Eigenvalue Bounds for Random Matrices via Zerofreeness

Sidhanth Mohanty[*]   Amit Rajaraman[†]

September 29, 2025

Abstract

We introduce a new technique to prove bounds for the spectral radius of a random matrix, based on using Jensen's formula to establish the zerofreeness of the associated characteristic polynomial in a region of the complex plane. Our techniques are entirely non-asymptotic, and we instantiate it in three settings:

(i) The spectral radius of *non-asymptotic Girko matrices*—these are asymmetric matrices $M \in \mathbb{C}^{n \times n}$ whose entries are independent and satisfy $\mathbf{E}\, M_{ij} = 0$ and $\mathbf{E}\, |M_{ij}^2| \leqslant \frac{1}{n}$.

(ii) The spectral radius of *non-asymptotic Wigner matrices*—these are symmetric matrices $M \in \mathbb{C}^{n \times n}$ whose entries above the diagonal are independent and satisfy $\mathbf{E}\, M_{ij} = 0$, $\mathbf{E}\, |M_{ij}^2| \leqslant \frac{1}{n}$, and $\mathbf{E}\, |M_{ij}^4| \leqslant \frac{1}{n}$.

(iii) The second eigenvalue of the adjacency matrix of a *random d-regular graph* on n vertices, as drawn from the configuration model.

In all three settings, we obtain constant-probability eigenvalue bounds that are tight up to a constant. Applied to specific random matrix ensembles, we recover classic bounds for Wigner matrices, as well as results of Bordenave–Chafaï–García-Zelada, Bordenave–Lelarge–Massoulié, and Friedman, up to constants.

1 Introduction

Given a matrix M, we use $\mathrm{Spec}(M)$ to denote its spectrum, and $\rho(M)$ to denote its *spectral radius*, defined by

$$\rho(M) := \max_{\lambda \in \mathrm{Spec}(M)} |\lambda|.$$

A central theme in random matrix theory is to understand the spectrum of a random matrix, and notably its *spectral radius*. Over the decades, numerous tools have been developed to understand the spectra of random matrices: the trace moment method [FK81], the method of resolvents [EKYY13], matrix concentration inequalities [Tro15], the polynomial method [CGVTvH24], chaining [Tal], and more.

In this work, we introduce a new technique to control the spectral radius of a random matrix, based on Jensen's formula from complex analysis. Our technique is inspired by the recent work of Bencs, Liu, and Regts [BLR25], which uses Jensen's formula to study the zeros of partition functions of spin glasses.

Concretely, as we will see in Section 2, the following fact is a straightforward consequence of Jensen's formula.

Theorem 1.1. *For any matrix M and $r \in \mathbb{R}_{>0}$, we have:*

$$\prod_{\lambda \in \mathrm{Spec}(M):\, \lambda > \tau} \left(\frac{|\lambda|}{\tau} \right)^2 \leqslant \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} \left| \det\left(I - \frac{e^{i\theta}}{\tau} \cdot M \right) \right|^2. \tag{1}$$

This suggests a natural approach for controlling the spectral radius, and the number of outliers of a random matrix M drawn from a "nice" ensemble $\mathcal{D}$: explicitly compute and bound the value of $\mathbf{E}_{M \sim \mathcal{D}} |\det(I - zM)|^2$. Concretely, in the setting of a random matrix, we have the following corollaries of Theorem 1.1.

Corollary 1.2. $\mathop{\mathbf{E}}_{M \sim \mathcal{D}} \rho(M)^2 \leqslant \tau^2 \cdot \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} \mathop{\mathbf{E}}_{M \sim \mathcal{D}} \left| \det\left(\mathrm{Id} - \frac{e^{i\theta}}{\tau} \cdot M \right) \right|^2.$

Corollary 1.3. *Given a matrix M and positive numbers τ and δ, let k_M be the number of "outlier" eigenvalues $\lambda \in \mathrm{Spec}(M)$ such that $|\lambda| > \tau\sqrt{1+\delta}$. Then,*

$$\mathop{\mathbf{E}}_{M \sim \mathcal{D}} (1 + \delta)^{k_M} \leqslant \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} \mathop{\mathbf{E}}_{M \sim \mathcal{D}} \left| \det\left(\mathrm{Id} - \frac{e^{i\theta}}{\tau} \cdot M \right) \right|^2.$$

One may then use Markov's inequality on the above bounds to obtain bounds on the spectral radius (or number of outlier eigenvalues) that hold with constant probability.

As it turns out, in many cases, the combinatorial calculations involved in obtaining a bound on this quantity are simpler than other methods, especially in the sparse regime.

We discuss below the spectral radius bounds we prove in this paper using this method.

Random matrices with independent entries. Our first result concerns the spectral radius of (asymmetric) matrices with independent entries. We say that a random matrix M is a *non-asymptotic Girko matrix* if it has independent (but not necessarily identically distributed) complex-valued entries: for every $i \neq j \in [n]$, we have $\mathbf{E}\, M_{i,j} = 0$ and $\mathbf{E}\, |M_{i,j}|^2 = \frac{1}{n}$, and $M_{i,i} = 0$ for all $i \in [n]$.

Theorem 1.4. *For an $n \times n$ non-asymptotic Girko matrix M, we have:*

- $\mathbf{E}_M\, \rho(M)^2 \leqslant C$ *for an absolute constant $C > 1$.*

- *For every $\varepsilon > 0$, $\mathbf{E}_M |\{\lambda : \lambda \in \mathrm{Spec}(M), |\lambda| > 1 + \varepsilon\}| \leqslant C'(\varepsilon)$ for an absolute constant $C'(\varepsilon)$ that depends only on ε.*

Remark 1.5. In the setting of *Girko matrices*, where one first fixes a mean-0 and variance-1 random variable X, and considers the ensemble of $n \times n$ random matrices M with independent copies of $X/\sqrt{n}$ for $n \to \infty$, it was recently proved by Bordenave, Chafaï, and García-Zelada [BCGZ22] that the spectral radius of M is $1 + o_n(1)$ with probability $1 - o_n(1)$. Observe that this setting does *not* encompass random matrix ensembles where the distribution of the entries is allowed

to depend on n, as is the case in sparse directed random graph models. While we work in a more general setting where the entries of M may depend on n in an arbitrary way, our results are necessarily weaker on two fronts: obtaining only an $O(1)$ bound on the spectral radius, rather than $1 + o(1)$, and guarantees that hold with constant probability rather than with high probability. See Remark 3.1 for examples of random matrix ensembles that witness these limitations. We consider it an interesting question to investigate whether there are settings where these constant-probability bounds can be boosted to high-probability bounds. For example, in the setting where M is a draw from the (normalized) Gaussian orthogonal ensemble, which we shortly discuss, a reasonably straightforward argument based on concentration of Lipschitz functions succeeds at doing so.

Remark 1.6. We further point out that the bound on the number of outlier eigenvalues in Theorem 1.4(b) is *better* than the bound one would obtain if they show that the empirical spectral distribution weakly converges to some limiting law—indeed, this would only imply that $o(n)$ of the eigenvalues are outliers, not $O(1)$.

Hermitian random matrices with independent entries. Our story in the setting where M is a Hermitian random matrix needs more setup. In the sequel, in addition to first and second moment assumptions, we also make a very mild assumption on the fourth moment; nothing interesting is true without such an assumption. We say that an $n \times n$ random matrix M is a *non-asymptotic Wigner matrix* if M is Hermitian, its entries above the diagonal are independent, and satisfy $\mathbf{E}\, M_{i,j} = 0$, $\mathbf{E}\,\big|M_{i,j}\big|^2 = \frac{1}{n}$, and $\mathbf{E}\,\big|M_{i,j}\big|^4 \leqslant \frac{1}{n}$. We further assume that the diagonal of M is zero to avoid uninteresting complications.

We control the spectral radius of M via its *nonbacktracking matrix* B_M, whose rows and columns are indexed by the set of directed edges of the n-vertex complete graph, defined as follows:

$$B_M[ij, k\ell] = \begin{cases} M_{k,\ell} & \text{if } j = k \text{ and } i \neq \ell \\ 0 & \text{otherwise} \end{cases}$$

We prove the following for symmetric independent ensembles.

Theorem 1.7. *Let M be an $n \times n$ non-asymptotic Wigner matrix. Then,*

- $\mathbf{E}_M\, \rho(B_M)^2 \leqslant C$ *for an absolute constant $C > 1$.*

- *For any $\varepsilon > 0$, there is a constant $C'(\varepsilon) > 1$ such that with probability at least $1 - \varepsilon$,*

$$\rho(M) \leqslant C'(\varepsilon) \cdot \left(1 + \max_{i \in [n]} \|M_i\|_2\right).$$

Remark 1.8. One can check that when M is a "nice" random matrix ensemble, such as when its entries are independent standard gaussians, $\max_{i \in [n]} \|M_i\|_2$ concentrates extremely tightly around 1, thus recovering the known bound of 2 on the operator norm up to a constant factor.

More generally, our bound is tight up to constant factors since for any Hermitian matrix M, $\rho(M) \geqslant \max_{i \in [n]} \|M_i\|_2$.

