The Annals of Probability
2021, Vol. 49, No. 6, 2961–3010
https://doi.org/10.1214/21-AOP1520

MULTIVARIATE NORMAL APPROXIMATION FOR TRACES OF RANDOM UNITARY MATRICES

BY KURT JOHANSSON[1] AND GAULTIER LAMBERT[2]

[1]*Department of Mathematics, KTH Royal Institute of Technology, kurtj@kth.se*
[2]*Institute of Mathematics, University of Zürich, gaultier.lambert@math.uzh.ch*

In this article we obtain a superexponential rate of convergence in total variation between the traces of the first m powers of a $n \times n$ random unitary matrices and a $2m$-dimensional Gaussian random variable. This generalizes previous results in the scalar case to the multivariate setting, and we also give the precise dependence on the dimensions m and n in the estimates with explicit constants. We are especially interested in the regime where m grows with n and our main result basically states that if $m \ll \sqrt{n}$, then the rate of convergence in the Gaussian approximation is $\Gamma(\frac{n}{m}+1)^{-1}$ times a correction. We also show that the Gaussian approximation remains valid for all $m \ll n^{2/3}$ without a fast rate of convergence.

CONTENTS

Received April 2020; revised February 2021.
MSC2020 subject classifications. 60B20, 60B15, 60F05.
Key words and phrases. Multivariate Gaussian approximation, Toeplitz determinants, Stein's method.

1. Introduction and main results.

1.1. *Introduction.* Let $\mathbf{U}$ be a random unitary matrix distributed, according to the normalized Haar measure $\mathbb{P}_n$, on the unitary group $U(n)$ of size $n \in \mathbb{N}$. In random matrix theory this is known as the circular unitary ensemble or CUE. The joint law of the eigenvalues $(e^{\mathrm{i}\theta_1}, \dots, e^{\mathrm{i}\theta_n})$ of $\mathbf{U}$, $\mathrm{i} = \sqrt{-1}$, $\theta_j \in [-\pi, \pi]$, under this probability measure has an explicit density given by the Weyl integration formula,

$$\frac{1}{(2\pi)^n n!} \prod_{1 \le k < j \le n} |e^{\mathrm{i}\theta_k} - e^{\mathrm{i}\theta_j}|^2 = \frac{\mho_n}{(2\pi)^n} \prod_{1 \le k < j \le n} \sin^2\left(\frac{\theta_k - \theta_j}{2}\right), \tag{1}$$

where $\mho_n = 2^{n(n-1)}/n!$. Consider the random variable

$$Z = \sum_{k=1}^{m} \xi_{2k-1} \sqrt{\frac{2}{k}} \Re \operatorname{Tr} \mathbf{U}^k + \xi_{2k} \sqrt{\frac{2}{k}} \Im \operatorname{Tr} \mathbf{U}^k,$$

where $\xi = (\xi_1, \dots, \xi_{2m}) \in \mathbb{R}^{2m}$ built from the traces of the unitary matrix $\mathbf{U}$. It is a well–known consequence of the strong Szegő's theorem ([28], see Theorem 2.1 below) that, for any fixed $m \in \mathbb{N}$, $Z \to \|\xi\| \mathcal{N}$ weakly as $n \to \infty$, where $\mathcal{N}$ is a standard Gaussian random variable and $\|\cdot\|$ denotes the Euclidean norm. This is a surprising result since the trace is the sum of n random variables and there is no normalization in n. This limit theorem is also a consequence of the striking fact proved by [12] that all joint moments of $\sqrt{\frac{2}{k}} \Re \operatorname{Tr} \mathbf{U}^k$ and $\sqrt{\frac{2}{k}} \Im \operatorname{Tr} \mathbf{U}^k$, up to a certain order, are identical to those of independent standard Gaussian random variables (see Theorem 7.1 below). Based on this result, Persi Diaconis [10] conjectured that the rate of convergence in total variation norm of Z to a normal random variable should be very fast, even superexponential. This was proved in [22], where it was shown that there are positive constants C and δ so that

$$\mathrm{d_{TV}}\big(Z, \|\xi\| \mathcal{N}\big) \le C n^{-\delta n}, \tag{2}$$

where $\mathrm{d_{TV}}$ denotes the total variation distance (see (7) below for a definition). No explicit expression for C or δ or their dependence on m and the parameters was given.

A related but separate problem is to consider the multivariate convergence of the random variables

$$\mathrm{X}_{2k-1} := \sqrt{\frac{2}{k}} \Re \operatorname{Tr} \mathbf{U}^k \quad \text{and} \quad \mathrm{X}_{2k} := \sqrt{\frac{2}{k}} \Im \operatorname{Tr} \mathbf{U}^k, \tag{3}$$

$1 \le k \le m$. We are interested in the law of the random vector $\mathbf{X} = (\mathrm{X}_1, \dots, \mathrm{X}_{2m})$ when the dimension of the matrix $\mathbf{U}$ is large. Let $\mathbf{G} = (\mathrm{G}_1, \dots, \mathrm{G}_{2m})$ be i.i.d. standard Gaussian random variables. For a fixed $m \in \mathbb{N}$, it again follows from the strong Szegő's theorem that $\mathbf{X} \to \mathbf{G}$ weakly as $n \to \infty$. Peter Sarnak [25] raised the following problem in connection with his work with M. Rubinstein on computing zeros of L-functions and under-determined matrix moment problems. How close is $\mathbf{X}$ to $\mathbf{G}$ in total variation distance as a function of m for a given n? Here, m can depend on n, for example, be a power of n. Is $\mathbf{X}$ still very close to $\mathbf{G}$? This is the main problem investigated in the present paper. Compared with [25], it would have been useful if the rate of convergence for the multivariate case had become slower quickly while increasing m. We show that this is not the case; Theorem 1.1 and Theorem 1.3 summarize our results. The other classical groups can also be considered; see [7]. Since we are mainly interested in the case when m is large, we assume that $m \ge 3$ throughout this paper, and we get a statement for m almost up to $\sqrt{n}$. In case $m = 1$, it is possible to get a more precise result, and this, together with results on single traces, will be considered for all the classical compact groups in a forthcoming publication, [8]. (A bound for $m = 2$ can be directly inferred from the case $m = 3$; a special treatment of this case would only give a slight improvement.) An important aspect of the present work is that, in contrast to [22], we keep explicit track of the constants and the dependence on m. We have also made an effort to optimize in the argument and get reasonable numerical constants.

Since $Z = \mathbf{X} \cdot \xi$, as a consequence of our multivariate results we can improve (2), for a fixed m and uniformly for all ξ, to

$$\mathrm{d}_{\mathrm{TV}}\big(Z, \|\xi\|\mathcal{N}\big) = \mathcal{O}\left(\frac{e^{\frac{n}{m}(\log(1+\log m)+\frac{1}{2})}}{\sqrt{n}\Gamma(\frac{n}{m}+1)}\right), \tag{4}$$

where the implied constant has an explicit dependence in $m \in \mathbb{N}$. Broadly speaking, we expect that the best possible estimate for the RHS of (4) is $\Gamma(\frac{n}{m}+1)^{-1}$ times some sub-exponential corrections. We can also let the degree m grow as $n \to +\infty$. From Proposition 1.6 we deduce the following estimate:

$$\sup_{m \le \frac{\sqrt{n}}{6.45(\log n)^{1/4}}} \mathrm{d}_{\mathrm{TV}}\big(Z, \|\xi\|\mathcal{N}\big) \le n^{3/4} \exp\big(-38 - 0.93\sqrt{n}(\log n)^{5/4}\big)$$

uniformly for all ξ when $n \ge 4322$.

Using Stein's method and the exact moment identities from [12], one can infer the following rate of convergence in the multivariate problem: for any $m \le 2n$,

$$\mathrm{W}_1(\mathbf{X}, \mathbf{G}) = \mathcal{O}\big(m^2/n\big), \tag{5}$$

where W_1 denotes the Wasserstein 1 distance between two probability measures on $\mathbb{R}^{2m}$; see [14], Theorem 1.1. By relying on the recent techniques from [24], we can improve on (5); see Theorem 1.5 below. See also [29] for an analogous multivariate result that applies to more general circular β-ensembles and Remark 1.1 below. Recently, rates of convergence to the Gaussian law have also been obtained for $\operatorname{Tr} f(\mathbf{M})$, where $f : \mathbb{R} \to \mathbb{R}$ is a real–analytic function and $\mathbf{M}$ is a random matrix from the Gaussian, Laguerre or Jacobi unitary ensembles by [3] using Riemann–Hilbert techniques. In contrast to the CUE, in these cases the optimal rates of convergence are expected to be polynomial in the dimension of the random matrix. A similar question has been considered in [18] for traces of a uniform random matrix from the unitary group over a finite field. The rate of convergence in this case is also superexponential.

The fast rate of convergence of $\mathbf{X}$ to $\mathbf{G}$ holds for $m \ll \sqrt{n}$ by Theorem 1.1, but we see from Theorem 1.5 below that we have convergence to the multivariate Gaussian for $m \ll n^{2/3}$. We have no conjecture concerning the threshold $m \in \mathbb{N}$ at which the Gaussian approximation fails. Also, we do not know whether there is some transition when varying m where we go from a fast convergence rate to some other rate of convergence.

1.2. *Main results.* For any $m \in \mathbb{N}$, we denote by $\Omega_m = \frac{\pi^m}{m!}$ the volume of the unit ball and by $\|x\| = \sqrt{x_1^2 + \cdots + x_{2m}^2}$ the Euclidean norm in $\mathbb{R}^{2m}$. It is straightforward to see that, for any $m, n \in \mathbb{N}$, the random vector $\mathbf{X}$ has a density on $\mathbb{R}^{2m}$ that we denote by $\wp_{n,m}$. For any $n, m \in \mathbb{N}$ and $k \in \mathbb{N}$, we define

$$\Delta_{n,m}^{(k)} := \left(\int_{\mathbb{R}^{2m}} \left|\wp_{n,m}(x) - \frac{e^{-\|x\|^2/2}}{(2\pi)^m}\right|^k \mathrm{d}x\right)^{1/k}. \tag{6}$$

In this paper we focus on getting (nonasymptotic) estimates for $\Delta_{n,m}^{(1)}$ and $\Delta_{n,m}^{(2)}$ with explicit constants which hold for large $n \in \mathbb{N}$ when $m \ll \sqrt{n}$. Let us observe that $\Delta_{n,m}^{(1)}$ controls the *total variation* distance between **X** and **G** (a standard Gaussian random variable on $\mathbb{R}^{2m}$). Namely, we have

$$\mathrm{d_{TV}}(\mathbf{X},\mathbf{G}) := \sup_{A\subset\mathbb{R}^{2m}} \left|\mathbb{P}_n[\mathbf{X}\in A] - \mathbb{P}[\mathbf{G}\in A]\right| \le \Delta^{(1)}_{n,m}, \tag{7}$$

where the supremum is taken over all Borel subsets $A \subset \mathbb{R}^{2m}$. Our main result, which is a quantitative generalization of the estimates (2) from [22] in a multidimensional setting, can be summarized as follows.

THEOREM 1.1. *For all $n, m \in \mathbb{N}$ such that $n \ge 1911$ and $N = n/m \ge 146.5m \times \sqrt{1+\log m}$, we have the following estimate in total variation distance*:

$$\mathrm{d_{TV}}(\mathbf{X},\mathbf{G}) \le 16 m^{\frac{5}{2}} \frac{2^m}{\pi^{\frac{m}{2}}\sqrt{m!}} e^{\frac{N}{2}+\frac{m^2}{4N}} \frac{(N\sqrt{\log N})^m (1+\log m)^N}{\sqrt{N}\,\Gamma(N+1)}.$$

We expect that, up to corrections, the factor $\Gamma(\frac{n}{m}+1)^{-1}$ is actually the correct order for the statistical distance between the random vectors **X** and **G** as long as $m \ll \sqrt{n}$. To clarify the meaning of this estimate in the regime where m grows with n, let us also give the following consequence when m is like n^α, $\alpha < 1/2$.

PROPOSITION 1.2. *Let $m = \lfloor n^\alpha \rfloor$ with $0 < \alpha < 1/2$, then, for all $n \ge n_\alpha$,*

$$\mathrm{d_{TV}}(\mathbf{X},\mathbf{G}) \le 1.4 \cdot 10^{-13} n^{3\alpha-\frac{3}{2}} \exp\bigl(-(1-\epsilon_n(\alpha)) n^{1-\alpha} \log(n^{1-\alpha})\bigr),$$

where $n_\alpha := \inf\{n \ge 17^{1/\alpha} : n^{1-2\alpha} \ge 20.4\sqrt{\log n}\}$ and $\epsilon_n(\alpha)$ is given by (56). In particular, it holds $1-\epsilon_n \ge 87 \cdot 10^{-3}$ and $\epsilon_n \to 0$ as $n \to \infty$.

The proof of Theorem 1.1 and Proposition 1.2 are given in Section 3.3. According to (7), these results are consequences of the following more precise bounds. We postpone the definition of $\Theta_{N,m}$ to Appendix B since it is rather involved.

THEOREM 1.3. *Let $\Theta_{N,m}$ be given by (114). For any $n, m \in \mathbb{N}$ such that $m \ge 3$ and $N = n/m > 4m$, we have*

$$\Delta^{(2)}_{n,m} \le \frac{8\sqrt{\Omega_m}}{(2\pi)^m} N^{\frac{m}{2}} \Theta_{N,m}. \tag{8}$$

If we assume that $\Theta_{N,m} \le m^{\frac{5}{4}} (\frac{e}{\sqrt{\pi}})^{-m} N^{-\frac{m}{2}}$, then

$$\Delta^{(1)}_{n,m} \le 2\bigl(8 \log \Delta^{(2)-1}_{n,m}\bigr)^{\frac{m}{2}} \Delta^{(2)}_{n,m}. \tag{9}$$

The proof of Theorem 1.3 is explained in Section 2, and it is given in Section 3. This shows that the parameter $\Theta_{N,m}$ controls the statistical distance between the random vectors **X** and **G**. We have made significant efforts to keep track carefully of the dependency in n, m of our estimate with reasonable numerical constants. Unfortunately, this leads to an expression for $\Theta_{N,m}$ which is rather involved; see Section B.1. In particular, there are several regimes depending on n and m where different contributions are relevant. Let us just point out that in the cases in which we are most interested, that is, when m is large and $N = \frac{n}{m}$ is sufficiently large compared to m, we obtain the following bounds which allow us to verify the second assumption in the formulation of Theorem 1.3.

PROPOSITION 1.4. *Fix an integer $M \geq 3$. It holds for all $m \geq M$ and $N = n/m \geq c(M)m\sqrt{1+\log m}$,*

$$\Theta_{N,m} \leq (1+\epsilon)m^{\frac{5}{2}} 2^{\frac{m}{2}} e^{\frac{m^2}{4N}} \frac{e^{\frac{N}{2}}(1+\log m)^N}{\sqrt{N}\Gamma(N+1)},$$

with $\epsilon \leq 25 \cdot 10^{-5}$ and $c(M)$ are explicit constants given in the table (125). We emphasize that $c(M)$ is nonincreasing in $M \in \mathbb{N}$ with $c(3) = 146.5$, as in Theorem 1.1 and $c(M) = 19.4$ for $M \geq 70$.

The proof of Proposition 1.4 involves rather technical numerical estimates (which have been obtained with *Mathematica*), and it is given in Appendix B.2.

Our next result shows that it is still possible to approximate $\mathbf{X}$ by a Gaussian random vector when $m \gg \sqrt{n}$. It is an interesting question whether the approximation also holds for the total variation distance. Recall that the Kantorovich or Wasserstein distances between the random vectors $\mathbf{X}$ and the Gaussian $\mathbf{G}$ are defined by, for any $q \geq 1$,

$$\mathrm{W}_q(\mathbf{X}, \mathbf{G}) = \inf_{\mathbb{P}} (\mathbb{E}[\|\mathrm{x} - \mathrm{g}\|^q])^{1/q}, \tag{10}$$

where the infimum is taken over all probability measures on $\mathbb{R}^{2m} \times \mathbb{R}^{2m}$ such that the first marginal of $\mathbb{P}$, x has the same law as $\mathbf{X}$ and the second marginal of $\mathbb{P}$, g is a standard Gaussian on $\mathbb{R}^{2m}$.

THEOREM 1.5. *For any $n, m \in \mathbb{N}$ such that $n \geq 2m$, it holds*

$$\mathrm{W}_2(\mathbf{X}, \mathbf{G}) \leq (\sqrt{8} + \sqrt{2}) \frac{(m+1)\sqrt{m}}{3n}.$$

This shows that if $m \to +\infty$ in such a way that $m = o(n^{2/3})$, then the Kantorovich distance between the random vector $\mathbf{X}$ and a standard Gaussian $\mathbf{G}$ on $\mathbb{R}^{2m}$ converges to 0 as $n \to +\infty$. The proof of Theorem 1.5 is given in Section 7, and it relies on the normal approximation method from [24]; see Proposition 7.2 below. This result allows to turn the moments' identities of [12] into a quantitative statement about the rate of convergence to the normal distribution in the Kantorovich distance. Let us emphasize that the result from [24], which is used to prove Theorem 1.5, is inspired by Stein's method and is, therefore, completely unrelated to the techniques that we develop in Sections 2–6 to prove our main result.

REMARK 1.1. If we let for $k \geq 1$,

$$\mathrm{X}_{2k-1} = \frac{2}{\sqrt{\beta k}} \sum_{j=1}^{n} \cos(k\theta_j) \quad \text{and} \quad \mathrm{X}_{2k} = \frac{2}{\sqrt{\beta k}} \sum_{j=1}^{n} \sin(k\theta_j),$$

then, the counterpart of Theorem 1.5 also holds for the circular β–ensembles $\{\theta_1, \dots, \theta_n\}$. That is, for any $\beta > 0$, there exists a constant $C_\beta > 0$ such that, for all $n, m \in \mathbb{N}$ with $n \geq 2m$,

$$\mathrm{W}_2(\mathbf{X}, \mathbf{G}) \leq C_\beta \frac{m^{3/2}}{n}.$$

The proof is similar to that of Theorem 1.5, and it relies on Proposition 7.2 and Lemma 7.3. The only differences lie in that instead of using the moments' identities of [12], one can make use of the estimates from [20], Theorem 1. These estimates for the joint moments of $\mathbf{X}$ corresponds to the analogue for general $\beta > 0$ of Theorem 7.1 with constants, which are not sharp, and they are obtained by using the Jack functions instead of Schur functions as in the case of the unitary group ($\beta = 2$). Then, it is straightforward to control the errors as in Lemmae 7.4 and 7.5. Likewise, a similar result also holds for the other classical compact groups (i.e., for the circular orthogonal and symplectic ensembles) with the appropriate normalization.

Let us give a final application of Theorem 1.3 when m is close to $\sqrt{n}$ and the dimension n of the random matrix $\mathbf{U}$ is large. Namely, we obtain the following corollary.

PROPOSITION 1.6. *Let us assume that $n \geq 4322$. Then, it holds for any integer $m \leq \sqrt{\frac{n}{41.5\sqrt{\log n}}}$,*

$$\mathrm{d}_{\mathrm{TV}}(\mathbf{X}, \mathbf{G}) \leq n^{3/4} \exp\bigl(-38 - 0.93\sqrt{n}(\log n)^{5/4}\bigr).$$

The proof of Proposition 1.6 is also given in Section 3.3. We verify numerically that under the assumptions of Proposition 1.6, $\mathrm{d}_{\mathrm{TV}}(\mathbf{X}, \mathbf{G}) \leq 10^{-391}$ which is far below *machine epsilon* (of order of 10^{-33} for quad(ruple) precision decimal). In the Appendix B.3 we present further numerical plots which illustrate our estimates in the case $m = 3$.

2. Overview of the proof of Theorem 1.3. The core of the proof of Theorem 1.3 is to obtain the estimate (8) for the L^2 distance $\Delta^{(2)}_{n,m}$ between the density $\wp_{n,m}$ of the random vector $\mathbf{X}$ and the standard Gaussian density on $\mathbb{R}^{2m}$. Observe that, by Parseval's formula, we can rewrite for any $n, m \in \mathbb{N}$,

$$\Delta^{(2)}_{n,m} = \frac{1}{(2\pi)^m}\left(\int_{\mathbb{R}^{2m}} \left|F_{n,m}(\xi) - e^{-\|\xi\|^2/2}\right|^2 \mathrm{d}\xi\right)^{1/2}, \tag{11}$$

where $F_{n,m}$ denotes the characteristic function of the random vector $\mathbf{X}$. Like in the proof of [22], the general strategy is to obtain precise estimates for $F_{n,m}$, and we need to distinguish different *regimes* depending the parameters ξ, m and $N = n/m$. These *regimes* are explained in Section 2.4, and we use different methods to treat them. Compared with the arguments of [22], considerable improvement is needed. There are two new challenges that come up since we allow the degree $m \in \mathbb{N}$ to grow with n, and we want to keep track carefully of the constants. Let us also point out that the improvements of Theorem 1.3 come from new techniques, especially from using the *Borodin–Okounkov formula* that we recall in the next section. We also make a more careful use of the *change of variables method* from [22] that we review in Section 2.3. The main steps of the proof of the estimate (8) are presented in Section 2.4, while the details of the proof are given in Section 3.

2.1. *Notation.* In this section we collect the main notation that will be use throughout the rest of this paper.

We let $\mathbb{T} = \mathbb{R}/[2\pi]$ and view the CUE measure (1) as a probability measure on $\mathbb{T}^n$. For any $f : \mathbb{T} \to \mathbb{C}$, which is integrable, the random variable $\operatorname{Tr} f(\mathbf{U}) = \sum_{k=1}^{n} f(\theta_k)$ is well defined with $\mathbb{E}_n[\operatorname{Tr} f(\mathbf{U})] = n\widehat{f}_0$. Then, for any $\xi \in \mathbb{R}^{2m}$, we have $\mathbf{X} \cdot \xi = \operatorname{Tr} \mathrm{g}(\mathbf{U})$, where g is a real–valued trigonometric polynomial,

$$\mathrm{g}(\theta) = \sum_{\substack{|k| \leq m \\ k \neq 0}} \frac{\zeta_k}{\sqrt{2|k|}} e^{\mathrm{i}k\theta}, \tag{12}$$

with $\zeta_k = \xi_{2k-1} - \mathrm{i}\xi_{2k}$ and $\zeta_{-k} = \overline{\zeta_k}$ for all $k = 1, \dots, m$. In particular, the characteristic function of the random vector $\mathbf{X}$ can be written as

$$\begin{aligned} F_{n,m}(\xi) &:= \int_{\mathbb{R}^{2m}} e^{\mathrm{i}\xi \cdot x} \wp_{n,m}(x)\,\mathrm{d}x \\ &= \mathbb{E}_n\bigl[e^{\mathrm{i}\operatorname{Tr}\mathrm{g}(\mathbf{U})}\bigr]. \end{aligned} \tag{13}$$

For any function $f \in L^1$, we define its Fourier coefficients for all $k \in \mathbb{Z}$,

$$\widehat{f}_k = \int_{\mathbb{T}} f(\theta) e^{-\mathrm{i}k\theta} \frac{\mathrm{d}\theta}{2\pi}.$$

Then, we define the following (semi)norm:

$$\|f\|_{H^{1/2}}^2 = \sum_{k\in\mathbb{Z}} |k||\widehat{f}_k|^2.$$

If $f \in H^{1/2}$, that is, if $f \in L^1$ and $\|f\|_{H^{1/2}}^2 < +\infty$, we let

$$\mathcal{A}(f) = \sum_{k\geq 1} k \widehat{f}_k \widehat{f}_{-k}. \tag{14}$$

If the real-valued function f lies in the Sobolev space H^1, we also verify that

$$\|f\|_{H^{1/2}}^2 = -\int f'(\theta)\mathcal{U}f(\theta)\,\frac{\mathrm{d}\theta}{2\pi},$$

where $\mathcal{U}f = -\sum_{k\in\mathbb{Z}} \mathrm{i}\,\mathrm{sgn}(k)\widehat{f}_k e^{\mathrm{i}k\theta}$ denotes the Hilbert tranform of f.

2.2. *Preliminaries*: *Toeplitz determinants and the Borodin–Okounkov formula*. Recall that the CUE refers to a random matrix $\mathbf{U}$, which is distributed according to the Haar measure on the unitary group $U(n)$, and that the eigenvalues of $\mathbf{U}$ have a joint law which is explicitly given by (1). One of the most remarkable feature of the CUE is the connection with Toeplitz determinants. Namely, for any integrable function $w = e^f$, $f : \mathbb{T} \to \mathbb{C}$ and $n \in \mathbb{N}$, if $\mathrm{Tr}\, f(\mathbf{U}) = \sum_{j=1}^n f(\theta_j)$, then we have

$$\mathbb{E}_n\big[e^{\mathrm{Tr}\, f(\mathbf{U})}\big] = \det_{n\times n}[\widehat{w}_{i-j}]. \tag{15}$$

Formula (15) implies that we can obtain the asymptotics of the Laplace transform of the random variable $\mathrm{Tr}\, f(\mathbf{U})$ by using Szegő's strong limit theorem.

THEOREM 2.1. *If $f \in H^{1/2}$, then, as $n \to +\infty$,*

$$\mathbb{E}_n\big[e^{\mathrm{Tr}\, f(\mathbf{U})}\big] = \exp\big(n\widehat{f}_0 + \mathcal{A}(f) + o(1)\big), \tag{16}$$

where $\mathcal{A}(f) = \sum_{k\geq 1} k\widehat{f}_k \widehat{f}_{-k} \in \mathbb{C}$.

The first version of Theorem 2.1 was first proved by [28] when $f \in C^{1,\alpha}$ is real-valued. The hypothesis from Theorem 2.1 are optimal, and this version was first obtained for real–valued f by [19] and [17]. We refer to the survey paper of [9] for a history of Szegő's strong limit theorem and its later generalizations and to the book [27], Chapter 6, for a detailed presentation of several proofs. A proof of Theorem 2.1, which holds for complex-valued f, can be found in [21].

