THE CHARACTERISTIC POLYNOMIAL OF SUMS OF RANDOM PERMUTATIONS AND REGULAR DIGRAPHS

SIMON COSTE, GAULTIER LAMBERT, AND YIZHE ZHU

ABSTRACT. Let A_n be the sum of d permutations matrices of size $n \times n$, each drawn uniformly at random and independently. We prove that $\det(\mathrm{I}_n - zA_n/\sqrt{d})$ converges when $n \to \infty$ towards a random analytic function on the unit disk. As an application, we obtain an elementary proof of the spectral gap of random regular digraphs. Our results are valid both in the regime where d is fixed and for d slowly growing with n.

1. INTRODUCTION

Spectral properties of non-Hermitian random matrices can have different behaviors depending on their degree of sparsity. These properties are now well understood for dense matrices with iid entries; a well-known example is the Circular Law [7], for which the optimal sparsity threshold is known [47, 4, 44]. However, when the matrices in question are very sparse, with a fixed number of non-zero entries on each row, including dependencies, the problem becomes different and more challenging. One of our goals in this paper is to understand these differences.

Sums-of-permutations. There are numerous ways to enforce sparsity in random matrices, and different ensembles are expected to behave differently. In this paper, we focus on *permutation matrices*, that is, matrices with exactly one non-zero entry on each row and each column; adding d independent permutation matrices chosen uniformly at random, we obtain a random matrix A with integer entries, whose row/column-sums are all exactly equal to d. Thus, A can be viewed as a typical matrix with fixed row sums and column sums. This model displays two important properties:

(i) The structure of A is very constrained (fixed row/column sums), hence the entries of A are not independent.
(ii) The rows and columns of A can be swapped while keeping the distribution of A invariant (invariance by permutations).

Random regular directed graphs. Sums of random permutations are of particular interest to random graph theory since they are a popular proxy for the (adjacency) matrix of *random regular digraphs*. A digraph is *regular* when each node has the same number of in-neighbors and out-neighbors; consequently, the matrix A defined above is the matrix of a d-regular digraph, possibly with multiple edges. It turns out that for fixed d, conditioning on A having no entry greater than 1, the graph represented by A is nearly uniform among all the d-regular directed graphs (the two models are *contiguous* [29], and both are contiguous with the configuration model). Our analysis thus provides results on the eigenvalues of random regular digraphs, notably a new proof for the directed version of Friedman's second eigenvalue theorem [21, 6] which relates to important graph-theoretical notions such as graph expansion and random walks on digraphs [42, 16].

Spectral properties. In general, studying the eigendecomposition of non-Hermitian matrices can be challenging. For an $n \times n$ permutation matrix ($d = 1$), the eigenvalues of A belong to the n-th roots of unity, and their multiplicity is given by the cycle decomposition theorem; if c_k denotes the number of cycles of length k (thus $c_1 + 2c_2 + \cdots + nc_n = n$), then the multiplicity of a root of unity ω will be the sum of the c_k for which $\omega^k = 1$.

This basic structure allows to study even non-uniform permutation matrices,[27, 5]. However, already for $d = 2$, this straightforward analysis breaks down because generically, the permutation matrices that we sum do not commute. In particular, a famous conjecture states that for fixed d as $n \to \infty$, the (spectral) empirical measure of A converges to the *oriented Kesten-McKay* distribution [7]. In contrast, in the regime where $d \to \infty$ as $n \to \infty$, one expects to recover the circular law and there are already several results in this direction [3, 10, 33].

Date: April 1, 2022.

Our contributions. In this paper, we study asymptotics of the characteristic polynomial outside of the spectral support. The results we obtain are analogous to [8, 39]; we identify the distributional functional limit of the characteristic polynomial of A, away from the spectrum. This problem has also been considered in the Hermitian case for Gaussian β-ensembles [32]. In particular, the relationships between our results and the theory of multiplicative chaos are discussed in Section 2.2. Surprisingly, we obtain the same as in [15], which considers Wigner matrices with $\mathrm{Bernoulli}(d/n)$ entries — a proxy for sparse Erdős-Rényi digraphs. This result is rather unexpected since in the Hermitian case, the spectral properties of random regular graphs and sparse Erdős-Rényi graphs are radically different for fixed d. In Appendix B, we also report on several observations regarding the sum of Ewens-distributed random permutations, which can be of independent interest.

As corollaries of our convergence results for characteristic polynomials (Theorems 2.2 and 2.4 stated in Section 2), we obtain two spectral gap theorems covering different regimes.

Theorem 1.1 (spectral gap for d-regular digraphs, d fixed)**.** *Let A_n be one of the following random matrix models:*

(i) the sum of d independent uniform permutation matrices of size $n \times n$;
(ii) the sum of d independent $n \times n$ uniform permutation matrices conditioned on their graph being simple;
(iii) the adjacency matrix of a uniform random directed d-regular graph on n vertices.

Then, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|\lambda_2| > \sqrt{d} + \varepsilon) = 0.$$

where λ_i are the complex eigenvalues of A_n, ordered by decreasing modulus: $d = \lambda_1 \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$.

Theorem 1.1 gives an alternate proof of the spectral gap result obtained in [16]; it notably implies that there are no outliers outside the support of the oriented Kesten-McKay law, except the trivial eigenvalue $\lambda_1 = d$. It is not known for the moment how to prove the matching lower-bound on $|\lambda_2|$; regarding this problem (directed analogs of the Alon-Boppana inequality), we refer to the discussion in [15, Section 4].

When $d(n) \to \infty$, there is no contiguity result in the literature between the uniform regular digraphs and sums of random permutations (it was conjectured in [13] these two models are contiguous when $d(n) = O(\log n)$); therefore, our next Theorem only applies to the sums-of-permutations model.

Theorem 1.2 (spectral gap for sums of permutations, $d(n) \to \infty$)**.** *Let A_n be the sum of $d(n)$ independent permutation matrices. Assume that $d(n) \to \infty$ in such a way that $d(n) = n^{o(1)}$ as $n \to \infty$. Then, for any $\varepsilon > 0$,*

$$\mathbb{P}(|\lambda_2| > \sqrt{d(n)}(1+\varepsilon)) \to 0 \tag{1.1}$$

where λ_i are the complex eigenvalues of A, ordered by decreasing modulus: $d = \lambda_1 \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$.

Combining the circular law proved in [3] for the random matrix $A_n/\sqrt{d(n)}$, the bound (1.1) is sharp (at least in the regime $d(n) \geq \log^{12}(n)/(\log\log n)^4$ and $d(n) = n^{o(1)}$). Moreover, with high probability, there is no outlier eigenvalues besides the trivial one $\lambda_1 = \sqrt{d(n)}$. In Theorem 1.2, the condition $d(n) = n^{o(1)}$ is technical (due to our proof of Theorem thm:trace (2)) and it could be dispensed with some extra work.

Related work. The spectral properties of random permutations under Ewens distribution were investigated in several works, including the characteristic polynomial and linear statistics [27, 17, 5, 2]. For a uniform random permutation matrix, the maximum of the characteristic polynomial on the unit circle has been studied in [12]. It is also of interest to study this question for sums-of-permutations for general d and we intend to consider this problem in subsequent work.

Spectral gap of random d-regular digraphs with fixed d was studied in [16] using the high trace method, which is limited to fixed d. For uniform random regular digraphs, it was shown in [48] that $|\lambda_2| \leq C\sqrt{d(n)}$ for $d(n) = O(n^{2/3})$, and for $n^{\varepsilon} \leq d(n) \leq n/2$ in [46]. By the size-biased coupling method introduced in [11], one can prove a similar result to [11, Theorem 2.6] that $|\lambda_2| = O(\sqrt{d(n)})$ for the sum of permutations for any $2 \leq d(n) \leq n/2$. However, all of these results rely on the ε-net and the Kahn-Szemrédi argument [22], which yields an absolute constant far from optimal. Our approach gives a unified treatment for fixed and growing d in the random permutation model with a sharp constant, while being elementary on a technical level.

Lower bounds on the least singular value of random d-regular digraphs for fixed d is a problem of considerable interests since it is a crucial step towards proving the conjectured oriented Kesten-McKay law. The singularity probability was estimated in [26, 36] for fixed d, and in [14, 34] for growing $d(n)$. Quantitative estimates on the least singular values for growing $d(n)$ were obtained in [3, 10, 35, 28].

In the Hermitian case, the sum of random permutations and their transpose $\sum_{i=1}^{d}(P_i + P_i^T)$ can be seen as a model for random $2d$-regular multi-graphs, and their spectral properties have been investigated, including spectral gaps [21, 11] and linear eigenvalue statistics [18, 30, 25]. There is also a direct connection between random d-regular digraphs and random bipartite d-regular graphs. In particular, the eigenvalue fluctuation results of random bipartite regular graphs shown in [19] can be translated to singular value fluctuations of random regular digraphs. However, it is a very challenging problem to study the eigenvalue fluctuations of random regular digraphs.

Acknowledgements. G.L. and Y.Z. acknowledge support from NSF DMS-1928930 during their participation in the program "Universality and Integrability in Random Matrix Theory and Interacting Particle Systems" hosted by the Mathematical Sciences Research Institute in Berkeley, California during the Fall semester of 2021. G.L. is supported by the SNSF Ambizione grant S-71114-05-01. Y.Z. was partially supported by NSF-Simons Research Collaborations on the Mathematical and Scientific Foundations of Deep Learning. S.C. is supported by ENS-PSL.

2. Main Results

2.1. **Sums of random permutations.** For any $n \in \mathbb{N}$, let $d = d(n) \in \mathbb{N}$ and $P^{(1)}, \dots, P^{(d)}$ be a collection of i.i.d.random uniform $n \times n$ permutation matrices. We consider the random matrix

$$A_n := P^{(1)} + \cdots + P^{(d)}.$$

In the sequel, we will often drop the n subscript and simply write d, A, etc. We define the (rescaled) characteristic polynomial of A,

$$\widehat{\chi}_n(z) := \frac{1}{\sqrt{d}} \det\left(\mathrm{I}_n - z\frac{A}{\sqrt{d}}\right); \qquad z \in \mathbb{C}. \tag{2.1}$$

Note that A has a trivial eigenvalue d associated with the vector $\mathbf{1}^n$, so that with our convention, $\widehat{\chi}_n(1/\sqrt{d}) = 0$ almost surely. If fact, we will see that for any $n \in \mathbb{N}$,

$$\mathbb{E}[\widehat{\chi}_n(z)] = z - 1/\sqrt{d}; \qquad z \in \mathbb{C}.$$

The main results of this paper identify the weak $n \to \infty$ limit of the sequence $(\widehat{\chi}_n)$ in the unit disk either when the degree $d(n) = d$ is fixed, or when $d(n) \to \infty$ slowly with n. Up to a rescaling, the limit is the same as the function defined in [15]; but we adopt different conventions allowing for a unified treatment of the cases $d = O(1)$ and $d \to \infty$. We denote $\mathrm{D}_r = \{z \in \mathbb{C} : |z| < r\}$ the disk of radius $r > 0$ in the complex plane.

Definition 2.1. Fix $d \in \mathbb{N}$ and let $\{\Lambda_\ell\}_{\ell \in \mathbb{N}}$ be a family of independent random variables, with

$$\Lambda_\ell \sim \mathrm{Poisson}\left(\frac{d^\ell}{\ell}\right).$$

We also denote $\overline{\Lambda_\ell} = \Lambda_\ell - \mathbb{E}[\Lambda_\ell]$ for $\ell \in \mathbb{N}$, so that $\{\overline{\Lambda_\ell}\}_{\ell \in \mathbb{N}}$ are centered random variables. For any integer d, we define the following formal series:

$$Y_d(z) := \sum_{k \in \mathbb{N}} \frac{z^k}{k d^{k/2}} \sum_{\ell | k} \ell \overline{\Lambda_\ell}, \qquad X_d(z) := \sum_{k \in \mathbb{N}} \frac{z^k}{d^{k/2}} \overline{\Lambda_k}. \tag{2.2}$$

Proposition 2.5 implies that these functions are well-defined and analytic functions in the unit disk D_1. Let us now state our main result for d fixed.

Theorem 2.2. *Consider a fixed integer d and let $d(n) = d$ for every n; let Y_d be as in (2.2). Then,*

$$\widehat{\chi}_n(z) \xrightarrow[n \to \infty]{\text{law}} (z - 1/\sqrt{d}) \frac{\mathrm{e}^{-Y_d(z)}}{\mathbb{E}[\mathrm{e}^{-Y_d(z)}]}$$

for the topology of uniform convergence of compact subsets of D_1.

In contrast, if the degree d diverges, then we recover a Gaussian analytic function in the limit, as for dense matrices non-Hermitian (real-valued) Wigner matrices [8] — note that this function is real-analytic, the coefficients being *real* Gaussian variables, not complex ones. The proof of Theorem 2.2 and the subsequent results is outlined in Section 3.

Definition 2.3. Let $\{N_\ell\}_{\ell\in\mathbb{N}}$ be i.i.d. standard real Gaussian random variables and define the random real-analytic function

$$X_\infty(z) = \sum_{k\in\mathbb{N}} \frac{N_k}{\sqrt{k}} z^k, \qquad z \in \mathrm{D}_1. \tag{2.3}$$

Theorem 2.4. *Consider a sequence $d(n) \in \mathbb{N}$ such that $d(n) \to \infty$ and $d(n) = n^{o(1)}$ as $n \to \infty$. Let X_∞ be as in (2.3). Then, it holds for the topology of locally uniform convergence on D_1,*

$$\widehat{\chi}_n(z) \xrightarrow[n\to\infty]{\text{law}} z\sqrt{1-z^2}\mathrm{e}^{X_\infty(z)}.$$

The technical condition $d(n) = n^{o(1)}$ means that $\log d(n) = o(\log n)$ as $n \to \infty$.

2.2. **Log-correlated structure of the limiting random fields.** We gather here a few properties of the functions X_d, Y_d. In particular, we show that the boundary-values (on the unit circle $\partial\mathrm{D}_1$) of the functions Y_d and X_∞ are log-correlated fields and discuss some expected consequences.

Proposition 2.5. *Y_d and X_d are random (centered) real-analytic functions on D_1. Moreover,*

- *For $z \in \mathrm{D}_1$, we have*

$$\mathbb{E}[\mathrm{e}^{-Y_d(z)}] = \bigg(\prod_{\ell\in\mathbb{N}} \mathrm{f}_\ell(z/\sqrt{d})^{\frac{d^\ell}{\ell}}\bigg)^{-1}, \tag{2.4}$$

  where $\mathrm{f}_\ell(z) := (1-z^\ell)\mathrm{e}^{z^\ell}$. The infinite product converges uniformly on compact sets of D_1.
- *For $d \geq 2$, $Y_d = X_d + \Upsilon_d$ where $\mathbb{E}[\|\Upsilon_d\|^2_{\mathrm{L}^\infty(\mathrm{D}_1)}] \leq C/d^{1/4}$ for a numerical constant C. Then, Y_d has covariance kernel*

$$\mathbb{E}[Y_d(z)Y_d(w)] = \log(1-zw)^{-1} + O_d(1), \qquad z, w \in \mathrm{D}_1$$

  where the error term is a bi-analytic function which converges to 0 uniformly in $\overline{\mathrm{D}_1 \times \mathrm{D}_1}$ as $d \to \infty$.

One can check that for $d \in \mathbb{R}_+ \cup \{\infty\}$, the random series Y_d also converges in the Sobolev space of (Schwartz) distributions $\mathrm{H}^{-\epsilon}(\partial\mathrm{D}_1)$ of any $\epsilon > 0$. In particular, for $d \in \mathbb{N}$, $\big(Y_d(z) : z \in \partial\mathrm{D}_1\big)$ defines a non-Gaussian log-correlated field. From this perspective, it is natural to ask for the asymptotics[1] of $\max_{|z|=r} \Re X_d(z)$ as $r \to \infty$.

If $d = 1$, the leading order of the maximums of Y_d and $\log|\widehat{\chi}_n|$ have been studied in [12]; specifically, it was shown that $\max_{z\in\partial\mathrm{D}_1} \log|\widehat{\chi}_n| \sim x_0 \log n$ as $n \to \infty$ where $x_0 \simeq 0.652$. The significance of this result is that the usual prediction for the value of x_0 coming from the theory of log-correlated fields [23] does not apply to this problem and this is due to the tails of the field $\log|\widehat{\chi}_n|$ which are not Gaussian. A similar and more accessible question is the maximum of the characteristic polynomial of the CUE (circular unitary ensemble or Haar-distributed random matrices over the group $\mathrm{U}(n)$). This problem has been thoroughly considered and precise results about the asymptotics of the maximum are available [24, 1, 41, 9, 31]. We intend to consider the asymptotics for maximum of Y_d and the leading order of that of $\log|\widehat{\chi}_n|$ for general d in subsequent work. Another perspective is that $|\widehat{\chi}_n(z)|^\gamma \mathrm{d}z$ appropriately renormalized should converge to a multiplicative chaos[2] measure as $n \to \infty$ for $\gamma > 0$ sufficiently small (in the subcritical phase). In contrast to the CUE characteristic polynomial [40] it is not clear whether the critical value is the standard $\gamma = \sqrt{2}$ for this model or if it depends on the degree d.

One can also investigate the regularity of the random (Schwartz) distributions $\big(\mathrm{e}^{\sqrt{\theta}Y_d(z)} : z \in \partial\mathrm{D}_1\big)$ depending on $\theta > 0$. This question has just been considered in the CUE case[3] in [39] (for the so-called *holomorphic multiplicative chaos*)

[1] $\big(Y_d(z) : z \in \mathrm{D}_1\big)$ is the harmonic extension of the log-correlated field inside the disk and it a natural way to regularize this random generalized function. An alternative approach consists in truncating the Fourier series (2.2), considering the asymptotics of the maximum of $\Re\big(\sum_{k\leq N} \frac{z^k}{kd^{k/2}} \sum_{\ell|k} \ell\overline{\Lambda_\ell}\big)$ as $N \to \infty$. Theorem 2.2 shows that yet another regularization is given by the log characteristic polynomial $\log|\widehat{\chi}_n(z)|$ (after the appropriate centering). These procedures are expected to be all equivalent.

