Integrable structure of Ginibre's ensemble of real random matrices and a Pfaffian integration theorem

Gernot Akemann[1] **and Eugene Kanzieper**[2]

[1] Department of Mathematical Sciences and BURSt Research Centre, Brunel University West London, Uxbridge UB8 3PH, United Kingdom

[2] Department of Applied Mathematics, H.I.T. – Holon Institute of Technology, Holon 58102, Israel

E-mail: Gernot.Akemann@brunel.ac.uk, Eugene.Kanzieper@weizmann.ac.il

Abstract. In the recent publication [E. Kanzieper and G. Akemann, Phys. Rev. Lett. **95**, 230201 (2005)], an exact solution was reported for the probability $p_{n,k}$ to find exactly k real eigenvalues in the spectrum of an $n \times n$ real asymmetric matrix drawn at random from Ginibre's Orthogonal Ensemble (GinOE). In the present paper, we offer a detailed derivation of the above result by concentrating on the proof of the Pfaffian integration theorem, the key ingredient of our analysis of the statistics of real eigenvalues in the GinOE. We also initiate a study of the correlations of complex eigenvalues and derive a formula for the joint probability density function of all complex eigenvalues of a GinOE matrix restricted to have exactly k real eigenvalues. In the particular case of $k = 0$, all correlation functions of complex eigenvalues are determined.

PACS numbers: 02.10.Yn, 02.50.-r, 05.40.-a, 75.10.Nr

arXiv: math-ph/0703019 (submitted on March 05, 2007)

Contents

1. Introduction

1.1. Motivation

This study grew out of our attempt to answer the question raised by A. Edelman (1997): "What is the probability $p_{n,k}$ that an $n \times n$ random real Gaussian matrix has exactly k real eigenvalues?" In the physics literature, an ensemble of such random matrices is known as GinOE – Ginibre's Orthogonal Ensemble (Ginibre 1965). Looking into this particular problem, we have realised that no comprehensive solution for the probability $p_{n,k}$ can be found without undertaking an in-depth study (Kanzieper and Akemann 2005) of the integrable structure of GinOE. The results of our investigation are reported in the present paper.

1.2. Main results

For the benefit of the readers, we collect our main results into this easy to read subsection with pointers to the sections containing detailed derivations of each statement.

1.2.1. Real part of GinOE spectrum

(I) Probability of Exactly k Real Eigenvalues. Let $\boldsymbol{\mathcal{H}}$ be an $n \times n$ random real matrix whose entries are statistically independent random variables picked from a normal distribution $\mathsf{N}(0,1)$. Then, for $n-k=2\ell$ even, the probability $p_{n,k}$ of exactly k real eigenvalues † occurring is

$$p_{n,k} = p_{n,n-2\ell} = \frac{p_{n,n}}{\ell!}\, Z_{(1^\ell)}(p_1, \cdots, p_\ell), \tag{1.1a}$$

where $p_{n,n}$ is the probability $p_{n,n} = 2^{-n(n-1)/4}$ of having *all* n eigenvalues real (Edelman 1997). The *universal* multivariate functions $Z_{(1^\ell)}$, solely determined by the number ℓ of pairs of complex conjugated eigenvalues, are so-called *zonal polynomials* (Macdonald 1998) that can be written as a sum over all partitions ‡ $\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g})$ of the size $|\boldsymbol{\lambda}| = \ell$

$$Z_{(1^\ell)}(p_1, \cdots, p_\ell) = (-1)^\ell \ell! \sum_{|\boldsymbol{\lambda}|=\ell} \prod_{j=1}^{g} \frac{1}{\sigma_j!}\left(-\frac{p_{\ell_j}}{\ell_j}\right)^{\sigma_j}. \tag{1.1b}$$

A few first zonal polynomials are displayed in Table 1. The arguments p_j's of the zonal polynomials are *nonuniversal* §

$$p_j = \operatorname{tr}_{(0,\lfloor n/2\rfloor-1)} \hat{\boldsymbol{\varrho}}^{\,j}, \tag{1.1c}$$

† The number of complex eigenvalues $n-k=2\ell$ is always even since the complex part of the spectrum consists of ℓ pairs of complex conjugated eigenvalues.

‡ The notation $\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g})$ is known as the frequency representation of the partition $\boldsymbol{\lambda}$ of the size $|\boldsymbol{\lambda}| = \ell$. It implies that the part ℓ_j appears σ_j times so that $\ell = \sum_{j=1}^{g} \ell_j\, \sigma_j$, where g is the number of inequivalent parts of the partition. In particular, the partition $\boldsymbol{\lambda} = (1^\ell)$ equals $(\underbrace{1, \cdots, 1}_{\ell \text{ times}})$.

§ The notation $\operatorname{tr}_{(0,a)} \hat{\boldsymbol{M}}$ denotes the trace of an $(a+1)\times(a+1)$ matrix $\hat{\boldsymbol{M}}$ such that $\operatorname{tr}_{(0,a)} \hat{\boldsymbol{M}} = \sum_{j=0}^{a} \hat{M}_{jj}$. Also, $\lfloor x \rfloor$ stands for the floor function. In what follows, the ceiling function $\lceil x \rceil$ will be used as well.

Table 1. Explicit examples of zonal polynomials $Z_{(1^\ell)}(p_1, \cdots, p_\ell)$ as defined by the equation (1.1b). Another way to compute $Z_{(1^\ell)}$ is based on the recursion equation (Macdonald 1998)

$$Z_{(1^\ell)}(p_1, \cdots, p_\ell) = (\ell-1)! \sum_{r=0}^{\ell-1} \frac{(-1)^{\ell-r-1}}{r!}\, p_{\ell-r}\, Z_{(1^r)}(p_1, \cdots, p_r)$$

supplemented by the formal "boundary" condition $Z_{(1^0)} = 1$. They are also tabulated in the manuscript by H. Jack (1976).

$\boldsymbol{\lambda}$	$Z_{\boldsymbol{\lambda}}(p_1, \cdots, p_\ell)$
(1^1)	p_1
(1^2)	$p_1^2 - p_2$
(1^3)	$p_1^3 - 3p_1p_2 + 2p_3$
(1^4)	$p_1^4 + 8p_1p_3 - 6p_1^2p_2 + 3p_2^2 - 6p_4$
(1^5)	$p_1^5 - 10p_1^3p_2 + 20p_1^2p_3 + 15p_1p_2^2 - 30p_1p_4 - 20p_2p_3 + 24p_5$

for they depend on a *nonuniversal* matrix $\hat{\boldsymbol{\varrho}}$. For $n = 2m$ even, its entries are

$$\begin{aligned} \hat{\varrho}^{\rm even}_{\alpha,\,\beta} = \int_0^\infty dy\, y^{2(\beta-\alpha)-1}\, e^{y^2}\, {\rm erfc}(y\sqrt{2}) \Big[(2\alpha+1) \\ \times L_{2\alpha+1}^{2(\beta-\alpha)-1}(-2y^2) + 2y^2\, L_{2\alpha-1}^{2(\beta-\alpha)+1}(-2y^2) \Big] \end{aligned} \qquad (1.1d)$$

while for $n = 2m+1$ odd,

$$\hat{\varrho}^{\rm odd}_{\alpha,\,\beta} = \hat{\varrho}^{\rm even}_{\alpha,\,\beta} - (-4)^{m-\beta} \frac{m!}{(2m)!} \frac{(2\beta)!}{\beta!}\, \hat{\varrho}^{\rm even}_{\alpha,\,m}. \qquad (1.1e)$$

Here, the notation ${\rm erfc}(\phi)$ stands for the complementary error function,

$${\rm erfc}(\phi) = \frac{2}{\sqrt{\pi}} \int_\phi^\infty dt\, e^{-t^2},$$

while $L_j^\alpha(\phi)$ denote the generalised Laguerre polynomials

$$L_j^\alpha(\phi) = \frac{1}{j!} \phi^{-\alpha}\, e^{\phi}\, \frac{d^j}{d\phi^j} \left(\phi^{j+\alpha}\, e^{-\phi} \right).$$

The above result (1.1a) will be derived in Section 6.

(II) Generating Function For Probabilities $p_{n,k}$. The generating function $G_n(z)$ for the probabilities $p_{n,k}$ is

$$G_n(z) = \sum_{\ell=0}^{\lfloor n/2 \rfloor} z^\ell p_{n,n-2\ell} = p_{n,n} \det \left[\hat{\mathbf{1}} + z\, \hat{\boldsymbol{\varrho}} \right]_{\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor}. \qquad (1.2)$$

Equation (1.2) with the $\hat{\boldsymbol{\varrho}}$ of needed parity provides us with yet another way of computing the entire set of $p_{n,k}$'s at once! Table 2 contains a comparison of our analytic predictions with numeric simulations. The result (1.2) will be proven in Section 6.

(III) Probability $p_{n,n-2}$ of Exactly One Pair of Complex Conjugated Eigenvalues. For $k = n-2$, the probability function $p_{n,k}$ reduces to

$$p_{n,n-2} = 2\, p_{n,n} \int_0^\infty dy\, y\, e^{y^2} \operatorname{erfc}(y\sqrt{2})\, L_{n-2}^2(-2y^2). \tag{1.3a}$$

An alternative expression reads:

$$p_{n,n-2} = p_{n,n} \left[\sqrt{2} \sum_{j=0}^{\lfloor n/2 \rfloor - 1} 3^{j+\alpha_n/2}\, P_{2j+\alpha_n}\left(\frac{2}{\sqrt{3}}\right) - \lfloor n/2 \rfloor \right]. \tag{1.3b}$$

Here, $\alpha_n = \lceil n/2 \rceil - \lfloor n/2 \rfloor$, and $P_n(\phi)$ stands for the Legendre polynomials

$$P_j(\phi) = \frac{(-1)^j}{2^j j!} \frac{d^j}{d\phi^j} \left[(1-\phi^2)^j \right].$$

The leading large-n behaviour of the probability $p_{n,n-2}$ is given by

$$p_{n,n-2} \approx \frac{3^{n+1/2}}{8\sqrt{\pi\, n}}\, p_{n,n}. \tag{1.3c}$$

The above three results will be derived in Section 4.1, Section 7.1 and Section 7.2, respectively.

1.2.2. Complex part of GinOE spectrum

(IV) Joint Probability Density Function of All Complex Eigenvalues Given There Are k Real Eigenvalues. Let $\boldsymbol{\mathcal{H}}_k$ be an $n \times n$ random real matrix with k real eigenvalues such that its entries are statistically independent random variables picked from a normal distribution $\mathsf{N}(0,1)$. Then, the joint probability density function (j.p.d.f.) of its $2\ell = n-k$ complex eigenvalues is

$$\begin{aligned} P_{\boldsymbol{\mathcal{H}}_k}(z_1, \cdots, z_\ell) &= \frac{p_{n,n}}{\ell!} \left(\frac{2}{i}\right)^\ell \\ &\times \prod_{j=1}^{\ell} \operatorname{erfc}\left(\frac{z_j - \bar{z}_j}{i\sqrt{2}}\right) \operatorname{pf} \left[\begin{array}{cc} \mathcal{D}_n(z_i, z_j) & \mathcal{D}_n(z_i, \bar{z}_j) \\ \mathcal{D}_n(\bar{z}_i, z_j) & \mathcal{D}_n(\bar{z}_i, \bar{z}_j) \end{array} \right]_{2\ell \times 2\ell}. \end{aligned} \tag{1.4}$$

Here, pf denotes the Pfaffian. The above j.p.d.f. is supported for $(\operatorname{Re} z_1, \cdots, \operatorname{Re} z_\ell) \in \mathbb{R}^\ell$, and $(\operatorname{Im} z_1, \cdots, \operatorname{Im} z_\ell) \in (\mathbb{R}^+)^\ell$. The antisymmetric kernel $\mathcal{D}_n(z, z')$ is given explicitly by (3.12) – (3.18) of Section 3 where the statement (1.4) is proven.

(V) Correlation Functions of Complex Eigenvalues in The Spectra Free of Real Eigenvalues. Let $\boldsymbol{\mathcal{H}}_0$ be an $n \times n$ random real matrix with no real eigenvalues such that its entries are statistically independent random variables picked from a normal distribution $\mathsf{N}(0,1)$. Then, the p-point correlation function $(1 \le p \le \ell)$ of its complex eigenvalues, defined by (2.7), equals

$$\begin{aligned} R_{0,p}^{(\boldsymbol{\mathcal{H}}_0)}(z_1, \cdots, z_p; n) &= p_{n,n} \frac{\prod_{j=0}^{\ell-1} r_j}{\prod_{j=1}^{n} \Gamma(j/2)} \prod_{j=1}^{p} \operatorname{erfc}\left(\frac{z_j - \bar{z}_j}{i\sqrt{2}}\right) \exp\left(-\frac{z_j^2 + \bar{z}_j^2}{2}\right) \\ &\times \operatorname{pf} \left[\begin{array}{cc} \kappa_\ell(z_i, z_j) & \kappa_\ell(z_i, \bar{z}_j) \\ \kappa_\ell(\bar{z}_i, z_j) & \kappa_\ell(\bar{z}_i, \bar{z}_j) \end{array} \right]_{2p \times 2p}. \end{aligned} \tag{1.5a}$$

Table 2. Exact solution for $p_{12,k}$ (first and second column) compared to numeric simulations (third column) performed by direct diagonalisation of $1,000,000$ of 12×12 matrices.

k	Analytic solution: Exact	Analytic solution: Approximate	Numeric simulation
0	$\frac{29930323227453-20772686238032\sqrt{2}}{17592186044416}$	0.031452	0.031683
2	$\frac{3(1899624551312\sqrt{2}-2060941421503)}{4398046511104}$	0.426689	0.427670
4	$\frac{3(2079282320189-505722262348\sqrt{2})}{8796093022208}$	0.465235	0.464098
6	$\frac{252911550974\sqrt{2}-27511352125}{4398046511104}$	0.075070	0.075021
8	$\frac{15(1834091507-10083960\sqrt{2})}{17592186044416}$	0.001552	0.001526
10	$\frac{3(1260495\sqrt{2}-512)}{2199023255552}$	0.000002	0.000002
12	$\frac{1}{8589934592}$	0.000000	0.000000

Here, $n = 2\ell$ and the 'pre-kernel' κ_ℓ equals

$$\kappa_\ell(z, z') = i \sum_{j=0}^{\ell-1} \frac{1}{r_j} \Big[p_{2j}(z) p_{2j+1}(z') - p_{2j}(z') p_{2j+1}(z) \Big]. \tag{1.5b}$$

The polynomials $p_j(z)$ are skew orthogonal in the complex half-plane $\operatorname{Im} z > 0$,

$$\langle p_{2j+1}, p_{2k} \rangle_{\rm c} = -\langle p_{2k}, p_{2j+1} \rangle_{\rm c} = i\, r_j\, \delta_{jk}, \tag{1.5c}$$

$$\langle p_{2j+1}, p_{2k+1} \rangle_{\rm c} = \langle p_{2j}, p_{2k} \rangle_{\rm c} = 0, \tag{1.5d}$$

with respect to the skew product

$$\langle f, g \rangle_{\rm c} = \int_{\operatorname{Im} z > 0} d^2 z\, \operatorname{erfc}\left(\frac{z - \bar{z}}{i\sqrt{2}} \right) \exp\left(-\frac{z^2 + \bar{z}^2}{2} \right) \left[f(z) g(\bar{z}) - f(\bar{z}) g(z) \right]. \tag{1.5e}$$

For detailed derivation, a reader is referred to Section 8 which also addresses the problem of calculating the probability $p_{n,0}$ to find no real eigenvalues in the spectrum of GinOE.

1.2.3. How to integrate a Pfaffian?

All the results announced so far would have not been derived without a Pfaffian integration theorem that we consider to be a major technical achievement of our study. Conceptually, it is based on a new, topological, interpretation of the ordered Pfaffian expansion as introduced in Section 5.

(VI) The Pfaffian Integration Theorem. Let $d\pi(z)$ be any benign measure on $z \in \mathbb{C}$, and the function $Q_n(x,y)$ be an antisymmetric function of the form

$$Q_n(x,y) = \frac{1}{2} \sum_{j,k=0}^{n-1} q_j(x)\, \hat{\mu}_{jk}\, q_k(y) \tag{1.6a}$$

where q_j's are arbitrary polynomials of j-th order, and $\hat{\boldsymbol{\mu}}$ is an antisymmetric matrix. Then the integration formula

$$\left(\frac{2}{i}\right)^{\ell} \prod_{j=1}^{\ell} \int_{z_j \in \mathbb{C}} d\pi(z_j)\; \mathrm{pf} \left[\begin{array}{cc} Q_n(z_i, z_j) & Q_n(z_i, \bar{z}_j) \\ Q_n(\bar{z}_i, z_j) & Q_n(\bar{z}_i, \bar{z}_j) \end{array} \right]_{2\ell \times 2\ell}$$
$$= Z_{(1^\ell)} \left(\frac{1}{2} \mathrm{tr}_{(0,n-1)} \hat{\boldsymbol{\upsilon}}^1, \cdots, \frac{1}{2} \mathrm{tr}_{(0,n-1)} \hat{\boldsymbol{\upsilon}}^\ell \right) \tag{1.6b}$$

holds, provided the integrals in its l.h.s. exist. Here, $Z_{(1^\ell)}$ are zonal polynomials whose ℓ arguments are determined by a matrix $\hat{\boldsymbol{\upsilon}}$ with the entries

$$\hat{\upsilon}_{\alpha,\beta} = i \sum_{k=0}^{n-1} \hat{\mu}_{\alpha,k} \int_{z \in \mathbb{C}} d\pi(z) \left[q_k(z)\, q_\beta(\bar{z}) - q_\beta(z)\, q_k(\bar{z}) \right]. \tag{1.6c}$$

This theorem that can be viewed as a generalisation of the Dyson integration theorem (Dyson 1970, Mahoux and Mehta 1991) will be proven in Section 5.

Interestingly, the Pfaffian integration theorem is not listed in the classic reference book (Mehta 2004) on the Random Matrix Theory (RMT). Also, we are not aware of any other RMT literature reporting this result which may have implications far beyond the scope of the present paper.

1.3. A guide through the paper

Having announced the main results of our study, we defer plunging into formal mathematical proofs of the above six statements until Section 3. Instead, in Section 2, we deliberately draw the reader's attention to a comparative analysis of GinOE and two other representatives of non-Hermitean random matrix models known as Ginibre's Unitary (GinUE) and Ginibre's Symplectic (GinSE) Ensemble. Starting with the definitions of the three ensembles, we briefly discuss their diverse physical applications, pinpoint *qualitative* differences between their spectra, and present a detailed *comparative analysis* of major *structural* results obtained for GinUE, GinSE and GinOE since 1965. We took great pains to write a review-style Section 2 in order (i) to help the reader better appreciate a profound difference between GinOE and the two other non-Hermitean random matrix ensembles on both qualitative and structural levels as well as (ii) place our own work in a more general context.

A formal analysis starts with Section 3 devoted to a general consideration of statistics of real eigenvalues. Its first part, Section 3.1, summarises previously known analytic results (Edelman 1997) for the probability function $p_{n,k}$ of the fluctuating number of real eigenvalues in the spectrum of GinOE. Section 3.2 deals with the joint probability density function of complex eigenvalues of GinOE random matrices that have a given number of real eigenvalues. The Pfaffian representation (1.4) is the main outcome of Section 3.2. This result is further utilised in Section 3.3 where the probability function $p_{n,k}$ is put into the form of a 'Pfaffian integral' (3.19). The analysis of the latter expression culminates in concluding that the Dyson integration theorem, a standard tool of Random Matrix Theory, is inapplicable for treating the Pfaffian integral obtained. The latter task will be accomplished in Section 5.

Section 4 attacks the probability function $p_{n,k}$ for a few particular values of k. The probabilities $p_{n,n-2}$, $p_{n,n-4}$ and $p_{n,n-6}$ of one, two, and three pairs of complex conjugated eigenvalues occurring are treated in Section 4.1, Section 4.2 and 4.3, respectively. This is done by explicit calculation of the Pfaffian in (3.19) followed by a term-by-term integration of the resulting Pfaffian expansion. In Section 4.4, we briefly discuss a faster-than-exponential growth of the number of terms in this expansion caused by further decrease of k.

Section 5, devoted to the Pfaffian integration theorem, is central to the paper. Its main objective is to introduce a topological interpretation of the terms arising in a permutational expansion of the Pfaffian in the l.h.s. of (1.6b). Such a topological interpretation turns out to be the proper language in the subsequent proof of the Pfaffian integration theorem. In Section 5.1, the Pfaffian integration theorem is formulated and discussed in the light of the Dyson integration theorem. In Section 5.2, an ordered permutational Pfaffian expansion is defined and interpreted in *topological terms*. The notions of *strings*, *substrings*, *loop-like strings* and *loop-like substrings* for certain subsets of terms arising in an ordered Pfaffian expansion are introduced and illustrated on simple examples in Sections 5.2.1 and 5.2.2. Further, *equivalent strings* and *equivalent classes of strings* are defined and counted. The issue of decomposition of strings into a set of loop-like substrings is also considered in detail (Lemma 5.2). Section 5.2.3 is devoted to the counting of loop-like strings. In Section 5.2.4, the notion of *adjacent* strings is introduced and illustrated. Adjacent strings are counted in Lemma 5.4. Their relation to loop-like strings is discussed in Lemma 5.5. Section 5.2.5 deals with a characterisation of adjacent strings by their *handedness*; adjacent strings of a given handedness are counted there, too. In Section 5.2.6, *equivalent classes of adjacent strings* are defined, counted, and explicitly built. Section 5.3 and Section 5.4 are preparatory for Section 5.5, where the *Pfaffian integration theorem* is eventually proven. For the readers' benefit, a vocabulary of the topological terms we use is summarised in Table 3.

Section 6 utilises the Pfaffian integration theorem to obtain a general solution for the sought probability function $p_{n,k}$ (Section 6.1), derive a determinantal expression for the entire generating function of $p_{n,k}$'s (Section 6.2), and address the issue of integer moments of the fluctuating number of real eigenvalues in GinOE spectra (Section 6.3).

Section 7 is devoted to the probability $p_{n,n-2}$ for two complex conjugated pairs of eigenvalues to occur; a large-n analysis of $p_{n,n-2}$ is also presented there.

Section 8 discusses the Pfaffian structure of the p-point correlation functions of complex eigenvalues belonging to spectra of a subclass of GinOE matrices without real eigenvalues. The same section addresses the problem of calculating the probability $p_{n,0}$ to find no real eigenvalues in spectra of GinOE.

Section 9 contains conclusions with the emphasis placed on open problems. The most involved technical calculations are collected in four appendices.

2. Comparative analysis of GinOE, GinUE, and GinSE

2.1. Definition and consequences of violated Hermiticity

Ginibre's three random matrix models – GinOE, GinUE, and GinSE – were derived from the celebrated Gaussian Orthogonal (GOE), Gaussian Unitary (GUE), and

Gaussian Symplectic (GSE) random matrix ensembles in a purely formal way by dropping the Hermiticity constraint. Consequently, the non-Hermitean descendants of GOE, GUE and GSE share the *same* Gaussian probability density function

$$P_\beta[\boldsymbol{\mathcal{H}}] = (\pi a_\beta)^{-\beta n^2/2} \exp\left[-a_\beta^{-1} \operatorname{tr} \left(\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\dagger\right)\right], \quad \boldsymbol{\mathcal{H}}^\dagger \neq \boldsymbol{\mathcal{H}}, \quad \beta = 1, 2, 4, \tag{2.1}$$

for a matrix "Hamiltonian" $\boldsymbol{\mathcal{H}} \in \mathbb{T}_\beta(n)$ to occur; the constant a_β is chosen to be $a_\beta = 2 - \delta_{\beta,2}$ (this slightly differs from the convention used in the original paper by Ginibre). However, the spaces $\mathbb{T}_\beta$ on which the matrices $\boldsymbol{\mathcal{H}}$ vary are *different*: $\mathbb{T}_1(n)$, $\mathbb{T}_2(n)$, and $\mathbb{T}_4(n)$ span all $n \times n$ matrices with real (GinOE, $\beta = 1$), complex (GinUE, $\beta = 2$), and real quaternion (GinSE, $\beta = 4$) entries, respectively.

The violated Hermiticity, $\boldsymbol{\mathcal{H}}^\dagger \neq \boldsymbol{\mathcal{H}}$, brings about the two major phenomena: (i) complex-valuedness of the random matrix *spectrum* and (ii) splitting the random matrix *eigenvectors* into a bi-orthogonal set of *left* and *right* eigenvectors. Statistics of complex eigenvalues, including their joint probability density function, and statistics of left and right eigenvectors of a random matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}_\beta$ drawn from (2.1) are of primary interest.

Only the spectral statistics will be addressed in the present paper. (For studies of the eigenvector statistics in GinUE, the reader is referred to the papers by Chalker and Mehlig (1998), Mehlig and Chalker (2000), and Janik *et al* (1999). We are not aware of any results for eigenvector statistics in GinSE and GinOE.)

2.2. Physical applications

While the ensemble definition (2.1) was born out of pure mathematical curiosity ‖, non-Hermitean random matrices have surfaced in various fields of knowledge by E. Wigner's "miracle of the appropriateness" (Wigner 1960). From the physical point of view, non-Hermitean random matrices have proven to be as important as their Hermitean counterparts. (For a detailed exposition of physical applications of Hermitean RMT we refer to the review by Guhr *et al* (1998)).

Random matrices drawn from GinUE appear in the description of dissipative quantum maps (Grobe *et al* 1988, Grobe and Haake 1989) and in the characterisation of two-dimensional random space-filling cellular structures (Le Cäer and Ho 1990, Le Cäer and Delannay 1993).

Ginibre's Orthogonal Ensemble of random matrices arises in the studies of dynamics (Sommers *et al* 1988, Sompolinsky *et al* 1988) and of synchronisation effect (Timme *et al* 2002, Timme *et al* 2004) in random networks; GinOE is also helpful in the statistical analysis of cross-hemisphere correlation matrix of the cortical electric activity (Kwapień *et al* 2000) as well as in the understanding of inter-market financial correlations (Kwapień *et al* 2006).

All three Ginibre ensembles (GinOE, GinUE, GinSE) arise in the context of directed "quantum chaos" (Efetov 1997a, Efetov 1997b, Kolesnikov and Efetov 1999, Fyodorov *et al* 1997, Markum *et al* 1999). Their chiral counterparts (Stephanov 1996, Halasz *et al* 1997, Osborn 2004, Akemann 2005) help elucidate universal aspects of the phenomenon of spontaneous chiral symmetry breaking in quantum chromodynamics (QCD) with chemical potential: the presence or absence of real eigenvalues in the complex spectrum singles out different chiral symmetry breaking patterns. For a review of QCD applications of non-Hermitean random matrices with built-in chirality, the reader is referred to Akemann (2007).

‖ J. Ginibre, private communication.

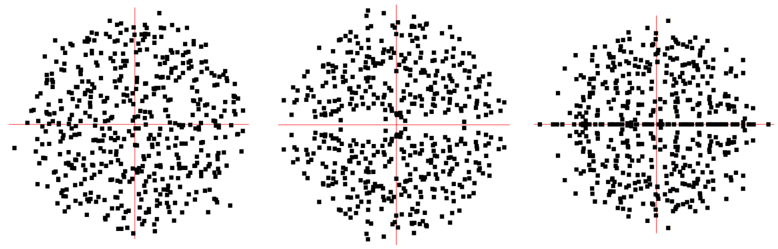

Figure 1. Numerically simulated distributions of complex eigenvalues in GinUE (left panel), GinSE (middle panel), and GinOE (right panel). Three different eigenvalue patterns are clearly observed. Eigenvalues are scattered almost uniformly in GinUE, depleted from the real axis in GinSE, and accumulated along the real axis in GinOE.

Other recent findings (Zabrodin 2003) associate statistical models of non-Hermitean normal random matrices with integrable structures of conformal maps and interface dynamics at both classical (Mineev-Weinstein *et al* 2000) and quantum (Agam *et al* 2002) scales. For a comprehensive review of these and other physical applications, the reader is referred to the survey paper by Fyodorov and Sommers (2003).

2.3. Spectral statistical properties of Ginibre's random matrices

A profound difference between spectral patterns of the three non-Hermitean random matrix models has been realised long ago. Anticipated in the early papers by Ginibre (1965) and Mehta and Srivastava (1966), it was further confirmed analytically by using varied techniques ¶ (Edelman 1997, Efetov 1997a, Efetov 1997b, Kolesnikov and Efetov 1999, Kanzieper 2002a, Kanzieper 2002b, Nishigaki and Kamenev 2002, Splittorff and Verbaarschot 2004, Kanzieper 2005, Akemann and Basile 2007).

Qualitatively, there is a general consensus that (i) the spectrum of GinUE is approximately characterised by a *uniform* density of complex eigenvalues. This is not the case for the two other ensembles. (ii) In GinSE, the density of complex eigenvalues is smooth but the probability density of real eigenvalues tends to zero. This corresponds to a *depletion* of the eigenvalues along the real axis. (iii) On the contrary, the density of eigenvalues in GinOE exhibits an *accumulation* of the eigenvalues along the real axis. *It is the latter phenomenon that will be quantified in our paper.*

Our immediate goal here is to highlight the inter-relation between these *qualitative features* of the complex spectra and the *formal structures* underlying Ginibre's random matrix ensembles. To this end, we present a brief comparative review of the major *structural* results obtained for all three Ginibre's ensembles (GinUE, GinSE, and GinOE) since 1965, in the order of increasing difficulty of their treatment.

Joint Probability Density Function of All n Eigenvalues. In this subsection, we collect explicit results for the joint probability density functions of all n complex eigenvalues of a random matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}_\beta(n)$ drawn from any of the three

¶ The difference in spectral patterns of non-Hermitean *chiral* random matrix models arising in the QCD context was first studied numerically by Halasz *et al* (1997). A review of recent theoretical developments can be found in Akemann (2007).

Ginibre random matrix ensembles.

- **GinUE:** The spectrum of a random matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}_2(n)$ drawn from GinUE consists of n complex eigenvalues $(z_1, \cdots z_n)$ whose joint probability density function mirrors (Ginibre 1965) that of GUE (see, e.g., Mehta 2004),

$$P_n^{(2)}(z_1, \cdots, z_n) = \left(\pi^n \prod_{\ell=1}^n \ell! \right)^{-1} \prod_{\ell_1 > \ell_2 = 1}^n |z_{\ell_1} - z_{\ell_2}|^2 \prod_{\ell=1}^n e^{-z_\ell \bar{z}_\ell}. \tag{2.2}$$

  Similarly to the electrostatic model introduced by E. Wigner (1957), J. Ginibre pointed out that the j.p.d.f. (2.2) can be thought of as that describing the distribution of the positions of charges of a two-dimensional Coulomb gas in an harmonic oscillator potential $U(z) = |z|^2/2$, at the inverse temperature $1/T = 2$.

- **GinSE:** The spectrum of a random matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}_4(n)$ drawn from GinSE consists of n pairs of complex conjugated eigenvalues $(z_1, \bar{z}_1, \cdots, z_n, \bar{z}_n)$. The corresponding joint probability density function was derived by Ginibre (1965),

$$\begin{aligned} P_n^{(4)}(z_1, \cdots, z_n) &= \left((2\pi)^n n! \prod_{\ell=1}^n (2\ell - 1)! \right)^{-1} \prod_{\ell_1 > \ell_2 = 1}^n |z_{\ell_1} - z_{\ell_2}|^2 |z_{\ell_1} - \bar{z}_{\ell_2}|^2 \\ &\times \prod_{\ell=1}^n |z_\ell - \bar{z}_\ell|^2 \exp(-z_\ell \bar{z}_\ell). \end{aligned} \tag{2.3}$$

  Notice that the factor $\prod_{\ell=1}^n |z_\ell - \bar{z}_\ell|^2$ in (2.3) is directly responsible for the *depletion* of the eigenvalues along the real axis. For GinSE, a physical analogy with a two-dimensional Coulomb gas is much less transparent; it has been discussed by P. Forrester (2005).

- **GinOE:** Contrary to GinUE and GinSE, the complex spectrum $(w_1, \cdots, w_n)$ of a random matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}_1(n)$ drawn from GinOE generically contains a *finite fraction* of real eigenvalues; the remaining complex eigenvalues always form complex conjugated *pairs*. Indeed, no other option is allowed by the *real* secular equation $\det(w - \boldsymbol{\mathcal{H}}) = 0$ determining the eigenvalues of $\boldsymbol{\mathcal{H}}$.

  This very peculiar feature of GinOE, that we call *accumulation* of the eigenvalues along the real axis, can conveniently be accommodated by dividing the entire space $\mathbb{T}_1(n)$ spanned by all real $n \times n$ matrices $\boldsymbol{\mathcal{H}} \in \mathbb{T}_1(n)$ into $(n+1)$ mutually exclusive sectors $\mathbb{T}_1(n/k)$ associated with the matrices $\boldsymbol{\mathcal{H}}_k \subset \boldsymbol{\mathcal{H}}$ having exactly k real eigenvalues, such that $\mathbb{T}_1(n) = \bigcup_{k=0}^n \mathbb{T}_1(n/k)$. The sectors $\mathbb{T}_1(n/k)$, characterised by the partial j.p.d.f.'s $P_{\boldsymbol{\mathcal{H}} \in \mathbb{T}_1(n/k)}(w_1, \cdots, w_n)$, can be explored separately because they contribute additively to the j.p.d.f. of all n eigenvalues of $\boldsymbol{\mathcal{H}}$ from $\mathbb{T}_1(n)$:

$$P_n^{(1)}(w_1, \cdots, w_n) = \sum_{k=0}^n P_{\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)}(w_1, \cdots, w_n). \tag{2.4}$$

  In entire generality, the partial j.p.d.f.'s have been determined by Lehmann and Sommers (1991) who proved, a quarter of a century after Ginibre's work, that the k-th partial j.p.d.f. ($0 \le k \le n$) equals

$$P_{\boldsymbol{\mathcal{H}}\in\mathbb{T}(n/k)}(\lambda_1,\cdots,\lambda_k;z_1,\cdots,z_\ell) = \frac{2^{\ell-n(n+1)/4}}{i^\ell\, k!\, \ell!\, \prod_{j=1}^n \Gamma(j/2)} \prod_{i>j=1}^k |\lambda_i-\lambda_j| \prod_{j=1}^k \exp(-\lambda_j^2/2) \times \prod_{j=1}^k \prod_{i=1}^\ell (\lambda_j - z_i)(\lambda_j-\bar z_i) \prod_{i>j=1}^\ell |z_i-z_j|^2 |z_i-\bar z_j|^2 \times \prod_{j=1}^\ell (z_j-\bar z_j)\,\mathrm{erfc}\left(\frac{z_j-\bar z_j}{i\sqrt{2}}\right) \exp\left(-\frac{z_j^2+\bar z_j^2}{2}\right). \tag{2.5}$$

Here, the parameterisation $(w_1,\cdots,w_n) = (\lambda_1,\cdots,\lambda_k;z_1,\bar z_1,\cdots,z_\ell,\bar z_\ell)$ was used to indicate that the spectrum is composed of k real and 2ℓ complex eigenvalues so that $k+2\ell = n$. The above j.p.d.f. is supported for $(\lambda_1,\cdots,\lambda_k)\in\mathbb{R}^k$, $(\mathrm{Re}\, z_1,\cdots,\mathrm{Re}\, z_\ell)\in\mathbb{R}^\ell$, and $(\mathrm{Im}\, z_1,\cdots,\mathrm{Im}\, z_\ell)\in(\mathbb{R}^+)^\ell$. Notice that since the complex eigenvalues come in ℓ conjugated pairs, the identity $n = 2\ell+k$ implies that half of the sets $\mathbb{T}_1(n/k)$ are empty: this happens whenever n and k are of different parity.

