THE EIGENVALUE POINT PROCESS FOR SYMMETRIC GROUP PERMUTATION REPRESENTATIONS ON $k$-TUPLES

BENJAMIN TSOU

Abstract. Equip the symmetric group $\mathfrak{S}_n$ with the Ewens distribution. We study the eigenvalue point process of the permutation representation of $\mathfrak{S}_n$ on $k$-tuples of distinct integers chosen from the set $\{1, 2, ..., n\}$. Taking $n \to \infty$, we find the limiting point process in the microscopic regime, i.e. when the eigenvalue point process is viewed at the scale of the mean eigenvalue spacing. A formula for the limiting eigenvalue gap probability in an interval is also given. In certain cases, a power series representation exists and a combinatorial procedure is given for computing the coefficients.

AMS MSC (2010): Primary 60B20, 60F99, 60C05; secondary 11K60, 11L15.

Keywords: random matrices; symmetric group; permutation representations; eigenvalue point processes; equidistribution; discrepancy; exponential sums.

1. Introduction

Historically, much of the attention in random matrix theory has focused on the spectrum of continuous matrix ensembles such as random Hermitian matrices or compact Lie groups (e.g. $O(n)$ or $U(n)$) equipped with Haar measure. Although less studied, there has been increasing interest in the asymptotic eigenvalue behavior of discrete matrix groups and in particular discrete (finite) subgroups of $U(n)$. Perhaps the most natural family of such subgroups to consider are the $n \times n$ permutation matrices, which can be thought of as a (faithful) representation of the symmetric group $\mathfrak{S}_n$ in $U(n)$. In [7], Diaconis and Shahshahani note that the spectrum of uniformly distributed permutation matrices (thought of as a random probability measure) converges weakly to a point mass at the Haar measure on the unit circle. Wieand [31] has shown that eigenvalue fluctuations of uniformly distributed permutation matrices satisfy an asymptotically Gaussian law. These results have been generalized via wreath products to more general families of finite groups. In [32], Wieand extends her limiting normality results to representations corresponding to the natural action of $(\mathbb{Z}_{K_1} \times ... \times \mathbb{Z}_{K_M}) \wr \mathfrak{S}_n$ and $\mathbb{T}^M \wr \mathfrak{S}_n$ on $\mathbb{T}^M \times \{1, ..., n\}$. Evans [9] studied the eigenvalue distribution of the natural permutation representation of the $n$-fold wreath product $G \wr G ... \wr G$ of a permutation group $G$ acting transitively on a set.

Rather than looking at representations arising from the natural permutation action of more complicated permutation groups, one can take a somewhat different perspective and instead consider higher dimensional representations of the symmetric group itself. In particular, let $\rho_{n,k}$ denote the permutation representation of $\mathfrak{S}_n$ on $k$-tuples of distinct integers chosen from the set $[n] := \{1, ..., n\}$ induced from the usual action of $\mathfrak{S}_n$ on $[n]$. In [29], the author studied the fluctuation of eigenvalue statistics of $\rho_{n,k}$ (as well as a few related representations) in a given arc of the unit circle as $n \to \infty$ when $\mathfrak{S}_n$ is equipped with the Ewens measure and showed convergence to a compactly supported limiting distribution. This corresponds to the macroscopic regime in which the arc being considered stays fixed and does not shrink as $n \to \infty$. In this paper, we instead investigate the microscopic regime and characterize the limiting behavior of the eigenvalue point process of $\rho_{n,k}$ when viewed at the scale of the average eigenvalue spacing.
To begin, we define the representation \( \rho_{n,k} \) precisely. Let \( Q_{n,k} \) be the set of ordered \( k \)-tuples \((t_1, \ldots, t_k)\) of distinct integers chosen from the set \([n]\). The symmetric group \( S_n \) acts naturally on this set by \( \sigma(t_1, \ldots, t_k) = (\sigma(t_1), \ldots, \sigma(t_k)) \). We can form the \( \frac{n!}{(n-k)!} \)-dimensional \( \mathbb{C} \)-vector space \( V_{n,k} \) with basis elements \( e_{(t_1, \ldots, t_k)} \). Then the action of \( S_n \) on \( Q_{n,k} \) induces the permutation representation \( \rho_{n,k} : S_n \to O(V_{n,k}) \) where \( O(V_{n,k}) \) is the orthogonal group on \( V_{n,k} \).

For \( \theta > 0 \), recall that the Ewens measure with parameter \( \theta \) is given by

\[
\nu_{n, \theta}(\sigma) = \frac{\theta K(\sigma)}{\theta(\theta + 1) \cdots (\theta + n - 1)}
\]

where \( K(\sigma) \) denotes the number of cycles of \( \sigma \). A random permutation chosen from this distribution is called \( \theta \)-biased. Note that the uniform distribution corresponds to the case \( \theta = 1 \). See [10] for historical motivation of this measure from population genetics.

Endowing \( S_n \) with the Ewens measure induces the corresponding eigenvalue point process of \( \rho_{n,k} \). When \( k = 1 \), the eigenvalues of \( \rho_{n,1}(\sigma) \) have a simple characterization. Since all the eigenvalues are of the form \( e^{2\pi i \phi} \), it will be convenient to refer to each eigenvalue \( e^{2\pi i \phi} \) by its eigenangle \( \phi \in [0, 1) \). For each \( \sigma \in S_n \), let \( C_j^{(n)}(\sigma) \) denote the number of cycles of \( \sigma \) of length \( j \). For \( \rho_{n,1} \), the set of eigenangles corresponding to a \( j \)-cycle of a permutation is \( \left\{ 0, \frac{1}{j}, \ldots, \frac{j-1}{j} \right\} \).

Thus, the point process of eigenangles is given by

\[
\Lambda_n = \sum_{j=1}^{n} C_j^{(n)} \sum_{i=0}^{j-1} \delta_{i/j}
\]

It is not hard to see (as noted in [7]) that the empirical spectral measure \( \Lambda_n/n \) converges weakly to the point mass at the Haar measure on the unit circle. To obtain a scaled point process limit, we renormalize by a factor of \( n \) and extend the process to the entire real line by noting that the angles are periodic with period 1. Then we are led to consider the convergence of the point process

\[
\Xi_n^0 := \sum_{j=1}^{n} C_j^{(n)} \sum_{i \in \mathbb{Z}} \delta_{ni/j}
\]

as \( n \to \infty \). To clarify the sense of convergence that will be used, we briefly review some notions from the theory of point processes. Daley and Vere-Jones [4, Chapt. 11] discuss various modes of convergence for point processes. Here, we just need a few basic facts.

**Definition 1.1.** Let \( \mathfrak{M} \) be the space of locally finite counting measures on \( \mathbb{R} \) and let \( \mathcal{N} \) be the smallest \( \sigma \)-algebra on \( \mathfrak{M} \) with respect to which the mappings \( N \mapsto N(A) \) are measurable for each \( A \in \mathcal{B}(\mathbb{R}) \). Then a point process \( \xi \) on \( \mathfrak{M} \) is a measurable map \( \xi : \Omega \to \mathfrak{M} \) from a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) to the measurable space \((\mathfrak{M}, \mathcal{N})\). Every point process can be represented as \( \xi = \sum_{i=1}^{N} \delta_{X_i} \), where \( \delta \) denotes Dirac measure, \( N \) is an integer-valued random variable, and \( X_i \) are real-valued random variables. The point process \( \xi \) is called simple if the \( X_i \) are a.s. distinct.

The space of measures \( \mathfrak{M} \) can be made into a completely separable metric space [4, pp. 402-406]. Under the induced metric topology, \( \mathfrak{M} \) is the Borel \( \sigma \)-algebra \( \mathcal{B}(\mathfrak{M}) \). Then weak convergence of point processes can be defined in the usual way: If \( \mu_n \) and \( \mu \) are probability measures on the measurable space \((\mathfrak{M}, \mathcal{B}(\mathfrak{M}))\) then \( \mu_n \to \mu \) weakly iff \( \int f d\mu_n \to \int f d\mu \) for all continuous and bounded functions \( f \) on \( \mathfrak{M} \).

**Definition 1.2.** Given a point process \( \xi \), a Borel set \( A \) is a stochastic continuity set for the process if \( \mathbb{P}\{\xi(\delta A) = 0\} = 1 \). The sequence of point processes \( \xi_n \) converges in the sense of
convergence of fidi distributions if for every finite family \( \{A_1, \ldots, A_k\} \) of bounded continuity sets \( A_i \in \mathcal{B}(\mathbb{R}) \), the joint distributions of \( \{\xi_n(A_1), \ldots, \xi_n(A_k)\} \) converge weakly in \( \mathcal{B}(\mathbb{R}^k) \) to the joint distribution of \( \xi(A_1), \ldots, \xi(A_k) \).

It is often easier to check the following conditions for weak convergence:

**Proposition 1.1** (c.f. [6] p. 137). Weak convergence of point processes is equivalent to convergence of the fidi distributions.

**Proposition 1.2** (c.f. [6] p. 138). Let \( \xi_n \) and \( \xi \) be point processes. Then \( \xi_n \rightarrow \xi \) weakly iff we have the weak convergence of random variables \( \int_{\mathbb{R}} f d\xi_n \rightarrow \int_{\mathbb{R}} f d\xi \) for all continuous and compactly supported \( f : \mathbb{R} \rightarrow \mathbb{R} \).

The Poisson-Dirichlet distribution \( PD(\theta) \) plays a central role in the limit of \( \mathcal{E}_n^0 \) and in general the (appropriately scaled) eigenvalue point process for \( \rho_{n,k} \) as \( n \rightarrow \infty \). Perhaps the most intuitive way to describe the \( PD(\theta) \) distribution is in terms of the \( GEM(\theta) \) distribution. Let \( U_1, U_2, \ldots \) be i.i.d. random variables with density \( \theta(1 - x)^{\theta - 1} \) on \([0, 1]\). Define \( V_j = U_j \prod_{i=1}^{j-1} (1 - U_i) \). By the Borel-Cantelli Lemma, \( \sum V_j = 1 \) with probability 1. The distribution of \( (V_1, V_2, \ldots) \) is called the \( GEM(\theta) \) distribution and there is a visually pleasing interpretation of this as a stick-breaking process. Then the decreasing order statistics of \( GEM(\theta) \) have the \( PD(\theta) \) distribution. For more information, see e.g. [4] pp. 38-49, [2] Sect. 5.7, [22], or [11].

Kingman [15] proved the following:

**Proposition 1.3** (Kingman). Let the random variable \( L_i^{(n)} \) denote the length of the \( i \)-th longest cycle in a \( \theta \)-biased random permutation of size \( n \). Then
\[
\lim_{n \to \infty} n^{-1}(L_i^{(n)}, L_2^{(n)}, \ldots) \overset{d}{\to} (L_1, L_2, \ldots)
\]
as \( n \rightarrow \infty \) where the limiting random vector is distributed \( PD(\theta) \).

Now define
\[
\mathcal{E}_\infty^0 := \sum_{i=1}^{\infty} \sum_{q \in \mathbb{Z}} \delta_{q/L_i}
\]
where \( (L_1, L_2, \ldots) \) is a random vector following the \( PD(\theta) \) distribution.

**Remark 1.1**. We could replace the \( PD(\theta) \)-distributed process \( (L_1, L_2, \ldots) \) with a \( GEM(\theta) \)-distributed process \( (V_1, V_2, \ldots) \) in the definition of \( \mathcal{E}_\infty^0 \) since this just corresponds to reordering the terms in the sum.

Note that \( \mathcal{E}_\infty^0 \) isn’t quite a point process since it has an infinite Dirac mass at 0. However, using Proposition 1.3 it is easy to see that \( \mathcal{E}_n^0 \rightarrow \mathcal{E}_\infty^0 \) weakly in the following sense: \( \int f d\mathcal{E}_n^0 \rightarrow \int f d\mathcal{E}_\infty^0 \) weakly for all continuous, compactly supported \( f \) such that \( f(0) = 0 \). Note that if \( f(0) > 0 \), the distribution of \( \int f d\mathcal{E}_n^0 \) converges weakly to the zero measure (i.e. mass escapes to \( \infty \)) and \( \int f d\mathcal{E}_\infty^0 = \infty \) almost surely. Najnudel and Nikeghbali [18] Prop. 4.1 in fact show a.s. convergence when all the symmetric groups are put on the same probability space through a construction known as virtual permutations.

The point process \( \mathcal{E}_n^\alpha \) can be thought of as the process \( \Lambda_n \) “zoomed in” at 0 by a factor of \( n \). “Zooming in” on different “windows” of the point process \( \Lambda_n \) can give very different limiting behaviors. To be more precise, define for any real number \( \alpha \) the shifted point process
\[
\mathcal{E}_n^\alpha := \sum_{j=1}^{n} C_j^{(n)} \sum_{i \in \mathbb{Z}} \delta_{n(i/j - \alpha)}
\]
(1.5)
which corresponds to $\Lambda_0$ zoomed in at $\alpha$. The number-theoretic properties of $\alpha$ play a very important role in determining the limiting process. Recall that the irrationality measure $\mu(r)$ of a real number $r$ is given by

$$\mu(r) = \inf \left\{ \lambda : \left| r - \frac{p}{q} \right| < \frac{1}{q^\lambda} \text{ has only finitely many integer solutions in } p \text{ and } q \right\}$$  \hspace{1cm} (1.6)

Then we can state the following convergence result, which we prove in Section 3.

**Theorem 1.1.** Let $U_1, U_2, \ldots$ be i.i.d. random variables distributed uniformly on the interval $[0, 1]$ and independent of the PD($\theta$) process $(L_1, L_2, \ldots)$. Then define the point process

$$\mathcal{E}_{\infty}^\alpha := \sum_{i=1}^{\infty} \sum_{q \in \mathbb{Z}} \delta_{(U_i + q)/L_i}$$  \hspace{1cm} (1.7)

If $\alpha \in (0, 1)$ is an irrational number with finite irrationality measure, then $\mathcal{E}_n^\alpha \rightarrow \mathcal{E}_{\infty}^\alpha$ weakly.

**Remark 1.2.** Najnudel and Nikeghbali [18] have studied the limiting eigenvalue distributions of “randomized” permutation matrices in which the 1’s are replaced by i.i.d. random variables taking values in $\mathbb{C}^*$. In the latter half of the paper, they study the eigenvalue point processes for these ensembles in the microscopic regime. Among other results, they show that if the 1’s are replaced by uniformly distributed entries on the complex unit circle, the limiting eigenangle process is also $\mathcal{E}_{\infty}^\ast$. Bahier [3] has recently proved asymptotic normality results on the number of points of the limiting point processes $\mathcal{E}_{\infty}^\ast$ and $\mathcal{E}_{\infty}^\ast$ lying in an interval of size tending to infinity. It is easy to see that $\mathcal{E}_{\infty}^\ast$ is a stationary simple point process with intensity 1, i.e. $\mathbb{E}[\mathcal{E}_n^\ast]$ is Lebesgue measure on $\mathbb{R}$.

In Theorem 1.1 we show that at rational angles $\alpha = s/t$ where $s$ and $t$ are relatively prime, $\mathcal{E}_n^\ast$ converges weakly to some limit $\mathcal{E}_\infty^\ast$. The point process $\mathcal{E}_{\infty}^\ast$ is defined essentially like $\mathcal{E}_{\infty}^\ast$ except that the $U_i$ are discrete random variables uniform over the $t$ fractions $1/t, 2/t, \ldots, 1$ instead of the interval $[0, 1]$.

Now let us consider the eigenvalue point process for general $k$. For $\sigma \in \mathfrak{S}_n$, let $\sigma_k$ be the permutation corresponding to $\rho_{n,k}(\sigma)$ and $C^{(n)}_{j,k}(\sigma)$ be the number of $\sigma_k$-cycles of length $j$. Note that $C^{(n)}_{j,k}(\sigma) > 0$ iff there exist $i_1, \ldots, i_k$ (not necessarily distinct) such that $\prod_{i=1}^{k} C^{(n)}_{i,i}(\sigma) > 0$ and $\text{lcm}(C^{(n)}_{i_1,i}(\sigma), \ldots, C^{(n)}_{i_k,i}(\sigma)) = j$. Then the eigenangle point process of $\rho_{n,k}$ is

$$\Lambda_{n,k} = \sum_{j} C^{(n)}_{j,k} \sum_{i=0}^{j-1} \delta_{i/j}$$  \hspace{1cm} (1.8)

As for the $k = 1$ case, it is easy to see that $\Lambda_{n,k}/n^k$ converges weakly to the point mass at Haar measure on the unit circle. To obtain a scaled point process limit, we renormalize by a factor of $n^k$ and extend the process to the real line by periodicity. Then we have the point process

$$\mathcal{E}_{n,k}^\alpha := \sum_{j} C^{(n)}_{j,k} \sum_{i \in \mathbb{Z}} \delta_{i/j}$$  \hspace{1cm} (1.9)

As $n \rightarrow \infty$, it is not hard to see that the number-theoretic properties of the largest cycle lengths of $\sigma \in \mathfrak{S}_n$ play a dominant role in the behavior of $\mathcal{E}_{n,k}^\alpha$. Thus, we introduce the following infinite array $X_{mi}$ of random variables giving the limiting prime factor distribution for a random integer in $[N]$ as $N \rightarrow \infty$.

**Definition 1.3.** Let $p_m$ be the $m$th prime. For positive integers $m$ and $i$, (by Kolmogorov’s extension theorem) define $X_{mi}$ to be independent geometrically distributed random variables with parameter $1 - 1/p_m$, i.e.

$$\mathbb{P}(X_{mi} = a) = (1 - 1/p_m)(1/p_m)^a$$  \hspace{1cm} (1.10)
Let $X_{n,i_{1},\ldots,i_{k}}$ denote the $l \times k$ matrix consisting of entries $X_{m,i_{j}}$ for $1 \leq m \leq l$ and $1 \leq j \leq k$ as defined in (1.10). We allow $l$ to be infinite.

We will also need:

**Definition 1.4.** For $k > 0$, let $e$ be an $\alpha \times k$ matrix whose entries $e_{mi}$ for $1 \leq m < \infty$ and $1 \leq i \leq k$ are non-negative integers. Then define the function

$$g_{k}(e) := \prod_{m=1}^{\infty} \frac{p_{m,1} + \cdots + p_{m,k}}{\max(p_{m,1}, \ldots, p_{m,k})}$$

(1.11)

We allow $g_{k}$ to take finite matrix arguments by filling in the remaining entries with 0’s.

By the Borel-Cantelli lemma, $\prod_{m=1}^{\infty} p_{m}^{\min(X_{m,1}, X_{m,2})}$ is a.s. finite. Then $g_{k}(X_{\infty,i_{1},\ldots,i_{k}})$ is finite for all $k$ and sequences $(i_{1}, \ldots, i_{k})$ almost surely.

Let the $PD(\theta)$ process $L_{1}, L_{2}, \ldots$ be independent from all the $X_{mi}$ defined in (1.10). Then we can define

$$\mathcal{E}_{n,k}^{0} := k! \sum_{i_{1}, \ldots, i_{k}} g_{k}(X_{\infty,i_{1},\ldots,i_{k}}) \sum_{q \in \mathbb{Z}} \delta_{g_{k}(X_{\infty,i_{1},\ldots,i_{k}}/(L_{i_{1}} \ldots L_{i_{k}}))}$$

(1.12)

The weak convergence $\mathcal{E}_{n,k}^{0} \to \mathcal{E}_{\infty,k}^{0}$ will follow as a special case of Theorem 9.1.

As for the $k = 1$ case, we define for each real number $\alpha$ the shifted point process

$$\mathcal{E}_{n,k}^{\alpha} := \sum_{j} C_{j,k}^{(n)} \sum_{i \in \mathbb{Z}} \delta_{n^{\alpha}(i/j - \alpha)}$$

(1.13)

and consider $\mathcal{E}_{n,k}^{\alpha}$ for rational and irrational $\alpha$ separately. Theorem 1.2 below is the main result of the paper. Section 5 discusses factorization classes and multi-dimensional discrepancy and well-distribution, which are used in the proof. In Section 6, we prove the theorem assuming the result in Theorem 7.1. Sections 7 and 8 are devoted to the proof of Theorem 7.1 which addresses the primary technical difficulty: showing that only the largest cycles in $\sigma$ make non-negligible contributions to the point process in the limit $n \to \infty$.

**Theorem 1.2.** Let $U_{i_{1},\ldots,i_{k}}$ for $i_{1} < \ldots < i_{k}$ be i.i.d. random variables distributed uniformly on $[0,1]$ and independent of the $PD(\theta)$ process $(L_{1}, L_{2}, \ldots)$. Let the variables $X_{mi}$ from (1.10) be independent of both $U_{i_{1},\ldots,i_{k}}$ and $(L_{1}, L_{2}, \ldots)$. Then define the point process

$$\mathcal{E}_{\infty,k}^{\alpha} := k! \sum_{i_{1}, \ldots, i_{k}} g_{k}(X_{\infty,i_{1},\ldots,i_{k}}) \sum_{q \in \mathbb{Z}} \delta_{g_{k}(X_{\infty,i_{1},\ldots,i_{k}}/(L_{i_{1}} \ldots L_{i_{k}}))}$$

(1.14)

If $\alpha \in (0,1)$ is an irrational number with finite irrationality measure, then $\mathcal{E}_{n,k}^{\alpha} \to \mathcal{E}_{\infty,k}^{\alpha}$ weakly.

The analog of Theorem 1.2 for rational $\alpha$ is stated in Theorem 9.1.

**Remark 1.3.** As before, the $PD(\theta)$-distributed vector $(L_{1}, L_{2}, \ldots)$ could be replaced by a GEM$(\theta)$-distributed vector $(V_{1}, V_{2}, \ldots)$ without changing the distribution of $\mathcal{E}_{\infty,k}^{\alpha}$. The point process $\mathcal{E}_{\infty,k}^{\alpha}$ is clearly stationary. If $\theta = 1$, the mean measure is

$$E(\mathcal{E}_{\infty,k}^{\alpha}) = k! \sum_{i_{1}, \ldots, i_{k}} E[L_{i_{1}} \ldots L_{i_{k}}] \lambda = \frac{1}{k!} \lambda$$

(1.15)

by equation (3.17) where $\lambda$ denotes Lebesgue measure on $\mathbb{R}$. Note that for $k > 1$ the mean intensity is less than 1, which reflects the fact that mass escapes to infinity as $\mathcal{E}_{n,k}^{\alpha} \to \mathcal{E}_{\infty,k}^{\alpha}$.

