The spherical ensemble and
uniform distribution of points on the sphere

Kasra Alishahi*   Mohammadsadegh Zamani†

Abstract

The spherical ensemble is a well-studied determinantal process with a fixed number of points on $\mathbb{S}^2$. The points of this process correspond to the generalized eigenvalues of two appropriately chosen random matrices, mapped to the surface of the sphere by stereographic projection. This model can be considered as a spherical analogue for other random matrix models on the unit circle and complex plane such as the circular unitary ensemble or the Ginibre ensemble, and is one of the most natural constructions of a (statistically) rotation invariant point process with repelling property on the sphere.

In this paper we study the spherical ensemble and its local repelling property by investigating the minimum spacing between the points and the area of the largest empty cap. Moreover, we consider this process as a way of distributing points uniformly on the sphere. To this aim, we study two "metrics" to measure the uniformity of an arrangement of points on the sphere. For each of these metrics (discrepancy and Riesz energies) we obtain some bounds and investigate the asymptotic behavior when the number of points tends to infinity. It is remarkable that though the model is random, because of the repelling property of the points, the behavior can be proved to be as good as the best known constructions (for discrepancy) or even better than the best known constructions (for Riesz energies).

1 Introduction

1.1 Background

The aim of this paper is to study the statistical properties of a natural point process on the sphere where the points exhibit repulsive behavior. This point process was introduced in [25] and is known as spherical ensemble; see [23] and [24]. The model was studied earlier in [12] and [15], but without observing the connection to random matrices. It was shown in [12, 15] that there exists a connection between this model and

---

*Sharif University of Technology, Tehran, Iran. E-mail: `alishahi@sharif.edu`
†Sharif University of Technology, Tehran, Iran. E-mail: `ms_zamani@mehr.sharif.edu`

the one-component plasma on the sphere for a special value of temperature. See [17] for further discussion.

Let A_n and B_n be independent $n \times n$ random matrices with independent and identically distributed standard complex Gaussian entries, and let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ denotes the set of eigenvalues of $A_n^{-1} B_n$. We can consider these eigenvalues as a (simple) random point process on complex plane $\mathbb{C}$. The point process $\{\lambda_1, \ldots, \lambda_n\}$ can be described using the k-*point correlation functions* $\rho_k^{(n)} : \mathbb{C}^k \to \mathbb{R}^{\geq 0}$, $1 \leq k \leq n$, defined in such a way that

$$\int_{\mathbb{C}^k} F(z_1, \ldots, z_k) \rho_k^{(n)}(z_1, \ldots, z_k) \, \mathrm{d}\mu(z_1) \ldots \mathrm{d}\mu(z_k) \tag{1.1}$$
$$= \mathbf{E} \sum_{\substack{i_1, \ldots, i_k \\ pairwise\ distinct}} F(\lambda_{i_1}, \ldots, \lambda_{i_k}),$$

for all continuous, compactly supported functions $F : \mathbb{C}^k \to \mathbb{C}$, where

$$\mathrm{d}\mu(z) := \frac{n}{\pi(1+|z|^2)^{n+1}} \mathrm{d}z$$

and $\mathrm{d}z$ denotes the Lebesgue measure on the complex plane $\mathbb{C}$.

Krishnapur [24] showed that this random point process is a determinantal point process on complex plane with kernel

$$K^{(n)}(z, w) = (1 + z\bar{w})^{n-1}$$

with respect to the background measure $\mathrm{d}\mu(z)$, i.e. we have

$$\rho_k^{(n)}(z_1, \ldots, z_k) = \det \left[K^{(n)}(z_i, z_j) \right]_{i,j=1}^k \tag{1.2}$$

for every $k \geq 1$ and $z_1, \ldots, z_k \in \mathbb{C}$. We note here that a random point process is said to be a *determinantal point process* if its k-point correlation functions have determinantal form similar to (1.2). The corresponding kernel $K^{(n)}(z, w)$ is called a correlation kernel of the determinantal point process. We refer to [22] or [23] and references therein for more information on deteminantal point processes.

Let $\mathbb{S}^2 = \{p \in \mathbb{R}^3 : |p| = 1\}$ be the unit two-dimensional sphere centred at the origin in three-dimensional Euclidean space $\mathbb{R}^3$. Also we let ν denotes the Lebesgue surface area measure on this sphere with total measure 4π. As mentioned in [23], these eigenvalues are best thought of as points on $\mathbb{S}^2$, using stereographic projection. Let g be the stereographic projection of the sphere $\mathbb{S}^2$ from the north pole $(0, 0, 1)$ onto the complex plane $\mathbb{C}$, seen as two-dimensional plane $\{(t_1, t_2, 0); t_1, t_2 \in \mathbb{R}\}$, i.e., for $(x_1, x_2, x_3) \in \mathbb{S}^2$,

$$g(x_1, x_2, x_3) = \frac{x_1 + ix_2}{1 - x_3}. \tag{1.3}$$

If we let $P_i = g^{-1}(\lambda_i)$ for $1 \leq i \leq n$ then the vector $(P_1, \ldots, P_n)$, in uniform random order, has the joint density

$$\text{Const.} \prod_{i<j} |p_i - p_j|^2$$

with respect to Lebesgue measure on $(\mathbb{S}^2)^n$ where $|p_i - p_j|$ denotes the Euclidean distance between two points p_i and p_j. Note that this density is similar to the circular unitary ensemble case and clearly this point process is invariant in distribution under the isometries of $\mathbb{S}^2$. Consider the point process on $\mathbb{S}^2$,

$$\mathcal{X}^{(n)} := \sum_{j=1}^n \delta_{P_j}.$$

We know that $\frac{1}{n}\mathcal{X}^{(n)}$ converges almost surely to the uniform measure on the sphere. In fact, it is also true in the more general case: Let A'_n and B'_n be independent $n \times n$ random matrices with i.i.d entries with mean 0 and variance 1, and let $\{\lambda'_1, \lambda'_2, \ldots, \lambda'_n\}$ denotes the set of eigenvalues of $A'^{-1}_n B'_n$. Based on the results of [35], Bordenave [8] shows that $\frac{1}{n}\sum_{j=1}^{n} \delta_{g^{-1}(\lambda'_j)}$ converges almost surely to uniform measure on $\mathbb{S}^2$ as $n \to +\infty$.

Moreover from the repulsive nature of determinantal point processes we expect that the points of process $\mathcal{X}^{(n)}$ are typically more evenly distributed than n independently chosen uniform points on the sphere. This repelling property is the common feature of many models in random matrix theory and has been comprehensively studied in some special models such as the Gaussian unitary or the circular unitary ensembles. For example, the distribution or the minimum or the maximum of the gaps between consecutive eigenvalues have been computed and compared to simpler models as a way to measure and understand the repulsive structure. One of the goals of this paper is to do the same computations for the spherical model. We specially focus on the minimum distance between the points, the area of the largest empty cap, the hole probability and the limiting distribution of the nearest neighbors distances as natural two-dimensional extensions of the so called metrics studied in one dimensional models.

On the other hand we apply this model and its properties to the classic problem of distributing points on the sphere. The problem of distributing a given number of points on the surface of a sphere "uniformly", is a challenging and old problem. Contrary to the one dimensional case (i.e. distributing points on a circle) where the most uniform arrangement clearly exists and is attained when the points are equidistributed, it seems that there is no arrangement on the sphere that can be considered as completely uniform, and the answer for the best arrangement depends on what criteria do we use to quantify the uniformity of an arrangement. Among the mostly used criteria are those related to the electrostatic potential energy and its generalizations where one tries to distribute the points in a way that minimizes some energy function. Another common metric is the discrepancy that measures the maximum deviance between the number of points and the expected number, in some class of regions in the underlying space (sphere, in our case). Both the energy and the discrepancy optimization problems, i.e. finding the most optimum arrangement or even obtaining some relatively sharp upper and lower bounds are open and challenging problems for the sphere. We study these metrics (discrepancy and Riesz energies) to measure the uniformity of points of $\mathcal{X}^{(n)}$. For each of these metrics we obtain some bounds and investigate the asymptotic behavior when the number of points tends to infinity. It is remarkable that though the model is random, because of the repelling property of the points, the behavior can be proved to be as good as the best known constructions (for discrepancy) or even better than the best known constructions (for Riesz energies).

The main results of the paper are stated in the next subsection, together with definition and some properties of the metrics discussed above.

1.2 Main results

Discrepancy

The geometrically most natural measure for the uniformity of the distribution of an n-point set on $\mathbb{S}^2$ is the *spherical cap discrepancy*. Let $\mathcal{P}_n = \{x_1, \ldots, x_n\}$ be an n-point set on $\mathbb{S}^2$. The spherical cap discrepancy of $\mathcal{P}_n$ is defined as

$$\mathbf{D}(\mathcal{P}_n) := \sup_{D \in \mathcal{A}} \left| \sum_{j=1}^{n} 1_D(x_j) - \frac{n|D|}{4\pi} \right|$$

where $\mathcal{A}$ is the set of all spherical caps on $\mathbb{S}^2$. A spherical cap is defined as the intersection of the sphere and a half-space. In [4], it was shown that there is constant $c > 0$, independent of n, such that for any n-point set $\mathcal{P}_n$ on $\mathbb{S}^2$ we have

$$cn^{1/4} \leq \mathbf{D}(\mathcal{P}_n).$$

On the other hand, using probabilistic methods it has been shown (see [5]) that for any $n \geq 1$, there exists n-point set $\mathcal{P}_n$ on $\mathbb{S}^2$ such that

$$\mathbf{D}(\mathcal{P}_n) \leq Cn^{1/4}\sqrt{\log n}.$$

The following theorem shows that the point process $\mathcal{X}^{(n)}$ has small spherical cap discrepancy.

Theorem 1.1. *Consider the point process $\mathcal{X}^{(n)} = \sum_{i=1}^{n} \delta_{P_i}$. For every $M > 0$ independent of n, one has*

$$\mathbf{D}(\{P_1, \ldots, P_n\}) = O\left(n^{1/4}\sqrt{\log n}\right)$$

with probability at least $1 - \frac{1}{n^M}$.

Note that for independent uniform points on sphere, the discrepancy is of order $\sqrt{n}$ (up to a logarithmic factor).

The key to the proof of Theorem 1.1 is an estimate on the variance of the number of points of $\mathcal{X}^{(n)}$ on a spherical cap. The asymptotic expansion of this variance and the proof of Theorem 1.1 will be presented in Section 2.

Largest empty cap

Given $\mathcal{P}_n = \{x_1, \ldots, x_n\} \subset \mathbb{S}^2$, define the covering radius of $\mathcal{P}_n$ as the infimum of the numbers $t > 0$ such that every point of $\mathbb{S}^2$ is within distance t of some x_j. If we let τ be the covering radius of $\mathcal{P}_n$, then the area of the largest spherical cap which does not contain any point of $\mathcal{P}_n$ in its interior is equal to $\pi\tau^2$ (note that, for fixed q, the area of the spherical cap $\{p \in \mathbb{S}^2 : |p - q| \leq r\}$ is πr^2). We will be interested in studying the asymptotic behavior of the area of the largest empty cap for the spherical ensemble.

Let M_n be the area of largest empty cap for random point process $\mathcal{X}^{(n)}$. In the following theorem, we derive first-order asymptotic for M_n.

Theorem 1.2. *Let*

$$X_n := \frac{n}{8\pi\sqrt{\log n}} M_n.$$

Then for any $s > 0$, X_n converges in L^s to one,

$$\mathbf{E}(|X_n - 1|^s) \to 0,$$

as $n \to +\infty$.

The proof of this theorem is given in Section 3. For the proof we need asymptotics of the *hole probability*, the probability that there are no points of $\mathcal{X}^{(n)}$ in a given spherical cap. The desired asymptotics of the hole probability will be given in Lemma 3.1. Then we will prove Theorem 1.2 using a method similar to the proof of Theorem 1.3 in [3]. Notice that for independent uniform points on $\mathbb{S}^2$, the area of the largest empty cap is of order $\frac{\log n}{n}$.

At the end of Section 3 we study the nearest neighbour statistics of spherical ensemble and show a connection between the local behavior of this model and the Ginibre ensemble.