In writing (2.5), we have used a representation due to Edelman (1997) who rediscovered the result by Lehmann and Sommers (1991) a few years later. The particular case $k=n$ of (2.5), corresponding to the matrices $\boldsymbol{\mathcal{H}}\in\mathbb{T}(n/n)$ with all eigenvalues real, was first derived by Ginibre (1965). No physical interpretation of the distribution (2.5) in terms of a two-dimensional Coulomb gas is known as yet.

Eigenvalue Correlation Functions and Inapplicability of the Dyson Integration Theorem to the Description of GinOE. Spectral statistical properties of random matrices can be retrieved from a set of spectral correlation functions defined as

$$R_p^{(\beta)}(z_1,\cdots,z_p;n) = \frac{n!}{(n-p)!}\prod_{j=p+1}^n \int_{\mathbb{C}} d^2 z_j\, P_n^{(\beta)}(z_1,\cdots,z_n), \quad \beta=2,\ 4 \tag{2.6}$$

for GinUE ($\beta=2$) and GinSE ($\beta=4$), and

$$R_{p,q}^{(\boldsymbol{\mathcal{H}}_k)}(\lambda_1,\cdots,\lambda_p;z_1,\cdots,z_q;n) = \frac{k!}{(k-p)!}\frac{\ell!}{(\ell-q)!} \times \prod_{j=p+1}^k \int_{\mathbb{R}} d\lambda_j \prod_{m=q+1}^\ell \int_{\mathbb{C}} d^2 z_m\, P_{\boldsymbol{\mathcal{H}}\in\mathbb{T}(n/k)}(\lambda_1,\cdots,\lambda_k;z_1,\cdots,z_\ell), \tag{2.7}$$

for GinOE ($\beta=1$). The GinOE correlation function refers to the spectrum of matrices $\boldsymbol{\mathcal{H}}_k\subset\boldsymbol{\mathcal{H}}$ having exactly k real eigenvalues.

The analytic calculation of the above correlation functions is one of the major operational tasks of the non-Hermitean RMT. Whenever feasible, such a calculation either explicitly rests on or can eventually be traced back to the *three concepts*: (i) a determinant (or Pfaffian) representation of the j.p.d.f.'s of all eigenvalues, (ii) a projection property of the kernel function associated with the aforementioned determinant representation, and (iii) the Dyson integration theorem (Dyson 1970, Mahoux and

Mehta 1991) that makes use of both (i) and (ii). Highly successful in the Hermitean RMT, the above three concepts are not always at work in the *non-Hermitean* RMT. Particularly, the Dyson integration theorem, being effective for GinUE and GinSE (see below), fails to work for GinOE.

It will be argued that a mixed character of the GinOE spectrum consisting of both complex and purely real eigenvalues is the direct cause of the failure. For the readers convenience as well as for the future reference, we cite the Dyson integration theorem below [+] (Mehta (1976); see also Theorem 5.1.4 in Mehta's book (2004)).

Theorem 2.1 (Dyson integration theorem). *Let $f(x,y)$ be a function with real, complex or quaternion values, such that*

$$\bar{f}(x,y) = f(y,x), \tag{2.8a}$$

where $\bar{f} = f$ if f is real, $\bar{f}$ is the complex conjugate if it is complex, and $\bar{f}$ is the dual of f if it is quaternion. Assume that

$$\int d\pi(y)\, f(x,y)\, f(y,z) \; = f(x,z) + \lambda\, f(x,z) - f(x,z)\,\lambda, \tag{2.8b}$$

where λ is a constant quaternion and $d\pi$ is a suitable measure. Let $[f(x_i,x_j)]_{n\times n}$ denote the $n\times n$ matrix with its (i,j) element equal to $f(x_i,x_j)$. Then,

$$\int d\pi(x_n)\det[f(x_i,x_j)]_{n\times n} \; = (c-n+1)\det[f(x_i,x_j)]_{(n-1)\times(n-1)} \tag{2.8c}$$

with

$$c = \int d\pi(x)\, f(x,x). \tag{2.8d}$$

When $f(x,y)$ is real or complex, the quaternion constant λ vanishes. For $f(x,y)$ taking quaternion values, det *should be replaced by* qdet*, the quaternion determinant (Dyson 1972).*

This theorem prompts the following definition.

Definition 2.1. *A function $f(x,y)$ satisfying the first and the second equation in the Dyson integration theorem is said to obey the projection property.*

Being equipped with the above reminder, we are ready to present, and discuss, a collection of formulae available for the p-point correlation functions in Ginibre's ensembles.

- **GinUE:** The joint probability density function (2.2) of all n eigenvalues is reducible to a *determinant* form (Ginibre 1965)

$$P_n^{(2)}(z_1,\cdots,z_n) = \frac{1}{n!}\det\left[K_n^{(2)}(z_k,z_\ell)\right]_{n\times n}\prod_{j=1}^{n} e^{-z_j\bar{z}_j} \tag{2.9}$$

[+] While the formulation in Mehta (2004) refers to the flat measure $d\pi(x) = dx$, the Theorem 2.1 stays valid for any benign measure $d\pi(x)$.

with $K_n^{(2)}(z,z')$ being a two-point *scalar kernel*

$$K_n^{(2)}(z,z') = \frac{1}{\pi}\sum_{\ell=0}^{n-1}\frac{(z\bar{z}')^\ell}{\Gamma(\ell+1)}. \tag{2.10}$$

Since it obeys the projection property, the Dyson Integration Theorem brings a *determinant* expression for the p-point correlation function:

$$R_p^{(2)}(z_1,\cdots,z_p;n) = \det\left[K_n^{(2)}(z_k,z_\ell)\right]_{p\times p}\prod_{j=1}^{p} e^{-z_j\bar{z}_j}. \tag{2.11}$$

These results, first derived by J. Ginibre (1965), provide a comprehensive description of spectral fluctuations in GinUE.

- **GinSE:** The joint probability density function (2.3) of all n eigenvalues is reducible to a *quaternion determinant* form (Mehta and Srivastava 1966)

$$P_n^{(4)}(z_1,\cdots,z_n) = \frac{1}{n!}\,\mathrm{qdet}\left[K_n^{(4)}(z_k,z_\ell)\right]_{n\times n}\prod_{j=1}^{n}(\bar{z}_j - z_j)\,e^{-z_j\bar{z}_j}, \tag{2.12}$$

where $K_n^{(4)}(z,z')$ is a quaternion whose 2×2 matrix representation reads (Kanzieper 2002a)

$$\Theta[K_n^{(4)}(z,z')] = \begin{pmatrix} -\kappa_n^{(4)}(\bar{z},z') & -\kappa_n^{(4)}(\bar{z},\bar{z}') \\ \kappa_n^{(4)}(z,z') & \kappa_n^{(4)}(z,\bar{z}') \end{pmatrix}. \tag{2.13}$$

Here,

$$\kappa_n^{(4)}(z,z') = \frac{1}{2\pi}\sum_{k=0}^{n-1}\sum_{\ell=0}^{k}\frac{z^{2k+1}(z')^{2\ell} - (z')^{2k+1}z^{2\ell}}{(2k+1)!!\,(2\ell)!!}. \tag{2.14}$$

Alternatively, but equivalently, (2.12) can be reduced to the Pfaffian form (Akemann and Basile 2007)

$$P_n^{(4)}(z_1,\cdots,z_n) = \frac{1}{n!}\,\mathrm{pf}\begin{bmatrix} \kappa_n^{(4)}(z_i,z_j) & \kappa_n^{(4)}(z_i,\bar{z}_j) \\ \kappa_n^{(4)}(\bar{z}_i,z_j) & \kappa_n^{(4)}(\bar{z}_i,\bar{z}_j) \end{bmatrix}_{2n\times 2n} \times \prod_{j=1}^{n}(\bar{z}_j - z_j)\,e^{-z_j\bar{z}_j} \tag{2.15}$$

which is instructive to compare with (1.4).

As soon as the quaternion kernel $K_n^{(4)}(z,z')$ satisfies the projection property, the p-point correlation functions take a *quaternion determinant/Pfaffian* form:

$$R_p^{(4)}(z_1,\cdots,z_p;n) = \mathrm{qdet}\left[K_n^{(4)}(z_k,z_\ell)\right]_{p\times p}\prod_{j=1}^{p}(\bar{z}_j - z_j)\,e^{-z_j\bar{z}_j}. \tag{2.16}$$

This result is due to Mehta and Srivastava (1966).

- **GinOE:** To the best of our knowledge, structural aspects of correlation functions in GinOE have never been addressed (see, however, a recent paper by Sinclair (2006)); consequently, no analogues of the above GinUE and GinSE formulae

[(2.11) and (2.16)] are available. This gap will partially be filled in the present paper where we derive a *quaternion determinant/Pfaffian* expression (1.4) for the j.p.d.f. of *all complex eigenvalues* of a random matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}_1(n/k)$. Importantly, the 2×2 kernel therein does *not* possess the projection property, hereby making the Dyson integration theorem inapplicable for the calculation of associated correlation functions. We reiterate that a mixed character of the GinOE spectrum, composed of both complex and purely real eigenvalues is behind the statement made *.

Quantifying the Qualitative Differences Between Spectra of GinUE, GinSE, and GinOE. The mean density of eigenvalues $R_1^{(\beta)}(z = x + iy; n)$ is the simplest spectral statistics exemplifying differences in spectral patterns of GinUE, GinSE, and

GinOE. Below we collect, and comment on, the exact and the large-n results for the mean density of eigenvalues of $\boldsymbol{\mathcal{H}} \in \mathbb{T}_\beta(n)$.

- **GinUE:** In accordance with (2.11) and (2.10), the exact result for the mean spectral density reads (Ginibre 1965)

$$R_1^{(2)}(z; n) = \frac{\Gamma(n, x^2 + y^2)}{\pi\,\Gamma(n)}. \tag{2.17}$$

Here, $\Gamma(a, \phi)$ is the upper incomplete gamma function

$$\Gamma(a, \phi) = \int_\phi^\infty dt\, t^{a-1} e^{-t}.$$

For a large-n GinUE matrix and $z = x + iy$ fixed ♯, the mean spectral density approaches the constant

$$R_1^{(2)}(z; n \gg 1) \simeq \frac{1}{\pi}. \tag{2.18}$$

This suggests that that the eigenvalues of a large-n GinUE matrix are distributed almost *uniformly* within the two-dimensional disk of the radius $\sqrt{n}$. More rigorously, this statement follows from the *macroscopic* limit of (2.17)

$$\lim_{n\to\infty} R_1^{(2)}(z = \hat{z}\sqrt{n}; n) = \frac{1}{\pi} \left\{ \begin{array}{ll} 1 & \text{if } |\hat{z}| < 1 \\ 0 & \text{if } |\hat{z}| > 1 \end{array} \right. \tag{2.19}$$

known as the Girko circular law (Girko 1984, Girko 1986, Bai 1997). The density of eigenvalues away from the disk is exponentially suppressed (Kanzieper 2005).

- **GinSE:** In accordance with (2.13) – (2.16), the exact result for the mean spectral density reads (Mehta and Srivastava 1966, Kanzieper 2002a)

$$R_1^{(4)}(z; n) = 2y\, e^{-(x^2+y^2)}\, \mathrm{Im}\, \kappa_n^{(4)}(x + iy, x - iy) \tag{2.20}$$

* Intriguingly, it will be shown in Section 8 that the matrices $\mathcal{H} \in \mathbb{T}_1(n/0)$ exhibit GinSE-like correlations; this contrasts the well known correlations of the GOE type (Ginibre 1965) for the matrices $\mathcal{H} \in \mathbb{T}_1(n/n)$.

♯ That is, z does not scale with $n \gg 1$.

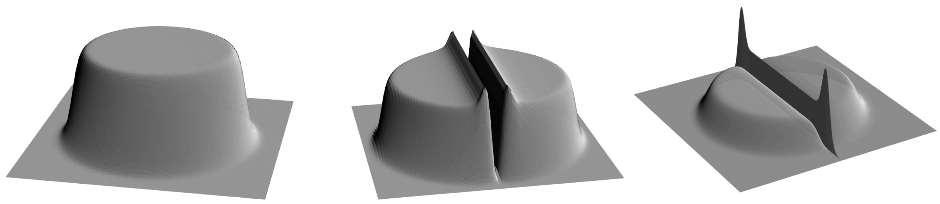

Figure 2. Profiles of eigenlevel densities $R_1^{(\beta)}(z;n)$, plotted as functions of the complex energy $z = x+iy$ for n fixed, show a nearly uniform eigenlevel distribution in GinUE (left panel), a depletion of eigenvalues along the real axis in GinSE as exemplified by the density drop at $y = 0$ (middle panel), and accumulation of real eigenvalues in GinOE displayed as a wall at $y = 0$ (right panel); the wall height imitates the density of real eigenvalues.

with $\kappa_n^{(4)}$ given by (2.14). For a large-n GinSE matrix and $z = x + iy$ fixed, it reduces to (Kanzieper, 2002a)

$$R_1^{(4)}(z; n \gg 1) \simeq \frac{y}{\sqrt{2\pi}}\, e^{-2y^2}\, \mathrm{erfi}\big(y\sqrt{2}\big). \tag{2.21}$$

Here, erfi (ϕ) is the imaginary error function

$$\mathrm{erfi}(\phi) = \frac{2}{\sqrt{\pi}} \int_0^\phi dt\, e^{t^2}.$$

Both results [(2.20) and (2.21)] suggest that the mean eigenvalue density is no longer uniform but exhibits a *depletion* of eigenvalues along the real axis. Similarly to the GinUE, the mean spectral density in GinSE is suppressed away from a disk of the radius $\sqrt{2n}$ as suggested by the circular law

$$\lim_{n\to\infty} R_1^{(4)}(\hat{z}\sqrt{2n}; n) = \frac{1}{2\pi} \left\{ \begin{array}{ll} 1 & \text{if } |\hat{z}| < 1 \\ 0 & \text{if } |\hat{z}| > 1 \end{array} \right. \tag{2.22}$$

due to Khoruzhenko and Mezzadri (2005).

- **GinOE:** A mixed character of the GinOE spectrum consisting of both complex and real eigenvalues makes the RMT techniques based on the Dyson integration theorem inapplicable to the description of spectral statistical properties of GinOE. To evaluate the mean eigenvalue density in the finite-n GinOE, a totally different approach has been invented by A. Edelman and co-workers (Edelman *et al* 1994, Edelman 1997). Starting directly with the definition (2.1) taken at $\beta = 1$ and applying the methods of multivariate statistical analysis (Muirhead 1982), these authors have separately determined the exact mean densities † of purely *real* eigenvalues (Edelman *et al* 1994)

† The two are related to the (p,q) correlation functions $R_{p,q}^{(\boldsymbol{\mathcal{H}}_k)}$ for the matrices $\boldsymbol{\mathcal{H}}_k$ restricted to have exactly k real eigenvalues [see the definition (2.7)] as follows:

$$R^{(1)}_{1,\mathrm{real}}(z;n) = \delta(y) \sum_{k=1}^{n} R_{1,0}^{(\boldsymbol{\mathcal{H}}_k)}(x;n), \quad R^{(1)}_{1,\mathrm{complex}}(z;n) = \sum_{\ell=1}^{\lfloor n/2 \rfloor} R_{0,1}^{(\boldsymbol{\mathcal{H}}_{n-2\ell})}(z;n).$$

$$R^{(1)}_{1,\text{real}}(z;n) = \frac{\delta(y)}{\sqrt{2\pi}}\left[\frac{\Gamma(n-1,x^2)}{\Gamma(n-1)} + \frac{2^{n/2-3/2}}{\Gamma(n-1)}\,|x|^{n-1}\,e^{-x^2/2}\,\gamma\left(\frac{n-1}{2},\frac{x^2}{2}\right)\right] \tag{2.23}$$

and of strictly *nonreal* ‡ eigenvalues (Edelman 1997):

$$R^{(1)}_{1,\text{complex}}(z;n) = \sqrt{\frac{2}{\pi}}\,\frac{\Gamma(n-1,x^2+y^2)}{\Gamma(n-1)}\;y\,e^{2y^2}\,\text{erfc}\left(y\sqrt{2}\right). \tag{2.24}$$

Here x and y are real and imaginary parts of $z = x + iy$. The function $\gamma(a,\phi)$ in (2.23) is the lower incomplete gamma function

$$\gamma(a,\phi) = \int_0^\phi dt\, t^{a-1} e^{-t}.$$

Importantly, *no* reference was made to the j.p.d.f. (2.4) and (2.5) in deriving (2.23) and (2.24).

For a large-n GinOE matrix and $z = x + iy$ fixed, the above two formulae yield the mean density of eigenvalues in the form

$$R^{(1)}_1(z;n\gg 1) \simeq \frac{1}{\sqrt{2\pi}}\delta(y) + \sqrt{\frac{2}{\pi}}\,y\,e^{2y^2}\,\text{erfc}\left(y\sqrt{2}\right). \tag{2.25}$$

Similarly to GinUE and GinSE, the circular law (Girko 1984, Sommers *et al* 1988, Bai 1997, Edelman 1997)

$$\lim_{n\to\infty} R^{(1)}_1(\hat{z}\sqrt{n};n) = \frac{1}{\pi}\begin{cases} 1 & \text{if } |\hat{z}|<1 \\ 0 & \text{if } |\hat{z}|>1 \end{cases} \tag{2.26}$$

holds.

2.4. Statistical description of the eigenvalue accumulation in GinOE

In essence, the approach culminating in the explicit formula (2.23) for the mean density of real eigenvalues represents the simplest possible quantitative description of the phenomenon of eigenvalue *accumulation* along the real axis. Complementarily, the one might be interested in the full statistics of the number $\mathcal{N}_r$ ($\mathcal{N}_c$) of real (complex) eigenvalues occurring in the GinOE spectrum. In the latter context, the result (2.23) can only supply the first moment of $\mathcal{N}_r$ – the expected number $E_n = \mathbb{E}[\mathcal{N}_r]$ of real eigenvalues. Indeed, integrating out the eigenlevel densities (2.23) and/or (2.24) over the entire complex plane,

$$E_n = \int_{\mathbb{C}} d^2z\, R^{(1)}_{1,\,\text{real}}(z;n) = n - \int_{\mathbb{C}\setminus\mathbb{R}} d^2z\, R^{(1)}_{1,\,\text{complex}}(z;n),$$

Edelman *et al* (1994) have obtained the remarkable result

$$E_n = \frac{1}{2} + \sqrt{2}\,\frac{{}_2F_1\left(1,-1/2;n;1/2\right)}{B(n,1/2)} \tag{2.27}$$

expressed in terms of the Gauss hypergeometric function and the Euler Beta function. As $n\to\infty$, it furnishes the asymptotic series

$$E_n = \sqrt{\frac{2n}{\pi}}\left(1 - \frac{3}{8n} - \frac{3}{128n^2} + \frac{27}{1024n^3} + \frac{499}{32768n^4} + \mathcal{O}(n^{-5})\right) + \frac{1}{2}. \tag{2.28}$$

‡ In the formulae, we use the subscript “complex” to identify eigenvalues with zero real part.

The nonperturbative $\sqrt{n}$-dependence of leading term in (2.28) had been earlier detected numerically by Sommers *et al* (1988).

What is about a more detailed statistical description of the number $\mathcal{N}_r$ ($\mathcal{N}_c$) of real (complex) eigenvalues? Can all the moments and the entire probability function of the discrete random variable $\mathcal{N}_r$ ($\mathcal{N}_c$) be determined? The preceding discussion, particularly highlighting a patchy knowledge of the spectral *correlations* in GinOE, suggests that both spectral characteristics, which can be expressed as §

$$\mathbb{E}[\mathcal{N}_r^p] = \sum_{k=1}^{p} S(p,k) \int_{\mathbb{C}} d^2 z_1 \cdots \int_{\mathbb{C}} d^2 z_k \, R^{(1)}_{k,\,\mathrm{real}}(z_1, \cdots, z_k; n) \tag{2.29}$$

and

$$\mathbb{E}[\mathcal{N}_c^p] = \sum_{\ell=1}^{p} S(p,\ell) \int_{\mathbb{C}\setminus\mathbb{R}} d^2 z_1 \cdots \int_{\mathbb{C}\setminus\mathbb{R}} d^2 z_\ell \, R^{(1)}_{\ell,\,\mathrm{complex}}(z_1, \cdots, z_\ell; n), \tag{2.30}$$

are not within immediate reach. First, only one correlation function out of p involved in either (2.29) or (2.30) is explicitly known. Second, even if all the multipoint correlation functions were readily available, the integrations in (2.29) and (2.30) would not be trivial either because of the anticipated inapplicability of the Dyson integration theorem as discussed in Section 2.3.

Some of the difficulties outlined here will be overcome in the present paper. For a list of our major results the reader is referred back to the Section 1.2. Proofs and derivations are given in the Sections to follow.

3. Statistics of real eigenvalues in GinOE spectra: A Pfaffian integral representation for the probability $p_{n,k}$

3.1. Generalities and known results

Instead of targeting the moments $\mathbb{E}[\mathcal{N}_r^p]$ of a fluctuating number of real eigenvalues as discussed in the previous section, we are going to directly determine the entire probability function $p_{n,k} = \mathrm{Prob}(\mathcal{N}_r = k)$. The definition of the $p_{n,k}$, describing the probability of finding exactly k real eigenvalues in the spectrum of an $n \times n$ real Gaussian random matrix, can be deduced from (2.4) and (2.5),

$$p_{n,k} = \mathrm{Prob}(\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)) = \prod_{i=1}^{k} \int_{\mathbb{R}} d\lambda_i \prod_{j=1}^{\ell} \int_{\mathrm{Im}\, z_j > 0} d^2 z_j \, P_{\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)}. \tag{3.1}$$

Here, ℓ is the number of pairs of complex conjugated eigenvalues in the spectrum of an $n \times n$ matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)$ having exactly k real eigenvalues, and $P_{\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)}$ is the j.p.d.f. of all n eigenvalues of such a matrix. Obviously, the identity $n = k + 2\ell$ holds.

Previous attempts to determine the probability function $p_{n,k}$ based on (3.1) brought no explicit formula for $p_{n,k}$ for generic k. The only analytic results available are due to Edelman (1997) who proved the following properties of the above probability function ‖:

Property 1. The probability of having all n eigenvalues real equals

$$p_{n,n} = 2^{-n(n-1)/4}. \tag{3.2}$$

§ The coefficient $S(p,k)$ is the Stirling number of the second kind.

‖ Table 2 provides a useful illustration of both properties.

Property 2. For all $0 \le k \le n$, the $p_{n,k}$ are of the form

$$p_{n,k} = r_{n,k} + s_{n,k}\sqrt{2}, \tag{3.3}$$

where $r_{n,k}$ and $s_{n,k}$ are rational.

The first result is a simple consequence of (3.1) which, at $k = n$, reduces to a known Selberg integral. All known examples suggest that this is the smallest probability out of all $p_{n,k}$'s. The second result is based on more involved considerations of (3.1) and (2.4).

3.2. Joint probability density function of complex eigenvalues of $\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)$

To evaluate the sought probability function $p_{n,k}$, we start with the definition (3.1). In a first step, we carry out the λ-integrations therein to assess the j.p.d.f. of all complex eigenvalues of a random matrix $\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)$ having exactly k real eigenvalues:

$$P_{\boldsymbol{\mathcal{H}}_k}(z_1, \cdots, z_\ell) = \prod_{i=1}^{k} \int_{\mathbb{R}} d\lambda_i \, P_{\boldsymbol{\mathcal{H}}\in\mathbb{T}(n/k)}(\lambda_1, \cdots, \lambda_k; z_1, \cdots, z_\ell). \tag{3.4}$$

To proceed, we consult (2.5) to spot that a part of it,

$$\mathcal{I}_{k,\ell}(\{z,\bar{z}\}) \equiv \prod_{i=1}^{k} \int_{\mathbb{R}} d\lambda_i \prod_{j=1}^{k} e^{-\lambda_j^2/2} \prod_{i>j}^{k} |\lambda_i - \lambda_j| \prod_{j=1}^{k}\prod_{i=1}^{\ell} (\lambda_j - z_i)(\lambda_j - \bar{z}_i), \tag{3.5}$$

coincides up to a prefactor to be specified below with the average characteristic polynomial ¶

$$\mathcal{P}_{k,\ell}(\{z,\bar{z}\}) = \left(\int D\boldsymbol{\mathcal{O}}\, e^{-\mathrm{tr}\,\boldsymbol{\mathcal{O}}^2/2} \right)^{-1} \times \int D\boldsymbol{\mathcal{O}} \prod_{j=1}^{\ell} \det(z_j - \boldsymbol{\mathcal{O}}) \det(\bar{z}_j - \boldsymbol{\mathcal{O}})\, e^{-\mathrm{tr}\,\boldsymbol{\mathcal{O}}^2/2} \tag{3.6}$$

of a $k \times k$ real symmetric matrix $\boldsymbol{\mathcal{O}} = \boldsymbol{\mathcal{O}}^T$ drawn from the GOE. More precisely,

$$\mathcal{I}_{k,\ell}(\{z,\bar{z}\}) = \mathfrak{s}_k\, k!\, \mathcal{P}_{k,\ell}(\{z,\bar{z}\}),$$

where $\mathfrak{s}_k$ is given by the Selberg integral (Mehta 2004)

$$\mathfrak{s}_k = \frac{1}{k!} \prod_{i=1}^{k} \int_{\mathbb{R}} d\lambda_i \prod_{j=1}^{k} e^{-\lambda_j^2/2} \prod_{i>j}^{k} |\lambda_i - \lambda_j| = 2^{k/2} \prod_{j=1}^{k} \Gamma(j/2). \tag{3.7}$$

Consequently, the j.p.d.f. of all complex eigenvalues of $\boldsymbol{\mathcal{H}} \in \mathbb{T}(n/k)$ is expressed in terms of the average GOE characteristic polynomial $\mathcal{P}_{k,\ell}(\{z,\bar{z}\})$ as

$$P_{\boldsymbol{\mathcal{H}}_k}(z_1, \cdots, z_\ell) = \frac{2^\ell}{i^\ell \ell!} \frac{\mathfrak{s}_k}{\mathfrak{s}_n} p_{n,n}\, \mathcal{P}_{k,\ell}(\{z,\bar{z}\}) \prod_{i>j=1}^{\ell} |z_i - z_j|^2 |z_i - \bar{z}_j|^2 \times \prod_{j=1}^{\ell} (z_j - \bar{z}_j)\, \mathrm{erfc}\left(\frac{z_j - \bar{z}_j}{i\sqrt{2}} \right) \exp\left(-\frac{z_j^2 + \bar{z}_j^2}{2} \right). \tag{3.8}$$

¶ Equations (3.4) and (3.5) suggest that $\mathcal{P}_{0,n} = 1$.

A little bit more spadework is needed to appreciate the beauty hidden in this representation. Borrowing the result due to Borodin and Strahov (2005), who discovered a Pfaffian formula for a general averaged GOE characteristic polynomial (see also an earlier paper by Nagao and Nishigaki (2001)), we may write $\mathcal{P}_{k,\ell}$ in the form

$$\mathcal{P}_{k,\ell}(\{z,\bar{z}\}) = \frac{\mathfrak{s}_n/\mathfrak{s}_k}{\Delta_{2\ell}(\{z,\bar{z}\})}\,(-1)^\ell \times \prod_{j=1}^{\ell} \exp\left(\frac{z_j^2+\bar{z}_j^2}{2}\right) \mathrm{pf}\left[\begin{array}{cc} \mathcal{D}_n(z_i,z_j) & \mathcal{D}_n(z_i,\bar{z}_j) \\ \mathcal{D}_n(\bar{z}_i,z_j) & \mathcal{D}_n(\bar{z}_i,\bar{z}_j) \end{array}\right]_{2\ell\times 2\ell}. \qquad (3.9)$$

Here, $\mathcal{D}_n(z_i,z_j)$ is the so-called D-part of the 2×2 GOE matrix kernel (Tracy and Widom 1998) to be defined in (3.12) and (3.13) below; $\Delta_{2\ell}(\{z,\bar{z}\})$ is the Vandermonde determinant [+]

$$\Delta_{2\ell}(\{z,\bar{z}\}) = \Delta_{2\ell}(z_1,\bar{z}_1,\cdots,z_\ell,\bar{z}_\ell) = \prod_{i>j}^{\ell}|z_i-z_j|^2 \prod_{i>j}^{\ell}|z_i-\bar{z}_j|^2 \prod_{i=1}^{\ell}(\bar{z}_i-z_i). \qquad (3.10)$$

Combining (3.8), (3.9) and (3.10), we obtain:

$$P_{\boldsymbol{\mathcal{H}}_k}(z_1\cdots,z_\ell) = \frac{p_{n,n}}{\ell!}\left(\frac{2}{i}\right)^\ell \times \prod_{j=1}^{\ell}\mathrm{erfc}\left(\frac{z_j-\bar{z}_j}{i\sqrt{2}}\right) \mathrm{pf}\left[\begin{array}{cc} \mathcal{D}_n(z_i,z_j) & \mathcal{D}_n(z_i,\bar{z}_j) \\ \mathcal{D}_n(\bar{z}_i,z_j) & \mathcal{D}_n(\bar{z}_i,\bar{z}_j) \end{array}\right]_{2\ell\times 2\ell} \qquad (3.11)$$

where $(\mathrm{Re}\, z_1,\cdots,\mathrm{Re}\, z_\ell)\in\mathbb{R}^\ell$ and $(\mathrm{Im}\, z_1,\cdots,\mathrm{Im}\, z_\ell)\in(\mathbb{R}^+)^\ell$ by derivation.

Equation (3.11) is the central result of this section. Announced in (1.4), it describes the j.p.d.f. of ℓ pairs of complex conjugated eigenvalues $\{z_j,\bar{z}_j\}$ of an $n\times n$ random matrix $\boldsymbol{\mathcal{H}}_k$ whose k remaining eigenvalues are real, $k+2\ell=n$.

To make the expression for the j.p.d.f. $P_{\boldsymbol{\mathcal{H}}_k}$ explicit, we have to specify the kernel function $\mathcal{D}_n(z_i,z_j)$. The latter turns out to be sensitive to the parity of n (see, e.g., Adler *et al* (2000)). For $n=2m$ even, the kernel function is given by

$$\mathcal{D}_{2m}(x,y) = \frac{1}{2}e^{-(x^2+y^2)/2}\sum_{j=0}^{m-1}\frac{q_{2j+1}(x)\,q_{2j}(y)-q_{2j}(x)\,q_{2j+1}(y)}{h_j} \qquad (3.12)$$

while for $n=2m+1$ odd, it equals

$$\mathcal{D}_{2m+1}(x,y) = \frac{1}{2}e^{-(x^2+y^2)/2}\sum_{j=0}^{m-1}\frac{\tilde{q}_{2j+1}(x)\,\tilde{q}_{2j}(y)-\tilde{q}_{2j}(x)\,\tilde{q}_{2j+1}(y)}{h_j}. \qquad (3.13)$$

Both representations (3.12) and (3.13) involve the polynomials $q_j(x)$ skew orthogonal on $\mathbb{R}$ with respect to the GOE skew product (Mehta 2004)

$$\langle f,\, g\rangle = \frac{1}{2}\int_{\mathbb{R}} dx\, e^{-x^2/2}\int_{\mathbb{R}} dy\, e^{-y^2/2}\mathrm{sgn}(y-x)\, f(x)\, g(y) \qquad (3.14)$$

such that

$$\langle q_{2k}, q_{2\ell+1}\rangle = -\langle q_{2k+1}, q_{2\ell}\rangle = h_k\delta_{k,\ell}, \quad \langle q_{2k}, q_{2\ell}\rangle = \langle q_{2k+1}, q_{2\ell+1}\rangle = 0. \qquad (3.15)$$

[+] Throughout the paper, we adopt the definition $\Delta_k(x)=\prod_{i>j}^k(x_i-x_j)$ which differs from that of Borodin and Strahov (2005) who use the same notation $\Delta_k(x)$ for the double product with $i<j$.

The skew orthogonal polynomials $q_j(x)$ can be expressed * in terms of Hermite polynomials as ♯

$$q_{2j}(x) = \frac{1}{2^{2j}}\, H_{2j}(x),$$
$$q_{2j+1}(x) = \frac{1}{2^{2j+1}}\left[H_{2j+1}(x) - 4j\, H_{2j-1}(x)\right] \tag{3.16}$$

while "tilded" polynomials $\tilde{q}_j(x)$ † entering (3.13) are related to $q_j(x)$ via

$$\tilde{q}_{2j}(x) = q_{2j}(x) - \frac{(2j)!}{2^{2j} j!}\,\frac{2^{2m} m!}{(2m)!}\, q_{2m}(x),$$
$$\tilde{q}_{2j+1}(x) = q_{2j+1}(x). \tag{3.17}$$

Specifying the normalisation

$$h_j = \langle q_{2j},\, q_{2j+1}\rangle = \frac{\sqrt{\pi}\,(2j)!}{2^{2j}} \tag{3.18}$$

completes our derivation of (3.11).

3.3. Probability function $p_{n,k}$ as a Pfaffian integral and inapplicability of the Dyson integration theorem for its calculation

The results obtained in Section 3.2 allow us to express the probability function $p_{n,k}$ in the form of an ℓ-fold integral

$$p_{n,k} = \frac{p_{n,n}}{\ell!}\left(\frac{2}{i}\right)^{\ell} \prod_{j=1}^{\ell} \int_{\mathrm{Im}\, z_j > 0} d^2 z_j$$
$$\times\, \mathrm{erfc}\left(\frac{z_j - \bar{z}_j}{i\sqrt{2}}\right) \mathrm{pf}\begin{bmatrix} \mathcal{D}_n(z_i, z_j) & \mathcal{D}_n(z_i, \bar{z}_j) \\ \mathcal{D}_n(\bar{z}_i, z_j) & \mathcal{D}_n(\bar{z}_i, \bar{z}_j) \end{bmatrix}_{2\ell\times 2\ell} \tag{3.19}$$

involving a Pfaffian. It can also be rewritten as a quaternion determinant

$$\mathrm{pf}\begin{bmatrix} \mathcal{D}_n(z_i, z_j) & \mathcal{D}_n(z_i, \bar{z}_j) \\ \mathcal{D}_n(\bar{z}_i, z_j) & \mathcal{D}_n(\bar{z}_i, \bar{z}_j) \end{bmatrix}_{2\ell\times 2\ell} = \mathrm{qdet}[\hat{\mathcal{D}}_n(z_i, z_j)]_{\ell\times\ell}$$

where the self-dual quaternion $\hat{\mathcal{D}}_n(z_i, z_j)$ has a 2×2 matrix representation

$$\Theta[\hat{\mathcal{D}}_n(z_i, z_j)] = \begin{pmatrix} -\mathcal{D}_n(\bar{z}_i, z_j) & -\mathcal{D}_n(\bar{z}_i, \bar{z}_j) \\ \mathcal{D}_n(z_i, z_j) & \mathcal{D}_n(z_i, \bar{z}_j) \end{pmatrix}.$$

The Pfaffian/quaternion determinant form of the integrand in (3.19), closely resembling the structure of both the j.p.d.f. (2.12) of all complex eigenvalues and the p-point correlation function (2.16) in GinSE, makes it tempting to attack the ℓ-fold integral (3.19) with the help of the Dyson integration theorem (see Section 2.3). Unfortunately, the key condition of this theorem – the projection property – is not fulfilled.

