SPECTRUM OF COMPLEX NETWORKS

DANIEL MONTEALEGRE AND VAN VU

ABSTRACT. The study of complex networks has been one of the most active fields in science in recent decades. Spectral properties of networks (or graphs that represent them) are of fundamental importance. Researchers have been investigating these properties for many years, and, based on numerical data, have raised a number of questions about the distribution of the eigenvalues and eigenvectors.

In this paper, we give the solution to some of these questions. In particular, we determine the limiting distribution of (the bulk of) the spectrum as the size of the network grows to infinity and show that the leading eigenvectors are strongly localized.

We focus on the preferential attachment graph, which is the most popular mathematical model for growing complex networks. Our analysis is, on the other hand, general and can be applied to other models.

1. INTRODUCTION

The study of complex networks has been an extremely active field in recent years (e.g., [2], [3], [4], [14], [18], [17], [22]). A fundamental goal of this study is to build and study mathematical models of networks that occur in nature.

A popular model in the field is the (random) preferential attachment model. In this model, the network grows continuously. At time $t = 1, 2, 3, \ldots$, a new vertex is born and attaches itself randomly to existing vertices, with respect to a distribution which favors vertices with higher degrees. This reflects the famous "rich get richer" phenomenon observed in real-life situations. As pointed out by many researchers, this model of random graphs is significantly different from the standard Erdős-Rényi model where all vertices play the same role. There is vast literature on this model and we refer the reader to the survey [5] as a starting point; see also [6, 1, 7, 19] and the references therein.

For any network (in fact for any data set which can be represented in matrix form), its spectral information is of fundamental interest. Based on numerical experiments, researchers have observed the following about the eigenvalues and eigenvectors of PA graphs:

Observation 1. The bulk of the spectrum has triangular shape. This is very different from the classical semi-circle law by Wigner which holds for Erdős-Rényi graph (e.g., [1], [10], [11], [20]). The edge eigenvalues seem to follow a power law. See Figure 1.
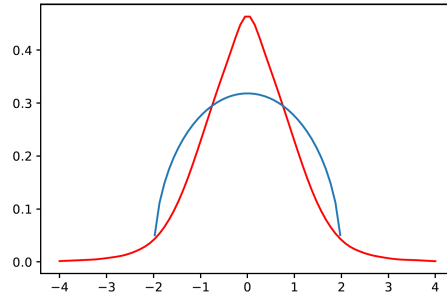

FIGURE 1. In red: The spectral measure of a random PA graph with $n = 6000$, $m = 5$ with a normalization factor of $1/\sqrt{np(1-p)}$ where $p = 2m/n$. In blue: the semicircle distribution $(1/2\pi)\sqrt{4 - x^2}$ for $x \in [-2, 2]$.

Observation 2. The leading eigenvectors are localized, namely, there are a few coordinates which capture a large amount of mass (e.g.,[1], [16], [11]).

Heuristics 3. The following heuristic has been mentioned in several papers [1, 11]: the power law of the edge eigenvalues is due to the localization of the corresponding eigenvectors.

These observations became well-known about twenty years ago. Among others, they have been discussed in Albert and Barabasi's widely circulated survey, which has been cited close to twenty thousand times; see Section 7 of [1]. However, there have been very few rigorous results. The most relevant paper is [12], which shows that for any fixed k, the first k eigenvalues of the graph satisfy a power law. As far as we know, there have been no results concerning the bulk (the triangle part) of the spectrum or the limit of the spectrum once the size of the graph tends to infinity. Likewise, we could not find any rigorous result concerning eigenvectors or any theoretical explanation of **Heuristic 3**. More is known about static models where the number of vertices is fixed and the edges are drawn independently (but with different probablities). However, from the spectral point of view, these models seem to behave differently from the PA model; see [8, 9] and the references there in.

In this paper, we report recent theoretical progress in the understanding of the bulk of the spectrum and the localization of leading eigenvectors. In particular, we will be able to determine the limiting spectral measure up to any precision. Next, we prove the localization phenomenon in a precise form, determining the exact mass of the localized coordinate. Our proofs also provide a satisfying explanation for **Heuristics 3**.

2. NEW RESULTS

2.1. **The model.** Let us first give the precise description of a PA graph, following [6, 5]. The key parameter here is a constant m, which is the number of links (edges) from a new vertex to the existing ones. Intuitively, one wants to link the new vertex to any existing vertex u (including itself) with probability proportional to the degree of u. The precise definition requires some care, since the degrees change after each link has been made.

First start with $m = 1$. The graph $G_{1,1}$ consists of 1 loop around vertex 1. Recursively define $G_{1,t}$ from $G_{1,t-1}$ as follows. Add a new vertex t (so our vertex set at time t will be $[t] := \{1, 2, \ldots, t\}$, and connect it to a random point $X_t \in [t]$ chosen with the following distribution

$$\mathbf{P}[X_t = i] = \begin{cases} \frac{d(i, G_{1,t-1})}{2t-1} & \text{if } 1 \leq i < t \\ \frac{1}{2t-1} & \text{if } i = t \end{cases}$$

where $d(u, G)$ is the degree of u. Note that this definition allows for loops.

Now we consider arbitrary m. We construct $G_{m,t}$ by partitioning the vertices $G_{1,mt}$ into t consecutive groups of size m and viewing each group as a vertex of $G_{m,t}$. Technically speaking, the vertices $\{(a-1)m+1, (a-1)m+2, \ldots, am\}$ of $G_{1,mt}$ get collapsed into vertex a of $G_{m,t}$; for $a = 1, \ldots, t$. This is equivalent to adding m edges at each time step where those edges are added one by one, counting the previous edges as well as the "half" contribution to the degrees. We denote by $G_{m,t}$ the graph at time t. For more details about this model, see [5].

Notation. The asymptotic notation such as O, Θ, o are used under the assumption that the size of the graph tends to infinity. $\mathbf{P}$ denotes probability and $\mathbf{E}$ denotes expectation. For a d dimensional vector v, $v^{(j)}, 1 \leq j \leq d$ are its coordinates. On the other hand, v_i denotes the eigenvectors of the graph under consideration.

2.2. **The moments of the adjacency matrix.** As the size of the graph tends to infinity, the most natural question is to find the limit of its spectrum. To be specific, at time n, the adjacency matrix of the graph has n eigenvalues $\lambda_1(n), \ldots, \lambda_n(n)$. We generate the spectral measure μ_n on the real line by defining $\mu_n(I) = \frac{1}{n}|\{i, \lambda_i(n) \in I\}|$ for any interval I. The question is to find the limit of μ_n (after a possible rescaling if necessary) as n tends to infinity. This limit, if it exists, is refereed to as the limiting spectral measure.

A natural (and popular) method to determine the spectral measure is to compute its moments. Wigner famously used this method to prove his classical semi-circle law [21]. Our first result provides the asymptotics of the moments in question.