Riesz and logarithmic energy

In Section 4, we compute the expectations of the logarithmic energy and Riesz s-energy of the random point process $\mathcal{X}^{(n)}$ on $\mathbb{S}^2$. The discrete logarithmic energy of n points $x_1, \ldots, x_n$ on $\mathbb{S}^2$ is given by

$$E_{\log}(x_1, \ldots, x_n) := \log \prod_{i \neq j} \frac{1}{|x_i - x_j|} = \sum_{i \neq j} \log \frac{1}{|x_i - x_j|}.$$

Also we define

$$\mathcal{E}_{\log}(n) := \min\{E_{\log}(x_1, \ldots, x_n); x_1, \ldots, x_n \in \mathbb{S}^2\}.$$

The n-tuples that minimize this energy are usually called elliptic Fekete Points. Define C_n by

$$\mathcal{E}_{\log}(n) = \left(\frac{1}{2} - \log 2\right) n^2 - \frac{1}{2} n \log n + C_n n.$$

In [28] it was shown that C_n satisfies the following estimates

$$-0.22553754 \cdots \leq \liminf_{n \to +\infty} C_n \leq \limsup_{n \to +\infty} C_n \leq -0.0469945 \ldots.$$

For a given s, the *Riesz s-energy* of n points $x_1, \ldots, x_n$ on $\mathbb{S}^2$ are defined as

$$E_s(x_1, \ldots, x_n) := \sum_{i \neq j} \frac{1}{|x_i - x_j|^s}.$$

Also, we consider the optimal n-point Riesz s-energy,

$$\mathcal{E}_s(n) := \begin{cases} \min\{E_s(x_1, \ldots, x_n); x_1, \ldots, x_n \in \mathbb{S}^2\} & \text{if } s \geq 0 \\[2mm] \max\{E_s(x_1, \ldots, x_n); x_1, \ldots, x_n \in \mathbb{S}^2\} & \text{if } s < 0. \end{cases}$$

The important special case $s = 1$ corresponds to electrostatic potential energy of electrons on $\mathbb{S}^2$ that repel each other with a force given by Coulomb's law. We remark that this problem is only interesting for $s > -2$. It is known that for the *potential-theoretic regime*, $-2 < s < 2$, we have

$$\lim_{n \to +\infty} \frac{\mathcal{E}_s(n)}{n^2} = \frac{1}{(4\pi)^2} \int_{\mathbb{S}^2 \times \mathbb{S}^2} \frac{1}{|p - q|^s} \, d\nu(p) \, d\nu(q) = \frac{2^{1-s}}{2-s}. \tag{1.4}$$

See e.g. [10].

Consider the difference $\mathcal{E}_s(n) - \frac{2^{1-s}}{2-s} n^2$. Wagner ([36] lower bound for $0 < s < 2$ and upper bound for $-2 < s < 0$, and [37] upper bound for $0 < s < 2$ and lower bound for $-2 < s < 0$) proved that

$$-c_1 n^{1+s/2} \leq \mathcal{E}_s(n) - \frac{2^{1-s}}{2-s} n^2 \leq -c_2 n^{1+s/2} \quad , \quad -2 < s < 2$$

where c_1 and c_2 are positive constants depending only on s. In [28], an alternative method which is based on partitioning $\mathbb{S}^2$ into regions of equal area and small diameter is used to prove the upper bound in the case $0 < s < 2$ (and lower bound in the case $-2 < s < 0$). Let $\varepsilon > 0$ be arbitrary, this method gives

$$\mathcal{E}_s(n) - \frac{2^{1-s}}{2-s} n^2 \leq -(2\sqrt{2\pi})^{-s}(1-\varepsilon) n^{1+s/2} \quad , \quad 0 < s < 2 \tag{1.5}$$

$$\mathcal{E}_s(n) - \frac{2^{1-s}}{2-s} n^2 \geq -(2\sqrt{2\pi})^{-s}(1+\varepsilon) n^{1+s/2} \quad , \quad -2 < s < 0 \tag{1.6}$$

for $n \geq n_0(\varepsilon, s)$. We show that the better bounds than (1.5) and (1.6) can be obtained by considering the expectation of Riesz s-energy of point process $\mathcal{X}^{(n)}$. It is conjectured (see [10], Conjecture 3) that the asymptotic expansion of the optimal Riesz s-energy for $-2 < s < 4$, $s \neq 2$ has the form

$$\mathcal{E}_s(n) = \frac{2^{1-s}}{2-s}n^2 + \frac{(\sqrt{3}/2)^{s/2}\zeta_{\Lambda_2}(s)}{(4\pi)^{s/2}}n^{1+s/2} + o(n^{1+s/2}) \quad n \to +\infty, \qquad (1.7)$$

where $\zeta_{\Lambda_2}(s)$ is the zeta function of the hexagonal lattice $\Lambda_2 = \{m(1,0) + n(1/2, \sqrt{3}/2) : m, n \in \mathbb{Z}\}$. See the survey [10] for more details and further discussion.

In the boundary case $s = 2$, we have (see Theorem 3 in [26])

$$\lim_{n \to +\infty} \frac{\mathcal{E}_2(n)}{n^2 \log n} = \frac{1}{4}.$$

Also, in [10] (see Proposition 3 and its remark therein), it is shown that

$$-\frac{1}{4}n^2 + O(n) \leq \mathcal{E}_2(n) - \frac{1}{4}n^2 \log n \leq \frac{1}{4}n^2 \log \log n + O(n^2). \qquad (1.8)$$

Considering the expectation of Riesz s-energy of point process $\mathcal{X}^{(n)}$, we are able to omit the $\log \log n$ term in this estimate. (See Conjecture 5 in [10] for the asymptotic expansion of the optimal Riesz 2-energy.) We remark that the first term of the asymptotics of $\mathcal{E}_s(n)$ for $s > 2$ is not known.

In the following theorem, we give the expectations of the logarithmic energy and Riesz s-energy of the random point process $\mathcal{X}^{(n)}$ on $\mathbb{S}^2$.

Theorem 1.3. *For the point process $\mathcal{X}^{(n)}$ on $\mathbb{S}^2$, $n \geq 2$, we have*

i) (Logarithmic energy)

$$\mathbf{E}\, E_{\log}(P_1, \ldots, P_n) = \left(\frac{1}{2} - \log 2\right)n^2 - \frac{1}{2}n \log n + \left(\log 2 - \frac{\gamma}{2}\right)n - \frac{1}{4} + O\left(\frac{1}{n}\right) \qquad (1.9)$$

Here, γ is the Euler constant.

ii) (Riesz s-energy: $s < 4$ and $s \neq 2$)

$$\mathbf{E}\, E_s(P_1, \ldots, P_n) = \frac{2^{1-s}}{2-s}n^2 - \frac{\Gamma(n)\Gamma(1-s/2)}{2^s\Gamma(n+1-s/2)}n^2 \qquad (1.10)$$

iii) (Riesz s-energy: $s = 2$)

$$\mathbf{E}\, E_2(P_1, \ldots, P_n) = \frac{1}{4}n^2 \log n + \frac{\gamma}{4}n^2 - \frac{n}{8} - \frac{1}{48} + O\left(\frac{1}{n^2}\right) \qquad (1.11)$$

As a corollary to Theorem 1.3, we obtain the following bounds for optimal Riesz s-energy.

Corollary 1.4. *for every $n \geq 2$,*

$$\mathcal{E}_s(n) - \frac{2^{1-s}}{2-s}n^2 \leq -\frac{\Gamma(1-s/2)}{2^s}n^{1+s/2} \quad , \quad 0 < s < 2 \qquad (1.12)$$

$$\mathcal{E}_s(n) - \frac{2^{1-s}}{2-s}n^2 \geq -\frac{\Gamma(1-s/2)}{2^s}n^{1+s/2} \quad , \quad -2 < s < 0 \qquad (1.13)$$

$$\mathcal{E}_2(n) - \frac{1}{4}n^2 \log n \leq \frac{\gamma}{4}n^2 \qquad (1.14)$$

Suppose that $x_1, \ldots, x_n$ are chosen randomly and independently on the sphere, with the uniform distribution. One can easily show that

$$\mathbf{E}\, E_{\log}(x_1, \ldots, x_n) = \left(\frac{1}{2} - \log 2\right)n^2 - \left(\frac{1}{2} - \log 2\right)n$$

and also for $s < 2$ (see (1.4))

$$\mathbf{E}\, E_s(x_1, \ldots, x_n) = \frac{2^{1-s}}{2-s} n^2 - \frac{2^{1-s}}{2-s} n.$$

Minimum spacing

Define the minimum spacing by

$$m_n := \min_{i \neq j} |P_i - P_j|.$$

We are interested in the asymptotic behavior of m_n as n tends to infinity. For $0 < t < 2$, set

$$G_{t,n} := \sum_{i<j} 1_{\{|P_i - P_j| \leq t\}} \tag{1.15}$$

to be the number of non-ordered pairs of distinct points at Euclidean distance at most t apart. Our result concerning distribution of $G_{t,n}$ is the following.

Theorem 1.5. *Let $G_{t,n}$ defined by (1.15) and assume that $t = \frac{x}{n^{3/4}}$ then $G_{t,n}$ converges in distribution to the Poisson random variable with mean $\frac{x^4}{64}$.*

The proof of this theorem is given in Section 5. As a consequence, since $\mathbf{P}(G_{t,n} = 0) = \mathbf{P}(m_n > t)$, Theorem 1.5 clearly implies that

Corollary 1.6. *For any $x > 0$,*

$$\lim_{n \to +\infty} \mathbf{P}(n^{3/4} m_n > x) = \exp\left(\frac{-x^4}{64}\right).$$

Suppose that $x_1, \ldots, x_n$ are chosen randomly and independently on the sphere, with the uniform distribution. It was shown that (see Theorem 2 of [11])

$$\lim_{n \to +\infty} \mathbf{P}(n \min_{i \neq j} |x_i - x_j| > x) = \exp\left(\frac{-x^2}{8}\right).$$

2 Discrepancy

Let D be a spherical cap on the sphere $\mathbb{S}^2$. Define

$$N_D := \mathcal{X}^{(n)}(D),$$

the number of points of $\mathcal{X}^{(n)}$ in D. Clearly, the expected value of N_D is equal to $\frac{n|D|}{4\pi}$. In order to prove Theorem 1.1 we need to control the variance of N_D. The following lemma gives the asymptotic behavior of the variance.

Lemma 2.1. *Let D be a spherical cap on the sphere (depending on n) such that $\frac{1}{|D|}, \frac{1}{|D^c|} = o(n)$. Then for any $\varepsilon > 0$*

$$\mathbf{Var}[N_D] = \frac{\sqrt{n}}{4\pi\sqrt{\pi}} \sqrt{|D||D^c|} + o\left(\log^{\frac{1}{2}+\varepsilon}(n|D||D^c|)\right) \tag{2.1}$$

where $|D|$ denotes the area of D and $D^c = \mathbb{S}^2 - D$.

Proof. The distribution of N_D is invariant under isometries of the sphere, so we may assume without loss of generality that $g(D)$ is a disk centred at the origin with radius r. From [22], Theorem 26, the set $\{|\lambda_k|^2 : 1 \leq k \leq n\}$ has the same distribution as set

$\{Q_k : 0 \le k \le n-1\}$ where the random variables Q_k are jointly independent and Q_k for $0 \le k \le n-1$ has density

$$\frac{n\binom{n-1}{k}q^k}{(1+q)^{n+1}} \quad , \quad q \ge 0$$

(Note that $K(z,w) = \sum_{k=0}^{n-1} \binom{n-1}{k}(z\bar{w})^k$ and μ has density $\phi(|z|)$ where $\phi(x) = \frac{n}{\pi(1+x^2)^{n-1}}$).
Thus Q_k has the beta prime distribution with parameters $k+1$ and $n-k$, i.e., $\frac{Q_k}{1+Q_k}$ has the beta distribution with parameters as before.