[2] According to Theorem 2.4, the limiting random measures should be Gaussian multiplicative chaos (GMC) only in the regime as $d(n) \to \infty$. For fixed d, we still expect that the limiting random measures have similar multi-fractal properties [20].

[3] The results of [9, 31, 39] are valid for general circular β-ensembles.

which exhibits a phase transition related to the asymptotics for the maximum of the CUE field. In summary, these are fascinating research questions which motivates the study of characteristic polynomials of permutation matrices and the corresponding limiting random fields.

Finally, we relate the Gaussian analytic function in (2.3) with the Y_d, X_d defined before. It is well known that X_∞ is an analytic function over D_1 with covariance kernel

$$\mathbb{E}[X_\infty(z)X_\infty(w)] = \log(1-zw)^{-1}, \qquad z, w \in \mathrm{D}_1.$$

Thanks to the second point in Proposition 2.5, we have the following convergence result.

Proposition 2.6. *With the notation of Proposition 2.5, it holds for the topology of locally uniform convergence on* D_1*,*

$$(X_d, \Upsilon_d) \xrightarrow[d\to\infty]{\text{law}} (X_\infty, 0).$$

The proofs of Propositions 2.5 and 2.6 will be given in Section 7. We also refer to [15, Section 2.3] for the computation of the generating function of the exponential moments of Y_d.

2.3. **On fluctuations outside the support of the Kesten-McKay density.** The empirical spectral density of sums of permutation matrices or random regular digraphs is conjectured to converge towards the Oriented Kesten-McKay law, whose density with respect to the Lebesgue measure on $\mathbb{C}$ is

$$\varrho_d(z) = \frac{d^2(d-1)}{\pi}\frac{\mathbf{1}_{|z|<\sqrt{d}}}{(d^2-|z|^2)^2}.$$

The reader will find many insights regarding this problem in [10, 37]. Our main result gives some information on the fluctuations around this law, in the spirit of [43]. The fluctuation around the oriented Kesten-McKay distribution is defined as the function

$$\Psi_n(z) = \log|\det(z\mathrm{I}_n - A_n)| - nU_d(z)$$

where U_d is the Log-potential of ϱ_d. It can be checked that

$$U_d(z) = \begin{cases} \log|z| \ \text{ for } |z| > \sqrt{d}, \\ -(d-1)\log\sqrt{d^2-|z|^2} + \alpha_d \text{ for } |z| \leqslant \sqrt{d} \end{cases} \tag{2.5}$$

where α_d is the numerical constant $(d-1)\log\sqrt{1-d^{-1}} + (d-1/2)\log(d)$. As a consequence, for every $|z| > \sqrt{d}$, the fluctuations are given by $\Psi_n(z) = \log|\det(\mathrm{I} - zA_n)|$ since the $n\log|z|$ terms cancel in Ψ_n. Our main result, Theorem 2.2, identifies the limit of the fluctuations Ψ_n outside the disk $\mathrm{D}_{\sqrt{d}}$ as $-\mathrm{Re}(Y_d(1/z))$. In figure 1 depicting Ψ_n, this corresponds to the smooth part of the picture. The rough part corresponds to the fluctuations inside the bulk of the oriented Kesten-McKay density; it is not clear what is going to be the correct definition of this generalized funtion.

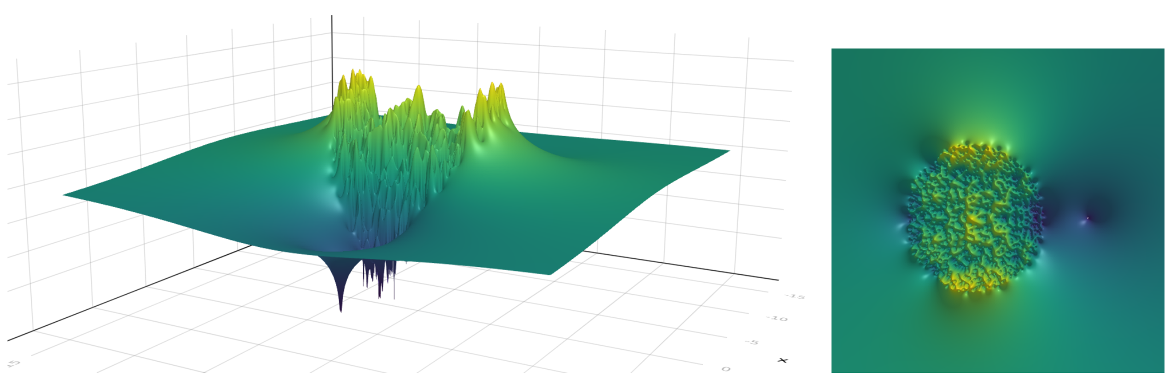


FIGURE 1. Picture of the values of $\log|\det(z - A_n)| - nU_d(z)$ for z in $[-5,5]^2$, from two different angles;Âăhere, A_n is a sum of 3 uniform permutation matrices, $n = 2000$ and U_d is the Log-potential (2.5), from two different angles. The logarithmic singularity of $\Psi_n(z)$ at $z = d$ is visible in the smooth (harmonic) part of the picture.

3. Outline of the paper

3.1. **Proof strategy.** The proofs of Theorems 2.2 and 2.4 are given in Section 8. They are based on the following three steps strategy;

(i) If $(d(n))_{n\in\mathbb{N}}$ is a sequence of degree such that $d(n) = o(\sqrt{n})$, then the family of random analytic function $\{\widehat{\chi}_n(z) : z \in \mathrm{D}_1\}_{n\in\mathbb{N}}$ is tight for the topology of locally uniform convergence on D_1; cf. Section 4.1 at page 7.

(ii) By expanding (2.1), for the principal branch of log, it holds for $k \in \mathbb{N}$ that

$$[z^k]\log\left(\frac{\widehat{\chi}_n(z)}{z+1/\sqrt{d}}\right) = \frac{(-1)^{k+1}}{k}\frac{\mathrm{tr}(A^k) - d^k}{d^{k/2}} \tag{3.1}$$

where $[z^k]X$ means the k-th coefficient of the analytic function X. Note that it is natural to consider log of $z \mapsto \frac{\widehat{\chi}_n(z)}{z+1/\sqrt{d}}$ to account for the trivial root of $\widehat{\chi}_n$ at $-1/\sqrt{d} \in \mathrm{D}_1$ if $d \geq 2$. We establish that

$$\frac{\mathrm{tr}(A^k) - d^k}{d^{k/2}} \xrightarrow[n\to\infty]{\text{law}} \mathrm{L}_k$$

in the sense of finite dimensional distributions for a sequence of independent random variables $(\mathrm{L}_k)_{k\in\mathbb{N}}$. This is the content of Theorem 3.1 thereafter, which is proved in Section 6 at page 10.

(iii) Suppose that $\sum_{k\in\mathbb{N}}|\mathrm{L}_k|r^k$ converges for any $r < 1$ almost surely. From (i) tightness and (ii) we conclude that the random analytic function

$$\left(\frac{\widehat{\chi}_n(z)}{z+1/\sqrt{d}}\right) \xrightarrow[n\to\infty]{\text{law}} \exp\left(\sum_{k\in\mathbb{N}}\frac{(-1)^{k+1}}{k}\mathrm{L}_k z^k\right)$$

in the topology of locally uniform convergence on D_1; cf. Section 8.

Step (ii) consists in a standard application of the trace method, without resorting to high traces. The main result is the following one.

Theorem 3.1. *(1) Suppose that $d \in \mathbb{N}$ is fixed and $\{\Lambda_\ell\}_{\ell\in\mathbb{N}}$ be as in Definition 2.1. Then for every $k \in \mathbb{N}$,*

$$\left(\mathrm{tr}(A), \ldots, \mathrm{tr}(A^k)\right) \xrightarrow[n\to\infty]{\text{law}} \left(\Lambda_1, \ldots, \sum_{\ell|k}\ell\Lambda_\ell\right). \tag{3.2}$$

(2) Consider a sequence $d(n) \in \mathbb{N}$ such that $d(n) \to \infty$ and $d(n) = n^{o(1)}$ as $n \to \infty$. Then for every $k \in \mathbb{N}$,

$$\left(\frac{\mathrm{tr}(A) - d}{\sqrt{d}}, \ldots, \frac{\mathrm{tr}(A^k) - d^k}{\sqrt{d^k}}\right) \xrightarrow[n\to\infty]{\text{law}} \left(N_1, \sqrt{2}N_2 + 1, \ldots, \sqrt{k}N_k + \mathbf{1}_{\{k \text{ is even}\}}\right). \tag{3.3}$$

The proof of Theorem 3.1 relies on the relationship between the random variables $\mathrm{tr}(A^k)$ and k-cycles on the (adjacency) digraph associated to the matrix A – note that A is generically non-Hermitan and d is the maximum degree of the digraph. The argument proceeds by the moment method, estimating the probabilities to observe some given collection of cycles for large n.

Remark 3.2. Our Theorem 3.1 implies eigenvalue fluctuation $\sum_{i=1}^n[f(\lambda_i) - \mathbb{E}f(\lambda_i)]$ for any polynomial f in z. To study eigenvalue fluctuations of A for any smooth test function f with the moment method, one needs to extend the results to polynomial test functions in z and $\overline{z}$ and approximate f with polynomials. This approach was developed in [43] for the Ginibre ensemble.

3.2. **Notational preliminaries.** We never indicate that d and the random variables defined in terms of A depend on n=size(A). We always consider limits as $n \to \infty$.

Sets. For any $k \in \mathbb{N}$, denote $[k] = \{1, \ldots, k\}$. If $\mathrm{I} \subset \mathbb{N}$ is a finite set, we denote by S_I, the group of permutations of elements in I. For $\sigma \in S_\mathrm{I}$, we let $\epsilon(\sigma)$ be the sign of this permutation. If I is a finite set, we denote by $|\mathrm{I}|$ its cardinal (number of elements). Tuples are denoted by **boldface letters**. If $\mathrm{I} \subseteq \mathbb{N}$ is a set, then

$$\mathrm{I}^k = \{\mathbf{i} = (i_1, \ldots, i_k) : i_1, \ldots, i_k \in \mathrm{I}\}$$

for $k \in \mathbb{N}$. For $k \in \mathbb{N}$, we denote

$$\mathcal{E}_k = \{\mathbf{i} \in [n]^k : i_1, \ldots, i_s \text{ are distinct}\}.$$

Matrices. In the sequel $A = (A_{i,j})_{i,j\in[n]}$ is a random matrix and we denote for any multi-index $\mathbf{i} \in [n]^k$,

$$A_{\mathbf{i}} := A_{i_1,i_2} \cdots A_{i_{k-1},i_k} A_{i_k,i_1}.$$

With this notation,

$$\mathrm{tr}(A^k) = \sum_{\mathbf{i}\in[n]^k} A_{\mathbf{i}}.$$

For any subset $\mathrm{I} \subset [n]$, we denote $A(\mathrm{I}) = (A_{i,j})_{i,j\in\mathrm{I}}$ the corresponding submatrix.
Directed graphs. The matrix A is an $n \times n$ matrix with integer entries. It represents the adjacency matrix of a (weighted) digraph, on the vertex set $V = [n]$. The edge set is $E = \{(i,j) \in [n] \times [n] : A_{i,j} \geq 1\}$ and the weights correspond to the entries of A. We insist on the fact that $G_A = (V,E)$ is directed and possibly has simple loops.

We can interpret $\mathbf{i} \in [n]^k$ as *cyclic paths* on the graph of A and $A_{\mathbf{i}}$ as the *weight* of this path. In addition, the paths $\mathbf{i} \in \mathcal{E}_k$ correspond to simple loops.
Random permutation. We define a random permutation matrix $P = (\mathbf{1}_{\pi(i)=j})_{i,j\in[n]}$ where π is a uniform random element of $S_{[n]}$. We will use that under this law, for any $k \in [n]$, any $\mathbf{i}, \mathbf{j} \in [n]^k$,

$$\mathbb{P}[\pi(i_1) = j_1, \ldots, \pi(i_k) = j_k] = \mathbb{P}[P_{i_1,j_1} = \cdots = P_{i_k,j_k} = 1] = \mathbf{1}_{\{\mathbf{i},\mathbf{j}\in\mathcal{E}_k\}} \frac{(n-k)!}{n!}.$$

We will also make several use of the following bounds; for any $d \in \mathbb{N}$, if $\mathbf{k} \in \mathbb{N}^d$ with $k_1 + \cdots + k_d = k \leq n$, then

$$\prod_{\delta=1}^{d} \frac{(n-k_\delta)!}{n!} \leq \frac{(n-k)!}{n!}. \tag{3.4}$$

Note that the following lower bound is essentially sharp if $k \ll \sqrt{n}$,

$$\prod_{\delta=1}^{d} \frac{(n-k_\delta)!}{n!} \geq \frac{(n-k)!}{n!} \left(\frac{n-k}{n}\right)^k \geq \frac{(n-k)!}{n!} \left(1 - \frac{k^2}{n}\right). \tag{3.5}$$

4. Tools for proving the tightness of $\{\chi_n\}$

4.1. A simple tightness criterion.

A family of analytic function $\{f_n\}$ on a domain $\mathrm{D} \subset \mathbb{C}$ is pre-compact (or normal) for the topology of locally uniform convergence if for any compact $\mathrm{K} \subset \mathrm{D}$,

$$\sup_n \|f_n\|_{\mathrm{L}^\infty(\mathrm{K})} < \infty.$$

In particular, $\{f_n\}$ is tight (in this topology) if for any compact $\mathrm{K} \subset \mathrm{D}$,

$$\left\{\|f_n\|_{\mathrm{L}^\infty(\mathrm{K})}\right\} \text{ is tight.}$$

In particular, it suffices to check that for $\alpha > 0$

$$\sup_n \mathbb{E}\|f_n\|^\alpha_{\mathrm{L}^\infty(\mathrm{K})} < \infty.$$

We will use this criterion to show that the characteristic polynomial of a random $n \times n$ matrix A,

$$\left\{\chi_n = \det(1 + \cdot A)\right\} \text{ forms a tight sequence in a disk } \mathrm{D}_r = \{|z| < r\}.$$

Note that the random matrices A need not be defined on the same probability spaces for different $n \in \mathbb{N}$.
Recall that for $z \in \mathbb{C}$

$$\chi_n(z) = \sum_{k\leq n} z^k \Delta_k^{(n)}$$

where Δ_k are called secular coefficients and are given by $\Delta_0 = 1$,

$$\Delta_k^{(n)} = \sum_{\mathrm{I}\subseteq[n],|\mathrm{I}|=k} \det(A(\mathrm{I})), \qquad k \in [n]. \tag{4.1}$$

For any $\theta < 1$, $r > 0$ and $\epsilon > 0$, by Jensen's inequality, it holds

$$\mathbb{E}\|\chi_n\|^{1+\epsilon}_{\mathrm{L}^\infty(\mathrm{D}_{r\theta})} \leq \mathbb{E}\Big(\sum_{k\leq n} r^k\theta^k|\Delta_k^{(n)}|\Big)^{1+\epsilon} \leq \frac{1}{(1-\theta)^\epsilon}\sum_{k\leq n} r^{k(1+\epsilon)}\mathbb{E}|\Delta_k^{(n)}|^{1+\epsilon}.$$

With the normalization (2.1), by scaling and choosing $\epsilon = 1$, it holds for $r > 0$ and $\theta < 1$,

$$\mathbb{E}\|\widehat{\chi}_n\|^2_{\mathrm{L}^\infty(\mathrm{D}_{r\theta})} \leq \frac{1}{1-\theta}\sum_{k\leq n}\frac{r^{2k}}{d^{k+1}}\mathbb{E}|\Delta_k^{(n)}|^2.$$

For the model $A = \sum_{q=1}^d P^{(q)}$ where $P^{(q)}$ are $d \in \mathbb{N}$ independent random permutation matrices, we will prove the following bound in the sequel. For any $n \geq 2$, it holds for all $0 < r < 1$ and $1 \leq d < \sqrt{\frac{n(1-r)}{r}}$,

$$\sum_{k\leq n}\frac{r^k}{d^{k+1}}\mathbb{E}|\Delta_k^{(n)}|^2 \leq \frac{r(1+4/d)}{(1-r-\frac{rd^2}{n})^2}$$

This immediately implies the following result.

