UC Davis
UC Davis Previously Published Works

Title
A Note on Cumulant Technique in Random Matrix Theory

Permalink

Journal

ISSN

Authors
Soshnikov, Alexander
Wu, Chutong

Publication Date
2023

DOI

Article

A Note on Cumulant Technique in Random Matrix Theory

Alexander Soshnikov *,† and Chutong Wu †

Department of Mathematics, University of California at Davis, Davis, CA 95616, USA; ctwu@ucdavis.edu
* Correspondence: soshniko@math.ucdavis.edu
† These authors contributed equally to this work.

Abstract: We discuss the cumulant approach to spectral properties of large random matrices. In particular, we study in detail the joint cumulants of high traces of large unitary random matrices and prove Gaussian fluctuation for pair-counting statistics with non-smooth test functions.

1. Introduction

Random Matrix Theory has its origins in the works of statisticians in the 1920s and nuclear physicists in the 1950s. In the pioneering papers [1–3], Eugene Wigner introduced an ensemble of random matrices that now have his name and computed the limiting spectral distribution. The main ingredient of the proof was the method of moments that allowed Wigner to study asymptotics of the traces of powers of a random symmetric (Hermitian) matrix with independent identically distributed (i.i.d.) components. Since then, the method of moments has been successfully used to study spectral properties of large random matrices on many occasions. It works particularly well when a random matrix has many independent components. We refer the reader to [4–13] and references therein.

The purpose of this paper is to discuss several applications of the cumulant technique in Random Matrix Theory. The cumulant approach is especially useful if point correlation functions are given by the determinantal or Pfaffian formulas.

The paper is organized as follows. In the Methods section, we give a brief overview of the known results. Several novel results related to the joint cumulants of traces of high powers and pair counting statistics of eigenvalues of a large random unitary matrix are formulated and proved in the Results section. The Discussion section is devoted to brief comments on the future lines of research.

Throughout the paper, the notation $a_N = O(b_N)$ means that the ratio a_N/b_N is bounded from above in absolute value. The notation $a_N = o(b_N)$ means that $a_n/b_N \to 0$ as $N \to \infty$. The notation $a_N = \Omega(b_N)$ means that the ratio a_N/b_N is bounded from above in absolute value by a power of $\log N$. Finally, we sometimes use the notation $a \vee b$ for the maximum of a and b.

2. Methods

2.1. Determinantal Point Process and Cluster Functions

The ideas of the cumulant approach in Random Matrix Theory go back at least to the 1995 paper [14] by Costin and Lebowitz, which studied a so-called sine random point process, namely, a determinantal random point process with the correlation kernel

$$K(x,y) = \frac{\sin(\pi(x-y))}{\pi(x-y)}. \tag{1}$$

In other words, the k-point correlation functions of the random point process are given by

$$\rho_k(x_1,\ldots,x_k) = \det[K(x_i,x_j)]_{1\le i,j\le k}, \quad k \ge 1. \tag{2}$$

We refer the reader to [15,16] for an introduction to determinantal random point processes. The sine random point process appears as a scaling limit of many ensembles of random matrices, including the Circular Unitary Ensemble defined below in (11).

The main result of [14] states that $\#[0, L]$, the number of particles in an interval $[0, L]$, has Gaussian fluctuation in the limit $L \to \infty$ with logarithmically growing variance. Namely,

$$\frac{\#[0, L] - L}{\sqrt{\log L}} \xrightarrow{\mathcal{D}} N(0, 1/\pi^2). \tag{3}$$

To study the limiting distribution of the counting random variable $\#[0, L]$, Costin and Lebowitz suggested using the so-called cluster (Ursell) k-point functions. For $k = 1, 2, 3$, the cluster functions are given by

$$
\begin{aligned}
r_1(x_1) &= \rho_1(x_1), \\
r_2(x_1, x_2) &= \rho_2(x_1, x_2) - \rho_1(x_1)\rho_1(x_2), \\
r_3(x_1, x_2, x_3) &= \rho_3(x_1, x_2, x_3) - \rho_2(x_1, x_2)\rho_1(x_3) - \rho_2(x_1, x_3)\rho_1(x_2) - \\
&\quad - \rho_2(x_2, x_3)\rho_1(x_1) + 2\rho_1(x_1)\rho_1(x_2)\rho_1(x_3).
\end{aligned}
$$

For arbitrary $k \geq 1$, one has

$$r_k(x_1, \ldots, x_k) = \sum_{\pi} (-1)^{|\pi|-1}(|\pi| - 1)! \prod_{j=1}^{|\pi|} \rho_{|B_j|}(x_i : i \in B_j), \tag{4}$$

where the sum at the r.h.s. of (4) is over all partitions π of the set $\{1, 2, \ldots, k\}$ into blocks $B_1, \ldots, B_m$, $|\pi| = m$ stands for the number of blocks in a partition π, and $|B_i|$ denotes the cardinality of a block B_i.

For determinantal random point processes, (2) and (4) imply

$$
\begin{aligned}
r_k(x_1, \ldots, x_k) &= \frac{(-1)^{k-1}}{k} \sum_{\sigma \in S_k} K(x_{\sigma(1)}, x_{\sigma(2)}) K(x_{\sigma(2)}, x_{\sigma(3)}) \ldots K(x_{\sigma(k)}, x_{\sigma(1)}) \\
&= (-1)^{k-1}\big(K(x_1, x_2)K(x_2, x_3) \ldots K(x_k, x_1) + \ldots\big),
\end{aligned}
$$

where the sum in the first line is over all permutations in the symmetric group S_k, and the sum in the second line is over all $(k-1)!$ cyclic permutations in S_k, with the first of such cyclic permutations being $1 \to 2 \to 3 \to \ldots \to k \to 1$.

Cluster point functions are closely related to the cumulants of $\#[0, L]$. Denote by $\mathcal{I}_k(L)$ the integral of the k-point cluster function over the k-dimensional cube $[0, L]^k$:

$$\mathcal{I}_k(L) = \int_{[0,L]^k} r_k(x_1, \ldots, x_k) dx_1 \ldots dx_k, \quad k \geq 1. \tag{5}$$

Furthermore, denote by $\kappa_j(L)$, $j = 1, 2, 3, \ldots$, the cumulants of the counting random variable $\#[0, L]$. Recall that the moment and cumulant generating functions are related by

$$\sum_{j=1}^{\infty} \frac{\kappa_j}{j!} z^j = \log\Big(\sum_{j=0}^{\infty} \frac{m_j}{j!} z^j\Big). \tag{6}$$

It follows from (4) and (6) that the j-th cumulant of $\#[0, L]$ can be written as a linear combination of the integrals $\mathcal{I}_k(L)$, $1 \leq k \leq j$. Namely, the following relation holds for the generating functions:

$$\sum_{j=1}^{\infty} \frac{\kappa_j(L)}{j!} z^j = \sum_{j=1}^{\infty} \frac{\mathcal{I}_j(L)}{j!} (e^z - 1)^j. \tag{7}$$

To prove the Central Limit Theorem for the normalized random variable $\frac{\#[0,L] - \mathbb{E}\#[0,L]}{\sqrt{\text{Var}\#[0,L]}}$, it is sufficient to show that

$$\text{Var}\#[0,L] \to \infty \text{ as } L \to \infty \text{ and } \kappa_j(L) = o((\text{Var}\#[0,L])^{j/2}) \text{ for } j > 2 \qquad (8)$$

since (8) would imply that all cumulants starting with the third one of the normalized counting random variable $\frac{\#[0,L] - \mathbb{E}\#[0,L]}{\sqrt{\text{Var}\#[0,L]}}$ go to zero as $L \to \infty$.