Theorem 2.1. *Let $G_{m,n}$ be a PA random graph, and $\mu_{m,n}$ its spectral measure. Let C_k denote the k-th spectral moment of $\mu_{m,n}$. Then*

$$C_0 = 1$$

$$C_1 = \Theta\left(\frac{\log n}{n}\right)$$

$$C_2 = (1 + o(1))2m$$

$$C_3 = \Theta\left(1/\sqrt{n}\right)$$

$$C_4 = (1 + o(1))2m(m+1)\log n.$$

Furthermore, for $k \geq 3$

$$(1) \qquad\qquad C_{2k} = \Theta(n^{k/2-1})$$

$$(2) \qquad\qquad C_{2k-1} = \Theta(n^{k/2-3/2}).$$

Despite this asymptotic result, one cannot use the moment method as in Wigner's case. The sequence $\{C_k\}$ simply *does not* determine a distribution. Since C_2 is a constant (as m is a constant), one does not need to do any rescaling. But without a rescaling, the higher moments tend to infinity.

The reason behind this surprising fact is that the spectrum of $G_{m,n}$ consists of two separate parts. The first part is formed by the edge eigenvalues and tends to follow a power law. The bulk of the spectrum (the triangle shape part) follows a different law. The moments are dominated by the edge eigenvalues (which are significantly larger than the rest) and thus reveal no information about the distribution of the eigenvalues in the bulk.

2.3. **Approximate measures.** Clearly, one needs a new idea to carry out further studies (to get information about the bulk of the spectrum, in particular). The idea we propose here is to study approximations of the spectral measure, rather than the measure itself.

To this end, we define the distance between two probability measure μ and η on the real line as

$$dist(\mu, \eta) = \sup_I |\mu(I) - \eta(I)|,$$

where I runs over the set of all intervals.

For any given precision $\varepsilon > 0$, we can define a sequence of (random) measures $\mu_{\varepsilon,m,n}$ and a deterministic measure $\mu_{\varepsilon,m,\infty}$ such that with probability one, $dist(\mu_{m,n}, \mu_{\varepsilon,m,n}) \leq \varepsilon$ and $\mu_{\varepsilon,m,n} \to \mu_{\varepsilon,m,\infty}$. We define the approximation $\mu_{\varepsilon,m,n}$ as the spectral measure of $G_{\varepsilon,m,n}$, where $G_{\varepsilon,m,n}$ is the graph obtained by discarding the first εn vertices of $G_{m,n}$ (we will refer to $G_{\varepsilon,m,n}$ as the truncated graph). We prove the following

Theorem 2.2. *With probability one,*

$$(3) \qquad\qquad dist(\mu_{m,n}, \mu_{\varepsilon,m,n}) \leq \varepsilon.$$

There is a deterministic measure $\mu_{\varepsilon,m,\infty}$ (uniquely defined by ε and m) such that $\{\mu_{\varepsilon,m,n}\}$ converges weakly in probability to $\mu_{\varepsilon,m,\infty}$. The moments of this limit can be computed explicitly for any given ε.

We designed the approximation scheme to avoid the troublesome large eigenvalues. The large eigenvalues in the spectrum are determined by the largest degrees in the graph, which, naturally, come from the vertices born earlier in the process. We eliminate exactly these vertices.

The estimate (3) is an immediate consequence of the interlacing law. The heart of the theorem is the existence and uniqueness of the limiting measure and the computation of its moments, which we now turn to.

2.4. **The approximate limit and its moments.** We can write down the approximate limit $\{\mu_{\varepsilon,m,\infty}\}$ through its moments. This distribution is symmetric around 0, so all odd moments are zero. For each even number k, the kth moment is a constant $C(k,\varepsilon,m)$, which depends only on k,ε and m. To give an explicit expression for this quantity, we need some preparation.

For the next discussion, it will be convenient to think of the edges in the graph as directed, with the direction going from the larger end to the smaller end. (e.g., if $\{1,3\}$ is an edge, then the direction goes from 3 to 1.) Let G be a graph on $V(G)$, we denote the in-degree of a vertex u by $d_{in}(u,G)$, and the out degree by $d_{out}(u,G)$. The set of vertices $u \in V(G)$ such that $d_{in}(u,G) > 0$ is denoted by $V^-(G)$ (the in-vertices of G). Similarly, we define $V^+(G)$ as the out-vertices of G. Henceforth, when we write an edge as (i,j) we tacitly imply $i \leq j$.

Definition 2.3. *Denote by $\mathcal{K}_{m,n}$ the multigraph on vertex set $[n]$ where between any $i \neq j$ there are m^2 parallel edges, and each vertex has $m(m+1)/2$ loops.*

An essential technicality with the PA model is the order in which the vertices appear.

Definition 2.4. *Define an ordered graph on t vertices, to be a graph H on vertex set $\{v_1,\ldots,v_t\}$ where we impose the ordering $v_i < v_{i+1}$. Note that we do not make the vertex set the integers since we want to emphasize this graph is not any specific subgraph of $\mathcal{K}_{m,n}$. A subgraph of $\mathcal{K}_{m,n}$ is said to be isomorphic to H if it is isomorphic in the classical sense and the mapping preserves the ordering.*

Definition 2.5. *Let G be a multigraph. A walk of length k is a sequence: $(v_1,e_1,v_2,\ldots,e_k,v_{k+1})$ where v_i are vertices of G, e_i are edges of G, and e_i is incident to v_i and v_{i+1}. If $v_{k+1} = v_1$, we call it a closed walk of length k. Let $\mathcal{W}_k(G)$ denote the set of all closed walks in G of length k.*

For a connected subgraph H of $\mathcal{K}$, let $\mathcal{M}_k(H)$ denote the number of elements of $\mathcal{W}_k(\mathcal{K})$ which yield H as a subgraph.

Definition 2.6. *Let $\mathcal{T}_k$ denote the set of ordered trees with at most k edges.*

Definition 2.7. *Let $D = (d_1, \ldots d_t)$ be a sequence of positive integers with $\sum_i d_i = 2(t-1)$. Let $\varepsilon > 0$ be a fixed constant between zero and one. Define*

$$(4) \qquad \psi(D, \varepsilon) = \frac{1}{(2m)^{t-1}} \int_\varepsilon^1 \int_\varepsilon^{y_t} \cdots \int_\varepsilon^{y_2} \prod_{i=1}^t \frac{1}{y_i^{d_i/2}} dy_1 \ldots dy_t.$$

For any given sequence D, it is simple to compute $\psi(D, \varepsilon)$ using standard calculus. Finally, for any graph H, set

$$(5) \qquad \varphi(H, m) = \left(\prod_{v \in V^-(H)} [m]^{d_{in}(v,H)} \right) \left(\prod_{v \in V^+(H)} [m]_{d_{out}(v,H)} \right),$$

where

$$[m]^r := m(m+1) \ldots (m + r - 1)$$

$$[m]_r := m(m-1) \ldots (m - r + 1)$$

are the upper and lower factorials. Now, we are ready to present our moment formula

$$C(2k, \varepsilon, m) = \frac{1}{1 - \varepsilon} \sum_{T \in \mathcal{T}_k} \varphi(T, m) \mathcal{M}_{2k}(T) \psi(D(T), \varepsilon)$$

$$C(2k + 1, \varepsilon, m) = 0.$$

We would like to emphasize that a sequence of constants may not correspond to the moments of any measure. Moreover, even if they do correspond to a measure, they might not uniquely define it. It was a non-trivial task to prove that the constants $C(k, \varepsilon, m)$ defined above determine a unique measure.