Actually, one can also obtain Theorem 2.1 as a consequence of the *Borodin–Okounkov formula*. This formula expresses the Toeplitz determinant (15) in terms of Fredholm determinant which is more amenable for asymptotic analysis. If $f : \mathbb{T} \to \mathbb{C}$ is an L^2 function, we denote

$$f^+(\theta) = \sum_{k\geq 1} \widehat{f}_k e^{\mathrm{i}k\theta}, \qquad f^-(\theta) = \sum_{k\geq 1} \widehat{f}_{-k} e^{-\mathrm{i}k\theta}.$$

Let $w : \mathbb{T} \to \mathbb{C}$ be an integrable function such that $\sum_{k\in\mathbb{Z}} |k||\widehat{w}_k|^2 < +\infty$, and define two Hankel operators,

$$H_+(w) = \begin{pmatrix} \widehat{w}_1 & \widehat{w}_2 & \widehat{w}_3 & \dots \\ \widehat{w}_2 & \widehat{w}_3 & \widehat{w}_4 & \dots \\ \widehat{w}_3 & \widehat{w}_4 & \widehat{w}_5 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{and} \quad H_-(w) = \begin{pmatrix} \widehat{w}_{-1} & \widehat{w}_{-2} & \widehat{w}_{-3} & \dots \\ \widehat{w}_{-2} & \widehat{w}_{-3} & \widehat{w}_{-4} & \dots \\ \widehat{w}_{-3} & \widehat{w}_{-4} & \widehat{w}_{-5} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{17}$$

Note that the condition $\sum_{k\in\mathbb{Z}} |k||\widehat{w}_k|^2 < +\infty$ guarantees that these operators are Hilbert–Schmidt on $L^2(\mathbb{N})$. We also denote by $(e_1, e_2, \dots)$, the standard basis of $L^2(\mathbb{N})$.

THEOREM 2.2. *Let* $f:\mathbb{T}\to\mathbb{C}$ *be a* L^∞ *function such that* $\sum_{k\in\mathbb{Z}}|k||\widehat{f}_k|^2<+\infty$ *and* $\widehat{f}_0=0$. *Let us also define*

$$K_f = H_+\big(e^{f^- - f^+}\big)H_-\big(e^{f^+ - f^-}\big). \tag{18}$$

The operator K_f *is trace class, and, for any* $n\in\mathbb{N}$,

$$\mathbb{E}_n\big[e^{\operatorname{Tr} f(\mathbf{U})}\big] = e^{\mathcal{A}(f)}\det[\mathrm{I}-K_f Q_n], \tag{19}$$

where Q_n *denotes the orthogonal projection with kernel* $\operatorname{span}(e_1,\dots,e_{n-1})$ *and the RHS is a Fredholm determinant on* $L^2(\mathbb{N})$.

Since the operator K_f is trace class, by definition of Q_n, we have $\det[\mathrm{I}-K_f Q_n]\to 1$, as $n\to+\infty$, so that Theorem 2.2 implies Szegő's strong limit theorem. The *Borodin–Okounkov formula* (19) (sometimes also known as *Geronimo–Case formula*) first appeared (formally) in [15]. [4] proved formula (19) in a different form when f is analytic using Gessel's Theorem which allows to express Toeplitz determinants as series in Schur functions, [16]. The version from Theorem 2.2 is due to [1]; see also [5] for a different proof. It is possible to remove the condition $f\in L^\infty$ from Theorem 2.2; see, for example, [27], Chapter 6.2.

Concerning our method, let us point out that, in order to obtain the *superexponential* rate of convergence in (2), [22] relied on exact formulae for Toeplitz determinants with certain specific symbols which are due to [2] and relates to the original proof of Szegő's strong theorem. Observe that, according to (12), we have $\mathcal{A}(\mathrm{ig})=-\|\zeta\|^2/2=-\|\xi\|^2/2$ so that by (13) and (19), we can rewrite, for all $n,m\in\mathbb{N}$ and $\xi\in\mathbb{R}^{2m}$,

$$F_{n,m}(\xi) = e^{-\|\xi\|^2/2}\det[\mathrm{I}-K_{\mathrm{ig}}Q_n]. \tag{20}$$

Hence, by controlling precisely how close the Fredholm determinant $\det[\mathrm{I}-K_{\mathrm{ig}}Q_n]$ is to 1, we are able to significantly improve the rate of convergence from [22]. Even though this might be difficult to verify, it is natural to expect that modulo corrections, $1/\Gamma(N+1)$ should be the true rate of convergence in Theorem 1.3 in the regime where $m\ll N=\frac{n}{m}$.

Throughout this article we also make crucial use of the following bound.

LEMMA 2.3. *Suppose that* $f\in\mathcal{C}(\mathbb{T})$ *is real valued with* $\mathcal{A}(f)<+\infty$, *where* $\mathcal{A}$ *is as in* (14). *Then, for any* $n\in\mathbb{N}$,

$$\mathbb{E}_n\big[e^{\operatorname{Tr} f(\mathbf{U})}\big]\le \exp\big(n\widehat{f}_0+\mathcal{A}(f)\big). \tag{21}$$

Let us recall that $\mathbb{E}_n[\operatorname{Tr} f(\mathbf{U})]=n\widehat{f}_0$ and that by Theorem 2.1, $\operatorname{Var}[\operatorname{Tr} f(\mathbf{U})]\to 2\mathcal{A}(f)$ as $n\to+\infty$, so that the estimate (21) is sharp. The upper bound (21) is classical, and it follows for instance from the monotonicity of Toeplitz determinants, [26]. For completeness, we show in the Appendix (Section A.1), how one can immediately deduce Lemma 2.3 from the *Borodin–Okounkov formula*.

2.3. *Change of variables.* In addition to the *Borodin–Okounkov formula* (Theorem 2.2) and Lemma 2.3, our main tool to prove Theorem 1.3 is the change of variables method introduced in [21]. More specifically, we rely on an estimate from the proof of [22], Proposition 2.8. Recall that, according to (13), $F_{n,m}$ denotes the characteristic function of the random variable $\operatorname{Tr} \mathrm{g}(\mathbf{U})$.

LEMMA 2.4. *Let* $\nu>0$ *and* $h:\mathbb{T}\to\mathbb{R}$ *be a* C^1 *function. Then, for any* $n,m\in\mathbb{N}$ *and* $\xi\in\mathbb{R}^{2m}$,

$$\big|F_{n,m}(\xi)\big|\le \mathbb{E}_n\Bigg[\prod_{1\le i<j\le n}\left|\frac{\sin\big(\frac{\theta_i-\theta_j}{2}+\mathrm{i}\nu\frac{h(\theta_i)-h(\theta_j)}{2n}\big)}{\sin\big(\frac{\theta_i-\theta_j}{2}\big)}\right|^2\prod_{j=1}^n\left|1+\mathrm{i}\frac{\nu}{n}h'(\theta_j)\right|e^{-\Im \mathrm{g}(\theta_j+\mathrm{i}\frac{\nu}{n}h(\theta_j))}\Bigg].$$

PROOF. For completeness, let us give the proof of Lemma 2.4. Using the explicit formula (1) for the joint law of the eigenvalues of the random matrix $\mathbf{U}$, we obtain

$$F_{n,m}(\xi) = \mho_n \int_{[-\pi,\pi]^n} \prod_{1\le i<j\le n} \sin^2\left(\frac{\theta_i-\theta_j}{2}\right) \prod_{k=1}^n e^{\mathrm{i}g(\theta_k)} \frac{\mathrm{d}\theta_k}{2\pi}.$$

If we regard θ_k as complex variables in the previous integral, since the integrand is an entire function, we can deform the contours of integration in the complex plane. Let $\boldsymbol{\gamma}$ be a positively oriented curve given by

$$\boldsymbol{\gamma} = \left\{\theta + \mathrm{i}\frac{\nu}{n}h(\theta) : \theta \in [-\pi,\pi]\right\}.$$

Since the functions g and $\sin^2(\cdot/2)$ are also 2π–periodic, we have, by Cauchy's theorem,

$$\begin{aligned} F_{n,m}(\xi) &= \mho_n \int_{\boldsymbol{\gamma}^n} \prod_{1\le i<j\le n} \sin^2\left(\frac{\theta_i-\theta_j}{2}\right) \prod_{k=1}^n e^{\mathrm{i}g(\theta_k)} \frac{\mathrm{d}\theta_k}{2\pi} \\ &= \mho_n \int_{[-\pi,\pi]^n} \prod_{1\le i<j\le n} \left(\sin\left(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{h(\theta_i)-h(\theta_j)}{2n}\right)\right)^2 \qquad (22) \\ &\quad \times \prod_{j=1}^n e^{\mathrm{i}g(\theta_j+\mathrm{i}\frac{\nu}{n}h(\theta_j))}\left(1+\mathrm{i}\frac{\nu}{n}h'(\theta_j)\right)\frac{\mathrm{d}\theta_j}{2\pi}. \end{aligned}$$

Hence, by (1), this implies that $|F_{n,m}(\xi)|$ is bounded by

$$\begin{aligned} &\mho_n \int_{[-\pi,\pi]^n} \prod_{1\le i<j\le n} \left|\sin\left(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{h(\theta_i)-h(\theta_j)}{2n}\right)\right|^2 \\ &\quad \times \prod_{j=1}^n e^{-\Im g(\theta_j+\mathrm{i}\frac{\nu}{n}h(\theta_j))}\left|1+\mathrm{i}\frac{\nu}{n}h'(\theta_j)\right|\frac{\mathrm{d}\theta_j}{2\pi} \\ &= \mathbb{E}_n\left[\prod_{1\le i<j\le n}\left|\frac{\sin(\frac{\theta_i-\theta_j}{2}+\mathrm{i}\nu\frac{h(\theta_i)-h(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right|^2 \prod_{j=1}^n\left|1+\mathrm{i}\frac{\nu}{n}h'(\theta_j)\right|e^{-\Im g(\theta_j+\mathrm{i}\frac{\nu}{n}h(\theta_j))}\right]. \qquad \square \end{aligned}$$

The key idea underlying this change of variables is that the eigenvalues of $\mathbf{U}$ are almost uniformly distributed on the unit circle (like the vertices of a regular n-gon). This means that at first order, we can approximate the empirical measure $\sum_{j=1}^n \delta_{\theta_j} \simeq n\frac{\mathrm{d}\theta}{2\pi}$. Choose $h = \mathcal{U}\mathrm{g}$, where $\mathcal{U}$ is the Hilbert transform,

$$h(\theta) = -\sum_{\substack{|k|\le m\\ k\ne 0}} \operatorname{sgn}(k)\frac{\mathrm{i}\zeta_k}{\sqrt{2|k|}}e^{\mathrm{i}k\theta}. \qquad (23)$$

By making the change of variables θ_j by $\theta_j + \mathrm{i}\frac{\nu}{n}h(\theta_j)$ in (22), we expect that, using first-order Taylor approximations,

$$\begin{aligned} F_{n,m}(\xi) &\simeq \mho_n \int_{[-\pi,\pi]^n} \prod_{1\le i<j\le n} \sin^2\left(\frac{\theta_i-\theta_j}{2}\right) e^{\frac{\nu^2}{n^2}H(\theta_i,\theta_j)} \prod_{j=1}^n e^{\mathrm{i}g(\theta_j)-\frac{\nu}{n}g'(\theta_j)h(\theta_j)+\frac{\nu^2}{n^2}h'(\theta_j)^2}\frac{\mathrm{d}\theta_j}{2\pi} \\ &= \mathbb{E}_n\left[e^{\frac{\nu^2}{2n^2}\sum_{i,j=1}^n H(\theta_i,\theta_j)} e^{\mathrm{i}\sum_{j=1}^n g(\theta_k)-\frac{\nu}{n}\sum_{j=1}^n g'(\theta_k)h(\theta_k)}\right], \end{aligned}$$

where

$$H(\theta,x)=\begin{cases}\left(\dfrac{h(\theta)-h(x)}{2\sin(\frac{\theta-x}{2})}\right)^2, & x\neq\theta,\\ h'(\theta)^2, & x=\theta.\end{cases}\tag{24}$$

Then, since $\widehat{g}_0=0$, we expect that

$$F_{n,m}(\xi)\simeq\exp\left(-\nu\int_{[0,2\pi]}g'(\theta)h(\theta)\,\frac{d\theta}{2\pi}+\frac{\nu^2}{2}\iint_{[0,2\pi]^2}H(\theta,x)\,\frac{d\theta}{2\pi}\,\frac{dx}{2\pi}\right).$$

Then, by Devinatz's formula [26], Proposition 6.1.10, since $h=-\mathcal{U}g$, we have

$$\|h\|^2_{H^{1/2}}=\int_{[0,2\pi]}g'(\theta)h(\theta)\,\frac{d\theta}{2\pi}=\iint_{[0,2\pi]^2}H(\theta,x)\frac{d\theta}{2\pi}\frac{dx}{2\pi}\tag{25}$$

and

$$\|h\|^2_{H^{1/2}}=\sum_{k\in\mathbb{Z}}|k||h_k|^2=\sum_{k=1}^{m}|\zeta_k|^2=\|\xi\|^2.\tag{26}$$

Whence it follows from this heuristic with $\nu=1$ that

$$F_{n,m}(\xi)\simeq e^{-\|\xi\|^2/2}.$$

To turn this heuristics rigorous, one needs to justify the approximation $\sum_{j=1}^{n}\delta_{\theta_j}\simeq n\,\frac{d\theta}{2\pi}$ and to control the errors coming from the Taylor expansions. This can be done by using *rigidity estimates* for the CUE eigenvalues (see [23]), but we present a different approach below (see (iii) Intermediate regime in the next section).

2.4. *Estimates for the function $F_{n,m}(\xi)$ in the different regimes.* Recall that we let $N=\frac{n}{m}$. All constants c_j used below, which can depend on m, are defined in Appendix B. Our main goal is to prove the following bound.

PROPOSITION 2.5. *For any $n,m\in\mathbb{N}$ such that $m\geq 3$ and $N=n/m>4m$, we have*

$$\Delta^{(2)}_{n,m}=\frac{1}{(2\pi)^m}\left(\int_{\mathbb{R}^{2m}}\left|F_{n,m}(\xi)-e^{-\|\xi\|^2/2}\right|^2 d\xi\right)^{1/2}\leq\frac{8}{(2\pi)^m}\sqrt{\Omega_m}N^{\frac{m}{2}}\Theta_{N,m},\tag{27}$$

where $\Theta_{N,m}$ is as in (111)–(114).

In this section we present the main estimates for the characteristic function $F_{n,m}(\xi)$ that are required to prove Proposition 2.5. We postpone the technical details of the arguments to Sections 4–6. Let us define

$$\Lambda_1=\frac{c_4 N}{\sqrt{1+\log m}}.\tag{28}$$

The proof consists in splitting the integral on the LHS of (27) in three different regimes, depending on whether: (i) $\|\xi\|\leq\Lambda_1$, (ii) $\|\xi\|\geq\Lambda_3$ or (iii) $\Lambda_1\leq\|\xi\|\leq\Lambda_3$ where $\Lambda_3\gg\Lambda_1$ is a parameter that we will choose later:

(i) Gaussian approximation for $\|\xi\|\leq\Lambda_1$. In this regime our goal is to compare the characteristic function $F_{n,m}$ with that of a $2m$-dimensional standard Gaussian by using the *Borodin–Okounkov formula* from Theorem 2.2. We obtain the following estimates.

PROPOSITION 2.6. *Under the assumptions of Proposition 2.5, we have, for all $\xi \in \mathbb{R}^{2m}$ such that $\|\xi\| \le \Lambda_1$,*

$$\left|F_{n,m}(\xi) - e^{-\|\xi\|^2/2}\right|^2 \le c_8{}^2 m^4 e^{2\sqrt{2(1+\log m)}\|\xi\|}\left(\frac{1+\log m}{2}\right)^{2N}\frac{\|\xi\|^{4N}}{\Gamma(N+1)^4}e^{-\|\xi\|^2}.$$

Let us point out that Proposition 2.6 gives the main contribution $\Theta^0_{N,m}$ to $\Delta^{(2)}_{n,m}$. We expect that the main error in the normal approximation should come from the regime where $\|\xi\|$ is not too large. The proof of Proposition 2.6 is given in Section 4. Let us observe that, according to formula (20), we have

$$\left|F_{n,m}(\xi) - e^{-\|\xi\|^2/2}\right|^2 = \left|1 - \det[\mathrm{I} - K_{\mathrm{ig}} Q_n]\right|^2 e^{-\|\xi\|^2},$$

and we expect that if both the degree m and $\|g\|^2_{H^{1/2}} = \|\xi\|^2$ are sufficiently small (depending on the dimension $n \in \mathbb{N}$ of the random matrix $\mathbf{U}$), then, by definition of the projection Q_n, the operator $K_{\mathrm{ig}} Q_n$ is also small (in trace norm) so that $\det[\mathrm{I} - K_{\mathrm{ig}} Q_n] \simeq 1$. This can be quantified by using the bound for Fredholm determinant from [27], Theorem 3.4,

$$\left|1 - \det[\mathrm{I} - K_{\mathrm{ig}} Q_n]\right| \le \|K_{\mathrm{ig}} Q_n\|_{\mathrm{J}_1} e^{1+\|K_{\mathrm{ig}} Q_n\|_{\mathrm{J}_1}}, \tag{29}$$

where $\|\cdot\|_{\mathrm{J}_1}$ denotes the Schatten 1-norm or trace norm of an operator. Then, in order to compute $\|K_{\mathrm{ig}} Q_n\|_{\mathrm{J}_1}$, we use the product structure of the operator K_{ig}, (18) and the Cauchy–Schwarz inequality (for the Hilbert–Schmidt norm $\|\cdot\|_{\mathrm{J}_2}$),

$$\|K_{\mathrm{ig}} Q_n\|_{\mathrm{J}_1} \le \left\|Q_n H_+(e^{2\Im g^+})\right\|_{\mathrm{J}_2} \left\|H_-(e^{-2\Im g^+}) Q_n\right\|_{\mathrm{J}_2}.$$

Moreover, since $H_\pm(\cdot)$ are Hankel operators (17), we can estimate the norms $\|H_\pm(e^{-2\Im g^\pm}) \times Q_n\|_{\mathrm{J}_2}$ by obtaining bounds for the Fourier coefficients of the symbols $e^{-2\Im g^\pm}$; see Lemma 4.1 below. To sum up, we show in Section 4 that $\|Q_n H_\pm(e^{2\Im g^\pm})\|_{\mathrm{J}_2} \ll 1/\Gamma(1+N)$, provided that $\|\xi\| \ll N$, and we use this estimate to deduce Proposition 2.6. Let us emphasize again that we expect that these estimates are of the right order and hold only in the regime where $\|\xi\| \ll N$.

(ii) Tail bound for large $\|\xi\|$. If $\|\xi\|$ is very large, we are not looking to compare $F_{n,m}$ with the characteristic function of a standard Gaussian but rather aiming at obtaining a *good* tail bound for $F_{n,m}$. By *good* we mean that we aim for estimates which yield errors that are smaller than $\Theta^0_{N,m}$ when N is sufficiently large. [22], Proposition 2.13, used the Hadamard's inequality

$$\left|F_{n,m}(\xi)\right|^2 \le \prod_{j=1}^{n}\sum_{i=1}^{n}\left|\left(\widehat{e^{\mathrm{ig}}}\right)_{j-i}\right|^2 \tag{30}$$

and an estimate for the Fourier coefficients of the function e^{ig} to obtain the tail bound $|F_{n,m}(\xi)|^2 \le \frac{C^n n^{\frac{3n}{2}}}{\|\xi\|^{\frac{N}{2}}}$ for a constant $C > 0$. By using (30) and a (sharp) *Van der Corput's inequality*, this estimate can be improved, and we obtain, for all $m, n \ge 3$ and $\xi \in \mathbb{R}^{2m}$,

$$\left|F_{n,m}(\xi)\right|^2 \le \frac{c^n n^n}{\|\xi\|^{\frac{n}{m+1}}}, \qquad c = 4\pi e(1 + 1/\sqrt{3}). \tag{31}$$

To obtain a multidimensional approximation valid for a growing number of traces, we need an estimate which does not involve the large factor n^n. One can obtain a different tail bound by relying on Lemma 2.4 with $h = \mathrm{g}'$. Choosing $\nu > 0$ appropriately, we obtain

$$\left|F_{n,m}(\xi)\right| \le e^{cn}\mathbb{E}_n\left[e^{-\gamma\sum_{j=1}^{n}\mathrm{g}'(\theta_j)^2}\right], \tag{32}$$

for a constant $c > 0$ and $\gamma \to 0$ as $n \to +\infty$; see Proposition 5.1 below for further details. Then, to estimate the RHS of (32), we use that

$$\mathbb{E}_n\big[e^{-\gamma \sum_{j=1}^n \mathrm{g}'(\theta_j)^2}\big] \le \frac{e^n}{\sqrt{2\pi n}}\bigg(\int_{\mathbb{T}} e^{-\gamma \mathrm{g}'(\theta)^2}\frac{d\theta}{2\pi}\bigg)^n, \tag{33}$$

and, since $\mathrm{g}' : \mathbb{T} \to \mathbb{R}$ is a trigonometric polynomial of degree $m \in \mathbb{N}$,

$$\int_{\mathbb{T}} e^{-\gamma \mathrm{g}'(\theta)^2}\frac{\mathrm{d}\theta}{2\pi} \le \frac{2e}{(2\gamma\|\mathrm{g}'\|_{L^2}^2)^{1/4m}}. \tag{34}$$

The estimate (33) is rather classical and its proof is given in the Appendix, Lemma 5.3, for completeness. On the other hand, (34) relies on an estimate of the measure of the set where a trigonometric polynomial is small by its L^2 norm which is taken from [6], Lemma 5.4 below. By combining these estimates, we obtain the following bound in Section 5.

PROPOSITION 2.7. *Fix $n, m \in \mathbb{N}$, and suppose that $N \ge 4m$. For any $\xi \in \mathbb{R}^{2m}$, we have*

$$\big|F_{n,m}(\xi)\big|^2 \le \Upsilon_3(m)^{N/2}\frac{c_{15}{}^{2n} N^{N/4}}{\|\xi\|^{N/2}}.$$

While this tail bound has a worse decay in $\|\xi\|$ than (31), the factor N^N is significantly better than n^n when the degree $m \in \mathbb{N}$ is large. Moreover, we see that this estimate will be useful in the proof of Proposition 2.5 in the regime where $\|\xi\| \gg N^c$ for a sufficiently large constant c. In the proof we will actually choose Λ_3 of order $e^{4c_1\frac{N}{1+\log m}}$.

(iii) Intermediate regime. It remains to deal with the intermediate regime, where $\Lambda_1 \le \|\xi\| \le \Lambda_3$. As we already pointed out, when $\|\xi\| \gg N$, we do not expect that the Fredholm determinant $\det[\mathrm{I} - K_{\mathrm{ig}} Q_n]$ is close to 1. However, we still expect that $|F_{n,m}(\xi)|^2 \ll 1/\Gamma(1+N)^2$ for all such $\xi \in \mathbb{R}^{2m}$. From a technical perspective this intermediate regime is the most challenging one because the direct estimates (e.g., the method used in [22], Section 2.2) lead to errors which are bigger than that of Proposition 2.6; see the estimate (40) below. Our final bounds are summarized in the next proposition. Define

$$\Lambda_2 = \frac{c_0{}^{-1}(1-c_{10})N\sqrt{m+1}}{8(1+\log m)^{3/4} c_{11}}, \tag{35}$$

where c_0, $c_{10}(m)$ and $c_{11}(m)$ are given by (105) and (107). We verify that both $c_{10}(m)$ and $c_{11}(m)$ are decreasing for $m \ge 3$. Since $c_{10}(3) \approx 0.0124$ and $c_{11}(3) \approx 1.583$, this shows that $\Lambda_2 \ge \Lambda_1$ and Λ_2 is increasing as a function of m for all $m \ge 3$.

PROPOSITION 2.8. *Fix $m, n \in \mathbb{N}$ with $m \ge 3$. We have, for all $\xi \in \mathbb{R}^{2m}$,*

$$\big|F_{n,m}(\xi)\big| \le \exp\bigg(c_9 - \frac{c_1(m)N^2}{1+\log m}\bigg) \quad \text{if } \|\xi\| \ge \Lambda_2, \tag{36}$$

and

$$\big|F_{n,m}(\xi)\big| \le \exp\bigg(c_9 - \frac{c_2(m)N^2}{\sqrt{m+1}(1+\log m)^{3/4}}\bigg) \quad \text{if } \Lambda_1 \le \|\xi\| \le \Lambda_2. \tag{37}$$

Let us observe that these bounds directly relate to the error terms $\Theta^1_{N,m}$ and $\Theta^2_{N,m}$ from (113) and (112), respectively. The proof of Proposition 2.8 is given in Section 6. The starting point of this proof is the change of variables and the heuristics described in Section 2.3. Namely, using Lemma 2.4 with $h = -\mathcal{U}\mathrm{g}$, as in (23), we obtain the following bound.

LEMMA 2.9. *Let $n, m \in \mathbb{N}$ and $\xi \in \mathbb{R}^{2m}$. We have, for any $\nu > 0$,*

$$|F_{n,m}(\xi)|^2 \le \mathbb{E}_n\left[\exp\left(\frac{\nu^2}{n^2}\sum_{i,j=1}^n H(\theta_i,\theta_j)\right)\right]\mathbb{E}_n\left[e^{-2\sum_{j=1}^n \Im g(\theta_j + \mathrm{i}\frac{\nu}{n}h(\theta_j))}\right], \tag{38}$$

where the function H is given by (24).