It is not hard to see that if a random point process is such that $\text{Var}\#[0,L]$ grows linearly in L and the cluster function integrals satisfy

$$\mathcal{I}_j(L) = o(L^{j/2}), \quad j > 2, \qquad (9)$$

a routine application of (7) and (8) finishes the proof of the Central Limit Theorem for the counting function. However, the sine random point process exhibits a more delicate behavior—the variance of the number of particles grows only logarithmically, namely:

$$\text{Var}\#[0,L] = \frac{1}{\pi^2} \log L + O(1). \qquad (10)$$

Thus, a more subtle analysis of the asymptotics of $\mathcal{I}_j(L)$ is required for $j > 2$. The proof ([14]) of the Central Limit Theorem (3) follows from the bounds

$$\int_0^L \int_0^L \frac{\sin(\pi(x_1 - x_2))}{\pi(x_1 - x_2)} \frac{\sin(\pi(x_2 - x_1))}{\pi(x_2 - x_1)} \, dx_1 dx_2 = L - \frac{1}{\pi^2} \log L + O(1),$$

$$\int_{[0,L]^k} \frac{\sin(\pi(x_1 - x_2))}{\pi(x_1 - x_2)} \frac{\sin(\pi(x_2 - x_3))}{\pi(x_2 - x_3)} \cdots \frac{\sin(\pi(x_k - x_1))}{\pi(x_k - x_1)} \, dx_1 \cdots dx_k = L + O(\log L).$$

Remarkably, the result can be generalized to the case of any determinantal random point field on a locally compact Hausdorff phase space E, equipped with a σ-finite measure μ on the Borel σ-algebra provided the correlation kernel is Hermitian and locally trace class and the variance goes to infinity. Namely, the following holds:

Theorem 1 ([17,18]). *Let $(X, \mathcal{F}, \mathbb{P}_n)$ be a family of determinantal random point fields with hermitian locally trace class correlation kernels K_n and I_n be a family of Borel subsets of E with compact closure. Then, if $\text{Var}_n\#(I_n) \to \infty$ as $n \to \infty$, the normalized counting random variable $\frac{\#(I_n) - \mathbb{E}\#(I_n)}{\sqrt{\text{Var}\#(I_n)}}$ converges in distribution to a standard normal $N(0,1)$.*

Remark 1. *The multivariate version of the theorem also holds (see [17,18]). It is worth noting that the univariate result allows a very nice probabilistic interpretation ([19]). Namely, one can show that the counting random variable $\#(I)$ is equal in distribution to a sum of independent non-identically distributed Bernoulli random variables with the probabilities of success given by the eigenvalues of the integral operator $K_I f(x) = \int_I K(x,y) f(y) \mu(dy)$ on $L^2(I, \mu)$. However, we are not aware of a simple probabilistic interpretation of the multivariate CLT result.*

Analogous results hold for linear statistics $L_f = \sum_i f(x_i)$, where f is a bounded measurable test function, e.g., with a compact support, and the summation is over all points of a random point field.

Theorem 2 ([17,18]). *Let $(X, \mathcal{F}, \mathbb{P}_n)$ be a family of determinantal random point fields with hermitian locally trace class correlation kernels K_n and f_n be a bounded, measurable, compactly supported function on the phase space E such that*

$$\text{Var} L_{f_n} \to \infty, \quad \sup |f_n(x)| = o((\text{Var} L_{f_n})^{\varepsilon}), \quad \mathbb{E} L_{|f|_n} = O((\text{Var} L_{f_n})^{\delta}),$$

for any $\varepsilon > 0$ and some $\delta > 0$ as $n \to \infty$. Then, the normalized linear statistic $\frac{L_{f_n} - \mathbb{E}L_{f_n}}{\sqrt{VarL_{f_n}}}$ converges in distribution to $N(0, 1)$ as $n \to \infty$.

For additional results in this direction, we refer the reader to [20–22]. Moreover, the connection between integrals of cluster point functions and cumulants can be exploited to study significantly more challenging problems regarding empirical spectral distribution of spacings and extreme spacings in determinantal random point processes (see, e.g., [23,24]).

2.2. Linear Statistics in Classical Compact Groups

For many ensembles of large random matrices, the variance of a linear statistic either stays finite or grows slower than any power of the dimension, provided the test function is sufficiently smooth. Therefore, Theorem 2 is not typically applicable even if the point correlation functions are determinantal. It is instructive to consider classical compact groups. In this paper, we focus our attention on random unitary matrices. However, many results can be extended to the orthogonal and symplectic matrices without much difficulty.

Let V be a unitary matrix chosen at random with respect to the Haar measure on the unitary group $U(N)$. We are interested in studying statistical properties of the eigenvalues of V, which we denote by $e^{i\theta_j}$, $1 \le j \le N$, $-\pi \le \theta_1, \ldots, \theta_N < \pi$. The joint probability density of the eigenvalues is given by the formula ([25]):

$$p_N(\theta_1, \ldots, \theta_N) = \frac{1}{(2\pi)^N N!} \prod_{1 \le j < k \le N} \left| e^{i\theta_j} - e^{i\theta_k} \right|^2. \tag{11}$$

The joint distribution of the eigenvalues is known as the Weyl measure, and the ensemble is known in Random Matrices as the Circular Unitary Ensemble ([26–28]). Remarkably, if a test function f on the unit circle is sufficiently smooth, the variance of the linear statistics $\sum_{j=1}^{N} f(\theta_j)$ stays finite as $N \to \infty$. Moreover,

Theorem 3 ([29,30]). *Let f be a real-valued function on the unit circle satisfying*

$$\sum_{-\infty}^{\infty} |\hat{f}_k|^2 |k| < \infty, \tag{12}$$

where

$$\hat{f}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx, \quad k \in \mathbb{Z} \tag{13}$$

are the Fourier coefficients of f. Then,

$$\sum_{j=1}^{N} f(\theta_j) - \hat{f}_0 N \xrightarrow{\mathcal{D}} N(0, \sum_{-\infty}^{\infty} |\hat{f}_k|^2 |k|). \tag{14}$$

Remark 2. *One can show that the result of Theorem 3 follows from the Strong Szego Theorem for Toeplitz determinants ([30]).*

Remark 3. *It follows from Theorem 3 that the real and imaginary parts of the traces of the powers of a random unitary matrix, $\mathrm{Tr}\, V^k = \sum_{j=1}^{N} e^{ik\theta_j}$, $k \ge 1$, converge to independent normal random variables with mean zero and variance $k/2$ as $N \to \infty$. In ([31]), Johansson proved that the rate of convergence to normal law is $O(N^{-\delta N})$ for some $\delta > 0$. Recently, the results have been further improved in [32,33].*

Many joint moments of the traces of powers of a random unitary matrix V coincide with the corresponding joint moments of the limiting normal random variables. Namely,

let Z_j, $j \geq 1$, be a sequence of independent standard complex normals. Then, for any $k \geq 1$ and non-negative integers $a_j, b_j, 1 \leq j \leq k$ satisfying

$$N \geq \left(\sum_{j=1}^{k} j a_j \right) \vee \left(\sum_{j=1}^{k} j b_j \right), \tag{15}$$

one has ([29]):

$$\mathbb{E}\left[\prod_{j=1}^{k} \mathrm{Tr}(V^j)^{a_j} \overline{\mathrm{Tr}(V^j)}^{b_j} \right] = \delta_{ab} \prod_{j=1}^{k} j^{a_j} a_j! = \mathbb{E}\left[\prod_{j=1}^{k} (\sqrt{j} Z_j)^{a_j} \overline{(\sqrt{j} Z_j)}^{b_j} \right], \tag{16}$$

where the delta symbol δ_{ab} is one if $a = (a_1, \ldots, a_k)$ coincides with $b = (b_1, \ldots, b_k)$ and is zero otherwise. For a beautiful survey of this and related results for random matrices from classical compact groups, we refer the reader to [34].

To better understand (16) and related phenomena, one can study the joint cumulants of the traces of powers of a random matrix. Let $\{X_{\alpha \in \mathcal{A}}\}$ be a family of random variables. The joint cumulants are defined as (see, e.g., [35]):

$$\kappa_{i_1, \ldots, i_n} := \kappa(X_{i_1}, \ldots, X_{i_n}) = \sum_{\pi} (|\pi| - 1)! (-1)^{|\pi|-1} \prod_{B \in \pi} \mathbb{E}\left(\prod_{i \in B} X_i \right), \tag{17}$$

where the sum is over all partitions π of the set $\{i_1, \ldots, i_n\}$, B goes over the list of all blocks of the partition π, and $|\pi|$ denotes the number of blocks in the partition.