Proposition 4.1. *For any sequence $d(n) \in \mathbb{N}$ with $d = o(\sqrt{n})$, the sequence of random analytic function $(\widehat{\chi}_n)_{n\in\mathbb{N}}$ admits a subsequence which converges weakly locally uniformly on compact subset of D_1.*

4.2. **Exchangeability.** By (4.1), it holds for any $k \in [n]$

$$\mathbb{E}\Delta_k^2 = \sum_{|\mathrm{I}|,|\mathrm{J}|=k}\mathbb{E}\det(A(\mathrm{I}))\det(A(\mathrm{J})).$$

where the sum is over all subsets $\mathrm{I}, \mathrm{J} \subseteq [n]$. Lemma 4.3 below allows to reduce this sum to subsets (I,J) which share at least $k-1$ elements;

$$\mathbb{E}\Delta_k^2 = \sum_{|\mathrm{I}|=k}\mathbb{E}\det(A(\mathrm{I}))^2 + \sum_{|\mathrm{I}|=|\mathrm{J}|=k, |\mathrm{I}\cap\mathrm{J}|=k-1}\mathbb{E}\det(A(\mathrm{I}))\det(A(\mathrm{J})). \tag{4.2}$$

The Lemmas of this section apply to any exchangeable random matrix model, that is, if the following invariance property holds; for any given permutation $\pi \in S_{[n]}$

$$\big(A_{i,\pi(j)}\big)_{i,j\in[n]} \overset{\text{law}}{=} \big(A_{\pi(i),j}\big)_{i,j\in[n]} \overset{\text{law}}{=} \big(A_{i,j}\big)_{i,j\in[n]}. \tag{4.3}$$

Lemma 4.2. *For any $k \in [n]$,*

$$\mathbb{E}\Delta_k^{(n)} = \mathbf{1}_{k=1}\mathbb{E}\mathrm{tr}A.$$

In particular, for the random matrix $A = \sum_{q=1}^d P^{(q)}$, we have $\mathbb{E}\Delta_1^{(n)} = d$ and

Proof. By (4.1), it holds for any $k \in [n]$

$$\mathbb{E}\Delta_k = \sum_{|\mathrm{I}|=k}\mathbb{E}\det(A(\mathrm{I})).$$

If the random matrix A satisfies (4.3), observe that by permuting two columns of A, $\mathbb{E}\det(A(\mathrm{I})) = -\mathbb{E}\det(A(\mathrm{I}))$, hence $\mathbb{E}\det(A(\mathrm{I})) = 0$ for any $\mathrm{I} \subset [n]$ provided that $|\mathrm{I}| \geq 2$. This shows that $\mathbb{E}\Delta_k = 0$ for all $k \geq 2$ and

$$\mathbb{E}\Delta_1 = \sum_{i\in[n]}\mathbb{E}A_{i,i} = \mathbb{E}\mathrm{tr}A. \qquad \square$$

Lemma 4.3. *Assume $|\mathrm{I}| = |\mathrm{J}| = k$. If $|\mathrm{I}\cap\mathrm{J}| \leq k-2$, then $\mathbb{E}\det(A(\mathrm{I}))\det(A(\mathrm{J})) = 0$.*

Proof. For ease of notation, we consider the case when $|\mathrm{I}\cap\mathrm{J}| = k-2$. Other cases follow in the same way. Let us denote $\{x,y\} = \mathrm{I}\setminus\mathrm{J}$. Given $\sigma \in S_\mathrm{I}$, $\tau \in S_\mathrm{J}$, using the invariance in law (4.3) applied to $\pi = (xy)$, we have

$$\mathbb{E}\left[\prod_{i\in\mathrm{I},j\in\mathrm{J}}A_{i,\sigma(i)}A_{j,\tau(j)}\right] = \mathbb{E}\left[\prod_{i\in\mathrm{I},j\in\mathrm{J}}A_{i,\sigma'(i)}A_{j,\tau(j)}\right]$$

where $\sigma' = \sigma \circ \pi$. Note that we used that $\tau \circ \pi = \tau$ since $x, y \notin \mathrm{J}$. Since the map $\sigma \to \sigma' = \sigma \circ \pi$ is a bijection on S_{I} with the property that $\epsilon(\sigma') = -\epsilon(\sigma)$, by summing over all $\sigma \in S_{\mathrm{I}}$, this implies that for any fixed $\tau \in S_{\mathrm{J}}$,

$$\sum_{\sigma \in S_{\mathrm{I}}} \epsilon(\sigma)\epsilon(\tau)\mathbb{E}\left[\prod_{i \in \mathrm{I}, j \in \mathrm{J}} A_{i,\sigma(i)} A_{j,\tau(j)}\right] = 0.$$

Summing over $\tau \in S_{\mathrm{J}}$ gives the desired result. $\square$

5. Proof of our main theorems in the simple case of a single permutation

When $d = 1$, the matrix A is indeed the permutation matrix of a uniform permutation of $[n]$ and the analysis is easier to perform. We include this special case because it gives all the ideas for the general proof.

5.1. Tightness of the determinants.

When $d = 1$, this reduces to the study of random permutation matrix. In this case we can calculate the second moment of secular coefficients exactly.

Lemma 5.1. *For all* $1 \leq k \leq n - 1$,

$$\mathbb{E}|\Delta_k|^2 = 2, \tag{5.1}$$

and

$$\mathbb{E}|\Delta_n|^2 = 1. \tag{5.2}$$

Proof. Let $\pi \in S_n$ be a uniform random permutation and A be the permutation matrix of π. By Lemma 4.3, we only need to evaluate

$$\sum_{|\mathrm{I}|=k} \mathbb{E}\det(A(\mathrm{I})^2) + \sum_{|\mathrm{I}|=|\mathrm{J}|=k, |\mathrm{I}\cap\mathrm{J}|=k-1} \mathbb{E}\det A(\mathrm{I})A(\mathrm{J}).$$

The summation in the second term is non-empty if and only if $k < n$. For the first term,

$$\begin{aligned}
\mathbb{E}\det(A(\mathrm{I})^2) &= \sum_{\sigma,\tau \in S(\mathrm{I})} \epsilon(\sigma)\epsilon(\tau)\mathbb{E}\prod_{i\in\mathrm{I}} A_{i,\sigma(i)}A_{i,\tau(i)} \\
&= \sum_{\sigma,\tau} \epsilon(\sigma)\epsilon(\tau)\mathbb{P}\left(\pi(i) = \sigma(i) = \tau(i), \forall i \in \mathrm{I}\right) \\
&= \sum_{\sigma} \mathbb{P}\left(\pi(i) = \sigma(i), \forall i \in \mathrm{I}\right) = \mathbb{P}(\pi : \mathrm{I} \to \mathrm{I}) = \frac{k!(n-k)!}{n!} = \frac{1}{\binom{n}{k}},
\end{aligned}$$

where $\pi : \mathrm{I} \to \mathrm{I}$ is the event that π restricted on I is a permutation on I. Hence

$$\sum_{|\mathrm{I}|=k} \mathbb{E}\det(A(\mathrm{I})^2) = 1. \tag{5.3}$$

By taking $k = n$, it gives $\mathbb{E}[|\Delta_n|^2] = 1$.

For the second term, with the assumption $|\mathrm{I}| = |\mathrm{J}| = k$, $|\mathrm{I} \cap \mathrm{J}| = k - 1$,

$$\mathbb{E}\det A(\mathrm{I})A(\mathrm{J}) = \sum_{\sigma\in S(\mathrm{I}),\tau\in S(\mathrm{J})} \epsilon(\sigma)\epsilon(\tau)\mathbb{E}\prod_{i\in\mathrm{I}} A_{i,\sigma(i)} \prod_{j\in\mathrm{J}} A_{j,\tau(j)} \tag{5.4}$$

$$= \sum_{\sigma\in S(\mathrm{I}),\tau\in S(\mathrm{J})} \epsilon(\sigma)\epsilon(\tau)\mathbb{P}(\pi(i) = \sigma(i), i \in \mathrm{I}, \pi(j) = \tau(j), j \in \mathrm{J}) \tag{5.5}$$

Without loss of generality, we assume $\mathrm{I} = \{1, \ldots, k\}$, $\mathrm{J} = \{2, \ldots, k+1\}$. To have a non-zero probability, we must have $\sigma(i) = \tau(i)$ for all $i \in \mathrm{I} \cap \mathrm{J}$. This forces σ, τ to be two permutations on $\mathrm{I} \cap \mathrm{J}$, and $\sigma(1) = 1, \tau(k+1) = k+1$. Let

$$\Omega = \{(\sigma,\tau) \in S(\mathrm{I}) \times S(\mathrm{J}) : \sigma(1) = 1, \tau(k+1) = k+1, \sigma(k) = \tau(k), \forall k \in \mathrm{I}\cap\mathrm{J}\}.$$

For any $(\sigma,\tau)\in\Omega$, both permutations have the same cycle types, hence $\epsilon(\sigma)=\epsilon(\tau)$. Therefore

$$\begin{aligned}\mathbb{E}\det A(\mathrm{I})A(\mathrm{J}) &= \sum_{(\sigma,\tau)\in\Omega}\mathbb{P}(\pi(i)=\sigma(i), i\in \mathrm{I}, \pi(j)=\tau(j), j\in\mathrm{J})\\ &= \mathbb{P}(\pi(1)=1, \pi(k+1)=k+1, \pi:\{2,\dots,k\}\to\{2,\dots,k\})\\ &= \frac{(k-1)!(n-k-1)!}{n!}.\end{aligned}$$

Since $\#\{\mathrm{I},\mathrm{J}\subset[n]:|\mathrm{I}|=|\mathrm{J}|=k, |\mathrm{I}\cap\mathrm{J}|=k-1\}=\binom{n}{k}k(n-k)$, we conclude that

$$\sum_{|\mathrm{I}|=|\mathrm{J}|=k,|\mathrm{I}\cap\mathrm{J}|=k-1}\mathbb{E}\det A(\mathrm{I})A(\mathrm{J})=\binom{n}{k}\frac{k!(n-k)!}{n!}=1.$$

$\square$

5.2. **Trace asymptotics for one random uniform permutation.** Recall that in case $d=1$, $A=P=(\mathbf{1}_{\pi(i)=j})_{i,j\in[n]}$ where π is a uniform random element of $S_{[n]}$. We start by giving an estimate on the probability of generic events. Recall that for $s\in\mathbb{N}$, we denote

$$\mathcal{E}_s=\{\mathbf{i}\in[n]^s : i_1,\dots,i_s \text{ are distinct}\}.$$

Proposition 5.2. *Let $r,s\in\mathbb{N}_0$ with $r+s<n$ and fix $\mathbf{i},\mathbf{j}\in\mathcal{E}_s$ and $\mathbf{k},\mathbf{q}\in[n]^r$. If*

$$\{(k_1,q_1),\cdots,(k_r,q_r)\}\cap\{(i_1,j_1),\cdots,(i_s,j_s)\}=\emptyset, \tag{5.6}$$

then there exists $0\le\varepsilon\leqslant\frac{r}{n-s}$ such that

$$\mathbb{P}\big(P_{i_1,j_1}=\ldots=P_{i_s,j_s}=1, P_{k_1,q_1}=\ldots=P_{k_r,q_r}=0\big)=(1-\varepsilon)\frac{(n-s)!}{n!}. \tag{5.7}$$

Otherwise, if $\mathbf{i}\notin\mathcal{E}_s$ or $\mathbf{j}\notin\mathcal{E}_s$, or the condition (5.6) fails, then the probability on the LHS of (5.7) equals 0.

Proof. Plainly, under the hypothesis (5.6), we have the upper-bound,

$$\mathbb{P}\big(P_{i_1,j_1}=\ldots=P_{i_s,j_s}=1, P_{k_1,q_1}=\ldots=P_{k_r,q_r}=0\big)\le\mathbb{P}\left(\pi(i_1)=j_1,\dots,\pi(i_s)=j_s\right)=\frac{(n-s)!}{n!}$$

valid for any $\mathbf{i},\mathbf{j}\in\mathcal{E}_s$. For the lower-bound, we may assume that $k_1,\dots,k_r\notin\mathbf{i}$ and $q_1,\dots,q_r\notin\mathbf{j}$. Otherwise, there are less constraints and the probability of the event in question is larger. In this case, using the invariance (**??**) of the uniform measure on $S_{[n]}$, we can write

$$\mathbb{P}\begin{pmatrix}P_{i_1,j_1}=\ldots=P_{i_s,j_s}=1\\ P_{k_1,q_1}=\ldots=P_{k_r,q_r}=0\end{pmatrix}=\frac{(n-s)!}{n!}\mathbf{Q}\left(\pi(1)\neq 1,\dots,\pi(r)\neq r\right)$$

where $\mathbf{Q}$ is the uniform measure on $S_{[n-s]}$. In particular, we have the simple lower-bound

$$\mathbf{Q}\left(\pi(1)\neq 1,\dots,\pi(r)\neq r\right)\ge 1-r\mathbf{Q}\big(\pi(1)=1\big)=1-\frac{r}{n-s}.$$

This proves the claim. $\square$

6. The general case: tightness and trace asymptotics

6.1. **Proving the tightness of the sequence (χ_n) with the second moment argument.** In the case of a sum of $d>1$ permutations, proving the tightness of (χ_n) using the criteria given in Subsection 4.1 is more involved and necessitates a detailed technical analysis of the determinant expansion of $\mathrm{I}-zA$. We recall that $A=\sum_{q=1}^d P^{(q)}$, where $P^{(q)}, q\in[d]$ are $n\times n$ independent random permutation matrices. The subsequent computations crucially rely on the exchangeability of A as in (4.3).

Lemma 6.1. *For any $n\ge 2$, it holds for all $0<r<1$ and $1\le d<\sqrt{\frac{n(1-r)}{r}}$,*

$$\sum_{k\ge1}\frac{r^k}{d^k}\sum_{|\mathrm{I}|=k}\mathbb{E}\det(A(\mathrm{I})^2)\le\frac{2r(1+\frac{d^2}{n})}{1-r(1+\frac{d^2}{n})}.$$

Proof. First observe that

$$\sum_{|\mathrm{I}|=k} \mathbb{E}\det(A(\mathrm{I})^2) = \sum_{|\mathrm{I}|=k} \sum_{\sigma,\tau\in S(\mathrm{I})} \epsilon(\sigma)\epsilon(\tau)\mathbb{E}\prod_{i\in\mathrm{I}} A_{i,\sigma(i)} \prod_{j\in\mathrm{I}} A_{i,\tau(i)}.$$

Then, by symmetry of the permutation model,

$$\begin{aligned}
\sum_{|\mathrm{I}|=k} \mathbb{E}\det(A(\mathrm{I})^2) &= \binom{n}{k} \sum_{\sigma,\tau\in S_k} \epsilon(\sigma)\epsilon(\tau) \prod_{i\in[k]} \mathbb{E}A_{i,\sigma(i)}A_{i,\tau(i)} \\
&= \binom{n}{k} \sum_{\sigma,\tau\in S_k} \epsilon(\sigma^{-1}\tau)\mathbb{E} \prod_{i\in[k]} A_{i,i}A_{i,\sigma^{-1}\tau(i)} \\
&= \binom{n}{k} k! \sum_{\tau\in S_k} \epsilon(\tau)\mathbb{E} \prod_{i\in[k]} A_{i,i}A_{i,\tau(i)} \\
&= \frac{n!}{(n-k)!} \sum_{q_1,\dots,q_k\in[d]} \sum_{\ell_1,\dots,\ell_k\in[d]} \sum_{\tau\in S_k} \epsilon(\tau)\mathbb{E}\prod_{i\in[k]} P_{ii}^{(q_i)} P_{i\tau(i)}^{(\ell_i)}.
\end{aligned} \tag{6.1}$$

where we used (4.3) to obtain (6.1).

Observe that for any fixed $q, \ell \in [d]^k$ and all $i \in [k]$, if $q_i = \ell_i$ then

$$P_{i,i}^{(q_i)} P_{i,\tau(i)}^{(\ell_i)} = P_{i,i}^{(q_i)} \mathbf{1}\{\tau(i) = i\},$$

and using independence,

$$\begin{aligned}
\mathbb{E}\prod_{i\in[k]} P_{ii}^{(q_i)} P_{i\tau(i)}^{(\ell_i)} &= \prod_{r\in[d]} \mathbb{E}\Big[\prod_{i:q_i=r} P_{ii} \prod_{i:\ell_i=r} P_{i\tau(i)}\Big] \\
&= \prod_{r\in[d]} \mathbf{1}\big\{\tau(i) = i; \forall i \in \{q_i = \ell_i = r\}\big\}\mathbb{E}\Big[\prod_{i:q_i=r} P_{ii} \prod_{i\in\mathrm{K}_r} P_{i\tau(i)}\Big]
\end{aligned}$$

where $\mathrm{K}_r = \mathrm{K}_r(\ell, q) = \{i : \ell_i = r,\ q_i \neq r\}$.

