*The Annals of Applied Probability*
2016, Vol. 26, No. 5, 2733–2753
DOI: 10.1214/15-AAP1160


# WHAT IS THE PROBABILITY THAT A LARGE RANDOM MATRIX HAS NO REAL EIGENVALUES?

BY EUGENE KANZIEPER[*,†,1], MIHAIL POPLAVSKYI[‡,2], CARSTEN TIMM[§], ROGER TRIBE[‡,2] AND OLEG ZABORONSKI[‡,3]

*Holon Institute of Technology*[*], *Weizmann Institute of Science*[†], *University of Warwick*[‡] *and Technische Universität Dresden*[§]

We study the large-$n$ limit of the probability $p_{2n,2k}$ that a random $2n \times 2n$ matrix sampled from the real Ginibre ensemble has $2k$ real eigenvalues. We prove that

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,2k} = \lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,0} = -\frac{1}{\sqrt{2\pi}}\zeta\left(\frac{3}{2}\right),$$

where $\zeta$ is the Riemann zeta-function. Moreover, for any sequence of non-negative integers $(k_n)_{n\geq 1}$,

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,2k_n} = -\frac{1}{\sqrt{2\pi}}\zeta\left(\frac{3}{2}\right),$$

provided $\lim_{n\to\infty}(n^{-1/2}\log(n))k_n = 0$.

**1. Introduction and the main result.** Our paper is dedicated to the study of the probability $p_{2n,2k}$ that a real $2n \times 2n$ random matrix with independent normal entries (the so-called "real Ginibre matrix") has $2k$ real eigenvalues. It has been known since [10] that a typical large $N \times N$ Ginibre matrix has $O(\sqrt{N})$ real eigenvalues. What is the probability of rare events consisting of such a matrix having either anomalously many or few real eigenvalues?

The former question has been addressed by many authors. Building on the original work by Ginibre [13], Edelman used the real Schur decomposition to prove that

$$p_{N,N} = \left(\frac{1}{2}\right)^{N(N-1)/4},$$

see [9]. In [2], Akemann and Kanzieper employed the method of skew-orthogonal polynomials to determine the probability that all but two eigenvalues of a real Ginibre matrix are real. In the large-$N$ limit, their result reads

$$p_{N,N-2} = e^{-(\log(2)/4)N^2+(\log(3\sqrt{2})/2)N+o(N)}, \tag{1.1}$$

Received April 2015.
[1]Supported by the Israel Science Foundation Grant No. 647/12.
[2]Supported by EPSRC Grant No. EP/K011758/1.
[3]Supported by a Leverhulme Trust Research Fellowship.
*MSC2010 subject classifications.* Primary 60B20; secondary 60F10.
*Key words and phrases.* Real Ginibre ensemble, large deviations.

where $\lim_{N\to\infty} o(N)/N = 0$. These answers were generalised in a very recent paper [7] where the large deviations principle of [3] was extended to prove that the probability that a real Ginibre matrix has $\alpha N$ (where $0 < \alpha < 1$) real eigenvalues is $p_{N,\alpha N} \overset{N\to\infty}{\sim} e^{-N^2 I_\alpha}$, where the symbol "$\sim$" denotes the logarithmic asymptotic equivalence and the constant $I_\alpha$ is characterised as the minimal value of an explicitly given rate functional; see Proposition 2 and formula (4) of [7].

In the present paper, we answer the question about the probability that a real Ginibre matrix has very few real eigenvalues.

THEOREM 1.1. *Let $G_{2n}$ be a random $2n \times 2n$ real matrix with independent $N(0, 1)$ matrix elements. Let $p_{2n,2k}$ be the probability that $G_{2n}$ has $2k$ real eigenvalues. Then for any fixed $k = 0, 1, 2, 3, \dots$,*

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,2k} = -\frac{1}{\sqrt{2\pi}} \zeta\left(\frac{3}{2}\right), \tag{1.2}$$

*where $\zeta$ is the Riemann zeta-function. Moreover,*

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,2k_n} = -\frac{1}{\sqrt{2\pi}} \zeta\left(\frac{3}{2}\right), \tag{1.3}$$

*where $(k_n)_{n\geq 1}$ is a sequence of nonnegative integers such that*

$$\lim_{n\to\infty} (n^{-1/2} \log(n)) k_n = 0.$$

In particular, the probability that a large $2n \times 2n$ Ginibre matrix has no real eigenvalue behaves as

$$p_{2n,0} \overset{n\to\infty}{\sim} e^{-\sqrt{n/\pi}\zeta(3/2)+o(\sqrt{n})}.$$

Notice that the answer (1.2) is qualitatively different from the results for the probability of having $O(n)$ real eigenvalues quoted above: the "cost" of having $O(n)$ real eigenvalues normalised by the total number of "anomalous" eigenvalues increases linearly with $n$, whereas the "cost" of removing all real eigenvalues from the real axis is constant per eigenvalue.

It is also worth noting that our result "almost" extends to the typical region $k \sim n^{1/2}$ (e.g., we can choose $k_n = [\sqrt{n}/\log^2 n]$ in (1.3)). It would be interesting to see if (1.3) survives for $k_n = [c\sqrt{n}]$ where $c \ll 1$.

The statement of the theorem can be *guessed* using existing results: in the limit $N \to \infty$, the unscaled law of real eigenvalues for the real Ginibre $N \times N$ ensemble converges. The limit coincides with the $t = 1$ law for the $A + A \to \varnothing$ interacting particle system on $\mathbb{R}$ [16]. The probability that an interval of length $s$ has no particles for $A + A \to \varnothing$ has been calculated formally by Derrida and Zeitak [8]. These two facts allowed Forrester [11] to conclude that the large-$N$ limit of the probability that there are no real eigenvalues in the interval $(a, a + s)$ should be given by

$$\text{Prob}\big[G_\infty \text{ has no eigenvalues in } (a, a+s)\big] \overset{s\to\infty}{\sim} e^{-(1/(2\sqrt{2\pi}))\zeta(3/2)s}. \tag{1.4}$$

Let us stress that equation (1.4) is valid for $N=\infty$ only. However, we know from the work of Borodin and Sinclair [6] and Forrester and Nagao [12] that the law of real eigenvalues for the real Ginibre ensemble is a Pfaffian point process for all values of $N \le \infty$. Convergence of the finite-$N$ kernel to the $N=\infty$ kernel is exponentially fast within the spectral radius. The spectral radius is $R_N = \sqrt{N} + O(1)$ [10]. We also know that the boundary effects for a large but finite matrix size $N$ are only felt in the boundary layer of the width of order 1 near the edge. Therefore, the simplest finite-$N$ guess for Prob[$G_N$ has no real eigenvalues] is

$$\begin{aligned}&\text{Prob}[G_N \text{ has no real eigenvalues}]\\&\qquad\approx \text{Prob}\big[G_N \text{ has no real eigenvalues in } (-R_N+L, R_N-L)\big]\\&\qquad\approx \text{Prob}\big[G_\infty \text{ has no real eigenvalues in } (-R_N, R_N)\big].\end{aligned}$$

Here, $L \gg 1$ is a large $N$-independent constant. The last probability in our heuristic chain of arguments can be approximated using (1.4) with $s = 2R_N$. This suggests

$$\text{Prob}[G_N \text{ has no real eigenvalues}] \approx e^{-(1/\sqrt{2\pi})\zeta(3/2)\sqrt{N}},$$

which agrees with the statement of Theorem 1.1.

The value of the constant which defines the rate of decay of $p_{2n,0}$ in (1.2) is

$$\frac{1}{\sqrt{2\pi}}\zeta(3/2) \approx 1.0422,$$

which is consistent with its numerical estimate; see Figure 1. The numerical analysis of the exact formula for $p_{2n,0}$ [see (2.2) below] also shows that under the assumption that the next-to-leading term in the large-$N$ expansion of $p_{N,0}$ is constant, the resulting coefficient ($\approx 0.06267$) is close to its exact counterpart from the large gap size expansion of the Derrida–Zeitak formula ($\approx 0.0627$). At the moment, we do not have a theory explaining this closedness.

Both the numerical simulations and the heuristic argument given above provide a strong hint in favour of Theorem 1.1.