2.5. **Remarks.** Our computation of the constants $C(k, \varepsilon, m)$ provides a quantitative explanation for the difference of the spectrum from Wigner's semi-circle law. In the computation with Erdős-Rényi graphs (which leads to the semi-circle law), the moments are essentially the number of closed walks that can be performed on trees of a given size [21]. For the PA graph, the story is more complicated. While the trees are still the family that we weight for the moment computation, one needs to consider: (1) trees of different sizes (2) distinct orderings on the trees (3) assignments of weights which depend on the trees' degree sequences.

Numerically, one can compute the approximate limit very quickly. For a given small ε, one can compute the first few moments and then use the inverse Fourier transform to obtain a distribution. The distribution matches the bulk (triangular) part of the distribution coming from the real graph closely; see Figure 2.
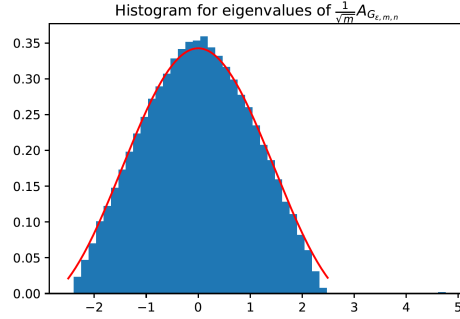
FIGURE 2. In blue: A histogram of the eigenvalues of $\frac{1}{\sqrt{m}} A_{G_{\varepsilon,m,n}}$ with $\varepsilon = 0.10$, $m = 15$ and $n = 4000$. In red: We computed $C(k, 0.10, 15)$ for $k \leq 6$, and using the Inverse Fourier transform to obtain an approximation to the distribution.
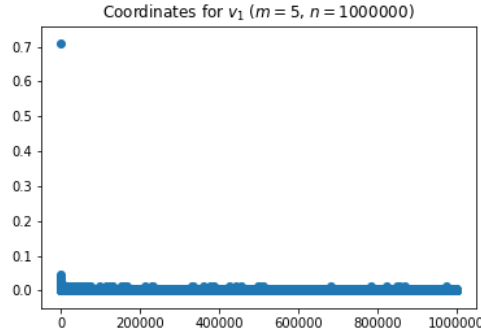

FIGURE 3. Coordinates of v_1: We can see $\|v_1\|_\infty \approx \frac{1}{\sqrt{2}}$ and all the other coordinates are small.

2.6. **Localization of the leading eigenvectors.** Our result concerning the eigenvectors is a rigorous theoretical justification of **Observation 2**. For a vector $v \in \mathbb{R}^n$, define

$$\|v\|_\infty = \max_{j \in [n]} |v(j)|.$$

Theorem 2.8. *Let $G_{m,n}$ be a PA random graph and K be a constant. With high probability[1],*

$$\|v_i\|_\infty = \frac{1}{\sqrt{2}} \pm o(1)$$

for all $1 \leq i \leq K$. Moreover, all the coordinates which do not realize the infinity norm are $o(1)$.

---

[1] An event $\mathcal{E}$ is said to hold with high probability if $\mathbf{P}[\mathcal{E}] = 1 - o(1)$.

We also obtained a new result regarding the power law of the edge eigenvalues. Let $\Delta_i, i = 1, 2, \ldots$ be the degrees in decreasing order.

Theorem 2.9. *Let $G_{m,n}$ be a PA random graph and $k = n^{1/25}$. Then for $i = 1, 2, \ldots, k$, we have, whp, that*

$$\lambda_i(G_{m,n}) = (1 \pm o(1))\sqrt{\Delta_i(G_{m,n})}.$$

This extends the main result of [12] which yields the statement for $k = O(1)$. It is an interesting open problem to find the threshold value for k.

Our analysis offers a more satisfying explanation of **Heuristics 3** mentioned in the introduction, which asserts that the power law of the leading eigenvalues is due to the localization of the corresponding eigenvectors. In fact, both properties (the power law and the localization) are consequences of the fact that the PA graph can be decomposed into a structural part and a random part. The structural part is a union of stars, whose leading eigenvalues satisfy a power law, and whose leading eigenvectors are localized. One can then prove that the impact of the random part on these eigenvalues and eigenvectors is negligible. Thus, the above properties still hold for the union of the two parts.

3. NOTATION AND COMBINATORIAL TOOLS

Fix $i < j$. To obtain an edge between these two points in $G_{m,n}$, by the definition of the graph process we must have had a point in $[(j-1)m+1, jm]$ joining a point in $[(i-1)m+1, im]$ in $G_{1,mn}$. There are clearly m^2 possible pairs. A *labeled* edge (i, j) for us will be one specific choice among the m^2 choices. A *labeled* graph will denote a collection of labeled edges. Since the out degree in $G_{1,mn}$ is exactly one, we only consider labeled graph with this property. Theorem 5.1 will give us an exact formula for the probability of a labeled graph to appear in $G_{m,n}$.

Let H be an ordered graph, and denote by $X_{m,n}(H)$ the number of subgraphs of $G_{m,n}$ isomorphic to H. That is,

$$(6) \qquad\qquad X_{m,n}(H) = \sum_{\substack{\tilde{H} \subset \mathcal{K}_{m,n} \\ \tilde{H} \cong H}} \mathbb{I}_{\tilde{H} \subset G_{m,n}}.$$

Let V be a subset of the vertices and $G[V]$ be the subgraph of G spanned by V. We will be interested in counting the number of copies of H is some fixed subset V. We extend the definition of graph counts as follows

$$(7) \qquad\qquad X_{m,n}(H, V) = \sum_{\substack{\tilde{H} \subset \mathcal{K}_{m,n}[V] \\ \tilde{H} \cong H}} \mathbb{I}_{\tilde{H} \subset G_{m,n}[V]}.$$

Note that in particular, $X_{m,n}(H, [n]) = X_{m,n}(H)$. Throughout this paper we will count graphs which have bounded size.

We are going to study the spectrum of a PA random graph. Since this is a multigraph, its adjacency matrix is not necessarily a $(0, 1)$ matrix.

Definition 3.1. *Let G be a multigraph on vertex set $[n]$. Define the adjacency matrix of G to be the matrix A_G, where for $i \neq j$, $(A_G)_{ij}$ is the number of edges between i and j, and for $i = j$, $(A_G)_{ii}$ is the number of loops at i.*

We will be using the following well known identity

Lemma 3.2. *Let G be a multigraph, and let A_G denote its adjacency matrix. Then*

$$\text{Tr}(A_G^k) = |\mathcal{W}_k(G)|.$$

Each walk in $\mathcal{W}_k(G)$ defines a connected subgraph of G, obtained by taking the union of the edges which appear in the walk.