Thus we have

$$N_D \overset{d}{\sim} |\{k : 0 \le k \le n-1 \,,\, Q_k < r^2\}|. \tag{2.2}$$

And so by the independence of Q_k, $0 \le k \le n-1$, we deduce that

$$N_D \overset{d}{\sim} \eta_0 + \eta_1 + \cdots + \eta_{n-1} \tag{2.3}$$

where the random variables η_k are independent and distributed as Bernoulli variables with $\mathbf{P}(\eta_k = 1) = \mathbf{P}(Q_k < r^2)$. By the properties of incomplete beta function (see e.g. [39]) we have

$$\mathbf{P}(Q_k < r^2) = \int_0^{r^2} \frac{n\binom{n-1}{k}q^k}{(1+q)^{n+1}}\,\mathrm{d}q = \frac{\sum_{j=k+1}^{n}\binom{n}{j}r^{2j}}{(1+r^2)^n} \tag{2.4}$$

Now suppose $B_1, \ldots, B_n$ are i.i.d. Bernoulli random variables with parameter $\alpha := \frac{r^2}{1+r^2}$. Let $S_n := B_1 + \cdots + B_n$. One can write the right-hand side of above equation as $\mathbf{P}(S_n > k)$. Therefore, we obtain for every $0 \le k \le n-1$

$$\mathbf{P}(\eta_k = 1) = \mathbf{P}(Q_k < r^2) = \mathbf{P}(S_n > k). \tag{2.5}$$

We also note that

$$|D| = \frac{4\pi r^2}{1+r^2} = 4\pi\alpha.$$

Now from (2.3) and (2.5) it follows that for every spherical cap D,

$$\mathbf{Var}[N_D] = \sum_{k=0}^{n-1} \mathbf{P}(S_n \le k)\,\mathbf{P}(S_n > k) \tag{2.6}$$

where S_n is a binomial random variable with parameters n and $\frac{|D|}{4\pi}$.

Fix $\varepsilon > 0$. Let

$$\sigma = \sqrt{\mathbf{Var}[S_n]} = \sqrt{n\alpha(1-\alpha)}$$

and

$$M = \sigma(\log\sigma)^{\frac{1}{2}+\varepsilon}.$$

Since $\mathbf{Var}[N_D] = \mathbf{Var}[N_{D^c}]$, by symmetry, we may assume that $\alpha \le 1/2$. Using the Bernstein's inequality we see that for any $t > 0$, one has

$$\mathbf{P}(|S_n - n\alpha| > t) \le 2\exp\left(-\min\left(\frac{t^2}{4\sigma^2}, \frac{t}{4}\right)\right) \tag{2.7}$$

(see for example Lemma 2 in [5] or Lemma 2.7.1 in [34]). We can use this to show that (Let $j = |k - n\alpha|$ and note that $\alpha \le 1/2$.)

$$\sum_{|k-n\alpha| \ge M} \mathbf{P}(S_n \le k)\,\mathbf{P}(S_n > k) \le \tag{2.8}$$

$$2\left[\sum_{M \le j \le \sigma^2} \exp\left(-\frac{j^2}{4\sigma^2}\right) + \sum_{\sigma^2 < j \le n(1-\alpha)} \exp(-j/4)\right].$$

The assumption of the lemma implies that $\sigma \to +\infty$ as $n \to +\infty$ and thus we have

$$\sum_{\sigma^2 < j \leq n(1-\alpha)} \exp(-j/4) = o(1). \tag{2.9}$$

Also, by comparing the integral of $\exp\left(\frac{-x^2}{4}\right)$ with its Riemann sum, we conclude that

$$\sum_{M+1 \leq j \leq \sigma^2} \exp\left(-\frac{j^2}{4\sigma^2}\right) \leq \sigma \int_{(\log \sigma)^{1/2+\varepsilon}}^{+\infty} \exp\left(\frac{-x^2}{4}\right) \mathrm{d}x$$

$$= \sigma \int_{(\log \sigma)^{1+2\varepsilon}}^{+\infty} \frac{1}{2\sqrt{x}} \exp(-x/4) \, \mathrm{d}x$$

$$\leq \frac{2\sigma}{(\log \sigma)^{1/2+\varepsilon}} \exp\left(-\frac{1}{4}(\log \sigma)^{1+2\varepsilon}\right) = o(1).$$

Thus, from (2.8), (2.9) and above estimate, we obtain

$$\sum_{|k-n\alpha| \geq M} \mathbf{P}(S_n \leq k)\,\mathbf{P}(S_n > k) = o(1).$$

Moreover, from the Berry-Esséen theorem we see that for some absolute constant C, one has (see [6])

$$\left|\mathbf{P}(S_n \leq k) - \Phi\left(\frac{k-n\alpha}{\sigma}\right)\right| \leq \frac{C}{\sigma}$$

where Φ is the cumulative distribution function of the standard normal distribution. From (2.6), and by using the above two estimates, we conclude that

$$\mathbf{Var}[N_D] = \sum_{|j|<M} \Phi\left(\frac{j}{\sigma}\right) \Phi\left(\frac{-j}{\sigma}\right) + o((\log \sigma)^{1/2+2\varepsilon}).$$

By considering the Riemann sum of $\Phi(x)\Phi(-x)$, Making the change of variable $x = (\log \sigma)^{1/2+\varepsilon} y$ and letting $\tilde{\Phi}(y) := \Phi\left((\log \sigma)^{1/2+\varepsilon} y\right)$, we have

$$\left|\sum_{|j|<M} \Phi\left(\frac{j}{\sigma}\right) \Phi\left(\frac{-j}{\sigma}\right) - \sigma \int_{-(\log \sigma)^{1/2+\varepsilon}}^{(\log \sigma)^{1/2+\varepsilon}} \Phi(x)\Phi(-x)\,\mathrm{d}x\right|$$

$$= \left|\sum_{|j|<M} \tilde{\Phi}\left(\frac{-j}{M}\right) \tilde{\Phi}\left(\frac{j}{M}\right) - M \int_{-1}^{1} \tilde{\Phi}(y)\tilde{\Phi}(-y)\,\mathrm{d}y\right|$$

$$\leq O(1) + \int_{-1}^{1} \left|(\tilde{\Phi}(y)\tilde{\Phi}(-y))'\right|\,\mathrm{d}y \leq O(1) + \int_{-\infty}^{\infty} \left|(\Phi(x)\Phi(-x))'\right|\,\mathrm{d}x = O(1).$$

Using the standard bound

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-t^2/2}\,\mathrm{d}t = \frac{1}{\sqrt{2\pi}} \int_{x^2}^{+\infty} \frac{1}{2\sqrt{t}} e^{-t/2}\,\mathrm{d}t < \frac{1}{\sqrt{2\pi}x} e^{\frac{-x^2}{2}}$$

for any $x > 0$, we have

$$\int_{\{x:|x|>(\log \sigma)^{1/2+\varepsilon}\}} \Phi(x)\Phi(-x)\,\mathrm{d}x \leq 2 \int_{(\log \sigma)^{1/2+\varepsilon}}^{+\infty} \Phi(-x)\,\mathrm{d}x$$

$$\leq \frac{2}{(\log \sigma)^{1/2+\varepsilon}} \Phi(-(\log \sigma)^{1/2+\varepsilon})$$

$$= o(1/\sigma).$$

Combining all these facts, we deduce that

$$\mathbf{Var}[N_D] = \sigma \int_{-\infty}^{+\infty} \Phi(x)\Phi(-x)\,\mathrm{d}x + o((\log\sigma)^{1/2+2\varepsilon}).$$

On the other hand, using integration by parts twice, we have

$$\int \Phi(x)\Phi(-x)\,\mathrm{d}x = x\Phi(x)\Phi(-x) - \frac{1}{\sqrt{2\pi}} \int xe^{-\frac{x^2}{2}}(\Phi(x) - \Phi(-x))\,\mathrm{d}x$$

$$= x\Phi(x)\Phi(-x) - \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{x^2}{2}}(\Phi(x) - \Phi(-x)) - \frac{1}{\sqrt{2\pi}} \int 2e^{-x^2}\,\mathrm{d}x \right)$$

$$= x\Phi(x)\Phi(-x) - \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}(\Phi(x) - \Phi(-x)) + \frac{1}{\sqrt{\pi}}\Phi(\sqrt{2}x)$$

Thus

$$\int_{-\infty}^{+\infty} \Phi(x)\Phi(-x)\,\mathrm{d}x = \frac{1}{\sqrt{\pi}}$$

and (2.1) follows. □

Remark 2.2. Clearly, the restriction of point process $\mathcal{X}^{(n)}$ to a spherical cap D is also a determinantal point process with kernel $1_{g(D)}(z)K^{(n)}(z,w)1_{g(D)}(w)$. Denote by $\mathscr{K}_{g(D)}^{(n)}$ the integral operator in $L^2(g(D))$ obtained by considering this kernel. Then the distribution of N_D is the sum of independent Bernoulli random variables, whose expectations are the eigenvalues of $\mathscr{K}_{g(D)}^{(n)}$ (see e.g. [2], Corollary 4.2.24). So from (2.3) we conclude that the non-zero eigenvalues of operator $\mathscr{K}_{g(D)}^{(n)}$ are equal to $\mathbf{P}(S_n > k)$, $0 \le k \le n-1$, where S_n is a binomial random variable with parameters n and $\frac{|D|}{4\pi}$.

From Lemma 2.1 and the general central limit theorem for determinantal point processes [31] (due to Costin and Lebowitz [13] in case of the sine kernel) we have the following theorem.

Theorem 2.3. *Let D be a spherical cap on $\mathbb{S}^2$ (depending on n) such that $\frac{1}{|D|}, \frac{1}{|D^c|} = o(n)$. Then*

$$\frac{N_D - \frac{n|D|}{4\pi}}{\frac{1}{2}\pi^{-3/4}n^{1/4}\sqrt[4]{|D||D^c|}} \xrightarrow{\mathrm{d}} N(0,1)$$

as $n \to +\infty$.

Next we prove Theorem 1.1.

Proof of Theorem 1.1. We use the notation from the proof of Lemma 2.1. As we have seen before

$$\mathbf{Var}[N_D] = \sum_{k=0}^{n-1} \mathbf{P}(S_n \le k)\,\mathbf{P}(S_n > k).$$

Similar to inequality (2.8), we have

$$\sum_{0 \le k \le n-1} \mathbf{P}(S_n \le k)\,\mathbf{P}(S_n > k) \le$$

$$2\left[\sum_{0 \le j \le \sigma^2} \exp\left(-\frac{j^2}{4\sigma^2}\right) + \sum_{j > \sigma^2} \exp(-j/4) \right].$$

Also, we see

$$\sum_{1 \le j \le \sigma^2} \exp\left(-\frac{j^2}{4\sigma^2}\right) \le \sigma \int_0^{+\infty} \exp(-x^2/4)\,\mathrm{d}x.$$

So we deduce that there exist absolute constants c_1, c_2 such that for any spherical cap D,

$$\mathbf{Var}[N_D] \leq c_1 \sigma + c_2 \tag{2.10}$$

and since $\sigma \leq \frac{\sqrt{n}}{2}$, we have for some absolute constant c_3,

$$\mathbf{Var}[N_D] \leq c_3 \sqrt{n}.$$

We know that for any spherical cap D, the random variable N_D has the same distribution as $\sum_k \eta_k$ where η_k are independent Bernoulli random variables (see equation (2.3)). From Bernstein's inequality (see Lemma 2.7.1 in [34]) we get a concentration estimate for N_D:

$$\mathbf{P}\left(|N_D - \mathbf{E}\, N_D| \geq t\right) \leq 2 \exp\left(-\min\left(\frac{t^2}{4\,\mathbf{Var}[N_D]}, \frac{t}{2}\right)\right) \quad , \quad t \geq 0.$$

Let $t = O(n^{1/4}\sqrt{\log n})$. Then for any $M > 0$ we have

$$N_D = \mathbf{E}\, N_D + O(n^{1/4}\sqrt{\log n}) \tag{2.11}$$

with probability at least $1 - n^{-M}$ where the implied constant in the $O()$ notation is independent of D. It can be easily shown that there is a collection $\mathcal{A}'$ of n^c spherical caps, for some absolute constant $c > 0$, such that for any spherical cap D there exist $D_1, D_2 \in \mathcal{A}'$ (D_1 can be empty) having the properties $D_1 \subset D \subset D_2$ and $|D_2 \backslash D_1| \leq \frac{4\pi}{n}$ and then we have

$$|N_D - \mathbf{E}\, N_D| \leq 1 + \max_{i=1,2} |N_{D_i} - \mathbf{E}\, N_{D_i}|.$$

Thus we obtain

$$\mathbf{D}(\{P_1, \ldots, P_n\}) \leq 1 + \max_{D' \in \mathcal{A}'} |N_{D'} - \mathbf{E}\, N_{D'}|.$$

So from the union bound, we see that the equation (2.11) is also true uniformly in D. This completes the proof of the theorem. $\qquad \square$

3 Hole probability and largest empty cap

3.1 Hole probability

In this subsection, we establish the asymptotic behavior of $\Delta_n(\alpha) := \mathbf{P}(N_D = 0)$, where D is a spherical cap on $\mathbb{S}^2$ such that $|D| = 4\pi\alpha$. We use the notation from the proof of Lemma 2.1. We know that $\Delta_n(\alpha)$ is equal to the Fredholm determinant of $\mathrm{Id} - \mathscr{K}_{g(D)}^{(n)}$. Also, from (2.3) and (2.5) we have

$$\Delta_n(\alpha) = \prod_{k=0}^{n-1} \mathbf{P}(S_n \leq k).$$

Proposition 3.1. *There exist a positive constant c' such that*

$$\log \Delta_n(\alpha) = \frac{n^2}{2}(\alpha + \log(1 - \alpha)) + O(n\alpha \log(n\alpha)) \tag{3.1}$$

uniformly in $c'/n < \alpha < 1 - c'/n$.