The proof of Lemma 2.9 is postponed to Section 6.2. Using Lemma 2.3, we can easily control the second factor in the RHS of (38). We obtain that there exists a constant $c > 0$ such that if $\|\xi\| \ll \Lambda_2$,

$$\mathbb{E}_n\left[e^{-2\sum_{j=1}^n \Im g(\theta_j + \mathrm{i}\frac{\nu}{n}h(\theta_j))}\right] \le e^{-c\nu\|\xi\|^2}; \tag{39}$$

see Lemma 6.1 below for further details. Moreover, using the deterministic bound $H(\theta_i,\theta_j) \le \|h'\|_\infty^2$ in (38) combined with $\|h'\|_\infty \le \sqrt{2}m\|\xi\|$, this implies that

$$|F_{n,m}(\xi)| \le \exp(-c\nu\|\xi\|^2 + 2\nu^2 m^2\|\xi\|^2).$$

If we optimize over $\nu > 0$, this leads to

$$|F_{n,m}(\xi)| \le \exp\left(-\frac{c^2\|\xi\|^2}{8m^2}\right). \tag{40}$$

In the regime where the degree $m \in \mathbb{N}$ depends on the dimension $n \in \mathbb{N}$ with $N \ge 4m$, the naive estimate (40) is not precise enough to lead to errors which are small compared with that of Proposition 2.6. Therefore, to prove Proposition 2.8, we need to introduce a new idea. One approach would be to use precise rigidity estimates from [23] to obtain a better estimate for the first term on the RHS of (38). But the method that we use consists in writing $\sum_{i,j=1}^n H(\theta_i,\theta_j)$ as a quadratic form in the random variables $\mathrm{T}_k = \operatorname{Tr}\mathbf{U}^k$,

$$\sum_{i,j=1}^n H(\theta_i,\theta_j) = n^2\iint_{[0,2\pi]^2} H(\theta,x)\frac{d\theta}{2\pi}\frac{dx}{2\pi} + n(\mathbf{a}^*\mathbf{T} + \mathbf{T}^*\mathbf{a}) + \mathbf{T}^*\mathbf{M}\mathbf{T}$$
$$\text{where } \mathbf{T} = \begin{pmatrix}\mathrm{T}_1\\ \vdots \\ \mathrm{T}_{2m-1}\end{pmatrix},$$

$\mathbf{a}(\xi) \in \mathbb{C}^{2m}$ is a deterministic vector and $\mathbf{M}(\xi) \in \mathbb{C}^{2m\times 2m}$ is a deterministic matrix which depend on $\xi \in \mathbb{R}^{2m}$; see Lemma 6.3 below. This allows us to express the Laplace transform of the random variable $\sum_{i,j=1}^n H(\theta_i,\theta_j)$ as an integral against a Gaussian measure on $\mathbb{C}^{2m}$. It is not at all clear that the matrix $\mathbf{M}(\xi)$ is positive definite, but *if* it were (see the Remark 6.1) by formulae (25)–(26), we would obtain, for any $\nu > 0$,

$$\mathbb{E}_n\left[e^{\frac{\nu^2}{n^2}\sum_{i,j=1}^n H(\theta_i,\theta_j)}\right] = \frac{e^{\nu^2\|\xi\|^2}}{\pi^{2m}\det(\mathbf{M})}\int_{\mathbb{C}^{2m-1}} e^{-\mathbf{z}^*\mathbf{M}^{-1}\mathbf{z}}\mathbb{E}_n\left[e^{\frac{\nu}{n}(\mathbf{z}^*\mathbf{T}+\mathbf{T}^*\mathbf{z}) + \frac{\nu^2}{n}(\mathbf{a}^*\mathbf{T}+\mathbf{T}^*\mathbf{a})}\right]d\mathbf{z},$$

where $d\mathbf{z}$ is the Lebesgue measure on $\mathbb{C}^{2m-1}$. The idea is now to use Lemma 2.3 to estimate the expectation on the RHS of the previous formula and then to evaluate the Gaussian integral. The details in the implementation of this idea, which requires a modification of $\mathbf{M}$, are somewhat involved and we defer the proof to Section 6.

3. Proof of the main result.

3.1. *Proof of Proposition* 2.5. In this section we give the proof of Proposition 2.5, relying on the estimates from Propositions 2.6, 2.7 and 2.8. Recall that we assume that $m \geq 3$ and $N = \frac{n}{m} > 4m$. Then, using the notation (107) and (115), we define

$$\Lambda_3 = e^{-4c_9/N} {c_{15}}^{4m} \left(\frac{N}{4m} - 1\right)^{2/N} \Upsilon_3(m)\sqrt{N} \exp\left(\frac{4c_1(m)N}{1+\log m}\right). \tag{41}$$

Let us also recall that Λ_1 is given by (28) and Λ_2 is given by (35) so that $\Lambda_3 \gg \Lambda_2 \gg \Lambda_1$, as $n \to +\infty$ (and possibly $m \to +\infty$). We will also need the following bound which is proved in the Appendix (Section A.2).

LEMMA 3.1. *For any $m \in \mathbb{N}$, if $\Lambda > \sqrt{m}$,*

$$\int_{\|\xi\| \geq \Lambda} e^{-\|\xi\|^2} \,\mathrm{d}\xi \leq \Omega_m \frac{m\Lambda^{2m}}{\Lambda^2 - m} e^{-\Lambda^2}.$$

By (28), since $\Lambda_1^2 \geq 2m$ for all $m \geq 3$, it follows from Lemma 3.1 that

$$\int_{\|\xi\| \geq \Lambda_1} e^{-\|\xi\|^2} \,\mathrm{d}\xi \leq \Omega_m \frac{m\Lambda_1^{2m}}{\Lambda_1^2 - m} e^{-\Lambda_1^2} \leq \Omega_m \frac{{c_4}^{2m} N^{2m}}{(1+\log m)^m} \exp\left(-\frac{N^2}{8(1+\log m)}\right). \tag{42}$$

Set $\Lambda_4 = +\infty$, and let us denote for $k = 1, 2, 3$,

$$\mathrm{J}_0^2 = \int_{\substack{\mathbb{R}^{2m} \\ \|\xi\| \leq \Lambda_1}} \left|F_{n,m}(\xi) - e^{-\|\xi\|^2/2}\right|^2 \mathrm{d}\xi, \qquad \mathrm{J}_k^2 = \int_{\substack{\mathbb{R}^{2m} \\ \Lambda_k \leq \|\xi\| \leq \Lambda_{k+1}}} \left|F_{n,m}(\xi)\right|^2 \mathrm{d}\xi.$$

Then, by splitting the integral in (11) and using the estimate (42), we obtain

$$(2\pi)^m \Delta_{n,m}^{(2)} \leq \mathrm{J}_0 + \mathrm{J}_1 + \mathrm{J}_2 + \mathrm{J}_3 + \sqrt{\Omega_m} \frac{{c_4}^m N^m}{(1+\log m)^{\frac{m}{2}}} \exp\left(-\frac{N^2}{16(1+\log m)}\right). \tag{43}$$

As we explain in Section 2.4, we expect that the main contribution in (43) comes from J_0. By Proposition 2.6 we obtain

$$\begin{aligned} \mathrm{J}_0^2 &= \int_{\|\xi\| \leq \Lambda_1} \left|e^{-\|\xi\|^2/2} - F_{n,m}(\xi)\right|^2 \mathrm{d}\xi \\ &\leq \frac{{c_8}^2 m^4 e^{2\sqrt{2(1+\log m)}\Lambda_1} (\frac{1+\log m}{2})^{2N}}{\Gamma(N+1)^4} \int_{\|\xi\| \leq \Lambda_1} \|\xi\|^{4N} e^{-\|\xi\|^2} \,\mathrm{d}\xi. \end{aligned}$$

Moreover, by going to polar coordinates, we have

$$\begin{aligned} \int_{\|\xi\| \leq \Lambda_1} \|\xi\|^{4N} e^{-\|\xi\|^2} \,\mathrm{d}\xi &= m\Omega_m \int_0^{\Lambda_1^2} u^{2N+m-1} e^{-u} \,\mathrm{d}u \\ &\leq m\Omega_m \Gamma(2N+m). \end{aligned}$$

For $3 \leq m \leq N$, we also have

$$\begin{aligned} \frac{\Gamma(2N+m)}{\Gamma(N+1)^2} &\leq e^{-m} \frac{(2N+m)^{2N+m}}{\sqrt{2\pi} N^{2N+1}} = \frac{4^N}{\sqrt{2\pi}} N^{m-1} \left(1 + \frac{m}{2N}\right)^{2N+m} \left(\frac{2}{e}\right)^m \\ &\leq \frac{4^N 2^m e^{\frac{m^2}{2N}}}{\sqrt{2\pi}} N^{m-1}. \end{aligned}$$

Here, we used the upper bound $\Gamma(k) \leq \sqrt{2\pi} k^k e^{-k}$, which holds for all integer $k \geq 9$, and the lower bound $\Gamma(N+1) \geq \sqrt{2\pi N} N^N e^{-N}$ which holds for all $N \in \mathbb{N}$. For the last step we used that $(1+x)^\alpha \le e^{\alpha x}$ for any $x, \alpha \ge 0$. By (28) it holds that $2\sqrt{2(1+\log m)}\Lambda_1 = N$, so we obtain

$$(44)\quad \begin{aligned} \mathrm{J}_0^2 &\le \frac{c_8{}^2}{\sqrt{2\pi}} m^5 \Omega_m 2^m \frac{e^{N+\frac{m^2}{2N}}(1+\log m)^{2N} N^{m-1}}{\Gamma(N+1)^2} \\ &= \big(c_3\sqrt{\Omega_m} N^{\frac{m}{2}} \Theta^0_{N,m}\big)^2, \end{aligned}$$

according to formula (111) and (107).

In the rest of the proof, we give bounds for the integrals J_k for $k = 1, 2, 3$. First, using the tail bound from Proposition 2.7, we have

$$\mathrm{J}_3^2 = \int_{\|\xi\|\ge\Lambda_3} \big|F_{n,m}(\xi)\big|^2 \,\mathrm{d}\xi \le \Upsilon_3(m)^{N/2} c_{15}{}^{2n} N^{N/4} \int_{\|\xi\|\ge\Lambda_3} \|\xi\|^{-N/2}\,\mathrm{d}\xi.$$

Hence, since we assume that $N > 4m$, the previous integral is finite, and we obtain

$$(45)\quad \mathrm{J}_3^2 \le \frac{2m\Omega_m}{N/2-2m} \frac{\Upsilon_3^{N/2} c_{15}{}^{2n} N^{N/4}}{\Lambda_3^{N/2-2m}}.$$

Second, by using the estimate (36) from Proposition 2.8, we also have

$$(46)\quad \mathrm{J}_2^2 = \int_{\Lambda_2\le\|\xi\|\le\Lambda_3} \big|F_{n,m}(\xi)\big|^2\,\mathrm{d}\xi \le \Omega_m \Lambda_3^{2m} \exp\Big(2c_9 - \frac{2c_1N^2}{1+\log m}\Big).$$

Hence, by combining the estimates (45) and (46), this implies that

$$(47)\quad \mathrm{J}_2 + \mathrm{J}_3 \le \sqrt{\Omega_m}\Lambda_3^m\bigg(\exp\Big(c_9 - \frac{c_1N^2}{1+\log m}\Big) + \frac{N^{N/8} c_{15}{}^n \Upsilon_3^{N/4}}{\sqrt{N/4m-1}\Lambda_3^{N/4}}\bigg).$$

Our choice of Λ_3 consists in optimizing[1] the RHS of (47). Namely, by choosing Λ_3 according to (41), we obtain, for all $N > 4m$,

$$\begin{aligned} \mathrm{J}_2 + \mathrm{J}_3 &\le \frac{\sqrt{\Omega_m}}{1-4m/N} \Lambda_3^m \exp\Big(c_9 - \frac{c_1N^2}{1+\log m}\Big) \\ &\le \Big(\frac{N}{4m}\Big)^{\frac{2m}{N}} \frac{e^{c_9(1-\frac{4m}{N})}\sqrt{\Omega_m}}{(1-4m/N)^{1-\frac{2m}{N}}} N^{\frac{m}{2}} c_{15}{}^{4m^2} \Upsilon_3^m \exp\Big(-\frac{c_1N(N-4m)}{1+\log m}\Big). \end{aligned}$$

Then, by using the estimate (116) and that $(\frac{N}{4m})^{\frac{2m}{N}} \le \sqrt{e}$, this implies that

$$\mathrm{J}_2 + \mathrm{J}_3 \le \frac{e^{3/2} e^{c_9(1-\frac{4m}{N})}\sqrt{\Omega_m}}{(1-4m/N)^{1-\frac{2m}{N}}} N^{\frac{m}{2}} c_7{}^m c_{15}{}^{4m^2} m^{\frac{5m}{2}} \exp\Big(-\frac{c_1N(N-4m)}{1+\log m}\Big).$$

According to the notation (107), (108) and (115), since $c_6 = 4\log c_{15}$, we have

$$e^{\Upsilon_1(m)} = e^{3/2} c_7{}^m c_{15}{}^{4m^2} m^{\frac{5m}{2}}.$$

Hence, according to formula (113), we have shown that, for all $N > 4m$,

$$(48)\quad \begin{aligned} \mathrm{J}_2 + \mathrm{J}_3 &\le \frac{e^{c_9(1-\frac{4m}{N})}\sqrt{\Omega_m}}{(1-4m/N)^{1-\frac{2m}{N}}} N^{\frac{m}{2}} \exp\Big(\Upsilon_1(m) - \frac{c_1N(N-4m)}{1+\log m}\Big) \\ &= c_3\sqrt{\Omega_m} N^{\frac{m}{2}} \Theta^1_{N,m}. \end{aligned}$$

[1]If $\alpha > m$, the minimum of the function $\Lambda^m\epsilon + C\Lambda^{m-\alpha}$ over all $\Lambda > 0$ is attained when $\Lambda^\alpha = (\frac{\alpha}{m} - 1)\epsilon^{-1}C$ and equals to $\frac{\epsilon\Lambda^m}{1-m/\alpha}$.

Third, by using the estimate (37) from Proposition 2.8, we also have the bound

$$\mathrm{J}_1^2 = \int_{\Lambda_1 \le \|\xi\| \le \Lambda_2} \left|F_{n,m}(\xi)\right|^2 \mathrm{d}\xi \le \Omega_m \Lambda_2^{2m} \exp\left(2c_9 - \frac{2c_2 N^2}{\sqrt{m+1}(1+\log m)^{3/4}}\right)$$

and, according to (35),

$$\Lambda_2^m \le \sqrt{e}\frac{(8c_0)^{-m} N^m m^{m/2}}{(1+\log m)^{3m/4}},$$

where we used that, by (107), $0 < c_{10}(m) < 1$, $c_{11}(m) \ge 1$ and $(m+1)^m \le \sqrt{e} m^{m/2}$ for all $m \ge 3$. According to (108), this shows that $\Lambda_2^m \le e^{\Upsilon_2(m)}$ so that, for all $m \ge 3$,

$$\begin{aligned} \mathrm{J}_1 &\le e^{c_9}\sqrt{\Omega_m} N^m \exp\left(\Upsilon_2(m) - \frac{c_2 N^2}{\sqrt{m+1}(1+\log m)^{3/4}}\right) \\ &= c_3 \sqrt{\Omega_m} N^{\frac{m}{2}} \Theta_{N,m}^2, \end{aligned} \tag{49}$$

according to formula (112).

Finally, by collecting the estimates (44), (48) and (49), we deduce from the decomposition (43) that, for any $m \ge 3$ and $N > 4m$,

$$\Delta_{n,m}^{(2)} \le \frac{c_3\sqrt{\Omega_m} N^{\frac{m}{2}}}{(2\pi)^m}\left(\Theta_{N,m}^0 + \Theta_{N,m}^1 + \Theta_{N,m}^2 + \frac{{c_3}^{-1} {c_4}^m N^{\frac{m}{2}}}{(1+\log m)^{\frac{m}{2}}} \exp\left(-\frac{N^2}{16(1+\log m)}\right)\right).$$

Replacing the last term by $\Theta_{N,m}^3$ according to (111) and using that $c_3 \le 8$, it completes the proof.

3.2. *Proof of Theorem* 1.3. The estimate (8) is just Proposition 2.5. Then, to prove the estimate (9), we need the following Gaussian tail bound which is a straightforward consequence of Lemma 2.3.

LEMMA 3.2 (Large deviation estimates). *For any $L > 0$, let $\square_L = [-\frac{L}{2}, \frac{L}{2}]^{2m}$. Then, we have, for any $n, m \in \mathbb{Z}_+$,*

$$\mathbb{P}_n[\mathbf{X} \notin \square_L] \le 4m e^{-L^2/8}$$

and

$$\int_{\mathbb{R}^{2m}\setminus\square_L} \frac{e^{-\|x\|^2/2}}{(2\pi)^m} \mathrm{d}x \le \frac{8m}{\sqrt{2\pi} L} e^{-L^2/8}.$$

PROOF. For any $k \ge 1$, it follows from Lemma 2.3 that, for any $t \in \mathbb{R}$,

$$\mathbb{E}_n\left[e^{t\mathrm{X}_{2k}}\right], \mathbb{E}_n\left[e^{t\mathrm{X}_{2k-1}}\right] \le e^{t^2/2}.$$

By Markov inequality, this implies that, for any $k \ge 1$ and $t > 0$,

$$\mathbb{P}_n\left[|\mathrm{X}_k| \ge L\right] \le 2e^{-tL + t^2/2}.$$

Choosing $t = L$, we obtain

$$\mathbb{P}_n\left[|\mathrm{X}_k| \ge L\right] \le 2e^{-L^2/2}.$$

Hence, by a union bound we obtain

$$\mathbb{P}_n\left[(\mathrm{X}_1, \dots, \mathrm{X}_{2m}) \notin \square_L\right] \le \sum_{k \le 2m} \mathbb{P}_n\left[|\mathrm{X}_k| \ge L/2\right] \le 4m e^{-L^2/8}.$$

By a similar union bound an analogous estimate holds in the Gaussian case. □

Recall the definitions (6), and let us split

$$\Delta^{(1)}_{n,m} = \left(\int_{\square_L} + \int_{\mathbb{R}^{2m}\setminus\square_L}\right)\left|\wp_{n,m}(x) - \frac{e^{-\|x\|^2/2}}{(2\pi)^m}\right| \mathrm{d}x$$
$$\leq L^m \Delta^{(2)}_{n,m} + \int_{\mathbb{R}^{2m}\setminus\square_L}\left(\wp_{n,m}(x) + \frac{e^{-\|x\|^2/2}}{(2\pi)^m}\right) \mathrm{d}x,$$

where we used the Cauchy–Schwarz inequality to bound the first integral. By Lemma 3.2 this implies that, for any $L \geq 2\sqrt{3}$,

$$\Delta^{(1)}_{n,m} \leq L^m \Delta^{(2)}_{n,m} + 5me^{-L^2/8}, \tag{50}$$

We choose the parameter L, which minimizes the RHS of (50), that is the (unique) solution of the equation

$$\Delta^{(2)}_{n,m} = \frac{5}{4} L^{2-m} e^{-L^2/8}. \tag{51}$$

Since $m \geq 3$, the function $L \mapsto L^{2-m}e^{-L^2/8}$ is decreasing, and it is bounded from below by $2^{2-m}m^{1-\frac{m}{2}}e^{-\frac{m}{2}}$ for $L \leq 2\sqrt{m}$, under the assumption that $\Delta^{(2)}_{n,m} \leq 5 \cdot 2^{-m}m^{1-\frac{m}{2}}e^{-\frac{m}{2}}$, the solution of equation (51) satisfies

$$2\sqrt{m} \leq L \leq \sqrt{8 \log \Delta^{(2)-1}_{n,m}}. \tag{52}$$

Hence, by (50) this implies that if $\Theta_{N,m} \leq m^{\frac{5}{4}}(\frac{e}{\sqrt{\pi}})^{-m}N^{-\frac{m}{2}}$ (which implies that $\Delta^{(2)}_{n,m} \leq 5 \cdot 2^{-m}m^{1-\frac{m}{2}}e^{-\frac{m}{2}}$ by (8)), then

$$\Delta^{(1)}_{n,m} \leq L^m\left(1 + \frac{4m}{L^2}\right)\Delta^{(2)}_{n,m}.$$

Finally, using the conditions (52) for L, we conclude that

$$\Delta^{(1)}_{n,m} \leq 2\big(8 \log \Delta^{(2)-1}_{n,m}\big)^{\frac{m}{2}} \Delta^{(2)}_{n,m}.$$

This completes the proof.

3.3. *Proof of Theorem* 1.1, *Proposition* 1.2 *and Proposition* 1.6. Throughout the proof we fix an integer $M \geq 3$. Using the estimates of Proposition B.3, it holds that, for $m \geq M$ and any $N \geq c(M)m\sqrt{1+\log m}$,

$$\Theta_{n,m} \leq (1+\epsilon)\Theta^0_{N,m} \leq N^{-\frac{m}{2}} \frac{\exp(-12m(\log m - 0.16))}{{c_{16}}^m}.$$

This shows that, under the hypothesis of Theorem 1.1, the condition $\Theta_{N,m} \leq m^{\frac{5}{4}}{c_{16}}^{-m}N^{-\frac{m}{2}}$ holds, and we can apply Theorem 1.3,

$$\Delta^{(1)}_{n,m} \leq 2\big(8 \log \Delta^{(2)-1}_{n,m}\big)^{\frac{m}{2}} \Delta^{(2)}_{n,m}$$
$$\leq \frac{2c_3\sqrt{\Omega_m}}{(2\pi)^m}\left(8N \log\left(\frac{2c_3\sqrt{\Omega_m}}{(2\pi)^m} N^{\frac{m}{2}}\Theta_{N,m}\right)^{-1}\right)^{\frac{m}{2}}\Theta_{N,m},$$

where we used that $\Delta^{(2)}_{n,m} < e^{-\frac{m}{2}}$ and the function $x \mapsto x(\log x^{-1})^{\frac{m}{2}}$ is nondecreasing for $x \in [0, e^{-\frac{m}{2}}]$. By (120) we also have $\frac{\Omega_m}{(2\pi)^{2m}} N^m \geq \frac{e^m (N/m)^m}{3(4\pi)^m \sqrt{m}}$ so that, according to formula (111), we obtain the (crude) lower bound

$$2c_3 \frac{\sqrt{\Omega_m}}{(2\pi)^m} N^{\frac{m}{2}} \Theta^0_{N,m} \geq N^{-N}.$$

This implies that if $m \geq M$ and $N \geq c(M) m \sqrt{1 + \log m}$, then

$$\Delta^{(1)}_{n,m} \leq \frac{2c_3 \sqrt{\Omega_m}}{(2\pi)^m} (N\sqrt{8 \log N})^m \Theta_{N,m}. \tag{53}$$

Using Proposition B.3 again and that $2c_3(1+\epsilon) \leq 16$, we obtain

$$\Delta^{(1)}_{n,m} \leq 16 \frac{\sqrt{\Omega_m}}{(2\pi)^m} (N\sqrt{8 \log N})^m \Theta^0_{N,m}. \tag{54}$$

By (7), (111) and replacing $\Omega_m = \frac{\pi^m}{m!}$, this completes the proof of Theorem 1.1. The condition $n \geq 1911$ follows from choosing $M = 3$ and $c(3) = 146.5$, according to the Table (125).

Now, let us choose $m = \lfloor n^\alpha \rfloor$ with $0 < \alpha < 1/2$, and let us assume that $m \geq 17$. By (54) or the estimate from Theorem 1.1, this implies that, for all integer $n \in \mathbb{N}$ such that $n^{1-2\alpha} \geq 20.4\sqrt{\log n}$,

$$\Delta^{(1)}_{n,m} \leq 16 \frac{2^m}{\pi^{m/2}\sqrt{m!}} n^{3\alpha - 1} e^{\frac{n^{3\alpha-1}}{4}} (N\sqrt{\log N})^{n^\alpha} \frac{e^{\frac{N}{2}} (\log N)^N}{\Gamma(N+1)}.$$

Here, we have used that $N \geq e^2 m$ and that $c(17) = 28.8 \leq 20.4\sqrt{2}$; see the Table (125). First, we have numerically, for all $m \in \mathbb{N}$,

$$\frac{16 \cdot 2^m}{\pi^{\frac{m}{2}} \sqrt{2\pi m!}} \leq 1.4 \cdot 10^{-13}. \tag{55}$$

Second, let us observe that the function $\frac{e^{\frac{N}{2}} (\log N)^N}{\Gamma(N+1)}$ is decreasing for $N \geq e^2$ so that by (120),

$$\frac{e^{\frac{N}{2}} (\log N)^N}{\Gamma(N+1)} \leq \frac{n^{\frac{\alpha-1}{2}}}{\sqrt{2\pi}} \exp\left(-n^{1-\alpha} \log(n^{1-\alpha})\left(1 - \frac{\log(\log\sqrt{n}) + 3/2}{\log(n^{1-\alpha})}\right)\right),$$

and these estimates imply that

$$\begin{aligned}\Delta^{(1)}_{n,m} &\leq 1.4 \cdot 10^{-13} n^{3\alpha - \frac{3}{2}} e^{\frac{n^{3\alpha-1}}{4}} (N\sqrt{\log N})^{n^\alpha} \\ &\quad \times \exp\left(-n^{1-\alpha} \log(n^{1-\alpha})\left(1 - \frac{\log(\log\sqrt{n}) + 3/2}{\log(n^{1-\alpha})}\right)\right).\end{aligned}$$

Now, let us also observe that $N \leq e^{0.0572} n^{1-\alpha}$ for $n \geq 17^{\alpha^{-1}}$ so that

$$(N\sqrt{\log N})^{n^\alpha} \leq \exp\left(n^\alpha \log(n^{1-\alpha})\left(1 + \frac{\log\log n + 0.1144}{2\log(n^{1-\alpha})}\right)\right).$$

This shows that if $n \geq 17^{\alpha^{-1}}$ (so that $m \geq 17$) and $n^{1-2\alpha} \geq 20.4\sqrt{\log n}$,

$$\Delta^{(1)}_{n,m} \leq 1.4 \cdot 10^{-13} n^{3\alpha - \frac{3}{2}} \exp(-(1-\epsilon_n) n^{1-\alpha} \log(n^{1-\alpha}))$$

with

$$\begin{aligned}\epsilon_n(\alpha) &:= \frac{\log\log(\sqrt{n}) + 3/2}{\log(n^{1-\alpha})} + n^{-(1-2\alpha)}\left(1 + \frac{\log\log n + 0.1144}{2\log(n^{1-\alpha})}\right) + \frac{n^{-2(1-2\alpha)}}{4\log(n^{1-\alpha})} \\ &\leq \frac{2(\log\log n + 0.8069)}{\log n} + \frac{0.065}{\sqrt{\log n}} + \frac{0.0012}{(\log n)^2}.\end{aligned} \tag{56}$$

Note that $\epsilon_n(\alpha)$ is increasing for α for $0 < \alpha < 1/2$ and to obtain the estimate (56), we have used that $n^{1-2\alpha} \geq 20.4\sqrt{\log n}$ and the numerical bound $1 + \frac{\log\log n + 0.1144}{\log n} \leq 20.4 \cdot 0.065$ for all $n \geq 17^2$ to obtain the estimate (56). We also deduce from (56) that $\epsilon_n \leq 1 - 75 \cdot 10^{-3}$ for all $n \geq 17^2$. Since $\mathrm{d_{TV}}(\mathbf{X}, \mathbf{G}) \leq \Delta^{(1)}_{n,m}$, this completes the proof of Proposition 1.2.