Hence, if G is a random graph which is always a subgraph of $\mathcal{K}$, then for any $W \in \mathcal{W}_k(\mathcal{K})$, the event "$W \subset G$" denotes the event that the subgraph given by W appears as a subgraph of G. Hence, we can re-write the above as

$$\text{Tr}(A_G^k) = \sum_{W \in \mathcal{W}_k(\mathcal{K})} \mathbb{I}_{W \subset G}.$$

For a connected subgraph H of $\mathcal{K}$, let $\mathcal{M}_k(H)$ denote the number of elements of $\mathcal{W}_k(\mathcal{K})$ which yield H as a subgraph. Hence, if we denote by $\mathcal{C}(\mathcal{K})$ the set of connected subgraphs of $\mathcal{K}$, we can re-write the above as

$$\text{Tr}(A_G^k) = \sum_{H \in \mathcal{C}(\mathcal{K})} \mathcal{M}_k(H) \mathbb{I}_{H \subset G}$$

Note that $\mathcal{C}(\mathcal{K})$ is a set of labeled graphs, so in order to write it as graph counts, we need to decompose it even further. Let $\mathcal{F}(\mathcal{K})$ denote the set of graphs which are isomorphic to a member of $\mathcal{C}(\mathcal{K})$. Note that by our definition of isomorphism, we have that $\mathcal{F}(\mathcal{K})$ is a set of ordered graphs. Furthermore, note that $\mathcal{M}_k(\cdot)$ was defined for labeled graphs, but since $H_1 \cong H_2$ implies $\mathcal{M}_k(H_1) = \mathcal{M}_k(H_2)$ then when we write $\mathcal{M}_k(H)$ for a $H \in \mathcal{F}(\mathcal{K})$ we refer to the obvious quantity. Hence, we arrive at the final form we will be using in this paper:

$$(8) \qquad \text{Tr}(A_G^k) = \sum_{H \in \mathcal{F}(\mathcal{K})} \mathcal{M}_k(H) X(H, G)$$

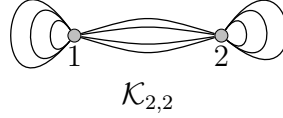where $X(H, G)$ is the number of copies of H in G.

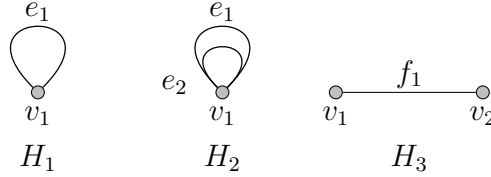Consider the following example: Suppose we sampled the following graph $G_{2,2,} \in \mathcal{G}_{2,2}$:

$$G_{2,2}$$
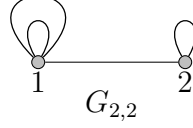
Then,

$$A_{G_{2,2}} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \qquad \text{and} \qquad \text{Tr}(A_{G_{2,2}}^2) = 7$$

Now we carry out the computation as suggested by equation (8). The complete graph $\mathcal{K}_{2,2}$ is given by:

$$\mathcal{K}_{2,2}$$

The only graphs $H \in \mathcal{F}(\mathcal{K}_{2,2})$ which have $X(H, G_{2,2}) > 0$ are the following:

The closed walks of length two that may have formed each are:

- H_1: $(v_1, e_1, v_1, e_1, v_1)$.
- H_2: $(v_1, e_1, v_1, e_2, v_1)$, and $(v_1, e_2, v_1, e_1, v_1)$.
- H_3: $(v_1, f_1, v_2, f_1, v_1)$, and $(v_2, f_1, v_1, f_1, v_2)$.

Hence, $\mathcal{M}_2(H_1) = 1$, and $\mathcal{M}_2(H_2) = \mathcal{M}_2(H_3) = 2$. Computing the graph counts we get: $X(H_1, G_{2,2}) = 3$, $X(H_2, G_{2,2}) = 1$, $X(H_3, G_{2,2}) = 1$. Hence,

$$\sum_{H \in \mathcal{F}(\mathcal{K}_{2,2})} \mathcal{M}_2(H) X(H, G_{2,2}) = 1(3) + 2(1) + 2(1) = 7 = \text{Tr}(A_{G_{2,2}}^2).$$

Definition 3.3. *Let A_n be an $n \times n$ symmetric matrix. Denote by $\{\lambda_i\}_{i=1}^n$ the eigenvalues of A_n. We define the empirical spectral distribution (ESD) to be the following measure:*

$$\mu_{A_n}(x) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\lambda_i = x}.$$

Usually when it is required, we will denote the eigenvalues by $\lambda_i(A_n)$, but if the matrix A_n is clear from the context, we will simply write λ_i. Moreover, it will always be assumed that

the eigenvalues are labeled to be non-increasing, i.e., $\lambda_i \geq \lambda_{i+1}$. We are interested in studying the measures when the matrices A_n follow the preferential attachment rule. We extend the definition of ESD to graphs.

Definition 3.4. *Let G be a graph on n vertices, and let A_G denote its adjacency matrix. Then the ESD of G which we denote by μ_G, is the ESD of the matrix A_G.*

Note that when the graph G is random, we will obtain a random measure. We are interested in studying the measures $\mu_{G_{m,n}}$ when $G_{m,n} \in \mathcal{G}_{m,n}$. In particular, we are interested in understanding their limiting behavior, which we formalize through the following notions.

Definition 3.5. *Let $\{\mu_n\}_{n=1}^\infty$ be a sequence of (deterministic) measures. We say that $\{\mu_n\}$ converges weakly to a measure μ, if for every bounded, continuous function f we have*

$$\int f d\mu_n \to \int f d\mu.$$

We use the following convention:

$$\mathbf{E}[f(\mu)] := \int f d\mu.$$

In the case when $f(x) = x^k$, we refer to $\mathbf{E}[\mu^k]$ as the k-th moment of the measure. If μ is determined by its moments, then convergence of the moments guarantees weak convergence:

Lemma 3.6. *Let $\{\mu_n\}$ be a sequence of (deterministic) measures. If μ is uniquely defined by its moments, and for every $k \geq 0$ we have*

$$\int x^k d\mu_n \to \int x^k d\mu$$

then $\{\mu_n\}$ converges weakly in probability to μ.

Proof. The proof is a standard exercise, see Appendix. $\square$

In the case when μ is the ESD of a graph, we refer to $\mathbf{E}[\mu_G^k]$ as the k-th spectral moment. Using the notation above we have,

$$\mathbf{E}[\mu_G^k] = \int x^k d\mu$$
$$= \frac{1}{|V(G)|} \sum_i \lambda_i^k$$
$$= \frac{1}{|V(G)|} \operatorname{Tr}(A_G^k)$$
$$(9) \qquad = \frac{1}{|V(G)|} \sum_{H \in \mathcal{F}(G)} \mathcal{M}_k(H) X(H, G).$$

Note that when $\{\mu_n\}$ are random measures, then $\mathbf{E}[f(\mu_n)]$ are random variables, so we need to specify the type of convergence one obtains in Definition 3.5. The first and most natural definition deals with the behavior of a "typical" measure μ_n.