To see this point, we represent the kernel function $\mathcal{D}_n(x, y)$ in the form

$$\mathcal{D}_n(x, y) = \frac{1}{2}\, e^{-(x^2+y^2)/2} \sum_{j,k=0}^{n-1} q_j(x)\, \hat{\mu}_{jk}\, q_k(y), \tag{3.20}$$

* The representation (3.16) is not unique; see, e.g., Eynard (2001).
♯ Equation (3.16) assumes that $H_{-1}(x) \equiv 0$.
† Note that the $\tilde{q}_{2j}(x)$ is no longer a polynomial of the degree $2j$.

see the primary definition (3.12) and (3.13). In (3.20), the real antisymmetric matrix $\hat{\boldsymbol{\mu}}$ of the size $n \times n$ depends on the parity of n. It equals

$$\hat{\mu}_{jk}^{\rm even} = \begin{pmatrix} -\hat{\boldsymbol{e}}_2 h_0^{-1} & & \\ & \ddots & \\ & & -\hat{\boldsymbol{e}}_2 h_{m-1}^{-1} \end{pmatrix} \tag{3.21}$$

and

$$\hat{\mu}_{jk}^{\rm odd} = \begin{pmatrix} -\hat{\boldsymbol{e}}_2 h_0^{-1} & & & -\hat{\boldsymbol{\varsigma}}_0^{\rm T} \\ & \ddots & & \vdots \\ & & -\hat{\boldsymbol{e}}_2 h_{m-1}^{-1} & -\hat{\boldsymbol{\varsigma}}_{m-1}^{\rm T} \\ \hat{\boldsymbol{\varsigma}}_0 & \cdots & \hat{\boldsymbol{\varsigma}}_{m-1} & 0 \end{pmatrix} \tag{3.22}$$

for $n = 2m$ and $n = 2m+1$, respectively. We remind that h_j is defined by (3.18); also we have used the notation

$$\hat{\boldsymbol{e}}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \hat{\boldsymbol{\varsigma}}_j = c_m \left(0, \frac{1}{j!}\right), \quad c_m = \frac{m!}{h_m}. \tag{3.23}$$

Actually, the representation (3.20) can be put into a more general form due to Tracy and Widom (1998) that would contain arbitrary, not necessarily skew orthogonal, polynomials upon a proper redefinition of the matrix $\hat{\boldsymbol{\mu}}$.

For the Dyson integration theorem to be applicable, the projection property for the self-dual quaternion $\hat{\mathcal{D}}_n(z_i, z_j)$ must hold. For this to be the case, the integral identity

$$\int d\alpha(w)\, \left[\mathcal{D}_n(z_1, w)\, \mathcal{D}_n(\bar{w}, z_2) - \mathcal{D}_n(z_2, w)\, \mathcal{D}_n(\bar{w}, z_1)\right] \stackrel{?}{=} -\mathcal{D}_n(z_1, z_2) \tag{3.24}$$

should be satisfied for the measure

$$d\alpha(w) = {\rm erfc}\left(\frac{w - \bar{w}}{i\sqrt{2}}\right) \theta({\rm Im}\, w)\, d^2 w. \tag{3.25}$$

Here, $\theta(\phi)$ is the Heaviside step function,

$$\theta(\phi) = \begin{cases} 1 & \text{if } \phi > 0 \\ 0 & \text{if } \phi < 0 \end{cases}.$$

Straightforward calculations based on (3.20) show that the integral on the l.h.s. of (3.24) equals

$$\frac{1}{2}\, e^{-(z_1^2 + z_2^2)/2} \sum_{j,k=0}^{n-1} q_j(z_1)\, (\hat{\boldsymbol{\mu}}\, \hat{\boldsymbol{\chi}}\, \hat{\boldsymbol{\mu}})_{jk}\, q_k(z_2), \tag{3.26}$$

where $n \times n$ matrix $\hat{\boldsymbol{\chi}}$ has the entries

$$\hat{\chi}_{jk} = \frac{1}{2} \int d\alpha(w)\, e^{-(w^2 + \bar{w}^2)/2} \left[q_j(w)\, q_k(\bar{w}) - q_j(\bar{w})\, q_k(w)\right]. \tag{3.27}$$

Since ‡ $(\hat{\boldsymbol{\mu}}\, \hat{\boldsymbol{\chi}}) \neq -\hat{\mathbf{1}}_n$, the l.h.s. of (3.24) given by (3.26) differs from the r.h.s. (3.24). As a result, the kernel function $\mathcal{D}_n(z_1, z_2)$ does not satisfy the projection property §. Consequently, the Dyson integration theorem is inapplicable for the calculation of $p_{n,k}$ in the form of the Pfaffian integral (3.19).

‡ That $-(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\chi}})$ cannot be a unit matrix $\hat{\mathbf{1}}$, follows from the fact that $\hat{\chi}_{jk}$ is purely imaginary [(3.27)] while $\hat{\mu}_{jk}$ are real valued [(3.21) and (3.22)]. See Appendix B for an explicit calculation.

§ As soon as the integral on the l.h.s. of (3.24) combines a D-part of the GOE 2×2 matrix kernel originally introduced for the GOE's *real* spectrum with the GinOE-induced measure $d\alpha(w)$ supported in the *complex* half-plane ${\rm Im}\, w > 0$, a violation of the projection property is not unexpected.

4. Probability function $p_{n,k}$: Sensing a structure through particular cases

Before turning to the derivation of the general formula for $p_{n,k}$ (see the results announced in Section 1.2), it is instructive to consider a few particular cases corresponding to low values of ℓ, the number of pairs of complex conjugated eigenvalues. Below, the cases of $\ell = 1$, 2 and 3 are treated explicitly.

4.1. What is the probability to find exactly one pair of complex conjugated eigenvalues?

As a first nontrivial application of the Pfaffian integral representation (3.19) for the probability function $p_{n,k}$, let us consider the next-to-the-simplest ‖ case of $\ell = 1$ corresponding to the occurrence of exactly one pair of complex conjugated eigenvalues. Since the D-kernel (3.20) is antisymmetric under exchange of its arguments,

$$\mathcal{D}_n(x,y) = -\mathcal{D}_n(y,x), \tag{4.1}$$

the Pfaffian in (3.19) reduces to

$$\mathrm{pf}\begin{bmatrix} 0 & \mathcal{D}_n(z,\bar{z}) \\ -\mathcal{D}_n(z,\bar{z}) & 0 \end{bmatrix} = \mathcal{D}_n(z,\bar{z})$$

resulting in

$$p_{n,n-2} = p_{n,n}\,\frac{2}{i}\int_{\mathrm{Im}\,z>0} d^2z\ \mathrm{erfc}\left(\frac{z-\bar{z}}{i\sqrt{2}}\right)\mathcal{D}_n(z,\bar{z}). \tag{4.2}$$

To calculate the integral

$$\mathcal{I}_1 = \int_{\mathrm{Im}\,z>0} d^2z\ \mathrm{erfc}\left(\frac{z-\bar{z}}{i\sqrt{2}}\right)\mathcal{D}_n(z,\bar{z}) = \int d\alpha(z)\,\mathcal{D}_n(z,\bar{z})$$

(see (3.25)), we rewrite it in a more symmetric manner

$$\frac{1}{2}\int d\alpha(z)\left[\mathcal{D}_n(z,\bar{z}) - \mathcal{D}_n(\bar{z},z)\right],$$

and make use of (3.27), (3.25) and (3.20) to deduce that it equals

$$-\frac{1}{2}\sum_{j,k=0}^{n-1}\mu_{jk}\chi_{kj} = -\frac{1}{2}\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\chi}}),$$

or, equivalently,

$$\mathcal{I}_1 = \int d\alpha(z)\,\mathcal{D}_n(z,\bar{z}) = \frac{i}{4}\,\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\sigma}}. \tag{4.3}$$

Here, $\hat{\boldsymbol{\sigma}}$ is given by $\hat{\boldsymbol{\sigma}} = 2i\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\chi}}$ (see also Appendix B). We therefore conclude that the probability sought equals

$$p_{n,n-2} = \frac{1}{2}\,p_{n,n}\,\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\sigma}}. \tag{4.4}$$

Due to the trace identity (C.5) proven in Appendix C, we eventually derive:

$$p_{n,n-2} = p_{n,n}\,\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}\hat{\boldsymbol{\varrho}}, \tag{4.5}$$

reducing the size of the matrix by two. Here, the smaller matrix $\hat{\boldsymbol{\varrho}}$ depends on the parity of n, as defined in the Section 1.2 (see also (C.3) and (C.4)).

‖ In the simplest case of $\ell = 0$, our representation (3.19) reproduces the result (3.2) first derived by Edelman (1997).

Remarkably, the trace in (4.5) can explicitly be calculated (see Appendix D) to yield a closed expression for the probability to find exactly one pair of complex conjugated eigenvalues:

$$p_{n,n-2} = 2\, p_{n,n} \int_0^\infty dy\, y\, e^{y^2}\, \mathrm{erfc}(y\sqrt{2})\, L_{n-2}^2(-2y^2). \tag{4.6}$$

Yet another, though equivalent, representation for the probability $p_{n,n-2}$ is given in Section 7 that addresses the large-n behaviour of $p_{n,n-2}$.

4.2. Two pairs of complex conjugated eigenvalues ($\ell = 2$)

For $\ell = 2$, the Pfaffian in (3.19)

$$\mathrm{pf}\begin{bmatrix} 0 & \mathcal{D}_n(z_1,\bar{z}_1) & \mathcal{D}_n(z_1,z_2) & \mathcal{D}_n(z_1,\bar{z}_2) \\ -\mathcal{D}_n(z_1,\bar{z}_1) & 0 & \mathcal{D}_n(\bar{z}_1,z_2) & \mathcal{D}_n(\bar{z}_1,\bar{z}_2) \\ -\mathcal{D}_n(z_1,z_2) & -\mathcal{D}_n(\bar{z}_1,z_2) & 0 & \mathcal{D}_n(z_2,\bar{z}_2) \\ -\mathcal{D}_n(z_1,\bar{z}_2) & -\mathcal{D}_n(\bar{z}_1,\bar{z}_2) & -\mathcal{D}_n(z_2,\bar{z}_2) & 0 \end{bmatrix} \tag{4.7}$$

reduces to

$$\mathcal{D}_n(z_1,\bar{z}_1)\,\mathcal{D}_n(z_2,\bar{z}_2) + \mathcal{D}_n(z_1,\bar{z}_2)\,\mathcal{D}_n(\bar{z}_1,z_2) - \mathcal{D}_n(z_1,z_2)\,\mathcal{D}_n(\bar{z}_1,\bar{z}_2), \tag{4.8}$$

so that

$$\begin{aligned} p_{n,n-4} = -2\, p_{n,n} &\int d\alpha(z_1) \int d\alpha(z_2) \\ &\times \Big[\mathcal{D}_n(z_1,\bar{z}_1)\,\mathcal{D}_n(z_2,\bar{z}_2) + \mathcal{D}_n(z_1,\bar{z}_2)\,\mathcal{D}_n(\bar{z}_1,z_2) - \mathcal{D}_n(z_1,z_2)\,\mathcal{D}_n(\bar{z}_1,\bar{z}_2)\Big]. \end{aligned} \tag{4.9}$$

Apart from a known integral taking the form of (4.3), a new integral

$$\mathcal{I}_2 = \int d\alpha(z_1) \int d\alpha(z_2) \Big[\mathcal{D}_n(z_1,\bar{z}_2)\,\mathcal{D}_n(\bar{z}_1,z_2) - \mathcal{D}_n(z_1,z_2)\,\mathcal{D}_n(\bar{z}_1,\bar{z}_2)\Big] \tag{4.10}$$

appears in (4.9). Somewhat lengthy but straightforward calculations based on (3.27), (3.25) and (3.20) result in

$$\mathcal{I}_2 = \frac{1}{8}\, \mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\sigma}}^2). \tag{4.11}$$

Combining (4.9), (4.3), (4.10) and (4.11), we obtain:

$$p_{n,n-4} = p_{n,n} \left[\frac{1}{8}\, \left(\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\sigma}}\right)^2 - \frac{1}{4}\, \mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\sigma}}^2)\right]. \tag{4.12}$$

Due to the trace identity (C.5) proven in Appendix C, the latter reduces to

$$p_{n,n-4} = \frac{1}{2}\, p_{n,n} \left[\left(\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}\hat{\boldsymbol{\varrho}}\right)^2 - \mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}(\hat{\boldsymbol{\varrho}}^2)\right]. \tag{4.13}$$

Interestingly, the expression in the parenthesis of (4.13) coincides with $Z_{(1^2)}(p_1,p_2)$ after the substitution $\mathrm{tr}\,\hat{\boldsymbol{\varrho}}^j = p_j$ (see Table 1).

4.3. Three pairs of complex conjugated eigenvalues ($\ell = 3$)

The complexity of the integrand in (3.19) grows rapidly with increasing ℓ. For $\ell = 3$, that is three pairs of complex conjugated eigenvalues in the matrix spectrum, the

Pfaffian of the 6×6 antisymmetric matrix

$$\begin{pmatrix} 0 & \mathcal{D}_n(z_1,\bar{z}_1) & \mathcal{D}_n(z_1,z_2) & \mathcal{D}_n(z_1,\bar{z}_2) & \mathcal{D}_n(z_1,z_3) & \mathcal{D}_n(z_1,\bar{z}_3) \\ -\mathcal{D}_n(z_1,\bar{z}_1) & 0 & \mathcal{D}_n(\bar{z}_1,z_2) & \mathcal{D}_n(\bar{z}_1,\bar{z}_2) & \mathcal{D}_n(\bar{z}_1,z_3) & \mathcal{D}_n(\bar{z}_1,\bar{z}_3) \\ -\mathcal{D}_n(z_1,z_2) & -\mathcal{D}_n(\bar{z}_1,z_2) & 0 & \mathcal{D}_n(z_2,\bar{z}_2) & \mathcal{D}_n(z_2,z_3) & \mathcal{D}_n(z_2,\bar{z}_3) \\ -\mathcal{D}_n(z_1,\bar{z}_2) & -\mathcal{D}_n(\bar{z}_1,\bar{z}_2) & -\mathcal{D}_n(z_2,\bar{z}_2) & 0 & \mathcal{D}_n(\bar{z}_2,z_3) & \mathcal{D}_n(\bar{z}_2,\bar{z}_3) \\ -\mathcal{D}_n(z_1,z_3) & -\mathcal{D}_n(\bar{z}_1,z_3) & -\mathcal{D}_n(z_2,z_3) & -\mathcal{D}_n(\bar{z}_2,z_3) & 0 & \mathcal{D}_n(z_3,\bar{z}_3) \\ -\mathcal{D}_n(z_1,\bar{z}_3) & -\mathcal{D}_n(\bar{z}_1,\bar{z}_3) & -\mathcal{D}_n(z_2,\bar{z}_3) & -\mathcal{D}_n(\bar{z}_2,\bar{z}_3) & -\mathcal{D}_n(z_3,\bar{z}_3) & 0 \end{pmatrix}$$

is getting involved. It can be calculated with some effort to give 15 terms which can be attributed to three different groups. The first group consists of the single term

$$G_1 = \mathcal{D}_n(z_1,\bar{z}_1)\,\mathcal{D}_n(z_2,\bar{z}_2)\,\mathcal{D}_n(z_3,\bar{z}_3). \tag{4.14}$$

The second group contains 6 terms,

$$\begin{aligned} G_2 &= \mathcal{D}_n(z_1,\bar{z}_1)\Big[\mathcal{D}_n(z_2,\bar{z}_3)\,\mathcal{D}_n(\bar{z}_2,z_3) - \mathcal{D}_n(z_2,z_3)\,\mathcal{D}_n(\bar{z}_2,\bar{z}_3)\Big] \\ &\quad + \mathcal{D}_n(z_2,\bar{z}_2)\Big[\mathcal{D}_n(z_1,\bar{z}_3)\,\mathcal{D}_n(\bar{z}_1,z_3) - \mathcal{D}_n(z_1,z_3)\,\mathcal{D}_n(\bar{z}_1,\bar{z}_3)\Big] \\ &\quad + \mathcal{D}_n(z_3,\bar{z}_3)\Big[\mathcal{D}_n(z_1,\bar{z}_2)\,\mathcal{D}_n(\bar{z}_1,z_2) - \mathcal{D}_n(z_1,z_2)\,\mathcal{D}_n(\bar{z}_1,\bar{z}_2)\Big], \end{aligned} \tag{4.15}$$

while the third group counts 8 terms:

$$\begin{aligned} G_3 &= \mathcal{D}_n(z_1,\bar{z}_3)\,\mathcal{D}_n(\bar{z}_2,z_3)\,\mathcal{D}_n(\bar{z}_1,z_2) - \mathcal{D}_n(\bar{z}_1,z_3)\,\mathcal{D}_n(z_2,\bar{z}_3)\,\mathcal{D}_n(z_1,\bar{z}_2) \\ &\quad + \mathcal{D}_n(z_1,\bar{z}_2)\,\mathcal{D}_n(z_2,z_3)\,\mathcal{D}_n(\bar{z}_1,\bar{z}_3) - \mathcal{D}_n(\bar{z}_1,z_2)\,\mathcal{D}_n(\bar{z}_2,\bar{z}_3)\,\mathcal{D}_n(z_1,z_3) \\ &\quad + \mathcal{D}_n(\bar{z}_1,\bar{z}_2)\,\mathcal{D}_n(z_2,\bar{z}_3)\,\mathcal{D}_n(z_1,z_3) - \mathcal{D}_n(z_1,z_2)\,\mathcal{D}_n(\bar{z}_2,z_3)\,\mathcal{D}_n(\bar{z}_1,\bar{z}_3) \\ &\quad + \mathcal{D}_n(z_1,z_2)\,\mathcal{D}_n(\bar{z}_2,\bar{z}_3)\,\mathcal{D}_n(\bar{z}_1,z_3) - \mathcal{D}_n(\bar{z}_1,\bar{z}_2)\,\mathcal{D}_n(z_2,z_3)\,\mathcal{D}_n(z_1,\bar{z}_3). \end{aligned} \tag{4.16}$$

In the above notation, the probability function $p_{n,n-6}$ takes the form

$$p_{n,n-6} = \frac{4i}{3}\,p_{n,n}\int d\alpha(z_1)\int d\alpha(z_2)\int d\alpha(z_3)\Big[G_1 + G_2 + G_3\Big]. \tag{4.17}$$

The integrals containing G_1 and G_2 can easily be performed with the help of (4.3), (4.10) and (4.11) to bring

$$\int d\alpha(z_1)\int d\alpha(z_2)\int d\alpha(z_3)\,G_1 = \mathcal{I}_1^3 = \left(\frac{i}{4}\,\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\sigma}}\right)^3 \tag{4.18}$$

and

$$\begin{aligned} \int d\alpha(z_1)\int d\alpha(z_2)\int d\alpha(z_3)\,G_2 &= 3\,\mathcal{I}_1\,\mathcal{I}_2 \\ &= 3\left(\frac{i}{4}\,\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\sigma}}\right)\left(\frac{1}{8}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\sigma}}^2)\right). \end{aligned} \tag{4.19}$$

The remaining integral involving G_3 can be evaluated similarly to $\mathcal{I}_1$ and $\mathcal{I}_2$, the result being

$$\mathcal{I}_3 = \int d\alpha(z_1)\int d\alpha(z_2)\int d\alpha(z_3)\,G_3 = \frac{i^3}{8}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\sigma}}^3). \tag{4.20}$$

Combining (4.17), (4.18), (4.19) and (4.20), we derive:

$$p_{n,n-6} = p_{n,n}\left[\frac{1}{48}\left(\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\sigma}}\right)^3 - \frac{1}{8}\,\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\sigma}}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\sigma}}^2) + \frac{1}{6}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\sigma}}^3)\right]. \tag{4.21}$$

Finally, we apply the trace identity (C.5) to end up with the formula

$$p_{n,n-6} = \frac{1}{6}\, p_{n,n} \Big[\big(\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}\hat{\boldsymbol{\varrho}}\big)^3 - 3\,\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}\hat{\boldsymbol{\varrho}}\,\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}(\hat{\boldsymbol{\varrho}}^2) + 2\,\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}(\hat{\boldsymbol{\varrho}}^3)\Big]. \tag{4.22}$$

The expression in the parenthesis of (4.22) is seen to coincide with $Z_{(1^3)}(p_1, p_2, p_3)$ after the substitution $\mathrm{tr}\, \hat{\boldsymbol{\varrho}}^j = p_j$ (see Table 1).

4.4. Higher ℓ

The three examples considered clearly demonstrate that the calculational complexity grows enormously with increasing ℓ, the number of complex conjugated eigenvalues in the random matrix spectrum. Indeed, the number of terms $\mathcal{N}_\ell$ in the expansion of the Pfaffian in (3.19) equals $\mathcal{N}_\ell = (2\ell-1)!!$ and exhibits a faster-than-exponential growth, $\mathcal{N}_\ell \approx 2^{\ell+1/2}\, e^{\ell\,(\ln \ell - 1)}$, for $\ell \gg 1$. For this reason, one has to invent a classification of the terms arising in the Pfaffian expansion that facilitates their effective computation. This will be done in Section 5, where we introduce a topological interpretation of the Pfaffian expansion, and prove the Pfaffian integration theorem which can be viewed as a generalisation of the Dyson integration theorem (see Section 2.3).

5. Topological interpretation of the Pfaffian expansion, and the Pfaffian integration theorem

5.1. Statement of the main result and its discussion

Before stating the main result of this section, the Pfaffian integration theorem, we wish to start with presenting a simple Corollary to the Dyson integration theorem.

Corollary 5.1. *Let $f(x,y)$ be a function with real, complex or quaternion values satisfying the conditions (2.8a) and (2.8b) of the Theorem 2.1, and $d\pi$ be a suitable measure. Then*

$$\int \prod_{j=1}^{\ell} d\pi(x_j)\, \det\left[f(x_i,x_j)\right]_{\ell\times\ell} = \frac{\Gamma(c+1)}{\Gamma(c+1-\ell)}, \tag{5.1a}$$

where

$$c = \int d\pi(x)\, f(x,x). \tag{5.1b}$$

For f taking quaternion values, the det *should be interpreted as* qdet*, the quaternion determinant (Dyson 1972).*

Proof. Repeatedly apply the Dyson integration theorem to the l.h.s. of (5.1a) to arrive at its r.h.s. ■

Importantly, the above Corollary exclusively applies to functions $f(x,y)$ satisfying the projection property as defined in Section 2.3 (see Definition 2.1 therein). However, guided by our study of the integrable structure of GinOE, we are going to ask if the integrals of the kind (5.1a) can explicitly be calculated if the projection property is relaxed. In general, the answer is positive. In particular, for $f(x,y)$ being a self-dual quaternion, the following integration theorem will be proven.

Theorem 5.1 (Pfaffian integration theorem). *Let $d\pi(z)$ be any benign measure on $z \in \mathbb{C}$, and the function $Q_n(x,y)$ be an antisymmetric function of the form*

$$Q_n(x,y) = \frac{1}{2}\sum_{j,k=0}^{n-1} q_j(x)\,\hat{\mu}_{jk}\,q_k(y) \tag{5.2a}$$

where the $q_j(x)$ are arbitrary polynomials of j-th order, and $\hat{\boldsymbol{\mu}}$ is an antisymmetric matrix. Then the integration formula

$$\int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j)\ \mathrm{pf}\left[\begin{array}{cc} Q_n(z_i,z_j) & Q_n(z_i,\bar{z}_j) \\ Q_n(\bar{z}_i,z_j) & Q_n(\bar{z}_i,\bar{z}_j) \end{array}\right]_{2\ell\times 2\ell}$$

$$= \left(\frac{i}{2}\right)^{\ell} Z_{(1^\ell)}\left(\frac{1}{2}\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\upsilon}}^1, \cdots, \frac{1}{2}\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\upsilon}}^\ell\right) \tag{5.2b}$$

holds, provided the integrals in its l.h.s. exist. Here, $Z_{(1^\ell)}$ are zonal polynomials whose ℓ arguments are determined by a matrix $\hat{\boldsymbol{\upsilon}}$ with the entries

$$\hat{\upsilon}_{\alpha,\beta} = i\sum_{k=0}^{n-1}\hat{\mu}_{\alpha,k}\int_{z\in\mathbb{C}} d\pi(z)\left[q_k(z)\,q_\beta(\bar{z}) - q_\beta(z)\,q_k(\bar{z})\right]. \tag{5.2c}$$

As we integrate over all variables, the Pfaffian integration theorem can be viewed as a generalisation of the Corollary 5.1 proven a few lines above, for the case of a kernel not satisfying the projection property. This follows from the identity

$$\mathrm{pf}\left[\begin{array}{cc} Q_n(z_i,z_j) & Q_n(z_i,\bar{z}_j) \\ Q_n(\bar{z}_i,z_j) & Q_n(\bar{z}_i,\bar{z}_j) \end{array}\right]_{2\ell\times 2\ell} = \mathrm{qdet}\left[\hat{Q}_n(z_i,z_j)\right]_{\ell\times\ell},$$

where the quaternion $\hat{Q}_n(z_i,z_j)$ has the 2×2 matrix representation:

$$f(z_i,z_j) = \Theta[\hat{Q}_n(z_i,z_j)] = \left(\begin{array}{cc} -Q_n(\bar{z}_i,z_j) & -Q_n(\bar{z}_i,\bar{z}_j) \\ Q_n(z_i,z_j) & Q_n(z_i,\bar{z}_j) \end{array}\right). \tag{5.3}$$

To see that the Pfaffian integration theorem reduces to the Corollary for the particular case of a kernel with restored projection property, we spot that the latter is equivalent to the statement

$$\hat{\boldsymbol{\upsilon}} = -2i\,\hat{\mathbf{1}}_n \tag{5.4}$$

as can be deduced from the discussion below (3.24), Section 3.3. As a result,

$$\mathrm{tr}_{\,(0,n-1)}\,\hat{\boldsymbol{\upsilon}}^{\,j} = n\left(\frac{2}{i}\right)^j, \tag{5.5}$$

and the r.h.s. of (5.2b) reduces to (Macdonald 1998)

$$\left(\frac{i}{2}\right)^{\ell} Z_{(1^\ell)}\left(\frac{n}{2}\left(\frac{2}{i}\right)^1, \dots, \frac{n}{2}\left(\frac{2}{i}\right)^\ell\right) = \left(\frac{i}{2}\frac{\partial}{\partial z}\right)^{\ell}\exp\left(\frac{n}{2}\sum_{r\geq 1}(-1)^{r-1}\frac{(-2i\,z)^r}{r}\right)\Bigg|_{z=0}$$

$$= \left(\frac{\partial}{\partial z}\right)^{\ell}(1+z)^{n/2}\Bigg|_{z=0} = \frac{\Gamma(n/2+1)}{\Gamma(n/2+1-\ell)}. \tag{5.6}$$

Finally, noticing from (5.3) that

$$\int d\pi(z) f(z,z) = \int d\pi(z)\left(\begin{array}{cc} Q_n(z,\bar{z}) & 0 \\ 0 & Q_n(z,\bar{z}) \end{array}\right)$$

$$= \frac{i}{4}\,\hat{\boldsymbol{e}}_0\,\mathrm{tr}_{\,(0,n-1)}\hat{\boldsymbol{\upsilon}} = (n/2)\,\hat{\boldsymbol{e}}_0 \tag{5.7}$$

Table 3. The vocabulary of topological terms defined to interpret the ordered Pfaffian expansion. The notation used is: D – Definition, E – Example, L – Lemma, T – Theorem, C – Corollary, F – Figure.

Term	Notation	Appearance
String	$\mathcal{S}_i$	D5.1, E5.1, F3, F4
Length of a string	$\\|\mathcal{S}_i\\|$	D5.2, E5.2
Equivalent strings	$\mathcal{S}_i \sim \mathcal{S}_j$	D5.3, E5.3, L5.1, F3
Equivalence class of strings	$\mathcal{C}_j$	D5.4, E5.4, L5.1, F3, F5
Size of equivalence class	$\\|\mathcal{C}_j\\|$	D5.4, L5.1, F3
Substring	$\mathcal{S}_i^{(p)}$	D5.5, E5.5
Length of a substring	$p = \\|\mathcal{S}_i^{(p)}\\|$	D5.5, E5.5
Loop-like substring	$\mathcal{S}_i^{(p)}$	D5.6, E5.6, L5.2, L5.5, F3, F4
Longest loop-like string	$\mathcal{S}_i^{(\ell)}$	F3, E5.7, L5.4
Adjacent loop-like string	$\mathcal{S}_i^{(p)}$	D5.7, E5.7, E5.9, E5.11, L5.4–L5.6, C5.2, F4, F5
Handedness of adjacent (sub)string	$\mathbb{H}(\alpha_L, \alpha_R)$	D5.8, E5.8, E5.9, E5.11, L5.6, C5.2, F5
Equivalent adjacent (sub)strings	$\mathcal{S}_i \sim \mathcal{S}_j$	D5.9, E5.10, E5.11, L5.7, F5
Equivalence class of adjacent (sub)strings	$\mathcal{AC}_j$	D5.10
Compound string	$\mathcal{S}_i$	D5.11, T5.3, F3, F6
Topology class	$\boldsymbol{\lambda}$	F6

with $\hat{\boldsymbol{e}}_0 = \mathrm{diag}(1,1)$, we conclude that the constant c in the Corollary [(5.1b)] equals $c = n/2$ so that the result (5.6) brought by the Pfaffian integration theorem is identically equivalent to the one [(5.1a)] following from the Dyson integration theorem. We stress that this is only true for the kernel $Q_n(z_i, z_j)$ satisfying the projection property.

To prove the Theorem 5.1, we will invent a formalism based on a topological interpretation of the ordered Pfaffian expansion. For the readers' benefit, a vocabulary of the topological terms to be defined and used in the following sections is summarised in Table 3.

5.2. Topological interpretation of the ordered Pfaffian expansion

To integrate the Pfaffian in (5.2b), we start with its *ordered* expansion

$$\mathrm{pf}\begin{bmatrix} Q_n(z_i, z_j) & Q_n(z_i, \bar{z}_j) \\ Q_n(\bar{z}_i, z_j) & Q_n(\bar{z}_i, \bar{z}_j) \end{bmatrix}_{2\ell\times 2\ell} = \frac{1}{2^\ell \ell!} \sum_{\sigma\in S_{2\ell}} \mathrm{sgn}(\sigma) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma(2j+1)}, w_{\sigma(2j+2)}). \tag{5.8}$$

Here, the summation extends over all permutations $\sigma \in S_{2\ell}$ of 2ℓ objects

$$\{w_1 = z_1, w_2 = \bar{z}_1, \cdots, w_{2\ell-1} = z_\ell, w_{2\ell} = \bar{z}_\ell\} \tag{5.9}$$

so that the total number of terms in (5.8) is $(2\ell)!$.

5.2.1. Strings and their equivalence classes

Definition 5.1. *Each term of the ordered Pfaffian expansion is called a* **string***. The i-th string $\mathcal{S}_i$ equals*

$$\mathcal{S}_i = \operatorname{sgn}(\sigma_i) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma_i(2j+1)}, w_{\sigma_i(2j+2)}). \tag{5.10}$$

where σ_i is the i-th permutation out of $(2\ell)!$ possible permutations $\sigma \in S_{2\ell}$. Notice that a sign is attached to each string.

Example 5.1. The ordered expansion of the Pfaffian for $\ell = 2$ [see (4.7)] contains $4! = 24$ strings that we assign to three different groups (their meaning will become clear below):

Group 1	Group 2	Group 3
$+(\mathbf{1\bar{1}})(\mathbf{2\bar{2}})$	$+(\mathbf{1\bar{2}})(\mathbf{\bar{1}2})$	$-(\mathbf{12})(\mathbf{\bar{1}\bar{2}})$
$-(1\bar{1})(\bar{2}2)$	$-(1\bar{2})(2\bar{1})$	$+(12)(\bar{2}\bar{1})$
$+(\bar{1}1)(\bar{2}2)$	$+(\bar{2}1)(2\bar{1})$	$-(21)(\bar{2}\bar{1})$
$-(\bar{1}1)(2\bar{2})$	$-(\bar{2}1)(\bar{1}2)$	$+(21)(\bar{1}\bar{2})$
$+(2\bar{2})(1\bar{1})$	$+(\bar{1}2)(1\bar{2})$	$-(\bar{1}\bar{2})(12)$
$-(2\bar{2})(\bar{1}1)$	$-(\bar{1}2)(\bar{2}1)$	$+(\bar{1}\bar{2})(21)$
$+(\bar{2}2)(\bar{1}1)$	$+(2\bar{1})(\bar{2}1)$	$-(\bar{2}\bar{1})(21)$
$-(\bar{2}2)(1\bar{1})$	$-(2\bar{1})(1\bar{2})$	$+(\bar{2}\bar{1})(12)$

(5.11)

For brevity, the obvious notation $\pm(pq)(\bar{p}\bar{q})$ was used to denote the string $\pm Q_n(z_p, z_q) Q_n(\bar{z}_p, \bar{z}_q)$. The three strings shown in bold are those that previously appeared in (4.8) when treating the probability $p_{n,n-4}$.

Definition 5.2. *The* **length** $\|\mathcal{S}_i\|$ *of a string $\mathcal{S}_i$ equals the number of kernels it is composed of.*

Example 5.2. The string $\mathcal{S}_i$ in (5.10) is of the length ℓ: $\|\mathcal{S}_i\| = \ell$. All strings in (5.11) are of the length 2.

Definition 5.3. *Two strings $\mathcal{S}_i$ and $\mathcal{S}_j$ of the ordered Pfaffian expansion are said to be* **equivalent strings***, $\mathcal{S}_i \sim \mathcal{S}_j$, if they can be obtained from each other by (i) permutation of kernels and/or (ii) permutation of arguments inside kernels (these will also be called intra-kernel permutations).*

Example 5.3. For $\ell = 2$, three different groups of equivalent strings can be identified as suggested by (5.11). The first group, exemplified by the string $+(\mathbf{1\bar{1}})(\mathbf{2\bar{2}})$ involves 8 equivalent strings; two other groups, each consisting of 8 strings as well, are represented by the strings $+(\mathbf{1\bar{2}})(\mathbf{\bar{1}2})$ and $-(\mathbf{12})(\mathbf{\bar{1}\bar{2}})$, respectively.