We now turn to the proof of Proposition 1.6. Let us choose $m = \lfloor \sqrt{\frac{n}{41.5\sqrt{\log n}}} \rfloor$ and suppose that $n \geq 4322$ so that $m \geq 6$, and we can use the estimate (54); we have $c(6) = 58.66 \leq 41.5\sqrt{2}$, according to the Table (125). As $N = \frac{n}{m} \geq e^2 m$, this implies that

$$\Delta^{(1)}_{n,m} \leq 16 m^{\frac{7}{2}} \frac{2^m}{\pi^{\frac{m}{2}}\sqrt{m!}} e^{\frac{m^3}{4n}} (N\sqrt{\log N})^m \frac{e^{\frac{3}{2}N}(\log N)^N}{\sqrt{2\pi} n N^N}.$$

Moreover, we verify that, as $N \geq \sqrt{41.5 n\sqrt{\log n}}$,

$$(\log N)^N N^{m-N} e^{\frac{3}{2}N} \leq \exp\biggl(-\frac{\sqrt{41.5}}{2}\sqrt{n}(\log n)^{5/4}\biggl(1 - \frac{1}{41.5\sqrt{\log n}} - \frac{2\log(\log\sqrt{n}) + 3}{\log n}\biggr)\biggr)$$

and, using the estimate (55), this shows that

$$\Delta^{(1)}_{n,m} \leq 1.4 \cdot 10^{-13} \frac{m^{\frac{7}{2}}}{n} e^{\frac{m^3}{4n}} (\log N)^{\frac{m}{2}} \\ \times \exp\biggl(-\frac{\sqrt{41.5}}{2}\sqrt{n}(\log n)^{5/4}\biggl(1 - \frac{1}{41.5\sqrt{\log n}} - \frac{2\log(\log\sqrt{n}) + 3}{\log n}\biggr)\biggr).$$

Moreover, since $(\log N)^{\frac{m}{2}} e^{\frac{m^3}{4n}} \leq \exp(\frac{\sqrt{n}}{4(41.5\sqrt{\log n})^{3/2}} + \frac{1}{2}\sqrt{\frac{n}{41.5\sqrt{\log n}}}\log\log n)$, we obtain

$$\Delta^{(1)}_{n,m} \leq 1.4 \cdot 10^{-13} \frac{n^{3/4}}{(\log n)^{\frac{7}{8}}(41.5)^{\frac{7}{4}}} \exp\biggl(-\frac{\sqrt{41.5}}{2}\sqrt{n}(\log n)^{5/4}(1 - \epsilon_n)\biggr),$$

where

$$\epsilon_n = \frac{1}{41.5\sqrt{\log n}} + \frac{3 - 2\log 2 + 2\log\log n}{\log n} + \frac{\log\log n}{41.5(\log n)^{\frac{3}{2}}} + \frac{1/2}{(41.5\log n)^2}.$$

We verify numerically that $\epsilon_n \leq 0.711$ for all $n \geq 4322$. In particular, this implies that

$$\Delta^{(1)}_{n,m} \leq n^{3/4}\exp(-38 - 0.93\sqrt{n}(\log n)^{5/4}).$$

Since $\Delta^{(1)}_{n,m}$ is nondecreasing in $m \in \mathbb{N}$, this completes the proof of Proposition 1.6.

4. Gaussian approximation: Proof of Proposition 2.6. Recall that $F_{n,m}$ denotes the characteristic function of the random vector $\mathbf{X}$ and that it is given by formula (20). In particular, it holds, for any $\xi \in \mathbb{R}^{2m}$,

$$\bigl|e^{-\|\xi\|^2/2} - F_{n,m}(\xi)\bigr|^2 = e^{-\|\xi\|^2}\bigl|1 - \det[\mathrm{I} - K_{\mathrm{ig}} Q_n]\bigr|^2, \tag{57}$$

where g is a trigonometric polynomial (12), Q_n is the orthogonal projection with kernel $\mathrm{span}(e_1, \dots, e_{n-1})$ and, according to formula (18),

$$K_{\mathrm{ig}} = H_+(e^{2\Im \mathrm{g}^+}) H_-(e^{2\Im \mathrm{g}^-}). \tag{58}$$

Recall that the operator K_{ig} is trace class, but observe that it is not selfadjoint, since, by (17), $K^*_{\mathrm{ig}} = H_+(e^{2\Im \mathrm{g}^-}) H_-(e^{2\Im \mathrm{g}^+})$ with $\Im \mathrm{g}^- = -\Im \mathrm{g}^+$ because the function g is real valued. As we explained in Section 2.4, in order to prove Proposition 2.6 we provide estimates for the

Fredholm determinant on the RHS of (57) in the regime where $\|\xi\| \ll N$ in order to guarantee that the Schatten norm $\|K_{\mathrm{ig}} Q_n\|_{\mathrm{J}_1}$ remains small.

The first step of the proof consists in obtaining a priori estimates on Fourier coefficients of the functions $e^{2\Im \mathrm{g}^{\pm}}$.

LEMMA 4.1. *Fix $m \in \mathbb{N}$ and $\xi \in \mathbb{R}^{2m}$. Let $\rho = \sqrt{\frac{1+\log m}{2}}\|\xi\|$. We have, for all integers $k > 2m\rho$,*

$$|(\widehat{e^{\pm 2\Im \mathrm{g}^{+}}})_k| \le 2e^{\rho} \frac{\rho^{\lceil k/m \rceil}}{\lceil k/m \rceil !}.$$

PROOF. Let us define $\phi_M(w) = \sum_{k=0}^{M} \frac{w^k}{k!}$ for $M \ge 1$. Since $\mathrm{g}^{+}(\theta) = \sum_{k=1}^{m} \frac{\zeta_k}{\sqrt{2k}} e^{\mathrm{i}k\theta}$ and $\mathrm{g}^{-} = \overline{\mathrm{g}^{+}}$, we have, for all integers $k > Mm$,

$$\int_{\mathbb{T}} \phi_M(-\mathrm{i}\mathrm{g}^{+}(\theta)) e^{\mathrm{i}\mathrm{g}^{-}(\theta) - \mathrm{i}k\theta} \frac{\mathrm{d}\theta}{2\pi} = 0.$$

This implies that, any $k > Mm$,

$$\begin{aligned} |(\widehat{e^{2\Im \mathrm{g}^{+}}})_k| &= \left| \int_{\mathbb{T}} e^{-\mathrm{i}\mathrm{g}^{+}(\theta) + \mathrm{i}\mathrm{g}^{-}(\theta) - \mathrm{i}k\theta} \frac{\mathrm{d}\theta}{2\pi} \right| \\ &\le \int_{\mathbb{T}} |e^{-\mathrm{i}\mathrm{g}^{+}(\theta)} - \phi_M(-\mathrm{i}\mathrm{g}^{+}(\theta))| e^{-\Im \mathrm{g}^{-}(\theta)} \frac{\mathrm{d}\theta}{2\pi}. \end{aligned} \tag{59}$$

Now, let us observe that, for any $|w| \le M/2$,

$$\begin{aligned} |e^{w} - \phi_M(w)| &\le \frac{|w|^{M+1}}{(M+1)!} \sum_{j \ge 0} \left(\frac{|w|}{M+2} \right)^{j} \\ &\le 2 \frac{|w|^{M+1}}{(M+1)!}. \end{aligned} \tag{60}$$

Moreover, by (12) and since $\sum_{k=1}^{m} |\zeta_k|^2 = \|\xi\|^2$, we also have

$$\|\mathrm{g}^{+}\|_{\infty} \le \sum_{k=1}^{m} \frac{|\zeta_k|}{\sqrt{2k}} \le \rho = \sqrt{\frac{1+\log m}{2}} \|\xi\|, \tag{61}$$

where we used that $\sum_{k=1}^{m} k^{-1} \le 1 + \log m$ for any $m \ge 1$ and the Cauchy–Schwarz inequality. Then, using the estimates (59), (60) and (61), we obtain, if both $M \ge 2\rho$ and $k > Mm$,

$$|(\widehat{e^{2\Im \mathrm{g}^{+}}})_k| \le 2e^{\rho} \frac{\rho^{M+1}}{(M+1)!}.$$

By choosing $M = \lfloor k/m \rfloor$, this implies the claim. Indeed, by the same argument we obtain the same bound for $|(\widehat{e^{-2\Im \mathrm{g}^{+}}})_k|$. $\square$

Now, let us use these estimates to bound the Schatten norm $\|K_{\mathrm{ig}} Q_n\|_{\mathrm{J}_1}$.

LEMMA 4.2. *Fix $m \in \mathbb{N}$, $\xi \in \mathbb{R}^{2m}$, and let $\rho = \sqrt{\frac{1+\log m}{2}}\|\xi\|$. If we assume that $N = \frac{n}{m} \ge {c_*}^{-1}\rho$ with $c_* < \frac{1}{2}$, then*

$$\|K_{\mathrm{ig}} Q_n\|_{\mathrm{J}_1} \le \frac{4m^2 e^{2\rho}}{(1-{c_*}^2)^2} \frac{\rho^{2N}}{\Gamma(N+1)^2}.$$

PROOF. Let us recall that the operators $H_\pm(e^{2\Im g^\pm})$ are Hilbert–Schmidt and that by (17) we have

$$\|Q_n H_\pm(e^{2\Im g^\pm})\|_{J_2}^2 \le \sum_{k\ge n}(k-n+1)\big|\big(\widehat{e^{2\Im g^\pm}}\big)_k\big|^2.$$

Moreover, by formula (58) and the Cauchy–Schwarz inequality (for the Schatten norms), since Q_n is a projection, we have

$$\|K_{\mathrm{ig}} Q_n\|_{J_1} \le \|H_+(e^{2\Im g^+})Q_n\|_{J_2}\|H_-(e^{2\Im g^-})Q_n\|_{J_2}.$$

Using the estimates from Lemma 4.1, this implies that if the dimension $n > 2m\rho$, then

$$\|K_{\mathrm{ig}} Q_n\|_{J_1} \le 4e^{2\rho}\sum_{k\ge n}(k-n+1)\frac{\rho^{2\lceil k/m\rceil}}{(\lceil k/m\rceil !)^2}. \tag{62}$$

Under the condition $N = \frac{n}{m} \ge {c_*}^{-1}\rho$, since $j! \ge \Gamma(N+1)N^{j-N}$ for all $j \ge N$, we obtain

$$\begin{aligned}\sum_{k\ge n}(k-n+1)\frac{\rho^{2\lceil k/m\rceil}}{(\lceil k/m\rceil !)^2} &\le m^2\sum_{j\ge N}(j+1-N)\frac{\rho^{2j}}{(j!)^2}\\ &\le m^2\frac{\rho^{2N}}{\Gamma(N+1)^2}\sum_{j\ge 0}(j+1)\Big(\frac{\rho}{N}\Big)^{2j}\\ &\le (1-{c_*}^2)^{-2}m^2\frac{\rho^{2N}}{\Gamma(N+1)^2}.\end{aligned}$$

Note that for the last bound, it suffices that $c_* < 1$. However, we impose that $c_* < \frac{1}{2}$ to guarantee that $n > 2m\rho$. Then, by combining the previous estimate with (62), this completes the proof. □

We are now ready to finish the proof of Proposition 2.6. First, let us observe that by Lemma 4.2 and using formula (120) for the Γ function, we obtain that, if $N \ge {c_*}^{-1}\rho$,

$$\|K_{\mathrm{ig}} Q_n\|_{J_1} \le \frac{2/\pi}{(1-{c_*}^2)^2}m^2 e^{2\rho}\frac{(\rho e)^{2N}}{N^{2N+1}} \le \frac{2/\pi}{(1-{c_*}^2)^2}\frac{m^2}{N}\big(c_* e^{c_*+1}\big)^{2N}. \tag{63}$$

If we choose $c_* = 1/4$, then $c_* e^{c_*+1} \le 0.873$ so that the RHS of (63) is very small for large N. Actually, in the regime where $N > 4m$ (in particular when $N \ge 13$), this implies that

$$\|K_{\mathrm{ig}} Q_n\|_{J_1} \le \frac{32}{225\cdot\pi}N\big(c_* e^{c_*+1}\big)^{2N} \le \frac{416}{225\cdot\pi}(0.873)^{26} \le \log(2.766) - 1,$$

where we obtained the last two bounds numerically. Hence, using the inequality (29) from [27], Theorem 3.4, we deduce from Lemma 4.2 with $c_* = 1/4$ and the previous estimate that if $N \ge 4(\rho \vee m)$,

$$\big|1 - \det[\mathrm{I} - K_{\mathrm{ig}} Q_n]\big|^2 \le \|K_{\mathrm{ig}} Q_n\|_{J_1}^2 e^{2(1+\|K_{\mathrm{ig}}Q_n\|_{J_1})} \le {c_8}^2 m^4 e^{4\rho}\frac{\rho^{4N}}{\Gamma(N+1)^4},$$

where $c_8 = 2.766\frac{4}{(1-{c_*}^2)^2}$, according to (107). If we combine this estimate with formula (57) and replace $\rho = \sqrt{(1+\log m)/2}\|\xi\|$, this implies that, for any $N \ge 4m$ and all $\|\xi\| \le \Lambda_1 = \frac{N}{4\sqrt{(1+\log m)/2}}$,

$$\big|e^{-\|\xi\|^2/2} - F_{n,m}(\xi)\big|^2 \le {c_8}^2 m^4 e^{4\rho}\frac{\rho^{4N}}{\Gamma(N+1)^4}e^{-\|\xi\|^2}.$$

This completes the proof.

5. Tail bound for large $\|\xi\|$: Proof of Proposition 2.7. Recall that the function g is given by (12), and let us observe that by choosing $h = \mathrm{g}'$ in Lemma 2.4, we obtain the following bound.

PROPOSITION 5.1. *Fix $m, n \in \mathbb{N}$, and let $N = \frac{n}{m}$. For any $\eta > 0$ and any $\xi \in \mathbb{R}^{2m}$, we have*

$$|F_{n,m}(\xi)| \le \exp\left(c_{20}\left(n + \frac{2}{\pi^2}\right)\right)\mathbb{E}_n\left[e^{-\gamma \sum_{j=1}^n \mathrm{g}'(\theta_j)^2}\right],$$

where $\gamma = \frac{\eta}{\sqrt{n}m(m+1)\|\xi\|}(1 - \frac{\eta^2 c_{21}}{n})$ and $c_{20} = \frac{\pi^2\eta^2}{8}$.

In order to prove Proposition 5.1, we need the following basic estimate which is proved in Appendix A.3.

LEMMA 5.2. *For any $y \in [-1, 1]$, $y \neq 0$ and $x \in \mathbb{R}$, we have*

$$1 + \left(\frac{\sinh(x)}{y}\right)^2 \le \exp\left(\frac{x}{y}\right)^2.$$

PROOF. We apply Lemma 2.4 with $h = \mathrm{g}'$ and $\nu = \frac{\eta\sqrt{n}}{m(m+1)\|\xi\|}$ where $\eta > 0$. We obtain

$$|F_{n,m}(\xi)| \le \mathbb{E}_n\left[\prod_{i<j}\left|\frac{\sin(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{\mathrm{g}'(\theta_i)-\mathrm{g}'(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right|^2 \prod_{j=1}^n\left|1 + \mathrm{i}\frac{\nu}{n}\mathrm{g}''(\theta_j)\right| e^{-\Im \mathrm{g}(\theta_j + \mathrm{i}\frac{\nu}{n}\mathrm{g}'(\theta_j))}\right]. \tag{64}$$

Moreover, by the Cauchy–Schwarz, we have

$$\|\mathrm{g}'\|_\infty \le \sum_{k=1}^m \sqrt{2k}|\zeta_k| \le \sqrt{2\sum_{k=1}^m |\zeta_k|^2 \sum_{k=1}^m k} = \sqrt{m(m+1)}\|\xi\|. \tag{65}$$

Observe that, with these choices, we have $\frac{\nu}{n}\|\mathrm{g}'\|_\infty \le \frac{\eta/m}{\sqrt{n(1+1/m)}}$ so that by Taylor's theorem, since g is real valued, we have, for $j \in \{1, \dots, n\}$,

$$\left|\Im \mathrm{g}\left(\theta_j + \mathrm{i}\frac{\nu}{n}\mathrm{g}'(\theta_j)\right) - \frac{\nu}{n}\mathrm{g}'(\theta_j)^2\right| \le \frac{1}{6}\left|\frac{\nu}{n}\mathrm{g}'(\theta_j)\right|^3 \sup_{|\Re z| \le \pi, |\Im z| \le \frac{\eta/m}{\sqrt{n(1+1/m)}}} |\mathrm{g}'''(z)|.$$

We also have $|\mathrm{g}'''(z)| \le \sum_{|k| \le m} |\zeta_k| \frac{|k|^{5/2}}{\sqrt{2}} e^{k|\Im z|}$ so that

$$\sup_{|\Re z| \le \pi, |\Im z| \le \frac{\eta/m}{\sqrt{n(1+1/m)}}} |\mathrm{g}'''(z)| \le 6c_{21}\big(m(m+1)\big)^{3/2}\|\xi\|, \quad \text{where } c_{21} = \frac{\exp(\eta/\sqrt{n(1+1/m)})}{6\sqrt{3}}$$

and we used that $\sum_{k=1}^m k^5 \le \frac{m^3(m+1)^3}{6}$. Then, using the estimate (65), the previous bounds imply that

$$\begin{aligned}\left|\Im \mathrm{g}\left(\theta_j + \mathrm{i}\frac{\nu}{n}\mathrm{g}'(\theta_j)\right) - \frac{\nu}{n}\mathrm{g}'(\theta_j)^2\right| &\le c_{21}\frac{\nu^3 m^2(m+1)^2\|\xi\|^2}{n^3}\mathrm{g}'(\theta_j)^2 \\ &= \frac{\nu}{n}\frac{\eta^2 c_{21}}{n}\mathrm{g}'(\theta_j)^2,\end{aligned}$$

where we used our choice for ν. This shows that

$$\prod_{j=1}^n e^{-\Im \mathrm{g}(\theta_j + \mathrm{i}\frac{\nu}{n}\mathrm{g}'(\theta_j))} \le \exp\left(-\frac{\nu}{n}\left(1 - \frac{\eta^2 c_{21}}{n}\right)\sum_{j=1}^n g'(\theta_j)^2\right). \tag{66}$$

Moreover, by Lemma 5.2 and since by convexity $\sin(u/2) \geq u/\pi$ for all $u \in [0, \pi]$, we obtain, for any $u \in [-\pi, \pi]$ and $\alpha > 0$,

$$1 + \left(\frac{\sinh(\alpha u/2)}{\sin(u/2)}\right)^2 \leq \exp\left(\frac{\alpha u}{2\sin(u/2)}\right)^2 \leq \exp\left(\frac{\pi\alpha}{2}\right)^2.$$

This estimate implies that, for all $i, j \in \{1, \dots, n\}$,

$$\begin{aligned}\left|\frac{\sin(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{g'(\theta_i)-g'(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right|^2 &= 1 + \left(\frac{\sinh(\nu\frac{g'(\theta_i)-g'(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right)^2 \\ &\leq 1 + \left(\frac{\sinh(\frac{\nu\|g''\|_\infty}{n}\frac{(\theta_i-\theta_j)}{2})}{\sin(\frac{\theta_i-\theta_j}{2})}\right)^2 \\ &\leq \exp\left(\frac{\nu\pi\|g''\|_\infty}{2n}\right)^2.\end{aligned}$$

Moreover, by the Cauchy–Schwarz inequality, we have

$$\|g''\|_\infty \leq \sum_{|k|\leq m}\frac{k^{3/2}}{\sqrt{2}}|\zeta_k| \leq \sqrt{\sum_{k=1}^m k^3 \sum_{|k|\leq m}|\zeta_k|^2} = \frac{m(m+1)}{\sqrt{2}}\|\xi\|.$$

Hence, this shows that

$$\prod_{1\leq i<j\leq n}\left|\frac{\sin(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{g'(\theta_i)-g'(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right|^2 \leq \exp\left(\frac{\nu m(m+1)\|\xi\|}{2\sqrt{2}/\pi}\right)^2 = e^{c_{20}n}, \tag{67}$$

where we used the definition of ν and set $c_{20} = \frac{\pi^2\eta^2}{8}$. Similarly, we have

$$\begin{aligned}\prod_{j=1}^n\left|1 + \mathrm{i}\frac{\nu}{n}g''(\theta_j)\right| &\leq \left(1 + \frac{\nu^2\|g''\|_\infty^2}{n^2}\right)^{n/2} \\ &\leq \exp\left(\frac{(\nu m(m+1)\|\xi\|)^2}{4n}\right) = \exp\left(\frac{2c_{20}}{\pi^2}\right),\end{aligned} \tag{68}$$

where we used that $1 + x \leq e^x$ for all $x \in \mathbb{R}$ to obtain the second estimate. By combining the estimates (66), (67), (68) with (64), we obtain that, for all $\xi \in \mathbb{R}^{2m}$,

$$|F_{n,m}(\xi)| \leq \mathbb{E}_n\left[\exp\left(c_{20}\left(n + \frac{2}{\pi^2}\right) - \gamma\sum_{j=1}^n g'(\theta_j)^2\right)\right],$$

where $\gamma = \frac{\nu}{n}(1 - \frac{\eta^2 c_{21}}{n}) = \frac{\eta}{\sqrt{n}m(m+1)\|\xi\|}(1 - \eta^2\frac{\exp(\frac{\eta}{\sqrt{n(1+1/m)}})}{6\sqrt{3}n})$. $\square$

Thus, in order to estimate $|F_{n,m}(\xi)|$ using Proposition 5.1, we need a bound for $\mathbb{E}_n[e^{-\gamma\sum_{j=1}^n g'(\theta_j)^2}]$. Let us point out that in the regime where $\|\xi\|$ is large, compared with N, we cannot use the bound from Lemma 2.3 to estimate this quantity. Indeed, we have $\|g'\|_{L^2}^2 \geq \|\xi\|^2$, while our basic estimate for $\mathcal{A}(g'^2)$ is of the form $\mathcal{A}(g'^2) \leq cm^5\|\xi\|^4$ for a numerical constant $c > 0$. Then, by optimizing over all $\gamma > 0$, we would obtain

$$\mathbb{E}_n[e^{-\gamma\sum_{j=1}^n g'(\theta_j)^2}] \leq \exp(-\gamma n\|g'\|_{L^2}^2 + \gamma^2\mathcal{A}(g'^2)) \leq \exp\left(-\frac{N^2}{4cm^3}\right).$$

This estimate is similar to those from Proposition 2.8, but it is not as good for large $m \in \mathbb{N}$. More importantly, it does not yield any decay as $\|\xi\| \to +\infty$. So, instead of Lemma 2.3, we will use the bound (33) which follows from the next lemma.

LEMMA 5.3. *For any function $f:\mathbb{T}\to\mathbb{R}$ such that e^{-f} is integrable, we have, for any $n\geq 2$,*

$$\mathbb{E}_n\big[e^{-\sum_{j=1}^n f(\theta_j)}\big]\leq \frac{e^n}{\sqrt{2\pi n}}\bigg(\int_{\mathbb{T}} e^{-f(\theta)}\frac{d\theta}{2\pi}\bigg)^n. \tag{69}$$

The proof of Lemma 5.3 is given in the Appendix (Section A.4), and it relies on the fact that the configurations, which minimize the *energy* associated with the probability measure $\mathbb{P}_n$, are uniformly distributed on $\mathbb{T}$ (like the vertices of a regular n-gon) so that we know explicitly the minimal energy as well as the partition function.

To complete the proof of Proposition 2.7, we also need [6], Lemma 2, in order to give an estimate for the integral on the RHS of (69).

LEMMA 5.4 ([6]). *Let $f:\mathbb{T}\to\mathbb{R}$ be a trigonometric polynomial of degree $m\in\mathbb{N}$, and let $\|f\|_{L^2}=\sqrt{\int_{\mathbb{T}} f(\theta)^2\,\mathrm{d}\mu}$, where $\mathrm{d}\mu=\frac{\mathrm{d}\theta}{2\pi}$ denotes the uniform measure on $\mathbb{T}$. If we let $\mathcal{T}_\lambda=\{\theta\in\mathbb{T}:|f(\theta)|\leq\lambda\}$, then we have, for any $\lambda>0$,*

$$\mu(\mathcal{T}_\lambda)\leq 2e\bigg(\frac{\lambda}{\sqrt{2}\|f\|_{L^2}}\bigg)^{1/2m}.$$

From Lemma 5.4 we deduce that, for any trigonometric polynomial $f:\mathbb{T}\to\mathbb{R}$ of degree, at most, $m\in\mathbb{N}$, we have

$$\begin{aligned}\int_{\mathbb{T}} e^{-f(\theta)^2}\frac{\mathrm{d}\theta}{2\pi}&=\int_{\mathbb{T}}\bigg(\int_0^{+\infty} e^{-\lambda}\mathbf{1}_{\{|f(\theta)|\leq\sqrt{\lambda}\}}\,\mathrm{d}\lambda\bigg)\mu(\mathrm{d}\theta)\\&=\int_0^{+\infty} e^{-\lambda}\mu(\mathcal{T}_{\sqrt{\lambda}})\,\mathrm{d}\lambda\\&\leq 2e\int_0^{+\infty} e^{-\lambda}\bigg(\frac{\lambda}{2\|f\|_{L^2}^2}\bigg)^{1/4m}\mathrm{d}\lambda\\&\leq\frac{2e}{(2\|f\|_{L^2}^2)^{1/4m}},\end{aligned}$$

where we used that $\Gamma(1+1/4m)=\int_0^{+\infty}e^{-\lambda}\lambda^{1/4m}\,\mathrm{d}\lambda\leq 1$ for any $m\in\mathbb{N}$ in the last step. Hence, by combining this estimate with (69), we obtain the following general bound.

PROPOSITION 5.5. *Let $f:\mathbb{T}\to\mathbb{R}$ be a trigonometric polynomial for degree $m\in\mathbb{N}$. We have, for any $n\geq 2$,*

$$\mathbb{E}_n\big[e^{-\sum_{j=1}^n f(\theta_j)^2}\big]\leq\frac{{c_{15}}^n}{\sqrt{2\pi n}(2\|f\|_{L^2}^2)^{N/4}},$$

where $N=\frac{n}{m}$, $c_{15}=2e^2$ and $\|f\|_{L^2}=\sqrt{\int_{\mathbb{T}} f(\theta)^2\frac{\mathrm{d}\theta}{2\pi}}$.

We are now ready to complete the proof of Proposition 2.7. By combining the estimates from Proposition 5.1 and Proposition 5.5 with $f=\sqrt{\gamma}\mathrm{g}'$, which is a real–valued[2] trigono-

[2]We verify that, for any $n,m\in\mathbb{N}$ and $\eta\in(0,1]$, $\gamma>0$.

metric polynomial of degree $m \in \mathbb{N}$, we obtain that, for any $n \geq 2$ and any $\eta \in (0, 1]$,

$$\begin{aligned} |F_{n,m}(\xi)| &\leq \exp\left(c_{20}\left(n + \frac{2}{\pi^2}\right)\right) \frac{c_{15}{}^n}{\sqrt{2\pi n}(2\gamma \|g'\|_{L^2}^2)^{N/4}} \\ &\leq \frac{1}{\sqrt{2\pi n}} \left(\frac{e^{m\eta^2(\frac{\pi^2}{2}+\frac{1}{n})}}{2\gamma\|\xi\|}\right)^{N/4} \frac{c_{15}{}^n}{\|\xi\|^{N/4}}, \end{aligned} \tag{70}$$

where we used that by definition we have $\|g'\|_{L^2}^2 = \sum_{k=1}^m k|\zeta_k|^2 \geq \|\xi\|^2$ and we replaced $c_{20} = \frac{\pi^2\eta^2}{8}$. We still have the freedom to choose the parameter $\eta \in (0, 1]$ in the estimate (70), and we choose it in such a way to minimize $\eta^{-1} e^{m\eta^2 \frac{\pi^2}{2}}$. That is, we choose $\eta = \frac{1/\pi}{\sqrt{m}}$, and, since $2\gamma\|\xi\| = \frac{2\eta}{\sqrt{nm(m+1)}}(1 - \frac{\eta^2 c_{21}}{n})$, this implies that

$$|F_{n,m}(\xi)| \leq \left(\frac{\pi e^{\frac{1}{2}(1+\frac{2}{\pi^2 n})}\sqrt{n} m^{3/2}(m+1)}{2(1 - \frac{c_{21}/\pi^2}{nm})}\right)^{N/4} \frac{c_{15}{}^n}{\|\xi\|^{N/4}}.$$

Finally, let us observe that in the regime where $n \geq 4m^2$ (note that it is the only place where we use this condition), this implies that

$$|F_{n,m}(\xi)|^2 \leq \Upsilon_3(m)^{N/2} \frac{c_{15}{}^{2n} n^{N/4}}{\|\xi\|^{N/2}},$$

where $\Upsilon_3(m) = \frac{\pi m^{3/2}(m+1)e^{\frac{1}{2}(1+\frac{1/2}{(\pi m)^2})}}{2(1-\frac{c_{21}}{4\pi^2 m^3})}$, according to (115). This completes the proof.