Definition 3.7. *Let A_n be a $n \times n$ random symmetric matrix, and let μ_n denote its ESD. We define the expected ESD, $\bar{\mu}_n$, via duality as follows:*

$$\int f d\bar{\mu}_n := \mathbf{E}\left[\int f d\mu_n\right]$$

Moreover, we say that $\{\mu_n\}$ converges in expectation to a fixed measure μ iff $\{\bar{\mu}_n\}$ converges weakly to μ.

A stronger type of convergence is given by the following notion.

Definition 3.8. *Let $\{\mu_n\}$ be a sequence of random measures. We say that $\{\mu_n\}$ converges weakly in probability to a fixed measure μ if for any bounded continuous function f and $\varepsilon > 0$ we have*

$$\lim_{n \to \infty} \mathbf{P}\left[\left|\int f d\mu_n - \int f d\mu\right| > \varepsilon\right] = 0.$$

Just as before, if we know that μ is determined by its moments, then we can apply the moment method:

Lemma 3.9. *Let $\{\mu_n\}$ be a sequence of random measures, and let μ be a fixed measure which is uniquely defined by its moments. If*

$$\int x^k d\mu_n \to \int x^k d\mu$$

in probability, then $\{\mu_n\}$ converges weakly in probability to μ.

Proof. This is a standard exercise, see Appendix.

$\square$

As mentioned before, when we perform the moment computation for the truncated graph, we will need to consider the degree sequence of trees. The following lemma gives the number of trees with a given degree sequence (see Lemma 1 in [15]):

Lemma 3.10. *If we have positive integers $d_1, \ldots, d_{k+1}$ such that $\sum d_i = 2k$, then the number of trees T on $k + 1$ vertices with degree sequence equal to $(d_1, \ldots, d_{k+1})$ is given by*

$$\binom{k-1}{d_1 - 1, \ldots, d_{k+1} - 1}.$$

4. Tools from Linear Algebra

In this section we present the linear algebra tools we will be using throughout the text.

Theorem 4.1 (Cauchy's interlacing theorem)**.** *Let A be an $n \times n$ symmetric matrix, and let B be an $m \times m$ (with $m < n$) principal submatrix (that is, B is obtained by deleting columns and their corresponding rows). Suppose A has eigenvalues $\lambda_n \leq \ldots \leq \lambda_1$ and B has eigenvalues $\beta_m \leq \ldots \leq \beta_1$. Then,*

$$\lambda_{m-k+1} \geq \beta_{m-k+1} \geq \lambda_{n-k+1}$$

for $k = 1, \ldots, m$.

Theorem 4.2 (Weyl's inequality)**.** *Let A, B be $n \times n$ symmetric matrices with $\|B\| \leq \delta$. Let $A' = A + B$. Then:*

$$|\lambda_i(A) - \lambda_i(A')| \leq \delta$$

for all i. (Again we assume that the eigenvalues are ordered increasingly.)

Lemma 4.3. *Let A be an $n \times n$ matrix with non-negative entries. Then for any positive $\{c_i\}$, $i = 1, \ldots, n$,*

$$(10) \qquad \|A\| \leq \max_{1 \leq i \leq n} \left\{ \frac{1}{c_i} \sum_{j=1}^{n} c_j a_{ij} \right\}.$$

Proof. Let C be a diagonal matrix with entries $\{c_1, c_2, \ldots, c_n\}$. Both A and $C^{-1}AC$ have the same eigenvalues. Since $c_i > 0$ and A is non-negative, then $C^{-1}AC$ is also non-negative. For a non-negative matrix, the largest eigenvalue is upper bounded by the max row sum, which implies the result. $\qquad\square$

Theorem 4.4 (Davis-Kahan, see Theorem 1 in [23])**.** *Let A and B be symmetric $n \times n$ matrices. Let $A = B + C$. Denote by $v_i(B), v_i(A)$ the eigenvectors corresponding to $\lambda_i(B)$ and $\lambda_i(A)$. Then,*

$$\sin(\angle(v_i(B), v_i(A))) \leq \frac{\|C\|}{\min(|\lambda_{i-1}(A) - \lambda_i(B)|, |\lambda_{i+1}(A) - \lambda_i(B)|)}.$$

5. Tools from Probability

In this section we state the probability bounds for graph counts that we will be using throughout the paper. Since we obtain $G_{m,n}$ by collapsing points in $G_{1,mn}$, it is natural to first study the probability of a labeled graph to appear in $G_{1,mn}$. Fortunately, this part was already done in [5]:

Theorem 5.1 (Theorem 13 in [5])**.** *Let $G_{1,n}$ be a random PA graph. Let S be a labeled graph with out-degree of every vertex at most 1. Define $C_S(t)$ to be the number of edges (i, j) in S*

such that $i \leq t$ and $t \leq j$. Then:

$$\mathbf{P}[S] = \left(\prod_{i \in V^-(S)} d_{in}(i, S) \right) \left(\prod_{i \in V^+(S)} \frac{1}{2i - 1} \right) \left(\prod_{i \notin V^+(H)} \left(1 + \frac{C_S(i)}{2i - 1} \right) \right)$$

Furthermore,

$$(11) \qquad \mathbf{P}[S] = \left(\prod_{i \in V^-(S)} d_{in}(i, S)! \right) \left(\prod_{(i,j) \in E(S)} \frac{1}{2\sqrt{ij}} \right) \exp\left(O\left(\sum_{i \in V(S)} C_S(i)^2 / i \right) \right).$$

We will be using only the second statement of the above theorem in this paper. Note that in particular when $S = (u, v)$, we obtain $\mathbf{P}[S] \approx 1/2\sqrt{uv}$. Hence, if S has more than one edge, then the second factor is exactly what one would obtain if the edges appeared independently. The third factor is a correction factor which tends to one. The first factor captures the correlation among the edges in the preferential attachment model. For instance if we have $a < b < c$, then we would obtain $\mathbf{P}[(a, b), (a, c)] \approx 2\mathbf{P}[(a, b)]\mathbf{P}[(a, c)]$ (since the in-degree of a is two).