We recall that the joint moments are expressed in terms of the cumulants as

$$\mathbb{E}(X_1, \ldots, X_n) := \mathbb{E} \prod_{1 \leq i \leq n} X_i = \sum_{\pi} \prod_{B \in \pi} \kappa(X_i : i \in B). \tag{18}$$

We will use the notation $T_{N,k} := \mathrm{Tr}(V^k)$ for the traces of powers of V and denote by $\kappa_p^{(N)}(k_1, \ldots, k_p)$ the joint cumulants of $T_{N,k_1}, \ldots, T_{N,k_p}$, i.e.,

$$\kappa_p^{(N)}(k_1, \ldots, k_p) := \kappa(T_{N,k_1}, \ldots, T_{N,k_p}). \tag{19}$$

For a determinantal random point process with a correlation kernel $K(x, y)$, the cumulants of a linear statistic $L_f = \sum_i f(x_i)$ can be computed as ([18]):

$$\kappa_p(L_f) = \sum_{m=1}^{p} \frac{(-1)^m}{m} \sum_{\substack{(p_1, \ldots, p_m): \\ p_1 + \ldots + p_m = p, \ p_1, \ldots, p_m \geq 1}} \frac{p!}{p_1! \cdots p_m!} \times \tag{20}$$

$$\int f^{p_1}(x_1) K(x_1, x_2) f^{p_2}(x_2) K(x_2, x_3) \cdots f^{p_m}(x_m) K(x_m, x_1) dx_1 \cdots dx_m. \tag{21}$$

For the CUE, the point correlation functions are given by the determinantal Formula (2) with the correlation kernel

$$K_N(x, y) = \frac{1}{2\pi} \sum_{j=0}^{N-1} e^{ij(x-y)}.$$

This allows one to obtain the following formula in the CUE case, provided the test function f is sufficiently smooth ([36], see also [37–39]):

$$
\kappa_p(L_f) = \sum_{\substack{k_1,\dots,k_p\in\mathbb{Z}\\k_1+\dots+k_p=0}} \prod_{j=1}^{p}\hat{f}_{k_j} \sum_{m=1}^{p} \frac{(-1)^{m-1}}{m} \sum_{\substack{(p_1,\dots,p_m):\\p_1+\dots+p_m=p,\ p_1,\dots,p_m\geq 1}} \frac{p!}{p_1!\cdots p_m!}\times
$$

$$
\max\Big(0, N - \max\big(0, \sum_{i=1}^{p_1}k_i,\dots,\sum_{i=1}^{p_1+\dots+p_{m-1}}k_i\big) - \max\big(0, -\sum_{i=1}^{p_1}k_i,\dots,-\sum_{i=1}^{p_1+\dots+p_{m-1}}k_i\big)\Big),
$$

that can be rewritten as

$$
\kappa_p(L_f) = \sum_{\substack{k_1,\dots,k_p\in\mathbb{Z}\\k_1+\dots+k_p=0}} \prod_{j=1}^{p}\hat{f}_{k_j} \sum_{m=1}^{p} \frac{(-1)^{m}}{m} \sum_{\substack{(p_1,\dots,p_m):\\p_1+\dots+p_m=p,\ p_1,\dots,p_m\geq 1}} \frac{1}{p_1!\cdots p_m!}\times
$$

$$
\sum_{\sigma\in S_p} J_N(p_1,\dots,p_m,k_{\sigma(1)},\dots,k_{\sigma(p)}), \tag{22}
$$

where J_N is a piece-wise linear function defined by the following formula

$$
J_N(p_1,\dots,p_m;k_1,\dots,k_p) :=
$$
$$
\min\Big(N,\ \max\big(0, \sum_{i=1}^{p_1}k_i, \sum_{i=1}^{p_1+p_2}k_i,\dots, \sum_{i=1}^{p_1+\dots+p_{m-1}}k_i\big) + \max\big(0, -\sum_{i=1}^{p_1}k_i,\dots, -\sum_{i=1}^{p_1+\dots+p_{m-1}}k_i\big)\Big) \tag{23}
$$

for positive integers $p_1,\dots,p_m \geq 1$, $p_1 + \dots + p_m = p$, and integers $k_1,\dots,k_p$, satisfying $\sum_{i=1}^{p} k_i = 0$. Since the joint cumulants are symmetric and multilinear, it follows from (22) and (23) that $\kappa_p^{(N)}(k_1,\dots,k_p)$ can be computed as

Lemma 1 ([40]).

    (i) If $p > 1$ and either $k_1 + \dots + k_p \neq 0$ or $\prod_i^{p} k_i = 0$, or both, then

$$
\kappa_p^{(N)}(k_1,\dots,k_p) = 0. \tag{24}
$$

    (ii) If $p > 1$, $\prod_i^{p} k_i \neq 0$, and $\sum_{i=1}^{p} k_i = 0$, then

$$
\kappa_p^{(N)}(k_1,\dots,k_p) = \sum_{m=1}^{p} \frac{(-1)^{m}}{m} \sum_{\substack{(p_1,\dots,p_m):\\p_1+\dots+p_m=p,\ p_1,\dots,p_m\geq 1}} \frac{1}{p_1!\cdots p_m!}\times \tag{25}
$$

$$
\sum_{\sigma\in S_p} J_N(p_1,\dots,p_m;k_{\sigma(1)},\dots,k_{\sigma(p)}).
$$

Remark 4. *Alternatively, one can rewrite the last formula as*

$$
\kappa_p^{(N)}(k_1,\dots,k_p) = \sum_{m=1}^{p} \frac{(-1)^{m}}{m} \sum_{\substack{(C_1,\dots,C_m):\\C_1\cup\dots\cup C_m=\{1,\dots,p\}}} \tag{26}
$$

$$
\min\Big(N,\ \max\big(0, \sum_{i\in C_1}k_i,\dots, \sum_{i\in C_1\cup\dots\cup C_{m-1}}k_i\big) + \max\big(0, -\sum_{i\in C_1}k_i,\dots, -\sum_{i\in C_1\cup\dots\cup C_{m-1}}k_i\big)\Big),
$$

where the summation is over all ordered collections of non-empty disjoint subsets $C_1,\dots,C_m$, $m \geq 1$, of the set $\{1,\dots,p\}$ such that $\cup_j C_j = \{1,\dots,p\}$.

It should be emphasized that the exact combinatorial structure manifested in Lemma 1 is specific to the classical compact groups and the sine determinantal point process.

To finish the proof of (14) and (16), using the cumulants, one observes that for

$$\sum_{i=1}^{p} |k_i| \le 2N \; and \; \sum_{i=1}^{p} k_i = 0 \tag{27}$$

the expression for $J_N(p_1, \ldots, p_m; k_1, \ldots, k_p)$ simplifies, namely:

$$J_N(p_1, \ldots, p_m; k_1, \ldots, k_p) =$$
$$\max\left(0, \sum_{i=1}^{p_1} k_i, \sum_{i=1}^{p_1+p_2} k_i, \ldots, \sum_{i=1}^{p_1+\ldots+p_{m-1}} k_i\right) + \max\left(0, \sum_{i=1}^{p_1}(-k_i), \ldots, \sum_{i=1}^{p_1+\ldots+p_{m-1}}(-k_i)\right). \tag{28}$$

One then uses a combinatorial lemma from [36] to show that all joint cumulants $\kappa_p^{(N)}(k_1, \ldots, k_p)$ of order three and higher ($p > 2$) are identically zero, provided $\sum_{i=1}^{p} |k_i| \le 2N$:

Lemma 2 ([36]). *Let $k_1, \ldots, k_p$, $p \ge 2$ be arbitrary real numbers such that $\sum_i k_i = 0$. Then, the sum*

$$\sum_{\sigma \in S_p} \sum_{m=1}^{p} \frac{(-1)^m}{m} \sum_{\substack{(p_1,\ldots,p_m): \\ p_1+\ldots+p_m=p, \; p_1,\ldots,p_m \ge 1}} \frac{1}{p_1! \cdots p_m!} \max\left(0, \sum_{i=1}^{p_1} k_{\sigma(i)}, \ldots, \sum_{i=1}^{p_1+\ldots+p_{m-1}} k_{\sigma(i)}\right)$$

equals $|k_1| = |k_2|$ if $p = 2$ and zero if $p > 2$.

Remark 5. *The lemma is related to the combinatorial proof of the Strong Szego Theorem. We refer the reader to [41–44] for recent applications of the cumulant method for determinantal processes to study linear statistics of Hermitian unitary invariant random matrices and free fermions processes.*

2.3. Multivariate Linear Statistics and Number Theory Connections

This subsection is devoted to the discussion of multivariate linear statistics of the form

$$S_N(f) = \sum_{1 \le j_1, j_2, \ldots, j_{n+1} \le N} f(L_N(\theta_{j_2} - \theta_{j_1}), \ldots, L_N(\theta_{j_{n+1}} - \theta_{j_1})), \tag{29}$$

where $\theta = (\theta_1, \ldots, \theta_N) \in \mathbb{T}^N$ comes from the CUE (or, in general, the Circular Beta Ensemble with arbitrary $\beta > 0$), the scaling factor L_N satisfies $1 \le L_N \le N$, and f is a sufficiently smooth test function defined on the unit circle (when $L_N = 1$) or the real line (when $L_N \to \infty$.)