Remark 1.9. The bound on $\mathbf{E}_M\, \rho(B_M)^2$ recovers the celebrated result of Bordenave, Lelarge, and Massoulié [BLM15] on the eigenvalues of nonbacktracking matrix of Erdős–Rényi graphs and stochastic block models of constant average degree up to absolute constant factors (cf. [FM17]).

Random regular graphs. Let G be a random d-regular graph on n vertices. The eigenvalues of the adjacency matrix A_G of G have been a subject of intense study in random matrix theory and the study of expander graphs. It was conjectured by Alon [Alo86] that the Alon–Boppana bound [Nil91] is tight for random d-regular graphs—besides the trivial eigenvalue of d corresponding to the all-ones vector, all eigenvalues of A_G are bounded in magnitude by $2\sqrt{d-1} + o_n(1)$. Friedman [Fri08] proved this conjecture two decades later in a highly technical tour de force using the trace moment method. Later, Bordenave [Bor19] gave a significantly simpler, though still far from easy, proof also based on the trace moment method. Notably, a year ago, Chen, Garza Vargas, Tropp and van Handel [CGVTvH24] gave a significantly simpler proof via a new method they pioneered based on connecting the spectra of random matrices to a series expansion in $1/n$ of trace moments of smooth functions of the adjacency matrix A_G. A parallel line of work [BHY19, HY24] culminating in a breakthrough of Huang, McKenzie, and Yau [HMY24] studied the *resolvent* to prove that all eigenvalues of A_G are at most $2\sqrt{d-1}$ (without the $o_n(1)$) with probability ≈ 0.69, along with other detailed information of the random matrix ensemble such as the distribution of the fluctuation of the largest eigenvalues, rigidity of the bulk eigenvalues, and delocalization of eigenvectors.

In this work, we use Theorem 1.1 to prove a constant probability version of Friedman's theorem with an eigenvalue bound that is tight up to a universal constant. While the result is quantitatively weaker than the results achieved by all of the aforementioned works, we believe our proof is simpler. In particular, rather loose arguments to control the expression on the right-hand side of Eq. (1) when M is the nonbacktracking matrix of a random regular graph lose only a constant factor in the spectral radius.

Theorem 1.10. *Let G be a random d-regular graph drawn from the configuration model. For every $\varepsilon > 0$, there is a constant $C(\varepsilon)$ such that with probability $1 - \varepsilon$, $\max\{\lambda_2(A_G), -\lambda_n(A_G)\}$ is at most $C(\varepsilon)\sqrt{d-1}$.*

Organization. In Section 2, we state and derive the basic complex analytic tools we will need. In Section 3, we will prove Theorem 1.4, our main result on the spectral radius of non-asymptotic Girko matrices. In Section 4, we describe the connection between the spectral radii of a matrix and its nonbacktracking matrix, and derive some properties of the determinant of a nonbacktracking matrix. Finally, in Section 5, we prove Theorem 1.7, our main result on non-asymptotic Wigner matrices, and in Section 6, we prove Theorem 1.10, our spectral radius bound on the random d-regular ensemble.

2 Complex analytic facts

The only tool we will need is Jensen's formula from complex analysis.

Theorem 2.1 (Jensen's formula [Ahl79, Eq. (44), Chapter 5.3.1]). *Let $f : \mathbb{C} \to \mathbb{C}$ be holomorphic, and let $a_1, \ldots, a_k$ be the zeros (with multiplicity) of f inside the scaled disk $r\mathbb{D}$, for some $r > 0$. Then,*

$$\mathop{\mathbf{E}}_{\theta \sim [0, 2\pi]} \log |f(re^{i\theta})| = \log |f(0)| + \sum_{t=1}^{k} \log\left(\frac{r}{|a_t|}\right).$$

In particular, using Jensen's inequality on the above easily yields the following.

Corollary 2.2. *Let $f : \mathbb{C} \to \mathbb{C}$ be holomorphic, and let $a_1, \ldots, a_k$ be the zeros (with multiplicity) of f inside the scaled disk $r\mathbb{D}$. Then,*

$$|f(0)|^2 \cdot \prod_{t=1}^{k} \left(\frac{r}{|a_t|} \right)^2 \leqslant \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} |f(re^{i\theta})|^2 .$$

Proof. By Jensen's inequality and Jensen's formula, we have:

$$\log \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} |f(re^{i\theta})|^2 \geqslant \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} \log |f(re^{i\theta})|^2 = 2 \cdot \left(\log |f(0)| + \sum_{t=1}^{k} \log \left(\frac{r}{|a_t|} \right) \right) .$$

We then obtain the desired statement by exponentiating the above inequality. □

Theorem 1.1, which we restate below for convenience, is a special case of Corollary 2.2 obtained by specializing $f(z) = \det(I - zM)$, and choosing $r = \frac{1}{\tau}$, since the roots of f are $\left\{ \frac{1}{\lambda} \right\}_{\lambda \in \mathrm{Spec}(M), \lambda \neq 0}$.

Theorem 1.1. *For any matrix M and $r \in \mathbb{R}_{>0}$, we have:*

$$\prod_{\lambda \in \mathrm{Spec}(M): \lambda > \tau} \left(\frac{|\lambda|}{\tau} \right)^2 \leqslant \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} \left| \det \left(I - \frac{e^{i\theta}}{\tau} \cdot M \right) \right|^2 . \tag{1}$$

Remark 2.3. The same inequality holds if we choose $f(z)$ as $\det(I - zM) \cdot g(z)$ for any holomorphic function g that is zero-free on $r\mathbb{D}$; in our case g is chosen as the constant function, but in some cases one may obtain mileage by a clever choice of g adapted to the ensemble at hand. Indeed, we believe that when studying symmetric matrix ensembles, passing to the nonbacktracking matrix implicitly involves choosing g as some suitable non-constant function.

3 Spectral radius bounds for non-asymptotic Girko matrices

In this section, we prove Theorem 1.4, restated below.

Theorem 1.4. *For an $n \times n$ non-asymptotic Girko matrix $\mathbf{M}$, we have:*

- $\mathbf{E}_\mathbf{M} \, \rho(\mathbf{M})^2 \leqslant C$ *for an absolute constant $C > 1$.*

- *For every $\varepsilon > 0$, $\mathbf{E}_\mathbf{M} \, |\{\lambda : \lambda \in \mathrm{Spec}(\mathbf{M}), |\lambda| > 1 + \varepsilon\}| \leqslant C'(\varepsilon)$ for an absolute constant $C'(\varepsilon)$ that depends only on ε.*

Proof of Theorem 1.4. By Corollaries 1.2 and 1.3, to prove both parts, it suffices to prove for any $\tau = 1 + \varepsilon$:

$$\mathop{\mathbf{E}}_{\mathbf{M}} \mathop{\mathbf{E}}_{\theta \sim [0,2\pi]} \left| \det \left(I - \frac{e^{i\theta}}{\tau} \cdot \mathbf{M} \right) \right|^2 \leqslant O_\varepsilon(1) .$$

Using the fact that the diagonal of $\mathbf{M}$ is equal to 0, we have for any z such that $|z| = \frac{1}{\tau}$:

$$\mathop{\mathbf{E}}_{\mathbf{M}} |\det(I - z \cdot \mathbf{M})|^2 = \mathop{\mathbf{E}}_{\mathbf{M}} \det (\mathrm{Id} - \mathbf{M}z) \cdot \det (\mathrm{Id} - \mathbf{M}^* z^*)$$

$$= \sum_{\pi, \sigma \in S_n} \mathop{\mathbf{E}}_{\mathbf{M}} \mathrm{sign}(\sigma) \mathrm{sign}(\pi) \cdot \left(\prod_{\substack{i \in [n] \\ \sigma(i) \neq i}} -\mathbf{M}_{i\sigma(i)} z \right) \cdot \left(\prod_{\substack{i \in [n] \\ \pi(i) \neq i}} -\mathbf{M}^*_{i\pi(i)} z^* \right) .$$

5

Observe that a term corresponding to (π, σ) in the above summation vanishes if there is some $i \in [n]$ such that $\sigma(i) \neq \pi(i)$. Therefore, denoting $\mathrm{NF}(\sigma) = \{i \in [n] : \sigma(i) \neq i\}$, we can continue the above chain of equalities as:

$$= \sum_{\pi \in S_n} \tau^{-2|\mathrm{NF}(\pi)|} \cdot \prod_{i \in \mathrm{NF}(\pi)} M_{i\pi(i)}^2$$

$$= \sum_{\pi \in S_n} \left(\tau^2 n\right)^{-\mathrm{NF}(\pi)}$$

$$= \sum_{k \geqslant 0} \left(\tau^2 n\right)^{-k} \cdot |\{\pi \in S_n : \mathrm{NF}(\pi) = k\}|$$

$$= \sum_{k \geqslant 0} \left(\tau^2 n\right)^{-k} \cdot \binom{n}{k} \cdot D_k ,$$

where D_k is the number of derangements on $[k]$, that is, the number of permutations $\sigma \in S_k$ such that $\sigma(i) \neq i$ for all i. Clearly, $D_k \leqslant k!$ for all k. Thus, the above can be bounded by

$$\leqslant \sum_{k \geqslant 0} \left(\tau^2 n\right)^{-k} \cdot \frac{n!}{(n-k)!}$$

$$\leqslant \sum_{k \geqslant 0} \tau^{-2k}$$

$$= \frac{\tau^2}{\tau^2 - 1} . \qquad \square$$

The above bound combined with Corollary 1.2 implies the desired bound on $\mathbf{E}_M \, \rho(M)^2$ by choosing τ as any constant larger than 1. The bound on $\mathbf{E}_M \, |\{\lambda : \lambda \in \mathrm{Spec}(M), |\lambda| > 1 + \varepsilon\}|$ follows from the above bound combined with Corollary 1.3 by choosing τ as $\sqrt{1 + \varepsilon}$, and $\delta = \varepsilon$.