Given an ordered graph H, and a subset of vertices V, we could have multiple copies of H in the vertex set V. For example:

Example 1. *Let H be a triangle, then the total number of copies of H that can occur in a given triplet of vertices $\{a, b, c\}$ is $m^5(m - 1)$. To see this, note that to obtain a triangle in $G_{m,n}$ we must have points $a_1, a_2 \in [(a - 1)m + 1, am]$, $b_1, b_2 \in [(b - 1)m + 1, bm]$ and $c_1, c_2 \in [(c - 1)m + 1, cm]$ such that during the construction of $G_{1,mn}$ we joined points $(a_1, b_1), (a_2, c_1)$ and (b_2, c_2). We have m^2 choices to choose a_1 and a_2 since they could be the same point, and similarly for b_1, b_2, but since the out-degree in $G_{1,mn}$ is 1, then we must have $c_1 \neq c_2$ giving $m(m - 1)$ options.*

Another ingredient one needs to keep in mind when performing graph counts is that all the $m^5(m - 1)$ triangles are not "weighted" the same. As we mentioned before, we see that when $a_1 = a_2$ we obtain an extra factor of two on the probability. Thus, to perform a graph count we need to keep in mind the number of ways the graph could have been obtained from the uncollapsed graph $G_{1,mn}$ and then assign it a weight according to the first factor in (11). As an illustration on how the count should be carried out, consider the following:

Example: We want to compute the expected number of paths of length two in a random $G_{m,n}$ graph. There are a total of three non-isomorphic paths of length two (see Definition 2.4):
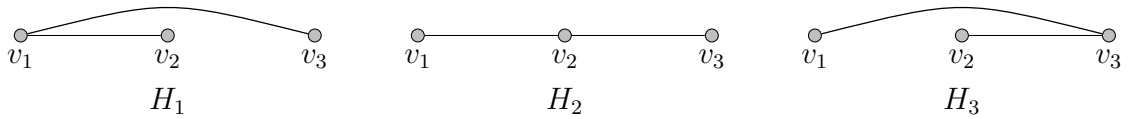
We will count each one separately. Recall the notation from equations (6) and (7). For H_1, we have:

$$\mathbf{E}[X_{m,n}(H_1)] = \sum_{1 \leq a < b < c \leq n} \mathbf{E}[X_{m,n}(H_1, \{a, b, c\})]$$

Let $A = [(a-1)m + 1, am]$, $B = [(b-1)m + 1, bm]$ and $C = [(c-1)m + 1, cm]$. Then all the possible contributions to $X_{m,n}(H_1, \{a, b, c\})$ come from having an edge (a_1, b_1) from B to A and an edge (a_2, c_1) from C to A (in the uncollapsed graph $G_{1,mn}$). If $a_1 = a_2$, then by Theorem 5.1, the probability of $(a_1, b_1), (a_1, c_1)$ is given by:

$$2 \cdot \frac{1}{4a_1\sqrt{b_1 c_1}} \exp\left(O\left(\frac{1}{a_1} + \frac{4}{b_1} + \frac{1}{c_1}\right)\right)$$

which can be rewritten as:

$$2 \cdot \frac{1}{4m^2 a\sqrt{bc}}\left(1 \pm O\left(\frac{1}{a}\right)\right)$$

There are m choices for a_1, b_1, and c_1, so we obtain an extra factor of m^3 to arrive at:

$$(12) \qquad \frac{2m}{4a\sqrt{bc}}\left(1 \pm O\left(\frac{1}{a}\right)\right)$$

If however a_1 and a_2 are distinct, then we do not have a factor of 2, and now there are $m(m-1)$ ways to choose them from A, so this case yields a contribution of

$$(13) \qquad \frac{m^3(m-1)}{4m^2 a\sqrt{bc}}\left(1 \pm O\left(\frac{1}{a}\right)\right)$$

Hence, by adding equations (12) and (13) we get:

$$(14) \qquad \mathbf{E}[X_{m,n}(H_1, \{a, b, c\})] = \frac{m(m+1)}{4a\sqrt{bc}}\left(1 \pm O\left(\frac{1}{a}\right)\right)$$

Taking the sum over all $1 \leq a < b < c \leq n$ gives:

$$(15) \qquad \mathbf{E}[X_{m,n}(H_1)] = (1 \pm o(1))\frac{m(m+1)n \log n}{2}$$

Similarly we can compute the contributions of $\mathbf{E}[X_{m,n}(H_2)]$ and $\mathbf{E}[X_{m,n}(H_3)]$, but they both turn out to be $O(n)$. Hence the expected number of paths of length two is given by:

$$(16) \qquad (1 \pm o(1))\frac{m(m+1)n \log n}{2}.$$

From the above example we see the importance of Definition 2.4. Since different orderings give completely different counts, it is natural to treat them separately. When we are concerned about the magnitude of the count, and do not care about the leading constant, the following theorem will give us a quick way to deal with such situations.

Theorem 5.2. *Let H be an ordered graph. Then:*

$$\mathbf{E}[X_{m,n}(H)] = O\left(n^{f(H)/2} \log^{g(H)} n\right) \tag{17}$$

where $f(H)$ denotes the number of vertices of H with degree one, and $g(H)$ denotes the number of vertices of H of degree 2. Moreover, if the ordering is such that $d(v_i, H) \geq d(v_{i+1}, H)$ then one can replace $O(\cdot)$ by $\Theta(\cdot)$.

Proof. Let t denote the number of vertices of H. Then:

$$\mathbf{E}[X_{m,n}(H)] = \sum_{V \subset [n]: |V|=t} \mathbf{E}[X_{m,n}(H,V)].$$

For $V = \{x_1 < \ldots < x_t\}$, we have:

$$\mathbf{E}[X_{m,n}(H,V)] = \Theta\left(\prod_{i=1}^{t} \frac{1}{x_i^{d(v_i,H)/2}}\right). \tag{18}$$

To see this, let $\mathcal{S}$ denote the set of labeled graphs in $G_{1,mn}$ which collapse to a copy of H on the vertex set V. Then:

$$\mathbf{E}[X_{m,n}(H,V)] = \sum_{S \in \mathcal{S}} \mathbf{P}[S \subset G_{1,mn}]$$

By Theorem 5.1 equation (11) we see that for each fixed S we have:

$$\mathbf{P}[S \subset G_{1,mn}] = \Theta(1)\left(\prod_{(i,j)\in E(S)} \frac{1}{2\sqrt{ij}}\right)$$

By taking sum over all possible $S \in \mathcal{S}$ gives equation (18). Since m, and $|E(H)|$ are constants we obtain:

$$\mathbf{E}[X_{m,n}(H)] = \sum_{1\leq x_1<\ldots<x_t\leq n} \mathbf{E}\left[X_{m,n}(H,\{x_1,\ldots,x_t\})\right]$$

$$= \sum_{1\leq x_1<\ldots<x_t\leq n} \Theta\left(\prod_{i=1}^{t} \frac{1}{x_i^{d(v_i,H)/2}}\right)$$

$$= O\left(\prod_{i=1}^{t}\left(\sum_{1\leq x_i\leq n} \frac{1}{x_i^{d(v_i,H)/2}}\right)\right)$$

$$= O(n^{f(H)/2} \log^{g(H)}).$$

For the moreover part of the theorem, see the Appendix. $\qquad \square$

As we will see later in the paper, we will sometimes be interested in counting ordered graphs where the smallest vertex is large. That is, given an ordered graph H, and $M = M(n)$ a function which tends to infinity, then we can consider the copies of H which appear on $[M, n]$.