Definition 5.4. *A group of equivalent strings is called the* **equivalence class of strings***. The j-th equivalence class to be denoted as $\mathcal{C}_j$ consists of $\|\mathcal{C}_j\|$ equivalent*

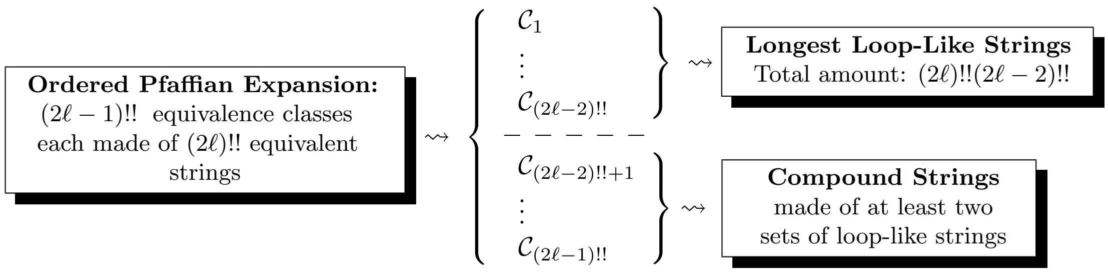

Figure 3. The $(2\ell)!$ terms, or strings, of an ordered Pfaffian expansion can be assigned to $(2\ell-1)!!$ equivalence classes $\{\mathcal{C}_1, \cdots, \mathcal{C}_{(2\ell-1)!!}\}$, each one containing $(2\ell)!!$ equivalent strings (Lemma 5.1). Any given string of the length ℓ can be decomposed into a set of loop-like substrings of respective lengths $\{\ell_j\}$ such that $\sum_j \ell_j = \ell$ (Lemma 5.2). The sets consisting of only one loop-like string are called *longest loop-like strings*; the sets made of more than one loop-like substring are called *compound strings*. An amount of longest loop-like strings is counted in Lemma 5.3.

strings. The number $\|\mathcal{C}_j\|$ is called the **size of equivalence class**.

Example 5.4. For $\ell = 2$, exactly three different equivalence classes of strings (5.11) can be identified in the ordered Pfaffian expansion (5.8).

In the context of the initiated classification of strings arising in the ordered Pfaffian expansion (5.8), a natural question to ask is this: Can the total number of equivalence classes and the number of of equivalent strings in each class be determined? The answer is provided by Lemma 5.1.

Lemma 5.1. *All terms of the ordered Pfaffian expansion can be assigned to* $(2\ell-1)!!$ *equivalence classes* $\{\mathcal{C}_1, \cdots, \mathcal{C}_{(2\ell-1)!!}\}$*, each containing* $(2\ell)!!$ *equivalent strings:* $\|\mathcal{C}_j\| = (2\ell)!!$ *for all* $j \in 1, \cdots, (2\ell-1)!!$.

Proof. Consider a string $\mathcal{S}_i$ belonging to the equivalence class $\mathcal{C}_j$ and composed of ℓ specific kernels, $\|\mathcal{S}_i\| = \ell$. There exist $\ell!$ possible permutations of kernels and 2^ℓ intra-kernel permutations of arguments. As a result, the total number of strings generated by these two operations from a given string $\mathcal{S}_i$ equals $\|\mathcal{C}_j\| = 2^\ell \ell! = (2\ell)!!$. Since the total number of strings in the ordered expansion is $(2\ell)!$, one concludes that there exist $(2\ell)!/(2\ell)!! = (2\ell-1)!!$ different equivalence classes. ■

The results of this and the two subsequent Sections are summarised in Fig. 3.

5.2.2. Decomposing strings into loop-like substrings

Having assigned all $(2\ell)!$ strings of the ordered Pfaffian expansion to $(2\ell-1)!!$ equivalence classes each containing $(2\ell)!!$ strings, we wish to concentrate on the structure of the strings themselves. Below, we shall prove that any string $\mathcal{S}_i$ (of the length $\|\mathcal{S}_i\| = \ell$) can be decomposed into a certain number (between 1 and ℓ) of *loop-like* *substrings*, see the Lemma 5.2. To prepare the reader to the definition of a *loop-like substring*, we first define the notion of a *substring* itself.

Definition 5.5. *A product* $\mathcal{S}_i^{(p)}$ *of* p *kernels* Q_n *is called a* **substring** *of the string*

$$\mathcal{S}_i = \mathrm{sgn}(\sigma_i) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma_i(2j+1)}, w_{\sigma_i(2j+2)}), \quad \|\mathcal{S}_i\| = \ell, \tag{5.12}$$

specified in Definition 5.1, if it takes the form

$$\mathcal{S}_i^{(p)} = \prod_{k=1}^{p} Q_n(w_{\sigma_i(2j_k+1)}, w_{\sigma_i(2j_k+2)}) \tag{5.13}$$

where $j_1 \neq j_2 \neq \cdots \neq j_p$. *The length* $\|\mathcal{S}_i^{(p)}\|$ *of the substring is* $\|\mathcal{S}_i^{(p)}\| = p$ *with* $1 \le p \le \ell$. *Notice that no sign is assigned to a substring.*

Example 5.5. The string

$$+(1\bar{1})(2\bar{3})(\bar{2}3)$$

arising in the $\ell = 3$ Pfaffian expansion (see the first term in (4.15)) can exhaustively be decomposed into seven substrings ¶

$$\underbrace{(1\bar{1}),\ (2\bar{3}),\ (\bar{2}3)}_{p=1},\ \underbrace{(1\bar{1})(2\bar{3}),\ (1\bar{1})(\bar{2}3),\ (2\bar{3})(\bar{2}3)}_{p=2},\ \underbrace{(1\bar{1})(2\bar{3})(\bar{2}3)}_{p=3}$$

of lengths $p = 1, 2$ and 3, respectively.

Definition 5.6. *A substring*

$$\mathcal{S}_i^{(p)} = \prod_{k=1}^{p} Q_n(w_{\sigma_i(2j_k+1)}, w_{\sigma_i(2j_k+2)})$$

is said to be a **loop-like substring** *of the length* $p = \|\mathcal{S}_i^{(p)}\|$, *if the two conditions are satisfied:*

(i) *The set of all* $2p$ *arguments*

$$\mathcal{W}_{2p} = \{w_{\sigma_i(2j_1+1)}, w_{\sigma_i(2j_1+2)}, \cdots, w_{\sigma_i(2j_p+1)}, w_{\sigma_i(2j_p+2)}\}$$

collected from the substring $\mathcal{S}_i^{(p)}$ *remains unchanged under the operation of complex conjugation*

$$\bar{\mathcal{W}}_{2p} = \{\bar{w}_{\sigma_i(2j_1+1)}, \bar{w}_{\sigma_i(2j_1+2)}, \cdots, \bar{w}_{\sigma_i(2j_p+1)}, \bar{w}_{\sigma_i(2j_p+2)}\}.$$

That means the two sets $\bar{\mathcal{W}}_{2p}$ *and* $\mathcal{W}_{2p}$ *of arguments are identical, up to their order. (This property will be referred to as invariance under complex conjugation.)*

(ii) *For all subsets* $\delta\mathcal{S}_i^{(q)}$ *consisting of* q *kernels* Q_n *with* $1 \le q \le p-1$, *the substring* $\mathcal{S}_i^{(p)} \backslash \delta\mathcal{S}_i^{(q)}$ *of the length* $p - q$ *obtained by removal of* $\delta\mathcal{S}_i^{(q)}$ *from* $\mathcal{S}_i^{(p)}$ *is not invariant under the operation of complex conjugation of its arguments.*

¶ It is easy to see that the number of substrings of the length p equals $\binom{\ell}{p}$ so that the total amount of all possible substrings of a string of the length ℓ is

$$\sum_{p=1}^{\ell} \binom{\ell}{p} = 2^\ell - 1.$$

Example 5.6. Out of seven substrings of the string $+(1\bar{1})(2\bar{3})(\bar{2}3)$ detailed in the Example 5.5, the two substrings

$$(1\bar{1}),\ \ (2\bar{3})(\bar{2}3)$$

are loop-like (of the lengths $p=1$ and 2, respectively). The remaining five substrings are not loop-like. Four of them,

$$(2\bar{3}),\ \ (\bar{2}3),\ \ (1\bar{1})(2\bar{3}),\ \ (1\bar{1})(\bar{2}3),$$

are not loop-like substrings because the property (i) of Definition 5.6 is not satisfied. The fifth substring (of the length $p=3$)

$$(1\bar{1})(2\bar{3})(\bar{2}3)$$

is not loop-like because the property (ii) of Definition 5.6 is violated. Indeed, there *does* exist a subset $\delta\mathcal{S}_i^{(1)}$ consisting of one kernel, represented by the pair of arguments $(1\bar{1})$, whose removal would *not* destroy the property (i) for the reduced substring $(2\bar{3})(\bar{2}3)$.

The example presented shows that a particular string of the ordered Pfaffian expansion could be decomposed into a set of loop-like substrings. Is such a decomposition possible in general? The answer is given by the following Lemma.

Lemma 5.2. *Any given string $\mathcal{S}_i$ of the length $\|\mathcal{S}_i\| = \ell$ from the ordered Pfaffian expansion can be decomposed into a set of loop-like substrings $\mathcal{S}_i^{(\ell_j)}$ of respective lengths $\|\mathcal{S}_i^{(\ell_j)}\| = \ell_j$,*

$$\mathcal{S}_i = \bigcup_j \mathcal{S}_i^{(\ell_j)}$$

such that $\sum_j \ell_j = \ell$.

Proof. We use induction to prove the above statement.

(i) *Induction Basis.* For $\ell = 1$, the Lemma obviously holds since the strings $(1,\bar{1})$ and $(\bar{1},1)$ are loop-like by Definition 5.6.

(ii) *Induction Hypothesis.* The Lemma is supposed to hold for any string $\mathcal{S}_i$ of the length $\|\mathcal{S}_i\| = \ell$:

$$\mathcal{S}_i = \bigcup_j \mathcal{S}_i^{(\ell_j)}, \quad \text{with} \quad \sum_j \ell_j = \ell. \tag{5.14}$$

(iii) *Induction Step.* Consider a given string $\tilde{\mathcal{S}}_i$ of the length $\|\tilde{\mathcal{S}}_i\| = \ell + 1$. Given the induction hypothesis, we are going to prove that such a string $\tilde{\mathcal{S}}_i$ can be decomposed into a set of loop-like substrings.

To proceed, we note that any given string $\tilde{\mathcal{S}}_i$ of the length $\|\tilde{\mathcal{S}}_i\| = \ell+1$ can be generated from some string $\mathcal{S}_i$ of length ℓ (see 5.12) by adding to it an additional pair of arguments $(z_{\ell+1}, \bar{z}_{\ell+1})$:

$$\tilde{\mathcal{S}}_i = (z_{\ell+1}, \bar{z}_{\ell+1}) \quad \otimes \quad \underbrace{\prod_{k=1}^{\ell} (w_{\sigma_i(2j_k+1)}, w_{\sigma_i(2j_k+2)})}_{\text{the string } \mathcal{S}_i \text{ with } \mathrm{sgn}(\sigma_i) \text{ dropped}} \tag{5.15}$$

with (or without) further exchange of either $z_{\ell+1}$ or $\bar{z}_{\ell+1}$ with one of the arguments belonging to the string $\mathcal{S}_i$ of length ℓ.

(a) If no exchange is made, the given string $\tilde{\mathcal{S}}_i$ is a unit of a single loop-like string $\Pi^{(1)} = (z_{\ell+1}, \bar{z}_{\ell+1})$ and of a string $\mathcal{S}_i$ admitting the decomposition (5.14). As a result, the string $\tilde{\mathcal{S}}_i$ of the length $(\ell+1)$ is decomposed into a set of loop-like substrings

$$\tilde{\mathcal{S}}_i = \Pi^{(1)} \bigcup \left(\bigcup_j \mathcal{S}_i^{(\ell_j)} \right). \tag{5.16}$$

(b) If either $z_{\ell+1}$ or $\bar{z}_{\ell+1}$ was swapped with one of the arguments belonging to a loop-like substring $S_i^{(\ell_{j_0})} \subset \mathcal{S}_i$ of the string $\mathcal{S}_i$, such an exchange will give rise to a new loop-like substring $S_i^{(\ell'_{j_0})} \subset \tilde{\mathcal{S}}_i$ of the length

$$\|S_i^{(\ell'_{j_0})}\| = \|S_i^{(\ell_{j_0})}\| + 1 = \ell_{j_0} + 1.$$

Consequently, the given string $\tilde{\mathcal{S}}_i$ of the length $(\ell+1)$ is then decomposed into a set of loop-like substrings

$$\tilde{\mathcal{S}}_i = \mathcal{S}_i^{(\ell'_{j_0})} \bigcup \left(\bigcup_{j \neq j_0} \mathcal{S}_i^{(\ell_j)} \right). \tag{5.17}$$

To prove that the substring $\mathcal{S}_i^{(\ell'_{j_0})}$ is indeed loop-like, two properties have to be checked in accordance with the Definition 5.6.

First, one has to show that the set of $2(\ell_{j_0}+1)$ arguments collected from the substring $\mathcal{S}_i^{(\ell'_{j_0})}$ is invariant under the operation of complex conjugation; this is obviously true because $\mathcal{S}_i^{(\ell_{j_0})}$ is loop-like. Second, one has to demonstrate that the removal of any subset $\delta\mathcal{S}$ from $\mathcal{S}_i^{(\ell'_{j_0})}$ will destroy the *invariance property of the remaining substring* $\mathcal{S}_i^{(\ell'_{j_0})} \backslash \delta\mathcal{S}$. Three different cases are to be considered here:

(b1) If the subset $\delta\mathcal{S}$ does not contain the fragments $(z_{\ell+1}, \cdots)$ and $(\cdots, \bar{z}_{\ell+1})$, it is also a subset of $S_i^{(\ell_{j_0})}$. Since the latter is loop-like, the invariance property is destroyed.

(b2) If the subset $\delta\mathcal{S}$ contains only one of the fragments $(z_{\ell+1}, \cdots)$ or $(\cdots, \bar{z}_{\ell+1})$, the invariance property is obviously destroyed.

(b3) If the subset $\delta\mathcal{S}$ contains both fragments $(z_{\ell+1}, u)$ and $(v, \bar{z}_{\ell+1})$, the invariance property is also destroyed. To prove it, we use the *reductio ad absurdum*. Indeed, let us assume that there exists a subset $\delta\mathcal{S} \subset \mathcal{S}_i^{(\ell'_{j_0})} \subset \tilde{S}_i$, containing both fragments $(z_{\ell+1}, u)$ and $(v, \bar{z}_{\ell+1})$, whose removal does not destroy the invariance property of $\mathcal{S}_i^{(\ell'_{j_0})} \backslash \delta\mathcal{S}$. The existence of such a subset $\delta\mathcal{S}$ implies the existence of yet another subset $\delta\mathcal{S}' \subset \mathcal{S}_i^{(\ell_{j_0})} \subset S_i$,

$$\delta\mathcal{S}' = \Big\{ \delta\mathcal{S} \backslash \{(z_{\ell+1}, u) \cup (v, \bar{z}_{\ell+1})\} \Big\} \cup \{(v, u)\}, \tag{5.18}$$

whose removal *does* destroy the invariance of $\mathcal{S}_i^{(\ell_{j_0})} \backslash \delta\mathcal{S}'$ under the operation of complex conjugation (this claim is obviously true because the substring

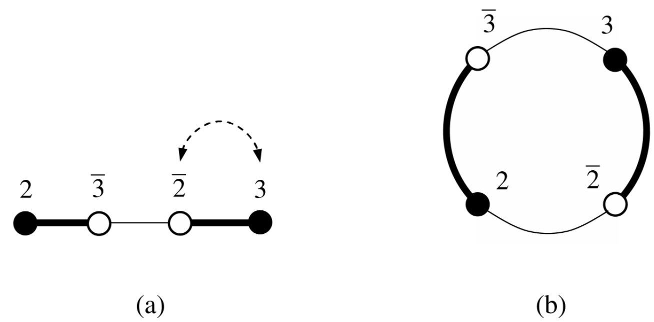

Figure 4. Graphic explanation of the notion "loop-like substring". (a) The substring $(2\bar{3})(\bar{2}3)$ can be transformed into the form of an adjacent substring (in the sense of Definition 5.7) by flipping arguments in the second kernel, $(\bar{2}3) \mapsto (3\bar{2})$. (b) The emerging adjacent substring $(2\bar{3})(3\bar{2})$ can clearly be depicted in the form of a loop by gluing the arguments 2 and $\bar{2}$ together.

$\mathcal{S}_i^{(\ell_{j_0})}$ is loop-like). Then, the identity

$$\mathcal{S}_i^{(\ell'_{j_0})} \backslash \delta\mathcal{S} = \mathcal{S}_i^{(\ell_{j_0})} \backslash \delta\mathcal{S}' \tag{5.19}$$

suggests that the removal of $\delta\mathcal{S}$ from $\mathcal{S}_i^{(\ell'_{j_0})}$ *must* destroy the invariance of $\mathcal{S}_i^{(\ell'_{j_0})} \backslash \delta\mathcal{S}$ as well. Since this contradicts to the assumption made, one concludes that the invariance property is indeed destroyed.

End of the proof. ■

We close this Section by providing a brief explanation of the origin of the term "loop-like substring" introduced in Definition 5.6. In Section 5.2.4, it will be proven that any loop-like substring can be brought to the form of an adjacent substring (Definition 5.7) by means of proper (i) permutation of kernels and/or (ii) permutation of intra-kernel arguments (Lemma 5.5). The latter can graphically be represented as a *loop*. Indeed, the loop-like substring $(2\bar{3})(\bar{2}3)$ considered in Example 5.6 can be transformed into the form of an adjacent substring (in the sense of Definition 5.7) by flipping arguments in the second kernel, $(\bar{2}3) \mapsto (3\bar{2})$:

$$\underbrace{(2\bar{3})(\bar{2}3)}_{\text{loop}-\text{like substring}} \quad \mapsto \quad \underbrace{(2\bar{3})(3\bar{2})}_{\text{adjacent substring}} \quad .$$

The resulting adjacent string $(2\bar{3})(3\bar{2})$ can be drawn in the form of a loop, see Fig. 4. This is precisely the reason why the substring $(2\bar{3})(\bar{2}3)$ is called loop-like.

5.2.3. Counting longest loop-like substrings $\mathcal{S}_i^{(\ell)}$ of the length ℓ

Although in this subsection, we are going to concentrate on the longest loop-like substrings [+] of the length ℓ, our main counting result given by Lemma 5.3 stays valid for loop-like substrings of a smaller length $1 \le p < \ell$.

Lemma 5.3. *The ordered Pfaffian expansion contains* $(2\ell)!!\,(2\ell-2)!!$ *longest loop-like strings of the length* ℓ.

[+] As soon as substrings of the longest possible length ℓ are considered, they are strings themselves.

Proof. Let $\mathcal{N}_L(\ell)$ be the total number of loop-like strings of the length ℓ in the ordered Pfaffian expansion and let $n_L(\ell)$ denote the number of equivalence classes all longest loop-like strings can be assigned to. Following Lemma 5.1, the two are related to each other as

$$\mathcal{N}_L(\ell) = (2\ell)!!\, n_L(\ell), \tag{5.20}$$

because each equivalence class $\mathcal{C}_j$ contains precisely $(2\ell)!!$ equivalent strings (see Definition 5.3). It thus remains to determine $n_L(\ell)$ that can equally be interpreted as a number of *inequivalent* * longest loop-like strings of the length ℓ.

The latter can be evaluated by counting the number of ways, $n_L(\ell+1)$, all *inequivalent* longest loop-like strings of the length $\ell+1$ can be generated from those longest of the length ℓ. Since both numbers, $n_L(\ell)$ and $n_L(\ell+1)$, refer to the *longest* loop-like strings, the two correspond to the Pfaffians of matrices of the size $\ell \times \ell$ and $(\ell+1)\times(\ell+1)$, correspondingly. In the language of strings, an increase of the matrix size by one leads to the appearance of an additional pair of arguments $(z_{\ell+1}, \bar{z}_{\ell+1})$ in a string of the length $\ell+1$.

We claim that

$$n_L(\ell+1) = 2\ell\, n_L(\ell). \tag{5.21}$$

To prove this, we concentrate on a given longest loop-like string of the length ℓ and add to it an additional pair of arguments $(z_{\ell+1}, \bar{z}_{\ell+1})$:

$$(z_{\ell+1}, \bar{z}_{\ell+1}) \quad \otimes \quad \underbrace{\prod_{k=1}^{\ell} (w_{\sigma_i(2j_k+1)}, w_{\sigma_i(2j_k+2)})}_{\text{longest loop}-\text{like string of the length } \ell} \quad . \tag{5.22}$$

Here, σ_i is a particular permutation of 2ℓ arguments (5.9) corresponding to a longest loop-like string of the length ℓ. The resulting string (5.22) is *not* a longest loop-like string of the length $\ell+1$ (rather, it is composed of two loop-like strings of the lengths 1 and ℓ, respectively). Since a loop-like string necessarily assumes the presence of a fragment $(z_{\ell+1}, \cdots)(\cdots, \bar{z}_{\ell+1})$ somewhere in the string, one has to exchange either $z_{\ell+1}$ or $\bar{z}_{\ell+1}$ with one of the arguments belonging to the original longest loop-like string of the length ℓ. Clearly, there exist 2ℓ exchange options for each argument, $z_{\ell+1}$ (or $\bar{z}_{\ell+1}$). As a result, we arrive at the relation (5.21). Given $n_L(1) = 1$, we derive the desired result by induction:

$$n_L(\ell) = (2\ell-2)!! \tag{5.23}$$

Combining it with (5.20) completes the proof. ♯ ■

5.2.4. Adjacent vs non-adjacent loop-like substrings

Further classification of loop-like substrings is needed in order to prepare ourselves to

* In view of Definition 5.3, the two strings are inequivalent if they cannot be reduced to each other by means of (i) permutation of kernels and/or (ii) intra-kernel permutation of arguments.

♯ It is instructive to turn to Example 5.1 that discusses the ordered Pfaffian expansion for $\ell = 2$. The Lemma 5.3 predicts existence of $\mathcal{N}_L(2) = 4!!\,2!! = 16$ longest loop-like strings that can be assigned to $n_L(2) = 2!! = 2$ equivalence classes, each composed of $(2\ell)!! = 4!! = 8$ equivalent strings. This is in line with direct counting (5.11): the two equivalence classes are represented by the longest loop-like strings $+(1\bar{2})(\bar{1}2)$ and $-(12)(\bar{1}\bar{2})$ (see the second and third column); each equivalence class contains 8 equivalent strings.

the proof of the Pfaffian integration theorem.

Definition 5.7. *A loop-like substring* $\mathcal{S}_i^{(p)}$ *of the length* $p = \|\mathcal{S}_i^{(p)}\|$ *is called* **adjacent loop-like substring**, *or simply a* **loop**, *if it is represented by a product*

$$\mathcal{S}_i^{(p)} = \prod_{k=1}^{p} Q_n(w_{\sigma_i(2j_k+1)}, w_{\sigma_i(2j_k+2)})$$

of p *kernels such that:*

(i) *The first argument of the first kernel and the second argument of the last,* p*-th kernel, are complex conjugate of each other,*

$$w_{\sigma_i(2j_1+1)} = \bar{w}_{\sigma_i(2j_p+2)}.$$

(ii) *For each pair of neighbouring kernels in the string, the second argument of the left kernel in the pair and the first argument of the right kernel in the pair are complex conjugate of each other,*

$$w_{\sigma_i(2j_k+2)} = \bar{w}_{\sigma_i(2j_{k+1}+1)}, \quad k = 1, \cdots, p-1.$$

Example 5.7. Out of 16 longest loop-like strings arising in the Pfaffian expansion for $\ell = 2$ †, the following eight are adjacent:

$$\begin{aligned} &-(1\bar{2})(2\bar{1}), \ -(\bar{2}1)(\bar{1}2), \ -(\bar{1}2)(\bar{2}1), \ -(2\bar{1})(1\bar{2}), \\ &+(12)(\bar{2}\bar{1}), \ +(21)(\bar{1}\bar{2}), \ +(\bar{1}\bar{2})(21), \ +(\bar{2}\bar{1})(12). \end{aligned}$$

Notice that although *longest* loop-like strings have been considered in the above example, the notion of an adjacent string is equally relevant for a loop-like string of *smaller* length.

Lemma 5.4. *Out of* $(2\ell)!!(2\ell-2)!!$ *longest loop-like strings of the length* ℓ *associated with an ordered Pfaffian expansion, exactly* $(2\ell)!!$ *are adjacent.*

Proof. To count the total number $\mathcal{N}_A(\ell)$ of all adjacent loop-like strings of the length ℓ, we consider a specific pair of adjacent loop-like strings of the length ℓ represented by the sequences of arguments

$$(z_{j_1}, \underbrace{\bar{z}_{j_2})(z_{j_2}}_{\text{pair}\,\natural\,1}, \underbrace{\bar{z}_{j_3})(z_{j_3}}_{\text{pair}\,\natural\,2}, \bar{z}_{j_4}) \cdots (z_{j_{p-1}}, \bar{z}_{j_p}) \cdots (z_{j_{\ell-1}}, \underbrace{\bar{z}_{j_\ell})(z_{j_\ell}}_{\text{pair}\,\natural\,(\ell-1)}, \bar{z}_{j_1})$$

and

$$(\bar{z}_{j_1}, \underbrace{\bar{z}_{j_2})(z_{j_2}}_{\text{pair}\,\natural\,1}, \underbrace{\bar{z}_{j_3})(z_{j_3}}_{\text{pair}\,\natural\,2}, \bar{z}_{j_4}) \cdots (z_{j_{p-1}}, \bar{z}_{j_p}) \cdots (z_{j_{\ell-1}}, \underbrace{\bar{z}_{j_\ell})(z_{j_\ell}}_{\text{pair}\,\natural\,(\ell-1)}, z_{j_1}).$$

Here, the mutually distinct j_p take the values from 1 to ℓ. The two strings are identical up to an exchange of the first and the last arguments $z_{j_1} \rightleftarrows \bar{z}_{j_1}$. The remaining $2(\ell-1)$ arguments are distributed between the kernels in such a way that an adjacent string is formed in accordance with the Definition 5.7; the $(\ell-1)$ underbraces identifying

† For an example, please refer to the second and third column in (5.11). Also, see Lemma 5.3 for the explanation of the number $16 = 4!!\,2!!$ and Definition 5.6 for the notion of a loop-like string.

$(\ell-1)$ pairs of complex conjugate arguments highlight the structure of an adjacent string.

The total number $\mathcal{N}_A(\ell)$ of all adjacent loop-like strings of the length ℓ equals the number of ways to generate those strings from the two depicted above. As soon as there are (i) ℓ ways to assign a number from 1 to ℓ to the label j_1, (ii) $(\ell-1)!$ ways to assign the remaining $(\ell-1)$ numbers to the $(\ell-1)$ pairs left (labelled by $j_2, \cdots, j_\ell$), and (iii) $2^{\ell-1}$ ways to exchange the arguments $z_{j_k} \rightleftarrows \bar{z}_{j_k}$ $(k=2,\cdots,\ell)$ within those $(\ell-1)$ pairs, we derive:

$$\mathcal{N}_A(\ell) = 2 \times \ell \times (\ell-1)! \times 2^{\ell-1} = (2\ell)!! \tag{5.24}$$

End of proof.‡ ■

Remark 5.1. Since the above reasoning holds for loop-like (sub)strings of any length $1 \le p \le \ell$, one concludes that the total number of adjacent (sub)strings of the length p equals $\mathcal{N}_A(p) = (2p)!!$.

Lemma 5.5. *Any loop-like (sub)string of the length p can be transformed into an adjacent (sub)string of the same length by means of proper (i) permutation of kernels and/or (ii) permutation of intra-kernel arguments.*

Proof. To be coherent with the notations used in the proof of the Lemma 5.3, we deal below with a loop-like string of the length ℓ. However, the very same argument applies to any loop-like (sub)string of the length $1 \le p \le \ell$ so that our proof (based on mathematical induction) holds generally.

(i) For $\ell = 1$ and $\ell = 2$, the Lemma's statement is obviously true. Indeed, a loop-like string of the length $\ell = 1$ is automatically an adjacent one. For $\ell = 2$, a loop-like string composed of two kernels is reduced to an adjacent string by utmost one intra-kernel permutation of arguments.

(ii) Now we assume the Lemma to hold for loop-like (sub)strings of the length ℓ (that is, that any loop-like string of the length ℓ can be reduced to an adjacent string by means of the two types of allowed operations).

(iii) Given the previous assumption, we have to show that a loop-like (sub)string of the length $(\ell+1)$ can also be reduced to an adjacent (sub)string. It follows from the proof of the Lemma 5.3 (see the discussion around (5.22)) that a loop-like string of the length $(\ell+1)$ can be generated from a loop-like string of a smaller length ℓ by adding an additional pair of arguments $(z_{\ell+1}, \bar{z}_{\ell+1})$ followed by exchange of either $z_{\ell+1}$ or $\bar{z}_{\ell+1}$ with one of the arguments of the original loop-like string of the length ℓ. Since, under the induction assumption (ii), the latter can be made adjacent,

$$(z_{\ell+1}, \bar{z}_{\ell+1}) \otimes \underbrace{[(z_{j_1}, \dots) \cdots (\dots, \bar{z}_{j_q})\,(z_{j_q}, \dots) \cdots (\dots, \bar{z}_{j_1})]}_{\text{adjacent string of the length } \ell}, \tag{5.25}$$

one readily concludes that we are only two steps away from forming an adjacent string of the length $(\ell+1)$ out of (5.25). Indeed, an exchange of arguments $\bar{z}_{\ell+1} \rightleftarrows \bar{z}_{j_q}$ brings (5.25) to the form

$$(z_{\ell+1}, \bar{z}_{j_q})\,(z_{j_1}, \dots) \cdots (\dots, \bar{z}_{\ell+1})\,(z_{j_q}, \dots) \cdots (\dots, \bar{z}_{j_1}) \tag{5.26}$$

‡ In particular, there should exist eight adjacent strings in the ordered Pfaffian expansion for $\ell = 2$. This is in concert with the explicit counting in Example 5.7.

which boils down to the required adjacent string

$$\underbrace{(z_{j_1},\ldots)\cdots(\ldots,\bar z_{\ell+1})\,(z_{\ell+1},\bar z_{j_q})\,(z_{j_q},\ldots)\cdots(\ldots,\bar z_{j_1})}_{\text{adjacent string of the length }(\ell+1)} \tag{5.27}$$

upon moving the pair $(z_{\ell+1}, \bar z_{j_q})$ through $(q-1)$ pairs on the right. Exchanging $\bar z_{\ell+1} \rightleftarrows z_{j_q}$ instead can be done in the same way. ■

Remark 5.2. The reduction of a loop-like (sub)string to an adjacent (sub)string is *not* unique. For instance, the loop-like string

$$+(1\bar 3)(\bar 2 3)(\bar 1 2)$$

(see the first term in (4.16)) can be reduced, by permutation of kernels and permutation of intra-kernel arguments, to one of the following adjacent strings:

$$\begin{aligned}&+(1\bar 3)(3\bar 2)(2\bar 1),\quad -(\bar 1 2)(\bar 2 3)(\bar 3 1),\\ &+(2\bar 1)(1\bar 3)(3\bar 2),\quad -(\bar 2 3)(\bar 3 1)(\bar 1 2),\\ &+(3\bar 2)(2\bar 1)(1\bar 3),\quad -(\bar 3 1)(\bar 1 2)(\bar 2 3).\end{aligned} \tag{5.28}$$

To handle the problem of the non-unique reduction of a loop-like string to an adjacent string, the notion of string *handedness* has to be introduced.

5.2.5. Handedness of an adjacent substring

Definition 5.8. *Close an adjacent substring $\mathcal{S}$ into a loop by "gluing" the right argument of the last kernel with the left argument of the first kernel below the chain as in Figure 4.,*

$$\mathcal{S}:\ \rightsquigarrow \bullet \underbrace{(w_{j_1},\bar w_{j_2})}_{\text{first kernel}}\cdot(w_{j_2},\bar w_{j_3})\cdots(w_{j_{q-1}},\bar w_{j_q})\cdots(w_{j_{p-1}},\bar w_{j_p})\cdot\underbrace{(w_{j_p},\bar w_{j_1})}_{\text{last kernel}}\bullet \rightsquigarrow$$

(here, a set of the arguments $(w_1,\cdots,w_{2p})$ is specified by (5.9) with ℓ set to p, and the symbol $\bullet$ denotes a gluing point). Read the arguments of a loop, one after the other, in a clockwise direction (as depicted by the symbol $\rightsquigarrow$), starting with $\bar w_{j_1}$ until you arrive at w_{j_p}. If α_R is the number of times an argument z_{j_q} is followed by its conjugate $\bar z_{j_q}$ for all $q \in (1,\cdots,p)$,

$$(\ldots, z_{j_q})\cdot(\bar z_{j_q},\ldots),$$

an adjacent substring $\mathcal{S}$ is said to have the **handedness** *$\mathbb{H}(\alpha_L,\alpha_R)$, where $\alpha_L = p-\alpha_R$.*

Example 5.8. The handedness of eight adjacent strings considered in the Example 5.7 with $p = l = 2$ is listed below:

$$\begin{aligned}&\mathbb{H}(2,0),\ \ \mathbb{H}(0,2),\ \ \mathbb{H}(0,2),\ \ \mathbb{H}(2,0),\\ &\mathbb{H}(1,1),\ \ \mathbb{H}(1,1),\ \ \mathbb{H}(1,1),\ \ \mathbb{H}(1,1).\end{aligned} \tag{5.29}$$

Lemma 5.6. *Out of $(2\ell)!!$ adjacent strings of the length ℓ arising in an ordered Pfaffian expansion, there are exactly*

$$\mathcal{N}_\alpha(\ell) = \ell!\binom{\ell}{\alpha} \tag{5.30}$$

strings with the handedness $\mathbb{H}(\alpha, \ell - \alpha)$.

