

LARGEST EIGENVALUE OF THE CONFIGURATION MODEL AND BREAKING OF ENSEMBLE EQUIVALENCE

PIERFRANCESCO DIONIGI, DIEGO GARLASCHELLI, RAJAT SUBHRA HAZRA,
AND FRANK DEN HOLLANDER

ABSTRACT. We analyse the largest eigenvalue of the adjacency matrix of the configuration model with large degrees, where the latter are treated as hard constraints. In particular, we compute the expectation of the largest eigenvalue for degrees that diverge as the number of vertices n tends to infinity, uniformly on a scale between 1 and \sqrt{n} , and show that a weak law of large numbers holds. We compare with what was derived in our earlier paper [14] for the Chung-Lu model, which in the regime considered represents the corresponding configuration model with soft constraints, and show that the expectation is shifted down by 1 asymptotically. This shift is a signature of breaking of ensemble equivalence between the hard and soft (also known as micro-canonical and canonical) versions of the configuration model. The latter result generalizes a previous finding [13] obtained in the case when all degrees are equal.

1. INTRODUCTION AND MAIN RESULTS

Motivation. Spectral properties of adjacency matrices in random graphs play a crucial role in various areas of network science. The largest eigenvalue is particularly sensitive to the graph architecture, making it a key focus. In this paper we focus on a random graph with a hard constraint on the degrees of nodes. In the homogeneous case (all degrees equal to d), it reduces to a random d -regular graph. In the heterogeneous case (different degrees), it is known as the configuration model. Our interest is characterizing the expected largest eigenvalue of the configuration model and comparing it with the same quantity for a corresponding random graph model where the degrees are treated as soft constraints.

The set of d -regular graphs on n vertices, with $d = d(n)$, is non-empty when $1 \leq d \leq n - 1$ and dn is even. Selecting a graph uniformly at random from this set results in what is known as the *random d -regular graph*, denoted by $G_{n,d}$. The spectral properties of $G_{n,d}$ are well-studied for $d \geq 2$ (for $d = 1$, the graph is trivially a set of disconnected edges). For instance, all eigenvalues of its adjacency matrix fall in the interval $[-d, d]$, with the largest eigenvalue being d . The computation of $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$, where λ_2 and λ_n are the second-largest and the smallest eigenvalue, respectively, has been challenging. It is well known that for fixed d the empirical distribution function of the eigenvalues of $G_{n,d}$ converges to the so-called

Date: December 10, 2023.

2000 Mathematics Subject Classification. 60B20, 60C05, 60K35.

Key words and phrases. Constrained random graphs; micro-canonical and canonical ensembles; adjacency matrix; largest eigenvalue; breaking of ensemble equivalence.

PD, RSH, FdH and MM are supported by the Netherlands Organisation for Scientific Research (NWO) through the Gravitation-grant NETWORKS-024.002.003. DG is supported by the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, The Netherlands) and by the European Union - NextGenerationEU - National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR), project ‘SoBig-Data.it - Strengthening the Italian RI for Social Mining and Big Data Analytics’ - Grant IR0000013 (n. 3264, 28/12/2021) (<https://pnrr.sobigdata.it/>).

Kesten-McKay law [22], the density of which is given by

$$f_{\text{KM}}(x) = \begin{cases} \frac{d\sqrt{4(d-1)-x^2}}{2\pi(d^2-x^2)} & \text{if } |x| \leq 2\sqrt{d-1}, \\ 0 & \text{otherwise.} \end{cases}$$

A lower bound of the type $\lambda \geq 2\sqrt{d-1}$ was expected from this, and was indeed proven in the celebrated Alon-Boppana theorem in [1], stating that $\lambda \geq 2\sqrt{d-1} + O_d(\log^{-2}(n))$, where the constant in the error term only depends on d . In the same paper an upper bound of the type $\lambda \leq 2\sqrt{d-1} + o(1)$ was conjectured. This conjecture was later proven in the pioneering work [17] (for a shorter proof, see [6]). When $d = d(n) \rightarrow \infty$, the spectral distribution converges to the semi-circular law [26]. It was proven in [10] that $\lambda = O(\sqrt{d})$ with high probability when $d = o(\sqrt{n})$. This result was extended in [3] to $d = o(n^{2/3})$. For recent results and open problems, we refer the reader to [30].

There has not been much work on the inhomogeneous setting where not all the degrees of the graph are the same. The natural extension of the regular random graph is the configuration model, where a degree sequence $\vec{d}_n = (d_1, \dots, d_n)$ is prescribed. Unless otherwise specified, we assume the degrees to be hard constraints, i.e. realized exactly on each configuration of the model (this is sometimes called the ‘microcanonical’ configuration model, as opposed to the ‘canonical’ one where the degrees are soft constraints realized only as ensemble averages [20, 24]). The empirical spectral distribution of the configuration model is known under some assumptions on the growth of the sum of the degrees. When $\sum_{i=1}^n d_i = O(n)$, the graph locally looks like a tree. It has been shown that the empirical spectral distribution exists and that the limiting measure is a functional of a size-biased Galton-Watson tree [7]. When $\sum_{i=1}^n d_i$ grows polynomially with n , the local geometry is no longer a tree. This case was recently studied in [12], where it is shown that the appropriately scaled empirical spectral distribution of the configuration model converges to the free multiplicative convolution of the semi-circle law with a measure that depends on the empirical degree sequence. The scaling of the largest and the second-largest eigenvalues has not yet been studied in full generality.

For the configuration model the behaviour of the largest eigenvalue is non-trivial. In the present paper, we consider the configuration model with large degrees, compute the expectation of the largest eigenvalue of its adjacency matrix, and prove a weak law of large numbers. When $d \rightarrow \infty$, the empirical distribution of $G_{n,d}$ (appropriately scaled) converges to the semi-circle law. Also, if we look at Erdős-Rényi random graphs on n vertices with connection probability d/n , then the appropriately scaled empirical distribution converges to the semi-circle law. It is known that the latter two random graphs exhibit different behaviour for the expected largest eigenvalue [13]. This can be understood in the broader perspective of *breaking ensemble equivalence* [20, 24], which we discuss later. In the inhomogeneous setting, the natural graph to compare the configuration model with is the Chung-Lu model. For the case where the sum of the degrees grows with n , it was shown in [11] that the empirical spectral distribution converges to the free multiplicative product of the semi-circle law with a measure that depends on the degree sequence, similar to [12]. In [14], we investigated the largest eigenvalue and its expectation for the Chung-Lu model (and even derived a central limit theorem). In the present paper we complete the picture by showing that, when the degrees are large, there is a *gap* between the expected largest eigenvalue of the Chung-Lu model and the configuration model. We refer the reader to the references listed in [13, 14] for more background.

Main theorem. For $n \in \mathbb{N}$, let $\text{CM}(\vec{d}_n)$ be the random graph on n vertices generated according to the *configuration model* with degree sequence $\vec{d}_n = (d_i)_{i \in [n]} \in \mathbb{N}^n$ [27, Chapter 7.2].

Define

$$m_0 = \min_{i \in [n]} d_i, \quad m_\infty = \max_{i \in [n]} d_i, \quad m_k = \sum_{i \in [n]} (d_i)^k, \quad k \in \mathbb{N}.$$

Throughout the paper we need the following assumptions on \vec{d}_n as $n \rightarrow \infty$.

Assumption 1.1.

(D1) **Bounded inhomogeneity:** $m_0 \asymp m_\infty$.

(D2) **Connectivity and sparsity:** $1 \ll m_\infty \ll \sqrt{n}$. ♠

Under these assumptions, $\text{CM}(\vec{d}_n)$ is with high probability *non-simple* [27, Chapter 7.4]. We write \mathbb{P} and \mathbb{E} to denote probability and expectation with respect to the law of $\text{CM}(\vec{d}_n)$ *conditional on being simple*, suppressing the dependence on the underlying parameters.

Let $A_{\text{CM}(\vec{d}_n)}$ be the adjacency matrix of $\text{CM}(\vec{d}_n)$. Let $(\lambda_i)_{i \in [n]}$ be the eigenvalues of $A_{\text{CM}(\vec{d}_n)}$, ordered such that $\lambda_1 \geq \dots \geq \lambda_n$. We are interested in the behaviour of λ_1 as $n \rightarrow \infty$. Our main theorem reads as follows.

Theorem 1.2. *Subject to Assumption 1.1,*

$$\mathbb{E}[\lambda_1] = \frac{m_2}{m_1} + \frac{m_1 m_3}{m_2^2} - 1 + o(1), \quad n \rightarrow \infty, \quad (1.1)$$

and

$$\frac{\lambda_1}{\mathbb{E}[\lambda_1]} \rightarrow 1 \text{ in } \mathbb{P}\text{-probability.}$$

In [14] we looked at an alternative version of the configuration model, called the *Chung-Lu model* $\text{CM}_n^*(\vec{d}_n)$, where the *average degrees*, rather than the degrees themselves, are fixed at \vec{d}_n . This is an ensemble with soft constraints; in the considered regime for the degrees, it coincides with a maximum-entropy ensemble, also called ‘canonical’ configuration model [24]. For this model we showed that, subject to Assumption 1.1,

$$\mathbb{E}^*[\lambda_1] = \frac{m_2}{m_1} + \frac{m_1 m_3}{m_2^2} + o(1), \quad n \rightarrow \infty, \quad (1.2)$$

and $\lambda_1/\mathbb{E}^*[\lambda_1] \rightarrow 1$ in \mathbb{P}^* -probability, where \mathbb{P}^* and \mathbb{E}^* denote expectation with respect to the law of $\text{CM}_n^*(\vec{d}_n)$ and λ_1 is the largest eigenvalue of the $A_{\text{CM}_n^*(\vec{d}_n)}$. The notable difference between (1.1) and (1.2) is the shift by -1 .

For the special case where all the degrees are equal to d , we have $m_0 = m_\infty = d$ and $m_k = nd^k$, and so $\mathbb{E}[\lambda_1] = d + o(1)$ and $\mathbb{E}^*[\lambda_1] = d + 1 + o(1)$. In fact, $\mathbb{P}(\lambda_1 = d) = 1$. Since in this model the degrees can fluctuate with the same law (*soft constraint* in the physics literature), this case reduces to the Erdős-Rényi random graph for which results on $\mathbb{E}^*[\lambda_1]$ were already well known in [19, 21] and further analyzed in [15].