The study of such multivariate linear statistics in [40,45–47] was motivated by connections between Random Matrices and Number Theory (see, e.g., [48–51] and references therein.) The local case ($L_N = N$) is of a particular interest for the CUE since the multivariate linear statistic

$$S_N(f) = \sum_{1 \le j_1, j_2, \ldots, j_{n+1} \le N} f(N(\theta_{j_2} - \theta_{j_1})_c, \ldots, N(\theta_{j_{n+1}} - \theta_{j_1})_c), \tag{30}$$

where $f \in C_c^\infty(\mathbb{R})$ and $(\theta - \phi)_c$ is the phase difference on the unit circle (i.e., the difference modulo 2π) and corresponds to multivariate linear statistics of the (rescaled) zeros of the Riemann zeta function studied by Montgomery [52,53], Hejhal [54], and Rudnick and Sarnak [55] (see, e.g., [40]). We also point out a recent related preprint [56], where multivariate linear statistics have been studied for the determinantal random process with the projection correlation kernel on the unit sphere S^d, $d > 1$.

To study the limiting distribution of (30), we write using the Fourier transform

$$S_N(f) = \frac{1}{(2\pi)^{n/2}N^n} \sum_{k \in \mathbb{Z}^n} \hat{f}\left(\frac{k_1}{N}, \ldots, \frac{k_n}{N}\right) \prod_{j=1}^{n+1} T_{N,k_j} \qquad (31)$$

where $k_{n+1} = -\sum_{j=1}^{n} k_j$, and $\hat{f}(t) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x)e^{-it \cdot x}dx$, $t = (t_1, \ldots, t_n)$.

The proof of the Gaussian fluctuation for $(S_N(f) - \mathbb{E}S_N(f))/\sqrt{N}$ presented in [40] relies on a detailed study of the joint cumulants $\kappa_p^{(N)}(k_1, \ldots, k_p)$, $p > 1$, for the arguments $k_i = O(N)$, $1 \le i \le p$. It is combinatorial in nature. One of the key ingredients of the proof is the fact that joint cumulants scale with N. Namely,

$$c_p(t_1, \ldots, t_p) := \frac{1}{N}\kappa_p^{(N)}(t_1 N, \ldots, t_p N) \qquad (32)$$

does not depend on N. Moreover, it is a piece-wise linear, bounded function of $t_1, \ldots, t_p$ on the hyperplane $t_1 + \ldots + t_p = 0$ (it is identically zero when $t_1 + \ldots + t_p \ne 0$.)

Lemma 3 ([40]). *Rescaled joint cumulants c_p defined in (32) can be written for $p > 1$ and $\sum_{i=1}^{p} t_i = 0$ as*

$$c_p(t_1, \ldots, t_p) = \sum_{m=1}^{p} \frac{(-1)^m}{m} \sum_{\substack{(p_1, \ldots, p_m): \\ p_1 + \ldots + p_m = p, \, p_1, \ldots, p_m \ge 1}} \frac{1}{p_1! \cdots p_m!} \sum_{\sigma \in S_p} j(p_1, \ldots, p_m; t_{\sigma(1)}, \ldots, t_{\sigma(p)}), \qquad (33)$$

where $j(p_1, \ldots, p_m; t_1, \ldots, t_p)$ is defined for positive integers $p_1, \ldots, p_m$, $p_1 + \ldots + p_m = p$, and real numbers $t_1, \ldots, t_p$, as

$$j(p_1, \ldots, p_m; t_1, \ldots, t_p) :=$$
$$\min\left(1, \max\left(0, \sum_{i=1}^{p_1} t_i, \sum_{i=1}^{p_1+p_2} t_i, \ldots, \sum_{i=1}^{p_1+\ldots+p_{m-1}} t_i\right) + \max\left(0, \sum_{i=1}^{p_1}(-t_i), \ldots, \sum_{i=1}^{p_1+\ldots+p_{m-1}}(-t_i)\right)\right). \qquad (34)$$

Moreover, the following holds:

(i) $c_p(t_1, \ldots, t_p)$, $p > 1$, *is a bounded symmetric piece-wise linear function on* $\sum_{i=1}^{p} t_i = 0$.
(ii) $c_p(t_1, \ldots, t_p) = 0$ *if* $p > 1$ *and* $\sum_{i=1}^{p} t_i \ne 0$.
(iii) $c_1(0) = 1$ *and* $c_1(t) = 0$ *for* $t \ne 0$.

We finish this section with a discussion of the multivariate linear statistics for non-smooth functions. For simplicity, we assume that $n = 1$ and restrict our attention to the global regime ($L_N = 1$). Then,

$$S_N(f) = \sum_{i,j=1}^{N} f(\theta_i - \theta_j). \qquad (35)$$

In [45], we proved that if f is an even real-valued function on the unit circle such that both f' belongs to $L^2(\mathbb{T})$, then

$$S_N(f) - \mathbb{E}S_N(f) \xrightarrow{\mathcal{D}} 2 \sum_{m=1}^{\infty} \hat{f}_m m(\varphi_m - 1), \qquad (36)$$

where φ_m are i.i.d. exponential random variables with $\mathbb{E}(\varphi_m) = 1$. We note that the condition $\sum_k |\hat{f}_k|^2 k^2 < \infty$ is necessary and sufficient for (36) to hold. If the series $\sum_k |\hat{f}_k|^2 k^2$ is slowly diverging so that the sequence of partial sums $V_N = \sum_{|k| \le N} |\hat{f}_k|^2 k^2$ is slowly varying in the sense of Karamata ([57]), then the Central Limit Theorem holds ([46]):

$$\frac{S_N(F) - \mathbb{E}S_N(f)}{\sqrt{V_N}} \xrightarrow{\mathcal{D}} N(0, 2).$$

The case of the linearly growing variance will be studied in the next section, where we will consider a class of even test functions f for which $\hat{f}_k k \to C \neq 0$ as $k \to \infty$. The values of the cumulant function $\kappa_p^{(N)}(k_1, \dots, k_p)$ for $k_i = O(N)$, $1 \le i \le p$ play an important role in the analysis.

3. Results

We study the Circular Unitary Ensemble of random matrices (11). In particular, we are interested in the joint distribution of the traces of high powers of a unitary matrix V. As before, we denote by $\kappa_p^{(N)}(k_1, \dots, k_p)$ the joint cumulants of $\mathrm{Tr}\, V^{k_1}, \dots, \mathrm{Tr}\, V^{k_p}$, $1 \le i \le p$. In what follows, $k_i = O(N)$, $1 \le i \le p$, and we are looking for scenarios when the joint cumulants vanish.

We start by considering the joint cumulants of $\mathrm{Tr}\, V^k$ and $\mathrm{Tr}\, V^{-k}$ for $k \ge N$. While the distribution in this case is well understood (see, e.g., [58]), it is instructive to consider the cumulant approach that will be further expanded later in this section. We are interested in the values of the joint cumulants $\kappa_p^{(N)}(k_1, \dots, k_p)$, $p > 1$, where

$$k_1 = \dots = k_q = k \text{ and } k_{q+1} = \dots = k_p = -k,\ 0 \le q \le p,\ k \ge N.$$

It follows from Lemma 1 that $\kappa_p^{(N)}(k, \dots, k, -k, \dots, -k) = 0$ unless p is even and $p = 2q$. We then have $k_1 = \dots k_q = k$ and $k_{q+1} = \dots = k_{2q} = -k$. We will use the Formula (26). Since $k \ge N$, one has

$$\min\left(N,\ \max\left(0, \sum_{i \in B_1} k_i, \dots, \sum_{i \in B_1 \cup \dots \cup B_{m-1}} k_i\right) + \max\left(0, -\sum_{i \in B_1} k_i, \dots, -\sum_{i \in B_1 \cup \dots \cup B_{m-1}} k_i\right)\right) = N \quad (37)$$

unless each subset B_j has an equal number of elements equal to k and $-k$, in which case the l.h.s. of (37) is zero. A routine combinatorial analysis produces

$$\kappa_{2q}^{(N)}(k, \dots, k, -k, \dots, -k) = N \sum_{m=1}^{q} \frac{(-1)^{m-1}}{m} \sum_{\substack{(q_1, \dots, q_m): \\ q_1 + \dots + q_m = q,\ q_1, \dots, q_m \ge 1}} \left(\frac{q!}{q_1! \cdots q_m!}\right)^2. \quad (38)$$

Therefore, the cumulant generating function $C_N(x, y)$ can be written as

$$\log\left(\mathbb{E}\exp(x\,\mathrm{Tr}\, V^k + y\,\mathrm{Tr}\, V^{-k})\right) = N \log\left(\sum_{m=0}^{\infty} \frac{1}{(m!)^2}(xy)^m\right) = N \log\left(I_0(\sqrt{4xy})\right), \quad (39)$$

where

$$I_0(u) = \frac{1}{2\pi} \int_0^{2\pi} e^{u \cos\theta} d\theta = \sum_{m=0}^{\infty} \frac{1}{(m!)^2 2^{2m}} u^{2m} \quad (40)$$

is the modified Bessel function of the first kind.