Proof. A string with the handedness $\mathbb{H}(\alpha, \ell - \alpha)$ closed into a loop (see Definition 5.8) contains α "left" fragments $(\dots, \bar{z}_{j_q}) \cdot (z_{j_q}, \dots)$ and $(\ell - \alpha)$ "right" fragments $(\dots, z_{j_q}) \cdot (\bar{z}_{j_q}, \dots)$ with the opposite order of complex conjugation in the nearest neighbouring kernels; each fragment is labelled by an integer number $j_q \in (1, \cdots, \ell)$. To count a total number of all strings with the handedness $\mathbb{H}(\alpha, \ell - \alpha)$, we notice that there exist $\binom{\ell}{\alpha}$ ways to distribute "left" and "right" fragments on the loop, and $\ell!$ ways to assign ℓ integer numbers (from 1 to ℓ) to the labels $j_1, \cdots, j_\ell$. Applying the combinatorial multiplication rule completes the proof. ■

Remark 5.3. It is instructive to realise that the Lemma 5.4 can be seen as a corollary to the Lemma 5.6. Indeed, the total number of all adjacent strings of the length ℓ is nothing but

$$\sum_{\alpha=0}^{\ell} \mathcal{N}_\alpha(\ell) = \ell! \sum_{\alpha=0}^{\ell} \binom{\ell}{\alpha} = (2\ell)!!$$

Not unexpectedly, this result is in concert with the Lemma 5.4.

Corollary 5.2. *The number of adjacent substrings of the length* p, $1 \le p < \ell$, *with the handedness* $\mathbb{H}(\alpha, p - \alpha)$ *equals*

$$\mathcal{N}_\alpha(p) = p! \binom{p}{\alpha}. \tag{5.31}$$

The total number of all adjacent substrings of the length p *is* $(2p)!!$

Proof. Follow the proof of the Lemma 5.6 and the Remark 5.3 with ℓ replaced by p. ■

Example 5.9. Out of 16 longest loop-like strings arising in the Pfaffian expansion for $\ell = 2$, there are eight adjacent as explicitly specified in Example 5.7. The handedness of those strings was considered in Example 5.8. In accordance with the Lemma 5.6, there must exist $\mathcal{N}_0(2) = 2$ strings of the handedness $\mathbb{H}(0, 2)$, $\mathcal{N}_2(2) = 2$ strings of the handedness $\mathbb{H}(2, 0)$, and $\mathcal{N}_1(2) = 4$ strings of the handedness $\mathbb{H}(1, 1)$. Direct counting (5.29) confirms that this is indeed the case.

5.2.6. Equivalence classes of adjacent (sub)strings

Having defined a notion of the handedness of an adjacent string, we are back to the issue of a *non-uniqueness* of reduction of a loop-like string to an adjacent one. To deal with the indicated non-uniqueness problem, we would like to define, and explicitly identify, all distinct *equivalence classes* for $(2\ell)!!$ adjacent strings arising in the context of an ordered Pfaffian expansion.

Definition 5.9. *Two adjacent strings,* $\mathcal{S}_i$ *and* $\mathcal{S}_j$, *are said to be* **equivalent adjacent strings**, $\mathcal{S}_i \sim \mathcal{S}_j$, *if they can be obtained from each other by (i) permutation of kernels and/or (ii) intra-kernel permutation of arguments.*

Example 5.10. The six adjacent strings (5.28) are equivalent to each other.

Definition 5.10. *A group of equivalent adjacent strings is called the* **equivalence class of adjacent strings**. *The j-th equivalence class to be denoted $\mathcal{AC}_j$ consists of $\|\mathcal{AC}_j\|$ equivalent adjacent strings.*

Lemma 5.7. *All $(2\ell)!!$ adjacent strings arising in the ordered Pfaffian expansion can be assigned to $(2\ell-2)!!$ equivalence classes $\{\mathcal{AC}_1, \cdots, \mathcal{AC}_{(2\ell-2)!!}\}$ of adjacent strings, each class containing 2ℓ equivalent adjacent strings: $\|\mathcal{AC}_j\| = 2\ell$ for all $j \in 1, \cdots, (2\ell-2)!!$.*

Proof. Let us concentrate on a given adjacent string $\mathcal{S}$ with the handedness $\mathbb{H}(\alpha, \ell-\alpha)$ that belongs to an equivalence class $\mathcal{AC}_j$ and count a number of ways to generate equivalent strings out of it by means of (i) permutation of kernels and/or (ii) intra-kernel permutation of arguments *without destroying the adjacency property*. Two complementary generating mechanisms exist.

- First way (M1): One starts with an adjacent string of the handedness $\mathbb{H}(\alpha, \ell-\alpha)$,

$$\rightsquigarrow \bullet \underbrace{(w_{j_1}, \bar{w}_{j_2})}_{\text{first kernel}} \cdot (w_{j_2}, \bar{w}_{j_3}) \cdots (w_{j_{p-1}}, \bar{w}_{j_p}) \cdots (w_{j_{\ell-1}}, \bar{w}_{j_\ell}) \cdot \underbrace{(w_{j_\ell}, \bar{w}_{j_1})}_{\text{last kernel}} \bullet \rightsquigarrow \tag{5.32}$$

  (the reader is referred to the Definition 5.8 for the notation used), simultaneously flips the intra-kernel arguments in all ℓ kernels,

$$\underbrace{(\bar{w}_{j_2}, w_{j_1})}_{\text{first kernel}} \cdot (\bar{w}_{j_3}, w_{j_2}) \cdots (\bar{w}_{j_p}, w_{j_{p-1}}) \cdots (\bar{w}_{j_\ell}, w_{j_{\ell-1}}) \cdot \underbrace{(\bar{w}_{j_1}, w_{j_\ell})}_{\text{last kernel}},$$

  and further permutes the kernels in a fan-like way so that the last ℓ-th kernel in the string becomes the first, the $(\ell-1)$-th kernel in the string becomes the second, etc.:

$$\rightsquigarrow \bullet \underbrace{(\bar{w}_{j_1}, w_{j_\ell})}_{\text{last kernel}} \cdot (\bar{w}_{j_\ell}, w_{j_{\ell-1}}) \cdots (\bar{w}_{j_p}, w_{j_{p-1}}) \cdots (\bar{w}_{j_3}, w_{j_2}) \cdot \underbrace{(\bar{w}_{j_2}, w_{j_1})}_{\text{first kernel}} \bullet \rightsquigarrow \tag{5.33}$$

  The so obtained *adjacent* string is equivalent to the initial one (5.32) but possesses the complementary handedness $\mathbb{H}(\ell-\alpha, \alpha)$.

- Second way (M2): One starts with an adjacent string of the handedness $\mathbb{H}(\alpha, \ell-\alpha)$, and permutes the ℓ kernels in a cyclic manner to generate $(\ell-1)$ additional (but equivalent) adjacent strings with the same handedness (to visualise the process, one may think of moving a gluing point $\bullet$ through the kernels in (5.32)).

The two mechanisms, M1 and M2, combined together bring up 2ℓ equivalent adjacent strings due to the combinatorial multiplication rule. As a result, we conclude that $\|\mathcal{AC}_j\| = 2\ell$. Consequently, the number of distinct equivalence classes equals $(2\ell)!!/(2\ell) = (2\ell-2)!!$. ■

Example 5.11. To illustrate the Lemma 5.7, we consider an ordered Pfaffian expansion for $\ell = 2$ as detailed in the Example 5.1. The eight adjacent strings of the length $\ell = 2$ were specified in the Example 5.7; their handedness was considered in the Example 5.8. In accordance with the Lemma 5.7, there should exist 2 distinct equivalence classes, each containing 4 adjacent strings. Indeed, one readily verifies that those two equivalence classes are

$$\mathcal{AC}_1: \quad -(1\bar{2})(2\bar{1}), \quad -(\bar{2}1)(\bar{1}2), \quad -(\bar{1}2)(\bar{2}1), \quad -(2\bar{1})(1\bar{2})$$

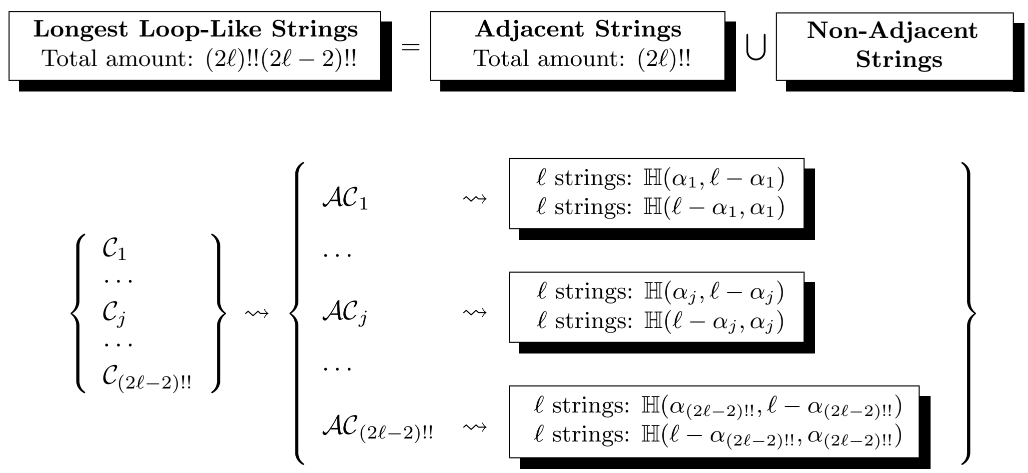

Figure 5. Any longest loop-like string from the equivalence classes $\{\mathcal{C}_1, \cdots, \mathcal{C}_{(2\ell-2)!!}\}$ can be reduced to an adjacent string belonging to one of the equivalence classes $\{\mathcal{AC}_1, \cdots, \mathcal{AC}_{(2\ell-2)!!}\}$ of adjacent strings by means of proper permutation of kernels and/or permutation of intra-kernel arguments (Lemma 5.5). This reduction is non-unique (Remark 5.2). Each equivalence class $\mathcal{AC}_j$ of adjacent strings contains ℓ adjacent strings of the handedness $\mathbb{H}(\alpha_j, \ell-\alpha_j)$ and ℓ adjacent strings of the complementary handedness $\mathbb{H}(\ell-\alpha_j, \alpha_j)$ (Remark 5.4). Notice that two distinct equivalence classes $\mathcal{AC}_i$ and $\mathcal{AC}_j$ (where $i \neq j$) may have the same values for α: $\alpha_i = \alpha_j$.

and

$$\mathcal{AC}_2: \quad +(12)(\bar{2}\bar{1}), \quad +(21)(\bar{1}\bar{2}), \quad +(\bar{1}\bar{2})(21), \quad +(\bar{2}\bar{1})(12).$$

Remark 5.4. In fact, a generic prescription can be given to build $(2\ell-2)!!$ distinct equivalence classes $\{\mathcal{AC}_1, \cdots, \mathcal{AC}_{(2\ell-2)!!}\}$ for $(2\ell)!!$ adjacent strings arising in the ordered Pfaffian expansion. Because of the "duality" between equivalent adjacent strings with complementary handedness $\mathbb{H}(\alpha, \ell-\alpha)$ and $\mathbb{H}(\ell-\alpha, \alpha)$ discussed in the proof of the Lemma 5.6, the adjacent strings whose handedness $\mathbb{H}(\alpha, \ell-\alpha)$ is restricted by the inequality $0 \le \alpha \le \lfloor \ell/2 \rfloor$ will form a natural basis in the consideration to follow.

- **The case $\ell = 2\lambda + 1$ odd.**

  (i) For $0 \le \alpha_1 \le \lambda$, (i.a) pick up an adjacent string with the handedness $\mathbb{H}(\alpha_1, \ell-\alpha_1)$, out of $\mathcal{N}_{\alpha_1}(\ell)$, and generate ℓ equivalent adjacent strings with the same handedness through the mechanism M2 of the Lemma 5.7. (i.b) Apply the mechanism M1 of the same Lemma to each of the ℓ adjacent strings generated in (i.a) to create ℓ more equivalent adjacent strings of the handedness $\mathbb{H}(\ell-\alpha_1, \alpha_1)$ hereby raising their total amount to 2ℓ. The strings generated in (i.a) and (i.b) are said to belong to the *equivalence class* $\mathcal{AC}_1$.

(ii) To generate the next equivalence class $\mathcal{AC}_2 \neq \mathcal{AC}_1$, pick up an adjacent string not belonging to $\mathcal{AC}_1$ with the handedness $\mathbb{H}(\alpha_2, \ell-\alpha_2)$ out of $(\mathcal{N}_{\alpha_1}(\ell)-\ell)$ left (again, α_2 is restricted to $0 \le \alpha_2 \le \lambda$), and repeat the actions described in (i.a) and (i.b) to generate another set of 2ℓ equivalent strings. These will belong to the equivalence class $\mathcal{AC}_2$ which is distinct from $\mathcal{AC}_1$.

(iii) To generate the j-th equivalence class $\mathcal{AC}_j$, one picks up an adjacent string out of $(\mathcal{N}_{\alpha_1}(\ell)-(j-1)\ell)$ left and repeats the actions sketched in (ii).

(iv) The procedure stops once there are no adjacent strings left. Obviously, the total number of equivalence classes is

$$\frac{1}{\ell}\sum_{\alpha=0}^{\lfloor \ell/2 \rfloor} \mathcal{N}_\alpha(\ell) = (2\lambda)! \sum_{\alpha=0}^{\lambda} \binom{2\lambda+1}{\alpha} = (2\lambda)!\, 2^{2\lambda} = (2\ell-2)!! \tag{5.34}$$

This is in concert with the Lemma 5.7.

- **The case $\ell = 2\lambda$ even.**

In this case, special care should be exercised for the set of adjacent strings with the handedness $\mathbb{H}(\lambda,\lambda)$ because these adjacent strings are *self*-complementary: the mechanism M1 applied to any of those adjacent strings generates a string with the *same*, not complementary, handedness. The latter circumstance can readily be accommodated when giving a prescription for building $(2\ell-2)!!$ distinct equivalence classes of adjacent strings.

(i) First, we separate all adjacent strings with the handedness $\mathbb{H}(\alpha, \ell-\alpha)$ where $0 \le \alpha \le \lambda - 1$ and apply a procedure identically equivalent to that described for the case ℓ odd to generate distinct equivalence classes of adjacent strings. The total amount of distinct equivalence classes built in this way equals

$$L_1 = \frac{1}{\ell}\sum_{\alpha=0}^{\lfloor \ell/2 \rfloor - 1} \mathcal{N}_\alpha(\ell) = (2\lambda-1)! \sum_{\alpha=0}^{\lambda-1} \binom{2\lambda}{\alpha} = (2\ell-2)!! - \frac{(2\lambda-1)!}{2}\binom{2\lambda}{\lambda}. \tag{5.35}$$

(ii) Second, having generated in the previous step L_1 distinct equivalence classes $\{\mathcal{AC}_1, \cdots, \mathcal{AC}_{L_1}\}$ of adjacent strings, we concentrate on the adjacent strings with the handedness $\mathbb{H}(\lambda,\lambda)$ not treated so far. To this end, we (ii.a) pick up an adjacent string out of $\mathcal{N}_\lambda(2\lambda)$ with the handedness $\mathbb{H}(\lambda,\lambda)$ and perform the operations M1 and M2 to generate $2\ell = 4\lambda$ equivalent strings with the *same* handedness. The 2ℓ equivalent adjacent strings will belong to a certain equivalence class, say, $\mathcal{AC}_{L_1+1}$. (ii.b) In the next step, we pick up an adjacent string with the handedness $\mathbb{H}(\lambda,\lambda)$ out of $(\mathcal{N}_\lambda(2\lambda)-2\ell)$ left, and perform the operations detailed in (ii.a) in order to generate yet another set of 2ℓ equivalent adjacent strings belonging to an equivalence class $\mathcal{AC}_{L_1+2}$. (ii.c) We proceed further on until the last available equivalence class composed of 2ℓ adjacent strings is formed, $\mathcal{AC}_{L_1+L_2}$, where L_2 equals

$$L_2 = \frac{1}{2\ell}\mathcal{N}_\lambda(\ell) = \frac{(2\lambda-1)!}{2}\binom{2\lambda}{\lambda}. \tag{5.36}$$

Hence, for $\ell = 2\lambda$, the total number of equivalence classes for adjacent strings equals

$$L_1 + L_2 = (2\ell - 2)!! \tag{5.37}$$

as expected from the Lemma 5.7.

5.3. Integrating out all longest loop-like strings of the length ℓ

More spadework is needed to prove the Pfaffian integration theorem. Below, we will be interested in calculating the contribution $C_L(\ell)$ of longest loop-like strings (of the length ℓ) into the sought integral (5.2*b*):

$$\begin{aligned} C_L(\ell) &= \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j) \left(\mathrm{pf} \left[\begin{array}{cc} Q_n(z_i, z_j) & Q_n(z_i, \bar{z}_j) \\ Q_n(\bar{z}_i, z_j) & Q_n(\bar{z}_i, \bar{z}_j) \end{array} \right]_{2\ell\times 2\ell} \right)_{\substack{\text{longest} \\ \text{loop}-\text{like strings}}} \\ &= \frac{1}{2^\ell \ell!} \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j) \sum_{\sigma\in S'_{2\ell}} \mathrm{sgn}(\sigma) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma(2j+1)}, w_{\sigma(2j+2)}). \end{aligned} \tag{5.38}$$

In the second line of (5.38), only that part of the ordered Pfaffian expansion (5.8) appears which corresponds to a set of all loop-like strings of the length ℓ. They are accounted for by picking up proper permutations $S'_{2\ell} \subset S_{2\ell}$ in the expansion (5.8),

$$S'_{2\ell} \;\mapsto\; (\text{longest loop} - \text{like strings of the length } \ell).$$

Although, in accordance with the Lemma 5.3, the number of terms in the expansion (5.38) equals $\mathcal{N}_L(\ell) = (2\ell)!!(2\ell-2)!!$, there is no need to integrate all of them out because various loop-like strings belonging to the same equivalence class yield identical contributions. The latter observation effectively reduces the number of terms in (5.38) so that

$$C_L(\ell) = \frac{(2\ell)!!}{2^\ell \ell!} \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j) \sum_{\sigma\in S''_{2\ell}} \mathrm{sgn}(\sigma) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma(2j+1)}, w_{\sigma(2j+2)}). \tag{5.39}$$

Here, the prefactor $(2\ell)!! = 2^\ell \ell!$ equals the number of longest loop-like strings in each equivalence class; the σ-series runs over the permutations $S''_{2\ell} \subset S'_{2\ell}$ corresponding to $n_L(\ell) = (2\ell-2)!!$ longest loop-like strings, each of them being *a* representative of one distinct equivalence class,

$$S''_{2\ell} \;\mapsto\; \{\mathcal{S}_1 \in \mathcal{C}_1, \cdots, \mathcal{S}_{(2\ell-2)!!} \in \mathcal{C}_{(2\ell-2)!!}\}, \tag{5.40}$$

see the Lemma 5.3. There are $(2\ell-2)!!$ terms in (5.39).

To perform the integration explicitly, one has to reduce the longest loop-like strings in (5.39) to the form of adjacent strings as discussed in the Lemma 5.5. In accordance with the Lemma 5.7, there exist $(2\ell-2)!!$ equivalence classes $\{\mathcal{AC}_1, \cdots, \mathcal{AC}_{(2\ell-2)!!}\}$ of adjacent strings, each of them containing 2ℓ equivalent adjacent strings (see also Fig. 5). This results in the representation

$$C_L(\ell) = \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j) \sum_{\sigma\in \tilde{S}''_{2\ell}} \mathrm{sgn}(\sigma) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma(2j+1)}, w_{\sigma(2j+2)}), \tag{5.41}$$

where the σ-series runs over the permutations $\tilde{S}''_{2\ell} \subset S'_{2\ell}$ corresponding to $(2\ell-2)!!$ adjacent loop-like strings of the length ℓ, each of them being *a* representative of each one of existing equivalence classes of *adjacent strings*,

$$\tilde{S}''_{2\ell} \;\mapsto\; \{\tilde{\mathcal{S}}_1 \in \mathcal{AC}_1, \cdots, \tilde{\mathcal{S}}_{(2\ell-2)!!} \in \mathcal{AC}_{(2\ell-2)!!}\}. \tag{5.42}$$

The number of terms in (5.41) is $(2\ell-2)!!$.

To proceed, we rewrite (5.41) in a more symmetric form that treats *all* adjacent strings on the same footing:

$$C_L(\ell) = \frac{1}{2\ell} \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j) \sum_{\sigma \in S'''_{2\ell}} \text{sgn}(\sigma) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma(2j+1)}, w_{\sigma(2j+2)}). \tag{5.43}$$

Here, the σ-series runs over the permutations $S'''_{2\ell} \subset S'_{2\ell}$ corresponding to *all* adjacent loop-like strings of the length ℓ:

$$S'''_{2\ell} \mapsto \left\{ \{\tilde{\mathcal{S}}_1^{(1)}, \cdots, \tilde{\mathcal{S}}_{2\ell}^{(1)}\} \in \mathcal{AC}_1, \cdots, \{\tilde{\mathcal{S}}_1^{((2\ell-2)!!)}, \cdots, \tilde{\mathcal{S}}_{2\ell}^{((2\ell-2)!!)}\} \in \mathcal{AC}_{(2\ell-2)!!} \right\}. \tag{5.44}$$

As soon as there exist 2ℓ equivalent adjacent strings in each equivalent class $\mathcal{AC}_i$ of adjacent strings, the prefactor $(2\ell)^{-1}$ was included into (5.43) to avoid the overcounting.

An advantage of the representation (5.43) can be appreciated with the help of the Lemma 5.6. According to it, the summation over the permutations $\sigma \in S'''_{2\ell}$ can be replaced with the summation over all longest adjacent strings with a given handedness $\mathbb{H}(\alpha, \ell-\alpha)$, for all $\alpha \in (0, \ell)$:

$$\sum_{\sigma \in S'''_{2\ell}} \text{sgn}(\sigma) \prod_{j=0}^{\ell-1} Q_n(w_{\sigma(2j+1)}, w_{\sigma(2j+2)}) = \sum_{\alpha=0}^{\ell} \sum_{i=1}^{\mathcal{N}_\alpha(\ell)} \mathcal{S}_i(\alpha). \tag{5.45}$$

Here, $\mathcal{S}_i(\alpha)$ denotes the i-th adjacent string of the length $\|\mathcal{S}_i(\alpha)\| = \ell$ with the handedness $\mathbb{H}(\alpha, \ell-\alpha)$. In accordance with the Lemma 5.6,

$$\mathcal{N}_\alpha(\ell) = \ell! \binom{\ell}{\alpha}. \tag{5.46}$$

Given (5.45), the integration in (5.43) can be performed explicitly. Due to the new representation

$$C_L(\ell) = \frac{1}{2\ell} \sum_{\alpha=0}^{\ell} \sum_{i=1}^{\mathcal{N}_\alpha(\ell)} \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j)\, \mathcal{S}_i(\alpha), \tag{5.47}$$

one has to calculate the contribution of a string $\mathcal{S}_i(\alpha)$ with the handedness $\mathbb{H}(\alpha, \ell-\alpha)$ to the integral:

$$I_\ell(\alpha) = \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j)\, \mathcal{S}_i(\alpha). \tag{5.48}$$

(i) The case $\alpha = 0$ is the simplest one. Having in mind the definition (5.2*a*) and introducing an auxiliary matrix $\hat{\boldsymbol{\varsigma}}$ with the entries

$$\varsigma_{jk} = \frac{1}{2} \int_{\mathbb{C}} d\pi(z)\, q_j(z)\, q_k(\bar{z}), \tag{5.49}$$

we straightforwardly derive:

$$I_\ell(0) = \text{sgn}(\sigma_0) \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j) \left[\prod_{j=1}^{\ell-1} Q_n(\bar{z}_j, z_{j+1}) \right] Q_n(\bar{z}_\ell, z_1) = -\text{tr}_{(0,n-1)} \left[(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}})^\ell \right]. \tag{5.50}$$

Here, the permutation sign, $\text{sign}(\sigma_0)$, is $\text{sgn}(\sigma_0) = -1$ (see (5.9)), while the trace $\text{tr}_{(0,n-1)}(\cdots)$ reflects the fact that the integrated adjacent string is loop-like. Importantly, the result of the integration (5.50) does *not* depend on a particular arrangement of the arguments z_j and $\bar{z}_j$ as far as the handedness $\mathbb{H}(0,\ell)$ is kept.

(ii) The case $\alpha = \ell$ associated with the adjacent loop-like strings of the handedness $\mathbb{H}(\ell, 0)$ can be treated along the same lines to bring

$$I_\ell(\ell) = (-1)^{\ell-1} \mathrm{tr}_{(0,n-1)} \left[(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*)^\ell \right]. \tag{5.51}$$

(iii) More care should be exercised for $0 < \alpha < \ell$. In this case, two adjacent strings with the *same* handedness $\mathbb{H}(\alpha, \ell-\alpha)$ may bring *different* contributions into the integral (5.48). For instance, the adjacent string §

$$\mathcal{S}_1(\alpha) = \mathrm{sgn}(\sigma_\alpha) \left[\prod_{j=1}^{\alpha-1} Q_n(z_j, \bar{z}_{j+1}) \right] Q_n(z_\alpha, z_{\alpha+1}) \left[\prod_{j=\alpha+1}^{\ell-1} Q_n(\bar{z}_j, z_{j+1}) \right] Q_n(\bar{z}_\ell, \bar{z}_1)$$

yields the contribution

$$I_\ell^{(1)}(\alpha) = (-1)^{\alpha-1} \mathrm{tr}_{(0,n-1)} \left[(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*)^\alpha (\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}})^{\ell-\alpha} \right]. \tag{5.52}$$

At the same time, the adjacent string

$$\mathcal{S}_{\binom{\ell}{\alpha}}(\alpha) = \mathrm{sgn}(\sigma_\alpha)$$
$$\times \left[\prod_{j=1}^{\ell-\alpha-1} Q_n(\bar{z}_j, z_{j+1}) \right] Q_n(\bar{z}_{\ell-\alpha}, \bar{z}_{\ell-\alpha+1}) \left[\prod_{j=\ell-\alpha+1}^{\ell-1} Q_n(z_j, \bar{z}_{j+1}) \right] Q_n(z_\ell, z_1),$$

possessing the very same handedness, yields

$$I_\ell^{(\binom{\ell}{\alpha})}(\alpha) = (-1)^{\alpha-1} \mathrm{tr}_{(0,n-1)} \left[(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}})^{\ell-\alpha} (\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*)^\alpha \right]. \tag{5.53}$$

In deriving the results (5.52) and (5.53), we have used the fact that the permutation sign $\mathrm{sign}(\sigma_\alpha)$ is $\mathrm{sign}(\sigma_\alpha) = (-1)^{\alpha-1}$.

It can readily be seen that the adjacent strings with a given handedness $\mathbb{H}(\alpha, \ell-\alpha)$ bring all possible $\binom{\ell}{\alpha}$ contributions, or *words*, that can be represented as a trace

$$W_j(\ell, \alpha) = -\mathrm{tr}_{(0,n-1)} \underbrace{\left[\cdots (-\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*) \cdots (\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}) \cdots (-\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*) \cdots \right]}_{\alpha \text{ letters } (-\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*) \text{ and } \ell-\alpha \text{ letters } (\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}})} \tag{5.54}$$

of a product of α matrices, or *letters*, $(-\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*)$ and $\ell - \alpha$ matrices (*letters*) $(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}})$ distributed in all possible $\binom{\ell}{\alpha}$ ways. Hence, the index j in (5.54) takes the values $1 \le j \le \binom{\ell}{\alpha}$. Importantly, each word $W_j(\ell, \alpha)$ appears exactly $\ell!$ times since there always exist $\ell!$ adjacent strings $\mathcal{S}_{i_k}(\alpha)$ $(k = 1, \cdots, \ell!)$ which are related to each other by a permutation of the integration variables in (5.47). As a result, the latter is reduced to

$$C_L(\ell) = \frac{1}{2\ell}\, \ell! \sum_{\alpha=0}^{\ell} \sum_{j=1}^{\binom{\ell}{\alpha}} W_j(\ell, \alpha). \tag{5.55}$$

Spotting that

$$\sum_{\alpha=0}^{\ell} \sum_{j=1}^{\binom{\ell}{\alpha}} W_j(\ell, \alpha) = -\mathrm{tr}_{(0,n-1)} \left[(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}} - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^*)^\ell \right], \tag{5.56}$$

§ The empty products are interpreted to be 1.

we obtain

$$C_L(\ell) = -\frac{1}{2}\,(\ell-1)!\,\mathrm{tr}_{(0,n-1)}\left[(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}} - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^{*})^{\ell}\right]. \tag{5.57}$$

Finally, noticing that the matrix $(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}} - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^{*})$ under the sign of trace is related to the matrix $\hat{\boldsymbol{\upsilon}}$ defined by (5.2c) as

$$\hat{\boldsymbol{\upsilon}} = 2i\,(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}} - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\varsigma}}^{*}), \tag{5.58}$$

we arrive at the remarkably compact result

$$C_L(\ell) = -\frac{1}{2}\,\frac{(\ell-1)!}{(2i)^{\ell}}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\upsilon}}^{\ell}). \tag{5.59}$$

Hence, we have proven the following Theorem.

Theorem 5.2. *Let $d\pi(z)$ be any benign measure on $z \in \mathbb{C}$, and the function $Q_n(x,y)$ be an antisymmetric function of the form*

$$Q_n(x,y) = \frac{1}{2}\sum_{j,k=0}^{n-1} q_j(x)\,\hat{\mu}_{jk}\,q_k(y)$$

where q_j's are arbitrary polynomials of j-th order, and $\hat{\boldsymbol{\mu}}$ is an antisymmetric matrix. Then the integration formula

$$C_L(\ell) = \prod_{j=1}^{\ell}\int_{z_j\in\mathbb{C}} d\pi(z_j)\,\left(\mathrm{pf}\begin{bmatrix} Q_n(z_i,z_j) & Q_n(z_i,\bar{z}_j) \\ Q_n(\bar{z}_i,z_j) & Q_n(\bar{z}_i,\bar{z}_j)\end{bmatrix}_{2\ell\times 2\ell}\right)_{\substack{\text{longest}\\ \text{loop}-\text{like strings}}}$$
$$= -\frac{1}{2}\,\frac{(\ell-1)!}{(2i)^{\ell}}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\upsilon}}^{\ell}) \tag{5.60}$$

holds, provided the integrals on the l.h.s. exist. Here, the matrix $\hat{\boldsymbol{\upsilon}}$ is determined by the entries

$$\hat{\upsilon}_{\alpha,\beta} = i\sum_{k=0}^{n-1}\hat{\mu}_{\alpha,k}\int_{z\in\mathbb{C}} d\pi(z)\,[q_k(z)\,q_\beta(\bar{z}) - q_\beta(z)\,q_k(\bar{z})]\,.$$

The following corollary holds.

Corollary 5.3. *Consider a set of $2p$ arguments*

$$\{w_1 = z_1, w_2 = \bar{z}_1, \cdots, w_{2p-1} = z_p, w_{2p} = \bar{z}_p\}, \tag{5.61}$$

all $(2p)!$ permutations of which are denoted by S_{2p}. Take the subset $S'_{2p} \subset S_{2p}$ of S_{2p} corresponding to all loop-like strings $\mathcal{S}_i^{(p)}$ of the length $\|\mathcal{S}_i^{(p)}\| = p$,

$$\mathcal{S}_i^{(p)} = \prod_{j=0}^{p-1} Q_n(w_{\sigma_i(2j+1)}, w_{\sigma_i(2j+2)}).$$

Here, σ_i labels the i-th permutation $\sigma_i \in S'_{2p}$. The Theorem 5.2 implies:

$$C_L(p) = \frac{1}{2^p p!}\int_{\mathbb{C}}\prod_{j=1}^{p} d\pi(z_j)\sum_{\sigma\in S'_{2p}}\mathrm{sgn}(\sigma)\prod_{j=0}^{p-1} Q_n(w_{\sigma(2j+1)}, w_{\sigma(2j+2)})$$
$$= -\frac{1}{2}\frac{(p-1)!}{(2i)^p}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\upsilon}}^{p}). \tag{5.62}$$

5.4. Integrating out compound strings

Having dealt with the longest loop-like strings arising in the ordered Pfaffian expansion (see Theorem 5.2) and having extended the result of this theorem to loop-like strings of a smaller length (see Corollary 5.3), we now turn to the treatment of the remaining $(2\ell)! - (2\ell)!!(2\ell-2)!!$ strings to be referred to as *compound strings*. Our final goal here is to calculate their contribution to the integral in the l.h.s. of (5.2b).

To give a definition of a *compound string*, we remind that any given string $\mathcal{S}_i$ of the length $\|\mathcal{S}_i\| = \ell$ from the ordered Pfaffian expansion can be decomposed into a set of loop-like substrings of smaller lengths (see Lemma 5.2). According to the Lemma 5.5, each of the above loop-like substrings $\mathcal{S}_i^{(\ell_j)}$ can further be reduced to the form of an adjacent substring. This leads us to the following definition:

Definition 5.11. *A string $\mathcal{S}_i$ of the length $\|\mathcal{S}_i\| = \ell$ from the ordered Pfaffian expansion is called a compound string if it is composed of a set of adjacent substrings $\mathcal{S}_i^{(\ell_j)}$ of respective lengths $\|\mathcal{S}_i^{(\ell_j)}\| = \ell_j$ such that $\mathcal{S}_i = \bigcup_j \mathcal{S}_i^{(\ell_j)}$ with $\sum_j \ell_j = \ell$.*

Remark 5.5. (i) This definition suggests that all compound strings can be classified in accordance with all possible patterns of *unordered partitions* $\boldsymbol{\lambda}$ of the size $|\boldsymbol{\lambda}| = \ell$ of an integer ℓ:

$$\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}). \tag{5.63}$$

The frequency representation (5.63) of the partition $\boldsymbol{\lambda}$ says that the part ℓ_j appears σ_j times so that

$$\ell = \sum_{j=1}^{g} \ell_j \sigma_j. \tag{5.64}$$

Here, g is the number of inequivalent parts of the partition $\boldsymbol{\lambda}$.

(ii) Alternatively, the partition (5.63) of an integer ℓ can be represented as

$$\boldsymbol{\lambda} = (\bar{\ell}_1, \cdots, \bar{\ell}_r), \tag{5.65}$$

where the order of $\bar{\ell}_j$'s is irrelevant, and some of them can be equal to each other. Obviously,

$$\sum_{j=1}^{r} \bar{\ell}_j = \ell. \tag{5.66}$$

Here, r is the length of unordered partition $\boldsymbol{\lambda}$.

(iii) The correspondence between compound strings of the length ℓ and unordered partitions $\boldsymbol{\lambda}$ of the size $|\boldsymbol{\lambda}| = \ell$ gives rise to a topological interpretation of compound strings, which can conveniently be represented in a diagrammatic form (see Fig. 6). The diagram for a generic compound string

$$\mathcal{S}_i = \bigcup_{j=1}^{r} \mathcal{S}_i^{(\bar{\ell}_j)}, \quad \sum_{j=1}^{r} \bar{\ell}_j = \ell,$$

consists of r loops, the j-th loop depicting an adjacent substring $\mathcal{S}_i^{(\bar{\ell}_j)}$ of the length $\bar{\ell}_j$. Such a diagram will be said to belong to a topology class $\{\bar{\ell}_1, \cdots, \bar{\ell}_r\} = \{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}$.

