Central limit theorems for the real eigenvalues of large Gaussian random matrices

Article (Accepted Version)

Central limit theorems for the real eigenvalues of large Gaussian random matrices

N. J. Simm[*]

Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK

October 6, 2016

Abstract

Let G be an $N \times N$ real matrix whose entries are independent identically distributed standard normal random variables $G_{ij} \sim \mathcal{N}(0,1)$. The eigenvalues of such matrices are known to form a two-component system consisting of purely real and complex conjugated points. The purpose of this note is to show that by appropriately adapting the methods of [19], we can prove a central limit theorem of the following form: if $\lambda_1, \ldots, \lambda_{N_\mathbb{R}}$ are the real eigenvalues of G, then for any even polynomial function $P(x)$ and even $N = 2n$, we have the convergence in distribution to a normal random variable

$$\frac{1}{\sqrt{\mathbb{E}(N_\mathbb{R})}} \left(\sum_{j=1}^{N_\mathbb{R}} P(\lambda_j/\sqrt{2n}) - \mathbb{E} \sum_{j=1}^{N_\mathbb{R}} P(\lambda_j/\sqrt{2n}) \right) \to \mathcal{N}(0, \sigma^2(P)) \qquad (0.1)$$

as $n \to \infty$, where $\sigma^2(P) = \frac{2-\sqrt{2}}{2} \int_{-1}^{1} P(x)^2 \, dx$.

1 Introduction

How many eigenvalues of a random matrix are real? This very natural and fundamental question was asked in 1994 by Edelman, Kostlan and Shub [7] who proved that if G is an $N \times N$ matrix of independent identically distributed standard normal variables, and $N_\mathbb{R}$ is the number of real eigenvalues of G, then

$$\mathbb{E}(N_\mathbb{R}) = \sqrt{2N/\pi} + O(1), \qquad N \to \infty. \qquad (1.1)$$

Note that we are not assuming G is symmetric, in the usual parlance we say that G belongs to the so-called Ginibre ensemble of real non-Hermitian random matrices, first considered by Ginibre in 1965 [16].

In addition to being of instrinsic mathematical interest, the statistics of non-Hermitian matrices also have important applications. The earliest such application is probably due to May [26] who showed that real random matrices describe the stability properties of large biological systems. Very recently it was shown [13] that the counting of the average number of equilibria in a non-linear analogue of May's model can be mapped to the problem of $N_\mathbb{R}$ and to the density of real eigenvalues in the Ginibre type ensembles. See also [15, 27] for further applications of $N_\mathbb{R}$ to the enumeration of equilibria in complex systems. The question of *fluctuations* in such contexts is usually extremely difficult and has only recently begun to receive attention [36].

---

[*]n.simm@warwick.ac.uk

The purpose of this article is to describe the asymptotic central limit theorem fluctuations around Edelman and company's estimate (1.1). In other words, thinking of (1.1) as a law of large numbers, what happens when one recenters $N_{\mathbb{R}}$ with respect to its expectation and studies the convergence in law of the fluctuating remainder?

Our approach to this problem is based on a formalism recently developed in [19], which allowed the authors to characterize the large deviation behaviour for the probability of an anomalously small number of real eigenvalues of G. We will show how it is possible to adapt their methods to prove a central limit theorem for the number of real eigenvalues, in addition to the following generalization. From now on let $N = 2n$ be even and denote the real eigenvalues of G by $\lambda_1, \lambda_2, \ldots, \lambda_{N_{\mathbb{R}}}$. The quantity,

$$X_n^{\mathbb{R}}(P) = \sum_{j=1}^{N_{\mathbb{R}}} P(\lambda_j/\sqrt{N}), \tag{1.2}$$

is known as a *linear statistic* (but crucially, note that we only sum the real eigenvalues). The count of real eigenvalues is the special case $X_n^{\mathbb{R}}(1) = N_{\mathbb{R}}$.

Theorem 1.1. *The variance of the total number of real eigenvalues of the standard $2n \times 2n$ Ginibre real random matrix is given by*

$$\mathrm{Var}(N_{\mathbb{R}}) = \frac{2\sqrt{2}}{\sqrt{\pi}} \sum_{k=1}^{n} \frac{\Gamma(2k - 3/2)}{\Gamma(2k - 1)} - \frac{2}{\pi} \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \frac{\Gamma(k_1 + k_2 - 3/2)^2}{\Gamma(2k_1 - 1)\Gamma(2k_2 - 1)} \tag{1.3}$$

and has $n \to \infty$ asymptotics given by

$$\mathrm{Var}(N_{\mathbb{R}}) = (2 - \sqrt{2})\mathbb{E}(N_{\mathbb{R}}) + O(1), \qquad n \to \infty \tag{1.4}$$

Let us note that the asymptotics (1.4) also appear in [9] and weaker variance estimates (without the constant $2 - \sqrt{2}$) were obtained in [38] for non-Gaussian matrices. The same asymptotics (including the constant) apply to the generalized eigenvalue problem of real Ginibre matrices [11]. Formulae (1.3) and (1.4) are proved in Section 2.2, including a generalization to the variance of (1.2) for P an even polynomial, see Proposition 2.5. We also have a central limit theorem for linear statistics:

Theorem 1.2. *Let $P(x)$ be any even polynomial with real coefficients and let $N = 2n$ be even. Then in the limit $n \to \infty$, we have the convergence in distribution*

$$\frac{1}{\sqrt{\mathbb{E}(N_{\mathbb{R}})}} (X_n^{\mathbb{R}}(P) - \mathbb{E}(X_n^{\mathbb{R}}(P)) \to \mathcal{N}(0, \sigma^2(P)) \tag{1.5}$$

where $\mathcal{N}(0, \sigma^2(P))$ denotes the normal distribution with mean 0 and variance

$$\sigma^2(P) := \frac{2 - \sqrt{2}}{2} \int_{-1}^{1} P(x)^2 \, dx \tag{1.6}$$

For *Hermitian* random matrices, results of this type continue to occupy a major industry in the field, since at least the 1980s [18] with work continuing unabated to the present day. A quite comprehensive treatment was given by Johansson [17], who proved that for a general class of Hermitian ensembles, the linear statistic (1.2) converges as $N \to \infty$, *without normalization*, to a normal random variable with finite variance. The lack of any normalization is usually interpreted as a consequence of strong correlations between the eigenvalues; indeed, for Hermitian matrices, the variance of (1.2) remains bounded in N. In the non-Hermitian case, including all complex eigenvalues in the sum (1.2) leads again to a bounded variance central limit theorem which is closely related to the Gaussian free field (GFF) [8, 31, 32, 2, 29], a log-correlated field of great importance in mathematical physics and probability, see [33] for a survey. See also [14, 23, 43] for further relations between linear statistics of random matrices and log-correlated fields. An

important question for future work could be to determine if there is a process interpolating between the Poisson fluctuations of Theorem 1.2 and the GFF obtained in [32].