There are several possible routes to the proof of the theorem. For example, one can try to use Forrester's observation, coupled with the knowledge of the rate of convergence of the Borodin–Sinclair–Forrester–Nagao kernel in the large-$N$ limit, to show that the errors in applying Derrida–Zeitak's formula to gaps of $N$-dependent sizes vanish as $N \to \infty$. There is however a problem with this approach: in the case we are interested in (annihilating Brownian motions or the 2-state Potts model) the infinite sums entering the gap formula converge only polynomially; see [8] for details. Therefore, a careful justification would be required for the validity of the interchange of summation and taking the large gap size limit. We feel that such a justification is best done in the context of a general theory of "Fredholm Pfaffians". In this paper, we will adopt *the spirit* of Derrida–Zeitak's

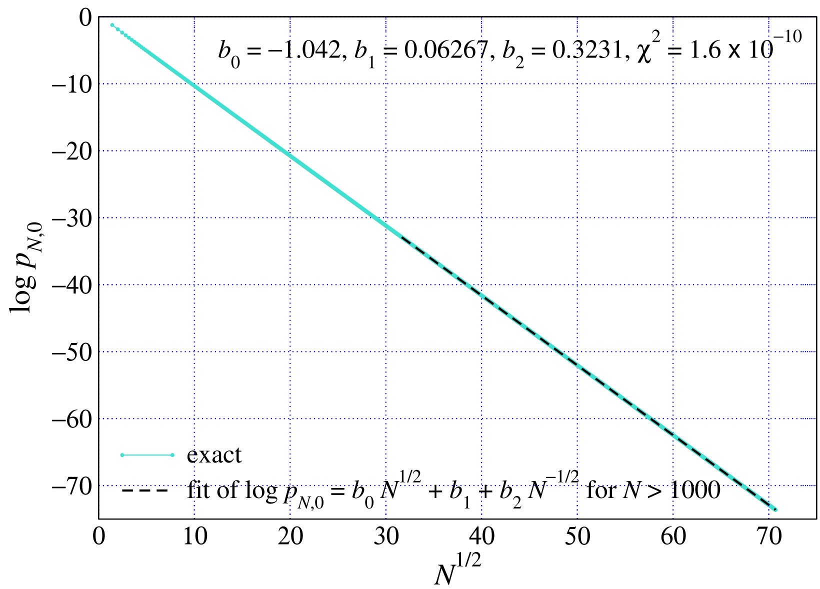


FIG. 1. *The logarithm of the probability* $p_{N,0}$ *that an* $N \times N$ *matrix of even size sampled from the real Ginibre ensemble does not have any real eigenvalues, as a function of* $\sqrt{N}$. *The leading coefficient extracted using the best fit is* $-1.042$, *the best fit for the next-to-leading constant is* 0.06267. *The "exact" curve is constructed using formula* (2.2) *of Lemma* 2.1 *below. The form of the* $b_2$*-term in the fitting curve was chosen to minimise the numerical goodness-of-fit* $\chi^2$.

calculation to construct rigorous asymptotics of a very compact and easy to use exact determinantal expression for the probability $p_{2n,2k}$ specific to the real Ginibre ensemble. This determinantal expression can be derived building upon the results of [14] and [12]; see Lemma 2.1 below. We hope of course that our very specialised proof will contribute to the general discussion of the theory of large deviations for Pfaffian point processes.

There is a drawback to our approach as well: even though we can now claim that (1.2) is true, we still do not know how a large Ginibre matrix without real eigenvalues *looks*. For example, is there a unique optimal configuration of complex eigenvalues for such matrices? What can be said about the overlaps between left and right eigenvectors of Ginibre matrices without real eigenvalues? To answer these questions, one has to develop a large deviations principle along the lines of [7] which will most likely use the picture of the "two-component" plasma consisting of one-dimensional and two-dimensional "gases" of eigenvalues discussed there.

Our paper is organised as follows: a reader who is satisfied by our heuristic argument and the numerics can stop here. Those interested in the mathematical proof are advised to read Section 2 and consult Appendix A for the proofs of the technical facts used in the proof of Theorem 1.1. Appendix B contains remarks on the numerical evaluation of $p_{2n,0}$ for large values of $n$.

## 2. The proof of Theorem 1.1.

Our starting point is the following exact determinantal representation for the generating function for the probabilities $p_{2n,2k}$.

LEMMA 2.1. *Let $n$ be a positive integer. Then*

$$\sum_{k=0}^{n} z^k p_{2n,2k} = \det_{j,k=1,n}\left[\delta_{j,k} + \frac{(z-1)}{\sqrt{2\pi}} \frac{\Gamma(j+k-3/2)}{\sqrt{\Gamma(2j-1)\Gamma(2k-1)}}\right]. \tag{2.1}$$

*In particular,*

$$p_{2n,0} = \det_{j,k=1,n}\left[\delta_{j,k} - \frac{1}{\sqrt{2\pi}} \frac{\Gamma(j+k-3/2)}{\sqrt{\Gamma(2j-1)\Gamma(2k-1)}}\right]. \tag{2.2}$$

We postpone to Appendix A the proofs of all lemmas used during the proof of the main theorem.

Notice that the expression (2.2) coincides (as it should) with the $s \to \infty$ limit of the probability that a $2n \times 2n$ real Ginibre matrix has no real eigenvalues in the interval $(-s, s)$ calculated by Forrester; see formula (3.48) of [11].

We will prove Theorem 1.1 in two steps: first, we will prove (1.2) for $k = 0$, then we will show that $\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,2k_n} = \lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,0}$, where $(k_n)_{n\geq 1}$ is a sequence of integers which grows with $n$ slower than $n^{1/2}/\log(n)$.

### 2.1. *The calculation of* $\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,0}$.

Let $M_n$ be an $n \times n$ symmetric matrix entering the statement of Lemma 2.1:

$$M_n(j,k) = \frac{1}{\sqrt{2\pi}} \frac{\Gamma(j+k-3/2)}{\sqrt{\Gamma(2j-1)\Gamma(2k-1)}}, \qquad 1 \leq j, k \leq n. \tag{2.3}$$

LEMMA 2.2. *$M_n$ is a positive definite matrix. Moreover, there exists a positive constant $\mu > 0$ and a natural number $N$ such that for any $n > N$,*

$$\lambda_{\max}(n) \leq 1 - \frac{\mu}{n}, \tag{2.4}$$

*where $\lambda_{\max}(n)$ is the maximal eigenvalue of $M_n$.*

Using Lemmas 2.1 and 2.2 we represent $p_{2n,0}$ as follows:

$$\begin{aligned} \frac{1}{\sqrt{2n}} \log p_{2n,0} &= \frac{1}{\sqrt{2n}} \operatorname{Tr} \log(I - M_n) \\ &= -\frac{1}{\sqrt{2n}} \sum_{m=1}^{K_n} \frac{1}{m} \operatorname{Tr} M_n^m - \frac{1}{\sqrt{2n}} R_n(K_n), \end{aligned} \tag{2.5}$$

where $K_n$ is a cut-off which increases with $n$ (chosen below) and $R_n$ is the remainder of the Taylor series for $\log(I - M_n)$ written in the integral form:

$$R_n(K) = \int_0^1 \mathrm{Tr}\left(\frac{M_n^{K+1}}{(1 - xM_n)^{K+1}}\right)(1-x)^K\,dx.$$

An upper bound on $|R_n(K)|$ follows from Lemma 2.2 by replacing all eigenvalues of $M_n$ with $\lambda_{\max}(n)$:

$$\begin{aligned}\left|R_n(K)\right| &\le n\lambda_{\max}^{K+1}(n)\int_0^1 \frac{(1-x)^K}{(1-\lambda_{\max}(n)x)^{K+1}}\,dx \\ &\le n\lambda_{\max}^{K}(n)\log\left(\frac{1}{1-\lambda_{\max}(n)}\right) \le n\log\left(\frac{n}{\mu}\right)\left(1-\frac{\mu}{n}\right)^K.\end{aligned}$$

So, if we choose

$$K_n = \lfloor n^\alpha \rfloor, \qquad \alpha > 1, \tag{2.6}$$

it is easy to check that

$$\lim_{n\to\infty} R_n(K_n) = 0. \tag{2.7}$$

The last step of the proof is the calculation of $\sum_{m=1}^{K_n} \frac{1}{m}\,\mathrm{Tr}\,M_n^m$. The relevant results can be summarised as follows.