Let $\mathrm{K} = \mathrm{K}(\ell, q) = \{i : q_i \neq \ell_i\} = \bigcup_{r=1}^d \mathrm{K}_r$ (disjoint) and for $r \in [d]$,

$$\Theta_r(\tau) = \mathbb{E}\Big[\prod_{i:q_i=r} P_{ii} \prod_{i\in\mathrm{K}_r} P_{i\tau(i)}\Big]\mathbf{1}\big\{\tau(i) \notin \{i : q_i = r\}; \forall i \in \mathrm{K}_r\big\}.$$

This implies that

$$\sum_{\tau\in S_k} \epsilon(\tau)\mathbb{E}\prod_{i\in[k]} P_{ii}^{(q_i)} P_{i\tau(i)}^{(\ell_i)} = \sum_{\tau\in S_\mathrm{K}} \epsilon(\tau) \prod_{r\in[d]} \Theta_r(\tau). \tag{6.2}$$

By (**??**), observe that for any permutation $\sigma \in S_\mathrm{K}$ which fixes the subsets $\{\mathrm{K}_r\}_{r\in[d]}$,

$$\prod_{r\in[d]} \Theta_r(\sigma\tau) = \prod_{r\in[d]} \Theta_r(\tau).$$

Let

$$\delta_r = \#\mathrm{K}_r = \#\{i : \ell_i = r, q_i \neq \ell_i\}.$$

If $\delta_r \geq 2$ for some $r \in [d]$, by a transposition, one has

$$\sum_{\tau\in S_\mathrm{K}} \epsilon(\tau) \prod_{r\in[d]} \Theta_r(\tau) = 0.$$

In particular, this shows that for any fixed $q, \ell \in [d]^k$,

$$\Big|\sum_{\tau\in S_\mathrm{K}} \epsilon(\tau) \prod_{r\in[d]} \Theta_r(\tau)\Big| \leq \Delta! \prod_{r\in[d]} \mathbf{1}\big\{\delta_r \in \{0,1\}\big\}\frac{(n - \#\{i : q_i = r\} - \delta_r)!}{n!}, \qquad \Delta = \sum_{r\in[d]} \delta_r = \#\mathrm{K}.$$

Observe that for any integers $k_1 \geq 1, k_2 \geq 0$, $\delta_1, \delta_2 \in \{0,1\}$,

$$(6.3) \qquad \frac{(n-k_1-\delta_1)!}{(n-k_1-k_2)!} \leq \begin{cases} \frac{n!}{(n-k_2-\delta_2)!}\frac{1}{n^{\delta_1+\delta_2}}, & \delta_1+\delta_2 \leq 1, \\ \frac{n!}{(n-k_2-\delta_2)!}\frac{2}{n^{\delta_1+\delta_2}}, & \delta_1=\delta_2=1. \end{cases}$$

Let $k_r(q) = \#\{i : q_i = r\}$. We have $k_1 + \cdots + k_r = k$. By symmetry, we can assume $k_1 \geq 1$, and by induction, we obtain

$$\begin{aligned} \prod_{r\in[d]} \frac{(n-k_r-\delta_r)!}{n!} &\leq \frac{2}{n^{\delta_1+\delta_2}} \frac{(n-k_1-k_2)!}{n!} \prod_{r\geq 3} \frac{(n-k_r-\delta_r)!}{n!} \\ &\leq \frac{(n-k)!}{n!}\frac{2}{n^{\Delta}} \end{aligned}$$

where $\Delta = \sum_{r\in[d]} \delta_r$. We conclude that

$$\Big| \sum_{\tau\in S_{\mathrm{K}}} \epsilon(\tau) \prod_{r\in[d]} \Theta_r(\tau) \Big| \leq \frac{2\Delta!(n-k)!}{n!n^{\Delta}} \prod_{r\in[d]} \mathbf{1}\{\delta_r \in \{0,1\}\}, \qquad \Delta = \sum_{r\in[d]} \delta_r = \#\mathrm{K}.$$

Going back to (6.1), using formula (6.2), this argument shows that

$$\sum_{|\mathrm{I}|=k} \mathbb{E}\det(A(\mathrm{I})^2) \leq 2 \sum_{q_1,\dots,q_k\in[d]} \sum_{\ell_1,\dots,\ell_k\in[d]} \frac{\Delta!}{n^{\Delta}} \prod_{r\in[d]} \mathbf{1}\{\delta_r \in \{0,1\}\}.$$

For $\ell \in [d]^k$, let us denote $k_r = k_r(\ell) = \#\{i : \ell_i = r\}$ for $r \in [d]$. Observe that given $k_1, \dots, k_r$, there are $\binom{k}{k_1,\dots,k_d}$ many choices of $\ell \in [d]^k$, and for $\delta \in \{0,1\}^d$, there are at most $d^{\Delta} \prod_{r=1}^d k_r^{\delta_r}$ configurations $q \in [d]^k$ which contributes to the previous sum. Hence,

$$\begin{aligned} \sum_{|\mathrm{I}|=k} \mathbb{E}\det(A(\mathrm{I})^2) &\leq 2 \sum_{k_1,\dots,k_d\geq 0} \binom{k}{k_1,\cdots,k_d} \sum_{\delta_1,\dots,\delta_d\in\{0,1\}} \prod_{r\in[d]} \left(\frac{k_r d^2}{n}\right)^{\delta_r} \\ &= 2 \sum_{k_1,\dots,k_d\geq 0} \binom{k}{k_1,\cdots,k_d} \prod_{r\in[d]} \left(1 + \frac{k_r d^2}{n}\right) \\ &= 2d^k \big(1 + \tfrac{d}{n}\partial_x\big)^d x^k \big|_{x=1}. \end{aligned} \qquad (6.4)$$

The last identity follows from the fact that the second display equals (up to a factor 2)

$$\begin{aligned} \sum_{k_1,\dots,k_d\geq 0} \binom{k}{k_1,\cdots,k_d} \prod_{r\in[d]} (1+\epsilon\partial_x) x^{k_r}\big|_{x=1} &= \prod_{r\in[d]} (1+\epsilon\partial_{x_r}) \Big(\sum_{k_1,\dots,k_d\geq 0} \binom{k}{k_1,\cdots,k_d} \prod_{r\in[d]} x_r^{k_r} \Big)\Big|_{x_r=1} \\ &= \prod_{r\in[d]} (1+\epsilon\partial_{x_r}) \big(x_1+\cdots+x_r\big)^k\big|_{x_1=\cdots=x_r=1} \\ &= \big(1+\epsilon\partial_x\big)^d x^k\big|_{x=d} \\ &= d^k \big(1+\tfrac{\epsilon}{d}\partial_x\big)^d x^k\big|_{x=1} \end{aligned}$$

by scaling, and taking $\epsilon = d^2/n$. By summing over all $k \in \mathbb{N}_0$, we conclude that for any $0 < r < 1$,

$$\begin{aligned}\sum_{k\in\mathbb{N}} \frac{r^k}{d^k} \sum_{|\mathrm{I}|=k} \mathbb{E}\det(A(\mathrm{I})^2) &\le 2\big(1+\tfrac{d}{n}\partial_x\big)^d \frac{1}{1-xr}\Big|_{x=1} \\ &= 2\sum_{\ell\ge 0}\binom{d}{\ell}\frac{d^\ell}{n^\ell}\partial_x^\ell \frac{1}{1-xr}\Big|_{x=1} \\ &\le 2\sum_{\ell=0}^{d} \frac{d!}{(d-\ell)!}\frac{(dr)^\ell}{n^\ell}\frac{1}{(1-r)^{\ell+1}} \\ &\le \frac{2}{1-r}\sum_{\ell=0}^{\infty}\left(\frac{d^2 r}{n(1-r)}\right)^\ell \le \frac{2}{1-r}\cdot\frac{1}{1-\frac{d^2r}{n(1-r)}}.\end{aligned}$$

The previous sum converges if $\frac{d^2 r}{n(1-r)} < 1$. This completes the proof. $\square$

Lemma 6.2. *For any $n \ge 2$, it holds for all $0 < r < 1$ and $1 \le d < \sqrt{\frac{n(1-r)}{r}}$,*

$$\sum_{k\ge 1}\frac{r^k}{d^{k+1}}\Big|\sum_{|\mathrm{I}|=|\mathrm{J}|=k,|\mathrm{I}\cap\mathrm{J}|=k-1}\mathbb{E}\det A(\mathrm{I})A(\mathrm{J})\Big| \le \frac{r}{(1-r-\frac{rd^2}{n})^2}.$$

Proof. Using the invariance (**??**) of the random matrix A, for $k \le n-1$, we can rewrite

$$\sum_{|\mathrm{I}|=|\mathrm{J}|=k,|\mathrm{I}\cap\mathrm{J}|=k-1}\mathbb{E}\det A(\mathrm{I})A(\mathrm{J}) = \binom{n}{k}k(n-k)\sum_{\substack{\tau\in S_k\otimes S_1\\ \sigma\in S_1\otimes S_k}}\epsilon(\sigma)\epsilon(\tau)\mathbb{E}\prod_{j=1}^{k}\prod_{i=2}^{k+1}A_{i,\sigma(i)}A_{j,\tau(j)} \tag{6.5}$$

where $S_k\otimes S_1 = \{\tau\in S_{k+1} : \tau(k+1) = k+1\}$. and $S_1\otimes S_k = \{\sigma \in S_{k+1} : \sigma(1) = 1\}$. Now (6.5) can be written as

$$\begin{aligned}&= \binom{n}{k}k(n-k)\sum_{\substack{\tau\in S_k\otimes S_1\\ \sigma\in S_1\otimes S_k}}\epsilon(\sigma)\epsilon(\tau)\sum_{\ell_1,\dots,\ell_k}\sum_{q_2,\dots,q_{k+1}}\mathbb{E}\prod_{i=2}^{k+1}P^{(q_i)}_{i,\sigma(i)}\prod_{j=1}^{k}P^{(\ell_j)}_{j,\tau(j)} \\ &= \binom{n}{k}k(n-k)\sum_{\ell_1,\dots,\ell_k}\sum_{q_2,\dots,q_{k+1}}\sum_{\substack{\tau\in S_k\otimes S_1\\ \sigma\in S_1\otimes S_k}}\epsilon(\sigma)\epsilon(\tau)\mathbf{1}_{\{\tau(j)=\sigma(j),\forall j\in[k]\setminus\mathrm{K}\}}\prod_{r\in[d]}\mathbb{E}\Big[\prod_{i\ge 2:q_i=r}P_{i,\sigma(i)}\prod_{j\in\mathrm{K}:\ell_j=r}P_{j\tau(j)}\Big],\end{aligned}$$

where $\mathrm{K} = \mathrm{K}(\ell,q) = \{j \in [k] : q_j \ne \ell_j\}$. By convention $1 \in \mathrm{K}$ and by symmetry we may assume that $l_1 = 1$ and multiply the sum with an extra factor d. Moreover, by the argument following (6.2), we must have for all $r \in \{2,\dots,d\}$,

$$\delta_r = \#\{j\in\mathrm{K} : \ell_j = r\} \in \{0,1\}$$

for otherwise the sum over all permutation $\tau \in S_k$ yields o by a transposition. Similarly, we must have

$$\delta_1 = \#\{j\in\mathrm{K} : \ell_j = 1\} = 1$$

since $1 \in \mathrm{K}$. Then, for any configuration $q,\ell \in [d]^k$ with $\ell_1 = 1$, following the argument in (6.3), it holds for any $\sigma,\tau \in S_{k+1}$

$$\begin{aligned}\prod_{r\in[d]}\mathbb{E}\Big[\prod_{i\ge 2:q_i=r}P_{i,\sigma(i)}\prod_{j\in\mathrm{K}:\ell_j=r}P_{j,\tau(j)}\Big] &\le \prod_{r\in[d]}\frac{(n-\#\{i\ge 2: q_i=r\}-\delta_r)!}{n!} \\ &= \frac{(n-k_1-1)!}{n!}\prod_{r\ge 2}\frac{(n-k_r-\delta_r)!}{n!} \\ &\le \frac{(n-k-1)!}{n!\,n^\Delta}\end{aligned}$$

where $\Delta = \sum_{r=2}^{d} \delta_r$ and $k_r = k_r(q) = \#\{i \geq 2 : q_i = r\}$ form a partition of k for $r \in [d]$. Moreover, for any $\Delta \leq d-1$ and any given $\sigma \in S_1 \otimes S_k$,

$$\#\big\{\tau \in S_k : \tau(j) = \sigma(j), \forall j \in [k] \setminus \mathrm{K}\big\} \leq (\Delta+1)! \leq d^{\Delta},$$

so that this argument shows that

$$\binom{n}{k} k(n-k) \sum_{\ell_1,\ldots,\ell_k} \sum_{q_2,\ldots,q_{k+1}} \sum_{\substack{\tau \in S_k \otimes S_1 \\ \sigma \in S_1 \otimes S_k}} \epsilon(\sigma)\epsilon(\tau) \mathbf{1}_{\{\tau(j)=\sigma(j), \forall j \in [k] \setminus \mathrm{K}\}} \prod_{r \in [d]} \mathbb{E}\Big[\prod_{i \geq 2 : q_i = r} P_{i,\sigma(i)} \prod_{j \in \mathrm{K} : \ell_j = r} P_{j\tau(j)}\Big]$$

$$\leq dk \sum_{q_2,\ldots,q_{k+1} \in [d]} \sum_{l_2,\ldots,l_k \in [d]} \frac{d^{\Delta}}{n^{\Delta}} \mathbf{1}\big\{\delta_1 = 1\big\} \prod_{r=2}^{d} \mathbf{1}\big\{\delta_r \in \{0,1\}\big\}.$$

Let $k_r = \#\{i \geq 2 : \ell_i = r\}$ for $r \in [d]$. There are $\binom{k-1}{k_1,\ldots,k_d}$ many choices of $\ell \in [d]^k$ with $\ell_1 = 1$. As in the proof of Lemma 6.1, for a given $l \in [d]^k$ with $\ell_1 = 1$, there are at most

$$d^{\Delta+1} \prod_{r=2}^{d} k_r^{\delta_r}$$

configurations $(q_2, \cdots, q_{k+1}) \in [d]^k$ which contribute to the previous sum; this is because there are $d^{\Delta} \prod_{r=2}^{d} k_r^{\delta_r}$ many choices for $(q_2, \ldots, q_k)$ and at most d choices for q_{k+1}. Hence,

$$\begin{aligned}
&dk \sum_{q_2,\ldots,q_{k+1} \in [d]} \sum_{\ell_2,\ldots,\ell_k \in [d]} \frac{d^{\Delta}}{n^{\Delta}} \mathbf{1}\big\{\delta_1 = 1\big\} \prod_{r=2}^{d} \mathbf{1}\big\{\delta_r \in \{0,1\}\big\} \\
&\leq d^2 k \sum_{k_1,\ldots,k_d \geq 0} \binom{k-1}{k_1,\ldots,k_d} \sum_{\delta_2,\ldots,\delta_d \in \{0,1\}} \prod_{r=2}^{d} \Big(\frac{k_r d^2}{n}\Big)^{\delta_r} \\
&\leq k d^{k+1} \big(1 + \tfrac{d}{n}\partial_x\big)^d x^{k-1}\big|_{x=1}.
\end{aligned}$$

The last inequality follows as in the proof of Lemma 6.1 (replacing k by $k-1$).

Summing over all $k \in \mathbb{N}$, we conclude that for any $0 < r < 1$,

$$\begin{aligned}
&\sum_{k \geq 1} \frac{r^k}{d^{k+1}} \Big| \sum_{|\mathrm{I}|=|\mathrm{J}|=k, |\mathrm{I} \cap \mathrm{J}|=k-1} \mathbb{E} \det A(\mathrm{I})A(\mathrm{J}) \Big| \\
&\leq \sum_{k \geq 1} k r^k \big(1 + \tfrac{d}{n}\partial_x\big)^d x^{k-1}\Big|_{x=1} = \big(1 + \tfrac{d}{n}\partial_x\big)^d \frac{r}{(1-rx)^2}\Big|_{x=1} \\
&= \sum_{\ell=0}^{d} \binom{d}{\ell} \frac{d^{\ell}}{n^{\ell}} \partial_x^{\ell} \frac{r}{(1-xr)^2}\Big|_{x=1} \\
&= r \sum_{\ell=0}^{d} \frac{d!}{(d-\ell)!\ell!} \frac{(rd)^{\ell}}{n^{\ell}} \frac{(\ell+1)!}{(1-r)^{\ell+2}} && (6.6) \\
&\leq \frac{r}{(1-r)^2} \sum_{\ell=0}^{\infty} (\ell+1) \Big(\frac{d^2 r}{n(1-r)}\Big)^{\ell} = \frac{r}{(1-r-\frac{rd^2}{n})^2}. && (6.7)
\end{aligned}$$

This completes the proof. $\square$

6.2. **Asymptotics of traces of sums of d independent uniform permutations.** We now turn to the identification of the limits of $\mathrm{trace}(A^k)$ stated in Theorem 3.1.

6.2.1. *Subgraph probability estimation.* Given integers $0 \leqslant k_1, \ldots, k_r \leqslant d$ with $k = k_1 + \cdots + k_r$, define

$$\mathscr{T}_{\mathbf{k}} = \left\{T \in \{0,1\}^{d\times r} : \textstyle\sum_{i=1}^{d} T_{ij} = k_j \text{ for all } j \in [r]\right\}.$$

Observe that

$$|\mathscr{T}_{\mathbf{k}}| = \binom{d}{k_1} \cdots \binom{d}{k_r} \leq \frac{d^k}{k_1! \cdots k_r!}. \tag{6.8}$$

Proposition 6.3. *Fix $\mathbf{i}, \mathbf{j} \in [n]^r$ for $r \in \mathbb{N}$ such that $(i_1, j_1), \cdots, (i_r, j_r)$ are distinct. Let $\mathbf{k} \in [d]^r$ and set $k = k_1 + \cdots + k_r$. Then,*

$$\mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) \leqslant |\mathscr{T}_{\mathbf{k}}| \frac{(n-k)!}{n!}. \tag{6.9}$$

Proof. Decomposing the event $\{A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r\}$ in terms of the permutations $P^1, \ldots, P^d$ leads to

$$\mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) = \sum_{T \in \mathscr{T}} \mathbb{P}\left(\begin{bmatrix} P^1_{i_1,j_1} & \cdots & P^1_{i_r,j_r} \\ \vdots & & \vdots \\ P^d_{i_1,j_1} & \cdots & P^d_{i_r,j_r} \end{bmatrix} = T\right).$$

Since $P^1, \ldots, P^d$ are independent, we obtain

$$\mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) = \sum_{T \in \mathscr{T}} \prod_{\delta=1}^{d} \mathbb{P}(P_{i_1,j_1} = T_{\delta,1}, \ldots, P_{i_r,j_r} = T_{\delta,r}).$$

Each term in this product falls under Proposition 5.2 – either it equals 0 or (5.7) holds. Let $s_\delta = T_{\delta,1} + \cdots + T_{\delta,r}$ for $\delta \in [d]$, this implies that there exists $0 \leq \varepsilon_\delta \leqslant \frac{r}{n-r}$ for $\delta \in [d]$ so that

$$\mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) \leq \sum_{T \in \mathscr{T}} \prod_{\delta=1}^{d} (1-\varepsilon_\delta) \frac{(n-s_\delta)!}{n!}, \quad \text{with equality if } \mathbf{i}, \mathbf{j} \in \mathcal{E}_r. \tag{6.10}$$

Note that by definition of $\mathscr{T}$, $\sum_{\delta=1}^{d} s_\delta = \sum_{j=1}^{r} k_j = k$, this yields the upper-bound,

$$\mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) \leq \sum_{T \in \mathscr{T}} \prod_{\delta=1}^{d} \frac{(n-s_\delta)!}{n!} \leq |\mathscr{T}| \frac{(n-k)!}{n!}$$

where we used (3.4). $\square$

Proposition 6.4. *Fix $\mathbf{i}, \mathbf{j} \in \mathcal{E}_r$ for $r \in \mathbb{N}$. Let $\theta = d/n$. There exists $|\varepsilon| \leq \frac{dr}{n-r} \wedge \frac{2r^2}{n}$ such that*

$$\mathbb{P}\big(A_{i_1,j_1} = \cdots = A_{i_r,j_r} = 1\big) = (1+\varepsilon)\theta^r. \tag{6.11}$$

Proof. By Proposition 6.3 and using the fact that $|\mathscr{T}_{\mathbf{k}}| = d^r$ for $\mathbf{k} = (1, \ldots, 1)$, we already have the upper-bound. Note that we can bound $\frac{(n-r)!}{n!} \leq \frac{1+2r^2/n}{n^r}$ for $r \leq n$. For the lower bound, by formula (6.10) in case $\mathbf{i}, \mathbf{j} \in \mathcal{E}_r$,

$$\mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) \geq (1-\varepsilon)^d \sum_{T \in \mathscr{T}} \prod_{\delta=1}^{d} \frac{(n-s_\delta)!}{n!}$$

with $\varepsilon = \frac{r}{n-r}$. By convexity, $(1-\varepsilon)^d \geq 1 - d\epsilon$. Then, using that in this case $\sum_{\delta=1}^{d} s_\delta = \sum_{j=1}^{r} k_j = r$, this implies that

$$\mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) \geq (1-d\varepsilon)|\mathscr{T}|n^{-r}.$$

This concludes the proof. $\square$

As a direct consequence of Proposition 6.4, we have the asymptotics of the probability of a simple cycle on the digraph of the random matrix A. We now turn to bounding joint moments of the entries of the random matrix A.