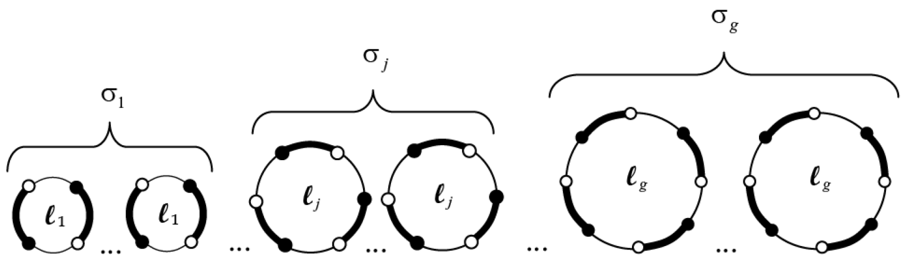

Figure 6. A diagrammatic representation of a compound string belonging to the topology class $\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g})$ with $\sum_{j=1}^{g} \ell_j \sigma_j = \ell$ and $\sum_{j=1}^{g} \sigma_j = r$. Here, g denotes the number of inequivalent parts of the partition $\boldsymbol{\lambda}$ whilst r equals the total number of loops.

Remark 5.6. The number of topologically distinct diagrams equals the number $p(\ell)$ of unordered partitions of an integer ℓ (see Remark 5.5). These are known to follow the sequence

$$p(\ell) = \{1, 2, 3, 5, 7, 11, 15, 22, 30, 42, \cdots\}. \tag{5.67}$$

An exact evaluation of $p(\ell)$ can be performed with the help of Euler's generating function (Andrews 1998)

$$\sum_{\ell=0}^{\infty} p(\ell)\, q^{\ell} = \prod_{m=1}^{\infty} \frac{1}{(1-q^m)} \equiv \frac{1}{(q)_\infty}. \tag{5.68}$$

The asymptotic behaviour of $p(\ell)$ for $\ell \gg 1$ was studied by Hardy and Ramanujan (1918),

$$p(\ell) \sim \frac{1}{4\ell\sqrt{3}} \exp\left(\pi\sqrt{2\ell/3}\right). \tag{5.69}$$

To calculate the contribution of compound strings to the integral in the l.h.s. of (5.2b), one has to determine (i) the contribution of a diagram belonging to a given topology class to the integral, and (ii) the number of diagrams within a given topology class.

Lemma 5.8. *The number of diagrams belonging to the topology class* $\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g})$ *equals*

$$\mathcal{N}_{\{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}} = \ell! \prod_{j=1}^{g} \frac{1}{(\ell_j!)^{\sigma_j}\, \sigma_j!}. \tag{5.70}$$

Proof. To determine the number $\mathcal{N}_{\{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}}$ of diagrams belonging to a given topology class $\{\bar{\ell}_1, \cdots, \bar{\ell}_r\} = \{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}$, we use the multiplication principle.

- First, we distribute the pairs of arguments $\{\bar{\ell}_1, \cdots, \bar{\ell}_r\}$ between r loops. This can be achieved by m_1 ways,

$$m_1 = \binom{\ell}{\bar{\ell}_1}\binom{\ell - \bar{\ell}_1}{\bar{\ell}_2} \cdots \binom{\ell - \bar{\ell}_1 - \cdots - \bar{\ell}_{r-2}}{\bar{\ell}_{r-1}}\binom{\ell - \bar{\ell}_1 - \cdots - \bar{\ell}_{r-1}}{\bar{\ell}_r} \times \frac{1}{r!}.$$

The factor $1/r!$ reflects the fact that the order of the r loops is irrelevant. (Indeed, a given topology class is associated with an *unordered* partition of an integer ℓ). Simple rearrangements show that m_1 simplifies down to

$$m_1 = \frac{\ell!}{\bar{\ell}_1! \cdots \bar{\ell}_r!} \frac{1}{r!}. \tag{5.71}$$

- Next, we shuffle all *inequivalent* loops (those with distinct lengths). This can be achieved by m_2 ways,

$$m_2 = \frac{(\sigma_1 + \cdots + \sigma_g)!}{\sigma_1! \cdots \sigma_g!} = \frac{r!}{\sigma_1! \cdots \sigma_g!}. \tag{5.72}$$

As a result, the total number of diagrams belonging to the topology class $\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g})$ equals

$$\mathcal{N}_{\{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}} = m_1 m_2 = \frac{\ell!}{\bar{\ell}_1! \cdots \bar{\ell}_r!} \frac{1}{\sigma_1! \cdots \sigma_g!}. \tag{5.73}$$

The observation $\prod_{j=1}^{r} \bar{\ell}_j! = \prod_{j=1}^{g} (\ell_j!)^{\sigma_j}$ ends the proof. ■

Lemma 5.9. *A diagram associated with a topology class $\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g})$ contributes*

$$C_{\{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}} = \prod_{j=1}^{g} C_L^{\sigma_j}(\ell_j) \tag{5.74}$$

to the integral in the l.h.s. of (5.2b). The function $C_L(p)$ is defined by (5.62) of the Corollary 5.3.

Proof. The above claim is a direct consequence of the Lemma 5.2, Definition 5.11, and Remark 5.5. End of proof. ■

Theorem 5.3. *All compound strings belonging to the topology class $\boldsymbol{\lambda} = (\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g})$ yield, after the integration in (5.2b), the contribution*

$$\tilde{C}_{\boldsymbol{\lambda}} = \frac{\ell!}{(2i)^\ell} \prod_{j=1}^{g} \left[\frac{1}{\ell_j^{\sigma_j} \sigma_j!} \left(-\frac{1}{2} \operatorname{tr}_{(0,n-1)}(\hat{\boldsymbol{\upsilon}}^{\ell_j}) \right)^{\sigma_j} \right]. \tag{5.75}$$

Proof. Observe that $\tilde{C}_{\boldsymbol{\lambda}} = \mathcal{N}_{\{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}} C_{\{\ell_1^{\sigma_1}, \cdots, \ell_g^{\sigma_g}\}}$, and make use of (5.62), (5.70) and (5.74) to derive (5.75). End of proof. ■

5.5. Proof of the Pfaffian integration theorem

Now we have all ingredients needed to complete the proof of the Pfaffian integration theorem announced in Section 5.1. Indeed, in accordance with the topological interpretation of the terms arising in the Pfaffian expansion (5.8), the integral on the l.h.s. of (5.2b) is given by the sum of contributions of adjacent strings of two types: the longest adjacent strings and the compound strings. The contribution of the former, $C_L(\ell)$, is given by the Theorem 5.2 while the contribution of the latter, $\tilde{C}_{\boldsymbol{\lambda}}$, is determined by the Theorem 5.3. As soon as

$$C_L(\ell) = \tilde{C}_{\boldsymbol{\lambda}=(\ell^1)},$$

one immediately concludes that the integrated Pfaffian equals the sum of $\tilde{C}_{\boldsymbol{\lambda}}$ over all unordered partitions $\boldsymbol{\lambda}$ of the size $|\boldsymbol{\lambda}| = \ell$:

$$\mathcal{I} = \int_{\mathbb{C}} \prod_{j=1}^{\ell} d\pi(z_j)\ \mathrm{pf}\left[\begin{array}{cc} Q_n(z_i,z_j) & Q_n(z_i,\bar{z}_j) \\ Q_n(\bar{z}_i,z_j) & Q_n(\bar{z}_i,\bar{z}_j)\end{array}\right]_{2\ell\times 2\ell} = \sum_{|\boldsymbol{\lambda}|=\ell} \tilde{C}_{\boldsymbol{\lambda}}$$
$$= \left(\frac{i}{2}\right)^{\ell} (-1)^{\ell}\ell! \sum_{|\boldsymbol{\lambda}|=\ell} \prod_{j=1}^{g} \left[\frac{1}{\sigma_j!}\left(-\frac{1}{2\ell_j}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\upsilon}}^{\ell_j})\right)^{\sigma_j}\right]. \tag{5.76}$$

Quite remarkably, the r.h.s. of (5.76) can be recognised to be a zonal polynomial (Macdonald 1998)

$$Z_{(1^\ell)}(p_1,\cdots,p_\ell) = (-1)^{\ell}\ell! \sum_{|\boldsymbol{\lambda}|=\ell} \prod_{j=1}^{g} \frac{1}{\sigma_j!}\left(-\frac{p_{\ell_j}}{\ell_j}\right)^{\sigma_j}$$

with the arguments

$$p_j = \frac{1}{2}\,\mathrm{tr}_{(0,n-1)}(\hat{\boldsymbol{\upsilon}}^j), \qquad j = 1,2,\cdots,\ell. \tag{5.77}$$

As a result, we conclude that

$$\mathcal{I} = \left(\frac{i}{2}\right)^{\ell} Z_{(1^\ell)}\left(\frac{1}{2}\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\upsilon}}^1,\cdots,\frac{1}{2}\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\upsilon}}^\ell\right). \tag{5.78}$$

This coincides with the statement (5.2b) of the Pfaffian integration theorem. ■

6. Probability function $p_{n,k}$: General solution and generating function

6.1. General solution

The general solution for the probability function $p_{n,k}$ of a fluctuating number of real eigenvalues in spectra of GinOE is now straightforward to derive. Indeed, it was shown in Section 3.3 that the probability function $p_{n,k}$ admits the representation

$$p_{n,k} = \frac{p_{n,n}}{\ell!}\left(\frac{2}{i}\right)^{\ell} \prod_{j=1}^{\ell} \int_{\mathrm{Im}\, z_j>0} d^2 z_j$$
$$\times\ \mathrm{erfc}\left(\frac{z_j-\bar{z}_j}{i\sqrt{2}}\right)\ \mathrm{pf}\left[\begin{array}{cc} \mathcal{D}_n(z_i,z_j) & \mathcal{D}_n(z_i,\bar{z}_j) \\ \mathcal{D}_n(\bar{z}_i,z_j) & \mathcal{D}_n(\bar{z}_i,\bar{z}_j)\end{array}\right]_{2\ell\times 2\ell} \tag{6.1}$$

with the kernel function $\mathcal{D}_n(x,y)$ given by (3.12) and (3.13).

The ℓ-fold integral in (6.1) can explicitly be performed by virtue of the Pfaffian integration theorem after the identification

$$d\pi(z) = e^{-(z^2+\bar{z}^2)/2}\mathrm{erfc}\left(\frac{z-\bar{z}}{i\sqrt{2}}\right)\theta(\mathrm{Im}\, z)\, d^2 z, \quad Q_n(x,y) = e^{(x^2+y^2)/2}\mathcal{D}_n(x,y).$$

Straightforward calculations bring

$$p_{n,k} = \frac{p_{n,n}}{\ell!}\, Z_{(1^\ell)}\left(\frac{1}{2}\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\upsilon}}^1,\cdots,\frac{1}{2}\mathrm{tr}_{(0,n-1)}\hat{\boldsymbol{\upsilon}}^\ell\right) \tag{6.2}$$

where the matrix $\hat{\boldsymbol{\upsilon}}$ is given by (5.2c). Combining the definition (5.2c) with (3.21), (3.22), (3.23), (3.25) and (3.27), one concludes that

$$\hat{\boldsymbol{\upsilon}} = 2i(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\chi}}) = \hat{\boldsymbol{\sigma}}, \tag{6.3}$$

see Appendices A and B. Finally, making use of the trace identity

$$\mathrm{tr}_{(0,n-1)} \hat{\boldsymbol{\sigma}}^j = 2\, \mathrm{tr}_{(0,\lfloor n/2\rfloor-1)} \hat{\boldsymbol{\varrho}}^j \tag{6.4}$$

proven in Appendix C, we end up with the exact formula

$$p_{n,k} = \frac{p_{n,n}}{\ell!} Z_{(1^\ell)} \left(\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)} \hat{\boldsymbol{\varrho}}^1, \cdots, \mathrm{tr}_{(0,\lfloor n/2\rfloor-1)} \hat{\boldsymbol{\varrho}}^\ell \right). \tag{6.5}$$

The entries of the matrix $\hat{\boldsymbol{\varrho}}$, calculated in Appendix C, are given by

$$\begin{aligned} \hat{\varrho}^{\mathrm{even}}_{\alpha,\beta} &= \int_0^\infty dy\, y^{2(\beta-\alpha)-1} e^{y^2} \,\mathrm{erfc}\,(y\sqrt{2}) \\ &\times \left[(2\alpha+1)\, L_{2\alpha+1}^{2(\beta-\alpha)-1}(-2y^2) + 2y^2\, L_{2\alpha-1}^{2(\beta-\alpha)+1}(-2y^2) \right] \end{aligned} \tag{6.6}$$

and

$$\hat{\varrho}^{\mathrm{odd}}_{\alpha,\beta} = \hat{\varrho}^{\mathrm{even}}_{\alpha,\beta} - (-4)^{m-\beta} \frac{m!}{(2m)!} \frac{(2\beta)!}{\beta!}\, \hat{\varrho}^{\mathrm{even}}_{\alpha,m} \tag{6.7}$$

for $n = 2m$ even and $n = 2m+1$ odd, respectively.

6.2. Generating function for $p_{n,k}$

Interestingly, the entire generating function

$$G_n(z) = \sum_{\ell=0}^{\lfloor n/2\rfloor} z^\ell p_{n,n-2\ell} \tag{6.8}$$

for the probabilities $p_{n,k}$ can explicitly be determined. To proceed, we make use of the summation formula

$$\sum_{r=0}^\infty \frac{z^r}{r!} Z_{(1^r)}(p_1, \cdots, p_r) = \exp\left(\sum_{r\ge 1} (-1)^{r-1} \frac{p_r z^r}{r} \right) \tag{6.9}$$

well known in the theory of symmetric functions (Macdonald 1998). With $p_r = \mathrm{tr}_{(0,\lfloor n/2\rfloor-1)} \hat{\boldsymbol{\varrho}}^r$, the r.h.s. transforms to

$$\begin{aligned} \exp\left(\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)} \sum_{r\ge 1} (-1)^{r-1} \frac{(z\hat{\boldsymbol{\varrho}})^r}{r} \right) &= \exp\left(\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)} \log(\hat{\mathbf{1}}_{\lfloor n/2\rfloor} + z\hat{\boldsymbol{\varrho}}) \right), \\ &= \det\left[\hat{\mathbf{1}} + z\hat{\boldsymbol{\varrho}} \right]_{\lfloor n/2\rfloor\times\lfloor n/2\rfloor}, \end{aligned} \tag{6.10}$$

resulting in an amazingly simple answer:

$$G_n(z) = \sum_{\ell=0}^{\lfloor n/2\rfloor} z^\ell p_{n,n-2\ell} = p_{n,n} \det\left[\hat{\mathbf{1}} + z\hat{\boldsymbol{\varrho}} \right]_{\lfloor n/2\rfloor\times\lfloor n/2\rfloor}. \tag{6.11}$$

6.3. Integer moments of the number of real eigenvalues

The result (6.11) allows us to formally determine any integer moment $\mathbb{E}[\mathcal{N}_r^q]$ of the fluctuating number $\mathcal{N}_r$ of real eigenvalues in the spectra of GinOE. Denoting the fluctuating number of complex eigenvalues through $2\mathcal{N}_c$, we derive:

$$\mathbb{E}[\mathcal{N}_r^q] = \mathbb{E}[(n-2\mathcal{N}_c)^q] = \sum_{j=0}^q \binom{q}{j} n^{q-j} (-2)^j \mathbb{E}[\mathcal{N}_c^j]. \tag{6.12}$$

Since

$$\mathbb{E}[\mathcal{N}_c^j] = \left(z\frac{\partial}{\partial z} \right)^j G_n(z)\Big|_{z=1}, \tag{6.13}$$

the formula (6.12) simplifies to

$$\mathbb{E}[\mathcal{N}_r^q] = p_{n,n} \left(n - 2z\frac{\partial}{\partial z} \right)^q \det \left[\hat{\mathbf{1}} + z\hat{\boldsymbol{\varrho}}\right]_{\lfloor n/2\rfloor \times \lfloor n/2 \rfloor} \Big|_{z=1}. \tag{6.14}$$

Of course, to make the formulae (6.11) and (6.14) explicit, the determinant

$$d_n(z) = \det \left[\hat{\mathbf{1}} + z\hat{\boldsymbol{\varrho}}\right]_{\lfloor n/2\rfloor \times \lfloor n/2 \rfloor}$$

has to be evaluated in a closed form. There are a few indications that this is a formidable task, but we have not succeeded in the calculation of $d_n(z)$ yet.

7. Asymptotic analysis of the probability $p_{n,n-2}$ to find exactly one pair of complex conjugated eigenvalues

To determine a qualitative behaviour of the probability function $p_{n,k}$, an asymptotic analysis of the exact solution (1.1a) is needed. In this section, the simplest probability function $p_{n,n-2}$,

$$p_{n,n-2} = 2\, p_{n,n} \int_0^\infty dy\, y\, e^{y^2} \operatorname{erfc}(y\sqrt{2})\, L_{n-2}^2(-2y^2), \tag{7.1}$$

is studied in the large-n limit. Our consideration is based on an alternative exact representation for $p_{n,n-2}$ (see the Theorem 7.1) which is more suitable for obtaining regular large-n asymptotics.

7.1. Alternative exact representation of $p_{n,n-2}$

Let us define the sequence

$$S_n = \int_0^\infty dy\, y\, e^{y^2} \operatorname{erfc}(y\sqrt{2})\, L_n^2(-2y^2), \quad n = 0, 1, \cdots, \tag{7.2}$$

such that

$$p_{n,n-2} = 2p_{n,n}\, S_{n-2}. \tag{7.3}$$

To find an exact alternative representation for S_n (and, hence, for $p_{n,n-2}$), we (i) introduce a generating function $\tau(z)$ in the form

$$\tau(z) = \sum_{n=0}^{\infty} S_n z^n \tag{7.4}$$

which is supposed to exist in some domain $\Omega_\tau \in \mathbb{R}$ of the real line $\mathbb{R}$ (to be specified later on), (ii) calculate $\tau(z)$ explicitly, and (iii) expand it back in $z \in \Omega_\tau$.

Lemma 7.1. *The generating function* $\tau(z)$ *reads:*

$$\tau(z) = \frac{1}{2(1-z^2)(1-z)} \left(-1 + \sqrt{2}\sqrt{\frac{1-z}{1-3z}} \right), \tag{7.5}$$

where

$$-1 < z < \frac{1}{3}. \tag{7.6}$$

Proof. The identity ‖

$$\sum_{n=0}^{\infty} L_n^{\lambda}(x)\, z^n = (1-z)^{-\lambda-1}\, \exp\left(\frac{xz}{z-1}\right), \quad |z|<1,$$

applied in the context of (7.2) and (7.4), gives rise to the representation

$$\tau(z) = \frac{1}{(1-z)^3} \int_0^{\infty} dy\, y\, \mathrm{erfc}\,(y\sqrt{2})\, \exp\left(y^2 \frac{1+z}{1-z}\right), \quad |z|<1. \qquad (7.7)$$

A change of the integration variable y to

$$\xi = y^2 \frac{1+z}{1-z} \qquad (7.8)$$

followed by integration by parts results in

$$\tau(z) = \frac{1}{2(1-z^2)(1-z)} \left[e^{\xi}\, \mathrm{erfc}\left(a_z \sqrt{\xi}\right) \Big|_{\xi=0}^{\xi=\infty} - \int_0^{\infty} d\xi\, e^{\xi} \frac{d}{d\xi}\, \mathrm{erfc}\left(a_z \sqrt{\xi}\right) \right], \qquad (7.9)$$

where

$$a_z = \sqrt{2}\sqrt{\frac{1-z}{1+z}}.$$

For the boundary term in (7.9) to nullify at $\xi = \infty$, the parameter z has to belong to the domain

$$\Omega_\tau : \quad -1 < z < \frac{1}{3}. \qquad (7.10)$$

Performing the remaining integral, we end up with (7.5). End of proof. ■

Having determined the generating function $\tau(z)$, we are going to Taylor-expand it around $z=0$ in order to arrive at an alternative formula for the sequence S_n. As $\tau(z)$ is a relatively simple function, we may expect that S_n obtained in this way will also have a relatively simple form.

Lemma 7.2. *The following formula holds:*

$$S_n = \frac{1}{\sqrt{2}} \sum_{j=0}^{\lfloor n/2 \rfloor} 3^{j+\alpha_n/2}\, P_{2j+\alpha_n}\left(\frac{2}{\sqrt{3}}\right) - \frac{1}{2}\Big(\lfloor n/2\rfloor + 1\Big). \qquad (7.11)$$

Here, $\alpha_n = \lceil n/2 \rceil - \lfloor n/2 \rfloor$.

Proof. To expand the function $\tau(z)$ given by (7.5) around $z=0$, we represent it in the form

$$\tau(z) = \frac{1}{2}\, \tau_1(z) \left(-1 + \sqrt{2}\, \tau_2(z)\right), \qquad (7.12)$$

where

$$\tau_1(z) = \frac{1}{(1-z^2)(1-z)}, \quad \tau_2(z) = \sqrt{\frac{1-z}{1-3z}}, \qquad (7.13)$$

and constantly use a variant of the Cauchy formula

$$\left(\sum_{k=0}^{\infty} a_k\, z^k\right) \cdot \left(\sum_{k=0}^{\infty} b_k\, z^k\right) = \sum_{n=0}^{\infty} c_n\, z^n, \quad c_n = \sum_{k=0}^{n} a_k\, b_{n-k}, \qquad (7.14)$$

‖ See Eq. (5.11.2.1) in Prudnikov, Brychkov and Marichev (1986).

where absolute convergence of the resulting series is assumed.

Expansion of $\tau_1(z)$. To determine the coefficients $c_k^{(1)}$ in the expansion

$$\tau_1(z) = \sum_{k=0}^{\infty} c_k^{(1)} z^k, \tag{7.15}$$

we notice that

$$\frac{1}{1-z} = \sum_{k=0}^{\infty} z^k, \quad \frac{1}{1-z^2} = \sum_{k=0}^{\infty} z^{2k}, \tag{7.16}$$

so that, in the notation of (7.14),

$$a_k = 1, \quad b_k = \frac{1+(-1)^k}{2}. \tag{7.17}$$

Straightforward application of the Cauchy formula yields

$$c_k^{(1)} = \sum_{j=0}^{k} a_j b_{k-j} = \frac{1}{2}\left[k + 1 + \frac{1+(-1)^k}{2}\right] = \lfloor k/2 \rfloor + 1. \tag{7.18}$$

Expansion of $\tau_2(z)$. To determine the coefficients $c_k^{(2)}$ in the expansion

$$\tau_2(z) = \sum_{k=0}^{\infty} c_k^{(2)} z^k, \tag{7.19}$$

we notice that

$$\sqrt{1-z} = \sum_{k=0}^{\infty} \binom{1/2}{k} (-z)^k, \quad \frac{1}{\sqrt{1-3z}} = \sum_{k=0}^{\infty} \binom{-1/2}{k} (-3z)^k, \tag{7.20}$$

so that, in the notation of (7.14),

$$a_k = \binom{1/2}{k} (-1)^k, \quad b_k = \binom{-1/2}{k} (-3)^k. \tag{7.21}$$

The Cauchy formula yields

$$c_k^{(2)} = \sum_{j=0}^{k} a_j b_{k-j} = 3^k \frac{(1/2)_k}{k!} \; {}_2F_1\left(-\frac{1}{2}, -k; \frac{1}{2} - k; \frac{1}{3}\right). \tag{7.22}$$

The latter can be expressed in terms of Legendre polynomials by means of the identity ¶

$${}_2F_1\left(-\frac{1}{2}, -k; \frac{1}{2} - k; w\right) = \frac{k!}{(1/2)_k} \, w^{k/2} \left[P_k\left(\frac{w+1}{2\sqrt{w}}\right) - \sqrt{w}\, P_{k-1}\left(\frac{w+1}{2\sqrt{w}}\right)\right] \tag{7.23}$$

that simplifies (7.22) to

$$c_k^{(2)} = 3^{k/2} P_k\left(\frac{2}{\sqrt{3}}\right) - 3^{(k-1)/2} P_{k-1}\left(\frac{2}{\sqrt{3}}\right). \tag{7.24}$$

¶ See Eq. (7.3.1.153) in Prudnikov, Brychkov and Marichev (1990)

Expansion of the product $\tau_1(z)\tau_2(z)$. To determine the coefficients $c_k^{(3)}$ in the expansion

$$\tau_1(z)\tau_2(z) = \sum_{k=0}^{\infty} c_k^{(3)} z^k, \tag{7.25}$$

we again use the Cauchy formula

$$c_k^{(3)} = \sum_{j=0}^{k} c_{k-j}^{(1)} c_j^{(2)}, \tag{7.26}$$

with $c_k^{(1)}$ and $c_k^{(2)}$ given by (7.18) and (7.24), respectively. Lengthy but straightforward calculations result in

$$c_k^{(3)} = \sum_{j=0}^{\lfloor k/2 \rfloor} 3^{j+\alpha_k/2} P_{2j+\alpha_k}\left(\frac{2}{\sqrt{3}}\right), \tag{7.27}$$

where $\alpha_k = \lceil k/2 \rceil - \lfloor k/2 \rfloor$.

The observation

$$S_n = \frac{1}{\sqrt{2}} c_n^{(3)} - \frac{1}{2} c_n^{(1)} \tag{7.28}$$

completes the proof. ■

Theorem 7.1. *The probability $p_{n,n-2}$ to find exactly one pair of complex conjugated eigenvalues in spectra of GinOE admits the following exact representation:*

$$p_{n,n-2} = p_{n,n} \left[\sqrt{2} \sum_{j=0}^{\lfloor n/2 \rfloor - 1} 3^{j+\alpha_n/2} P_{2j+\alpha_n}\left(\frac{2}{\sqrt{3}}\right) - \lfloor n/2 \rfloor \right]. \tag{7.29}$$

Here, $\alpha_n = \lceil n/2 \rceil - \lfloor n/2 \rfloor$, and P_n stands for Legendre polynomials.

Proof. Use the Lemma 7.2 and relation (7.3) to deduce (7.29). ■

7.2. Asymptotic analysis of $p_{n,n-2}$

The result (7.29), combined with the integral representation of Legendre polynomials

$$P_n(\phi) = \frac{1}{\pi} \int_0^{\pi} d\theta \left(\phi + \sqrt{\phi^2 - 1} \cos\theta\right)^n, \tag{7.30}$$

is particularly useful for carrying out an asymptotic analysis of the probability $p_{n,n-2}$ in the large-n limit. Indeed, (7.30) facilitates performing a summation in (7.29) leading to

$$p_{n,n-2} = p_{n,n} \left[\frac{\sqrt{2}}{\pi} \int_{-1}^{+1} \frac{dx}{\sqrt{1-x^2}} \frac{(x+2)^{\alpha_n} - (x+2)^n}{1-(x+2)^2} - \lfloor n/2 \rfloor \right]. \tag{7.31}$$

The large-n behaviour of the integral

$$J_n = \int_{-1}^{+1} \frac{dx}{\sqrt{1-x^2}} \frac{(x+2)^{\alpha_n} - (x+2)^n}{1-(x+2)^2} \tag{7.32}$$

is of our primary interest. A saddle-point analysis shows that the saddle point $x_{\rm sp} \approx -2$ lies away from the integration domain $x \in (-1, +1)$. As a result, the contribution of the end points of the integration domain, $x_{\rm L} = -1$ and $x_{\rm R} = +1$, should be examined. One can see that the close vicinity $x = 1 - \epsilon$ of $x_{\rm R} = +1$ dominates *exponentially* in $n \gg 1$. Indeed, the vicinity $\epsilon \in (0, c_0)$ yields

$$J_n \approx \int_0^{c_0} \frac{d\epsilon}{\sqrt{\epsilon(2-\epsilon)}} \frac{3^n(1-\epsilon/3)^n - 3^{\alpha_n}(1-\epsilon/3)^{\alpha_n}}{3^2(1-\epsilon/3)^2 - 1},$$

where c_0 is a proper cut-off. In the large-n limit, only a region of order n^{-1}, $\epsilon = \tau/n$, effectively contributes J_n reducing it to

$$J_n \approx \frac{3^n}{8\sqrt{2n}} \int_0^\infty \frac{d\tau}{\sqrt{\tau}}\, e^{-\tau/3} = \frac{3^{n+1/2}}{8}\sqrt{\frac{\pi}{2n}}. \tag{7.33}$$

Combined with (7.31) and (7.32), this estimate leads to the following theorem.

Theorem 7.2. *The leading large-n behaviour of the probability $p_{n,n-2}$ is given by the formula*

$$p_{n,n-2} \approx \frac{3^{n+1/2}}{8\sqrt{\pi\, n}}\, p_{n,n} \tag{7.34}$$

where $p_{n,n} = 2^{-n(n-1)/4}$.

Remark 7.1. The Theorem 7.2 implies the inequality $p_{n,n-2} \gg p_{n,n}$.

8. Correlations of complex eigenvalues of a matrix without real eigenvalues

8.1. GOE correlations in GinOE spectra

One of the earliest results on eigenlevel statistics in GinOE is due to Ginibre (1965) who spotted that spectra of random real matrices which happened to have *no complex eigenvalues* exhibit the famous GOE behaviour. Indeed, for $\boldsymbol{\mathcal{H}}_n \in \mathbb{T}(n/n)$, the j.p.d.f. (2.5) reduces to [+]

$$P_{\boldsymbol{\mathcal{H}}_n}(\lambda_1, \cdots, \lambda_n) = \frac{2^{-n(n+1)/4}}{n! \prod_{j=1}^n \Gamma(j/2)} \prod_{i>j=1}^n |\lambda_i - \lambda_j| \prod_{j=1}^n \exp(-\lambda_j^2/2). \tag{8.1}$$

The GOE spectral correlations readily follow (Mehta 2004).

8.2. GinSE-like correlations in GinOE spectra

Below, we concentrate on just the opposite case of random real matrices $\boldsymbol{\mathcal{H}}_0 \in \mathbb{T}(n/0)$ whose spectrum occasionally contains *no real eigenvalues*. The j.p.d.f. of all complex

[+] As a side remark, we notice that the explicit formula (3.2) for the probability $p_{n,n}$ can easily be derived by integrating $P_{\boldsymbol{\mathcal{H}}_n}$ over all of its arguments. Due to Selberg's integral (Mehta 2004)

$$\prod_{j=1}^n \int_{\mathbb{R}} d\lambda_j\, e^{-\lambda_j^2/2} \prod_{i>j=1}^n |\lambda_i - \lambda_j| = 2^{n/2} n! \prod_{j=1}^n \Gamma(j/2),$$

one obtains $p_{n,n} = 2^{-n(n-1)/4}$.

eigenvalues of $\boldsymbol{\mathcal{H}}_0$ can also be deduced from (2.5), the result being

$$P_{\boldsymbol{\mathcal{H}}_0}(z_1, \cdots, z_\ell) = \frac{2^{-n(n-1)/4}\, i^{n/2}}{(n/2)!\, \prod_{j=1}^{n} \Gamma(j/2)} \prod_{i>j=1}^{\ell} |z_i - z_j|^2 |z_i - \bar{z}_j|^2 \\ \times \prod_{j=1}^{\ell} (\bar{z}_j - z_j)\, \mathrm{erfc}\left(\frac{z_j - \bar{z}_j}{i\sqrt{2}}\right) \exp\left(-\frac{z_j^2 + \bar{z}_j^2}{2}\right) \tag{8.2}$$

with n even, $n = 2\ell$. Remarkably, while the above j.p.d.f. resembles the j.p.d.f. of complex eigenvalues in GinSE (2.3), it is manifestly different from the latter.

Is it possible to determine the correlation functions for the new complex eigenvalue model (8.2)? The answer is positive.

Lemma 8.1. *Let $\boldsymbol{\mathcal{H}}_0$ be an $n \times n$ random real matrix with no real eigenvalues such that its entries are statistically independent random variables picked from a normal distribution* $\mathsf{N}(0,1)$*. Then, the p-point correlation function ($1 \le p \le \ell$) of its complex eigenvalues equals*

$$R_{0,p}^{(\boldsymbol{\mathcal{H}}_0)}(z_1, \cdots, z_p; n) = p_{n,n} \frac{\prod_{j=0}^{\ell-1} r_j}{\prod_{j=1}^{n} \Gamma(j/2)} \prod_{j=1}^{p} \mathrm{erfc}\left(\frac{z_j - \bar{z}_j}{i\sqrt{2}}\right) \exp\left(-\frac{z_j^2 + \bar{z}_j^2}{2}\right) \\ \times \mathrm{pf} \begin{bmatrix} \kappa_\ell(z_i, z_j) & \kappa_\ell(z_i, \bar{z}_j) \\ \kappa_\ell(\bar{z}_i, z_j) & \kappa_\ell(\bar{z}_i, \bar{z}_j) \end{bmatrix}_{2p \times 2p}. \tag{8.3}$$

Here, $n = 2\ell$ and the 'pre-kernel' κ_ℓ is

$$\kappa_\ell(z, z') = i \sum_{j=0}^{\ell-1} \frac{1}{r_j} \Big[p_{2j}(z) p_{2j+1}(z') - p_{2j}(z') p_{2j+1}(z) \Big]. \tag{8.4}$$

The polynomials $p_j(z)$ in (8.4) are skew orthogonal in the complex half-plane $\mathrm{Im}\, z > 0$,

$$\langle p_{2j+1}, p_{2k} \rangle_{\mathrm{c}} = -\langle p_{2k}, p_{2j+1} \rangle_{\mathrm{c}} = i\, r_j\, \delta_{jk}, \tag{8.5}$$

$$\langle p_{2j+1}, p_{2k+1} \rangle_{\mathrm{c}} = \langle p_{2j}, p_{2k} \rangle_{\mathrm{c}} = 0, \tag{8.6}$$

with respect to the skew product

$$\langle f, g \rangle_{\mathrm{c}} = \int_{\mathrm{Im}\, z > 0} d^2 z\, \mathrm{erfc}\left(\frac{z - \bar{z}}{i\sqrt{2}}\right) \exp\left(-\frac{z^2 + \bar{z}^2}{2}\right) \left[f(z) g(\bar{z}) - f(\bar{z}) g(z) \right]. \tag{8.7}$$

Proof. By definition (2.7), the p-point correlation function is

$$R_{0,p}^{(\boldsymbol{\mathcal{H}}_0)}(z_1, \cdots, z_p; n) = \frac{\ell!}{(\ell - p)!} \prod_{j=p+1}^{\ell} \int_{\mathrm{Im} z_j > 0} d^2 z_j\, P_{\boldsymbol{\mathcal{H}}_0}(z_1, \cdots, z_\ell). \tag{8.8}$$

Since the $n \times n$ real matrix with no real eigenvalues has ℓ pairs of complex conjugated eigenvalues, it holds that $n = 2\ell$.

Conceptually, the proof to be presented consists of three parts. First, we concentrate on the j.p.d.f. $P_{\boldsymbol{\mathcal{H}}_0}(z_1, \cdots, z_\ell)$ and show that it can be represented in terms of a certain quaternion determinant. Second, we prove that the quaternion matrix under the quaternion determinant enjoys the projection property (see Definition 2.1). Third, we apply the Dyson integration theorem to carry out all $(\ell - p)$ integrations in (8.8).