Proof. We can write

$$\log \Delta_n(\alpha) = \sum_{k=0}^{n-1} \log \mathbf{P}(S_n \leq k). \tag{3.2}$$

First, we estimate $\log \mathbf{P}(S_n \leq k)$ using large deviation principle. By the exponential Chebyshev inequality (See the proof of Theorem I.4, Cramer's Theorem, in [21].), we have

$$\mathbf{P}(S_n \leq k) \leq e^{-nI(\frac{k}{n})} \tag{3.3}$$

for any $k \leq n\alpha$ where $I(x) = \sup_{t \in \mathbb{R}}(tx - \log(\mathbf{E}\,e^{tB_1}))$ is the Legendre transform of the cumulant generating function of B_1. One easily computes that $I(x) = x \log \frac{x}{\alpha} + (1 - x) \log \frac{1-x}{1-\alpha}$. On the other hand, from Stirling's formula we know $\{n! e^n n^{n+1/2}\}$ is bounded away from zero and infinity and thus we have

$$\mathbf{P}(S_n \leq k) \geq \binom{n}{k} \alpha^k (1 - \alpha)^{n-k} \geq c_0 \frac{n^{n+1/2}}{k^{k+1/2}(n-k)^{n-k+1/2}} \alpha^k (1-\alpha)^{n-k}$$

$$\geq c_0 \left(\frac{n}{k(n-k)} \right)^{1/2} e^{-nI(\frac{k}{n})}$$

for $0 < k \leq n\alpha$ and some absolute constant c_0. Therefore, for all $0 < k \leq n\alpha$, we have

$$\log \mathbf{P}(S_n \leq k) \geq -nI(k/n) - \frac{1}{2}\log(n\alpha) + \log c_0. \tag{3.4}$$

From (3.3) and (3.4), it follows that

$$\left| \sum_{0 \leq k \leq n\alpha} \log \mathbf{P}(S_n \leq k) + n \sum_{0 \leq k \leq n\alpha} I(k/n) \right| \leq n\alpha \left(\frac{1}{2}\log(n\alpha) - \log c_0 \right) \tag{3.5}$$

for sufficiently large n (for $k = 0$, note that $\log \mathbf{P}(S_n = 0) = -nI(0)$).

Since the median of binomial distribution is either $\lfloor n\alpha \rfloor$ or $\lceil n\alpha \rceil$, this implies that $\mathbf{P}(S_n > k) \leq 1/2$ when $k > n\alpha$. Using the bound $-\log(1 - x) < 2x$, $0 \leq x \leq 1/2$, we then obtain

$$\sum_{n\alpha < k \leq n-1} -\log \mathbf{P}(S_n \leq k) < 2 \sum_{n\alpha < k \leq n-1} \mathbf{P}(S_n > k).$$

From (2.6), the right hand side of above inequality is smaller than $4\,\mathbf{Var}[N_D]$, and then from (2.10) we have

$$\sum_{n\alpha < k \leq n-1} -\log \mathbf{P}(S_n \leq k) \leq c_1 \sqrt{n\alpha(1-\alpha)} + c_2 \tag{3.6}$$

for some constants c_1, c_2. Also, by Riemann integration, we have

$$\left| \sum_{0 \leq k \leq n\alpha} I(k/n) - n \int_0^\alpha I(x)\,\mathrm{d}x \right| \leq I(0) - I(\alpha) = -\log(1-\alpha) \tag{3.7}$$

(Since I is a decreasing function on $[0, \alpha]$ and $I(\alpha) = 0$) From (3.2), (3.5), (3.6) and (3.7), we have

$$\left| \log \Delta_n(\alpha) + n^2 \int_0^\alpha I(x)\,\mathrm{d}x \right| \leq$$
$$\frac{1}{2}n\alpha \log(n\alpha) - n\log(1-\alpha) - n\alpha \log c_0 + c_1 \sqrt{n\alpha(1-\alpha)} + c_2.$$

So we conclude that for sufficiently large constant c'

$$\left| \log \Delta_n(\alpha) - \frac{n^2}{2}(\alpha + \log(1-\alpha)) \right| = O(n\alpha \log(n\alpha))$$

uniformly in $c'/n < \alpha < 1 - c'/n$. $\qquad \square$

Fix $x > 0$ and let $\alpha = \frac{x\sqrt{\log n}}{n}$. From Taylor expansion (Note that $\log(1 - \alpha) = -\alpha - \frac{\alpha^2}{2} + O(\alpha^3)$ for, say, $\alpha \leq 1/2$.) and Proposition 3.1 we have

$$\log \Delta_n(\alpha) = -\frac{x^2}{4}\log n + o(\log n)$$

or equivalently

$$\Delta_n(\alpha) = n^{-\frac{x^2}{4} + o(1)}. \tag{3.8}$$

3.2 largest empty cap

Proof of Theorem 1.2. We follow the proof of Theorem 1.3 in [3]. Let $\varepsilon > 0$ be arbitrary and recall that

$$X_n = \frac{n}{8\pi\sqrt{\log n}}M_n.$$

We easily see the inequality

$$\mathbf{E}(|X_n - 1|^s) \leq \epsilon^s + \mathbf{P}(X_n < 1 - \epsilon) + \mathbf{E}(|X_n - 1|^s 1_{\{X_n > 1 + \epsilon\}}).$$

Hence, it suffices to show that the last two terms go to zero as n tends to infinity. Integrating by parts, we get

$$\mathbf{E}(|X_n - 1|^s I_{\{X_n > 1 + \epsilon\}}) = \int_{1+\epsilon}^{+\infty} s(u - 1)^{s-1} \mathbf{P}(X_n > u)\,\mathrm{d}u + \epsilon^s\,\mathbf{P}(X_n > 1 + \epsilon).$$

Next, we give an upper bound for $\mathbf{P}(X_n > u)$. We can choose on $\mathbb{S}^2$ at most $4n$ points $q_1, \ldots, q_m$ so that every spherical cap with area $4\pi/n$ contains at least one of these points. (We can find $m \leq 4n$ spherical caps with area π/n such that there exists no spherical cap with this area that does not intersect these m spherical caps. Let $q_1, \ldots, q_m$ be the center of these m spherical caps. One can check that $q_1, \ldots, q_m$ have desired properties.) Let Y_i, $1 \leq i \leq m$, be the area of the largest empty cap centred at q_i. Note that there exists some q_j within distance $2/\sqrt{n}$ of the center of the largest empty cap. From this, we can conclude that for any $u \geq 1 + \varepsilon$ and sufficiently large n,

$$\mathbf{P}(X_n > u) \leq \mathbf{P}\left(M_n > \frac{8\pi\sqrt{\log n}}{n}(1 + \varepsilon)\right) \tag{3.9}$$

$$\leq \sum_{i=1}^{m} \mathbf{P}\left(Y_i > \frac{8\pi\sqrt{\log n}}{n}(1 + \varepsilon/2)\right).$$

Thus, for sufficiently large n

$$\mathbf{P}(X_n > u) \leq 4n\Delta_n\left(\frac{2\sqrt{\log n}}{n}(1 + \varepsilon/2)\right). \tag{3.10}$$

Using (3.8) and (3.10), we conclude that there is some $\delta > 0$ such that

$$\mathbf{P}(X_n > u) = o(n^{-\delta}) \tag{3.11}$$

uniformly for $u \geq 1 + \varepsilon$. Also, we can write

$$\int_{1+\varepsilon}^{+\infty} s(u - 1)^{s-1}\mathbf{P}(X_n > u)\,\mathrm{d}u = \int_{1+\varepsilon}^{\log n} s(u - 1)^{s-1}\,\mathbf{P}(X_n > u)\,\mathrm{d}u$$

$$+ \int_{\log n}^{\frac{n}{2\sqrt{\log n}}} s(u - 1)^{s-1}\,\mathbf{P}(X_n > u)\,\mathrm{d}u.$$

The first integral on the right hand side goes to zero as $n \to +\infty$, thanks to (3.11). Using a similar argument as in (3.9) and the crude upper bound

$$\Delta_n(\alpha) \leq \mathbf{P}(S_n = 0) = (1 - \alpha)^n$$

we conclude that for any fixed $C > 0$, one has

$$\mathbf{P}(X_n > u) = o(n^{-C})$$

uniformly for $u \geq \log n$. Thus the second integral also goes to zero as $n \to +\infty$.

To prove Theorem 1.2, it suffices to show that $\mathbf{P}(X_n < 1 - \varepsilon)$ converges to zero. The following lemma is similar to Lemma 3.3 in [3]. The proof is based on negative association property for the events such as $\{N_{D_1} = 0\}$ and $\{N_{D_2} = 0\}$ where D_1, D_2 are two disjoint spherical caps on $\mathbb{S}^2$. Note that negative association property holds in this case. See Theorem 1.4 in [18] or proof of Lemma 3.8 in [3].

Lemma 3.2. *Consider a set of disjoint spherical caps on the sphere. Let F_n be the number of such caps free of points $P_1, ..., P_n$. Then* $\mathbf{Var}(F_n) \leq \mathbf{E}(F_n)$.

Now consider $c_n = \Omega\left(n/\sqrt{\log n}\right)$ disjoint spherical caps $D_1, D_2, \ldots, D_{c_n}$ with area $\frac{(1-\epsilon)8\pi\sqrt{\log n}}{n}$. If F_n be the number of such caps free of points $P_1, ..., P_n$ then from previous lemma and Chebyshev's inequality one has

$$\mathbf{P}(X_n < 1 - \epsilon) \leq \mathbf{P}(F_n = 0) \leq \mathbf{P}(|F_n - \mathbf{E}\,F_n| \geq \mathbf{E}\,F_n)$$
$$\leq \frac{\mathbf{Var}(F_n)}{\mathbf{E}(F_n)^2} \leq \frac{1}{\mathbf{E}(F_n)}.$$

Also, using (3.8) we have

$$\mathbf{E}(F_n) = c_n \Delta_n \left(\frac{2(1-\epsilon)\sqrt{\log n}}{n}\right) = n^{1-(1-\epsilon)^2 + o(1)},$$

which implies that $\mathbf{P}(X_n < 1 - \epsilon) \to 0$ as desired. $\square$

3.3 Nearest neighbour statistics

Consider the random point process $\mathcal{X}^{(n)} = \sum_{j=1}^n \delta_{P_j}$. Define for $1 \leq j \leq n$

$$d_j := \min_{i \neq j} |P_i - P_j|$$

the minimum distance from P_j to the remaining points. We define, as in [9], the nearest neighbour spacing measure $\mu(P_1, \ldots, P_n)$ on $[0, +\infty)$ by

$$\mu(P_1, \ldots, P_n) := \frac{1}{n} \sum_{j=1}^n \delta_{\frac{n}{4} d_j^2}. \tag{3.12}$$

Let

$$Q(x) = -\frac{\mathrm{d}}{\mathrm{d}x} E_\infty(x)$$

where

$$E_\infty(x) = \lim_{n \to +\infty} E_n(x) \quad , \quad E_n(x) = \prod_{k=1}^{n-1} e^{-x} \sum_{j=0}^k \frac{x^j}{j!}.$$

We want to show that,

Theorem 3.3. *As $n \to +\infty$,*

$$\mu(P_1, \ldots, P_n) \longrightarrow Q(x)\,\mathrm{d}x. \tag{3.13}$$

One can easily check that for independent uniform points on sphere, this measure converges to $e^{-x} \, \mathrm{d}x$ as n tends to infinity. See 1.

