Probab. Theory Relat. Fields 130, 337–367 (2004)
Digital Object Identifier (DOI) 10.1007/s00440-004-0355-x

B. Rider

Deviations from the Circular Law

Received: 19 August 2003 / Revised version: 13 February 2004 /
Published online: 29 April 2004 –

Abstract. Consider Ginibre's ensemble of $N \times N$ non-Hermitian random matrices in which all entries are independent complex Gaussians of mean zero and variance $\frac{1}{N}$. As $N \uparrow \infty$ the normalized counting measure of the eigenvalues converges to the uniform measure on the unit disk in the complex plane. In this note we describe fluctuations about this *Circular Law*. First we obtain finite N formulas for the covariance of certain linear statistics of the eigenvalues. Asymptotics of these objects coupled with a theorem of Costin and Lebowitz then result in central limit theorems for a variety of these statistics.

1. Introduction

A fundamental non-Hermitian ensemble of Random Matrix Theory (RMT) is that of $N \times N$ matrices with independent complex Gaussian entries of mean zero and variance $\frac{1}{N}$. This model is typically attributed to Ginibre who derived ([14]) the joint distribution of eigenvalues $\{z_k\}$ lying in the complex plane $\mathbb{C}$. That object is given by

$$dP_N(z_1, z_2, \ldots, z_N) = \frac{1}{Z_N} \prod_{1 \le \ell < k \le N} |z_\ell - z_k|^2 \mu_N(dz_1) \cdots \mu_N(dz_N) \qquad (1.1)$$

in which $\mu_N(dz) = e^{-N|z|^2} d\Re(z) d\Im(z)$ and Z_N is the appropriate normalizer. A straightforward Large Deviation analysis will take you from (1.1) to the *Circular Law*: for $N \uparrow \infty$ the measure $\frac{1}{N}\sum_{k=1}^{N} \delta_{z_k}$ tends weakly to the uniform measure on the disk $|z| \le 1$.[1] In other words, the mean number of eigenvalues falling in some subset of the disk is well approximated by $N\times$ the normalized area. The goal here is to describe the associated fluctuations.

Our study is centered around linear eigenvalue statistics of the form

$$X(f) = \sum_{k=1}^{N} f(z_k),$$

B. Rider: Department of Mathematics, Duke University, Box 90320, Durham, NC 27708, USA. e-mail: rider@math.duke.edu

*This work was supported in part by the grant NSF-9983320.

[1] Experts know this to hold under far more general conditions on the underlying distribution, see [1] for the best result.

and the behavior of their covariances

$$\mathrm{Cov}_N\Big(X(f), X(g)\Big) = \int_{\mathbb{C}^N} X(f)X(g)dP_N - \int_{\mathbb{C}^N} X(f)dP_N \int_{\mathbb{C}^N} X(g)dP_N$$

as $N \uparrow \infty$. In particular, if the statistic is one of either the moduli, $f(z_k) = f(|z_k|)$, or the angles, $f(z_k) = f(\arg(z_k))$, we obtain exact formulas for the covariance at finite N which are amenable to a precise asymptotic analysis. Afterward, noting that the Vandermonde component in (1.1) identifies the present ensemble as an *determinantal point field* [2], we may apply a basic result of that theory due to Costin and Lebowitz along with our asymptotics to conclude a central limit theorem for $X(f)$ in several cases.

There is already a vast body of work on fluctuations of linear statistics in random matrix ensembles. The majority of these concern either Hermitian ensembles (for example, [2], [3], [4], [7], [15], [16], [18], [27], and [29]) or the ensembles defined by Haar measure on the classical compact groups ([9], [20], [19], [28], and [32] to name a few). Of course, works such as [30] which focus on determinantal point fields take on a more general point of view. Nevertheless, in the non-Hermitian setting it appears this type of fluctuation question is only directly considered in [12], and that more or less from a physics standpoint. In fact, one motivation for the present paper was the increased interest in the physics community in non-Hermitian ensembles ([12] and [13] offer guides to that literature). A second motivation lies in the discovery that Ginibre's complex Gaussian ensemble possesses a structure allowing for exact formulas that in turn may be studied via rather straightforward mathematical machinery (basically Laplace-type asymptotics). Finally, it is expected that the results here should provide some indication as to the order and shape of fluctuations in more general non-Hermitian ensembles.[3]

We begin with a description of our covariance formulas. The case when f and g are functions of the spectral moduli alone is far the simpler of the two. The observation is that:

Theorem 1.1. *For statistics of the moduli, let f and g be bounded over $\mathbb{R}_+$ then*

$$\begin{aligned}\mathrm{Cov}_N\Big(X(f), X(g)\Big) = \sum_{\ell=1}^{N}\Big\{&E\Big[f\Big(\sqrt{\tfrac{1}{N}s_\ell}\Big)g\Big(\sqrt{\tfrac{1}{N}s_\ell}\Big)\Big] \\ &- E\Big[f\Big(\sqrt{\tfrac{1}{N}s_\ell}\Big)\Big]E\Big[g\Big(\sqrt{\tfrac{1}{N}s_\ell}\Big)\Big]\Big\} \qquad (1.2)\end{aligned}$$

where $s_\ell = \eta_1 + \cdots + \eta_\ell$ for $\{\eta_\ell\}$ a sequence of independent exponential random variables of mean one.

[2] Also known as *fermionic fields*, these objects were introduced by Macchi in [23]; an excellent introduction to their applications in RMT and beyond is contained in [31].

[3] Part of the sequel carries over to Ginibre's Quaternion Gaussian ensemble, see Section 5. However, describing fluctuations in the real Gaussian ensemble will require a new approach due to the intricacies of that eigenvalue density, see [21] or [11].

The angular statistics possess a more interesting structure, requiring the setting down of some notation. The angles or arguments of the $\{z_k\}$ are indexed to lie in the circle $\mathbb{T} = [-\pi, \pi]$, the latter normalized throughout according to $\int_{\mathbb{T}} d\theta = 1$. We recall the Dirichlet kernel

$$D_\ell(\theta) = \sum_{k=-\ell}^{\ell} e^{ik\theta} = \frac{\sin\Big((\ell + \frac{1}{2})\theta\Big)}{\sin\Big(\frac{1}{2}\theta\Big)},$$

which acts on functions f of $\mathbb{T}$ by convolution: $(D_\ell * f)(\theta) = \int_{\mathbb{T}} D_\ell(\theta - \theta') f(\theta')\, d\theta'$. As is well known, this action produces an *approximation of the identity*, that is, $\lim_{k\uparrow\infty}(D_k * f)(\theta) = f(\theta)$ at points of continuity of f. Figuring more prominently in what follows is a different kernel with that property. We define

$$C_\ell(\theta) = 2^{2\ell}\frac{(\ell!)^2}{(2\ell)!}\cos^{2\ell}(\theta) + 2^{2\ell+1}\frac{(\Gamma(\ell+\frac{3}{2}))^2}{(2\ell+1)!}\cos^{2\ell+1}(\theta) \tag{1.3}$$

in which Γ is the usual Gamma function. The result is:

Theorem 1.2. *For statistics of the spectral angles, let $f, g \in L^2(\mathbb{T})$. We then have*

$$\begin{aligned}\mathrm{Cov}_N\Big(X(f), X(g)\Big) &= N\Big((f * \tilde{g})(0) - \frac{1}{N}\sum_{\ell=0}^{N-1}(C_\ell * f * \tilde{g})(0)\Big) \\ &\quad - \sum_{\ell=\lceil \frac{N}{2}\rceil}^{N-1}\Big((C_\ell - D_{2N-2\ell-2} * C_\ell) * (f * \tilde{g})\Big)(0)\end{aligned} \tag{1.4}$$

in which $\tilde{g}(\theta) = g(-\theta)$.

Theorem 1.2 should be compared with the results of [6] which discusses the allied question for Haar distributed eigenvalues in the Unitary group $U(N)$. As in [6], it is striking that the covariance depends on the test functions f and g only through $f * \tilde{g}$. Further, with the Cesaró averages $\frac{1}{N}\sum_{\ell=0}^{N-1} C_\ell(\theta)$ also forming approximation to the identity, the content of (1.4) is that the large N properties of the covariance are tied to the error in that approximation. In [6] fluctuations for $U(N)$ are shown to be related in the same way to the error in Fejér approximation.

Next we turn to the asymptotics. It is of interest to connect the growth of the covariance in N to the smoothness imposed on the underlying test functions f and g. For example, the $O(1)$ fluctuations of $X(f)$ for f a bit better than twice differentiable plus a growth condition in the Hermitian case (see [18]) and for $f \in H_2^{\frac{1}{2}}(\mathbb{T})$ in $U(N)$ (see [9] or [19]) are well known manifestations of the rigidity of random matrix ensembles. In line with those results, an immediate consequence of Theorem 1.1 is the following.

Theorem 1.3. *Let $f(z) = f(|z|)$ and $g(z) = g(|z|)$ lie in $C^{2,\delta}$ for $r = |z| \in (0, 1+\varepsilon)$ with some $\delta > 0$ and $\varepsilon > 0$ and otherwise be bounded. Then, as $N \uparrow \infty$, $\mathrm{Cov}_N(X(f), X(g)) = \frac{1}{2}\int_0^1 f'(r)g'(r)r\,dr + o(1)$. Moreover, the centered (but unnormalized) random variable $X(f) - E[X(f)]$ converges to a mean zero Gaussian with the corresponding variance.*

Actually, the central limit theorem for $X(f(|z|))$ described in Theorem 1.3 has already been discussed in [12]. We in fact follow the ideas therein; a proof is included here for the sake of completeness and rigor. While [12] employs the same product structure behind Theorem 1.1, the probabilistic interpretation which guides our results and allows for more to be accomplished (see below) was apparently not noticed. Also in [12], the author conjectures (and provides heuristics based on the associated log-gas) that the general (smooth) linear statistic should have order one Gaussian fluctuations in the large N limit. Our next result shows that this is not the case.

Theorem 1.4. *Let the function f of the phase have Fourier coefficients $\widehat{f}(k) = \int_{\mathbb{T}} e^{-ik\theta} f(\theta) d\theta$ satisfying $\sum_{-\infty}^{\infty} k^2 |\widehat{f}(k)|^2 < \infty$ and likewise for g. Then*

$$\mathrm{Cov}_N\Big(X(f), X(g)\Big) = \frac{\log N}{4} \int_{\mathbb{T}} f'(\theta) g'(\theta)\, d\theta + o(\log N).$$

If a bit more smoothness on either test function is assumed, in particular if $\sum_{-\infty}^{\infty} |k|^{2+\delta} |\widehat{f}(k)|^2 < \infty$ for some $\delta > 0$, the $o(\log N)$ error may be replaced by an $O(1)$.

Of course, an $O(\log N)$ fluctuation is rigid compared to the $O(N)$ characterizing an ensemble of independent particles. Still, this marked difference between radial and angular statistics was not anticipated.

Changing focus, an important class of non-smooth test functions to consider are indicators of given sets within the spectrum, thus providing an eigenvalue count. Setting this case of the *number statistic* apart, let us denote $\#\Theta[\alpha, \beta] = X(\mathbb{1}_{\{\arg(z_k) \in [\alpha,\beta]\}})$. and $\#\mathcal{I}[a, b] = X(\mathbb{1}_{\{|z_k| \in [a,b]\}})$. The next result concerns the growth of the variances of $\#\Theta[\alpha, \beta]$ and $\#\mathcal{I}[a, b]$, with added attention to the *mesoscopic scales*, that is, when $|\alpha - \beta|$ (or $|a - b|$) $\downarrow 0$ as $N \uparrow \infty$. It is found there is a break in the behavior depending on whether $\sqrt{N}|\alpha - \beta|$ (or $\sqrt{N}|a - b|$) tends to ∞ or not.

For the angular number statistic the result is:

Theorem 1.5. *Note first by rotation invariance it is enough to consider the symmetric interval $[-\alpha/2, \alpha/2]$. If $\alpha > 0$ remains fixed while $N \uparrow \infty$, then*

$$\mathrm{Var}_N\Big(\#\Theta[-\frac{\alpha}{2}, \frac{\alpha}{2}]\Big) = \sqrt{N}\, \frac{1}{\pi^{3/2}} + O(\log N). \tag{1.5}$$

If instead $\alpha = \alpha(N) \downarrow 0$, $N\alpha(N) \uparrow \infty$, there are the following cases:

$$\mathrm{Var}_N\Big(\#\Theta[-\frac{\alpha}{2}, \frac{a}{2}]\Big) = \begin{cases} \sqrt{N} \frac{1}{\pi^{3/2}} (1 + o(1)) & \textit{if } \lim_{N\uparrow\infty} \sqrt{N}\alpha = \infty \\ \sqrt{N} \frac{1}{\pi^{3/2}} I_{arg}(\beta) + O(\log N) & \textit{if } \lim_{N\uparrow\infty} \sqrt{N}\alpha = \beta > 0 \\ N\alpha(1 + o(1)) & \textit{if } \lim_{N\uparrow\infty} \sqrt{N}\alpha = 0, \end{cases} \tag{1.6}$$

where, $I_{arg}(\beta) \geq 0$ and is such that $lim_{\beta\downarrow 0} I_{arg}(\beta) = 0$ while $\lim_{\beta\uparrow\infty} I_{arg}(\beta) = 1$; its definition may be found below in (3.11).