Part 1.—As the Vandermonde structure of (8.2) mimics that of GinSE (2.3), it is tempting to employ the identity

$$\prod_{i>j=1}^{\ell} |z_i - z_j|^2 |z_i - \bar{z}_j|^2 \prod_{j=1}^{\ell} (\bar{z}_j - z_j) = \det \left[z_j^{i-1},\, \bar{z}_j^{i-1} \right]_{\substack{i=1,\cdots,2\ell \\ j=1,\cdots,\ell}} \tag{8.9}$$

that helps us reduce the j.p.d.f. $P_{\mathcal{H}_0}$ to the form

$$\begin{aligned} P_{\mathcal{H}_0}(z_1, \cdots, z_\ell) &= \frac{2^{-n(n-1)/4}\, i^\ell}{(n/2)! \prod_{j=1}^{n} \Gamma(j/2)} \det \left[p_{i-1}(z_j),\, p_{i-1}(\bar{z}_j) \right]_{\substack{i=1,\cdots,2\ell \\ j=1,\cdots,\ell}} \\ &\times \prod_{j=1}^{\ell} \operatorname{erfc}\left(\frac{z_j - \bar{z}_j}{i\sqrt{2}} \right) \exp\left(-\frac{z_j^2 + \bar{z}_j^2}{2} \right). \end{aligned} \tag{8.10}$$

Here, $p_k(z)$ are *arbitrary* monic polynomials of degree k.

The very structure of the matrix under the determinant in (8.10),

$$\det \begin{bmatrix} p_0(z_1) & p_0(\bar{z}_1) & \cdots & \cdots & p_0(z_\ell) & p_0(\bar{z}_\ell) \\ p_1(z_1) & p_1(\bar{z}_1) & \cdots & \cdots & p_1(z_\ell) & p_1(\bar{z}_\ell) \\ \vdots & \vdots & & & \vdots & \vdots \\ \vdots & \vdots & & & \vdots & \vdots \\ p_{2\ell-2}(z_1) & p_{2\ell-2}(\bar{z}_1) & \cdots & \cdots & p_{2\ell-2}(z_\ell) & p_{2\ell-2}(\bar{z}_\ell) \\ p_{2\ell-1}(z_1) & p_{2\ell-1}(\bar{z}_1) & \cdots & \cdots & p_{2\ell-1}(z_\ell) & p_{2\ell-1}(\bar{z}_\ell) \end{bmatrix}$$

suggests that we introduce a set of quaternions $\{\boldsymbol{\psi}_0(z), \cdots, \boldsymbol{\psi}_{\ell-1}(z)\}$,

$$\begin{aligned} \boldsymbol{\psi}_j(z) &= \frac{p_{2j}(z) + p_{2j+1}(\bar{z})}{2}\, \hat{\boldsymbol{e}}_0 + \frac{p_{2j}(z) - p_{2j+1}(\bar{z})}{2i}\, \hat{\boldsymbol{e}}_1 \\ &+ \frac{p_{2j}(\bar{z}) - p_{2j+1}(z)}{2}\, \hat{\boldsymbol{e}}_2 + \frac{p_{2j}(\bar{z}) + p_{2j+1}(z)}{2i}\, \hat{\boldsymbol{e}}_3 \end{aligned} \tag{8.11}$$

whose 2×2 matrix representation reads *

$$\Theta[\boldsymbol{\psi}_j(z)] = \begin{bmatrix} p_{2j}(z) & p_{2j}(\bar{z}) \\ p_{2j+1}(z) & p_{2j+1}(\bar{z}) \end{bmatrix}. \tag{8.12}$$

As a result, the above determinant can equivalently be written as

$$\det \left[p_{i-1}(z_j),\, p_{i-1}(\bar{z}_j) \right]_{\substack{i=1,\cdots,2\ell \\ j=1,\cdots,\ell}} = \det \left[\Theta[\boldsymbol{\psi}_{i-1}(z_j)] \right]_{\substack{i=1,\cdots,\ell \\ j=1,\cdots,\ell}}. \tag{8.13}$$

The latter can be put into a quaternion determinant form (see Corollary 5.1.3 in Mehta's book (2004)):

$$\det \left[\Theta[\boldsymbol{\psi}_{i-1}(z_j)] \right]_{\substack{i=1,\cdots,\ell \\ j=1,\cdots,\ell}} = \operatorname{qdet} \left[\boldsymbol{\mathcal{A}} \bar{\boldsymbol{\mathcal{A}}} \right]_{\ell \times \ell} = \operatorname{qdet} \left[\bar{\boldsymbol{\mathcal{A}}} \boldsymbol{\mathcal{A}} \right]_{\ell \times \ell}. \tag{8.14}$$

Here, $\boldsymbol{\mathcal{A}}$ is an $\ell \times \ell$ quaternion matrix with the entries

$$\boldsymbol{\mathcal{A}}_{ij} = \boldsymbol{\psi}_{i-1}(z_j) \tag{8.15}$$

* The 2×2 matrices $\hat{\boldsymbol{e}}_j$ are defined as follows:

$$\hat{\boldsymbol{e}}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \hat{\boldsymbol{e}}_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \hat{\boldsymbol{e}}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \hat{\boldsymbol{e}}_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

and $\bar{\boldsymbol{\mathcal{A}}}$ is the dual quaternion matrix whose entries

$$\bar{\boldsymbol{\mathcal{A}}}_{ij} = \bar{\boldsymbol{\psi}}_{j-1}(z_i) \tag{8.16}$$

are determined by the dual quaternion $\bar{\boldsymbol{\psi}}_j(z)$,

$$\begin{aligned} \bar{\boldsymbol{\psi}}_j(z) = & \frac{p_{2j}(z) + p_{2j+1}(\bar{z})}{2}\, \hat{\boldsymbol{e}}_0 - \frac{p_{2j}(z) - p_{2j+1}(\bar{z})}{2i}\, \hat{\boldsymbol{e}}_1 \\ & - \frac{p_{2j}(\bar{z}) - p_{2j+1}(z)}{2}\, \hat{\boldsymbol{e}}_2 - \frac{p_{2j}(\bar{z}) + p_{2j+1}(z)}{2i}\, \hat{\boldsymbol{e}}_3 \end{aligned} \tag{8.17}$$

such that

$$\Theta[\bar{\boldsymbol{\psi}}_j(z)] = \hat{\boldsymbol{e}}_2\, \Theta[\boldsymbol{\psi}_j(z)]^{\mathrm{T}}\, \hat{\boldsymbol{e}}_2^{-1} = \left[\begin{array}{cc} p_{2j+1}(\bar{z}) & -p_{2j}(\bar{z}) \\ -p_{2j+1}(z) & p_{2j}(z) \end{array} \right]. \tag{8.18}$$

Combining (8.13) and (8.14) into

$$\det\, [p_{i-1}(z_j),\, p_{i-1}(\bar{z}_j)]_{\substack{i=1,\cdots,2\ell \\ j=1,\cdots,\ell}} = \mathrm{qdet}\, \left[\bar{\boldsymbol{\mathcal{A}}}\boldsymbol{\mathcal{A}}\right]_{\ell\times\ell}, \tag{8.19}$$

we obtain

$$\det\, [p_{i-1}(z_j),\, p_{i-1}(\bar{z}_j)]_{\substack{i=1,\cdots,2\ell \\ j=1,\cdots,\ell}} = i^{-\ell} \prod_{j=0}^{\ell-1} r_j\; \mathrm{qdet}\, [\hat{\boldsymbol{\kappa}}_\ell(z_i, z_j)]_{\ell\times\ell}\,. \tag{8.20}$$

Here, the (self-dual) *quaternion kernel* $\hat{\boldsymbol{\kappa}}_\ell$ admits a 2×2 matrix representation

$$\Theta\, [\hat{\boldsymbol{\kappa}}_\ell(z_i, z_j)] = \left(\begin{array}{cc} -\kappa_\ell(\bar{z}_i, z_j) & -\kappa_\ell(\bar{z}_i, \bar{z}_j) \\ \kappa_\ell(z_i, z_j) & \kappa_\ell(z_i, \bar{z}_j) \end{array} \right) \tag{8.21}$$

with

$$\kappa_\ell(z, w) = i \sum_{j=0}^{\ell-1} \frac{1}{r_j} \Big[p_{2j}(z) p_{2j+1}(w) - p_{2j}(w) p_{2j+1}(z) \Big]. \tag{8.22}$$

The set of constants $\{r_j\}$ is not fixed so far.

The above consideration results in the following expression for the j.p.d.f. $P_{\boldsymbol{\mathcal{H}}_0}$ [Eq. (8.10)]

$$\begin{aligned} P_{\boldsymbol{\mathcal{H}}_0}(z_1, \cdots, z_\ell) \;\; & = \frac{2^{-n(n-1)/4}}{(n/2)!}\, \frac{\prod_{j=0}^{\ell-1} r_j}{\prod_{j=1}^{n} \Gamma(j/2)}\, \mathrm{qdet}\, [\hat{\boldsymbol{\kappa}}_\ell(z_i, z_j)]_{\ell\times\ell} \\ & \times \prod_{j=1}^{\ell} \mathrm{erfc} \left(\frac{z_j - \bar{z}_j}{i\sqrt{2}} \right)\, \exp\left(-\frac{z_j^2 + \bar{z}_j^2}{2} \right), \end{aligned} \tag{8.23}$$

that serves as a proper starting point for evaluating the p-point correlation function $R_{0,p}^{(\boldsymbol{\mathcal{H}}_0)}(z_1, \cdots, z_p; n)$ specified in Eq. (8.3).

Part 2.—Now, we are going to prove that the quaternion $\hat{\boldsymbol{\kappa}}_\ell$ satisfies the projection property Definition 2.1. To simplify the consideration to follow, we set the so far arbitrary monic polynomials $p_k(z)$ to be skew-orthogonal in the complex half-plane, $\mathrm{Im}\, z > 0$, as defined by (8.5) – (8.7).

Having imposed the skew-orthogonality on $p_k(z)$, with $r_j \in \mathbb{R}$, we are in the position to verify whether or not the projection property for the quaternion $\hat{\boldsymbol{\kappa}}_\ell$ is fulfilled. In accordance with the Definition 2.1, one has to consider the integral

$$I = \int_{\mathrm{Im}\, w > 0} d^2 w\, \mathrm{erfc} \left(\frac{w - \bar{w}}{i\sqrt{2}} \right)\, \exp\left(-\frac{w^2 + \bar{w}^2}{2} \right)\, \Theta\, [\hat{\boldsymbol{\kappa}}_\ell(z_1, w)]\; \Theta\, [\hat{\boldsymbol{\kappa}}_\ell(w, z_2)] \tag{8.24}$$

that equals

$$I = \left(\begin{array}{cc} \delta_\ell(\bar{z}_1, z_2) & \delta_\ell(\bar{z}_1, \bar{z}_2) \\ -\delta_\ell(z_1, z_2) & -\delta_\ell(z_1, \bar{z}_2) \end{array} \right). \tag{8.25}$$

Here, the function $\delta_\ell(z_1, z_2)$ is defined by the integral

$$\delta_\ell(z_1, z_2) = \int_{\mathrm{Im}\, w>0} d^2 w \,\mathrm{erfc}\left(\frac{w-\bar{w}}{i\sqrt{2}}\right) \exp\left(-\frac{w^2+\bar{w}^2}{2}\right) \times \left[\kappa_\ell(z_1, w)\kappa_\ell(\bar{w}, z_2) - \kappa_\ell(z_2, w)\kappa_\ell(\bar{w}, z_1)\right]. \tag{8.26}$$

Its evaluation, based on (8.22), (8.7), (8.5) and (8.6), is straightforward, the result being

$$\delta_\ell(z_1, z_2) = -\kappa_\ell(z_1, z_2) \tag{8.27}$$

so that

$$I = \Theta\left[\hat{\boldsymbol{\kappa}}_\ell(z_1, z_2)\right]. \tag{8.28}$$

Put differently, the *quaternion kernel* $\hat{\boldsymbol{\kappa}}_\ell$ satisfies the projection property in the form

$$\int_{\mathrm{Im}\, w>0} d^2 w \,\mathrm{erfc}\left(\frac{w-\bar{w}}{i\sqrt{2}}\right) \exp\left(-\frac{w^2+\bar{w}^2}{2}\right) \hat{\boldsymbol{\kappa}}_\ell(z_1, w)\, \hat{\boldsymbol{\kappa}}_\ell(w, z_2) = \hat{\boldsymbol{\kappa}}_\ell(z_1, z_2). \tag{8.29}$$

This is precisely (2.8b) of the Dyson integration theorem with a quaternion $\lambda = 0$.

Part 3.—The above proof of the projection property for the quaternion kernel $\hat{\boldsymbol{\kappa}}_\ell$ in (8.23) paves the way for carrying out the $(\ell - p)$ integrations in (8.8). Indeed, the integrations therein can be performed by virtue of the Dyson integration theorem (Theorem 2.1) since

$$\Theta\left[\bar{\hat{\boldsymbol{\kappa}}}_\ell(z_1, z_2)\right] \stackrel{\mathrm{def}}{=} \hat{\boldsymbol{e}}_2\, \Theta\left[\hat{\boldsymbol{\kappa}}_\ell(z_1, z_2)\right]^{\mathrm{T}} \hat{\boldsymbol{e}}_2^{-1} = \Theta\left[\hat{\boldsymbol{\kappa}}_\ell(z_2, z_1)\right]. \tag{8.30}$$

To this end, one has to determine the constant c defined by the integral [see (2.8d)]

$$c\, \boldsymbol{e}_0 = \int_{\mathrm{Im}\, z>0} d^2 z \,\mathrm{erfc}\left(\frac{z-\bar{z}}{i\sqrt{2}}\right) \exp\left(-\frac{z^2+\bar{z}^2}{2}\right) \hat{\boldsymbol{\kappa}}_\ell(z, z) \tag{8.31}$$

yielding

$$c = \int_{\mathrm{Im}\, z>0} d^2 z \,\mathrm{erfc}\left(\frac{z-\bar{z}}{i\sqrt{2}}\right) \exp\left(-\frac{z^2+\bar{z}^2}{2}\right) \kappa_\ell(z, \bar{z}) = \ell. \tag{8.32}$$

The projection property (8.29) combined with the result (8.32) brings the key integration identity:

$$\frac{1}{(\ell-p)!} \prod_{j=p+1}^{\ell} \int_{\mathrm{Im}\, z_j>0} d^2 z_j \,\mathrm{erfc}\left(\frac{z_j-\bar{z}_j}{i\sqrt{2}}\right) \exp\left(-\frac{z_j^2+\bar{z}_j^2}{2}\right) \mathrm{qdet}\left[\hat{\boldsymbol{\kappa}}_\ell(z_i, z_j)\right]_{\ell\times\ell} = \mathrm{qdet}\left[\hat{\boldsymbol{\kappa}}_\ell(z_i, z_j)\right]_{p\times p} = \mathrm{pf}\left[\begin{array}{cc} \kappa_\ell(z_i, z_j) & \kappa_\ell(z_i, \bar{z}_j) \\ \kappa_\ell(\bar{z}_i, z_j) & \kappa_\ell(\bar{z}_i, \bar{z}_j) \end{array} \right]_{2p\times 2p}. \tag{8.33}$$

Applied to (8.8) and (8.23), it results in

$$R_{0,p}^{(\boldsymbol{\mathcal{H}}_k)}(z_1, \cdots, z_p; n) = p_{n,n} \frac{\prod_{j=0}^{\ell-1} r_j}{\prod_{j=1}^{n} \Gamma(j/2)} \prod_{j=1}^{p} \mathrm{erfc}\left(\frac{z_j-\bar{z}_j}{i\sqrt{2}}\right) \exp\left(-\frac{z_j^2+\bar{z}_j^2}{2}\right) \times \mathrm{pf}\left[\begin{array}{cc} \kappa_\ell(z_i, z_j) & \kappa_\ell(z_i, \bar{z}_j) \\ \kappa_\ell(\bar{z}_i, z_j) & \kappa_\ell(\bar{z}_i, \bar{z}_j) \end{array} \right]_{2p\times 2p}, \tag{8.34}$$

where $n = 2\ell$. This completes the proof of the Lemma 8.1. ■

8.3. Probability $p_{n,0}$ to find no real eigenvalues

The technique exposed in the previous subsection allows us to establish the following structural result for the probability to find no real eigenvalues in GinOE spectra.

Corollary 8.1. *The probability $p_{n,0}$ to find no real eigenvalues in spectra of GinOE equals*

$$p_{n,0} = \frac{p_{n,n}}{\prod_{j=1}^{n} \Gamma(j/2)} \prod_{j=0}^{\ell-1} r_j. \tag{8.35}$$

Here, $n = 2\ell$; the constants r_j are defined by (8.5).

Proof. For $n = 2\ell$ even, the definition (3.1) translates to

$$p_{n,0} = \prod_{j=1}^{\ell} \int_{\mathrm{Im}\, z_j > 0} d^2 z_j \, P_{\boldsymbol{\mathcal{H}}_0}(z_1, \cdots, z_\ell). \tag{8.36}$$

Given $P_{\boldsymbol{\mathcal{H}}_0}$ in the form (8.23), the above probability reduces to

$$p_{n,0} = \frac{p_{n,n}}{\ell!} \frac{\prod_{j=0}^{\ell-1} r_j}{\prod_{j=1}^{n} \Gamma(j/2)} \prod_{j=1}^{\ell} \int_{\mathrm{Im}\, z_j > 0} d^2 z_j \; \mathrm{qdet} \left[\hat{\boldsymbol{\kappa}}_\ell(z_i, z_j) \right]_{\ell \times \ell} \times \prod_{j=1}^{\ell} \mathrm{erfc} \left(\frac{z_j - \bar{z}_j}{i\sqrt{2}} \right) \exp \left(-\frac{z_j^2 + \bar{z}_j^2}{2} \right). \tag{8.37}$$

Due to the projection property (8.29) of the quaternion kernel $\hat{\boldsymbol{\kappa}}_\ell$, the integration in (8.37) can readily be performed [see, e.g., the identity (8.33)] bringing

$$\prod_{j=1}^{\ell} \int_{\mathrm{Im}\, z_j > 0} d^2 z_j \; \mathrm{qdet} \left[\hat{\boldsymbol{\kappa}}_\ell(z_i, z_j) \right]_{\ell \times \ell} \prod_{j=1}^{\ell} \mathrm{erfc} \left(\frac{z_j - \bar{z}_j}{i\sqrt{2}} \right) \exp \left(-\frac{z_j^2 + \bar{z}_j^2}{2} \right) = \ell! \tag{8.38}$$

This establishes the result (8.35) thus completing the proof. ■

Remark 8.1. To make the formula (8.35) explicit, one has to know the normalisation constants r_j which we have failed to determine explicitly so far.

9. Conclusions and open problems

To summarise, an exact formula was derived for the probability $p_{n,k}$ to find precisely k real eigenvalues in the spectrum of an $n \times n$ random matrix drawn from GinOE. Based on the Pfaffian integration theorem (that can be seen as an extension of the Dyson integration theorem to kernels that do not possess the projection property), the solution found expresses the probability function $p_{n,k}$ in terms of zonal polynomials ♯. This links the integrable structure of GinOE to the theory of symmetric functions (Macdonald 1998). Undoubtedly, much more effort is needed to accomplish the spectral theory of GinOE. Below, we list some of the open problems that have to be addressed.

♯ To the best of our knowledge, this is the first ever random-matrix-theory *observable* admitting a representation in terms of symmetric functions.

(I) Probability Function $p_{n,k}$ and Associated Generating Function $G_n(z)$. The exact solution for the probability $p_{n,k}$ expresses the probability in terms of zonal polynomials of some complicated arguments. Does a more explicit representation (like the one for $p_{n,n-2}$, see (7.1) and (7.29)) exist for $p_{n,k}$? A similar question arises in the context of the solution for the determinantal generating function $G_n(z)$ given by (6.11). Can the determinant (6.11) be calculated in a closed form as a function of k, n and z? We have a few indications that this is a formidable task.

(II) Asymptotic Analysis of $p_{n,k}$ in the Large-n Limit. The problem we wish to pose here concerns the large-n behaviour of $p_{n,k}$ in various scaling limits: (a) $k \sim n^0$, (b) $n - k \sim n^0$, and (c) $k \sim n^{1/2}$ (this scaling is prompted by the result (2.28) for an average number of real eigenvalues). If available, the large-n formulae of this kind would facilitate a comparison of our exact theory with existing numerical and experimental data, as reported † by Halasz *et al* (1997) and Kwapień *et al* (2000).

(III) Correlation Functions of Complex Eigenvalues and $p_{n,0}$. The solution presented for these two spectral characteristics involves specific polynomials which are skew orthogonal in the complex plane with respect to a somewhat unusual weight function containing the complementary error function. Can those be determined explicitly to eventually bring closed formulae for both the correlation functions and the probability $p_{n,0}$? Their large-n analysis would be of great interest, too.

(IV) Correlation Functions of Both Real and Complex Eigenvalues and a Generalised Pfaffian Integration Theorem. The calculation of all partial (p,q)-point correlation functions $R_{p,q}^{(\boldsymbol{\mathcal{H}}_k)}$ for GinOE matrices with a given number (k) of real eigenvalues (defined as an integral of the j.p.d.f. over all but p real and q complex eigenvalues, see (2.7)) is yet another important problem to tackle. Also, can the unconditional (p,q)-point correlation functions $R_{p,q} = \sum_k R_{p,q}^{(\boldsymbol{\mathcal{H}}_k)}$ be explicitly determined? We believe that progress in this direction can be achieved through a proper extension of the Pfaffian integration theorem:

$$\prod_{j=p+1}^{\ell} \int_{z_j \in \mathbb{C}} d\pi(z_j) \ \mathrm{pf} \left[\begin{array}{cc} Q_n(z_i, z_j) & Q_n(z_i, \bar{z}_j) \\ Q_n(\bar{z}_i, z_j) & Q_n(\bar{z}_i, \bar{z}_j) \end{array} \right]_{2\ell \times 2\ell} = ? \tag{9.1}$$

Here, the notation of the Theorem 5.1 was used. Notice that one should not assume the projection property for the kernel Q_n.

The above list of open problems calls for further research of GinOE that will eventually unveil the rich mathematical structures underlying this classical but still largely unexplored non-Hermitean random matrix model.

Acknowledgements

The authors appreciate a clarifying correspondence with A. Borodin (Caltech), V. B. Kuznetsov (Leeds) and G. Olshanski (IITP, Moscow) and useful discussions with V. Al. Osipov (H.I.T.). A part of this work was done during the visits to Brunel University West London (E.K.), University of Warwick (E.K. and G.A.), and H.I.T.

† The numerics by Halasz *et al* (1997) refers to the chiral counterparts of the Ginibre ensembles.

– Holon Institute of Technology (G.A.) supported by a BRIEF grant from Brunel University. The work of G.A. is supported by EU network ENRAGE MRTN-CT-2004-005616 and by EPSRC grant EP/D031613/1. The research of E.K. is supported by the Israel Science Foundation through the grant No 286/04.

Appendix A. The $n \times n$ matrix $\hat{\boldsymbol{\chi}}$

A1. Symmetries of $\hat{\boldsymbol{\chi}}$ *and useful expansions.*—The definition (3.27) of the matrix $\hat{\boldsymbol{\chi}}$ suggests that it is antisymmetric,

$$\hat{\chi}_{jk} = -\hat{\chi}_{kj}, \tag{A.1}$$

and purely imaginary

$$\bar{\hat{\chi}}_{jk} = -\hat{\chi}_{jk} \tag{A.2}$$

as confirmed by the formula

$$\hat{\chi}_{jk} = i \int d\alpha(w)\, e^{-(w^2+\bar{w}^2)/2}\, \left[\operatorname{Re} q_j(w) \operatorname{Im} q_k(\bar{w}) + \operatorname{Im} q_j(w) \operatorname{Re} q_k(\bar{w}) \right]. \tag{A.3}$$

Due to (3.16) that relates the skew orthogonal polynomials $q_j(w)$ to Hermite polynomials $H_j(w)$, the following expansions‡ are useful for even/odd order Hermite polynomials taken at $w = x + iy$:

$$\operatorname{Re} H_{2m}(w) = \sum_{j=0}^{m} \frac{(-1)^j 2^{2j}}{(2j)!} (-2m)_{2j}\, y^{2j} H_{2m-2j}(x), \tag{A.4}$$

$$\operatorname{Im} H_{2m}(w) = \sum_{j=0}^{m-1} \frac{(-1)^{j-1} 2^{2j+1}}{(2j+1)!} (-2m)_{2j+1}\, y^{2j+1} H_{2m-2j-1}(x), \tag{A.5}$$

and

$$\operatorname{Re} H_{2m+1}(w) = \sum_{j=0}^{m} \frac{(-1)^j 2^{2j}}{(2j)!} (-2m-1)_{2j}\, y^{2j} H_{2m+1-2j}(x), \tag{A.6}$$

$$\operatorname{Im} H_{2m+1}(w) = \sum_{j=0}^{m} \frac{(-1)^{j-1} 2^{2j+1}}{(2j+1)!} (-2m-1)_{2j+1}\, y^{2j+1} H_{2m-2j}(x). \tag{A.7}$$

Here, $(-a)_n$ is the Pochhammer symbol ($a > 0$),

$$(-a)_n = (-1)^n \frac{\Gamma(a+1)}{\Gamma(a-n+1)}.$$

A2. Calculation of the matrix elements $\hat{\chi}_{2\alpha,2\beta}$ *and* $\hat{\chi}_{2\alpha+1,2\beta+1}$.—We claim that

$$\hat{\chi}_{2\alpha,2\beta} = \hat{\chi}_{2\alpha+1,2\beta+1} = 0 \tag{A.8}$$

for all $\alpha = 0, 1, \cdots$ and $\beta = 0, 1, \cdots$. To prove this statement, let us consider $\hat{\chi}_{2\alpha,2\beta}$. In accordance with (A.3), this matrix element is related to the integral containing

$$e^{-(w^2+\bar{w}^2)/2} \Big[\operatorname{Re} q_{2\alpha}(w) \operatorname{Im} q_{2\beta}(\bar{w}) + \operatorname{Im} q_{2\alpha}(w) \operatorname{Re} q_{2\beta}(\bar{w}) \Big].$$

Having in mind (3.16), (A.4) and (A.5), we observe that the above expression, being integrated over the x-part of the integration measure

$$d\alpha(w) = dx\, \cdot \operatorname{erfc}(y\sqrt{2})\, \theta(y)\, dy, \tag{A.9}$$

nullifies due to the products of Hermite polynomials $H_{\mathrm{even}\,\natural}(x)\, H_{\mathrm{odd}\,\natural}(x)$ of even and odd orders. Thus, we conclude that $\hat{\chi}_{2\alpha,2\beta} = 0$. By the same token, the matrix element $\hat{\chi}_{2\alpha+1,2\beta+1}$ is zero as well.

‡ http://functions.wolfram.com/05.01.19.0001.01 and .../05.01.19.0002.01

A3. Calculation of the matrix elements $\hat{\chi}_{2\alpha+1,2\beta}$ *and* $\hat{\chi}_{2\alpha,2\beta+1}$.—We claim that

$$\hat{\chi}_{2\alpha+1,2\beta} = \frac{i}{2}(-1)^{\beta-\alpha} h_\alpha \int_0^\infty dy\, y^{2(\beta-\alpha)-1} e^{y^2} \operatorname{erfc}(y\sqrt{2}) \times \left[(2\alpha+1) L_{2\alpha+1}^{2(\beta-\alpha)-1}(-2y^2) + 2y^2 L_{2\alpha-1}^{2(\beta-\alpha)+1}(-2y^2)\right]. \quad (A.10)$$

In accordance with the symmetry relation (A.1), it holds

$$\hat{\chi}_{2\alpha,2\beta+1} = -\hat{\chi}_{2\beta+1,2\alpha}. \quad (A.11)$$

To prove the result (A.10), we start with the definition (A.3),

$$\hat{\chi}_{2\alpha+1,2\beta} = i \int d\alpha(w)\, e^{-(w^2+\bar{w}^2)/2} \times \left[\operatorname{Re} q_{2\alpha+1}(w) \operatorname{Im} q_{2\beta}(\bar{w}) + \operatorname{Im} q_{2\alpha+1}(w) \operatorname{Re} q_{2\beta}(\bar{w})\right], \quad (A.12)$$

where the skew orthogonal polynomials $q_j(w)$ are given by (3.16). Substituting them into (A.12), representing w as $w = x + iy$ and making use of the expansions (A.4) to (A.7), one can carry out the integration over $x \in \mathbb{R}$ straightforwardly, due to the factorised integration measure (A.9) and the orthogonality of Hermite polynomials on $\mathbb{R}$ with respect to the weight $\exp(-x^2)$. Lengthy but straightforward calculations result in

$$\hat{\chi}_{2\alpha+1,2\beta} = \frac{i\,\alpha!}{2}(\tilde{\gamma}_{\alpha,\beta} - \tilde{\gamma}_{\alpha-1,\beta}) \quad (A.13)$$

where $\tilde{\gamma}_{\alpha,\beta}$ is

$$\tilde{\gamma}_{\alpha,\beta} = (-1)^{\beta-\alpha}(2\alpha+1)\frac{h_\alpha}{\alpha!}\int_0^\infty dy\, y^{2(\beta-\alpha)-1} e^{y^2} \operatorname{erfc}(y\sqrt{2}) \times (2\beta)! \sum_{j=0}^{2\alpha+1} \frac{(2y^2)^j}{j!\,(2\alpha+1-j)!\,(j+2(\beta-\alpha)-1)!}. \quad (A.14)$$

Notice that for $\alpha \le -1$, the sum in (A.14) is void so that $\tilde{\gamma}_{\alpha\le-1,\beta} = 0$. Otherwise, the series can be summed up in terms of Laguerre polynomials,

$$(2\beta)! \sum_{j=0}^{2\alpha+1} \frac{(2y^2)^j}{j!\,(2\alpha+1-j)!\,(j+2(\beta-\alpha)-1)!} = L_{2\alpha+1}^{2(\beta-\alpha)-1}(-2y^2), \quad (A.15)$$

to yield

$$\tilde{\gamma}_{\alpha,\beta} = (-1)^{\beta-\alpha}(2\alpha+1)\frac{h_\alpha}{\alpha!}\int_0^\infty dy\, y^{2(\beta-\alpha)-1} e^{y^2} \operatorname{erfc}(y\sqrt{2})\, L_{2\alpha+1}^{2(\beta-\alpha)-1}(-2y^2). \quad (A.16)$$

From now on, the Laguerre polynomials of "negative order" are interpreted to be zeros. Substituting it back to (A.13), we end up with (A.10).

Finally, the formula for $\hat{\chi}_{2\alpha,2\beta+1}$ follows from the symmetry relation (A.11).

A4. The structure of $\hat{\boldsymbol{\chi}}$.—To summarise the calculations of Appendix A, the structure of the $n \times n$ matrix $\hat{\boldsymbol{\chi}}$ is presented below. For $n = 2m$ even, the matrix $\hat{\boldsymbol{\chi}}$ is

$$\hat{\chi}_{ij}^{\text{even}} = \begin{pmatrix} 0 & \hat{\chi}_{01} & \cdots & 0 & \hat{\chi}_{0,2m-1} \\ -\hat{\chi}_{01} & 0 & \cdots & \hat{\chi}_{1,2m-2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & -\hat{\chi}_{1,2m-2} & \cdots & 0 & \hat{\chi}_{2m-2,2m-1} \\ -\hat{\chi}_{0,2m-1} & 0 & \cdots & -\hat{\chi}_{2m-2,2m-1} & 0 \end{pmatrix}. \quad (A.17)$$

For $n = 2m+1$ odd, the structure of $\hat{\boldsymbol{\chi}}$ follows the pattern

$$\hat{\chi}_{ij}^{\rm odd} = \begin{pmatrix} 0 & \hat{\chi}_{01} & \cdots & \hat{\chi}_{0,2m-1} & 0 \\ -\hat{\chi}_{01} & 0 & \cdots & 0 & \hat{\chi}_{1,2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\hat{\chi}_{0,2m-1} & 0 & \cdots & 0 & \hat{\chi}_{2m-1,2m} \\ 0 & -\hat{\chi}_{1,2m} & \cdots & -\hat{\chi}_{2m-1,2m} & 0 \end{pmatrix}. \tag{A.18}$$

Appendix B. The $n \times n$ matrix $\hat{\boldsymbol{\sigma}} = 2i\,(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\chi}})$

B1. Calculation of the entries $\hat{\sigma}_{\alpha,\beta} = 2i\,(\hat{\mu}\hat{\chi})_{\alpha,\beta}$ for n even.—The matrices $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\chi}}$ and thus $(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\chi}})$ are sensitive to the parity of n, the matrix dimensionality. For $n = 2m$, we make use of (3.21), (A.8), (A.11) and (A.17) to derive

$$(\hat{\mu}\hat{\chi})^{\rm even}_{2\alpha,2\beta+1} = (\hat{\mu}\hat{\chi})^{\rm even}_{2\alpha+1,2\beta} = 0, \tag{B.1}$$

$$(\hat{\mu}\hat{\chi})^{\rm even}_{2\alpha,2\beta} = -\frac{1}{h_\alpha}\,\hat{\chi}_{2\alpha+1,2\beta} = \frac{1}{h_\alpha}\,\hat{\chi}_{2\beta,2\alpha+1}, \tag{B.2}$$

$$(\hat{\mu}\hat{\chi})^{\rm even}_{2\alpha+1,2\beta+1} = \frac{1}{h_\alpha}\,\hat{\chi}_{2\alpha,2\beta+1} = -\frac{1}{h_\alpha}\,\hat{\chi}_{2\beta+1,2\alpha} = \frac{h_\beta}{h_\alpha}\,(\hat{\mu}\hat{\chi})^{\rm even}_{2\beta,2\alpha}. \tag{B.3}$$

Here, $\alpha = 0, 1, \cdots$ and $\beta = 0, 1, \cdots$.

An integral representation for the entries $\hat{\sigma}^{\rm even}_{2\alpha,2\beta}$ and $\hat{\sigma}^{\rm even}_{2\alpha+1,2\beta+1}$ can be read off from (A.10). Explicitly, we have

$$\begin{aligned}\hat{\sigma}^{\rm even}_{2\alpha,2\beta} = (-1)^{\beta-\alpha} \int_0^\infty dy\, y^{2(\beta-\alpha)-1}\, e^{y^2}\, {\rm erfc}\,(y\sqrt{2}) \\ \times \Big[(2\alpha+1)\, L_{2\alpha+1}^{2(\beta-\alpha)-1}(-2y^2) + 2y^2\, L_{2\alpha-1}^{2(\beta-\alpha)+1}(-2y^2)\Big]\end{aligned} \tag{B.4}$$

and

$$\hat{\sigma}^{\rm even}_{2\alpha+1,2\beta+1} = \frac{h_\beta}{h_\alpha}\,\hat{\sigma}^{\rm even}_{2\beta,2\alpha}. \tag{B.5}$$

All other entries are zeros.