Proposition 6.5. *For $r \in \mathbb{N}$, fix $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^r$ and let $\beta = \beta_1 + \cdots + \beta_r$. Let $\theta = d/n$ and assume that $d \leq \sqrt{n}$. Then, for any $\mathbf{i}, \mathbf{j} \in [n]^r$,*

$$\mathbb{E}\big[A^{\alpha_1}_{i_1,j_1} \cdots A^{\alpha_r}_{i_r,j_r} \mathbf{1}_{\{A_{i_\ell,j_\ell} \geq \beta_\ell, \ell \in [r]\}}\big] = O_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\theta^\beta).$$

Proof. Without loss of generality, we may assume that $(i_1, j_1), \cdots, (i_r, j_r)$ are distinct. By Proposition 6.3 and (6.8), we have

$$\begin{aligned}\mathbb{E}\big[A_{i_1,j_1}^{\alpha_1} \cdots A_{i_r,j_r}^{\alpha_r} \mathbf{1}_{\{A_{i_\ell,j_\ell} \geq \beta_\ell, \ell \in [r]\}}\big] &= \sum_{\mathbf{k} \in [d]^r, \mathbf{k} \geq \boldsymbol{\beta}} k_1^{\alpha_1} \cdots k_r^{\alpha_r} \mathbb{P}(A_{i_1,j_1} = k_1, \ldots, A_{i_r,j_r} = k_r) \\ &\leq \sum_{\mathbf{k} \in [d]^r, \mathbf{k} \geq \boldsymbol{\beta}} \frac{k_1^{\alpha_1} \cdots k_r^{\alpha_r}}{k_1! \cdots k_r!} d^k \frac{(n-k)!}{n!}; \qquad k = k_1 + \cdots + k_r.\end{aligned}$$

We can bound $\frac{(n-k)!}{n!} \leq \frac{\mathrm{e}^{2k^2/n}}{n^k}$ for $k \leq n$. Since r is fixed, this yields the estimate valid for $d \leq \sqrt{n}$,

$$\mathbb{E}\big[A_{i_1,j_1}^{\alpha_1} \cdots A_{i_r,j_r}^{\alpha_r}\big] \leq \mathrm{e}^{2r^2} \sum_{\mathbf{k} \geq \boldsymbol{\beta}} \prod_{j=1}^{r} \frac{k_j^{\alpha_j}}{k_j!} \theta^{k_j}$$

Now, we can use the bound for $\theta \leq 1$,

$$\sum_{k \geq b} \frac{k^a}{k!} \theta^k \leq C_{a,b} \theta^b,$$

to conclude the proof. $\square$

6.2.2. *Reducing traces to cycle counts in the digraph of A.* Recall that

$$\mathrm{tr}(A^k) = \sum_{\mathbf{i} \in [n]^k} A_{\mathbf{i}}.$$

In particular, by Proposition 6.5 applied with $\alpha = k$,

$$\mathbb{E}\mathrm{tr}(A^k) \leq C_k n^k \theta^k = C_k d^k.$$

In particular, in the regime where d is fixed, for any $k \in \mathbb{N}$, the non-negative random variables $\big(\mathrm{tr}(A^k)\big)_{n \in \mathbb{N}}$ are tight.

For every $\mathbf{i} = (i_1, \ldots, i_k) \in [n]^k$, we can associate a digraph

$$\begin{aligned} V(\mathbf{i}) &= \{i_1, \ldots, i_k\} & v(\mathbf{i}) &= \#V(\mathbf{i}) \\ E(\mathbf{i}) &= \{(i_1, i_2), (i_2, i_3), \ldots, (i_{k-1}, i_k), (i_k, i_1)\} & \mathrm{e}(\mathbf{i}) &= \#E(\mathbf{i}). \end{aligned}$$

One can interpret $v(\mathbf{i})$ as the number of 'vertices' of $[n]$ that are visited by $\mathbf{i}$, and $\mathrm{e}(\mathbf{i})$ as the number of distinct 'edges' of $\mathbf{i}$. The digraph $(V(\mathbf{i}), E(\mathbf{i}))$ contains at least a loop, so necessarily $v(\mathbf{i}) \leqslant \mathrm{e}(\mathbf{i}) \leq k$.

For integers $v \leqslant \mathrm{e} \leqslant k$, we define

$$\mathscr{E}_k(v, \mathrm{e}) = \{\mathbf{i} \in [n]^k : v(\mathbf{i}) = v, \mathrm{e}(\mathbf{i}) = \mathrm{e}\}. \tag{6.12}$$

Let us decompose

$$\mathrm{tr}(A^k) = \sum_{\substack{\mathbf{i} \in \mathscr{E}_k(v,\mathrm{e}) \\ v = \mathrm{e}}} \mathbf{1}_{A_{\mathbf{i}} = 1} + \sum_{\substack{\mathbf{i} \in \mathscr{E}_k(v,\mathrm{e}) \\ v = \mathrm{e}}} A_{\mathbf{i}} \mathbf{1}_{A_{\mathbf{i}} > 1} + \sum_{\substack{\mathbf{i} \in \mathscr{E}_k(v,\mathrm{e}) \\ v < \mathrm{e}}} A_{\mathbf{i}} =: T_k + R_k + S_k. \tag{6.13}$$

First, let us show that both R_k and S_k can be treated as negligible errors.

Lemma 6.6. *Suppose that $d = n^{o(1)}$ as $n \to \infty$. For any fixed $k \in \mathbb{N}$, $(R_k + S_k) \to 0$ in probability as $n \to \infty$.*

Proof. For any $\mathbf{i} \in [n]^k$, by Proposition 6.5 with k fixed,

$$\mathbb{E}A_{\mathbf{i}} \leq C_k \theta^{\mathrm{e}(\mathbf{i})}, \qquad \theta = d/n,$$

since there are e different entries and we can take $\boldsymbol{\beta} = \underbrace{(1, \ldots, 1)}_{\times \mathrm{e}}$.

Since $|\mathscr{E}_k(v, \mathrm{e})| \leq k! n^v$, this shows that

$$\mathbb{E}S_k \leq \sum_{v < \mathrm{e} \leq k} C_k n^v \theta^{\mathrm{e}} \leq C_k \frac{d^k}{n}.$$

We conclude that if t $d = n^{o(1)}$, then $\mathbb{E}S_k \to 0$ as $n \to \infty$.

Similarly, by Proposition 6.5,

$$\mathbb{E}A_{\mathbf{i}}\mathbf{1}_{A_{\mathbf{i}}>1} \le C_k\theta^{\mathrm{e}(\mathbf{i})+1}$$

so that

$$\mathbb{E}R_k \le \sum_{v\le k} C_k n^v\theta^{v+1} \le C_k\frac{d^{k+1}}{n}.$$

Hence, also $\mathbb{E}R_k \to 0$ as $n\to\infty$. □

Lemma 6.7. *$\mathscr{E}_k(v,v)$ is empty if v is not a divisor of k. Otherwise, if $k = vq$, then the elements of $\mathscr{E}_k(v,v)$ are exactly the sequences*

(6.14) $$(i_1,i_2,\ldots,i_v,i_1,\ldots,i_v,\ldots,i_1,\ldots,i_v)$$

where the subsequence $\mathbf{i}' = (i_1,\ldots,i_v) \in \mathcal{E}_v$ is repeated q times. Moreover, the events $\{A_{\mathbf{i}}=1\} = \{A_{\mathbf{i}'}=1\}$.

Proof. Same as in [15, Lemma 9.3]. □

Lemma 6.7 implies that according to (6.13),

$$T_k = \sum_{v|k}\sum_{\mathbf{i}'\in\mathscr{E}_v}\mathbf{1}_{A_{\mathbf{i}'}=1}$$

Let us denote for $\ell\in[n]$,

(6.15) $$\mathcal{C}_\ell := \big\{\mathbf{i} = (i_1,\ldots,i_\ell) \text{ modulo cyclic permutation} : i_k\in[n],\ \text{distincts}\big\}$$

and the random variables

(6.16) $$Q_\ell := \sum_{\mathbf{i}\in\mathcal{C}_\ell}\mathbf{1}_{A_{\mathbf{i}}=1}.$$

With this new notation,

(6.17) $$T_k = \sum_{\ell|k}\ell Q_\ell.$$

The interpretation is that Q_ℓ is the number of (oriented) ℓ-cycles on the digraph defined by the random matrix A and T_k is a good approximation for $\mathrm{tr}(A^k)$.

6.2.3. *Asymptotics of joint moments of cycle counts.* If $Q\in\mathbb{N}$, we denote its falling factorials by

$$(Q)_r = Q(Q-1)\cdots(Q-r+1), \qquad r\in\mathbb{N}.$$

Our interest in these quantities stems from the fact that if $Q = |\mathcal{C}|$ for a finite set $\mathcal{C}$, then

(6.18) $$(Q)_r = \big|\{(j_1,\ldots,j_r) : j_k\in\mathcal{C} \text{ distincts}\}\big|.$$

and the following basic probabilistic result.

Lemma 6.8. *For $\lambda>0$, $\Lambda\sim\mathrm{Poisson}(\lambda)$ if and only if for all $r\in\mathbb{N}$,*

$$\mathbb{E}(\Lambda)_r = \lambda^r.$$

Proposition 6.9. *Recall the notation (6.16) and fix $k\in\mathbb{N}$ and $\boldsymbol{\alpha}\in\mathbb{N}_0^k$. Then, as $n\to\infty$*

$$\mathbb{E}\left[(Q_1)_{\alpha_1}\cdots(Q_k)_{\alpha_k}\right] = \left(\frac{d^1}{1}\right)^{\alpha_1}\cdots\left(\frac{d^m}{m}\right)^{\alpha_m}\big(1+O_{\boldsymbol{\alpha}}(d/n)\big).$$

The proof of Proposition 6.9 will be given in Section 6.3.2. It relies on Proposition 6.4 and basic combinatorial arguments counting certain collections of (distinct) cycles on the digraph of the random matrix A. To illustrate the argument, we compute the first and second moment of the random variable Q_ℓ in Section 6.3.1.

According to Lemma 6.8, Proposition 6.9 directly implies that for a fixed $d\in\mathbb{N}$, for any fixed $\boldsymbol{\ell},\boldsymbol{\alpha}\in\mathbb{N}^k$,

(6.19) $$\lim_{n\to\infty}\mathbb{E}\left[(Q_{\ell_1})_{\alpha_1}\cdots(Q_{\ell_k})_{\alpha_k}\right] = \mathbb{E}\left[(\Lambda_{\ell_1})_{\alpha_1}\cdots(\Lambda_{\ell_k})_{\alpha_k}\right]$$

where the Poisson random variables $\{\Lambda_\ell\}_{\ell\in\mathbb{N}}$ are given by Definition 2.1.

Proof of (3.2). The joint convergence of the factorial moments described in (6.19) implies that for every $k \in \mathbb{N}$,

$$(Q_1, \dots, Q_k) \xrightarrow[n\to\infty]{\text{law}} (\Lambda_1, \dots, \Lambda_k)$$

and in the sense of moments. Then, by the Cramér-Wold theorem, for every $k \in \mathbb{N}$,

$$\Big(T_1 = Q_1, \cdots, T_k = \sum_{\ell | k} \ell Q_\ell\Big) \xrightarrow[n\to\infty]{\text{law}} \Big(\Lambda_1, \dots, \sum_{\ell|k} \ell \Lambda_\ell\Big).$$

By Lemma 6.6, for every $k \in \mathbb{N}$ and any $\epsilon > 0$,

$$\mathbb{P}\big(|\mathrm{tr}(A) - T_1| + \cdots + |\mathrm{tr}(A^k) - T_k| \geq \epsilon\big) \xrightarrow[n\to\infty]{} 0 \tag{6.20}$$

This establishes that for any fixed $d \in \mathbb{N}$ and $k \in \mathbb{N}$

$$\big(\mathrm{tr}(A), \cdots, \mathrm{tr}(A^k)\big) \xrightarrow[n\to\infty]{\text{law}} \Big(\Lambda_1, \dots, \sum_{\ell|k} \ell \Lambda_\ell\Big).$$

Proof of (3.3). In the regime where the degree $d = d(n) \to \infty$ so that $d = n^{o(1)}$ as $n \to \infty$, the error term in Proposition 6.9 is so good that we can directly show that the random variables $(Q_\ell)_{\ell\in\mathbb{N}}$, if suitably normalized, converge to independent Gaussians in the sense of finite dimensional distributions.

Namely, the asymptotics from Proposition 6.9 yields the conditions (A.4)–(A.5) with

$$\lambda_{i,n} = d(n)^i / i \qquad \text{and} \qquad \epsilon_{\mathbf{k},n} = O_{\mathbf{k}}(d(n)/n).$$

In particular since for every $k \in \mathbb{N}$, $d(n)^k/n \to 0$ as $n \to \infty$, by Lemma A.1, we conclude that for any $\ell \in \mathbb{N}$

$$\left(\frac{Q_1 - d}{\sqrt{d/1}}, \dots, \frac{Q_\ell - d^\ell/\ell}{\sqrt{d^\ell/\ell}}\right) \xrightarrow[n\to\infty]{\text{law}} \big(N_1, \dots, N_\ell\big), \tag{6.21}$$

as well as in the sense of moments, where $\{N_\ell\}_{\ell\in\mathbb{N}}$ are as in Definition 2.3.

Observe that according to (6.17), we can write for $k \in \mathbb{N}$,

$$\frac{T_k - d^k}{d^{k/2}} = \sum_{\ell|k} \sqrt{\ell} d^{-(k-\ell)/2} \left(\frac{Q_\ell - d^\ell/\ell}{\sqrt{d^\ell/\ell}}\right) + \sum_{\ell|k, \ell<k} d^{\ell-k/2} \tag{6.22}$$

Hence, by (6.21) and the Cramér-Wold theorem, for every $k \in \mathbb{N}$,

$$\left(\frac{T_1 - d}{\sqrt{d}}, \dots, \frac{T_{2k} - d^{2k}}{\sqrt{d^{2k}}}\right) \xrightarrow[n\to\infty]{\text{law}} \big(N_1, \sqrt{2}N_2 + 1, \sqrt{3}N_3, \dots, \sqrt{2k}N_{2k} + 1\big).$$

Note that in (6.22), as $d(n) \to \infty$, only the term $\ell = k$ contributes to the random and the term $\ell = k/2$ contributes to the mean if $k \in 2\mathbb{N}$. In particular, once normalized the weak limit of T_k are still independent Gaussians.

To finish the proof of (3.3), by (6.13) and Lemma 6.6 which still holds in the regime where $d = n^{o(1)}$, for a fixed $k \in \mathbb{N}$,

$$\left(\frac{\mathrm{tr}(A^1) - d}{\sqrt{d}}, \dots, \frac{\mathrm{tr}(A^{2k}) - d^{2k}}{\sqrt{d^{2k}}}\right) = \left(\frac{T_1 - d}{\sqrt{d}}, \dots, \frac{T_{2k} - d^{2k}}{\sqrt{d^{2k}}}\right) + \underset{n\to\infty}{o(1)}$$

where the error is controlled as in (6.20). This completes the proof of Theorem 3.1.

6.3. **Proof of Theorem 6.9.** We now prove Theorem 6.9. We first prove the theorem only for first and second moments, to give a flavour of the proof. The complete proof is in Subsection 6.3.2.