Breaking of ensemble equivalence. The shift by -1 was proven in [13] for the homogeneous case with equal degrees and is a *spectral signature* of *breaking of ensemble equivalence* [24]. Indeed, a d -regular graph is the ‘micro-canonical’ version of a random graph where all degrees are equal and ‘hard’, and the Erdős-Rényi random graph is the corresponding ‘canonical’ version where all degrees are equal and ‘soft’. More in general, $\text{CM}(\vec{d}_n)$ is the micro-canonical configuration model where the constraint on the degrees is ‘hard’, while $\text{CM}_n^*(\vec{d}_n)$ is the canonical version where the constraint is ‘soft’. We refer the reader to [20] for the precise definition of these two configuration model *ensembles* and for the proof that they are not asymptotically equivalent *in the measure sense* [25]. This means that the *relative entropy per node* of \mathbb{P} with respect to \mathbb{P}^* has a strictly positive limit as $n \rightarrow \infty$. This shows that the choice of constraint matters, not only on a microscopic scale but also on a

macroscopic scale. Indeed, for non-equivalent ensembles one expects the existence of certain macroscopic properties that have different expectation values in the two ensembles (*macrostate (in)equivalence* [25]). The fact that the largest eigenvalue picks up this discrepancy is interesting. What is remarkable is that the shift by -1 , under the hypotheses considered, holds true also in the case of heterogeneous degrees and remains the same *irrespective* of the scale of the degrees and of the distribution of the degrees on this scale.

Outline. The remainder of this paper is organised as follows. In Section 2 we look at the issue of simplicity of the graph. In Section 3 we bound the spectral norm of the matrix

$$H = A_{\text{CM}(\vec{d}_n)} - \mathbb{E}[A_{\text{CM}(\vec{d}_n)}]. \quad (1.3)$$

We use the proof of [10] to show that $\|H\| = o(m_\infty)$ with high probability. Using the latter we show that

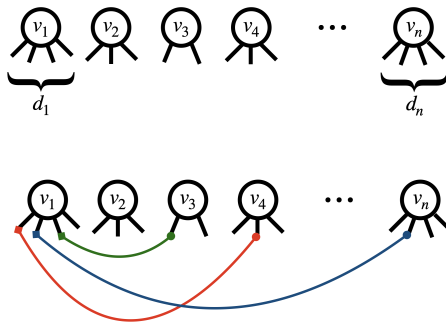
$$\lambda_1 \sim \lambda_1 \left(\mathbb{E}[A_{\text{CM}(\vec{d}_n)}] \right) \sim \frac{m_2}{m_1}$$

with high probability. In Section 4 we use the estimates in Section 3 to prove Theorem 1.2.

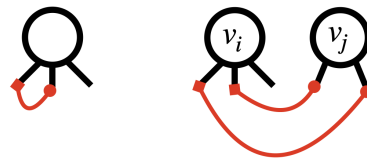
2. CONFIGURATION MODEL AND SIMPLE GRAPHS WITH A GIVEN DEGREE SEQUENCE

In the context of random graphs with hard constraints on the degrees, we work with both the *configuration model* (a random *multi-graph* with a prescribed degree sequence) and with the *conditional configuration model* (a random *simple* graph with a prescribed degree sequence). We view the second as a special case of the first.

Configuration model. The configuration model with degree sequence $\vec{d} = (d_1, \dots, d_n)$, $\text{CM}(\vec{d}_n)$, generates a graph through a *perfect matching* P of the set of half-edges $\mathcal{E} = \cup_{i \in [n]} \{i\} \times [d_i]$, i.e., $P: \mathcal{E} \rightarrow \mathcal{E}$ such that P is an isomorphism and an involution, and $P(\alpha) \neq \alpha$ for all $\alpha \in \mathcal{E}$.



(A) Pairing procedure.



(B) The configuration model generates multi-graphs with self-loops (left) and multiple edges (right).

It turns out that a pairing scheme, such as matching half-edges from left to right or selecting pairs uniformly at random, is considered admissible as long as it yields the correct probability for a perfect matching (see, for example, [27, Lemma 7.6]). An important property of each admissible pairing scheme is that

$$\mathbb{P}_{\text{CM}(\vec{d}_n)}(\alpha \sim \beta) = \frac{1}{m_1 - 1} \quad \forall \alpha, \beta \in \mathcal{E}. \quad (2.1)$$

We can therefore endow \mathcal{E} with the lexicographic order. An element $\alpha \in \mathcal{E}$, $\alpha = (i, \ell)$, is associated with a vertex $v(\alpha) = i$ through the first component of α . An edge e is an element of $\mathcal{E} \times \mathcal{E}$ given by the pairing $e = \{\alpha, \beta\} = \{\alpha, P(\alpha)\}$. We define the configuration \mathcal{C} as the set of all such e . Given the lexicographic order, we may assume that $\alpha < \beta$ and identify arbitrarily the head and the tail of the edge: $t(e) = v(\alpha)$, $h(e) = v(\beta)$. We can then order the edges in their order of appearance via $t(e)$, forming a list $\{e_i, i = 1, \dots, m_1/2\}$. These properties of the configuration model will be used in Subsection 3.1. For the configuration model it is easy to check that¹

$$\mathbb{E}[a_{ii}] = \frac{d_i(d_i - 1)}{m_1 - 1}, \quad \mathbb{E}[a_{ij}] = \frac{d_i d_j}{m_1 - 1}, \quad i \neq j. \quad (2.2)$$

Indeed every matrix element a_{ij} , can be expressed as

$$\sum_{k=1}^{d_i} \sum_{h=1}^{d_j} \mathbb{1}(\alpha \sim \beta, \alpha = (i, k), \beta = (j, h)),$$

where $\mathbb{1}(E)$ is the indicator function of the event E . Taking expectations and using (2.1), we get 2.2.

Simple graphs with a given degree sequence. A linked but different model is the one that samples uniformly at random a simple graph with a given degree sequence \vec{d} , $\mathcal{G}(\vec{d})$ ². An immediate question is whether there is any relation between $\text{CM}(\vec{d})$ and $\mathcal{G}(\vec{d})$. It turns out that the following is true (see [27] for reference):

Lemma 2.1. *Let be G_n a graph generated via the configuration model $\text{CM}(\vec{d}_n)$ with degree sequence \vec{d} . Then*

$$\mathbb{P}_{\text{CM}(\vec{d}_n)}(G_n \text{ is simple}) = \exp \left[-O \left(\frac{m_1^2}{n^2} \right) \right] \quad (2.3)$$

and

$$\mathbb{P}_{\text{CM}(\vec{d}_n)}(G_n = \cdot \mid G_n \text{ is simple}) \text{ is uniform.} \quad (2.4)$$

Thus, we can identify the law of the *uniform random graph with a given degree sequence* with the law of the *configuration model conditional on simplicity*, i.e.,

$$\mathbb{P}_{\mathcal{G}(\vec{d})}(\cdot) = \mathbb{P}_{\text{CM}(\vec{d})}(\cdot \mid G \text{ is simple}).$$

It follows that the expectation under the uniform random graph with a given degree sequence can be expressed as a the expectation of the configuration model conditional on simplicity. As explained in [28, Remark 1.14], we have

$$\begin{aligned} \mathbb{E}_{\mathcal{G}(\vec{d})}[a_{ij}] &= \mathbb{P}_{\mathcal{G}(\vec{d})}(i \sim j) = (1 + o(1)) \frac{d_i d_j}{m_1 + d_i d_j} \\ &\leq \frac{d_i d_j}{m_1(1 + O(1/m_\infty))} = \frac{d_i d_j}{m_1} + O(1/m_\infty). \end{aligned} \quad (2.5)$$

¹We adopt the convention that a self-loop contributes 2 to the degree, i.e., a_{ii} is twice the number of self-loops attached to vertex i . This convention is useful because it yields $\sum_j a_{ij} = d_i$.

²With a little abuse of notation we permit our simple graphs to have self-loops. We maintain the convention explained in the previous footnote.

3. BOUND ON $\|H\|$

In this section we derive the following bound on the spectral norm of the matrix H defined in (1.3). This bound will play a crucial role in the proof of Theorem 1.2 in Section 4.

Theorem 3.1. *For every $K > 0$, there exists a constant $C > 0$ such that*

$$\|H\| \leq C\sqrt{m_\infty} \text{ with probability } 1 - o(n^{-K}),$$

where the constant C depends on $\frac{m_\infty}{m_0}$ and K only.

The proof is given in Section 3.1 and we follow the proof of Lemma 16 of [10]. In Sections 3.2–3.3 we derive some estimates that are needed in Section 3.1.

We have

$$\|H\| = \max\{\lambda_1(H), |\lambda_n(H)|\}.$$

Recall (1.3). In Lemma 3.3 below we will see that $\mathbb{E}[A_{\text{CM}(\vec{d}_n)}]$ is asymptotically rank 1 (more precisely, the other eigenvalues are of order $O(n^{-1})$). This means that $\lambda_i(A_{\text{CM}(\vec{d}_n)}) = O(n^{-1})$, $i \neq 1$. Therefore, by Weyl's inequality, it is easy to see that $\lambda_1(H)$ and $\lambda_n(H)$ are bounded above and below by $\lambda_2(A_{G_n})$ and $\lambda_n(A_{G_n})$ with error $O(n^{-1})$, while $\lambda_1(A_{G_n})$ is in a window of size $\sqrt{m_\infty}$ around $\frac{m_2}{m_1-1}$. We therefore have the following Corollary, which will be needed in Section 4:

Corollary 3.2. *For every $K > 0$,*

$$\lambda_1(A_{G_n}) - \lambda_1(\mathbb{E}[A_{G_n}]) = O(\sqrt{m_\infty}) \text{ with probability } 1 - o(n^{-K}).$$

3.1. Spectral estimates.

Notation. Abbreviate $G_n = \text{CM}(\vec{d}_n)$. Let U be uniformly distributed on $[n]$, let d_U be the degree of a vertex that is picked uniformly at random, and let

$$\omega_n = \mathbb{E}[d_U]$$

be the average of the empirical degree distribution. Under Assumption 1.1, $\omega_n \rightarrow \infty$ and $\omega_n = o(n)$. We define the *normalised degree sequence* $(\hat{d}_i)_{i \in [n]}$ as

$$\hat{d}_i = \frac{d_i}{\omega_n} \tag{3.1}$$

and the *normalised adjacency matrix* \hat{A}_{G_n} as

$$\hat{A}_{G_n} = \frac{A_{G_n}}{\sqrt{\omega_n}}.$$

In the following we will need multiples of the vector

$$\tilde{e}_i = \frac{d_i}{\sqrt{m_1 - 1}}, \quad 1 \leq i \leq n, \tag{3.2}$$

which is the eigenvector corresponding to the rank-1 approximation of $\mathbb{E}[A_{G_n}]$, as it is easy to check using (2.2).