The $N \to \infty$ fate of the sum (1.2) is therefore quite different to that typically encountered in random matrix theory, requiring a normalization of order $N^{-1/4}$ to ensure distributional convergence. Furthermore, linear statistics of random matrix eigenvalues involving a *random number of terms* have not been studied so widely. However, the Poissonian structure of the limiting Gaussian process can be guessed at in the following way. Viewed as a point process, it is known [41, 40, 42] that the unscaled law of the real Ginibre eigenvalues converges as $N \to \infty$ to a system of annihilating Brownian motions taken at time $t = 1$. Due to the relatively weak correlations in the bulk between blocks of particles of macroscopic size, combined with Edelman's law (1.1) we may expect that (1.2) is close to a sum of $O(\sqrt{N})$ independent random variables, for which the classical central limit theorem is applicable. These heuristics are enough to guess (1.5), but do not seem to explain the constant[1] $2 - \sqrt{2}$ in (1.6).

For finite N, the real spectrum of a Ginibre matrix is not completely independent and therefore (1.5) requires its own proof. The results of [10] and [3] indicate that the real eigenvalues have quite interesting statistics, with linear repulsion at close range and Poisson behaviour at large spacings. Specifically, it is shown in [10] that if $p_{\mathrm{GinOE}}(s)$ is the probability density of real eigenvalue spacings, then

$$
\begin{aligned}
p_{\mathrm{GinOE}}(s) &\sim c_0 s, & s &\to 0 \\
p_{\mathrm{GinOE}}(s) &\sim c_1^2 e^{-c_1 s}, & s &\to \infty
\end{aligned}
\tag{1.7}
$$

where $c_0 = 1/(2\sqrt{2\pi})$ and $c_1 = \zeta(3/2)/c_0$. This should be contrasted with the case of random symmetric matrices which have the Wigner-Dyson form (see [28])

$$
\begin{aligned}
p_{\mathrm{GOE}}(s) &\sim (\pi^2/6)s, & s &\to 0 \\
p_{\mathrm{GOE}}(s) &\sim e^{-(\pi s)^2/16}, & s &\to \infty
\end{aligned}
\tag{1.8}
$$

In [3], the real eigenvalues of non-Hermitian matrices are shown to characterize level crossings in a superconducting quantum dot. Although not of the Ginibre type, the ensembles considered in [3] seem to share the same 'mermaid statistics' as (1.7).

Finally, as noted in [9], the real eigenvalues of Ginibre matrices bare a close analogy to the study of real roots of random polynomials of high degree. For a quite general class of random polynomials, variance estimates and central limit theorems for the number of real roots were obtained by Maslova [25, 24]. See [37] for further references and recent progress in the field of random polynomials. An ensemble of random polynomials closely related to the present study are the $SO(2)$ polynomials defined by $p(x) = \sum_{j=0}^{N} c_j x^j$ where c_j are i.i.d. Gaussian variables with mean zero and variance $\binom{N}{j}$. As for the Ginibre ensemble, the mean and variance of the number of real roots scale as $\sqrt{N}$ [4]

$$
\mathrm{Var}(N_{\mathbb{R}}^{\mathrm{SO}(2)}) \sim c\sqrt{N}
\tag{1.9}
$$

where the constant $c = 0.57173\ldots$ is close to the Ginibre constant $2 - \sqrt{2} = 0.5857\ldots$ in (1.5). We do not yet have a good explanation for this closeness.

To prove Theorem 1.2, we rely on the fact that the Ginibre ensemble is a Pfaffian point process. This means that all real and complex correlation functions of the eigenvalues can be written as a Pfaffian [5, 9, 12, 35], in addition to the class of ensemble averages described in [34]. These results rely on the explicit knowledge of the joint probability density function of real and complex eigenvalues [22, 6]. In fact, for f even, the moment generating function of the random variable (1.2) is actually a *determinant* of size $n \times n$. In general, if f is not even it is a Pfaffian of size $2n \times 2n$ that seems more difficult to analyze. From the determinantal formulae, the

---

[1]Interestingly, in the parlance of log-gases, the $2-\sqrt{2}$ prefactor has the physical interpretation as the *compressibility* of the particle system [10].

cumulants of (1.2) can easily be extracted, and further analysis of their asymptotic behaviour is made possible by appropriately modifying the method used in [19].

Note: During the preparation of this article, the arXiv submission [21] appeared, which proves Theorem 1.2 under the (different) condition that P is compactly supported inside $(-1, 1)$ (so does not cover the case of $N_\mathbb{R}$ or the polynomial statistics of Theorem 1.2). It is likely that combining the methods of [21] and the present article would yield an improved regularity condition on P.

2 Proof of the main result

In the first section we compute the joint cumulant generating function of linear statistics of real and complex eigenvalues. In the second section we calculate the variance and prove Theorem 1.1. In the final section we bound the higher order cumulants and establish our main result, Theorem 1.2.

2.1 Pfaffian and determinantal structures

The first step towards proving (1.5) is to calculate the moment generating function of the statistic (1.2). A key role (see [20] and [34]) is played by the real and complex integrals

$$A[h(x)h(y)]_{jk} = \frac{1}{2} \int_\mathbb{R} dx \int_\mathbb{R} dy \, h(x)h(y)e^{-x^2/2-y^2/2}P_{j-1}(x)P_{k-1}(y)\mathrm{sign}(y-x) \tag{2.1}$$

$$B[g(z)g(\overline{z})]_{jk} = -2i \int_\mathbb{C} g(z)g(\overline{z})P_{j-1}(z)P_{k-1}(\overline{z})\mathrm{sign}(\Im(z))e^{-z^2/2-\overline{z}^2/2}\mathrm{erfc}(\sqrt{2}|\Im(z)|)\,d^2z \tag{2.2}$$

where $\{P_j(x)\}_{j\geq 0}$ are a family of degree j monic polynomials. We will choose them to be skew-orthogonal with respect to (2.1) and (2.2), as in [9] where they were calculated to be

$$P_{2j}(x) = x^{2j}, \qquad P_{2j+1}(x) = x^{2j+1} - 2jx^{2j-1} \tag{2.3}$$

With these polynomials specified, the following skew-orthogonality relation is satisfied:

$$A[1] + B[1] = \mathrm{diag}\left\{ \begin{pmatrix} 0 & r_{j-1} \\ -r_{j-1} & 0 \end{pmatrix} \right\}_{j=1}^n \tag{2.4}$$

where $r_{j-1} = \sqrt{2\pi}\Gamma(2j-1)$.

Proposition 2.1. *Let $f \in L^2(\mathbb{R})$ and $g \in L^2(\mathbb{C})$ be integrable functions and consider the linear statistics*

$$X_N^\mathbb{R}(f) = \sum_{j=1}^{N_\mathbb{R}} f(\lambda_j), \qquad X_N^\mathbb{C}(g) = \sum_{j=1}^{N_\mathbb{C}} g(z_j) \tag{2.5}$$

Then the joint cumulant generating function of (2.5) is given by:

$$\log \mathbb{E}\left(\exp\left(sX_N^\mathbb{R}(f) + tX_N^\mathbb{C}(g) \right) \right) = \frac{1}{2} \log \det \left(I_{2n} + M^\mathbb{R}[e^{sf(x)+sf(y)} - 1] + M^\mathbb{C}[e^{tg(z)+tg(\overline{z})} - 1] \right) \tag{2.6}$$

where I_{2n} is the $2n \times 2n$ identity matrix and $M^{\mathbb{R}/\mathbb{C}}[h(x,y)]$ are $2n \times 2n$ block matrices, where block (j,k) is given by

$$\begin{aligned}
M^\mathbb{R}[h(x,y)]_{jk} &:= \frac{1}{\sqrt{2\pi}\Gamma(2j-1)} \begin{pmatrix} -A[h(x,y)]_{2j,2k-1} & -A[h(x,y)]_{2j,2k} \\ A[h(x,y)]_{2j-1,2k-1} & A[h(x,y)]_{2j-1,2k.} \end{pmatrix} \\
M^\mathbb{C}[g(z,\overline{z})]_{jk} &:= \frac{1}{\sqrt{2\pi}\Gamma(2j-1)} \begin{pmatrix} -B[g(z,\overline{z})]_{2j,2k-1} & -B[g(z,\overline{z})]_{2j,2k} \\ B[g(z,\overline{z})]_{2j-1,2k-1} & B[g(z,\overline{z})]_{2j-1,2k} \end{pmatrix}
\end{aligned} \tag{2.7}$$

4

Remark 2.2. The resulting structure of Proposition 2.1 is reminiscent of formula (3.1) in Tracy and Widom [39], which proved to be extremely useful for the $\beta = 2$ Hermitian ensembles.