6.3.1. *First and second moment.* Let us first prove that

$$\mathbb{E}[Q_i] = \frac{d^i}{i}(1 + O_i(d/n)) \tag{6.23}$$

and that

$$\mathbb{E}[(Q_i)_2] = \left(\frac{d^i}{i}\right)^2 (1 + O_i(d/n)). \tag{6.24}$$

By (6.16) and since $|\mathcal{C}_\ell| = (n)_\ell/\ell$, according to Proposition 6.4, there exists $\varepsilon = O_\ell(d/n)$ so that

$$\mathbb{E}Q_\ell = |\mathcal{C}_\ell|(1+\varepsilon)(d/n)^\ell.$$

For fixed $\ell \in \mathbb{N}$, this shows that as $n \to \infty$,

$$\mathbb{E}Q_\ell = d^\ell/\ell\big(1 + O_\ell(d/n)\big).$$

For the second (factorial) moment, by (6.18),

$$\mathbb{E}(Q_\ell)_2 = \sum_{\mathbf{i},\mathbf{j}\in\mathcal{C}_\ell, \mathbf{i}\neq\mathbf{j}} \mathbb{P}\big(A_\mathbf{i} = 1, A_\mathbf{j} = 1\big).$$

If we identify $\mathbf{i}, \mathbf{j}$ with ℓ-cycles, observe that the condition $\mathbf{i} \neq \mathbf{j}$ implies that the digraph $\{\mathbf{i}, \mathbf{j}\}$ obtained by concatenating $\mathbf{i}, \mathbf{j}$,

$$\begin{aligned} &V(\{\mathbf{i},\mathbf{j}\}) = \{i_1, \dots, i_\ell, j_1, \dots, j_\ell\} \\ &E(\{\mathbf{i},\mathbf{j}\}) = \{(i_1,i_2),(i_2,i_3),\dots,(i_{k-1},i_\ell),(i_\ell,i_1),(j_1,j_2),(j_2,j_3),\dots,(j_{k-1},j_\ell),(j_\ell,j_1)\} \end{aligned} \tag{6.25}$$

satisfies

$$\mathbf{i} \cap \mathbf{j} = \emptyset \quad \text{if and only if} \quad |V(\{\mathbf{i},\mathbf{j}\})| = |E(\{\mathbf{i},\mathbf{j}\})|. \tag{6.26}$$

In particular, we can split

$$\mathbb{E}(Q_\ell)_2 = \sum_{\mathbf{i},\mathbf{j}\in\mathcal{C}_\ell, \mathbf{i}\cap\mathbf{j}=\emptyset} \mathbb{P}\big(A_\mathbf{i} = 1, A_\mathbf{j} = 1\big) + \sum_{\mathbf{i},\mathbf{j}\in\mathcal{C}_\ell, |V|<|E|} \mathbb{P}\big(A_\mathbf{i} = 1, A_\mathbf{j} = 1\big)$$

where the second term refers to (6.26). By Proposition 6.3, we have for $\mathbf{i}, \mathbf{j} \in \mathcal{C}_\ell$,

$$\mathbb{P}\big(A_\mathbf{i} = 1, A_\mathbf{j} = 1\big) \le C_\ell (d/n)^E$$

and $\big|\{\mathbf{i},\mathbf{j} \in \mathcal{C}_\ell, |V| = v\}\big| \le C_\ell n^v$ for $v \in \{\ell, \cdots, 2\ell\}$, so that

$$\sum_{\mathbf{i},\mathbf{j}\in\mathcal{C}_\ell, |V|<|E|} \mathbb{P}\big(A_\mathbf{i} = 1, A_\mathbf{j} = 1\big) \le C_\ell \frac{d^{2\ell}}{n}.$$

On the other hand, also by Proposition 6.4, for $\mathbf{i}, \mathbf{j} \in \mathcal{C}_\ell$ disjoint, there exists $\varepsilon = O_\ell(d/n)$ so that

$$\mathbb{P}\big(A_\mathbf{i} = 1, A_\mathbf{j} = 1\big) = (1+\varepsilon)(d/n)^{2\ell}.$$

Moreover, for a fixed $\ell \in \mathbb{N}$, we have $\big|\{\mathbf{i},\mathbf{j} \in \mathcal{C}_\ell, \mathbf{i}\cap\mathbf{j} = \emptyset\}\big| = \frac{n^{2\ell}}{\ell^2}\big(1 - O_\ell(1/n)\big)$ as $n \to \infty$. This shows that as $n \to \infty$

$$\mathbb{E}(Q_\ell)_2 = \frac{d^{2\ell}}{\ell^2}\big(1 + O_\ell(d/n)\big). \tag{6.27}$$

6.3.2. *Proof of Proposition 6.9.* Fix $k \in \mathbb{N}$ and $\boldsymbol{\alpha} \in \mathbb{N}_0^k$. Recall (6.15) and let us denote

$$\Gamma_{\boldsymbol{\alpha}} = \big\{\vec{\mathbf{i}} = \{\mathbf{i}_{r,\ell}\}_{r\in[\alpha_\ell],\ell\in[k]} : \mathbf{i}_{1,\ell},\dots,\mathbf{i}_{\alpha_\ell,\ell} \in \mathcal{C}_\ell, \text{ distincts, for all } \ell \in [k]\big\}.$$

An element $\vec{\mathbf{i}} \in \Gamma_{\boldsymbol{\alpha}}$ is a collection of distinct cycles of length $(\underbrace{1,\cdots,1}_{\times\alpha_1},\cdots,\underbrace{k,\cdots,k}_{\times\alpha_k})$. Then, by (6.18), we have

$$\mathbb{E}\left[(Q_{\ell_1})_{\alpha_1}\cdots(Q_{\ell_k})_{\alpha_k}\right] = \sum_{\vec{\mathbf{i}}\in\Gamma_{\boldsymbol{\alpha}}} \mathbb{P}\big(A_{\mathbf{i}_{r,\ell}} = 1, \forall r \in [\alpha_\ell], \forall \ell \in [k]\big).$$

Generalizing the notation (6.25), the digraph $\vec{\mathbf{i}}$ satisfies

$$|V(\vec{\mathbf{i}})| = |E(\vec{\mathbf{i}})| \quad \text{if and only if} \quad \text{the cycles } \mathbf{i}_{r,\ell} \text{ are all disjoint.}$$

This follows e.g. by induction on (6.26). Hence, we split

$$\mathbb{E}\left[(Q_{\ell_1})_{\alpha_1}\cdots(Q_{\ell_k})_{\alpha_k}\right] = \sum_{\vec{\mathbf{i}}\in\Gamma_{\boldsymbol{\alpha}},\, \mathbf{i}_{r,\ell}\text{ disjoint}} \mathbb{P}\big(A_{\mathbf{i}_{r,\ell}} = 1, \forall r \in [\alpha_\ell], \forall s \in [k]\big) \tag{6.28}$$

$$+ \sum_{\vec{\mathbf{i}}\in\Gamma_{\boldsymbol{\alpha}},\, |V(\vec{\mathbf{i}})|<|E(\vec{\mathbf{i}})|} \mathbb{P}\big(A_{\mathbf{i}_{r,\ell}} = 1, \forall r \in [\alpha_\ell], \forall s \in [k]\big). \tag{6.29}$$

By Proposition 6.4, for any $\vec{\mathbf{i}} \in \Gamma_{\boldsymbol{\alpha}}$, there exists $\varepsilon(\vec{\mathbf{i}}) = O(d/n)$ so that

$$\mathbb{P}\big(A_{\mathbf{i}_{r,\ell}} = 1, \forall r \in [\alpha_\ell], \forall \ell \in [k]\big) = (1+\varepsilon(\vec{\mathbf{i}}))(d/n)^{E(\vec{\mathbf{i}})}.$$

Since $\big|\{\vec{\mathbf{i}} \in \Gamma_{\boldsymbol{\alpha}}, |V| = v\}\big| = O(n^v)$ for $v \in \{1,\cdots,|\boldsymbol{\alpha}|\}$ where $|\boldsymbol{\alpha}| = 1\cdot\alpha_1 + \cdots + k\cdot\alpha_k$, we obtain

$$(6.29) = O\big(d^{|\boldsymbol{\alpha}|}/n\big).$$

Now, if $\mathbf{i}_{r,\ell}$ are all disjoint, $|V(\vec{\mathbf{i}})| = |E(\vec{\mathbf{i}})| = |\boldsymbol{\alpha}|$ and

$$\big|\{\vec{\mathbf{i}} \in \Gamma_{\boldsymbol{\alpha}},\, \mathbf{i}_{r,\ell}\text{ disjoint}\}\big| = \frac{n^{|\boldsymbol{\alpha}|}}{1^{\alpha_1}\cdots k^{\alpha_k}}\big(1 - O(1/n)\big)$$

This implies that

$$(6.28) = \frac{d^{|\boldsymbol{\alpha}|}}{1^{\alpha_1}\cdots k^{\alpha_k}}\big(1 + O(d/n)\big).$$

Note that the implied constants above depend only on $\boldsymbol{\alpha}$. This completes the proof of Proposition 6.9.

7. Poisson analytic functions

Proof of Proposition 2.5. The function Y_d is centred and it is almost surely analytic in D_1 since $\|Y_d\|_{\mathrm{L}^\infty(\mathrm{D}_r)} < \infty$ for any $r < 1$. Indeed, by Jensen's inequality, it holds for $r < 1$,

$$\mathbb{E}\|Y_d\|^2_{\mathrm{L}^\infty(\mathrm{D}_r)} \le \mathbb{E}\bigg(\sum_{k\in\mathbb{N}} \frac{r^k}{kd^{k/2}}\Big|\sum_{\ell|k}\ell\overline{\Lambda_\ell}\Big|\bigg)^2 \le \frac{r}{1-r}\sum_{k\in\mathbb{N}}\frac{r^k}{k^2d^k}\mathbb{E}\Big|\sum_{\ell|k}\ell\overline{\Lambda_\ell}\Big|^2$$

By independence of $\{\Lambda_\ell\}_{\ell\in\mathbb{N}}$, it holds for $k \in \mathbb{N}$

$$\frac{1}{k^2d^k}\mathbb{E}\Big|\sum_{\ell|k}\ell\overline{\Lambda_\ell}\Big|^2 = \frac{1}{k^2d^k}\sum_{\ell|k}\ell^2\mathbb{E}\big|\overline{\Lambda_\ell}\big|^2 = \frac{1}{k^2d^k}\sum_{\ell|k}\ell d^\ell \le 1$$

where we used that $\operatorname{var}\Lambda_\ell = d^\ell/\ell$ for $\ell \in \mathbb{N}$. This shows that for any $r < 1$,

$$\mathbb{E}\|Y_d\|^2_{\mathrm{L}^\infty(\mathrm{D}_r)} \le \frac{r^2}{(1-r)^2}. \tag{7.1}$$

Repeating this argument clearly leads to the same estimate for the random analytic function X_d.

This proves that Y_d is well-defined, then by rearranging the series, we can write

$$Y_d(z) = \sum_{\ell\in\mathbb{N}} \overline{\Lambda_\ell} \sum_{k\in\mathbb{N}} \frac{z^{k\ell}}{kd^{k\ell/2}} = -\sum_{\ell\in\mathbb{N}} \overline{\Lambda_\ell} \log\left(1-(z/\sqrt{d})^\ell\right) \tag{7.2}$$

for the principle branch of $\log(1+z)$ which is analytic for $z \in \mathrm{D}_1$.

Recall that the Laplace transform of a random variable $\Lambda \sim \mathrm{Poisson}(\lambda)$ satisfies

$$\mathbb{E}\exp(z(\Lambda - \mathbb{E}\Lambda)) = \exp\left(\lambda(\mathrm{e}^z - 1 - z)\right), \qquad z \in \mathbb{C}. \tag{7.3}$$

In particular, this can be used to give an alternative proof that the series (7.2) is (almost surely) absolutely convergent and to justify rearranging this sum. Namely, it holds for all $0 \le t \le \sqrt{\lambda}$,

$$\mathbb{E}\exp\left(t|\Lambda - \mathbb{E}\Lambda|/\sqrt{\lambda}\right) \le 2\exp\left(t^2\right).$$

For $d \ge 2$, by Markov's inequality, we obtain the large deviation estimate,

$$\mathbb{P}\left[|\overline{\Lambda_\ell}| \ge \sqrt{d^\ell}\right] \le \exp\left(-\ell/4\right)$$

Hence, almost surely, $|\overline{\Lambda_\ell}| \le \sqrt{d^\ell}$ for all $\ell \in \mathbb{N}$ sufficiently large and (7.2) is absolutely convergent.

Remark 7.1. In the special case $d = 1$, observe that

$$\mathbb{P}\left[\Lambda_\ell \ge 2 \text{ infinitely often}\right] \le \lim_{n\to\infty} \sum_{\ell \ge n} \left(\mathbb{P}\left[\Lambda_\ell \ge 2\right]\right) \le \left(\lim_{n\to\infty} \sum_{\ell\ge n} \frac{1/2}{\ell^2}\right) = 0.$$

In particular, the random analytic function $\sum_{k\in\mathbb{N}} \frac{z^k}{k} \sum_{\ell|k} \ell\Lambda_\ell$ converges almost surely for $z \in \mathrm{D}_1$. In contrast, for $d \ge 2$, it is necessary to re-center the Poisson random coefficients to define a random analytic function in D_1.

Let us denote by $\mathrm{f}_\ell(z) := (1 - z^\ell)\mathrm{e}^{z^\ell}$ for $\ell \in \mathbb{N}$ so that $\log \mathrm{f}_\ell$ is analytic in D_1 (using the principal branch). By independence of $\{\Lambda_\ell\}_{\ell\in\mathbb{N}}$ and (7.3), it follows from the expansion (7.2) that

$$\mathbb{E}\mathrm{e}^{-Y_d(z)} = \left(\prod_{\ell\in\mathbb{N}} \mathrm{f}_\ell(z/\sqrt{d})^{\frac{d^\ell}{\ell}}\right)^{-1}.$$

This infinite product converges since $\left|\log \mathrm{f}_\ell(z)\right| \le \frac{r^{2\ell}}{1-r}$ for $z \in \mathrm{D}_r$ and we used that

$$\mathrm{e}^{\log(1-z^\ell)} - 1 - \log(1-z^\ell) = -\log \mathrm{f}_\ell(z).$$

This proves (2.4).

Now, if we repeat the computation leading to (7.1), then for any $r < d^{1/4}$,

$$\begin{aligned}
\mathbb{E}\|\Upsilon_d\|^2_{\mathrm{L}^\infty(\mathrm{D}_r)} = \mathbb{E}\left\| \sum_{k\in\mathbb{N}} \frac{z^k}{kd^{k/2}} \sum_{\ell|k,\ell<k} \ell\overline{\Lambda_\ell}\right\|^2_{\mathrm{L}^\infty(\mathrm{D}_r)} &\le \frac{r}{d^{1/4}-r} \sum_{k\in\mathbb{N}} \frac{r^k}{k^2 d^{3k/4}} \mathbb{E}\left|\sum_{\ell|k,\ell<k} \ell\overline{\Lambda_\ell}\right|^2 \\
&\le \frac{r/2}{d^{1/4}-r} \sum_{k\in\mathbb{N}} \frac{r^k}{d^{k/4}} = \frac{r^2/2}{(d^{1/4}-r)^2}
\end{aligned}$$

where we used that $\sum_{\ell|k,\ell<k} \ell d^\ell \le \frac{k^2}{2} d^{k/2}$. This shows that the random analytic function Υ_d converges almost surely in the disk $\mathrm{D}_{d^{1/4}}$, which is strictly larger than D_1 if the degree $d \ge 2$. On the other-hand, by independence of $\{\Lambda_\ell\}_{\ell\in\mathbb{N}}$, the covariance kernel of X_d is

$$\mathbb{E}X_d(z)X_d(w) = \sum_{k\in\mathbb{N}} \frac{z^k w^k}{d^k} \mathbb{E}\overline{\Lambda_k}^2 = \sum_{k\in\mathbb{N}} \frac{z^k w^k}{k} = \log(1-zw)^{-1}, \qquad z, w \in \mathrm{D}_1.$$

Remarkably, this kernel is independent of $d \in \mathbb{N}$ and $z \in \partial\mathrm{D}_1 \mapsto X_d(z)$ is a non-Gaussian log-correlated field. If $d \ge 2$, so is the fields $z \in \partial\mathrm{D}_1 \mapsto Y_d(z)$. ☐

Proof of Proposition 2.6. We have the following multi-dimensional CLT, for any $k \in \mathbb{N}$,

$$\left(\overline{\Lambda_1}/\sqrt{d}, \sqrt{2}\overline{\Lambda_2}/d, \cdots, \sqrt{k}\overline{\Lambda_k}/\sqrt{d^k}\right) \xrightarrow[d\to\infty]{\text{law}} (N_1, \cdots, N_k); \tag{7.4}$$

cf. Section A. In particular, we can choose a coupling, where almost surely, $\sqrt{\ell}\overline{\Lambda_\ell}/\sqrt{d^\ell} \to (-1)^{\ell+1}N_\ell$ for all $\ell \in \mathbb{N}$ as $d \to \infty$. Within this coupling, we verify that almost surely, $d \to \infty$

$$X_d \to X_\infty \qquad \text{and} \qquad \Upsilon_d \to 0$$

uniformly on compact subsets of D_1. $\square$

8. Proofs of Theorem 2.2 and Theorem 2.4

We use the following lemma from [8], which is inspired from [45].

Lemma 8.1 (Lemma 3.2 in [8])**.** *Let $\{f_n : \mathrm{D}_1 \to \mathbb{C}\}$ be a sequence of random analytic functions, $f_n(z) = \sum_{k=0}^\infty a_k^{(n)} z^k$ for $n \in \mathbb{N}$. Assume that $\{f_n\}$ is tight and the process $(a_k^{(n)})_{k\in\mathbb{N}_0} \to (a_k)_{k\in\mathbb{N}_0}$ converges in finite-dimensional distributions; for every $k \geq 0$,*

$$\left(a_0^{(n)}, \ldots, a_k^{(n)}\right) \xrightarrow[n\to\infty]{\text{law}} (a_0, \ldots, a_k), \tag{8.1}$$

then $f = \sum_{k=0}^\infty a_k z^k$ is converges almost surely in D_1 and

$$f_n \xrightarrow[n\to\infty]{\text{law}} f.$$

We also need the following non-trivial observation.