Remark 3.1 (On tightness of bounds). Most past study on Girko matrices has been in the setting where the distributions of the entries of the matrix do not depend on n. More concretely, they assume some ensemble $(a_{ij})_{i,j \geqslant 1}$ of iid centered unit variance random variables, and set $A = \left(\frac{1}{\sqrt{n}} a_{ij}\right)_{1 \leqslant i, j \leqslant n}$. In particular, it was shown that the empirical spectral distribution of A weakly converges to the circular law $\mathrm{Unif}(\mathbb{D})$ [Gir85, Gir18, TVK10]. Later, in [BCGZ22], it was proved that in fact, for any $\varepsilon > 0$, $\mathbf{Pr}\left[|\rho(A) - 1| > \varepsilon\right] = o(1)$. The latter is significantly stronger than the weak "Markov-esque" tails we obtain in Theorem 1.4.

However, such statements are *false* in the generality we assume, where the a_{ij} are further allowed to depend on n. For example, consider the matrix M whose entries M_{ij} are independently distributed as

$$M_{ij} = \begin{cases} 0, & \text{w.p. } 1 - 2^{-n} , \\ \frac{1}{\sqrt{n}} \cdot 2^{n/2}, & \text{w.p. } 2^{-n-1} , \\ -\frac{1}{\sqrt{n}} \cdot 2^{n/2}, & \text{w.p. } 2^{-n-1} . \end{cases}$$

Clearly, the entries of M are centered and have variance $\frac{1}{n}$. However, with $1 - o(1)$ probability, $M = 0$, so its empirical spectral distribution in fact converges to δ_0, the Dirac measure on 0. At the level of generality we work in, it is also not true that for every $\varepsilon > 0$, $\mathbf{Pr}\left[\rho(A) > 1 + \varepsilon\right] = o(1)$.

For example, consider the appropriately scaled centered adjacency matrix of a draw G from the (directed) Erdős–Rényi graph ensemble, that is,

$$\frac{1}{\sqrt{2}} \cdot M_{ij} = \begin{cases} -\frac{1}{2n}, & \text{w.p. } 1 - \frac{1}{2n}, \\ 1 - \frac{1}{2n}, & \text{w.p. } \frac{1}{2n}. \end{cases}$$

The entries of this matrix have variance approximately $\frac{1}{n}$. It is easy to check that with $\Omega(1)$ probability, $\rho(M) \gtrsim \sqrt{2}$. Indeed, with $\Omega(1)$ probability, there is a 3-cycle abc in G, such that a, b, c further have no edges to any other vertices in G. In this case, the spectral radius of M is bounded from below by the spectral radius of the principal submatrix on $\{a, b, c\}$, which is given by

$$\sqrt{2} \begin{pmatrix} 0 & 1 - \frac{1}{2n} & -\frac{1}{2n} \\ -\frac{1}{2n} & 0 & 1 - \frac{1}{2n} \\ 1 - \frac{1}{2n} & -\frac{1}{2n} & 0 \end{pmatrix} .$$

However, for sufficiently large n, the spectral radius of this matrix is $\approx \sqrt{2} > 1$, so the spectral radius of M is also bounded away from 1 with constant probability.

4   Hermitian matrices and the nonbacktracking matrix

In this section, we review the well-understood connection between the spectral radius of a Hermitian matrix and its nonbacktracking matrix, and also establish some properties of nonbacktracking matrices that we use in our proof.

In particular, we use the following fact, which may be proved using the remarkable Ihara–Bass formula [Iha66, Bas92].

Lemma 4.1 (Consequence of [BGBK20, Theorem 2.2]). *Let M be an $n \times n$ Hermitian matrix, and let B_M be its nonbacktracking matrix. Then:*

$$\rho(M) \leqslant 2\rho(B_M) + 9 \cdot \max_i \|M_i\|_2 .$$

Remark 4.2. The translation between the spectral radius and that of its nonbacktracking matrix applies to any matrix and hence is crude, and lossy by a constant factor. For many "nice" random matrix ensembles, there is essentially no loss incurred in passing to the nonbacktracking matrix.

Remark 4.3. While the above statement only translates bounds on the spectral radius of the nonbacktracking matrix to bounds on the spectral radius of M, the Ihara–Bass formula may similarly be used to translate bounds on the number of outlier eigenvalues (of the form in Theorem 1.4(b)). We omit the details; such bounds immediately follow from all the computations we perform here, when used in conjunction with Corollary 1.3.

For a random matrix M in the wild, it is often easy to control $\max_i \|M_i\|_2$, and the challenging part is in getting a handle on $\rho(B_M)$. Today, we will use Corollary 1.2 to control $\rho(B_M)$ by studying $\det(I - B_M z)$ where M is either a non-asymptotic Wigner matrix, or the adjacency matrix of a random regular graph.

We use $\vec{E}$ to denote the set of directed edges on the n-vertex complete graph, and $S_{\vec{E}}$ to denote the set of all permutations on $\vec{E}$. We will use $\mathrm{NBP}_{\vec{E}} \subseteq S_{\vec{E}}$ to refer to the set of all *nonbacktracking permutations*, i.e., permutations $\pi \in S_{\vec{E}}$ such that for any $ij \in \vec{E}$, $\pi(ij)$ is either equal to ij, or is equal to jk for some $k \neq i$. For a permutation π, we use $\mathrm{ntcyc}(\pi)$ to denote the number of nontrivial cycles in the permutation π, i.e., cycles of length at least 2, and we use $\mathrm{cyc}(\pi)$ to denote the number of cycles in π (including trivial ones).

We have:

$$\det (\mathrm{Id} - B_M z) = \sum_{\pi \in S_{\vec{E}}} \mathrm{sign}(\pi) \prod_{e \in \mathrm{NF}(\pi)} (-B_{e\pi(e)} z)$$

$$= \sum_{\pi \in \mathrm{NBP}_{\vec{E}}} \mathrm{sign}(\pi) \cdot \prod_{e \in \mathrm{NF}(\pi)} (-B_{e\pi(e)} z)$$

$$= \sum_{\pi \in \mathrm{NBP}_{\vec{E}}} (-1)^{\mathrm{ntcyc}(\pi)} \cdot z^{|\mathrm{NF}(\pi)|} \cdot \prod_{e \in \mathrm{NF}(\pi)} M_{\pi(e)} .$$

Using $\mathcal{H}$ to denote the set of all directed subgraphs on $[n]$ (that is, subsets of $\vec{E}$), and for $H \in \mathcal{H}$, using NBP_H to denote the set of all $\pi \in \mathrm{NBP}_{\vec{E}}$ such that $\mathrm{NF}(\pi) = H$, we may write the above as:

$$\det(\mathrm{Id} - B_M z) = \sum_{H \in \mathcal{H}} z^{e(H)} \cdot \prod_{e \in H} M_e \cdot \sum_{\pi \in \mathrm{NBP}(H)} (-1)^{\mathrm{ntcyc}(\pi)} \tag{2}$$

We first substantially simplify the sum over elements in $\mathrm{NBP}(H)$. Towards doing so, let us introduce some terminology.

Definition 4.4. For $H \in \mathcal{H}$, we say that $e \in H$ is a *doubleton* if the reverse of e also occurs in H, and we say e is a *singleton* otherwise. For a vertex $v \in [n]$, we use $\mathrm{In}_H(v)$ to denote the set of all incoming edges to v in H, and $\mathrm{Out}_H(v)$ to denote the set of all outgoing edges. Let $\widetilde{\mathcal{H}} \subseteq \mathcal{H}$ be the set of all $H \in \mathcal{H}$ such that every vertex has at most one incoming singleton edge and at most one outgoing singleton edge.

A simple yet important observation is the following.

Observation 4.5. Let $\pi \in \mathrm{NBP}(H)$. By the fact that π is a nonbacktracking permutation, for every vertex v in $[n]$, $d_v^H := |\mathrm{In}_H(v)| = |\mathrm{Out}_H(v)|$, and $\pi(\mathrm{In}_H(v)) = \mathrm{Out}_H(v)$.

We prove the following simplification for the determinant.

Lemma 4.6. *For any matrix M, we have*

$$\det(\mathrm{Id} - B_M z) = \sum_{H \in \widetilde{\mathcal{H}}} z^{e(H)} \cdot \prod_{e \in H} M_e \cdot \sum_{\pi \in T(H)} (-1)^{\mathrm{ntcyc}(\pi)}$$

for some $T(H) \subseteq \mathrm{NBP}(H)$ with $|T(H)| \leq \prod_{v \in [n]} d_v^H$.