6. Intermediate regime. The goal of this section is to prove Proposition 2.8. Recall that the polynomial g is given by (12) and that $h = -\mathcal{U}g$ is the Hilbert transform of the function $-g$; see (23). We will make use of the following basic estimates. We have, for any $\xi \in \mathbb{R}^{2m}$,

$$\|g\|_\infty, \|h\|_\infty \leq \sqrt{2(1 + \log m)}\|\xi\|. \tag{71}$$

Similarly, for any $\xi \in \mathbb{R}^{2m}$ and any integer $\kappa \geq 0$,

$$\|h^{(\kappa+1)}\|_\infty \leq \sum_{k \leq m} |\zeta_k| k^\kappa \sqrt{2k} \leq C_\kappa \|\xi\| \big(m(m+1)\big)^{\frac{\kappa+1}{2}} \tag{72}$$

with $C_0 = 1$, $C_1 = 1/\sqrt{2}$, $C_2 = 1/\sqrt{3}$ and

$$\|h^{(\kappa+1)}\|_{L^2} = \sqrt{\sum_{k \leq m} |\zeta_k|^2 k^{2\kappa+1}} \leq m^{\kappa+1/2}\|\xi\|. \tag{73}$$

We will also make use of Lemma 2.9, which is proved in Section 6.2, and we fix (throughout this section) the parameter

$$\nu = \frac{\nu_* N}{\sqrt{m+1}(1 + \log m)^{1/4}\|\xi\|}, \tag{74}$$

where $N = \frac{n}{m}$ and $0 < \nu_* \leq c_0$, as in (105). This last condition is necessary for our proof of Proposition 6.2 below, and we will optimize over the parameter ν_* in the proof of Proposition 2.8 which is given in the next section. This proof relies crucially on the following two estimates.

PROPOSITION 6.1. *Let $n, m \in \mathbb{Z}_+$ and $\xi \in \mathbb{R}^{2m}$. If $\nu > 0$ is given by* (74), *then*

$$\mathbb{E}_n\big[e^{-2\sum_{j=1}^n \Im g(\theta_j + \mathrm{i}\frac{\nu}{n}h(\theta_j))}\big] \le \exp\left(-2\nu\|\xi\|^2\left(1 - c_{10} - \frac{4c_{11}\nu_*\|\xi\|}{N\sqrt{m+1}}(1+\log m)^{3/4}\right)\right).$$

PROPOSITION 6.2. *Let $n, m \in \mathbb{Z}_+$ with $m \ge 3$, $\xi \in \mathbb{R}^{2m}$, and suppose that the parameter ν is given by* (74) *with $0 < \nu_* \le c_0$. If H is given by* (24), *we have*

$$\mathbb{E}_n\left[\exp\left(\frac{\nu^2}{n^2}\sum_{i,j=1}^n H(\theta_i, \theta_j)\right)\right] \le \exp\left(2c_9 + \frac{\nu_*^2 N^2(1+\epsilon_0)}{(m+1)\sqrt{1+\log m}}\right).$$

The proof of Proposition 6.1 is given in Section 6.3 while the proof of Proposition 6.2 is given in Section 6.4. Now that we are equipped with these two estimates, we can proceed with the proof of Proposition 2.8.

6.1. *Proof of Proposition* 2.8. Let us recall that the parameter ν is chosen according to (74), and we assume that $0 < \nu_* \le c_0 = \sqrt{\frac{1}{6\sqrt{2}}}$. By combining Lemma 2.9, Lemma 6.1 and Proposition 6.2, we obtain

$$\begin{aligned}|F_{n,m}(\xi)|^2 \le \exp\bigg(2c_9 &+ \frac{\nu_*^2 N^2(1+\epsilon_0)}{(m+1)\sqrt{1+\log m}} \\ &- \frac{2\nu_* N\|\xi\|}{\sqrt{m+1}(1+\log m)^{1/4}}\left(1 - c_{10} - \frac{4c_{11}\nu_*\|\xi\|}{N\sqrt{m+1}}(1+\log m)^{3/4}\right)\bigg).\end{aligned} \tag{75}$$

Let Λ_2 be as in (35), that is,

$$\Lambda_2 = \frac{c_0{}^{-1}(1-c_{10})N\sqrt{m+1}}{8(1+\log m)^{3/4}c_{11}}.$$

In order to maximize the polynomial $\nu_*(1 - c_{10} - 4\nu_* c_{11}\frac{\|\xi\|}{N\sqrt{m+1}}(1+\log m)^{3/4})$, we choose $\nu_* = \frac{(1-c_{10})N\sqrt{m+1}}{8\|\xi\|(1+\log m)^{3/4}c_{11}}$. Then, we verify that in the regime where $\|\xi\| \ge \Lambda_2$, we have $\nu_* \le c_0$ so that we are allowed to use the estimate (75). We obtain

$$|F_{n,m}(\xi)|^2 \le \exp\left(2c_9 + \frac{c_0{}^2 N^2(1+\epsilon_0)}{(m+1)\sqrt{1+\log m}} - \frac{(1-c_{10})^2 N^2}{8(1+\log m)c_{11}}\right).$$

If $c_1 = \frac{(1-c_{10})^2}{16c_{11}} - c_0{}^2(1+\epsilon_0)\frac{\sqrt{1+\log m}}{2(m+1)}$, according to (107), it follows from the previous formula that in the regime where $\|\xi\| \ge \Lambda_2$,

$$|F_{n,m}(\xi)|^2 \le \exp\left(2c_9 - \frac{2c_1 N^2}{1+\log m}\right).$$

This proves the estimate (36).

On the other hand, in the regime where $\|\xi\| \le \Lambda_2$ if we choose $\nu_* = c_0$ in the estimate (75), by (35) we verify that

$$\begin{aligned}|F_{n,m}(\xi)|^2 &\le \exp\left(2c_9 + c_0{}^2\left(\frac{N^2(1+\epsilon_0)}{(m+1)\sqrt{1+\log m}} - 8c_{11}\frac{\sqrt{1+\log m}}{m+1}\|\xi\|(2\Lambda_2 - \|\xi\|)\right)\right) \\ &\le \exp\left(2c_9 + c_0{}^2\left(\frac{N^2(1+\epsilon_0)}{(m+1)\sqrt{1+\log m}} - 8c_{11}\frac{\sqrt{1+\log m}}{m+1}\Lambda_1(2\Lambda_2 - \Lambda_1)\right)\right),\end{aligned}$$

where we used that the minimum of the function $\xi \mapsto \|\xi\|(2\Lambda_2 - \|\xi\|)$ for $\Lambda_1 \le \|\xi\| \le \Lambda_2$ equals

$$\begin{aligned}\Lambda_1(2\Lambda_2 - \Lambda_1) &= \frac{c_4 N}{\sqrt{1+\log m}}\left(\frac{{c_0}^{-1}(1-c_{10})N\sqrt{m+1}}{4c_{11}(1+\log m)^{3/4}} - \frac{c_4 N}{\sqrt{1+\log m}}\right) \\ &= \frac{{c_0}^{-1}c_4\sqrt{m+1}}{4c_{11}(1+\log m)^{5/4}}N^2\left(1 - c_{10} - \frac{4c_4c_0c_{11}(1+\log m)^{1/4}}{\sqrt{m+1}}\right).\end{aligned}$$

Hence, if $c_2 = c_0c_4(1-c_{10}-\frac{4c_4c_0c_{11}(1+\log m)^{1/4}}{\sqrt{m+1}}) - {c_0}^2\frac{(1+\epsilon_0)(1+\log m)^{1/4}}{2\sqrt{m+1}}$, according to (107), it follows from the previous formulae that in the regime where $\Lambda_1 \le \|\xi\| \le \Lambda_2$,

$$\left|F_{n,m}(\xi)\right|^2 \le \exp\left(2c_9 - \frac{2c_2(m)N^2}{\sqrt{m+1}(1+\log m)^{3/4}}\right).$$

This proves the estimate (37), and it completes the proof. It just remains to prove Lemma 2.9 as well as Propositions 6.1 and 6.2 which is the task that we undertake in the next sections.

6.2. *Proof of Lemma* 2.9. Let us recall that by Lemma 2.4, we have, for any $\nu > 0$,

$$\left|F_{n,m}(\xi)\right| \le \mathbb{E}_n\left[\prod_{1\le i<j\le n}\left|\frac{\sin(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{h(\theta_i)-h(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right|^2\prod_{j=1}^n\left|1+\mathrm{i}\frac{\nu}{n}h'(\theta_j)\right|e^{-\Im g(\theta_j+\mathrm{i}\frac{\nu}{n}h(\theta_j))}\right].$$

By Lemma 5.2, we obtain, for all $\theta_i, \theta_j \in \mathbb{T}$ with $\theta_i \ne \theta_j$,

$$\begin{aligned}\left|\frac{\sin(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{h(\theta_i)-h(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right|^2 &= 1 + \left(\frac{\sinh(\nu\frac{h(\theta_i)-h(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right)^2 \\ &\le \exp\left(\nu\frac{h(\theta_i)-h(\theta_j)}{2n\sin(\frac{\theta_i-\theta_j}{2})}\right)^2 \\ &= \exp\left(\frac{\nu^2}{n^2}H(\theta_i,\theta_j)\right),\end{aligned}$$

where the function H is as in (24). Moreover, we also have

$$\prod_{j=1}^n\left|1+\mathrm{i}\frac{\nu}{n}h'(\theta_j)\right|^2 \le \exp\left(\frac{\nu^2}{n^2}\sum_{j=1}^n H(\theta_j,\theta_j)\right).$$

Combining these bounds, we obtain, for any $\theta_1,\dots,\theta_n \in \mathbb{T}$ distinct and any $\nu > 0$,

$$\prod_{1\le i<j\le n}\left|\frac{\sin(\frac{\theta_i-\theta_j}{2} + \mathrm{i}\nu\frac{h(\theta_i)-h(\theta_j)}{2n})}{\sin(\frac{\theta_i-\theta_j}{2})}\right|^2\prod_{j=1}^n\left|1+\mathrm{i}\frac{\nu}{n}h'(\theta_k)\right| \le \exp\left(\frac{1}{2}\frac{\nu^2}{n^2}\sum_{i,j=1}^n H(\theta_i,\theta_j)\right).$$

Hence, by the Cauchy–Schwarz inequality this implies that

$$\left|F_{n,m}(\xi)\right|^2 \le \mathbb{E}_n\left[\exp\left(\frac{\nu^2}{n^2}\sum_{i,j=1}^n H(\theta_i,\theta_j)\right)\right]\mathbb{E}_n\left[e^{-2\sum_{j=1}^n \Im g(\theta_j+\mathrm{i}\frac{\nu}{n}h(\theta_j))}\right].$$

6.3. *Proof of Proposition* 6.1. Recall that, according to (74), we assume that $\nu = \frac{\nu_* n}{m\sqrt{m+1}(1+\log m)^{1/4}\|\xi\|}$ for a constant $\nu_* > 0$. Using the estimate (71), this implies that

$\frac{\nu}{n}\|h\|_\infty \le \sqrt{2}\nu_* \frac{(1+\log m)^{1/4}}{m\sqrt{m+1}}$. Then, since both functions g, h are real valued on $\mathbb{T}$ and g is an analytic function, we have

$$\left|\Im g\left(\theta_j + i\frac{\nu}{n}h(\theta_j)\right) - \frac{\nu}{n}g'(\theta_j)h(\theta_j)\right| \le \frac{\nu^3}{6n^3}|h(\theta_j)|^3 \sup_{\substack{z\in\mathbb{C}:\\ |\Re z|\le\pi, |\Im z|\le\sqrt{2}\nu_*\frac{(1+\log m)^{1/4}}{m\sqrt{m+1}}}} |g'''(z)|.$$

Moreover, by (12), we have, for any $z \in \mathbb{C}$,

$$g'''(z) = \frac{-i}{\sqrt{2}} \sum_{|k|\le m} |k|^{5/2}\zeta_k e^{ikz}$$

so that if $|\Re z| \le \pi$, $|\Im z| \le \sqrt{2}\nu_* \frac{(1+\log m)^{1/4}}{m\sqrt{m+1}}$, then

$$|g'''(z)| \le \sqrt{2}e^{\sqrt{2}\nu_* \frac{(1+\log m)^{1/4}}{\sqrt{m+1}}} \sum_{k=1}^m |\zeta_k|k^{5/2} \le 3\sqrt{2}c_{19}\|\xi\| m^{3/2}(m+1)^{3/2},$$

where $c_{19}(m) = \frac{1}{3\sqrt{6}}e^{\sqrt{2}c_0\frac{(1+\log m)^{1/4}}{\sqrt{m+1}}}$ and we used that $\sum_{k=1}^m k^5 \le \frac{m^3(m+1)^3}{6}$. These bounds and the estimate (71) show that

$$\begin{aligned}
&\mathbb{E}_n\left[\exp\left(-2\sum_{j=1}^n \Im g\left(\theta_j + i\frac{\nu}{n}h(\theta_j)\right)\right)\right] \\
&\le \mathbb{E}_n\left[\exp\left(-\frac{2\nu}{n}\sum_{j=1}^n g'(\theta_j)h(\theta_j) + 2c_{19}\|\xi\|^2 \frac{\nu^3 m^{3/2}(m+1)^{3/2}}{n^3}\sqrt{1+\log m}\sum_{j=1}^n h(\theta_j)^2\right)\right].
\end{aligned}$$

Let us denote $\gamma = c_{19}\|\xi\|^2 \frac{\nu^2 m^{3/2}(m+1)^{3/2}}{n^2}\sqrt{1+\log m}$ and $f = g' - \gamma h$. By Lemma 2.3 this implies that

$$\begin{aligned}
\mathbb{E}_n\left[\exp\left(-2\sum_{j=1}^n \Im g\left(\theta_j + i\frac{\nu}{n}h(\theta_j)\right)\right)\right] &\le \mathbb{E}_n\left[\exp\left(-\frac{2\nu}{n}\sum_{j=1}^n f(\theta_j)h(\theta_j)\right)\right] \\
&\le \exp\left(-2\nu\int_{\mathbb{T}} f(\theta)h(\theta)\frac{d\theta}{2\pi} + \frac{4\nu^2}{n^2}\mathcal{A}(fh)\right).
\end{aligned}$$

First, observe that, since we have chosen $h = -\mathcal{U}g$, we have

$$\int_{\mathbb{T}} h(\theta)^2 \frac{d\theta}{2\pi} \le \|\xi\|^2, \tag{76}$$

and by formulae (25)-(26) we obtain

$$\mathbb{E}_n\left[\exp\left(-2\sum_{j=1}^n \Im g\left(\theta_j + i\frac{\nu}{n}h(\theta_j)\right)\right)\right] \le \exp\left(-2\nu(1-\gamma)\|\xi\|^2 + \frac{4\nu^2}{n^2}\mathcal{A}(fh)\right). \tag{77}$$

It remains to estimate the quantities $\mathcal{A}(fh)$ where the seminorm $\mathcal{A}$ is given by (14) and $f = g' - \gamma h$. To that end, we may use the bound $\mathcal{A}(u) \le \|u\|_{L^2}\|u'\|_{L^2}$ which holds for any smooth function $u : \mathbb{T} \to \mathbb{C}$. First, we have

$$\|fh\|_{L^2} \le \|h\|_\infty \|f\|_{L^2} \le \sqrt{2(1+\log m)}(\sqrt{m}+\gamma)\|\xi\|^2,$$

where we used the estimates (71), (73) and (76). Second, we have

$$\begin{aligned}\|(fh)'\|_{L^2} &\le \|h\|_\infty \|f'\|_{L^2} + \|f\|_\infty \|h'\|_{L^2}\\ &\le \big(\sqrt{2m(1+\log m)}(m+\gamma) + \big(\sqrt{m(m+1)} + \gamma\sqrt{2(1+\log m)}\big)\sqrt{m}\big)\|\xi\|^2\\ &= m\sqrt{2m(1+\log m)}\bigg(1 + \frac{2\gamma}{m} + \sqrt{\frac{1+1/m}{2(1+\log m)}}\bigg)\|\xi\|^2.\end{aligned}$$

Here we used that $\|g^{(\kappa)}\|_{L^2} = \|h^{(\kappa)}\|_{L^2}$ for any $\kappa \ge 0$ since h is the Hilbert transform of g and the estimates (71)–(73). Combining all these estimates, we deduce from formula (77) that

$$\begin{aligned}&\mathbb{E}_n\bigg[\exp\bigg(-2\sum_{j=1}^n \Im\Big\{g\Big(\theta_j + \mathrm{i}\frac{\nu}{n}h(\theta_j)\Big)\Big\}\bigg)\bigg]\\ &\le \exp\bigg(-2\nu\|\xi\|^2\bigg(1-\gamma-\frac{4\nu\|\xi\|^2m^2}{n^2}(1+\log m)\Big(1+\frac{\gamma}{\sqrt{m}}\Big)\bigg(1+\frac{2\gamma}{m}+\sqrt{\frac{1+1/m}{2(1+\log m)}}\bigg)\bigg)\bigg).\end{aligned}$$

To complete the proof, it remains to observe that, by (74) and (107), we have

$$\begin{aligned}\gamma &= c_{19}\|\xi\|^2\frac{\nu^2(m+1)^{3/2}}{N^2\sqrt{m}}\sqrt{1+\log m} = c_{19}\nu_*^2\sqrt{1+1/m} \le c_{10}(m)\\ &= c_0{}^2\frac{\sqrt{1+1/m}}{3\sqrt{6}}e^{c_0\frac{(1+\log m)^{1/4}}{\sqrt{(m+1)/2}}}\end{aligned}$$

after replacing $c_{19} = \frac{1}{3\sqrt{6}}e^{c_0\frac{(1+\log m)^{1/4}}{\sqrt{(m+1)/2}}}$ and using that $\nu_* \le c_0$. Moreover, by (74) we also have $\frac{\nu\|\xi\|^2m^2}{n^2}(1+\log m) = \frac{\nu_*\|\xi\|}{N\sqrt{m+1}}(1+\log m)^{3/4}$ and $c_{11} = (1+\frac{c_{10}}{\sqrt{m}})(1+\frac{2c_{10}}{m}+\sqrt{\frac{1+1/m}{2(1+\log m)}})$.

6.4. *Proof of Proposition* 6.2. Let us denote, for any $k \in \mathbb{Z}$,

$$\mathrm{T}_k = \operatorname{Tr}\mathbf{U}^k = \sum_{j=1}^n e^{\mathrm{i}k\theta_j} = \sqrt{\frac{k}{2}}(\mathrm{X}_{2k-1} + \mathrm{i}\mathrm{X}_{2k}).$$

The idea of the proof is to view $\sum_{i,j=1}^n H(\theta_i,\theta_j)$ as a quadratic form in the random variables $(\mathrm{T}_k)_{k\in\mathbb{Z}}$ and to use this observation to express the Laplace transform of the random variable $\sum_{i,j=1}^n H(\theta_i,\theta_j)$ as a (multivariate) Gaussian integral, as explained at the end of Section 2.4.

LEMMA 6.3. *We have the identity*

$$\sum_{i,j=1}^n H(\theta_i,\theta_j) = \frac{1}{2}\Re\bigg\{\sum_{p,q\in\mathbb{Z}} A_{pq}\mathrm{T}_p\mathrm{T}_q + \sum_{p,q\in\mathbb{Z}} B_{pq}\mathrm{T}_p\mathrm{T}_q\bigg\},$$

where

$$A_{pq} = \sum_{1\le k\le \ell\le m}(\mathbf{1}_{1\le p-k+1\le p+q-\ell\le m} + \mathbf{1}_{1\le q-k+1\le p+q-\ell\le m})\frac{\zeta_\ell\zeta_{p+q-\ell}}{\sqrt{\ell(p+q-\ell)}}$$

and

$$B_{pq} = \sum_{1\le k\le \ell\le m}(\mathbf{1}_{1\le k-p\le \ell-p-q\le m} + \mathbf{1}_{1\le k-q\le \ell-p-q\le m})\frac{\zeta_\ell\zeta_{p+q-\ell}}{\sqrt{\ell(\ell-p-q)}}.$$

PROOF. An elementary computation gives that, for any $\ell \in \mathbb{Z}$,

$$\frac{e^{\mathrm{i}\ell\theta} - e^{\mathrm{i}\ell x}}{2\mathrm{i}\sin(\frac{\theta - x}{2})} = \sum_{k=1}^{\ell} e^{\mathrm{i}(k-1/2)\theta} e^{\mathrm{i}(\ell-k+1/2)x}, \quad x, \theta \in \mathbb{T}.$$

By (23)–(24), this directly implies that, for any $i, j = 1, \dots, n$,

$$H(\theta_i, \theta_j) = \Re\Bigg\{ \sum_{1\le k\le \ell\le m} \sum_{1\le r\le s\le m} \frac{\zeta_\ell \zeta_s}{\sqrt{\ell s}} e^{\mathrm{i}(k-1/2)\theta_i} e^{\mathrm{i}(\ell-k+1/2)\theta_j} e^{\mathrm{i}(r-1/2)\theta_i} e^{\mathrm{i}(s-r+1/2)\theta_j} \Bigg\}$$
$$+ \Re\Bigg\{ \sum_{1\le k\le \ell\le m} \sum_{1\le r\le s\le m} \frac{\zeta_\ell \overline{\zeta_s}}{\sqrt{\ell s}} e^{\mathrm{i}(k-1/2)\theta_i} e^{\mathrm{i}(\ell-k+1/2)\theta_j} e^{-\mathrm{i}(r-1/2)\theta_i} e^{-\mathrm{i}(s-r+1/2)\theta_j} \Bigg\}.$$

Then, summing over all variables θ_i, θ_j, we obtain

$$\sum_{1\le i,j\le n} H(\theta_i, \theta_j) = \Re\Bigg\{ \sum_{1\le k\le \ell\le m} \sum_{1\le r\le s\le m} \frac{\zeta_\ell \zeta_s}{\sqrt{\ell s}} \mathrm{T}_{k+r-1} \mathrm{T}_{\ell+s-k-r+1} \Bigg\} \tag{78}$$

$$+ \Re\Bigg\{ \sum_{1\le k\le \ell\le m} \sum_{1\le r\le s\le m} \frac{\zeta_\ell \overline{\zeta_s}}{\sqrt{\ell s}} \mathrm{T}_{k-r} \mathrm{T}_{\ell-s+r-k} \Bigg\}. \tag{79}$$

In (78) we make the change of variables $(r, s) \leftrightarrow (p, q)$ given by $r = p - k + 1$ and $s = q + p - \ell$. Similarly, in (79) we make the change of variables $(r, s) \leftrightarrow (p, q)$ given by $r = k - p$ and $s = \ell - q - p$. This implies that

$$\sum_{1\le i,j\le n} H(\theta_i, \theta_j) = \Re\Bigg\{ \sum_{1\le k\le \ell\le m} \sum_{p,q\in\mathbb{Z}} \frac{\zeta_\ell \zeta_{q+p-\ell}}{\sqrt{\ell(q+p-\ell)}} \mathbf{1}_{1\le p-k+1\le q+p-\ell\le m} \mathrm{T}_p \mathrm{T}_q \Bigg\} \tag{80}$$

$$+ \Re\Bigg\{ \sum_{1\le k\le \ell\le m} \sum_{p,q\in\mathbb{Z}} \frac{\zeta_\ell \overline{\zeta_{\ell-q-p}}}{\sqrt{\ell(\ell-q-p)}} \mathbf{1}_{1\le k-p\le \ell-q-p\le m} \mathrm{T}_p \mathrm{T}_q \Bigg\}. \tag{81}$$

To finish the proof, it remains to symmetrize the previous formula over (p, q) and use that $\overline{\zeta_{-j}} = \zeta_j$ for all $j = 1, \dots, m$. Then, (80) corresponds to $\frac{1}{2}\Re\{\sum_{p,q\in\mathbb{Z}} A_{pq} \mathrm{T}_p \mathrm{T}_q\}$, and (81) corresponds to $\frac{1}{2}\Re\{\sum_{p,q\in\mathbb{Z}} B_{pq} \mathrm{T}_p \mathrm{T}_q\}$. □

Let us observe that, in the notation of Lemma 6.3, $A_{pq} \neq 0$ only if $1 \le p, q \le 2m - 1$ and $B_{pq} \neq 0$ only if $|p|, |q| \le m - 1$, so we may view $\mathbf{A} = (A_{pq})_{p,q=1}^{2m-1}$ and $\mathbf{B} = (B_{pq})_{1\le|p|,|q|<m}$ as symmetric matrix-valued functions of the parameters $(\zeta_k)_{k=1}^m$. In the following, we denote:

$$\mathfrak{Q}_{\mathbf{A}} = \Re\Bigg\{ \sum_{p,q\in\mathbb{Z}} A_{pq} \mathrm{T}_p \mathrm{T}_q \Bigg\}, \qquad \mathfrak{Q}_{\mathbf{B}} = \Re\Bigg\{ \sum_{\substack{p,q\in\mathbb{Z}\\ p,q\neq 0}} B_{pq} \mathrm{T}_p \mathrm{T}_q \Bigg\}$$

and

$$\mathfrak{L} = \Re\Bigg\{ n B_{00} + 2 \sum_{\substack{p\in\mathbb{Z}\\ p\neq 0}} B_{p0} \mathrm{T}_p \Bigg\}.$$

We introduce this decomposition because $\mathrm{T}_0 = n$ is not a random variable and should be treated individually. By the Cauchy–Schwarz inequality, Lemma 6.3 implies that

$$\mathbb{E}_n\Bigg[\exp\Bigg(\delta \sum_{i,j=1}^{n} H(\theta_i, \theta_j)\Bigg)\Bigg] \le \big(\mathbb{E}_n[\exp(2\delta \mathfrak{Q}_{\mathbf{A}})] \mathbb{E}_n[\exp(2\delta \mathfrak{Q}_{\mathbf{B}})]\big)^{1/4} \sqrt{\mathbb{E}_n[\exp(\delta n \mathfrak{L})]}, \tag{82}$$

where $\delta = (\frac{\nu}{n})^2$. Our first observation is that $\mathfrak{L}$ is a linear statistic associated with the trigonometric polynomial $f = nB_{00} + 2\Re\{\sum_{p\neq 0} B_{p0}e^{\mathrm{i}p\theta}\}$ so that, by Lemma 2.3, we have the estimate

$$\mathbb{E}_n[\exp(\delta n\mathfrak{L})] \le \exp\biggl(\delta n^2 B_{00} + (\delta n)^2 \sum_{0<p<m} p|B_{p0} + \overline{B_{-p0}}|^2\biggr).$$

In combination with Lemma 6.4 below, this implies that

$$\mathbb{E}_n[\exp(\delta n\mathfrak{L})] \le \exp\biggl(2\delta n^2 \|\xi\|^2 + (\delta n)^2 \frac{4m^3}{3}\|\xi\|^4\biggr). \tag{83}$$

LEMMA 6.4. *In the notation of Lemma 6.3, we have* $B_{00} = 2\|\xi\|^2$ *and*

$$\sum_{0<p<m} p|B_{p0} + \overline{B_{-p0}}|^2 \le \frac{4m^3}{3}\|\xi\|^4.$$

PROOF. First of all, by definition we have

$$\frac{B_{00}}{2} = \sum_{1\le k\le \ell\le m} \frac{\zeta_\ell \zeta_{-\ell}}{\ell} = \sum_{1\le \ell\le m} |\zeta_\ell|^2 = \|\xi\|^2.$$

Second, we also have, for any $p \in \mathbb{Z}$,

$$B_{p0} = \sum_{1\le k\le \ell\le m} (\mathbf{1}_{1\le k-p\le \ell-p\le m} + \mathbf{1}_{1\le k\le \ell-p\le m}) \frac{\zeta_\ell \zeta_{p-\ell}}{\sqrt{\ell(\ell-p)}}.$$

This shows that, for $p \ge 1$,

$$\begin{aligned} |B_{p0}| &\le \|\xi\| \sum_{1\le k\le \ell\le m} (\mathbf{1}_{p+1\le k\le \ell\le m} + \mathbf{1}_{1\le k\le \ell-p\le m}) \frac{|\zeta_\ell|}{\sqrt{\ell(\ell-p)}} \\ &= 2\|\xi\| \sum_{p<\ell\le m} \sqrt{1-p/\ell}\,|\zeta_\ell|, \end{aligned}$$

where at the second step we computed the sum over k. By the Cauchy–Schwarz inequality, this shows that

$$|B_{p0}| \le 2\sqrt{m-p}\|\xi\|^2.$$

This estimate implies that

$$\sum_{p=1}^{m-1} p|B_{p0}|^2 \le 4\|\xi\|^4 \sum_{p=1}^{m-1} p(m-p) \le \frac{2m^3}{3}\|\xi\|^4.$$

Similarly, we can show that $|B_{-p0}| \le 2\sqrt{m-p}\|\xi\|^2$ for any $p \ge 1$ so that we also have $\sum_{p=1}^{m-1} p|B_{-p0}|^2 \le \frac{2m^3}{3}\|\xi\|^4$. This completes the proof. □

In the remainder of this section, our task is to bound the terms which involve the quadratic forms $\mathfrak{Q}_{\mathbf{A}}$ and $\mathfrak{Q}_{\mathbf{B}}$ on the RHS of (82). In order to do this, we need a priori estimates for the norms of the corresponding matrices $\mathbf{A}$ and $\mathbf{B}$.