Two lemmas. The matrix $\mathbb{E}[A_{G_n}]$ is asymptotically rank 1:

Lemma 3.3. *Define*

$$\tilde{e} = (\tilde{e}_1, \dots, \tilde{e}_n)^t$$

and

$$A_{\text{shift}} = \mathbb{E}[A_{G_n}] - |\tilde{e}\rangle\langle\tilde{e}| = -\frac{1}{m_1 - 1} \text{diag}(d_1, \dots, d_n).$$

Then

$$\|A_{\text{shift}}\| = O(n^{-1}).$$

Proof. Using (2.2), we have A_{shift} is a diagonal matrix and hence $\|A_{\text{shift}}\| \leq \frac{m_\infty}{m_1 - 1} = O(n^{-1})$ under Assumption 1.1. Note that when all the degrees are equal to d , then this bound is sharp, i.e., $\|A_{\text{shift}}\| = \frac{d}{dn-1}$, and $\mathbb{E}[A_{G_n}]$ is exactly rank 1. \square

It is shown in [12] that the empirical spectral distribution of A_{G_n} has a deterministic limit given by $\mu = \mu_{sc} \boxtimes \mu_{\hat{D}}$, where μ_{sc} is the standard Wigner semicircle law, $\mu_{\hat{D}}$ is the distribution of \hat{d}_U , and \boxtimes is the free product defined in [2, 4, 29].

Lemma 3.4. *Let $(\hat{d}_i)_{i \in [n]}$ be the normalised degree sequence, and suppose that Assumption 1.1 holds and*

$$\frac{1}{n} \sum_{i=1}^n \delta_{\hat{d}_i} \Rightarrow \mu_{\hat{D}}. \quad (3.3)$$

Then $\mu_{\hat{D}}$ is compactly supported.

Proof. By Assumption 1.1 D(2),

$$0 < c \leq \liminf_{n \rightarrow \infty} \frac{\min_i \hat{d}_i}{\max_i \hat{d}_i} \leq \limsup_{n \rightarrow \infty} \frac{\max_i \hat{d}_i}{\min_i \hat{d}_i} \leq C < \infty.$$

Hence the support of \hat{d}_U is contained in a multiple of $[c, C]$, and (3.3) implies that $\mu_{\hat{D}}$ is compactly supported. \square

Key estimates. From [12, Theorem 1] we know that, under Assumption 1.1, the empirical spectral distribution of \hat{A}_{G_n} , written

$$\mu_{\hat{A}_{G_n}},$$

converges weakly to $\mu_{sc} \boxtimes \mu_{\hat{D}}$. The fact that μ_{sc} and $\mu_{\hat{D}}$ are compactly supported implies that $\mu_{sc} \boxtimes \mu_{\hat{D}}$ is compactly supported too. However, this still allows that H has $o(n)$ outliers, which possibly control $\|H\|$. It is hard to analyse $\|H\|$ for sparse graphs, because it is related to expansion properties of the graph, mixing times of random walks on the graph, and more. To prove that $\|H\| = O(\sqrt{m_\infty})$, we will use the argument described in [10], which is an adaptation of the argument in [18]. We are only after the order of $\|H\|$, not sharp estimates.

For random regular graphs with fixed degree, the problem of computing $\|H\|$ was solved in [16]. We are not aware of any proof concerning the second largest eigenvalue of configuration models with bounded degrees, but any sharp bound must come from techniques of the type employed in [5, 16]. In what follows we prove Theorem 3.1 based on [18] and its adaptation to the inhomogeneous setting in [10]. In the following, we present the proof as outlined in the latter paper, adapting it to our case to make our paper self-contained. This adjustment is necessary because the result we need is somewhat obscured in the context in which it appears in [10].

Let $D_n = \text{diag}(d_1, \dots, d_n)$. The transition kernel of a random walk on the graph G_n is given by $P_n = D_n^{-1}A_{G_n}$. The *random* matrix P_n has as principal normalised eigenvector $\vec{1} = (1, \dots, 1)/\sqrt{n}$ with eigenvalue 1. Define

$$\lambda^* = \max\{\lambda_2(P_n), |\lambda_n(P_n)|\}.$$

Note that the matrices P_n and $S_n = D_n^{1/2}P_nD_n^{-1/2}$ have the same spectrum by a similarity transformation. Hence we can write the Rayleigh formula

$$\lambda^* = \max_{z: \langle z, \vec{1} \rangle = 0} \frac{|\langle P_n z, z \rangle|}{\|z\|^2} = \max_{x: \langle x, \sqrt{\vec{d}} \rangle = 0} \frac{|\langle S_n x, x \rangle|}{\|x\|^2},$$

where $\sqrt{\vec{d}} = (\sqrt{d_1}, \dots, \sqrt{d_n})$. Define $M_n = D_n^{-1}A_nD_n^{-1}$. Let $\tilde{\lambda}$ be the second largest eigenvalue of M_n . Then

$$\langle S_n x, x \rangle = \langle D_n^{1/2}M_nD_n^{1/2}x, x \rangle = \langle M_nD_n^{1/2}x, D_n^{1/2}x \rangle.$$

Since

$$\frac{\langle x, x \rangle}{\|x\|^2} \geq \frac{1}{m_\infty} \frac{\langle D_n^{1/2}x, D_n^{1/2}x \rangle}{\|x\|^2},$$

putting $y = D_n^{1/2}x$ we see that

$$\lambda^* \leq m_\infty \max_{\langle y, \vec{1} \rangle = 0} \frac{|\langle M_n y, y \rangle|}{\|y\|^2} = m_\infty \tilde{\lambda}, \quad (3.4)$$

which gives a bound of the type $\|H\| = O(m_\infty^2 \tilde{\lambda})$. Note that the matrix elements of M_n can be expressed as

$$(M_n)_{ij} = \frac{a_{ij}}{d_i d_j}, \quad (3.5)$$

where a_{ij} counts the number of edges between vertices i and j , with the convention for the diagonal elements stated earlier. In view of (3.4), in order to obtain a bound on λ^* we must focus on $\tilde{\lambda}$. In fact, we must show that

$$\tilde{\lambda} = O(m_\infty^{-3/2}), \quad (3.6)$$

in order to obtain the desired bound $\|H\| = O(\sqrt{m_\infty})$. To achieve this we proceed in steps:

- We reduce the computation of $\tilde{\lambda}$ to the analysis of two terms.
- In (3.7) below we identify the leading order of $\mathbb{E}[\tilde{\lambda}]$ from these two terms, which turns out to be $O(m_\infty^{-3/2})$.
- We show that the other terms are of higher order and therefore are negligible.
- We prove concentration around the mean through concentration of the leading order term in Lemma 3.8 below.
- Recalling the denominator of (3.5), we get a bound on $\|H\|$ after multiplying by m_∞^2 .

To prove (3.6), we reduce the problem to a maximisation problem in a simpler space. Namely, let $\varepsilon \in (0, 1)$, and

$$T = \left\{ x \in \left(\frac{\varepsilon}{\sqrt{n}} \mathbb{Z} \right)^n : \sum_{i \in [n]} x_i = 0, \sum_{i \in [n]} x_i^2 \leq 1 \right\}.$$

Then, using the formula for the volume of the n -dimensional ball, we have

$$|T| \leq \left(\frac{(2 + \varepsilon)\sqrt{n}}{2\varepsilon} \right)^n \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} \leq \left(\frac{(2 + \varepsilon)\sqrt{2\pi e}}{2\varepsilon} \right)^n.$$

The maximisation over \mathbb{R}^n can be reduced to over T and this leads to an error that depends on ε .

Lemma 3.5. *Let*

$$\lambda = \max\{|\langle x, M_n y \rangle| : x, y \in T\}.$$

Then

$$\tilde{\lambda} \leq (1 - \varepsilon)^{-2} \lambda.$$

Proof. Let be $\mathcal{S} = \{x \in \mathbb{R}^n : \sum_{i \in [n]} x_i = 0, \|x\| \leq 1\}$. We want to show that for every $x \in \mathcal{S}$ there is a vector $y \in T$ such that $\|x - y\| \leq \varepsilon$ and $\sum_{i \in [n]} (x_i - y_i) = 0$. Let us write the components of x as

$$x_i = \varepsilon \frac{m_i}{\sqrt{n}} + f_i, \quad i \in [n],$$

where $m_i \in \mathbb{Z}$ and $f_i \in [0, \varepsilon n^{-1/2})$ is an error term. Because $\sum_{i \in [n]} x_i = 0$, we choose m_i such that $\sum_{i \in [n]} f_i v = v \varepsilon f n^{-1/2}$, where f is a non-negative integer smaller than n . We relabel the indices in such a way that $m_i \leq m_j$ when $i \leq j$. Now consider the vector y given by

$$y_i = \begin{cases} \varepsilon \frac{m_i+1}{\sqrt{n}} & \text{if } i \leq f, \\ \varepsilon \frac{m_i}{\sqrt{n}} & \text{if } i > f. \end{cases}$$

It follows that $\sum_{i \in [n]} y_i = 0$ and $\|y\| \leq 1$, and therefore $y \in T$. Furthermore, because $|x_i - y_i| \leq \varepsilon n^{-1/2}$ by construction, we have the required property $\|x - y\| \leq 1$. Iterating the previous argument, we can express every vector $x \in \mathcal{S}$ in terms of a sequence of vectors $(y^{(i)})_{i \in [n]}$ in T such that

$$x = \sum_{i \in [n]} \varepsilon^i y^{(i)}.$$

Therefore, because (3.4) is maximized on \mathcal{S} , we have

$$\langle x, M_n x \rangle = \sum_{i, j \in [n] \times [n]} \varepsilon^{i+j} \langle x^{(i)} M_n x^{(j)} \rangle \leq \frac{1}{(1 - \varepsilon)^2} \max\{|\langle y, M_n z \rangle| : y, z \in T\},$$

from which the claim follows. \square

Our next goal is to show that $|\langle x, M_n y \rangle| = o(m_\infty^{-3/2})$ for all $x, y \in T$ with a suitably high probability. This can be done in the following way. Fix $x, y \in T$ and define the random variable

$$X = \sum_{i, j \in [n] \times [n]} x_i (M_n)_{ij} y_j.$$

Define the set of indices

$$\mathcal{B} = \left\{ (i, j) \in [n] \times [n] : 0 < |x_i y_j| < \frac{\sqrt{m_\infty}}{n} \right\}.$$

Then we can rewrite $X = X' + X''$ with

$$X' = \sum_{(i, j) \in \mathcal{B}} x_i (M_n)_{ij} y_j, \quad X'' = \sum_{(i, j) \notin \mathcal{B}} x_i (M_n)_{ij} y_j.$$

In Section 3.2 we show that $\mathbb{E}[X']$ is of the correct order and that X' is well concentrated around its mean. In Section 3.3 we analyse X'' , which is of a different nature and requires that we exclude subgraphs in the configuration model that are too dense.