Proof. For $\pi \in \mathrm{NBP}(H)$ and $v \in [n]$, we use $\pi_v : \mathrm{In}_H(v) \to \mathrm{Out}_H(v)$ to refer to the restriction of π to the edges incident to v. We first observe that π is uniquely determined by a collection of local bijections $(\pi_v)_{v \in [n]}$. In particular, π is forced to be the permutation where $\pi(uv) = \pi_v(uv)$. We also observe that if each local permutation obeys the nonbacktracking constraint, π is in $\mathrm{NBP}(H)$. We now set up some notation.

8

- Let Ω_v be the collection of all nonbacktracking local bijections from $\mathrm{In}_H(v)$ and $\mathrm{Out}_H(v)$.

- Let $\Omega := \prod_{v \in [n]} \Omega_v$, and let $\Omega_{-v} = \prod_{u \in [n] \setminus v} \Omega_u$.

- Let π_{-v} denote $(\pi_u)_{u \in [n] \setminus v}$.

- Let $\mathrm{ntcyc}_v(\pi)$ refer to the number of nontrivial cycles in π that pass through v, and let $\mathrm{ntcyc}_{-v}(\pi)$ refer to the number of cycles in π that do not pass through v.

For any $v \in [n]$, we can write:

$$\sum_{\pi \in \mathrm{NBP}(H)} (-1)^{\mathrm{ntcyc}(\pi)} = \sum_{\pi_1, \ldots, \pi_n \in \Omega} (-1)^{\mathrm{ntcyc}(\pi)}$$

$$= \sum_{\pi_{-v} \in \Omega_{-v}} \sum_{\pi_v \in \Omega_v} (-1)^{\mathrm{ntcyc}_{-v}(\pi) + \mathrm{ntcyc}_v(\pi)}$$

$$= \sum_{\pi_{-v} \in \Omega_{-v}} (-1)^{\mathrm{ntcyc}_{-v}(\pi_{-v})} \sum_{\pi_v \in \Omega_v} (-1)^{\mathrm{ntcyc}_v(\pi)},$$

where in the last equality uses the fact that $\mathrm{ntcyc}_{-v}(\pi)$ does not depend on π_v. Observe that one can extract from π_{-v} a bijection $\sigma : \mathrm{Out}_H(v) \to \mathrm{In}_H(v)$ where $\sigma(vw)$ is equal to $w'v$ obtained by following the path starting at vw described by π_{-v} until it hits v next. We may treat σ and π_v as permutations on $[d_v^\pi]$ by arbitrarily labeling the elements of $\mathrm{Out}_H(v)$ and $\mathrm{In}_H(v)$ with elements of $[d_v^\pi]$. Next, observe that $\mathrm{ntcyc}_v(\pi) = \mathrm{cyc}(\sigma \circ \pi_v)$, and so,

$$(-1)^{\mathrm{ntcyc}_v(\pi)} = (-1)^{\mathrm{cyc}(\sigma \circ \pi_v)} = (-1)^{d_v^\pi} \cdot \mathrm{sign}(\sigma \circ \pi_v) = (-1)^{d_v^\pi} \cdot \mathrm{sign}(\sigma) \cdot \mathrm{sign}(\pi_v).$$

Since σ depends only on π_{-v}, we get:

$$\sum_{\pi \in \mathrm{NBP}(H)} (-1)^{\mathrm{ntcyc}(\pi)} = \sum_{\pi_{-v} \in \Omega_{-v}} (-1)^{\mathrm{ntcyc}_{-v}(\pi_{-v})} \cdot \mathrm{sign}(\sigma) \cdot (-1)^{d_v^\pi} \cdot \sum_{\pi_v \in \Omega_v} \mathrm{sign}(\pi_v). \qquad (3)$$

The innermost sum is taken over all permutations subject to the constraint that π_v cannot map an in-edge to the out-edge that reverses it. Thus, for the $d_v^\pi \times d_v^\pi$ matrix R where R_{ij} is 0 if the i-th in-edge is the reversal of the j-th outedge, and place 1 otherwise, we have that

$$\sum_{\pi_v \in \Omega_v} \mathrm{sign}(\pi_v) = \det(R).$$

Observe that R is a $\{0, 1\}$-matrix where every row and column has at most a single 0, and so we can understand its determinant based on the following cases:

- *Number of rows that are all-ones is at least* 2. In this case, $\det(R) = 0$ since the matrix is singular by virtue of having repeated rows.

- *Number of rows that are all-ones is exactly* 1. In this case, $\det(R) \in \{\pm \det(\mathbb{1}\mathbb{1}^\top - \mathrm{Id} + E_{11})\}$, which is equal to $\{\pm 1\}$ via explicit calculation of eigenvalues.

- *Every row/column has a* 0. In this case, $\det(R) \in \{\pm \det(\mathbb{1}\mathbb{1}^\top - \mathrm{Id})\}$, which is equal to $\{\pm(d_v^\pi - 1)\}$ via explicit calculation of eigenvalues.

9

The above casework has a few upshots. First, if H is not in $\widetilde{\mathcal{H}}$, then we choose v to be a vertex with either more than one incoming singleton or outgoing singleton edge. For such a vertex, we must have $\sum_{\pi_v \in \Omega_v} \text{sign}(\pi_v) = 0$, and so the entire term corresponding to H in Eq. (2) vanishes. Thus, we may assume H is in $\widetilde{\mathcal{H}}$. Since $\left| \sum_{\pi_v \in \Omega_v} \text{sign}(\pi_v) \right| \leqslant d_v^\pi$, and each term in the summand is a sign, we may fix a subset $T_v \subseteq \Omega_v$ of at most d_v^π elements such that $\sum_{\pi_v \in T_v} \text{sign}(\pi_v) = \sum_{\pi_v \in \Omega_v} \text{sign}(\pi_v)$. As a consequence, we have:

$$\sum_{\pi \in \text{NBP}(H)} (-1)^{\text{ntcyc}(\pi)} = \sum_{\pi_{-v} \in \Omega_{-v}} \sum_{\pi_v \in T_v} (-1)^{\text{ntcyc}(\pi)}.$$

We may iteratively replace each Ω_v with T_v and obtain:

$$\sum_{\pi \in \text{NBP}(H)} (-1)^{\text{ntcyc}(\pi)} = \sum_{\pi_1, \dots, \pi_n \in T_1 \times \cdots \times T_n} (-1)^{\text{ntcyc}(\pi)}.$$

The desired statement follows from choosing $T(H) = T_1 \times \cdots \times T_n$, and our bound on all $|T_v|$. □

5 Spectral radius bound for the nonbacktracking matrix

In this section, we prove that for a random matrix M satisfying fairly reasonable bounds on mixed moments in its entries, we can obtain a spectral radius bound on B_M via Corollary 1.2. Our bounds on the spectral radius of non-asymptotic Wigner matrices and random regular graphs will then immediately follow from verifying that they satisfy the requisite condition on their mixed moments.

For the sequel we will need the following notation.

Definition 5.1 (Notation for undirecting a directed graph)**.** Given a directed graph H, we use $\text{UD}(H)$ to denote the undirected (multi-)graph obtained by taking each directed edge uv in H, and replacing it with an undirected edge between u and v. We use $\text{SUD}(H)$ to denote the *simple* undirected graph obtained by taking $\text{UD}(H)$ and replacing each multi-edge with a single edge.

Assumption 5.2. For every directed subgraph $H \in \widetilde{\mathcal{H}}$ (for $\widetilde{\mathcal{H}}$ defined in Definition 4.4), and every assignment of "multiplicities" $m : E(\text{SUD}(H)) \to \{1, 2, 3, 4\}$ to the edges of $\text{SUD}(H)$, we have:

$$\mathop{\mathbf{E}}_{M} \prod_{ij \in E(\text{SUD}(H))} M_{ij}^{m(ij)} \leqslant \left(\frac{C}{n} \right)^{|E(H)|}$$

for an absolute constant $C > 1$ independent of n.

Theorem 5.3. *There exists an absolute constant $\alpha > 0$ such that for any $n \times n$ random matrix M satisfying* Assumption 5.2*, we have:*

$$\mathop{\mathbf{E}}_{M} \rho(M)^2 \leqslant \alpha.$$

Remark 5.4. Theorem 1.7 follows immediately from Theorem 5.3 since a non-asymptotic Wigner matrix can be readily verified to satisfy Assumption 5.2 using independence. Verifying Assumption 5.2 for the random d-regular ensemble is a more challenging task, and we do so in Section 6.

We now prove Theorem 5.3 below.