**Remark 1.4.** Numbers that have infinite irrationality measure are called Liouville numbers and are known to have Hausdorff dimension 0. It seems unlikely that any nice limiting behavior exists for $\alpha$ an arbitrary Liouville number.
In contrast to the $\rho_{n,k}$ representations, recall that the $U(n)$ eigenvalue point processes are translation invariant and the limiting process when zooming in at any point on the unit circle is a determinantal point process characterized by the Dyson sine kernel.

Finally, we use Theorem [1.2] to compute the limiting eigenvalue gap probability for $\mathcal{E}^a_{n,k}$ as $n \to \infty$.

**Definition 1.5.** Let $\alpha$ be irrational with finite irrationality measure. If $\sigma \in \mathfrak{S}_n$ is a randomly chosen permutation from the Ewens measure with parameter $\theta$, then let $P^\theta_{n,k}(y_1, y_2)$ denote the limit as $n \to \infty$ of the gap probability that there are no eigenangles of $\rho_{n,k}(\sigma)$ in the interval $\left(\alpha + \frac{y_1}{n^k}, \alpha + \frac{y_2}{n^k}\right)$.

By Proposition [1.1] the limiting eigenvalue gap probability is the gap probability of the limiting process $\mathfrak{E}^{a,k}_\infty$. Since the $U_{i_1, \ldots, i_k}$ are i.i.d. random variables distributed uniformly on $[0, 1]$, it is easy to see that conditional on the PD($\theta$) process $(L_1, L_2, \ldots)$ and independent variables $X_{mi}$, the gap probability is

$$\lim_{r \to \infty} \prod_{1 \leq i_1 < \cdots < i_k \leq r} \left(1 - \min \left(\frac{(y_2 - y_1) \prod_{u=1}^k L_{iu}}{g_k(\mathbf{X}_\infty^{i_1, \ldots, i_k})}, 1\right)\right)$$

The limit clearly exists since this is a decreasing sequence in $r$. Taking the expectation, we get

**Corollary 1.1** (Formula for eigenvalue gap probability). Let the random variables $X_{mi}$ be independent from the PD($\theta$) process $(L_1, L_2, \ldots)$. Then the eigenvalue gap probability is given by

$$P^\theta_{k}(y_1, y_2) = \mathbb{E} \left[ \prod_{i_1 < \cdots < i_k} \left(1 - \min \left(\frac{(y_2 - y_1) \prod_{u=1}^k L_{iu}}{g_k(\mathbf{X}_\infty^{i_1, \ldots, i_k})}, 1\right)\right) \right] \quad (1.16)$$

Note that $\prod_{u=1}^k L_{iu} < \frac{1}{k^k}$ almost surely. Therefore, if $y_2 - y_1 \leq k^k$, we can remove the min and get

$$P^\theta_{k}(y_1, y_2) = \mathbb{E} \left[ \prod_{i_1 < \cdots < i_k} \left(1 + \frac{(y_1 - y_2) \prod_{u=1}^k L_{iu}}{g_k(\mathbf{X}_\infty^{i_1, \ldots, i_k})}\right) \right] \quad (1.17)$$

Expanding out the product, $P^\theta_{k}(y_1, y_2)$ has a power series representation in $y_1 - y_2$. For $\theta = 1$, nicer formulas can be obtained and we will provide a combinatorial procedure for computing the series coefficients in Section [13].

In the next section we begin by discussing joint cycle length statistics, which play an important role in understanding the eigenvalue point processes. In essence, we want to sum over the events $(C_1^{(n)}, \ldots, C_n^{(n)}) = (c_1, \ldots, c_n)$ such that $\sum_{i=1}^n ic_i = n$ and determine when the $\sigma_k$-cycle of length $\text{lcm}(C_{i_1}, \ldots, C_{i_k})$ contains an eigenangle in a given interval. Thus, the proofs of Theorems [1.1] and [1.2] involve combining asymptotic point probability estimates of cycle length distributions with probabilistic number theory and Diophantine approximation methods.

2. Joint Cycle Length Statistics

We borrow much of the notation in this section from Arratia, Barbour, and Tavare’s book [2]. This monograph contains essentially all the material we need about joint cycle length distributions of permutations drawn from the Ewens distribution.

The Conditioning Relation [2] (4.1) states that the distribution of cycle lengths $C_j^{(n)}$ behaves the same as independent Poisson random variables conditioned on the necessary restriction that
the size of the permutation is \( n \). This relation in fact holds for a broad class of classical combinatorial structures such as mappings, partitions, and trees. See [2] for details. To state the Conditioning Relation formally, we define the following random variables.

**Definition 2.1.** Let \( Z_j \) be independent Poisson variables with mean \( \theta/j \). Then we define

\[
T_{bn} := \sum_{j=b+1}^{n} jZ_j
\]  

(2.1)

For any subset \( B \subset [n] \), we define

\[
T_B := \sum_{j \in B} jZ_j
\]  

(2.2)

**Proposition 2.1** (Conditioning Relation). For any vector \( c = \langle c_1, ..., c_n \rangle \in \mathbb{Z}_+^n \), we have

\[
P\{C_1^{(n)} = c_1, ..., C_n^{(n)} = c_n\} = P\{Z_1 = c_1, ..., Z_n = c_n \mid T_{0n} = n\}
\]  

(2.3)

Using the Conditioning Relation, one can compute

**Proposition 2.2** (Ewens Sampling Formula). Under the Ewens measure with parameter \( \theta \),

\[
P\{(C_1^{(n)}, ..., C_n^{(n)}) = (b_1, ..., b_n)\} = \frac{n!}{\theta(\theta + 1)(\theta + n - 1)} \prod_{j=1}^{n} \frac{\theta}{j} \frac{b_j}{b_j!} \frac{1+...+n}{\sum_{i=1}^{n} ib_i = \sum_{i=1}^{n} i b_i = n}
\]  

(2.4)

By [2] Thm. 4.6, \( n^{-1}T_{0n} \overset{d}{\rightarrow} X_\theta \) with Laplace transform given by

\[
\mathbb{E}[e^{-sX_\theta}] = \exp\left(-\int_{0}^{1} \frac{1 - e^{-sx}}{x} dx\right)
\]  

(2.5)

and \( \mathbb{E}[X_\theta] = \theta \).

Equation (2.5) can be used to find the continuous density function \( p_\theta \) of \( X_\theta \) for \( x > 0 \) (see [2] Lemma 4.7). We won’t need the exact form in the sequel. However, for \( 0 \leq x \leq 1 \), we have the simple expression (see [2] Cor. 4.8)

\[
p_\theta(x) = \frac{e^{-\gamma\theta}x^{\theta-1}}{\Gamma(\theta)}
\]  

(2.6)

Now we can state the Poisson-Dirichlet density (c.f. [2] (5.40)). Here, \( \Gamma(z) \) is the Gamma function and \( \gamma \) is the Euler-Mascheroni constant.

**Proposition 2.3.** Let \( (L_1, L_2, ...) \) follow the Poisson-Dirichlet distribution \( PD(\theta) \). Then the density of the finite-dimensional distributions \( (L_1, ..., L_r) \) are given by

\[
f_\theta^{(r)}(x_1, ..., x_r) = \frac{e^{\gamma\theta} \theta^{\theta-1}}{x_1...x_r} p_\theta\left(\frac{1-x_1-...-x_r}{x_r}\right)
\]  

(2.7)

and supported in the region given by \( 0 < x_r < ... < x_1 < 1 \) and \( x_1 + ... + x_r < 1 \).

The following uniform local limit theorem for the large cycle distributions will play a fundamental role in our proofs. We follow the proof in the local limit theorem [2] Thm. 5.10. There, pointwise convergence for each choice of \( x_i \) is established. For our purposes, this needs to be strengthened to uniform convergence.

**Lemma 2.1.** Let \( 0 < \lambda < 1 \) and \( r \geq 1 \). For \( 1 \leq i \leq r \), let \( \Delta \) be the set of integer tuples \( \langle m_1, ..., m_r \rangle \) such that \( \lambda n < m_r < m_{r-1} < ... < m_1 \leq n \) and \( 0 < m = m_1 + ... + m_r < (1-\lambda)n \).

Set \( x_i = \frac{m_i}{n} \). Then for each \( \varepsilon > 0 \) and for sufficiently large \( n \),

\[
\left| n^{-\varepsilon} p\{L_i^{(n)} = m_i, 1 \leq i \leq r\} - f^{(r)}(x_1, ..., x_r) \right| < \varepsilon
\]
for all \((m_1, ..., m_r) \in \Delta\).

Proof. Let \(A_n(C^{(n)}) = A_n(C^{(n)}; m_1, ..., m_r)\) denote the event such that \(C^{(n)}_{m_i} = 1\) for \(1 \leq j \leq r-1\) and \(C^{(n)}_i = 0\) for all \(m_r + 1 \leq i \leq n\). Then

\[
P\{L_i^{(n)} = m_i, 1 \leq i \leq r\} = P\{A_n(C^{(n)}), C^{(n)}_{m_r} = 1\} + \sum_{l \geq 2} P\{A_n(C^{(n)}), C^{(n)}_{m_r} = l\}
\]

(2.8)

As shown in [2, Thm. 5.10], the second term is \(o(m_r^{-r}) = o(n^{-r})\). By the Conditioning Relation (2.3), the first term is

\[
P\{A_n(Z), Z_{m_r} = 1 \mid T_{0n} = n\} = P\{A_n(Z)\} P\{Z_{m_r} = 1\} \frac{P\{T_{0,m_r-1} = n-m\}}{P\{T_{0n} = n\}}
\]

(2.9)

which reduces to

\[
\theta e^{-\theta(h(n+1)-h(m_r))} \frac{P\{T_{0,m_r-1} = n-m\}}{P\{T_{0n} = n\}}
\]

(2.10)

where \(h(n) = 1 + \frac{1}{2} + \ldots + \frac{1}{n-1}\).

The size biasing equation [2, (4.8)] states that for any \(B \subset [n]\), we have

\[
m P\{T_B = m\} = \theta \sum_{l \in B} P\{T_B = m - l\}
\]

(2.11)

Then by (2.11),

\[
(n-m) P\{T_{0,m_r-1} = n-m\} = \theta P\left\{\frac{n-m}{m_r} - 1 \leq \frac{T_{0,m_r-1}}{m_r} < \frac{n-m}{m_r}\right\}
\]

(2.12)

and we see that the local point probabilities can be expressed in terms of interval probabilities.

If distribution functions \(F_n\) converge weakly to a continuous distribution \(F\), then the convergence is in fact uniform on the real line. This applies to the convergence \(\frac{T_{0n}}{n} \xrightarrow{d} X_\theta\). By [2, (4.23)], \(xp(n) = \theta \int_{x-1}^x p(n) \, du\) Thus, for any \(\varepsilon > 0\) and for sufficiently large \(n\),

\[
\left|\frac{m_r}{n-m} \theta P\left\{\frac{n-m}{m_r} - 1 \leq \frac{T_{0,m_r-1}}{m_r} < \frac{n-m}{m_r}\right\} - \theta e^{-\gamma} \frac{n-m}{m_r}\right| < \varepsilon
\]

The uniform bound follows from the fact that \(\lim_{n \to \infty} n P\{T_{0n} = n\} = P(1) = \frac{e^{-\gamma}}{\Gamma(\gamma)}\).

\[\square\]

3. Eigenvalue process for random permutation matrices at angles of finite irrationality measure

In this section, we prove Theorem 1.1. First, we introduce some notation for the discrepancy of a sequence.

**Definition 3.1.** Let \(J\) be the set of \(d\)-dimensional boxes of the form \(\prod_{i=1}^{d} [a_i, b_i]\) where \(0 \leq a_i < b_i \leq 1\) and let \(\mathcal{P}_n\) be a multiset or sequence of \(n\) elements of the \(d\)-dimensional torus \(T^d\). The multidimensional discrepancy is given by

\[
D(\mathcal{P}_n) := \sup_{B \in J} \left| \frac{A(B; n)}{n} - \prod_{i=1}^{d} (b_i - a_i) \right|
\]

where \(A(B; n)\) counts the number of elements of \(\mathcal{P}_n\) in the box \(B\).
Define the intervals $$r$$ for some error function $$\varepsilon$$.

Then by Lemma 2.1, for $$\varepsilon > 0$$,

$$D_{i,n}(x) = O(n^{-\frac{1}{15} + \varepsilon})$$ \hspace{1cm} (3.1)

**Definition 3.3.** For any real number $$\alpha$$, let $$\psi_\alpha(j)$$ be the minimum positive element of the set $$\{q - j\alpha : q \in \mathbb{Z}\}$$.

The following lemma shows that in the limit, the eigenangles associated to the largest cycles are randomly situated in the window centered at $$\alpha$$.

**Lemma 3.1.** Let $$\alpha \in \mathbb{R}$$ be irrational and $$(L_1, L_2, \ldots)$$ be a PD($$\theta$$) process. Then for any positive integer $$r$$ and i.i.d. random variables $$U_i$$ chosen independently from the vector $$(L_1, L_2, \ldots)$$ and distributed uniformly on $$[0, 1]$$, we have the distributional convergence

$$(L_1^{(n)}/n, \ldots, L_r^{(n)}/n, \psi_\alpha(L_1^{(n)}), \ldots, \psi_\alpha(L_r^{(n)})) \xrightarrow{d} (L_1, \ldots, L_r, U_1, \ldots, U_r)$$

**Proof.** Define the intervals

$$I_{m,b} = ((m - 1)n/b, mn/b] \cap \mathbb{Z}$$ \hspace{1cm} (3.2)

This defines a partition $$[n] = \bigcup_{m=1}^b I_{m,b}$$. For each $$\lambda > 0$$, $$f_\theta^{(r)}(x_1, \ldots, x_r)$$ is uniformly continuous on the set of tuples

$$\{(x_1, \ldots, x_r) : \lambda < x_r \leq \ldots \leq x_1 < 1) \land (x_1 + \ldots + x_r < 1 - \lambda)\}$$.

Then by Lemma 2.1 for $$\lambda b \leq k_r \leq \ldots \leq k_1$$ and $$k_1 + \ldots + k_r \leq (1 - \lambda)b$$ and $$m_i \in I_{k_i,b}$$ distinct, we have the bound

$$\left| f_\theta^{(r)} \left( \frac{k_1}{b}, \ldots, \frac{k_r}{b} \right) - n^r \mathbb{P}\{L_i^{(n)} = m_i, 1 \leq i \leq r\} \right| < \varepsilon(n,b)$$ \hspace{1cm} (3.3)

for some error function $$\varepsilon(n,b) \to 0$$ as $$n,b \to \infty$$. Let $$\varepsilon^*(n) = \mathbb{P}\left( \bigcup_{i=1}^r \{L_i^{(n)} > \lambda n\} \right)$$, i.e. the probability that there is a duplicate cycle length among the $$r$$ largest cycles. Clearly, $$\varepsilon^*(n) \to 0$$ as $$n \to \infty$$.

Since $$f_\theta^{(r)}(x_1, \ldots, x_r)$$ integrates to 1, for every $$\lambda > 0$$ we have

$$\lim_{n \to \infty} \mathbb{P}\{L_r^{(n)} < \lambda n\} + \mathbb{P}\{L_1^{(n)} + \ldots + L_r^{(n)} > (1 - \lambda)n\} < \varepsilon(\lambda)$$

for some error function $$\varepsilon(\lambda)$$ tending to 0 as $$\lambda \to 0$$. Finally, note that $$\psi_\alpha(j) \leq x$$ iff $$\{j\alpha\} \geq 1 - x$$.

Then for $$c_i, d_i \in [0, 1]$$,

$$\mathbb{P}\left( \bigcap_{i=1}^r \{L_i^{(n)} < c_i\} \cap \{\psi_\alpha(L_i^{(n)}) < d_i\} \right) \leq \sum_{1 \leq m_i < \ldots < m_i \leq n} \mathbb{P}\left( \bigcap_{i=1}^r \{L_i^{(n)} = m_i\} \cap \{\psi_\alpha(m_i) < d_i\} \right) + \varepsilon^*(n)$$

$$\leq \sum_{\lambda b \leq k_1 \leq \ldots \leq k_1 \leq b} \left( f_\theta^{(r)} \left( \frac{k_1}{b}, \ldots, \frac{k_r}{b} \right) + \varepsilon(n,b) \right) \frac{1}{b^r} \prod_{i=1}^r \left( d_i + D_{\phi_{b\lambda}}(\{j\alpha\}) \right)$$

$$+ \mathbb{P}\{L_r^{(n)} < \lambda n\} + \mathbb{P}\{L_1^{(n)} + \ldots + L_r^{(n)} > (1 - \lambda)n\} + \varepsilon^*(n)$$
Taking $b = o(n)$ so that $n/b \to \infty$,

$$
\limsup_{n \to \infty} \mathbb{P}\left( \bigcap_{i=1}^{r} \left\{ L_i^{(n)} / n < c_i \right\} \cap \{ \psi_\alpha(L_i^{(n)}) < d_i \} \right) \leq \int f^{(r)}(x_1, \ldots, x_r) dx_i \prod_{i=1}^{r} d_i + \varepsilon(\lambda)
$$

Similarly,

$$
\liminf_{n \to \infty} \mathbb{P}\left( \bigcap_{i=1}^{r} \left\{ L_i^{(n)} / n < c_i \right\} \cap \{ \psi_\alpha(L_i^{(n)}) < d_i \} \right) \geq \int f^{(r)}(x_1, \ldots, x_r) dx_i \prod_{i=1}^{r} d_i
$$

Taking $\lambda \to 0$ completes the proof. \hfill \Box

Now we are ready to prove the eigenvalue point process convergence for the $k = 1$ case of random permutation matrices at irrational $\alpha$ with finite irrationality measure.

**Proof of Theorem 1.1.** Let $r > 1$ be a positive integer. To isolate the contribution from the $r$ largest cycles, we define

$$
\xi^r_n = \sum_{i=1}^{r} \sum_{q \in \mathbb{Z}} \delta_{n/q(L_i^{(n)} - \alpha)}
$$

and we let $\eta^r_n = \mathcal{E}^\alpha_n - \xi^r_n$ denote the remaining measure. Similarly, define

$$
\xi^r_\infty = \sum_{i=1}^{r} \sum_{q \in \mathbb{Z}} \delta_{(U_i + q)/L_i}
$$

and set $\eta^r_\infty = \mathcal{E}^*_{\infty} - \xi^r_\infty$.

Let $f$ be a continuous function with support contained in some interval $(-T, T)$. Then

$$
\int f d\xi^r_n = \sum_{i=1}^{r} \sum_{q \in \mathbb{Z}} \int f\left( \frac{q - L_i^{(n)}}{L_i^{(n)}/n} \right) = \sum_{i=1}^{r} \sum_{q \in [-T+\epsilon, T+\epsilon]} \int f\left( \frac{q + \psi_\alpha(L_i^{(n)})}{L_i^{(n)}/n} \right)
$$

and

$$
\int f d\xi^r_\infty = \sum_{i=1}^{r} \sum_{q \in [-T+\epsilon, T+\epsilon]} \int f\left( \frac{U_i + q}{L_i} \right)
$$

By Lemma 3.1 and the continuous mapping theorem,

$$
\int f d\xi^r_n \overset{d}{\to} \int f d\xi^r_\infty
$$

Now we turn to the remainder measures $\eta^r_n$ and $\eta^r_\infty$. Since $\mathbb{E}[\sum_i L_i] = \mathbb{E}[1] = 1$, by the Borel-Cantelli lemma $\int f d\eta^r_\infty \overset{a.s.}{\to} 0$. Note that this just verifies that $\mathcal{E}^*_{\infty}$ is in fact a point process.

For $\eta^r_n$, we need to show that

$$
\lim_{r \to \infty} \limsup_{n \to \infty} \mathbb{P}\left\{ \int f |d\eta^r_n| > 0 \right\} = 0
$$

Note that $L_i^{(n)} < n/r$ for $i > r$ with probability 1. Thus, we have the inequality of measures

$$
\eta^r_n < \sum_{j=1}^{[n/r]} C_j^{(n)} \sum_{i \in \mathbb{Z}} \delta_{n(i/j - \alpha)}
$$

Define the set

$$
J(n) = \{ j \in [n] : (\exists q \in \mathbb{Z}) [\alpha - T/n < q/j < \alpha + T/n] \}
$$

(3.7)
Specializing Watterson’s factorial moment formula \[30\] gives

\[
E[C_j^{(n)}] = \frac{\theta^{j-1}}{j!} \prod_{i=1}^{j-1} \frac{n - i}{\theta + n - i - 1} \leq \frac{\theta^n}{j n - j + \theta} \tag{3.8}
\]

Then

\[
\mathbb{P}\left( \int |f| d\mathbb{P}_n > 0 \right) \leq \mathbb{P}\left( \bigcup_{j \in J^{(n)}[r]} \{C_j^{(n)} > 0\} \right) \leq \sum_{j \in J^{(n)}[r]} E(C_j^{(n)}) \leq \sum_{j \in J^{(n)}[r]} \frac{2\theta}{j} \tag{3.9}
\]

Looking at each interval \( I_{m,b} \) (defined in (3.2)) separately,

\[
|J \cap I_{m,b}| \leq \frac{n}{b} \left( \min \left( \frac{2mT}{b}, 1 \right) + D_{\frac{(m-1)n}{x}} \left( \{j\alpha\} \right) \right) \tag{3.10}
\]

Recall \( \alpha \) has some finite irrationality measure \( \mu(\alpha) \). Now set \( b = n^{1-1/(\mu(\alpha)+1)} \). Then the first interval \( I_{1,b} \) contains the integers \( 0 < j \leq n^{1/(\mu(\alpha)+1)} \). Setting \( \varepsilon = 1/2 \) (say), we have

\[
\left| \alpha - \frac{p}{j} \right| < \frac{C}{n^{1-1/(2\mu(\alpha)+2)}}
\]

for some absolute constant \( C \) and all \( j \in I_{1,b} \). Since the length of the interval \((\alpha - T/n, \alpha + T/n)\) is \( O(1/n) \), the intersection \( J \cap I_{1,b} \) is empty for sufficiently large \( n \). Then

\[
\sum_{j \in J^{(n)}[r]} \frac{1}{j} \leq \sum_{m=2}^{[b/r]} \frac{b}{(m-1)n} \left( \min \left( \frac{2mT}{b}, 1 \right) + D_{\frac{(m-1)n}{x}} \left( \{j\alpha\} \right) \right) \tag{3.11}
\]

By Proposition 3.1 if \( 1/r < 1/(2T) \) we have

\[
\limsup_{n \to \infty} \sum_{j \in J^{(n)}[r]} \frac{1}{j} \leq 4T/r
\]

and (3.10) follows. Recall we want to show

\[
\int f d\mathbb{P}^*_n \overset{d}{\rightarrow} \int f d\mathbb{P}^*_\infty.
\]

For ease of notation, set \( X_n = \int f d\mathbb{P}^*_n \), \( X_\infty = \int f d\mathbb{P}^*_\infty \), and \( X_r^* = \int f d\mathbb{P}^*_r \), \( X^*_\infty = \int f d\mathbb{P}^*_{\infty} \). Finally, let \( \mathbb{E}_n^* = \int f d\mathbb{P}^*_n \). Then it is sufficient to show that \( \lim_{n \to \infty} \mathbb{E}[g(X_n)] = \mathbb{E}[g(X_\infty)] \) for all continuous and compactly supported \( g \).