A more elementary way to derive (39) relies on the observation by Rains ([58]) that the kth powers of the eigenvalues of V are i.i.d. random variables uniformly distributed on the unit circle provided $k \ge N$. Thus, $\mathrm{Tr}\, V^k$, for $k \ge N$, is given by the sum of N i.i.d. uniform random variables on the unit circle. The traces of different powers are still correlated for finite N. However, a significant portion of the joint cumulants vanishes.

Proposition 1. *Let s_i, $1 \le i \le l$ be $l > 1$ distinct positive integers and a_i, b_i be non-negative integers such that $a_i + b_i \ge 1$ for all i. Suppose that*

$$|\sum_{i=1}^{l} n_i s_i| \ge N \ \forall n_1, \ldots, n_l \in \mathbb{Z} : -b_i \le n_i \le a_i, \ 1 \le i \le l, \ \sum_{i=1}^{l} |n_i| > 0. \tag{41}$$

Let $p = \sum_1^l (a_i + b_i)$ and $(k_1, \ldots, k_p) \in \mathbb{Z}^p$ be such that the first a_1 coordinates of the vector are equal to s_1, the next b_1 coordinates are equal to $-s_1$, the further next a_2 coordinates are equal to s_2, the following b_2 coordinates are equal to $-s_2$, etc, ..., and the last b_l coordinates are equal to $-s_l$.
Then,

$$\kappa_p^{(N)}(k_1, \ldots, k_p) = 0. \tag{42}$$

Moreover, for $0 \le n_i \le a_i$, $0 \le m_i \le b_i$, $1 \le i \le l$, one has

$$\mathbb{E}\left[\prod_{i=1}^{l}(\operatorname{Tr} V^{s_i})^{n_i}(\operatorname{Tr} V^{-s_i})^{m_i}\right] = \delta_{nm} \prod_{i=1}^{l} \mathbb{E}\left[|\sum_{j=1}^{N} \varphi_j|^{2n_i}\right], \tag{43}$$

where $\{\varphi_i\}$ are i.i.d. uniform random variables on the unit circle, and δ_{nm} is the delta symbol.

Proof of Proposition 1. While the proof does not depend on the value of l, we will assume that $l = 2$ in order to simplify the notations. It follows from (41) that the joint cumulant is zero unless

$$a_i = b_i, \ 1 \le i \le 2, \tag{44}$$

since otherwise $\sum_1^p k_i \ne 0$. Therefore, from now on, we can assume that (44) holds. As before, we will use (26) to compute the joint cumulants. We call a subset B_j balanced if the number of times s_1 appears in it is equal to the number of times $-s_1$ appears in it, and the same holds for s_2 and $-s_2$. It follows from (41) that unless unless each subset B_j is balanced, the l.h.s. of (37) is N. Otherwise, the l.h.s. of (37) is zero. We obtain that for $a_1 = b_1 = c$ and $a_2 = b_2 = d$,

$$\kappa_p^{(N)}(s_1, \ldots, s_1, -s_1, \ldots, -s_1, s_2, \ldots, s_2, -s_2, \ldots, -s_2) = N \sum_{m \ge 1} \frac{(-1)^{m-1}}{m}$$

$$\sum_{\substack{c_1 + \ldots + c_m = c, \ c_1, \ldots c_m \ge 0 \\ d_1 + \ldots + d_m = d, \ d_1, \ldots d_m \ge 0 \\ c_i + d_i \ge 1, \ 1 \le i \le m}} \left(\frac{c!}{c_1! \cdots c_m!}\right)^2 \left(\frac{d!}{d_1! \cdots d_m!}\right)^2, \quad p = 2(c + d).$$

Therefore, the coefficient in front of $(xy)^c(zw)^d$ in the cumulant generating function

$$C_N(x, y, z, w) := \log\left(\mathbb{E}\exp(x \operatorname{Tr} V^{s_1} + y \operatorname{Tr} V^{-s_1} + z \operatorname{Tr} V^{s_2} + w \operatorname{Tr} V^{-s_2})\right)$$

coincides with the corresponding coefficient in the power series expansion of

$$N \log\left(\sum_{m=0}^{\infty} \frac{1}{(m!)^2}(xy)^m \sum_{m=0}^{\infty} \frac{1}{(n!)^2}(zw)^n\right) = N \log\left(I_0(\sqrt{4xy})I_0(\sqrt{4zw})\right)$$

$$= N \log\left(I_0(\sqrt{4xy})\right) + N \log\left(I_0(\sqrt{4zw})\right)$$

$$= \log\left(\mathbb{E}\exp(x \sum_{j=1}^{N} \phi_j + y \sum_{j=1}^{N} \overline{\phi_j} + z \sum_{j=1}^{N} \eta_j + w \sum_{j=1}^{N} \overline{\eta_j})\right),$$

where $\{\phi_i\}$, $\{\eta_i\}$ are i.i.d. uniform random variables on the unit circle, and $\overline{\phi}$ denotes the complex conjugate. The identity (43) follows from the derived identity for the cumulants and the Formula (18), expressing moments in terms of cumulants. $\square$

To illustrate the utility of the cumulant approach, we formulate and prove below the Central Limit Theorem for pair-counting statistics for a class of non-smooth test functions. We consider

$$S_N(f) = \sum_{i,j=1}^{N} f(\theta_i - \theta_j), \tag{45}$$

where $(\theta_1, \dots, \theta_N)$ is a random CUE point configuration, and f is an even real-valued function on the unit circle such that

$$\hat{f}_k k \to C \neq 0 \text{ as } k \to \infty. \tag{46}$$

Remark 6. $f(\theta) = \log|\sin(\theta/2)|$ *satisfies the above test function requirements.*

Theorem 4. *Let f be a real even function on the unit circle, satisfying (46). Then,*

$$\frac{S_N(f) - \mathbb{E}S_N(f)}{\sqrt{N}} \xrightarrow{\mathcal{D}} N(0, \sigma^2(C)), \tag{47}$$

where the limiting variance can be computed as

$$\sigma^2(C) = 4C^2 \left(3 - 2 \int_0^1 \frac{1}{y} \left((1 + 1/y) \log(1 + y) - 1 \right) dy \right) \tag{48}$$

$$-4C^2 \left(2 \int_0^1 \frac{1}{y} (1 + y) \log(1 + y) dy + \int_0^1 \frac{1 - y}{y} \log(1 - y) dy \right). \tag{49}$$

Proof of Theorem 4. We start with the following formula for the variance of $S_N(f)$:

Proposition 2 ([45]). *Let f be a real even function on the unit circle such that $f \in L^2(\mathbb{T})$. Then,*

$$Var[S_N(f)] = 4 \left(\sum_{1 \leq s \leq N-1} s^2 (\hat{f}(s))^2 + N^2 \sum_{N \leq s} (\hat{f}(s))^2 - N \sum_{N \leq s} (\hat{f}(s))^2 \right) \tag{50}$$

$$-4 \left(\sum_{\substack{1 \leq s,t \\ 1 \leq |s-t| \leq N-1 \\ N \leq \max(s,t)}} (N - |s - t|) \hat{f}(s) \hat{f}(t) + \sum_{\substack{1 \leq s,t \leq N-1 \\ N+1 \leq s+t}} ((s + t) - N) \hat{f}(s) \hat{f}(t) \right). \tag{51}$$

Since the test function f satisfies (46), we have

$$\sum_{1 \leq s \leq N-1} s^2 (\hat{f}(s))^2 = N(1 + o(1)), \quad \sum_{N \leq s} (\hat{f}(s))^2 = \frac{1}{N} + O(\frac{1}{N^2}),$$

and the two double sums in (51), up to a negligible error, are equal to the Riemann sums of converging integrals. As a result,

$$Var[S_N(f)] = 4C^2 N \left(2 - \int_{D_1} (1 - |x - y|) \frac{1}{xy} dxdy - \int_{D_2} (x + y - 1) \frac{1}{xy} dxdy \right) (1 + o(1)), \tag{52}$$

where

$$D_1 = \{(x,y) : x, y > 0, |x - y| \leq 1, \max(x,y) \geq 1\}, D_2 = \{(x,y) : 0 < x, y \leq 1, x + y \geq 1\}.$$

The integrals in (52) can be trivially simplified as

$$\int_{D_1}(1-|x-y|)\frac{1}{xy}dxdy = -2 + 2\int_0^1 \frac{1}{y}(1+y)\log(1+y)dy$$

$$+2\int_1^\infty \frac{1}{y}((1+y)\log(1+1/y)-1)dy, \tag{53}$$

$$\int_{D_2}(x+y-1)\frac{1}{xy}dxdy = 1 + \int_0^1 \frac{1-y}{y}\log(1-y)dy,$$

which leads to (48) and (49).