LEMMA 2.3. *For any fixed integer $m > 0$,*

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}}\,\mathrm{Tr}\,M_n^m = \sqrt{\frac{1}{2\pi m}}. \tag{2.8}$$

*Moreover, for any positive integers $m, n$*

$$\mathrm{Tr}\,M_n^m \le \sqrt{\frac{n}{\pi m}}(1+n^{-1}) + \frac{1}{4} + \frac{1}{8}\sqrt{\frac{m}{\pi n}}(1+2n^{-1}). \tag{2.9}$$

Let us stress that formula (2.8) alone is not enough for the calculation of the $\lim_{n\to\infty} n^{-1/2}\log p_{2n,0}$ using (2.5) since the limits $n \to \infty$ and $m \to \infty$ do not necessarily commute. Instead, let us fix an arbitrary integer $K > 0$. For a sufficiently large $n$ (so that $K_n > K$), relation (2.9) gives

$$\begin{aligned}&\frac{1}{\sqrt{2n}}\sum_{m=1}^{K}\frac{1}{m}\,\mathrm{Tr}\,M_n^m \\ &\quad\le \frac{1}{\sqrt{2n}}\sum_{m=1}^{K_n}\frac{1}{m}\,\mathrm{Tr}\,M_n^m \\ &\quad\le \frac{(1+n^{-1})}{\sqrt{2\pi}}\sum_{m=1}^{K_n} m^{-3/2} + \frac{1}{4\sqrt{2n}}\sum_{k=1}^{K_n}\frac{1}{m} + \frac{(1+2n^{-1})}{8\sqrt{2\pi}n}\sum_{k=1}^{K_n}\frac{1}{\sqrt{m}}.\end{aligned} \tag{2.10}$$

In writing the above double inequality, we used the fact that $M_n$ is positive definite, which implies that $\operatorname{Tr} M_n^m > 0$ for all values of $m, n$. Let us choose $K_n$ in the form (2.6) with $\alpha < 2$ and take $n \to \infty$ in (2.10). As $K$ is $n$-independent, we can use formula (2.8) to compute the limit of the left-hand side. On the right-hand side, the last two sums vanish in the limit [as $\log(n)/\sqrt{n}$ and $n^{\alpha/2-1}$ correspondingly]. The first sum converges to

$$\frac{1}{\sqrt{2\pi}} \sum_{m=1}^{\infty} m^{-3/2} = \frac{1}{\sqrt{2\pi}} \zeta(3/2),$$

where $\zeta(x) = \sum_{m=1}^{\infty} m^{-x}$ is the Riemann zeta-function.

We have found that for any positive integer $K$,

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \sum_{m=1}^{K} m^{-3/2} &\le \liminf_{n\to\infty} \frac{1}{\sqrt{2n}} \sum_{m=1}^{K_n} \frac{1}{m} \operatorname{Tr} M_n^m \\ &\le \limsup_{n\to\infty} \frac{1}{\sqrt{2n}} \sum_{m=1}^{K_n} \frac{1}{m} \operatorname{Tr} M_n^m \\ &\le \frac{1}{\sqrt{2\pi}} \zeta(3/2). \end{aligned}$$

As $K$ is arbitrary, we conclude that

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \sum_{m=1}^{K_n} \operatorname{Tr} M_n^m = \frac{1}{\sqrt{2\pi}} \zeta(3/2). \tag{2.11}$$

So we proved that both (2.7) and (2.11) hold provided the cut-off is taken in the form (2.6) for any fixed $\alpha \in (1, 2)$.

Finally, we can take the $n \to \infty$ limit in (2.5). Employing (2.7) and (2.11), we find that

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,0} = -\frac{1}{\sqrt{2\pi}} \zeta(3/2).$$

Theorem 1.1 is proved for $k = 0$.

2.2. *The calculation of* $\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \log p_{2n,2k}$ *for* $k > 0$. It follows from Lemma 2.1 that

$$p_{2n,2k} = \frac{1}{k!} \left(\frac{d}{dz}\right)^k \det\bigl(I + (z-1)M_n\bigr)\Big|_{z=0}.$$

Equivalently,

$$p_{2n,2k} = \frac{p_{2n,0}}{k!} \left(\frac{d}{dz}\right)^k \det(I + zP_n)\Big|_{z=0}, \tag{2.12}$$

where $P_n = (I - M_n)^{-1} M_n$. Recall that

$$\det(I + zP_n) = \sum_{k=0}^{n} z^k e_k(\nu),$$

where $\nu = (\nu_1, \nu_2, \ldots, \nu_n)$ are the eigenvalues of $P_n$ and $e_k$ is the degree-$k$ elementary symmetric polynomial in $n$ variables [15],

$$e_k(\nu) = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \nu_{i_1} \nu_{i_2} \cdots \nu_{i_k}.$$

Therefore,

$$p_{2n,2k} = p_{2n,0} e_k(\nu) \qquad \text{for } k = 0, 1, \ldots, n. \tag{2.13}$$

Let us enumerate the eigenvalues of $M_n$ and $P_n$ as follows:

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0,$$
$$\nu_1 \ge \nu_2 \ge \cdots \ge \nu_n > 0.$$

By the definition of $P_n$, $\nu_i = \frac{\lambda_i}{1-\lambda_i}$. Note that $\nu_i$ is a monotonically increasing function of $\lambda_i$. Combining this remark with the spectral bound of Lemma 2.2, we get the following bound on the elementary symmetric polynomials:

$$e_k(\nu) \le \nu_1^k e_k(1, 1, \ldots, 1) \le \left(\frac{\lambda_1}{1 - \lambda_1}\right)^k n^k \le \left(\frac{n}{\mu}\right)^k n^k. \tag{2.14}$$

Substituting (2.14) into (2.13), we obtain the following upper bound on $\log p_{2n,2k}$:

$$\log p_{2n,2k} \le \log p_{2n,0} + k \log\left(\frac{n^2}{\mu}\right). \tag{2.15}$$

Next, we derive a lower bound on $\log p_{2n,2k}$. By positive definiteness, $\nu_i \ge \lambda_i$ and, therefore, $e_k(\nu) \ge e_k(\lambda)$. Let us fix a positive integer $k$. Due to (2.8), for any $\varepsilon > 0$ there is a positive integer $N_\varepsilon$ such that for any $n > N_\varepsilon$

$$\sqrt{\frac{n}{\pi}}(1 - \varepsilon) \le \operatorname{Tr} M_n \le \sqrt{\frac{n}{\pi}}(1 + \varepsilon). \tag{2.16}$$

On the other hand,

$$\begin{aligned} \operatorname{Tr} M_n &= (\lambda_1 + \cdots + \lambda_{k-1}) + (\lambda_k + \cdots + \lambda_n) \\ &\le (k - 1) + (n - k + 1)\lambda_k, \end{aligned} \tag{2.17}$$

where the inequality is due to (2.4) and the chosen ordering of $\lambda$'s.

Combining (2.16) and (2.17), we obtain the following bound on the $k$th largest eigenvalue of $M_n$:

$$\lambda_k \ge \frac{\sqrt{n/\pi}(1 - \varepsilon) - k + 1}{n - k + 1}, \tag{2.18}$$

which holds for $n > N_\varepsilon$. Inequality (2.18) leads to the desired bound for $e_k(\nu)$:

$$e_k(\nu) \geq e_k(\lambda) \geq \lambda_1\lambda_2\cdots\lambda_k \geq \lambda_k^k \geq \left(\frac{\sqrt{n/\pi}(1-\varepsilon)-k+1}{n-k+1}\right)^k.$$

Substituting this result into (2.13), we find that

$$\log p_{2n,2k} \geq \log p_{2n,0} + k\log\left(\frac{\sqrt{n/\pi}(1-\varepsilon)-k+1}{n-k+1}\right). \tag{2.19}$$

Combining (2.15) and (2.19), we find that

$$\lim_{n\to\infty}\frac{1}{\sqrt{2n}}\log p_{2n,2k} = \lim_{n\to\infty}\frac{1}{\sqrt{2n}}\log p_{2n,0}. \tag{2.20}$$

Relations (2.20) and (2.1) imply that formula (1.2) of Theorem 1.1 is proved for any fixed integer $k > 0$.