Proof of Theorem 3.3. For fixed $x > 0$ let

$$Y_n = \left| \left\{ P_j : 1 \le j \le n, \frac{n}{4} d_j^2 < x \right\} \right|$$

To prove (3.13) it suffices to show that

$$\frac{Y_n}{n} \xrightarrow{\text{a.s.}} 1 - E_\infty(x). \tag{3.14}$$

First, we compute the expectation of Y_n. Let $\tilde{\Delta}_n(\alpha)$ denote the probability that the spherical cap D with area $4\pi\alpha$ centred at P_j has no other points of $\{P_1, \ldots, P_n\}$ in its interior. With this definition we obtain

$$\mathbf{E}\, Y_n = n(1 - \tilde{\Delta}_n(x/n)). \tag{3.15}$$

We first show that

$$\tilde{\Delta}_n(\alpha) = \prod_{k=1}^{n-1} \mathbf{P}(S_n \le k) \tag{3.16}$$

where S_n is a binomial random variable with parameters n and α. Note that the law of the point process conditioned to have a point at P_j is independent of P_j due to rotational symmetry, so we can assume it to be conditioned at $g^{-1}(0)$ and D to be the spherical cap (with area $4\pi\alpha$) centred at $g^{-1}(0)$. It follows from [29], Theorem 1.7, that the palm measure of a determinantal process is also determinantal and its associated kernel on $\mathbb{C}$, $K^{\{0\}}(z, w)$, is given by

$$K^{\{0\}}(z,w) = \frac{1}{K(0,0)} \det \begin{pmatrix} K(z,w) & K(z,0) \\ K(0,w) & K(0,0) \end{pmatrix} = (1 + z\bar{w})^{n-1} - 1$$

$$= \sum_{k=1}^{n-1} \binom{n-1}{k} (z\bar{w})^k$$

with respect to the measure $\mathrm{d}\mu(z)$. Therefore, $\tilde{\Delta}_n(\alpha)$ is equal to the Fredholm determinant of $I - \mathscr{K}_{g(D)}^{\{0\}}$, where $\mathscr{K}_{g(D)}^{\{0\}}$ is the integral operator with the kernel $K^{\{0\}}(z, w)$ on the disk $g(D)$. Let r be the radius of $g(D)$, as in the proof of Lemma 2.1. Since the functions z^k are orthogonal on $g(D)$, we get for $1 \le k \le n-1$ and $z \in g(D)$,

$$\int_{g(D)} K^{\{0\}}(z,w) w^k \, \mathrm{d}\mu(w) = z^k \int_{g(D)} \frac{n \binom{n-1}{k} |w|^{2k}}{\pi (1 + |w|^2)^{n+1}} \, \mathrm{d}w$$

$$= z^k \int_0^{r^2} \frac{n \binom{n-1}{k} q^k}{(1 + q)^{n+1}} \, \mathrm{d}q.$$

Thus, from (2.4) and (2.5), we conclude that the non-zero eigenvalues of operator $\mathscr{K}_{g(D)}^{\{0\}}$ are equal to $\mathbf{P}(S_n > k)$, $1 \le k \le n-1$, which implies (3.16) as desired (see the remark of lemma 2.1).

Next, we show that

$$|\tilde{\Delta}_n(x/n) - E_n(x)| = O\left(\frac{\log n}{n} \right). \tag{3.17}$$

For this, we will use identity (3.16) and Poisson approximation. Let $\alpha = x/n$. From Theorem 1 of [19] we have

$$\sup_{0 \le k \le n} \left| \mathbf{P}(S_n \le k) - e^{-x} \sum_{j=0}^{k} \frac{x^j}{j!} \right| \le \frac{2x^2}{n} = O(n^{-1}). \tag{3.18}$$

From (2.7) we see that there exists $M = O(\log n)$ such that for every $k \geq M$ we have

$$\mathbf{P}(S_n \leq k) \geq 1 - \frac{1}{n^2} \quad , \quad e^{-x} \sum_{j=0}^{k} \frac{x^j}{j!} \geq 1 - \frac{1}{n^2} \tag{3.19}$$

where, for the second inequality, note that from Taylor's theorem

$$\left| 1 - e^{-x} \sum_{j=0}^{k} \frac{x^j}{j!} \right| \leq e^{-x} \left| e^x - \sum_{j=0}^{k} \frac{x^j}{j!} \right| \leq \frac{x^{k+1}}{(k+1)!}$$

and then use Stirling's formula. Therefore, using the inequality

$$
\begin{aligned}
|\tilde{\Delta}_n(x/n) - E_n(x)| &= \left| \prod_{k=1}^{n-1} \mathbf{P}(S_n \leq k) - \prod_{k=1}^{n-1} e^{-x} \sum_{j=0}^{k} \frac{x^j}{j!} \right| \\
&\leq \sum_{k=1}^{n-1} \left| \mathbf{P}(S_n \leq k) - e^{-x} \sum_{j=0}^{k} \frac{x^j}{j!} \right|,
\end{aligned}
$$

(3.18) and (3.19) we get (3.17). From (3.15) and (3.17) we deduce that

$$\lim_{n \to +\infty} \frac{\mathbf{E}\, Y_n}{n} = 1 - E_\infty(x).$$

In [27] (See Theorem 3.5 therein) it has been shown that for every 1-Lipschitz function on finite counting measures on $\mathbb{S}^2$ such as f, (This means that deleting or adding one point to a configuration of points on $\mathbb{S}^2$ changes f by at most 1. The point process on $\mathbb{S}^2$ can also be viewed as a random counting measure on $\mathbb{S}^2$.) satisfies the following concentration inequality:

$$\mathbf{P}(|f - \mathbf{E}\, f| \geq a) \leq 5 \exp \left(-\frac{a^2}{16(a + 2n)} \right).$$

Fix $r > 0$ and let f count the number of points of point process $\{P_1, \ldots, P_n\}$ such that $d_j < r$. It is simple to check that f is Lipschitz with some constant $c' > 0$. (If one point is added then f increases by at most c'. Using a simple geometric argument, one can choose $c' = 7$.) Applying previous inequality to f/c', we conclude that

$$\mathbf{P}(|Y_n - \mathbf{E}\, Y_n| \geq a) \leq 5 \exp \left(-\frac{a^2}{16c'(a + 2c'n)} \right).$$

Using the Borel-Cantelli Lemma, this gives (3.14) as desired. $\qquad\square$

Remark 3.4. The infinite Ginibre ensemble is a translation invariant determinantal point process on the complex plane $\mathbb{C}$ with kernel $K(z, w) = e^{z\bar{w}}$ with respect to the Gaussian measure $\frac{1}{\pi} e^{-|z|^2}\, dz$. It can be viewed as the local limit of the law of eigenvalues of random matrices from the complex Ginibre ensemble (see [23] for more details). From (3.17) we deduce that

$$\lim_{n \to +\infty} \Delta_n(x/n) = \exp(-x) E_\infty(x).$$

The right-hand side of above equation is equal to the probability that a disk of radius $\sqrt{x}$ in the complex plane contains no points of infinite Ginibre ensemble (see e.g. [16]). Also, if we consider the complex Ginibre ensemble, then $E_n(x)$ is the conditional probability that if one eigenvalue (of an $n \times n$ random matrix with i.i.d. standard complex Gaussian entries) lies at the origin, all $n - 1$ others are further away than $\sqrt{x}$. Compare with the definition of $\tilde{\Delta}$ and see (3.17).
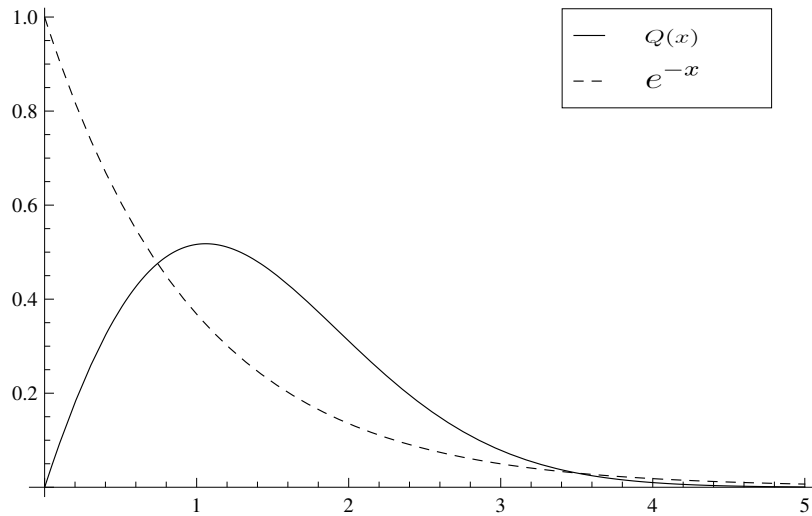
Figure 1: Comparison between the density of nearest neighbour spacing measure μ (see (3.12)) in the case that the points are independently chosen as uniform distribution (on sphere) and spherical ensemble case in the limit $n \to +\infty$.

4 Riesz and logarithmic energy

The purpose of this section is to establish Theorem 1.3 and Corollary 1.4. We start with computing the correlation functions of $\mathcal{X}^{(n)}$ on $\mathbb{S}^2$. Let $\mathring{\rho}_k^{(n)}(p_1, \ldots, p_k)$, $k \geq 1$, be the correlation functions of the point process $\mathcal{X}^{(n)}$ with respect to the surface area measure $\mathrm{d}\nu(p)$. Set $g(p) = z$ and $g(q) = w$. Since $\mathrm{d}\nu(p) = \frac{4}{(1+|z|^2)^2} \, \mathrm{d}z$ we conclude that

$$\mathring{\rho}_k^{(n)}(p_1, \ldots, p_k) = \det(\mathring{K}^{(n)}(p_i, p_j))_{1 \leq i,j \leq k}$$

where

$$\mathring{K}^{(n)}(p, q) = \frac{n}{4\pi} \frac{K^{(n)}(z, w)}{(1 + |z|^2)^{\frac{n-1}{2}}(1 + |w|^2)^{\frac{n-1}{2}}}.$$

Also, we have

$$|p - q| = \frac{2|z - w|}{\sqrt{1 + |z|^2}\sqrt{1 + |w|^2}}. \tag{4.1}$$

(To see this, one can verify that the right-hand side of (4.1) is invariant under the map $(z, w) \mapsto (\phi(z), \phi(w))$ where transformation $\phi(z)$ is defined by

$$\phi(z) = \frac{\alpha z + \beta}{-\bar{\beta}z + \bar{\alpha}}$$

such that $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. (It suffices to show this for $\phi(z) = \frac{1+z}{1-z}$ and $\phi(z) = \alpha z$ where $|\alpha| = 1$.) These transformations correspond via stereographic projection to the rotations of $\mathbb{S}^2$, see [23]. So it suffices to check (4.1) for $z, w \in \mathbb{C}$ such that $|z| = |w| = 1$. But in this case, from definition of g it is obvious that $|g^{-1}(z) - g^{-1}(w)| = |z - w|$)