To prove the theorem, we will study asymptotics of the higher moments of $S_N(f)$. We start by truncating f. Namely, we replace it by $\tilde{f}(x) = \sum_{|k|\leq N\log N}\hat{f}_k e^{ikx}$. It follows from (50) and (51) that $\mathrm{Var}[S_N(f)-S_N(\tilde{f})] = \mathrm{Var}[S_N(f-\tilde{f})] = o(N)$. Therefore, without a loss of generality, we can assume that the Fourier coefficients $\hat{f}_k$ are zero for $|k| > N\log N$. Whenever it does not lead to ambiguity, we will use the notation f for the truncated version of a test function. Recall that

$$S_N(f) = \sum_k \hat{f}_k|T_{N,k}|^2 = \sum_k \hat{f}_k T_{N,k}T_{N,-k}, \tag{54}$$

where we use the notation $T_{N,k} = \mathrm{Tr}\, V^k$. Let us fix a positive integer $m > 2$. We have

$$\mathbb{E}(S_N(f) - \mathbb{E}S_N(f))^m = \sum_{k_1,\ldots,k_m\in\mathbb{Z}} \hat{f}_{k_1}\cdots\hat{f}_{k_m}\, \mathbb{E}\left[\prod_{i=1}^m\left(T_{N,k_i}T_{N,-k_i} - \mathbb{E}T_{N,k_i}T_{N,-k_i}\right)\right].$$

Let us denote $k_{m+i} = -k_i$, $1 \leq i \leq m$. Applying

Lemma 4. *9.2 from [45] (see also Lemma 4.1 in [40]), one rewrites*

$$\mathbb{E}\left[\prod_{i=1}^m\left(T_{N,k_i}T_{N,-k_i} - \mathbb{E}T_{N,k_i}T_{N,-k_i}\right)\right] = \mathbb{E}\left[\prod_{i=1}^m\left(T_{N,k_i}T_{N,k_{m+i}} - \mathbb{E}T_{N,k_i}T_{N,k_{m+i}}\right)\right] \tag{55}$$

as $\sum_\pi^* \prod_{B\in\pi}\kappa(T_{N,k_i}:i\in B)$, where the sum is over all partitions π of $\{1,\ldots,2m\}$ that do not contain atoms and two-element subsets of the form $\{i, i+m\}$, $i = 1,\ldots,m$. Therefore,

$$\mathbb{E}(S_N(f) - \mathbb{E}S_N(f))^m = \sum_\pi^* \sum_{k_1,\ldots,k_m\neq 0} \hat{f}_{k_1}\cdots\hat{f}_{k_m} \prod_{B\in\pi}\kappa(T_{N,k_i}:i\in B). \tag{56}$$

For a given partition π, denote

$$\Sigma_\pi = \sum_{k_1,\ldots,k_m\neq 0} \hat{f}_{k_1}\cdots\hat{f}_{k_m} \prod_{B\in\pi}\kappa(T_{N,k_i}:i\in B) \tag{57}$$

and

$$\mathcal{S}_\pi = \sum_{\substack{k_1,\ldots,k_m\neq 0\\ |k_i|\leq N\log N}} \frac{1}{|k_1|}\cdots\frac{1}{|k_m|}\prod_{B\in\pi}\left[\min(N,\sum_{i\in B}|k_i|)1_{\sum_{i\in B}k_i=0}\right]. \tag{58}$$

Clearly,

$$\Sigma_\pi = O(\mathcal{S}_\pi). \tag{59}$$

Following [40,45], we denote

$$[i] := \{i, i+m\}, \, 1 \leq i \leq m, \tag{60}$$

and introduce the equivalence relation $\sim_\pi$ on the set $\{[1], \ldots, [m]\}$:

$$[i] \sim_\pi [j] \tag{61}$$

if and only if there is a sequence of blocks $B_1, \ldots, B_q$, $q \geq 1$, of the partition π such that

$$B_1 \cap [i] \neq \varnothing, \ B_1 \cap [i_1] \neq \varnothing; \ B_2 \cap [i_1] \neq \varnothing, \ B_2 \cap [i_2] \neq \varnothing; \ldots; B_q \cap [i_{q-1}] \neq \varnothing, \ B_q \cap [j] \neq \varnothing, \tag{62}$$

for some $1 \leq i_1, \ldots, i_{q-1} \leq m$. We call a partition π optimal if the cardinality of every equivalence class is 2. We claim that for each optimal partition π, the corresponding sum $\Sigma_\pi = \sigma^m N^{m/2} \times (1 + o(1))$, and for each suboptimal partition, $\Sigma_\pi = o(N^{m/2})$.

The first part of the claim immediately follows from the variance computations above, as the $m-$dimensional sum (57) then factorizes into the product of $m/2$ two-dimensional sums, each equal $\sigma^2 N (1 + o(1))$.

Next, we will show that the suboptimal partitions give negligible contributions $o(N^{m/2})$. One of the key ingredients is the following upper bound on $\mathcal{S}_\pi$.

Lemma 5. *Let π be a partition of $\{1, \ldots, 2m\}$ that does not contain atoms and two-element subsets of the form $\{i, i+m\}$, $i = 1, \ldots, m$. Then,*

$$\mathcal{S}_\pi = \Omega(N^{m/2}), \tag{63}$$

i.e., $\mathcal{S}_\pi \leq N^{m/2} (\log N)^d$ for some d and all sufficiently large N. Moreover, if π contains at least one block of size less than 4, then

$$\mathcal{S}_\pi = \Omega(N^{\frac{m-1}{2}}). \tag{64}$$

Proof of Lemma 4. The proof uses mathematical induction on m. Without a loss of generality, we can assume that the equivalence relation $\sim_\pi$ has only one equivalence class. Otherwise, the sum $\mathcal{S}_\pi$ factorizes over the equivalence classes, and the argument is applicable to each equivalence class separately.

If π does not contain an equivalence class of size less than 4, the bound (63) follows since the number of terms in the product on the r.h.s. of (58) is not bigger than $m/2$, each cumulant is $O(N)$, and the partial sums of the harmonic series grow logarithmically.

Suppose now that π contains a block of size 2, say $B_1 = \{i, j + m\}, 1 \leq i \neq j \leq m$. Then, $\{i + m, j\}$ cannot be a block of the partition. Consider the block of the partition that contains $i + m$. Without a loss of generality, we can assume that this block is B_2, i.e., $i + m \in B_2$. Construct a partition π' of the set $\{1, \ldots, 2m\} \setminus \{i, i + m\}$, where we discard B_1 and replace $i + m$ in B_2 by $j + m$. It follows from the construction that $S_\pi \leq S_{\pi'}$ and by the inductive hypothesis, $S_{\pi'} = \Omega(N^{\frac{m-1}{2}})$.

Now, suppose that π has a block of size 3, say $B_1 = \{i, j, l + m\}$, $1 \leq i, j, l \leq m$, $l \neq i, j$. Consider the block of π that contains l, say $l \in B_2$. Construct a partition π' of the set $\{1, \ldots, 2m\} \setminus \{l, l + m\}$, where we discard B_1 and replace l in B_2 by i and j. Then, $S_\pi \leq S_{\pi'}$ and, again, applying the inductive hypothesis, $S_{\pi'} = \Omega(N^{\frac{m-1}{2}})$. $\quad\square$

Remark 7. *A similar argument shows that if π contains a block of size 4 of the form $\{i, j, l, l + m\}$, $1 \leq i, j, l \leq m$, then $S_\pi = \Omega(N^{\frac{m-1}{2}})$.*

Finally, if π contains a block of size 4 of the form $B_1 = \{i, j, n + m, l + m\}$, $1 \leq i, j, n, l \leq m$, such that $\{i, j\} \neq \{n, l\}$, we consider two cases, namely, $|k_n| > |k_l|$ and $|k_n| \leq |k_l|$. In the first case, we construct a partition π' of the set $\{1, \ldots, 2m\} \setminus \{n, n + m\}$, where we discard B_1 and replace n in its block by $i, j,$ and $l + m$. In the second case, we do the same procedure with l instead of n. In both cases, the sum is bounded from above by a sum of dimension one less, and we again obtain $S_\pi = \Omega(N^{\frac{m-1}{2}})$ by the inductive assumption.

