_{i,j} a_{ij} \left(\frac{E G_{ij} G_{jj}(S^2)_{ij} (T) (U)}{N^2} E(U)^2 E G_{ij} G_{jj}(S^2)_{ij} (T) \right) \\
+ \frac{1}{N} E G_{ij} G_{jj}(S^2)_{ij} (T) (U)^2 + \frac{2}{N^2} E U T^2 G_{ij} G_{jj}(S^2)_{ij} (T) + \frac{2}{N^2} (G^2 U)_{ij} G_{ij} (S^2)_{ij} (T) \\
+ \frac{2}{N^2} (G^2 U)_{jj} G_{ij} (S^2)_{ij} (T) + \frac{4}{N^2} (S^3 U)_{ij} G_{ij} (T) - 4R^{(2,ij)} \right). \tag{5.16}
\]
We can rewrite the first term in the above to
\[ -\frac{1}{s_1^4 N^{7/2}} \sum_{i,j} a_{ij} \mathbb{E} \left( (G_{ii} - m(z_1))(G_{jj} - m(z_1)) + m(z_1)G_{ii} - m(z_1) \right) + m(z_1)(G_{jj} - m(z_1)) + m(z_1)^2 \right) (S^2)_{ij} \langle T \rangle \langle U \rangle^2, \]
and it is bounded by \( O_s(\mathcal{E}) \) using Lemma 5.2. Similarly, the second and third terms on RHS of (5.16) are also bounded by \( O_s(\mathcal{E}) \). By (5.14) and Lemma 5.2, the fourth to sixth terms on RHS of (5.16) are bounded by \( O_s(\mathcal{E}) \). Similar as in Lemma 5.3, we can apply Lemma 5.2 to show that
\[ \frac{1}{s_1^4 N^{7/2}} \sum_{i,j} a_{ij} \mathcal{R}^{(2,ij)} \ll \mathcal{E}. \]
Thus we conclude the estimate of \( L_{2,1}/s_1 \). The estimate for other terms in \( L_2 \) is similar, and we omit the details.

**Lemma 5.6.** We have
\[ L_{3}/s_1 = -\frac{2}{s_1 N^2} \sum_{i,j} C_4(H_{ij}) \mathbb{E} G_{ii} G_{jj} \left( (S^2)_{ii} S_{jj} \langle T \rangle \langle U \rangle + (T^2)_{ii} T_{jj} \langle S \rangle \langle U \rangle + (U^2)_{ii} U_{jj} \langle T \rangle \right) + O_s(\mathcal{E}). \]

**Proof.** By applying the differentials carefully, it is easy to reveal the leading terms on RHS of the above. In addition, we see that the most dangerous error term is
\[ L_{3,1}/s_1 := -\frac{1}{s_1 N^2} \sum_{i,j} C_4(H_{ij}) \mathbb{E} (G_{ii}^2 G_{jj}^2) \langle S \rangle \langle T \rangle \langle U \rangle. \]
By writing \( z_1 G_{ii} = (HG)_{ii} - 1 \) and cumulant expansion, we have
\[ L_{3,1}/s_1 = \frac{1}{(z_1 + \mathbb{E} G_{ii}) s_1 N^2} \sum_{i,j} C_4(H_{ij}) \mathbb{E} \left( (G_{ii}^2 G_{jj}^2) \langle G \rangle + 3N^{-1}(G_{ii}^2) G_{jj} + 5N^{-1}G_{ij}(G_{ii}^2)_{ij} \langle S \rangle \langle T \rangle \langle U \rangle) + 2N^{-2}G_{ii}^2 G_{jj}^2 \langle T \rangle \langle U \rangle + 2N^{-2}(GT^2)_{ii} G_{ii} G_{jj} \langle S \rangle \langle U \rangle + 2N^{-2}(GU)^2_{ii} G_{ii} G_{jj} \langle S \rangle \langle T \rangle - 4\mathcal{R}^{(3,ij)} \right). \]
By \( 1/|z_1 + \mathbb{E} G_{ii}| = O(1) \) and Lemma 5.2, one readily checks that
\[ \frac{-4}{(z_1 + \mathbb{E} G_{ii}) s_1 N^2} \sum_{i,j} C_4(H_{ij}) \mathcal{R}^{(3,ij)} \ll \mathcal{E}, \]
and other terms in \( L_{3,1}/s_1 \) satisfy the same bound. Other terms in \( L_3/s_1 \) can be estimated as in the proofs of Lemmas 5.4 and 5.5; we omit the details.