Since \( X^*_n \overset{d}{\rightarrow} X^*_\infty \) and \( X^*_\infty \overset{d}{\rightarrow} X^*_r \), we already have \( \mathbb{E}[g(X^*_n)] \to \mathbb{E}[g(X^*_\infty)] \) and \( \mathbb{E}[g(X^*_\infty)] \to \mathbb{E}[g(X^*_r)] \) and therefore \( \lim_{n,r \to \infty} \mathbb{E}[g(X^*_n)] = \mathbb{E}[g(X^*_\infty)] \).

By (3.6),

\[
\lim_{n \to \infty} \mathbb{E}[g(X_n)] = \lim_{n,r \to \infty} \mathbb{E}[g(X^*_n)] = \mathbb{E}[g(X^*_\infty)] \tag{3.12}
\]

For \( q \geq 3 \), the \( q \)-point correlation function of the limiting point process \( \mathbb{P}^*_\infty \) is not absolutely continuous with respect to Lebesgue measure since \( \mathbb{P}^*_\infty \) almost surely has 3 points \( x, y, z \in \mathbb{R} \) such that \( z - y = y - x \). However, the 2-point correlation function is absolutely continuous. Najnudel and Nikeghbali [13 Prop. 5.6] show that

**Proposition 3.2.** The 2-point correlation function for the point process \( \mathbb{P}^*_\infty \) is given by \( \rho_2(x, y) = \phi_\theta(x - y) \) where

\[
\phi_\theta(x) = \frac{\theta}{\theta + 1} + \frac{\theta}{x^2} \sum_{a \in \mathbb{N}, a \leq |x|} a \left( 1 - \frac{a}{|x|} \right)^{\theta - 1} \tag{3.13}
\]
Proposition 3.3. For all \( x \in \mathbb{R} \), set

\[
H(x) := \mathbb{1}(x > 0)x^{\theta-1}P_1^\theta(0,x)
\]

Then \( H \) is integrable and satisfies

\[
xH(x) = \theta \int_0^1 (1-y)H(x-y)dy.
\] (3.14)

If \( y_2 - y_1 \leq 1 \), we can use (1.17) to obtain a simple power series expression for \( P_1^\theta(y_1, y_2) \) in the special case \( \theta = 1 \). We have

\[
P_1^1(y_1, y_2) = \mathbb{E} \left[ \prod_{i=1}^\infty \left( 1 + (y_1 - y_2)L_i \right) \right] = \sum_{m=0}^\infty (y_1 - y_2)^m \sum_{i_1 < \ldots < i_m} \mathbb{E}[L_{i_1} \ldots L_{i_m}]
\]

Griffiths [12] computed the mixed moments of the random variables \( L_i \):

\[
\mathbb{E}[L_{i_1} \ldots L_{i_j}] = \frac{\theta^\gamma \Gamma(\theta)}{\Gamma(\theta+j)} \int_{y_1 > \ldots > y_r > 0} y_1^{i_1-1} \ldots y_r^{i_r-1} \exp \left( -\sum_{i=1}^r y_i - \theta E_1(y_r) \right) dy_1 \ldots dy_r
\] (3.15)

where \( j = j_1 + \ldots + j_r \) and \( E_1(s) = \int_1^\infty e^{-sx} dx \). Taking \( \theta = 1 \) in (3.13) and using symmetry in the domain,

\[
\mathbb{E}[L_{i_1} \ldots L_{i_m}] = \frac{1}{m!} \int_0^\infty \int_{x_m}^\infty \ldots \int_{x_2}^\infty e^{-(E_1(x_1) + \ldots + E_1(x_m))} dx_1 \ldots dx_m
\]

We have

\[
\sum_{i_1, \ldots, i_m = 0}^\infty \frac{E_1(x_1)^{i_1} \ldots E_1(x_m)^{i_m}}{i_1! \ldots i_m!} = e^{E_1(x_1) + \ldots + E_1(x_m)}
\]

Thus,

\[
\sum_{i_1 < \ldots < i_m} \mathbb{E}[L_{i_1} \ldots L_{i_m}] = \frac{1}{m!} \int_0^\infty \int_{x_m}^\infty \ldots \int_{x_2}^\infty e^{-x_1} \ldots e^{-x_m} dx_1 \ldots dx_m = \frac{1}{(m!)^2}
\] (3.17)

and therefore

\[
P_1^1(y_1, y_2) = \sum_{m=0}^\infty \frac{(y_1 - y_2)^m}{(m!)^2}
\] (3.18)

for \( y_2 - y_1 \leq 1 \).

Remark 3.1. For \( y_2 - y_1 \leq 1 \), the gap probability \( P_1^1(y_1, y_2) = J_0(2\sqrt{y_2 - y_1}) \) where for real \( \beta \),

\[
J_\beta(x) = \sum_{m=0}^\infty \frac{(-1)^m}{m!(m + \beta + 1)} \left( \frac{x}{2} \right)^{2m+\beta}
\]

are the Bessel functions of the first kind.

Remark 3.2. One can check that \( \sum_{m=0}^\infty \frac{(-x)^m}{(m!)^2} \) does indeed satisfy the integral equation (3.14) for \( 0 < x < 1 \).
4. Eigenvalue process for random permutation matrices at rational angles

Equation (1.4) states the limiting process $E_{\infty}^n$ when zooming in at the origin. In general, it is not difficult to obtain the limiting process $\lim_{n \to \infty} E_{\alpha}^n$ at any rational angle $\alpha = s/t$. As when zooming in at 0, $\lim_{n \to \infty} E_{\alpha}^n$ will have an infinite dirac mass at 0 since there will be an eigenangle at $\alpha$ corresponding to every $j$-cycle with $j \in t\mathbb{Z}$ and the number of such cycles is approximately $(\log n)/t$ for large $n$. This follows from the following result by Arratia, Barbour, Tavare [1] (for a method involving generating functions see [25]).

**Proposition 4.1** (Arratia, Barbour, Tavare). Let $Z_i$ be independent Poisson distributed random variables with parameter $\theta/i$. Then as $n \to \infty$ and for $b = o(n)$,

$$(C_{1}^{(n)}, \ldots, C_{b}^{(n)}) \xrightarrow{TV} (Z_1, \ldots, Z_b)$$

where the convergence is in total variation.

Let $f$ be continuous with support contained in the interval $(-T, T)$ and let $J(n)$ be defined as in (3.7). Then we have the set equality

$$J \cap [n/(tT)] = t\mathbb{Z} \cap [n/(tT)], \quad (4.1)$$

Also, if $j \in t\mathbb{Z}$ and $j < n/(tT)$, then the interval $(\alpha - T/n, \alpha + T/n)$ contains exactly one eigenangle corresponding to the $j$-cycle (namely $\alpha$). To isolate the mass at 0, we define $\xi_{n}^{tT}$ and $\eta_{n}^{tT}$ as in (3.3). Then

$$\int f \xi_{n}^{tT} = \sum_{1 \leq i \leq tT, q \in \mathbb{Z}} f(n(q/L_{i}^{(n)} - \alpha)) \quad (4.2)$$

and if $f(0) \geq 0$, since $L_{r}^{(n)} < n/(tT)$ for $r > tT$, we have

$$f(0) \left( \sum_{1 \leq i \leq n/(tT), i \neq j} C_{j}^{(n)} - tT \right) \leq \int f \eta_{n}^{tT} \leq f(0) \sum_{1 \leq i \leq n/(tT)} C_{i}^{(n)} \quad (4.3)$$

Let $U_1, U_2, \ldots$ be i.i.d. random variables distributed uniformly on the set $\{1/t, 2/t, \ldots, (t-1)/t, 1\}$ and define the process

$$E_{\infty}^{t} = \sum_{i=1}^{\infty} \sum_{q \in \mathbb{Z}} \delta_{(U_{i} + q)/L_{i}} \quad (4.4)$$

Define $\xi_{\infty}^{tT}$ and $\eta_{\infty}^{tT}$ as in equation (3.5) with $E_{\infty}^{t}$ in place of $E_{\infty}^{n}$ and with our new definition of the variables $U_i$. Then

$$\int f \xi_{\infty}^{tT} = \sum_{1 \leq i \leq tT, q \in \mathbb{Z}} f((U_{i} + q)/L_{i}) \quad (4.5)$$

and since $L_{r} < 1/(tT)$ for $r > tT$, we have

$$\int f \eta_{\infty}^{tT} = f(0) \sum_{i > tT} 1(U_{i} = 1) \quad (4.6)$$

which is infinite almost surely if $f(0) > 0$.

Let $\mathbb{R} = \mathbb{R} \cup \{\pm \infty\}$ denote the set of extended real numbers. Then if $f(0) > 0$, $\int f \xi_{\infty}^{t}$ converges weakly to the $\mathbb{R}$-valued random variable $\int f \xi_{\infty}^{tT}$, which is infinite almost surely.

Now we can assume $f(0) = 0$. By partitioning $[n]$ into intervals of size $t$, it is easy to see that Lemma 3.1 holds for $\alpha = s/t$ and $U_i$ as defined in this section instead of uniform on $[0, 1]$. This proves
**Theorem 4.1.** Let $\alpha = \frac{n}{l}$ be a rational number in reduced form. Then $\mathcal{E}_n^\alpha \to \mathcal{E}_t^\infty$ weakly in the sense that we have the weak convergence $\int f \, d\mathcal{E}_n^\alpha \xrightarrow{d} \int f \, d\mathcal{E}_t^\infty$ of $\mathbb{R}$-valued random variables for all continuous, compactly supported functions $f$.

5. Factorization classes and well-distribution

To set the stage for the multidimensional case, we now define factorization classes and establish well-distribution for polynomial sequences of a given factorization class in the torus. The notation $|\cdot|$ is shorthand for “exactly divides,” i.e. a prime power $p^r \mid \alpha$ if $p^r \mid \alpha$ and $p^{r+1} \not\mid \alpha$.

**Definition 5.1.** Let $p_1 < \ldots < p_l$ be the first $l$ primes and let $(e_1, \ldots, e_l)$ be an $l$-tuple of non-negative integers. The factorization class $\mathcal{P}_{e_1, \ldots, e_l}$ is defined to be the set

$\{m \in \mathbb{N} : p_j^{e_j} \mid m, 1 \leq j \leq l\}$. For a matrix $A = [a_1, \ldots, a_r] \in \mathbb{N}^{l \times r}$, let $\mathcal{P}_A = \prod_{i=1}^r \mathcal{P}_{a_i}$. Then $\mathcal{P}_A$ will also be called a factorization class.

The following simple lemma gives the asymptotic density of each factorization class.

**Lemma 5.1.** Let $I_n$ be an interval of $n$ consecutive positive integers and let $(e_1, \ldots, e_l)$ be a vector of non-negative integers. Then as $n \to \infty$, the fraction of integers $a \in I_n$ in the factorization class $\mathcal{P}_{e_1, \ldots, e_l}$ approaches $\prod_{m=1}^l (1 - 1/p_m)(1/p_m)^{e_m}$.

**Proof.** Let $N = \prod_{m=1}^l p_m^{e_m+1}$ and apply the Chinese Remainder Theorem to $\mathbb{Z}/N\mathbb{Z}$. Then consider multiples $aN$ as $a \to \infty$. \hfill \square

We will need the following notion of equidistribution and well-distribution for multi-indexed multidimensional sequences.

**Definition 5.2.** Let $(s_{j_1, \ldots, j_r})$ for $(j_1, \ldots, j_r) \in \mathbb{N}^r$ be a multi-indexed sequence of elements in the $d$-dimensional torus $\mathbb{T}^d$. Let $\mathcal{P}_N^{t_1, \ldots, t_r} = \{s_{j_1, \ldots, j_r} : t_i \leq j_i \leq t_i + N\}$ as a multiset. Then we say that $(s_{j_1, \ldots, j_r})$ is equidistributed if $\lim_{N \to \infty} D(\mathcal{P}_N^{t_1, \ldots, t_r}) = 0$. The sequence $(s_{j_1, \ldots, j_r})$ is well-distributed if $\lim_{N \to \infty} D(\mathcal{P}_N^{t_1, \ldots, t_r}) = 0$ uniformly over the tuples $(t_1, \ldots, t_r) \in \mathbb{N}^r$.

Before stating the main result Lemma 5.2 of this section, we define the following sets.

**Definition 5.3.** Let the intervals $I_{m,b}$ be as in (3.2), let $A \in \mathbb{N}^{l \times r}$ be a matrix of non-negative integers, and let $\alpha = (\alpha_{i_1, \ldots, i_r})_{1 \leq i_1 < \ldots < i_k \leq r}$ be a vector of irrational numbers. Define the multiset

$\mathcal{P}_{(k_1, \ldots, k_r)}^{\beta_{(k_1, \ldots, k_r)}} = \left\{(m_{i_1} \ldots m_{i_k} \alpha_{i_1, \ldots, i_k})_{1 \leq i_1 < \ldots < i_k \leq r} : (m_{i_1}, \ldots, m_{i_r}) \in \mathcal{P}_A \cap \prod_{i=1}^r I_{k_i, b}\right\}$ (5.1)

of tuples in $\mathbb{T}^l$.

**Lemma 5.2.** The discrepancies $D(\mathcal{P}_{(k_1, \ldots, k_r)}^{\beta_{(k_1, \ldots, k_r)}}) \to 0$ as $n/b \to \infty$ uniformly over the tuples $(k_1, \ldots, k_r)$.

To prove this, we use a result about well-distribution in the torus stated in Lemma 5.4 below. We briefly review the essential ingredients needed, which are given as exercises in Tao [27] pp. 12-13. To simplify notation involving exponential sums, we set $e(x) := e^{2\pi ix}$. 


Proposition 5.1 (Multi-dimensional Weyl Criterion for well-distribution). The sequence 
\( (s_{j_1},...,j_m) \) is well-distributed in \( \mathbb{T}^d \) if and only if for each non-zero \( h \in \mathbb{Z}^d \),
\[
\lim_{N \to \infty} \frac{1}{N^m} \sum_{(j_1,...,j_m) \in B} e(h \cdot s_{j_1},...,j_m) = 0
\]
uniformly over all cubes \( B \) of side length \( N \) with vertices in \( \mathbb{N}^m \).

To apply the Weyl Criterion, we need a uniform version of the multidimensional van der 
Corput Lemma.

Lemma 5.3 (Multidimensional van der Corput Lemma). Fix \( k \in \mathbb{Z}^d \). Let \( x: \mathbb{N}^m \to \mathbb{T}^d \) be such 
that for every \( h \) outside of a hyperplane of \( \mathbb{R}^m \), the difference sequence \( \partial_h x: n \to x(n+h) - x(n) \)
satisfies
\[
\frac{1}{N^m} \left| \sum_{n \in B} e(k \cdot \partial_h x(n)) \right| \to 0 \text{ uniformly over all cubes } B \text{ of side length } N \text{ with vertices}
\]
lying in \( \mathbb{N}^m \). Then \( \frac{1}{N^m} \left| \sum_{n \in B} e(k \cdot x(n)) \right| \to 0 \text{ uniformly over all } B \text{ as well.}

Proof. This is given as exercise 1.1.12 (without the condition on uniformity) in [27]. The argument remains unaffected by adding this condition in both the hypothesis and conclusion. \( \square \)

Using Lemma 5.3 we get

Lemma 5.4. Let 
\[
P(n_1,...,n_m) = \sum_{s_1,...,s_m \geq 0, \ s_1 + ... + s_m \leq s} \alpha_{s_1,...,s_m} n_1^{s_1}...n_m^{s_m}
\]
be a polynomial map from \( \mathbb{N}^m \) to \( \mathbb{T}^d \) of degree \( s \), where \( \alpha_{s_1,...,s_m} \in \mathbb{T}^d \) are coefficients. Suppose 
there does not exist a non-zero \( k \in \mathbb{Z}^d \) such that \( k \cdot \alpha_{s_1,...,s_m} = 0 \) for all \( (s_1,...,s_m) \neq 0 \). Then 
\( P \) is well-distributed in the torus.

Proof. This is given as exercise 1.1.13 in [27], but in terms of equidistribution rather than well-
distribution. The idea of the proof is to use Weyl’s criterion for well distribution and proceed by 
induction on the degree of the polynomial. The inductive step uses Lemma 5.3. The reason that 
we can upgrade from equidistribution to well-distribution is that in the base case, the sequence 
\( \{na\} \) is not only equidistributed, but well-distributed (because of the summation formula for 
geometric series). \( \square \)

Proof of Lemma 5.2. For \( 1 \leq i \leq r \), set \( N_i = \prod_{m=1}^{l} p_m^{A_m + 1} \) and let \( s_i \) be an element of \( \mathcal{P}_a \cap [N_i] \).
Define the sequences \( s^i(t) = s_i + tN_i \) where \( t \) runs over the non-negative integers. Note that 
\( s^i(t) \in \mathcal{P}_a \) for all \( t \). Thus, we can partition each factorization class \( \mathcal{P}_a \) into congruence classes modulo \( N_i \). Then it suffices to show that the multidimensional polynomial sequence
\[
P(t_1,...,t_r) := \left( \alpha_{i_1,...,i_k} \prod_{j=1}^{k} s^{i_j}(t_{i_j}) \right)_{1 \leq i_1 < ... < i_k \leq r}
\]
is well-distributed. This follows from Lemma 5.4 since the polynomials \( \prod_{j=1}^{k} s^{i_j}(t_{i_j}) \) are linearly
independent. \( \square \)
6. Eigenvalue process for $\rho_{n,k}$ at angles of finite irrationality measure

In this section, we prove Theorem 1.2 assuming the main technical result Theorem 7.1; we will prove in the following sections.

Remark 6.1. Let $Y_{n,k}$ be the scaled eigenvalue fluctuation statistics of $\rho_{n,k}$ studied in [29]. In that case, the non-negligible contributions as $n \to \infty$ come from $\sigma_k$-cycles containing tuples $(t_1, \ldots, t_k)$ such that $t_1, \ldots, t_k$ are all in the same $\sigma$-cycle. For the eigenvalue point process on the other hand, we will show that the non-negligible contributions to $C_{n,k}$ all come from $\sigma_k$-cycles containing tuples $(t_1, \ldots, t_k)$ such that $t_1, \ldots, t_k$ are all in different $\sigma$-cycles.

Definition 6.1. Let $p_m$ be the $m$th prime and $r$ be a positive integer. For each positive integer $a$, let $e_m(a)$ satisfy $p_m(a) \parallel a$. Then let $E_t(a_1, \ldots, a_r)$ denote the $l \times r$ matrix whose $m^{th}$ entry is $e_m(a_i)$.

The following lemma shows that the lcm of $k$-subsets of the largest $r$ cycle lengths can be expressed in terms of the first $l$ primes with a small error.

Lemma 6.1. With the notation defined above,

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{1 \leq i_1 < \cdots < i_k \leq r} \left\{ \frac{L_{i_1}^{(n)} \cdots L_{i_k}^{(n)}}{\lcm(L_{i_1}^{(n)}, \ldots, L_{i_k}^{(n)})} \neq g_k(E_t(L_{i_1}^{(n)}, \ldots, L_{i_k}^{(n)})) \right\} \leq \left( \frac{r}{2} \right) \frac{1}{p_l}$$

Proof. Let $I_{m,b}$ be as in (3.2). Define $\varepsilon(n,b), \varepsilon^*(n)$ and $\varepsilon(\lambda)$ as in Lemma 3.3. Set $b = o(n)$. For $1 \leq k_1, k_2 \leq b$, let $\delta_{k_1,k_2}(l,n)$ be the fraction of tuples $(m_1, m_2) \in I_{k_1,b} \times I_{k_2,b}$ such that $p \mid m_1$ and $p \mid m_2$ for some prime $p > p_l$. Let $\delta_{k_1,\ldots,k_r}(l,n)$ denote the fraction of tuples $(l,n)$ such that $1 \leq i_1 < \cdots < i_k \leq r$. Then

$$\limsup_{n \to \infty} \delta_{k_1,\ldots,k_r}(l,n) \leq \left( \frac{r}{2} \right) \limsup_{n \to \infty} \frac{n^{-1}}{P_m} \sum_{m=l+1}^{n} \frac{3b}{p_m n^2} + \frac{3b^2}{n^2} \leq \left( \frac{r}{2} \right) \frac{1}{p_l}.$$

Therefore,

$$\mathbb{P} \left( \bigcup_{1 \leq i_1 < \cdots < i_k \leq r} \left\{ \frac{L_{i_1}^{(n)} \cdots L_{i_k}^{(n)}}{\lcm(L_{i_1}^{(n)}, \ldots, L_{i_k}^{(n)})} \neq g_k(E_t(L_{i_1}^{(n)}, \ldots, L_{i_k}^{(n)})) \right\} \right) \leq \sum_{1 \leq m_1 < \cdots < m_k \leq n} \mathbb{P} \left( \bigcap_{i=1}^{k} \left\{ L_{i}^{(n)} = m_i \right\} \bigcap_{1 \leq i_1 < \cdots < i_k \leq r} \left\{ \frac{m_1 \cdots m_k}{\lcm(m_1, \ldots, m_k)} \neq g_k(E_t(m_1, \ldots, m_k)) \right\} \right) \leq \sum_{k_1 + \cdots + k_r \leq n} \mathbb{P} \left( f^{(r)} \left( \frac{k_1}{b}, \ldots, \frac{k_r}{b} \right) + \varepsilon(n,b) \right) \left( \frac{1}{b} \right)^r \delta_{k_1,\ldots,k_r}(l,n) \leq \sum_{k_1 + \cdots + k_r \leq n} \mathbb{P} \left( L_r^{(n)} < \lambda n \right) + \mathbb{P} \left( L_r^{(n)} > (1-\lambda)n \right)$$

The result follows on taking $n \to \infty$ and $\lambda \to 0$. $\square$

We establish the following generalization of Lemma 3.3.