The radial case has a related structure:

Theorem 1.6. *For fixed* $[a, b] \in (0, 1)$ *we have that,*

$$\mathrm{Var}_N\Big(\#\mathcal{I}[a,b]\Big) = \sqrt{N}\left(\frac{a+b}{\sqrt{\pi}}\right) + O(1). \tag{1.7}$$

If rather $(b-a) \downarrow 0$ *while* $N(b-a) \uparrow \infty$*, then*

$$\mathrm{Var}_N\Big(\#\mathcal{I}[a,b]\Big) = \begin{cases} \sqrt{N}\frac{a}{\sqrt{\pi}}(1+o(1)) & \text{if } \lim_{N\uparrow\infty}\sqrt{N}(b-a) = \infty \\ \sqrt{N}\frac{a}{\sqrt{\pi}} I_{mod}(c) + O(1) & \text{if } \lim_{N\uparrow\infty}\sqrt{N}(b-a) = c > 0 \\ N(b^2-a^2)(1+o(1)) & \text{if } \lim_{N\uparrow\infty}\sqrt{N}(b-a) = 0, \end{cases} \tag{1.8}$$

in which, similar to the above, I_{mod} *(defined in (4.13)) is non-negative and satisfies* $\lim_{c\downarrow 0} I_{mod}(c) = 0$, $\lim_{c\uparrow\infty} I_{mod}(c) = 1$.

Remark. Note that in the radial case the variance asymptotics are modulated by the limiting radial distribution, the corresponding angular distribution being uniform.

We may now invoke the Costin-Lebowitz Theorem (see Section 5). First proved in [7] for the sine kernel and extended by A. Soshnikov in [29] and [30], in the present setting the theorem implies that: with $N\alpha \uparrow \infty$ (respectively $N(b-a) \uparrow \infty$)

$$\frac{\#\Theta[-\alpha,\alpha] - E_N[\#\Theta[-\alpha,\alpha]]}{\sqrt{\mathrm{Var}_N(\#\Theta[-\alpha,\alpha])}} \quad \left(\text{respectively } \frac{\#\mathcal{I}[a,b] - E_N[\#\mathcal{I}[a,b]]}{\sqrt{\mathrm{Var}_N(\#\mathcal{I}[a,b])}}\right) \tag{1.9}$$

tends to a Gaussian random variable of mean zero and variance one. Since we may also compute $E_N[\#\Theta[-\alpha,\alpha]] = N(2\alpha)$ and $E_N[\#\mathcal{I}[a,b]] = N(b-a) + O(1)$, an interesting observation regarding the mesoscopic scales is that when $\lim_{N\uparrow\infty}\sqrt{N}\alpha < \infty$ (or $\lim_{N\uparrow\infty}\sqrt{N}(b-a) < \infty$) we have that $\mathrm{Var}_N[\#\Theta[-\alpha,\alpha]] = O(E_N[\#\Theta[-\alpha,\alpha]])$ (or $\mathrm{Var}_N[\#\mathcal{I}[a,b]] = O(E_N[\#\mathcal{I}[a,b]])$) as in the case of independent particles. Contrariwise, if $\lim_{N\uparrow\infty}\sqrt{N}\alpha = \infty$ (and $\lim_{N\uparrow\infty}\sqrt{N}(b-a) = \infty$) the slow growth of the variance compared to the mean indicates rigidity. These remarks are suggestive of a possible asymptotic independence taking place on the small scales. In this context we add that if one considers the outlying or edge eigenvalues (which may be considered a small scale) it is found that $\mathrm{Var}_N[\#\mathcal{I}[1,\infty]] = O(E_N[\#\mathcal{I}[1,\infty]]) = O(\sqrt{N})$. Further evidence of an asymptotic independence of the edge is contained in [25] and [26] which show the largest eigenvalues in absolute value respond to a limit theorem shared by independent sequences (see also Section 5).

Finally, consider the limiting Gaussian field $\Theta_\infty[\alpha,\beta]$ on $\mathbb{T}$ (or $\mathcal{I}_\infty[a,b]$ on $[0,1]$) resulting from the properly normalized statistics (1.9) for fixed arguments (α,β) (or (a,b)) as $N \uparrow \infty$. The corresponding correlations are anticipated to be such that, taking the first case, $\Theta_\infty[\alpha,\beta]$ and $\Theta_\infty[\delta,\gamma]$ are independent if either $[\alpha,\beta]$ and $[\delta,\gamma]$ are disjoint or one is properly contained in the other; those intervals can positively or negatively correlated when they share a single endpoint. This rather odd structure was first observed by Wieand (see [32]) for $U(N)$ and has since been rediscovered in this and other ensembles by a variety of methods. If we wanted only to obtain this limit, an extension of Costin-Lebowitz could be employed (see [30]). However, another advantage of the formulas obtained in Theorems 1.1 and 1.2 is that we can directly describe the correlations at all large but finite values of N. Our last result is then as follows: note the shape of the second order terms which speak to the differing degree of rigidity in the two cases,

Theorem 1.7. *The covariance of* $\#\Theta[\alpha, \beta]$ *and* $\#\Theta[\delta, \gamma]$ *satisfies*

$$\mathrm{Cov}_N\left(\frac{\#\Theta[\alpha,\beta]}{(N\pi^3)^{1/4}}, \frac{\#\Theta[\delta,\gamma]}{(N\pi^3)^{1/4}}\right) = \begin{cases} \pm O(\frac{1}{\sqrt{N}}) & \textit{if } [\alpha,\beta] \subset [\delta,\gamma] \textit{ or } [\alpha,\beta] \cap [\delta,\gamma] = \varnothing, \\ +\frac{1}{2} + O(\frac{\log N}{\sqrt{N}}) & \textit{if } \alpha = \delta\ (\beta \neq \gamma) \textit{ or } \beta = \gamma\ (\alpha \neq \delta), \\ -\frac{1}{2} + O(\frac{\log N}{\sqrt{N}}) & \textit{if } \beta = \delta\ (\alpha \neq \gamma) \textit{ or } \alpha = \gamma\ (\beta \neq \delta). \end{cases}$$

For $\#\mathcal{I}[a, b]$ *and* $\#\mathcal{I}[c, d]$ *the result is*

$$\mathrm{Cov}_N\left(\frac{\#\mathcal{I}[a,b]}{(N\pi)^{1/4}}, \frac{\#\mathcal{I}[c,d]}{(N\pi)^{1/4}}\right) = \begin{cases} \pm O(e^{-c_1 N}) & \textit{if } [a,b] \subset [c,d] \textit{ or } [a,b] \cap [c,d] = \varnothing, \\ +\frac{a}{2} + O(\frac{1}{\sqrt{N}}) & \textit{if } a = c\ (b \neq d) \textit{ or } b = d\ (a \neq c), \\ -\frac{b}{2} + O(\frac{1}{\sqrt{N}}) & \textit{if } b = c\ (a \neq d) \textit{ or } a = d\ (b \neq c) \end{cases}$$

where $c_1 > 0$ *and depends on* a, b, c, d.

As for the remainder of this note: in the next section we derive our basic formulas for the covariance (Theorems 1.1 and 1.2), the verification of the asymptotic statements are found in Section 3 (angular case) and Section 4 (radial case). Section 5 extends the present results in the radial setting to Ginibre's Gaussian quaternion ensemble. For the reader's convenience Section 6 serves as a brief appendix on the Costin-Lebowitz theorem.

Remark. In a recent extension of Costin-Lebowitz, [29] shows that (modulo technicalities) linear statistics in general determinantal point fields satisfy a central limit theorem as long as the variance grows faster than a small positive power of the expectation. The estimates below then imply a central limit theorem at rate $N^{1/4}$ for $X(f)$ (either in the angular or radial case) when f has anywhere a jump discontinuity. Unfortunately, the $O(\log N)$ growth of the variance in Theorem 1.4 is not fast enough from this point of view. A proof of the central limit theorem in the case of smooth angular statistics will appear elsewhere.

2. Covariance Formulas

Appropriate row and column operations in (1.1) allow the eigenvalue density to be re-expressed as in

$$dP_N(z_1, z_2, \ldots z_N) = \frac{1}{N!} \det\Big[K_N(z_\ell, \bar{z_k})\Big]_{1 \le \ell, k \le N} \mu_N(dz_1) \cdots \mu_N(dz_N) \quad (2.1)$$

with the Hermitian kernel

$$K_N(z,\bar{w}) = \sum_{\ell=0}^{N-1} \frac{N^{\ell+1}}{\pi \ell!} z^\ell \bar{w}^\ell,$$

see [24]. As an operator on $L^2(\mathbb{C}, \mu_N(dz))$, $K_N(z,\bar{w})$ projects onto the span of the first N polynomials orthogonal with respect to that weight (those being just the monomials $1, z, z^2$ etc.). This fact explains the rules $\int_{\mathbb{C}} K_N(z,\bar{z})\mu_N(dz) = N$ and $\int_{\mathbb{C}} K_N(z,\bar{v})K_N(v,\bar{w})\mu_N(dv) = K_N(z,\bar{w})$, from which it follows that the marginal densities of P_N are given by:

$$\begin{aligned} &P_N^k(z_1,\dots,z_k) \qquad (2.2)\\ &:= \Big\{ \int_{\mathbb{C}^{N-k}} \frac{1}{N!} \det\Big[K_N(z_i,\bar{z_j})\Big]_{1\le i,j\le N} \mu_N(dz_{k+1})\cdots\mu_N(dz_N) \Big\} \\ &\quad \mu_N(dz_1)\cdots\mu_N(dz_k) \\ &= \frac{(N-k)!}{N!} \det\Big[K_N(z_i,\bar{z_j})\Big]_{1\le i,j\le k} \mu_N(dz_1)\cdots\mu_N(dz_k). \end{aligned}$$

(normalized differently these objects are also known as the correlation functions). Applying the integrating-out rules behind (2.2) along with the symmetries of the integrands below we find

$$E_N\Big[X(f)\Big] = N \int_{\mathbb{C}^N} f(z_1)dP_N = \int_{\mathbb{C}} f(z)K_N(z,\bar{z})\mu_N(dz) \qquad (2.3)$$

and

$$\begin{aligned} \mathrm{Cov}_N\Big(X(f), X(g)\Big) &= \int_{\mathbb{C}} f(z)g(z)K_N(z,\bar{z})\mu_N(dz) \qquad (2.4)\\ &\quad - \int_{\mathbb{C}}\int_{\mathbb{C}} f(z)g(w)K_N(z,\bar{w})K_N(w,\bar{z})\mu_N(dz)\mu_N(dw). \end{aligned}$$

With that our expressions for the covariance may be established.

Proof of Theorem 1.1. In polar coordinates, $z = (r,\alpha)$, $w = (s,\beta)$, you find that

$$\int_{\mathbb{C}} f(|z|)g(|z|)K_N(z,\bar{z})\mu_N(dz) = \sum_{\ell=0}^{N-1} N^{\ell+1}\frac{2}{\ell!}\int_0^\infty f(r)g(r)r^{2\ell+1}e^{-Nr^2}\,dr,$$

while in the second integral of (2.4) only the diagonal term survives:

$$\begin{aligned} &\int_{\mathbb{C}}\int_{\mathbb{C}} f(|z|)g(|w|)|K_N(z,\bar{w})|^2\mu_N(dz)\mu_N(dw) \\ &= \sum_{\ell=0}^{N-1}\Big[4\frac{N^{2\ell+2}}{(\ell!)^2}\int_0^\infty\int_0^\infty f(r)g(s)r^{2\ell+1}s^{2\ell+1}e^{-Nr^2}e^{-Ns^2}\,dr\,ds\Big]. \end{aligned}$$

The connection with (1.2) is then made after a change of variables: the distribution $p(dr) := (N^{\ell+1}/\ell!)r^\ell e^{-Nr}\,dr$ being recognized as that for a sum of $(\ell+1)$ independent exponential random variables, each of mean $\frac{1}{N}$. The proof is finished.

Proof of Theorem 1.2. From here on we denote $\phi(\theta) = (f * \tilde{g})(\theta)$. Also, in the present derivation we consider N to be even just to have things set; the formulas are easily adjusted for odd values of N.

To arrive at (1.4), it is now the radial component which is integrated out. Doing so in the first term of (2.4) is immediate:

$$\int_{\mathbb{C}} (fg)(\arg(z)) K_N(z,\bar{z}) \mu_N(dz) = \int_{\mathbb{T}} f(\theta) g(\theta) d\theta \int_0^\infty K_N(r,r) r e^{-Nr^2} dr = N\phi(0).$$

While for the second term we have

$$\int_{\mathbb{C}}\int_{\mathbb{C}} f(\arg(z)) g(\arg(w)) |K_N(z,\bar{w})|^2 \mu_N(dz)\mu_N(dw) = \int_{\mathbb{T}}\int_{\mathbb{T}} f(\alpha) g(\beta) \times \Big[\sum_{0\le k,\ell\le N-1} e^{i\alpha(k-\ell)} e^{-i\beta(k-\ell)} \frac{(\int_0^\infty r^{k+\ell+1} e^{-Nr^2} dr)^2}{k!\ell!} N^{k+\ell+2} \Big] d\alpha d\beta.$$

With $\widehat{\phi}(k) = \int_{\mathbb{T}} e^{ik\theta}\phi(\theta)d\theta = \widehat{f}(k)\widehat{g}(-k)$ the k-th Fourier coefficient of ϕ, this produces

$$\mathrm{Cov}_N\Big(X(f), X(g)\Big) = N\phi(0) - \sum_{0\le k,\ell\le N-1} \Big[\Gamma\Big(\frac{k+\ell}{2}+1\Big)\Big]^2 \frac{\widehat{\phi}(k-\ell)}{k!\ell!} := N\phi(0) - S_N^0(\phi). \tag{2.5}$$

as a preliminary form of our covariance formula. The obvious next step is make the substitution $(k+\ell, k-\ell) = (2n, 2m)$ for k and ℓ of the same parity and $(k+\ell, k-\ell) = (2n+1, 2m-1)$ otherwise. The result of that move is

$$S_N^0(\phi) = \sum_{(m,n)\in\mathcal{D}_N} \frac{(n!)^2\widehat{\phi}(2m)}{(n-m)!(n+m)!} + \sum_{(m,n)\in\mathcal{D}_N'} \frac{(\Gamma(n+\frac{3}{2}))^2\widehat{\phi}(2m-1)}{(n+m)!(n-m+1)!} \tag{2.6}$$

in which the sets $\mathcal{D}_N$ and $\mathcal{D}_N'$ in the (m,n)-plane are described as follows. $\mathcal{D}_N$ consists of two triangles: the lower defined by the points $(0,0)$, $(N/2-1, N/2-1)$ and $(-N/2+1, N/2-1)$, and the upper defined by $(0, N-1)$, $(N/2-1, N/2)$ and $(-N/2+1, N/2)$. The set $\mathcal{D}_N' = \mathcal{D}_N \cup \mathcal{D}_N''$ where the latter consists of all integer points lying on the lines between $(1,0)$ and $(N/2+1, N/2-1)$ and between $(N/2+1, N/2)$ and $(1, N-1)$.