Moreover, it is evident from (2.15) and (2.19) that equality (2.20) generalises to

$$\lim_{n\to\infty}\frac{1}{\sqrt{2n}}\log p_{2n,2k_n} = \lim_{n\to\infty}\frac{1}{\sqrt{2n}}\log p_{2n,0}, \tag{2.21}$$

where $(k_n)_{n\geq 1}$ is a sequence of natural numbers such that

$$\lim_{n\to\infty}(n^{-1/2}\log(n))k_n = 0.$$

This proves the last claim of Theorem 1.1.

REMARK. Our proof of the $k > 0$ part of the theorem is a simple consequence of positive-definiteness of $M_n$, the spectral bound and the fact that $\mathrm{Tr}(M_n) \overset{n\to\infty}{\sim} \sqrt{n/\pi}$. It is interesting that the proof does not rely on any detailed knowledge of the spectrum of $M_n$.

## APPENDIX A: PROOFS FOR THE LEMMAS

**A.1. Proof of Lemma 2.1.** To prove the lemma, we start with the exact formula due to Kanzieper and Akemann [14] which expresses the probabilities $p_{2n,2k}$ in terms of elementary symmetric functions:

$$p_{2n,2k} = p_{2n,2n}e_{n-k}(t_1,\ldots,t_{n-k}), \tag{A.1}$$

where $t_j$'s are given by

$$t_j = \tfrac{1}{2}\,\mathrm{Tr}\left(\mathbf{A}^{-1}\mathbf{B}\right)^j. \tag{A.2}$$

Here, $\mathbf{A}$ and $\mathbf{B}$ are $2n \times 2n$ antisymmetric matrices whose entries

$$\mathbf{A}_{jk} = \langle q_{j-1}, q_{k-1}\rangle_{\mathrm{R}}, \tag{A.3}$$

$$\mathbf{B}_{jk} = \langle q_{j-1}, q_{k-1}\rangle_{\mathrm{C}}, \tag{A.4}$$

are defined in terms of skew products

$$\langle f, g\rangle_{\mathrm{R}} = \frac{1}{2}\int_{\mathbb{R}^2} dx\, dy e^{-(x^2+y^2)/2}\,\mathrm{sgn}(y-x) f(x) g(y) \tag{A.5}$$

and

$$\langle f, g\rangle_{\mathrm{C}} = i\int_{\mathrm{Im} z>0} d^2 z e^{-(z^2+\bar{z}^2)/2}\,\mathrm{erfc}\biggl(\frac{z-\bar{z}}{i\sqrt{2}}\biggr)\bigl[f(z)g(\bar{z}) - g(z) f(\bar{z})\bigr]. \tag{A.6}$$

Let us stress that (A.1) is valid for an *arbitrary* choice of monic polynomials $q_j(x)$ of degree $j$, provided matrix $\mathbf{A}$ is invertible.

Substituting equations (A.1) and (A.2) into the generating function

$$g_{2n}(z) = \sum_{k=0}^{n} z^k p_{2n,2k} \tag{A.7}$$

and making use of the summation formula [15]

$$\sum_{\ell=0}^{\infty} z^\ell e_\ell(t_1, \ldots, t_\ell) = \exp\Biggl(\sum_{j=1}^{\infty} (-1)^{j-1} t_j \frac{z^j}{j}\Biggr), \tag{A.8}$$

we obtain the Pfaffian representation [2, 5, 14]:

$$g_{2n}(z) = p_{2n,2n}\,\mathrm{Pf}\bigl(-\mathbf{A}^{-1}\bigr)\,\mathrm{Pf}(z\mathbf{A}+\mathbf{B}); \tag{A.9}$$

see remark 1.3 of [5] justifying the transition from square roots of determinants to Pfaffians. Since $g_{2n}(1) = 1$, $p_{2n,2n} = (\mathrm{Pf}(-\mathbf{A}^{-1})\,\mathrm{Pf}(z\mathbf{A}+\mathbf{B}))^{-1}$ and (A.9) simplifies to

$$g_{2n}(z) = \frac{\mathrm{Pf}(z\mathbf{A}+\mathbf{B})}{\mathrm{Pf}(\mathbf{A}+\mathbf{B})}. \tag{A.10}$$

Next, we will use the fact that expression (A.10) for the generating function does not depend on a particular choice of monic polynomials $q_j(x)$ in (A.3) and (A.4) to simplify it even further. Namely, we will choose $q_j(x)$'s in such a way that the matrix $\mathbf{A}+\mathbf{B}$ is block diagonal. Clearly, such polynomials should be skew-orthogonal with respect to the skew product

$$\langle f, g\rangle = \langle f, g\rangle_{\mathrm{R}} + \langle f, g\rangle_{\mathrm{C}}, \tag{A.11}$$

that is

$$\begin{aligned} \langle q_{2j}, q_{2k+1}\rangle &= -\langle q_{2k+1}, q_{2j}\rangle = r_j \delta_{j,k}, \\ \langle q_{2j}, q_{2k}\rangle &= \langle q_{2j+1}, q_{2k+1}\rangle = 0. \end{aligned} \tag{A.12}$$

These were first calculated in the paper [12]:

$$\begin{aligned} q_{2j}(x) = x^{2j}, \qquad & q_{2j+1}(x) = x^{2j+1} - 2jx^{2j-1}, \\ r_j &= \sqrt{2\pi}\,\Gamma(2j+1). \end{aligned} \tag{A.13}$$

Given the choice of $q_j$'s described above:

(a) the matrix $\mathbf{A}+\mathbf{B}$ acquires a block-diagonal form, $\mathbf{A}+\mathbf{B}=\mathbf{r}\otimes\mathbf{J}$, where

$$\mathbf{r}=\operatorname{diag}(r_0,\dots,r_{n-1}),\qquad \mathbf{J}=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \tag{A.14}$$

which leads to

$$g_{2n}(z)=\frac{\operatorname{Pf}(\mathbf{r}\otimes\mathbf{J}+(z-1)\mathbf{A})}{\operatorname{Pf}(\mathbf{r}\otimes\mathbf{J})}. \tag{A.15}$$

(b) the matrix $\mathbf{A}$ is given by

$$\mathbf{A}_{2j,2k}=\mathbf{A}_{2j+1,2k+1}=0,\qquad \mathbf{A}_{2j-1,2k}=\Gamma\big(j+k-\tfrac{3}{2}\big). \tag{A.16}$$

Notice that matrix elements of both $\mathbf{r}\otimes\mathbf{J}$ and $\mathbf{A}$ labeled by a pair of indexes of the same parity vanish. Therefore, the $2n\times 2n$ Pfaffians in the numerator and the denominator of (A.15) are reduced to $n\times n$ determinants:

$$g_{2n}(z)=\frac{\det[r_{j-1}\delta_{jk}+(z-1)\mathbf{A}_{2j-1,2k}]_{1\le j,k\le n}}{\det[r_{j-1}\delta_{jk}]_{1\le j,k\le n}}. \tag{A.17}$$

Finally, we apply the formula $\det(U)/\det(V^2)=\det(V^{-1}UV^{-1})$ to perform division in (A.17). With the help of the explicit formulae (A.13) and (A.16) we get

$$g_{2n}(z)=\det\left[\delta_{jk}+\frac{(z-1)}{\sqrt{2\pi}}\frac{\Gamma(j+k-3/2)}{\sqrt{\Gamma(2j-1)\Gamma(2k-1)}}\right]_{1\le j,k\le n}. \tag{A.18}$$

Lemma 2.1 is proved.

**A.2. Proof of Lemma 2.2.** The proofs of Lemmas 2.2, 2.3 are based on the following integral representation for the matrix elements (2.3) of matrix $M_n$:

$$M_n(j,k)=\frac{1}{\sqrt{2\pi}}\int_0^\infty\frac{dx}{x^{5/2}}e^{-x}\frac{x^j}{\sqrt{\Gamma(2j-1)}}\frac{x^k}{\sqrt{\Gamma(2k-1)}}, \tag{A.19}$$

$1\le j,k\le n$, which can be obtained by representing $\Gamma(j+k-3/2)$ in (2.3) as an integral.