3.2. Estimate of first contribution. **1.** Use (2.2) and (3.5) to write out $\mathbb{E}[m_{ij}] = \mathbb{E}[(M_n)_{ij}]$. This gives

$$\mathbb{E}[X'] = \sum_{(i,j) \in \mathcal{B}} \frac{x_i y_j}{m_1 - 1} - \sum_{(i,i) \in \mathcal{B}} \frac{x_i y_i}{d_i^2 (m_1 - 1)}.$$

In view of the bound on $|x_i y_i|$, the last term gives a contribution

$$\left| \sum_{(i,i) \in \mathcal{B}} \frac{x_i y_i}{d_i^2 (m_1 - 1)} \right| = O\left(\frac{1}{m_1 m_\infty^{3/2}}\right).$$

Since $x, y \in T$, we have that $\sum_{i \in [n]} x_i = 0$ and $\sum_{j \in [n]} y_j = 0$, and therefore $\sum_{(i,j) \in [n] \times [n]} x_i y_j = 0$. Hence

$$\left| \sum_{(i,j) \in \mathcal{B}} x_i y_j \right| = \left| \sum_{(i,j) \notin \mathcal{B}} x_i y_j \right|,$$

where we can bound the right-hand side as

$$\left| \sum_{(i,j) \notin \mathcal{B}} x_i y_j \right| \leq \sum_{(i,j): |x_i y_j| \geq \frac{\sqrt{m_\infty}}{n}} |x_i y_j| \leq \sum_{(i,j): |x_i y_j| \geq \frac{\sqrt{m_\infty}}{n}} \frac{x_i^2 y_j^2}{|x_i y_j|} \leq \frac{n}{\sqrt{m_\infty}} \sum_{(i,j)} x_i^2 y_j^2 \leq \frac{n}{\sqrt{m_\infty}}.$$

We can therefore conclude that

$$|\mathbb{E}[X']| \leq \frac{n}{(m_1 - 1)\sqrt{m_\infty}} + O\left(\frac{1}{m_1 m_\infty^{3/2}}\right). \quad (3.7)$$

2. To prove that X' is concentrated around its mean, we use an argument originally developed in [9, 23] and used for the configuration model in [10, 18]. This argument employs the *martingale structure* of the configuration model *conditional on partial pairings*. Define

$$\chi(x) = \begin{cases} x, & \text{if } |x| < \frac{\sqrt{m_\infty}}{n}, \\ 0, & \text{otherwise.} \end{cases}$$

We can then express X' as

$$X' = \sum_{e \in \mathcal{C}} \frac{\chi(x_{t(e)}) y_{h(e)}}{d_{t(e)} d_{h(e)}} + \sum_{e \in \mathcal{C}} \frac{\chi(x_{h(e)}) y_{t(e)}}{d_{t(e)} d_{h(e)}} = X'_a + X'_b.$$

We divide the set of pairings \mathcal{C} into three sets

$$\begin{aligned} \mathcal{C}_1 &= \left\{ e \in \mathcal{C} : |x_{t(e)}| > \frac{1}{\varepsilon \sqrt{n}} \right\}, \\ \mathcal{C}_2 &= \left\{ e \in \mathcal{C} : |y_{t(e)}| > \frac{1}{\varepsilon \sqrt{n}}, |x_{t(e)}| \leq \frac{1}{\varepsilon \sqrt{n}} \right\}, \\ \mathcal{C}_3 &= \left\{ e \in \mathcal{C} : |y_{t(e)}| \leq \frac{1}{\varepsilon \sqrt{n}}, |x_{t(e)}| \leq \frac{1}{\varepsilon \sqrt{n}} \right\}, \end{aligned}$$

and write

$$X'_a = X_1 + X_2 + X_3, \quad X_i = \sum_{e \in \mathcal{C}_i} \frac{\chi(x_{t(e)}) y_{h(e)}}{d_{t(e)} d_{h(e)}}.$$

We can do a similar decomposition for X'_b .

3. The following martingale lemma from [10, 18] is the core estimate that we want to apply to each of the X_i 's.

Lemma 3.6. *Let be G_1 and G_2 two graphs generated via the configuration model through perfect matchings P_1 and P_2 . Let $\{e_i^{(1)}\}_{i \geq 1}$ be the edges of G_1 and $\{e_i^{(2)}\}_{i \geq 1}$ be the edges of G_2 , ordered as above. Define an equivalence relation on the probability space Ω by setting $G_1 \equiv_k G_2$ when $\{e_i^{(1)}\}_{i=1}^k = \{e_i^{(2)}\}_{i=1}^k$, i.e., the first k pairings match. Let Ω_k be the set of equivalence classes, and \mathcal{F}_k the corresponding σ -algebra with $\mathcal{F}_0 = \mathcal{F}$. Consider a bounded measurable $f: \mathcal{G}(\vec{d}) \rightarrow \mathbb{R}$, and set $Y_k = \mathbb{E}[f \mid \mathcal{F}_k]$. Note that*

$$(Y_k)_{0 \leq k \leq m_1/2} \text{ is a Doob martingale with } \mathbb{E}[Y_k \mid \mathcal{F}_{k-1}] = Y_{k-1} \\ \text{and } Y_0 = \mathbb{E}[f], Y_{m_1/2} = f.$$

Define $Z_k = Y_k - Y_{k-1}$, and suppose that there exist functions $(g_k(\zeta))_{1 \leq k \leq \frac{1}{2}m_1}$ such that

$$\mathbb{E} \left[e^{\zeta^2 Z_k^2} \mid \mathcal{F}_{k-1} \right] \leq g_k(\zeta), \quad 1 \leq k \leq m_1/2.$$

Then, for all $t \geq 0$ and $\zeta > 0$,

$$\mathbb{P}(|f - \mathbb{E}[f]| \geq t) \leq 2e^{-\zeta t} \prod_{k=1}^{m_1/2} g_k(\zeta). \quad (3.8)$$

Remark 3.7. The condition on the existence $g_k(\zeta)$ can be rephrased as the existence of a random variable W_k that stochastically dominates Z_k on $(\Omega_{k-1}, \mathcal{F}_{k-1})$ (see [18]). \spadesuit

Lemma 3.8. [10, Lemma 15] *There exist constants $B_\ell > 0$, $\ell = 1, 2, 3$, depending only on the ratio $\frac{m_\infty}{m_0}$, such that*

$$\mathbb{P} \left(|X_\ell - \mathbb{E}[X_\ell]| \geq \frac{t}{m_\infty^{3/2}} \right) \leq 2e^{-tn + B_\ell n}, \quad \ell = 1, 2, 3. \quad (3.9)$$

Proof. While the properties in \mathcal{C}_3 allow us to apply standard martingale arguments to capture X_3 (see below), the properties in \mathcal{C}_1 and \mathcal{C}_2 force us to use Lemma 3.6 to capture X_1 and X_2 .

i. From the definition of \mathcal{C}_1 and \mathcal{C}_2 it follows that a bound on X_1 implies by symmetry a bound on X_2 (with, possibly, different constants). We will therefore focus on X_1 , the result for X_2 carrying through trivially. Without loss of generality we may reorder the indices in such a way that $|x_i| \geq |x_{i+1}|$ and for each $e_i = \{\alpha_i, \beta_i\}$ use the lexicographic order $\alpha_{i+1} > \alpha_i$ and $\beta_i > \alpha_i$ (i.e., we perform the pairing sequentially from left to right; see [27, Lemma 7.6]). It follows that $v(\alpha_i) \leq v(\alpha_{i+1})$ and $|x_{v(\alpha_i)}| \geq |x_{v(\alpha_{i+1})}|$. Define

$$\hat{\chi}(x, y) = \begin{cases} xy, & \text{if } |xy| < \frac{\sqrt{m_\infty}}{n} \text{ and } |x| > \frac{1}{\epsilon\sqrt{n}}, \\ 0, & \text{otherwise.} \end{cases}$$

For each $e = \{\alpha, \beta\}$ in the configuration \mathcal{C} we have

$$X_1 = \sum_{e \in \mathcal{C}} q(e)$$

with $q(e) = \hat{\chi}(x_{v(\alpha)}, y_{v(\beta)})/d_{v(\alpha)}d_{v(\beta)}$.

ii. Next, take $Y_k = \mathbb{E}[X_1 \mid \mathcal{F}_k]$, with $Y_0 = \mathbb{E}[X_1]$ and $Y_m = X_1$. Then $(Y_k)_{k \in \mathbb{N}_0}$ is a Doob martingale and, by the definition of the configuration model, we can write $Z_k = Y_k - Y_{k-1}$ as

$$Z_k(\mathcal{C}) = \frac{2^{\frac{1}{2}m_1 - k} (\frac{1}{2}m_1 - k)!}{(m_1 - 2k)!} \left(\sum_{\mathcal{C}' \equiv_k \mathcal{C}} X_1(\mathcal{C}') - \frac{1}{m_1 - 2k + 1} \sum_{\mathcal{C}'' \equiv_{k-1} \mathcal{C}} X_1(\mathcal{C}'') \right).$$

Now that we have an expression for Z_k , we can use the method of switching (see, for example, [31]). Indeed, given a $\mathcal{C}' \equiv_k \mathcal{C}$, we can define a quantity \mathcal{C}'_η as follows. Given the first k

pairings of \mathcal{C} , let I be the set of points already paired, and let $\{\alpha, \beta\}$ be the k -th pair. Put $\eta \notin I - \{\beta\}$ and $\{\eta, \gamma\} \in \mathcal{C}'$. Then \mathcal{C}'_η is the pairing obtained from \mathcal{C}' by mapping

$$\{\alpha, \beta\}, \{\eta, \gamma\} \rightarrow \{\alpha, \eta\}, \{\gamma, \beta\}.$$

Is easy to see that $\mathcal{C}'_\eta \equiv_{k-1} \mathcal{C}$ and that $\{\{\mathcal{C}'_\eta: \eta \notin I - \{\beta\}\} \mid \mathcal{C}' \equiv_k \mathcal{C}\}$ is a partition of $\{\mathcal{C}'' \mid \mathcal{C}'' \equiv_{k-1} \mathcal{C}\}$. We can therefore rewrite

$$\begin{aligned} Z_k(\mathcal{C}) &= \frac{2^{\frac{1}{2}m_1-k} (\frac{1}{2}m_1 - k)!}{(m_1 - 2k)!} \sum_{\mathcal{C}' \equiv_k \mathcal{C}} \sum_{\eta \notin I} (X_1(\mathcal{C}') - X_1(\mathcal{C}'_\eta)) \\ &= \sum_{\eta \notin I} \sum_{\gamma \notin I, \gamma \neq \eta} \frac{q(\{\alpha, \beta\}) + q(\{\eta, \gamma\}) - q(\{\alpha, \eta\}) - q(\{\gamma, \beta\})}{(2m - 2k + 1)(2m - 2k - 1)}. \end{aligned}$$