Recall that $K^{(n)}(z, w) = (1 + z\bar{w})^{n-1}$. A short computation then shows that

$$
\begin{aligned}
|\mathring{K}^{(n)}(p, q)|^2 &= \left(\frac{n}{4\pi}\right)^2 \left(\frac{1 + |zw|^2 + 2\mathrm{Re}(z\bar{w})}{(1 + |z|^2)(1 + |w|^2)}\right)^{n-1} \\
&= \left(\frac{n}{4\pi}\right)^2 \left(1 - \frac{|z - w|^2}{(1 + |z|^2)(1 + |w|^2)}\right)^{n-1} \\
&= \left(\frac{n}{4\pi}\right)^2 \left(1 - \frac{|p - q|^2}{4}\right)^{n-1}.
\end{aligned}
\tag{4.2}
$$

Thus from above equation, the 2-point correlation function $\mathring{\rho}_2^{(n)}$ is given by

$$
\mathring{\rho}_2^{(n)}(p, q) = \begin{vmatrix} \mathring{K}^{(n)}(p, p) & \mathring{K}^{(n)}(p, q) \\ \mathring{K}^{(n)}(q, p) & \mathring{K}^{(n)}(q, q) \end{vmatrix} = \left(\frac{n}{4\pi}\right)^2 \left(1 - \left(1 - \frac{|p - q|^2}{4}\right)^{n-1}\right).
\tag{4.3}
$$

Proof of Theorem 1.3. We begin with part ii). Similar to (1.1), we have for a suitable test function F

$$
\mathbf{E} \sum_{i \neq j} F(P_i, P_j) = \int_{\mathbb{S}^2 \times \mathbb{S}^2} F(p, q) \mathring{\rho}_2^{(n)}(p, q) \, \mathrm{d}\nu(p) \, \mathrm{d}\nu(q).
\tag{4.4}
$$

Thus by setting $F(p, q) = |p - q|^{-s}$ we obtain

$$
\mathbf{E} E_s(P_1, \ldots, P_n) = \int_{\mathbb{S}^2 \times \mathbb{S}^2} |p - q|^{-s} \mathring{\rho}_2^{(n)}(p, q) \, \mathrm{d}\nu(p) \, \mathrm{d}\nu(q).
$$

Notice that the point process is invariant in distribution under isometries of the sphere and so by Fubini's theorem we conclude that

$$
\mathbf{E} E_s(P_1, \ldots, P_n) = 4\pi \int_{\mathbb{S}^2} |p - q|^{-s} \mathring{\rho}_2^{(n)}(p, q) \, \mathrm{d}\nu(p)
$$

where $q = g^{-1}(0)$. Thus we can write

$$
\mathbf{E} E_s(P_1, \ldots, P_n) = \frac{n^2}{2^s \pi} \int_{\mathbb{C}} |z|^{-s} \frac{(1 + |z|^2)^{n-1} - 1}{(1 + |z|^2)^{n+1-s/2}} \, \mathrm{d}z.
$$

(This can also be obtained directly from Equation 1.1, letting

$$
F(z, w) = \left(\frac{2|z - w|}{\sqrt{1 + |z|^2}\sqrt{1 + |w|^2}}\right)^{-s}
$$

and using suitable linear fractional transformations corresponding to the rotations of $\mathbb{S}^2$ by stereographic projection). Changing to polar coordinates, we get

$$
\mathbf{E} E_s(P_1, \ldots, P_n) = \frac{n^2}{2^{s-1}} \int_0^{+\infty} \frac{1}{r^{s-1}} \frac{(1 + r^2)^{n-1} - 1}{(1 + r^2)^{n+1-s/2}} \, \mathrm{d}r.
\tag{4.5}
$$

Making the change of variable $u = r^2$, we see that

$$
\begin{aligned}
\mathbf{E} E_s(P_1, \ldots, P_n) &= \frac{n^2}{2^s} \sum_{j=1}^{n-1} \binom{n-1}{j} \int_0^{+\infty} \frac{1}{u^{s/2 - j}(1 + u)^{n+1-s/2}} \, \mathrm{d}u \\
&= \frac{n^2}{2^s} \sum_{j=1}^{n-1} \frac{\Gamma(n)\Gamma(j + 1 - s/2)}{\Gamma(n + 1 - s/2)\Gamma(j + 1)}.
\end{aligned}
\tag{4.6}
$$

In the last line, we have used the beta function identity

$$\int_0^{+\infty} \frac{1}{u^{s/2-j}(1+u)^{n+1-s/2}} \, \mathrm{d}u = \int_0^1 t^{j-s/2}(1-t)^{n-j-1} \, \mathrm{d}t$$
$$= \frac{\Gamma(j+1-s/2)\Gamma(n-j)}{\Gamma(n+1-s/2)}.$$

By using induction on n, one can show that for $s \neq 2$,

$$\sum_{j=0}^{n-1} \frac{\Gamma(j+1-s/2)}{\Gamma(j+1)} = \frac{\Gamma(n+1-s/2)}{(1-s/2)\Gamma(n)}. \tag{4.7}$$

Note that Γ satisfies the functional equation $x\Gamma(x) = \Gamma(x+1)$ which implies that

$$\frac{\Gamma(n+1-s/2)}{(1-s/2)\Gamma(n)} + \frac{\Gamma(n+1-s/2)}{\Gamma(n+1)} = \frac{(n+1-s/2)\Gamma(n+1-s/2)}{(1-s/2)\Gamma(n+1)}$$
$$= \frac{\Gamma(n+2-s/2)}{(1-s/2)\Gamma(n+1)}$$

Combining (4.7) with (4.6), we obtain (1.10) as required.

For $s = 2$, from (4.6) we have

$$\mathbf{E} \, E_2(P_1, \ldots, P_n) = \frac{n^2}{4} \sum_{j=1}^{n-1} \frac{1}{j}.$$

Also, from Euler-Maclaurin Summation Formula, we know (see the first part of the proof of Theorem 3 and equation (2.1) in [7])

$$\sum_{j=1}^n \frac{1}{j} = \log n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + O\left(\frac{1}{n^4}\right), \tag{4.8}$$

and hence

$$\mathbf{E} \, E_2(P_1, \ldots, P_n) = \frac{1}{4}n^2 \log n + \frac{\gamma}{4}n^2 - \frac{n}{8} - \frac{1}{48} + O\left(\frac{1}{n^2}\right).$$

It remains to prove (1.9). By differentiation right hand side of (4.5) with respect to s at $s = 0$, we conclude that

$$\mathbf{E} \, E_{\log}(P_1, \ldots, P_n) = \frac{\mathrm{d}}{\mathrm{d}s} \, \mathbf{E} \, E_s(P_1, \ldots, P_n)\Big|_{s=0^+}$$
$$= \frac{\mathrm{d}}{\mathrm{d}s} \left[\frac{2^{1-s}}{2-s}n^2 - \frac{\Gamma(n)\Gamma(1-s/2)}{2^s\Gamma(n+1-s/2)}n^2\right]\Big|_{s=0^+}.$$

Thus,

$$\mathbf{E} \, E_{\log}(P_1, \ldots, P_n) = \left(\frac{1}{2} - \log 2\right)n^2 -$$
$$-\frac{-\frac{1}{2}\Gamma'(1)\Gamma(n+1) - \left(\Gamma(n+1)\log 2 - \frac{1}{2}\Gamma'(n+1)\right)}{\Gamma(n+1)^2}\Gamma(n)n^2$$

Moreover, we have (see, e.g. 6.3.1-2 of [1])

$$\Gamma'(n) = \Gamma(n)\left(-\gamma + \sum_{j=1}^{n-1} \frac{1}{j}\right) \quad, \quad \Gamma'(1) = -\gamma.$$

Using above identities, it follows that

$$\mathbf{E}\, E_{\log}(P_1, \ldots, P_n) = \left(\frac{1}{2} - \log 2 \right) n^2 + \left(-\frac{\gamma}{2} + \log 2 - \frac{1}{2} \left(-\gamma + \sum_{j=1}^{n} \frac{1}{j} \right) \right) n.$$

Finally, from (4.8) we obtain (1.9). □

Proof of Corollary 1.4. For $0 < s < 2$, from Theorem 1.3, we conclude that there exist n-point set $\{x_1, \ldots, x_n\}$ such that

$$E_s(x_1, \ldots, x_n) \leq \frac{2^{1-s}}{2-s} n^2 - \frac{\Gamma(n)\Gamma(1-s/2)}{2^s \Gamma(n+1-s/2)} n^2.$$

Thus, by definition

$$\mathcal{E}_s(n) - \frac{2^{1-s}}{2-s} n^2 \leq -\frac{\Gamma(n)\Gamma(1-s/2)}{2^s \Gamma(n+1-s/2)} n^2. \tag{4.9}$$

It is well known that (see e.g. (1) of [38] or 6.1.46 of [1])

$$\lim_{n \to +\infty} \frac{\Gamma(n) n^{1-s/2}}{\Gamma(n+1-s/2)} = 1.$$

Also, For $0 < \beta < 1$ and $x > 0$, we have (see (7) of [38])

$$\frac{x}{(x+\beta)^{1-\beta}} \leq \frac{\Gamma(x+\beta)}{\Gamma(x)} \leq x^{\beta}. \tag{4.10}$$

Hence, for $0 < s < 2$,

$$\frac{\Gamma(n+1-s/2)}{\Gamma(n)} \leq n^{1-s/2}.$$

Consequently, from (4.9) and above inequality we obtain for $0 < s < 2$

$$\mathcal{E}_s(n) - \frac{2^{1-s}}{2-s} n^2 \leq -\frac{\Gamma(1-s/2)}{2^s} n^{1+s/2},$$

as desired. If we set $\beta = s/2$ and $x = 1 - s/2$ in (4.10), we get for $0 < s < 2$, $\Gamma(1-s/2) > 1$. Thus, for $0 < s < 2$, we have

$$\frac{\Gamma(1-s/2)}{2^s} > (2\sqrt{2\pi})^{-s}$$

and this shows that the bound (1.12) is better than (1.5). See Figure 2.
Similarly, for $-2 < s < 0$, from (4.10), we have

$$\frac{\Gamma(n+1-s/2)}{\Gamma(n)} = (n-s/2)\frac{\Gamma(n-s/2)}{\Gamma(n)} \geq n(n-s/2)^{-s/2} \geq n^{1-s/2}. \tag{4.11}$$

Therefore, using similar argument as above, we get (1.13). If we set $\beta = -s/2$ and $x = 1$ in (4.10), we then have for $-2 < s < 0$, $\Gamma(1-s/2) < 1$, thus

$$\frac{\Gamma(1-s/2)}{2^s} < (2\sqrt{2\pi})^{-s}$$

which implies that the bound (1.13) is better than (1.6). See Figure 2.

From (1.11), we easily obtain (1.14). To control the error term $O(1/n^2)$ in (1.11) see the proof of Theorem 3 in [7]. This term is positive and bounded by $\frac{1}{480n^2}$. □
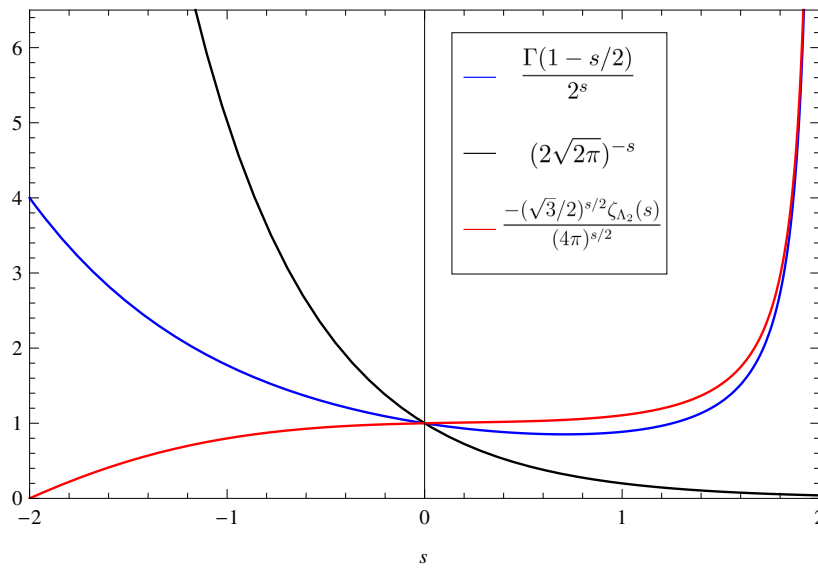
Figure 2: Blue curve: The bound for $(n^2 - \mathcal{E}_s(n))/n^{1+s/2}$ given by Corollary 1.4. Black curve: The asymptotic bound for $(n^2 - \mathcal{E}_s(n))/n^{1+s/2}$ given by 1.5 and 1.6. Red curve: conjectured value of $(n^2 - \mathcal{E}_s(n))/n^{1+s/2}$ as $n \to +\infty$.