**Lemma 5.7.** We have
\[ K/s_1 = -\frac{\sigma^2 - 2}{48s_1 N^3} \sum_i \mathbb{E} \left( G_{ii} (S^2)_{ii} \langle T \rangle \langle U \rangle + \mathbb{E} G_{ii} (T^2)_{ii} \langle S \rangle \langle U \rangle + \mathbb{E} G_{ii} (U^2)_{ii} \langle T \rangle \right) + O_s(\mathcal{E}). \]

**Proof.** We have
\[ K/s_1 = -\frac{\sigma^2 - 2}{48s_1 N^2} \sum_i \left( \mathbb{E} (G_{ii}^3) \langle S \rangle \langle T \rangle \langle U \rangle + \frac{1}{N} \mathbb{E} G_{ii} (S^2)_{ii} \langle T \rangle \langle U \rangle + \frac{1}{N} \mathbb{E} G_{ii} (T^2)_{ii} \langle S \rangle \langle U \rangle + \frac{1}{N} \mathbb{E} G_{ii} (U^2)_{ii} \langle S \rangle \langle T \rangle \right). \] (5.17)
By Theorem 1.2 and Lemma 5.2, the first term on RHS of (5.17) can be bounded by

\[ O_\prec \left( \frac{1}{u_1 N^2} \cdot N \cdot \sqrt{\frac{k}{N u_1 N^3 u_2 u_3 u_4}} \right) = O_\prec (\mathcal{E}). \]

This completes the proof. \(\square\)

5.3. Proof of Proposition 5.1. Now we insert Lemmas 5.3 - 5.7 to (5.5), and after rearranging the terms, we get

\[
\begin{align*}
\mathbb{E}(G \langle S \rangle \langle T \rangle \langle U \rangle) &= \mathbb{E}(G S^2 + 4 \sum_{i,j} C_4(H_{ij}) G_{ii} G_{jj} (S^2)_{ii} S_{jj} + \frac{\sigma^2 - 2}{2N} \sum_i G_{ii} (S^2)_{ii}) \langle T \rangle \langle U \rangle \\
&- \frac{1}{2s_1 N^2} \mathbb{E}(G T^2 + 4 \sum_{i,j} C_4(H_{ij}) G_{ii} G_{jj} (T^2)_{ii} T_{jj} + \frac{\sigma^2 - 2}{2N} \sum_i G_{ii} (T^2)_{ii}) \langle S \rangle \langle U \rangle \\
&- \frac{1}{2s_1 N^2} \mathbb{E}(G T^2 + 4 \sum_{i,j} C_4(H_{ij}) G_{ii} G_{jj} (U^2)_{ii} U_{jj} + \frac{\sigma^2 - 2}{2N} \sum_i G_{ii} (U^2)_{ii}) \langle S \rangle \langle T \rangle + O_\prec (\mathcal{E}).
\end{align*}
\]

The terms on RHS of (5.18) can be further computed using the following lemma. The proof is again by cumulant expansion and Lemma 5.2. We omit the details.

Lemma 5.8. We have

\[
\begin{align*}
- \frac{1}{2s_1 N^2} \mathbb{E}(G S^2 + 4 \sum_{i,j} C_4(H_{ij}) G_{ii} G_{jj} (S^2)_{ii} S_{jj} + \frac{\sigma^2 - 2}{2N} \sum_i G_{ii} (S^2)_{ii}) \langle T \rangle \langle U \rangle &= \frac{1}{4s_1 s_3 N^4} \mathbb{E}(G S^2 + 4 \sum_{i,j} C_4(H_{ij}) G_{ii} G_{jj} (S^2)_{ii} S_{jj} + \frac{\sigma^2 - 2}{2N} \sum_i G_{ii} (S^2)_{ii}) \\
&\times \left( T U^2 + 4 \sum_{i,j} C_4(H_{ij}) T_{ii} T_{jj} (U^2)_{ii} U_{jj} + \frac{\sigma^2 - 2}{2N} \sum_i T_{ii} (U^2)_{ii} \right) + O_\prec (\mathcal{E}) \\
&= f(z_1, z_2) f(z_3, z_4) + O_\prec (\mathcal{E}),
\end{align*}
\]

where \( f(\cdot, \cdot) \) is defined as in (4.5).