Because $\sum_i x_u^2 \leq 1$, there are at most $\varepsilon^2 n$ indices of x such that $|x_i| > 1/(\varepsilon\sqrt{n})$. By the definition of X_1 , the lexicographic ordering of $\{\alpha_i, \beta_i\}$ and the ordering of $|x_j| > |x_{j+1}|$, there exists a \tilde{k} such that $Z_{\tilde{k}} = 0$. Take $\tilde{k} = \varepsilon^2 m_\infty n$. For $k > \tilde{k}$, we have that

$$m_1 - 2k - 1 \geq m_1 - 2\varepsilon^2 m_\infty n - 1 \geq m_0 n,$$

where the free parameter ε has to be fixed such that the last inequality holds. (Note that, because there exists a constant θ such that $\frac{m_\infty}{m_0} < \theta$, we can always choose an ε small enough so that this holds). Hence we can bound

$$|Z_k(\mathcal{C})| \leq \frac{1}{(m_0 n)^2} \sum_{\eta \notin I} \sum_{\gamma \notin I, \gamma \neq \eta} (|q(\{\alpha, \beta\})| + |q(\{\eta, \gamma\})| + |q(\{\alpha, \eta\})| + |q(\{\gamma, \beta\})|).$$

iii. Define

$$y^\alpha = \frac{1}{|x_{v(\alpha)}|} \min \left\{ |y_{x_{v(\alpha)}}|, \frac{\sqrt{m_\infty}}{n} \right\}.$$

Because $\alpha < \beta$, we can bound

$$q(\{\alpha, \beta\}) = \frac{\hat{\chi}(x_{v(\alpha)}, y_{v(\beta)})}{d_{v(\alpha)} d_{v(\beta)}} \leq \frac{y_{v(\beta)}^\alpha |x_{v(\alpha)}|}{m_0^2}.$$

A similar bound holds for $\{\alpha, \eta\}$ for the same reason. For the other two edges, $\{\eta, \gamma\}$ and $\{\gamma, \beta\}$, we need to upper bound with a symmetric term, because we do not know whether $\gamma > \eta$ or $\gamma < \eta$. Thus, we have the upper bound

$$q(\{\eta, \gamma\}) \leq \frac{1}{m_0^2} \left(y_{v(\eta)}^\gamma |x_{v(\gamma)}| + y_{v(\gamma)}^\eta |x_{v(\eta)}| \right)$$

(the same bound holds for $\{\gamma, \beta\}$). Moreover, by the lexicographic order, $x_{v(\alpha)}$ bounds all the other components, and therefore $y_{v(\beta)}^\alpha |x_{v(\alpha)}| \leq y_{v(\beta)}^\alpha |x_{v(\alpha)}|$. Now note that $\sum_i |y_i| \leq \sqrt{n}$ (because $\sum_i y_i^2 \leq 1$), and so

$$\sum_{\eta \notin I} y_{v(\eta)}^\alpha \leq \sum_{\eta \notin I} |y_{v(\eta)}| \leq m_\infty \sum_i |y_i| \leq m_\infty \sqrt{n}.$$

By the previous considerations, substituting into the expression for $Z(\mathcal{C})$, we have

$$\begin{aligned} |Z_k(\mathcal{C})| &\leq \frac{1}{m_0^2} \left(y_{v(\beta)}^\alpha |x_{v(\alpha)}| + \left(y_{v(\eta)}^\alpha |x_{v(\alpha)}| + y_{v(\gamma)}^\alpha |x_{v(\alpha)}| \right) + y_{v(\eta)}^\alpha |x_{v(\alpha)}| + \left(y_{v(\beta)}^\alpha |x_{v(\alpha)}| + y_{v(\gamma)}^\alpha |x_{v(\alpha)}| \right) \right) \\ &\leq \frac{4m_\infty^2}{m_0^4} |x_{v(\alpha)}| \left(y_{v(\beta)}^\alpha + \frac{1}{\sqrt{n}} \right). \end{aligned}$$

iv. From the above bounds we are able to obtain an upper bound for $\mathbb{E}[\exp(\zeta^2 Z_k^2) \mid \mathcal{F}_{k-1}]$ and then use lemma 3.6. Indeed,

$$\mathbb{E} \left[e^{\zeta^2 Z_k^2} \mid \mathcal{F}_{k-1} \right] \leq \frac{1}{m_1 - 2k - 1} \sum_{\omega \notin I \setminus \beta} \exp \left[16 \zeta^2 m_\infty^4 m_0^{-8} (x_{v(\alpha)})^2 \right] \left(y_{v(\omega)}^\alpha + \frac{1}{\sqrt{n}} \right)^2.$$

Looking at (3.8), we see that we have to fix $\zeta = m_\infty^{3/2} n$ in order to achieve the required bound. Using that $x_{v(\alpha)} \leq \sqrt{m_\infty}/(\epsilon\sqrt{n})$ (because $x_{v(\alpha)} y_{v(\beta)}^\alpha \leq \sqrt{m_\infty} n$ and $y_{v(\beta)}^\alpha \geq \epsilon/\sqrt{n}$), the exponent in the previous display is bounded by $64\theta^8 \epsilon^{-2}$. Using that $e^x \leq 1 + xe^x$, $x \geq 0$, and putting $B = 64\theta^8 \epsilon^{-2}$, we have

$$\begin{aligned} \mathbb{E} \left[e^{\zeta^2 Z_k^2} \mid \mathcal{F}_{k-1} \right] &\leq 1 + \frac{B}{m_1 - 2k - 1} \sum_{\omega \notin I \setminus \beta} \zeta^2 m_\infty^4 m_0^{-8} (x_{v(\alpha)})^2 \left(y_{v(\omega)}^\alpha + \frac{1}{\sqrt{n}} \right)^2 \\ &\leq 1 + B \zeta^2 m_\infty^4 m_0^{-9} (x_{v(\alpha)})^2 \sum_{\omega} \left(y_{v(\omega)}^2 + 2y_{v(\omega)} \frac{1}{\sqrt{n}} + \frac{1}{n} \right) \\ &\leq \exp \left[4B \frac{\zeta^2 m_\infty^5}{m_0^9 n} (x_{v(\alpha)})^2 \right]. \end{aligned}$$

Next, let us pick an index $i(k)$ such that, for all $C \in \Omega$,

$$\mathbb{E} \left[e^{\zeta^2 Z_k^2} \mid \mathcal{F}_{k-1} \right] \leq \exp \left[4B \frac{\zeta^2 m_\infty^5}{m_0^9 n} (x_{i(k)})^2 \right].$$

One possible choice is to take $i(k) = \lceil k/m_\infty \rceil$. We finally get

$$\mathbb{P} \left(\left| X_1 - \mathbb{E}[X_1] \right| \geq \frac{t}{m_\infty^{3/2}} \right) \leq 2e^{-\frac{\zeta t}{m_\infty^{3/2}}} e^{-\sum_{k=1}^{\frac{1}{2}m_1} 4B \frac{\zeta^2 m_\infty^5}{m_0^9 n} (x_{i(k)})^2} \leq 2e^{-tn + 4B \frac{m_\infty^9}{m_0^9} n},$$

which proves what was said at the beginning of the proof (3.9) for $\ell = 1, 2$.

v. Finally, consider X_3 . In view of the bounds in \mathcal{C}_3 , this case can be dealt with via classical martingale arguments (see, for example, the McDiarmid inequality and its generalizations in [8]). Considering the variables $Y_k = \mathbb{E}[X_3 \mid \mathcal{C}_3]$, we have that $|Y_k - Y_{k-1}| \leq 4/(\epsilon^2 n m_0^2)$. Thus, given our choice of $\epsilon = \epsilon(\theta)$ being constant, we have

$$\mathbb{P} \left(\left| X_3 - \mathbb{E}[X_3] \right| \geq t m_\infty^{-3/2} \right) \leq 2e^{-\frac{t^2 \epsilon^4 n^2 m_0^4}{16 m_\infty^3 m_1}} \leq 2e^{-C(\epsilon, \theta) t^2 n},$$

where $C(\epsilon, \theta) > 0$. □

4. Combining the results for $\ell = 1, 2, 3$ in the above lemma, we get an exponential bound on X'_a of the type

$$\mathbb{P} \left(\left| X'_a - \mathbb{E}[X'_a] \right| \geq \frac{t}{m_\infty^{3/2}} \right) \leq 2e^{-tn + B_a n},$$

where B_a is a suitable constant, possibly different from any of the B_ℓ . By symmetry, the same holds for X'_b , which proves the concentration around $m_\infty^{-3/2}$ of X' .

Remark 3.9. Up to now we have worked with multi-graphs, so we have to pass to simple graphs with a prescribed degree sequence. It is easy to see that, because of (2.3), we can express the final result as saying that, conditional on the event that the graph is simple, there exist two constants $\hat{\xi} < \xi$ in $(0, 1)$ and a constant $K(\theta, \xi) > 0$ such that

$$\mathbb{P} \left(\left| X' - \mathbb{E}[X'] \right| \geq K m_\infty^{-3/2} \right) \leq e^{O(d^2) \hat{\xi}^n} \leq \xi^n.$$



3.3. Estimate of second contribution. It remains to show that the pairs with $x, y \notin \mathcal{B}$ give a bounded contribution of the order $O(m_\infty^{-3/2})$ with sufficiently high probability. For this it suffices to show that a simple random graph G with a prescribed degree sequence \vec{d} cannot have too dense subgraphs (see [18, Lemma 2.5] and [10, Lemma 16]):

Lemma 3.10. *Let G be a simple random graph of size n drawn uniformly at random with a given degree sequence \vec{d} . Let $A, B \subseteq [n]$ be two subsets of the vertex set, and let $e(A, B)$ be the set of edges $e = \{\alpha, \beta\}$ such that either $\alpha \in A, \beta \in B$ or $\alpha \in B, \beta \in A$. Since $\mu(A, B) = \theta|A||B|^{\frac{m_\infty}{n}}$ with $\theta > m_\infty/m_0$ a sufficiently large constant, for any $K > 0$ there exist a constant $C = C(\theta, K)$ such that with probability $1 - o(n^{-K})$ any pair A, B with $|A| \leq |B|$ satisfies at least one of the following:*

$$e(A, B) \leq C\mu(A, B) \tag{3.10}$$

$$e(A, B) \log \left(\frac{e(A, B)}{\mu(A, B)} \right) \leq C|B| \log \left(\frac{n}{|B|} \right). \tag{3.11}$$

The above lemma has the following corollary.

Corollary 3.11. *For all $x, y \in T$,*

$$X'' = O \left(\frac{1}{m_\infty^{3/2}} \right) \quad \text{with probability at least } 1 - O(n^{-K}),$$

where K is the constant in Lemma 3.10.

Proof. Fix $x, y \in T$ and define

$$S_i(x) = \left\{ \ell: \frac{\varepsilon^{2-i}}{\sqrt{n}} \leq |x_\ell| < \frac{\varepsilon^{1-i}}{\sqrt{n}} \right\}, \quad i \in I,$$

$$S_j(y) = \left\{ \ell: \frac{\varepsilon^{2-i}}{\sqrt{n}} \leq |y_\ell| < \frac{\varepsilon^{1-i}}{\sqrt{n}} \right\}, \quad j \in J,$$

where $I = \{i | S_i(x) \neq \emptyset\}$ (J is defined similarly). Then, for $x \in T$, write

$$x_u|_S = \begin{cases} x_u, & \text{if } u \in S, \\ 0, & \text{otherwise.} \end{cases}$$

In order to apply Lemma 3.10, define $A_i = S_i(x)$ and $B_j = S_j(y)$, and let a_i and b_j be their cardinality, respectively. Divide the set of indices into two groups:

$$\mathcal{E} = \{(i, j): i, j > 0, \varepsilon^{2-i-j} > \sqrt{m_\infty}, a_i \leq b_j\},$$

$$\mathcal{E}' = \{(i, j): i, j > 0, \varepsilon^{2-i-j} > \sqrt{m_\infty}, a_i > b_j\}.$$

By the definition of X'' and the set \mathcal{B} , we have

$$X'' = \sum_{x_i y_j > \sqrt{m_\infty}/n} x_i A_{ij} y_j = \sum_{(i,j) \in \mathcal{E}} (x|_{A_i})^t A y|_{B_j} + \sum_{(i,j) \in \mathcal{E}'} (x|_{A_i})^t A y|_{B_j}.$$

It suffices to show that either of the contributions coming from \mathcal{E} or \mathcal{E}' are $O(m_\infty^{-3/2})$ (the other will follow by symmetry). Focus on \mathcal{E} . Putting $e_{i,j} = e(A_i, B_j)$ and $\mu_{i,j} = \mu(A_i, B_j)$, we see that the bound can be rewritten as

$$\frac{1}{n} \sum_{(i,j) \in \mathcal{E}} \frac{e_{i,j}}{\varepsilon^{i+j}} = O(\sqrt{m_\infty}).$$

Divide \mathcal{E} into the union of \mathcal{E}_a and \mathcal{E}_b , where \mathcal{E}_a satisfies (3.10) and \mathcal{E}_b satisfies (3.11). Clearly $\mathcal{E} = \mathcal{E}_a \cup \mathcal{E}_b$ and

$$\frac{1}{n} \sum_{(i,j) \in \mathcal{E}} \frac{e_{i,j}}{\varepsilon^{i+j}} \leq \frac{1}{n} \sum_{(i,j) \in \mathcal{E}_a} \frac{e_{i,j}}{\varepsilon^{i+j}} + \frac{1}{n} \sum_{(i,j) \in \mathcal{E}_b} \frac{e_{i,j}}{\varepsilon^{i+j}}.$$

If we are able to show that both contributions from \mathcal{E}_a and \mathcal{E}_b are $O(\sqrt{m_\infty})$, then the theorem follows. It is easy to see that \mathcal{E}_a gives a bounded contribution. Indeed,

$$\frac{1}{n} \sum_{(i,j) \in \mathcal{E}_a} \frac{e_{i,j}}{\varepsilon^{i+j}} \leq \frac{1}{n^2} \sum_{(i,j) \in \mathcal{E}_a} \frac{Ca_i b_j \theta m_\infty}{\varepsilon^{i+j}} \leq \frac{C' m_\infty}{n^2} \sum_{(i,j) \in \mathcal{E}_a} \frac{a_i b_j}{\varepsilon^{2(i+j)} \sqrt{m_\infty}} = O(\sqrt{m_\infty}),$$

where in the last step we use that, because $\sum_i x_i^2 \leq 1$,

$$\sum_{i \in I} \frac{a_i}{\varepsilon^{2(i-2)}} \leq n, \quad \sum_{j \in J} \frac{b_j}{\varepsilon^{2(i-2)}} \leq n.$$

It remains to show that

$$\frac{1}{n} \sum_{(i,j) \in \mathcal{E}_b} \frac{e_{i,j}}{\varepsilon^{i+j}} = O(\sqrt{m_\infty}).$$

In order to do so, we divide \mathcal{E}_b into subsets $\mathcal{E}_b^{(\ell)}$, $\ell = 1, \dots, 5$, having the following properties:

- (1) $\varepsilon^j \geq \varepsilon^i \sqrt{m_\infty}$.
- (2) $e_{i,j} \leq \frac{\mu_{i,j}}{\varepsilon^{i+j} \sqrt{m_\infty}}$.
- (3) $\log \left(\frac{e_{i,j}}{\mu_{i,j}} \right) \geq \frac{1}{4} \log \left(\frac{n}{b_j} \right)$.
- (4) $\frac{n}{b_j} \leq e^{-4j}$.
- (5) $\frac{n}{b_j} > e^{-4j}$.

For $j > i$ we have that $\mathcal{E}_b^{(\ell)} \not\subseteq \mathcal{E}_b^{(\ell)}$ and $\mathcal{E}_b = \cup_\ell \mathcal{E}_b^{(\ell)}$. Thus it suffices to show a bound of $O(\sqrt{n})$ for each of the quantities $S_\ell = 1/n \sum_{\mathcal{E}_b^{(\ell)}} e_{i,j}/\varepsilon^{i+j}$.

- For $\ell = 1$ we note that, since $e_{i,j} \leq a_i m_\infty$,

$$S_1 \leq \frac{1}{n} \sum_i \sum_{j | \varepsilon^j \geq \varepsilon^i \sqrt{m_\infty}} \frac{a_i m_\infty}{\varepsilon^{i+j}} = bO \left(\frac{1}{n} \sum_i \frac{a_i \sqrt{m_\infty}}{\varepsilon^{2i}} \right) = O(\sqrt{m_\infty}).$$

- For $\ell = 2$ we obtain

$$S_2 \leq \frac{1}{n} \sum_{ij} \frac{\mu_{ij}}{\varepsilon^{2(i+j)} \sqrt{m_\infty}} = O \left(\frac{\sqrt{m_\infty}}{n^2} \sum_{ij} \frac{a_i b_j}{\varepsilon^{2(i+j)}} \right) = O(\sqrt{n}).$$

- For $\ell = 3$, because the pairs $(i, j) \in \mathcal{E}_b$ have property (3.11), it follows easily that $e_{ij} = O(b_j)$. Furthermore, because $\mathcal{E}_b^{(3)} \not\subseteq \mathcal{E}_b^{(1)}$, we have that $\forall (i, j) \in \mathcal{E}_b^{(3)}$, $\varepsilon^j < \varepsilon^i \sqrt{m_\infty}$. It follows that

$$S_3 = O \left(\frac{1}{n} \sum_j \sum_{i | \varepsilon^i > \varepsilon^j / \sqrt{m_\infty}} \frac{b_j}{\varepsilon^{i+j}} \right) = O \left(\frac{1}{n} \sum_j \frac{\sqrt{m_\infty} b_j}{\varepsilon^{2j}} \right) = O(\sqrt{m_\infty}).$$

- For $\ell = 4$ we take advantage of the fact that $(i, j) \in \mathcal{E}_b^{(4)}$ do not belong to $\mathcal{E}_b^{(3)}$ and $\mathcal{E}_b^{(2)}$. This implies that

$$\frac{e_{ij}}{m_{ij}} \leq \frac{1}{e^j} \quad \frac{e_{ij}}{\mu_{ij}} \geq \frac{1}{\varepsilon^{i+j}\sqrt{m_\infty}}.$$

Hence we have $\varepsilon^{-i} \leq \sqrt{m_\infty}$ and, by (3.11), also $e_{ij} = O(jb_j)$. We can therefore conclude that

$$S_4 = O\left(\frac{1}{n} \sum_j \sum_{i|\varepsilon^{-i} \leq \sqrt{m_\infty}} \frac{jb_j}{\varepsilon^{i+j}}\right) = O\left(\frac{\sqrt{m_\infty}}{n} \sum_j \frac{jb_j}{\varepsilon^j}\right) = O(\sqrt{m_\infty}),$$

where in the last equality we use that $\sum_{j \in J} b_j / (n\varepsilon^2) = O(1)$.

- For $\ell = 5$, using the property in (5) and (3.11), we have

$$e_{ij} \leq Cn\varepsilon^{4j} \log \varepsilon^{-4j} = O(nj\varepsilon^{4j}).$$

Also, using that $\mathcal{E}_b^{(4)} \not\subseteq \mathcal{E}_b^{(4)}$, we have $\varepsilon^j < \varepsilon^i \sqrt{m_\infty}$, from which we conclude that

$$S_5 = O\left(\sum_j \sum_{i|\varepsilon^i > \varepsilon^j / \sqrt{m_\infty}} j\varepsilon^{3j-i}\right) = O\left(\sqrt{m_\infty} \sum_j j\varepsilon^{2j}\right) = O(\sqrt{m_\infty}).$$

This completes the proof. \square

4. PROOF OF THE MAIN THEOREM

Expansion. Throughout this section we abbreviate $A = A_{G_n}$ and condition on the event that G_n is simple. Recall Lemma 3.3. Compute

$$\begin{aligned} Av_1 &= \lambda_1 v_1, \\ (H + |\tilde{e}\rangle\langle\tilde{e}| + (\mathbb{E}[A] - |\tilde{e}\rangle\langle\tilde{e}|)) v_1 &= \lambda_1 v_1, \\ \left(H + |\tilde{e}\rangle\langle\tilde{e}| - \text{diag}\left(\frac{d_1}{m_1 - 1}, \dots, \frac{d_n}{m_1 - 1}\right)\right) v_1 &= \lambda_1 v_1. \end{aligned}$$

Rewriting the equation we have,

$$\langle\tilde{e}, v_1\rangle\tilde{e} = \left(\lambda_1 \mathbb{I} + \text{diag}\left(\frac{d_1}{m_1 - 1}, \dots, \frac{d_n}{m_1 - 1}\right) - H\right) v_1.$$