Proof of Theorem 5.3. By Corollary 1.2, it suffices to prove an upper bound on

$$\mathop{\mathbf{E}}_{M} \left| \det(\mathrm{Id} - z \cdot B_M) \right|^2$$

for $|z| = 1/\tau$ for some sufficiently large constant τ. We apply Lemma 4.6 and obtain:

$$\mathop{\mathbf{E}}_{M} \left| \det(\mathrm{Id} - z \cdot B_M) \right|^2 = \mathop{\mathbf{E}}_{M} \det(\mathrm{Id} - z \cdot B_M) \cdot \det(\mathrm{Id} - z^* \cdot B_{M^*})$$

$$\leqslant \sum_{H,H'\in\widetilde{\mathcal{H}}} \tau^{-e(H)-e(H')} \cdot \left| \mathop{\mathbf{E}}_{M} \prod_{e\in\vec{E}(H)} M_e \cdot \prod_{e\in\vec{E}(H')} M_e \right| \cdot \left| \sum_{\substack{\pi\in T(H) \\ \pi'\in T(H')}} (-1)^{\mathrm{ntcyc}(\pi)+\mathrm{ntcyc}(\pi')} \right|$$

$$\leqslant \sum_{H,H'\in\widetilde{\mathcal{H}}} \left(\frac{C}{\tau n} \right)^{e(\mathrm{SUD}(H\cup H'))} \cdot |T(H)| \cdot |T(H')|$$

where the third inequality used Assumption 5.2. We may continue the above chain of inequalities below.

$$\leqslant \sum_{k\geqslant 0} \sum_{\substack{H,H'\in\widetilde{\mathcal{H}} \\ e(\mathrm{SUD}(H,H'))=k}} \left(\frac{C}{\tau n} \right)^k \cdot |T(H)| \cdot |T(H')| \qquad (\star)$$

First, observe that for any $H, H' \in \widetilde{\mathcal{H}}$, all non-isolated vertices of the graph $R := \mathrm{SUD}(H, H')$ have degree at least 2, and thus $e(R) \geqslant v(R)$. We relax the above sum to enumerate over *all* k-edge graphs R where all non-isolated vertices of R have degree at least 2, and over *all* pairs of directed graphs H, H' such that $\mathrm{SUD}(H \cup H') = R$. For every edge $\{u, v\} \in E(R)$, given a subset of the statements $\{uv \in H, vu \in H, uv \in H', vu \in H'\}$, it is possible to recover H and H'. Thus, given R, there are at most 16^k choices for H and H'. Next, we use the bound on $|T(H)|$ and $|T(H')|$ from Lemma 4.6 to observe that:

$$|T(H)| \cdot |T(H')| \leqslant \prod_{v\in V(R)} \deg_R(v)^2 \leqslant \left(\frac{\sum_v \deg_R(v)}{r} \right)^{2r} = \left(\frac{2k}{r} \right)^{2r} \leqslant O(1)^k,$$

since $(2k/r)^{r/2k}$ is uniformly bounded. Consequently, using $\mathcal{R}(r, k)$ to denote the set of all k-edge subgraphs of the complete graph with r nonisolated vertices, we may write

$$(\star) \leqslant \sum_{r\geqslant 0} \sum_{k\geqslant r} \sum_{R\in\mathcal{R}(r,k)} \left(\frac{O(1)}{\tau n} \right)^k$$

$$\leqslant \sum_{r\geqslant 0} \sum_{k\geqslant r} |\mathcal{R}(r,k)| \left(\frac{O(1)}{\tau n} \right)^k$$

$$\leqslant \sum_{k\geqslant 0} \sum_{0\leqslant r\leqslant\min\{k,n\}} \binom{n}{r} \binom{r^2/2}{k} \left(\frac{O(1)}{\tau n} \right)^k$$

$$\leqslant \sum_{k\geqslant 0} \tau^{-k} \cdot O(1)^k \cdot \sum_{0\leqslant r\leqslant\min\{k,n\}} \left(\frac{en}{r} \right)^r \cdot \left(\frac{er^2}{2k} \right)^k \cdot \frac{1}{n^k}$$

$$\leqslant \sum_{k \geqslant 0} \tau^{-k} \cdot O(1)^k \cdot \sum_{0 \leqslant r \leqslant \min\{k,n\}} \left(\frac{r}{n}\right)^{k-r}$$

$$\leqslant \sum_{k \geqslant 0} \tau^{-k} \cdot O(1)^k,$$

where the third inequality used that one can construct an r-vertex k-edge graph by picking r vertices and then choosing any subset of k edges, and the second last inequality used $r \leqslant k$. Choosing τ as a sufficiently large constant so as to induce geometric decay in the sum completes the proof. $\square$

6 On mixed moments of entries of random regular graphs

In this section, we shall prove Assumption 5.2 in the setting where M is the normalized and centered adjacency matrix of a random d-regular graph. Since we are okay losing constant factors in the spectral radius bound, we may assume that d is some sufficiently large constant. The specific model of random d-regular graphs we study is the configuration model.

Definition 6.1 (Configuration model). Given $n \geqslant 1$ and $d \geqslant 2$, a draw G from the corresponding configuration model is defined as follows. Let $\mathcal{G}$ be a uniformly random perfect matching on $[n] \times [d]$, and let G be the multigraph on $[n]$ obtained by adding an edge ij for each edge in $\mathcal{G}$ between the corresponding "clouds" $\{i\} \times [d]$ and $\{j\} \times [d]$. Let A be the adjacency matrix of G, and M the random matrix defined by

$$M_{ij} = \begin{cases} \frac{1}{\sqrt{d}} \cdot \left(A_{ij} - \frac{d}{n}\right), & i \neq j, \\ 0, & i = j. \end{cases}$$

Theorem 6.2. *Assumption 5.2 holds for M defined as above.*

Proof of Theorem 1.10. Theorems 5.3 and 6.2 yield a constant-probability bound on the spectral radius of the nonbacktracking matrix B_M associated to M of an absolute constant C. It remains to prove that the maximum row norm of M is bounded by $O(1)$ with high probability; this would translate back to a bound of $O(1)$ on the largest eigenvalue of M via Lemma 4.1.

First, observe that $\|M_i\| \leqslant \frac{\|(A_G)_i\|}{\sqrt{d}}$, and so it suffices to prove $\|(A_G)_i\| \leqslant O(\sqrt{d})$ with probability at least $1 - o(1/n)$. We split into cases: when $d \leqslant \sqrt{n}$, the bound of the maximum row norm follows from the fact that A_G has at most d nonzero entries per row, and that all entries of A_G are at most 5 with probability $1 - o(1)$.

We now treat the case when $n \geqslant d > \sqrt{n}$. The random variable x describing a row of A_G can be modeled as the histogram resulting from tossing d balls into n bins. We can couple x with a random variable y obtained by tossing Poisson($2d$) balls into n bins (or equivalently a vector of n independent draws from Poisson($2d/n$)) such that $y \geqslant x$ with probability $1 - o(1/n)$. One may prove a bound on $\|y\|$ of $O(\sqrt{d})$ that holds with probability $1 - o(1/n)$ by proving $\mathbf{E}(\|y\|^2 - \mathbf{E}\|y\|^2)^k \leqslant k! \cdot d^{k/2}$; we omit the details. $\square$

Key to the proof will be establishing concrete wins that can be gained when many of the multiplicities $m(ij)$ are equal to 1—indeed, in the Wigner setting from the previous section, any monomial with a variable that appears with degree-1 (henceforth referred to as a *singleton* edges) is zero in expectation.

Our proof relies on estimates for various centered (and uncentered) expectations in the configuration model, starting with a relatively trivial bound that does *not* see any of the extra wins from singleton edges. In the next few lemmas, one must think of $N = nd$.

Lemma 6.3. *Let S be an arbitrary subgraph on vertex set* $[N]$. *Let* $m_e \geqslant 1$ *be some constant independent of* N *for each* $e \in S$. *Then, for* $\mathcal{G}_{[N]}$ *a uniformly random perfect matching on* $[N]$, *we have*

$$\left| \mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}_{[N]}} - \frac{1}{N} \right)^{m_e} \right| \leqslant O\left(\frac{1}{N}\right)^{|S|} .$$

Proof. We may simply expand out the product to write

$$\left| \mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}_{[N]}} - \frac{1}{N} \right)^{m_e} \right| \leqslant \sum_{R \subseteq S} \left(\frac{1}{N}\right)^{|S \backslash R|} \cdot \mathbf{E} \, \mathbf{1}_{R \subseteq \mathcal{G}_{[N]}}$$

$$\leqslant \sum_{R \subseteq S} \left(\frac{1}{N}\right)^{|S \backslash R|} \cdot \frac{(N - 2|R| - 1)!!}{(N-1)!!} = O\left(\frac{1}{N}\right)^{|S|} .$$

$\square$

Now, on the other end of the spectrum, let us demonstrate the wins that appear when every edge is a singleton in the configuration model. To do so, we shall use the Laplace method.

Lemma 6.4 (Laplace's Method, [But07, Chapter 2])**.** *Let* $f, g : \mathbb{R} \to \mathbb{R}$ *be such that g has a unique global maximizer* $x^\star$ *with g analytic in a neighborhood of* $x^\star$, *with* $f(x^\star) \neq 0$ *and* $g''(x^\star) \neq 0$. *Then,*

$$\int_{-\infty}^{\infty} f(x) e^{Ng(x)} \, \mathrm{d}x = f(x^\star) e^{Ng(x^\star)} \sqrt{\frac{2\pi}{Ng''(x^\star)}} \left(1 + O\left(\frac{1}{N}\right)\right) .$$

Here, the $O\left(\frac{1}{N}\right)$ *hides factors depending on the third and fourth derivatives of f and g at* $x^\star$.