To evaluate (2.6), first consider the sum over the lower half of $\mathcal{D}_N$. In particular start with a fixed $n \le N/2-1$ so that the corresponding sum over m ranges fully between $-n$ and $+n$. This object may be computed by the following observation. Denote by $\zeta_1, \zeta_2, \dots$ a sequence of independent ± 1 Bernoulli random variables with mean E, and notice that: with $c(n) = 2^{2n}\binom{2n}{n}^{-1}$,

$$\sum_{m=-n}^{n} \frac{(n!)^2 \widehat{\phi}(2m)}{(n-m)!(n+m)!} = c(n) \sum_{m=-n}^{n} \widehat{\phi}(2m) \binom{2n}{n+m} 2^{-2n} \tag{2.7}$$
$$= c(n) E\Big[\widehat{\phi}(\zeta_1 + \cdots + \zeta_{2n})\Big] = c(n) \int_{\mathbb{T}} \Big(E\Big[e^{\sqrt{-1}\zeta_1\theta}\Big]\Big)^{2n} \phi(\theta) d\theta$$
$$= c(n) \int_{\mathbb{T}} \phi(\theta) \cos^{2n}(\theta)\, d\theta.$$

In the same way we have

$$\sum_{m=-n}^{n+1} \frac{(\Gamma(n+3/2))^2 \widehat{\phi}(2m+1)}{(n-m)!(n+m+1)!} = 2^{2n+1} \frac{(\Gamma(n+3/2))^2}{(2n+1)!} \int_{\mathbb{T}} \phi(\theta) \cos^{2n+1}(\theta)\, d\theta \tag{2.8}$$

for the typical term in lower half of the $\mathcal{D}'_N$ sum. Pairing the final expressions in (2.7) and (2.8) in n brings out the sum of $(C_n * \phi)(0)$ for n from 0 to $N/2 - 1$.

The stated form of the covariance is then obtained by extending the summation over m in the upper reaches of the n variables (extending to a "big triangle"), in order to produce a summation of C_n over all $n \le N - 1$. That is, we write

$$S_N^0(\phi) := S_N(\phi) + S'_N(\phi)$$

with $S_N(\phi) = \sum_{n=0}^{N-1} (C_n * \phi)(0)$ and $S'_N(\phi)$ the error involved in throwing too much into the mix. To complete the picture, we spell out the contribution to $S'_N(\phi)$ connected to the equal-parity, or $\mathcal{D}_N$, sum. That reads: with $p_{2n}(m) = \binom{2n}{n+m} 2^{-2n}$ and $c(n)$ as before,

$$\sum_{n=\frac{N}{2}}^{N-1} c(n) \sum_{m=-n}^{n} \widehat{\phi}(2m) \Big[1 - \mathbb{1}_{\{|m| \le N-1-n\}}\Big] p_{2n}(m)$$
$$= \sum_{n=\frac{N}{2}}^{N-1} c(n) \sum_{m=-n}^{n} \Big[\Big(\phi - (D_{2N-2n-2} * \phi)\Big)\widehat{\ }(2m)\Big] p_{2n}(m)$$
$$= \sum_{n=\frac{N}{2}}^{N-1} 2^{2n} \binom{2n}{n}^{-1} \int_{\mathbb{T}} \cos^{2n}(\theta) \Big[\phi(\theta) - (D_{2N-2n-2} * \phi)(\theta)\Big] d\theta.$$

The sum corresponding to the $\mathcal{D}'_N$ part of our expression is similar; that its outcome is as advertised should be clear. The proof is finished.

3. Angular Statistics

That the kernels C_ℓ form an approximation of the identity is seen from: (1) the evaluation $\int_{\mathbb{T}} C_\ell(\theta)\, d\theta = 1$ (based on $\int_{\mathbb{T}} \cos^{2\ell}(\theta)\, d\theta = 2^{-2\ell}\binom{2\ell}{\ell}$ and $\int_{\mathbb{T}} \cos^{2\ell+1}(\theta)\, d\theta = 0$), and (2), the basic estimate

$$\int_{-\varepsilon}^{\varepsilon} C_\ell(\theta) d\theta = \sqrt{\frac{\ell}{\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-\ell\theta^2} d\theta + o(1) = 1 + o(1) \tag{3.1}$$

valid for any small $\varepsilon > 0$. The proofs of the main theorems hinge on the justification of exactly this sort of standard Laplace asymptotics, though it is necessary to extract the precise form of the error terms. For the applications in mind, C_ℓ acts on one of two classes of test functions. We restrict our attention to those cases:

Lemma 3.1. *Let $h(\theta)$ be twice continuously differentiable on $\mathbb{T}$. Then, as $\ell \uparrow \infty$,*

$$(C_\ell * h)(0) = \int_{\mathbb{T}} h(\theta) C_\ell(\theta)\, d\theta = h(0) + \frac{h''(0)}{4\ell} + \frac{\nu(\ell, h)}{\ell} + O\Big(\frac{1}{\ell^{3/2}}\Big). \quad (3.2)$$

In general, $\nu(\ell, h) = o(1)$ for large values of ℓ. This may be improved to $\nu(\ell, h) = O(\ell^{-\delta/2})$ if $h''(\theta)$ happens to be Hölder continuous of order $\delta > 0$ in a neighborhood of the origin.

Let instead $h(\theta)$ be a continuous, piecewise-linear function of $\theta \in \mathbb{T}$. In that case,

$$(C_\ell * h)(0) = h(0) + \frac{h'_+(0)}{2\sqrt{\pi\ell}} - \frac{h'_-(0)}{2\sqrt{\pi\ell}} + O\Big(\frac{1}{\ell^{3/2}}\Big) \quad (3.3)$$

is the appropriate estimate for $\ell \uparrow \infty$. Here h'_+ and h'_- indicate right and left derivatives.

A comment is perhaps in order at this point as to the comparison between the kernel arising and being analyzed here and the better known Dirichlet kernel. The content of the second equality in (3.1) is that $C_\ell(\theta)$ concentrates its mass in a neighborhood of width $O(\ell^{-1/2})$. The concentration of the Dirichlet is sharper, taking place in a neighborhood of order ℓ^{-1}. Returning to the discussion after Theorem 1.2 this explains why the fluctuations for the angles in Ginibre's ensemble may be expected to be larger than those for $U(N)$.

Proof of Lemma 3.1. For h in either class, it may be assumed from the start that $h(0) = 0$ ($\int_{\mathbb{T}} C_\ell(\theta)\, d\theta = 1$). Also, throughout the proof we will make use of the simple estimates

$$2^{2\ell} \binom{2\ell}{\ell}^{-1} = \sqrt{\pi\ell}\Big(1 + \frac{1}{8\ell} + O\Big(\frac{1}{\ell^2}\Big)\Big) \text{ and } \frac{\Gamma(\ell + \frac{1}{2})}{\ell!} = \frac{1}{\sqrt{\ell}}\Big(1 - \frac{1}{8\ell} + O\Big(\frac{1}{\ell^2}\Big)\Big), \quad (3.4)$$

which follow from Stirling's approximation in the form $\Gamma(\ell) = \ell^{\ell - \frac{1}{2}} e^{-\ell} \sqrt{2\pi}(1 + \frac{1}{12}\ell^{-1} + O(\ell^{-2}))$.

Making the abbreviation

$$\begin{aligned} C_\ell(\theta) &= \Big[2^{2\ell} \frac{(\ell!)^2}{(2\ell)!}\Big] \cos^{2\ell}(\theta) + \Big[2^{2\ell+1} \frac{\Gamma^2(\ell + \frac{3}{2})}{(2\ell+1)!}\Big] \cos^{2\ell+1}(\theta) \\ &:= a(\ell) \cos^{2\ell}(\theta) + b(\ell) \cos^{2\ell+1}(\theta), \end{aligned}$$

the rule $\Gamma(k) = (k-1)\Gamma(k-1)$ provides the identity $b(\ell) = a(\ell) \times (\ell + \frac{1}{2}) \times (\Gamma(\ell+\frac{1}{2})/\ell!)^2$, and from (3.4) we note that $a(\ell) \simeq b(\ell) \simeq \sqrt{\pi \ell}$. The basic behavior of C_ℓ for $\ell \uparrow \infty$ may then be described as a sum of two (approximate) point masses of weight $\simeq 1/2$ at $\theta = 0$ and the difference of two such masses at the common point $\theta = \pm\pi$. Our first step is to dispense of the cancellation which must take place at the latter point.

Consider in particular the integral $\int_{\pi/2}^{\pi} h(\theta) C_\ell(\theta)\, d\theta$, with that over $\theta \in [-\pi, -\pi/2]$ being treated similarly. Denoting $g(\theta) = h(\theta) - h(\pi)$ we write

$$\int_{\pi/2}^{\pi} C_\ell(\theta) h(\theta)\, d\theta = h(\pi) \int_{\pi/2}^{\pi} C_\ell(\theta) d\theta + (a(\ell) - b(\ell)) \int_{\pi/2}^{\pi} \cos^{2\ell}(\theta) g(\theta)\, d\theta$$
$$+ b(\ell) \int_{\pi/2}^{\pi} \cos^{2\ell}(\theta) g(\theta)(1+\cos(\theta))\, d\theta. \tag{3.5}$$

This object is to be dominated by a constant multiple of $\ell^{-3/2}$. The first term on the right hand side responds to an exact calculation plus by an appeal to (3.4):

$$\int_{\pi/2}^{\pi} C_\ell(\theta) d\theta = a(\ell) \int_0^{\pi/2} \cos^{2\ell}(\theta) d\theta - b(\ell) \int_0^{\pi/2} \cos^{2\ell+1}(\theta) d\theta$$
$$= \frac{\pi}{2} - \frac{1}{2}\Big(2^{2\ell} \binom{2\ell}{\ell}^{-1} \times \frac{\Gamma(\ell+\frac{1}{2})}{\ell!}\Big)^2 = O\Big(\frac{1}{\ell^2}\Big).$$

As for terms two and three, our regularity assumptions on h imply that $|g(\theta)| \le c_1|\pi - \theta|$ and so $|g(\theta)(1+\cos(\theta))| \le c_2 |\pi - \theta|^3$ for $\theta \in [\pi - \alpha, \pi]$ with a small $\alpha > 0$. Also on that same range we have the bound $\cos^{2\ell}(\theta) \le e^{-c_3 \ell (\theta - \pi)^2}$ for $c_3 = c_3(\alpha) > 0$, while a fixed distance away from $\theta = \pi$ (or $\theta = 0$) $\cos^{2\ell}(\theta)$ is exponentially small. It follows that

$$\Big| \int_{\pi/2}^{\pi} \cos^{2\ell}(\theta) g(\theta) d\theta \Big| \le c_4 \frac{1}{\ell} \ \text{ and } \ \Big| \int_{\pi/2}^{\pi} \cos^{2\ell}(\theta) g(\theta)(1+\cos(\theta))\, d\theta \Big| \le c_5 \frac{1}{\ell^2}.$$

Using again (3.4) shows that

$$a(\ell) - b(\ell) = 2^{2\ell} \binom{2\ell}{\ell}^{-1} \Big[1 - (\ell + \frac{1}{2}) \Big(\frac{\Gamma(\ell+\frac{1}{2})}{\ell!}\Big)^2\Big] = O\Big(\frac{1}{\sqrt{\ell}}\Big)$$

and completes our consideration of (3.5).