Note that in some cases, this quantity will be asymptotically equal to the graph count in the whole graph. To perform these counts more accurately, we present Lemma 5.3:

Lemma 5.3. *Let H be an ordered loopless graph on t vertices. Let $V \subset [M, n]$ be a subset of t vertices $\{x_1 < \ldots < x_t\}$. Then:*

$$(19) \qquad \mathbf{E}[X_{m,n}(H, V)] = \left(1 \pm O\left(\frac{1}{M}\right)\right) \frac{\varphi(H, m)}{(2m)^{|E(H)|}} \prod_{i=1}^{t} \frac{1}{x_i^{d(v_i, H)/2}}$$

where we recall the definition of φ:

$$\varphi(H, m) = \left(\prod_{v \in V^-(H)} [m]^{d_{in}(v, H)}\right)\left(\prod_{v \in V^+(H)} [m]_{d_{out}(v, H)}\right).$$

Note that the above lemma takes care of a couple of trivial cases as well: If a graph H has out-degree more than m, then we know it can never appear in our process, and the lower factorials do indeed give zero if any out degree is more than m. Furthermore, it saves a lot of work when computing graph counts: If we revisit the situation where $H = \{(v_1, v_2), (v_1, v_2)\}$, then note that $\varphi(H, m) = m^3(m + 1)$, so we obtain equation (14) immediately.

Proof of lemma 5.3. Let $\mathcal{S}(H, V)$ be the set of graphs in $G_{1,mn}$ such that they form a copy of H in the vertex set V after being collapsed. We want to compute the probability $S \subset G_{1,mn}$ for $S \in \mathcal{S}(H, V)$, by Theorem 5.1 this is given by :

$$\mathbf{P}[S] = \left(\prod_{i \in V^-(S)} d_{in}(i, S)!\right)\left(\prod_{(i,j) \in E(S)} \frac{1}{2\sqrt{ij}}\right) \exp\left(O\left(\sum_{i \in V(S)} C_S(i)^2/i\right)\right)$$

the last factor will be always $(1 \pm O(1/M))$. The second factor we can write as:

$$\frac{1}{(2m)^{|E(H)|}} \prod_{i=1}^{t} \frac{1}{x_i^{d(v_i, H)/2}} \left(1 \pm O\left(\frac{1}{M}\right)\right)$$

Thus,

$$\mathbf{E}[X_{m,n}(H, V)] = \left(1 \pm O\left(\frac{1}{M}\right)\right) \frac{1}{(2m)^{|E(H)|}} \prod_{i=1}^{t} \frac{1}{x_i^{d(v_i, H)/2}} \sum_{S \in \mathcal{S}(H,V)} \left(\prod_{i \in V^-(S)} d_{in}(i, S)!\right).$$

We focus only on the last factor. Let r be the number of edges of H, and denote these edges by $e_j = \{(a_j, b_j)\}$. Let X_j denote the subset of $\{a_1, \ldots, a_r, b_1 \ldots, b_r\}$ which corresponds to vertex x_j, in particular note that $|X_j| = d(v_j, H)$. We will further split each X_j into two types: In-type and out-type which will correspond to whether they were an in-vertex or an out vertex, and call these sets $I(X_j)$ and $O(X_j)$ respectively. For example, if we consider H to be the following ordered graph with the given labeling:

$$H$$
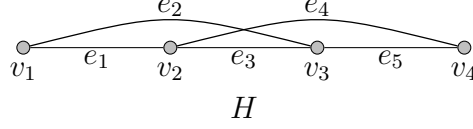
Then

  • $X_1 = \{a_1, a_2\}$, $I(X_1) = \{a_1, a_2\}$ and $O(X_1) = \emptyset$.
  • $X_2 = \{b_1, a_3, a_4\}$, $I(X_2) = \{a_3, a_4\}$ and $O(X_2) = \{b_2\}$.
  • $X_3 = \{b_3, b_2, a_5\}$, $I(X_3) = \{a_5\}$, and $O(X_3) = \{b_2, b_3\}$.
  • $X_4 = \{b_4, b_5\}$, $I(X_4) = \emptyset$, and $O(X_4) = \{b_4, b_5\}$.

Note that we can think of $S \in \mathcal{S}(H, V)$ as maps σ such that for a point $u \in X_j$ we have $\sigma(u) \in [(x_j - 1)m + 1, x_j m]$. Furthermore, given any mapping σ, which takes $\sigma : X_j \to [(x_j - 1)m + 1, x_j m]$ which satisfies $|\sigma(O(X_j))| = d_{out}(v_j, H)$ (this is the condition that the out vertices have degree one), then we can construct a graph.

For a map σ, we can define a function $w(\sigma)$ which equals $\prod_{j \in V^-(S)} d_{in}(j, S)!$ where S is the graph corresponding to σ. Let Ω be the set of all such mappings. Then, the factor we are interested in equals

$$\sum_{\sigma \in \Omega} w(\sigma)$$

Note that every σ is uniquely defined by its action on X_j. Hence, we can break them up as $\sigma = \sigma_1 \times \ldots \times \sigma_t$. Let Ω_j be the set of all mappings σ_j from X_j to $[(x_j - 1)m + 1, x_j m]$ which satisfies $|\sigma(O(X_j))| = d_{out}(v_j, H)$. Given one σ_j, we can define

$$w(\sigma_j) = \prod_{u \in Im(\sigma_j)} |\sigma_j^{-1}(u) \cap I(X_j)|!$$

Note that

$$w(\sigma) = \prod_{i=1}^{t} w(\sigma_j)$$

where σ_j is the restriction of σ to X_j. Thus, we obtain:

$$\sum_{S \in \mathcal{S}(H,V)} \prod_{j \in V^-(S)} d_{in}(j, S)! = \sum_{(\sigma_1, \ldots, \sigma_t) \in \Omega_1 \times \ldots \times \Omega_t} \prod_{j=1}^{t} w(\sigma_j)$$

$$= \prod_{j=1}^{t} \sum_{\sigma_j \in \Omega_j} w(\sigma_j)$$

Fix j, and for the remainder of the proof write d_j instead of $d_{in}(v_j, H)$. First of all note that the only restriction on where to map $O(X_j)$ is that they must map to distinct points. This will gives us a factor of $[m]_{d_{out}(v_j, H)}$. To map $I(X_j)$ and weight the maps accordingly, we have:

- If we choose the images to be d_j distinct points, then we have $[m]_{d_j}$ many maps with $w(\sigma_j) = 1$.
- If we choose the images to be $d_j - 1$ distinct points, then we have $\binom{d_j-1}{2}[m]_{d_j-1}$ many maps with $w(\sigma_j) = 2$.

$\vdots$

- If we choose the images to be 1 distinct point, then we have $[m]_1$ many maps with $w(\sigma_j) = d_j!$.