Lemma 6.5. *Let S be a matching with k edges on vertex set* $[N]$, *and* $\beta \geqslant 1$. *For simplicity, assume that* $k < \frac{N}{4}$. *If* $\mathcal{G}$ *is a uniformly random perfect matching on* $[N]$, *we have*

$$\left| \mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}} - \frac{1}{\beta N} \right) \right| \leqslant O\left(\frac{1}{N}\right)^k \cdot \left(\frac{\beta - 1}{\beta} + \sqrt{\frac{2k}{\beta N}} \right)^k .$$

Proof. We start by expanding

$$\mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}} - \frac{1}{\beta N} \right) = \sum_{T \subseteq S} \mathbf{E} \, \mathbf{1}_{T \subseteq \mathcal{G}} \cdot \left(-\frac{1}{\beta N} \right)^{k - |T|}$$

By the definition of $\mathcal{G}$, we have

$$\mathbf{Pr}\left[T \subseteq \mathcal{G}\right] = \frac{(N - 2|T| - 1)!!}{(N-1)!!} .$$

Thus, the above expectation may be written as

$$\mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}} - \frac{1}{N} \right) = \sum_{T \subseteq S} \mathbf{Pr}\left[T \subseteq \mathcal{G}\right] \cdot \left(-\frac{1}{\beta N} \right)^{k - |T|}$$

$$= \sum_{T \subseteq S} \frac{(N - 2|T| - 1)!!}{(N-1)!!} \cdot \left(-\frac{1}{\beta N}\right)^{k-|T|}$$

$$= \sum_{0 \leqslant r \leqslant k} \binom{k}{r} \cdot \frac{(N - 2|T| - 1)!!}{(N-1)!!} \cdot \left(-\frac{1}{\beta N}\right)^{k-|T|}$$

$$= \frac{1}{(N-1)!!} \sum_{0 \leqslant r \leqslant k} \binom{k}{r} \cdot (N - 2r - 1)!! \cdot \left(-\frac{1}{\beta N}\right)^{k-r}$$

$$= \frac{1}{(N-1)!!} \mathop{\mathbf{E}}_{g \sim \mathcal{N}(0,1)} \sum_{0 \leqslant r \leqslant k} \binom{k}{r} \cdot g^{N-2r} \cdot \left(-\frac{1}{\beta N}\right)^{k-r}$$

$$= \frac{1}{(N-1)!!} \mathop{\mathbf{E}}_{g \sim \mathcal{N}(0,1)} g^{N-2k} \left(1 - \frac{g^2}{\beta N}\right)^k$$

$$= \frac{1}{(N-1)!!} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} x^{N-2k} \left(1 - \frac{x^2}{\beta N}\right)^k \mathrm{d}x$$

$$= \frac{N^{\frac{N+1}{2}-k}}{2\sqrt{2\pi}(N-1)!!} \int_0^{\infty} e^{-Nt/2} t^{\frac{N-1}{2}-k} \left(1 - \frac{t}{\beta}\right)^k \mathrm{d}t , \tag{4}$$

where in the final step we substituted $x = \sqrt{Nt}$. To deal with the integral expression (4), we will use Laplace's method Lemma 6.4. Set $\alpha = \frac{2k}{N} \in [0, \frac{1}{2}]$, and consider the integral

$$\int_0^{\infty} e^{N\left(-\frac{t}{2} + \left(\frac{1-\alpha}{2}\right) \log t + \frac{\alpha}{2} \log\left|1 - \frac{t}{\beta}\right|\right)} , \tag{5}$$

obtained by ignoring a factor of t in the integral of interest. It turns out that (4) $\leqslant O(1) \cdot$ (5), and so it suffices to bound (5). Indeed, this will follow from the fact that the maximizer $t^\star$ of the function of t in the exponent is achieved at a value that is $\Omega(1)$ by using an appropriate version of Lemma 6.4 where we bring some factors of $t^\star$ out as another function f, and allowing g to depend (very slightly) on N—we omit the details, and refer the curious reader to the usual proof of the legitimacy of Laplace's method from, e.g., [But07].

Let us find the maximizers of the exponent of this function, given by $g(t) = -\frac{t}{2} + \left(\frac{1-\alpha}{2}\right) \log t + \frac{\alpha}{2} \log\left|1 - \frac{t}{\beta}\right|$. Setting the derivative to 0 for $t < \beta$ gives

$$-\frac{1}{2} + \left(\frac{1-\alpha}{2}\right) \cdot \frac{1}{t} - \frac{\alpha}{2} \cdot \frac{1}{\beta - t} = 0$$

$$(1 - \alpha) \cdot \frac{1}{t} - \alpha \cdot \frac{1}{\beta - t} = 1$$

$$(1 - \alpha) \cdot (\beta - t) - \alpha t = \beta t - t^2$$

$$t^2 - (1 + \beta)t + \beta(1 - \alpha) = 0 .$$

This yields the local maximizer $t^\star = \frac{\beta+1}{2} - \sqrt{\frac{(\beta+1)^2}{4} - \beta(1 - \alpha)}$. Observe that because $\alpha < \frac{1}{2}$, $t^\star = \Omega(1)$. Setting the derivative to 0 for $t > \beta$ gives

$$-\frac{1}{2} + \left(\frac{1-\alpha}{2}\right) \cdot \frac{1}{t} + \frac{\alpha}{2} \cdot \frac{1}{\beta - t} = 0$$

14

$$(1-\alpha) \cdot \frac{1}{t} + \alpha \cdot \frac{1}{\beta - t} = 1$$

$$(1-\alpha) \cdot (\beta - t) + \alpha t = \beta t - t^2$$

$$t^2 + (\beta + 1 - 2\alpha)t + \beta(1-\alpha) = 0.$$

It is not difficult to see that this has no positive roots. Therefore, the original function is decreasing on (β, ∞), and $t^\star$ is in fact the global maximizer. Therefore, Laplace's method Lemma 6.4 yields that

$$\int_0^\infty e^{Ng(t)} \, \mathrm{d}t = e^{Ng(t^\star)} \cdot \sqrt{\frac{2\pi}{-Ng''(t^\star)}} \left(1 + O\left(\frac{1}{N}\right)\right).$$

Indeed, the third and fourth derivatives of g at $t^\star$ are bounded due to the upper bound on α. As a result,

$$\mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}} - \frac{1}{\beta N}\right) \lesssim \frac{N^{\frac{N}{2}-k}}{(N-1)!!} \cdot \frac{e^{Ng(t^\star)}}{\sqrt{-g''(t^\star)}}$$

$$= \frac{N^{\frac{N}{2}-k}}{(N-1)!!} \cdot \frac{e^{-Nt^\star/2}}{\sqrt{-g''(t^\star)}} \cdot (t^\star)^{N(1-\alpha)/2} \cdot \left(1 - \frac{t^\star}{\beta}\right)^{N\alpha/2}.$$

We have

$$1 - \frac{t^\star}{\beta} = \frac{1}{\beta}\left(\frac{\beta - 1}{2} + \sqrt{\frac{(\beta-1)^2}{4} + \alpha\beta}\right)$$

$$\leqslant \frac{1}{\beta}\left(\beta - 1 + \sqrt{\alpha\beta}\right) = \frac{\beta - 1}{\beta} + \sqrt{\frac{\alpha}{\beta}}.$$

Using Eq. (4) $\leqslant O(1) \cdot$ Eq. (5) and plugging this into the above gives

$$\mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}} - \frac{1}{\beta N}\right) \lesssim \frac{N^{\frac{N}{2}-k}}{(N-1)!!} \cdot \frac{e^{-Nt^\star/2}}{\sqrt{-g''(t^\star)}} \cdot (t^\star)^{N(1-\alpha)/2} \cdot \left(\frac{\beta - 1}{\beta} + \sqrt{\frac{2k}{\beta N}}\right)^k.$$

Stirling's approximation allows us to write

$$\mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}} - \frac{1}{\beta N}\right) \lesssim \frac{(N - 2k - 1)!!}{(N-1)!!} \cdot e^{\frac{N(1-\alpha)}{2}} \cdot \frac{e^{-Nt^\star/2}}{\sqrt{-g''(t^\star)}} \cdot \left(\frac{t^\star}{1-\alpha}\right)^{N(1-\alpha)/2} \cdot \left(\frac{\beta - 1}{\beta} + \sqrt{\frac{2k}{\beta N}}\right)^k.$$

We now observe that

$$-t^\star + (1-\alpha)\left(1 + \ln \frac{t^\star}{1-\alpha}\right) \leqslant 0.$$

Indeed, $1 + \ln \frac{t^\star}{1-\alpha} \leqslant \frac{t^\star}{1-\alpha}$. As a result,

$$\mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}} - \frac{1}{\beta N}\right) \lesssim \frac{(N - 2k - 1)!!}{(N-1)!!} \cdot \frac{1}{\sqrt{-g''(t^\star)}} \cdot \left(\frac{\beta - 1}{\beta} + \sqrt{\frac{2k}{\beta N}}\right)^k.$$

Finally, let us deal with the second derivative term. As observed already, we have $\beta - t^\star \leqslant \beta - 1 + \sqrt{\alpha\beta}$. We have $t^\star \leqslant \frac{\beta+1}{2} \cdot \frac{4\beta(1-\alpha)}{(\beta+1)^2} \lesssim 1 - \alpha$. Consequently,

$$-2g''(t^\star) = \frac{1-\alpha}{(t^\star)^2} + \frac{\alpha}{(\beta - t^\star)^2} \geqslant \frac{1-\alpha}{(t^\star)^2} \gtrsim 1 \,.$$

To conclude, we use Stirling's approximation to bound

$$\frac{(N - 2k - 1)!!}{(N-1)!!} \lesssim O(1)^k \cdot \frac{(N - 2k)^{(N-2k)/2}}{N^{N/2}}$$

$$\leqslant O\left(\frac{1}{N}\right)^k \,. \qquad \square$$

Let us next describe a bound that puts together the above two, and demonstrates the singleton wins when there are other edges that appear with multiplicity greater than 1.