LEMMA 6.5. *Let* $\|\mathbf{A}\| = \max_{1\le p<2m} \sum_{q=1}^{2m-1} |A_{pq}|$ *and* $\|\mathbf{B}\| = \max_{1\le |p|<m} \sum_{|q|=1}^{m-1} |B_{pq}|$. *We have*

$$\|\mathbf{A}\|, \|\mathbf{B}\| \le \sqrt{2m(m+1)(1+\log m)}\|\xi\|^2.$$

PROOF. By definition we have

$$\sum_{q=1}^{2m-1} |A_{pq}| \le 2 \sum_{1\le k\le \ell\le m} \frac{|\zeta_\ell|}{\sqrt{\ell}} \sum_{q=1}^{2m-1} \mathbf{1}_{1\le p+q-\ell\le m} \frac{|\zeta_{p+q-\ell}|}{\sqrt{p+q-\ell}}.$$

The last sum is bounded by $\sum_{r=1}^{m} \frac{|\zeta_r|}{\sqrt{r}}$, so we obtain

$$\sum_{q=1}^{2m-1} |A_{pq}| \le 2 \sum_{k,r=1}^{m} \sqrt{\frac{\ell}{r}} |\zeta_\ell||\zeta_r|.$$

By the Cauchy–Schwarz inequality, this shows that

$$\sum_{q=1}^{2m-1} |A_{pq}| \le 2\|\zeta\|^2 \sqrt{\sum_{k,r=1}^{m} \frac{\ell}{r}} \le \sqrt{2m(m+1)(1+\log m)}\|\zeta\|^2.$$

Since $\|\zeta\| = \|\xi\|$, this gives the estimate for $\|\mathbf{A}\|$; the argument for $\|\mathbf{B}\|$ is exactly the same. □

Let us define new objects. For $\delta_1, \delta_2 > 0$, we set

$$\mathbf{M} = \begin{pmatrix} \mathbf{I}_{2m-1} & \delta_2\mathbf{A}^* \\ \delta_2\mathbf{A} & \mathbf{I}_{2m-1} \end{pmatrix} \quad \text{and} \quad \mathbf{v} = \sqrt{\delta_1} \begin{pmatrix} \mathrm{T}_1 \\ \vdots \\ \mathrm{T}_{2m-1} \end{pmatrix}. \tag{84}$$

REMARK 6.1. As explained in Section 2.4, it is not clear whether the matrices $\mathbf{A}$ and $\mathbf{B}$ are positive definite. This issue is resolved by bounding the quadratic from $\mathfrak{Q}_{\mathbf{A}}$, using the matrix $\mathbf{M}$ (see the estimate (86)), and by choosing the parameter δ_2 small enough to guarantee that $\mathbf{M}$ is positive definite and the Gaussian integral (87) is convergent.

Since $\mathbf{A}$ is a symmetric matrix, we have $\|\mathbf{A}^*\| = \|\mathbf{A}\|$ and

$$\|\mathbf{M} - \mathbf{I}_{4m-2}\| = \max_{1\le p<4m-1} \sum_{q=1}^{4m-2} |M_{pq} - \mathbf{1}_{pq}| = \delta_2\|\mathbf{A}\|.$$

Hence, by Lemma 6.5, if $\delta_2 \le \frac{1}{3\sqrt{2m(m+1)(1+\log m)}\|\xi\|^2}$, then

$$\|\mathbf{M} - \mathbf{I}_{4m-2}\| \le \frac{1}{3} \tag{85}$$

so that the matrix $\mathbf{M}$ is invertible with $\mathbf{M}^{-1} = \sum_{k=0}^{+\infty}(\mathbf{I}_{4m-2} - \mathbf{M})^k$ (convergent Neumann series). This also implies that $\mathbf{M}$ is positive definite, and we have

$$\begin{pmatrix} \mathbf{v}^* & \overline{\mathbf{v}}^* \end{pmatrix} \begin{pmatrix} \mathbf{I} & \delta_2\mathbf{A}^* \\ \delta_2\mathbf{A} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix} = 2\delta_2\Re\{\mathbf{v}^{\mathrm{t}}\mathbf{A}\mathbf{v}\} + 2|\mathbf{v}|^2.$$

Thus, if we set $\delta = \delta_1\delta_2$ and use the notation (84), this shows that

$$2\delta\mathfrak{Q}_{\mathbf{A}} = 2\delta\Re\Big\{\sum_{p,q\in\mathbb{Z}} A_{pq}\mathrm{T}_p\mathrm{T}_q\Big\} \le \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}^* \mathbf{M} \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}. \tag{86}$$

Then, in order to estimate the quantity $\mathbb{E}_n[\exp(2\delta\mathfrak{Q}_{\mathbf{A}})]$, we may use the identity

$$\begin{aligned} &\pi^{4m-2}\det(\mathbf{M})\exp\left(\begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}^* \mathbf{M} \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}\right) \\ &\quad = \int_{\mathbb{C}^{4m-2}} \exp\left(-\mathbf{z}^*\mathbf{M}^{-1}\mathbf{z} + \mathbf{z}^*\begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix} + \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}^*\mathbf{z}\right) \mathrm{d}^2\mathbf{z}, \end{aligned} \tag{87}$$

where $\mathbf{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_{4m-2} \end{pmatrix}$ and $\mathrm{d}^2\mathbf{z} = \prod_{k=1}^{4m-2} \mathrm{d}\Re(z_k)\,\mathrm{d}\Im(z_k)$ denotes the Lebesgue measure on $\mathbb{C}^{4m-2}$. Formula (87) is a simple Gaussian integration on $\mathbb{C}^{4m-2}$, and it makes sense since we have seen that the matrix $\mathbf{M}$ is positive definite by (85). Moreover, it is useful since

$$\begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}^* \mathbf{z} = \sqrt{\delta_1} \sum_{k=1}^{2m-1} (z_k \overline{\mathrm{T}_k} + z_{2m-1+k}\mathrm{T}_k)$$

is a (mean-zero) linear statistic of a trigonometric polynomial so that, by Lemma 2.3, we have

$$\begin{aligned} \mathbb{E}_n\left[\exp\left(\mathbf{z}^* \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix} + \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}^* \mathbf{z}\right)\right] &\le \exp\left(\delta_1 \sum_{k=1}^{2m-1} k|\overline{z_k} + z_{2m-1+k}|^2\right) \\ &\le \exp(2\delta_1 \mathbf{z}^* \mathbf{C} \mathbf{z}), \end{aligned} \tag{88}$$

where $\mathbf{C}$ is a diagonal matrix given by

$$\mathbf{C} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 2m-1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 2m-1 \end{pmatrix}.$$

Hence, taking expectation in formula (87) and using the bound (88), we obtain

$$\begin{aligned} \mathbb{E}_n\left[\exp\left(\begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}^* \mathbf{M} \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}\right)\right] &\le \frac{1}{\pi^{4m-2}\det(\mathbf{M})} \int_{\mathbb{C}^{4m-2}} \exp(-\mathbf{z}^*(\mathbf{M}^{-1} - 2\delta_1\mathbf{C})\mathbf{z})\,\mathrm{d}^2\mathbf{z} \\ &= \frac{\det(\mathbf{M}^{-1} - 2\delta_1\mathbf{C})^{-1}}{\det(\mathbf{M})} = \frac{1}{\det(\mathbf{I} - 2\delta_1\mathbf{M}\mathbf{C})}. \end{aligned} \tag{89}$$

Here, we used that the matrix $\mathbf{M}^{-1} - 2\delta_1\mathbf{C}$ is also positive definite. Indeed, it follows from the above discussion (in particular from the estimate (85)) that if $\delta_2 \le \frac{1}{3\sqrt{2m(m+1)(1+\log m)}\|\xi\|^2}$ and $\delta_1 \le \frac{1}{2m^{3/2}\sqrt{m+1}}$, then, for any $m \ge 3$,

$$\|\mathbf{M}^{-1} - \mathbf{I}_{4m-2}\| \le \sum_{k=1}^{+\infty} \|\mathbf{M} - \mathbf{I}_{4m-2}\|^k \le \frac{1}{2} \quad \text{and} \quad 2\delta_1\|\mathbf{C}\| \le \frac{2m-1}{m^{3/2}\sqrt{m+1}} \le \frac{5}{6\sqrt{3}} < \frac{1}{2}.$$

Note that the condition $m \ge 3$ is crucial in order to obtain the second estimate. Moreover, since $\mathbf{M}$, $\mathbf{C}$ are Hermitian matrices with $\|\mathbf{M}\| \le 4/3$, it follows that, for all $m \ge 3$,

$$\det(\mathbf{I} - 2\delta_1\mathbf{M}\mathbf{C}) \ge \left(1 - \frac{4(2m-1)}{3m^{3/2}\sqrt{m+1}}\right)^{2(2m-1)} \ge e^{-c_{17}},$$

where $c_{17} = \frac{32}{3}(1 + \frac{(2-1/m)^3}{3(m+1)})$ and we used that $1 - x \ge e^{-x-x^2}$ for $0 \le x \le 2/3$. Hence, by formulae (86) and (89), we obtain, for $m \ge 3$,

$$\mathbb{E}_n[\exp(2\delta\mathfrak{Q}_{\mathbf{A}})] \le \mathbb{E}_n\left[\exp\left(\begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}^* \mathbf{M} \begin{pmatrix} \mathbf{v} \\ \overline{\mathbf{v}} \end{pmatrix}\right)\right] \le e^{c_{17}}. \tag{90}$$

In an analogous way, let us denote

$$\mathbf{N}=\begin{pmatrix}\mathbf{I} & \delta_2\mathbf{B}^*\\ \delta_2\mathbf{B} & \mathbf{I}\end{pmatrix},\qquad \mathbf{D}=\begin{pmatrix} m-1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & m-1\end{pmatrix}\quad\text{and}\quad \mathbf{w}=\sqrt{\delta_1}\begin{pmatrix}\mathrm{T}_{-m+1}\\ \vdots\\ \mathrm{T}_{-1}\\ \mathrm{T}_1\\ \vdots\\ \mathrm{T}_{m-1}\end{pmatrix}.$$

Then, we have

$$2\delta\mathfrak{Q}_{\mathbf{B}}=2\delta\Re\Bigg\{\sum_{\substack{p,q\in\mathbb{Z}\\ p,q\neq 0}}B_{pq}\mathrm{T}_p\mathrm{T}_q\Bigg\}\le\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}^*\mathbf{N}\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}.$$

By Lemma 6.5, if $\delta_2\le\frac{1}{3\sqrt{2m(m+1)(1+\log m)}\|\xi\|^2}$, then $\|\mathbf{N}-\mathbf{I}_{4m-2}\|\le 1/3$ so that both $\mathbf{N}$ and $\mathbf{N}^{-1}-2\delta_1\mathbf{D}$ are positive definite matrices. Like in our previous computations, this implies that

$$\begin{aligned}
&\mathbb{E}_n\left[\exp\left(\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}^*\mathbf{N}\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}\right)\right]\\
&\quad=\frac{1}{\pi^{4m-4}\det(\mathbf{N})}\int_{\mathbb{C}^{4m-4}}\exp(-\mathbf{z}^*\mathbf{N}^{-1}\mathbf{z})\mathbb{E}_n\left[\exp\left(\mathbf{z}^*\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}+\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}^*\mathbf{z}\right)\right]\mathrm{d}^2\mathbf{z}\\
&\quad\le\frac{1}{\pi^{4m-4}\det(\mathbf{N})}\int_{\mathbb{C}^{4m-4}}\exp(-\mathbf{z}^*(\mathbf{N}^{-1}-2\delta_1\mathbf{D})\mathbf{z})\,\mathrm{d}^2\mathbf{z}=\frac{1}{\det(\mathbf{I}-2\delta_1\mathbf{N}\mathbf{D})},
\end{aligned}$$

where at the second step we used an estimate analogous to (88). Moreover, since $\mathbf{N}$, $\mathbf{D}$ are Hermitian matrices with $\|\mathbf{N}\|\le 4/3$ and $\|\mathbf{D}\|=m-1$ for $m\ge 3$, if $\delta_1\le\frac{1}{2m^{3/2}\sqrt{m+1}}$, we have

$$\det(\mathbf{I}-2\delta_1\mathbf{N}\mathbf{D})\ge\left(1-\frac{4(m-1)}{3m^{3/2}\sqrt{m+1}}\right)^{2(m-1)}\ge e^{-c_{18}},$$

where $c_{18}=\frac{8}{3}(1+\frac{4(1-1/m)^3}{3(m+1)})$ and we used that $1-x\ge e^{-x-x^2}$ for $0\le x\le 2/3$. Combining these estimates, this implies that, for $m\ge 3$,

$$\mathbb{E}_n[\exp(2\delta\mathfrak{Q}_{\mathbf{B}})]\le\mathbb{E}_n\left[\exp\left(\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}^*\mathbf{N}\begin{pmatrix}\mathbf{w}\\ \overline{\mathbf{w}}\end{pmatrix}\right)\right]\le e^{c_{18}}. \tag{91}$$

Now, let us recall that we must have $\delta=(\frac{\nu}{n})^2=\delta_1\delta_2$. Hence, if we choose

$$\delta_1=\frac{1}{2m^{3/2}\sqrt{m+1}}\quad\text{and}\quad\delta_2=\frac{2\nu^2m^{3/2}\sqrt{m+1}}{n^2}=\frac{2\nu_*^2}{\sqrt{m(m+1)(1+\log m)}\|\xi\|^2},$$

according to (74), then we have $\delta_2\le\frac{1}{3\sqrt{2m(m+1)(1+\log m)}\|\xi\|^2}$, as required, provided that $\nu_*^2\le\frac{1}{6\sqrt{2}}$. Observe that this explains our choice for c_0 as the maximum admissible value for ν_*. In the end, if we combine all our estimates (82), (83), (90) and (91), if the parameter ν is given by (74) and $m\ge 3$, then we obtain

$$\mathbb{E}_n\left[\exp\left(\delta\sum_{i,j=1}^n H(\theta_i,\theta_j)\right)\right]\le\exp\left(2\frac{c_{17}+c_{18}}{8}+\delta n^2\|\xi\|^2+(\delta n)^2\frac{2m^3}{3}\|\xi\|^4\right). \tag{92}$$

By definitions we have $\frac{c_{17}+c_{18}}{8} = \frac{1}{3}(5 + 4\frac{(1-1/m)^3+(2-1/m)^3}{3(m+1)})$. This function attains its maximum over the positive integers for $m = 3$ so that $\frac{c_{17}+c_{18}}{4} \le c_9 = \frac{538}{243}$. Hence, if we replace $\delta = \frac{\nu^2}{n^2}$ in formula (92) and use (74), we conclude that

$$\begin{aligned}
&\mathbb{E}_n\left[\exp\left(\frac{\nu^2}{n^2}\sum_{i,j=1}^{n} H(\theta_i,\theta_j)\right)\right] \\
&\quad \le \exp\left(2c_9 + \nu^2\|\xi\|^2 + \frac{2\nu^4 m^3}{3n^2}\|\xi\|^4\right) \\
&\quad = \exp\left(2c_9 + \frac{\nu_*^2 N^2}{(m+1)\sqrt{1+\log m}} + \frac{2\nu_*^4 N^2}{3(m+1)(1+1/m)(1+\log m)}\right).
\end{aligned}$$

By definition of $\epsilon_0(m) \ge 0$ (106), this completes the proof.

7. Proof of Theorem 1.5. The method used in this section relies on the formalism introduced in [24] which provides a normal approximation result for certain observable of a Gibbs-type distribution and the following moment identities from [12]. According to (3), we let, for any $k \ge 1$,

$$\mathrm{T}_k = \sqrt{\frac{2}{k}}\,\mathrm{Tr}\,\mathbf{U}^k = \mathrm{X}_{2k} + \mathrm{i}\mathrm{X}_{2k+1}. \tag{93}$$

THEOREM 7.1 ([12]). *Fix $m \in \mathbb{N}$, and let $\mathbf{a}, \mathbf{b} \in \{0, 1, \dots\}^m$. Then, for all $n \ge \sum_{k=1}^m k a_k \vee \sum_{k=1}^m k b_k$,*

$$\mathbb{E}_n\left[\prod_{k=1}^{m} \mathrm{T}_k^{a_k}\overline{\mathrm{T}_k^{b_k}}\right] = \mathbb{E}\left[\prod_{k=1}^{m} \mathrm{Z}_k^{a_k}\overline{\mathrm{Z}_k^{b_k}}\right],$$

where $\mathrm{Z}_k = \mathrm{G}_{2k} + \mathrm{i}\mathrm{G}_{2k+1}$ for all $k \ge 1$ and G_k are i.i.d. standard Gaussian random variables.

Note that the hypothesis of Theorem 7.1 is incorrectly stated in [12] and we refer instead to [11] for a correct version of this Theorem as well as several applications to the asymptotic distributions of linear statistics of the eigenvalues of the CUE.

One can interpret the the law (1) of the eigenvalues of the CUE as a Gibbs distribution[3] on $\mathbb{T}^n$ with energy $\Phi(\boldsymbol{\theta}) := \sum_{1\le i<j\le n} \log|2\sin(\frac{\theta_i-\theta_j}{2})|^{-2}$. This implies that, formally, (1) is the stationary measure of a diffusion with generator

$$\mathrm{L} = -\Delta + \nabla\Phi\cdot\nabla = -\sum_{j=1}^{n}\left(\partial_{jj} + \sum_{i\ne j}\frac{\partial_j}{\tan(\frac{\theta_j-\theta_i}{2})}\right). \tag{94}$$

We view the vector $\mathbf{X} : \mathbb{T}^n \to \mathbb{R}^{2m}$ as a smooth function in $L^\infty(\mathbb{P}_n)$ so that we can define the vector $\mathrm{L}\mathbf{X}$ and the $2m \times 2m$ matrix

$$\boldsymbol{\Gamma}_{k,\ell} = \nabla\mathrm{X}_k \cdot \nabla\mathrm{X}_\ell. \tag{95}$$

Recall also the definition of the Kantorovich or Wasserstein distance (10). Then, by applying [24], Corollary 2.4, to the random variable $\mathbf{X}$ we obtain the following result.

[3]This means that the probability measure (1) also describes a 2d Coulomb gas of N point charges confined on the unit circle at inverse temperature $\beta = 2$.

PROPOSITION 7.2. *For all $n, m \in \mathbb{N}$ and for any positive definite diagonal matrix $\mathbf{K}$ of size $2m \times 2m$, we have*

$$\mathrm{W}_2(\mathbf{X}, \mathbf{G}) \le \sqrt{\mathbb{E}_n\big[|\mathbf{K}^{-1}\mathrm{L}\mathbf{X} - \mathbf{X}|^2\big]} + \sqrt{\mathbb{E}_n\big[\|\mathbf{I} - \mathbf{K}^{-1}\boldsymbol{\Gamma}\|^2\big]}, \tag{96}$$

where $\|\cdot\|$ denotes the Hilbert–Schmidt norm.

The reason the RHS of (96) is small is because the random variables $\mathrm{T}_1, \mathrm{T}_2, \dots$ are *approximate eigenfunctions* of the generator L, and the matrix $\mathbf{K}$ records the corresponding eigenvalues. The following Lemma makes this claim precise. Observe that $\boldsymbol{\xi}$ is small compared to $\mathbf{K}$ which is of order n.

LEMMA 7.3. *For all $n, m \in \mathbb{N}$, we have $\mathrm{L}\mathbf{X} = \mathbf{K}\mathbf{X} + \boldsymbol{\xi}$, where*

$$\mathbf{K} = n \cdot \mathrm{diag}(1, 1, 2, 2, \dots, m, m),$$

$$\boldsymbol{\xi} = (\Re\zeta_1, \Im\zeta_1, \Re\zeta_2, \Im\zeta_2, \dots, \Re\zeta_m, \Im\zeta_m),$$

and, for all $k \ge 1$,

$$\zeta_k = \sqrt{\frac{k}{2}} \sum_{\ell=1}^{k-1} \sqrt{\ell(k-\ell)}\,\mathrm{T}_\ell \mathrm{T}_{k-\ell}.$$

PROOF. The Lemma follows from the fact that $\mathrm{T}_k = \sqrt{\frac{2}{k}} \sum_{j=1}^n e^{\mathrm{i}k\theta_j}$ and explicit computations. Let us fix $k \in \mathbb{N}$ and observe that

$$\Delta \mathrm{T}_k = -k^2 \mathrm{T}_k. \tag{97}$$

Second, since $\tan(\frac{\theta_j - \theta_i}{2}) = -\mathrm{i}\frac{e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_i}}{e^{\mathrm{i}\theta_j} + e^{\mathrm{i}\theta_i}}$ for any $i, j = 1, \dots, n$, we have

$$\sum_{i \ne j} \frac{\partial_j \mathrm{T}_k}{\tan(\frac{\theta_j - \theta_i}{2})} = -\sqrt{2k} \sum_{i \ne j} \frac{e^{\mathrm{i}\theta_j} + e^{\mathrm{i}\theta_i}}{e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_i}} e^{\mathrm{i}k\theta_j}.$$

By symmetry this implies that

$$\begin{aligned}
\sum_{i \ne j} \frac{\partial_j \mathrm{T}_k}{\tan(\frac{\theta_j - \theta_i}{2})} &= -\sqrt{2k} \sum_{i \ne j} \frac{e^{\mathrm{i}k\theta_j} - e^{\mathrm{i}k\theta_i}}{e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_i}} e^{\mathrm{i}\theta_j} \\
&= -\sqrt{2k} \sum_{i \ne j} \sum_{\ell=1}^{k} e^{\mathrm{i}\ell\theta_j} e^{\mathrm{i}(k-\ell)\theta_i} \\
&= -\sqrt{2k} \sum_{i,j} \sum_{\ell=1}^{k} e^{\mathrm{i}\ell\theta_j} e^{\mathrm{i}(k-\ell)\theta_i} + k^2 \mathrm{T}_k,
\end{aligned} \tag{98}$$

where we have used that $\frac{e^{\mathrm{i}k\theta_j} - e^{\mathrm{i}k\theta_i}}{e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_i}} = \sum_{\ell=1}^k e^{\mathrm{i}(\ell-1)\theta_j} e^{\mathrm{i}(k-\ell)\theta_i}$. Note that, in the sum on the RHS of (98), the term $\ell = k$ equals to $-nk\mathrm{T}_k$ while the other terms can be expressed in terms of the variables $(\mathrm{T}_\ell)_{\ell=1}^{k-1}$. Hence, according to formula (94) and by combining formulae (97) and (98), this shows that, for any $k \ge 1$,

$$\mathrm{L}\mathrm{T}_k = nk\mathrm{T}_k + \zeta_k. \tag{99}$$

Taking real and imaginary parts of equation (99), this completes the proof. □

Remarkably, with Lemma 7.3 and Theorem 7.1 we can exactly compute the error terms on the RHS of the estimate (96). We obtain the following results.