Proof. This follows from a result of Sinclair [34] combined with an important observation of Forrester and Nagao [9]. Namely, we apply Theorem 2.1 of [34] but as was observed in [9] the proof continues to hold separately for the real and complex eigenvalues. Namely, if we define

$$X_N^{\mathbb{C}}(g) = \sum_{j=1}^{N_{\mathbb{C}}} g(z_j) \tag{2.8}$$

where z_j are the purely complex eigenvalues, then one has a slightly more general statement

$$\mathbb{E}(\exp(sX_N^{\mathbb{R}}(f) + tX_N^{\mathbb{C}}(g))) = \frac{\mathrm{Pf}(A[e^{sf(x)+sf(y)}] + B[e^{tg(z)+tg(\overline{z})}])}{2^{N(N+1)/4}\prod_{j=1}^{N}(\Gamma(j/2))}. \tag{2.9}$$

By normalization of the generating function and linearity of the scalar products A and B, we have

$$\mathbb{E}(\exp(sX_N^{\mathbb{R}}(f) + tX_N^{\mathbb{C}}(g))) = \frac{\mathrm{Pf}(A[1] + B[1] + A[e^{sf(x)+sf(y)} - 1] + B[e^{tg(z)+tg(\overline{z})} - 1])}{\mathrm{Pf}(A[1] + B[1])} \tag{2.10}$$

Due to the skew-orthogonality of the $P_j's$, the matrix $A[1] + B[1]$ is block diagonal and skew-symmetric:

$$A[1] + B[1] = \mathbf{r} \otimes J, \qquad \mathbf{r} = \mathrm{diag}(r_0, \ldots, r_{n-1}), \qquad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.11}$$

with $r_j = \sqrt{2\pi}\Gamma(2j + 1)$. Taking logarithms and writing the Pfaffians as square roots of determinants gives (2.6) after elementary algebra. $\qquad\square$

Remark 2.3. A further simplification occurs whenever the functions f and g are both even. In this case the Pfaffian has a checkerboard structure of zeros and the Pfaffians reduce to determinants of half the size. We then have a *bona fide* determinant

$$\mathbb{E}(\exp(sX_N^{\mathbb{R}}(f) + tX_N^{\mathbb{C}}(g))) = \det\left\{\delta_{jk} + \frac{(A[e^{sf(x)+sf(y)-1}] + B[e^{tg(z)+tg(\overline{z})} - 1])_{2j-1,2k}}{\sqrt{2\pi\Gamma(2j-1)\Gamma(2k-1)}}\right\}_{j,k=1}^{n} \tag{2.12}$$

which is a generalization of formula (6) in [19] (setting $g = 0$ and $f = 1$). See also [12] for similar calculations.

To proceed, we will focus our attention on the real eigenvalues and set $g \equiv 0$ from now on. To prove the central limit theorem we will calculate the cumulants of $X_N^{\mathbb{R}}(f)$, for which the determinantal formula (2.6) is quite well-suited.

Lemma 2.4. *The lth order cumulant κ_l of any even linear statistic $X_N^{\mathbb{R}}(f)$ is given by*

$$\kappa_l(f) = l! \sum_{m=1}^{l} \frac{(-1)^{m+1}}{m} \sum_{\substack{\nu_1+\ldots+\nu_m=l \\ \nu_i \geq 1}} \frac{\mathrm{Tr}M^{(\nu_1)}[f]\ldots M^{(\nu_m)}[f]}{\nu_1!\ldots\nu_m!} \tag{2.13}$$

where

$$M^{(\nu)}[f]_{jk} := \frac{A[(f(x) + f(y))^{\nu}]_{2j-1,2k}}{\sqrt{2\pi\Gamma(2j-1)\Gamma(2k-1)}} \tag{2.14}$$

and $A[f(x, y)]$ is given by (2.1).

5

Proof. From formula (2.12) with $g = 0$, we get

$$[s^l] \log \mathbb{E}(\exp(sX_N^{\mathbb{R}}(f))) = [s^l] \log \det \left\{ \delta_{jk} + \frac{(A[e^{sf(x)+sf(y)} - 1]_{2j-1,2k}}{\sqrt{2\pi\Gamma(2j-1)\Gamma(2k-1)}} \right\}_{j,k=1}^n \tag{2.15}$$

$$= [s^l] \text{Tr} \log \left\{ \delta_{jk} + \frac{(A[e^{sf(x)+sf(y)} - 1]_{2j-1,2k}}{\sqrt{2\pi\Gamma(2j-1)\Gamma(2k-1)}} \right\}_{j,k=1}^n \tag{2.16}$$

$$= [s^l] \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \text{Tr} \left(\left\{ \frac{A[e^{sf(x)+sf(y)} - 1]_{2j-1,2k}}{\sqrt{2\pi\Gamma(2j-1)\Gamma(2k-1)}} \right\}_{j,k=1}^n \right)^m \tag{2.17}$$

Expanding the term $e^{s(f(x)+f(y))} - 1$ in a Taylor series and re-ordering the sum gives (2.13). $\quad\square$

2.2 The covariance

The main purpose of this section is to prove the following

Proposition 2.5. *Let $P(x)$ and $Q(x)$ be any even polynomials with real coefficients. Then the covariance of the linear statistics $X_n^{\mathbb{R}}[P]$ and $X_n^{\mathbb{R}}[Q]$ satisfies the asymptotic formula*

$$\lim_{n\to\infty} \text{Cov}\left\{ n^{-1/4} X_n^{\mathbb{R}}[P], n^{-1/4} X_n^{\mathbb{R}}[Q] \right\} = \frac{(2-\sqrt{2})}{\sqrt{\pi}} \int_{-1}^1 P(x)Q(x)\,dx \tag{2.18}$$

To compute the covariance of a general polynomial linear statistic, it suffices to just consider the case of monomials

$$C_{p,q} := \text{Cov}(X_n^{\mathbb{R}}(\lambda^p), X_n^{\mathbb{R}}(\lambda^q)) \tag{2.19}$$