Lemma 6.6. *Let S be an arbitrary subgraph on vertex set $[N]$, and $T \subseteq S$ such that T and $S \setminus T$ are vertex-disjoint, and T forms a matching. Further suppose that $|S| \leqslant \frac{N}{K}$ for some $K \geqslant 2$. Let $m_e \geqslant 1$ be some constant independent of N for each $e \in S$, with $m_e = 1$ for all $e \in T$. Then, for $\mathcal{G}_{[N]}$ a uniformly random perfect matching on $[N]$, we have*

$$\left| \mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}_{[N]}} - \frac{1}{N} \right)^{m_e} \right| \leqslant O\left(\frac{1}{N}\right)^{|S|} \cdot \left(\frac{1}{K} + \sqrt{\frac{|T|}{N}} \right)^{|T|} \,.$$

Proof. Let us expand out the non-T terms in the product. This yields an expression of the form

$$\left| \mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}_{[N]}} - \frac{1}{N} \right)^{m_e} \right| \leqslant \sum_{R \subseteq S \setminus T} \frac{1}{N^{|S \setminus (T \cup R)|}} \left| \mathbf{E} \left[\prod_{e \in T} \left(\mathbf{1}_{e \in \mathcal{G}_{[N]}} - \frac{1}{N} \right)^{m_e} \cdot \mathbf{1}_{R \subseteq \mathcal{G}_{[N]}} \right] \right| \,.$$

Observe now that if R consists of any edges that are not vertex-disjoint, the corresponding term on the right is zero. In case all these edges are vertex-disjoint (so they form a matching), we have $\mathbf{Pr}\left[R \subseteq \mathcal{G}_{[N]} \right] = \frac{(N - 2|R| - 1)!!}{(N-1)!!} = O\left(\frac{1}{N}\right)^{|R|}$. Thus,

$$\left| \mathbf{E} \prod_{e \in S} \left(\mathbf{1}_{e \in \mathcal{G}_{[N]}} - \frac{1}{N} \right)^{m_e} \right| \leqslant \sum_{R \subseteq S \setminus T} \frac{1}{N^{|S \setminus (T \cup R)|}} \cdot O\left(\frac{1}{N}\right)^{|R|} \cdot \left| \mathbf{E} \left[\prod_{e \in T} \left(\mathbf{1}_{e \in \mathcal{G}_{[N]}} - \frac{1}{N} \right)^{m_e} \mid R \subseteq \mathcal{G}_{[N]} \right] \right|$$

$$= O\left(\frac{1}{N}\right)^{|S \setminus T|} \cdot \sum_{R \subseteq S \setminus T} \left| \mathbf{E} \left[\prod_{e \in T} \left(\mathbf{1}_{e \in \mathcal{G}_{[N] \setminus \cup_{e \in R} e}} - \frac{1}{N} \right)^{m_e} \right] \right|$$

We may now apply Lemma 6.5 to the summand to continue the chain of inequalities as follows.

$$\leqslant O\left(\frac{1}{N}\right)^{|S \setminus T|} \cdot \sum_{R \subseteq S \setminus T} O\left(\frac{1}{N\left(1 - \frac{1}{K}\right)} \right)^{|T|} \cdot \left(\frac{1}{K} + \sqrt{\frac{|T|}{N}} \right)^{|T|}$$

$$\leqslant O\left(\frac{1}{N}\right)^{|S|} \cdot \left(\frac{1}{K} + \sqrt{\frac{|T|}{N}} \right)^{|T|} \qquad \square$$

Finally, let us put the pieces together.

Lemma 6.7. *Let H^1, H^2 be a pair of disjoint undirected graphs on n vertices, where the maximum degree in H^1 is 2, and $m_e \geqslant 2$ an n-independent constant for each edge e in H^2. Then,*

$$\left| \mathbf{E} \prod_{e \in H^1} \left(\mathbf{1}_{e \in G} - \frac{d}{n} \right) \prod_{e \in H^2} \left(\mathbf{1}_{e \in G} - \frac{d}{n} \right)^{m_e} \right| \leqslant O\left(\frac{1}{n}\right)^{e(H^1)+e(H^2)} \cdot d^{\frac{1}{2}e(H^1)+e(H^2)},$$

where G is, as usual, a random d-regular graph on n vertices, drawn from the configuration model.

Proof of Theorem 6.2. Let us set up some notation to make things easier. Let H^1, H^2 be a pair of disjoint undirected graphs on n vertices, where the maximum degree of H^1 is 2, and $2 \leqslant m_e \leqslant 4$ some constant for each $e \in H^2$. Denote $k_1 = e(H^1)$, $k_2 = e(H^2)$, and $k = k_1 + k_2$. Also let $\widetilde{k}_2 = \sum_{e \in H^2} m_e$. Observe that $k_1 \leqslant 2n$.

In the terminology of Definition 4.4, H^1 indicates the singleton part of H and H^2 indicates the rest of H. Our goal is to show that

$$\left| \mathbf{E} \left[\prod_{e \in H^1} M_e \cdot \prod_{e \in H^2} M_e^{m_e} \right] \right| \leqslant O\left(\frac{1}{n}\right)^{e(H^1)+e(H^2)}.$$

Let us start by moving to the configuration model, so we may apply the bounds we have proved. For an edge e, let $u_e, v_e \in [n]$ be the two endpoints of e, ordered arbitrarily. Let $\mathcal{G}$ be a draw from the configuration model resulting in G. Then,

$$\mathbf{E} \prod_{e \in H^1} M_e \prod_{e \in H^2} M_e^{m_e}$$

$$= \frac{1}{d^{\frac{1}{2}k_1 + \frac{1}{2}\widetilde{k}_2}} \mathbf{E} \prod_{e \in H^1} \left(\sum_{(i_e, j_e) \in [d]^2} \mathbf{1}_{(u_e, i_e)(v_e, j_e) \in \mathcal{G}} - \frac{1}{dn} \right) \prod_{e \in H^2} \left(\sum_{(i_e, j_e) \in [d]^2} \mathbf{1}_{(u_e, i_e)(v_e, j_e) \in \mathcal{G}} - \frac{1}{dn} \right)^{m_e}$$

$$= \frac{1}{d^{\frac{1}{2}k_1 + \frac{1}{2}\widetilde{k}_2}} \sum_{\substack{\{(i_e, j_e)\}_{e \in H^1} \\ \{(i_e^\ell, j_e^\ell)\}_{e \in H^2}^{1 \leqslant \ell \leqslant m_e}}} \mathbf{E} \prod_{e \in H^1} \left(\mathbf{1}_{(u_e, i_e)(v_e, j_e) \in \mathcal{G}} - \frac{1}{dn} \right) \prod_{e \in H^2} \prod_{\ell \leqslant m_e} \left(\mathbf{1}_{(u_e, i_e^\ell)(v_e, j_e^\ell) \in \mathcal{G}} - \frac{1}{dn} \right)$$

$$= d^{\frac{3}{2}k_1 + \frac{3}{2}\widetilde{k}_2} \cdot \mathop{\mathbf{E}}_{\substack{\{(i_e, j_e)\}_{e \in H^1} \sim [d]^2 \\ \{(i_e^\ell, j_e^\ell)\}_{e \in H^2}^{1 \leqslant \ell \leqslant m_e} \sim [d]^2}} \mathbf{E} \prod_{e \in H^1} \left(\mathbf{1}_{(u_e, i_e)(v_e, j_e) \in \mathcal{G}} - \frac{1}{dn} \right) \prod_{e \in H^2} \prod_{\ell \leqslant m_e} \left(\mathbf{1}_{(u_e, i_e^\ell)(v_e, j_e^\ell) \in \mathcal{G}} - \frac{1}{dn} \right)$$

$$\tag{$\dagger$}$$

To understand the above quantity, we split into cases based on $e(H^2)$.

Case $e(H^2) \geqslant n\sqrt{d}$. First, consider the simpler situation where $e(H^2) \geqslant n\sqrt{d}$. In this case, a simpler argument will suffice where we only win a $\frac{d}{n}$ for each $e \in H^1 \cup H^2$, ignoring the extra wins we get from the singleton edges in H^1. Consider the multiplicity of each edge $e \in H^2$, that is, the number of distinct (i_e^ℓ, j_e^ℓ). It is equal to $1 \leqslant r_e \leqslant m_e$ with probability $O\left(\frac{d^{2r_e}}{d^{2m_e}}\right)$. Further observe that this multiplicity is independent for each edge. Therefore, in this scenario, we may use Lemma 6.3 (using the aforementioned tail bound on the number of distinct edges) to bound this as

$$(\dagger) \leqslant O(1)^k \cdot d^{\frac{3}{2}e(H^1)+\frac{3}{2}\widetilde{k}_2} \cdot \sum_{(r_e)_{e \in H^2}} \left(\prod_{e \in H^2} \frac{d^{2r_e}}{d^{2m_e}} \right) \cdot \left(\frac{1}{dn}\right)^{e(H^1)} \left(\frac{1}{dn}\right)^{\sum r_e}$$

$$= O(1)^k \cdot \left(\frac{\sqrt{d}}{n}\right)^{e(H^1)} \cdot \frac{1}{d^{\frac{1}{2}\widetilde{k}_2}} \cdot \sum_{(r_e)_{e \in H^2}} \left(\frac{d}{n}\right)^{\sum r_e}$$

$$\leqslant O(1)^k \cdot \left(\frac{\sqrt{d}}{n}\right)^{k_1} \cdot \frac{1}{d^{k_2}} \cdot \left(\frac{d}{n}\right)^{k_2} = O\left(\frac{1}{n}\right)^k \cdot d^{\frac{1}{2}k_1}.$$

To conclude here, note that $d^{\frac{1}{2}k_1} \leqslant O(1)^{n \log d} \leqslant O(1)^k$, since $k_2 \geqslant n\sqrt{d}$.