Remark 4.1. For $s \geq 4$, the expected value of Riesz s-energy for $\mathcal{X}^{(n)}$ is infinite. Compare with the points that are chosen randomly and independently on the sphere, with the uniform distribution. in this case, the expected value of Riesz s-energy is infinite for $s \geq 2$.

Remark 4.2. Another interesting random point process on $\mathbb{S}^2$ is the roots of random polynomials via the stereographic projection. Let $f(z) = \sum_{j=0}^n a_j z^j$ where the coefficients a_j are independent complex Gaussian random variables with mean 0 and variance $\binom{n}{j}$. Let $z_1, \ldots, z_n$ are the complex zeros of $f(z)$. In [14] the expectation of the logarithmic energy for point process $\{g^{-1}(z_1), \ldots, g^{-1}(z_n)\}$ was computed. It was shown that

$$\mathbf{E}\, E_{\log}(g^{-1}(z_1), \ldots, g^{-1}(z_n)) = \left(\frac{1}{2} - \log 2\right) n^2 - \frac{1}{2} n \log n - \left(\frac{1}{2} - \log 2\right) n.$$

Compare with (1.9).

Remark 4.3. Another criterion for the quality of the distribution of points on the sphere is the spherical cap L^2-discrepancy, which is given by

$$\mathbf{D}_2(\{x_1, \ldots, x_n\}) = \left(\int_{-1}^1 \frac{1}{4\pi} \int_{\mathbb{S}^2} \left| \sum_{j=1}^n 1_{D(p,t)}(x_j) - \frac{n|D(p,t)|}{4\pi} \right|^2 \mathrm{d}\nu(p)\, \mathrm{d}t \right)^{1/2}.$$

Here $D(p, t) = \{q \in \mathbb{S}^2 : \langle p, q \rangle \leq t\}$. Stolarsky's invariance theorem says that (see [33])

$$\frac{1}{2} \sum_{i,j=1}^n |x_i - x_j| + \mathbf{D}_2^2(\{x_1, \ldots, x_n\}) = \frac{2}{3} n^2$$

Consequently, by letting $s = -1$ in (1.10) we see that

$$\mathbf{E}\, \mathbf{D}_2^2(P_1, \ldots, P_n) = \frac{\Gamma(3/2)\Gamma(n)n^2}{\Gamma(n+3/2)}.$$

From (4.11) and taking $s = -1$, we have $\Gamma(n + 3/2) \geq n^{3/2}\Gamma(n)$ and thus

$$\mathbf{E}\, \mathbf{D}_2^2(P_1, \ldots, P_n) \leq \Gamma(3/2)\sqrt{n}.$$

By the Cauchy-Schwarz inequality and above inequality, we then obtain, for $n \geq 2$,

$$\mathbf{E}\, \mathbf{D}_2(P_1, \ldots, P_n) \leq \sqrt{\Gamma(3/2)}\,\sqrt[4]{n}.$$

5 Minimum spacing

Point pair statistics

Recall the definition (1.15) of the function $G_{t,n}$. We first compute the expectation of $G_{t,n}$. The argument is similar to what was done in the proof of Theorem 1.3 in Section 4. Let $F(p, q) = 1_{\{|p-q| \leq t\}}$. From (4.4) and Fubini's theorem we have

$$2\,\mathbf{E}\, G_{t,n} = \int_{\mathbb{S}^2 \times \mathbb{S}^2} F(p, q)\mathring{\rho}_2^{(n)}(p, q)\, \mathrm{d}\nu(p)\, \mathrm{d}\nu(q) \tag{5.1}$$

$$= 4\pi \int_{\mathbb{S}^2} F(p, q)\mathring{\rho}_2^{(n)}(p, q')\, \mathrm{d}\nu(p)$$

$$= 4\pi \int_{\{p:|p-q'| \leq t\}} \mathring{\rho}_2^{(n)}(p, q')\, \mathrm{d}\nu(p).$$

where $q' = g^{-1}(0)$. Since $\mathrm{d}\nu(p) = \frac{4}{(1+|z|^2)^2}\mathrm{d}z$, from (4.1) and (4.3) we obtain that

$$\mathbf{E}\, G_{t,n} = \frac{n^2}{2\pi} \int_{\{z \in \mathbb{C}:|z| \leq \frac{t}{\sqrt{4-t^2}}\}} \left(\frac{1}{(1 + |z|^2)^2} - \frac{1}{(1 + |z|^2)^{n+1}} \right) \mathrm{d}z$$

$$= n^2 \int_0^{\frac{t}{\sqrt{4-t^2}}} r \left(\frac{1}{(1 + r^2)^2} - \frac{1}{(1 + r^2)^{n+1}} \right) \mathrm{d}r$$

$$= \frac{n^2 t^2}{8} - \frac{n}{2}\left(1 - (1 - t^2/4)^n \right).$$

If we set $t = o(\frac{1}{\sqrt{n}})$ then we see that

$$\mathbf{E}\, G_{t,n} = \frac{n^3 t^4}{64}(1 + o(1))$$

and for $t = \frac{x}{n^{3/4}}$, where $x > 0$ is fixed, we get

$$\lim_{n \to +\infty} \mathbf{E}\, G_{t,n} = \frac{x^4}{64}. \tag{5.2}$$

The above equation shows that the correct scaling for the minimum spacing is $n^{-3/4}$. To prove Theorem 1.5, similar to [3], we will use a modification of the method from Soshnikov. This method has been used in [3] and [32] (see also [30]) to solve similar problems in one-dimensional case.

The following two lemmas will be used frequently in the proof.

Lemma 5.1. *For $p, q \in \mathbb{S}^2$ such that $|p - q| = O(n^{-3/4})$ we have*

$$\mathring{\rho}_2^{(n)}(p, q) = O(n^{3/2}). \tag{5.3}$$

Proof. The proof is immediate from equation (4.3). $\qquad\square$

The next lemma will be used to control the k-point correlation function in terms of the lower order correlation functions. For the proof see [20].

Lemma 5.2. *(Hadamard-Fischer inequality) Let M be an $n \times n$ (Hermitian) positive definite matrix and let $\omega \subset \{1, 2, \ldots, n\}$ be an index set. Let M_ω be the submatrix of M using rows and columns numbered in ω. Then*

$$\det(M) \leq \det(M_\omega) \det(M_{\bar{\omega}})$$

where $\bar{\omega} = \{1, 2, \ldots, n\}\backslash\omega$.

Fix $x > 0$ and let $t = \frac{x}{n^{3/4}}$. Define for $p \in \mathbb{S}^2$

$$C^{(n)}(p) = \{q \in \mathbb{S}^2 : |p - q| \leq t\}$$

Consider the random point process

$$\mathcal{X}^{(n)} = \sum_{i=1}^{n} \delta_{P_i}.$$

We define a new point process $\tilde{\mathcal{X}}^{(n)}$. First consider all unordered pairs (P_i, P_j) such that $|P_i - P_j| < t$ and

$$C^{(n)}(P_i) = \{P_j\} \quad , \quad C^{(n)}(P_j) = \{P_i\}.$$

Then from each pair select independently with probability $\frac{1}{2}$ one of the two items, and consider all this points as $\tilde{\mathcal{X}}^{(n)}$. (Compare this with the modified processes that have been used in [3, 32] in similar cases.) Then let $Z_n = \tilde{\mathcal{X}}^{(n)}(\mathbb{S}^2)$.

Lemma 5.3. *For fixed $x > 0$, let $t = \frac{x}{n^{3/4}}$. Then*

$$G_{t,n} - Z_n \xrightarrow{d} 0.$$

Proof. We show that $G_{t,n} - Z_n \xrightarrow{L^1} 0$. First note that we have

$$\mathbf{E}\,|G_{t,n} - Z_n| \leq \int_{\mathbb{S}^2} \int_{C^{(n)}(p) \times C^{(n)}(p)} \mathring{\rho}_3^{(n)}(p, q_1, q_2)\, \mathrm{d}\nu(q_1)\, \mathrm{d}\nu(q_2)\, \mathrm{d}\nu(p). \qquad (5.4)$$

To see this, observe that $G_{t,n} - Z_n$ is equal to number of pairs (P_i, P_j) such that $0 < |P_i - P_j| < t$ and there exist some point P_k where $0 < |P_i - P_k| < t$ or $0 < |P_j - P_k| < t$. By considering the triples (P_i, P_j, P_k) or (P_j, P_i, P_k) we see that

$$G_{t,n} - Z_n \leq \sum_{\substack{i,j,k \\ \text{pairwise distinct}}} 1_{\{P_j \in C^{(n)}(P_i), P_k \in C^{(n)}(P_i)\}}.$$

Thus, taking expectation gives (5.4).

If $q_1, q_2 \in C^{(n)}(p)$ then $|q_1 - q_2| = O(n^{-3/4})$ and using Hadamard-Fischer inequality and (5.3) we obtain

$$\mathring{\rho}_3^{(n)}(p, q_1, q_2) \leq \mathring{\rho}_1^{(n)}(p)\mathring{\rho}_2^{(n)}(q_1, q_2) = \frac{n}{4\pi}\mathring{\rho}_2^{(n)}(q_1, q_2) = O(n^{5/2}).$$

By integrating on the domain of area $4\pi|C^{(n)}(p)|^2 = O(n^{-3})$, we conclude $\mathbf{E}\,|G_{t,n} - Z_n| \to 0$ as n goes to infinity. $\qquad \square$

Our goal is to prove $G_{t,n} \xrightarrow{d} \text{Poisson}(\frac{x^4}{64})$ where $t = \frac{x}{n^{3/4}}$. Thanks to the Lemma 5.3, it thus suffices to show that $Z_n \xrightarrow{d} \text{Poisson}(\frac{x^4}{64})$. Denote $\tilde{\rho}_k^{(n)}(p_1, \ldots, p_k)$ as the k-point correlation function $\tilde{\mathcal{X}}^{(n)}$. From (1.1) we have

$$\mathbf{E}\,\frac{Z_n!}{(Z_n - k)!} = \int_{(\mathbb{S}^2)^k} \tilde{\rho}_k^{(n)}(p_1, \ldots, p_k)\, \mathrm{d}\nu(p_1) \ldots \mathrm{d}\nu(p_k).$$

So using the moment method it suffices to show that for every $k \geq 1$

$$\int_{(\mathbb{S}^2)^k} \tilde{\rho}_k^{(n)}(p_1, \ldots, p_k) \, \mathrm{d}\nu(p_1) \ldots \mathrm{d}\nu(p_k) \underset{n \to \infty}{\longrightarrow} \left(\frac{x^4}{64} \right)^k. \tag{5.5}$$

(The k-th factorial moment of the Poisson distribution with mean λ is equal to λ^k.)