Lemma 8.2. *According to Definition 2.1,*

$$\mathbb{E}\mathrm{e}^{-Y_d(z)} = \exp\left(\sum_{k\in\mathbb{N}} \frac{z^k}{kd^{k/2}} \sum_{\ell|k,\ell<k} \ell\,\mathbb{E}\Lambda_\ell\right)$$

where the sum converges absolutely in D_1.

Proof. Recall that $\mathrm{f}_\ell(z) := (1-z^\ell)\mathrm{e}^{z^\ell}$ for $\ell \in \mathbb{N}$. These functions does not vanishes in D_1 so $\log \mathrm{f}_\ell$ is well-defined for the principal branch of log and

$$-\log \mathrm{f}_\ell(z) = \sum_{j\geq 2} \frac{z^{j\ell}}{j}; \qquad z \in \mathrm{D}_1.$$

Then, since $\mathbb{E}\Lambda_\ell = d^\ell/\ell$ for $\ell \in \mathbb{N}$, we compute

$$\begin{aligned}\sum_{k\in\mathbb{N}} \frac{z^k}{kd^{k/2}} \sum_{\ell|k,\ell<k} \ell\,\mathbb{E}\Lambda_\ell &= \sum_{\ell\in\mathbb{N}} \frac{d^\ell}{\ell} \sum_{j\geq 2} \frac{z^{j\ell}}{jd^{j\ell/2}} \\ &= -\sum_{\ell\in\mathbb{N}} \frac{d^\ell}{\ell} \log \mathrm{f}_\ell(z/\sqrt{d})\end{aligned}$$

where all sums converge absolutely in D_1; e.g. using the bound $\left|\log \mathrm{f}_\ell(z)\right| \leq \frac{r^{2\ell}}{1-r^\ell}$ valid for $z \in \mathrm{D}_r$. Taking exp and using formula (2.4), this proves the claim. $\square$

The characteristic polynomial of a $n \times n$ random matrix A satisfies for $z \in \mathbb{C}$,

$$\chi_n(z) = \det(1 - zA) = \sum_{k\leq n} z^k \Delta_k^{(n)}$$

where $\Delta_0 = 1$ and for $k \geq 1$,

$$\Delta_k = p_k\left(-\operatorname{tr}(A), \cdots, -\operatorname{tr}(A^k)\right) \tag{8.2}$$

where p_k is a (multivariate) polynomial of degree k.

These polynomials arise by (formally) identifying the power series

$$\exp\left(\sum_{k\in\mathbb{N}} \frac{z^k}{k} x_k\right) = 1 + \sum_{k\in\mathbb{N}} p_k(x_1,\cdots,x_k) z^k, \qquad (x_1,x_2,\cdots) \in \mathbb{C}^\infty. \tag{8.3}$$

In particular, they have the following property for every $k \in \mathbb{N}$;

$$p_k\big(zx_1,\cdots,z^k x_k\big) = z^k p_k\big(x_1,\cdots,x_k\big) \qquad \text{for } z\in\mathbb{C}$$

and

$$\big|p_k(x_1,\cdots,x_k)\big| \le p_k(|x_1|,\cdots,|x_k|).$$

Hence, if $\sum_{k\in\mathbb{N}} \frac{|x_k|}{k} r^k < \infty$ for $r>0$, then both sums in (8.3) are absolutely convergent for $z \in \overline{\mathrm{D}_r}$.

The underlying idea is that for $d\in\mathbb{N}$ fixed, by formula (8.2), Theorem 3.1 (1) and the continuous mapping theorem, for every $k\in\mathbb{N}$

$$\big(\Delta_1^{(n)},\dots,\Delta_k^{(n)}\big) \xrightarrow[n\to\infty]{\text{law}} \big(P_1,\dots,P_k\big)$$

where

$$P_k = p_k\bigg(-\Lambda_1,\dots,-\sum_{\ell|k}\ell\Lambda_\ell\bigg).$$

Moreover, by Proposition 4.1 below, the characteristic polynomial $\{\chi_n\}$ is tight in $\mathrm{D}_{1/\sqrt{d}}$. Thus, by Lemma 8.1,

$$\chi_n(z) \xrightarrow[n\to\infty]{\text{law}} 1 + \sum_{k\in\mathbb{N}} P_k z^k$$

locally uniformly for $z\in\mathrm{D}_{1/\sqrt{d}}$. However, we cannot directly identify the limit from this arguments since the series $\sum_{k\in\mathbb{N}} \frac{z^k}{k}\sum_{\ell|k}\ell\Lambda_\ell$ does not converge for $z \in \mathrm{D}_{1/\sqrt{d}}$ (cf. Remark 7.1). Hence, we give a modified version of this argument which also applies in the regime where $d(n)\to\infty$.

Proof of Theorem 2.2. Instead of the characteristic polynomial, we consider the function

$$f_n(z) = \frac{\widehat{\chi}_n(z)}{z-1/\sqrt{d}}\mathbb{E}\mathrm{e}^{-Y_d(z)}, \qquad z\in\mathrm{D}_1, n\in\mathbb{N}.$$

By Lemma 8.2 and since the rescaled characteristic polynomial has a trivial root at $z = 1/\sqrt{d}$, these are still well-defined random analytic functions. Moreover, by Proposition 4.1, $\mathbb{E}\|\widehat{\chi}_n\|_{\mathrm{L}^\infty(\mathrm{D}_{\mathrm{e}^{-t}})} \le C_t$ for $t>0$ so that by Cauchy's formula,

$$\mathbb{E}\|f_n\|_{\mathrm{L}^\infty(\mathrm{D}_{\mathrm{e}^{-2t}})} \le \int_0^{2\pi} \frac{\mathbb{E}\big|\widehat{\chi}_n(\mathrm{e}^{-t+i\theta})\mathbb{E}\mathrm{e}^{-Y_d(\mathrm{e}^{-t+i\theta})}\big|}{|\mathrm{e}^{-t}-1/\sqrt{d}||\mathrm{e}^{-t}-\mathrm{e}^{-2t}|}\frac{d\theta}{2\pi}$$

which is bounded uniformly for $n\in\mathbb{N}$. This shows that $\{f_n\}$ is tight.

We can expand

$$f_n(z) = 1 + \sum_{k\in\mathbb{N}} a_k^{(n)} z^k; \qquad a_k^{(n)} = p_k\bigg(-\frac{\mathrm{tr}(A)-\mathbb{E}\Lambda_1}{\sqrt{d}},\dots,-\frac{\mathrm{tr}(A^k)-\sum_{\ell|k}\ell\mathbb{E}\Lambda_\ell}{\sqrt{d^k}}\bigg).$$

This follows from (8.3) and the facts that

$$\widehat{\chi}_n(z) = \exp\bigg(-\sum_{k\ge1}\frac{z^k}{kd^{k/2}}\mathrm{tr}(A^k)\bigg)/\sqrt{d}$$

(the sum converges at least for z in a small neighborhood of 0 since we have the crude bound $|\mathrm{tr}(A^k)| \le (nd)^k$ because the entries of A are bounded by d) and according to Lemma 8.2,

$$\frac{\mathbb{E}\mathrm{e}^{-Y_d(z)}}{1-z\sqrt{d}} = \exp\bigg(\sum_{k\in\mathbb{N}}\frac{z^k}{kd^{k/2}}\sum_{\ell|k}\ell\,\mathbb{E}\Lambda_\ell\bigg), \qquad z\in\mathrm{D}_{1/\sqrt{d}}.$$

Hence, for a fixed $d \in \mathbb{N}$, by Theorem 3.1 (1) and the continuous mapping theorem, it holds for every $k \in \mathbb{N}$

$$\big(a_1^{(n)}, \dots, a_k^{(n)}\big) \xrightarrow[n\to\infty]{\text{law}} (a_1, \dots, a_k),$$

where

$$a_k = d^{-k/2} p_k\bigg(-\overline{\Lambda_1}, \dots, -\sum_{\ell | k} \ell \overline{\Lambda_\ell} \bigg).$$

By Lemma 8.1, we conclude that

$$f_n(z) = \frac{\widehat{\chi}_n(z)}{z - 1/\sqrt{d}} \mathbb{E} \mathrm{e}^{-Y_d(z)} \xrightarrow[n\to\infty]{\text{law}} f(z) = 1 + \sum_{k\in\mathbb{N}} a_k z^k$$

Moreover, from the proof of Proposition 2.5, it holds almost surely $\sum_{k\in\mathbb{N}} \frac{r^k}{kd^{k/2}} \big| \sum_{\ell|k} \ell \overline{\Lambda_\ell} \big| < \infty$ for any $r < 1$. By (8.3) and the subsequent observation, it follows that

$$\mathrm{e}^{-Y_d(z)} = 1 + \sum_{k\in\mathbb{N}} a_k z^k$$

is the weak-limit of the sequence $\{f_n\}$. This completes the proof. □

Proof of Theorem 2.4. This is a variant of the previous argument. We consider the function

$$f_n(z) = \frac{\widehat{\chi}_n(z)}{z - 1/\sqrt{d}}, \qquad z \in \mathrm{D}_1, n \in \mathbb{N}.$$

By Proposition 4.1 and Cauchy's formula, $\{f_n\}$ is tight (the fact that $1/\sqrt{d(n)} \to 0$ is no relevant).

In this case, we have the expansion

$$f_n(z) = 1 + \sum_{k\in\mathbb{N}} a_k^{(n)} z^k; \qquad a_k^{(n)} = p_k\bigg(-\frac{\mathrm{tr}(A) - d}{\sqrt{d}}, \dots, -\frac{\mathrm{tr}(A^k) - d^k}{\sqrt{d^k}} \bigg).$$

Hence, by Theorem 3.1 (2) and the continuous mapping theorem, in the regime where $d(n) \to \infty$ and $d(n) = n^{o(1)}$, it holds for every $k \in \mathbb{N}$

$$\big(a_1^{(n)}, \dots, a_k^{(n)}\big) \xrightarrow[n\to\infty]{\text{law}} (a_1, \dots, a_k),$$

where

$$p_k\left(N_1, \sqrt{2}N_2 - 1, \dots, \sqrt{k}N_k - \mathbf{1}_{\{k \text{ is even}\}} \right).$$

using the symmetry of N_k.

Now, we verify that for $z \in \mathrm{D}_1$,

$$\sqrt{1 - z^2}\mathrm{e}^{X_\infty(z)} = \exp\left(\sum_{k\in\mathbb{N}} \frac{z^k}{k} \big(\sqrt{k}N_k - \mathbf{1}_{\{k \text{ is even}\}}\big) \right)$$

where the sum converges absolutely for $z \in \overline{\mathrm{D}_r}$ for any $r < 1$. Hence, we conclude that if $d(n) \to \infty$ and $d(n) = n^{o(1)}$, then

$$f_n(z) = \frac{\widehat{\chi}_n(z)}{z - 1/\sqrt{d}} \xrightarrow[n\to\infty]{\text{law}} \sqrt{1 - z^2}\mathrm{e}^{X_\infty(z)}$$

This proves the claim. □

9. Proof of spectral gaps

There are two popular models to generate simple random regular digraphs:

(1) Permutation model: Let $A = \sum_{i=1}^{d} P^{(i)}$ be a sum of d independent uniformly random permutations. Sample A conditioned on A being an adjacency matrix of a simple digraph.

(2) Uniform model: Sample a d-regular digraph uniformly from all d-regular digraphs with n vertices.

It was proved in [29] and [38] that the two models are *contiguous*, which means if a sequence of events happens asymptotically almost surely in model, it holds for the other model as well.

For fixed d, our Theorem 2.2 can be applied to show a spectral gap result for the two models. The proof follows the same way as in [8, Theorem 1.1] and [15, Theorem 2.6].

We will use the following proposition from [45, Proposition 2.3] We identify multisets with integer-valued Radon measures as in [45] and endow the space of multisets with the topology of vague convergence, and the space of random multisets with the topology of weak convergence with respect to the vague topology.

Proposition 9.1. *Let f_n be a sequence of random functions in D_r converging in law to a non-zero function f. Let Φ_n and Φ be the random multisets of the zeros of f_n, f in D_r, respectively. Then Φ_n converges in law towards Φ.*

Proof of Theorem 1.1. We first prove the statement for case (i): the sum of random permutations. Let $f(z)$ be the limiting function of $\widehat{\chi}_n/\sqrt{d}$. Then from Theorem 2.2, on D_r for any $0 < r < 1$, $f(z)$ has only one zero in D_r at $-1/\sqrt{d}$. By Proposition 9.1, with high probability $\widehat{\chi}_n(z)$ has only one zero at $-1/\sqrt{d}$, which implies $A/\sqrt{d}$ has no eigenvalue outside $\mathrm{D}_{1/r}$ except $\sqrt{d}$.

For any $\varepsilon > 0$, take $\frac{1}{1+\varepsilon/\sqrt{d}} < r < 1$, then

$$\mathbf{P}(|\lambda_2| > \sqrt{d} + \varepsilon) \le \mathbf{P}\left(\frac{|\lambda_2|}{\sqrt{d}} > \frac{1}{r}\right) \to 0.$$

as $n \to \infty$.

It was shown in [38] that the probability that A is simple is bounded from below when d is fixed. Therefore $|\lambda_2| > \sqrt{d}$ holds asymptotically almost surely for case (ii). By contiguity between model (ii) and model (iii), it holds for uniform random d regular digraphs as well. $\square$

When $d(n) \to \infty$, there is no contiguity result in the literature between the two models (1) and (2), and we are not aware of any result to provide a constant lower bound on the probability that $A = \sum_{i=1}^{d(n)} P^{(i)}$ being simple. So our Theorem 2.4 only gives a result for sums of random permutations.

Proof of Theorem 1.2. In Theorem 2.4, the limiting function $z\sqrt{1-z^2}\mathrm{e}^{X_\infty(z)}$ has only one root at 0. From Proposition 9.1, with high probability $\widehat{\chi}_n(z)$ has one root converging to zero and no other roots in D_r for any $0 < r < 1$. Then with high probability, $A/\sqrt{d(n)}$ has no eigenvalue outside $\mathrm{D}_{1/r}$ except for $\sqrt{d(n)}$, which implies (1.1). $\square$

Appendix A. Appendix: a Poisson CLT

Let $\Lambda \sim \mathrm{Poisson}(\lambda)$. Recall that the Laplace transform of Λ is

$$\mathbb{E}\exp(z(\Lambda - \lambda)) = \exp\big(\lambda(\mathrm{e}^z - 1 - z)\big), \qquad z \in \mathbb{C}.$$

Let $\mu_k(\lambda) = \mathbb{E}(\Lambda - \lambda)^k$ be the central moments of Λ for $k \in \mathbb{N}$. We have for $k \in \mathbb{N}$

$$\mu_k(\lambda)/k! = [z^k]\exp\big(\lambda(\mathrm{e}^z - 1 - z)\big).$$

In particular, $\mu_1 = 0$ and for $k \in \mathbb{N}$,

$$\begin{cases} \mu_{2k}(\lambda) = m_k\lambda^k + p_k(\lambda), & p_k \in \mathcal{P}_{k-1} \\ \mu_{2k+1} \in \mathcal{P}_k \end{cases}$$

where $m_k = \frac{(2k)!}{2^k k!}$ and $\mathcal{P}_k = \{\text{polynomials of degree } \le k\}$.

From this fact, we deduce that for every $k \in \mathbb{N}$

$$\frac{\mu_k(\lambda)}{\lambda^{k/2}} = \mathbb{E}\left(\frac{\Lambda - \lambda}{\sqrt{\lambda}}\right)^k \xrightarrow[\lambda\to\infty]{} \mathbb{E}N^k = \begin{cases} m_{k/2} & k \text{ even} \\ 0 & k \text{ odd} \end{cases} \tag{A.1}$$

where N is a standard Gaussian. This is a CLT in the sense of moments for a Poisson random variable.

The goal of this appendix is to extend this results to a collection of asymptotically Poisson random variables with large parameters.

First recall that we can expand for $k \in \mathbb{N}$

$$(x-\lambda)^k = \textstyle\sum_{j=0}^{k} (x)_{k-j} q_{k,j}(\lambda); \qquad q_{k,j} \in \mathcal{P}_j. \tag{A.2}$$

Then, according to Lemma 6.8, we have

$$\mu_k(\lambda) = \sum_{j\le k} q_{k,j}(\lambda)\mathbb{E}(\Lambda)_{k-j} = \sum_{j\le k} q_{k,j}(\lambda)\lambda^{k-j}. \tag{A.3}$$

Remarkably the RHS is a polynomial of degree $\le k/2$.

Lemma A.1. *Fix $\ell \in \mathbb{N}$ and for $j \in [\ell]$, let $(\lambda_{j,n})_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}_+$ such that $\lambda_{j,n} \to \infty$ as $n \to \infty$. Let $(Q_{j,n})_{j\in[\ell],n\in\mathbb{N}}$ be a sequence of random vectors such that its joint factorial moments satisfy for any $\mathbf{k} \in \mathbb{N}^\ell$,*

$$\mathbb{E}\big[(Q_{1,n})_{k_1}\cdots(Q_{\ell,n})_{k_\ell}\big] = \lambda_{1,n}^{k_1}\cdots\lambda_{\ell,n}^{k_\ell}(1+\epsilon_{\mathbf{k},n}) \tag{A.4}$$

where $\epsilon_{\mathbf{k},n} \in \mathbb{R}$ and for $\mathbf{k} \in \mathbb{N}^\ell$,

$$\lambda_{1,n}^{k_1/2}\cdots\lambda_{\ell,n}^{k_\ell/2} \cdot \max_{\mathbf{j}\le\mathbf{k}} |\epsilon_{\mathbf{j},n}| \to 0. \tag{A.5}$$

Then

$$\left(\frac{Q_{1,n}-\lambda_{1,n}}{\sqrt{\lambda_{1,n}}},\dots,\frac{Q_{\ell,n}-\lambda_{\ell,n}}{\sqrt{\lambda_{\ell,n}}}\right) \xrightarrow[n\to\infty]{\text{law}} (N_1,\dots,N_\ell)$$

and in the sense of moments, where $(N_i)_{i\in[\ell]}$ are i.i.d. standard Gaussians.