Lemma 6.2. Let $\alpha \in \mathbb{R}$ be irrational and fix positive integers $r$ and $l$ with $r \geq k$. For $1 \leq i_1 < \cdots < i_k \leq l$, let $U_{i_1,\ldots,i_k}$ be i.i.d. random variables distributed uniformly on $[0,1]$ and independent of the $PD(\theta)$ process $(L_1, L_2, \ldots)$. For $1 \leq m \leq l$ and $1 \leq i \leq r$, let the random variables $X_{m,i}$ from (1,10) be independent of both $U_{i_1,\ldots,i_k}$ and $(L_1, L_2, \ldots)$. Then we have the distributional convergence.
\[
\left( E_{i}(L_{1}^{(n)}), ..., L_{r}^{(n)} \right), L_{1}^{(n)}/n, ..., L_{r}^{(n)}/n, \left( \psi_{\alpha} \left( \frac{L_{i_{1}}^{(n)} ... L_{i_{k}}^{(n)}}{g_{k}(E_{i}(L_{i_{1}}^{(n)}, ..., L_{i_{k}}^{(n)}))} \right) \right)_{1 \leq i_{1} < ... < i_{k} \leq r} \xrightarrow{d} \left( X_{i_{1}}^{1}, ..., X_{i_{k}}^{r}, U_{i_{1},1}, ..., i_{k} \right)_{1 \leq i_{1} < ... < i_{k} \leq r}
\]

**Proof.** Let the intervals \( I_{m,b} \) be as in (3.2). Define \( \varepsilon(n, b), \varepsilon^{*}(n) \) and \( \varepsilon(\lambda) \) as in the proof of Lemma 3.1. Recall \( \psi_{\alpha}(j) \leq x \) if \( \{ j \alpha \} \geq 1 - x \).

For \( 1 \leq i_{1} < ... < i_{k} \leq r \) and \( 1 \leq i \leq r \), let \( c_{i}, d_{i}, ..., i_{k} \in [0, 1] \) be real numbers. Let \( A = [a_{1}, ..., a_{r}] \in \mathbb{N}^{r} \). Then \( E_{i}(m_{1}, ..., m_{r}) = \mathbb{A} \) if \( (m_{1}, ..., m_{r}) \in \mathbb{A} \). Define the fraction

\[
\gamma_{(k),b}^{\mathbb{A}} := \prod_{i=1}^{r} \frac{|I_{k_{i},b} \cap \mathbb{A}|}{|I_{k_{i},b}|}
\]  

(6.1)

Finally, let \( \alpha = \left( \frac{\alpha}{g_{k}(a_{1}, ..., a_{k})} \right)_{1 \leq i_{1} < ... < i_{k} \leq r} \). Then

\[
P \left( \bigcap_{i=1}^{r} \{ L_{i}^{(n)}/n < c_{i} \} \cap \{ E_{i}(L_{1}^{(n)}, ..., L_{r}^{(n)}) = \mathbb{A} \} \cap \right.
\]

\[
\prod_{1 \leq i_{1} < ... < i_{k} \leq r} \left\{ \psi_{\alpha} \left( \frac{L_{i_{1}}^{(n)} ... L_{i_{k}}^{(n)}}{g_{k}(E_{i}(L_{i_{1}}^{(n)}, ..., L_{i_{k}}^{(n)}) \right)} < d_{i_{1}, ..., i_{k}} \right) \right)
\]

\[
\leq \sum_{1 \leq m_{r} < ... < m_{1} \leq n, m_{i}/n < c_{i}} \prod_{1 \leq i_{1} < ... < i_{k} \leq r} \left\{ \psi_{\alpha} \left( \frac{m_{i_{1}} ... m_{i_{k}}}{g_{k}(a_{1}, ..., a_{k})} \right) < d_{i_{1}, ..., i_{k}} \right\} + \varepsilon^{*}(n)
\]

\[
\leq \sum_{\lambda b \leq k_{1} \leq ... \leq k_{1} \leq \lambda b, k_{1} + ... + k_{r} \leq (1 - \lambda) b, \lambda b/k_{1} \in (c_{i})_{i \leq c_{i}}} \int f_{b}^{(r)}(x_{1}, ..., x_{r}) dx_{1} ... dx_{r} \prod_{i=1}^{r} \prod_{m=1}^{l} \left( 1 - 1/p(m) \right) (1/p(m))^{A_{m}^{r}} \prod_{1 \leq i_{1} < ... < i_{k} \leq r} d_{i_{1}, ..., i_{k}} + \varepsilon(\lambda)
\]

Then by Lemma 5.1, \( \gamma_{(k),b}^{\mathbb{A}} \to \prod_{i=1}^{r} \prod_{m=1}^{l} \left( 1 - 1/p(m) \right) (1/p(m))^{A_{m}^{r}} \) uniformly as \( n/b \to \infty \) and by Lemma 5.2, \( D(P_{(k),b}^{\mathbb{A}, \alpha}) \to 0 \) uniformly as \( n/b \to \infty \). Taking \( b = o(n) \) so that \( n/b \to \infty \), we can bound (6.2) in the limit by

\[
\int \int f_{b}^{(r)}(x_{1}, ..., x_{r}) dx_{1} ... dx_{r} \prod_{i=1}^{r} \prod_{m=1}^{l} \left( 1 - 1/p(m) \right) (1/p(m))^{A_{m}^{r}} \prod_{1 \leq i_{1} < ... < i_{k} \leq r} d_{i_{1}, ..., i_{k}} + \varepsilon(\lambda)
\]

We obtain the analogous expression (without the \( \varepsilon(\lambda) \) term) for the lower bound.

Taking \( \lambda \to 0 \), the limiting probability is

\[
P \{ L_{i} < c_{i}, 1 \leq i \leq r \} P \{ X_{i_{1}}^{1}, ..., X_{i_{k}}^{r} = \mathbb{A} \} \prod_{1 \leq i_{1} < ... < i_{k} \leq r} P \{ U_{i_{1},1}, ..., i_{k} < d_{i_{1}, ..., i_{k}} \}
\]

\[\square\]
Corollary 6.1. Recall the notation from Lemma 6.2 and as usual, allow the multi-dimensional indices \((i_1, ..., i_k)\) to run over the range \(1 \leq i_1 < ... < i_k \leq r\). Then we have the distributional convergence

\[
\left(\frac{L^{(n)}_{i_1}...L^{(n)}_{i_k}}{\text{lcm}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k})}, (\psi_n(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k}))\right) \xrightarrow{d} \left((g_k(X^{i_1,...,i_k}_n)), (L_{i_1}...L_{i_k}), (U_{i_1},...,U_{i_k})\right)
\]

Proof. By Lemma 6.2 and the continuous mapping theorem,

\[
\left((g_k(\mathbf{E}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k}))), (L^{(n)}_{i_1}...L^{(n)}_{i_k})/n^k, \psi_n(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k}))\right) \\
\xrightarrow{d} \left((g_k(X^{i_1,...,i_k}_n)), (L_{i_1}...L_{i_k}), (U_{i_1},...,U_{i_k})\right)
\]

By Lemma 6.1 there exists an error function \(\varepsilon(l) \to 0\) as \(l \to \infty\) such that

\[
\limsup_{n \to \infty} \mathbb{P}\left\{\left(\frac{L^{(n)}_{i_1}...L^{(n)}_{i_k}}{\text{lcm}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k})}\right) \neq (g_k(\mathbf{E}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k})))\right\} < \varepsilon(l)
\]

and

\[
\mathbb{P}\{(g_k(X^{i_1,...,i_k}_n)) \neq (g_k(X^{i_1,...,i_k}_\infty))\} < \varepsilon(l).
\]

Taking \(l \to \infty\), we obtain the desired convergence. \(\square\)

Proof of Theorem 1.2. Let \(r\) be a positive integer. To isolate the contribution from the \(r\) largest cycles, define

\[
\xi_{n,k}^r = k! \sum_{1 \leq i_1 < ... < i_k \leq r} \frac{L^{(n)}_{i_1}...L^{(n)}_{i_k}}{\text{lcm}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k})} \sum_{q \in \mathbb{Z}} \delta_{n^k(q/\text{lcm}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k}))}(\xi_{i_1,...,i_k})
\]

and let \(\eta_{n,k}^r = \mathbf{c}_{n,k}^r - \xi_{n,k}^r\) denote the remaining measure. Similarly, define

\[
\xi_{\infty,k}^r = k! \sum_{1 \leq i_1 < ... < i_k \leq r} g_k(X^{i_1,...,i_k}_\infty) \sum_{q \in \mathbb{Z}} \delta_{q+(U_{i_1},...,U_{i_k})} g_k(X^{i_1,...,i_k}_\infty)/(L_{i_1}...L_{i_k})
\]

and set \(\eta_{\infty,k}^r = \mathbf{c}_{\infty,k}^r - \xi_{\infty,k}^r\).

Let \(f\) be a continuous function with support contained in some interval \((-T, T)\). Then

\[
\int f d\xi_{n,k}^r = k! \sum_{1 \leq i_1 < ... < i_k \leq r} \frac{L^{(n)}_{i_1}...L^{(n)}_{i_k}}{\text{lcm}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k})} \sum_{q \in \mathbb{Z}} f\left(\frac{q - \text{lcm}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k}))}{\text{lcm}(L^{(n)}_{i_1}, ..., L^{(n)}_{i_k})/n^k}\right)
\]

and

\[
\int f d\xi_{\infty,k}^r = k! \sum_{1 \leq i_1 < ... < i_k \leq r} g_k(X^{i_1,...,i_k}_\infty) \sum_{q \in [-T+1,T+1]} f\left(\frac{q + U_{i_1},...,U_{i_k}) g_k(X^{i_1,...,i_k}_\infty)}{L_{i_1}...L_{i_k}}\right)
\]

By Corollary 6.1 and the continuous mapping theorem,

\[
\int f d\xi_{n,k}^r \xrightarrow{d} \int f d\xi_{\infty,k}^r
\]
Now we turn to the remainder measures $\eta_{n,k}^r$ and $\eta_{n,k}^c$. Since $E[(\sum_i L_i)^k] = E[1^k] = 1$, by the Borel-Cantelli lemma \[ \int f d\eta_{n,k}^r \xrightarrow{a.s.} 0 \text{ as } r \to \infty. \] Note that this just verifies that $E_{\ast}^\infty_{n,k}$ is in fact a point process.

Showing that \[ \lim_{r \to \infty} \limsup_{n \to \infty} P \left\{ \int |f| d\eta_{n,k}^r > 0 \right\} = 0 \] is considerably more involved than in the $k = 1$ case. We split the point process $\eta_{n,k}^r$ into two parts:

1. points coming from $\sigma_k$-cycles containing tuples $(t_1, ..., t_k)$ such that $t_1, ..., t_k$ are all in different $\sigma$-cycles, not all of which are chosen from the $r$ largest cycles.
2. points coming from $\sigma_k$-cycles containing tuples $(t_1, ..., t_k)$ such that for some $i \neq j$, $t_i$ and $t_j$ are in the same $\sigma$-cycle.

More formally, we make the following:

**Definition 6.2.** For $0 \leq c \leq s \leq k$ and each $\sigma \in S_n$, define $H_{n,s}^{c,s}(\sigma)$ to be the set of all integers \[ \text{lcm}(L_{i_1}^{(n)}(\sigma), ..., L_{i_s}^{(n)}(\sigma)) \] such that $L_{i_1}^{(n)}(\sigma)$ are all distinct where $1 \leq i_1 < ... < i_c \leq r < i_{c+1} < ... < i_s$.

Define the corresponding point process \[ \nu_{n,s}^{c,s} = \sum_{j \in H_{n,s}^{c,s}} \sum_{q \in \mathbb{Z}} \delta_{n^k(q/j-\alpha)} \] (6.6)

It is easy to check that \[ \eta_{n,k}^* = \left( \sum_{c=0}^{k-1} \nu_{n,s}^{c,s} + \sum_{s=1}^{k-1} \sum_{c=0}^{k-1} \nu_{n,s}^{c,s} \right)^* \] (6.7)

where $\xi^*$ denotes the corresponding simple point process of a point process $\xi$. Then \[ P \left\{ \int |f| d\eta_{n,k}^* > 0 \right\} \leq \sum_{c=0}^{k-1} P \left\{ \int |f| d\nu_{n,s}^{c,s} > 0 \right\} + \sum_{s=1}^{k-1} \sum_{c=0}^{k-1} P \left\{ \int |f| d\nu_{n,s}^{c,s} > 0 \right\} \] (6.8)

If $c = s$, then $H_{n,s}^{c,s} \subset \{ \text{lcm}(L_{i_1}^{(n)}(\sigma), ..., L_{i_s}^{(n)}(\sigma)) : 1 \leq i_1 < ... < i_s \leq r \}$. Note that $\nu_{n,s}^{c,s}$ is the simple point process $\xi_{n,s}^c$ dilated by a factor of $n^{k-s}$. In other words,

\[ \int |f(x)| d\xi_{n,s}^c(x) \leq \int |f(xn^{k-s})| d\xi_{n,s}^c(x) \] (6.9)

The support of $f(xn^{k-s})$ is contained in $(-T/n^{k-s}, T/n^{k-s})$. Let $g_c$ be a continuous function with support contained in a small interval $(\varepsilon, \varepsilon)$. The mean intensity of $E_{\ast}^\infty_{n,s}$ is some finite constant $\lambda_s$. We have

\[ \limsup_{n \to \infty} P \left\{ \int |g_c| d\xi_{n,s}^c > 0 \right\} < 2\varepsilon \lambda_s \]

Thus, for $s < k$ and for each $r$, \[ \lim_{n \to \infty} P \left\{ \int |f| d\nu_{n,s}^{r,s} > 0 \right\} = 0 \text{ by (6.5).} \]

Define the set \[ J_k(n) := \{ j \in [n^k] : (\exists q \in \mathbb{Z}) [\alpha - T/n^k < q/j < \alpha + T/n^k] \} \] (6.10)

Then by Theorem 7.4 below, for $1 \leq s \leq k$ and $0 \leq c < s$,

\[ \lim_{r \to \infty} \limsup_{n \to \infty} P \left\{ \int |f| d\nu_{n,s}^{r,c} > 0 \right\} \leq \lim_{r \to \infty} \limsup_{n \to \infty} P \left\{ H_{n,s}^{r,c} \cap J_k(n) \neq \emptyset \right\} = 0 \] (6.11)

By (6.8),

\[ \lim_{r \to \infty} \limsup_{n \to \infty} P \left\{ \int |f| d\eta_{n,k} > 0 \right\} = 0. \]
Clearly, $\mathbb{P}\{f|dn_{n,k}^r > 0\} = \mathbb{P}\{f|dn_{n,k}^{r*,k} > 0\}$. The rest of the proof follows the proof of Theorem 1.1 identically. □

Remark 6.2. For $\alpha$ of finite irrationality measure, let $\text{sub}^{\alpha}_{n,k}$ and $\text{irrep}^{\alpha}_{n,k}$ be the corresponding versions of the scaled eigenvalue point processes $\mathbf{E}_{n}^{\alpha}$ for the $k$-subset permutation representations and $S_{(n-k,1^k)}$ irreducible representations of the symmetric group $\mathfrak{S}_n$ (see [29] for details). Then it is easy to see that with $\xi_{n,k}^r$ and $\eta_{n,k}^r$ as defined above,

$$\text{sub}^{\alpha}_{n,k} = \xi_{n,k}^r / k! + \text{sub}^{\alpha}_{n,k}$$

and

$$\text{irrep}^{\alpha}_{n,k} = \xi_{n,k}^r / k! + \text{irrep}^{\alpha}_{n,k}$$

where the supports of $\text{sub}^{\alpha}_{n,k}$ and $\text{irrep}^{\alpha}_{n,k}$ are both contained in the support of $\eta_{n,k}^r$ almost surely. Thus, both $\text{sub}^{\alpha}_{n,k}$ and $\text{irrep}^{\alpha}_{n,k}$ converge weakly to $\frac{1}{k!} \mathbf{E}_{\infty}^\alpha$.

7. Bound on small cycle contribution

We devote this section to the task of establishing the technical result Theorem 7.1 used in 6.11 for the proof of Theorem 1.2. We will use a key discrepancy bound proved in Theorem 8.1 of the next section.

Theorem 7.1. For $1 \leq s \leq k$ and $0 \leq c < s$,

$$\lim_{r \to \infty} \limsup_{n \to \infty} \mathbb{P}\{H_{n,s}^{r,c} \cap J_k(n) \neq \emptyset\} = 0$$  (7.1)

In the sequel, we fix some choice of $c,s$, and $k$ (according to the restrictions of the theorem) for the sake of concreteness. Recall the definition of factorization classes from Definition 5.1. The following notation will be useful:

Definition 7.1. For $n > 10$, let $l = \lfloor (\log n)^{2s} \rfloor$. Define the set of $l$-tuples

$$Q_n = \{(e_1, ..., e_l) \in \mathbb{N}^l : (0 \leq e_j < 2s \log \log n) \wedge (\{|j : e_j > 0| < (\log \log n)^2\})\}$$  (7.2)

We also define the (disjoint) union

$$\mathcal{U}_n = \bigcup_{(e_1, ..., e_l) \in Q_n} \mathcal{P}_{e_1, ..., e_l}$$  (7.3)

Remark 7.1. The choice $l = \lfloor (\log n)^{2s} \rfloor$ is motivated by Lemma 7.7 below. All three parameters $\lfloor (\log n)^{2s} \rfloor$, $2s \log \log n$, and $(\log \log n)^2$ are chosen so that we can obtain the discrepancy bound in Theorem 8.1.

The first major task is to bound the probability $\mathbb{P}\{H_{n,s}^{r,c} \cap J_k(n) \neq \emptyset\}$ by a sum involving factorization classes in $\mathcal{U}_n$ and to replace the lcm in the definition of $H_{n,s}^{r,c}$ with a product involving the functions $g_s$ from (1.11), which will be constant over each factorization class. This is essential so that we can apply the equidistribution theory of polynomials.

To help streamline the presentation of inequalities, we will use the following conventions:

Definition 7.2. For two non-negative functions $f(n)$ and $g(n)$, we say that

$$f \ll g$$

if $f \leq Cg$ for some absolute constant $C$ that does not depend on $n$. We say that

$$f \lesssim g$$

if $f \leq Cg + h$ for some absolute constant $C$ and function $h(n)$ such that $\lim_{n \to \infty} h(n) = 0$. 
In the following, we allow the indices of $j_v$ to run over $c + 1 \leq v \leq s$ instead of $1 \leq v \leq s - c$ for notational convenience. For $i \geq 1$, let $Z_i$ be independent Poisson random variables with mean $\theta/i$. We will let $1(E)$ denote the indicator function of the event $E$, i.e. it takes the value 1 if $E$ holds and 0 otherwise. To state the next theorem, we introduce the functions

$$ f_{j_v}^{n, r, A + 1, \ldots, u_s}(x_1, \ldots, x_r) := \frac{e^{\theta} \Gamma(\theta) \theta^{r-1}}{x_1 \ldots x_r} \left( 1 - \sum_{i=1}^{r} x_i - \sum_{v=c+1}^{s} \frac{u_v j_v}{n} \right) / x_r 
$$

for positive integers $u_v$ and $j_v$ with domain

$$ D := \{ (x_1, \ldots, x_r) : 0 < x_r \leq \ldots \leq x_1 < 1 \} \land \left( x_1 + \ldots + x_r < 1 - \sum_{v=c+1}^{s} \frac{u_v j_v}{n} \right) \}
$$

This definition is motivated by the form of the density $f_{j_v}^{(r)}(x_1, \ldots, x_r)$ in (7.4). By scaling linearly, one can check that

$$ \int_{D} f_{j_v}^{n, u_v}(x_1, \ldots, x_r) dx_1 \ldots dx_r = \left( 1 - \sum_{v=c+1}^{s} \frac{u_v j_v}{n} \right)^{r-1} \quad (7.5) $$

**Theorem 7.2.** Let $Q_n$ be as defined in (7.2). Fix $0 < \lambda < 1$ and $0 < \varepsilon < 1$. Then as $n \to \infty$,

$$ \mathbb{P} \left\{ H_{n, c}^{r, \xi} \cap J_k(n) \neq \emptyset \right\} \leq \sum_{c \leq s} \sum_{1 \leq j_v < \ldots < j_s \leq [n/r]} \sum_{\lambda n \leq m_v \leq \ldots \leq m_s \leq \lambda n} \sum_{j_v \in \mathcal{P}_{u_v}} \frac{1}{n^r} \prod_{v=c+1}^{s} \frac{1}{j_v} \left( \prod_{j_v \in [n/m_v]} \prod_{j_v \in [n/m_v]} J_k \right) \sum_{1 \leq u_v \leq \lambda \varepsilon} f_{j_v}^{n, u_v}(x_1, \ldots, x_r) \left( \frac{m_1}{n}, \ldots, \frac{m_r}{n} \right) + \varepsilon(\lambda) \quad (7.6) $$

where the error term $\varepsilon(\lambda) \to 0$ as $\lambda \to 0$ and the implied absolute constant does not depend on $r$.