As a result, we have

\[
\mathbb{E}(G \langle S \rangle \langle T \rangle \langle U \rangle) = f(z_1, z_2) f(z_3, z_4) + f(z_1, z_3) f(z_2, z_4) + f(z_1, z_4) f(z_2, z_3) + O_\prec (\mathcal{E}).
\]

By repeating the above proof, one also gets that

\[
\mathbb{E}(G \langle S \rangle) = f(z_1, z_3) + O_\prec (k t_{1,2} N^{-2-\zeta})
\]

and

\[
\mathbb{E}(T \langle U \rangle) = f(z_3, z_4) + O_\prec (k t_{1,2} N^{-2-\zeta}).
\]

The above three relations conclude the proof of (5.1).

5.4. The explicit shift. In this section we replace the centered random variables \( \langle G(z_i) \rangle \) in (5.1) by the explicit shift \( G(z_i) - m(z_i) \). We have the following result.

Lemma 5.9. Recall \( g(\cdot) \) from (4.6). We have

\[
\mathbb{E} G - m(z_1) = g(z_1) + O_\prec \left( \frac{k}{N^{3/2} u_1^2} \right).
\]

uniformly for \( z_1 \in \gamma_1 \).
Proof. By \( z_1E G = \mathbb{E} H G - 1 \) and cumulant expansion formula, we have

\[
(\mathbb{E} G)^2 + 4z_1 E G + 4 + \frac{1}{N} E G^2 + \mathbb{E}(G)^2 - 4L_2^{(4)} - 4L_3^{(4)} - 4K^{(4)} - 4R^{(4)} = 0,
\]

where

\[
L_n^{(4)} := \frac{1}{n!} \frac{1}{N} \sum_{i,j} C_{n+1}(H_{ij}) E \partial^n_{ij} G_{ji}, \quad K^{(4)} = -\frac{\sigma^2 - 2}{4N^2} \sum_i E G_{ii}^2
\]

and \( R^{(4)} \) is the remainder term. By Lemma 5.2 we have

\[
-4L_3^{(4)} = \frac{4e_4}{N} (m(z_1))^4 + O\left(\frac{k}{N^{2u_1}}\right),
\]

\[
-4K^{(4)} = \frac{\sigma^2 - 2}{N} (m(z_1))^2 + O\left(\frac{k^{1/2}}{N^{3/2u_1/2}}\right),
\]

and

\[
-4L_2^{(4)} - 4R^4 \prec \left(\frac{k}{N^{1/2u_1}}\right)^{3/2} \frac{1}{\sqrt{N}} + \frac{1}{N^{3/2}}.
\]

Similarly, by cumulant expansion and Lemma 5.2, we have

\[
\mathbb{E}(G)^2 = \frac{1}{-4z - 2E G} \left( \mathbb{E}(G)^3 + \frac{1}{N} \mathbb{E}(G) (G^2) + \frac{2}{N^2} E G^3 - 4R^{(5)} \right) \prec \frac{k}{N^{3/2u_1^2}}
\]

and

\[
E G^2 = \frac{1}{-4z - 2E G} \left( 4E G + \frac{1}{N} E G^3 + 2E G (G^2) - 4R^{(6)} \right) = m'(z_1) + O\left(\frac{k}{N^{1/2u_1}}\right).
\]

By (5.23) – (5.27), we can rewrite (5.22) to

\[
(\mathbb{E} G - m(z_1)) E (S) (T) (U) = O\prec (E),
\]

which implies

\[
E(m_N^{(z_1)} m_N^{(z_2)} m_N^{(z_3)} m_N^{(z_4)}) - E(G) E(S) E(T) E(U)
\]

implies

\[
= \frac{1}{2} \sum_{a,b,c,d=1,2,3,4} (E(G)(z_a) - m(z_a))(E(G)(z_b) - m(z_b)) E(G)(z_c) E(G)(z_d)
\]

\[
+ (E(G) - m(z_1))(E(S) - m(z_2)) (E(T) - m(z_3)) (E(U) - m(z_4)) + O\prec (E),
\]

and

\[
E(m_N^{(z_1)} m_N^{(z_2)}) E(m_N^{(z_3)} m_N^{(z_4)}) - E(G) E(S) E(T) E(U)
\]

implies

\[
= (E(G) - m(z_1))(E(S) - m(z_2)) E(T) E(U) + (E(T) - m(z_3)) (E(U) - m(z_4)) E(G) E(S)
\]

\[
+ (E(G) - m(z_1))(E(S) - m(z_2)) (E(T) - m(z_3)) (E(U) - m(z_4)) + O\prec (E).
\]