Proof of Theorem 1.5. Let $p_1, \ldots, p_k$ be fixed distinct elements in $(\mathbb{S}^2)^k$. First we show that

$$\tilde{\rho}_k^{(n)}(p_1, \ldots, p_k) \underset{n \to \infty}{\longrightarrow} \left(\frac{x^4}{256\pi} \right)^k. \tag{5.6}$$

By taking n large enough we can assume that for $i \neq j$,

$$|p_i - p_j| > 4t.$$

From inclusion-exclusion argument, we have (see [32])

$$\tilde{\rho}_k^{(n)}(p_1, \ldots, p_k) =$$

$$\frac{1}{2^k} \sum_{m=0}^{n-2k} \frac{(-1)^m}{m!} \int_{C^{(n)}(p_k)} \cdots \int_{C^{(n)}(p_1)} \int_{\left(\bigsqcup_{i=1}^k C^{(n)}(p_i) \cup C^{(n)}(q_i) \right)^m}$$

$$\mathring{\rho}_{2k+m}^{(n)}(p_1, q_1, \ldots, p_k, q_k, r_1, \ldots, r_m) \mathrm{d}\nu(r_1) \ldots \mathrm{d}\nu(r_m) \, \mathrm{d}\nu(q_1) \ldots \mathrm{d}\nu(q_k). \tag{5.7}$$

First consider the $m = 0$ case. From determinantal formula, we have

$$\mathring{\rho}_{2k}^{(n)}(p_1, q_1, \ldots, p_k, q_k) = \det_{1 \leq i,j \leq k} \begin{pmatrix} \mathring{K}^{(n)}(p_i, p_j) & \mathring{K}^{(n)}(p_i, q_j) \\ \mathring{K}^{(n)}(q_i, p_j) & \mathring{K}^{(n)}(q_i, q_j) \end{pmatrix}. \tag{5.8}$$

Consider the (i, j)-th 2×2 block of above determinant. First assume that $i \neq j$. Since p_i and p_j are fixed and distinct we conclude that if $q_i \in C^{(n)}(p_i)$ and $q_j \in C^{(n)}(p_j)$, then for sufficiently large n, $|p_i - p_j|, |p_i - q_j|, |q_i - p_j|$ and $|q_i - q_j|$ are greater than a positive constant (say $|p_i - p_j|/2$). Thus from (4.2) we deduce that all terms of this block are exponentially small in n.

Now if $i = j$, then from (5.3) the determinant of (i, i)-th 2×2 block is

$$\mathring{\rho}_2^{(n)}(p_i, q_i) = \det \begin{pmatrix} \frac{n}{4\pi} & \mathring{K}^{(n)}(p_i, q_i) \\ \mathring{K}^{(n)}(q_i, p_i) & \frac{n}{4\pi} \end{pmatrix} = O(n^{3/2})$$

where $q_i \in C^{(n)}(p_i)$. So from the expansion of the determinant in (5.8) over all permutations of length $2k$ we conclude that only the terms contain the entries in the diagonal 2×2 blocks can have a non-zero limit (note that $|\mathring{K}^{(n)}| \leq \frac{n}{4\pi}$). Thus for $q_i \in C^{(n)}(p_i), 1 \leq i \leq k$, we get

$$\mathring{\rho}_{2k}^{(n)}(p_1, q_1, \ldots, p_k, q_k) = o(1) + \prod_{i=1}^k \mathring{\rho}_2^{(n)}(p_i, q_i)$$

where $o(1)$ term is exponentially small. Since the integration domain of $\mathring{\rho}_{2k}^{(n)}$ in the $m = 0$ case of (5.7) has size $O(n^{\frac{-3k}{2}})$ then from above equation, (5.1) and (5.2) we obtain

$$\int_{C^{(n)}(p_k)} \cdots \int_{C^{(n)}(p_1)} \mathring{\rho}_{2k}^{(n)}(p_1, q_1, \ldots, p_k, q_k) \, \mathrm{d}\nu(q_1) \ldots \mathrm{d}\nu(q_k)$$

$$= o(1) + \prod_{i=1}^k \int_{C^{(n)}(p_i)} \mathring{\rho}_2^{(n)}(p_i, q_i) \, \mathrm{d}\nu(q_i)$$

$$= o(1) + \left(\frac{\mathbf{E}\, G_{t,n}}{2\pi} \right)^k \to \left(\frac{x^4}{128\pi} \right)^k$$

as $n \to +\infty$. By Hadamard-Fischer inequality we conclude that the contribution of the terms corresponding to $m \geq 1$ in (5.7) is bounded by

$$
\left(\int_{C^{(n)}(p_k)} \dots \int_{C^{(n)}(p_1)} \mathring{\rho}_{2k}^{(n)}(p_1, q_1, \dots, p_k, q_k) \, \mathrm{d}\nu(q_1) \dots \mathrm{d}\nu(q_k) \right) \times
$$

$$
\sum_{m=1}^{n-2k} \frac{1}{m!} \left(\int_{\bigsqcup_{i=1}^{k} \left(C^{(n)}(p_i) \cup C^{(n)}(q_i) \right)} \mathring{\rho}_{1}^{(n)}(r_1) \mathrm{d}\nu(r_1) \right)^m.
$$

The integration domain $\bigsqcup_{i=1}^{k} \left(C^{(n)}(p_i) \cup C^{(n)}(q_i) \right)$ has size $O(n^{-3/2})$ and $\mathring{\rho}_{1}^{(n)}(r_1) = \frac{n}{4\pi}$. Thus the second term of the above product goes to zero as $n \to +\infty$. Also the first factor of the above product is just the $m = 0$ case, which we show converges. Thus the whole expression goes to zero as $n \to +\infty$ and (5.6) is obtained.

Also for all $(p_1, \dots, p_k) \in (\mathbb{S}^2)^k$ we have

$$
\tilde{\rho}_k^{(n)}(p_1, \dots, p_k) \leq \int_{C^{(n)}(p_k)} \dots \int_{C^{(n)}(p_1)} \mathring{\rho}_{2k}^{(n)}(p_1, q_1, \dots, p_k, q_k) \, \mathrm{d}\nu(q_1) \dots \mathrm{d}\nu(q_k)
$$

$$
\leq \prod_{i=1}^{k} \int_{C^{(n)}(p_i)} \mathring{\rho}_{2}^{(n)}(p_i, q_i) \, \mathrm{d}\nu(q_i) = \left(\frac{x^4}{128\pi} \right)^k.
$$

Finally, From (5.6) and the dominated convergence theorem, one obtains (5.5), and the claim follows. $\qquad\square$

References

[1] *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Dover Publications, Inc., New York, 1992. Edited by Milton Abramowitz and Irene A. Stegun, Reprint of the 1972 edition. MR-1225604

[2] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. MR-2760897

[3] Gérard Ben Arous and Paul Bourgade. Extreme gaps between eigenvalues of random matrices. *Ann. Probab.*, 41(4):2648–2681, 2013. MR-3112927

[4] József Beck. Sums of distances between points on a sphere—an application of the theory of irregularities of distribution to discrete geometry. *Mathematika*, 31(1):33–41, 1984. MR-0762175

[5] József Beck. Some upper bounds in the theory of irregularities of distribution. *Acta Arith.*, 43(2):115–130, 1984. MR-0736726

[6] Andrew C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.*, 49:122–136, 1941. MR-0003498

[7] R. P. Boas, Jr. Growth of partial sums of divergent series. *Math. Comp.*, 31(137):257–264, 1977. MR-0440862

[8] Charles Bordenave. On the spectrum of sum and product of non-Hermitian random matrices. *Electron. Commun. Probab.*, 16:104–113, 2011. MR-2772389

[9] J. Bourgain, Z. Rudnick, P. Sarnak, Local statistics of lattice points on the sphere, arXiv:1204.0134.

[10] J. S. Brauchart, D. P. Hardin, and E. B. Saff. The next-order term for optimal Riesz and logarithmic energy asymptotics on the sphere. In *Recent advances in orthogonal polynomials, special functions, and their applications*, volume 578 of *Contemp. Math.*, pages 31–61. Amer. Math. Soc., Providence, RI, 2012. MR-2964138

[11] Tony Cai, Jianqing Fan, and Tiefeng Jiang. Distributions of angles in random packing on spheres. *J. Mach. Learn. Res.*, 14:1837–1864, 2013. MR-3104497

[12] Caillol, J. M. Exact results for a two-dimensional one-component plasma on a sphere. *J. Phys. Lett-Paris.*, 42, no. 12 (1981): 245-247.

[13] Ovidiu Costin and Joel L. Lebowitz. Gaussian fluctuation in random matrices. *Phys. Rev. Lett.*, 75(1):69–72, 1995. MR-3155254

[14] Diego Armentano, Carlos Beltrán, and Michael Shub. Minimizing the discrete logarithmic energy on the sphere: the role of random polynomials. *Trans. Amer. Math. Soc.*, 363(6):2955–2965, 2011. MR-2775794

[15] P. J. Forrester, B. Jancovici, and J. Madore. The two-dimensional Coulomb gas on a sphere: exact results. *J. Statist. Phys.*, 69(1-2):179–192, 1992. MR-1184774

[16] P. J. Forrester. Some statistical properties of the eigenvalues of complex random matrices. *Phys. Lett. A*, 169(1-2):21–24, 1992. MR-1181356

[17] P.J. Forrester, Log-gases and random matrices (LMS-34), Princeton University Press, 2010. MR-2641363

[18] S .Ghosh, Determinantal processes and completeness of random exponentials: the critical case, arXiv:1211.2435

[19] J. L. Hodges, Lucien Le Cam. The Poisson Approximation to the Poisson Binomial Distribution. *Ann. Math. Stat.* 737-740, 1960. MR-0117812

[20] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1985. MR-0832183

[21] Frank den Hollander. *Large deviations*. Vol. 14. American Mathematical Soc., 2008. MR-1739680

[22] J. Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. Determinantal processes and independence. *Probab. Surv.*, 3:206–229, 2006. MR-2216966

[23] J. Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2009. MR-2552864

[24] Manjunath Krishnapur. From random matrices to random analytic functions. *Ann. Probab.*, 37(1):314–346, 2009. MR-2489167

[25] M. Krishnapur, Zeros of random analytic functions, Ph.D. thesis, U.C. Berkeley (2006). Preprint available at arXiv:math/0607504v1 [math.PR]. MR-2709142

[26] A. B. J. Kuijlaars and E. B. Saff. Asymptotics for minimal discrete energy on the sphere. *Trans. Amer. Math. Soc.*, 350(2):523–538, 1998. MR-1458327

[27] Robin Pemantle and Yuval Peres. Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. *Combin. Probab. Comput.*, 23(1):140–160, 2014. MR-3197973

[28] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou. Minimal discrete energy on the sphere. *Math. Res. Lett.*, 1(6):647–662, 1994. MR-1306011

[29] T. Shirai, Y. Takahashi. Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes. *J. Funct. Anal.*, 205, no. 2 (2003): 414-463. MR-2018415

[30] Alexander Soshnikov. Level spacings distribution for large random matrices: Gaussian fluctuations. *Ann. of Math. (2)*, 148(2):573–617, 1998. MR-1668559

[31] Alexander Soshnikov. Gaussian limit for determinantal random point fields. *Ann. Probab.*, 30(1):171–187, 2002. MR-1894104

[32] Alexander Soshnikov. Statistics of extreme spacing in determinantal random point processes. *Mosc. Math. J.*, 5(3):705–719, 744, 2005. MR-2241818

[33] Kenneth B. Stolarsky. Sums of distances between points on a sphere. II. *Proc. Amer. Math. Soc.*, 41:575–582, 1973. MR-0333995

[34] M. Talagrand. The Generic Chaining: Upper and Lower Bounds of Stochastic Processes. *Springer Monogr. Math.*, Springer-Verlag, Berlin, 2005. MR-2133757

[35] Terence Tao and Van Vu. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, 38(5):2023–2065, 2010. With an appendix by Manjunath Krishnapur. MR-2722794

[36] Gerold Wagner. On means of distances on the surface of a sphere (lower bounds). *Pacific J. Math.*, 144(2):389–398, 1990. MR-1061328

[37] Gerold Wagner. On means of distances on the surface of a sphere. II. Upper bounds. *Pacific J. Math.*, 154(2):381–396, 1992. MR-1159518

[38] J. G. Wendel. Note on the gamma function. *Amer. Math. Monthly*, 55:563–564, 1948. MR-0029448

[39] M. E. Wise. The incomplete beta function as a contour integral and a quickly converging series for its inverse. *Biometrika*, 37:208–218, 1950. MR-0040622

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])

- Non profit, sponsored by IMS[2], BS[3], PKP[4]

- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] OJS: Open Journal Systems http://pkp.sfu.ca/ojs/
[2] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[3] BS: Bernoulli Society http://www.bernoulli-society.org/
[4] PK: Public Knowledge Project http://pkp.sfu.ca/
[5] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm