ON THE SPECTRAL RADIUS AND THE CHARACTERISTIC POLYNOMIAL OF A RANDOM MATRIX WITH INDEPENDENT ELEMENTS AND A VARIANCE PROFILE

Walid Hachem* **Michail Louvaris***

January 2, 2025

ABSTRACT

In this paper, it is shown that with large probability, the spectral radius of a large non-Hermitian random matrix with a general variance profile does not exceed the square root of the spectral radius of the variance profile matrix. A minimal moment assumption is considered and sparse variance profiles are covered. Following an approach developed recently by Bordenave, Chafaï and García-Zelada, the key theorem states the asymptotic equivalence between the reverse characteristic polynomial of the random matrix at hand and a random analytic function which depends on the variance profile matrix. The result is applied to the case of a non-Hermitian random matrix with a variance profile given by a piecewise constant or a continuous non-negative function, the inhomogeneous (centered) directed Erdős–Rényi model, and more.

MSC2020 subject classification: 60B20, 15A18

1 Problem description and results

Let $(W_{ij})_{ij\geq 1}$ be an infinite array of complex-valued independent and identically distributed random variables such that $\mathbb{E}W_{11} = 0$ and $\mathbb{E}|W_{11}|^2 = 1$. For each integer $n > 0$, let $S^{(n)} = \left[s_{ij}^{(n)}\right]_{i,j=1}^n$ be a $n \times n$ deterministic matrix with non-negative elements. Consider the $\mathbb{C}^{n\times n}$–valued random matrix $X^{(n)} = \left[X_{ij}^{(n)}\right]_{i,j=1}^n$ which elements are defined as

$$X_{ij}^{(n)} = \sqrt{s_{ij}^{(n)}} W_{ij}.$$

The purpose of this paper is to study the large–n behavior of the spectral radius $\rho(X^{(n)})$ under general assumptions on the sequence of matrices $(S^{(n)})$ that cover the sparse cases. These assumptions stand as follows:

Assumption 1.1. The following hold true.

(i) There exists a constant $C_S > 0$ such that at least one of the following bound holds:

$$\left\|\left\| S^{(n)} \right\|\right\| \leq C_S \quad \text{or} \quad \left\|\left\| (S^{(n)})^\top \right\|\right\| \leq C_S,$$

where $|||\cdot|||$ is the max row norm of a matrix.

*CNRS, Laboratoire d'informatique Gaspard Monge (LIGM / UMR 8049), Université Gustave Eiffel, ESIEE Paris, France. E-mails: walid.hachem,michail.louvaris@univ-eiffel.fr. M.L. received funding by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 101034255.

(ii) There exists a positive sequence $(K_n)_{n\geq 1}$ converging to infinity, and there exists a constant $C'_S > 0$ such that

$$s_{ij}^{(n)} \leq \frac{C'_S}{K_n}$$

for all n and all $i, j \in [n]$.

Assumption 1.2. For each $\varepsilon \in (0,1]$, it holds that

$$\liminf_n \min_{\gamma\in[0,1-\varepsilon]} \det\left(I_n - \gamma S^{(n)}\right) > 0.$$

This assumption can be re-expressed as

$$\limsup_n \rho(S^{(n)}) \leq 1, \quad \text{and} \tag{1.1a}$$

$$\forall \varepsilon \in (0,1],\ \liminf_n \det\left(I_n - (1-\varepsilon) S^{(n)}\right) > 0. \tag{1.1b}$$

Indeed, when Assumption 1.2 holds true, (1.1b) is obvious. Moreover, since $S^{(n)}$ has non-negative elements, it has an eigenvalue which is equal to $\rho(S^{(n)})$ [1, Th. 8.3.1], hence (1.1a). Conversely, assume the conditions (1.1) are satisfied. Since $\limsup \rho(S^{(n)}) \leq 1$, the series $\sum_{k\geq 1} \gamma^k \operatorname{tr}(S^{(n)})^k / k$ is convergent for each $\gamma \in [0,1)$ and each large enough n, and we can write

$$\forall \gamma \in [0,1), \quad \det(I_n - \gamma S^{(n)}) = \exp\left(-\sum_{k=1}^{\infty} \gamma^k \frac{\operatorname{tr}(S^{(n)})^k}{k}\right). \tag{1.2}$$

This shows that $\gamma \mapsto \det(I_n - \gamma S^{(n)})$ is a non-negative decreasing function on $[0,1)$, and (1.1b) implies Assumption 1.2.

The following theorem is established in this paper:

Theorem 1.3. *Let Assumptions 1.1 and 1.2 hold true. Then, it holds that*

$$\forall \varepsilon > 0, \quad \mathbb{P}\left[\rho\left(X^{(n)}\right) \geq 1 + \varepsilon\right] \xrightarrow[n\to\infty]{} 0.$$

Let us sketch some application examples to shed some light on the assumptions and the result. These will be clarified and detailed in the next section. To begin with, assume that $S^{(n)}$ is block variance profile matrix with a fixed number of rectangular blocks which dimensions are of order n, and which elements are of order $1/n$. Then, $S^{(n)}$ satisfies Assumption 1.1 with $K_n = n$. Assume that the spectral radius of $S^{(n)}$ is of order one. By normalizing this matrix with its spectral radius, Condition (1.1a) is satisfied. Furthermore, since the rank of $S^{(n)}$ is bounded by a constant, Condition (1.1b) is also satisfied, and Theorem 1.3 asserts that with high probability, $\rho(X^{(n)})$ cannot be smaller and away of $\sqrt{\rho(S^{(n)})}$ before the normalization.

A similar conclusion can be obtained when $S^{(n)}$ is obtained by a regular sampling of a continuous non-negative function on the rectangle $[0,1]^2$. Assumption 1.1 will still be satisfied with $K_n = n$. Let us turn to Condition (1.1b). Here, the rank of $S^{(n)}$ is no more necessarily bounded. However, this condition will still be satisfied after the proper normalization because, roughly speaking, $S^{(n)}$ will have only a few non-negligible eigenvalues. This is due to the fact that $S^{(n)}$ is a discrete approximation of a continuous function on $[0,1]^2$, which is as is well known a compact trace-class operator on the Banach space $C([0,1])$ of the continuous functions on $[0,1]$.

Of particular interest are the situations where $K_n = o(n)$, that we refer to as the "sparse" cases, where, typically, the number of non zero elements of $S^{(n)}$ per row belongs to the interval $[cK_n, CK_n]$ where $0 < c < C < \infty$, and these elements are bounded by C'/K_n for $C' > 0$. Examples of these cases where Assumption 1.2 is satisfied will be detailed below.

Theorem 1.3 only provides an upper bound on the spectral radius $\rho(X^{(n)})$. One can expect the stronger result $\mathbb{P}\left[|\rho(X) - 1| \geq \varepsilon\right] \to 0$, which requires showing that $\mathbb{P}\left[\rho\left(X^{(n)}\right) \leq 1 - \varepsilon\right] \to 0$ when $\rho(S^{(n)})$ is close to one. One way of establishing this last result is to establish a so-called global law on the spectral measure of $X^{(n)}$, showing that this spectral measure can be approximated for all large n with a distribution supported by the closed unit disk. For matrices with variance profiles, global laws were established in the literature in some non-sparse situations where

$K_n = n$ and where the ℓ^1 norms of all the rows and the columns of $S^{(n)}$ are of order one. The first of such results was revealed by Girko [2]. A global law was rigorously established by Alt *et.al.* in [3] under moment and density assumptions and in the case where all the numbers $ns_{ij}^{(n)}$ belong to a compact interval lying away from zero (the so-called "flat" variance profile). We note here that beyond the global law, the large-n behavior of the spectral radius is also controlled in [3] by means of establishing local law results. Finer results regarding the spectral radius under the same kind of assumptions are established by the same authors in [4]. Beyond the flat variance profile in the non-sparse case, the global law was established in [5] under a so-called robust irreducibility assumption on the variance profile, and a moment assumption on the matrix entries.

Regarding the applications, the control of the spectral radius of $X^{(n)}$ is essential in the study of many dynamical systems that arise in the fields of control theory, natural or artificial neural networks, theoretical biology and ecology, and others. For instance, in neural networks, $X^{(n)}$ is the matrix that represents the couplings between n neurons [6]. In theoretical ecology, $X^{(n)}$ is used to model the random food interactions between n living species that coexist within an ecosystem [7]. In these situations, the inhomogeneity of the matrix model represented by the variance profile, or the sparsity of its non-zero elements are often advocated to model realistic situations. The transition of such systems from stationary to chaotic dynamics is often driven by $\rho(X^{(n)})$.

To obtain Theorem 1.3, we use the approach based on the reverse characteristic polynomial of $X^{(n)}$ that Bordenave, Chafaï and García-Zelada developed in [8] to deal with the case $s_{ij}^{(n)} = 1/n$, and that was partially inspired by the article [9] devoted to a different problem. Observe that no moment of the random variables W_{ij} beyond the second moment is required in the statement of Theorem 1.3. This is a prominent feature of the approach of [8], which improves upon the older literature such as [10, 11, 12]. In the recent literature, the approach of [8] for controlling the spectral radius was applied for the Elliptic Ginibre model in [13]. In the same vein, the recent papers [14] and [15], consider the characteristic polynomial of sparse Bernoulli matrices and sums of random permutations and regular digraphs.

Let us denote as $\mathbb{H}$ the space of holomorphic functions on the open unit-disk $D(0,1)$ of $\mathbb{C}$ equipped with the topology of the uniform convergence on the compacts of $D(0,1)$. As is well known, this space is a Polish space.

Let us consider the reverse characteristic polynomial of the matrix $X^{(n)}$, which is defined as

$$q_n(z) = \det\left(I_n - zX^{(n)}\right).$$

Obviously, q_n is a $\mathbb{H}$–valued random variable. Our paper is mainly devoted towards studying the asymptotic behavior of the probability distribution of q_n on $\mathbb{H}$. Here, a notation is in order. Let (U_n) and (V_n) be two sequences of random variables valued in some metric space. For each n, let μ_n and ν_n be the probability distributions of U_n and V_n respectively. We shall use the notation

$$U_n \sim_n V_n$$

to refer to the facts that the sequences (μ_n) and (ν_n) are relatively compact, and that

$$\int f d\mu_n - \int f d\nu_n \xrightarrow[n\to\infty]{} 0$$

for each bounded continuous real function f on the metric space. We shall say then that (U_n) and (V_n) are "asymptotically equivalent". Note that (μ_n) and (ν_n) do not necessarily converge narrowly to some probability distribution. This setting is well-suited to describe the asymptotics of our sequence (q_n) because without an additional assumption on the construction of the sequence $(S^{(n)})$, the distribution of q_n has no reason to converge narrowly to a limit probability measure on $\mathbb{H}$. These asymptotics are described by the following theorem, which will be proven in Section 3:

Theorem 1.4. *Let Assumptions 1.1 and 1.2 hold true. Then, for all large n, the function*

$$\kappa_n(z) = \sqrt{\det(I - z^2 \mathbb{E}W_{1,1}^2 S^{(n)})}$$

is a well-defined element of $\mathbb{H}$ with the square root being the one for which $\kappa_n(0) = 1$. The function

$$F_n(z) = \sum_{k=1}^{\infty} z^k Z_k \sqrt{\frac{\operatorname{tr}(S^{(n)})^k}{k}},$$

where $(Z_k)_{k=1,2,\dots}$ is a sequence of independent complex Gaussian random variables such that

$$\mathbb{E}Z_k = 0, \quad \mathbb{E}|Z_k|^2 = 1, \quad \textit{and} \quad \mathbb{E}Z_k^2 = (\mathbb{E}W_{11}^2)^k.$$

is a well-defined $\mathbb{H}$–valued random variable. The sequence (F_n) is tight in $\mathbb{H}$, and the sequence (κ_n) satisfies for each compact set $\mathcal{K} \subset D(0,1)$:

$$0 < \liminf_n \min_{z \in \mathcal{K}} |\kappa_n(z)| \le \limsup_n \max_{z \in \mathcal{K}} |\kappa_n(z)| < \infty. \tag{1.3}$$

Finally, it is true that

$$q_n(z) \sim_n \kappa_n(z) \exp(-F_n(z)) \tag{1.4}$$

as $\mathbb{H}$–valued random variables.

Theorem 1.3 can be deduced from Theorem 1.4 by an argument provided in Section 3.6.

2 Case studies

In this section, we describe some matrix models for which Assumptions 1.1 and 1.2 are satisfied. The proofs related with this section are provided in Section 4.

2.1 Block variance profile

Fix $d, m \in \mathbb{N}$, and let $\boldsymbol{A}$ be a $d \times m$ matrix with positive entries. Set $p = md$. For each $n > 0$, define the matrix $A^{(pn)}$ as

$$A^{(pn)} = \frac{1}{pn} \boldsymbol{A} \otimes (\mathbf{1}_{mn} \mathbf{1}_{dn}^\top) \in \mathbb{R}^{pn \times pn}, \tag{2.1}$$

where $\otimes$ denotes the Kronecker product of matrices, and $\mathbf{1}_m$ is the vector of all ones in $\mathbb{R}^m$. Thus, $A^{(pn)}$ consists in md rectangular blocks of positive numbers, and the dimensions of each block scale with n.

Proposition 2.1. *In the above setting, $\rho(A^{(pn)}) > 0$, and the matrix $S^{(pn)} = \rho(A^{(pn)})^{-1} A^{(pn)}$ satisfies Assumptions 1.1 and 1.2.*

Proof. We check these assumptions with $K_n = n$. Since all the elements of $\boldsymbol{A}$ are positive, the minimum row sum of $A^{(pn)}$ lies in a compact interval of $\mathbb{R}_+$ away from zero. Thus, $\rho(A^{(pn)}) \ge c$ for some constant $c > 0$ [1, Th. 8.1.22]. As a consequence, $S^{(pn)}$ exists and complies with Assumptions 1.1. Furthermore, $\rho(S^{(pn)}) = 1$, and the rank r of $S^{(pn)}$ is upper bounded by $\min(d, m)$. Assumption 1.2 follows from the inequality $\det(I - \gamma S^{(pn)}) \ge (1 - \gamma)^r$ for $\gamma \in [0, 1)$. □

We thus obtain from Theorem 1.3 that $\mathbb{P}[\rho(X^{(pn)}) \ge 1 + \varepsilon] \to_n 0$ for each $\varepsilon > 0$. If we take out the normalization by $\rho(A^{(pn)})$ in the construction of $S^{(pn)}$, we of course obtain that $\mathbb{P}[\rho(X^{(pn)}) \ge \sqrt{\rho(A^{(pn)})} + \varepsilon] \to_n 0$ for each $\varepsilon > 0$.

2.2 Sampling a continuous variance profile

It is well known that any continuous function $\boldsymbol{S} : [0,1]^2 \to \mathbb{R}_+$, seen as an integral operator on the Banach space $C([0,1])$, is a compact trace-class operator [16]. If its spectral radius $\rho(\boldsymbol{S})$ is positive, we normalize our operator with $\rho(\boldsymbol{S})$ which amounts to assuming that $\rho(\boldsymbol{S}) = 1$. If $\rho(\boldsymbol{S}) = 0$, then, we replace $\boldsymbol{S}$ with $C_{\boldsymbol{S}} \boldsymbol{S}$ where $C_{\boldsymbol{S}} > 0$ is an arbitrarily large constant.

For $n > 0$, our variance profile matrix $S^{(n)}$ will be obtained by sampling regularly the function $\boldsymbol{S}$ on the rectangle $[0,1]^2$, namely, by setting $s_{ij}^{(n)} = n^{-1} \boldsymbol{S}(i/n, j/n)$.

It is obvious that Assumption 1.1 is satisfied by $S^{(n)}$ with $K_n = n$. The validity of Assumption 1.2 is the object of the following proposition which proof follows from standard arguments. We include it in Section 4 for completeness.

Proposition 2.2. *In the setting described above, it holds that*

$$\lim_n \min_{\gamma \in [0,1-\varepsilon]} \det(I_n - \gamma S^{(n)}) = \lim_n \det(I_n - (1-\varepsilon) S^{(n)}) = \det(\boldsymbol{I} - (1-\varepsilon)\boldsymbol{S}) > 0$$

for each $\varepsilon \in (0,1]$*, where* $\boldsymbol{I}$ *is the identity operator on the Banach space* $C([0,1])$*, and* $\det(\boldsymbol{I} - (1-\varepsilon)\boldsymbol{S})$ *is a Fredholm determinant.*

Thus, Assumption 1.2 holds true, and Theorem 1.3 follows. Consequently, if we get back to our original operator $\boldsymbol{S}$ (before the multiplication by $\rho(\boldsymbol{S})^{-1}$ or by $C_{\boldsymbol{S}}$), we obtain that $\mathbb{P}[\rho(X^{(n)}) \geq \sqrt{\rho(\boldsymbol{S})} + \varepsilon] \to_n 0$ for each $\varepsilon > 0$. In particular, $\rho(X^{(n)}) \xrightarrow{\mathcal{P}} 0$ if $\rho(\boldsymbol{S}) = 0$.

2.3 Random sparse sampling à la Erdős–Rényi of a continuous variance profile

We now provide an example of a (semi-)sparse model covered by our result. In this example, the sequence of matrices $(S^{(n)})$ will be random and independent of the array $(W_{ij})_{ij \geq 1}$. We shall show that Assumptions 1.1 and 1.2 will be satisfied with high probability (to be made precise below). In these conditions, Theorem 1.3 will be obtained by conditioning on an appropriate event which indicator is $S^{(n)}$–measurable.

Let $\boldsymbol{S} : [0,1]^2 \to [0,1]$ be a continuous function as in the previous section. Assume for simplicity that the spectral radius $\rho(\boldsymbol{S})$ of $\boldsymbol{S}$, seen as an operator, is positive. Let us consider that $\rho(\boldsymbol{S}) = 1$. Our variance profile matrix $S^{(n)}$ will be obtained by randomly sampling the function $\boldsymbol{S}$. Define the matrix

$$\mathsf{S}^{(n)} = \left[\mathsf{S}^{(n)}_{ij}\right]_{i,j=1}^{n} = \frac{1}{n}\left[\boldsymbol{S}(i/n, j/n)\right]_{i,j=1}^{n}.$$

Let (K_n) be a sequence of positive numbers such that $K_n \to \infty$ and $K_n = o(n)$. For some large enough integer $n_0 > 0$, define the sequence of random matrices $(B^{(n)} = [B^{(n)}_{ij}])_{n \geq n_0}$ as follows:

$$B^{(n)}_{ij} = \begin{cases} 1 & \text{with probability } K_n \mathsf{S}^{(n)}_{ij} \\ 0 & \text{with probability } 1 - K_n \mathsf{S}^{(n)}_{ij}, \end{cases}$$

and the random variables $\{B^{(n)}_{ij}\}_{i,j \in [n]}$ are independent. Let

$$S^{(n)} = \frac{1}{K_n} B^{(n)}.$$

Trivially, $K_n \| S^{(n)} \|_\infty \leq 1$ for each elementary event, where $\| \cdot \|_\infty$ is the max norm. Therefore, Assumption 1.1–(ii) is satisfied. Moreover,

Proposition 2.3. *Assume that* $K_n \geq \log n$*. Then, there exists a constant* $C_S > 0$ *such that* $\limsup_n \left|\!\left|\!\left| S^{(n)} \right|\!\right|\!\right| \leq C_S$ *w.p. 1.*

Furthermore, for each $\varepsilon \in (0,1]$*, it holds that*

$$\min_{\gamma \in [0,1-\varepsilon]} \det\left(I_n - \gamma S^{(n)}\right) \xrightarrow[n \to \infty]{\mathcal{P}} \det(\boldsymbol{I} - (1-\varepsilon)\boldsymbol{S}) > 0, \tag{2.2}$$

where $\xrightarrow{\mathcal{P}}$ *is the convergence in probability.*

Corollary 2.4. *Assume that* $K_n \geq \log n$*. Then,* $\mathbb{P}\left[\rho(X^{(n)}) \geq 1 + \varepsilon\right] \to_n 0$ *for each* $\varepsilon > 0$*.*

Remark 2.5. If $W_{1,1}$ is a Rademacher random variable, then the matrix $X^{(n)}$ can be considered the centered adjacency matrix of a directed inhomogeneous Erdős–Rényi graph. In recent years, there has been tremendous attention on the spectrum of undirected inhomogeneous Erdős–Rényi models (see, for example, [17], [18], [19], and [20]). A similar result to Corollary 2.4, is proven in Theorem 3.4 of [21].

Remark 2.6. The condition $K_n \geq \log n$ in the statement of Proposition 2.3 is required to obtain that $\limsup_n \left|\!\left|\!\left| S^{(n)} \right|\!\right|\!\right| \leq C_S$ w.p.1., which leads to Corollary 2.4 by the conditioning on the event $\mathcal{E}_n$ that we make in Section 4.3. We believe that this condition is not necessary to obtain Theorem 1.3. Indeed, it is possible to obtain a an analogue of Theorem 1.4 by including all the randomness of our model within the matrix $X^{(n)}$ (without conditioning), and by simply taking $\mathsf{S}^{(n)}$ as the variance profile matrix. We shall not develop this issue here.

Let us provide a simple example where the condition $K_n \geq \log n$ is avoided while still using our Theorem 1.4 to obtain the spectral radius confinement stated by Theorem 1.3.

2.4 Random sparse sampling of a continuous variance profile with a fixed outer degree

We still consider a operator represented by a continuous function $\boldsymbol{S} : [0,1]^2 \to [0,1]$ such that $\rho(\boldsymbol{S}) = 1$. Our variance profile matrix $S^{(n)}$ is now obtained by randomly sampling the function $\boldsymbol{S}$ as follows. Let (K_n) be a sequence of positive numbers such that $K_n \to \infty$ and $K_n = o(n)$. Let $\mathcal{I}^{(n)}$ a random sub-set of $[n]$ which is uniformly distributed among the $\binom{n}{K_n}$ sub-sets of $[n]$ with cardinality K_n. Let $\mathcal{I}_1^{(n)}, \ldots, \mathcal{I}_n^{(n)}$ be i.i.d. subsets of $[n]$ such that $\mathcal{I}_1^{(n)}$ is equal to $\mathcal{I}^{(n)}$ in distribution. Define the $\{0,1\}^{n\times n}$–valued random matrix $R^{(n)} = [R_{ij}^{(n)}]$ as

$$R_{ij}^{(n)} = \begin{cases} 1 & \text{if } j \in \mathcal{I}_i^{(n)} \\ 0 & \text{if } j \notin \mathcal{I}_i^{(n)} \end{cases}$$

Finaly, let $S^{(n)} = [S_{ij}^{(n)}]$ be defined as

$$S_{ij}^{(n)} = \frac{n}{K_n} \mathsf{S}_{ij}^{(n)} R_{ij}^{(n)}.$$

Trivially, $K_n \|S^{(n)}\|_\infty = \|\boldsymbol{S}\|_\infty$ and $\left|\left|\left| S^{(n)} \right|\right|\right| \leq \|\boldsymbol{S}\|_\infty$, where $\|\boldsymbol{S}\|_\infty$ is the norm of $\boldsymbol{S}$ on $C([0,1])$. Regarding Assumption 1.2, we have:

Proposition 2.7. *For $\varepsilon \in (0,1]$, it holds that*

$$\min_{\gamma \in [0,1-\varepsilon]} \det\left(I_n - \gamma S^{(n)}\right) \xrightarrow[n\to\infty]{\mathcal{P}} \det\left(\boldsymbol{I} - (1-\varepsilon)\boldsymbol{S}\right) > 0.$$

We close this section with a final remark.

Remark 2.8. One can show that in the last three application examples, Theorem 1.4 can be reformulated by stating that the sequence (q_n) converges in distribution. We state without further comment the expression of the limit in distribution $\boldsymbol{q} \in \mathbb{H}$, which is the same in the three cases. This limit reads:

$$\boldsymbol{q}(z) = \sqrt{\det(\boldsymbol{I} - z^2 \mathbb{E} W_{1,1}^2 \boldsymbol{S})} \exp\left(- \sum_{k=1}^{\infty} z^k Z_k \sqrt{\frac{\operatorname{tr} \boldsymbol{S}^k}{k}} \right), \quad z \in D(0,1),$$

where

$$\operatorname{tr} \boldsymbol{S}^k = \int_{[0,1]^k} \boldsymbol{S}(x_1,x_2)\boldsymbol{S}(x_2,x_3)\ldots\boldsymbol{S}(x_k,x_1) \prod_{i=1}^{k} dx_i. \tag{2.3}$$

3 Proof of Theorems 1.3 and 1.4

In all the remainder, $C > 0$ is a generic constant independent of n that can change from a display to another. In the proofs, the superscript $^{(n)}$ such as in $X^{(n)}$ will be often removed for notational simplicity. Given a matrix $M \in \mathbb{C}^{n\times n}$ and a set $I \subset [n]$ with cardinality $|I|$, we denote as M_I the $\mathbb{C}^{|I|\times|I|}$ sub-matrix of M consisting of the rows and columns which indices belong to I. Given a function $f : [0,1]^2 \to \mathbb{R}$, we denote as $f\begin{pmatrix} x_1 & x_2 & \cdots & x_k \\ x_1 & x_2 & \cdots & x_k \end{pmatrix}$ the $k\times k$ matrix which element (i,j) is $f(x_i,x_j)$.

We start with the proof of Theorem 1.4.

3.1 Proof of Theorem 1.4: preliminary results on $\mathbb{H}$–valued random variables.

Before entering the proof of Theorem 1.4, it will be useful to recall first some basic results on the convergence in distribution of $\mathbb{H}$–valued random variables. The reader is referred to *e.g.* [22] (see also [8]) for more details on this subject.

Proposition 3.1. *Let (f_n) be a sequence of random elements valued in $\mathbb{H}$. If, for each compact set $\mathcal{K} \subset D(0,1)$, the sequence of random variables $(\max_{z\in\mathcal{K}} |f_n(z)|)_n$ is tight, then, (f_n) is tight. For this condition to hold, it is enough that $\mathbb{E}|f_n(z)|^p \leq g(z)$ for $p \geq 1$, where $g(z)$ is bounded on the compacts of $D(0,1)$.*

Proposition 3.2. *Let (f_n) be a tight sequence of random elements valued in $\mathbb{H}$. Denote as $f_n(z) = \sum_{k=0}^{\infty} a_k^{(n)} z^k$ the power series representation of f_n in $D(0,1)$. Assume that there exists a sequence $a_0, a_1, \dots$ of random variables such that for each positive integer m, the m–uple $(a_0^{(n)}, \dots, a_m^{(n)})$ converges in distribution to $(a_0, \dots, a_m)$ as $n \to \infty$. Then, the function $f(z) = \sum_{k=0}^{\infty} a_k z^k$ is well-defined as a random element valued in $\mathbb{H}$, and (f_n) converges in distribution to f.*

This proposition can be easily modified to obtain the following result, which is better suited to our context:

Proposition 3.3. *Let (f_n) and (g_n) be two tight sequences of random elements valued in $\mathbb{H}$. Denote as $f_n(z) = \sum_{k=0}^{\infty} a_k^{(n)} z^k$ and $g_n(z) = \sum_{k=0}^{\infty} b_k^{(n)} z^k$ the power series representations of f_n and g_n in $D(0,1)$ respectively. If for each fixed positive integer m, it holds that $(a_0^{(n)}, \dots, a_m^{(n)}) \sim_n (b_0^{(n)}, \dots, b_m^{(n)})$, then $f_n \sim_n g_n$.*

To establish Theorem 1.4, we start by writing $q_n(z)$ as

$$q_n(z) = \det(1 - zX^{(n)}) = 1 + \sum_{k=1}^{n} (-z)^k P_k^{(n)}, \tag{3.1}$$

where

$$P_k^{(n)} = \sum_{I \subset [n] : |I| = k} \det X_I^{(n)}.$$

3.2 Tightness of (q_n)

Our first result pertains to the tightness of the sequence of $\mathbb{H}$–valued random variables (q_n):

Proposition 3.4. *The sequence (q_n) is tight.*

To prove this proposition, we need the following result.

Lemma 3.5. *Let Assumption 1.2 hold true. Then,*

$$\forall \varepsilon > 0, \ \sup_n \operatorname{perm}\left(I + (1-\varepsilon) S^{(n)}\right) < \infty,$$

where $\operatorname{perm} M$ *is the permanent of the matrix* M.

Proof. Given $n > 0$, we identify the matrix $S^{(n)}$ with an integral kernel to which we apply the Fredholm permanent theory developed in [23]. Our kernel $\boldsymbol{S}^{(n)} : [0,1) \times [0,1) \to \mathbb{R}_+$ is defined as $\boldsymbol{S}^{(n)}(x,y) = nS^{(n)}(i,j)$ when $(x,y) \in \left[\frac{i-1}{n}, \frac{i}{n}\right) \times \left[\frac{j-1}{n}, \frac{j}{n}\right), i,j \in [n]$. Following [23], define the function $\boldsymbol{p}^{(n)} : \mathbb{C} \to \mathbb{C}$ through the power series

$$\boldsymbol{p}^{(n)}(w) = 1 + \sum_{k=1}^{\infty} \boldsymbol{p}_k^{(n)} w^k,$$

with

$$\boldsymbol{p}_k^{(n)} = \frac{1}{k!} \int_0^1 \int_0^1 \cdots \int_0^1 \operatorname{perm} \boldsymbol{S}^{(n)} \begin{pmatrix} x_1 & x_2 & \cdots x_k \\ x_1 & x_2 & \cdots x_k \end{pmatrix} dx_1 dx_2 \dots dx_k.$$

Notice that $|\boldsymbol{p}_k^{(n)}| \leq (n\|S^{(n)}\|_\infty)^k$. Thus, the radius of convergence $R^{(n)}$ of this series satisfies $R^{(n)} \geq 1/(n\|S^{(n)}\|_\infty) > 0$, which shows that there exists a centered open disk where $\boldsymbol{p}^{(n)}(w)$ is well-defined and analytic. Let $\boldsymbol{d} : \mathbb{C} \to \mathbb{C}$ be given by the series

$$\boldsymbol{d}^{(n)}(w) = 1 + \sum_{k=1}^{\infty} (-1)^k \boldsymbol{d}_k^{(n)} w^k,$$

with

$$\boldsymbol{d}_k^{(n)} = \frac{1}{k!} \int_0^1 \int_0^1 \cdots \int_0^1 \det \boldsymbol{S}^{(n)} \begin{pmatrix} x_1 & x_2 & \cdots x_k \\ x_1 & x_2 & \cdots x_k \end{pmatrix} dx_1 dx_2 \dots dx_k.$$

It is easy to see that for each $k \in [n]$, it holds that

$$\boldsymbol{d}_k^{(n)} = \int_0^1 dx_1 \int_0^{x_1} dx_2 \cdots \int_0^{x_{k-1}} dx_k \det \boldsymbol{S}^{(n)} \begin{pmatrix} x_1 & x_2 & \cdots x_k \\ x_1 & x_2 & \cdots x_k \end{pmatrix} = \sum_{I \subset [n], |I| = k} \det S_I^{(n)},$$

and $\boldsymbol{d}_k^{(n)} = 0$ for $k > n$. Therefore, $\boldsymbol{d}^{(n)}(w)$ coincides with the reverse characteristic polynomial

$$\boldsymbol{d}^{(n)}(w) = \det\left(I - wS^{(n)}\right).$$

Theorem 4.4 (a) of [23] states that

$$\boldsymbol{d}^{(n)}(w)\boldsymbol{p}^{(n)}(w) = 1 \tag{3.2}$$

for w in the open disk of radius $R^{(n)}$, and by analytic continuation, on the whole $\mathbb{C}$. Since $\boldsymbol{p}_k^{(n)} \geq 0$ for each k, the spectral radius $R^{(n)}$ is a singular point of $\boldsymbol{p}^{(n)}(w)$ (see [24, Fact 7.21]), and thus, it is a zero of $\boldsymbol{d}^{(n)}(w)$ by the previous identity. By Assumption (1.2), we then obtain that $\liminf_n R^{(n)} \geq 1$, and by Identity (3.2) again, it holds that

$$\forall \varepsilon > 0, \quad \sup_n \boldsymbol{p}^{(n)}(1-\varepsilon) < \infty.$$

Let $p^{(n)}(w) = \operatorname{perm}\left(I + wS^{(n)}\right)$. As is well known (see, *e.g.*, [25, Th. 1.4]), $p^{(n)}(w) = 1 + \sum_{k=1}^n p_k^{(n)} w^k$ with

$$p_k^{(n)} = \sum_{I \subset [n], |I| = k} \operatorname{perm} S_I^{(n)} \quad \text{for } k \in [n].$$

Writing

$$\boldsymbol{p}_k^{(n)} = \int_0^1 dx_1 \int_0^{x_1} dx_2 \cdots \int_0^{x_{k-1}} dx_k \operatorname{perm} \boldsymbol{S}^{(n)} \begin{pmatrix} x_1 & x_2 & \cdots x_k \\ x_1 & x_2 & \cdots x_k \end{pmatrix},$$

it is easy to see that $p_k^{(n)} \leq \boldsymbol{p}_k^{(n)}$ for each $k \in [n]$, therefore,

$$\forall \varepsilon > 0, \quad \sup_n p^{(n)}(1-\varepsilon) < \infty,$$

which is the required result. □

Proof of Proposition 3.4. To prove our proposition, we bound $\mathbb{E}|q_n(z)|^2$ and use Proposition 3.1.

Denoting as $\mathfrak{S}_I$ the group of permutations over a set $I \subset [n]$, and $\operatorname{sign}(\sigma)$ the signature of a permutation σ, we first observe that

$$\mathbb{E} \det X_I = \sum_{\sigma \in \mathfrak{S}_I} \operatorname{sign}(\sigma) \mathbb{E} \prod_{i \in I} X_{i,\sigma(i)} = 0,$$

since the entries of X are centered.

Similarly for $J, I \subset [n]$ such that $I \neq J$ it is true that

$$\mathbb{E} \det X_I \overline{\det X_J} = 0.$$

Lastly, for any $I \subset [n]$,

$$\mathbb{E}|\det X_I|^2 = \mathbb{E} \det X_I \overline{\det X_I} = \sum_{\sigma \in \mathfrak{S}_I} \prod_{i \in I} s_{i,\sigma(i)} \mathbb{E}|W_{i,\sigma(i)}|^2 = \operatorname{perm} S_I.$$

We therefore have

$$\mathbb{E}|q_n(z)|^2 = \mathbb{E}\left|1 + \sum_{k=1}^n (-z)^k \sum_{I \subseteq [n] : |I| = k} \det X_I\right|^2 = \operatorname{perm}\left(I + |z^2| S^{(n)}\right),$$

which is bounded by the previous lemma on the compacts of $D(0,1)$. □

3.3 Asymptotics of the finite-dimensional distributions when W_{11} is bounded

Having established the tightness of (q_n), it remains to examine the distributional large–n properties of the random vector $(P_1^{(n)}, \ldots, P_k^{(n)})$ for each fixed integer $k > 0$, and apply Proposition 3.3 above. To this end, we temporarily assume that the random variables W_{ij} are bounded by a constant. We also rely on the fact that in order to study the distribution of $(P_1^{(n)}, \ldots, P_k^{(n)})$, it is enough to study the distribution of $(\operatorname{tr} X^{(n)}, \ldots, \operatorname{tr}(X^{(n)})^k)$ for large n, a much easier task. Specifically, for $z \in \mathbb{C}$, the series $\sum_{k=1}^{\infty}(z^k/k)(X^{(n)})^k$ is well-defined for $|z|$ small enough, and we can express $q_n(z)$ as

$$q_n(z) = \exp\left(-\sum_{k=1}^{\infty} \operatorname{tr}((X^{(n)})^k)\frac{z^k}{k}\right) \tag{3.3}$$

for $|z|$ small enough. Recalling the identity (3.1), we obtain that the k–uple $(P_1^{(n)}, \ldots, P_k^{(n)})$ is a polynomial function of $(\operatorname{tr} X^{(n)}, \ldots, \operatorname{tr}(X^{(n)})^k)$ independent of the dimension n, by an expansion to a power series on both sides of (3.3) and by examining at the first k terms of the expansion. Thus, we end up that in order to prove the asymptotic equivalence in (1.4), it is sufficient to examine the large-n distributions of the vectors $(\operatorname{tr} X^{(n)}, \ldots, \operatorname{tr}(X^{(n)})^k)$ for any integer $k > 0$. We shall analyze the distributions of these vectors with the help of the moment method, which explains why the assumed boundedness of the W_{ij} is important in our proof. Of course, the finiteness of all their moments would have been enough.

Proposition 3.6. *Assume that the random variables W_{ij} are bounded by a constant. Consider the sequence of independent complex-valued Gaussian random variables $(Z_\ell)_{\ell \geq 1}$ defined in the statement of Theorem 1.4. For each integers $n, \ell > 0$, define $\boldsymbol{m}_\ell^{(n)}$ as*

$$\boldsymbol{m}_\ell^{(n)} = \begin{cases} \left(\mathbb{E}W_{11}^2\right)^{\ell/2} \operatorname{tr}((S^{(n)})^{\ell/2}) & \text{if } \ell \text{ is even,} \\ 0 & \text{if } \ell \text{ is odd.} \end{cases}$$

Then, for each fixed integer $k > 0$, the asymptotic equivalence

$$\left(\operatorname{tr} X^{(n)}, \ldots, \operatorname{tr}(X^{(n)})^k\right) \sim_n \left(\sqrt{\operatorname{tr} S^{(n)}} Z_1 + \boldsymbol{m}_1^{(n)}, \ldots, \sqrt{k \operatorname{tr}(S^{(n)})^k} Z_k + \boldsymbol{m}_k^{(n)}\right) \tag{3.4}$$

holds true.

Most of the remainder of this section is devoted to the proof of this proposition. We start with a simple lemma.

Lemma 3.7. *Let Assumptions 1.1 and 1.2 hold true. Then*

$$\forall k > 0, \exists C > 0, \ \operatorname{tr} S^k \leq C \text{ and } \|S^k\|_\infty \leq C/K_n.$$

Proof. Using Assumption 1.2 and recalling the development (1.2), we obtain the first bound by setting, *e.g.*, $\gamma = 1/2$.

Assumption 1.1–(ii) asserts that $\|S\|_\infty \leq C/K_n$, thus, the second bound is effective for $k = 1$. Assume without generality loss that $|||S||| \leq C$ from Assumption 1.1–(i). For $k > 1$, we have $\|S^k\|_\infty \leq |||S||| \|S^{k-1}\|_\infty \leq \cdots \leq |||S|||^{k-1} \|S\|_\infty \leq C/K_n$. □

Given a k–uple $\boldsymbol{I} = (i_1, \ldots, i_k) \subset [n]^k$, we write

$$X_{\boldsymbol{I}} = X_{i_1 i_2} X_{i_2 i_3} \ldots X_{i_{k-1} i_k} X_{i_k i_1}.$$

As in [8], we write $\operatorname{tr}(X^k) = \sum_{\boldsymbol{I} \in [n]^k} X_{\boldsymbol{I}}$ as

$$\operatorname{tr}(X^{(n)})^k = R_k^{(n)} + Q_k^{(n)},$$

where, denoting as $\mathcal{D}_k$ the sub-set of $[n]^k$ defined as

$$\mathcal{D}_k = \{(i_1, \ldots, i_k) \in [n]^k \ : \ \forall j \neq \ell \in [k], i_j \neq i_\ell\},$$

we set

$$R_k = \sum_{\boldsymbol{I} \in \mathcal{D}_k} X_{\boldsymbol{I}}, \quad \text{and} \quad Q_k = \sum_{\boldsymbol{I} \in [n]^k \setminus \mathcal{D}_k} X_{\boldsymbol{I}}.$$

It is obvious that $\mathbb{E}R_k = 0$. The analogues of R_k and Q_k are called in [8] the "random term" and the "deterministic term" respectively. We shall deal with these two terms separately. The following two lemmas are proven in Section 3.3 below.

Lemma 3.8. *Let the m–uple $(k_1, \ldots, k_m)$ be as in the statement of Proposition 3.6. Given $x \in \mathbb{C}$, use the notation $x^s = x$ when $s = \cdot$ and $x^s = \bar{x}$ when $s = *$. Let $s_1, \ldots, s_m \in \{\cdot, *\}$.*

If m is even, and if there exists at least one partition P of $[m]$ into pairs such that $k_\ell = k_{\ell'}$ if ℓ and ℓ' are a pair (notation $\{\ell, \ell'\} \in P$), then,

$$\mathbb{E}\left[R_{k_1}^{s_1} R_{k_2}^{s_2} \ldots R_{k_m}^{s_m}\right] - \sum_{P \in \boldsymbol{P}} \prod_{\{\ell,\ell'\} \in P} \left(k_\ell \left(\mathbb{E} W_{11}^{s_\ell} W_{11}^{s_{\ell'}} \right)^{k_\ell} \operatorname{tr} S^{k_\ell} \right) \xrightarrow[n\to\infty]{} 0,$$

where $\boldsymbol{P}$ is the set of such partitions. Otherwise,

$$\mathbb{E}\left[R_{k_1}^{s_1} R_{k_2}^{s_2} \ldots R_{k_m}^{s_m}\right] \xrightarrow[n\to\infty]{} 0.$$

Lemma 3.9. *It holds that*

$$Q_k - \boldsymbol{m}_k \xrightarrow[n\to\infty]{\mathcal{P}} 0.$$

Proof of Proposition 3.6.

By Lemma 3.7, for each sequence (n) of integers, there exists a sub-sequence such that for every integer $\ell > 0$, $\operatorname{tr}(S^{(n)})^\ell$ converges to some real number $\boldsymbol{s}_\ell$ along this sub-sequence. Fix an integer $k > 0$. Lemma 3.8 along with the Isserlis/Wick theorem show that $(R_1^{(n)}, \ldots, R_k^{(n)})$ converges in distribution along this sub-sequence to $(\sqrt{\boldsymbol{s}_1} Z_1, \ldots \sqrt{k \boldsymbol{s}_k} Z_k)$. By Lemma 3.9, for $\ell \in [k]$, $Q_\ell^{(n)}$ converges in probability along this sub-sequence to $\left(\mathbb{E} W_{11}^2\right)^{\ell/2} \boldsymbol{s}_{\ell/2}$ if ℓ is even and to zero if ℓ is odd. The result stated by Proposition 3.6 follows.

Proofs of Lemmas 3.8 and 3.9

The following preliminary result will be needed.

Lemma 3.10. *Let Assumptions 1.1 and 1.2 hold true. Let $k_1, \ldots, k_m$ be positive integers, and write $k = k_1 + \cdots + k_m$. Decomposing a k–uple $\boldsymbol{I} \in [n]^k$ as $\boldsymbol{I} = (\boldsymbol{I}_1, \ldots, \boldsymbol{I}_m)$ where $\boldsymbol{I}_j \in [n]^{k_j}$, it holds that there exists $C > 0$ such that*

$$0 \le \operatorname{tr} S^{k_1} \ldots \operatorname{tr} S^{k_m} - \sum_{\boldsymbol{I} \in \mathcal{D}_k} S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m} \le C / K_n.$$

Proof. Observe first that $\operatorname{tr} S^{k_1} \ldots \operatorname{tr} S^{k_m} = \sum_{\boldsymbol{I} \in [n]^k} S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m}$. If $k = 1$, the result is trivial. Assume not. The indicator function $\mathbb{1}_{\mathcal{D}_k}(\boldsymbol{I})$ with $\boldsymbol{I} = (i_1, \ldots, i_k)$ can be encoded into the product of the $k(k-1)/2$ indicators of the type $\mathbb{1}_{i_j \neq i_\ell}$ for $j \neq \ell \in [k]$. Let us order the constraints $i_j \neq i_\ell$ in some way from 1 to $k(k-1)/2$, and let us write $\mathbb{1}^{(m)}(\boldsymbol{I})$ as the product of the indicators on the first m constraints, so that $\mathbb{1}^{(k(k-1)/2)}(\boldsymbol{I}) = \mathbb{1}_{\mathcal{D}_k}(\boldsymbol{I})$. Writing $\mathbb{1}^{(0)} \equiv 1$, we have

$$\sum_{\boldsymbol{I} \in [n]^k} S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m} - \sum_{\boldsymbol{I} \in \mathcal{D}_k} S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m} = \sum_{m=0}^{k(k-1)/2-1} \sum_{\boldsymbol{I} \in [n]^k} (\mathbb{1}^{(m)}(\boldsymbol{I}) - \mathbb{1}^{(m+1)}(\boldsymbol{I})) S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m}.$$

Write $\mathbb{1}^{(m)}(\boldsymbol{I}) - \mathbb{1}^{(m+1)}(\boldsymbol{I}) = \mathbb{1}^{(m)}(\boldsymbol{I}) \mathbb{1}_{i_j = i_\ell}$ for some $j \neq \ell \in [k]$. Assuming, *e.g.*, $k_1 \ge 2$, $j = 1$, and $\ell \le k_1$, we obtain

$$\begin{aligned}
\sum_{\boldsymbol{I} \in [n]^k} (\mathbb{1}^{(m)}(\boldsymbol{I}) - \mathbb{1}^{(m+1)}(\boldsymbol{I})) S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m} &\le \sum_{\boldsymbol{I} \in [n]^k} \mathbb{1}_{i_1 = i_\ell} S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m} \\
&= \sum_{i_1, \ldots, i_{\ell-1}, i_{\ell+1}, \ldots, i_{k_1}} s_{i_1 i_2} \ldots s_{i_{\ell-1} i_1} \, s_{i_1 i_{\ell+1}} \ldots s_{i_{k_1} i_1} \, \operatorname{tr} S^{k_2} \ldots \operatorname{tr} S^{k_m} \\
&\le C \sum_{i_1} \left[S^{\ell-1}\right]_{i_1 i_1} \left[S^{k_1 - \ell + 1}\right]_{i_1 i_1} \le \frac{C}{K_n} \operatorname{tr} S^{k_1 - \ell + 1} \le \frac{C}{K_n}
\end{aligned}$$

by Lemma 3.7. The cases where the indices j and ℓ belong to two different t–uples $\boldsymbol{I}_r$ are treated similarly. □

Proof of Lemma 3.8. We deal with the expression

$$R_{k_1}^{s_1} R_{k_2}^{s_2} \ldots R_{k_m}^{s_m} = \sum_{(\boldsymbol{I}_1, \ldots, \boldsymbol{I}_m) \in \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_m}} X_{\boldsymbol{I}_1}^{s_1} \ldots X_{\boldsymbol{I}_m}^{s_m}.$$

We introduce some new notation. We let $k = k_1 + \cdots k_m$, and we write

$$\boldsymbol{I} = (\boldsymbol{I}_1, \ldots, \boldsymbol{I}_m) = \big((i_1^1, \ldots, i_{k_1}^1), \ldots, (i_1^m, \ldots, i_{k_m}^m)\big) \in [n]^k.$$

In what follows, it is always meant that $i^\ell_{k_\ell + j} = i^\ell_j$ for $j = 1, \ldots, k_\ell - 1$.

When $k = 1$ $(= m)$, it is obvious that $\mathbb{E} R_1^{s_1} = 0$, thus the lemma is true. Assume that $k > 1$, and define the two sets $\mathcal{A}, \mathcal{B} \subset \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_m}$ as

$$\begin{aligned} \mathcal{A} = \Big\{ \boldsymbol{I} \in \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_m} \ &: \text{each couple } (i_j^\ell, i_{j+1}^\ell) \text{ appears exactly twice in } \boldsymbol{I}, \\ &\quad \text{each index } i_j^\ell \text{ appears exactly twice in } \boldsymbol{I} \Big\}, \\ \mathcal{B} = \Big\{ \boldsymbol{I} \in \mathcal{D}_{k_1} \times \cdots \times \mathcal{D}_{k_m} \ &: \text{each couple } (i_j^\ell, i_{j+1}^\ell) \text{ appears at least twice in } \boldsymbol{I}, \\ &\quad \text{there exists an index } i_j^\ell \text{ that appears three times at least in } \boldsymbol{I} \Big\}. \end{aligned}$$

Since the elements of the matrix X are centered, $\mathbb{E} X^{s_1}_{\boldsymbol{I}_1} \ldots X^{s_m}_{\boldsymbol{I}_m}$ is equal to zero if there exists a couple (i_j^ℓ, i_{j+1}^ℓ) that appears only once within $\boldsymbol{I}$. This implies that

$$\mathbb{E} R_{k_1}^{s_1} R_{k_2}^{s_2} \ldots R_{k_m}^{s_m} = \sum_{\boldsymbol{I} \in \mathcal{A}} \mathbb{E} X^{s_1}_{\boldsymbol{I}_1} \ldots X^{s_m}_{\boldsymbol{I}_m} + \sum_{\boldsymbol{I} \in \mathcal{B}} \mathbb{E} X^{s_1}_{\boldsymbol{I}_1} \ldots X^{s_m}_{\boldsymbol{I}_m} \tag{3.5}$$

We now show that

$$\left| \sum_{\boldsymbol{I} \in \mathcal{B}} \mathbb{E} X^{s_1}_{\boldsymbol{I}_1} \ldots X^{s_m}_{\boldsymbol{I}_m} \right| \leq \frac{C}{\sqrt{K_n}}. \tag{3.6}$$

There are $2^{k(k-1)/2}$ ways of constructing an indicator function on $[n]^k$ defined as a product of indicators of the type $\mathbb{1}_{i_i^\ell = i_{i'}^{\ell'}}$ and indicators of the type $\mathbb{1}_{i_i^\ell \neq i_{i'}^{\ell'}}$, where this product involves all the $k(k-1)/2$ sets of the type $\{(\ell, i), (\ell', i')\}$ with cardinality 2. There is a sub-set of these functions that completely describes the set $\mathcal{B}$ in the sense that we can write

$$\mathbb{1}_{\mathcal{B}}(\boldsymbol{I}) = \sum_{r=1}^{C_B} f_r(\boldsymbol{I})$$

where the functions f_r are chosen appropriately in the family that we just defined, and where $C_B = C_B(k_1, \ldots, k_m)$ is the number of these functions. To establish (3.6), we show that for each $r \in [C_B]$, it holds that $\sum_{\boldsymbol{I} \in [n]^k} |\mathbb{E} X^{s_1}_{\boldsymbol{I}_1} \ldots X^{s_m}_{\boldsymbol{I}_m}| f_r(\boldsymbol{I}) \leq C/\sqrt{K_n}$. Relying on the boundedness of the elements of X, we write

$$\sum_{\boldsymbol{I} \in [n]^k} |\mathbb{E} X^{s_1}_{\boldsymbol{I}_1} \ldots X^{s_m}_{\boldsymbol{I}_m}| f_r(\boldsymbol{I}) \leq C \sum_{\boldsymbol{I} \in [n]^k} \sqrt{S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m}} f_r(\boldsymbol{I}). \tag{3.7}$$

To deal with this expression, we need to introduce some new notations. Given two d–uples $\boldsymbol{J} = (j_1, \ldots, j_d) \in [n]^d$ and $\boldsymbol{a} = (a_1, \ldots, a_d)$ with $a_i \in \{1, 3/2, 2, 5/2, \ldots\}$, we write $|\boldsymbol{J}| = |\boldsymbol{a}| = d$,

$$S^{\boldsymbol{a}}_{\boldsymbol{J}} = s^{a_1}_{j_1, j_2} s^{a_2}_{j_2, j_3} \ldots s^{a_d}_{j_{d-1}, j_d} s^{a_d}_{j_d, j_1} \quad \text{and} \quad S^{\boldsymbol{a}}_{\overset{\circ}{\boldsymbol{J}}} = s^{a_1}_{j_1, j_2} s^{a_2}_{j_2, j_3} \ldots s^{a_d}_{j_{d-1}, j_d}.$$

Of course, $S_{\boldsymbol{J}} = S^{\boldsymbol{1}_d}_{\boldsymbol{J}}$. We also write $S_{\overset{\circ}{\boldsymbol{J}}} = S^{\boldsymbol{1}_d}_{\overset{\circ}{\boldsymbol{J}}}$. With these notations, the right-hand side of (3.7) can be re-expressed as follows. By merging all the couples (i^ℓ_j, i^ℓ_{j+1}) that are forced to be identical in the encoding by f_r, and keeping after a merger the couple (i^ℓ_j, i^ℓ_{j+1}) with the smallest value of ℓ, we can observe after a possible index renumbering that there exists:

- An integer $p > 0$, t–uples $\boldsymbol{J}_1, \ldots, \boldsymbol{J}_p$, and $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p$ such that $|\boldsymbol{J}_\ell| = |\boldsymbol{a}_\ell| \in \{k_1, \ldots, k_m\}$ for $\ell \in [p]$,
- An integer $q \geq 0$, and, when $q > 0$, t–uples $\boldsymbol{J}_{p+1}, \ldots, \boldsymbol{J}_{p+q}$ and $\boldsymbol{a}_{p+1}, \ldots, \boldsymbol{a}_{p+q}$ with $|\boldsymbol{J}_{p+\ell}| = |\boldsymbol{a}_{p+\ell}|$,

such that

$$\sum_{\boldsymbol{I} \in [n]^k} \sqrt{S_{\boldsymbol{I}_1} \ldots S_{\boldsymbol{I}_m}} f_r(\boldsymbol{I}) = \sum_{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_p, \boldsymbol{J}_{p+1}, \ldots, \boldsymbol{J}_{p+q}} S^{\boldsymbol{a}_1}_{\boldsymbol{J}_1} \cdots S^{\boldsymbol{a}_p}_{\boldsymbol{J}_p} S^{\boldsymbol{a}_{p+1}}_{\overset{\circ}{\boldsymbol{J}}_{p+1}} \cdots S^{\boldsymbol{a}_{p+q}}_{\overset{\circ}{\boldsymbol{J}}_{p+q}} g(\boldsymbol{J}_1, \ldots, \boldsymbol{J}_{p+q}) := \chi, \tag{3.8}$$

an expression that we now explain. The function $g(\boldsymbol{J}_1, \ldots, \boldsymbol{J}_{p+q})$ is a product of indicators that encodes the residual constraints after the merger of the couples. Let us see an initial t–uple $\boldsymbol{I}_\ell$ as constituting a cycle $i_1^\ell \to i_2^\ell \to \cdots \to i_{k_\ell}^\ell \to i_1^\ell$. After the merger, $\boldsymbol{I}_1$ gives rise to $\boldsymbol{J}_1$. The cycle in $\boldsymbol{J}_1$ is not broken when performing the sum in (3.8), since the couple (i_j^ℓ, i_{j+1}^ℓ) with the smallest value of ℓ is kept after a merger. Pursuing, the t–uples $\boldsymbol{J}_2, \ldots, \boldsymbol{J}_p$ give rise to unbroken cycles in the expression (3.8). The other $\boldsymbol{J}_{p+\ell}$'s, when they exist, correspond to broken cycles. Let us consider $\boldsymbol{J}_{p+1}$. An important feature of this t–uple is that its extremities are connected to the $\boldsymbol{J}_\ell$'s for $\ell \in [p]$ by the merger procedure. In other words, writing in the remainder $\boldsymbol{J}_\ell = (j_1^\ell, \ldots, j_{|\boldsymbol{J}_\ell|}^\ell)$, there exists within the function g a product of indicators of the type $\mathbb{1}_{j_1^{p+1} = \times} \mathbb{1}_{j_{|\boldsymbol{J}_{p+1}|}^{p+1} = \times'}$ where the indices $\times$ and $\times'$ belong to the $\boldsymbol{J}_\ell$ for $\ell \in [p]$. Pursuing this process, the extremities of the t–uple $\boldsymbol{J}_{p+q}$ are connected to the $\boldsymbol{J}_\ell$'s for $\ell \in [p+q-1]$.

We now use these observations to bound χ. We shall repeatidly use the bound $S_{\boldsymbol{J}_\ell}^{\boldsymbol{a}_\ell} \leq C S_{\boldsymbol{J}_\ell}$ and $S_{\overset{\circ}{\boldsymbol{J}}_\ell}^{\boldsymbol{a}_\ell} \leq C S_{\overset{\circ}{\boldsymbol{J}}_\ell}$ due to $\|S\|_\infty \leq 1$ for all large n. According to the form of the function f_r, at least one of the three following situations occurs:

- $q > 0$. Assume for the sake of example that $q = 1$, and write $j_1^{p+1} = j_i^\ell$ and $j_{|\boldsymbol{J}_{p+1}|}^{p+1} = j_{i'}^{\ell'}$ where j_i^ℓ and $j_{i'}^{\ell'}$ are found in $\boldsymbol{J}_1, \ldots, \boldsymbol{J}_p$. Using Lemma 3.7, we have
$$\chi \leq C \sum_{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_p, \boldsymbol{J}_{p+1}} S_{\boldsymbol{J}_1} \cdots S_{\boldsymbol{J}_p} S_{\overset{\circ}{\boldsymbol{J}}_{p+1}} \mathbb{1}_{j_1^{p+1} = j_i^\ell} \mathbb{1}_{j_{|\boldsymbol{J}_{p+1}|}^{p+1} = j_{i'}^{\ell'}} = C \sum_{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_p} S_{\boldsymbol{J}_1} \cdots S_{\boldsymbol{J}_p} \left[S^{|\boldsymbol{J}_{p+1}|-1} \right]_{j_i^\ell j_{i'}^{\ell'}}$$
$$\leq \frac{C}{K_n} \operatorname{tr} S^{|\boldsymbol{J}_1|} \ldots \operatorname{tr} S^{|\boldsymbol{J}_p|} \leq \frac{C}{K_n}.$$
For general q, we get the bound C/K_n^q by iterating this argument backwards, starting with $\boldsymbol{J}_{p+q}$.
- $q = 0$, and there exists an exponent within the $\boldsymbol{a}_\ell$ that is $\geq 3/2$. Here it is easy to observe that $\chi \leq C/\sqrt{K_n}$.
- $q = 0$ and all the vectors $\boldsymbol{a}_\ell$ are made of ones. Since there is an index that appears at least three times in $f_r(\boldsymbol{I})$ by the definition of $\mathcal{B}$, there is an index that appears at least two times in the $\boldsymbol{J}_\ell$'s. Say this index is j_1^1 with $j_1^1 = j_1^2$. Then,
$$\chi \leq \sum_{\boldsymbol{J}_1, \ldots, \boldsymbol{J}_p} S_{\boldsymbol{J}_1} \cdots S_{\boldsymbol{J}_p} \mathbb{1}_{j_1^1 = j_1^2} = C \sum_{\boldsymbol{J}_1, \boldsymbol{J}_2} S_{\boldsymbol{J}_1} S_{\boldsymbol{J}_2} \mathbb{1}_{j_1^1 = j_1^2} \leq C \sum_{j_1} \left[S^{|\boldsymbol{J}_1|-1} \right]_{j_1 j_1} \left[S^{|\boldsymbol{J}_2|-1} \right]_{j_1 j_1} \leq \frac{C}{K_n}.$$

This establishes Inequality (3.6).

Getting back to (3.5), we now deal with the term $\sum_{\boldsymbol{I} \in \mathcal{A}} \mathbb{E} X_{\boldsymbol{I}_1}^{s_1} \ldots X_{\boldsymbol{I}_m}^{s_m}$. Here, one can check that a necessary condition for $\mathcal{A}$ to be nonempty is that m is even and the t–uples $\boldsymbol{I}_1$, ..., $\boldsymbol{I}_m$ can be grouped into pairs of equal length. In other words, the set $\boldsymbol{P}$ of pair partitions P of $[m]$ as specified in the statement of the lemma is not empty. The case being, we have
$$\mathbb{1}_{\mathcal{A}}(\boldsymbol{I}) = \sum_{P \in \boldsymbol{P}} h_P(\boldsymbol{I}) \prod_{\{\ell, \ell'\} \in P} \left(\prod_{j=1}^{k_\ell} \mathbb{1}_{i_j^\ell = i_j^{\ell'}} + \prod_{j=1}^{k_\ell} \mathbb{1}_{i_{j+1}^\ell = i_j^{\ell'}} + \cdots + \prod_{j=1}^{k_\ell} \mathbb{1}_{i_{j+k_\ell-1}^\ell = i_j^{\ell'}} \right),$$
where $h_P(\boldsymbol{I}) \in \{0,1\}$ forces the indices $i_1^\ell, \ldots, i_{k_\ell}^\ell$ within the pair $\{\ell, \ell'\} \in P$ to be different, and to be different from the indices within all the other pairs. To better understand the previous formula, let us see once again the t–uples $\boldsymbol{I}_\ell$ as cycles. When $\{\ell, \ell'\} \in P$, we need to make the cycles associated to $\boldsymbol{I}_\ell$ and $\boldsymbol{I}_{\ell'}$ coincide, and there are k_ℓ ways to do this. This corresponds to the sum of products within the parenthesis of the last display.

With this identity, we have
$$\begin{aligned}
&\sum_{\boldsymbol{I} \in \mathcal{A}} \mathbb{E} X_{\boldsymbol{I}_1}^{s_1} \ldots X_{\boldsymbol{I}_m}^{s_m} \\
&= \sum_{P \in \boldsymbol{P}} \sum_{\boldsymbol{I} \in [n]^k} h_P(\boldsymbol{I}) \prod_{\{\ell, \ell'\} \in P} \left(\prod_{j=1}^{k_\ell} \mathbb{1}_{i_j^\ell = i_j^{\ell'}} + \prod_{j=1}^{k_\ell} \mathbb{1}_{i_{j+1}^\ell = i_j^{\ell'}} + \cdots + \prod_{j=1}^{k_\ell} \mathbb{1}_{i_{j+k_\ell-1}^\ell = i_j^{\ell'}} \right) \mathbb{E} X_{\boldsymbol{I}_1}^{s_1} \ldots X_{\boldsymbol{I}_m}^{s_m} \\
&= \sum_{P \in \boldsymbol{P}} \left(\prod_{\{\ell, \ell'\} \in P} k_\ell \left(\mathbb{E} W_{11}^{s_\ell} W_{11}^{s_{\ell'}} \right)^{k_\ell} \right) \sum_{\neq} \prod_{\{\ell, \ell'\} \in P} S_{\boldsymbol{I}_\ell},
\end{aligned}$$

where $\sum_{\neq}$ is the sum over all the $m/2$ t–uples $\boldsymbol{I}_\ell$ such that $\{\ell, \ell'\} \in P$, with the constraint that all the indices that belong to these t–uples are different. Thanks to Lemma 3.10, we obtain that

$$\sum_{\boldsymbol{I} \in \mathcal{A}} \mathbb{E} X_{\boldsymbol{I}_1}^{s_1} \ldots X_{\boldsymbol{I}_m}^{s_m} = \sum_{P \in \boldsymbol{P}} \prod_{\{\ell,\ell'\} \in P} \left(k_\ell \left(\mathbb{E} W_{11}^{s_\ell} W_{11}^{s_{\ell'}} \right)^{k_\ell} \operatorname{tr} S^{k_\ell} \right) + \varepsilon,$$

where $|\varepsilon| \le C/K_n$. Recalling (3.6), our lemma is proven. □

Proof of Lemma 3.9. We first evaluate the asymptotics of $\mathbb{E} Q_k$. Assuming k is even, let us focus on the case where the indices of $\boldsymbol{I} = (i_1, \ldots, i_k)$ satisfy the constraints $i_j = i_{k/2+j}$ for all $j \in [k/2]$ and $|\{i_1, \ldots, i_{k/2}\}| = k/2$, generating the double cycle $i_1 \to \cdots i_{k/2} \to i_1 \to \cdots i_{k/2} \to i_1$. The expectation of the sum over the $\boldsymbol{I}$ with these constraints is $\sum_{\boldsymbol{J} \in \mathcal{D}_{k/2}} \left(\mathbb{E} W_{11}^2 \right)^{k/2} S_{\boldsymbol{J}} = \left(\mathbb{E} W_{11}^2 \right)^{k/2} \operatorname{tr} S^{k/2} + \mathcal{O}(1/K_n) = \boldsymbol{m}_k + \mathcal{O}(1/K_n)$ thanks to Lemma 3.10. When they exist, all other possibly non-zero contributions to $\mathbb{E} Q_k$, including k being odd, correspond to the couples (i_j, i_{j+1}) appearing two times at least and one index appearing three times at least. By treating these cases similarly to what we did for the term $\sum_{\boldsymbol{I} \in \mathcal{B}} \cdots$ in the previous lemma, we can show that these cases are bounded by $C/\sqrt{K_n}$. We thus have

$$\mathbb{E} Q_k - \boldsymbol{m}_k \xrightarrow[n \to \infty]{} 0.$$

To establish the result of the lemma, we now show that the variance $\mathbb{V}\mathrm{ar}(Q_k)$ of Q_k converges to zero. Write

$$\mathbb{V}\mathrm{ar}\, Q_k = \sum_{\boldsymbol{I}_1 = (i_1^1, \ldots, i_k^1), \boldsymbol{I}_2 = (i_1^2, \ldots, i_k^2) \in [n]^k \setminus \mathcal{D}_k} \mathbb{E}(X_{\boldsymbol{I}_1} - \mathbb{E} X_{\boldsymbol{I}_1})(X_{\boldsymbol{I}_2} - \mathbb{E} X_{\boldsymbol{I}_2}).$$

We observe here that

$$|\mathbb{E}(X_{\boldsymbol{I}_1} - \mathbb{E} X_{\boldsymbol{I}_1})(X_{\boldsymbol{I}_2} - \mathbb{E} X_{\boldsymbol{I}_2})| \le C \sqrt{s_{i_1^1 i_2^1} \cdots s_{i_k^1 i_1^1}} \sqrt{s_{i_1^2 i_2^2} \cdots s_{i_k^2 i_1^2}}.$$

Moreover, the left hand side is equal to zero unless every couple (i_j^ℓ, i_{j+1}^ℓ) appears at least twice in $\boldsymbol{I} = (\boldsymbol{I}_1, \boldsymbol{I}_2)$, and at least one of these couples is common to $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$. In a manner similar to what we did in the proof of the previous lemma, these constraints can be encoded into functions that have the form of products of the type $\mathbb{1}_{i_j^\ell = i_{j'}^{\ell'}}$. The number of such functions does not depend on n. Fixing one of these functions, we merge the identical couples within $\boldsymbol{I}_1$ and within $\boldsymbol{I}_2$, keeping the lowest indices, and then we merge what remains in $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$, keeping the lowest exponent. By doing so, we get an expression similar to (3.8) in the proof of the previous lemma. Re-using the notations of that proof and repeating the argument there, the case where $q > 0$ and the case $q = 0$ with an exponent $\ge 3/2$ have negligible contributions. Let us deal with the case where $q = 0$ and where all the exponents are equal to 1. In this case, each couple (i_j^ℓ, i_{j+1}^ℓ) appears exactly twice, and we recall that there is one couple common to $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$. Also observe that since $\boldsymbol{I}_1, \boldsymbol{I}_2 \in [n]^k \setminus \mathcal{D}_k$, there is at least one index repetition within $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$. In these conditions, one can check that the only available possibilities are of the form $\sum_{\boldsymbol{J}_1, \boldsymbol{J}_2} S_{\boldsymbol{J}_1} S_{\boldsymbol{J}_2} \mathbb{1}_\times$, where $\mathbb{1}_\times$ links an index in $\boldsymbol{J}_1$ to an index in $\boldsymbol{J}_2$. This leads to a negligible contribution. Lemma 3.9 is proven. □

3.4 Proof of Theorem 1.4 when W_{11} is bounded

We begin by establishing the properties of the functions κ_n and F_n provided in the statement of Theorem 1.4. Recalling that $\limsup \rho(S^{(n)}) \le 1$, the function κ_n is well-defined as an element of $\mathbb{H}$ for all large n. Moreover, using Assumption (1.2) and recalling the development (1.2), we obtain that for all small $\varepsilon > 0$,

$$\max_{\gamma \in [0, 1-\varepsilon]} \sum_{k=1}^\infty \gamma^k \frac{\operatorname{tr}(S^{(n)})^k}{k} < \infty.$$

Therefore, for each compact $\mathcal{K} \subset D(0,1)$, it holds that

$$\begin{aligned}
\limsup_n \max_{z \in \mathcal{K}} \left| \log \det(I_n - z^2 \mathbb{E} W_{11}^2 S^{(n)}) \right| &= \limsup_n \max_{z \in \mathcal{K}} \left| \sum_{k=1}^\infty (z^2 \mathbb{E} W_{11}^2)^k \frac{\operatorname{tr}(S^{(n)})^k}{k} \right| \\
&\le \limsup_n \max_{z \in \mathcal{K}} \sum_{k=1}^\infty |z|^{2k} \frac{\operatorname{tr}(S^{(n)})^k}{k} < \infty,
\end{aligned}$$

and the bounds in (1.3) hold true. Regarding $F_n(z)$, we can check by, *e.g.*, a moment calculation that for each $n > 0$, it holds that

$$\limsup_{k\to\infty} \frac{|Z_k|^{1/k}(\operatorname{tr}(S^{(n)})^k)^{1/(2k)}}{k^{1/(2k)}} \leq 1 \quad \text{w.p. } 1$$

therefore, $F_n(z)$ is well-defined as a $\mathbb{H}$–valued random variable. Moreover, we can see that $\mathbb{E}|F_n(z)|^2 = \sum_{k\geq 1} |z|^{2k} \operatorname{tr}(S^{(n)})^k/k = -\log\det(1-|z|^2 S^{(n)})$ which is upper bounded on the compacts of $D(0,1)$. Therefore (F_n) is tight.

As a consequence, the function $g_n(z) = \kappa_n(z)\exp(-F_n(z))$ is a well-defined $\mathbb{H}$-valued random variable, and the sequence (g_n) is tight. Write $g_n(z) = 1 + \sum_{k\geq 1} G_k^{(n)}(-z)^k$, and recall that $q_n(z) = 1 + \sum_{k=1}^n P_k^{(n)}(-z)^k$. We need to show that for each fixed integer $k > 0$, the asymptotic equivalence

$$(P_1^{(n)}, \ldots, P_k^{(n)}) \sim_n (G_1^{(n)}, \ldots, G_k^{(n)}) \tag{3.9}$$

holds true. By applying Propositions 3.4 and 3.3, Theorem 1.4 will then be proven.

Recalling the discussion that precedes Proposition 3.6, we can write $(P_1^{(n)}, \ldots, P_k^{(n)}) = Q(\operatorname{tr} X^{(n)}, \ldots, \operatorname{tr}(X^{(n)})^k)$ for some polynomial Q independent of n. Notice also that g_n can be written as

$$g_n(z) = \exp\left(-\sum_{k=1}^{\infty} z^k \left(Z_k\sqrt{\frac{\operatorname{tr}(S^{(n)})^k}{k}} + \frac{\boldsymbol{m}_k^{(n)}}{k}\right)\right).$$

Therefore, by applying the same argument as for q_n, we obtain that

$$(G_1^{(n)}, \ldots, G_k^{(n)}) = Q\Big(Z_1\sqrt{\operatorname{tr} S^{(n)}} + \boldsymbol{m}_1^{(n)}, \ldots, Z_k\sqrt{k\operatorname{tr}(S^{(n)})^k} + \boldsymbol{m}_k^{(n)}\Big),$$

for the same polynomial Q, and (3.9) follows from Proposition 3.6.

3.5 Releasing the boundedness assumption on W_{11}. End of the proof of Theorem 1.4

To finish the proof of Theorem 1.4, all what remains to prove is the truth of the asymptotic equivalence (3.9) when W_{11} has a second moment without any additional assumption. As in [8], we truncate the W_{ij}'s by writing $W_{ij}^{(M)} = \mathbb{1}_{|W_{i,j}|\leq M} W_{i,j} - \mathbb{E}\mathbb{1}_{|W_{i,j}|\leq M} W_{i,j}$ for some $M > 0$, and we show that the truncation error is negligible when M is large. One of the ideas of [8] is that it is much easier to control the effect of this truncation on the coefficients $P_k^{(n)}$ rather than on the traces $\operatorname{tr}(X^{(n)})^k$, as it is frequently done in random matrix theory.

Lemma 3.11. *For $M > 0$, let $W_{ij}^{(M)}$ be defined as*

$$W_{ij}^{(M)} = \mathbb{1}_{|W_{i,j}|\leq M} W_{i,j} - \mathbb{E}\mathbb{1}_{|W_{i,j}|\leq M} W_{i,j}.$$

Define the matrix $X^{(n,M)} = \left[X_{i,j}^{(n,M)}\right]_{i,j\in[n]}$ as $X_{i,j}^{(n,M)} := \sqrt{s_{i,j}^{(n)}}W_{i,j}^{(M)}$ for $i,j \in [n]$. For $k \in [n]$, let $P_k^{(n,M)}$ be given as

$$P_k^{(n,M)} = \sum_{I\subset[n]:|I|=k} \det X_I^{(n,M)} \quad \text{for } k \in [n].$$

Then, for each fixed integer $k > 0$, the bound

$$\sup_n \mathbb{E}|P_k^{(n)} - P_k^{(n,M)}|^2 \leq \varepsilon_M$$

holds true, with $\varepsilon_M \to 0$ as $M \to \infty$.

Proof. Recalling that the polynomial $p^{(n)}(w) = \operatorname{perm}\left(I + wS^{(n)}\right)$ introduced in the proof of Lemma 3.5 can be written as $p^{(n)}(w) = 1 + \sum_{k=1}^{n} p_k^{(n)} w^k$ with $p_k^{(n)} = \sum_{I\subset[n],|I|=k} \operatorname{perm} S_I^{(n)}$ for $k \in [n]$, we write

$$\begin{aligned}
\mathbb{E}|P_k^{(n)} - P_k^{(n,M)}|^2 &= \sum_{I\subset[n]:|I|=k} \mathbb{E}\left|\det X_I^{(n)} - \det X_I^{(n,M)}\right|^2 \\
&= \sum_{I\subset[n]:|I|=k} \sum_{\sigma\in\mathfrak{S}_I} \mathbb{E}\left|\prod_{i\in I} X_{i,\sigma(i)}^{(n)} - \prod_{i\in I} X_{i,\sigma(i)}^{(n,M)}\right|^2 \\
&= \mathbb{E}\left|W_{1,1}\cdots W_{1,k} - W_{1,1}^{(M)}\cdots W_{1,k}^{(M)}\right|^2 \sum_{I\subset[n]:|I|=k} \sum_{\sigma\in\mathfrak{S}_I} \prod_{i\in I} s_{i,\sigma(i)} \\
&= \mathbb{E}\left|W_{1,1}\cdots W_{1,k} - W_{1,1}^{(M)}\cdots W_{1,k}^{(M)}\right|^2 p_k^{(n)}.
\end{aligned}$$

Observing that

$$\sup_n p_k^{(n)} \le \frac{\sup_n p^{(n)}(1/2)}{(1/2)^k} < \infty$$

by Lemma 3.5, we can clearly choose ε_M as

$$\varepsilon_M = \mathbb{E}\left|W_{1,1}\cdots W_{1,k} - W_{1,1}^{(M)}\cdots W_{1,k}^{(M)}\right|^2 \sup_n p_k^{(n)}.$$

□

Fix an integer $k > 0$, and let $\varphi : \mathbb{C}^k \to \mathbb{R}$ be a Lipschitz and bounded function. Define the sequence of independent Gaussian random variables $(Z_\ell^{(M)})_{\ell\ge1}$ as $\mathbb{E}Z_\ell^{(M)} = 0$, $\mathbb{E}|Z_\ell^{(M)}|^2 = \mathbb{E}|W_{11}^{(M)}|^2$, and $\mathbb{E}[(Z_\ell^{(M)})^2] = \mathbb{E}[(W_{11}^{(M)})^2]^\ell$. By a slight modification of the proof of Proposition 3.6, we obtain an asymptotic equivalence similar to (3.4), namely, for $M > 0$ large enough,

$$\left(\operatorname{tr} X^{(n,M)}, \dots, \operatorname{tr}(X^{(n,M)})^k\right) \sim_n \left(\sqrt{\operatorname{tr} S^{(n)}} Z_1^{(M)} + \boldsymbol{m}_1^{(n,M)}, \dots, \sqrt{k \operatorname{tr}(S^{(n)})^k} Z_k^{(M)} + \boldsymbol{m}_k^{(n,M)}\right).$$

where $\boldsymbol{m}_k^{(n,M)}$ has the same expression as $\boldsymbol{m}_k^{(n)}$ with W_{11} being replaced with $W_{11}^{(M)}$. As in the last sub-section, we have $(P_1^{(n,M)}, \dots, P_k^{(n,M)}) = Q(\operatorname{tr} X^{(n,M)}, \dots, \operatorname{tr}(X^{(n,M)})^k)$, where Q is a polynomial independent of n. Therefore,

$$\mathbb{E}\varphi(P_1^{(n,M)}, \dots, P_k^{(n,M)}) - \mathbb{E}\varphi \circ Q\Big(\sqrt{\operatorname{tr} S^{(n)}} Z_1^{(M)} + \boldsymbol{m}_1^{(n,M)}, \dots, \sqrt{k \operatorname{tr}(S^{(n)})^k} Z_k^{(M)} + \boldsymbol{m}_k^{(n,M)}\Big) \xrightarrow[n\to\infty]{} 0.$$

We also know that

$$\mathbb{E}\varphi(G_1^{(n)}, \dots, G_k^{(n)}) = \mathbb{E}\varphi \circ Q\Big(Z_1\sqrt{\operatorname{tr} S^{(n)}} + \boldsymbol{m}_1^{(n)}, \dots, Z_k\sqrt{\frac{\operatorname{tr}(S^{(n)})^k}{k}} + \boldsymbol{m}_k^{(n)}\Big).$$

Now, by the previous lemma, it holds that

$$\sup_n \left|\mathbb{E}\varphi(P_1^{(n,M)}, \dots, P_k^{(n,M)}) - \mathbb{E}\varphi(P_1^{(n)}, \dots, P_k^{(n)})\right| \xrightarrow[M\to\infty]{} 0.$$

Moreover, using that the traces $\operatorname{tr}(S^{(n)})^\ell$ are bounded by numbers independent of n, we also have

$$\begin{aligned}
\sup_n \Big|&\mathbb{E}\varphi \circ Q\Big(\sqrt{\operatorname{tr} S^{(n)}} Z_1 + \boldsymbol{m}_1^{(n)}, \dots, \sqrt{k \operatorname{tr}(S^{(n)})^k} Z_k + \boldsymbol{m}_k^{(n)}\Big) \\
&- \mathbb{E}\varphi \circ Q\Big(\sqrt{\operatorname{tr} S^{(n)}} Z_1^{(M)} + \boldsymbol{m}_1^{(n,M)}, \dots, \sqrt{k \operatorname{tr}(S^{(n)})^k} Z_k^{(M)} + \boldsymbol{m}_k^{(n,M)}\Big)\Big| \xrightarrow[M\to\infty]{} 0.
\end{aligned}$$

It results that

$$\mathbb{E}\varphi(P_1^{(n)}, \dots, P_k^{(n)}) - \mathbb{E}\varphi(G_1^{(n)}, \dots, G_k^{(n)}) \xrightarrow[n\to\infty]{} 0,$$

and the asymptotic equivalence (3.9) is established when W_{11} satisfies our general assumptions. This concludes the proof of Theorem 1.4.

3.6 Proof of Theorem 1.3 from Theorem 1.4

Denoting as $\overline{D}(0,r)$ the closed centered disk of $\mathbb{C}$ of radius r, the probability event $[\rho(X^{(n)}) \geq 1+\varepsilon]$ satisfies

$$\left[\rho(X^{(n)}) \geq 1+\varepsilon\right] = \left[\min_{z\in\overline{D}(0,r)} |q_n(z)| = 0\right],$$

where $r = 1/(1+\varepsilon)$. The function $\mathbb{H} \to [0,\infty), f \mapsto \min_{z\in\overline{D}(0,r)} |f(z)|$ is continuous. Therefore, recalling the notation $g_n = \kappa_n \exp(-F_n)$, we obtain from Theorem 1.4 that

$$\min_{z\in\overline{D}(0,r)} |q_n(z)| \sim_n \min_{z\in\overline{D}(0,r)} |g_n(z)|. \tag{3.10}$$

By Theorem 1.4, the deterministic sequence (κ_n) is precompact in $\mathbb{H}$ by the normal family theorem, and the random sequence (F_n) is tight. Take a sub-sequence, call it (n), along which (g_n) converges in distribution towards $g_\infty = \kappa_\infty \exp(-F_\infty)$, where $\kappa_\infty \in \mathbb{H}$ and F_∞ is a $\mathbb{H}$–valued random variable. From (3.10), we obtain that $\min_{z\in\overline{D}(0,r)} |q_n(z)|$ converges in distribution to $\min_{z\in\overline{D}(0,r)} |g_\infty(z)|$ along the sub-sequence (n). But $g_\infty(z) = 0$ if and only if $\kappa_\infty(z) = 0$ for each $z \in D(0,1)$, and furthermore, the equation $\kappa_\infty(z) = 0$ has no solution on $D(0,1)$ by (1.3). Therefore,

$$\lim_n \mathbb{P}\left[\min_{z\in\overline{D}(0,r)} |q_n(z)| > 0\right] = 1,$$

and Theorem 1.3 is proven.