Therefore componentwise we have the following inequality,

$$\left(\lambda_1 + \frac{m_0}{m_1 - 1}\right) \left(\mathbb{I} - \frac{H}{\lambda_1 + \frac{m_0}{m_1 - 1}}\right) v_1 \leq \langle\tilde{e}, v_1\rangle\tilde{e} \leq \left(\lambda_1 + \frac{m_\infty}{m_1 - 1}\right) \left(\mathbb{I} - \frac{H}{\lambda_1 + \frac{m_\infty}{m_1 - 1}}\right) v_1,$$

If x, y and e are non-negative vector (the non-negativity of v_1 follows from Perron-Frobenius theory) with $x \geq y$, then $\langle e, x \rangle \geq \langle e, y \rangle$, we can use Corollary 3.2 and invert the matrix multiplying v_1 . Indeed, given that $\|H\| = O(\sqrt{m_\infty})$ and $\lambda_1 \sim m_2/m_1$ on the event with probability $1 - o(n^{-K})$ of Corollary 3.2 and Theorem 3.1, we can invert and expand

$$\left(\mathbb{I} - \frac{H}{\lambda_1 + \frac{m_0}{m_1 - 1}}\right)^{-1} = \sum_{k \in \mathbb{N}_0} \frac{H^k}{\left(\lambda_1 + \frac{m_0}{m_1 - 1}\right)^k},$$

and similarly for m_∞ . Thus,

$$\begin{aligned} \lambda_1 &\leq \frac{m_2}{m_1 - 1} - \frac{m_0}{m_1 - 1} + \sum_{k \in \mathbb{N}} \frac{\langle\tilde{e}, H^k \tilde{e}\rangle}{\left(\lambda_1 + \frac{m_0}{m_1 - 1}\right)^k} \\ \lambda_1 &\geq \frac{m_2}{m_1 - 1} - \frac{m_\infty}{m_1 - 1} + \sum_{k \in \mathbb{N}} \frac{\langle\tilde{e}, H^k \tilde{e}\rangle}{\left(\lambda_1 + \frac{m_\infty}{m_1 - 1}\right)^k}. \end{aligned}$$

Our final goal is to determine the expectation $\mathbb{E}[\lambda_1]$, which splits as

$$\mathbb{E}[\lambda_1] = \mathbb{E}[\lambda_1|\mathcal{E}] \mathbb{P}(\mathcal{E}) + \mathbb{E}[\lambda_1|\mathcal{E}^c] \mathbb{P}(\mathcal{E}^c). \quad (4.1)$$

The event \mathcal{E}^c has probability at most n^{-K} , where K is a large arbitrary constant. Thus, given the deterministic bound $\lambda_1 \leq n$, we may focus on $\mathbb{E}[\lambda_1|\mathcal{E}] \mathbb{P}(\mathcal{E})$. In order to do this, we need to be able to handle terms of the type

$$\frac{m_2}{m_1 - 1} + \frac{\mathbb{E}[\langle \tilde{e}, H \tilde{e} \rangle]}{\frac{m_2}{m_1 - 1} (1 + o(1))} + \frac{\mathbb{E}[\langle \tilde{e}, H^2 \tilde{e} \rangle]}{\left(\frac{m_2}{m_1 - 1}\right)^2 (1 + o(1))} + \sum_{k \in \mathbb{N} \setminus \{1, 2\}} \frac{\mathbb{E}[\langle \tilde{e}, H^k \tilde{e} \rangle]}{\left(\frac{m_2}{m_1 - 1}\right)^k (1 + o(1))}$$

Since

$$\frac{\mathbb{E}[\langle \tilde{e}, H^k \tilde{e} \rangle]}{\left(\frac{m_2}{m_1 - 1}\right)^k (1 + o(1))} \leq \frac{m_\infty^{k/2}}{\left(\frac{m_2}{m_1 - 1}\right)^{k-1}} = o\left(\frac{1}{m_0^{k/2-1}}\right),$$

the last sum is $o(1/\sqrt{m_\infty})$, which is an error term. It therefore remains to study $\langle \tilde{e}, H^k \tilde{e} \rangle$, $k = 1, 2$. The study of these moments for the configuration model is more involved than for random regular graphs.

Case $k = 1$. Compute

$$\begin{aligned} \langle \tilde{e}, H \tilde{e} \rangle &= \langle \tilde{e}, (A - \mathbb{E}[A]) \tilde{e} \rangle \\ &= \frac{1}{m_1 - 1} \left(\sum_{ij} d_i d_j a_{ij} - \frac{1}{m_1 - 1} \sum_{ij} d_i^2 d_j^2 \right) = \frac{\sum_j d_j (\sum_{i \sim j} d_i)}{m_1 - 1} - \frac{m_2^2}{(m_1 - 1)^2}. \end{aligned}$$

Since

$$\mathbb{E} \left[\sum_j d_j \sum_{i \sim j} d_i \right] = \sum_{ij} d_j d_i \mathbb{E} a_{ij} = \frac{m_2^2}{m_1 - 1} - \frac{m_3}{m_1 - 1},$$

where the last term comes from the presence of selfloops. It follows that $\mathbb{E}[\langle \tilde{e}, H \tilde{e} \rangle] = O(1/\sqrt{n})$.

Case $k = 2$. Compute

$$\begin{aligned} \langle \tilde{e}, H^2 \tilde{e} \rangle &= \langle \tilde{e}, (A - \mathbb{E}[A])^2 \tilde{e} \rangle \\ &= \frac{1}{m_1 - 1} \left(\sum_{ijk} d_i d_k a_{ij} a_{jk} - \frac{1}{m_1 - 1} \sum_{ijk} d_i d_k^2 a_{ij} d_j \right. \\ &\quad \left. - \frac{1}{m_1 - 1} \sum_{ijk} d_i^2 d_j d_k a_{jk} + \frac{1}{(m_1 - 1)^2} \sum_{ijk} d_i^2 d_j^2 d_k^2 \right). \end{aligned}$$

Write

$$\frac{1}{m_1 - 1} \sum_{ijk} d_i d_k^2 a_{ij} d_j = \frac{m_2}{m_1 - 1} \sum_{ijk} d_i a_{ij} d_j = \frac{m_2}{m_1 - 1} \sum_i d_i \left(\sum_{i \sim j} d_j \right)$$

(by symmetry the third term is equal) and

$$\sum_{ijk} d_i d_k a_{ij} a_{jk} = \sum_j \left(\sum_{i \sim j} d_i \right)^2 = \sum_j \sum_{i \sim j} d_i^2 + \sum_k \sum_{\substack{i, j \sim k \\ i \neq j}} d_i d_j = m_3 + \sum_k \sum_{\substack{i, j \sim k \\ i \neq j}} d_i d_j.$$

Indeed in $\sum_j \sum_{i \sim j} d_i^2$, the summand d_i^2 appears exactly d_i times, because the node i has exactly d_i neighbours, and so $\sum_j \sum_{i \sim j} d_i^2 = m_3$. Putting the terms together, we get

$$\langle \tilde{\epsilon}, H^2 \tilde{\epsilon} \rangle = \frac{1}{m_1 - 1} \left(m_3 + \sum_k \sum_{\substack{i, j \sim k \\ i \neq j}} d_i d_j - 2 \frac{m_2}{m_1 - 1} \sum_i d_i \left(\sum_{i \sim j} d_j \right) + \frac{m_2^3}{(m_1 - 1)^2} \right).$$

Taking expectations, we get

$$\mathbb{E}[\langle \tilde{\epsilon}, H^2 \tilde{\epsilon} \rangle] = \frac{1}{m_1 - 1} \left(m_3 + \mathbb{E} \left[\sum_k \sum_{\substack{i, j \sim k \\ i \neq j}} d_i d_j - \frac{m_2^3}{(m_1 - 1)^2} \right] \right).$$

Note that $\mathbb{E}[\sum_k \sum_{i, j \sim k, i \neq j} d_i d_j]$ is a sum over the wedges centered at vertex k , summed all k . We can swap the summation over pairs of vertices, and choose a third neighbour to form a wedge, which gives

$$\sum_k \sum_{\substack{i, j \sim k \\ i \neq j}} d_i d_j = \sum_{\substack{i, j \\ i \neq j}} d_i d_j \sum_k \mathbb{1}_{(k \sim i, k \sim j)}.$$

Compute

$$\begin{aligned} \mathbb{E} \left[\sum_{\substack{i, j \\ i \neq j}} d_i d_j \sum_k \mathbb{1}_{(k \sim i, k \sim j)} \right] &= \sum_{\substack{i, j \\ i \neq j}} d_i d_j \sum_k \mathbb{E} [\mathbb{1}_{(k \sim i, k \sim j)}] \\ &= \sum_{\substack{i, j \\ i \neq j}} d_i d_j \sum_k \left(\frac{d_i d_j d_k (d_k - 1)}{(m_1 - 1)(m_1 - 2)} \mathbb{1}_{k \neq i, k \neq j} + \frac{d_i d_j (d_k - 1)(d_k - 2)}{(m_1 - 1)(m_1 - 2)} \mathbb{1}_{k=i \text{ or } k=j} \right) \\ &= \sum_{\substack{i, j \\ i \neq j}} d_i d_j \sum_k \left(\frac{d_i d_j d_k (d_k - 1)}{(m_1 - 1)(m_1 - 2)} - \frac{2 d_i d_j (d_k - 1)}{(m_1 - 1)(m_1 - 2)} \mathbb{1}_{k=i \text{ or } k=j} \right) \\ &= \frac{1}{(m_1 - 1)(m_1 - 2)} \sum_{\substack{i, j \\ i \neq j}} d_i^2 d_j^2 \sum_k (d_k (d_k - 1) - 2(d_k - 1) \mathbb{1}_{k=i \text{ or } k=j}) \\ &= \frac{1}{(m_1 - 1)(m_1 - 2)} \left((m_2 - m_1) \left(\sum_{i, j} d_i^2 d_j^2 - \sum_i d_i^4 \right) - 2 \sum_{\substack{i, j \\ i \neq j}} d_i^2 d_j^3 - 2 \sum_{\substack{i, j \\ i \neq j}} d_i^3 d_j^2 + 4 \sum_{\substack{i, j \\ i \neq j}} d_i^2 d_j^2 \right) \\ &= \frac{1}{(m_1 - 1)(m_1 - 2)} \left((m_2 - m_1) (m_2^2 - m_4) - 4 \left(\sum_{i, j} d_i^2 d_j^3 - \sum_i d_i^5 \right) + 4 \left(\sum_{i, j} d_i^2 d_j^2 - \sum_i d_i^4 \right) \right) \\ &= \frac{1}{(m_1 - 1)(m_1 - 2)} ((m_2 - m_1) (m_2^2 - m_4) - 4(m_3 m_2 - m_5) + 4(m_2^2 - m_4)) \\ &= \frac{m_2^3 - m_2 m_4 - m_1 m_2^2 + m_1 m_4 - 4 m_3 m_2 + 4 m_5 + 4 m_2^2 - 4 m_4}{(m_1 - 1)(m_1 - 2)}. \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{E}[\langle \tilde{\epsilon}, H^2 \tilde{\epsilon} \rangle] &= \frac{1}{m_1 - 1} \left(m_3 + \frac{m_2^3 - m_2 m_4 - m_1 m_2^2 + m_1 m_4 - 4 m_3 m_2 + 4 m_5 + 4 m_2^2 - 4 m_4}{(m_1 - 1)(m_1 - 2)} - \frac{m_2^3}{(m_1 - 1)^2} \right). \end{aligned}$$