Our goal in what follows will be to first find an exact formula for $C_{p,q}$ in Lemmas 2.6 and 2.7, and then compute the large-n asymptotics, which is done in Proposition 2.10. Throughout the paper we will make use of the notation

$$f_{j,k}^{(r,s)} := A[x^r y^s]_{jk}. \tag{2.20}$$

Lemma 2.6. *The covariance of two even monomial linear statistics is given for any even matrix dimension $N = 2n$ by the formula:*

$$\begin{aligned}
C_{p,q} = n^{-(p+q)/2} \sum_{k_1=1}^n \frac{f_{2k_1-2,2k_1-1}^{(p,q)} + f_{2k_1-2,2k_1-1}^{(q,p)} + f_{2k_1-2,2k_1-1}^{(0,p+q)} + f_{2k_1-2,2k_1-1}^{(p+q,0)}}{\sqrt{2\pi}\Gamma(2k_1-1)} \\
- n^{-(p+q)/2} \sum_{k_1,k_2=1}^n \frac{f_{2k_1-2,2k_2-1}^{(0,p)} f_{2k_2-2,2k_1-1}^{(0,q)} + f_{2k_1-2,2k_2-1}^{(p,0)} f_{2k_2-2,2k_1-1}^{(q,0)}}{2\pi\Gamma(2k_1-1)\Gamma(2k_2-1)}
\end{aligned} \tag{2.21}$$

Proof. This follows from expressing $C_{p,q}$ in terms of variances using the identity

$$2C_{p,q} = \kappa_2(\lambda^p + \lambda^q) - \kappa_2(\lambda^p) - \kappa_2(\lambda^q) \tag{2.22}$$

The variance terms are then calculated from Lemma 2.4 with $l = 2$. $\quad\square$

The coefficients in (2.20) appearing in (2.21) can be evaluated in the following convenient form.

Lemma 2.7. *For any even p and q, the following exact formula holds:*

$$f_{2k_1-2,2k_2-1}^{(p,q)} = \Gamma(k_1 + k_2 + (p+q)/2 - 3/2) + qE(k_1 + p/2, k_2 + q/2 - 1) \tag{2.23}$$

where

$$E(j,k) := (k-1)! 2^{k-1} \sum_{i=0}^{k-1} \frac{\Gamma(i + j - 1/2)}{2^i i!}. \tag{2.24}$$

The second term is an error term that satisfies the inequality

$$E(k_1 + p/2, k_2 + q/2 - 1) \le c(k_2 + q/2 - 2)! 2^{k_2} \sum_{i=0}^{\infty} \frac{\Gamma(i + k_1 + p/2 - 1/2)}{2^i i!}$$

$$\le c\sqrt{n} 2^{k_1+k_2} \Gamma(k_2 + q/2 - 3/2) \Gamma(k_1 + p/2 - 1/2) \tag{2.25}$$

where c is a constant independent of k_1, k_2 and n.

Remark 2.8. The first term in $\Gamma(k_1 + k_2 + (p+q)/2 - 3/2)$ in (2.23) is a natural generalization of the case $p = q = 0$ found in [19] and will play just as important a role here in determining the $n \to \infty$ asymptotics.

Proof. From the identities $P_{2k_1-2}(x)x^p = P_{r+2k_1-2}(x)$ and $P_{2k_2-1}(y)y^q = P_{q+2k_2-1} + qy^{q+2k_2-3}$ we have

$$f^{(p,q)}_{2k_1-2,2k_2-1} = f^{(0,0)}_{2k_1+p-2,2k_2+q-1} + q f^{(2k_1+p-2,2k_2+q-3)}_{0,0} \tag{2.26}$$

The proof is completed by verifying the following identities which are a simple integration exercise:

$$f^{(0,0)}_{2k_1+p-2,2k_2+q-1} = \Gamma(k_1 + k_2 + (p+q)/2 - 3/2)$$

$$f^{(2k_1+p-2,2k_2+q-3)}_{0,0} = E(k_1 + p/2, k_2 + q/2 - 1) \tag{2.27}$$

$\square$

Remark 2.9. The key point is that to prove Theorem 1.2, it will suffice to only consider the contribution from the first term in (2.23). This is proved more generally for all cumulants in Proposition 2.12.

To extract the asymptotics based on just the first term in (2.23), we have the following

Proposition 2.10. *Consider the sum*

$$S_{p,q} := N^{-(p+q+1)/2} \sum_{k_1,k_2=1}^{N} \frac{\Gamma(k_1 + k_2 + \frac{q}{2} - 3/2)\Gamma(k_1 + k_2 + \frac{p}{2} - 3/2)}{\Gamma(2k_1 - 1)\Gamma(2k_2 - 1)} \tag{2.28}$$

Then the following limit holds:

$$\lim_{n \to \infty} S_{p,q} = \sqrt{\pi} \frac{2^{(p+q+1)/2}}{p + q + 1} \tag{2.29}$$

Proof. Our strategy will be to bound the sum from above and below. An upper bound can be obtained by extending the k_2 range of summation to ∞:

$$S_{p,q} \le n^{-(p+q+1)/2} \sum_{k_1=1}^{n} \frac{\Gamma(k_1 + (p-1)/2)\Gamma(k_1 + (q-1)/2)}{\Gamma(2k_1 - 1)}$$

$$\times {}_2F_1([k_1 + (p-1)/2, k_1 + (q-1)/2], [1/2], 1/4) \tag{2.30}$$

where ${}_2F_1$ is the classical Gauss hypergeometric function. Since the summand is independent of n, it suffices to substitute the $k_1 \to \infty$ asymptotics in (2.30). Hence we need the asymptotics of the hypergeometric function with fixed argument and large parameters. These were calculated by several authors using the method of steepest descent, see e.g. [30]. Indeed, the main result in Section 4 of [30] and Stirling's formula imply that

$$\frac{\Gamma(k_1 + (p-1)/2)\Gamma(k_1 + (q-1)/2)}{\Gamma(2k_1 - 1)} {}_2F_1([k_1 + (p-1)/2, k_1 + (q-1)/2], [1/2], 1/4)$$

$$\sim \sqrt{\pi}(2k_1)^{(p+q-1)/2} = \sqrt{\pi}(2k_1)^{(p+q-1)/2}, \qquad k_1 \to \infty \tag{2.31}$$

Inserting this into the summand of (2.30) shows that

$$\lim_{n\to\infty} S_{p,q} \leq \lim_{n\to\infty} n^{-(p+q+1)/2}\sqrt{\pi}\sum_{k_1=1}^{n}(2k_1)^{(p+q-1)/2}$$

$$= \sqrt{\pi}\frac{2^{(p+q+1)/2}}{p+q+1} \tag{2.32}$$

To obtain a lower bound, we will use the techniques of [19]. The main idea is to write the Gamma functions in the numerator of (2.28) as Gaussian integrals. For $a \geq 0$ even, we have

$$\Gamma(k_1+k_2+a/2-3/2) = 2\int_{\mathbb{R}_+} x^a\,x^{2k_1-2}\,x^{2k_2-2}\,e^{-x^2}\,dx \tag{2.33}$$