Turning our attention to the contribution to the integral from the vicinity of $\theta = 0$, take first the case that $h \in C^2$ and use Taylor's theorem with remainder to write

$$h(\theta) = h'(0)\theta + \frac{1}{2} h''(0)\theta^2 + \int_0^\theta \Big(h''(\alpha) - h''(0)\Big)(\theta - \alpha) d\alpha, \tag{3.6}$$

(recall $h(0) = 0$). Since $C_\ell(\theta)$ is even there is no contribution from the first term on the right hand side: $\int_{-\pi/2}^{\pi/2} C_\ell(\theta)\theta\, d\theta = 0$. Continuing to the quadratic term, one may again cut down the range of integration. Needing to be more precise this time around we restrict to $|\theta| \le \ell^{-1/4}$, the ensuing error still exponentially small. Further, on that range it holds that $|1-\cos(\theta)| \le \theta^2$ and $|1-e^{\ell\theta^2}\cos^{2\ell}(\theta)| \le 4\ell\theta^4$. Using these estimates in $\int_{-\pi/2}^{\pi/2} C_\ell(\theta)\theta^2\,d\theta = \int_{-\ell^{-1/4}}^{\ell^{-1/4}} C_\ell(\theta)\theta^2 d\theta + O(e^{-c_6\ell})$ and then restoring the limits of integration over the whole line, we have

$$\begin{aligned}\int_{-\pi/2}^{\pi/2} C_\ell(\theta)\theta^2 &= \frac{a(\ell)}{2\pi}\int_{-\pi/2}^{\pi/2}\theta^2(1+\cos(\theta))\cos^{2\ell}(\theta)\,d\theta \\ &\quad + \frac{a(\ell)-b(\ell)}{4\pi}\int_{-\pi/2}^{\pi/2}\theta^2\cos^{2\ell+1}(\theta)\,d\theta \\ &= \sqrt{\frac{\ell}{\pi}}\int_{-\infty}^{\infty}\theta^2 e^{-\ell\theta^2}\,d\theta \\ &\quad + O\left(\int_{-\infty}^{\infty}(\ell^{-1/2}\theta^2+\ell^{1/2}\theta^4+\ell^{3/2}\theta^6)e^{-\ell\theta^2}\,d\theta\right) \\ &= \frac{1}{2\ell} + O\left(\frac{1}{\ell^{5/2}}\right).\end{aligned}$$

There remains the contribution due the last term on the right of (3.6). The kernel C_ℓ is to be integrated against

$$\left|\int_0^\theta \Big(h''(\alpha)-h''(0)\Big)(\theta-\alpha)d\alpha\right| \le \theta^2 \max_{-1\le\alpha\le1}|h''(\theta\alpha)-h''(0)| := \theta^2\tilde{\nu}(\theta,h) \tag{3.7}$$

from $-\pi/2 \le \theta \le \pi/2$. By continuity, the nonnegative function $\tilde{\nu}$ satisfies $\tilde{\nu} \downarrow 0$ with $\theta \downarrow 0$. Running through the above arguments will explain why the integral in question is bounded by a constant multiple of

$$\sqrt{\ell}\int_{-\ell^{-1/4}}^{\ell^{-1/4}}\theta^2\nu(\tilde{\theta,h})e^{-\ell\theta^2}\,d\theta \le \frac{2}{\ell}\max_{|\theta|\le\ell^{-1/4}}\tilde{\nu}(\theta,h) := \frac{1}{\ell}\nu(\ell,h)$$

with $\lim_{\ell\uparrow\infty}\nu(\ell,h)=0$ as promised in (3.2). The comment following that display stems from the fact that if h'' is Hölder continuous with exponent $\delta>0$, the right hand side of (3.7) may be replaced by a constant times $\theta^{2+\delta}$. The ensuing integral in the last display would then decay like $\ell^{-(1+\delta/2)}$.

The details behind (3.3) are much the same. One may assume that $h(\theta) = h_+\theta^+ - h_-\theta^-$ in some neighborhood of the origin and the leading order comes out of $\int_0^\infty \theta e^{-\ell\theta^2}d\theta = 1/2\ell$. The proof is finished.

Proof of Theorem 1.4. As f and $\tilde{g}$ each have one L^2 derivative, their convolution $\phi(\theta) = (f*\tilde{g})(\theta)$ is twice continuously differentiable. Recalling the covariance formula

$$\mathrm{Cov}_N\Big(X(f),X(g)\Big) = N\phi(0) - S_N(\phi) - S_N'(\phi)$$

for

$$S_N(\phi) = \sum_{\ell=0}^{N-1} \int_{\mathbb{T}} C_\ell(\theta)\phi(\theta)d\theta \ \text{ and } \ S'_N(\phi) = \sum_{N/2\le \ell \le N} C_\ell * (\phi - D_{2N-2\ell-2}\phi)(0),$$

we are in the setting of the first half of Lemma 3.1. The statement (3.2) implies that

$$S_N(\phi) = N\phi(0) + \frac{1}{4}\phi''(0)\sum_{\ell=1}^{N}\frac{1}{\ell} + \sum_{\ell=1}^{N}\frac{\nu(\ell,\phi)}{\ell} + O(1). \tag{3.8}$$

(Here, the $O(1)$ term includes both the sum of the order $\ell^{-3/2}$ errors in (3.2) as well as the sum of first several terms $(C_\ell * \phi)(0)$ where the estimate is of no affect.) The second term produces the $\log N$ figuring into the leading order growth of covariance. Generally $\nu(\ell,\phi) = o(1)$ and we can conclude only that the third term $\sum_{\ell=1}^{N}\frac{\nu(\ell,\phi)}{\ell} = o(\log N)$. The point of the closing remark in the theorem's statement is that the proposed additional decay on $|\widehat{f}(k)|$ (or $|\widehat{g}(k)|$) for $k \uparrow \infty$ implies Hölder continuity of ϕ''. In particular, for positive δ satisfying $\delta/2 < 1$,

$$\begin{aligned}\left|\phi''(\theta) - \phi''(0)\right| &\le \sum_{k=-\infty}^{\infty} k^2|\widehat{f}(k)||\widehat{g}(k)|\left|1 - e^{ik\theta}\right| \\ &\le 4|\theta|^{\delta/2}\sum_{k=-\infty}^{\infty}|k|^{2+\delta/2}|\widehat{f}(k)||\widehat{g}(k)|\end{aligned}$$

in which the last sum is finite by Schwartz's inequality. By the related remark in the statement of Lemma 3.1 it follows that the term $\ell^{-1}\nu(\ell,\phi)$ is summable in this case. In either case, $N\phi(0) - S_N(\phi)$ already accounts for the advertised asymptotic behavior of the covariance (note that $\phi''(0) = -\int_{\mathbb{T}} f'(\theta)g'(\theta)d\theta$).

The remaining (error) term $S'_N(\phi)$ is seen to be of constant order under the present smoothness assumptions. We write

$$S'_N(\phi) = \sum_{N/2\le\ell\le N}\ \sum_{|k|>2N-2\ell-2}\widehat{C}_\ell(k)\widehat{\phi}(k), \tag{3.9}$$

and introduce the simple facts: $\widehat{C}_\ell(k) = 0$ for $k > 2\ell+1$, and, in general, $|\widehat{C}_\ell(k)| \le 2$. It follows that

$$\begin{aligned}|S'_N(\phi)| &\le c_1 \sum_{0\le\ell\le N/2}\ \sum_{N-2\ell\le k\le N+2\ell}\Big(|\widehat{\phi}(k)| + |\widehat{\phi}(-k)|\Big) \\ &\le c_2\sum_{0\le k\le N}|k|\Big(|\widehat{\phi}(k)| + |\widehat{\phi}(-k)|\Big) \le c_3\Big(\sum_{k=-\infty}^{\infty}|k|^4|\widehat{\phi}(k)|^2\Big)^{1/2} < \infty\end{aligned}$$

after changing variables and then the order of summation in inequalities one and two, an application of Cauchy-Schwartz for the third, and at last the given fact: $\sum|k|^4|\widehat{\phi}(k)|^2 \le (\sum|m|^2|\widehat{f}(m)|^2)\,(\sum|n|^2|\widehat{g}(-n)|^2) < \infty$. The proof is finished.

Proof of Theorem 1.5. When $f(\theta) = g(\theta) = \mathbb{1}_{[-\alpha/2,\alpha/2]}(\theta)$, the convolution := $\phi_\alpha(\theta)$ is the tent function

$$\phi_\alpha(\theta) = \frac{1}{2\pi}\int_{-\frac{\alpha}{2}}^{\frac{\alpha}{2}} \mathbb{1}_{[-\frac{\alpha}{2},\frac{\alpha}{2}]}(\theta' - \theta)\,d\theta' = \frac{1}{2\pi}\left[\alpha - |\theta|\right]\mathbb{1}_{[-\alpha,\alpha]}(\theta),$$

and the main contribution to the variance is then

$$N\phi_\alpha(0) - S_N(\phi_\alpha) = N\frac{\alpha}{2\pi} - \frac{\alpha}{4\pi^2}\sum_{\ell=0}^{N-1}\int_{-\alpha}^{\alpha} C_\ell(\theta)\,d\theta + \frac{1}{4\pi^2}\sum_{\ell=0}^{N-1}\int_{-\alpha}^{\alpha}|\theta| C_\ell(\theta)\,d\theta. \tag{3.10}$$

Let us first consider this object for the various situations, α fixed or $\alpha \downarrow 0$ as $N \uparrow \infty$, and return to the error term $S'_N(\phi_\alpha)$ at the end.

From the proof of Lemma 3.1 it follows that for $\alpha > 0$ fixed $\frac{1}{2\pi}\int_{-\alpha}^{\alpha} C_\ell(\theta)d\theta = 1 + O(\ell^{-3/2})$, and so the first two terms in (3.10) contribute something of constant order. On the other hand, from the conclusion of the second lemma we have that

$$\frac{1}{4\pi^2}\sum_{\ell=0}^{N-1}\int_{-\alpha}^{\alpha}|\theta| C_\ell(\theta)d\theta = \frac{1}{2\pi}\sum_{\ell=1}^{N-1}\left[\frac{1}{\sqrt{\pi\ell}} + O\left(\frac{1}{\ell^{3/2}}\right)\right] + O(1) = \frac{1}{\pi^{3/2}}\sqrt{N} + O(1),$$

which is the advertised result (1.5) for order one regions of the phase.

If $\alpha \downarrow 0$ with $N \uparrow \infty$, we start in the borderline case in which $\alpha(N) = \beta/\sqrt{N}$ with a fixed $\beta > 0$. At this scale the quantities in (3.10) go over into approximating Riemann sums. First there is,

$$\begin{aligned}\frac{1}{4\pi^2}\sum_{\ell=1}^{N}\int_{-\alpha(N)}^{\alpha(N)}|\theta| C_\ell(\theta)d\theta &= \frac{\sqrt{N}}{2\pi^{3/2}} \times \frac{1}{N}\sum_{\ell=1}^{N}\sqrt{\frac{N}{\ell}}\int_{-\beta\sqrt{\frac{\ell}{N}}}^{\beta\sqrt{\frac{\ell}{N}}}|\theta| e^{-\theta^2}d\theta + O(1)\\ &= \sqrt{N}\frac{1}{\pi^{3/2}}\int_0^1 \frac{1}{\sqrt{x}}\int_0^{\beta\sqrt{x}}\theta e^{-\theta^2}\,d\theta dx\\ &\quad + O\left(\sqrt{N}\sum_{\ell=1}^{N}\int_{\frac{\ell-1}{N}}^{\frac{\ell}{N}}\left[F_\beta(\ell/N) - F_\beta(x)\right]dx\right) + O(1)\end{aligned}$$

for $F_\beta(x) = \frac{1-e^{-\beta x}}{\sqrt{x}}$, and that error term is controlled (for whatever β) as in

$$\begin{aligned}&\sqrt{N}\sum_{\ell=1}^{N}\int_{\frac{\ell-1}{N}}^{\frac{\ell}{N}}\left[F_\beta(\ell/N) - F_\beta(x)\right]dx\\ &\le \frac{1}{\sqrt{N}}\int_0^1 |F'_\beta(x)|dx = O\left(\frac{1}{\sqrt{N}}\right).\end{aligned}$$

(The difference between an integral and its Riemann sum being less than $1/N\times$ the total variation of the integrand.) In a similar way we find that

$$N\frac{\alpha(N)}{2\pi} - \frac{\alpha(N)}{4\pi^2}\sum_{\ell=0}^{N-1}\int_{-\alpha(N)}^{\alpha(N)} C_\ell(\theta)\,d\theta$$
$$= \sqrt{N}\left[\frac{\beta}{2\pi} - \frac{\beta}{2\pi}\int_0^1\int_{-\beta\sqrt{x}}^{\beta\sqrt{x}} \frac{e^{-\theta^2}}{\sqrt{\pi}}\,d\theta dx\right] + O(1).$$

Combined, the last three displays translate as,

$$\begin{aligned} N\phi_{\frac{\beta}{\sqrt{N}}}(0) - S_N(\phi_{\frac{\beta}{\sqrt{N}}}) &= \sqrt{N}\frac{1}{\pi^{3/2}}\left[\int_0^1 \frac{1-e^{-\beta x^2}}{2\sqrt{x}}\,dx + \beta\int_0^1\int_{\beta\sqrt{x}}^{\infty} e^{-\theta^2}\,d\theta dx\right] \\ &\quad + O(1) \\ &:= \sqrt{N}\frac{1}{\pi^{3/2}} I_{arg}(\beta) + O(1). \end{aligned} \tag{3.11}$$

Next, if $\sqrt{N}\alpha(N)\uparrow\infty$ one can clearly apply the same steps with the point of view that $\beta=\beta(N)\uparrow\infty$. Noting that

$$I_{arg}(\beta) = 1 - \frac{c_1}{\beta^{1/4}} + \frac{c_2}{\beta}$$

as $\beta\uparrow\infty$ completes the explanation of that case. Finally, if $\sqrt{N}\alpha(N)\downarrow 0$ while maintaining $N\alpha(N)\uparrow\infty$, things are a bit different. In the range $\ell\le N$ and $\sqrt{N}\theta = o(1)$,

$$C_\ell(\theta) = 2\sqrt{\pi\ell}\left(1-\ell\theta^2 + O(\ell^2\theta^4)\right)\left(1+O\left(\frac{1}{\ell}\right)\right) + O\left(\frac{1}{\ell^{3/2}}\right)$$

Substituting the above into (3.10) we find that

$$\frac{\alpha}{4\pi^2}\sum_{\ell=0}^{N-1}\int_{-\alpha}^{\alpha} C_\ell(\theta)d\theta - \frac{1}{4\pi^2}\sum_{\ell=0}^{N-1}\int_{-\alpha}^{\alpha}|\theta|C_\ell(\theta)d\theta = N\alpha\left[\frac{\sqrt{N}\alpha}{3\pi^{3/2}} - \frac{(\sqrt{N}\alpha)^3}{30\pi^{3/2}}\right] + O(1),$$

which anyway is $o(N\alpha)$, identifying the leading order in this case as $N(\alpha/2\pi)$.