LEMMA 7.4. *For any $n, m \in \mathbb{N}$ such that $m \le n$,*

$$\mathbb{E}_n[|\mathbf{K}^{-1}\boldsymbol{\xi}|^2] = \frac{(2m+5)m(m-1)}{9n^2}.$$

LEMMA 7.5. *For any $n, m \in \mathbb{N}$ such that $m \le n/2$, we have*

$$\mathbb{E}_n[\|\mathbf{I} - \mathbf{K}^{-1}\boldsymbol{\Gamma}\|^2] = \frac{(8m+7)(m+1)m}{6n^2}.$$

Using Lemmas 7.4 and 7.5, we can complete the proof of Theorem 1.5. According to Lemma 7.3, we have

$$\mathbf{K}^{-1}\mathrm{L}\mathbf{X} - \mathbf{X} = \mathbf{K}^{-1}\boldsymbol{\xi}$$

so that, for all $m, n \in \mathbb{N}$ such that $m \le n/2$,

$$\begin{aligned}
&\sqrt{\mathbb{E}_n[|\mathbf{K}^{-1}\mathrm{L}\mathbf{X} - \mathbf{X}|^2]} + \sqrt{\mathbb{E}_n[\|\mathbf{I} - \mathbf{K}^{-1}\boldsymbol{\Gamma}\|^2]} \\
&\quad \le \frac{\sqrt{(2m+5)m(m-1)} + \sqrt{(8m+7)(m+1)9m/6}}{3n} \le (\sqrt{8} + \sqrt{2})\frac{(m+1)\sqrt{m}}{3n}.
\end{aligned}$$

By Proposition 7.2 we obtain the required bound for the Kantorovich distance $\mathrm{W}_2(\mathbf{X}, \mathbf{G})$ between the random vector $\mathbf{X}$ and a standard Gaussian random variable on $\mathbb{R}^{2m}$. Thus, to complete the proof it remains to prove Lemmas 7.4 and 7.5.

PROOF OF LEMMA 7.4. According to the notation of Lemma 7.3, we have

$$|\mathbf{K}^{-1}\boldsymbol{\xi}|^2 = \sum_{k=1}^{m} \frac{|\zeta_k|^2}{n^2k^2} = \sum_{1\le \ell,\ell' < k \le m} \frac{\sqrt{\ell(k-\ell)}\sqrt{\ell'(k-\ell')}}{2kn^2} \mathrm{T}_\ell \mathrm{T}_{k-\ell} \overline{\mathrm{T}_{\ell'}\mathrm{T}_{k-\ell'}}.$$

Moreover, according to Theorem 7.1, if $m \le n$, then it holds, for any integers $1 \le \ell, \ell' < k \le m$,

$$\begin{aligned}
\mathbb{E}_n[\mathrm{T}_\ell \mathrm{T}_{k-\ell} \overline{\mathrm{T}_{\ell'}\mathrm{T}_{k-\ell'}}] &= (\mathbf{1}_{\{\ell=\ell',\ell\ne k/2\}} + \mathbf{1}_{\{\ell=k-\ell',\ell\ne k/2\}})\mathbb{E}_n[|\mathrm{Z}_\ell|^2 |\mathrm{Z}_{k-\ell}|^2] \\
&\quad + \mathbf{1}_{\ell=\ell'=k/2}\mathbb{E}_n[|\mathrm{Z}_\ell|^4] \\
&= 4(\mathbf{1}_{\{\ell=\ell',\ell\ne k/2\}} + \mathbf{1}_{\{\ell=k-\ell',\ell\ne k/2\}}) + 8\mathbf{1}_{\ell=\ell'=k/2}.
\end{aligned}$$

This implies that

$$\mathbb{E}_n[|\mathbf{K}^{-1}\boldsymbol{\xi}|^2] = \frac{4}{n^2} \sum_{1\le\ell<k\le m} \frac{\ell(k-\ell)}{k} = \frac{(2m+5)m(m-1)}{9n^2},$$

where we have used that $\sum_{1\le\ell<k} \frac{\ell(k-\ell)}{k} = \frac{k^2-1}{6}$. □

PROOF OF LEMMA 7.5. Let us decompose $\boldsymbol{\Gamma} = \widetilde{\boldsymbol{\Gamma}} + \boldsymbol{\Delta}$, where $\boldsymbol{\Delta} = \operatorname{diag}(\boldsymbol{\Gamma})$. The point is that

$$\|\mathbf{I} - \mathbf{K}^{-1}\boldsymbol{\Gamma}\|^2 = \|\mathbf{I} - \mathbf{K}^{-1}\boldsymbol{\Delta}\|^2 + \|\mathbf{K}^{-1}\widetilde{\boldsymbol{\Gamma}}\|^2. \tag{100}$$

Since $\mathrm{X}_{2k-1} = \sqrt{\frac{2}{k}}\sum_{j=1}^n \cos(k\theta_j)$ and $\mathrm{X}_{2k} = \sqrt{\frac{2}{k}}\sum_{j=1}^n \sin(k\theta_j)$, by (95) we have, for any $k = 1, \ldots, m$,

$$\begin{aligned}
\boldsymbol{\Gamma}_{2k-1,2k-1} &= 2k\sum_{j=1}^n \sin^2(k\theta_j) = nk - \frac{k^{3/2}}{\sqrt{2}}\Re(\mathrm{T}_{2k}), \\
\boldsymbol{\Gamma}_{2k,2k} &= 2k\sum_{j=1}^n \cos^2(k\theta_j) = nk + \frac{k^{3/2}}{\sqrt{2}}\Re(\mathrm{T}_{2k}).
\end{aligned} \tag{101}$$

According to the notation of Lemma 7.3, this shows that

$$\|\mathbf{I}-\mathbf{K}^{-1}\boldsymbol{\Delta}\|^2 = \sum_{k=1}^{m} \frac{k}{n^2}\Re(\mathrm{T}_{2k})^2. \tag{102}$$

It remains to compute the second term on the RHS of (100). Let $\mathbf{K}^{1/2}$ be the positive square root of the diagonal matrix $\mathbf{K}$, and observe that, by definition of the Hilbert Schmidt norm,

$$\begin{aligned}\|\mathbf{K}^{-1}\widetilde{\boldsymbol{\Gamma}}\| &= \|\mathbf{K}^{-1/2}\widetilde{\boldsymbol{\Gamma}}\mathbf{K}^{-1/2}\|^2 \\ &= \frac{2}{n^2}\biggl(\sum_{1\le k<\ell\le m} \frac{\boldsymbol{\Gamma}_{2\ell,2k}^2+\boldsymbol{\Gamma}_{2\ell-1,2k-1}^2}{k\ell} + \sum_{1\le k\le \ell\le m} \frac{\boldsymbol{\Gamma}_{2\ell-1,2k}^2+\boldsymbol{\Gamma}_{2\ell,2k-1}^2}{k\ell}\biggr).\end{aligned} \tag{103}$$

Like (101), we can compute the coefficients on the RHS of (103). We check that, for any $1\le k\le \ell\le m$,

$$\begin{aligned}\boldsymbol{\Gamma}_{2\ell,2k} &= 2\sqrt{k\ell}\sum_{j=1}^{n}\cos(k\theta_j)\cos(\ell\theta_j) \\ &= \sqrt{k\ell\Bigl(\frac{\ell-k}{2}\Bigr)}\Re(\mathrm{T}_{\ell-k}) + \sqrt{k\ell\Bigl(\frac{\ell+k}{2}\Bigr)}\Re(\mathrm{T}_{\ell+k}), \\ \boldsymbol{\Gamma}_{2\ell-1,2k-1} &= 2\sqrt{k\ell}\sum_{j=1}^{n}\sin(k\theta_j)\sin(\ell\theta_j) \\ &= \sqrt{k\ell\Bigl(\frac{\ell-k}{2}\Bigr)}\Re(\mathrm{T}_{\ell-k}) - \sqrt{k\ell\Bigl(\frac{\ell+k}{2}\Bigr)}\Re(\mathrm{T}_{\ell+k}), \\ \boldsymbol{\Gamma}_{2\ell-1,2k} &= -2\sqrt{k\ell}\sum_{j=1}^{n}\cos(k\theta_j)\sin(\ell\theta_j) \\ &= -\sqrt{k\ell\Bigl(\frac{\ell-k}{2}\Bigr)}\Im(\mathrm{T}_{\ell-k}) - \sqrt{k\ell\Bigl(\frac{\ell+k}{2}\Bigr)}\Im(\mathrm{T}_{\ell+k}), \\ \boldsymbol{\Gamma}_{2\ell,2k-1} &= -2\sqrt{k\ell}\sum_{j=1}^{n}\sin(k\theta_j)\cos(\ell\theta_j) \\ &= +\sqrt{k\ell\Bigl(\frac{\ell-k}{2}\Bigr)}\Im(\mathrm{T}_{\ell-k}) - \sqrt{k\ell\Bigl(\frac{\ell+k}{2}\Bigr)}\Im(\mathrm{T}_{\ell+k}).\end{aligned}$$

By (103) this implies that

$$\begin{aligned}\|\mathbf{K}^{-1}\widetilde{\boldsymbol{\Gamma}}\|^2 &= \frac{2}{n^2}\biggl(\sum_{1\le k<\ell\le m}\bigl((\ell-k)\Re(\mathrm{T}_{\ell-k})^2+(\ell+k)\Re(\mathrm{T}_{\ell+k})^2\bigr) \\ &\quad + \sum_{1\le k\le \ell\le m}\bigl((\ell-k)\Im(\mathrm{T}_{\ell-k})^2+(\ell+k)\Im(\mathrm{T}_{\ell+k})^2\bigr)\biggr) \\ &= \frac{2}{n^2}\biggl(\sum_{1\le k<\ell\le m}\bigl((\ell-k)|\mathrm{T}_{\ell-k}|^2+(\ell+k)|\mathrm{T}_{\ell+k}|^2\bigr) + 2\sum_{k=1}^{m} k\Im(\mathrm{T}_{2k})^2\biggr).\end{aligned}$$

Combining the previous formula with (100) and (102), we obtain

$$\mathbb{E}_n\big[\|\mathbf{I}-\mathbf{K}^{-1}\mathbf{\Gamma}\|^2\big] = \frac{2}{n^2}\Big(\sum_{1\le k<\ell\le m} (\ell-k)\mathbb{E}_n\big[|\mathrm{T}_{\ell-k}|^2\big] + (\ell+k)\mathbb{E}_n\big[|\mathrm{T}_{\ell+k}|^2\big]\Big) + \frac{5}{n^2}\sum_{k=1}^{m} k\mathbb{E}_n\big[\Re(\mathrm{T}_{2k})^2\big],$$

where we used that the random variables $\Re(\mathrm{T}_k)$ and $\Im(\mathrm{T}_k)$ have the same law for all $k \ge 1$. Hence, by Theorem 7.1 we conclude that, if $m \le n/2$,

$$\mathbb{E}_n\big[\|\mathbf{I}-\mathbf{K}^{-1}\mathbf{\Gamma}\|^2\big] = \sum_{1\le k<\ell\le m}\frac{4\ell}{n^2} + \sum_{k=1}^{m}\frac{5k}{n^2} = \frac{(8m+7)(m+1)m}{6n^2}.$$

This completes the proof. □

APPENDIX A: ADDITIONAL PROOFS

A.1. Proof or Lemma 2.3. Without loss of generality, we assume that $\widehat{f_0}=0$, then, by (19), we have

$$\mathbb{E}_n\big[\exp \operatorname{Tr} f(\mathbf{U})\big] = e^{\mathcal{A}(f)} \det[\mathrm{I}-K_f Q_n], \tag{104}$$

where, according to (18), if we let $w = e^{-2\mathrm{i}\Im(f^+)}$, the kernel K_f is given by

$$K_f = H_+(w)H_-(\overline{w}) = H_+(w)H_+(w)^*,$$

where $H_+(w)^*$ is the adjoint of $H_+(w)$. Therefore, $Q_n K_f Q_n \ge 0$ as a trace-class operator; this implies that $\det[\mathrm{I}-K_f Q_n] \le 1$ for any $n \in \mathbb{Z}_+$. On the other hand, we see from formula (104) that $\det[\mathrm{I}-K_f Q_n] > 0$ for any $n \in \mathbb{Z}_+$. This yields the bound.

A.2. Proof of Lemma 3.1. Recall that, for any $m \in \mathbb{N}$ and $\Lambda > 0$, we let

$$g_m(\Lambda) = e^{-\Lambda^2}\sum_{1\le k\le m} m^k \Lambda^{-2k}.$$

First, by going to polar coordinates and making the change of variable $u = \|\xi\|^2$, we have, for any $\Lambda > 0$,

$$\int_{\substack{\mathbb{R}^{2m}\\ \|\xi\|\ge\Lambda}} e^{-\|\xi\|^2}\,\mathrm{d}\xi = \Omega_m \int_{\Lambda^2}^{+\infty} e^{-u}\,\mathrm{d}\big(u^m\big).$$

The integral on the RHS corresponds to the incomplete Gamma function (see [13], Formula (8.2.2)) and repeated integrations by parts give, for any $\lambda > 0$,

$$\int_{\lambda}^{+\infty} e^{-u}\,\mathrm{d}\big(u^m\big) = m!e^{-\lambda}\sum_{0\le k<m}\frac{\lambda^k}{k!}.$$

Using the bound $(m-k)! \ge \frac{m!}{m^k}$ valid for all $k = 1,\dots,m$, this implies that

$$\int_{\substack{\mathbb{R}^{2m}\\ \|\xi\|\ge\Lambda}} e^{-\|\xi\|^2}\,\mathrm{d}\xi \le \Omega_m e^{-\Lambda^2}\Lambda^{2m}\sum_{1\le k\le m} m^k\Lambda^{-2k} = \Omega_m\Lambda^{2m} g_m(\Lambda).$$

Finally, if $\Lambda^2 > m$, by summing the geometric sum, we obtain $g_m(\Lambda) \le \frac{me^{-\Lambda^2}}{\Lambda^2-m}$.

A.3. Proof of Lemma 5.2. By symmetry it suffices to prove that, for all $y \in (0,1]$ and $x \geq 0$,

$$1 + \left(\frac{\sinh(x)}{y}\right)^2 \leq \exp\left(\frac{x}{y}\right)^2.$$

We have, for any fixed $y \in (0,1]$ and $x \geq 0$,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\left(\left(1 + \left(\frac{\sinh(x)}{y}\right)^2\right)\exp\left(-\frac{x^2}{y^2}\right)\right) \\
&= -\frac{2}{y^2}\exp\left(-\frac{x^2}{y^2}\right)\left(x\left(1 + \left(\frac{\sinh(x)}{y}\right)^2\right) - \sinh(x)\cosh(x)\right) \\
&\leq -\frac{2}{y^2}\exp\left(-\frac{x^2}{y^2}\right)\left(x\left(1+\sinh(x)^2\right) - \sinh(x)\cosh(x)\right) \\
&\leq -\frac{2}{y^2}\exp\left(-\frac{x^2}{y^2}\right)\cosh(x)\big(x\cosh(x) - \sinh(x)\big).
\end{aligned}$$

Since $x\cosh(x) - \sinh(x) \geq 0$, this shows that, for any fixed $y \in (0,1]$ and $x \geq 0$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\left(1 + \left(\frac{\sinh(x)}{y}\right)^2\right)\exp\left(-\frac{x^2}{y^2}\right)\right) \leq 0.$$

This implies that, for any $y \in (0,1]$,

$$\max_{x>0}\left\{\left(1 + \left(\frac{\sinh(x)}{y}\right)^2\right)\exp\left(-\frac{x^2}{y^2}\right)\right\} = 1.$$

Since the RHS is independent of $y \in (0,1]$, this completes the proof.

A.4. Proof of Lemma 5.3. Let us define the function $\Phi(\boldsymbol{\theta}) = \sum_{1 \leq i < j \leq n} \log|e^{\mathrm{i}\theta_i} - e^{\mathrm{i}\theta_j}|^{-2}$ for $\boldsymbol{\theta} \in \triangle$, where $\triangle := \{\boldsymbol{\theta} \in \mathbb{R}^n : \theta_1 = 0 < \theta_2 < \cdots < \theta_n < 2\pi\}$ is a convex set. Observe that, by symmetry, we have

$$\max_{\theta_1,\dots,\theta_n \in \mathbb{T}}\left(\prod_{1 \leq i < j \leq n} |e^{\mathrm{i}\theta_i} - e^{\mathrm{i}\theta_j}|^2\right) = \max_{\boldsymbol{\theta} \in \triangle}\left(e^{-\Phi(\boldsymbol{\theta})}\right) = e^{-\min_{\boldsymbol{\theta} \in \triangle}\Phi(\boldsymbol{\theta})}.$$

Since function Φ is smooth on $\triangle$, by computing its Hessian (with respect to $\theta_2, \dots, \theta_n$), we verify that Φ is strictly convex.[4] Moreover, if we let $\boldsymbol{\vartheta} = (0, \frac{2\pi}{n}, \dots, \frac{2\pi(n-1)}{n})$, we see that by symmetry, for any $j = 2, \dots, n$,

$$\nabla_j \Phi(\boldsymbol{\vartheta}) = \sum_{i \neq j} \frac{1}{\tan(\frac{\vartheta_i - \vartheta_j}{2})} = \sum_{i \neq j} \frac{1}{\tan(\pi\frac{i-j}{n})} = 0.$$

This implies that $\boldsymbol{\vartheta}$ is the only critical point of Φ inside $\triangle$, and, since $\Phi = +\infty$ on $\partial\triangle$, we have

$$\min_{\boldsymbol{\theta} \in \triangle} \Phi(\boldsymbol{\theta}) = \Phi(\boldsymbol{\vartheta}).$$

Moreover, by definition of the Vandermonde determinant,

$$e^{-\Phi(\boldsymbol{\vartheta})} = \prod_{1 \leq i < j \leq n} |e^{\mathrm{i}\vartheta_i} - e^{\mathrm{i}\vartheta_j}|^2 = \big|\det_{n \times n}\big(e^{\mathrm{i}(j-1)\vartheta_i}\big)\big|^2 = \big|\det_{n \times n}\big(e^{\mathrm{i}2\pi\frac{(j-1)(i-1)}{n}}\big)\big|^2.$$

---

[4]This follows from the fact that the Hessian $\nabla^2\Phi$ has a strictly dominant diagonal with positive entries on $\triangle$.

This shows that, for any $n \geq 2$,

$$\max_{\theta_1,\dots,\theta_n \in \mathbb{T}} \prod_{1\leq i<j\leq n} |e^{\mathrm{i}\theta_i} - e^{\mathrm{i}\theta_j}|^2 = e^{-\Phi(\boldsymbol{\vartheta})} = n^n |\det_{n\times n} A|^2,$$

where $A_{ij} = \frac{e^{\mathrm{i}2\pi \frac{(j-1)(i-1)}{n}}}{\sqrt{n}}$. We easily verify that the columns of the matrix A are orthonormal so that A is a unitary matrix and $|\det_{n\times n} A| = 1$. This proves that, for any integer $n \geq 2$,

$$\max_{\theta_1,\dots,\theta_n \in \mathbb{T}} \prod_{1\leq i<j\leq n} |e^{\mathrm{i}\theta_i} - e^{\mathrm{i}\theta_j}|^2 = n^n.$$

We immediately deduce from this fact and formula (1) for the joint density of $\mathbb{P}_n$ that

$$\begin{aligned}
\mathbb{E}_n[e^{-\sum_{j=1}^n f(\theta_j)}] &= \frac{1}{n!}\int_{\mathbb{T}^n} \prod_{1\leq i<j\leq n} |e^{\mathrm{i}\theta_i} - e^{\mathrm{i}\theta_j}|^2 e^{-\sum_{j=1}^n f(\theta_j)} \frac{d\theta_1}{2\pi}\cdots\frac{d\theta_n}{2\pi} \\
&\leq \frac{e^n}{\sqrt{2\pi n}}\left(\int_{\mathbb{T}} e^{-f(\theta)}\frac{d\theta}{2\pi}\right)^n,
\end{aligned}$$

where we used that $\frac{n^n}{n!} \leq \frac{e^n}{\sqrt{2\pi n}}$ by (120).

APPENDIX B: CONSTANTS AND NUMERICAL APPROXIMATIONS

As we pointed out in the Introduction, one of the main challenge of the proof of Theorem 1.3 is to try to optimize and keep track of all the constants involved in our different estimates. For the convenience of the readers, these constants as well as the error terms in Theorem 1.3 are collected in this section. The constants are denoted by $c_j = c_j(m)$, $\epsilon_j = \epsilon_j(m)$ and $\Upsilon_j = \Upsilon_j(m)$ for $j \in \mathbb{N}_0$ since they are allowed to depend on m but not on the dimension n the random matrix. They are positive for all $m \geq 3$, and we use the following conventions:

- $c_j(m) \to \widehat{c_j}$ as $m \to +\infty$ where $\widehat{c_j} > 0$.
- $\epsilon_j(m) \to 0$ as $m \to +\infty$.
- $\Upsilon_j(m) \to +\infty$ as $m \to +\infty$ and $\Upsilon_j(m)$ is a regularly varying function.

When relevant, we also provide a numerical approximation or an estimate for these constants.

B.1. Errors in Theorem 1.3.

We let

(105) $$c_0 = \sqrt{\frac{1}{6\sqrt{2}}} \approx 0.343$$

and

(106) $$\epsilon_0(m) = \frac{2{c_0}^2}{3(1+1/m)\sqrt{1+\log m}}.$$

We note that $\epsilon_0(m) \leq 0.041$ for all $m \geq 3$.

The constants, which are directly involved in $\Theta_{N,m}$ from Theorem 1.3, are given by

$$c_1(m) = \frac{(1-c_{10})^2}{16c_{11}} - c_0{}^2(1+\epsilon_0)\frac{\sqrt{1+\log m}}{2(m+1)},$$

$$c_2(m) = c_0c_4\left(1 - c_{10} - \frac{4c_4c_0c_{11}(1+\log m)^{1/4}}{\sqrt{m+1}}\right) - c_0{}^2\frac{(1+\epsilon_0)(1+\log m)^{1/4}}{2\sqrt{m+1}},$$

$$c_3 = c_8/(2\pi)^{1/4} \approx 7.98,$$

$$c_4 = \frac{1}{2\sqrt{2}} \approx 0.354,$$

$$c_5 = c_3{}^{-1}e^{c_9} \approx 1.147,$$

$$c_6 = 4(2+\log 2) \approx 10.78,$$

$$c_7 = \frac{\pi\sqrt{e}}{2} \approx 2.59, \tag{107}$$

$$c_8 = \frac{4\cdot 2.766}{(1-1/16)^2} \approx 12.63,$$

$$c_9 = \frac{538}{243} \approx 2.21,$$

$$c_{10}(m) = c_0{}^2c_{19}(m)\sqrt{1+1/m} = c_0{}^2\frac{\sqrt{1+1/m}}{3\sqrt{6}}e^{c_0\frac{(1+\log m)^{1/4}}{\sqrt{(m+1)/2}}},$$

$$\widehat{c_{10}} = \frac{c_0{}^2}{3\sqrt{6}} \approx 0.016,$$

$$c_{11}(m) = \left(1+\frac{c_{10}}{\sqrt{m}}\right)\left(1+\frac{2c_{10}}{m}+\sqrt{\frac{1+1/m}{2(1+\log m)}}\right).$$

We verify that both $c_1(m)$ and $c_2(m)$ are increasing for $m \geq 3$, and we have

$$\widehat{c_1} = (1-\widehat{c_{10}})^2/16 \approx 0.0605, \qquad \widehat{c_2} = 4c_0c_4\sqrt{\widehat{c_1}} \approx 0.119.$$

Note also that the convergence is slow since $c_j = \widehat{c_j} + \mathcal{O}(\frac{(1+\log m)^{1/4}}{\sqrt{(m+1)}})$ for $j = 1, 2$ as $m \to +\infty$.

Let us also define for all $m \geq 1$,

$$\Upsilon_1(m) = c_6m^2 + \frac{5}{2}m\log m + \log(c_7)m + \frac{3}{2},$$
$$\Upsilon_2(m) = \frac{1}{2}m\log m - \frac{3}{4}m\log(1+\log m) - m\log(8c_0) + \frac{1}{2}. \tag{108}$$

It will turn out that we need the following estimates for the functions $\Upsilon_1(m)$ and $\Upsilon_2(m)$. We have, for all $m \geq 3$,

$$\sqrt{c_1(m)^{-1}(1+\log m)\Upsilon_1(m)} \geq 34m \tag{109}$$

and

$$\frac{c_1(m)\Upsilon_2(m)\sqrt{m+1}}{c_2(m)(1+\log m)^{1/4}} \leq \frac{\Upsilon_1(m)}{1500}. \tag{110}$$

The numerical constants in (109) and (110) are not optimal, but they suffice for our applications.

For any $N, m \geq 1$, let us define the following functions:

$$\Theta^0_{N,m} = m^{\frac{5}{2}} 2^{\frac{m}{2}} e^{\frac{m^2}{4N}} \frac{e^{\frac{N}{2}}(1+\log m)^N}{\sqrt{N}\Gamma(N+1)},$$

(111)

$$\Theta^3_{N,m} = \frac{{c_3}^{-1}{c_4}^m N^{\frac{m}{2}}}{(1+\log m)^{\frac{m}{2}}} \exp\biggl(-\frac{N^2}{16(1+\log m)}\biggr),$$

(112)
$$\Theta^2_{N,m} = c_5 N^{\frac{m}{2}} \exp\biggl(\Upsilon_2(m) - \frac{c_2(m)N^2}{\sqrt{m+1}(1+\log m)^{\frac{3}{4}}}\biggr)$$

and, if $N > 4m$,

(113)
$$\Theta^1_{N,m} = \frac{c_5 e^{-c_9 \frac{4m}{N}}}{(1-\frac{4m}{N})^{1-\frac{2m}{N}}} \exp\biggl(\Upsilon_1(m) - c_1(m)\frac{N(N-4m)}{(1+\log m)}\biggr).$$

Then, the error in Theorem 1.3 is given by

(114)
$$\Theta_{N,m} = \Theta^0_{N,m} + \Theta^1_{N,m} + \Theta^2_{N,m} + \Theta^3_{N,m}.$$

One should keep in mind that $\Theta^0_{N,m}$ is the main term; the term $\Theta^3_{N,m}$ is always negligible, while $\Theta^1_{N,m}$ and $\Theta^2_{N,m}$ are corrections which become negligible when $m \ll N$. This is quantified by Proposition B.1 below.