Take any $v=(v_1,v_2,\dots,v_n)\in\mathbb{R}^n\setminus\{0\}$. It follows from (A.19) that

$$\langle v,M_nv\rangle=\frac{1}{\sqrt{2\pi}}\int_0^\infty\frac{dx}{x^{5/2}}e^{-x}\left(\sum_{j=1}^n\frac{v_jx^j}{\sqrt{\Gamma(2j-1)}}\right)^2>0. \tag{A.20}$$

So, $M_n$ is positive definite by definition.

Next, let us prove bound (2.4) on the spectral radius of $M_n$. Let $\lambda_1,\lambda_2,\dots,\lambda_n>0$ be the eigenvalues of $M_n$. Then

$$\lambda_{\max}(n)=\big(\lambda_{\max}^n(n)\big)^{1/n}\le\left(\sum_{k=1}^n\lambda_k^n\right)^{1/n}=\big(\operatorname{Tr}M_n^n\big)^{1/n}. \tag{A.21}$$

It follows from the upper bound (2.9) of Lemma 2.3 that for any $\varepsilon > 0$, there is $N_\varepsilon$ such that for any $n > N_\varepsilon$,

$$\operatorname{Tr} M_n^n \le \sqrt{\frac{1}{\pi}} + \frac{1}{4} + \frac{1}{8}\sqrt{\frac{1}{\pi}} + \varepsilon = 1 - c_1 + \varepsilon,$$

where $c_1 \in (0, 1)$. Therefore, we can choose $\varepsilon$ small enough so that

$$\operatorname{Tr} M_n^n \le 1 - \mu,$$

where $\mu \in (0, 1)$. Using this estimate in (A.21) for $n > N_\varepsilon$ we get

$$\lambda_{\max}(n) \le (1-\mu)^{1/n} \le 1 - \frac{\mu}{n}. \tag{A.22}$$

Lemma 2.2 is proved for $N = N_\varepsilon$.

REMARK. The spectral properties of $M_n$ seem quite interesting. For instance, in the large-$n$ limit there is a concentration of eigenvalues near 1 such that the restriction of $M_n$ to the space spanned by the corresponding eigenvectors is close to an identity operator perturbed by an elliptic linear differential operator. Formal analysis of this perturbation suggests the asymptotic $\lambda_{\max}(n) = 1 - \mu_0 n^{-1} + o(n^{-1})$ for suitable $\mu_0 > 0$.

**A.3. Proof of Lemma 2.3.** The integral representation (A.19) for the matrix elements of $M_n$ leads to the following integral representation for the trace of a power of $M_n$:

$$\begin{aligned}\operatorname{Tr} M_n^m = {} & \int_0^\infty \frac{dx_1}{\sqrt{2\pi x_1}} \int_0^\infty \frac{dx_2}{\sqrt{2\pi x_2}} \cdots \int_0^\infty \frac{dx_m}{\sqrt{2\pi x_m}} e^{-x_1 - x_2 - \cdots - x_m} \\ & \times \cosh_{n-1}(\sqrt{x_m x_1}) \cosh_{n-1}(\sqrt{x_1 x_2}) \cdots \cosh_{n-1}(\sqrt{x_{m-1} x_m}),\end{aligned} \tag{A.23}$$

where $\cosh_n(x) = \sum_{k=0}^n \frac{x^{2k}}{(2k)!}$ is the degree-$2n$ Taylor polynomial generated by the hyperbolic cosine. Performing the change of variables $x_k = y_k^2$ in (A.23), we can rewrite the integral representation for $\operatorname{Tr} M_n^m$ as follows:

$$\begin{aligned}\operatorname{Tr} M_n^m = {} & \left(\frac{2}{\pi}\right)^{m/2} \int_{\mathbb{R}_+^m} dy e^{-\sum_{k=1}^m y_k^2} \\ & \times \cosh_{n-1}(y_m y_1) \cosh_{n-1}(y_1 y_2) \cdots \cosh_{n-1}(y_{m-1} y_m).\end{aligned} \tag{A.24}$$

Here, $\mathbb{R}_+^m = \{(y_1, y_2, \dots, y_m) \in \mathbb{R}^m | y_k \ge 0, k = 1, 2, \dots, m\}$ is the first "quadrant" of $\mathbb{R}^m$ and $dy$ is a shorthand notation for Lebesgue measure on $\mathbb{R}^m$. As the integrand of (A.24) is symmetric with respect to reflection $y_i \to -y_i$ for any $i = 1, 2, \dots, m$, we can rewrite $\operatorname{Tr} M_n^m$ as an integral over $\mathbb{R}^m$:

$$\begin{aligned}\operatorname{Tr} M_n^m = {} & \left(\frac{1}{2\pi}\right)^{m/2} \int_{\mathbb{R}^m} dy e^{-\sum_{k=1}^m y_k^2} \\ & \times \cosh_{n-1}(y_m y_1) \cosh_{n-1}(y_1 y_2) \cdots \cosh_{n-1}(y_{m-1} y_m).\end{aligned} \tag{A.25}$$

To prove Lemma 2.3, we will establish an upper and a lower bound on $\operatorname{Tr} M_n^m$ and then compute the large-$n$ limit of each of these bounds.

A.3.1. *An upper bound for* $\operatorname{Tr} M_n^m$. A good starting point for the calculation is formula (A.25). For any $x \in \mathbb{R}$, $\cosh_{n-1}(x) \leq \cosh(x)$. Also,

$$\cosh_{n-1}(x) = \oint \frac{dz}{2\pi i z} \frac{1 - z^{-2n}}{1 - z^{-2}} e^{zx}, \tag{A.26}$$

where the integral is anti-clockwise around a circle of radius smaller than 1 centred at the origin in the complex plane. Replacing all but one $\cosh_{n-1}$ with cosh we get:

$$\begin{aligned} \operatorname{Tr} M_n^m &\leq \left(\frac{1}{2\pi}\right)^{m/2} \int_{\mathbb{R}^m} dy e^{-\sum_{k=1}^m y_k^2} \cosh_{n-1}(y_m y_1) \cosh(y_1 y_2) \cdots \\ &\quad \times \cosh(y_{m-1} y_m) \\ &= \left(\frac{1}{2\pi}\right)^{m/2} \mathbb{E}_{\alpha_1 \alpha_2 \cdots \alpha_{m-1}} \int_{\mathbb{R}^m} dy e^{-\sum_{k=1}^m y_k^2} \\ &\quad \times \cosh_{n-1}(y_m y_1) e^{\sum_{l=1}^{m-1} \alpha_l y_l y_{l+1}}, \end{aligned} \tag{A.27}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ are independent identically distributed random variables which take values $\pm 1$ with probability $1/2$. Representing the remaining $\cosh_{n-1}$ with the help of (A.26) and then computing resulting Gaussian integral over $\mathbb{R}^m$, we find

$$\operatorname{Tr} M_n^m \leq \left(\frac{1}{2}\right)^{m/2} \mathbb{E}_{\alpha_1 \alpha_2 \cdots \alpha_{m-1}} \oint \frac{dz}{2\pi i z} \frac{1 - z^{-2n}}{1 - z^{-2}} \left[D_m^{(\alpha)}(z)\right]^{-1/2}, \tag{A.28}$$

where

$$D_m^{(\alpha)}(z) = \det \begin{pmatrix} 1 & -\frac{\alpha_1}{2} & 0 & 0 & \ldots & 0 & -\frac{z}{2} \\ -\frac{\alpha_1}{2} & 1 & -\frac{\alpha_2}{2} & 0 & 0 & \ldots & 0 \\ 0 & -\frac{\alpha_2}{2} & 1 & -\frac{\alpha_3}{2} & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & -\frac{\alpha_{m-3}}{2} & 1 & -\frac{\alpha_{m-2}}{2} & 0 \\ 0 & \ldots & 0 & 0 & -\frac{\alpha_{m-2}}{2} & 1 & -\frac{\alpha_{m-1}}{2} \\ -\frac{z}{2} & 0 & \ldots & 0 & 0 & -\frac{\alpha_{m-1}}{2} & 1 \end{pmatrix}. \tag{A.29}$$

The determinant can be calculated recursively in $m$, yielding $D_1^{(\alpha)}(z) = 1 - z$ and

$$D_m^{(\alpha)}(z) = -(m-1)\frac{1}{2^m}(z - A_m)\left(z + A_m\frac{m+1}{m-1}\right) \qquad \text{for } m \geq 2, \tag{A.30}$$

where $A_m = \prod_{k=1}^{m-1} \alpha_k$. Note that (A.30) implies that all principal minors of the matrix under the sign of the determinant in (A.29) are positive for $z = 0$. Therefore, the matrix itself is positive definite for $z = 0$. By continuity, the real part of this matrix remains positive definite for $z \neq 0$ provided $|z|$ is small enough. Therefore, the real part of the quadratic form which determines the Gaussian integral in (A.27) is positive definite, which justifies the interchange of integrals leading to (A.28) provided the contour is taken to be a circle around the origin of a sufficiently small radius.