To complete the proof, we need to track the growth of S_N'. A bit of good fortune is that one may compute $\widehat{\phi}_\alpha(k) = \frac{\sin^2(k\alpha)}{\pi k^2} \ge 0$ which leaves us to control

$$|S_N'(\phi_\alpha)| \le c_1 \sum_{N/2\le\ell\le N}\ \sum_{2N-2\ell\le|k|\le 2\ell} |\widehat{C}_\ell(k)|\left(\frac{\sin^2(k\alpha)}{k^2}\right), \tag{3.12}$$

see (3.9). Simply bounding $|\widehat{C}_\ell(k)|\sin^2(k\alpha)$ by a constant will produce the estimate $S_N'(\phi_\alpha) = O(\log N)$ stated for the cases where $\liminf_{N\uparrow\infty}\sqrt{N}\alpha > 0$.

If however, $\alpha = o(1/\sqrt{N})$ we need to, and can, do better. Bring in the additional estimate $|\widehat{C}_\ell(k)| \le c_2\sqrt{\ell}|k|^{-1}$ courtesy Lemma 3.1 which has bite for $|k| >> \ell$.

With this in mind the sum is split as follows: with $\ell^*(N) \simeq N - \frac{1}{2}\sqrt{N}$ the point where $2N - 2\ell = \sqrt{\ell}$ over $\ell > N/2$,

$$
\begin{aligned}
|S_N'(\phi_\alpha)| \; &\le \left[\sum_{\substack{N/2 \le \ell \le \ell^*(N) \\ 2N-2\ell \le |k| \le 2\ell}} + \sum_{\substack{\ell^*(N) \le \ell \le N \\ \sqrt{\ell} \le |k| \le 2\ell}} + \sum_{\substack{\ell^*(N) \le \ell \le N \\ 1 \le |k| \le \sqrt{\ell}}} \right] |\widehat{C}_\ell(k)| \left(\frac{\sin^2(k\alpha)}{k^2} \right) \\
&:= \mathcal{A}_N + \mathcal{B}_N + \mathcal{C}_N .
\end{aligned}
$$

Moving from left to right, and first bounding the summand by $\sqrt{\ell}|k|^{-3}$ we have

$$
\mathcal{A}_N \le c_3 \sum_{N/2 \le \ell \le N - \frac{1}{4}\sqrt{N}} \frac{\sqrt{\ell}}{(N-\ell)^2} \le c_4 \sqrt{N} \sum_{\frac{1}{4}\sqrt{N} \le \ell \le N/2} \frac{1}{\ell^2} = O(1),
$$

as well as

$$
\mathcal{B}_N \le c_5 \sum_{N-\sqrt{N} \le \ell \le N} \; \sum_{\sqrt{\ell} \le k \le \infty} \frac{\sqrt{\ell}}{k^3} \le c_6 \sum_{N-\sqrt{N} \le \ell \le N} \frac{1}{\sqrt{\ell}} = O(1).
$$

For $\mathcal{C}_N$ we finally use the fact that $\alpha \downarrow 0$: as $k^{-2}\sin^2(\alpha k) \le 2\alpha^2$ for $\alpha = o(1/\sqrt{N})$ and $k \le \sqrt{N}$ we conclude

$$
\mathcal{C}_N \le c_7 \, \alpha^2 \sum_{N-\sqrt{N} \le \ell \le N} \sqrt{\ell} \le c_8 \, (N\alpha^2) = o(1)
$$

as needed. In closing we note that a careful review of the $\mathcal{C}_N$ will demonstrate that the stated $O(\log N)$ error term in the cases where $\liminf_{N\uparrow\infty} \sqrt{N}\alpha > 0$ cannot be improved. The proof is finished.

Proof of Theorem 1.7 (angular case). In each on the four cases, the covariance of the pair $(\#\Theta[\alpha,\beta], \#\Theta[\delta,\gamma])$ is connected to a $\phi_k(\theta)$ $(k = 1,2,3,4)$ which is either constant in a neighborhood of the origin or possesses a corner at that point. In particular, when the intervals in question are either disjoint or else $[\alpha,\beta]$ sits inside of $[\delta,\gamma]$ we have

$$
\begin{aligned}
\phi_{1,2}(\theta) &:= \frac{1}{2\pi} \int_{\delta-\theta}^{\gamma-\theta} \mathbb{1}_{[\alpha,\beta]}(\theta')\, d\theta' \\
&= \begin{cases} 0 & \text{on } [\beta-\delta, 2\pi+\alpha-\gamma] \text{ when } \alpha < \beta < \delta < \gamma \\ \frac{\beta-\alpha}{2\pi} & \text{on } [\delta-\alpha, \gamma-\beta] \text{ when } \delta < \alpha < \beta < \gamma \end{cases}
\end{aligned}
$$

The relevant computation is then: for $\phi(\theta) = \phi(0)$ over $-\varepsilon \le \theta \le \varepsilon$,

$$
\begin{aligned}
N\phi(0) - \sum_{\ell=0}^{N-1} \int_{\mathbb{T}} C_\ell(\theta)\phi(\theta)\, d\theta &= N\phi(0) \\
&\quad - \phi(0) \sum_{\ell=0}^{N-1} \left[\int_{-\ell\varepsilon}^{\ell\varepsilon} e^{-\theta^2} \frac{d\theta}{\sqrt{\pi}} + O(\ell^{-3/2}) \right] = O(1).
\end{aligned}
$$

As for $S'_N(\phi_{1,2})$, we will estimate $f_m^{1,2}(\theta) := (\phi_{1,2} - D_m * \phi_{1,2})(\theta)$ (which is evaluated at $m = 2N - 2l - 2$ and then summed over ℓ) near the origin. In particular, we have already seen that by adding a constant if necessary we can assume that $\phi_{1,2}(\theta) = 0$ in $[-\varepsilon, \varepsilon]$, and we want to show that f_m can be made uniformly small throughout say $[-\varepsilon/2, \varepsilon/2]$. Now,

$$\begin{aligned} f_m^{1,2}(\theta) &= \int_T \frac{\phi_{1,2}(\theta - \theta')}{\sin(\frac{1}{2}\theta')} \sin\left((m + \frac{1}{2}\right)\theta')\, d\theta' \\ &:= \int_T g_{1,2}(\theta', \theta) \sin\left((m + \frac{1}{2})\theta'\right)\, d\theta', \end{aligned}$$

and if $|\theta| \le \varepsilon/2$ then $\phi_{1,2}(\theta - \theta') = 0$ for all $|\theta'| \le \varepsilon/2$. As otherwise $\phi_{1,2}$ is piece-wise linear, the second derivative of $g_{1,2}(\theta')$ will certainly be integrable away from $[-\varepsilon/2, \varepsilon/2]$. Therefore, integrating by parts twice will show that $f_m^{1,2}(\theta) = O(m^{-2})$ throughout $|\theta| \le \varepsilon/2$. This in turn will imply that

$$\left(C_\ell * f_{2N-2\ell-2}^{1,2}\right)(0) = O\left(\frac{1}{(N-\ell)^2}\right) + O(\ell^{-3/2})$$

which is summable over the appropriate range of ℓ: $S'_N(\phi_{1,2})$ is of constant order. The de-correlation at the proposed rate is thus established.

Taking next the situation that $\alpha < \beta = \delta < \gamma$, it is enough to note that there exists an $\varepsilon > 0$ with

$$\phi_3(\theta) := \frac{1}{2\pi} \int_{\alpha-\theta}^{\beta-\theta} \mathbb{1}_{[\beta,\gamma]}(\theta')\, d\theta' = \begin{cases} \frac{-\theta}{2\pi} & \text{for } \theta \in [-\varepsilon, 0] \\ 0 & \text{for } \theta \in [0, \varepsilon], \end{cases}$$

whereas if $\alpha = \delta$ and $\beta \ne \gamma$ (take $\beta < \gamma$ for convenience) we may say that

$$\phi_4(\theta) := \frac{1}{2\pi} \int_{\alpha-\theta}^{\beta-\theta} \mathbb{1}_{[\alpha,\gamma]}(\theta')\, d\theta' = \begin{cases} \frac{\beta-\alpha}{2\pi} & \text{for } \theta \in [-\varepsilon', 0] \\ \frac{\beta-\alpha-\theta}{2\pi} & \text{for } \theta \in [0, \varepsilon'] \end{cases}$$

for some $\varepsilon' > 0$. It will then follow from the second half of Lemma 3.1 that

$$-\sum_{\ell=0}^{N} \int_{\mathbb{T}} C_\ell(\theta)\phi_3(\theta)\, d\theta = +\frac{1}{2}\sqrt{\frac{N}{\pi^3}} + O(1) \quad \text{and} \quad -\sum_{\ell=0}^{N} \int_{\mathbb{T}} C_\ell(\theta)\phi_4(\theta)\, d\theta$$
$$= -\frac{1}{2}\sqrt{\frac{N}{\pi^3}} + O(1).$$

Furthermore, re-tooling that part of the previous proof (for the variance estimates) related to the S'_N term will show that $S'_N(\phi_{3,4}) = O(\log N)$; the corner at the origin being the cause of this increased growth. The proof is finished.

4. Radial Statistics

We begin with the proof of Theorem 1.3 which relies on the product formula

$$E\left[\prod_{k=1}^{N} f(|z_k|)\right] = \prod_{k=1}^{N} E\left[f\left(\sqrt{\frac{1}{N} s_k}\right)\right]. \tag{4.1}$$

where you will recall that $s_k = \eta_1 + \cdots + \eta_k$ with the η's independent exponentials of mean one. While related to the covariance formula (1.2), it is easier to see (4.1) as direct consequence of (5.1) derived in the next section.

Proof of Theorem 1.3. Let h satisfy the assumed boundedness and regularity. For convenience we consider the test function $r \to h(r^2)$ with a Gaussian limit for $\sum_{k=1}^{N} h(|z_k|^2) - E[\sum_{k=1}^{N} h(|z_k|^2)]$ proved by computing Laplace transforms. In particular, the formula (4.1) is applied with $f(r) = \exp[\lambda h(r^2)]$ and λ a parameter lying in a fixed neighborhood of the origin to find that

$$\begin{aligned} E\left[\exp\{\lambda[X(h) - EX(h)]\}\right] &= E\left[\exp\left\{\lambda[\sum_{k=1}^{N}(h(|z_k|^2) - Eh(|z_k|^2))]\right\}\right] \\ &= \prod_{k=1}^{N} E\left[\exp\{\lambda[h(s_k/N) - Eh(s_k/N)]\}\right]. \end{aligned} \tag{4.2}$$

We wish to expand each appearance of $h(s_k/N)$ as in $h(s_k/N) = h(k/N) + h'(k/N)((s_k - k)/N) + \frac{1}{2}h''(k/N)((s_k - k)/N)^2 + R_h(s_k/N, k/N)$. Toward this, make the definitions $\bar{h}(s_k/N) = h(s_k/N) - E[h(s_k/N)]$,

$$\mathcal{M}_{N,k}(h) := \exp\left\{\lambda h'(k/N)\left(\frac{s_k - k}{N}\right)\right\},$$

and

$$\mathcal{R}_{N,k}(h) := \exp\left\{\frac{\lambda}{2N^2} h''(k/N)\left((s_k - k)^2 - k\right) + \lambda \bar{R}_h(s_k/N, k/N)\right\}$$

in which $\bar{R}_h(s_k/N, k/N) = R_h(s_k/N, k/N) - E[R_h(s_k/N, k/N)]$, and the right hand side of (4.2) may be continued as the product of

$$\begin{aligned} E\left[\exp[\lambda\bar{h}(s_k/N)]\right] &= E\left[\mathcal{M}_{N,k}(h)\right] \\ &+ E\left[\left(e^{\lambda\bar{h}(s_k/N)} - \mathcal{M}_{N,k}(h)\right)\mathbb{1}_{\{\frac{s_k}{N} \geq 1+\varepsilon\}}\right] \\ &+ E\left[\mathcal{M}_{N,k}(h)\left(\mathcal{R}_{N,k}(h) - 1\right)\mathbb{1}_{\{\frac{s_k}{N} \leq 1+\varepsilon\}}\right] \end{aligned} \tag{4.3}$$

from $k = 1$ to N.

Computing the product of the first term in the right of (4.3) produces the Gaussian structure: choosing λ so that $|\lambda| \max_{r\in[0,1]} |h'(r)| < 1/2$,

$$\log \prod_{k=1}^{N} E\left[\mathcal{M}_{N,k}(h)\right] = -\sum_{k=1}^{N}\left(\lambda h'(k/N)\frac{k}{N} + k\log\left(1 - \frac{\lambda h'(k/N)}{N}\right)\right)$$
$$= -\frac{1}{2}\sum_{k=1}^{N} k\left(\lambda h'(k/N)\frac{1}{N}\right)^2 + O\left(\sum_{k=1}^{N}\frac{k}{N^3}\right) = -\frac{1}{2}\lambda^2\int_0^1 (h'(r))^2\, r dr + o(1).$$

Now similar estimates to those used in the last display will show that $E\left[\mathcal{M}_{N,k}(h)\right]$ is bounded above and below by positive constants independent of k or N, and so the proof is completed by establishing that

$$\sum_{k=1}^{N}\Big\{E\left[\left(e^{\lambda\bar{h}(s_k/N)} + \mathcal{M}_{N,k}(h)\right)\mathbb{1}_{\{\frac{s_k}{N}\geq 1+\varepsilon\}}\right]$$
$$+E\left[\mathcal{M}_{N,k}(h)(\mathcal{R}_{N,k}(h)-1)\mathbb{1}_{\{\frac{s_k}{N}\leq 1+\varepsilon\}}\right]\Big\} \to 0$$

as $N \uparrow \infty$.