Thus we have

\[
Cov(m_N^{(z_1)} m_N^{(z_2)} m_N^{(z_3)} m_N^{(z_4)}) - Cov(G) Cov(S) Cov(T) Cov(U)
\]

implies

\[
= (E(G) - m(z_1))(E(T) - m(z_3)) E(S) E(U) + (E(T) - m(z_3)) (E(U) - m(z_4)) E(G) E(S)
\]

\[
+ (E(S) - m(z_2)) (E(U) - m(z_4)) E(G) E(T) + (E(S) - m(z_2)) (E(T) - m(z_3)) E(G) E(U) + O\prec (E).
\]
By computing $\mathbb{E}(G(z_i)|G(z_j))$ as in (5.21) and applying Lemma 5.9, we have

$$\text{Cov}(m_N^N(z_1)m_N^N(z_2), m_N^N(z_3)m_N^N(z_4)) - \text{Cov}(\langle G(S), T(U) \rangle)$$

$$= g(z_1)g(z_3)f(z_2, z_4) + g(z_1)g(z_4)f(z_2, z_3)$$

$$+ g(z_2)g(z_4)f(z_1, z_3) + g(z_2)g(z_3)f(z_1, z_4) + O_\prec(\mathcal{E}) \cdot (5.28)$$

Using the above relation and Proposition 5.1, we conclude the proof of (4.8).

6. Further Discussion

In this section, we make some further remarks on the CvM statistics.

- Shortcoming of the CvM statistics

From application point of view, although the original CvM statistic $A_N$ is a robust statistic, it has its own shortcoming. For instance, the statistic $A_N$ is not sensitive to the strength of a low rank deformation of Wigner matrix. By Cauchy interlacing property, a rank one deformation can only cause a change of order $1/N$ for $F_N(t)$. This fact does not depend on the strength of the deformation of the Wigner matrix. Hence, the power of the statistic $A_N$ will not be significantly good even if the rank one deformation is very large, if we use $A_N$ to test the existence of the deformation, say. In other words, the statistic $A_N$ is not sensitive to the possible outlier of the spectrum. The same problem exists for our MCvM $A_{N,\omega}$, since we use $\tilde{\lambda}$ instead of $\lambda_i$ in the definition of $F_{N,\omega}$ in (1.7). However, from the proof of the main result Theorem 1.2, it is clear that we indeed have the same convergence as Theorem 1.2 for the following partial sum

$$\tilde{A}_{N,\omega} = \frac{2}{\pi^4} \sum_{k=1}^{n_\omega} \frac{1}{k^2} t_k(H)^2 - \frac{2}{\pi^2} \sum_{k=1}^{n_\omega} \frac{1}{k(k+2)} t_k(H) t_{k+2}(H) - \frac{1}{\pi^2} (t_1(H))^2 \cdot (6.1)$$

as long as $n_\omega \leq N^{\delta - \epsilon}$, where $\tilde{H}$ is replaced by $H$. From the application point of view, if we use the statistic $\tilde{A}_{N,\omega}$, it is expected to be sensitive to the strength of the low rank deformation since now we have $H$ instead of $\tilde{H}$. In this sense, the partial sum statistic such as $\tilde{A}_{N,\omega}$ has its own advantage in contrast to the original $A_N$. Further, we do not expect that the statistics $\tilde{A}_{N,\omega}$ and $A_N$ have the same fluctuation (approximately) in case of $n_\omega \gg N^{1/3}$. In this regime, we expect that the edge eigenvalues of $H$ will dominate the fluctuation of (6.1).