4 Proofs for Section 2

4.1 Proof of Proposition 2.2

Let $\log$ be the branch of the logarithm that is analytical on $\mathbb{C} \setminus \mathbb{R}_-$, and observe that $|\log(1+z)| = |\sum_{i=1}^\infty z^i/i| \leq |z|/(1-|z|)$ for $|z| < 1$. Denoting as $\{\lambda_1, \lambda_2, \ldots\}$ the (eigenvalue) spectrum of the compact operator $\boldsymbol{S}$, and recalling that $\boldsymbol{S}$ is trace-class with $\rho(\boldsymbol{S}) \leq 1$, it holds that

$$\forall\gamma \in [0,1), \quad \sum_{k=1}^\infty |\log(1-\gamma\lambda_k)| \leq \sum_{k=1}^\infty \frac{\gamma|\lambda_k|}{1-\gamma|\lambda_k|} \leq \frac{\gamma}{1-\gamma}\sum_{k=1}^\infty |\lambda_k| < \infty.$$

Therefore, the Fredholm determinant $\det(\boldsymbol{I} - \gamma\boldsymbol{S})$ can be written for $\gamma \in [0,1)$ as

$$\det(\boldsymbol{I} - \gamma\boldsymbol{S}) = \prod_{k=1}^\infty (1-\gamma\lambda_k),$$

and satisfies $\min_{\gamma\in[0,1-\varepsilon]} \det(\boldsymbol{I} - \gamma\boldsymbol{S}) > 0$ for $\varepsilon \in (0,1]$.

We can also express $\det(\boldsymbol{I} - z\boldsymbol{S})$ as an analytic expansion for $z \in \mathbb{C}$. Namely, consider the series

$$f(z) = 1 + \sum_{k=1}^\infty (-1)^k \boldsymbol{d}_k z^k, \quad z \in \mathbb{C},$$

where

$$\boldsymbol{d}_k = \frac{1}{k!}\int_0^1\int_0^1 \cdots \int_0^1 \det \boldsymbol{S}\begin{pmatrix} x_1 & x_2 & \cdots x_k \\ x_1 & x_2 & \cdots x_k \end{pmatrix} dx_1 dx_2 \ldots dx_k.$$

Since, by Hadamard's inequality, $|\boldsymbol{d}_k| \leq \|\boldsymbol{S}\|_\infty^k/k!$, the function f is entire, and it is well-known to coincide with $\det(\boldsymbol{I} - z\boldsymbol{S})$. Now, as in the proof of Lemma 3.5, we interpret the matrix $S^{(n)}$ as a piece-wise constant approximation $\boldsymbol{S}^{(n)} : [0,1]^2 \to \mathbb{R}_+$ of the operator $\boldsymbol{S}$, by writing $\boldsymbol{S}^{(n)}(x,y) = ns_{ij}^{(n)} = \boldsymbol{S}(i/n, j/n)$ when $(x,y) \in \left[\frac{i-1}{n}, \frac{i}{n}\right) \times \left[\frac{j-1}{n}, \frac{j}{n}\right)$, $i,j \in [n]$ (completions on the right and upper borders irrelevant). With this, we have

$$\det\left(I - zS^{(n)}\right) = \det\left(\boldsymbol{I} - z\boldsymbol{S}^{(n)}\right) = 1 + \sum_{k=1}^n (-1)^k \boldsymbol{d}_k^{(n)} z^k$$

where

$$\boldsymbol{d}_k^{(n)} = \sum_{I\subset[n],|I|=k} \det S_I^{(n)} = \frac{1}{k!}\int_0^1\int_0^1\cdots\int_0^1 \det \boldsymbol{S}^{(n)}\begin{pmatrix} x_1 & x_2 & \cdots x_k \\ x_1 & x_2 & \cdots x_k\end{pmatrix} dx_1dx_2\ldots dx_k.$$

Similarly to $\boldsymbol{d}_k$, it holds that $|\boldsymbol{d}_k^{(n)}| \leq \|\boldsymbol{S}\|_\infty^k/k!$. Furthermore, $\boldsymbol{d}_k^{(n)} \to_n \boldsymbol{d}_k$ for each k thanks to the continuity of the kernel $\boldsymbol{S}$. With this at hand, one can check that the sequence of polynomials $\left(\det\left(I - zS^{(n)}\right)\right)_n$ is bounded on the compacts of $\mathbb{C}$ and converges point-wise to $\det(\boldsymbol{I} - z\boldsymbol{S})$. Thus, this convergence is uniform on the compacts of $\mathbb{C}$, and Proposition 2.2 follows.

4.2 Proof of Proposition 2.3

Write $S^{(n)} = [S_{ij}^{(n)}]_{i,j=1}^n$. For $i\in[n]$, define $b_i^{(n)}$ as

$$b_i^{(n)} = \sum_{j=1}^n \mathbb{E}S_{ij}^{(n)} = \frac{1}{n}\sum_{j=1}^n \boldsymbol{S}(i/n, j/n) \leq \|\boldsymbol{S}\|_\infty.$$

By Chernoff's theorem [26, Th. 2.3.1], it holds that

$$\mathbb{P}\left[\sum_{j=1}^n S_{ij}^{(n)} \geq t\right] = \mathbb{P}\left[\sum_{j=1}^n B_{ij}^{(n)} \geq tK_n\right] \leq \left(\frac{eb_i^{(n)}}{t}\right)^{tK_n} \leq \left(\frac{e\|\boldsymbol{S}\|_\infty}{t}\right)^{tK_n}$$

for $t > 0$ large enough. By the union bound, we therefore have

$$\mathbb{P}\left[\left|\!\left|\!\left|S^{(n)}\right|\!\right|\!\right| \geq t\right] \leq \exp\left(\log n + tK_n\log(e\|\boldsymbol{S}\|_\infty/t)\right).$$

Using that $K_n \geq \log n$, choosing t large enough and invoking the Borel-Cantelli lemma, we obtain the first assertion of Proposition 2.3.

To prove the second assertion, we work on the reverse characteristic polynomials

$$q_n^S(z) = \det(I_n - zS^{(n)}).$$

Following the general canvas of the proof of Theorem 1.4, we consider (q_n^S) as a sequence of $\mathbb{H}$–valued random variables. We first establish the tightness of this sequence. Second, we show that for each fixed integer $k > 0$, it holds that

$$\operatorname{tr}(S^{(n)})^k - \operatorname{tr}(\mathsf{S}^{(n)})^k \xrightarrow[n\to\infty]{\mathcal{P}} 0 \tag{4.1}$$

By an obvious continuity argument, we further know that $\operatorname{tr}(\mathsf{S}^{(n)})^k \to_n \operatorname{tr}\boldsymbol{S}^k$, where $\operatorname{tr}\boldsymbol{S}^k$ is given by (2.3). As a consequence,

$$\operatorname{tr}(S^{(n)})^k \xrightarrow[n\to\infty]{\mathcal{P}} \operatorname{tr}\boldsymbol{S}^k.$$

The tightness and these convergences for each $k > 0$ lead to the convergence $q_n^S(z) \xrightarrow{\mathcal{P}} \det(\boldsymbol{I} - z\boldsymbol{S})$ in $\mathbb{H}$ following the approach developed in the proof of Theorem 1.4, and the convergence (2.2) follows.

To establish the tightness of (q_n^S), we write as in (3.1)

$$q_n^S(z) = 1 + \sum_{k=1}^n (-z)^k \sum_{I\subset[n],|I|=k} \det S_I^{(n)}.$$

By the triangle inequality,

$$\mathbb{E}|q_n^S(z)| \leq 1 + \sum_{k=1}^n |z|^k \frac{1}{K_n^k} \sum_{I\subset[n],|I|=k}\sum_{\sigma\in\mathfrak{S}_I} \mathbb{E}\prod_{i\in I} B_{i,\sigma(i)}^{(n)} = \operatorname{perm}(I_n + |z|\mathsf{S}^{(n)}).$$

Remembering that $\min_{\gamma\in[0,1-\varepsilon]}\det(\boldsymbol{I} - \gamma\boldsymbol{S}) > 0$ and using an argument similar to the proof of Lemma 3.5, we obtain that $\operatorname{perm}(I_n + |z|\mathsf{S}^{(n)})$ is uniformly bounded on the compacts of $D(0,1)$. By Proposition 3.1, we obtain that (q_n^S) is tight.

It remains to establish the convergence (4.1) to finish the proof.

In what follows, let $\mathbf{P}_k$ denote the isomorphic classes (in the classical graph-theoretic sense) of closed walks which use k edges. We allow the walks to have multiple edges and loops. For each $P \in \mathbf{P}_k$ and $n \geq k$, we shall denote $P([n])$ the set of walks belonging to the isomorphic class of P and having vertices in $[n]$. Given a matrix $M = [M_{ij}]$, recall the notation $M_{\boldsymbol{J}} = M_{i_1 i_2} \dots M_{i_k i_1}$ for $\boldsymbol{J} = (i_1, \dots, i_k)$. We start by writing

$$\mathbb{E}\,\mathrm{tr}(S^{(n)})^k = \frac{1}{K_n^k} \sum_{P \in \mathbf{P}_k} \sum_{\boldsymbol{J} \in P([n])} \mathbb{E} B_{\boldsymbol{J}}^{(n)}.$$

Now, for any walk $P \in \mathbf{P}_k$, we have:

$$\frac{1}{K_n^k} \sum_{\boldsymbol{J} \in P([n])} \mathbb{E} B_{\boldsymbol{J}}^{(n)} \leq \frac{C}{K_n^k} n^{|\{\text{vertices of } P\}|} \left(\frac{K_n}{n}\right)^{|\{\text{distinct edges of } P\}|},$$

where the $|\{\text{distinct edges of } P\}|$ counts each edge of P once, ignoring repetitions.

Furthermore, the number of vertices of P is bounded by the number of distinct edges of P. This is true because we can start a closed walk in P and assign each distinct edge to a vertex, starting from the first edge during the walk that starts from that vertex. Since the number of distinct edges of P is clearly bounded by k, we conclude that for a walk P to have a non-negligible contribution in the above sum, the number of distinct edges should equal the number of vertices. Thus, the non-negligible paths P should satisfy $k =$ number of vertices of $P =$ number of distinct edges of P. Thus, denoting as $\mathcal{D}_k$ the sub-set of $[n]^k$ defined as

$$\mathcal{D}_k = \{(i_1, \dots, i_k) \in [n]^k : \forall l \neq j \in [k], i_l \neq i_j\},$$

we obtain that

$$\mathbb{E}\,\mathrm{tr}(S^{(n)})^k = \sum_{\boldsymbol{J} \in \mathcal{D}_k} \mathsf{S}_{\boldsymbol{J}}^{(n)} + o_n(1).$$

Now, we know that the matrix $\mathsf{S}^{(n}$ complies with Assumptions 1.1 and 1.2 by replacing the K_n in the statement of Assumption 1.1 with n. Therefore, by Lemma 3.10, it holds that

$$0 \leq \mathrm{tr}(\mathsf{S}^{(n)})^k - \sum_{\boldsymbol{J} \in \mathcal{D}_k} \mathsf{S}_{\boldsymbol{J}}^{(n)} \leq C/n.$$

Thus:

$$\mathbb{E}\,\mathrm{tr}(S^{(n)})^k = \mathrm{tr}(\mathsf{S}^{(n)})^k + o_n(1).$$

To obtain (4.1), it remains to show that $\mathbb{V}\mathrm{ar}\,\mathrm{tr}(S^{(n)})^k \to_n 0$. We have here

$$\mathbb{V}\mathrm{ar}\,\mathrm{tr}(S^{(n)})^k = \frac{1}{K_n^{2k}} \sum_{P_1, P_2 \in \mathbf{P}_k} \sum_{\boldsymbol{J}_1 \in P_1([n]), \boldsymbol{J}_2 \in P_2([n])} \mathbb{E}[(B_{\boldsymbol{J}_1}^{(n)} - \mathbb{E}B_{\boldsymbol{J}_1}^{(n)})(B_{\boldsymbol{J}_2}^{(n)} - \mathbb{E}B_{\boldsymbol{J}_2}^{(n)})].$$

Clearly, the summand is zero when the walks P_1 and P_2 have no common vertex. When these walks have a common vertex, we have by a similar argument as above that

$$\begin{aligned} &\frac{1}{K_n^{2k}} \sum_{\boldsymbol{J}_1 \in P_1([n]), \boldsymbol{J}_2 \in P_2([n])} \mathbb{E}[(B_{\boldsymbol{J}_1}^{(n)} - \mathbb{E}B_{\boldsymbol{J}_1}^{(n)})(B_{\boldsymbol{J}_2}^{(n)} - \mathbb{E}B_{\boldsymbol{J}_2}^{(n)})] \\ &\leq \frac{C}{K_n^{2k}} \left(\frac{K_n}{n}\right)^{|\{\text{distinct edges of } P_1\}| + |\{\text{distinct edges of } P_2\}|} n^{|\{\text{vertices of } P_1\}| + |\{\text{vertices of } P_2\}| - 1} \\ &\leq \frac{C}{n}. \end{aligned}$$

Thus, $\mathbb{V}\mathrm{ar}\,\mathrm{tr}(S^{(n)})^k \to_n 0$, and Proposition 2.3 is proven.