Given the event \mathcal{E} , by (4.1) and Corollary 3.2, we have that $\mathbb{E}[\lambda_1]$ concentrates around m_2/m_1 with an $O(\sqrt{m_\infty})$ error, and so we can write

$$\mathbb{E} \left[\frac{\langle \tilde{e}, H^2 \tilde{e} \rangle}{\lambda_1^2} \right] = \frac{\mathbb{E} [\langle \tilde{e}, H^2 \tilde{e} \rangle]}{\left(\frac{m_2}{m_1-1}\right)^2} (1 + o(1)) + o(1).$$

We run through the various contributions separately (using Assumption 1.1). Noting that $nm_0^k \leq m_k \leq nm_\infty^k$ and that there are positive constants c, C such that $c \leq \frac{m_\infty}{m_0} \leq C$, we have

$$\begin{aligned} \frac{m_2^3}{m_2^2(m_1-1)(m_1-2)} &= \Theta\left(\frac{1}{n}\right), \\ \frac{m_2 m_4}{m_2^2(m_1-2)} &= o\left(\frac{1}{\sqrt{n}}\right), \quad \frac{m_1 m_4}{m_2^2(m_1-2)} = \Theta\left(\frac{1}{n}\right), \quad \frac{4m_3 m_2}{m_2^2(m_1-2)} = o\left(\frac{1}{n}\right), \\ \frac{4m_5}{m_2^2(m_1-2)} &= o\left(\frac{1}{n^{3/2}}\right), \quad \frac{4m_2^2}{m_2^2(m_1-2)} = o\left(\frac{1}{n}\right), \quad \frac{4m_4}{m_2^2(m_1-2)} = o\left(\frac{1}{n^2}\right). \end{aligned}$$

Therefore

$$\frac{\mathbb{E} [\langle \tilde{e}, H^2 \tilde{e} \rangle]}{(m_2/(m_1-1))^2} = \frac{m_3 m_1}{m_2^2} - 1 + o\left(\frac{1}{\sqrt{n}}\right),$$

which settles (1.1).

Weak law of large numbers. We want to show that

$$\frac{\lambda_1}{\mathbb{E}\lambda_1} \rightarrow 1$$

in \mathbb{P} -probability. Using Corollary 3.2 and the Weyl interlacing inequality, we have that, with probability $1 - n^{-K}$ for $K > 0$,

$$\frac{m_2}{m_1} - O(\sqrt{m_\infty}) \leq \lambda_1 \leq \frac{m_2}{m_1} + O(\sqrt{m_\infty}).$$

By (1.1),

$$\frac{\frac{m_2}{m_1} - O(\sqrt{m_\infty})}{\frac{m_2}{m_1}(1 + o(1))} \leq \frac{\lambda_1}{\mathbb{E}[\lambda_1]} \leq \frac{\frac{m_2}{m_1} + O(\sqrt{m_\infty})}{\frac{m_2}{m_1}(1 + o(1))}.$$

It follows that

$$1 - O\left(\frac{1}{\sqrt{m_\infty}}\right) \leq \frac{\lambda_1}{\mathbb{E}[\lambda_1]} \leq 1 + O\left(\frac{1}{\sqrt{m_\infty}}\right)$$

with probability $1 - n^{-K}$, and hence the claim follows.

REFERENCES

- [1] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, June 1986.
- [2] O. Arizmendi E. and V. Pérez-Abreu. The \$\$\$-transform of symmetric probability measures with unbounded supports. *Proceedings of the American Mathematical Society*, 137(09):3057–3057, Sept. 2009.
- [3] R. Bauerschmidt, J. Huang, A. Knowles, and H.-T. Yau. Edge rigidity and universality of random regular graphs of intermediate degree. *Geom. Funct. Anal.*, 30(3):693–769, 2020.
- [4] H. Bercovici and D. Voiculescu. Free Convolution of Measures with Unbounded Support. *Indiana University Mathematics Journal*, 42(3), 1993.
- [5] C. Bordenave. A new proof of Friedman’s second eigenvalue Theorem and its extension to random lifts, Mar. 2019. arXiv:1502.04482 [math].
- [6] C. Bordenave. A new proof of Friedman’s second eigenvalue theorem and its extension to random lifts. *Ann. Sci. Éc. Norm. Supér. (4)*, 53(6):1393–1439, 2020.
- [7] C. Bordenave and M. Lelarge. Resolvent of large random graphs. *Random Structures Algorithms*, 37(3):332–352, 2010.

- [8] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities: a nonasymptotic theory of independence*. Oxford University Press, Oxford, 1st ed edition, 2013. OCLC: ocn818449985.
- [9] A. Broder and E. Shamir. On the second eigenvalue of random regular graphs. In *28th Annual Symposium on Foundations of Computer Science (sfcs 1987)*, pages 286–294, Los Angeles, CA, USA, Oct. 1987. IEEE.
- [10] A. Z. Broder, A. M. Frieze, S. Suen, and E. Upfal. Optimal Construction of Edge-Disjoint Paths in Random Graphs. *SIAM Journal on Computing*, 28(2):541–573, Jan. 1998.
- [11] A. Chakrabarty, R. S. Hazra, F. den Hollander, and M. Sfragara. Spectra of adjacency and Laplacian matrices of inhomogeneous Erdos-Renyi random graphs. *Random Matrices Theory Appl.*, 10(1):Paper No. 2150009, 34, 2021.
- [12] A. Dembo, E. Lubetzky, and Y. Zhang. Empirical Spectral Distributions of Sparse Random Graphs. In M. E. Vares, R. Fernández, L. R. Fontes, and C. M. Newman, editors, *In and Out of Equilibrium 3: Celebrating Vladas Sidoravicius*, volume 77, pages 319–345. Springer International Publishing, Cham, 2021. Series Title: Progress in Probability.
- [13] P. Dionigi, D. Garlaschelli, F. d. Hollander, and M. Mandjes. A spectral signature of breaking of ensemble equivalence for constrained random graphs. *Electronic Communications in Probability*, 26(none), Jan. 2021. arXiv:2009.05155 [math-ph].
- [14] P. Dionigi, D. Garlaschelli, R. Subhra Hazra, F. den Hollander, and M. Mandjes. Central limit theorem for the principal eigenvalue and eigenvector of Chung–Lu random graphs. *Journal of Physics: Complexity*, 4(1):015008, Mar. 2023.
- [15] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. *The Annals of Probability*, 41(3B), May 2013.
- [16] J. Friedman. A proof of Alon’s second eigenvalue conjecture and related problems, May 2004. arXiv:cs/0405020.
- [17] J. Friedman. A proof of Alon’s second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.*, 195(910):viii+100, 2008.
- [18] J. Friedman, J. Kahn, and E. Szemerédi. On the second eigenvalue of random regular graphs. In *Proceedings of the twenty-first annual ACM symposium on theory of computing*, STOC ’89, pages 587–598, New York, NY, USA, 1989. Association for Computing Machinery. Number of pages: 12 Place: Seattle, Washington, USA.
- [19] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1:233–241, 1981.
- [20] D. Garlaschelli, F. den Hollander, and A. Roccaverde. Covariance Structure Behind Breaking of Ensemble Equivalence in Random Graphs. *Journal of Statistical Physics*, 173(3-4):644–662, Nov. 2018.
- [21] F. Juhász. On the spectrum of a random graph. *Algebraic methods in graph theory*, 1, 1981.
- [22] B. D. McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra and its applications*, 40:203–216, 1981.
- [23] E. Shamir and J. Spencer. Sharp concentration of the chromatic number on random graphs $G(n, p)$. *Combinatorica*, 7(1):121–129, Mar. 1987.
- [24] T. Squartini, J. de Mol, F. den Hollander, and D. Garlaschelli. Breaking of ensemble equivalence in networks. *Physical review letters*, 115(26):268701, 2015.
- [25] H. Touchette. Equivalence and nonequivalence of ensembles: Thermodynamic, macrostate, and measure levels. *Journal of Statistical Physics*, 159(5):987–1016, 2015.
- [26] L. V. Tran, V. H. Vu, and K. Wang. Sparse random graphs: eigenvalues and eigenvectors. *Random Structures Algorithms*, 42(1):110–134, 2013.
- [27] R. v. d. van der Hofstad. *Random Graphs and Complex Networks*. Cambridge University Press, 1 edition, Nov. 2016.
- [28] R. v. d. van der Hofstad. *Random Graphs and Complex Networks vol 2*, volume 2. 2023.
- [29] D. Voiculescu. Multiplication of Certain Non-Commuting Random Variables. *Journal of Operator Theory*, 18(2):223–235, 1987.
- [30] V. H. Vu. Recent progress in combinatorial random matrix theory. *Probab. Surv.*, 18:179–200, 2021.
- [31] N. C. Wormald. Models of Random Regular Graphs. In J. D. Lamb and D. A. Preece, editors, *Surveys in Combinatorics, 1999*, pages 239–298. Cambridge University Press, 1 edition, July 1999.

MATHEMATICAL INSTITUTE, LEIDEN UNIVERSITY, P.O. Box 9512, 2300 RA LEIDEN, THE NETHERLANDS.

Email address: `p.dionigi@math.leidenuniv.nl`

LORENTZ INSTITUTE FOR THEORETICAL PHYSICS, LEIDEN UNIVERSITY, P.O. Box 9504, 2300 RA LEIDEN,
THE NETHERLANDS & IMT SCHOOL FOR ADVANCED STUDIES, PIAZZA S. FRANCESCO 19, 55100 LUCCA,
ITALY & INDAM-GNAMPA ISTITUTO NAZIONALE DI ALTA MATEMATICA, ITALY

Email address: `garlaschelli@lorentz.leidenuniv.nl`

MATHEMATICAL INSTITUTE, LEIDEN UNIVERSITY, P.O. Box 9512, 2300 RA LEIDEN, THE NETHERLANDS

Email address: `r.s.hazra@math.leidenuniv.nl`

Email address: `denholla@math.leidenuniv.nl`