- On expectation of $A_N$

According to Remark 1.4 and the simulation result in Figure 1, we can formulate the following conjecture

**Conjecture 6.1.** Under the assumption of Theorem 1.2, we conjecture that the following holds

$$N^2(A_N - \mathbb{E}A_N) \overset{d}{\to} \frac{1}{\beta \pi^2} \sum_{k=1}^{\infty} \left( \frac{1}{k} (Z_k^2 - 1) - \frac{1}{\sqrt{k(k+2)}} Z_k Z_{k+2} \right)$$

$$+ \frac{2 - \beta}{\sqrt{\beta \pi}} \sum_{k=1}^{\infty} \left( \frac{k+2}{4k^{3/2}(k+1)} Z_{2k} - \frac{1}{4k^{1/2}k+1} Z_{2k+2} \right) + a_\beta,$$

where $a_\beta$ is defined in Theorem 1.2.

In addition to the proof of the above conjecture (in case it is true), for application purpose, it is also necessary to identify $N^2 \mathbb{E}A_N$ up to the constant order, since the RHS is an order
1 random variable. According to the definition of $A_N$ in (1.1) and the rigidity in (2.5), one can cut the integral in (1.1) to the domain $t \in [-1 + N^{-\varepsilon}, 1 - N^{-\varepsilon}]$. Hence, roughly speaking, in order to identify $N^2 \mathbb{E} A_N$ up to the constant order, it would be enough to estimate $\mathbb{E}(F_N(t) - F(t))^2$ for $t \in [-1 + N^{-\varepsilon}, 1 - N^{-\varepsilon}]$, up to the order of $\frac{1}{N^2}$. Write

$$\mathbb{E}(F_N(t) - F(t))^2 = \text{Var}(F_N(t))^2 + (\mathbb{E}F_N(t) - F(t))^2.$$  

Even in the case of GUE/GOE, only the first order term of $\text{Var}(F_N(t))^2$ is known (see [20, 29]), which is of order $\frac{\log N}{N}$ in the bulk. The subleading order is not available so far. Starting from the representation in (1.4), for $N^2 \mathbb{E} A_N$, one can also turn to identify $N^2 \mathbb{E}(\lambda_i - \gamma)^2$ up to the constant order. However, again, from the reference such as [20, 29], only the leading term of order $\log N$ is precisely available. It is worth mentioning that in the recent work [28], a precise estimate on the constant order of $N^2(\text{Var}(\lambda_i(H)) - \text{Var}(\lambda_i(\text{GOE})))$ is obtained in the bulk, see Theorem 1.4 therein. Here $\lambda_i(H)$ means the $i$-th largest eigenvalue of a general Wigner matrix $H$ and $\lambda_i(\text{GOE})$ means the counterpart for a GOE. Unfortunately, the constant order term of $N^2\text{Var}(\lambda_i(\text{GOE}))$ itself is not available so far.

- **Sample covariance matrices counterpart**

  We remark here that all the discussion and result of the current work for Wigner matrix can be adapted to the sample covariance matrices. Considering the importance of covariance matrices in statistics theory, the CvM statistic and the mesoscopic approximations are potentially useful in many hypothesis testing problems. Especially, as we mentioned earlier, such a statistic has been used in [27] for testing the structure of the covariance matrices. Specifically, for an analogue of the result in this paper for covariance matrix, one needs to consider the expansion of the spectral distribution using the basis of the shifted Chebyshev polynomials diagonalize the covariance structure of the linear spectral statistics of the sample covariance matrices. The detailed derivation for the counterpart of the current result for sample covariance matrices and the discussion for its applications will be considered in a future work.

**Acknowledgement:** We would like to thank Jiang Hu for discussion and simulation. We also thank Yan Fyodorov, Gaultier Lambert, Dong Wang and Lun Zhang for reference and helpful comments.