**Lemma 7.1.** Fix $0 < \lambda < 1$ and a positive integer $m$. For $1 \leq j_1 < \ldots < j_m \leq n$, let $c_{j_v}$ be non-negative integers such that $\sum_{v=1}^{m} j_v c_{j_v} < (1 - \lambda)n$. Then

$$ \mathbb{P} \left( \bigcap_{v=1}^{m} \{ C_{j_v}^{(n)} = c_{j_v} \} \right) \leq \mathbb{P} \left( \bigcap_{v=1}^{m} \{ Z_{j_v} = c_{j_v} \} \right) \quad (7.7) $$

where the absolute constant does not depend on $j_v$ or $c_{j_v}$.

**Proof.** Let $B = [n] \setminus \{ j_1, \ldots, j_m \}$. Recall from Definition 2.1 that $T_B := \sum_{j \in B} j Z_j$.

By the size biasing equation (2.11), we have the pair of equations

$$ i \mathbb{P} \{ T_B = i \} = \theta \sum_{l \in B} \mathbb{P} \{ T_B = i - l \} \quad (7.8) $$

$$ i \mathbb{P} \{ T_0 = i \} = \theta \sum_{l=1}^{i} \mathbb{P} \{ T_0 = i - l \} \quad (7.9) $$

Since

$$ \mathbb{P} \{ T_B = 0 \} \leq \prod_{i=m+1}^{n} \frac{1}{e^{\theta i}} \leq (3m)^{\theta} \mathbb{P} \{ T_0 = 0 \}, $$

we get

$$ \mathbb{P} \{ T_B = l \} \leq (3m)^{\theta} \mathbb{P} \{ T_0 = l \} \quad (7.10) $$
for all \( l \geq 0 \) by comparing (4.8) and (4.9) inductively. By using (4.10), we have (see e.g. 2)

\[
P\{T_{0n} = l\} = \exp(-\theta h(n + 1)) \frac{\theta(\theta + 1)...(\theta + l - 1)}{l!}
\]

(7.11)

for \( 0 < l \leq n \) where \( h(n + 1) \) as before is the \( n \)th harmonic number. Then

\[
nP\{T_{0n} = l\} \to \frac{e^{-\gamma} \theta^{\theta - 1}}{\Gamma(\theta)}
\]

(7.12)

uniformly over \( \lambda n < l \leq n \) and \( l/n \to x \in [\lambda, 1] \). Hence, by the Conditioning Relation 2.3,

\[
P\left( \bigcap_{v=1}^{m} \left\{ C_{j_v}^{(n)} = c_{j_v} \right\} \right) = P\left( \bigcap_{v=1}^{m} \left\{ Z_{j_v} = c_{j_v} \right\} \mid T_{0n} = n \right)
\]

\[
= P\left( \bigcap_{v=1}^{m} \left\{ Z_{j_v} = c_{j_v} \right\} \right) \frac{P\{T_{0n} = n - \sum j_v c_{j_v}\}}{P\{T_{0n} = n\}} \leq (3m)^{\theta} \frac{P\left( \bigcap_{v=1}^{m} \left\{ Z_{j_v} = c_{j_v} \right\} \right)}{P\{T_{0n} = n\}}
\]

\[
\ll P\left( \bigcap_{v=1}^{m} \left\{ Z_{j_v} = c_{j_v} \right\} \right)
\]

□

Lemma 7.2. Fix \( \varepsilon > 0 \). Then

\[
\sum_{j > n^2} P\{C_{j}^{(n)} > \lceil 2/\varepsilon \rceil\} \ll \frac{1}{n}
\]

(7.13)

Proof. By using the Conditioning Relation along with (7.10) and (7.11), we see that

\[
P\{C_{j}^{(n)} = l\} \ll n^{\max(1 - \theta, 0)} (\theta/j)^l
\]

(7.14)

uniformly over \( j \) and \( l \). Let \( a \) be a positive integer. Then

\[
P\{C_{j}^{(n)} > a\} = \sum_{l=a+1}^{\infty} P\{C_{j}^{(n)} = l\} \ll \frac{n^{\max(1 - \theta, 0)}}{a}
\]

(7.15)

Setting \( a = \lceil 2/\varepsilon \rceil \) completes the proof.

□

Lemma 7.3. Fix real numbers \( 0 < \lambda < 1 \) and \( 0 < \varepsilon < 1 \) and positive integers \( u_{c+A+1}, \ldots, u_s \) where \( A \) is an integer such that \( 0 \leq A \leq s - c \). Let \( j_{c+1}, \ldots, j_s \) be integers such that \( 1 \leq j_{c+1} < \ldots < j_{c+A} < n^2 < j_{c+A+1} < \ldots < j_s \leq [n/r] \). Let \( T_{j_{c+A+1}, \ldots, j_s} \) be the set of tuples \( (m_1, \ldots, m_r) \) such that \( m = m_1 + \ldots + m_r \leq (1 - \lambda)n - \sum u_v j_v \) and \( \max(j_v, 1/2) < m_r < \ldots < m_1 < n \). Set \( x_i = m_i/n \). Then for all \( (m_1, \ldots, m_r) \in T_{j_{c+A+1}, \ldots, j_s} \),

\[
n^r P\left( \bigcap_{v=c+1}^{c+A} \left\{ C_{j_v}^{(n)} > 0 \right\} \cap \bigcap_{v=c+A+1}^{s} \left\{ C_{j_v}^{(n)} = u_v \right\} \cap \bigcap_{i=1}^{r} \left\{ L_i^{(n)} = m_i \right\} \right)
\]

\[
\ll n \frac{\theta^r \theta^{(1 - \varepsilon)/2}}{\lambda^{n^{(1 - \varepsilon)/2}}} |n^{(1 - \varepsilon)/2}|
\]

where the absolute constant does not depend on \( r \) or the choice of integers \( j_v \) or \( m_i \).

Proof. Following analogous calculations to the steps in (2.8), (2.9) and (2.10), we see that

\[
P\left( \bigcap_{v=c+1}^{c+A} \left\{ C_{j_v}^{(n)} > 0 \right\} \cap \bigcap_{v=c+A+1}^{s} \left\{ C_{j_v}^{(n)} = u_v \right\} \cap \bigcap_{i=1}^{r} \left\{ L_i^{(n)} = m_i \right\} \right) = \frac{\theta^r e^{-\theta h(n+1) - h(m_v)}}{m_1 \ldots m_r} \Delta_n
\]
where

\[
\Delta_n = \frac{\Pr\left\{ T_{0,m-1} = n - m \right\} \cap \bigcap_{v=c+1}^{c+A} \{ Z_{j_v} > 0 \} \cap \bigcap_{v=c+A+1}^{s} \{ Z_{j_v} = u_v \}}{\Pr\{ T_{0n} = n \}}
\]

(7.16)

Let \( B = [m, -1] \setminus \{ j_{c+1}, \ldots, j_s \} \). If \( Z_{j_v} \) is less than \( n^{(1-\epsilon)/2} \) for \( c + 1 \leq v \leq c + A \), then for sufficiently large \( n \), \( T_{0,m-1} = n - m \) implies that \( T_B = n - m - \sum_{v=c+A+1}^{s} u_v j_v - o(n) \geq \lambda n / 2 \) (say).

The proof of Lemma 2.1 shows that the point probabilities \( n \Pr\{ T_{0,m-1} = l \} \) and \( n \Pr\{ T_{0,n} = l \} \) converge uniformly for \( l \geq \lambda n / 2 \). Then since \( \Pr\{ T_B = l \} \leq (3m)^{\theta} \Pr\{ T_{0,m-1} = l \} \) by (7.10),

\[
\Delta_n \ll \Pr\left( \bigcap_{v=c+1}^{s} \{ Z_{j_v} > 0 \} \right) \frac{q_n}{|n^\delta|!} \text{ for any } \delta > 0.
\]

\[\square\]

We define the following two sets:

\[\mathcal{B}_1 = \{ q \in \mathbb{N} : (3m, e \in \mathbb{N})[(m \leq (\log n)^{2s}) \land (e > 2s \log \log n) \land (p_{c+1}^s | q)]\} \quad (7.17)\]

\[\mathcal{B}_2 = \{ q \in \mathbb{N} : \left| \{ j : e_j(q) > 0 \}_{1 \leq j \leq \lfloor \log(\log n) \rfloor} > (\log n)^2 \right\} \quad (7.18)\]

where recall \( e_j(q) \) is the exponent of the \( p_j \)th prime in the prime factorization of \( q \). Then \( \mathbb{N} \setminus \mathcal{Y}_n = \mathcal{B}_1 \cup \mathcal{B}_2 \). The next two lemmas show that as \( n \to \infty \), the contribution to the probability for \( j_v \notin \mathcal{Y}_n \) is negligible.

**Lemma 7.4.** Let \( \mathcal{B}_1 \) be as defined in (7.17). Then for \( 0 < \lambda < 1 \),

\[
\lim_{n \to \infty} \Pr\left( \bigcap_{v=c+1}^{s} \{ C_{j_v}^{(n)} > 0 \} \cap \left\{ \sum_{v=c+1}^{s} j_v C_{j_v}^{(n)} < (1 - \lambda)n \right\} \right) = 0
\]

where \( j_{c+1} \neq \ldots \neq j_s \) is shorthand for \( j_{c+1}, \ldots, j_s \) all distinct.

**Proof.** By Lemma 7.1

\[
\leq \sum_{m=1}^{(\log n)^{2s}} \left( \sum_{1 \leq j_{c+1}, \ldots, j_s \leq n} \Pr\left( \bigcap_{v=c+1}^{s} \{ C_{j_v}^{(n)} > 0 \} \cap \left\{ \sum_{v=c+1}^{s} j_v C_{j_v}^{(n)} < (1 - \lambda)n \right\} \right) \right)
\]

\[
\leq \sum_{m=1}^{(\log n)^{2s}} \left( \sum_{1 \leq j_{c+1}, \ldots, j_s \leq n} \frac{1}{p_{m}^{2s} \log \log n} \sum_{i=1}^{n} \frac{1}{ip_{m}^{2s} \log \log n} \right)
\]

\[
\leq \sum_{m=1}^{(\log n)^{2s}} \frac{1}{p_{m}^{2s} \log \log n}
\]

\[
\leq \frac{1}{p_{m}^{2s} \log \log n}
\]
(\log n)^s \frac{1}{2^{2s} \log \log n - 1} \to 0 \text{ as } n \to \infty.

Now we turn to the second set \(B_2\). In words, \(B_2\) is the set of integers with more than \((\log \log n)^2\) distinct prime factors among the first \((\log n)^2\) primes. Let \(\omega(m)\) denote the number of distinct prime factors of an integer \(m\). The Erdős-Kac theorem (see e.g. [8] or [28]) tells us that if \(x\) is a randomly chosen integer from \([n]\), then

\[\frac{\omega(x) - \log \log x}{\sqrt{\log \log x}} \xrightarrow{d} \mathcal{N}(0, 1)\text{ as } n \to \infty.\]

Thus, the number of prime factors of \(m\) is concentrated about \(\log \log m\) and one would expect \(\omega(m) > (\log \log m)^2\) to be a rare event. We need a large deviation inequality to bound this probability.

Let \(P_{a,b}\) be the uniform distribution on the integers in \((a, b]\) and let \(\mu_m = \mathbb{E}[\omega(x)]\) where the expectation is taken under the measure \(P_{0,m}\). Let \(\delta_p(m) = 1\) if \(m\) is divisible by \(p\) and \(0\) otherwise. Then (see e.g. [13, Chapt. 22])

\[
\mu_m = \mathbb{E}[\omega(x)] = \mathbb{E}\left[\sum_{p \leq m} \delta_p(x)\right] = \sum_{p \leq m} \frac{m/p}{m} = \log \log m + O(1) \quad (7.19)
\]

Srinivasan [26] has derived an explicit exponential Chernoff-type bound for large deviations of \(\omega(m)\):

**Proposition 7.1** (Srinivasan). For \(m \geq 2\) and any \(\delta > 0\),

\[
\mathbb{P}_{0,m}\{\omega(x) \geq \mu_m(1 + \delta)\} \leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^{\mu'_m} \quad (7.20)
\]

where \(\mu'_m = \sum_{p \leq m} \frac{1}{p} = \log \log m + O(1)\).

Note that this is a bound for each \(m\) rather than an asymptotic bound. For more information on Erdős-Kac large deviation results, see for instance [17, 24, 14, 23].

**Definition 7.3.** The \(n\)\(^{th}\) primorial \(p_n\#\) is the product of the first \(n\) primes, i.e.

\[p_n\# := \prod_{m=1}^n p_m\]

**Remark 7.2.** By the Prime Number Theorem, we have the well-known asymptotic

\[p_n\# = \exp((1 + o(1))n \log n) \quad (7.21)\]

**Lemma 7.5.** Let \(B_2\) be as defined in (7.18). Then for \(0 < \lambda < 1\),

\[
\lim_{n \to \infty} \sum_{1 \leq j_1 < \ldots < j_s \leq n} \mathbb{P}\left(\bigcap_{j_s \in B_2} \{C_{j_s}(n) > 0\} \cap \left\{\sum_{v=c+1}^s j_v C_{j_v}(n) < (1 - \lambda)n\right\}\right) = 0
\]

**Proof.** By (7.21),

\[c_n := \prod_{m=1}^{(\log \log n)^2} p_m = \exp((1 + o(1))2(\log \log n)^2 \log \log \log n)\]

For each positive integer \(m\), define the interval \(I_m = (2^{m-1}c_n, 2^m c_n]\). Let \(I_M\) be the interval that contains the integer \(n\). Then \(M \leq 2 \log n\).
We set $\delta = \log \log n$ in Proposition 7.1. For $c_n < m \leq n$ and sufficiently large $n$,

$$\mathbb{P}_{0,m}\{\omega(x) \geq (\log \log n)^2\} \leq \frac{1}{(\log \log n)(\log \log \log n)}$$

and therefore for $m \leq M$,

$$\mathbb{P}_{1,n}\{\omega(x) \geq (\log \log n)^2\} \leq \frac{2}{(\log \log n)(\log \log \log n)} \leq \frac{1}{(\log n)^2s}$$

for sufficiently large $n$.

Then

$$\sum_{1 \leq j \leq n} \frac{j}{1} = \sum_{c_0 \leq j \leq n} \frac{1}{(\log \log n)^2} \leq \sum_{m=1}^{M} \sum_{j \in I_m} \frac{1}{(\log \log n)^2} \leq 2 \log n \frac{1}{(\log n)^2s}.$$

The result then follows from Lemma 7.1.

The following is immediate.

**Corollary 7.1.** If $c > 0$, let $1 \leq i_1 < \ldots < i_c \leq r$. Then

$$\lim_{n \to \infty} \mathbb{P}\left( \bigcup_{1 \leq u \leq c} \{L^{(n)}_{i_u} \notin \mathcal{Y}_n\} \right) = 0.$$

**Definition 7.4.** Let $l = [(\log n)^{2s}]$. For each $t \in \mathbb{N}$, we define the set of tuples

$$\mathcal{X}^t_n = \{(m_1, \ldots, m_t) \in \mathbb{N}^t : \gcd(m_i, m_j) = g_2(E_l(m_i, m_j)) \text{ for } 1 \leq i < j \leq t\} \quad (7.22)$$

Lemmas 7.4 and 7.5 and Corollary 7.1 provided bounds on the probability of avoiding the "good set" $\mathcal{Y}_n$. The next two lemmas deal with the "good sets" $\mathcal{X}^t_n$ from (7.22). In words, $\mathcal{X}^t_n$ consists of the tuples $(m_1, \ldots, m_t)$ of integers that are pairwise relatively prime with respect to primes $p_m$ such that $m > (\log n)^{2s}$.

**Lemma 7.6.** If $c > 0$, let $1 \leq i_1 < \ldots < i_c \leq r$. Then

$$\lim_{n \to \infty} \mathbb{P}\{(L^{(n)}_{i_1}, \ldots, L^{(n)}_{i_c}) \notin \mathcal{X}^{c}_n\} = 0.$$

**Proof.** This follows from Lemma 6.5.

**Lemma 7.7.** If $c > 0$, let $1 \leq i_1 < \ldots < i_c \leq r$. If $s > 1$, let $l = [(\log n)^{2s}]$. Define the sets (where appropriate)

$$\mathcal{A}_1 = \bigcup_{c+1 \leq u \leq s} \{\gcd(L^{(n)}_{i_u}, j_v) \neq g_2(E_l(L^{(n)}_{i_u}, j_v))\}$$

$$\mathcal{A}_2 = \bigcup_{c+1 \leq u \leq s} \{\gcd(j_u, j_v) \neq g_2(E_l(j_u, j_v))\}$$

Then for $0 < \lambda < 1$,

$$\lim_{n \to \infty} \sum_{1 \leq j_{c+1}, \ldots, j_s \leq n} \mathbb{P}\left( \bigcap_{v=c+1}^{s} \{C^{(n)}_{j_v} > 0\} \cap \bigcap_{i=1}^{r} \{C^{(n)}_{L^{(n)}_{i}} = 1\} \cap \left(\sum_{i=1}^{s} L^{(n)}_{i} + \sum_{v=c+1}^{s} j_v C^{(n)}_{j_v} < (1-\lambda)n\right) \cap (\mathcal{A}_1 \cup \mathcal{A}_2) \right) = 0 \quad (7.23)$$
Proof. First, consider the set $\mathcal{A}_1$. By Lemma 7.1, the expression in (7.23) with $\mathcal{A}_1 \cup \mathcal{A}_2$ replaced by $\mathcal{A}_1$ can be bounded by

$$\ll \sum_{1 \leq j_1+1 \cdots j_{s-1} \leq n \atop m > (\log n)^{2s}} \sum_{j_{s+1} \not\equiv \cdots \not\equiv j_{s-1} \atop p_m \mid j_{s+1}} \sum_{m > (\log n)^{2s}} \frac{1}{m} \left( \sum_{j_1+1 \cdots j_{s+1} \leq n} \frac{1}{p_m} \right) \ll (\log n)^{s+1} \sum_{m > (\log n)^{2s}} \frac{1}{p_m^2} \ll (\log n)^{s+1} \frac{1}{(\log n)^{2s}} \to 0$$

as $n \to \infty$. The bound for $\mathcal{A}_2$ is immediate. $\square$

Now we are ready to put the ingredients together to obtain the bound in (7.6).

Proof of Theorem 7.2. Let $l = \lfloor (\log n)^{2s} \rfloor$. If $c = 0$, the appropriate summations and events in the expressions below are ignored. By Corollary 7.1 and Lemma 7.6

$$\mathbb{P}\left( \bigcup_{c_1 < \cdots < c_r \leq n} \left\{ \text{lcm}(L_{c_1}^{(n)}, \ldots, L_{c_r}^{(n)}) \in J_k \right\} \right) \leq \sum_{1 \leq i_1 \cdots i_s \leq r \atop r < i_{s+1} \cdots < i_t} \sum_{u=1}^{c} \mathbb{P}\left( \bigcup_{c_1 < \cdots < c_r \leq n} \left\{ \text{lcm}(L_{c_1}^{(n)}, \ldots, L_{c_r}^{(n)}) \in J_k \right\} \cap \bigcup_{u=1}^{c} \left\{ L_{c_u}^{(n)} \neq L_{i_u}^{(n)} \right\} \cap \left\{ L_{r}^{(n)} > \lambda n \right\} \right)$$

where $\varepsilon(\lambda) \to 0$ as $\lambda \to 0$. Applying Lemma 7.2, this is bounded by

$$\ll \sum_{1 \leq i_1 < \cdots < i_s \leq r \atop r < i_{s+1} \cdots < i_t} \sum_{u=1}^{c} \mathbb{P}\left( \bigcup_{c_1 < \cdots < c_r \leq n} \left\{ \text{lcm}(L_{c_1}^{(n)}, \ldots, L_{c_r}^{(n)}) \in J_k \right\} \cap \bigcup_{u=1}^{c} \left\{ L_{c_u}^{(n)} \neq L_{i_u}^{(n)} \right\} \cap \left\{ L_{r}^{(n)} > \lambda n \right\} \cap \left\{ \left( \sum_{i=1}^{r} L_{i}^{(n)} + \frac{4}{\varepsilon} \sum_{u=1}^{c} j_u < (1-\lambda)n \right) \right\} \right)$$

Then applying Lemmas 7.1, 7.6 and 7.2, this expression is bounded by

$$\ll \sum_{1 \leq i_1 < \cdots < i_s \leq r \atop r < i_{s+1} \cdots < i_t} \sum_{u=1}^{c} \mathbb{P}\left( \bigcup_{c_1 < \cdots < c_r \leq n} \left\{ \text{lcm}(L_{c_1}^{(n)}, \ldots, L_{c_r}^{(n)}) \in J_k \right\} \cap \bigcup_{u=1}^{c} \left\{ L_{c_u}^{(n)} \neq L_{i_u}^{(n)} \right\} \cap \left\{ \left( \sum_{i=1}^{r} L_{i}^{(n)} + \frac{4}{\varepsilon} \sum_{u=1}^{c} j_u < (1-\lambda)n \right) \right\} \right)$$
Finally, applying Lemma 7.3 completes the proof.

Recall that in the proof of the one-dimensional $k = 1$ case, the contribution from the remainder measure $\eta'_n$ was bounded by a sum of the form $\sum_{j=1}^{n/r} \frac{1}{j} \mathbf{1}(j \in J_1)$. This sum was in turn dealt with in (3.10) by partitioning the range of summation into intervals $I_m,b$ defined in (3.2) and using equidistribution to estimate the number of $j$ in $|J_1 \cap I_m,b|$. To bound the sum in (7.9), we similarly proceed by partitioning the range of summation into boxes and estimate for each tuple $(e_1, \ldots, e_s)$ of vectors in $Q_n$ the number of points $(j_v, m_{i_v})$ in each box such that $j_v \in \mathcal{P}_{e_i}$ and $m_{i_v} \in \mathcal{P}_{e_i}$ and $\prod_{j_v \in \mathcal{P}_{e_i}} m_{i_v} \in J_k$. The definition of the intervals used in this multidimensional partition are somewhat more involved than in the $k = 1$ case. For the $m_i$ indices, we will use the intervals $I_{m,b}$ where we set $b = \log n$ (any choice of $b$ such that $b, n/b \to \infty$ will suffice). For the $j_v$ indices, we define new intervals $T_m$ as follows:

**Definition 7.5.** Let $\mu = \mu(\alpha)$ be the irrationality measure of $\alpha$. Set $\varepsilon = 1/(100k^2\mu^3)$ and $\delta = 10k\mu^3$. Let $M'$ be the smallest integer that satisfies the inequality $2^{M'} n^{\varepsilon} > n/r$ and let $M'' = \lfloor M' - \log \log n \rfloor$. Then for $1 \leq m \leq M''$, define $T_m = \lfloor 2^{m-1} n^{\varepsilon}, 2^{m} n^{\varepsilon} \rfloor$. For $m > M''$, define $T_m$ to be consecutive intervals of integers all of length $2^{M''} n^{\varepsilon}$ starting with $T_{M''+1} = \lfloor 2^{M''+1} n^{\varepsilon}, 2^{M''+1} n^{\varepsilon} \rfloor$. Define $M$ to be the smallest integer that satisfies the inequality $2^{M''} n^{\varepsilon} > n^\delta$ and define $M^*$ to be the integer such that $T_{M^*}$ contains $[n/r]$. For $1 \leq m \leq M^*$, write $T_m = (F_m, F_{m+1})$ where $F_m \in \mathbb{Z}$.