For the first part, let $c_1 = \sup_{0\leq r<\infty} h(r)$ and write

$$E\left[\left(e^{\lambda\bar{h}(s_k/N)} + \mathcal{M}_{N,k}(h)\right)\mathbb{1}_{\{\frac{s_k}{N}\geq 1+\varepsilon\}}\right]$$
$$\leq e^{2|\lambda|c_1} P\left(\frac{s_k}{N}\geq 1+\varepsilon\right) + E\left[e^{s_k/N}\mathbb{1}_{\{\frac{s_k}{N}\geq 1+\varepsilon\}}\right]$$
$$\leq e^{2|\lambda|c_1} P\left(\frac{s_k}{N}\geq 1+\varepsilon\right) + \left(\frac{N}{N-1}\right)^{k/2}\sqrt{P\left(\frac{s_k}{N}\geq 1+\varepsilon\right)}$$

in which the last line is easily bounded by a constant multiple of $e^{-\gamma N}$ for $\gamma > 0$ depending on ε but not on k. For the remaining part, note first that on $\{s_k/N \leq 1+\varepsilon\}$ we have that $\mathcal{M}_{N,k}(h) \leq \exp(1+\varepsilon)$ and also that the exponent in $\mathcal{R}_{N,k}(h)$ may be bounded uniformly in k and N. Since $|e^a - 1| \leq |a|e^{|a|}$ we have that

$$E\left[\mathcal{M}_{N,k}(h)|\mathcal{R}_{N,k}(h)-1|\mathbb{1}_{\{\frac{s_k}{N}\leq 1+\varepsilon\}}\right]$$
$$\leq c_2\frac{k}{N^2}\, E\left[\,\left|\left(\frac{s_k-k}{\sqrt{k}}\right)^2 - 1\right|\,\right] + c_3 E\left[|R_h(s_k/N, k/N)|\right] \qquad (4.4)$$

with constants c_2 and c_3. Clearly $E[|(\frac{s_k-k}{\sqrt{k}})^2 - 1|] = o(1)$ for $k \uparrow \infty$. It follows that the first term on the right of (4.4), summed from $k = 1$ to N, is also $o(1)$ as $N \uparrow \infty$. For the final term in (4.4), by the regularity of h we may assume that $|R_h(s_k/N, k/N)| \leq c_4|\frac{s_k-k}{N}|^{2+\alpha}$ (having applied this estimate while the restriction $\{s_k/N \leq 1+\varepsilon\}$ was still in place). The upshot is that

$$\sum_{k=1}^{N} E\left[|R_h(s_k/N, k/N)|\right] \le \frac{c_4}{N^{2+\alpha}} \sum_{k=1}^{N} E\left[|s_k - k|^{2+\alpha}\right]$$
$$\le c_5 \frac{1}{N^{2+\alpha}} \sum_{k=1}^{N} k^{\frac{2+\alpha}{2}} \simeq N^{-\alpha/2}$$

as desired. Here, the second inequality may be arrived at rather directly in the present setting of exponential random variables. More generally, an inequality due to Marcinkiewicz and Zygmund (which may be found in [8]) will explain why $E[|\sum_{\ell=1}^{k} x_\ell|^p] \simeq k^{p/2}$ for $p > 2$ and $\{x_\ell\}$ independent copies of a mean-zero x_1 for which $E[|x_1|^p] < \infty$.

Finally, the statement concerning the asymptotic behavior of the covariance may be gleaned from the above by setting $h = \beta f + \gamma g$ for parameters β and γ. Having controlled moment generating functions, we have adequate tightness to draw conclusions on the limiting covariance as well as higher moments. The proof is finished.

Moving on to the covariance of the number statistic $\#\mathcal{I}[\cdot,\cdot]$, the basic error estimate required has in fact already been done for us. It is embodied in the classical Edgeworth expansion providing corrections to the local central limit theorem. The statement is: with $p(a, M)$ the density of the random variable $\frac{1}{\sqrt{M}} \sum_{\ell=1}^{M} (\eta_\ell - 1)$ at a,

$$\sup_{-\infty < a < \infty} \left| p(a, M) - \frac{1}{\sqrt{2\pi}} e^{-a^2/2} - \frac{1}{\sqrt{M}} \frac{1}{\sqrt{\pi/2}} a^3 e^{-a^2/2} \right| = O(M^{-1}), \quad (4.5)$$

see for example [5], Corollary 19.4.

Proof of Theorem 1.6. Start with a fixed interval $[a, b]$ $(0 < a \le b < 1)$ of the moduli for $N \uparrow \infty$. Denote by $p_{N,k}(A) = P(\frac{1}{N} s_k \in A)$. The variance of $\#\mathcal{I}[a, b]$ then reads

$$\begin{aligned} \mathrm{Var}_N\,(\#\mathcal{I}[a, b]) &= \sum_{k=1}^{N} p_{N,k}([a^2, b^2]) \left(1 - p_{N,k}([a^2, b^2])\right) \\ &:= \sum_{k=1}^{N} p_{N,k}([a^2, b^2]) p_{N,k}([a^2, b^2]^C), \end{aligned} \quad (4.6)$$

and the (classical) central limit theorem explains why it should be only $O(\sqrt{N})$ neighborhoods of $k = Na^2$ and $k = Nb^2$ that contribute to the sum on the right. Breaking up the sum in accordance with that observation, we write:

$$\mathrm{Var}_N\,(\#\mathcal{I}[a, b]) = \left[\sum_{|k - Na^2| \le \delta_N \sqrt{N}} + \sum_{|k - Nb^2| \le \delta_N \sqrt{N}} + \sum{}' \right] p_{N,k}([a^2, b^2]) \, p_{N,k}([a^2, b^2]^C).$$

Here $\sum'$ has in general three components ($k \in [0, Na^2 - \delta_N\sqrt{N}] \cup [Na^2 + \delta_N\sqrt{N}, Nb^2 - \delta_N\sqrt{N}] \cup [Nb^2 + \delta_N\sqrt{N}, N]$), and δ_N is some rate chosen so that this sum is negligible for $N \uparrow \infty$. We spell out the estimate for the first component, the rest being nearly identical.

With $\{\lambda_k\}$ a sequence of positive constants less than one to be determined, we have:

$$\begin{aligned}\sum_{k\le Na^2-\delta_N\sqrt{N}} p_{N,k}([a^2,b^2])p_{N,k}([a^2,b^2]^C) &\le \sum_{k\le Na^2-\delta_N\sqrt{N}} P(s_k \ge Na^2)\\ \le \sum_{\delta_N\sqrt{N}\le k\le Na^2} \exp\left[-N\left(a^2\lambda_k + (a^2 - \frac{k}{N})\log(1-\lambda_k)\right)\right]\\ \le \sum_{\delta_N\sqrt{N}\le k\le Na^2} e^{-k^2/2Na^2} \le c_1\sqrt{N}\int_{\delta_N}^{\infty} e^{-c^2/2a^2}\,dc \le c_2\frac{\sqrt{N}e^{-\delta_N^2/2a^2}}{\delta_N}.\end{aligned} \tag{4.7}$$

The second inequality invokes Chernov's bound plus the change of variables $k \to \lfloor Na^2\rfloor - k$. The third follows from the choice of $\lambda_k = \frac{k}{Na^2}$, and the fourth from a Taylor expansion. The conclusion of the last line is that letting $\delta_N = \sqrt{\log N}$ implies $\sum' p_{N,k}(1-p_{N,k}) = o(1)$.

Turning to the sums of consequence (over $|k - Na^2| \le \delta_N\sqrt{N}$ and $|k - Nb^2| \le \delta_N\sqrt{N}$), it is evident that they will have equal contributions. Consider the first sum and rewrite the typical appearance of $p_{N,k}([a^2,b^2])$ as: with $k = Na^2 + \ell$ and ℓ ranging between $\pm\delta_N\sqrt{N}$,

$$\begin{aligned}&P\left(\frac{1}{N}s_{Na^2+\ell} \in [a^2,b^2]\right)\\ &= P\left(\frac{1}{\sqrt{\lfloor Na^2+\ell\rfloor}}\sum_{1\le m\le Na^2+\ell}(\eta_m - 1) \ge -\frac{\ell}{\sqrt{N}a} + \varepsilon_N(\ell,a)\right)\\ &- O\left(e^{-c_{a,b}N}\right).\end{aligned} \tag{4.8}$$

Here

$$\begin{aligned}|\varepsilon_N(\ell,a)| &\le \left|\frac{\ell}{\sqrt{Na^2}} - \frac{\ell}{\sqrt{Na^2+\ell}}\right| + \frac{1}{\sqrt{N}}\left(Na^2 + \ell - \lfloor Na^2+\ell\rfloor\right)\\ &\le c_1\frac{1}{\sqrt{N}} + c_2\frac{\ell^2}{N^{3/2}},\end{aligned}$$

and the (exponentially small) error term stems from a standard large deviation estimate. Applying (4.5), we may then continue (4.8) as in

$$
\begin{aligned}
&p_{N,Na^2+\ell}([a^2,b^2])\\
&= \frac{1}{\sqrt{2\pi}}\int_{-\frac{\ell}{\sqrt{N}a}+\varepsilon_N}^{\infty} e^{-x^2/2}dx + \sqrt{\frac{2}{\pi N}}\int_{-\frac{\ell}{\sqrt{N}a}+\varepsilon_N}^{\infty} x^3 e^{-x^2/2}dx + O\left(\frac{\log N}{N}\right)\\
&= \frac{1}{\sqrt{2\pi}}\int_{-\frac{\ell}{\sqrt{N}a}}^{\infty} e^{-x^2/2}dx + \sqrt{\frac{2}{\pi N}}\int_{-\frac{\ell}{\sqrt{N}a}}^{\infty} x^3 e^{-x^2/2}dx\\
&+ O\left(\frac{\ell^2}{N^{3/2}}e^{-\ell^2/N}\right) + O\left(\frac{\log N}{N}\right).
\end{aligned}
\tag{4.9}
$$

The first equality is achieved by first cutting down the integration to $|c| \le \log N$ and then applying the Edgeworth expansion. The error produced by this process of limiting the range of integration and later restoring it up to $+\infty$ may be controlled by the same large deviation procedure used already twice before and the Gaussian estimate $\int_a^\infty e^{-x^2/2}dx \le a^{-1}e^{-a^2/2}$. The second equality is self-explanatory.

The leading order of the variance is then read off by subtracting the square of (4.9) and summing in ℓ:

$$
\begin{aligned}
&\sum_{-\sqrt{N}\delta_N \le \ell \le \sqrt{N}\delta_N} p_{N,Na^2+\ell}([a^2,b^2])(1 - p_{N,Na^2+\ell}([a^2,b^2]))\\
&= \sqrt{N}\sum_{-\sqrt{N}\delta_N \le \ell \le \sqrt{N}\delta_N} \frac{1}{\sqrt{N}}\left\{\int_{\frac{\ell}{\sqrt{N}a}}^{\infty} e^{-x^2/2}\frac{dx}{\sqrt{2\pi}} - \left[\int_{\frac{\ell}{\sqrt{N}a}}^{\infty} e^{-x^2/2}\frac{dx}{\sqrt{2\pi}}\right]^2\right\}\\
&+ \mathcal{E}_1(N)\\
&= \sqrt{N}\int_{-\infty}^{\infty}\left\{\int_{\frac{x'}{a}}^{\infty} e^{-x^2/2}\frac{dx}{\sqrt{2\pi}} - \left[\int_{\frac{x'}{a}}^{\infty} e^{-x^2/2}\frac{dx}{\sqrt{2\pi}}\right]^2\right\}dx' + \mathcal{E}_1(N) + \mathcal{E}_2(N).
\end{aligned}
\tag{4.10}
$$

The error involved in taking the sum over to the (full) integral in the second line is easily controlled:

$$
\begin{aligned}
\mathcal{E}_2(N) &\le c_3\sqrt{N}\int_{\delta_N}^{\infty}\int_{x'}^{\infty} e^{-x^2/2}dx\,dx' + c_4\sqrt{N}\sum_{|\ell|\le\sqrt{N}\delta_N}\int_{\frac{\ell}{\sqrt{N}}}^{\frac{\ell+1}{\sqrt{N}}}\left|\int_{\frac{\ell}{\sqrt{N}}}^{x'} e^{-x^2/2}dx\right|dx'\\
&\le c_5\frac{\sqrt{N}}{\delta_N^2}e^{-\delta_N^2/2} + c_6\frac{1}{\sqrt{N}}\sum_{|\ell|\le\sqrt{N}\delta_N} e^{-\ell^2/2N} = o(1).
\end{aligned}
\tag{4.11}
$$

It remains to show that $\mathcal{E}_1(N) = O(1)$. Terms three and four in (4.9) and their squares, summed over the range $\pm\sqrt{N}\log N$ clearly remain finite as $N \uparrow \infty$. As for the second term, note that it is already of the form $N^{-1/2}$ times an integrable function of $\ell N^{-1/2}$. That is, it is already properly normalized so that the corresponding sums go over into (convergent) integrals as $N \uparrow \infty$: an instance of this being,

$$\frac{1}{\sqrt{N}} \sum_{|\ell| \le \sqrt{N}\delta_N} \int_{\frac{\ell}{\sqrt{N}}}^{\infty} x^3 e^{-x^2/2} \le c_7 \frac{1}{\sqrt{N}} \sum_{|\ell| \le \sqrt{N}\delta_N} (1 + \frac{\ell^2}{N}) e^{-\ell^2/2N}$$
$$\le c_8 \int_{-\infty}^{\infty} (1 + x^2) e^{-x^2/2} dx \tag{4.12}$$

after an integration by parts. Finally, the integral in line three of (4.10) is evaluated as $= \frac{a}{\sqrt{\pi}}$, and adding the corresponding term for $k \simeq Nb^2$ completes the proof of (1.7).