Now we are ready to finish the proof of the theorem. Recall that we could assume without a loss of generality that the equivalence relation $\sim_\pi$ has only one equivalence class. Note that (59) and (63), Lemma 4 and Remark 7 prove $\Sigma_\pi = o(N^{m/2})$ for all suboptimal partitions π, except the ones that comprise the blocks of size 4 of the form $\{i, j, j + m, l + m\}$, $1 \leq i, j, l \leq m$, $i \neq l$. Since $k_i + k_j + k_{j+m} + k_{l+m} = 0$, and $k_{j+m} = -k_j$, we obtain $k_{l+m} = -k_i$. It follows from the direct computations (see, e.g., [40]) that

$$\kappa_4^{(N)}(k_i, k_j, -k_j, -k_i) = \begin{cases} 0, & \text{if } 1 \leq |k_i| = |k_j| \leq N/2, \\ N - 2|k_i|, & \text{if } N/2 < |k_i| = |k_j| \leq N, \\ -N, & \text{if } |k_i| = |k_j| \geq N, \\ ||k_i| - |k_j|| - N, & \text{if } 1 \leq ||k_i| - |k_j|| \leq N - 1, \ N \leq \max(|k_i|, |k_j|), \\ N - |k_i| - |k_j|, & \text{if } 1 \leq |k_i| \neq |k_j| \leq N - 1, N + 1 \leq |k_i| + |k_j|, \\ 0, & \text{else.} \end{cases} \quad (65)$$

In particular, $|\kappa_4^{(N)}(k_i, k_j, -k_j, -k_i)| \leq \min(|k_i|, |k_j|)$. Without a loss of generality, we can assume that $\pi = \{B_1, \ldots, B_{m/2}\}$, where

$$B_1 = \{1, 2, 2 + m, 3 + m\}, \ B_2 = \{3, 4, 4 + m, 5 + m\}, \ldots, \quad (66)$$

$$B_{m/2-1} = \{m - 3, m - 2, 2m - 2, 2m - 1\}, \ B_{m/2} = \{m - 1, m, 2m, m + 1\}. \quad (67)$$

Then,

$$\Sigma_\pi = \sum_{k_1, \ldots, k_m \neq 0} \hat{f}_{k_1} \cdots \hat{f}_{k_m} \prod_{i=1}^{m/2} \kappa_4^{(N)}(k_{2i-1}, k_{2i}, -k_{2i}, -k_{2i+1}). \quad (68)$$

Since the fourth cumulant function vanishes unless the sum of the arguments is zero, one has $k_1 = k_3 = k_5 = \ldots = k_{m-1}$. Moreover, the fourth cumulant function vanishes unless the sum of the absolute values of the arguments is at least $2N$. Therefore, $|k_1| + |k_{2i}| \geq N$, $1 \leq i \leq m$, and

$$|\Sigma_\pi| \leq Const \sum_{\substack{k_1 \neq 0, \\ |k_1| \leq N \log N}} \sum_{\substack{k_2, k_4, k_6, \ldots, k_m \neq 0 \\ |k_{2i}| \leq N \log N}} \frac{1}{|k_1|^{m/2}} \prod_{i=1}^{m/2} \frac{1}{|k_{2i}|} \prod_{i=1}^{m/2} \min(|k_1|, |k_{2i}|) 1_{|k_1| + |k_{2i}| \geq N}. \quad (69)$$

Splitting the summation with respect to k_1 into two parts, corresponding to $|k_1| \leq \sqrt{N}$ and $|k_1| > \sqrt{N}$ correspondingly, one arrives at the bound:

$$|\Sigma_\pi| \leq 2 \times Const \left(\sum_{0 < k_1 \leq \sqrt{N}} 1 \sum_{\substack{k_2, k_4, k_6, \ldots, k_m \neq 0 \\ |k_{2i}| \leq N \log N}} \prod_{i=1}^{m/2} \frac{1}{|k_{2i}|} + \sum_{|k_1| \geq \sqrt{N}} \frac{1}{|k_1|^{m/2}} \sum_{\substack{k_2, k_4, k_6, \ldots, k_m \neq 0 \\ |k_{2i}| \leq N \log N}} 1 \right).$$

The right-hand side is $= o(N^{m/2})$ for $m > 2$. Therefore, we have shown that $\Sigma_\pi = o(N^{m/2})$ for all suboptimal partitions.

If $m = 2q + 1$ is odd, then clearly there are no optimal partitions, which implies

$$\mathbb{E}(S_N(f) - \mathbb{E}S_N(f))^{2q+1} = o(N^{q+1/2}). \quad (70)$$

If $m = 2q$ is even, then there are exactly $(2q - 1)!!$ optimal partitions, and for each such partition π, one has $\Sigma_\pi = \sigma^{2q}(C)N^q(1 + o(1))$. When we combine these results together, we obtain

$$\mathbb{E}(S_N(f) - \mathbb{E}S_N(f))^{2q} = (2q - 1)!!\sigma^{2q}(C)N^q(1 + o(1)). \quad (71)$$

The bounds (70) and (71) finish the proof. $\square$

4. Discussion

We have discussed several applications of the cumulant technique in Random Matrix Theory, specifically for the ensembles with determinantal k-point correlation functions. The

suggested approach to the fluctuations of multivariate linear statistics of the eigenvalues of random unitary matrices can be extended to other classes of test functions and other classical groups.

Author Contributions: Methodology, A.S.; formal analysis, A.S. and C.W.; investigation, A.S. and C.W.; writing—original draft preparation, A.S. and C.W.; writing—review and editing, A.S. and C.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data sharing not applicable to this article, as no datasets were generated or analyzed during the current study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wigner, E. Characteristic Vectors of Bordered Matrices with Infinite Dimension. I. *Ann. Math.* **1955**, *62*, 548–564. [CrossRef]
2. Wigner, E. Characteristic Vectors of Bordered Matrices with Infinite Dimension. II. *Ann. Math.* **1957**, *65*, 203–207. [CrossRef]
3. Wigner, E. On the Distribution of the Roots of Certain Symmetric Spaces. *Ann. Math.* **1958**, *67*, 325–327. [CrossRef]
4. Bai, Z.D. Methodologies in Spectral Analysis of Large Dimensional Random Matrices, a Review. *Stat. Sin.* **1999**, *9*, 611–677.
5. Bose, A.; Hazra, R.S.; Saha, K. Patterned Random Matrices and Method of Moments. In Proceedings of the International Congress of Mathematicians, Hyderabad, India, 19–27 August 2010; pp. 2203–2231.
6. Diaconu, S. More Limiting Distributions for Eigenvalues of Wigner Matrices. *arXiv* **2022**, arXiv:2203.08712.
7. Diaconu, S. Finite Rank Perturbations of Heavy-Tailed Wigner Matrices. *arXiv* **2022**, arXiv:2208.02756.
8. Feldheim, O.; Sodin, S. A Universality Result for the Smallest Eigenvalues of Certain Sample Covariance Matrices. *Geom. Funct. Anal.* **2010**, *20–21*, 88–123. [CrossRef]
9. Peche, S.; Soshnikov, A. Wigner Random Matrices with Non-Symmetrically Distributed Entries. *J. Stat. Phys.* **2007**, *129*, 857–884. [CrossRef]
10. Sinai, Y.; Soshnikov, A. A Refinement of Wigner's Semicircle Law in a Neighborhood of the Spectrum Edge for Random Symmetric Matrices. *Funct. Anal. Its Appl.* **1998**, *32*, 114–131. [CrossRef]
11. Sodin, S. Several Applications of the Moment Method in Random Matrix Theory. In Proceedings of the International Congress of Mathematicians, Seoul, Republic of Korea, 13–21 August 2014; pp. 451–475.
12. Soshnikov, A. Universality at the Edge of the Spectrum in Wigner Random Matrices. *Commun. Math. Phys.* **1999**, *207*, 697–733. [CrossRef]
13. Yin, Y.Q.; Bai, Z.D.; Krishnaiah, P.R. On the Limit of the Largest Eigenvalue of the Large Dimensional Sample Covariance Matrix. *Probab. Theory Relat. Fields* **1988**, *78*, 509–521. [CrossRef]
14. Costin, O.; Lebowitz, J. Gaussian Fluctuations in Random Matrices. *Phys. Rev. Lett.* **1995**, *75*, 69–72. [CrossRef] [PubMed]
15. Borodin, A. Determinantal Point Processes. In *The Oxford Handbook of Random Matrix Theory*; Akemann, B., Baik, J., Di Francesco, P., Eds.; Oxford University Press: Oxford, UK, 2011; pp. 231–249.
16. Soshnikov, A. Determinantal Random Point Fields. *Russ. Math. Surv.* **2000**, *55*, 923–975. [CrossRef]
17. Soshnikov, A. Gaussian Fluctuation in Airy, Bessel, Sine and Other Determinantal Random Point Fields. *J. Stat. Phys.* **2000**, *100*, 491–522. [CrossRef]
18. Soshnikov, A. Gaussian Limit for Determinantal Random Point Fields. *Ann. Probab.* **2002**, *30*, 171–187. [CrossRef]
19. Hough, J.B.; Krishnapur, M.; Peres, Y.; Virag, B. Determinantal Processes and Independence. *Probab. Surv.* **2006**, *3*, 206–229. [CrossRef]
20. Ameur, Y.; Hedenmalm, H.; Makarov, N. Fluctuations of Eigenvalues of Random Normal Matrices. *Duke Math. J.* **2011**, *159*, 31–81. [CrossRef]
21. Haimi, A.; Romero, J.L. Normality of Smooth Statistics for Planar Determinantal Point Processes. *arXiv* **2022**, arXiv:2210.10588.
22. Rider, B.; Virag, B. Complex Determinantal Processes and H1 Noise. *Electron. J. Probab.* **2007**, *12*, 1238–1257. [CrossRef]
23. Soshnikov, A. Level Spacings Distribution for Large Random Matrices: Gaussian Fluctuations. *Ann. Math.* **1998** *148*, 573–617. [CrossRef]
24. Soshnikov, A. Statistics of Extreme Spacings in Determinantal Random Point Processes. *Mosc. Math. J.* **2005**, *5*, 705–719. [CrossRef]
25. Weyl, H. *The Classical Groups: Their Invariants and Representations*, 2nd ed.; Princeton University Press: Princeton, NJ, USA, 1997.
26. Dyson, F.J. Statistical Theory of the Energy Levels of Complex Systems. III. *J. Math. Phys.* **1962**, *3*, 1191–1198. [CrossRef]
27. Meckes, E.S. *The Random Matrix Theory of the Classical Compact Groups*; Cambridge University Press: Cambridge, UK, 2019.
28. Mehta, M.L. *Random Matrices*, 3rd ed.; Elsevier Ltd.: New York, NY, USA, 2004.
29. Diaconis, P.; Shahshahani, M. On Eigenvalues of Random Matrices. *J. Appl. Probab.* **1994**, *31A*, 49–62. [CrossRef]