## Appendix A. A toy model: CUE

In this appendix, we consider a Cramér-von Mises type statistic for the Circular Unitary Ensemble (CUE). Let $U$ be a $N$-dimensional CUE, which is a Haar distributed unitary matrix. And we denote its eigenvalues by $e^{i\theta}, 1 \leq i \leq N$. Let $F_N(x) = \sharp \{i : 0 \leq \theta_i \leq x \}/N, x \in [0, 2\pi]$. Since the eigenvalues of CUE are on a unit circle and all the eigenvalues shall be regarded as “bulk” eigenvalues, we shall modify the definition of $A_N$ to avoid the accumulation of the fluctuation of the eigenvalues around the origin. Hence, we choose the statistic

$$A_N^{\text{CUE}} := \int_0^{2\pi} \int_0^{2\pi} \left( (F_N(y) - F_N(x)) - \frac{y - x}{2\pi} \right)^2 dydx. \quad (A.1)$$

One can also consider the Rains’ statistic in [32], or the following Watson’s statistic

$$\int \left( F_N(x) - F(x) - \int (F_N(t) - F(t))dF(t) \right)^2 dF(x) \quad (A.2)$$
which is independent of the choice of the origin, see [33]. In the sequel, we discuss \( \mathcal{A}_{N}^{\text{CUE}} \) in (A.1) only.

Recall the Fourier transform of the indicator function
\[
1_{[\alpha, \beta]}(\theta) \mapsto \frac{1}{2\pi} (\beta - \alpha) + \frac{1}{2\pi i} \sum_{j=1}^{\infty} \frac{e^{-ij\alpha} - e^{-ij\beta}}{j} e^{ij\theta}.
\]
Therefore, for \( x \leq y \), we have
\[
F_N(y) - F_N(x) - \frac{y - x}{2\pi} \mapsto \frac{1}{2\pi N \pi} \sum_{j=1}^{\infty} \frac{e^{-ijx} - e^{-ijy}}{j} \text{Tr} U^j + \frac{1}{2\pi N \pi} \sum_{j=1}^{\infty} \frac{e^{ijy} - e^{ijx}}{j} \text{Tr} U^j.
\]
Hence, we have
\[
\mathcal{A}_{N}^{\text{CUE}} = -\frac{1}{2N^2\pi^2} \int_0^{2\pi} \left( \int_0^y \left( \sum_{j=1}^{\infty} \frac{e^{-ijx} - e^{-ijy}}{j} \text{Tr} U^j + \sum_{j=1}^{\infty} \frac{e^{ijy} - e^{ijx}}{j} \text{Tr} U^j \right)^2 dx \right) dy
\]
\[
= \frac{4}{N^2} \sum_{j=1}^{\infty} \frac{1}{j^2} \text{Tr} U^j \text{Tr} U^j = \frac{4}{N^2} \sum_{j=1}^{N} \frac{1}{j^2} \text{Tr} U^j \text{Tr} U^j + \frac{4}{N^2} \sum_{j=N+1}^{\infty} \frac{1}{j^2} \text{Tr} U^j \text{Tr} U^j.
\]
To study the expectation and fluctuation of \( \mathcal{A}_{N}^{\text{CUE}} \), we need the following result from [9].

**Theorem A.1** (Theorem 2.1, [9]). (a) Consider \( a = (a_1, \ldots, a_k) \) and \( b = (b_1, \ldots, b_k) \) with \( a_j, b_j \in \{0, 1, \ldots\} \). Let \( Z_1, Z_2, \ldots, Z_k \) be independent standard complex normal random variables. Then \( N \geq (\sum_{j=1}^{k} j a_j) \vee (\sum_{j=1}^{k} j b_j) \),
\[
\mathbb{E} \left( \prod_{j=1}^{k} (\text{Tr} U^j)^{a_j} (\text{Tr} U^j)^{b_j} \right) = \mathbb{E} \left( \prod_{j=1}^{k} (\sqrt{j} Z_j)^{a_j} (\sqrt{j} Z_j)^{b_j} \right) = \delta_{ab} \prod_{j=1}^{k} j^{a_j} a_j!
\]
(b) For any \( j, k \),
\[
\mathbb{E} \left( \text{Tr} U^j \text{Tr} U^k \right) = \delta_{jk} (j \wedge N).
\]

Based on Theorem A.1, we have
\[
\mathbb{E} \mathcal{A}_{N}^{\text{CUE}} = \frac{4}{N^2} \sum_{j=1}^{N} \frac{1}{j} + \frac{4}{N^2} \sum_{j=N+1}^{\infty} \frac{1}{j^2} = \frac{4}{N^2} (\log N + \gamma + 1) + O \left( \frac{1}{N^3} \right)
\]
Next, we need to identify the limiting distribution of \( Q \) after centering. We write
\[
\mathcal{A}_{N}^{\text{CUE}} - \mathbb{E} \mathcal{A}_{N}^{\text{CUE}} = \frac{4}{N^2} \sum_{j=1}^{M} \frac{1}{j^2} \left( \text{Tr} U^j \text{Tr} U^j - j \right) + \frac{4}{N^2} \sum_{j=M+1}^{N} \frac{1}{j^2} \left( \text{Tr} U^j \text{Tr} U^j - j \right) + \frac{4}{N^2} \sum_{j=N+1}^{\infty} \frac{1}{j^2} \left( \text{Tr} U^j \text{Tr} U^j - N \right),
\]
where we can choose \( M \) to be sufficiently large. Our aim is to show that the last two terms in (A.3) are negligible in probability when \( M \) is large. It would be sufficient to show that their second moments are negligible. We claim that
\[
\text{Cov} \left( \text{Tr} U^j \text{Tr} U^j, \text{Tr} U^k \text{Tr} U^k \right)
= \begin{cases} 
  j^2 \delta_{jk}, & \text{if } 1 \leq j + k \leq N \\
  j^2 \delta_{jk} + N - j - k, & \text{if } j + k \geq N + 1, \ j, k \leq N \\
  N^2 \delta_{jk} - (N - |k - j|) \vee 0, & \text{if } j \text{ or } k \geq N
\end{cases}
\]
The above identity can be checked by elementary calculation with the following kernel of the determinantal point process \( \{e^{i\theta_k}\}_{k=1}^N \) for CUE
\[
K_N(\theta, \theta') = \frac{1}{2\pi} e^{i\varepsilon_N(\theta-\theta')} \sin \frac{N(\theta-\theta')}{2} = \frac{1}{2\pi} \sum_{k=0}^{N-1} e^{ik(\theta-\theta')}.
\]

More specifically, using Theorem A.1, we can first conclude that \( \{ \sqrt{\frac{1}{M}} \text{Tr} U_j \}_{j=1}^M \) converges to \( N_C(0, 1) \) variables for any large fixed \( M \). Further, using (A.4) to the second and the third parts in (A.3), one can easily check that these two parts are negligible (in \( M \)) in probability when \( M \) is large. Similarly to the discussions in (3.12) – (3.14), we can conclude that
\[
\frac{N^2}{4} \left( A_N^{\text{CUE}} - E A_N^{\text{CUE}} \right) \implies \sum_{j=1}^{\infty} \frac{1}{j} (|Y_j|^2 - 1),
\]
where \( Y_i \)'s are i.i.d. \( N_C(0, 1) \).

References

[1] Ajanki, O. H., Erdős, L., Krüger, T. (2017). Universality for general Wigner-type matrices. Probability Theory and Related Fields, 169(3-4), 667-727.
[2] Anderson, T. W., Darling, D. A. Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. The annals of mathematical statistics, 25(2), 193-212 (1952).
[3] Arguin, L.-P., Belius, D., Bourgade, P. Maximum of the Characteristic Polynomial of Random Unitary Matrices. Comm. Math. Phys., 349:703-751, 2017.
[4] Bai, Z.D., Yao, J.F.: On the convergence of the spectral empirical process of Wigner matrices. Bernoulli 11(6), 1059-1092 (2005).
[5] Bourgade, P., Mody, K. Gaussian fluctuations of the determinant of Wigner Matrices. arXiv:1811.06815 (2018).
[6] Chhaibi, R., Madaule, T., Najnudel, J.: On the maximum of the CβE field. Duke Math. J. 167, no. 12 (2018), 2243-2345.
[7] Claeys, T., Fals, B., Lambert, G., Webb, C. How much can the eigenvalues of a random Hermitian matrix fluctuate?. arXiv:1906.01561, (2019).
[8] Cook, N., Zeitouni, O. Maximum of the characteristic polynomial for a random permutation matrix. arXiv:1806.07549 (2018).
[9] Diaconis, P., Evans, S.: Linear functionals of eigenvalues of random matrices. Transactions of the American Mathematical Society, 353(7), 2615-2633 (2001).
[10] Dumitriu, I., Edelman, A. Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models. Journal of Mathematical Physics, 47(6), 063302 (2006).
[11] Erdős, L., Knowles, A., and Yau, H.-T.: Averaging fluctuations in resolvents of random band matrices. Annales Henri Poincaré. Vol. 14. No. 8. Springer Basel, 2013.
[12] Erdős, L., Knowles, A., Yau, H. T., Yin, J. Spectral statistics of Erdős-Rényi graphs I: local semicircle law. The Annals of Probability, 41(3B), 2279-2375 (2013).
[13] Erdős, L., Knowles, A., Yau, H. T., Yin, J. The local semicircle law for a general class of random matrices. Electron. J. Probab 18, 1-58 (2013).
[14] Erdős, L., Yan, H. T., Yin, J. Rigidity of eigenvalues of generalized Wigner matrices. Advances in Mathematics, 229(3), 1435-1515. (2012).
[15] Fyodorov, Y. V., Haar G. A., Keating, J. P. Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta function. Phys. Rev. Lett., 108:170601, Apr 2012.
[16] Fyodorov, Y. V., Keating, J. P. Freezing transitions and extreme values: random matrix theory, and disordered landscapes. Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372(2007):20120503, 32, 2014.
[17] Fyodorov, Y. V., Khoruzhenko, B. A., and Simm, N. J. Fractional Brownian motion with Hurst index \( H = 0 \) and the Gaussian unitary ensemble. The Annals of Probability, 44(4), 2980-3031 (2016).
[18] Fyodorov, Y. V., Simm, N. J.: On the distribution of the maximum value of the characteristic polynomial of GUE random matrices. Nonlinearity 29 (Sep. 2016), p. 2837.
[19] F. Götze and A. Tikhomirov.: On the Rate of Convergence to the Semi-circular Law. Progress in Probability, 139. (2013)
[20] Gustavsson, J.: Gaussian fluctuations of eigenvalues in the GUE. Ann I H Poincare-PR. 41(2),151-178 (2005).
[21] He, Y., Knowles, A.: Mesoscopic linear statistics of Wigner matrices. Annals of Applied Probability, 2017.
[22] He, Y., Knowles, A., Rosenthal, R.: Isotropic self-consistent equations for mean-field random matrices. Probability Theory and Related Fields, 171 (2018), 203-249.
[23] Johansson, K.: On fluctuations of eigenvalues of random Hermitian matrices. Duke mathematical journal, 91(1), 151-204 (1998).
[24] Knowles, A., Yin, J.: The Isotropic Semicircle Law and Deformation of Wigner Matrices. Communications on Pure and Applied Mathematics, 2013 (66), 1663-1749.
[25] Lambert, G. The law of large numbers for the maximum of the characteristic polynomial of the Ginibre ensemble. arXiv:1902.01983 (2019).
[26] Lambert, G., Ostrovsky, D., Simm, N. Subcritical Multiplicative Chaos for Regularized Counting Statistics from Random Matrix Theory. Commun. Math. Phys. 360, 154 (2018)
[27] Lambert, G., Paquette, E. The law of large numbers for the maximum of almost Gaussian log-correlated fields coming from random matrices. Probability Theory and Related Fields, 173(1-2), 157-209 (2019).
[28] Landon, B., Sasoe, P. Applications of mesoscopic CLTS in random matrix theory. arXiv:1811.05915 (2018).
[29] O’Rourke, S. Gaussian fluctuations of eigenvalues in Wigner random matrices. Journal of Statistical Physics, 138(6), 1045-1066 (2010).
[30] Pan, G. M., Gao J. T.: Asymptotic theory for sample covariance matrix under cross-sectional dependence. Preprint (2012).
[31] Paquette, E., Zeitouni, O. The maximum of the CUE field. International Mathematics Research Notices, 2018(16), 5928-5119 (2017).
[32] Rains, E. M. High powers of random elements of compact Lie groups. Probability theory and related fields, 107(2), 219-241 (1997).
[33] Stephens, M. A.: Asymptotic results for goodness-of-fit statistics with unknown parameters. The Annals of Statistics, 1976: 357-369.
[34] Wang, L. L. , and Debashis, P. Limiting spectral distribution of renormalized separable sample covariance matrices when $p/n \to 0$. Journal of Multivariate Analysis 126 : 25-52 (2014).
[35] Wigner, E. P. On the distribution of the roots of certain symmetric matrices. Ann. Math, 67(2), 325-327, (1958).