Proof. Let us denote $\mu_{\mathbf{k},n} = \mathbb{E}\big[(Q_{1,n}-\lambda_{1,n})^{k_1}\cdots(Q_{\ell,n}-\lambda_{\ell,n})^{k_\ell}\big]$ for $\mathbf{k}\in\mathbb{N}^\ell$. Using (A.2) and (A.4), we can write

$$\begin{aligned}\mu_{\mathbf{k},n} &= \sum_{\mathbf{j}\le\mathbf{k}}\prod_{i\le\ell} q_{k_i,j_i}(\lambda_{i,n})\mathbb{E}\big[(Q_{1,n})_{k_1-j_1}\cdots(Q_{\ell,n})_{k_\ell-j_\ell}\big]\\ &= \sum_{\mathbf{j}\le\mathbf{k}}\prod_{i\le\ell} q_{k_i,j_i}(\lambda_{i,n})\lambda_{i,n}^{k_i-j_i}(1+\epsilon_{\mathbf{k}-\mathbf{j},n}).\end{aligned}$$

Then, the conditions (A.5) imply that for any fixed $\mathbf{k}\in\mathbb{N}^\ell$,

$$\mu_{\mathbf{k},n} = \prod_{i\le\ell}\Big(\sum_{j\le k_i} q_{k_i,j}(\lambda_{i,n})\lambda_{i,n}^{k_i-j}\Big) + \underset{n\to\infty}{o_{\mathbf{k}}}\big(\lambda_{1,n}^{k_1/2}\cdots\lambda_{\ell,n}^{k_\ell/2}\big).$$

Here we used that $q_{k,j}\in\mathcal{P}_j$. Hence using the combinatorial identity (A.3),

$$\mu_{k,n} = \mu_{k_1}(\lambda_{1,n})\cdots\mu_{k_\ell}(\lambda_{\ell,n}) + \underset{n\to\infty}{o_{\mathbf{k}}}\big(\lambda_{1,n}^{k_1/2}\cdots\lambda_{\ell,n}^{k_\ell/2}\big).$$

From (A.1), we conclude that for every $\mathbf{k}\in\mathbb{N}^\ell$

$$\frac{\mathbb{E}\big[(Q_{1,n}-\lambda_{1,n})^{k_1}\cdots(Q_{\ell,n}-\lambda_{\ell,n})^{k_\ell}\big]}{\lambda_{1,n}^{k_1/2}\cdots\lambda_{\ell,n}^{k_\ell/2}} \xrightarrow[n\to\infty]{} \mathbb{E}\big[N_1^{k_1}\cdots N_\ell^{k_\ell}\big].$$

This completes the proof, since the law of $(N_i)_{i\in[\ell]}$ is characterized by its joint moments. □

Appendix B. Appendix: sums of non-uniform random permutations

In this paper, we studied random digraphs arising as sums of $n \times n$ uniform permutations. A crucial technical argument in our analysis was that the rows of the matrix $A = P^{(1)} + \cdots + P^{(d)}$ are exchangeable; but this property is in fact very strong. It would not hold if the permutation matrices $P^{(i)}$ were not uniformly distributed on $\mathfrak{S}_n$. This is typically what happens when they are skewed towards having more or less short cycles. For example, the Ewens distribution with parameter θ is defined as follows; we say that a random permutation π is $\mathrm{Ewens}(\theta)$-distributed if

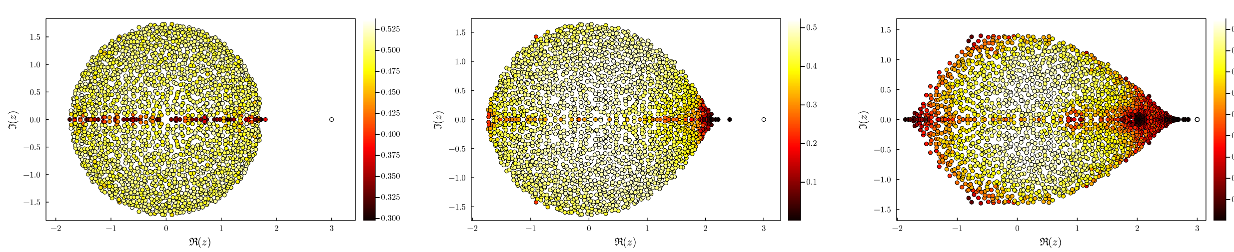
$$\forall \sigma \in \mathfrak{S}_n, \qquad \mathbb{P}(\pi = \sigma) = \frac{\theta^{\mathrm{cyc}(\sigma)}}{\theta(\theta+1)\cdots(\theta+n-1)} \tag{B.1}$$

where $\mathrm{cyc}(\sigma)$ is the number of cycles in the cycle decomposition of σ. When $\theta = 1$, the random permutation π is uniform; when $\theta > 1$ (respectively, < 1), it is skewed towards having more cycles (respectively, less), resulting in a directed graph with a different local structure. Crucially, Ewens-distributed random permutations with $\theta \neq 1$ are invariant by *conjugation* by any permutation matrix, but not invariant by multiplication by a permutation matrix.

We were able to identify the asymptotics of the traces of $A = P^{(1)} + \cdots + P^{(d)}$ is this model, with θ fixed and not depending on n; indeed, Theorem 3.1 still holds, with the limiting random variables Λ_ℓ being replaced with the following ones:

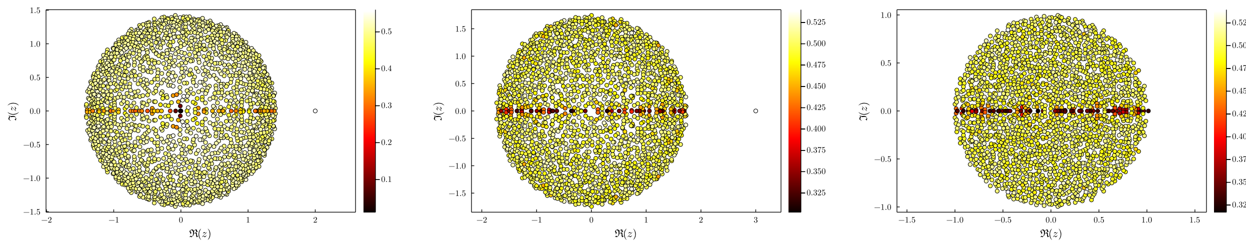
$$\Theta_\ell \sim \mathrm{Poisson}\left(\frac{d^\ell + d(\theta-1)}{\ell}\right).$$

A proof is available on demand; however, the proof of the tightness of the sequence $\det(\mathrm{I} - zA_n)$ is still under exploration. We mention this fact, since the asymptotic behaviour of the spectrum of A in the Ewens case displays some unusual features; when θ is sufficiently large (this might depend on n), the whole limiting shape of the eigenvalues seems not rotation-invariant, with a previously unseen pattern of eigenvector localization — see Figure 2. For comparison, Figure 3 displays the eigenvalues of our model of sums-of-permutation matrices, together with an example of a real Ginibre spectrum.



$\theta = 0.0001$ $\qquad$ $\theta = 100$ $\qquad$ $\theta = 500$

Figure 2. These figures display the $n = 2000$ complex eigenvalues of sums of $d = 3$ permutations matrices, the underlying permutations being $\mathrm{Ewens}(\theta)$-distributed on $\mathfrak{S}_{2000}$ for various θ. Each eigenvalue is coloured according to the degree of localization of its corresponding right-eigenvector; here, we measure the localisation of a vector $\varphi \in \mathbb{C}^n$ using the Inverse Participation Ratio $\mathrm{IPR}(\varphi) = |\varphi|_2^4/n|\varphi|_4^4 \leqslant 1$. An IPR equal to 1 means that φ is constant up to phases (pure delocalization); an IPR equal to $1/n$ means that φ is a multiple of a Dirac (pure localization). Here, we see that the eigenvectors with eigenvalues close to the real axis are more localized, a phenomenon already visible for classical matrix ensembles such as Ginibre (see Figure 3).



$d = 2$ $\qquad$ $d = 3$ $\qquad$ Ginibre ensemble

FIGURE 3. These figures display the $n = 2000$ complex eigenvalues of sums of d uniform permutations matrices, with eigenvector localization depicted as the Figure 2. We plotted in the third panel the eigenvalues of a Real Ginibre matrix (the entries are iid $\mathscr{N}_{\mathbb{R}}(0, 1/n)$) for comparison.

REFERENCES

[1] Louis-Pierre Arguin, David Belius, and Paul Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Communications in Mathematical Physics*, 349(2):703–751, 2017.

[2] Valentin Bahier and Joseph Najnudel. On smooth mesoscopic linear statistics of the eigenvalues of random permutation matrices. *Journal of Theoretical Probability*, pages 1–22, 2021.

[3] Anirban Basak, Nicholas Cook, and Ofer Zeitouni. Circular law for the sum of random permutation matrices. *Electronic Journal of Probability*, 23:1–51, 2018.

[4] Anirban Basak and Mark Rudelson. The circular law for sparse non-Hermitian matrices. *The Annals of Probability*, 47(4):2359–2416, 2019.

[5] Gérard Ben Arous and Kim Dang. On fluctuations of eigenvalues of random permutation matrices. *Annales de l'IHP Probabilités et statistiques*, 51(2):620–647, 2015.

[6] Charles Bordenave. A new proof of Friedman's second eigenvalue theorem and its extension to random lifts. *Annales Scientifiques de l'École Normale Supérieure*, 4(6):1393–1439, 2020.

[7] Charles Bordenave and Djalil Chafaï. Around the circular law. *Probability surveys*, 9:1–89, 2012.

[8] Charles Bordenave, Djalil Chafaï, and David García-Zelada. Convergence of the spectral radius of a random matrix through its characteristic polynomial. *Probability Theory and Related Fields*, pages 1–19, 2021.

[9] Reda Chhaibi, Thomas Madaule, and Joseph Najnudel. On the maximum of the cβe field. *Duke Mathematical Journal*, 167(12):2243–2345, 2018.

[10] Nicholas Cook. The circular law for random regular digraphs. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(4):2111–2167, 2019.

[11] Nicholas Cook, Larry Goldstein, and Tobias Johnson. Size biased couplings and the spectral gap for random regular graphs. *The Annals of Probability*, 46(1):72–125, 2018.

[12] Nicholas Cook and Ofer Zeitouni. Maximum of the characteristic polynomial for a random permutation matrix. *Communications on Pure and Applied Mathematics*, 73(8):1660–1731, 2020.

[13] Nicholas A Cook. Discrepancy properties for random regular digraphs. *Random Structures & Algorithms*, 50(1):23–58, 2017.

[14] Nicholas A Cook. On the singularity of adjacency matrices for random regular digraphs. *Probability Theory and Related Fields*, 167(1):143–200, 2017.

[15] Simon Coste. Sparse matrices: convergence of the characteristic polynomial seen from infinity. *arXiv preprint arXiv:2106.00593*, 2021.

[16] Simon Coste. The spectral gap of sparse random digraphs. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(2):644–684, 2021.

[17] Kim Dang and Dirk Zeindler. The characteristic polynomial of a random permutation matrix at different points. *Stochastic Processes and their Applications*, 124(1):411–439, 2014.

[18] Ioana Dumitriu, Tobias Johnson, Soumik Pal, and Elliot Paquette. Functional limit theorems for random regular graphs. *Probability Theory and Related Fields*, 156(3):921–975, 2013.

[19] Ioana Dumitriu and Yizhe Zhu. Global eigenvalue fluctuations of random biregular bipartite graphs. *arXiv preprint arXiv:2008.11760*, 2020.

[20] Bertrand Duplantier, Rémi Rhodes, Scott Sheffield, and Vincent Vargas. Log-correlated gaussian fields: an overview. *Geometry, analysis and probability*, pages 191–216, 2017.

[21] Joel Friedman. *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*. Memoirs of the American Mathematical Society. American Mathematical Society, 2008.

[22] Joel Friedman, Jeff Kahn, and Endre Szemeredi. On the second eigenvalue of random regular graphs. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 587–598, 1989.

[23] Yan V Fyodorov and Olivier Giraud. High values of disorder-generated multifractals and logarithmically correlated processes. *Chaos, Solitons & Fractals*, 74:15–26, 2015.

[24] Yan V Fyodorov, Ghaith A Hiary, and Jonathan P Keating. Freezing transition, characteristic polynomials of random matrices, and the riemann zeta function. *Physical review letters*, 108(17):170601, 2012.

[25] Shirshendu Ganguly and Soumik Pal. The random transposition dynamics on random regular graphs and the gaussian free field. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 56(4):2935–2970, 2020.

[26] Jiaoyang Huang. Invertibility of adjacency matrices for random d-regular graphs. *Duke Mathematical Journal*, 170(18):3977–4032, 2021.

[27] Christopher Hughes, Joseph Najnudel, Ashkan Nikeghbali, and Dirk Zeindler. Random permutation matrices under the generalized ewens measure. *The Annals of Applied Probability*, 23(3):987–1024, 2013.
[28] Vishesh Jain, Ashwin Sah, and Mehtaab Sawhney. The smallest singular value of dense random regular digraphs. *International Mathematics Research Notices*, 09 2021. rnab247.
[29] Svante Janson. Random regular graphs: asymptotic distributions and contiguity. *Combinatorics, Probability and Computing*, 4(4):369–405, 1995.
[30] Tobias Johnson and Soumik Pal. Cycles and eigenvalues of sequentially growing random regular graphs. *The Annals of Probability*, 42(4):1396–1437, 2014.
[31] Gaultier Lambert. Mesoscopic central limit theorem for the circular β-ensembles and applications. *Electronic Journal of Probability*, 26:1–33, 2021.
[32] Gaultier Lambert and Elliot Paquette. Strong approximation of gaussian beta-ensemble characteristic polynomials: the hyperbolic regime. *arXiv preprint arXiv:2001.09042*, 2020.
[33] Alexander E Litvak, Anna Lytova, Konstantin Tikhomirov, Nicole Tomczak-Jaegermann, and Pierre Youssef. Circular law for sparse random regular digraphs. *Journal of the European Mathematical Society*, 23(2):467–501, 2020.
[34] Alexander E Litvak, Anna Lytova, Konstantin Tikhomirov, Nicole Tomczak-Jaegermann, and Pierre Youssef. Adjacency matrices of random digraphs: singularity and anti-concentration. *Journal of Mathematical Analysis and Applications*, 445(2):1447–1491, 2017.
[35] Alexander E Litvak, Anna Lytova, Konstantin Tikhomirov, Nicole Tomczak-Jaegermann, and Pierre Youssef. The smallest singular value of a shifted d-regular random square matrix. *Probability Theory and Related Fields*, 173(3):1301–1347, 2019.
[36] András Mészáros. The distribution of sandpile groups of random regular graphs. *Transactions of the American Mathematical Society*, 373(9):6529–6594, 2020.
[37] Fernando Lucas Metz, Izaak Neri, and Tim Rogers. Spectral theory of sparse non-hermitian random matrices. *Journal of Physics A: Mathematical and Theoretical*, 52(43):434003, 2019.
[38] Michael S. O. Molloy, Hanna Robalewska, Robert W. Robinson, and Nicholas C. Wormald. 1-factorizations of random regular graphs. *Random Structures & Algorithms*, 10(3):305–321, 1997.
[39] Joseph Najnudel, Elliot Paquette, and Nick Simm. Secular coefficients and the holomorphic multiplicative chaos. *arXiv preprint arXiv:2011.01823*, 2020.
[40] Miika Nikula, Eero Saksman, and Christian Webb. Multiplicative chaos and the characteristic polynomial of the cue: The l1 phase. *Transactions of the American Mathematical Society*, 373(6):3905–3965, 2020.
[41] Elliot Paquette and Ofer Zeitouni. The maximum of the cue field. *International Mathematics Research Notices*, 2018(16):5028–5119, 2018.
[42] Ori Parzanchevski. Ramanujan graphs and digraphs. *Analysis and Geometry on Graphs and Manifolds*, 461:344, 2020.
[43] Brian Rider and Bálint Virág. The noise in the circular law and the gaussian free field. *International Mathematics Research Notices*, 2007(9):rnm006, 2007.
[44] Mark Rudelson and Konstantin Tikhomirov. The sparse circular law under minimal assumptions. *Geometric and Functional Analysis*, 29(2):561–637, 2019.
[45] Tomoyuki Shirai. Limit theorems for random analytic functions and their zeros: Dedicated to the late professor yasunori okabe (functions in number theory and their probabilistic aspects). *RIMS Kokyuroku Bessatsu*, 34:335–359, 2012.
[46] Konstantin Tikhomirov and Pierre Youssef. The spectral gap of dense random regular graphs. *The Annals of Probability*, 47(1):362–419, 2019.
[47] Philip Matchett Wood. Universality and the circular law for sparse random matrices. *The Annals of Applied Probability*, 22(3):1266–1300, 2012.
[48] Yizhe Zhu. On the second eigenvalue of random bipartite biregular graphs. *arXiv preprint arXiv:2005.08103*, 2020.

Département d'informatique, ENS - PSL University
Email address: simon.coste@ens.fr

Institute of Mathematics, University of Zurich
Email address: glambert@kth.se

Department of Mathematics, University of California, Irvine
Email address: yizhe.zhu@uci.edu