Case $e(H^2) \leqslant n\sqrt{d}$. Let us thus return to (†), assuming now that $e(H^2) \leqslant n\sqrt{d}$. We must introduce some more notation for the rest of this proof. For a specific choice of the labels $(i_e, j_e)_{e \in K}$ for some subgraph K on $[n]$, let us refer to the corresponding subgraph in $[n] \times [d]$ as the "label-extension" of K. Let H^1_{ab} be the set of all edges e in H^1 such that some endpoint of e has degree at least $\sqrt{d}$ in H^2. Let H^1_{nor} be $H^1 \setminus H^1_{ab}$. We will be interested in the value of the following quantity: what is the size of the largest subgraph of H^1_{nor} that is

(a) a matching, and

(b) isolated from the rest of the label-extension of H^2 and H^1_{nor}.

For a given choice of labels, let this subgraph be H^1_*. Observe that because the endpoints of every edge e in H^1_{nor} are incident on at most $O(\sqrt{d})$ edges in $H^1 \cup H^2$, the label-extension of e is in H^1_* with probability at least $1 - O\left(\frac{1}{\sqrt{d}}\right)$—it must simply avoid the $O(\sqrt{d})$ edges arising from (the label-extension of) $(H^1 \cup H^2) \setminus \{e\}$. In particular, $e(H^1_*) = e(H^1_{nor}) - s$ with probability at most $O\left(\frac{1}{d}\right)^{s/2}$. As in the previous argument, let r_e be the multiplicity of the edge $e \in H^2$ after label-extending. We can then use Lemma 6.6 to remove the $\sum r_e + s$ edges arising from $(H^1 \cup H^2) \setminus H^1_*$. Observe that by our assumption that $k = O(n\sqrt{d})$, the constant K in Lemma 6.6 can be taken to be $\Omega(\sqrt{d})$. Consequently, recalling again that $e(H^1_{nor}) \leqslant e(H^1) \leqslant 2n$,

$$(\dagger) \leqslant O(1)^k \cdot d^{\frac{3}{2}k_1 + \frac{3}{2}\widetilde{k}_2} \cdot \sum_{\substack{(r_e)_{e \in H^2} \geqslant 1 \\ 0 \leqslant s \leqslant k}} \underbrace{\left(\prod \frac{d^{2r_e}}{d^{2m_e}}\right)}_{\text{probability of } (r_e)} \cdot \underbrace{\left(\frac{1}{d}\right)^{s/2}}_{\text{probability of } s}$$

$$\cdot \underbrace{\left(\frac{1}{dn}\right)^{\sum r_e + e(H^1_{ab}) + s} \cdot \left(\frac{1}{dn}\right)^{e(H^1_{nor}) - s} \cdot \left(\frac{1}{\sqrt{d}} + \sqrt{\frac{e(H^1_{nor}) - s}{dn}}\right)^{e(H^1_{nor}) - s}}_{\text{from Lemma 6.6}}$$

$$\leqslant O(1)^k \cdot d^{\frac{3}{2}e(H^1) + \frac{3}{2}\widetilde{k}_2} \cdot \sum_{\substack{(r_e)_{e \in H^2} \geqslant 1 \\ 0 \leqslant s \leqslant k}} \left(\prod \frac{d^{2r_e}}{d^{2m_e}}\right) \cdot \left(\frac{1}{d}\right)^{s/2} \cdot \left(\frac{1}{dn}\right)^{\sum r_e + e(H^1_{ab}) + s} \cdot \left(\frac{1}{dn}\right)^{e(H^1_{nor}) - s} \cdot \left(\frac{1}{\sqrt{d}}\right)^{e(H^1_{nor}) - s}$$

$$= O(1)^k \cdot \sum_{\substack{(r_e)_{e \in H^2} \geqslant 1 \\ 0 \leqslant s \leqslant k}} \left(\frac{1}{n}\right)^{k_1 + \sum r_e} \cdot d^{\sum r_e - \frac{1}{2}\widetilde{k}_2 - e(H^1_{ab}) - \frac{3}{2}e(H^1_{nor}) + \frac{3}{2}k_1}$$

$$= O\left(\frac{1}{n}\right)^k \cdot d^{k_2 - \frac{1}{2}\widetilde{k}_2 + \frac{1}{2}e(H^1_{ab})} \leqslant O\left(\frac{1}{n}\right)^k \cdot d^{\frac{1}{2}e(H^1_{ab})}$$

To conclude, we note that much like the simpler setting discussed earlier, $d^{e(H^1_{ab})} = O(1)^{e(H^2)}$, since nearly by definition, $e(H^1_{ab}) \leqslant O(\sqrt{d}) \cdot e(H^2)$. This completes the proof, and brings us to the end of the paper. □

Acknowledgments

We would like to thank Kuikui Liu for insightful and inspiring discussions. We are also grateful to Brice Huang, Daniel Lee, and Francisco Pernice for early discussions about our techniques.

References

[Ahl79]   Lars Valerian Ahlfors. *Complex analysis*, volume 3. McGraw-Hill New York, 1979. 4

[Alo86]   Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. 4

[Bas92]   Hyman Bass. The Ihara-Selberg zeta function of a tree lattice. *International Journal of Mathematics*, 3(06):717–797, 1992. 7

[BCGZ22]   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, 2022. 2, 6

[BGBK20]   Florent Benaych-Georges, Charles Bordenave, and Antti Knowles. Spectral radii of sparse random matrices. *Annales de l'Institut Henri Poincaré*, 2020. 7

[BHY19]   Roland Bauerschmidt, Jiaoyang Huang, and Horng-Tzer Yau. Local Kesten–McKay law for random regular graphs. *Communications in Mathematical Physics*, 369(2):523–636, 2019. 4

[BLM15]   Charles Bordenave, Marc Lelarge, and Laurent Massoulié. Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1347–1357. IEEE, 2015. 3

[BLR25]   Ferenc Bencs, Kuikui Liu, and Guus Regts. On zeros and algorithms for disordered systems: mean-field spin glasses. *arXiv preprint arXiv:2507.15616*, 2025. 2

[Bor19]   Charles Bordenave. A new proof of Friedman's second eigenvalue theorem and its extension to random lifts. Technical Report 1502.04482v4, arXiv, 2019. To appear in Annales scientifiques de l'École normale supérieure. 4

[But07]   Ronald W Butler. *Saddlepoint approximations with applications*, volume 22. Cambridge University Press, 2007. 13, 14

[CGVTvH24] Chi-Fang Chen, Jorge Garza-Vargas, Joel A Tropp, and Ramon van Handel. A new approach to strong convergence. *Annals of Mathematics*, 2024. 1, 4

[EKYY13]   László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. Delocalization and diffusion profile for random band matrices. *Communications in Mathematical Physics*, 323(1):367–416, 2013. 1

[FK81]   Zoltán Füredi and János Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981. 1

[FM17]   Zhou Fan and Andrea Montanari. How well do local algorithms solve semidefinite programs? In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 604–614, 2017. 3

[Fri08]   Joel Friedman. *A proof of Alon's second eigenvalue conjecture and related problems*. American Mathematical Soc., 2008. 4

[Gir85]   Vyacheslav L Girko. Circular law. *Theory of Probability & Its Applications*, 29(4):694–706, 1985. 6

[Gir18]   Vyacheslav L Girko. From the first rigorous proof of the circular law in 1984 to the circular law for block random matrices under the generalized Lindeberg condition. *Random Operators and Stochastic Equations*, 26(2):89–116, 2018. 6

[HMY24]   Jiaoyang Huang, Theo McKenzie, and Horng-Tzer Yau. Ramanujan Property and Edge Universality of Random Regular Graphs. *arXiv preprint arXiv:2412.20263*, 2024. 4

[HY24]   Jiaoyang Huang and Horng-Tzer Yau. Spectrum of Random d-regular Graphs Up to the Edge. *Communications on Pure and Applied Mathematics*, 77(3):1635–1723, 2024. 4

[Iha66]   Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over p-adic fields. *Journal of the Mathematical Society of Japan*, 18(3):219–235, 1966. 7

[Nil91]   Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991. 4

[Tal]   Michel Talagrand. Upper and lower bounds for stochastic processes. 1

[Tro15]   Joel A Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230, 2015. 1

[TVK10]   Terence Tao, Van Vu, and Manjunath Krishnapur. Random matrices: Universality of ESDs and the circular law. 2010. 6