In Sections 2–6, the following constants will come in play:

(115)

$$c_{12} = \frac{1+\sqrt{290}}{17} \approx 1.06,$$

$$c_{13} = \frac{(\log 108)\sqrt{1+\log 3}}{68\sqrt{108}c_2(m)},$$

$$c_{14} = \frac{\sqrt{13}}{1500} \approx 0.0024,$$

$$c_{15} = 2e^2 \approx 14.78,$$

$$c_{16} = e/\sqrt{\pi} \approx 1.534,$$

$$c_{17} = \frac{32}{3}\biggl(1 + \frac{(2-1/m)^3}{3(m+1)}\biggr),$$

$$c_{18} = \frac{8}{3}\biggl(1 + \frac{4(1-1/m)^3}{3(m+1)}\biggr),$$

$$c_{19}(m) = \frac{1}{3\sqrt{6}} e^{\sqrt{2}c_0 \frac{(1+\log m)^{1/4}}{\sqrt{m+1}}},$$

$$c_{20}(\eta) = \frac{\pi^2\eta^2}{8}, \qquad \eta = \frac{1/\pi}{\sqrt{m}},$$

$$c_{21}(m,\eta) = \frac{\exp(\frac{\eta}{2\sqrt{m(m+1)}})}{6\sqrt{3}}, \qquad \eta = \frac{1/\pi}{\sqrt{m}},$$

$$\Upsilon_3(m) = \frac{\pi m^{3/2}(m+1)e^{\frac{1}{2}(1+\frac{1/2}{(\pi m)^2})}}{2(1-\frac{c_{21}}{4\pi^2 m^3})}.$$

Observe that $\Upsilon_3(m) = c_7 m^{\frac{5}{2}}(1 + \mathcal{O}(m^{-1}))$ as $m \to +\infty$. Moreover, as $c_{21} \leq \frac{1}{12}$, we verify that, for all $m \geq 3$,

$$\Upsilon_3{}^m \leq c_7{}^m m^{\frac{5m}{2}} \frac{e^{\frac{1}{4\pi^2 m}}(1+1/m)^m}{(1 - \frac{m^{-3}}{48\pi^2})^m} \leq e c_7{}^m m^{\frac{5m}{2}}. \tag{116}$$

B.2. Estimates for errors—Proof of Proposition 1.4.

PROPOSITION B.1. *For all $N, m \geq 3$ such that $N \geq \gamma\sqrt{c_1(m)^{-1}(1+\log m)\Upsilon_1(m)}$ and $\gamma > c_{12} = \frac{1+\sqrt{290}}{17}$, we have the estimates*

$$\Theta^1_{N,m} \leq c_5 \exp\left(-\left(1 - \frac{2\gamma^{-1}}{17} - \gamma^{-2}\right)\frac{c_1(m)N^2}{1+\log m}\right) \tag{117}$$

and

$$\begin{aligned}\Theta^2_{N,m} &\leq c_5 N^{\frac{m}{2}} \exp\left(-\frac{(1-\frac{\gamma^{-2}}{1500})c_2(m)N^2}{\sqrt{m+1}(1+\log m)^{\frac{3}{4}}}\right) \\ &\leq c_5 \exp\left(-(\sqrt{13\gamma} - c_{13}\gamma^{-1} - c_{14}\gamma^{-\frac{3}{2}})\frac{c_2(m)N^{\frac{3}{2}}}{\sqrt{1+\log m}}\right).\end{aligned} \tag{118}$$

Moreover, we have the lower bounds, $c_1(m) \geq 0.0148$, $c_2(m) \geq 0.077$ *and* $c_{13} \leq 0.125$ *for all* $m \geq 3$.

PROOF. Since $e^{c_9} \geq 4$, we verify that the function $x \mapsto e^{-c_9 x}(1-x)^{1-2x}$ is decreasing on $[0, \frac{1}{2}]$ so that we deduce from (113) that, for all $N \geq 8m$,

$$\Theta^1_{N,m} \leq c_5 \exp\left(\Upsilon_1(m) - c_1(m)\frac{N(N-4m)}{(1+\log m)}\right).$$

Then, we verify that, if the condition $N \geq \gamma\sqrt{c_1(m)^{-1}(1+\log m)\Upsilon_1(m)}$ holds with $\gamma > 0$,

$$c_1(m)\frac{N(N-4m)}{(1+\log m)} - \Upsilon_1(m) \geq \frac{c_1(m)N^2}{(1+\log m)}\left(1 - \frac{2\gamma^{-1}}{17} - \gamma^{-2}\right), \tag{119}$$

where we used the lower bound (109). The RHS of (119) is positive so long as $\gamma > c_{12} = \frac{1+\sqrt{290}}{17}$, and this yields the estimate (117). For the estimate (118), let us also observe that, according to (110), we have, for all $N, m \geq 3$ such that $N \geq \gamma\sqrt{c_1(m)^{-1}(1+\log m)\Upsilon_1(m)}$,

$$c_2(m)^{-1}\Upsilon_2(m)\sqrt{m+1}(1+\log m)^{3/4} \leq \frac{\gamma^{-2}}{1500}N^2.$$

This implies that

$$\Theta^2_{N,m} \leq c_5 N^{\frac{m}{2}} \exp\left(-\left(1 - \frac{\gamma^{-2}}{1500}\right)\frac{c_2(m)N^2}{\sqrt{m+1}(1+\log m)^{\frac{3}{4}}}\right).$$

Moreover, we also verify that, for all $m \geq 3$,

$$(m+1)^2 \leq \frac{c_1(m)^{-1}\Upsilon_1(m)}{13^2}$$

so that under our hypothesis,

$$\begin{aligned}\frac{N^2}{\sqrt{m+1}(1+\log m)^{\frac{3}{4}}} &\geq \frac{\sqrt{13}N^2}{\sqrt{1+\log m}(c_1(m)^{-1}(1+\log m)\Upsilon_1(m))^{\frac{1}{4}}} \\ &\geq \frac{\sqrt{13\gamma}N^{3/2}}{\sqrt{1+\log m}}.\end{aligned}$$

Hence, if we agree to loose the Gaussian decay in N of $\Theta^2_{N,m}$, we obtain that, for all $N, m \geq 3$ such that $N \geq \gamma\sqrt{c_1(m)^{-1}(1+\log m)\Upsilon_1(m)}$,

$$\Theta^2_{N,m} \leq c_5 N^{\frac{m}{2}} \exp\biggl(-\biggl(\sqrt{13\gamma} - \frac{\sqrt{13}}{1500}\gamma^{-\frac{3}{2}}\biggr)\frac{c_2(m)N^{\frac{3}{2}}}{\sqrt{1+\log m}}\biggr).$$

Finally, it follows from (109) that, under our hypothesis, $N \geq 34\gamma m$ with $\gamma > c_{12}$ so that $N \geq 108$ and

$$\begin{aligned}N^{\frac{m}{2}} &\leq \exp\biggl(\frac{N\log N}{68\gamma}\biggr) \\ &\leq \exp\biggl(\frac{N^{\frac{3}{2}}}{\sqrt{1+\log m}}\frac{\log N\sqrt{1+\log(N/36)}}{68\gamma\sqrt{N}}\biggr) \\ &\leq \exp\biggl(\frac{c_{13}c_2(m)N^{\frac{3}{2}}}{\gamma\sqrt{1+\log m}}\biggr),\end{aligned}$$

where $c_{13} = \frac{(\log 108)\sqrt{1+\log 3}}{68\sqrt{108}c_2(m)}$. This yields the estimate (118). Since c_1, c_2 are increasing functions for $m \geq 3$, we obtain the numerical estimates for c_1, c_2 and c_{13} by evaluating these functions for $m = 3$ on *Mathematica*. This completes the proof. □

We will also need the following basic estimates for the main error term $\Theta^0_{N,m}$.

LEMMA B.2. *For all $m, N \in \mathbb{N}$ such that $m \geq 3$ and $N \geq 5m$, it holds*

$$\Theta^0_{N,m} \leq \frac{1}{\sqrt{\pi}}\exp\biggl(-N\log m\biggl(1 - \frac{\log(1+\log m)}{\log m}\biggr)\biggr).$$

PROOF. Let us recall from [13], Formula (5.6.1), that, for any $x > 0$,

$$\Gamma(x+1) = \sqrt{2\pi x}x^x \exp\biggl(-x + \frac{\theta_x}{12x}\biggr) \quad \text{where } \theta_x \in (0,1). \tag{120}$$

In addition, since $e^{\frac{3}{2}} \leq 5$, let us observe that we have, for all $N \geq 5m$,

$$\frac{m^{\frac{3}{2}}2^{\frac{m}{2}}e^{\frac{m^2}{4N}}e^{\frac{3N}{2}}}{\sqrt{2\pi}5^{N+1}} \leq \frac{m^{\frac{3}{2}}e^{-cm}}{5\sqrt{2\pi}} \leq \frac{1}{\sqrt{\pi}},$$

where we used that $c = 5\log 5 - \frac{15}{2} - \frac{1}{20} - \frac{\log 2}{2} \geq 0.15$. By (120) this implies that, for all $N \geq 5m$,

$$\Theta^0_{N,m} \leq \frac{m^{\frac{3}{2}} 2^{\frac{m}{2}} e^{\frac{m^2}{4N}} e^{\frac{3N}{2}}}{\sqrt{2\pi}5^{N+1}} \frac{(1+\log m)^N}{m^N} \leq \frac{1}{\sqrt{\pi}} \exp\Bigl(-N\log m\Bigl(1 - \frac{\log(1+\log m)}{\log m}\Bigr)\Bigr). \quad \square$$

Using the previous estimates, we are now ready to prove Proposition 1.4.

PROPOSITION B.3. *Fix $M \geq 3$. For all $m \geq M$ and $N \geq c(M)m\sqrt{1+\log m}$, we have*

$$\Theta^1_{N,m} + \Theta^2_{N,m} + \Theta^3_{N,m} \leq 0.015 \frac{(1+\log m)^N e^{\frac{N}{2}}}{\sqrt{N}\Gamma(N+1)} \leq \epsilon \Theta^0_{N,m}, \tag{121}$$

where $\epsilon \leq 34 \cdot 10^{-5}$ and the constant $c(M)$ are explicitly given by the Table (125) below. Moreover, under the same conditions we also have $\Theta^0_{N,m} \leq N^{-\frac{m}{2}} \frac{\exp(-12m(\log m - 0.16))}{\sqrt{\pi}{c_{16}}^m}$.

PROOF. First, observe that, for all $N \geq \gamma\sqrt{c_1(m)^{-1}(1+\log m)\Upsilon_1(m)}$, if $\theta(m) := \sqrt{\frac{c_1(m)\Upsilon_1(m)}{1+\log m}} \geq (\gamma - \frac{2}{17} - \gamma^{-1})^{-1}$, then it holds

$$\frac{e^{-3/2}\cdot N}{1+\log m} \exp\Bigl(-\frac{(1 - \frac{2\gamma^{-1}}{17} - \gamma^{-2})c_1 N}{1+\log m}\Bigr) \leq {c_1}^{-1}\gamma\theta e^{-3/2-(\gamma-2/17-\gamma^{-1})\theta}.$$

Let us suppose that $\gamma \leq 5.12$. This shows that if we choose γ, depending on $m \geq 3$, in such a way that

$$(\gamma - 2/17 - \gamma^{-1})\theta \geq \log(5.12{c_1}^{-1}\theta) - 1.48 > 0; \tag{122}$$

then,

$$\frac{e^{-3/2}\cdot N}{1+\log m} \exp\Bigl(-\frac{(1 - \frac{2\gamma^{-1}}{17} - \gamma^{-2})c_1 N}{1+\log m}\Bigr) \leq e^{-0.02}.$$

We can choose a (numerical) solution $\gamma(m)$ of (122) which is nonincreasing,

(123)

m	3	4	5	6	8	12	17	23	30	40	≥ 70
$\gamma(m)$	5.119	3.806	3.149	2.754	2.30	1.882	1.65	1.507	1.413	1.334	1.230

This solution satisfies our requirements $c_{12} < \gamma \leq 5.12$ and $(\gamma - 2/17 - \gamma^{-1})\theta \geq 1$, and we also check that the function $m^{-1}\sqrt{c_1(m)^{-1}\Upsilon_1(m)}$ is decreasing for $m \geq 3$. Then, by (117) this implies that, for any $M \geq 3$, if $m \geq M$ and $N \geq c(M)m\sqrt{1+\log m}$,

$$\begin{aligned} \Theta^1_{N,m} &\leq c_5 \exp\Bigl(-\frac{(1 - \frac{2\gamma^{-1}}{17} - \gamma^{-2})c_1 N^2}{1+\log m}\Bigr) \\ &\leq c_5 e^{-0.02N}\Bigl(\frac{e^{-3/2}\cdot N}{1+\log m}\Bigr)^{-N}, \end{aligned} \tag{124}$$

where $c(M) = \gamma(M)M^{-1}\sqrt{c_1(M)^{-1}\Upsilon_1(M)}$ is a decreasing function given by

(125)

M	3	4	5	6	8	12	17	23	30	40	70
$c(M)$	146.5	93.8	71.1	58.66	45.5	34.5	28.8	25.5	23.4	21.64	19.4

With this choice we check numerically that $N \geq 44m$ and $N \geq 575$. Then, in the regime that we consider

$$2.51 c_5 N e^{-0.02N} \leq 14.63 \cdot 10^{-3}$$

so that by (124) and using that according to (120), $\Gamma(N+1) \leq 2.51 N^{N+1/2} e^{-N}$, we obtain for $m \geq M$ and $N \geq c(M)m\sqrt{1+\log m}$,

$$\Theta^1_{N,m} \leq 14.63 \cdot 10^{-3} \frac{(1+\log m)^N e^{\frac{N}{2}}}{\sqrt{N}\Gamma(N+1)}. \tag{126}$$

By a similar argument using (111) and that in the regime that we consider $\frac{N}{1+\log m} \geq \min_{M \geq 3} \frac{c(M)M}{\sqrt{1+\log M}} \geq 207$, we obtain

$$\Theta^3_{N,m} \leq \frac{{c_3}^{-1} {c_4}^3}{(1+\log 3)^{3/2}} \left(\frac{N}{1+\log m}\right)^{\frac{N}{88}} \exp\left(-\frac{N^2}{16(1+\log m)}\right).$$

Then, since $\min_{x \geq 207}\{\frac{x}{16} - \frac{89}{88}\log x\} \geq 7.544$, this implies that

$$\Theta^3_{N,m} \leq 1.83 \cdot 10^{-3} N e^{-7.544N} \frac{(1+\log m)^N e^{\frac{N}{2}}}{\sqrt{N}\Gamma(N+1)}. \tag{127}$$

Using the estimate (118) and the fact that, according to the Table (123)

$$\min_{m \geq 3}\{(\sqrt{13\gamma(m)} - c_{13}/\gamma(m) - c_{14}/\gamma(m)^{\frac{3}{2}})c_2(m)\} \geq 0.422,$$

we obtain the estimate

$$\Theta^2_{N,m} \leq c_5 \exp\left(-\frac{0.422 \cdot N^{\frac{3}{2}}}{\sqrt{1+\log m}}\right).$$

Since we have seen that $\frac{N}{1+\log m} \geq 207$, this implies that

$$\frac{N}{1+\log m} \exp\left(-\frac{0.422\sqrt{N}}{\sqrt{1+\log m}}\right) \leq \max_{x \geq \sqrt{207}}\{x^2 e^{-0.422x}\} \leq e^{-.7388}.$$

So, using the same argument once more, we obtain that, for any $m \geq M$ and $N \geq c(M)m\sqrt{1+\log m}$,

$$\Theta^2_{N,m} \leq 3N e^{-1.2388N} \frac{(1+\log m)^N e^{\frac{N}{2}}}{\sqrt{N}\Gamma(N+1)}. \tag{128}$$

By combining the estimates (126), (127) and (128), we easily verify that, for any $m \geq M$ and for all $N \geq c(M)m\sqrt{1+\log m}$,

$$\Theta^1_{N,m} + \Theta^2_{N,m} + \Theta^3_{N,m} \leq 0.015 \frac{(1+\log m)^N e^{\frac{N}{2}}}{\sqrt{N}\Gamma(N+1)}.$$

Then, from (111) we deduce the bound (121) with $\epsilon \leq 0.011 \cdot 3^{-\frac{5}{2}} 2^{-\frac{3}{2}} \leq 34 \cdot 10^{-5}$. Finally, it remains to obtain the upper bound for $\Theta^0_{N,m}$. According to Lemma B.2, we verify that, for all $m \geq 3$ and $N \geq 5m$,

$$N^{\frac{m}{2}} \Theta^0_{N,m} \leq \frac{1}{\sqrt{\pi}} \exp(-0.3N \log m + 0.5m \log N).$$

Since the function $N \mapsto 0.3N \log m - 0.5m \log N$ is increasing and $\min_{M \geq 3} c(M) \times \sqrt{1 + \log M} \geq 44$, this implies that, for $N \geq 44m$,

$$N^{\frac{m}{2}} \Theta^0_{N,m} \leq \frac{\exp(-12m \log m + \frac{\log 42}{2} m)}{\sqrt{\pi}} \leq \frac{\exp(-13.2m(\log m - 0.16))}{\sqrt{\pi} c_{16}{}^m},$$

where we used that $\frac{\log(44 c_{16})}{26.2} \leq 0.16$. $\square$

B.3. Numerics for $m = 3$.

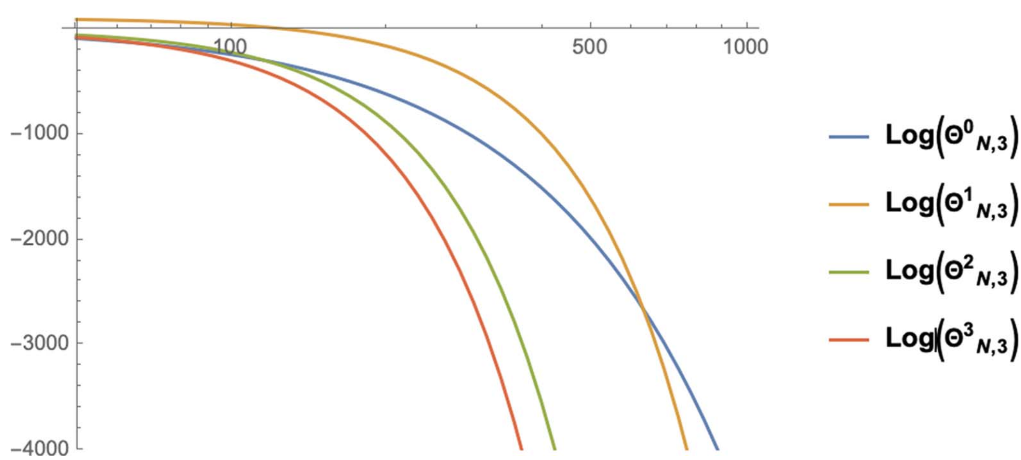

FIG. 1. *Log-plot of the errors* (111)–(112) *for $m = 3$ as functions of $N = n/3$ where n is the dimension of a random unitary matrix. We observe that $\Theta^0_{N,3} \geq \Theta^1_{N,3}$ when $N \geq 631$ which is consistent with the threshold* $3c(3)\sqrt{1 + \log 3} \approx 637$ *from Proposition* B.3. *Note that if m is small, then the error $\Theta^1_{N,3}$ coming from the* tail regime *is most significant. On the other hand, for large m, the error $\Theta^2_{N,3}$ coming from the* intermediate regime *becomes the most relevant.*

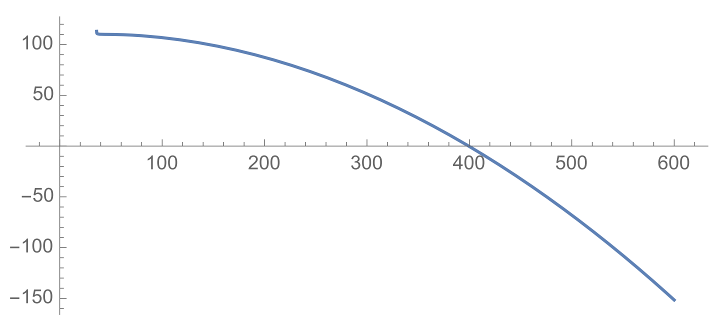

FIG. 2. *To illustrate our estimate for $\Delta^{(2)}_{n,3}$, we plot the RHS of* (8) *for $m = 3$ that is,* $\log(\frac{8\sqrt{\Omega_m}}{(2\pi)^m}(\frac{n}{m})^{\frac{m}{2}})$ *as a function of the dimension n of a random unitary matrix. By Theorem* 1.3, *this controls the total variation distance between* $\mathbf{X}$ *and a standard Gaussian vector in $\mathbb{R}^6$. This quantity is large for small n because of the error $\Theta^1_{N,3}$ and the fact that $\Upsilon_1(3) \simeq 110$,* (108). *We observe that our estimates become relevant as soon as $n \geq 400$ which can still be considered a small size random matrix.*

Acknowledgments. The authors would like to thank the anonymous referee for her/his constructive comments.

Funding. K.J. was supported by the grant KAW 2015.0270 from the Knut and Alice Wallenberg Foundation and the Swedish Research Council grant 2015-0487. G.L.'s research is supported by the SNSF Ambizione grant S-71114-05-01.

REFERENCES

[1] BASOR, E. L. and WIDOM, H. (2000). On a Toeplitz determinant identity of Borodin and Okounkov. *Integral Equations Operator Theory* **37** 397–401. MR1780119 https://doi.org/10.1007/BF01192828
[2] BAXTER, G. (1961). Polynomials defined by a difference system. *J. Math. Anal. Appl.* **2** 223–263. MR0126125 https://doi.org/10.1016/0022-247X(61)90033-6
[3] BEREZIN, S. and BUFETOV, A. I. (2021). On the rate of convergence in the central limit theorem for linear statistics of Gaussian, Laguerre, and Jacobi ensembles. *Pure Appl. Funct. Anal.* **6** 57–99. MR4213299
[4] BORODIN, A. and OKOUNKOV, A. (2000). A Fredholm determinant formula for Toeplitz determinants. *Integral Equations Operator Theory* **37** 386–396. MR1780118 https://doi.org/10.1007/BF01192827
[5] BÖTTCHER, A. (2002). On the determinant formulas by Borodin, Okounkov, Baik, Deift and Rains. In *Toeplitz Matrices and Singular Integral Equations* (*Pobershau*, 2001). *Oper. Theory Adv. Appl.* **135** 91–99. Birkhäuser, Basel. MR1935759
[6] CHAHKIEV, M. A. (2008). On oscillatory integral Hilbert transformation with trigonometric polynomial phase. *Anal. Math.* **34** 177–185. MR2434672 https://doi.org/10.1007/s10476-008-0302-7
[7] COURTEAUT, K. and JOHANSSON, K. Multivariate normal approximation for traces of orthogonal and symplectic matrices. Available at abs/2103.03791.
[8] COURTEAUT, K., JOHANSSON, K. and LAMBERT, G. In preparation.
[9] DEIFT, P., ITS, A. and KRASOVSKY, I. (2013). Toeplitz matrices and Toeplitz determinants under the impetus of the Ising model: Some history and some recent results. *Comm. Pure Appl. Math.* **66** 1360–1438. MR3078693 https://doi.org/10.1002/cpa.21467
[10] DIACONIS, P. (1994). Personal communication.
[11] DIACONIS, P. and EVANS, S. N. (2001). Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** 2615–2633. MR1828463 https://doi.org/10.1090/S0002-9947-01-02800-8
[12] DIACONIS, P. and SHAHSHAHANI, M. (1994). On the eigenvalues of random matrices *J. Appl. Probab.* **31A** 49–62.
[13] DLMF. Nist digital library of mathematical functions. http://dlmf.nist.gov/.
[14] DÖBLER, C. and STOLZ, M. (2011). Stein's method and the multivariate CLT for traces of powers on the classical compact groups. *Electron. J. Probab.* **16** 2375–2405. MR2861678 https://doi.org/10.1214/EJP.v16-960
[15] GERONIMO, J. S. and CASE, K. M. (1979). Scattering theory and polynomials orthogonal on the unit circle. *J. Math. Phys.* **20** 299–310. MR0519213 https://doi.org/10.1063/1.524077
[16] GESSEL, I. M. (1990). Symmetric functions and P-recursiveness. *J. Combin. Theory Ser. A* **53** 257–285. MR1041448 https://doi.org/10.1016/0097-3165(90)90060-A
[17] GOLINSKIĬ, B. L. and IBRAGIMOV, I. A. (1971). A limit theorm of G. Szegő. *Izv. Ross. Akad. Nauk Ser. Mat.* **35** 408–427. MR0291713
[18] GORODETSKY, O. and RODGERS, B. Traces of powers of matrices over finite fields. *Trans. Amer. Math. Soc.* To appear. Available at pdf/1909.03666.pdf.
[19] IBRAGIMOV, I. A. (1968). A theorem of Gabor Szegő. *Mat. Zametki* **3** 693–702. MR0231114
[20] JIANG, T. and MATSUMOTO, S. (2015). Moments of traces of circular beta-ensembles. *Ann. Probab.* **43** 3279–3336. MR3433582 https://doi.org/10.1214/14-AOP960
[21] JOHANSSON, K. (1988). On Szegő's asymptotic formula for Toeplitz determinants and generalizations. *Bull. Sci. Math.* (2) **112** 257–304. MR0975365
[22] JOHANSSON, K. (1997). On random matrices from the compact classical groups. *Ann. of Math.* (2) **145** 519–545. MR1454702 https://doi.org/10.2307/2951843
[23] LAMBERT, G. (2021). Mesoscopic central limit theorem for the circular β-ensembles and applications. *Electron. J. Probab.* **26** Paper No. 7, 33. MR4216520 https://doi.org/10.1214/20-ejp559
[24] LAMBERT, G., LEDOUX, M. and WEBB, C. (2019). Quantitative normal approximation of linear statistics of β-ensembles. *Ann. Probab.* **47** 2619–2685. MR4021234 https://doi.org/10.1214/18-AOP1314
[25] SARNAK, P. (2019). Memorial conference for I. Piatetski-Shapiro: An underdetermined matrix moment problem and its applications to computing zeros of l-functions. http://publications.ias.edu/sarnak/paper/2705.

[26] SIMON, B. (2005). *Orthogonal Polynomials on the Unit Circle*. *Part* 1: *Classical Theory*. *American Mathematical Society Colloquium Publications* **54**. Amer. Math. Soc., Providence, RI. MR2105088 https://doi.org/10.1090/coll054.1

[27] SIMON, B. (2005). *Trace Ideals and Their Applications*, 2nd ed. *Mathematical Surveys and Monographs* **120**. Amer. Math. Soc., Providence, RI. MR2154153 https://doi.org/10.1090/surv/120

[28] SZEGÖ, G. (1952). On certain Hermitian forms associated with the Fourier series of a positive function. *Comm*. *Sém*. *Math*. *Univ*. *Lund* [*Medd*. *Lunds Univ*. *Mat*. *Sem*.] **1952** 228–238. MR0051961

[29] WEBB, C. (2016). Linear statistics of the circular β-ensemble, Stein's method, and circular Dyson Brownian motion. *Electron*. *J*. *Probab*. **21** Paper No. 25, 16. MR3485367 https://doi.org/10.1214/16-EJP4535