Substituting this expression for the numerator in (2.28) and summing over k_1 and k_2 leads to an integral representation

$$S_{p,q} = n^{-(p+q+1)/2}4\int_{\mathbb{R}_+^2} dx_1\,dx_2\,x_1^p x_2^q \cosh_{n-1}(x_1x_2)^2 e^{-x_1^2-x_2^2} \tag{2.34}$$

where we have employed the hyperbolic cosine series $\cosh_{n-1}(x) = \sum_{k=0}^{n-1}\frac{x^{2k}}{(2k)!}$. By Lemma 4 of [19], we have the lower bound

$$\cosh_{n-1}(x_1x_2n) \geq h_n e^{x_1x_2n}1(x_1x_2 < T_n) \tag{2.35}$$

where $\lim_{n\to\infty} T_n = 2$ and $\lim_{n\to\infty} h_n = 1/2$. Changing variables $x_i \to \sqrt{n}x_i$ for $i = 1,2$ in (2.28) and inserting (2.35), we get

$$\begin{aligned}
S_{p,q} &\geq 4\sqrt{n}h_n^2\int_{\mathbb{R}_+^2} dx_1\,dx_2\,x_1^p x_2^q 1(x_1x_2 < S_n)e^{-n(x_1-x_2)^2}\\
&\geq 4\sqrt{n}h_n^2\int_0^{\sqrt{T_n}}\int_0^{\sqrt{T_n}} dx_1\,dx_2\,x_1^p x_2^q e^{-n(x_1-x_2)^2}\\
&= 4\sqrt{n}h_n^2\frac{1}{2}\int_0^{\sqrt{T_n}} dR\int_{-R}^R dz\left(\left(\frac{R+z}{2}\right)^p\left(\frac{R-z}{2}\right)^q\right.\\
&\qquad\qquad\left. + \left(\frac{2\sqrt{2}-R-z}{2}\right)^p\left(\frac{2\sqrt{2}-R+z}{2}\right)^q\right)e^{-nz^2}\\
&\sim 4\sqrt{n}h_n^2\frac{1}{2}\int_0^{\sqrt{T_n}} dR\left((R/2)^{p+q}+((2\sqrt{2}-R)/2)^{p+q}\right)\int_{-R}^R dz\,e^{-nz^2}\\
&\sim \sqrt{\pi}\frac{2^{(p+q+1)/2}}{(p+q+1)}
\end{aligned} \tag{2.36}$$

where we used that the domain $\{x_1x_2 < T_n\}\cap\mathbb{R}_+^2$ contains the square $[0,\sqrt{T_n}]^2$. The subsequent estimates follow from integration by parts. $\qquad\square$

To complete the proof of Proposition 2.5, it is enough to observe that the first line of (2.21) is asymptotic to

$$4\sum_{k_1=1}^{n}\frac{\Gamma(2k_1+\frac{p+q}{2}-3/2)}{\Gamma(2k_1-1)\sqrt{2\pi}} \sim \frac{2\sqrt{2}(2n)^{(p+q+1)/2}}{p+q+1} \tag{2.37}$$

By Proposition 2.10, the second line is asymptotic to $\frac{(2n)^{(p+q+1)/2}2}{\sqrt{\pi}(p+q+1)}$. The difference of these two terms divided by the normalizing factor $(2n)^{(p+q+1)/2}$ is equal to $\frac{(\sqrt{2}-1)}{\sqrt{\pi}}\int_{-1}^1 x^{p+q}\,dx$. The fact that nothing contributes from the second term in (2.23) is proved for all cumulants in Proposition 2.12.

8

2.3 Higher cumulants and Gaussian fluctuations

Specialising now to the case $f \equiv P$ of an even polynomial, we will prove in this section that the cumulants of (1.2) with any order $l \geq 3$ are $O(\sqrt{n})$ as $n \to \infty$. Due to the normalization of order $n^{-1/4}$ in (1.5), this bound will be sufficient to conclude the central limit theorem and completes the proof of our main result, Theorem 1.2.

By Lemma 2.4 it will suffice to prove that the trace in (2.13) satisfies the bound

$$\mathrm{Tr} M^{(\nu_1)}[P] \dots M^{(\nu_m)}[P] = O(\sqrt{n}), \qquad n \to \infty \tag{2.38}$$

If P is an even polynomial, the above trace is a finite linear combination of terms of the form

$$\mathcal{Z}_{n,m} := n^{-\mathcal{M}_m} \sum_{k_1,\dots,k_m} f^{(2r_1,2s_1)}_{2k_1-2,2k_2-1} \cdots f^{(2r_{m-1},2s_{m-1})}_{2k_{m-1}-2,2k_m-1} f^{(2r_m,2s_m)}_{2k_m-2,2k_1-1} \tag{2.39}$$

where $\mathcal{M}_m = \sum_{i=1}^m (r_i + s_i)$ and as before we have $f^{(2r_i,2s_i)}_{2k-2,2j-1} = A[x^{2r_i} y^{2s_i}]_{2k-1,2j}$ which are explicitly evaluated in (2.23). This follows by definition of the trace and expanding P in a basis of monomials. We will prove (2.38) with two Propositions. First we estimate the leading term in (2.39) by substituting just the first factor from (2.23), denoted $\Gamma^{(2r_i,2s_i)}_{2k_i-2,2k_{i+1}-1}$. Then in the second Proposition we deal with the error term in (2.23) using the bound (2.25).

Proposition 2.11. *Define*

$$\Gamma^{(2r_i,2s_i)}_{2k_i-2,2k_{i+1}-1} = \Gamma(k_i + k_{i+1} + r_i + s_i - 3/2) \tag{2.40}$$

for $i = 1, \dots, m$, where $k_{m+1} \equiv k_1$ and define exponents

$$\mathcal{M}_m = \sum_{i=1}^m (r_i + s_i) \tag{2.41}$$

Then the sum

$$\mathcal{Z}^{(0)}_{n,m} := \sum_{k_1,\dots,k_m} \Gamma^{(2r_1,2s_1)}_{2k_1-2,2k_2-1} \cdots \Gamma^{(2r_{m-1},2s_{m-1})}_{2k_{m-1}-2,2k_m-1} \Gamma^{(2r_m,2s_m)}_{2k_m-2,2k_1-1} \tag{2.42}$$

is $O(n^{1/2+\mathcal{M}_m})$ as $n \to \infty$.

Proof. As for the covariance calculation, we write the Gamma factors as Gaussian integrals:

$$\Gamma^{(r_i,s_i)}_{2k_i-2,2k_{i+1}-1} = \int_{\mathbb{R}} dx\, x^{2r_i+2s_i}\, x^{2k_i-2}\, x^{2k_{i+1}-2}\, e^{-x^2} \tag{2.43}$$

which shows that

$$Z^{(0)}_{n,m} = n^{-\mathcal{M}_m} \int_{\mathbb{R}^m} \prod_{j=1}^m dx_j\, x_j^{2r_j+2s_j} \cosh_{n-1}(x_j x_{j+1}) e^{-x_j^2}. \tag{2.44}$$

We now use the obvious bound $\cosh_{n-1}(x) \leq \cosh(x)$ on every factor (2.44) except one, which we write as a contour integral:

$$\cosh_{n-1}(x_j x_1) = \oint_{\mathcal{C}} \frac{dz}{2\pi i} \frac{z^{-2n+1}}{1-z^2} e^{z x_j x_1} \tag{2.45}$$

where $\mathcal{C}$ is a small loop around $z = 0$. Writing the other cosh factors as exponentials leads to a finite linear combination of terms of the form

$$Z^{(0)}_{n,m} \leq c_m \oint_{\mathcal{C}} \frac{dz}{2\pi i} \frac{z^{-2n+1}}{1-z^2} \int_{\mathbb{R}^m} \left(\prod_{j=1}^m dx_j\, x_j^{2r_j+2s_j} \right) \exp\left(-\mathbf{x}^{\mathrm{T}} A(z) \mathbf{x}\right) \tag{2.46}$$