Substituting (A.30) into (A.28) and changing the integration variable $z \to A_m z$, we find that the integrand no longer depends on $\alpha$'s. Averaging over $\alpha$'s becomes trivial and we get the following integral upper bound:

$$\operatorname{Tr} M_n^m \leq \oint \frac{dz}{2\pi z}\frac{z^{-2n}-1}{z^{-2}-1}\frac{1}{\sqrt{1-z}}\frac{1}{\sqrt{(m-1)z+m+1}}. \tag{A.31}$$

The rest of the calculation is slightly different depending on whether $m = 1$ or $m > 1$. Here, present the calculation for $m > 1$ only, the (simpler) case of $m = 1$ can be treated along similar lines. We calculate the integral in the right-hand side of (A.31) as follows. First, we replace $z^{-2n} - 1$ with $z^{-2n}$ in the integrand on the right-hand side of (A.31), since this does not change the value of the integral as the omitted term is analytic inside of the contour of integration. Next, we deform the contour away from the singularity at zero and out to infinity, leading to integrals around the other singularities of the (modified) integrand: a simple pole at $z = -1$, a branch cut singularity along the real line from 1 to $+\infty$, and a branch cut singularity along the real line from $-\frac{m+1}{m-1}$ to $-\infty$. The contribution from the integral over the large circle at infinity is zero. The contribution from the pole at $z = -1$ is easily evaluated as $1/4$. Evaluating the integral around the branch from 1 to $+\infty$ it is convenient first integrate by parts, so that the singularity at $z = 1$ is integrable. The integrals along the two branch cuts lead to two real integrals whose asymptotics are controlled by the integrand $(1 + y)^{-2n}$. Changing variable $y \to y/2n$, and making some simple estimates on terms that do not affect the leading asymptotics, we are led to

$$\begin{aligned}\operatorname{Tr} M_n^m \leq \frac{1}{4} &+ \sqrt{\frac{n}{\pi m}}\int_0^\infty \frac{dy}{\sqrt{\pi y}}\left(1 + \frac{y}{2n}\right)^{-2n} \\ &+ \frac{1}{\sqrt{2\pi n}}\frac{m+1}{2\sqrt{m-1}}\left(\frac{m+1}{2m}\right)^{3/2}\left(\frac{m-1}{m+1}\right)^{2n+1} \\ &\times \int_0^\infty \frac{dy}{\sqrt{\pi y}}\left(1 + \frac{y}{2n}\right)^{-2n+1}.\end{aligned} \tag{A.32}$$

Both integrals in the above expression can be estimated using the following bound:

$$I_M = \int_0^\infty \frac{dy}{\sqrt{\pi y}}\left(1+\frac{y}{M}\right)^{-M} \le 1+\frac{2}{M}, \tag{A.33}$$

which follows by evaluating the integral, using the substitution $t = (1+\frac{y}{M})^{-1}$, in terms of the beta function as

$$I_M = \sqrt{\frac{M}{\pi}}B\left(M-\frac{1}{2},\frac{1}{2}\right) = \sqrt{M}\frac{\Gamma(M-3/2)}{\Gamma(M-1/2)}$$

and using bounds on the Gamma function. Using this in (A.32), the final result is

$$\begin{aligned} \operatorname{Tr} M_n^m &\le \frac{1}{4} + \sqrt{\frac{n}{\pi m}}\left(1+\frac{1}{n}\right) + \frac{1}{8}\sqrt{\frac{m}{\pi n}}\left(1+\frac{2}{n}\right)\left(\frac{m-1}{m+1}\right)^{2n-3/2} \\ &\le \frac{1}{4} + \sqrt{\frac{n}{\pi m}}\left(1+\frac{1}{n}\right) + \frac{1}{8}\sqrt{\frac{m}{\pi n}}\left(1+\frac{2}{n}\right), \end{aligned} \tag{A.34}$$

which coincides with the claim (2.9) of Lemma 2.3.

Dividing both sides of (A.34) by $\sqrt{2n}$ and taking the large $n$ limit, we find that

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2n}} \operatorname{Tr} M_n^m \le \sqrt{\frac{1}{2\pi m}}. \tag{A.35}$$

A.3.2. *The limit* $\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \operatorname{Tr} M_n^m$. The strategy is to derive an integral lower bound for $\operatorname{Tr} M_n^m$ and calculate the large $n$-limit of the bound. Our starting point is the relation (A.24) and the following estimate for the polynomial $\cosh_{n-1}$.

LEMMA A.1. *There exist two sequences* $(h_n)_{n\ge1}, (S_n)_{n\ge1} \subset \mathbb{R}$ *such that*

$$\begin{gathered} \lim_{n\to\infty} h_n = \frac{1}{2}, \qquad \lim_{n\to\infty} S_n = 2, \\ e^{-ny}\cosh_{n-1}(ny) \ge h_n \mathbb{1}(y < S_n) \qquad \text{for } y \ge 0, n \ge 1. \end{gathered} \tag{A.36}$$

*Here*, $\mathbb{1}(y < S_n)$ *is the indicator function of the set* $[0, S_n)$.

In fact, as $n \to \infty$, $e^{-ny}\cosh_{n-1}(ny)$ converges almost everywhere to $\frac{1}{2}\mathbb{1}(y < 2)$ for $y \ge 0$, but here we only need the lower bound. The proof of Lemma A.1 is given in Section A.4.

Using the bound (A.36) in (A.24), we find that

$$\begin{aligned} \operatorname{Tr} M_n^m \ge{}& h_n^m \left(\frac{2}{\pi}\right)^{m/2} n^{m/2} \int_{\mathbb{R}_+^m} dy \\ &\times \prod_{l=1}^m \mathbb{1}(y_l y_{l+1} < S_n) e^{-(n/2)\sum_{k=1}^m (y_{k+1}-y_k)^2}, \end{aligned} \tag{A.37}$$

where $y_{m+1} := y_1$. It is straightforward to verify that the domain of integration for the integral in (A.37) contains the hypercube $(0, \sqrt{S_n})^m$,

$$(0, \sqrt{S_n})^m \subset \{y \in R_+^m | y_k y_{k+1} < S_n, k = 1, 2, \dots, m\}.$$

Therefore,

$$\prod_{l=1}^{n} \mathbb{1}(y_l < \sqrt{S_n}) \le \prod_{l=1}^{m} \mathbb{1}(y_l y_{l+1} < S_n), \qquad y \in \mathbb{R}_+^m \tag{A.38}$$

Substituting (A.38) in (A.37) and changing the integration variables according to

$$R = y_1 + y_2 + \cdots + y_m,$$
$$z_k = y_{k+1} - y_k, \qquad k = 1, 2, \dots, m-1,$$

we get the following lower bound:

$$\begin{aligned} \operatorname{Tr} M_n^m \ge{}& \frac{h_n^m}{m} \left(\frac{2}{\pi}\right)^{m/2} n^{m/2} \int_0^{m\sqrt{S_n}} dR \\ &\times \int_{P_{m-1}(R)} dz_1 \cdots dz_{m-1} e^{-(n/2)[\sum_{k=1}^{m-1} z_k^2 + (\sum_{k=1}^{m-1} z_k)^2]}, \end{aligned} \tag{A.39}$$

where $P_{m-1}(R)$ is the intersection of the hypercube $(0, \sqrt{S_n})^m$ and the hyperplane

$$\{y \in R_+^m | y_1 + y_2 + \cdots + y_m = R\}.$$

In the derivation of (A.39), we used the fact that the Jacobian of the transformation $y \to (R, z)$ is equal to $1/m$.

The large-$n$ limit of the right-hand side of (A.39) can be evaluated by arguing as in the Laplace method:

$$\begin{aligned} \liminf_{n\to\infty} &\frac{1}{\sqrt{2n}} \operatorname{Tr} M_n^m \\ &\ge \lim_{n\to\infty} \frac{h_n^m}{\sqrt{2n}m} \left(\frac{2}{\pi}\right)^{m/2} n^{m/2} \int_0^{m\sqrt{S_n}} dR \\ &\quad \times \int_{\mathbb{R}^{m-1}} dz_1 \cdots dz_{m-1} e^{-(n/2)[\sum_{k=1}^{m-1} z_k^2 + (\sum_{k=1}^{m-1} z_k)^2]} \\ &= \lim_{n\to\infty} \sqrt{\frac{S_n}{2}} h_n^m \left(\frac{2}{\pi}\right)^{m/2} n^{(m-1)/2} \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} \\ &\quad \times \int_{\mathbb{R}^m} dz_1 \cdots dz_m e^{i\lambda \sum_{k=1}^m z_k} e^{-(n/2)\sum_{k=1}^m z_k^2} \\ &= \lim_{n\to\infty} \sqrt{\frac{S_n}{2}} h_n^m \left(\frac{2}{\pi}\right)^{m/2} n^{(m-1)/2} \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} \left(\int_{-\infty}^{\infty} dz e^{i\lambda z - (n/2)z^2}\right)^m \end{aligned}$$

$$= \lim_{n\to\infty} \sqrt{\frac{S_n}{2} h_n^m \left(\frac{2}{\pi}\right)^{m/2} n^{(m-1)/2} \left(\frac{2\pi}{n}\right)^{m/2} \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} e^{-(m/(2n))\lambda^2}}$$

$$= \lim_{n\to\infty} \sqrt{\frac{S_n}{2} h_n^m \left(\frac{2}{\pi}\right)^{m/2} n^{(m-1)/2} \left(\frac{2\pi}{n}\right)^{m/2} \sqrt{\frac{n}{2\pi m}}} = \sqrt{\frac{1}{2\pi m}}.$$

The crucial, albeit very standard, first step in the above derivation consists of verifying that extending the integration space for the $z$-integral from $P_{m-1}(R)$, when $R \in (0, 2)$, to $\mathbb{R}^{m-1}$ does not change the large $n$-limit.

We conclude that

$$\liminf_{n\to\infty} \frac{1}{\sqrt{2n}} \operatorname{Tr} M_n^m \geq \sqrt{\frac{1}{2\pi m}},$$

and in combination with (A.35) this gives

$$\lim_{n\to\infty} \frac{1}{\sqrt{2n}} \operatorname{Tr} M_n^m = \sqrt{\frac{1}{2\pi m}}.$$

Statement (2.8) of Lemma 2.3 is proved.

**A.4. Proof of Lemma A.1.** Let $\{\alpha_n\}_{n=1}^{\infty}$ be an arbitrary sequence of positive real numbers which diverges as $n \to \infty$ slower than $n^{1/2}$, that is $\lim_{n\to\infty} \alpha_n = \infty$, but $\lim_{n\to\infty} \alpha_n n^{-1/2} = 0$. We will show that there exists $N_0 > 0$ such that for any $n > N_0$ and $x \geq 0$

$$e^{-nx} \cosh_n(nx) \geq \left(\frac{1}{2} - \frac{1}{\sqrt{4\pi}} \alpha_n^{-1} e^{-\alpha_n^2/4}\right) \mathbb{1}\left(x \leq 2 - \alpha_n n^{-1/2}\right). \tag{A.40}$$

The statement of Lemma A.1, where $\cosh_n(nx)$ is replaced by $\cosh_{n-1}(nx)$, is easily deduced from equation (A.40).

Our proof builds on the ideas of [4] dedicated to the study of sections of exponential series (Taylor polynomials generated by exp). Let $e_n$ be a section of exponential series defined by

$$e_n(x) = \sum_{j=0}^{n} \frac{x^j}{j!}.$$

Consider also

$$e_n^{(+)}(x) = e^{-nx} e_n(nx), \qquad e_n^{(-)}(x) = e^{-nx} e_n(-nx).$$

Then the function we are interested in can be written as

$$f_n(x) := e^{-nx} \cosh_n(nx) = \frac{1}{2}\left(e_{2n}^{(+)}\left(\frac{x}{2}\right) + e_{2n}^{(-)}\left(\frac{x}{2}\right)\right).$$

First, we show that $e_{2n}^{(-)}(x) > 0$ for $x \geq 0$. One can check that

$$\frac{d}{dx}\big(e^{2nx}e_{2n}(-2nx)\big) = \frac{1}{(2n-1)!}(2nx)^{2n}e^{2nx} \geq 0, \tag{A.41}$$

and $e^{2nx}e_{2n}(-2nx)|_{x=0} = 1$. So $e_{2n}^{(-)}(x) \geq e^{-4nx} > 0$ for $x \geq 0$. The next step is to show that $f_n(x)$ is a decreasing function. However,

$$f_n'(x) = -ne_{2n}^{(-)}\left(\frac{x}{2}\right),$$

which is negative by (A.41).

The fact that $f_n(x)$ is decreasing and the positivity of $e_{2n}^{(-)}(x)$ imply that for any nonnegative $x$

$$\begin{aligned} f_n(x) &\geq f_n(x)\mathbb{1}\big(x \leq 2 - \alpha_n n^{-1/2}\big) \\ &\geq f_n\big(2 - \alpha_n n^{-1/2}\big)\mathbb{1}\big(x \leq 2 - \alpha_n n^{-1/2}\big) \\ &\geq \frac{1}{2}e_{2n}^{(+)}\left(1 - \frac{\alpha_n}{2}n^{-1/2}\right)\mathbb{1}\big(x \leq 2 - \alpha_n n^{-1/2}\big). \end{aligned}$$

Therefore, it remains to prove that

$$e_{2n}^{(+)}\left(1 - \frac{\alpha_n}{2}n^{-1/2}\right) \geq 1 - \sqrt{\frac{1}{\pi}}\alpha_n^{-1}e^{-\alpha_n^2/4}, \tag{A.42}$$

for all $n > N_0$, where $N_0$ is chosen to satisfy $\alpha_n n^{-1/2} < 2$ for all $n > N_0$.

We start with a differential equation satisfied by $e_n^{(+)}$. As it is easy to check,

$$\frac{d}{dx}e_n^{(+)}(x) = -\frac{1}{(n-1)!}(nx)^n e^{-nx}. \tag{A.43}$$

So $e_n^{(+)}(x)$ is a decreasing function on $\mathbb{R}_+$.