**Remark 7.3.** The seemingly arbitrary definition of the parameters $\varepsilon$ and $\delta$ and the intervals $T_m$ is motivated by the need for them to satisfy several properties. Most crucially, they need to be defined so that a satisfactory discrepancy bound can be obtained uniformly over all the boxes in the partition (proved in the next section). The parameters $\varepsilon$ and $\delta$ are chosen so that the multi-dimensional Weyl inequality in Lemma 7.4 can be applied to obtain a non-trivial bound. For this, we need $\delta > 2\varepsilon k \mu$ while being small enough so that the interval $1 \leq j \leq n^{\delta \varepsilon}$ does not contain any $j \in J_k$ for sufficiently large $n$. Also, we need that $|T_m| \geq n^\delta$ for $m > M$. We see that these two properties are indeed satisfied. The first follows from the definition of irrationality measure of $\alpha$ and the second is a result of the geometric progression in the interval lengths up to $M''$. We taper the interval lengths for $m > M''$ so that all the $T_m$ have cardinality at most $2^{M''} n^{\varepsilon} \leq n/\sqrt{\log n} = o(n)$ so that (7.24) below is satisfied.

**Proof of Theorem 7.7.** By Theorem 7.2 $\mathbb{P}\left\{ H^{r,c}_{n,c} \cap J_k(n) \neq \emptyset \right\}$ is bounded by a sum $\sum_{0 \leq A \leq c} S_A$ where each $S_A$ is a sum over “small” indices $j_v$ for $c+1 \leq v \leq c+A$ which are all bounded by $n^{\varepsilon}$. Splitting further, we can write this as a sum $\sum_{A,B,C>0} S_{A,B,C}$ where each sum $S_{A,B,C}$ is a sum over the following ranges: for $c+1 \leq v \leq c+A$ we have $1 \leq j_v \leq n^\varepsilon$; for $c+A+1 \leq v \leq c+A+B$ we have $j_v \in T_{l_v}$ where $1 \leq l_v < M$; for $c+A+B+1 \leq v \leq s$ we have $j_v \in T_{l_v}$ where $l_v \geq M$.

Note that the ranges of the variables $A,B,C$, and $u_{c+A+1}, \ldots, u_s$ depend only on the dimension $k$ and irrationality measure $\mu$. Therefore, it suffices to consider a fixed choice of these parameters. If $C + c = 0$, then $\prod_{j_v \in T_{l_v}} m_{i_v} \leq n^{\delta \varepsilon}$ and hence $\prod_{j_v \in T_{l_v}} m_{i_v} \notin J_k$ for sufficiently large $n$. Thus, we may assume that $C + c \geq 1$.

Let $\Delta_{u_{c+A+1}, \ldots, u_s}^{m_{i_v}} (x_1, \ldots, x_s)$ be as defined in Lemma 7.3. Since $|T_{l_v}| = o(n)$, the collection of functions $f_{j_v \in T_{l_v}, \ldots, j_s}^{n, u_{c+A+1}, \ldots, u_s} (x_1, \ldots, x_s)$ satisfy

$$\left| f_{j_v \in T_{l_v}, \ldots, j_s}^{n, u_{c+A+1}, \ldots, u_s} \left( \frac{k_1}{b}, \ldots, \frac{k_s}{b} \right) - f_{j_v \in T_{l_v}, \ldots, j_s}^{n, u_{c+A+1}, \ldots, u_s} \left( \frac{m_1}{n}, \ldots, \frac{m_s}{n} \right) \right| < \varepsilon(n) \quad (7.24)$$
for $j_v \in T_{e_v}$, $(m_1, \ldots, m_r) \in \Delta^c_{j_v+1} \cdots k_v$ and $m_i \in I_{k_i,b}$, and for some error function $\varepsilon(n) \to 0$ as $n \to \infty$.

Fix a choice of $1 \leq i_1 < \ldots < i_c \leq r$. We have (where the condition $j_v \in T_{e_v}$ is over the range $c + A + 1 \leq v \leq s$ and the condition $j_v \in \mathcal{P}_{e_v}$ is over the range $c + 1 \leq v \leq s$)

$$
\sum_{e_1, \ldots, e_s \in Q_n} \sum_{1 \leq j_v < \ldots < j_r \leq [n/r]} \lambda^{n \leq m_v < \ldots < m_1 \leq n} j_v \in \mathcal{P}_{e_v} \frac{1}{n^r} \prod_{v=c+1}^s \frac{1}{j_v} \left( \prod_{v=c+1}^s \frac{j_v \prod_{u=m_v}^n} {g_s(e_1, \ldots, e_s)} \in J_k \right) \left( \frac{m_1}{n}, \ldots, \frac{m_r}{n} \right)
$$

Now for each choice of $l_{c+A+1}, \ldots, l_s$ and $k_1, \ldots, k_r$, define the fraction

$$
\gamma(k_1, \ldots, k_r) := \prod_{v=c+1}^s \frac{|T_{e_v} \cap \mathcal{P}_{e_v}|}{|T_{e_v}|} \prod_{u=1}^c \frac{|I_{k_u, b} \cap \mathcal{P}_{e_u}|}{|I_{k_u, b}|}
$$

and the multiset

$$
\mathcal{P}_{(k_1, \ldots, k_r)} := \left\{ \prod_{u=1}^c m_{i_u} \prod_{v=c+1}^s j_v \prod_{u=m_v+1}^{s} \frac{j_v}{g_s(e_1, \ldots, e_s)} : (k_1, \ldots, k_r) \in T : \bigwedge_{v=c+1}^s \{m_{i_u} \in (I_{k_u, b} \cap \mathcal{P}_{e_u}) \} \right\}
$$

Recall that $j \in J_k$ if and only if $\{j\alpha \} \in [0, jT/n^k) \cup (1 - jT/n^k, 1]$ and therefore we are interested in the fraction of points $\{j\alpha \}$ in $\mathcal{P}_{(k_1, \ldots, k_r)}$ that satisfy that property. By definition, this fraction is bounded by

$$
\frac{2T}{n^k} \prod_{u=1}^c \frac{k_i}{b} \prod_{v=c+1}^{c+A} \frac{j_v}{n} \prod_{u=m_v+1}^{s} \frac{F_{i_u+1}}{n} \left( \frac{F_{i+1}}{n} \right)
$$

up to the discrepancy $D(\mathcal{P}_{(k_1, \ldots, k_r)})$. Then the expression (7.26) is bounded by

$$
\sum_{e_1, \ldots, e_s \in Q_n} \sum_{1 \leq l_{c+A+1} \leq \ldots \leq l_s \leq A + B} \left| T_{e_v} \right| \left( \frac{2T}{n^k} \prod_{u=1}^c \frac{k_i}{b} \prod_{v=c+1}^{c+A} \frac{j_v}{n} \prod_{u=m_v+1}^{s} \frac{F_{i_u+1}}{n} \right) + D(\mathcal{P}_{(k_1, \ldots, k_r)})
$$

$$
\prod_{v=c+1}^s \frac{1}{j_v} \left( \frac{m_1, u_{c+A+1}, \ldots, u_s}{k_1, \ldots, k_r} \right) + \varepsilon(n)
$$
We split (7.30) into two terms:

\[ A_1 := \sum_{e_1, \ldots, e_r \in Q_n} \sum_{1 \leq j_w \leq \frac{2T}{n^{k_w}} \sum_{u=1}^{c} k_u} \frac{F_{l,v}}{n^{c}} \prod_{w=c+1}^{s-t} \prod_{w=c+1}^{s-t} \left( f_{l,v}^{n, u_1 + \cdots + u_s} \left( \frac{k_1}{b}, \ldots, \frac{k_s}{b} \right) + \varepsilon(n) \right) \]  

and

\[ A_2 := \sum_{e_1, \ldots, e_r \in Q_n} \sum_{1 \leq j_w \leq \frac{2T}{n^{k_w}} \sum_{u=1}^{c} k_u} \frac{F_{l,v}}{n^{c}} \prod_{w=c+1}^{s-t} \prod_{w=c+1}^{s-t} \left( f_{l,v}^{n, u_1 + \cdots + u_s} \left( \frac{k_1}{b}, \ldots, \frac{k_s}{b} \right) + \varepsilon(n) \right) \]  

For the first term \( A_1 \), note that \( \frac{F_{l,v}}{n^{c}} \leq 2 \) and that \( \sum_{u=1}^{c} k_u \leq \frac{1}{11}, \ldots, c \). Then summing over \( e_u, k_i, j_w \) and \( l_v \), we see that

\[ A_1 \ll \frac{1}{n^{k_w}} \frac{1}{\frac{1}{11}, \ldots, c} \frac{1}{r^{c}} \]  

where of course the implicit constant does not depend on \( r \). (Here, if \( c = 0 \) then the \( \frac{1}{11, \ldots, c} \) term is omitted.)

For the second term \( A_2 \), note that each index \( l_v \) has a maximum range going from 1 to \( M^* \ll \log n \). Then using the bound on \( D(P_{(k_i), (l_i), (j_w)}) \) proved in Theorem 8.1 below, we have

\[ A_2 \ll \left( \frac{100s \log \log \log n}{\log \log n} \right)^{\log \log n} (\log n)^s \rightarrow 0 \]  

Putting this all together, the RHS of (7.30) is bounded by

\[ \ll \sum_{1 \leq l_1 < \cdots < l_c \leq r} \frac{1}{\frac{1}{11}, \ldots, c} \frac{1}{r^{c}} + \varepsilon(\lambda) \ll \frac{(\log r)^c}{r^{c}} + \varepsilon(\lambda) \]  

Taking \( \lambda \rightarrow 0 \) completes the proof. \( \square \)

8. Bound for Discrepancy

It remains to obtain a bound for the discrepancy \( D(P_{(k_i), (l_i), (j_w)}) \) of the set defined in (7.28). The following bound is not intended to be tight.

**Theorem 8.1.** For all choices of \( A, B, \) and \( C \) (such that \( C + c \geq 1 \)), the discrepancy satisfies the asymptotic bound given by

\[ D(P_{(k_i), (l_i), (j_w)}) \ll \left( \frac{100s \log \log \log n}{\log \log n} \right)^{\log \log n} \]  

uniformly over all \( j_w \in \mathcal{P}_{e_w} \) and factorization classes \( \mathcal{P}_{e_{1}, \ldots, e_r} \) and boxes given by the subscripts \( k_i \) and \( l_v \).
Recall that in the one-dimensional case, the discrepancy asymptotic from Proposition 3.1 played an important role in bounding the sum in (3.10). One can prove Proposition 3.1 by analyzing the corresponding exponential sums and using the Erdős-Turan inequality, a quantitative version of Weyl’s Criterion. We will also use this inequality to prove Theorem 8.1.

**Remark 8.1.** Throughout this section, we will use \( \varepsilon' \) to denote an arbitrary positive number in an asymptotic or bound. For example, writing \( f(n) = o(n^{\varepsilon'}) \) means that \( f(n) = o(n^{\varepsilon'}) \) for all \( \varepsilon' > 0 \).

For each \( 1 \leq j \leq s \), write \( \mathbf{e}_j = (e_{1j}, ..., e_{lj}) \). Let \( q_j = \prod_{i=1}^{l} p_i^\varepsilon_{ij} \). Then from the definition of \( Q_n \) in Definition 7.1

\[
q_j \ll (4s(\log n)^{2s}(\log \log n))^{2s(\log \log n)^3} = o(n^{\varepsilon'})
\]

First, we establish a lower bound on the cardinality of the multiset \( |P_{(k_i),(l_i),(j_w)}| \).

**Lemma 8.1.**

\[
|P_{(k_i),(l_i),(j_w)}| \geq \prod_{u=c+A}^{s} |T_{v} \cap \mathcal{P}_{e_u}| \prod_{u=1}^{c} |I_{k_{i_u}, b} \cap \mathcal{P}_{e_u}|
\]

**Proof.** By definition,

\[
|P_{(k_i),(l_i),(j_w)}| = \prod_{v=c+A}^{s} |T_{v} \cap \mathcal{P}_{e_v}| \prod_{u=1}^{c} |I_{k_{i_u}, b} \cap \mathcal{P}_{e_u}|
\]

Note that \( |T_{v} \cap \mathcal{P}_{e_v}| \) is the number of integers in the interval \( \frac{E_u}{q_v} \leq x_v \leq \frac{F_{l+1}}{q_v} \) that are not divisible by the primorial \( \prod_{m=1}^{[\log n]/2^s} p_m \). Counting the number of such integers corresponds to “sieving out” by primes \( p_{m} \) with \( m < (\log n)^{2s} \). By the “fundamental lemma of sieve theory” (c.f. [28 p. 60]), if \( w = y^{1/u} \) then the fraction of integers in \( [x, x+y] \) that are not divisible by any prime less than \( w \) is \( \sim \prod_{p < w} (1 - \frac{1}{p}) (1 + O(u^{-u})) \). As \( u \to \infty \), we get the expected asymptotic which by Mertens’s 3rd theorem (c.f. [28 p. 17]) is \( \frac{e^{-\gamma} + o(1)}{\log w} \) where \( w = (\log n)^{2s} \). For our purposes, any finite \( u \) suffices and we see that
By the same reasoning,
\[
\frac{|I_{k_{iu}, b} \cap \mathcal{P}_{e_u}|}{|I_{k_{iu}, b}|/q_u} \gg \frac{1}{\log \log n}
\]  
(8.8)

The bound then follows from (8.4).

Set \( G_m = (m - 1)n/\log n \). Then \( I_{m,b} = (G_m, G_{m+1}) \). Let \( \sum_j^u \) formally denote summation over the range
\[
\sum_{m_1 < \ldots < m_j < (\log n)^2 s \frac{G_{k_{iu}}}{p_{m_1} \ldots p_{m_j}}}^u \sum_{x_u} \left( \frac{G_{k_{iu}+1}}{p_{m_1} \ldots p_{m_j}} \right)
\]  
(8.9)

and let \( \sum_j^v \) formally denote summation over the range
\[
\sum_{m_1 < \ldots < m_j < (\log n)^2 s \frac{F_{k_{iv}}}{p_{m_1} \ldots p_{m_j}}}^v \sum_{x_v} \left( \frac{F_{k_{iv}+1}}{p_{m_1} \ldots p_{m_j}} \right)
\]  
(8.10)

By the inclusion-exclusion principle, (8.3) can be rewritten as an alternating sum over rectangular lattices:
\[
\frac{1}{|\mathcal{P}_{(k_i), (l_i), (j_u)}|} \prod_{u=1}^{c} \left( \sum_{j=0}^{(\log n)^2 s} (-1)^j \sum_{j}^u \right) \prod_{v=c+1}^{s} \left( \sum_{j=0}^{(\log n)^2 s} (-1)^j \sum_{j}^v \right)
\]
\[
\prod_{u=1}^{c} x_u \prod_{v=c+1}^{s} x_v + O(1)
\]  
(8.11)

where here the symbols \( \prod_{u=1}^{c} \) and \( \prod_{v=c+1}^{s} \) are to be interpreted as a formal product of the summation operators. By Bonferroni’s inequalities, we can truncate the alternating sums over \( j \) at some even index \( q \) to obtain upper bounds on the absolute value of (8.11). To be precise, for each even index \( q \) there exist multisets \( A^u_q \) for \( 1 \leq u \leq c \) such that we have the equality of summation ranges
\[
\sum_{j=0}^{q} (-1)^j \sum_{j}^u = \sum_{j=0}^{q-1} (-1)^j \sum_{j}^u + \sum_{x_u \in A^u_q}
\]  
(8.12)

where
\[
\sum_{x_u \in A^u_q} \sum_{1 \leq q \leq q} \sum_{1 \leq m_1 < \ldots < m_q < (\log n)^2 s} \left( \frac{|I_{k_{iu}, b}|/q_u}{p_{m_1} \ldots p_{m_q}} + O(1) \right)
\]  
(8.13)

Similarly, there exist multisets \( A^v_q \) for \( c + A + 1 \leq v \leq s \) with the analogous properties. For all \( q \) and \( r \), we have the inequality
\[
\sum_{1 \leq m_1 < \ldots < m_q \leq r} \frac{1}{p_{m_1} \ldots p_{m_q}} \leq \left( \frac{c}{q} \right)^q \left( \frac{r}{\sum_{i=1}^{q} p_{m_i}} \right)^q
\]  
(8.14)
For $q = 2\lfloor \log \log n \rfloor$ and $r = (\log n)^{2s}$, this gives a bound of $\ll \left( \frac{10s \log \log n}{\log \log n} \right)^{\log \log n}$. Thus, truncating at the index $q = 2\lfloor \log \log n \rfloor$, we can bound \((8.11)\) by
\[
\ll \frac{1}{|P_{\mathbb{Z}_{q_1}, ..., q_s}|} \prod_{u=1}^{c} \left( \sum_{j=0}^{2\lfloor \log \log n \rfloor - 1} (-1)^j \sum_{t}^{u} \right) \prod_{v=c+A+1}^{s} \left( \sum_{j=0}^{2\lfloor \log \log n \rfloor - 1} (-1)^j \sum_{t}^{v} \right) e \left( \frac{ht}{u} x_u q_u \prod_{v=c+A+1}^{s} x_v q_v \right) + E \tag{8.15}
\]
where
\[
E = \frac{10s \log \log n}{\log \log n} \log \log n \left( \sum_{v} \frac{|T_{I_v}|}{q_v T_{I_v} \cap \mathcal{F}_{e_v}} + \sum_{u} \frac{|I_{k_u,b}|}{q_v |I_{k_u,b} \cap \mathcal{F}_{e_v}|} \right) \tag{8.16}
\]
By \((8.7)\) and \((8.8)\),
\[
E \ll \log n \left( \frac{10s \log \log n}{\log \log n} \right)^{\log \log n} \ll \left( \frac{20s \log \log n}{\log \log n} \right)^{\log \log n} \tag{8.17}
\]
Expanding out the alternating inclusion-exclusion sum, we see that the exponential sum in \((8.15)\) can be written as the sum of
\[
\ll \left( \sum_{j=0}^{2\lfloor \log \log n \rfloor - 1} \left( \frac{\lfloor \log n \rfloor^{2s}}{j} \right) \right)^s \ll (2 \log n)^s (\log n)^{4s^2 (\log \log n)} = o(n^{s'}) \tag{8.18}
\]
terms where each term is an exponential sum of the form
\[
S_D = \sum_{(x_v, x_u) \in D} e \left( b \zeta \prod_{u=1}^{c} x_u \prod_{v=c+A+1}^{s} x_v \right) \tag{8.19}
\]
where the domain of summation is given by
\[
D := \prod_{v=c+A+1}^{s} \left[ \frac{F_v}{r_v}, \frac{F_{v+1}}{r_v} \right] \prod_{u=1}^{c} \left[ \frac{G_{k_u}}{r_u}, \frac{G_{k_u+1}}{r_u} \right] \tag{8.20}
\]
and $z = t \prod_{u=1}^{c} q_u \prod_{v=c+A+1}^{s} q_v$ is an integer. Here, by \((8.4)\) we see that $r_u = o(n^{s'})$ and $r_v = o(n^{s'})$ for $1 \leq u \leq c$ and $c + A + 1 \leq v \leq s$. Also, $z = o(n^{sA+s'})$ since $t = O(n^{sA})$. The domain of summation has cardinality satisfying
\[
\frac{1}{n^s} \sum_{v=c+A+1}^{s} |T_{k_v}| \prod_{u=1}^{c} |I_{k_u,b}| \ll |D| \leq \sum_{v=c+A+1}^{s} |T_{k_v}| \prod_{u=1}^{c} |I_{k_u,b}| \tag{8.21}
\]
Note that compared to \((8.3)\), the sum in \((8.19)\) is much simpler since the domain of summation is just a box of integer lattice points. We will use a multivariate version of Weyl’s inequality stated in Lemma \((8.4)\) to bound this sum. In [20, Lemma 2.1], Parsell gives a bound for the exponential sum when the range of summation is a box with equal side lengths. The proof goes through unchanged to give the inequality in Lemma \((8.2)\) below for boxes with varying side lengths. The basic idea is to apply coordinate-wise Weyl differencing (Cauchy’s inequality) to the polynomial until we are left with a linear polynomial which can be summed directly as a geometric series. Note that the advantage of Weyl’s method is its simplicity. It only gives a savings in the exponent that is exponentially small (but which is sufficient for our purposes). Much stronger bounds on multidimensional exponential sums with polynomial savings in the exponent can be attained by following Vinogradov’s mean value method (cf. [19]). Most recently, adapting Wooley’s method
of efficient congruencing in the 1-D case (see [21]), near-optimal bounds have been obtained (see [21]).

Let \( N := \prod_{i=1}^{d} [1, N_i] \) be a \( d \)-dimensional box of integer lattice points and let \( j_1, \ldots, j_d \) be non-negative integers such that \( j_1 + \ldots + j_d = j \). Let

\[
P(n_1, \ldots, n_d) = \alpha_{j_1, \ldots, j_d} n_1^{j_1} \ldots n_d^{j_d} + \sum_{\substack{i_1, \ldots, i_d \geq 0 \\sum_{i_1 + \ldots + i_d \leq j-1}}} \alpha_{i_1, \ldots, i_d} n_1^{i_1} \ldots n_d^{i_d} \quad (8.22)
\]

be a multivariable polynomial of degree \( j \). (The assumption of a unique highest degree term is simply for convenience.)