9

where $\mathbf{x} = (x_1, \ldots, x_m)$ and

$$A(z) = \begin{pmatrix} 1 & -\alpha_1/2 & 0 & 0 & \ldots & 0 & -z/2 \\ -\alpha_1/2 & 1 & -\alpha_2/2 & 0 & 0 & \ldots & 0 \\ 0 & -\alpha_2/2 & 1 & -\alpha_3/2 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & -\alpha_{m-3}/2 & 1 & -\alpha_{m-2}/2 & 0 \\ 0 & \ldots & 0 & 0 & -\alpha_{m-2}/2 & 1 & -\alpha_{m-1}/2 \\ -z/2 & 0 & \ldots & 0 & 0 & -\alpha_{m-1}/2 & 1 \end{pmatrix} \quad (2.47)$$

where $\alpha_i \in \{1, -1\}$. According to Wick's formula, the integral (2.46), which is essentially the moments of a multivariate Gaussian, can be evaluated explicitly in terms of the determinant and inverse of $A(z)$. We have

$$\det(A(z)) = -2^{-m}(z - A_m)((m-1)z + (m+1)A_m) \quad (2.48)$$

and

$$\sigma_{jk}(z) := (A^{-1}(z))_{jk} = \frac{a_{jk}z^2 + b_{jk}z + c_{jk}}{\det(A(z))} \quad (2.49)$$

where $A_m = \pm 1$ and a_{jk}, b_{jk} and c_{jk} are constants. The calculation of σ_{jk} is given in Lemma A.1. As follows from (2.48), all principal minors of $\Re\{A(0)\}$ have positive determinant, implying by Sylvester's criterion that $\Re\{A(0)\}$ is positive definite. By continuity the same is true for $\Re\{A(z)\}$ for sufficiently small $|z| \leq \epsilon$. Hence we take the integration contour to be a sufficiently small circle centered at 0 of radius $\epsilon > 0$, denoted $\mathcal{C}_\epsilon$.

Now let $\mathcal{P}_2$ be the set of all pairings of elements of the set $\{1, 2, \ldots, 2\mathcal{M}_m\}$. Then Wick's formula tells us that

$$\int_{\mathbb{R}^m} \left(\prod_{j=1}^m dx_j \, x_j^{2r_j + 2s_j} \right) \exp\left(-\mathbf{x}^{\mathrm{T}} A(z) \mathbf{x}\right) = \det^{-1/2}\{A(z)\} \sum_{\pi \in \mathcal{P}_2} \prod_{(r,s) \in \pi} \sigma_{\chi(r),\chi(s)}(z) \quad (2.50)$$

where $\chi(r), \chi(s)$ are some indices which will not be important for our considerations and we use the principal branch of the square root throughout. Inserting (2.50) into (2.46), it is apparent that one can set $A_m = 1$ in (2.48), as can be seen by changing variables $z \to zA_m$. Of key importance for estimating the asymptotics is the fact that the number of terms in the product in (2.50) is precisely $\mathcal{M}_m$, so that the integral (2.46) can be bounded by

$$Z_{n,m}^{(0)} \leq c_m \left| \oint_{\mathcal{C}_\epsilon} \frac{dz}{2\pi i} \frac{z^{-2n+1}}{1+z} \frac{K(z)}{(1-z)^{\mathcal{M}_m + 3/2}((m-1)z + (m+1))^{\mathcal{M}_m + 1/2}} \right| \quad (2.51)$$

for some other constant $c_m > 0$. Here $K(z)$ is a polynomial of degree $2\mathcal{M}_m$ with no dependence on n. As in [19], we deform $\mathcal{C}_\epsilon$ to a large circle of radius R centered at zero (say $\mathcal{C}_R$) and notice that the integral on $\mathcal{C}_R$ is exponentially small for large n due to the factor z^{-2n+1} in (2.51). Standard methods of contour integration show that the integral on $\mathcal{C}_R$ can be decomposed as four separate integrals I_j, $j = 1, \ldots, 4$ according to the singularities of the integrand in (2.51). The first integral I_1 is on a small circle around $z = 0$ which is precisely the desired integral in (2.51). Secondly I_2 is an integral along a small circle centered at the simple pole $z = -1$, giving a contribution $K(-1)2^{-(2(\mathcal{M}_m + 1))}$ so that $I_2 = O(1)$ as $n \to \infty$. Thirdly I_3 is an integral encircling the branch cut $[1, R]$ (but avoiding 0) and finally I_4 is an integral encircling the branch cut $[-R, -(m+1)/(m-1)]$ (but avoiding -1). Hence $I_1 = -I_2 - I_3 - I_4 + O(e^{-cRn})$ as $n \to \infty$ for some constant $c > 0$.

To evaluate I_3 we employ a standard technique for complex integrals involving a branch cut, which is to introduce a 'keyhole contour' (see [1] for several examples) of outer radius $R > 0$ (not intersecting $\Re\{z\} = 0$) with a small semi-circle of radius $\epsilon > 0$ around $z = 1$. Along this

contour, we integrate by parts $p := \mathcal{M}_m + 1$ times and take the limit $\epsilon \to 0$ followed by $R \to \infty$. This shows that

$$\lim_{R \to \infty} I_{3,R} = c \int_1^\infty \frac{dy}{\sqrt{y-1}} \frac{d^p}{dy^p} \left[\frac{y^{-2n+1} K(y)}{(1+y)((m-1)y+(m+1))^{\mathcal{M}_m+1/2}} \right] \tag{2.52}$$

for some constant c. To evaluate this integral one can expand out the derivatives and, ignoring the term $(1+y)^{-1}((m-1)y+(m+1))^{-\mathcal{M}_m-1/2}$ and its derivatives (which are uniformly bounded on $[1,\infty)$) we obtain integrals of the generic form

$$n^{p_1} \int_1^\infty dy \frac{y^{-2n+p_2}}{\sqrt{y-1}} = n^{p_1} \frac{\Gamma(2n - p_2 - 1/2)\Gamma(1/2)}{\Gamma(2n - p_2)} \tag{2.53}$$

where p_2 is fixed. The exponent $p_1 \leq p$ comes from taking derivatives of y^{-2n+1} and gives the biggest contribution when $p_1 = p = \mathcal{M}_m + 1$. Inserting this into (2.53) and applying Stirling's formula gives a bound of order $O(n^{p-1/2}) = O(n^{\mathcal{M}_m+1/2})$. The treatment of the integral I_4 is completely analogous and gives rise to a smaller contribution of order $O(n^{\mathcal{M}_m-1/2})$. This ultimately shows that (2.51) is $O(n^{\mathcal{M}_m+1/2})$ and completes the proof of the Proposition. $\qquad\square$

It remains to show that the error terms in (2.25) only give rise to sub-leading contributions in the summation (2.39).