4.3 Corollary 2.4: Sketch of proof

Given a small $\varepsilon > 0$, let $\delta > 0$ be such that $\det(\boldsymbol{I} - (1-\varepsilon/2)\boldsymbol{S}) \geq \delta$, and define the probability event

$$\mathcal{E}_n = \left[\left|\!\left|\!\left| S^{(n)} \right|\!\right|\!\right| \leq 2C_S \text{ and } \min_{\gamma\in[0,1-\varepsilon/2]} \det\left(I_n - \gamma S^{(n)}\right) \geq \delta/2 \right].$$

Then we have

$$\mathbb{P}\left[\rho(X^{(n)}) \geq 1+\varepsilon\right] \leq \mathbb{P}\left[[\rho(X^{(n)}) \geq 1+\varepsilon] \mid \mathcal{E}_n\right] + \mathbb{P}\left[\mathcal{E}_n^{\mathrm{c}}\right].$$

By modifying the statement and the proof of Theorem 1.4 in such a way that $\mathbb{H}$ is replaced with the set of holomorphic functions on the open centered disk of $\mathbb{C}$ with radius $1-\varepsilon/2$, we obtain by a slight modification of the proof of Theorem 1.3 that $\mathbb{P}\left[[\rho(X^{(n)}) \geq 1+\varepsilon] \mid \mathcal{E}_n\right] \to_n 0$. It is clear by Proposition 2.3 that $\mathbb{P}\left[\mathcal{E}_n^{\mathrm{c}}\right] \to_n 0$, and we obtain the spectral confinement result $\mathbb{P}\left[\rho(X^{(n)}) \geq 1+\varepsilon\right] \to_n 0$.

4.4 Proposition 2.7: Sketch of proof

As for the previous proposition, we show that $q_n^S(z) = \det(I_n - zS^{(n)})$ converges in probability to $\det(\boldsymbol{I} - z\boldsymbol{S})$ in $\mathbb{H}$.

Using that the rows of $R^{(n)}$ are independent and that $\mathbb{E}R_{ij}^{(n)} = K_n/n$, thus, $\mathbb{E}S_{ij}^{(n)} = \mathsf{S}_{ij}^{(n)}$, we obtain that $\mathbb{E}|q_n^S(z)| \leq \operatorname{perm}(I + |z|\mathsf{S}^{(n)})$, hence the tightness of (q_n^S).

We now quickly show that $\operatorname{tr}(S^{(n)})^k \xrightarrow{\mathcal{P}} \operatorname{tr}\boldsymbol{S}^k$ for each fixed integer $k > 0$. Let $E^{(n)} = [E_{ij}^{(n)}]_{i,j=1}^n$ be a random matrix with i.i.d. Bernoulli elements such that $\mathbb{E}E_{11}^{(n)} = K_n/n$, and define the random matrix $Y^{(n)} = [Y_{ij}^{(n)}]_{i,j=1}^n$ as $Y_{ij}^{(n)} = (n/K_n)\mathsf{S}_{ij}^{(n)}E_{ij}^{(n)}$. Along the principle of the proof of Proposition 2.3, one can prove that $\mathbb{E}\operatorname{tr}(Y^{(n)})^k \to_n \operatorname{tr}\boldsymbol{S}^k$ and $\mathbb{V}\mathrm{ar}\operatorname{tr}(Y^{(n)})^k \to_n 0$. To show that $\operatorname{tr}(S^{(n)})^k \xrightarrow{\mathcal{P}} \operatorname{tr}\boldsymbol{S}^k$, it is enough to show that $\mathbb{E}\operatorname{tr}(S^{(n)})^k - \mathbb{E}\operatorname{tr}(Y^{(n)})^k \to_n 0$ and $\mathbb{E}(\operatorname{tr}(S^{(n)})^k)^2 - \mathbb{E}(\operatorname{tr}(Y^{(n)})^k)^2 \to_n 0$.

Given a fixed integer $\ell > 0$ and integers $i_1, \ldots i_\ell \in [n]$ which are all different, it holds that

$$\mathbb{E}R_{1i_1}^{(n)} \ldots R_{1i_\ell}^{(n)} = \mathbb{P}[(R_{1i_1}^{(n)}, \ldots, R_{1i_\ell}^{(n)}) = (1, \ldots, 1)] = \frac{\binom{n-\ell}{K_n-\ell}}{\binom{n}{K_n}} = \frac{(K_n - \ell + 1) \times \cdots \times K_n}{(n-\ell+1) \times \cdots \times n},$$

while

$$\mathbb{E}E_{1i_1}^{(n)} \ldots E_{1i_\ell}^{(n)} = \left(\frac{K_n}{n}\right)^\ell.$$

We therefore have after a small calculation that

$$0 \leq \mathbb{E}E_{1i_1}^{(n)} \ldots E_{1i_\ell}^{(n)} - \mathbb{E}R_{1i_1}^{(n)} \ldots R_{1i_\ell}^{(n)} \leq \frac{C_\ell}{K_n}\mathbb{E}E_{1i_1}^{(n)} \ldots E_{1i_\ell}^{(n)}, \tag{4.2}$$

where C_ℓ depends on ℓ only. Re-using the notations of the proof of Proposition 2.3, we now have

$$\begin{aligned} 0 \leq \mathbb{E}\operatorname{tr}(Y^{(n)})^k - \mathbb{E}\operatorname{tr}(S^{(n)})^k &= \frac{n^k}{K_n^k} \sum_{P\in\mathbf{P}_k} \sum_{\boldsymbol{J}\in P([n])} \mathsf{S}_{\boldsymbol{J}}^{(n)} \left(\mathbb{E}E_{\boldsymbol{J}}^{(n)} - \mathbb{E}R_{\boldsymbol{J}}^{(n)}\right) \\ &\leq \frac{C}{K_n}\frac{n^k}{K_n^k} \sum_{P\in\mathbf{P}_k} \sum_{\boldsymbol{J}\in P([n])} \mathsf{S}_{\boldsymbol{J}}^{(n)}\mathbb{E}E_{\boldsymbol{J}}^{(n)} \\ &= \frac{C}{K_n}\mathbb{E}\operatorname{tr}(Y^{(n)})^k, \end{aligned}$$

where, to obtain the inequality, we decomposed $\mathbb{E}R_{\boldsymbol{J}}^{(n)}$ into a product of expectations over the different rows of $R^{(n)}$, and we used the bound (4.2). Since $\mathbb{E}\operatorname{tr}(Y^{(n)})^k \to_n \operatorname{tr}\boldsymbol{S}^k$, we obtain that $\mathbb{E}\operatorname{tr}(S^{(n)})^k - \mathbb{E}\operatorname{tr}(Y^{(n)})^k \to_n 0$. The proof that $\mathbb{E}(\operatorname{tr}(S^{(n)})^k)^2 - \mathbb{E}(\operatorname{tr}(Y^{(n)})^k)^2 \to_n 0$ is similar.

References

[1] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, Cambridge, 1990, corrected reprint of the 1985 original.

[2] V. Girko, *Theory of stochastic canonical equations. Vol. I and II*, ser. Mathematics and its Applications. Kluwer Academic Publishers, Dordrecht, 2001, vol. 535. [Online]. Available: https://doi.org/10.1007/978-94-010-0989-8

[3] J. Alt, L. Erdős, and T. Krüger, “Local inhomogeneous circular law,” *Ann. Appl. Probab.*, vol. 28, no. 1, pp. 148–203, 2018. [Online]. Available: https://doi.org/10.1214/17-AAP1302

[4] ——, “Spectral radius of random matrices with independent entries,” *Probab. Math. Phys.*, vol. 2, no. 2, pp. 221–280, 2021. [Online]. Available: https://doi.org/10.2140/pmp.2021.2.221

[5] N. Cook, W. Hachem, J. Najim, and D. Renfrew, “Non-Hermitian random matrices with a variance profile (I): deterministic equivalents and limiting ESDs,” *Electron. J. Probab.*, vol. 23, pp. Paper No. 110, 61, 2018. [Online]. Available: https://doi.org/10.1214/18-ejp230

[6] H. Sompolinsky, A. Crisanti, and H.-J. Sommers, “Chaos in random neural networks,” *Phys. Rev. Lett.*, vol. 61, no. 3, pp. 259–262, 1988. [Online]. Available: https://doi.org/10.1103/PhysRevLett.61.259

[7] I. Akjouj, M. Barbier, M. Clenet, W. Hachem, M. Maïda, F. Massol, J. Najim, and V. C. Tran, “Complex systems in ecology: a guided tour with large Lotka-Volterra models and random matrices,” *Proc. A.*, vol. 480, no. 2285, pp. Paper No. 20 230 284, 35, 2024. [Online]. Available: https://doi.org/10.1098/rspa.2023.0284

[8] C. Bordenave, D. Chafaï, and D. García-Zelada, “Convergence of the spectral radius of a random matrix through its characteristic polynomial,” *Probab. Theory Related Fields*, vol. 182, no. 3-4, pp. 1163–1181, 2022. [Online]. Available: https://doi.org/10.1007/s00440-021-01079-9

[9] A. Basak and O. Zeitouni, “Outliers of random perturbations of Toeplitz matrices with finite symbols,” *Probab. Theory Related Fields*, vol. 178, no. 3-4, pp. 771–826, 2020. [Online]. Available: https://doi.org/10.1007/s00440-020-00990-x

[10] Z. D. Bai and Y. Q. Yin, “Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang,” *Probab. Theory Related Fields*, vol. 73, no. 4, pp. 555–569, 1986. [Online]. Available: https://doi.org/10.1007/BF00324852

[11] S. Geman, “The spectral radius of large random matrices,” *Ann. Probab.*, vol. 14, no. 4, pp. 1318–1328, 1986. [Online]. Available: http://links.jstor.org/sici?sici=0091-1798(198610)14:4<1318:TSROLR>2.0.CO;2-O&origin=MSN

[12] C. Bordenave, P. Caputo, D. Chafaï, and K. Tikhomirov, “On the spectral radius of a random matrix: an upper bound without fourth moment,” *Ann. Probab.*, vol. 46, no. 4, pp. 2268–2286, 2018. [Online]. Available: https://doi.org/10.1214/17-AOP1228

[13] Q. François and D. García-Zelada, “Asymptotic analysis of the characteristic polynomial for the Elliptic Ginibre ensemble,” *arXiv preprint arXiv:2306.16720*, 2023.

[14] S. Coste, “Sparse matrices: convergence of the characteristic polynomial seen from infinity,” *Electron. J. Probab.*, vol. 28, pp. Paper No. 8, 40, 2023. [Online]. Available: https://doi.org/10.1214/22-ejp875

[15] S. Coste, G. Lambert, and Y. Zhu, “The characteristic polynomial of sums of random permutations and regular digraphs,” *International Mathematics Research Notices*, vol. 2024, no. 3, pp. 2461–2510, 08 2023. [Online]. Available: https://doi.org/10.1093/imrn/rnad182

[16] B. Simon, *Trace ideals and their applications*, 2nd ed., ser. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2005, vol. 120. [Online]. Available: https://doi.org/10.1090/surv/120

[17] L. Avena, R. S. Hazra, and N. Malhotra, “Limiting spectra of inhomogeneous random graphs,” *arXiv preprint arXiv:2312.02805*, 2023.

[18] F. Benaych-Georges, C. Bordenave, and A. Knowles, “Largest eigenvalues of sparse inhomogeneous Erdős-Rényi graphs,” *Ann. Probab.*, vol. 47, no. 3, pp. 1653–1676, 2019. [Online]. Available: https://doi.org/10.1214/18-AOP1293

[19] A. Chakrabarty, S. Chakraborty, and R. Hazra, "Eigenvalues outside the bulk of inhomogeneous Erdős-Rényi random graphs," *J. Stat. Phys.*, vol. 181, no. 5, pp. 1746–1780, 2020. [Online]. Available: https://doi.org/10.1007/s10955-020-02644-7

[20] A. Chakrabarty, R. Hazra, F. den Hollander, and M. Sfragara, "Spectra of adjacency and Laplacian matrices of inhomogeneous Erdős-Rényi random graphs," *Random Matrices Theory Appl.*, vol. 10, no. 1, pp. Paper No. 2 150 009, 34, 2021. [Online]. Available: https://doi.org/10.1142/S201032632150009X

[21] F. Benaych-Georges, C. Bordenave, and A. Knowles, "Spectral radii of sparse random matrices," 2020.

[22] T. Shirai, "Limit theorems for random analytic functions and their zeros," in *Functions in number theory and their probabilistic aspects*, ser. RIMS Kôkyûroku Bessatsu. Res. Inst. Math. Sci. (RIMS), Kyoto, 2012, vol. B34, pp. 335–359.

[23] D. Kershaw, "Permanents in Fredholm's theory of integral equations," *J. Integral Equations*, vol. 1, no. 4, pp. 281–289, 1979.

[24] E. C. Titchmarsh, *The theory of functions*, 2nd ed. Oxford University Press, Oxford, 1939.

[25] H. Minc, *Permanents*, ser. Encyclopedia of Mathematics and its Applications. Addison-Wesley Publishing Co., Reading, MA, 1978, vol. 6, with a foreword by Marvin Marcus.

[26] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*. Cambridge university press, 2018, vol. 47.

[27] E. M. Wermuth, "Some elementary properties of infinite products," *The American mathematical monthly*, vol. 99, no. 6, pp. 530–537, 1992.