**Lemma 8.2** (Parssell). Let \( 1 \leq m \leq j-1 \) and \( i_1, \ldots, i_d \) be non-negative integers such that \( i_1 + \ldots + i_d = m \). Then

\[
\left| \sum_{n \in N} e(P(n)) \right| \ll \left( \prod_{i=1}^{d} N_i \right)^{2m-1} \left( \prod_{i=1}^{d} N_i^h \right)^{-1} \sum_{h_i \in (-N_i, N_i)} \left| \sum_{n \in B_h(N)} e(h_1 \ldots h_m P_{i_1, \ldots, i_d}(n, h)) \right|
\]

where \( B_h(N) \) is a box contained in \( N \) for each \( h = (h_1, \ldots, h_m) \) and \( P_{i_1, \ldots, i_d}(n, h) \) is a polynomial with leading term

\[
\alpha_{i_1, \ldots, i_d} \frac{j_1! \cdots j_d!}{(j_1 - i_1)! \cdots (j_d - i_d)!} n_1^{j_1 - i_1} \ldots n_d^{j_d - i_d}
\]

To apply Lemma 8.2 we will choose \( i_1, \ldots, i_d \) so that \( P_{i_1, \ldots, i_d}(n, h) \) is a linear polynomial in \( n \). Then we can evaluate the inner sum over \( n \in B_h(N) \) via the geometric series estimate

\[
\sum_{n=1}^{N} e(\lambda n) \ll \min(N, \|\lambda\|^{-1}) \quad (8.23)
\]

for any non-integer \( \lambda \) where \( \| \cdot \| \) denotes distance to the nearest integer. Since our goal is to obtain bounds that are uniform over polynomials with leading coefficient \( h z \alpha \), we will need the following result to address the outer sum over \( h \).

**Lemma 8.3.** Let \( p, q \) be integers such that \( |\alpha - p/q| \leq q^{-2} \) where \( (p, q) = 1 \) and \( q > 0 \). For any positive integers \( M, N, \) and \( m \), we have

\[
\sum_{i=1}^{M} \min(N, \|i m \alpha\|^{-1}) \ll (M + 1) (m N + q \log q) \quad (8.24)
\]

**Proof.** We divide the sum over \( M \) terms into \( O(M/q + 1) \) blocks of size \( q \) (plus one possibly partial block) where each block has the form

\[
\sum_{i=0}^{q-1} \min(N, \|(i_0 + i) m \alpha\|^{-1}) \quad (8.25)
\]

and \( i_0 \) is the first number in the block. Since \( |\alpha - p/q| \leq q^{-2} \), we have \( (i_0 + i) m \alpha = i_0 m \alpha + \text{imp} \) where \( \|u\| \leq m/q \). As \( i \) runs over the set \( [q] \), \text{imp} runs over the residues \( r \) \( \text{gcd}(m, q) \mod q \) for \( 1 \leq r \leq q/\text{gcd}(m, q) \) with multiplicity \( \text{gcd}(m, q) \). Let \( b \) be the nearest integer to \( qi_0 m \alpha \). Then \( \|(i_0 + i) m \alpha\| = \|((b + r \text{gcd}(m, q))/q + u) \| \) where \( \|u\| \leq 3m/q \). Note that there are \( O(m) \) integers \( i \) such that \( \|(i_0 + i) m \alpha\| \leq \|m/q\| \). Therefore, (8.25) is bounded by

\[
\ll m N + \sum_{r=1}^{q} \frac{q}{r} \ll m N + q \log q \quad (8.26)
\]

which proves the lemma.
Lemma 8.4. Let $p, q$ be integers such that $|\alpha - p/q| \leq q^{-2}$ where $(p, q) = 1$ and $q > 0$. Let $P(n) = m(n_1 - a_1)\ldots(n_d - a_d)$ for some integers $m$ and $a_1, \ldots, a_d$ where $m \geq 1$ and $d \geq 2$. Assume that $N_1 \geq \ldots \geq N_d$. Then

\[
\sum_{n \in \mathbb{N}} e(P(n)) \ll \left( \prod_{l=1}^{d} N_l \right)^{1 + \varepsilon} \left( \frac{m}{q} + \frac{1}{N_d} + \frac{m}{\prod_{l=1}^{d-1} N_l} + \frac{q}{\prod_{l=1}^{d} N_l} \right)^{2^{1-d}}
\]

where the absolute constant does not depend on $a_1, \ldots, a_d$.

Proof. Let $i_l = 1$ for $1 \leq l \leq d - 1$ and $i_d = 0$. Then by Lemma 8.2

\[
\left| \sum_{n \in \mathbb{N}} e(P(n)) \right|^{2^{d-1}} \ll \left( \prod_{l=1}^{d} N_l \right)^{2^{d-1}} \left( \prod_{l=1}^{d-1} N_l \right)^{-1} \sum_{h_i \in \{-N_i, N_i\}} \sum_{n \in B_h(N)} e(h_1\ldots h_{d-1}m\sigma_{N_d})
\]

Since the divisor function $\tau(n)$ satisfies $\tau(n) \ll n^{\varepsilon}$, the number of ways to write an integer $1 \leq n \leq \prod_{i=1}^{d-1} N_i$ as a product $h_1\ldots h_{d-1}$ is $\ll N_1^{\varepsilon}$. Using this fact and (8.23) and separating out the terms where some $h_i$ is, we have

\[
\left| \sum_{n \in \mathbb{N}} e(P(n)) \right|^{2^{d-1}} \ll \left( \prod_{l=1}^{d} N_l \right)^{2^{d-1} - 1} (N_{d-1})^{-1} \left( \prod_{l=1}^{d} N_l \right)^{2^{d-1} - 1 + \varepsilon} \sum_{1 \leq i \leq \prod_{l=1}^{d-1} N_l} \min(N_d, |ima|^{-1})
\]

Finally, by Lemma 8.3 this is

\[
\ll \left( \prod_{l=1}^{d} N_l \right)^{2^{d-1} + \varepsilon} \left( \frac{1}{N_{d-1}} + \frac{m}{q} + \frac{1}{N_d} + \frac{m}{\prod_{l=1}^{d-1} N_l} + \frac{q}{\prod_{l=1}^{d} N_l} \right)^{2^{1-d}} \quad (8.27)
\]

Raising to the power $2^{1-d}$ completes the proof.

We are now ready to obtain an estimate for $S_D$ from (8.19) and apply the Erdős-Turan inequality to prove the discrepancy bound.

Proof of Theorem 8.7. Let $d = s - A = B + C + c$ be the degree of the polynomial in $S_D$ with leading coefficient $h\alpha$. First assume $d > 1$. The successive continued fraction convergents of $\alpha$ satisfy

\[
\frac{1}{2q_i q_{i+1}} \leq |\alpha - p_i/q_i| \leq \frac{1}{q_i q_{i+1}} \quad (8.28)
\]

By the definition of irrationality measure, $|\alpha - p/q| > 1/\delta$ for all integers $p$ and $q$ sufficiently large. Therefore $q_{i+1} \leq q_i^{\alpha-1+\varepsilon}$ for $i$ sufficiently large.

Assume that $1 \leq h \leq n^\varepsilon$. Then $hz = o(n^{\varepsilon(A+2)})$. Recall that $C + c \geq 1$, and therefore

\[
\prod_{i=1}^{c} |T_{i_1}| \prod_{i=1}^{c} |I_{k_{i_1}}| \geq n^{\delta}. \quad \text{Choose $q$ to be a convergent such that} \quad n^{\varepsilon(A+3)} \leq q \leq n^{\varepsilon(A+3)},
\]

By Lemma 8.3

\[
S_D \ll \sum_{i=C+1}^{s} |T_{i_1}| \prod_{i=1}^{c} |I_{k_{i_1}}| n^{\varepsilon(A+3)} \left[ \frac{n^{\varepsilon(A+2)}}{n^{\varepsilon(A+3)}} + \frac{n^{\varepsilon(A+2)+\varepsilon}}{n^{\delta}} + \frac{n^{\varepsilon(A+3)+\varepsilon}}{n^{\delta}} \right]^{2^{1-d}} \quad (8.29)
\]
Then using (8.31), we can bound (8.15) by
\[
n^r \left( \frac{n^{A+2}}{n^A} + \frac{n^{A+2} \epsilon^r}{n^A} + \frac{n^{A+2} \epsilon^r}{n^A} \right) ^{2^d - 1} + \left( \frac{20 \log \log \log n}{\log n} \right) ^{\log \log n} \tag{8.30}
\]
Choosing \( m = n^r \) in the Erdös-Turan inequality (Proposition 8.1), we obtain the bound
\[
D(P^{(k_1), (k_2), (\omega)}) \ll \frac{1}{n^r} + \frac{20 \log \log \log n}{\log n} ^{\log \log n} \tag{8.31}
\]
for \( d > 1 \). If the degree \( d = 1 \), either \( c = 1 \) or \( C = 1 \) since \( C + c \geq 1 \). By (8.20), we have
\[
S_D \ll ||h\alpha||^{-1} \text{ where } z = o(n^{A+\epsilon}). \text{ Then}
\]
\[
\frac{n^r |S_D|}{|P^{(k_1), (k_2), (\omega)}|} \ll \frac{n^r}{n^d} |h\alpha| \tag{8.32}
\]
By [16] Lemma 2.3.3, for any integer \( m \)
\[
\sum _{h=1} ^m \frac{1}{h ||h\alpha||} \leq \sum _{h=1} ^{m^z} \frac{1}{h ||h\alpha||} \ll (m^z)^{\mu - 2 + \epsilon} \tag{8.33}
\]
Then choosing \( m = n^r \) in the Erdös-Turan inequality as before completes the proof. \( \square \)

9. Eigenvalue process at rational angles

To obtain the limiting eigenvalue point process \( \lim \mathcal{E}_{n, k} ^a \) at a rational angle \( \alpha = s/t \), we need the following distributional convergence analogous to Lemma 6.2 for irrational \( \alpha \).

**Lemma 9.1.** Let \( \alpha = s/t \) be a rational number in reduced form where \( t = \prod _{m=1} ^M p_m ^{e_m(t)} \). Fix positive integers \( r \) and \( l \) such that \( r \geq k \) and \( l \geq M \). For \( 1 \leq i_1 < \ldots < i_k \leq r \), define \( \mathbb{Z}/t\mathbb{Z} \)-valued exchangeable random variables
\[
V_{i_1, \ldots, i_k} ^l = \prod _{m=1} ^M \frac{x_{m_i_1} + \ldots + x_{m_i_k}}{g_k (\mathbf{X} _{i_1, \ldots, i_k})} U_{i_1, \ldots, i_k} \tag{9.1}
\]
where \( U_1, U_2, \ldots \) are i.i.d. random variables distributed uniformly on the group of units \( \mathbb{Z}/t\mathbb{Z} ^\times \).

Let \( U_{i_1, \ldots, i_k} ^l = V_{i_1, \ldots, i_k} ^l / t \in \mathbb{Q} \) where we identify \( \mathbb{Z}/t\mathbb{Z} \) with \( \{t \} \subset \mathbb{Z} \).

Then we have the distributional convergence
\[
\left( E_l (L_1 ^{(n)}, \ldots, L_r ^{(n)}), L_1 ^{(n)} / n, \ldots, L_r ^{(n)} / n, \left( \psi _\alpha \left( \frac{L_1 ^{(n)} \ldots L_r ^{(n)}}{g_k (\mathbf{E}_l (L_1 ^{(n)}, \ldots, L_r ^{(n)}))} \right) \right) \right) \quad \xrightarrow{d} \quad \left( \mathbf{X} _{i_1, \ldots, i_k} ^{l-r}, L_1, \ldots, L_r, (U_{i_1, \ldots, i_k} ^l) \right)
\]

**Proof.** Following the proof of Lemma 6.2 fix some matrix \( A \in \mathbb{N} ^{l \times r} \) and corresponding factorization class \( \mathcal{P}_A \). The analog of computing the discrepancy \( D(\mathcal{P}_A ^{\mathcal{P}_A, \alpha}) \) in Lemma 6.2 is finding the distribution of \( \psi _\alpha \left( \frac{m_{i_1} \ldots m_{i_k}}{g_k (\mathbf{E}_l (m_{i_1}, \ldots, m_{i_k}))} \right) \) such that \( (m_1, \ldots, m_r) \) are chosen uniformly from the set
\[
\{(m_1, \ldots, m_r) : (\mathbf{E}_l (m_1, \ldots, m_r) = A) \wedge \bigwedge _{i=1} ^r \{m_i \in I_{k_i} ^i \} \} \tag{9.2}
\]
It is sufficient to find the distribution of \( \left( \frac{m_{i_1} \cdots m_{i_k}}{g_k(\mathbf{E} [m_{i_1} \cdots m_{i_k}])} \right) \mod t \) and then apply the function \( \psi_\alpha \).

For \( 1 \leq i \leq r \), let \( N_i = \prod_{m=1}^{l} p_m^{A_{m+1} + e_m(i)} \) and let \( s_i \) be an element of \( \mathcal{P}_{a_i} \cap [N_i] \). Define the sequences \( s^i(q) = s_i + qN_i \) where \( q \) runs over the non-negative integers. Note that \( s^i(q) \in \mathcal{P}_{a_i} \) for all \( q \). Thus, we can partition each factorization class \( \mathcal{P}_{a_i} \) into congruence classes modulo \( N_i \). Also, for \( 1 \leq i_1 < \ldots < i_k \leq r \), all elements in \( \left\{ s^{i_1}(q_{i_1}) \ldots s^{i_k}(q_{i_k}) : q_{i_j} \in \mathbb{N}\right\} \) lie in the same congruence class mod \( t \). Thus, it is sufficient to restrict our attention to \( (m_1, \ldots, m_r) \in [N_1] \times \ldots \times [N_r] \) and compute the distribution of the multiset

\[
V^A := \left\{ \left( \frac{m_{i_1} \cdots m_{i_k}}{g_k(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k})} \right) \in (\mathbb{Z}/t\mathbb{Z})^{\mathbb{N}} : m_{i_j} \in \mathbb{Z}/N_i \mathbb{Z} \cap \mathcal{P}_{a_j} \right\}
\]

(9.3)

Here, \( V^A \) is well-defined since there is a natural projection map \( \mathbb{Z}/N_i \mathbb{Z} \to \mathbb{Z}/t\mathbb{Z} \). By the Chinese remainder theorem,

\[
\mathbb{Z}/N_i \mathbb{Z} \cong \prod_{m=1}^{l} \mathbb{Z}/p_m^{A_{m+1} + e_m(i)} \mathbb{Z}
\]

Then a little modular arithmetic shows that the \( V^A \) can be written as

\[
V^A_{i_1, \ldots, i_k} = \prod_{m=1}^{M} \frac{X_{m_{i_1} + \ldots + m_{i_k}}}{g_k(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k})} U_{i_1} \ldots U_{i_k}
\]

(9.4)

where \( U_1, \ldots, U_r \) are i.i.d. random variables distributed uniformly on \( (\mathbb{Z}/t\mathbb{Z})^\times \). Here, we use the fact that \( p_i^{A_{i_j}} \) are invertible in \( \mathbb{Z}/t\mathbb{Z} \) for \( l > M \) and hence can be absorbed into the random variable \( U_j \) by translation invariance. The rest of the proof follows as for Lemma 6.2 \( \square \)

**Corollary 9.1.** Recall the notation from Lemma 9.1 and as usual, allow the multi-dimensional indices \( (i_1, \ldots, i_k) \) to run over the range \( 1 \leq i_1 < \ldots < i_k \leq r \). Define \( \mathbb{Z}/t\mathbb{Z} \)-valued exchangeable random variables

\[
V_{i_1, \ldots, i_k} = \prod_{m=1}^{M} \frac{X_{m_{i_1} + \ldots + m_{i_k}}}{g_k(\mathbf{X}_{\mathbb{N}}, \ldots, \mathbf{X}_{\mathbb{N}})} U_{i_1} \ldots U_{i_k}
\]

where \( U_1, U_2, \ldots \) are i.i.d. random variables distributed uniformly on \( (\mathbb{Z}/t\mathbb{Z})^\times \). Let \( U_{i_1, \ldots, i_k} = V_{i_1, \ldots, i_k}/t \in \mathbb{Q} \) where we identify \( \mathbb{Z}/t\mathbb{Z} \) with \( \{ t \} \subset \mathbb{Z} \). Then we have the distributional convergence

\[
\left( \left( \frac{L_{i_1}^{(n)} \ldots L_{i_k}^{(n)}}{\text{lcm}(L_{i_1}^{(n)} \ldots L_{i_k}^{(n)})} \right), \left( L_{i_1}^{(n)} \ldots L_{i_k}^{(n)} / n^k \right), \left( \psi_\alpha(\text{lcm}(L_{i_1}^{(n)} \ldots L_{i_k}^{(n)})) \right) \right) \to_d \left( (g_k(\mathbf{X}_{\mathbb{N}}^{i_1, \ldots, i_k})), (L_{i_1} \ldots L_{i_k}), (U_{i_1, \ldots, i_k}) \right)
\]

**Proof.** This follows from the continuous mapping theorem and Lemma 6.1 \( \square \)

**Theorem 9.1.** Let \( \alpha = s/t \) be a rational number in reduced form where \( t = \prod_{m=1}^{M} p_m^{e_m(i)} \). For \( i_1 < \ldots < i_k \), let \( U_{i_1, \ldots, i_k} \) be as defined in Corollary 9.1. Define the process

\[
\mathcal{E}_{\infty, k}^t = k! \sum_{i_1 < \ldots < i_k} g_k(\mathbf{X}_{\infty}^{i_1, \ldots, i_k}) \sum_{q \in \mathbb{Z}} \delta(q + U_{i_1, \ldots, i_k}) g_k(\mathbf{X}_{\mathbb{N}}^{i_1, \ldots, i_k}) / (L_{i_1} \ldots L_{i_k})
\]

Then \( \mathcal{E}_{\infty, k}^\alpha \to \mathcal{E}_{\infty, k}^t \) weakly in the sense that we have the weak convergence of \( \mathcal{R} \)-valued random variables

\[
\int f d\mathcal{E}_{\infty, k}^\alpha \to \int f d\mathcal{E}_{\infty, k}^t
data all continuous, compactly supported functions \( f \).
Proof. Let \( f \) have support contained in the interval \((-T, T)\) and define \( J_k(n) \) as in (6.10). Just as in equation (4.1), we have the set equality

\[
J_k \cap [n^k/(tT)] = t\mathbb{Z} \cap [n^k/(tT)].
\] (9.5)

If \( j \in t\mathbb{Z} \cap [n^k/(tT)] \), the interval \((\alpha - T/n^k, \alpha + T/n^k)\) contains exactly one eigenangle corresponding to the \( \sigma_k \)-cycle of length \( j \) (namely \( \alpha \)).

Write \( \xi^T_{n,k} = \xi^T_{n,k} + \eta^T_{n,k} \) where \( \xi^T_{n,k} \) and \( \eta^T_{n,k} \) are as defined in (6.3). Then

\[
\int f \xi^T_{n,k} = k! \sum_{1 \leq i_1 < \ldots < i_k \leq tT} \frac{L^{(n)}_{i_1} \ldots L^{(n)}_{i_k}}{\text{lcm}(L^{(n)}_{i_1}, \ldots, L^{(n)}_{i_k})} \sum_{q \in \mathbb{Z}} f(n^k(q/\text{lcm}(L^{(n)}_{i_1}, \ldots, L^{(n)}_{i_k}) - \alpha))
\] (9.6)

If \( r > tT \), then \( L^{(n)}_r < n/(tT) \) and therefore \( \text{lcm}(L^{(n)}_{i_1}, L^{(n)}_{i_2}, \ldots, L^{(n)}_{i_k}) < n^k/(tT) \). Then if \( f(0) \geq 0 \),

\[
f(0) \left( \sum_{1 \leq j \leq n^k/(tT)} \frac{C^{(n)}_{j,k}}{\binom{[tT]}{k}} \right) \leq \int fd\eta^T_{n,k} \leq f(0) \sum_{1 \leq j \leq n^k/(tT)} \frac{C^{(n)}_{j,k}}{\binom{[tT]}{k}}
\] (9.7)

Let \( \xi^T_{\infty,k} = \xi^T_{\infty,k} + \eta^T_{\infty,k} \) where \( \xi^T_{\infty,k} \) and \( \eta^T_{\infty,k} \) are defined as in (6.4) with \( \xi^T_{\infty,k} \) in place of \( \xi^T_{\infty,k} \) and with the new definition of the variables \( U_{i_1, \ldots, i_k} \). Then

\[
\int f \xi^T_{\infty,k} = k! \sum_{1 \leq i_1 < \ldots < i_k \leq tT} g_k(X^{i_1, \ldots, i_k}_\infty) \sum_{q \in \mathbb{Z}} f((U_{i_1, \ldots, i_k} + q)g_k(X^{i_1, \ldots, i_k}_\infty)/(L_{i_1} \ldots L_{i_k}))
\] (9.8)

and since \( L_r < 1/(tT) \) for \( r > tT \), we have

\[
\int fd\eta^T_{\infty,k} = k! f(0) \sum_{1 \leq i_1 < \ldots < i_k \leq tT} g_k(X^{i_1, \ldots, i_k}_\infty) \mathbb{1}(U_{i_1, \ldots, i_k} = 1)
\] (9.9)

If \( f(0) > 0 \), then \( \int fd\xi^T_{\infty,k} \) converges weakly to the \( \mathbb{F} \)-valued random variable \( \int fd\xi^T_{\infty,k} \), which is infinite almost surely by the Borel-Cantelli lemma.

Now we can assume \( f(0) = 0 \). By Corollary 9.1 and the continuous mapping theorem,

\[
\int f \xi^T_{n,k} \overset{d}{\to} \int f \xi^T_{\infty,k}
\]

which proves the theorem. \( \square \)

Remark 9.1. At rational \( \alpha = s/t \), the corresponding eigenvalue point processes for the \( k \)-subset and \( S^{(n-k,k^k)} \) representations both converge weakly to \( \frac{1}{k!} \xi^T_{\infty,k} \) by the reasoning in Remark 6.2.