Proposition 2.12. *Consider the sum $Z_{n,m}$ in (2.39), the summands of which consist of a product of m factors. Suppose that $1 \leq c \leq m$ factors are replaced with the error bound in (2.25), while the remaining factors are replaced with the leading Γ-factor in (2.23). Denoting the resulting sum by $Z_{n,m}^{(c)}$, we have*

$$Z_{n,m}^{(c)} = O(n^{\mathcal{M}_m}), \qquad n \to \infty. \tag{2.54}$$

Proof. Due to the factorized form of (2.25), the sum (2.39) is a product of c terms of the form

$$E_{v,\sigma} := \sum_{k_1,\ldots,k_v} \frac{b_{2k_1-1}^{(2s_{\sigma(1)})} \Gamma_{2k_1-2,2k_2-1}^{(2s_{\sigma(2)},2r_{\sigma(2)})} \ldots \Gamma_{2k_v-1-2,2k_v-1}^{(2s_{\sigma(v)},2r_{\sigma(v)})} a_{2k_v-2}^{(2r_{\sigma(v+1)})}}{\Gamma(2k_1-1)\Gamma(2k_2-1)\ldots\Gamma(2k_{v+1}-1)} \tag{2.55}$$

for some permutation σ (corresponding to a re-labelling of the $k_i's$) and $1 \leq v \leq m$. The boundary terms a and b come directly from the error term (2.25) and are given by

$$a_{2k_{v+1}-2}^{(2r_{\sigma(v+1)})} = \sqrt{n} 2^{k_{v+1}} \Gamma\left(k_{v+1} + r_{\sigma(v+1)} - 3/2\right) \tag{2.56}$$

$$b_{2k_1-1}^{(2s_{\sigma(1)})} = 2^{k_1} \Gamma\left(k_1 + s_{\sigma(1)} - 1/2\right) \tag{2.57}$$

The asymptotic behaviour of $E_{v,\sigma}$ as $n \to \infty$ can be estimated according to the programme already outlined for the leading term. We get

$$|E_{v,\sigma}| \leq c\sqrt{n} \left| \oint \frac{dz}{2\pi i} \frac{z^{-2n+1}}{1-z^2} \int_{\mathbb{R}^{v+1}} x_1^{2s_{\sigma(1)}} x_{v+1}^{2r_{\sigma(v+1)}-2} \right.$$

$$\left. \times \prod_{i=2}^v x_i^{2s_{\sigma(i)}+2r_{\sigma(i)}} \exp\left(-\mathbf{x}^{\mathrm{T}} \tilde{A}(z)\mathbf{x}\right) dx_1 \ldots dx_{v+1} \right| \tag{2.58}$$

This time $\tilde{A}(z)$ is symmetric and tridiagonal:

$$\tilde{A}(z) = \begin{pmatrix} 1 & -\sqrt{2}z/2 & 0 & 0 & \ldots & 0 & 0 \\ -\sqrt{2}z/2 & 1 & -\alpha_1/2 & 0 & 0 & \ldots & 0 \\ 0 & -\alpha_1/2 & 1 & -\alpha_2/2 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & -\alpha_{v-3}/2 & 1 & -\alpha_{v-2}/2 & 0 \\ 0 & \ldots & 0 & 0 & -\alpha_{v-2}/2 & 1 & -\sqrt{2}\alpha_{v-1}/2 \\ 0 & 0 & \ldots & 0 & 0 & -\sqrt{2}\alpha_{v-1}/2 & 1 \end{pmatrix} \tag{2.59}$$

11

The Gaussian integral (2.58) is again evaluated by Wick's theorem. It's easy to show that

$$
\det(\tilde{A}(z)) = (1 - z^2)2^{1-v}
$$
$$
\tilde{A}^{-1}(z)_{jk} = \frac{a_{jk} + b_{jk}z + c_{jk}z^2}{1 - z^2}
\tag{2.60}
$$

where a_{jk}, b_{jk} and c_{jk} are constants independent of n and z. Therefore (2.58) gives a finite linear combination of terms of the form

$$
\sqrt{n} \oint_{\mathcal{C}} \frac{dz}{2\pi i} \frac{z^{-2n+1} P(\alpha z)}{(1 - z^2)^{3/2 + \mathcal{E}_v}}
\tag{2.61}
$$

where $\mathcal{E}_v = s_{\sigma(1)} + \sum_{i=2}^{v}(s_{\sigma(i)} + r_{\sigma(i)}) + r_{\sigma(v+1)} - 1$ and $P(\alpha z)$ is an n-independent polynomial. The asymptotics are now dominated by the two branch cuts along $(\pm 1, \infty)$ and one easily sees that the integrals along these cuts are both $O(n^{\mathcal{E}_v + 1})$ as $n \to \infty$ (this includes the $\sqrt{n}$ factor in (2.61)). Taking the product over all such factors gives a bound of order $O(n^{\mathcal{M}_m})$, which is what we wanted to show. □

A   Miscellaneous Lemmas

In [19] the determinant of the matrix $A(z)$ in (2.47) was evaluated explicitly. Due to the application of Wick's theorem, we need the inverse too.

Lemma A.1. *The inverse of the cyclic tridiagonal matrix $A(z)$ in (2.47) is given by*

$$
(A(z)^{-1})_{rs} := \sigma_{rs}(z) = \sigma(0)_{rs} - \frac{p_{rs}(z)}{D_m(z)}
\tag{A.1}
$$

where $D_m(z) = \det(A(z))$, $A_s = \prod_{j=1}^{s} \alpha_j$ and

$$
\sigma(0)_{rs} = (-1)^{r+s} \frac{A_s}{A_r} 2r \frac{j - s + 1}{j + 1}, \qquad r \leq s
\tag{A.2}
$$

and

$$
p_{rs}(z) = 2z^2 \sigma(0)_{mm} \sigma(0)_{1s} \sigma(0)_{r1} - 2z(z\sigma(0)_{m1} - 2)\sigma(0)_{rm}\sigma(0)_{1s}
\tag{A.3}
$$

Proof. The matrix $A(z)$ in (2.47) is called a cyclic tri-diagonal matrix. Its inverse can be calculated by noting that it is a rank 2 perturbation of the tridiagonal matrix $A(0)$:

$$
A(z) = A(0) + R(z)S^{\mathrm{T}}
\tag{A.4}
$$

where $R(z)$ is a $j \times 2$ matrix of zeros except the corners $R(z)_{12} = R(z)_{j2} = -z/2$. Similarly S is a $j \times 2$ matrix of zeros except the corners $S_{11} = 1$ and $S_{j2} = 1$. The inverse now follows from the algebraic identity

$$
A(z)^{-1} = A(0)^{-1} - A(0)^{-1}R(z)(I_2 + S^{\mathrm{T}}A(0)^{-1}R(z))^{-1}S^{\mathrm{T}}A(0)^{-1}
\tag{A.5}
$$

The important part for us is the 2×2 matrix

$$
F := (I_2 + S^{\mathrm{T}}A(0)^{-1}R(z))^{-1}
$$
$$
= \frac{1}{\mathcal{D}_j(z)} \begin{pmatrix} z(A(0)^{-1})_{j1} - 2 & -z(A(0)^{-1})_{11} \\ -z(A(0)^{-1})_{jj} & z(A(0)^{-1})_{j1} - 2 \end{pmatrix}
\tag{A.6}
$$

where $D_j(z) = z^2((A(0)^{-1})_{11}(A(0)^{-1})_{jj} - (A(0)^{-1})_{j,1}^2) + 4z(A(0)^{-1})_{j,1} - 4$. The inverse of the tridiagonal matrix $A(0)$ can be calculated via classical recurrence relations which can be solved explicitly in this case:

$$
(A(0))_{rs}^{-1} =: \sigma(0)_{rs} = (-1)^{r+s} \frac{A_s}{A_r} 2r \frac{j - s + 1}{j + 1}, \qquad r \leq s
\tag{A.7}
$$

This completes the proof of the Lemma. □

12

References

[1] M. J. Ablowitz and A. S. Fokas. *Complex Variables: Introduction and Applications*. Cambridge texts in Applied Mathematics, 2003.