30. Johansson, K. On Szego's Asymptotic Formula for Toeplitz Determinants and Generalizations. *Bull. Des Sci. Math.* **1988**, *112*, 257–304.
31. Johansson, K. On Random Matrices from the Compact Classical Groups. *Ann. Math.* **1997**, *145*, 519–545. [CrossRef]
32. Courteaut, K.; Johansson, K.; Lambert, G. From Berry-Esseen to Super-exponential. *arXiv* **2022**, arXiv:2204.03282.
33. Johansson, K.; Lambert, G. Multivariate Normal Approximation for Traces of Random Unitary Matrices. *Ann. Probab.* **2021**, *49*, 2961–3010. [CrossRef]
34. Diaconis, P. Patterns in Eigenvalues. *Bull. New Ser. Am. Math. Soc.* **2003**, *40*, 155–178. [CrossRef]
35. Malyshev, V.A.; Minlos, R.A. *Gibbs Random Fields. Cluster Expansions*; Springer: Dordrecht, The Netherlands, 1991.
36. Soshnikov, A. Central Limit Theorem for Local Linear Statistics in Classical Compact Groups and Related Combinatorial Identities. *Ann. Probab.* **2000**, *28*, 1353–1370. [CrossRef]
37. Haake, F.; Kus, M.; Sommers, H.J.; Schomerus, H.; Zyczkowski, K. Secular Determinants of Random Unitary Matrices. *J. Phys. A* **1996**, *29*, 3641–3658. [CrossRef]
38. Haake, F.; Sommers, H.-J.; Weber, J. Fluctuations and Ergodicity of the Form Factor of Quantum Propagators and Random Unitary Matrices. *J. Phys. A* **1999**, *32*, 6903–6913. [CrossRef]
39. Spohn, H. Interacting Brownian Particles: A Study of Dyson Model. In *Hydrodynamic Behavior and Interacting Particle Systems*; Papanicolau, G., Ed.; Springer: Dordrecht, The Netherlands, 1987; pp. 151–179.
40. Soshnikov, A. Gaussian Fluctuation for Smoothed Local Correlations in CUE. *J. Stat. Phys.* **2023**, *190*, 13. [CrossRef]
41. Breuer, J.; Duits, M. Central Limit Theorems for Biorthogonal Ensembles and Asymptotics of Recurrence Coefficients. *J. Am. Math. Soc.* **2017**, *30*, 27–66. [CrossRef]
42. Deleporte, A.; Lambert, G. Universality for Free Fermions and the Local Weyl Law for Semiclassical Schrödinger Operators. *arXiv* **2021**, arXiv:2109.02121.
43. Lambert, G. CLT for Biorthogonal Ensembles and Related Combinatorial Identities. *Adv. Math.* **2018**, *329*, 590–648. [CrossRef]
44. Lambert, G. Mesoscopic Fluctuations for Unitary Invariant Ensembles. *Electron. J. Probab.* **2018**, *23*, 1–33. [CrossRef]
45. Aguirre, A.; Soshnikov, A.; Sumpter, J. Pair Dependent Linear Statistics for CβE. *Random Matrices Theory Appl.* **2021**, *10*, 2150035. [CrossRef]
46. Aguirre, A.; Soshnikov, A. A Note on Pair Dependent Linear Statistics with Slowly Growing Variance. *Theor. Math. Phys.* **2020**, *205*, 502–512. [CrossRef]
47. Sumpter, J.C. Pair Dependent Linear Statistics for Circular Random Matrices. Ph.D. Thesis, University of California at Davis, Davis, CA, USA, 2021.
48. Chhaibi, R.; Najnudel, J.; Nikeghbali, A. The Circular Unitary Ensemble and the Riemann Zeta Function: The Microscopic Landscape and a New Approach to Ratios. *Invent. Math.* **2017**, *207*, 23–113. [CrossRef]
49. Conrey, J.B. L-Functions and Random Matrices. In *Mathematics Unlimited—2001 and Beyond*; Engquist, B., Schmid, W., Eds.; Springer: Berlin/Heidelberg, Germany, 2000; pp. 331–352.
50. Keating, J. Random Matrices and Number Theory: Some Recent Themes. In *Stochastic Processes and Random Matrices. Lecture Notes of the Les Houches Summer School: Volume 104*; Schehr, G., Altland, A., Fyodorov, Y., O'Connell, N., Cugliandolo, L., Eds.; Oxford University Press: Oxford, UK, 2017; pp. 348–381.
51. Mezzadri, F.; Snaith, N.C. (Eds.) *Recent Perspectives in Random Matrix Theory and Number Theory*; London Mathematical Society Lecture Note Series 322; Cambridge University Press: Cambridge, UK, 2005.
52. Montgomery, H.L. On Pair Correlation of Zeros of the Zeta Function. *Proc. Symp. Pure Math.* **1973**, *24*, 181–193.
53. Montgomery, H.L. Distribution of the Zeros of the Riemann Zeta Function. In Proceedings of the International Congress of Mathematicians, Vancouver, DC, Canada, 21–29 August 1974; pp. 379–381.
54. Hejhal, D. On the Triple Correlation of the Zeros of the Zeta Function. *Int. Math. Res. Notes* **1994**, *7*, 293–302. [CrossRef]
55. Rudnick, Z.; Sarnak, P. Zeros of Principal L-functions and Random Matrix Theory. *Duke Math. J.* **1996**, *81*, 269–322. [CrossRef]
56. Feng, R.; Goetze, F.; Yao, D. Determinantal Point Processes on Spheres: Multivariate Linear Statistics. *arXiv* **2023**, arXiv:2301.09216.
57. Bingham, N.H.; Goldie, C.M.; Teugels, J.L. *Regular Variation*; Cambridge University Press: Cambridge, UK, 1987.
58. Rains, E. High Powers of Random Elements of Compact Lie Groups. *Probab. Theory Relat. Fields* **1997**, *107*, 219–241. [CrossRef]