Equation (A.43) has to be solved with a boundary condition $\lim_{x\to\infty} e_n^{(+)}(x) = 0$, which follows from the definition of $e_n^{(+)}$. The solution is

$$e_n^{(+)}(x) = \frac{n^n}{(n-1)!}\int_x^\infty t^n e^{-nt}\,dt. \tag{A.44}$$

Let

$$\phi_n = \frac{\sqrt{2\pi n}(n/e)^n}{n!}.$$

By the Stirling approximation formula, $\phi_n = 1 + O(n^{-1})$ for $n \to \infty$ and $\phi_n < 1$. Define

$$\tau(t) = t - 1 - \log t \geq 0 \qquad \text{for } t \in \mathbb{R}_+. \tag{A.45}$$

In terms of $\phi_n$ and $\tau$, expression (A.44) acquires the following form:

$$e_n^{(+)}(x) = \sqrt{\frac{n}{2\pi}}\phi_n \int_x^\infty e^{-n\tau(t)}\,dt. \tag{A.46}$$

The integral in the right-hand side can be analysed using the Laplace method. It follows from the definition that

$$1 = e_n^{(+)}(0) = \sqrt{\frac{n}{2\pi}}\phi_n \int_0^\infty e^{-n\tau(t)}\,dt.$$

Therefore, (A.46) can be rewritten as follows:

$$e_n^{(+)}(x) = 1 - \sqrt{\frac{n}{2\pi}}\phi_n \int_0^x e^{-n\tau(t)}\,dt =: 1 - r_n(x).$$

Let us estimate the remainder $r_n(x)$. Evidently, $r_n(x) \geq 0$. An application of Taylor's theorem with the Lagrange form of the remainder reveals that for $0 < t \leq x \leq 1$,

$$\tau(t) \geq \frac{\tau''(x)}{2}(t-x)^2 + \tau'(x)(t-x) + \tau(x). \tag{A.47}$$

Noticing that $\tau'(x) = -\frac{1-x}{x}$ and $\tau''(x) = \frac{1}{x^2}$ we can use the above bound on $\tau(t)$ to obtain the following upper bound on $r_n$:

$$\begin{aligned} r_n(x) &\leq \sqrt{\frac{n}{2\pi}}\phi_n e^{-n\tau(x)} \int_0^x e^{-(n/(2x^2))(t-x)^2 + (n(1-x)/x)(t-x)}\,dt \\ &= \frac{\phi_n}{2} x e^{-n(\tau(x)-(1-x)^2/2)}\left(\operatorname{erfc}\left(\sqrt{\frac{n}{2}}(1-x)\right) - \operatorname{erfc}\left(\sqrt{\frac{n}{2}}(2-x)\right)\right) \\ &\leq \frac{\phi_n}{2} x e^{-n\tau(x)} \operatorname{erfcx}\left(\sqrt{\frac{n}{2}}(1-x)\right), \end{aligned}$$

where erfc and erfcx are complementary and scaled complementary error functions correspondingly. Finally, applying the classical estimate $\operatorname{erfcx}(x) \leq \frac{1}{x\sqrt{\pi}}$ valid for any $x > 0$ (see, e.g., [1]), we obtain

$$r_n(x) \leq \frac{\phi_n}{\sqrt{2n\pi}} \frac{x}{1-x} e^{-n\tau(x)} < \frac{1}{\sqrt{2n\pi}} \frac{x}{1-x} e^{-n\tau(x)},$$

where we used that $\phi_n < 1$. Therefore,

$$r_{2n}\left(1 - \frac{\alpha_n}{2} n^{-1/2}\right) \leq \sqrt{\frac{1}{\pi}} \alpha_n^{-1} e^{-2n\tau(1-(\alpha_n/2)n^{-1/2})}.$$

Using (A.47) for $x = 1$ and $t = 1 - \frac{\alpha_n}{2} n^{-1/2}$, we obtain

$$r_{2n}\left(1 - \frac{\alpha_n}{2} n^{-1/2}\right) \leq \sqrt{\frac{1}{\pi}} \alpha_n^{-1} e^{-\alpha_n^2/4},$$

which leads to the desired bound (A.42) for $e_{2n}^{(+)}$. Lemma A.1 is proved.

## APPENDIX B: ON THE NUMERICAL EVALUATION OF $p_{2n,0}$

It is clear from the proof of Lemma 2.1 that the final form of the Pfaffian or determinantal expression for the probability that an $n \times n$ real Ginibre matrix has no real eigenvalues is strongly influenced by the choice of skew orthogonal polynomials used in the derivation. And even though the final exact result does not depend on the choice of the skew orthogonal polynomials, its numerical stability is highly sensitive to the choice.

For example, the determinantal formula (2.2) is highly suitable for numerical evaluations since the condition number of the corresponding matrix $I - M_n$ grows at most linearly with $n$. Indeed, its largest eigenvalue is smaller than unity since $M_n$ is positive definite in virtue of the first part of Lemma 2.2. On the other hand, its smallest eigenvalue is separated from zero by an interval of length of order $O(n^{-1})$ due to the result of Lemma 2.2 concerning the largest eigenvalue of $M_n$.

This should be contrasted to the determinantal formula derived in [14]. The condition number of the matrix $\rho$ appearing in this formula grows exponentially with $n$, forcing one to use high-precision numerics and leading to computation times growing exponentially with $n$.

**Acknowledgement.** C. Timm acknowledges useful discussions with K. Nestmann.

## REFERENCES

[1] ABRAMOWITZ, M. and STEGUN, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series* **55**. U.S. Government Printing Office, Washington, D.C. MR0167642

[2] AKEMANN, G. and KANZIEPER, E. (2007). Integrable structure of Ginibre's ensemble of real random matrices and a Pfaffian integration theorem. *J. Stat. Phys.* **129** 1159–1231. MR2363393

[3] BEN AROUS, G. and ZEITOUNI, O. (1998). Large deviations from the circular law. *ESAIM Probab. Stat.* **2** 123–134 (electronic). MR1660943

[4] BLEHER, P. and MALLISON, R. JR. (2006). Zeros of sections of exponential sums. *Int. Math. Res. Not. IMRN* Art. ID 38937, 49. MR2264712

[5] BORODIN, A. and KANZIEPER, E. (2007). A note on the Pfaffian integration theorem. *J. Phys. A* **40** F849–F855. MR2396206

[6] BORODIN, A. and SINCLAIR, C. D. (2009). The Ginibre ensemble of real random matrices and its scaling limits. *Comm. Math. Phys.* **291** 177–224. MR2530159

[7] DEL MOLINO, L. C. G., PAKDAMAN, K., TOUBOUL, J. and WAINRIB, G. (2015). The Ginibre ensemble with $k = O(n)$ real eigenvalues. Available at arXiv:1501.03120v1.

[8] DERRIDA, B. and ZEITAK, R. (1996). Distribution of domain sizes in the zero temperature Glauber dynamics of the one-dimensional Potts model. *Phys. Rev. E* **54** 2513–2525.

[9] EDELMAN, A. (1997). The probability that a random real Gaussian matrix has $k$ real eigenvalues, related distributions, and the circular law. *J. Multivariate Anal.* **60** 203–232. MR1437734

[10] EDELMAN, A., KOSTLAN, E. and SHUB, M. (1994). How many eigenvalues of a random matrix are real? *J. Amer. Math. Soc.* **7** 247–267. MR1231689

[11] FORRESTER, P. J. (2015). Diffusion processes and the asymptotic bulk gap probability for the real Ginibre ensemble. *J. Phys. A* **48** 324001, 14. MR3376016

[12] FORRESTER, P. J. and NAGAO, T. (2007). Eigenvalue statistics of the real Ginibre ensemble. *Phys. Rev. Lett.* **99** 050603.

[13] GINIBRE, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** 440–449. MR0173726

[14] KANZIEPER, E. and AKEMANN, G. (2005). Statistics of real eigenvalues in Ginibre's ensemble of random real matrices. *Phys. Rev. Lett.* **95** 230201, 4. MR2185860

[15] MACDONALD, I. G. (1995). *Symmetric Functions and Hall Polynomials*, 2nd ed. Clarendon Press, Oxford. MR1354144

[16] TRIBE, R. and ZABORONSKI, O. (2011). Pfaffian formulae for one dimensional coalescing and annihilating systems. *Electron. J. Probab.* **16** 2080–2103. MR2851057

E. KANZIEPER
DEPARTMENT OF APPLIED MATHEMATICS
HOLON INSTITUTE OF TECHNOLOGY
HOLON 5810201
ISRAEL
AND
DEPARTMENT OF PHYSICS OF COMPLEX SYSTEMS
WEIZMANN INSTITUTE OF SCIENCE
REHOVOT 7610001
ISRAEL
E-MAIL: eugene.kanzieper@hit.ac.il

M. POPLAVSKYI
R. TRIBE
O. ZABORONSKI
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WARWICK
COVENTRY CV4 7AL
UNITED KINGDOM
E-MAIL: m.poplavskyi@warwick.ac.uk
r.p.tribe@warwick.ac.uk
olegz@maths.warwick.ac.uk

C. TIMM
INSTITUTE OF THEORETICAL PHYSICS
TECHNISCHE UNIVERSITÄT DRESDEN
01062 DRESDEN
GERMANY
E-MAIL: carsten.timm@tu-dresden.de