[2] Yacin Ameur, Håkan Hedenmalm, and Nikolai Makarov. Fluctuations of eigenvalues of random normal matrices. *Duke Math. J.*, 159(1):31–81, 2011.

[3] C.W.J. Beenakker, J.M. Edge, J.P. Dahlhaus, D.I. Pikulin, Shuo Mi, and M. Wimmer. Wigner-Poisson Statistics of Topological Transitions in a Josephson Junction. *Phys. Rev. Lett.*, 111:037001, 2013.

[4] Pavel Bleher and Xiaojun Di. Correlations between zeros of a random polynomial. *J. Statist. Phys.*, 88(1-2):269–305, 1997.

[5] A. Borodin and C. D. Sinclair. The Ginibre ensemble of real random matrices and its scaling limits. *Comm. Math. Phys.*, 291(1):177–224, 2009.

[6] A. Edelman. The Probability that a Random Real Gaussian Matrix Has k Real Eigenvalues, Related Distributions, and the Circular Law. *Journal of Multivariate Analysis*, 60:203–232, 1997.

[7] Alan Edelman, Eric Kostlan, and Michael Shub. How many eigenvalues of a random matrix are real? *J. Amer. Math. Soc.*, 7(1):247–267, 1994.

[8] P. J. Forrester. Fluctuation formula for complex random matrices. *J. Phys. A*, 32(13):L159–L163, 1999.

[9] P. J. Forrester and T. Nagao. Eigenvalue Statistics of the Real Ginibre Ensemble. *Phys. Rev. Lett.*, 99:050603, 2007.

[10] Peter J. Forrester. Diffusion processes and the asymptotic bulk gap probability for the real Ginibre ensemble. *J. Phys. A*, 48(32):324001, 14, 2015.

[11] Peter J. Forrester and Anthony Mays. Pfaffian point process for the Gaussian real generalised eigenvalue problem. *Probab. Theory Related Fields*, 154(1-2):1–47, 2012.

[12] Peter J. Forrester and Taro Nagao. Skew orthogonal polynomials and the partly symmetric real Ginibre ensemble. *J. Phys. A*, 41(37):375003, 19, 2008.

[13] Y. V. Fyodorov and B. A. Khoruzhenko. A Nonlinear Analogue of May-Wigner Instability Transition. *Proc. Natl. Acad. Sci. USA*, 113:6827–6832, 2016.

[14] Y. V. Fyodorov, B. A. Khoruzhenko, and N. J. Simm. Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian Unitary Ensemble. *Ann. Probab.*, 44(4):2980–3031, 2016.

[15] Yan V. Fyodorov and Pierre Le Doussal. Topology trivialization and large deviations for the minimum in the simplest random optimization. *J. Stat. Phys.*, 154(1-2):466–490, 2014.

[16] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Mathematical Phys.*, 6:440–449, 1965.

[17] Kurt Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998.

[18] Dag Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, 12(1):1–38, 1982.

[19] E. Kanzieper, M. Poplavskyi, C. Timm, R. Tribe, and O. Zaboronski. What is the probability that a large random matrix has no real eigenvalues? eprint = `arXiv:1503.07926`, 2015.

[20] Eugene Kanzieper and Gernot Akemann. Statistics of real eigenvalues in Ginibre's ensemble of random real matrices. *Phys. Rev. Lett.*, 95(23):230201, 4, 2005.

[21] P. Kopel. Linear Statistics of Non-Hermitian matrices Matching the Real or Complex Ginibre Ensemble to Four Moments. eprint = `1510.02987`, 2015.

[22] N. Lehmann and H-J. Sommers. Eigenvalue statistics of random real matrices. *Physical Review Letters*, 67:941–944, 1991.

[23] A. Lodhia and N.J. Simm. Mesoscopic linear statistics of Wigner matrices. eprint = `arXiv:1503.03533`, 2015.

[24] N. B. Maslova. The distribution of the number of real roots of random polynomials. *Teor. Verojatnost. i Primenen.*, 19:488–500, 1974.

[25] N. B. Maslova. The variance of the number of real roots of random polynomials. *Teor. Verojatnost. i Primenen.*, 19:36–51, 1974.

[26] R. M. May. Will a Large Complex System be Stable? *Nature*, 238:413–414, 1972.

[27] D. Mehta, J. D. Hauenstein, M. Niemerg, N. J. Simm, and D. A. Stariolo. Energy landscape of the finite-size mean-field 2-spin spherical model and topology trivialization. *Phys. Rev. E*, 91:022133, 2015.

[28] Madan Lal Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[29] Sean O'Rourke and David Renfrew. Central Limit Theorem for Linear Eigenvalue Statistics of Elliptic Random Matrices. *J. Theoret. Probab.*, 29(3):1121–1191, 2016.

[30] R. B. Paris. Asymptotics of the Gauss hypergeometric function with large parameters, I. *Journal of Classical Analysis*, 2(2):183–203, 2013.

[31] B. Rider and Jack W. Silverstein. Gaussian fluctuations for non-Hermitian random matrix ensembles. *Ann. Probab.*, 34(6):2118–2143, 2006.

[32] Brian Rider and Bálint Virág. The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not. IMRN*, (2):Art. ID rnm006, 33, 2007.

[33] Scott Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields*, 139(3-4):521–541, 2007.

[34] Christopher D. Sinclair. Averages over Ginibre's ensemble of random real matrices. *Int. Math. Res. Not. IMRN*, (5):Art. ID rnm015, 15, 2007.

[35] H-J. Sommers and W. Wieczorek. General eigenvalue correlations for the Ginibre ensemble. *J. Phys. A: Math. Theor.*, 41(40), 2008.

[36] E. Subag. The complexity of spherical p-spin models - a second moment approach. eprint = `1504.02251`, 2015.

[37] Terence Tao and Van Vu. Local universality of zeroes of random polynomials. *Int. Math. Res. Not. IMRN*, (13):5053–5139, 2015.

[38] Terence Tao and Van Vu. Random matrices: universality of local spectral statistics of non-Hermitian matrices. *Ann. Probab.*, 43(2):782–874, 2015.

[39] Craig A. Tracy and Harold Widom. Correlation functions, cluster functions, and spacing distributions for random matrices. *J. Statist. Phys.*, 92(5-6):809–835, 1998.

[40] Roger Tribe, Siu Kwan Yip, and Oleg Zaboronski. One dimensional annihilating and coalescing particle systems as extended Pfaffian point processes. *Electron. Commun. Probab.*, 17:no. 40, 7, 2012.

[41] Roger Tribe and Oleg Zaboronski. Pfaffian formulae for one dimensional coalescing and annihilating systems. *Electron. J. Probab.*, 16:no. 76, 2080–2103, 2011.

[42] Roger Tribe and Oleg Zaboronski. The Ginibre evolution in the large-N limit. *J. Math. Phys.*, 55(6):063304, 26, 2014.

[43] C. Webb. On the logarithm of the characteristic polynomial of the Ginibre ensemble. eprint = 1507.08674, 2015.