Yet another representation in terms of $\tilde{\gamma}_{\alpha,\beta}$ given by (A.16) is useful. Equations (A.13), (B.2) and (B.3) yield:

$$\hat{\sigma}^{\rm even}_{2\alpha,2\beta} = \frac{\alpha!}{h_\alpha}\,(\tilde{\gamma}_{\alpha,\beta} - \tilde{\gamma}_{\alpha-1,\beta})\,, \tag{B.6}$$

$$\hat{\sigma}^{\rm even}_{2\alpha+1,2\beta+1} = \frac{\beta!}{h_\alpha}\,(\tilde{\gamma}_{\beta,\alpha} - \tilde{\gamma}_{\beta-1,\alpha})\,. \tag{B.7}$$

All other entries nullify in accordance with (B.1).

B2. Calculation of the entries $\hat{\sigma}_{\alpha,\beta} = 2i\,(\hat{\mu}\hat{\chi})_{\alpha,\beta}$ for n odd.—For $n = 2m+1$, we make use of (3.22), (A.8), (A.11), (A.18), (B.2) and (B.3) to derive

$$(\hat{\mu}\hat{\chi})^{\rm odd}_{2\alpha,2\beta+1} = (\hat{\mu}\hat{\chi})^{\rm odd}_{2\alpha+1,2\beta} = 0, \tag{B.8}$$

$$(\hat{\mu}\hat{\chi})^{\rm odd}_{2\alpha,2\beta} = (1-\delta_{\alpha,m})(\hat{\mu}\hat{\chi})^{\rm even}_{2\alpha,2\beta} - \frac{m!}{h_m}\,\delta_{\alpha,m} \sum_{j=0}^{m-1} \frac{h_j}{j!}\,(\hat{\mu}\hat{\chi})^{\rm even}_{2j,2\beta}, \tag{B.9}$$

$$(\hat{\mu}\hat{\chi})^{\rm odd}_{2\alpha+1,2\beta+1} = \frac{h_\beta}{h_\alpha}\,\left[(\hat{\mu}\hat{\chi})^{\rm even}_{2\beta,2\alpha} - \frac{m!}{h_m}\,\frac{h_\alpha}{\alpha!}(\hat{\mu}\hat{\chi})^{\rm even}_{2\beta,2m}\right]. \tag{B.10}$$

Here, $\alpha = 0, 1, \cdots$ and $\beta = 0, 1, \cdots$.

The above formulae simplify if expressed in terms of $\tilde{\gamma}_{\alpha,\beta}$ akin to (B.6) and (B.7). Straightforward calculations lead to the following result for the entries of $\hat{\boldsymbol{\sigma}}^{\rm odd}$:

$$\hat{\sigma}^{\rm odd}_{2\alpha,2\beta} = \frac{\alpha!}{h_\alpha}\Big[(1-\delta_{\alpha,m})(\tilde{\gamma}_{\alpha,\beta} - \tilde{\gamma}_{\alpha-1,\beta}) - \delta_{\alpha,m}\,\tilde{\gamma}_{m-1,\beta}\Big], \tag{B.11}$$

$$\hat{\sigma}^{\rm odd}_{2\alpha+1,2\beta+1} = \frac{\beta!}{\alpha!}\Big[\frac{\alpha!}{h_\alpha}(\tilde{\gamma}_{\beta,\alpha} - \tilde{\gamma}_{\beta-1,\alpha}) - \frac{m!}{h_m}(\tilde{\gamma}_{\beta,m} - \tilde{\gamma}_{\beta-1,m})\Big]. \tag{B.12}$$

Finally, the integral representations for (B.11) and (B.12) can be obtained from (A.16).

Appendix C. The $\lfloor n/2\rfloor \times \lfloor n/2\rfloor$ matrix $\hat{\boldsymbol{\varrho}}$ and the trace identity

C1. The definition of $\hat{\boldsymbol{\varrho}}$.—Since the $n\times n$ matrix $\hat{\boldsymbol{\sigma}}$ has half of its entries vanishing, it is useful to define a reduced, $\lfloor n/2\rfloor \times \lfloor n/2\rfloor$ matrix $\hat{\boldsymbol{\varrho}}$, such that

$$\hat{\varrho}^{\rm even}_{\alpha,\beta} = (-1)^{\beta-\alpha}\,\hat{\sigma}^{\rm even}_{2\alpha,2\beta} = (-1)^{\beta-\alpha}\,\frac{\alpha!}{h_\alpha}\,(\tilde{\gamma}_{\alpha,\beta} - \tilde{\gamma}_{\alpha-1,\beta}) \tag{C.1}$$

for $n = 2m$ even, and

$$\begin{aligned}\hat{\varrho}^{\rm odd}_{\alpha,\beta} &= (-1)^{\beta-\alpha}\frac{h_\beta}{h_\alpha}\,\hat{\sigma}^{\rm odd}_{2\beta+1,2\alpha+1} = (-1)^{\beta-\alpha}\left[\hat{\sigma}^{\rm even}_{2\alpha,2\beta} - \frac{m!}{h_m}\,\frac{h_\beta}{\beta!}\hat{\sigma}^{\rm even}_{2\alpha,2m}\right]\\ &= \hat{\varrho}^{\rm even}_{\alpha,\beta} - (-1)^{\beta-m}\,\frac{m!}{h_m}\,\frac{h_\beta}{\beta!}\,\hat{\varrho}^{\rm even}_{\alpha,m}\end{aligned} \tag{C.2}$$

for $n = 2m+1$ odd. Explicitly,

$$\begin{aligned}\hat{\varrho}^{\rm even}_{\alpha,\beta} = &\int_0^\infty dy\, y^{2(\beta-\alpha)-1}\, e^{y^2}\,{\rm erfc}\,(y\sqrt{2})\\ &\times\,\left[(2\alpha+1)\,L^{2(\beta-\alpha)-1}_{2\alpha+1}(-2y^2) + 2y^2\,L^{2(\beta-\alpha)+1}_{2\alpha-1}(-2y^2)\right]\end{aligned} \tag{C.3}$$

and

$$\hat{\varrho}^{\rm odd}_{\alpha,\,\beta} = \hat{\varrho}^{\rm even}_{\alpha,\,\beta} - (-4)^{m-\beta}\,\frac{m!}{(2m)!}\,\frac{(2\beta)!}{\beta!}\,\hat{\varrho}^{\rm even}_{\alpha,\,m}. \tag{C.4}$$

The above formulae have been obtained from the definitions (C.1) and (C.2) with the help of (B.4) and (3.18).

C2. The trace identity.—It turns out that the probability function $p_{n,k}$ is most economically expressed in terms of $\hat{\boldsymbol{\varrho}}$. To realise this, it is instructive to prove the following trace identity.

Lemma C1. *Let* $\hat{\boldsymbol{\sigma}}$ *be an* $n\times n$ *matrix with the entries given by either (B.6), (B.7) or (B.11), (B.12) depending on the parity of* n*. Then, the trace identity*

$${\rm tr}_{\,(0,n-1)}\hat{\boldsymbol{\sigma}}^j = 2\,{\rm tr}_{\,(0,\lfloor n/2\rfloor-1)}\hat{\boldsymbol{\varrho}}^j \tag{C.5}$$

holds for all $j = 1, 2, \cdots$ *and the matrix* $\hat{\boldsymbol{\varrho}}$ *defined by (C.1) and (C.2).*

Proof. Since the matrix elements $\hat{\sigma}_{2\alpha,2\beta+1}$ and $\hat{\sigma}_{2\alpha+1,2\beta}$ are zeros, the trace of $\hat{\boldsymbol{\sigma}}^j$ $(j = 1, 2, \cdots)$ can always be separated into two pieces:

$${\rm tr}_{\,(0,n-1)}\hat{\boldsymbol{\sigma}}^j = {\rm tr}_{\,(0,\lceil n/2\rceil-1)}\hat{\boldsymbol{a}}^{\,j} + {\rm tr}_{\,(0,\lfloor n/2\rfloor-1)}\hat{\boldsymbol{b}}^{\,j}. \tag{C.6}$$

The matrices $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ are defined through their entries,

$$\hat{a}_{\alpha,\beta} = \hat{\sigma}_{2\alpha,2\beta}, \tag{C.7}$$

$$\hat{b}_{\alpha,\beta} = \hat{\sigma}_{2\alpha+1,2\beta+1}, \tag{C.8}$$

and explicitly depend on the parity of n in accordance with the discussion in Appendix B.

(i) *The case $n = 2m$ even* is the simplest one because of the relation (B.5). Indeed, its straightforward use results in

$$\mathrm{tr}_{\,(0,m-1)}\hat{\boldsymbol{b}}^{\,j} = \mathrm{tr}_{\,(0,m-1)}\hat{\boldsymbol{a}}^{\,j}. \tag{C.9}$$

This can further be simplified due to(C.9), (C.7), (C.6) and (C.1),

$$\mathrm{tr}_{\,(0,2m-1)}\hat{\boldsymbol{\sigma}}^{j} = 2\,\mathrm{tr}_{\,(0,m-1)}\hat{\boldsymbol{a}}^{\,j} = 2\,\mathrm{tr}_{\,(0,m-1)}\hat{\boldsymbol{\varrho}}^{\,\mathrm{even}\,j}. \tag{C.10}$$

This completes the proof of (C.5) for $n = 2m$ even.

(ii) *The case $n = 2m + 1$ odd* is a bit more complicated since a simple analogue of (B.5) does not exist. Instead, we have [(C.7) and (C.8)]

$$\hat{a}_{\alpha,\beta} = c_\alpha \Big[(1-\delta_{\alpha,m})(\tilde{\gamma}_{\alpha,\beta} - \tilde{\gamma}_{\alpha-1,\beta}) - \delta_{\alpha,m}\,\tilde{\gamma}_{m-1,\beta}\Big], \tag{C.11}$$

$$\hat{b}_{\alpha,\beta} = c_\alpha\,(\tilde{\gamma}_{\beta,\alpha} - \tilde{\gamma}_{\beta-1,\alpha}) - c_m\,(\tilde{\gamma}_{\beta,m} - \tilde{\gamma}_{\beta-1,m}), \tag{C.12}$$

where $c_\alpha = \alpha!/h_\alpha$ has been defined in (3.23). In writing (C.12), we dropped the prefactor $\beta!/\alpha!$ appearing in (B.12) since it does not affect the value of the second trace in (C.6). To prove (C.5), we will start with (C.6) in order to demonstrate that

$$\mathrm{tr}_{\,(0,m-1)}\hat{\boldsymbol{b}}^{\,j} = \mathrm{tr}_{\,(0,m)}\hat{\boldsymbol{a}}^{\,j}. \tag{C.13}$$

This will be followed by a proof that either of these traces reduces to $\mathrm{tr}_{\,(0,m-1)}\hat{\boldsymbol{\varrho}}^{\mathrm{odd}\,j}$.

Let us prove (C.13) by focusing on the eigenvalues $\{\lambda\}$ of the matrices $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ which are the roots of the secular equations

$$\det\,\big[\hat{\boldsymbol{a}} - \lambda\,\hat{\mathbf{1}}\big]_{(m+1)\times(m+1)} = 0, \tag{C.14}$$

$$\det\,\big[\hat{\boldsymbol{b}} - \lambda\,\hat{\mathbf{1}}\big]_{m\times m} = 0. \tag{C.15}$$

We claim that (C.13) holds because exactly *one* out of $(m+1)$ eigenvalues of the matrix $\hat{\boldsymbol{a}}$ is *zero* whilst the remaining m eigenvalues of $\hat{\boldsymbol{a}}$ coincide with m eigenvalues of $\hat{\boldsymbol{b}}$. Put differently, we are going to prove that

$$\det\,\big[\hat{\boldsymbol{a}} - \lambda\,\hat{\mathbf{1}}\big]_{(m+1)\times(m+1)} = -\lambda\,\det\,\big[\hat{\boldsymbol{b}} - \lambda\,\hat{\mathbf{1}}\big]_{m\times m}. \tag{C.16}$$

The proof consists of four steps.

Step ♮1. Consider the $(m+1)\times(m+1)$ matrix $\hat{\boldsymbol{a}} - \lambda\hat{\mathbf{1}}$ under the determinant in the secular equation (C.14),

$$\begin{pmatrix} c_0\,\tilde{\gamma}_{0,0} - \lambda & c_0\,\tilde{\gamma}_{0,1} & \cdots & c_0\,\tilde{\gamma}_{0,m-1} & c_0\,\tilde{\gamma}_{0,m} \\ c_1\,\Gamma_{1,0} & c_1\,\Gamma_{1,1} - \lambda & \cdots & c_1\,\Gamma_{1,m-1} & c_1\,\Gamma_{1,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{m-1}\,\Gamma_{m-1,0} & c_{m-1}\,\Gamma_{m-1,1} & \cdots & c_{m-1}\,\Gamma_{m-1,m-1} - \lambda & c_{m-1}\,\Gamma_{m-1,m} \\ -c_m\,\tilde{\gamma}_{m-1,0} & -c_m\,\tilde{\gamma}_{m-1,1} & \cdots & -c_m\,\tilde{\gamma}_{m-1,m-1} & -c_m\,\tilde{\gamma}_{m-1,m} - \lambda \end{pmatrix}, \tag{C.17}$$

where $\Gamma_{\alpha,\beta} = \tilde{\gamma}_{\alpha,\beta} - \tilde{\gamma}_{\alpha-1,\beta}$. Let us perform a number of operations with rows and columns that will leave the value of the secular determinant intact. First, we multiply the content of the first row by c_1/c_0 and add it to the second row. Having done this, we multiply a new content of the second row by c_2/c_1 and add it to the content of the third row. We go on with this procedure until we arrive at the modified mth row whose content is multiplied by c_m/c_{m-1} and further added to the last, $(m+1)$th, row. While not affecting a value of the determinant in (C.14), the above sequence of transformations brings (C.17) to the form

$$\begin{pmatrix} c_0\,\tilde{\gamma}_{0,0} - \lambda & c_0\,\tilde{\gamma}_{0,1} & \cdots & c_0\,\tilde{\gamma}_{0,m} \\ c_1\,\tilde{\gamma}_{1,0} - (c_1/c_0)\lambda & c_1\,\tilde{\gamma}_{1,1} - \lambda & \cdots & c_1\,\tilde{\gamma}_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-1}\,\tilde{\gamma}_{m-1,0} - (c_{m-1}/c_0)\lambda & c_{m-1}\,\tilde{\gamma}_{m-1,1} - (c_{m-1}/c_1)\lambda & \cdots & c_{m-1}\,\tilde{\gamma}_{m-1,m} \\ -(c_m/c_0)\lambda & -(c_m/c_1)\lambda & \cdots & -(c_m/c_m)\lambda \end{pmatrix}. \tag{C.18}$$

The last row of this equation suggests that a factor λ can be taken out of the secular determinant. In other words, $\lambda = 0$ is always an eigenvalue of the $(m+1)\times(m+1)$ matrix $\hat{\boldsymbol{a}}$.

Step ♮2. Next, we multiply the content of the first row in (C.18) by c_1/c_0 and subtract it from the content of the second row; having done that, we multiply the content of the modified second row by c_2/c_1 and subtract it from the third row; going on with this procedure, we arrive at the $(m-1)$th row, multiply it by c_{m-1}/c_{m-2} to subtract this from the content of the mth row. We do not touch the last, $(m+1)$th row. This set of transformations yields

$$\begin{pmatrix} c_0\,\tilde{\gamma}_{0,0} - \lambda & c_0\,\tilde{\gamma}_{0,1} & \cdots & c_0\,\tilde{\gamma}_{0,m-1} & c_0\,\tilde{\gamma}_{0,m} \\ c_1\,\Gamma_{1,0} & c_1\,\Gamma_{1,1} - \lambda & \cdots & c_1\,\Gamma_{1,m-1} & c_1\,\Gamma_{1,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{m-1}\,\Gamma_{m-1,0} & c_{m-1}\,\Gamma_{m-1,1} & \cdots & c_{m-1}\,\Gamma_{m-1,m-1} - \lambda & c_{m-1}\,\Gamma_{m-1,m} \\ -(c_m/c_0)\lambda & -(c_m/c_1)\lambda & \cdots & -(c_m/c_{m-1})\lambda & -(c_m/c_m)\lambda \end{pmatrix}. \tag{C.19}$$

Note, that all but the last row of the matrix (C.19) coincide with those in (C.17).

Step ♮3. Now, let us factor out c_0 from the first row, c_1 from the second row, ..., c_{m-1} from the mth row to obtain

$$\prod_{j=0}^{m-1} c_j$$

times

$$\begin{pmatrix} \tilde{\gamma}_{0,0} - \lambda/c_0 & \tilde{\gamma}_{0,1} & \cdots & \tilde{\gamma}_{0,m-1} & \tilde{\gamma}_{0,m} \\ \Gamma_{1,0} & \Gamma_{1,1} - \lambda/c_1 & \cdots & \Gamma_{1,m-1} & \Gamma_{1,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Gamma_{m-1,0} & \Gamma_{m-1,1} & \cdots & \Gamma_{m-1,m-1} - \lambda/c_{m-1} & \Gamma_{m-1,m} \\ -(c_m/c_0)\lambda & -(c_m/c_1)\lambda & \cdots & -(c_m/c_{m-1})\lambda & -(c_m/c_m)\lambda \end{pmatrix}. \tag{C.20}$$

Next, we multiply the first *column* by c_0, the second column by c_1, ..., the mth column by c_{m-1}, and do not alter the last, $(m+1)$th column. This leads us to

$$\begin{pmatrix} c_0\,\tilde{\gamma}_{0,0}-\lambda & c_1\,\tilde{\gamma}_{0,1} & \cdots & c_{m-1}\,\tilde{\gamma}_{0,m-1} & \tilde{\gamma}_{0,m} \\ c_0\,\Gamma_{1,0} & c_1\,\Gamma_{1,1}-\lambda & \cdots & c_{m-1}\,\Gamma_{1,m-1} & \Gamma_{1,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_0\,\Gamma_{m-1,0} & c_1\,\Gamma_{m-1,1} & \cdots & c_{m-1}\,\Gamma_{m-1,m-1}-\lambda & \Gamma_{m-1,m} \\ -c_m\lambda & -c_m\lambda & \cdots & -c_m\lambda & -\lambda \end{pmatrix}. \tag{C.21}$$

Step ♮4. Now, we multiply the last, $(m+1)$th column by c_m and subtract it from the first, second, ..., the mth column to reduce (C.21) to

$$\begin{pmatrix} c_0\,\tilde{\gamma}_{0,0}-c_m\tilde{\gamma}_{0,m}-\lambda & \cdots & \cdots & \tilde{\gamma}_{0,m} \\ c_0\,\Gamma_{1,0}-c_m\,\Gamma_{1,m} & \cdots & \cdots & \Gamma_{1,m} \\ \vdots & \ddots & \vdots & \vdots \\ c_0\,\Gamma_{m-1,0}-c_m\,\Gamma_{m-1,m} & \cdots & \cdots & \Gamma_{m-1,m} \\ 0 & \cdots & 0 & -\lambda \end{pmatrix}. \tag{C.22}$$

The determinant of the latter matrix can be calculated via expanding with respect to its last row. As a result, the secular equation (C.14) is reduced to

$$-\lambda\cdot\det\left[c_\beta\left(\tilde{\gamma}_{\alpha,\beta}-\tilde{\gamma}_{\alpha-1,\beta}\right)-c_m\left(\tilde{\gamma}_{\alpha,m}-\tilde{\gamma}_{\alpha-1,m}\right)-\lambda\,\delta_{\alpha,\beta}\right]_{(\alpha,\beta)=0,\cdots,m-1}. \tag{C.23}$$

A comparison with (C.12) allows us to rewrite (C.23) in the form

$$-\lambda\,\det\left[\hat{\boldsymbol{b}}^{\mathrm{T}}-\lambda\,\hat{\mathbf{1}}\right]_{m\times m}. \tag{C.24}$$

This establishes (C.13).

Finally, it remains to show that, for all $j=1,2,\cdots$, the identity

$$\mathrm{tr}_{(0,m-1)}\hat{\boldsymbol{b}}^{\,j}=\mathrm{tr}_{(0,m-1)}\hat{\boldsymbol{\varrho}}^{\,\mathrm{odd}\,j}, \tag{C.25}$$

holds. That (C.25) is indeed true, follows from (C.8) and (C.2). This completes our proof of the Lemma. ■

Appendix D. Calculation of the trace $\mathrm{tr}_{(0,\lfloor n/2\rfloor-1)}\hat{\boldsymbol{\varrho}}$

Since the matrix $\hat{\boldsymbol{\varrho}}$ is sensitive to the parity of n, two separate calculations are needed.

(i) *The case $n=2m$ even.*—To calculate the trace, we make use of (C.3) to write down

$$\mathrm{tr}_{(0,m-1)}\hat{\boldsymbol{\varrho}}^{\mathrm{even}}=\sum_{\alpha=0}^{m-1}\hat{\varrho}_{\alpha,\alpha}=\int_0^\infty dy\,y^{-1}\,e^{y^2}\,\mathrm{erfc}(y\sqrt{2})$$
$$\times\sum_{\alpha=0}^{m-1}\Big[(2\alpha+1)\,L_{2\alpha+1}^{-1}(-2y^2)+2y^2\,L_{2\alpha-1}^{1}(-2y^2)\Big]. \tag{D.1}$$

First, to put (D.1) into a more tractable form, we apply the identity §

$$L_n^{-m}(w)=\frac{w^m}{(-n)_m}\,L_{n-m}^m(w) \tag{D.2}$$

§ `http://functions.wolfram.com/05.08.17.0009.01` which, in the context of (D.1), reads

$$L_{2\alpha+1}^{-1}(w) = -\frac{w}{2\alpha+1}\, L_{2\alpha}^{1}(w). \tag{D.3}$$

The use of (D.3) reduces the integrand of (D.1) to

$$(2\alpha+1)\, L_{2\alpha+1}^{-1}(-2y^2) + 2y^2\, L_{2\alpha-1}^{1}(-2y^2) = 2y^2 \Big[L_{2\alpha}^{1}(-2y^2) + L_{2\alpha-1}^{1}(-2y^2) \Big]. \tag{D.4}$$

Second, we spot that the transformation

$$L_{\nu}^{\lambda-1}(w) = L_{\nu}^{\lambda}(w) - L_{\nu-1}^{\lambda}(w) \tag{D.5}$$

applied to (D.4), yields

$$\underbrace{L_{2\alpha}^{1}(-2y^2)}_{L_{2\alpha}^{2} - L_{2\alpha-1}^{2}} + \underbrace{L_{2\alpha-1}^{1}(-2y^2)}_{L_{2\alpha-1}^{2} - L_{2\alpha-2}^{2}} = L_{2\alpha}^{2}(-2y^2) - L_{2\alpha-2}^{2}(-2y^2). \tag{D.6}$$

As a consequence, the summation in (D.1) can be performed explicitly,

$$\begin{aligned} \sum_{\alpha=0}^{m-1} & \Big[(2\alpha+1)\, L_{2\alpha+1}^{-1}(-2y^2) + 2y^2\, L_{2\alpha-1}^{1}(-2y^2) \Big] \\ &= 2y^2 \sum_{\alpha=0}^{m-1} \Big[L_{2\alpha}^{2}(-2y^2) - L_{2\alpha-2}^{2}(-2y^2) \Big] \\ &= 2y^2\, L_{2m-2}^{2}(-2y^2), \end{aligned} \tag{D.7}$$

resulting in a remarkably simple formula

$$\mathrm{tr}_{(0,m-1)}\, \hat{\boldsymbol{\varrho}}^{\mathrm{even}} = 2 \int_0^\infty dy\, y\, e^{y^2}\, \mathrm{erfc}(y\sqrt{2})\, L_{2m-2}^{2}(-2y^2). \tag{D.8}$$

(ii) *The case $n = 2m+1$ odd.*—To calculate the trace, we combine (C.1) and (C.4) into

$$\hat{\varrho}_{\alpha,\alpha}^{\mathrm{odd}} = \hat{\varrho}_{\alpha,\alpha}^{\mathrm{even}} - \frac{m!}{h_m}\, (\tilde{\gamma}_{\alpha,m} - \tilde{\gamma}_{\alpha-1,m}). \tag{D.9}$$

Summing it up, we derive

$$\mathrm{tr}_{(0,m-1)}\, \hat{\boldsymbol{\varrho}}^{\mathrm{odd}} = \mathrm{tr}_{(0,m-1)}\, \hat{\boldsymbol{\varrho}}^{\mathrm{even}} - \frac{m!}{h_m}\, \tilde{\gamma}_{m-1,m}. \tag{D.10}$$

Further use of (D.8) and (A.16) yields

$$\mathrm{tr}_{(0,m-1)}\, \hat{\boldsymbol{\varrho}}^{\mathrm{odd}} = 2 \int_0^\infty dy\, y\, e^{y^2}\, \mathrm{erfc}(y\sqrt{2})\, \left[L_{2m-2}^{2}(-2y^2) + L_{2m-1}^{1}(-2y^2) \right]. \tag{D.11}$$

With the help of the identity (D.5), this eventually simplifies to

$$\mathrm{tr}_{(0,m-1)}\, \hat{\boldsymbol{\varrho}}^{\mathrm{odd}} = 2 \int_0^\infty dy\, y\, e^{y^2}\, \mathrm{erfc}(y\sqrt{2})\, L_{2m-1}^{2}(-2y^2). \tag{D.12}$$

The formulae (D.8) and (D.12) can be unified into a single equation, holding for n of arbitrary parity:

$$\mathrm{tr}_{(0,\lfloor n/2 \rfloor -1)}\, \hat{\boldsymbol{\varrho}} = 2 \int_0^\infty dy\, y\, e^{y^2}\, \mathrm{erfc}(y\sqrt{2})\, L_{n-2}^{2}(-2y^2). \tag{D.13}$$

References

Adler M, Forrester P J, Nagao T, van Moerbeke P 2000 Classical skew orthogonal polynomials and random matrices *J. Stat. Phys.* **99** 141

Agam O, Bettelheim E, Wiegmann P B, and Zabrodin A 2002 Viscous fingering and the shape of an electronic droplet in the quantum Hall regime *Phys. Rev. Lett.* **88** 236801

Akemann 2005 The Complex Laguerre Symplectic Ensemble of Non-Hermitean Matrices *Nucl. Phys.* B **730** 253

Akemann G and Basile F 2007 Massive partition functions and complex eigenvalue correlations in matrix models with symplectic symmetry *Nucl. Phys.* B **766** 150

Akemann G 2007 Matrix Models and QCD with Chemical Potential *arXiv: hep-th/0701175 [Int. J. Mod. Phys. A (in press)]*

Andrews G E 1998 *The Theory of Partitions* (Cambridge: Cambridge University Press)

Bai Z D 1997 Circular law *Ann. Probab.* **25** 494

Borodin A and Strahov E 2005 Averages of characteristic polynomials in random matrix theory *Commun. Pure and Appl. Math.* **LVIII** 0001

Chalker J T and Mehlig B 1998 Eigenvector statistics in non-Hermitean random matrix ensembles *Phys. Rev. Lett.* **81** 3367

Dyson F J 1970 Correlations between eigenvalues of a random matrix *Commun. Math. Phys.* **19** 235

Dyson F J 1972 Quaternion determinants *Helv. Phys. Acta* **49** 289

Edelman A, Kostlan E, and Shub M 1994 How many eigenvalues of a random matrix are real? *J. Amer. Math. Soc.* **7** 247

Edelman A 1997 The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law *J. Mult. Analysis* **60** 203

Efetov K B 1997a Directed quantum chaos *Phys. Rev. Lett.* **79** 491

Efetov K B 1997b Quantum disordered systems with a direction *Phys. Rev.* B **56** 9630

Eynard B 2001 Asymptotics of skew orthogonal polynomials *J. Phys. A: Math. Gen.* **34** 7591

Forrester P J 2005 *Log-Gases and Random Matrices (web-book)*

Fyodorov Y V, Khoruzhenko B, and Sommers H-J 1997 Almost Hermitean random matrices: Crossover from Wigner-Dyson to Ginibre eigenvalue statistics *Phys. Rev. Lett.* **79** 557

Fyodorov Y V and Sommers H-J 2003 Random matrices close to Hermitean or unitary: Overview of methods and results *J. Phys. A: Math. Gen.* **36** 3303

Ginibre J 1965 Statistical ensembles of complex, quaternion, and real matrices *J. Math. Phys.* **19** 133

Girko V L 1984 Circle law *Theory Probab. Appl.* **29** 694

Girko V L 1986 Elliptic law *Theory Probab. Appl.* **30** 677

Grobe R, Haake F, and Sommers H-J 1988 Quantum distinction of regular and chaotic dissipative motion *Phys. Rev. Lett.* **61** 1899

Grobe R and Haake F 1989 Universality of cubic-level repulsion for dissipative quantum chaos *Phys. Rev. Lett.* **62** 2893

Guhr T, Müller-Groeling A, and Weidenmüller H A 1998 Random matrix theories in quantum physics: Common concepts *Phys. Reports* **299** 189

Halasz M A, Osborn J C, and Verbaarschot J J M 1997 Random matrix triality at nonzero chemical potential *Phys. Rev.* D **56** 7059

Hardy G H, Ramanujan S 1918 Asymptotic formulae in combinatory analysis *Proc. London Math. Soc.* B **17** 75

Jack H 1976 A class of polynomials in search of a definition, or the symmetric group parameterized *Jack, Hall-Littlewood and Macdonald polynomials (Contemporary Mathematics Series)* ed V B Kuznetsov and S Sahi (Providence: AMS 2006)

Janik R A, Nörenberg W, Nowak M A, Papp G, and Zahed I 1999 Correlations of eigenvectors for non-Hermitean random-matrix models *Phys. Rev.* E **60** 2699

Kanzieper E 2002a Eigenvalue correlations in non-Hermitean symplectic random matrices *J. Phys. A: Math. Gen.* **35** 6631

Kanzieper E 2002b Replica field theories, Painlevé transcendents, and exact correlation functions *Phys. Rev. Lett.* **89** 250201

Kanzieper E 2005 Exact replica treatment of non-Hermitean complex random matrices *Frontiers in Field Theory* ed O Kovras (New York: Nova Science Publishers) p 23

Kanzieper E and Akemann G 2005 Statistics of real eigenvalues in Ginibre's ensemble of random real matrices *Phys. Rev. Lett.* **95** 230201

Khoruzhenko B A and Mezzadri F 2005 *Private communication*

Kolesnikov A V and Efetov K B 1999 Distribution of complex eigenvalues for symplectic ensembles

of non-Hermitean matrices *Waves Random Media* **9** 71
Kwapień J, Drożdż S, and Ioannides A A 2000 Temporal correlations versus noise in the correlation matrix formalism: An example of the brain auditory response *Phys. Rev.* E **62** 5557
Kwapień J, Drożdż S, Górski A Z, and Oświęcimka P 2006 Asymmetric matrices in an analysis of financial correlations *Acta Phys. Polonica* **B37**, 3039
Le Caër G and Ho J S 1990 The Voronoi tessellation generated from eigenvalues of complex random matrices *J. Phys. A: Math. Gen.* **23** 3279
Le Caër G and Delannay R 1993 Topological models of 2D fractal cellular structures *J. Phys. I France* **3** 1777
Lehmann N and Sommers H J 1991 Eigenvalue statistics of random real matrices *Phys. Rev. Lett.* **67** 941
Macdonald I G 1998 *Symmetric Functions and Hall Polynomials* (Oxford: Oxford University Press)
Mahoux G and Mehta M L 1991 A method of integration over matrix variables *J. Phys. I (France)* **1** 1093
Markum H, Pullirsch R, and Wettig T 1999 Non-Hermitean random matrix theory and lattice QCD with chemical potential *Phys. Rev. Lett.* **83** 484
Mehlig B and Chalker J T 2000 Statistical properties of eigenvectors in non-Hermitean Gaussian random matrix ensembles *J. Math. Phys.* **41** 3233
Mehta M L and Srivastava P K 1966 Correlation functions for eigenvalues of real quaternion matrices *J. Math. Phys.* **7** 341
Mehta M L 1976 A note on certain multiple integrals *J. Math. Phys.* **17** 2198
Mehta M L 2004 *Random Matrices* (Amsterdam: Elsevier)
Mineev-Weinstein M, Wiegmann P B, and Zabrodin A 2000 Integrable structure of interface dynamics *Phys. Rev. Lett.* **84** 5106
Muirhead R J 1982 *Aspects of Multivariate Statistical Theory* (New York: John Wiley & Sons)
Nagao T and Nishigaki S M 2001 Massive random matrix ensembles at $\beta = 1$ and 4: QCD in three dimensions *Phys. Rev.* D **63** 045011
Nishigaki S M and Kamenev A 2002 Replica treatment of non-Hermitean disordered Hamiltonians *J. Phys. A: Math. Gen.* **35** 4571
Osborn J C 2004 Universal results from an alternate random matrix model for QCD with a baryon chemical potential *Phys. Rev. Lett.* **93** (2004) 222001
Prudnikov A P, Brychkov Yu A, and Marichev O I 1986 *Integrals and Series* vol 2 (New York: Gordon and Breach)
Prudnikov A P, Brychkov Yu A, and Marichev O I 1990 *Integrals and Series* vol 3 (New York: Gordon and Breach)
Sinclair C D 2006 Averages over Ginibre's ensemble of random real matrices *arXiv: math-ph/0605006*
Sommers H J, Crisanti A, Sompolinsky H, and Stein Y 1988 Spectrum of large random asymmetric matrices *Phys. Rev. Lett.* **60** 1895
Sompolinsky H, Crisanti A, Sommers H J 1988 Chaos in random neural networks *Phys. Rev. Lett.* **61** 259
Splittorff K and Verbaarschot J J M 2004 Factorization of Correlation Functions and the Replica Limit of the Toda Lattice Equation *Nucl. Phys.* B **683** 467
Stephanov M 1996 Random matrix model of QCD at finite density and the nature of the quenched limit *Phys. Rev. Lett.* **76** 4472
Timme M, Wolf F, and Geisel T 2002 Coexistence of regular and irregular dynamics in complex networks of pulse-coupled oscillators *Phys. Rev. Lett.* **89** 258701
Timme M, Wolf F, and Geisel T 2004 Topological speed limits to network synchronization *Phys. Rev. Lett.* **92** 074101
Tracy C A and Widom H 1998 Correlation functions, cluster functions, and spacing distributions for random matrices *J. Stat. Phys.* **92** 809
Wigner E P 1957 Statistical properties of real symmetric matrices with many dimensions *Proc. 4th Can. Math. Cong. (Toronto)* p 174
Wigner E P 1960 The unreasonable effectiveness of mathematics in natural sciences *Commun. Pure and Applied Mathematics* **13** 1
Zabrodin A 2003 New applications of non-Hermitean random matrices *Proceedings of the International Conference on Theoretical Physics (TH-2002)* ed D Iagolnitzer, V Rivasseau, and J Zinn-Justin (Basel: Birkhäuser Verlag)