10. Power series representation for gap probability

Corollary 10.1 gives a formula for the limiting eigenvalue gap probability \( P_k^\theta(y_1, y_2) \) for irrational \( \alpha \) with finite irrationality measure. In this section, we give a procedure to compute the expectation explicitly when \( y_2 - y_1 \leq k^k \) and \( \theta = 1 \). To simplify the resulting combinatorics, it is convenient to borrow some terminology from graph theory. The following definitions introduce basic notions about hypergraphs.

Definition 10.1. A hypergraph is a generalization of a graph in which an edge can connect any number of vertices. Formally, a hypergraph \( G = (V, E) \) where \( V \) is a set of vertices and \( E \) is a multiset of non-empty subsets of \( V \). A subset of \( V \) of size \( j \) is called a \( j \)-edge. Since \( E \) is a multiset, multi-edges or multiple edges on the same set of vertices in hypergraphs are allowed. A hypergraph isomorphism between \( (V_1, E_1) \) and \( (V_2, E_2) \) is a bijection between vertex sets \( V_1 \) and \( V_2 \) that respects the edge multisets \( E_1 \) and \( E_2 \). Given a vertex subset \( W = \{v_1, \ldots, v_{|W|}\} \subseteq V \),
we define the induced subhypergraph $G^W = (W, \{ e \cap W : e \in E(G), e \cap W \neq \emptyset \})$. Given two hypergraphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, the union is given by $G_1 \cup G_2 = (V_1 \cup V_2, E_1 \cup E_2)$. (Note that $V_1 \cup V_2$ is a union of sets while $E_1 \cup E_2$ is a union of multisets).

**Definition 10.2.** A labeled hypergraph is a hypergraph whose vertices have been assigned distinct labels in $\mathbb{N}$. We define (nonstandard terminology) a label isomorphism, $\phi : G_1 \rightarrow G_2$, between labeled hypergraphs $G_1$ and $G_2$ to be a hypergraph isomorphism that respects ordering of the vertices. In other words, if we have two vertices $a, b \in V(G_1)$ such that $a < b$, then $\phi(a) < \phi(b)$. This is denoted by $G_1 \cong^l G_2$.

**Definition 10.3.** The $j$-degree of a vertex of a hypergraph is the total number of $j$-edges incident to the vertex. The unlabeled $j$-degree sequence is the non-decreasing sequence of its vertex $j$-degrees. The labeled $j$-degree sequence is the list of vertex $j$-degrees by labelling.

**Definition 10.4.** Let $\mathcal{B}_m^j$ be the set of labelled hypergraphs whose edge set consists of $m$ distinct $j$-edges and which do not have isolated vertices, i.e. every vertex lies in some edge $e$. Let $\mathcal{A}_m^j \subseteq \mathcal{B}_m^j$ be the finite subset of labelled hypergraphs whose vertex label set is of the form $\{1, 2, ..., r\}$ for some $r$.

We can think of the random variables $L_1, L_2, ..., L_r$ as vertices of a labeled hypergraph $G$ and each product $\prod_{n=1}^k L_{i_n}$ as a $k$-edge containing the vertices $L_{i_1}, ..., L_{i_k}$. Then each term in the product expansion of (1.17) corresponds to a unique labelled hypergraph $G \in \mathcal{B}_m^k$ for some $m$.

For $G \in \mathcal{B}_m^k$, let

$$\psi(G) = E \left[ \prod_{(i_1, \ldots, i_k) \in E(G)} \frac{1}{g_k(x_{i_1}, \ldots, x_{i_k})} \right]$$

(10.1)

Note that if $H \in \mathcal{B}_m^k$ is isomorphic to $G$, then $\psi(H) = \psi(G)$ since $\psi(H)$ only depends on the isomorphism class of $H$ (doesn’t care about the labelling of the vertices).

For a hypergraph $H = (V, E) \in \mathcal{B}_m^k$ with labeled $k$-degree sequence $(d_1, ..., d_{|V|})$, define

$$L^H := L_1^{d_1} L_2^{d_2} ... L_{|V|}^{d_{|V|}}$$

(10.2)

With this notation, it is easy to see the following:

**Lemma 10.1.** Let $y_2 - y_1 \leq k^k$. Write $P^0(y_1, y_2) = f(y_1 - y_2)$ where $f(x) = \sum_{m=0}^\infty c_m x^m$. The coefficient for the $m^{th}$ term of the power series representation is given by

$$c_m = \sum_{H \in \mathcal{B}_m^k} \psi(H) E[L^H] = \sum_{G \in \mathcal{A}_m^k} \psi(G) \sum_{H \sim^l G} E[L^H]$$

(10.3)

Let us first compute $\sum_{H \sim^l G} E[L^H]$ for each $G \in \mathcal{A}_m^k$. We’ve essentially already seen how to calculate this sum in Section 3. Let $i_1, ..., i_v$ be the vertices of $H$ with $k$-degrees $d_1, ..., d_v$. If $\theta = 1$, then from the moments formula (3.14) we have

$$E[L^{d_{i_1}}_{i_1} L^{d_{i_2}}_{i_2} ... L^{d_{i_v}}_{i_v}] = \frac{d!}{d_{i_1}! d_{i_2}! ... d_{i_v}!} \int_0^\infty \int_0^\infty \ldots \int_0^\infty \frac{E(x_1)^{i_1-1}}{(i_1-1)!} \frac{(E(x_2) - E(x_1))^{i_2-1}}{(i_2 - i_1 - 1)!} \ldots \frac{(E(x_v) - E(x_{v-1}))^{i_v-1}}{(i_v - i_{v-1} - 1)!} d_{i_1}! d_{i_2}! ... d_{i_v}!$$
By (3.11),
\[ \sum_{H \subseteq G} \mathbb{E}[L^H] = \sum_{i_1 < \cdots < i_v} \mathbb{E}[L_{i_1}^{d_1} L_{i_2}^{d_2} \cdots L_{i_v}^{d_v}] \]
\[ = \frac{1}{d!} \int_0^\infty \int_{x_v} \cdots \int_{x_1} \frac{e^{-s} \prod e^{-x_1} \cdots e^{-x_v} dx_1 \cdots dx_v}{s^\infty} \] (10.4)

This is an elementary integral and evaluates to a rational number for any choice of \( d_1, \ldots, d_v \).
One can compute the integrals successively by making use of the identity
\[ \Gamma(s, x) = (s - 1) e^{-x} \sum_{j=0}^{s-1} \frac{x^j}{j!} \] (10.5)

for positive integers \( s \) where \( \Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt \) is the incomplete Gamma function.

We now give a procedure to compute \( \psi(G) \) for each \( G \in \mathcal{H}_k^m \). First, for an arbitrary hypergraph \( G \) with vertex set \( V = \{v_1, \ldots, v_{|V|}\} \), define
\[ S_G^p = \sum_{e_1, \ldots, e_{|V|} \in \mathcal{E}(G)} \prod_{i=1}^{p} \frac{1}{p^{e_{v_i}} (v_1, \ldots, v_{|V|}) \in E(G)} \frac{1}{p^{\max(e_{v_1}, \ldots, e_{v_{|V|}})}} \] (10.6)

Then for \( G \in \mathcal{H}_k^m \), summing over \( A \in \mathbb{N}^{\times |V|} \), we have
\[ \psi(G) = \mathbb{E} \left[ \prod_{(i_1, \ldots, i_k) \in \mathcal{E}(G)} \frac{1}{g_k(X_{i_1} \ldots, i_k)} \right] \]
\[ = \lim_{l \to \infty} \sum_A \left( \prod_{m=1}^l \left( 1 - \frac{1}{p^m} \right) \mathbb{E} \left[ \prod_{i=1}^{\max(e_{i_1}, \ldots, e_{i_k})} \frac{1}{p^{e_{v_i}} (i_1, \ldots, i_k) \in E(G)} \right] \frac{1}{g_k(a_{i_1}, \ldots, a_{i_k})} \right) \]
\[ = \prod_p \left( 1 - \frac{1}{p} \right)^{|V|} \sum_{e_1, \ldots, e_{|V|} \in \mathcal{E}(G)} \prod_{i=1}^{\max(e_{i_1}, \ldots, e_{i_k})} \frac{1}{p^{e_{v_i}} (v_1, \ldots, v_{|V|}) \in E(G)} \frac{1}{p^{\max(e_{v_1}, \ldots, e_{v_{|V|}})}} \]
\[ = \prod_p \left( 1 - \frac{1}{p} \right)^{|V|} S_G^p \] (10.7)

For any hypergraph \( G \), we have the following recursive formula for computing \( S_G^p \):

**Lemma 10.2.** Let \( G = (V, E) \) be a hypergraph and \( U \subset V \) be a vertex subset. Then
\[ S_G^p = \left( 1 - \frac{1}{p^{|V| + \sum (j-1)|E_j^U|}} \right)^{-1} \sum_{U \subseteq V} \frac{1}{p^{|U| + \sum (j-1)|E_j^U|}} S_{GU}^p \] (10.8)

where \( E_j^U \) is the \( j \)-edge set of the subhypergraph \( G^U \) induced by the vertex subset \( U \subset V \) (counting multiedges). The base case is \( S_G = 1 \).

**Proof.** Partition the domain of summation \((e_{v_1}, \ldots, e_{v_{|V|}}) \in \mathbb{N}^{\times |V|}\) of \( S_G^p \) into blocks \( B_c \) such that \( \min(e_{v_1}, \ldots, e_{v_{|V|}}) = c \). In each block \( B_c \), we “factor out” \( p^c \) from each vertex, meaning we take out a total factor of
\[ \frac{1}{p^{|V| + \sum (j-1)|E_j^U|}} \prod_{(v_1, \ldots, v_{|V|}) \in E(G)} \frac{1}{p^{e_{v_i}}} \] (10.9)

from
\[ \sum_{(e_{v_1}, \ldots, e_{v_{|V|}}) \in B_c} \prod_{i=1}^{\max(e_{v_1}, \ldots, e_{v_{|V|}})} \frac{1}{p^{e_{v_i}} (v_1, \ldots, v_{|V|}) \in E(G)} \frac{1}{p^{\max(e_{v_1}, \ldots, e_{v_{|V|}})}} \] (10.10)
Summing over all \( c \in \mathbb{N} \) yields the \( \left( 1 - \frac{1}{p^{n+\left(k_{\text{max}}\right)}} \right)^{-1} \) term. After this factoring, at least one vertex will be left with an exponent \( e_{v_i} = 0 \). Thus, we are reduced to summing over all subhypergraphs \( G^U \) with \( U = (v_1, \ldots, v_{\left| V \right|}) \subseteq V \) and exponents \((e_{v_1}, \ldots, e_{v_{\left| V \right|}})\) such that \( e_{v_i} > 0 \). In each subhypergraph \( G^U \), we have room to factor out at least one factor of \( p \) from each vertex. Putting this all together yields the formula.

With Lemma \( \text{10.2} \), we now have a procedure to compute \( \psi(G) = \prod_p \left( 1 - \frac{1}{p} \right)^{\left| V \right|} S_G^p \) for all \( G \in \mathcal{B}_k^m \).

The following examples tabulate \( S_G^p \) for a few small hypergraphs \( G \). For ease of notation, we leave off the superscript \( p \) and set \( q = 1/p \). Note that if the hypergraph \( G \) is not connected, then this recursive formula should be applied on each component since if \( G = G_1 \cup G_2 \) where \( G_1 \) and \( G_2 \) are disjoint, then \( S_G = S_{G_1} S_{G_2} \) and \( \psi(G) = \psi(G_1) \psi(G_2) \).

**Example 10.1** (Complete (nonhyper) graphs).

\[
S_{K_n} = \left( 1 - \frac{1}{p^{n+\left(\frac{k}{2}\right)}} \right)^{-1} \sum_{W \subseteq V} \frac{1}{p^{\left| W \right| + \left| E^W \right|}} S_{G^W} = \left( 1 - \frac{1}{p^{n+\left(\frac{k}{2}\right)}} \right)^{-1} \left( 1 + \sum_{i=1}^{n-1} \binom{n}{i} p^{-i(k+1)} S_{K_i} \right)
\]

\[
S_{K_1} = \frac{1}{1-q} \\
S_{K_2} = \frac{1+q}{(1-q)(1-q^3)} \\
S_{K_3} = \frac{1+2q+2q^3+q^4}{(1-q)(1-q^3)(1-q^6)} \\
S_{K_4} = \frac{1+3q+5q^3+3q^4+3q^6+5q^7+3q^9+q^{10}}{(1-q)(1-q^3)(1-q^6)(1-q^{10})}
\]

**Example 10.2** (Hypergraphs \( G_k \in \mathcal{B}_k^m \)).

\[
S_{G_k} = \left( 1 - \frac{1}{p^{2k-1}} \right)^{-1} \left( 1 + \sum_{i=1}^{k-1} \binom{k}{i} \frac{1}{p^{2i-1}} S_{G_i} \right)
\]

\[
S_{G_1} = \frac{1}{1-q} \\
S_{G_2} = \frac{1+q}{(1-q)(1-q^3)} \\
S_{G_3} = \frac{1+2q+2q^3+q^4}{(1-q)(1-q^3)(1-q^5)} \\
S_{G_4} = \frac{1+3q+5q^3+3q^4+3q^6+5q^7+3q^9+q^{10}}{(1-q)(1-q^3)(1-q^5)(1-q^7)}
\]

**Example 10.3** (\( H_1 \cup H_2 \) where \( H_1 \in \mathcal{B}_{h_1}^h \) and \( H_2 \in \mathcal{B}_{h_2}^h \)). Any such hypergraph \( G = H_1 \cup H_2 \) is characterized by three numbers: \( h_1, h_2, \) and \( h_3 := |V(H_1) \cap V(H_2)| \). We denote the corresponding sum \( S_G \) by \( S_{h_1,h_2,h_3} \). Note that if the vertex sets \( V(H_1) \) and \( V(H_2) \) are disjoint, \( S_{h_1,h_2,0} = S_{h_1} S_{h_2} \).

\[
S_{1,1,0} = (S_{G_1})^2 = \frac{1}{(1-q)^2}
\]
Proof. Applying Lemma 10.1, we first compute the term of the power series, there is only one graph $G$ for $\psi$. These give

$$S_{2,1,0} = (S_G_2)(S_G_1) = \frac{1 + q}{(1 - q)^2(1 - q^3)}$$

$$S_{2,1,1} = S_G_2 = \frac{1 + q}{(1 - q)(1 - q^3)}$$

$$S_{2,2,0} = (S_G_2)^2 = \frac{(1 + q)^2}{(1 - q)^2(1 - q^3)^2}$$

$$S_{2,2,1} = \frac{1}{1 - q^9}(1 + 3qS_G_1 + 2q^3S_G_2 + q^2(S_G_1)^2)$$

$$= \frac{1}{1 - q^9} \left( 1 + 3 \frac{q}{1 - q} + 2q^3 \frac{1 + q}{(1 - q)(1 - q^3)} + \frac{q^2}{(1 - q)^2} \right)$$

$$S_{2,2,2} = \frac{1}{1 - q^9}(1 + 2qS_G_1) = \frac{1 - q^2}{(1 - q)^2(1 - q^4)}$$

This process can be continued as far as desired.

Out of curiosity, we record the first few terms of the power series $P_2^1(y_1, y_2)$ when $k = 2$.

Corollary 10.1. For $y_2 - y_1 \leq 2^2 = 4$, we have

$$P_2^1(y_1, y_2) \approx 1 - 0.18269(y_2 - y_1) + 0.01448(y_2 - y_1)^2 + O((y_2 - y_1)^3) \quad (10.11)$$

Proof. Applying Lemma 10.1, we first compute $\psi(G)$ for $G \in A_2^1$ and $G \in A_2^2$. For the linear term of the power series, there is only one graph $G_2 \in A_2^1$ and

$$\psi(G_2) = \prod_p (1 - q)^2 \frac{1 + q}{(1 - q)(1 - q^3)} = \prod_p \frac{1 - q^2}{1 - q^3} = \frac{\zeta(3)}{\zeta(2)}$$

For the quadratic term, there are 2 nonisomorphic graphs $H_1$ and $H_2$ (ignoring labelling) in $A_2^2$. These give

$$\psi(H_1) = \prod_p (1 - q)^3 S_{2,2,1} = \prod_p (1 - q)^3 \frac{1}{1 - q^9} \left( 1 + 3 \frac{q}{1 - q} + 2q^3 \frac{1 + q}{(1 - q)(1 - q^3)} + \frac{q^2}{(1 - q)^2} \right)$$

$$= \prod_p (1 - q)^2(1 + 2q + q^2 + 2q^3 + q^4) \approx 0.561356$$

$$\psi(H_2) = (\psi(G_2))^2 = \left( \frac{\zeta(3)}{\zeta(2)} \right)^2 \approx 0.534015$$

Now let us compute $\sum_{H \cong G} E[L^H]$ for $G$ in $A_2^1$ or $A_2^2$ using (10.2).

For the linear term, $\sum_{H \cong G_2} E[L^H] = \sum_{i_1 < i_2} E[L_{i_1} L_{i_2}] = \left( \frac{1}{2!} \right)^2$

For the quadratic term,

$$\sum_{i_1 < i_2 < i_3} E[L_{i_1}^2 L_{i_2} L_{i_3}] = \frac{11}{864}$$

$$\sum_{i_1 < i_2 < i_3} E[L_{i_1} L_{i_2}^2 L_{i_3}] = \frac{5}{864}$$
\[ \sum_{i_1 < i_2 < i_3} E[L_{i_1} L_{i_2} L_{i_3}^2] = \frac{1}{432} \]
\[ \sum_{i_1 < i_2 < i_3 < i_4} E[L_{i_1} L_{i_2} L_{i_3} L_{i_4}] = \left( \frac{1}{4!} \right)^2 \]

Putting this together, the coefficient in front of the \((y_1 - y_2)\) term is \(\frac{\zeta(3)}{4 \zeta(2)} \approx 0.18269\).

The coefficient in front of the \((y_1 - y_2)^2\) term is
\[
\sum_{G \in A^2} \psi(G) \sum_{H \cong H} E[L^H] = \psi(H_1) \left( \frac{11}{864} + \frac{5}{864} + \frac{1}{432} \right) + 3 \psi(H_2) \left( \frac{1}{4!} \right)^2
\]
\[
\approx 0.561356 \left( \frac{11}{864} + \frac{5}{864} + \frac{1}{432} \right) + 3 \cdot 0.534015 \cdot \frac{1}{576} \approx 0.01448
\]

\[ \square \]

Acknowledgements
The author wishes to thank his PhD advisor Steven Evans for helpful discussions and comments on this work.

References
[1] R. Arratia, A. Barbour, and S. Tavare. Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.*, 2(3):519–535, 1992.
[2] R. Arratia, A. Barbour, and S. Tavare. *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society, Zurich, 2003.
[3] V. Bahier. On a limiting point process related to modified permutation matrices. arXiv:1803.03546, 2018.
[4] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Wiley-Interscience, 2 edition, 1999.
[5] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume I*. Probability and Its Applications. Springer-Verlag, 2 edition, 2003.
[6] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume II*. Probability and Its Applications. Springer-Verlag, 2 edition, 2008.
[7] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31A:49–62, 1994.
[8] R. Durrett. *Probability: Theory and Examples*. Duxbury Press, 4th edition, 2005.
[9] S. Evans. Eigenvalues of random wreath products. *Electron. J. Probab.*, 7(9):1–15, 2002.
[10] W. Ewens. The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, 3:87–112, 1972.
[11] S. Feng. *The Poisson-Dirichlet Distribution and Related Topics*. Probability and Its Applications. Springer-Verlag, 2010.
[12] R. Griffiths. On the distribution of allele frequencies in a diffusion model. *Theoretical Population Biology*, 15:140–158, 1979.
[13] G. Hardy and E. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, 1959.
[14] A. Hildebrand and G. Tenenbaum. On the number of prime factors of an integer. *Duke Mathematical Journal*, 56(3):471–501, 1988.
[15] J. Kingman. The population structure associated with the Ewens sampling formula. *Theoretical Population Biology*, 11:274–283, 1977.
[16] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences*. Dover Books on Mathematics. Dover Publications, 2006.
[17] B. Mehrdad and Lingjiong Zhu. Moderate and large deviations for Erdős-Kac theorem. *The Quarterly Journal of Mathematics*, 67:147–160, 2016.
[18] J. Najnudel and A. Nikeghbali. The distribution of eigenvalues of randomized permutation matrices. *Annales de L’Institut Fourier*, 63(3):773–838, 2013.
[19] S. T. Parsell. A generalization of Vinogradov’s mean value theorem. *Proc. London Math. Soc.*, 91:132, 2005.
[20] S. T. Parsell. Hua-type iteration for multidimensional Weyl sums. *Mathematika*, 58:209224, 2012.
[21] S.T. Parsell, S.M. Prendiville, and T.D. Wooley. Near-optimal mean value estimates for multidimensional Weyl sums. *Geom. Funct. Anal.*, 23:1962–2024, 2013.
[22] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25(2):855–900, 1997.
[23] C. Pomerance. On the distribution of round numbers. In *Lecture Notes in Mathematics*, volume 1122, pages 173–200. Number Theory Proceedings, Ootacamund, India 1984, K. Alladi (Ed.), 1985.
[24] M. Radziwill. On large deviations of additive functions. arXiv:0909.5274, 2009.
[25] L. Shepp and S. Lloyd. Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.*, 121(2):340–357, 1966.
[26] A. Srinivasan. A note on the distribution of the number of prime factors of the integers. *Information Processing Letters*, 109(2):133–135, 2008.
[27] T. Tao. *Higher Order Fourier Analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Soc., 2012.
[28] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1995.
[29] B. Tsou. Eigenvalue fluctuations of symmetric group permutation representations on k-tuples and k-subsets. arXiv:1810.11904, 2018.
[30] G. A. Watterson. The sampling theory of selectively neutral alleles. *Advances in Applied Probability*, 5(3):463–488, 1974.
[31] K. Wieand. Eigenvalue distributions of random permutation matrices. *Ann. Probab.*, 28(4):1563–1587, 2000.
[32] K. Wieand. Permutation matrices, wreath products, and the distribution of eigenvalues. *J. Theoret. Probab.*, 16(3):599–623, 2003.
[33] T.D. Wooley. Vinogradov’s mean value theorem via efficient congruencing. *Annals of Math.*, 175:1575–1627, 2012.