The proof for mesoscopic scales (1.8) builds on the above. Let us set, $\sqrt{N}(b - a) = c$, and consider first the cases that c is fixed (the critical case) or $c \uparrow \infty$ with N. Also, we take the interval in question to be $[a, a + c/\sqrt{N}]$ for a fixed a. While one may also consider symmetric intervals shrinking to a at $N = \infty$, the proof is more in lines with the above if we make the present convention. In particular, retracing those arguments through the first line of (4.10), it is evident that

$$\mathrm{Var}_N \left(\#\mathcal{I}\left[a, a + \frac{c}{\sqrt{N}}\right]\right)$$
$$= \sum_{|\ell| \le \sqrt{N}\delta_N} \left\{ \int_{\frac{\ell}{\sqrt{N}a}}^{2c + \frac{\ell}{\sqrt{N}a}} e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} - \left[\int_{\frac{\ell}{\sqrt{N}a}}^{2c + \frac{\ell}{\sqrt{N}a}} e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} \right]^2 \right\} + \mathcal{E}_1(c, N).$$

The error $\mathcal{E}_1(c, N)$ plays the same role as $\mathcal{E}_1(N)$ in (4.10) and is amenable to the same arguments. While certain aspects such as the estimate in (4.12) are now more elaborate due to the contribution of the upper limit in the integral, everything goes through just as easily if one keeps in mind that c is at most $o(\sqrt{N})$. Next the outer sum is replaced by an integral after an estimate nearly identical to (4.11). That is, the previous display may be continued, and the variance given by

$$\frac{\sqrt{N}a}{\sqrt{\pi}} \times \sqrt{\pi} \int_{-\infty}^{\infty} \left\{ \int_{x'}^{2c + x'} e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} - \left[\int_{x'}^{2c + x'} e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} \right]^2 \right\} dx' \tag{4.13}$$
$$:= \left(\frac{\sqrt{N}a}{\sqrt{\pi}} \right) I_{mod}(c)$$

up to order one errors. Recall that this holds for either c fixed, completing the proof in that case, or $c = c(N) \uparrow \infty$. In the latter setting, one may easily check that $I_{mod}(c)$ increases to its value $1 = I_{mod}(\infty)$ which explains the statement at those scales.

We finish with the case that $c \downarrow 0$. Here the Edgeworth expansion is used slightly differently and the growth of the variance is accounted for by

$$V_N(a, c) := \sum_{-N \le \ell \le N} \left\{ \int_0^{2c} e^{-\frac{1}{2}(x + \frac{\ell}{\sqrt{N}a})^2} \frac{dx}{\sqrt{2\pi}} - \left[\int_0^{2c} e^{-\frac{1}{2}(x + \frac{\ell}{\sqrt{N}a})^2} \frac{dx}{\sqrt{2\pi}} \right]^2 \right\}, \tag{4.14}$$

as may be understood upon further review of (4.10). In this expression it is only the first sum which plays a role in the asymptotics. Integrating by parts in that term we find that

$$\begin{aligned}
&\sum_{-N\le\ell\le N}\int_0^{2c} e^{-\frac{1}{2}(x+\frac{\ell}{\sqrt{N}a})^2}\frac{dx}{\sqrt{2\pi}} \\
&= 2c\sum_{-N\le\ell\le N}\frac{e^{-\frac{1}{2}(c+\frac{\ell}{\sqrt{N}a})^2}}{\sqrt{2\pi}} + O\left(\sqrt{N}c\int_0^{2c}\sum_{-N\le\ell\le N} e^{-\frac{1}{2}(x+\frac{\ell}{\sqrt{N}a})^2}dx\right) \\
&= 2c\sqrt{N}\left[\int_{-\infty}^{\infty} e^{-x^2/2a^2}\frac{dx}{\sqrt{2\pi}}\right](1+o(1)) + O(Nc^2) \\
&= \sqrt{N}(2ac)(1+o(1)) = N(b^2-a^2)(1+o(1))
\end{aligned}$$

after substituting back $c = \sqrt{N}(b-a)$. The same computation shows that the second, or sum of squares, term in (4.14) is indeed of lower order. The proof is finished.

Proof of Theorem 1.7 (radial case). Whatever the relation among a, b, c and d,

$$\begin{aligned}
&\mathrm{Cov}_N\,(\#\mathcal{I}[a,b],\#\mathcal{I}[c,d]) \\
&\quad = \sum_{k=1}^{N}\left[p_{N,k}([a,b]\cap[c,d]) - p_{N,k}([a,b])p_{N,k}([c,d])\right],
\end{aligned}$$

and the estimates obtained in the course of the previous proof may be revisited. Beginning with the situation $[a^2,b^2]\cap[c^2,d^2]=\varnothing$, the covariance reads as

$$\mathrm{Cov}_N\,(\#\mathcal{I}[a,b],\#\mathcal{I}[c,d]) = -\sum_{k=1}^{N} p_{N,k}([a^2,b^2])p_{N,k}([c^2,d^2]). \tag{4.15}$$

Taking for definiteness $b<c$, this sum may be bounded in a simple way as in

$$\begin{aligned}
\sum_{k=1}^{N} p_{N,k}([a^2,b^2])p_{N,k}([c^2,d^2]) \le &\sum_{1\le k\le N(c^2+b^2)/2} P(s_k\ge Nc^2) \\
&+\sum_{N(c^2+b^2)/2\le k\le N} P(s_k\le Nb^2)
\end{aligned}$$

with the first term on the right (being indicative of either) subject to

$$\sum_{k=1}^{N(c^2+b^2)/2} P(s_k\ge Nc^2)\le e^{-Nc^2}\sum_{k=1}^{N(c^2+b^2)/2}\left(\frac{Nc^2}{k}\right)^k e^k \le c_1 N e^{-Nc^2(1-c'+c'\log c')}.$$

with $c'=(c^2+b^2)/2c^2<1$, ensuring that $1-c'+c'\log c'>0$. This second bound just maximizes the summands, while the first is the usual trick of optimizing in Chebychev's inequality (as used in (4.7)). That the right hand side of (4.15) vanishes exponentially fast in N is established. The case when $[a^2, b^2] \subset [c^2, d^2]$ follows in kind. Then

$$\mathrm{Cov}_N\,(\#\mathcal{I}[a,b], \#\mathcal{I}[c,d]) = \sum_{k=1}^{N} p_{N,k}([a^2,b^2])\,(1 - p_{N,k}([c^2,d^2])), \quad (4.16)$$

but since now $[a, b]$ and $[c, d]^C$ are disjoint this expression is really quite the same as (4.15) and thus is also exponentially small (though positive) as $N \uparrow \infty$.

The last two cases in which the intervals share a boundary point, either $a < b = c < d$ or $a = c$ while say $b < d$, also go hand in hand. In the former we have: with $\delta_N = \sqrt{\log N}$,

$$\begin{aligned}
\mathrm{Cov}_N\,(\#\mathcal{I}[a,b], \#\mathcal{I}[b,d]) &= -\sum_{k=1}^{N} p_{N,k}([a^2,b^2])\,p_{N,k}([b^2,d^2]) \\
&= -\sum_{|k-Nb^2|\le\delta_N\sqrt{N}} p_{N,k}([0,b^2])\,p_{N,k}([b^2,\infty)) \\
&\quad + O\left(\sum_{|k-Nb^2|\ge\delta_N\sqrt{N}} p_{N,k}([0,b^2])\,p_{N,k}([b^2,\infty)\right) \\
&= -\frac{b}{2}\sqrt{N} + o\left(\frac{1}{\sqrt{\log N}}\right). \quad (4.17)
\end{aligned}$$

Here the choice of δ_N and the last line are just reprises of the proof of Theorem 1.6. Finally, if $a = c$ the covariance is

$$\begin{aligned}
\mathrm{Cov}_N\,(\#\mathcal{I}[a,b], \#\mathcal{I}[a,d]) &= \sum_{k=1}^{N} p_{N,k}([0,a^2])\,p_{N,k}([a^2,b^2]) \\
&\quad + \sum_{k=1}^{N} p_{N,k}([a^2,b^2])\,p_{N,k}([d^2,\infty]),
\end{aligned}$$

where we see that the first sum is of similar type to that in the first line of (4.17) and so produces the claimed $\frac{a}{2}\sqrt{N} + o(1)$ as $N \uparrow \infty$. The second term does not affect this outcome; having the same form as either (4.15) or (4.16) it is exponentially small. The proof is finished.

5. Extensions to the Quaternion Ensemble

Replacing the complex Gaussian entries in the definition of P_N with real quaternion Gaussians produces another solvable ensemble P_N^Q first by studied Ginibre. In this case the eigenvalue density has the form

$$\begin{aligned}
dP_N^Q(z_1, \cdots z_N) = \frac{1}{Z_N^Q} &\prod_{1\le\ell<k\le N} |z_\ell - z_k|^2 |z_\ell - \bar{z}_k|^2 \\
&\prod_{1\le k\le N} |z_k - \bar{z}_k|^2\, d\tilde{\mu}_N(z_1)\cdots d\tilde{\mu}_N(z_N)
\end{aligned}$$

where now $d\tilde{\mu}(z) = e^{-2N|z|^2} d\Re(z) d\Im(z)$ is the appropriately scaled weight to keep the limiting density of states uniform on $|z| \leq 1$. The introduction of the additional differences-squared terms breaks the rotation symmetry of P_N. It also breaks some of its analytic structure. While there again exist explicit formulas for the finite dimensional marginals or correlation functions, these are now expressed in terms of quaternion determinants or pfaffians. That is, P_N^Q is not a determinantal point field. Still, part of the previous discussion extends readily to this model for the reason that the radial components of the eigenvalues under P_N^Q have a nearly identical statistical description to those in the P_N ensemble.

The fact is that in either P_N or P_N^Q one may integrate out the angles from the start, producing a full joint density on the moduli $\{r_k = |z_k|, k = 1, \ldots, N\}$ (see [13] for related remarks). The resulting densities possess their own structure, and for certain analysis it is not important whether or not they derived from a determinantal setup. We start with P_N. Returning to the original expression (1.1), we simply expand the Vandermonde determinants to find that: with σ and ψ permutations on N letters and $\theta_k = \arg(z_k)$,

$$\begin{aligned} & dP_N(r_1, \cdots, r_N) \\ &= \frac{1}{Z_N} \sum_{\sigma,\psi} sgn(\sigma) sgn(\psi) \left\{ \prod_{k=1}^{N} \int_{-\pi}^{\pi} z_k^{\sigma_k - 1} \bar{z}_k^{\psi_k - 1} d\theta_k \right\} \times \prod_{k=1}^{N} e^{-N r_k^2} r_k \, dr_k \\ &= \frac{1}{N!} \sum_{\sigma} p_N^{(1)}(r_{\sigma_1}) p_N^{(2)}(r_{\sigma_2}) \cdots p_N^{(N)}(r_{\sigma_N}) \, dr_1 \cdots dr_N \end{aligned} \tag{5.1}$$

in which $p_N^{(\ell)}(r) = r^{2\ell-1} e^{-Nr^2}$, normalized to be a probability density on $r > 0$. For P_N^Q use is made of the identity

$$\det \left[z_\ell^{k-1} \;\; \bar{z}_\ell^{k-1} \right]_{1 \leq k \leq 2N, 1 \leq \ell \leq N} = \prod_{1 \leq \ell < k \leq N} |z_\ell - z_k|^2 |z_\ell - \bar{z}_k|^2 \prod_{1 \leq k \leq N} (z_k - \bar{z}_k).$$

Expanding the left hand side, multiplying by $\prod (z_k - \bar{z}_k) e^{-N|z_k|^2}$ and integrating over the $\{\theta_k\}$ variables produces: with ψ now a permutation on $2N$ letters,

$$\begin{aligned} dP_N^Q(r_1, \cdots, r_N) = \frac{1}{Z_N^Q} \sum_{\psi} sgn(\psi) \prod_{k=1}^{N} \left\{ \int_{-\pi}^{\pi} z_k^{\psi_{2k-1} - 1} \bar{z}_k^{\psi_{2k} - 1} (z_k - \bar{z}_k) d\theta_k \right\} \\ \times \prod_{k=1}^{N} e^{-2N r_k^2} r_k \, dr_k. \end{aligned}$$

Noting that only terms with $|\psi_{2k-1} - \psi_{2k}| = 1$ for all k contribute, we continue the computation as in

$$dP_N^Q(r_1, \cdots, r_N) = \frac{1}{N!} \sum_{\sigma} p_{2N}^{(2\cdot 1)}(r_{\sigma_1}) p_{2N}^{(2\cdot 2)}(r_{\sigma_2}) \cdots p_{2N}^{(2\cdot N)}(r_{\sigma_N}) \, dr_1 \cdots dr_N. \tag{5.2}$$

Here we are back to an N^{th} order permutation in σ and the $p_M^{(k)}$ have precisely the same definition as in their appearance in (5.1). For background on this and related computations we refer to [24], Appendix 18.

Certainly the covariance formula (1.2) may be derived directly from (5.1), circumventing the need for computing correlations. It follows from (5.2) that the covariances of general linear statistics on the moduli in the P_N^Q ensemble will satisfy the same formula, so long as each appearance of $\frac{1}{N}s_\ell$ is replaced with $\frac{1}{2N}s_{2\ell}$. The asymptotics follow suit with identical proofs:

Corollary 5.1. *Under P_N^Q the covariances of linear statistics of the radial components have the identical $N \uparrow \infty$ asymptotic behavior to their counterparts in the P_N ensemble described in Theorem 1.6 and the latter half of Theorem 1.7.*

Further, it should be clear that, by repeating the proof of Theorem 1.3, we may obtain Gaussian fluctuations for smooth radial linear statistics in P_N^Q. More interesting, for indicator type test function, while we may no longer quote the Costin-Lebowitz result directly, we do have the following.

Theorem 5.1. *The properly normalized radial number statistics $\#\mathcal{I}[a,b]$ under P_N^Q also satisfy a central limit theorem as long as $N(b-a) \uparrow \infty$ as $N \uparrow \infty$.*

The proof of Theorem 5.1 is actually a slight reworking of Costin-Lebowitz; for that reason it is postponed until the next section.

Remark. While given by a *permanent*, point processes of type (5.1) or (5.2) are not the bosonic fields appearing in the mathematical physics literature (see [10] and references therein). They are in fact simpler objects; it would be be interesting to know if something of their kind arose in other contexts.

6. Appendix

For completeness we include the statement and a sketch of the proof of the Costin and Lebowitz theorem. In keeping with the setting and notation used above, we say a measure P_N on $\mathbb{C}^N$ is a determinantal point field if all ℓ-dimensional correlation functions $\rho_{N,\ell}$ may be written as

$$\rho_{N,\ell}(z_1, z_2, \cdots, z_\ell) = \frac{N!}{(N-\ell)!} \det \left[K_N(z_i, \bar{z_j})\right]_{1 \le i,j \le \ell} \mu_N(dz_1) \cdots \mu_N(dz_\ell)$$

(compare (2.2)) in which K_N is a Hermitian kernel ($K_N(z,w) = \bar{K}_N(w,z)$) satisfying $||K_N|| \le 1$ as an operator on $L^2(\mathbb{C}, \mu_N)$ for some measure μ_N.

Theorem 6.1 (O. Costin, J. Lebowitz). *Let P_N be a determinantal random point field and let $\#A$ equal the number of points lying in $A \subset \mathbb{C}$ under that measure. As long as $Var_N[\#A] \uparrow \infty$ with $N \uparrow \infty$ then $(\#A - E[\#A])/\sqrt{Var_N[\#A]}$ converges in distribution to a Gaussian random variable of mean zero and variance one as $N \uparrow \infty$.*

Proof. Following [28] as well as the original [7], the proof checks that the cumulants $C_\ell^N(\#A)$ fall into line as N gets large. Recall that for any random variable X its ℓ-th cumulant $C_\ell(X)$ is defined through

$$\log E[e^{itX}] = \sum_{\ell=1}^{\infty} C_\ell(X) \frac{(it)^\ell}{\ell!}$$

and that X is Gaussian if and only if $C_\ell(X) = 0$ for all $\ell \geq 2$ and $C_2(X) > 0$ (see [22]).

Next define $A_N = K_N \mathbb{1}_A$, the restriction of K_N to $L^2(A, \mu_N)$ and notice that (compare 2.4) $E_N[\#A] = C_1^N(\#A) = \mathrm{Trace}(A_N)$ and $\mathrm{V}ar_N[\#A] = C_2^N(\#A) = \mathrm{Trace}(A_N - A_N^2)$. For the higher cumulants, there is the following recursion:

$$C_\ell^N(\#A) = (-1)^\ell(\ell-1)!\,\mathrm{Trace}(A_N - A_N^\ell) + \sum_{k=2}^{\ell-2} \alpha_{k\ell} C_k^N(\#A). \tag{6.1}$$

in which $\{\alpha_{k\ell}\}$ are certain combinatorial factors. The origins of (6.1) lie in the connection between the cumulants and the so-called cluster (or Ursell) functions. The latter are defined in terms of the correlation functions. In particular, with $u_{N,k}$ denoting the k-point cluster function,

$$u_{N,k}(z_1, z_2, \ldots, z_k) = \sum (-1)^{k-m}(m-1)!\,\rho_{N,|S_1|} \cdots \rho_{N,|S_m|} \tag{6.2}$$

where the sum is over $m \geq 1$ and, for each m, all partitions of $\{1, 2, \cdots, k\}$ into (nonempty) subsets S_1 etc. (also, $\rho_{|S_1|,N}$ indicated the $|S_1|$-point function in the variables corresponding to that partition). Defining also

$$U_k^N(A) = \int_A \cdots \int_A u_{N,k}(z_1, z_2, \ldots, z_k)\, dx_1 dy_1 \cdots dx_k dy_k, \tag{6.3}$$

the advertised connection is provided by the identity

$$\sum_{k=1}^{\infty} C_k^N(\#A) \frac{z^k}{k!} = \sum_{k=1}^{\infty} \frac{1}{k!} U_k^N(A)(e^z - 1).$$

This last display implies that

$$C_\ell^N(\#A) = \sum_{k=2}^{\ell-2} \alpha_{k\ell} C_k^N(\#A) + (-1)^\ell(\ell-1)!\, C_1^N(\#A) + U_k^N(A). \tag{6.4}$$

And finally, the fact is that, for determinantal point fields, $U_\ell^N(A) = (-1)^{\ell-1}(\ell - 1)!\,\mathrm{Trace}(A_N^\ell)$, explaining (6.1).

Now note the simple estimate

$$0 \leq \mathrm{Trace}(A_N - A_N^\ell) \leq (\ell-1)\mathrm{Trace}(A_N - A_N^2) \tag{6.5}$$

which may be checked by writing $\mathrm{Trace}(A_N - A_N^\ell) = \sum_{j=1}^{\ell-1} \mathrm{Trace}\,(A_N^j - A_N^{j+1})$ $\leq \sum_{j=1}^{\ell-1} ||A_N^{j-1}||\,\mathrm{Trace}(A_N - A_N^2)$ and recalling that $||K_N|| \leq 1$. Substituting

(6.5) into (6.1) we have that $C_\ell^N(\#A) = O(C_2^N(\#A))$ for all $\ell \geq 2$. And since $C_2^N(\#A) \uparrow \infty$ as $N \uparrow \infty$, we also have that

$$C_\ell^N\left(\frac{\#A - E_N[\#A]}{\sqrt{Var_N[\#A]}}\right) = \frac{C_\ell^N(\#A)}{(C_2^N(\#A))^{\ell/2}} \to 0$$

for $\ell \geq 2$. The proof is finished.

Proof of Theorem 5.1. Consider more generally a set of probability densities $\{p_{N,k}\}$ $(k = 1, 2, \ldots N)$ on $\mathbb{C}$ and the point process defined for each N by joint density

$$P_N(z_1, \ldots, z_N) = \frac{1}{N!} \sum_{permutations\ \sigma} p_{N,\sigma(1)}(z_1) p_{N,\sigma(2)}(z_2) \cdots p_{N,\sigma(N)}(z_N).$$

We show that the number statistic for such an ensemble satisfies a central limit theorem if its variance grows without bound as $N \uparrow \infty$. This may then by applied to the particular case of (5.2) in which a choice of $p_{N,k}$ has been made $(= p_{2N}^{(2k)}$ in that notation), things are restricted to the positive reals rather than all of $\mathbb{C}$, and the needed growth of the variance has already been checked.

This proof also rests on the behavior of the cumulants. While not determinantal, the correlation functions now have the simpler form:

$$\rho_{\ell,N}(z_1, \cdots, z_\ell) = \sum_{1 \leq k_1, k_2, \ldots k_\ell \leq N} p_{N,k_1}(z_1) p_{N,k_2}(z_2) \cdots p_{N,k_\ell}(z_\ell).$$

So, returning to the Costin-Lebowitz proof, we make use not of (6.1), but instead the preliminary cumulant formula (6.4). With $\rho_{\ell,N}$ as in the previous display, a calculation in (6.3) and (6.2) shows that

$$U_\ell^N(A) = (-1)^{\ell-1}(\ell - 1)! \sum_{k=1}^{N} [p_{N,k}(A)]^\ell$$

in the present case. That is, the growth of the ℓ^{th} cumulant is now tied to $\sum_{k=1}^N (p_{N,k}(A) - p_{N,k}^\ell(A))$. But, $p - p^\ell \leq (\ell - 1)(p - p^2)$ for $\ell > 2$ and any $p \in [0, 1]$, which is to say that ℓ^{th} cumulant is dominated by a constant multiple of the second. The proof is finished.

Acknowledgements. A serious debt of gratitude is due to the referee. In addition to helping correct various small glitches, their thoroughness identified some important errors in an earlier version of this paper.

References

[1] Bai, Z.D.: Circular Law. Annals of Probability **25**, 494–529 (1997)
[2] Bai, Z.D., Silverstein, J.: CLT of linear spectral statistics of large dimensional sample covariance matrices. To appear, Annals of Probability, (2003)
[3] Basor, E.: Distribution functions for random variables for ensembles of positive Hermitian matrices. Comm. Mathematical Physics **188**, 327–350 (1997)

[4] Basor, E., Widom, H.: Determinants of Airy operators and applications to random matrices. Journal of Statistical Physics **96**, 1–20 (1999)
[5] Bhattacharya, R.N., Ranga Rao, R.: Normal Approximation and Asymptotic Expansions. Wiley, New York, (1976)
[6] Bump, D., Diaconis, P., Keller, J.: Unitary Correlations and the Fejér Kernel. Mathematical Physics, Analysis, and Geometry **5**, 101–123 (2002)
[7] Costin, A., Lebowitz, J.L.: Gaussian fluctuations in random matrices. Physical Review Letters **75**, 69–72 (1995)
[8] Chow, Y.S., Teicher, H.: Probability Theory: Independence, Interchangeability, Martingales (Second edition), Springer-Verlag (1988)
[9] Diaconis, P., Evans, S.N.: Linear functionals of eigenvalues of random matrices. Transactions AMS **353**, 2615–2633 (2001)
[10] Diaconis, P., Evans, S.N.: Immanants and finite point processes. J. Comb. Theory Ser. A. **91**, 305–321 (2000)
[11] Edelman, A.: The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law. Journal of Multivariate Analysis **60**, 203–232 (1997)
[12] Forrester, P.J.: Fluctuation formula for complex random matrices. Journal of Physics A: Mathematical and General **32**, 159–163 (1999)
[13] Fyodorov, Y.V., Sommers, H.-J.: Random matrices close to Hermitian and unitary: overview of methods and results. Journal of Physics A: Mathematical and General **36**, 3303–3347 (2003)
[14] Ginibre, J.: Statistical ensembles of complex, quaternion, and real matrices. Journal of Mathematical Physics **6**, 440–449 (1965)
[15] Guionnet, A.: Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices. Ann. Inst. H. Poincaré Probab. Statist. **38**, 341–384 (2002)
[16] Israelsson, S.: Asymptotic fluctuations of a particle system with singular interaction. Stoch. Processes and their Appl. **93**, 25–56 (2001)
[17] Khorunzhy, A. M., Khoruzhenko, B.A., Pastur, L.A.: Asymptotic properties of large random matrices with independent entries. Journal of Mathematical Physics **37**, 5033–5059 (1996)
[18] Johansson, K.: On the fluctuation of eigenvalues of random Hermitian matrices. Duke Math Journal **91**, 151–204 (1998)
[19] Johansson, K.: On random matrices from the the classical compact groups. Annals of Mathematics **145**, 519–545 (1997)
[20] Keating, J.P., Snaith, N.C.: Random matrix theory and $\zeta(1/2 + it)$. Communications in Mathematical Physics **214**, 57–89 (2000)
[21] Lehmann, N., Sommers, H.J.: Eigenvalue statistics of random real matrices. Physics Review Letters **67**, 941–944 (1991)
[22] Lukacs, E.: Characteristic Functions (Second edition), Griffin, London (1970)
[23] Macchi, O.: The coincidence approach to stochastic point processes. Advances in Applied Probability **7**, 83–122 (1975)
[24] Mehta, M.L.: Random Matrices (Second edition), Academic Press, Boston, 1991
[25] Rider, B.: A limit theorem at the edge of a non-Hermitian random matrix ensemble. Journal of Physics A: Mathematical and General **36**, 3401–3409 (2003)
[26] Rider, B.: Order Statistics and Ginibre's Ensembles. Journal of Statistical Physics **114**, 1139–1149 (2004)
[27] Sinai, Ya., Soshnikov, A.: Central limit theorem for traces of large random symmetric matrices with independent entries. Bol. Soc. Brasil. Mat. **29**, 1–24 (1998)

[28] Soshnikov, A.: Central limit theorems for local linear statistics in classical compact groups and related combinatorial identities. Annals of Probability **28**, 1353–1370 (2000)
[29] Soshnikov, A.: Gaussian fluctuations for Airy, Bessel and and other determinantal random point fields. Journal of Statistical Physics **100**, 491–522 (2000)
[30] Soshnikov, A.: Gaussian limits for determinantal random point fields. Annals of Probability **30**, 171–181 (2002)
[31] Soshnikov, A.: Determinantal random point fields. Russian Mathematical Surveys **55**, 923–975 (2000)
[32] Wieand, K.: Eigenvalue distributions of random unitary matrices. Probability Theory and Related Fields **123**, 202–224 (2002)