In general, the contribution of adding over maps σ with exactly k points in the image is given by:

$$\frac{d_j!\binom{d_j-1}{k-1}}{k!}[m]_k.$$

Indeed, First we permute the points. Then we partition them into k non-empty sets, and for the first partition we have m ways to select its image, for the second partition we have $m-1$ ways to select its image and so on. Lastly, we divide by $k!$ to account for the double counting given by the ways we can re-arrange the partition amongst themselves. The key point here is that we do not further divide by the number of ways we can re-arrange the inside of these partitions. This will make each mapping σ_j be counted a total of $\prod_{u \in Im(\sigma_j)} |\sigma_j^{-1}(u) \cap I(X_j)|!$ many times (which is exactly the factor we wanted to weight it by). Hence,

$$
\begin{aligned}
\sum_{\sigma_j \in \Omega_j} w(\sigma_j) &= [m]_{d_{out}(v_j,H)} \sum_{k=1}^{d_j} \frac{d_j!\binom{d_j-1}{k-1}}{k!}[m]_k \\
&= [m]_{d_{out}(v_j,H)} d_j! \sum_{k=1}^{d_j} \binom{d_j-1}{k-1}\binom{m}{k} \\
&= [m]_{d_{out}(v_j,H)} d_j! \sum_{k=0}^{d_j-1} \binom{d_j-1}{d_j-1-k}\binom{m}{k+1} \\
&= [m]_{d_{out}(v_j,H)} d_j! \binom{m+d_j-1}{d_j} \\
&= [m]_{d_{out}(v_j,H)} [m]^{d_j} \\
&= [m]_{d_{out}(v_j,H)} [m]^{d_{in}(v_j,H)}
\end{aligned}
$$

That is,

$$
\begin{aligned}
\sum_{S \in \mathcal{S}(H,V)} \prod_{j \in V^-(S)} d_{in}(j,S)! &= \prod_{j=1}^{t} \sum_{\sigma_j \in \Omega_j} w(\sigma_j) \\
&= \prod_{j=1}^{t} [m]_{d_{out}(v_j,H)} [m]^{d_{in}(v_j,H)}
\end{aligned}
$$

$$= \varphi(H, m)$$

as defined in the theorem. This concludes the proof. □

Lastly, we will be using the following lemma:

Lemma 5.4 (Lemma 3 in [6])**.** *Let H_1 and H_2 be labeled graphs (that is, some specific graphs of $\mathcal{K}_{m,n}$), which are vertex disjoint. Then:*

$$\mathbf{P}[H_1 \cup H_2] \leq \mathbf{P}[H_1]\mathbf{P}[H_2].$$

That is, disjoint graphs are negatively correlated.

The above lemma is of course no surprise. When we condition on one of the graphs appearing, this reduces the expected degree of the other one (since we know that some of the edges "missed" connections), and thus they are negatively correlated.

Recall from above that we will be describing the approximate limit of the spectrum via its moments. There are two natural issues with this strategy: Is there a distribution whose moments match the given constants $C(k, \varepsilon, m)$, and if so, is it unique?

The first question is known in the literature as the Hamburger moment problem: Given a sequence $\{m_k\}_{k \geq 0}$, we say that it is solvable if there is a positive Borel measure μ on the real line such that the moments of μ are given by $\{m_k\}$. If the sequence is indeed solvable, the question of whether the measure is unique is known as M-determinacy (short for moment-determinacy).

Theorem 5.5 (Hamburger Theorem)**.** *A sequence $\{m_k\}_{k \geq 0}$ is solvable if and only if for an arbitrary sequence $\{c_j\}_{j \geq 0}$ of complex numbers with finite support, one has:*

$$(20) \qquad\qquad \sum_{r,s \geq 0} m_{r+s} c_r \bar{c}_s \geq 0.$$

Theorem 5.6 (Carleman's condition)**.** *Let μ be a measure on $\mathbb{R}$ such that all the moments*

$$m_k = \int x^k d\mu$$

are finite. If

$$(21) \qquad\qquad \limsup_{k \to \infty} \frac{m_{2k}^{1/2k}}{2k} < \infty$$

then μ is the only measure on $\mathbb{R}$ with $\{m_k\}$ as its sequence of moments. That is, $\{m_k\}$ is M-determinate.

6. Spectral moments

In this section we find the spectral moments of a random PA graph. Note that in [20] they conjectured that in order to carry out the moment computation for a random preferential attachment graph, it suffices to consider the edges as if they appeared independently (see their Conjecture 1). Although their model is slightly different than the one used here, we can see that the conjecture does not hold since the preferential attachment rule by which we add edges gives us the extra factor in equation (11).

Theorem 6.1. *Let $G_{m,n}$ be a PA random graph, and let $\mu_{m,n}$ denote its ESD. Let $\bar{\mu}_{m,n}$ be the expected ESD. Then the first spectral moments are:*

$$\mathbf{E}[\bar{\mu}_{m,n}^0] = 1$$

$$\mathbf{E}[\bar{\mu}_{m,n}^1] = \Theta\left(\frac{\log n}{n}\right)$$

$$\mathbf{E}[\bar{\mu}_{m,n}^2] = (1 + o(1))2m$$

$$\mathbf{E}[\bar{\mu}_{m,n}^3] = \Theta\left(1/\sqrt{n}\right)$$

$$\mathbf{E}[\bar{\mu}_{m,n}^4] = (1 + o(1))2m(m + 1)\log n$$

and for $k \geq 3$ we have:

$$(22) \qquad \mathbf{E}[\bar{\mu}_{m,n}^{2k}] = \Theta(n^{k/2-1})$$

$$(23) \qquad \mathbf{E}[\bar{\mu}_{m,n}^{2k-1}] = \Theta(n^{k/2-3/2}).$$

Remark 6.2. *Note that in particular we obtain the same first, second and third moments, but the 4th-moment is different in [20]. The extra factor in equation (11) which they do not use changes the leading coefficient from m^2 to $m(m + 1)$.*

Proof of Theorem 6.1. For short hand, we are going to write M_k instead of $\mathbf{E}[\bar{\mu}_{m,n}^k]$. Using equation (9) we have:

$$M_k = \frac{1}{n}\mathbf{E}\left[\sum_{H \in \mathcal{F}(\mathcal{K}_{m,n})} \mathcal{M}_k(H)X_{m,n}(H)\right]$$

$$= \frac{1}{n}\sum_{H \in \mathcal{F}(\mathcal{K}_{m,n})} \mathcal{M}_k(H)\mathbf{E}[X_{m,n}(H)].$$

We mention a few words on moments M_k for $k \leq 4$ and then handle the remaining cases. We have to treat them separately since they are corner cases.

- M_0 follows from the definition of a pdf.
- M_1 comes from loops. By Theorem 5.2 the expected number of loops is $\Theta(\log n)$.
- M_2 comes from edges and loops. The contribution from loops is again $\Theta(\log n)$, and the number of edges is mn. Since we can transverse any edge in two ways we get $M_2 = (1 + o(1))2m$.

- M_3 will be dealt with in the general case below.
- M_4 the contributions come from: A collection of loops around a vertex (contribution of $\Theta(\log n)$), a collection of loops and an edge (contribution of $\Theta(\sqrt{n})$ by using Theorem 5.2 and placing the vertex with the loop before the leaf), an edge which gives a contribution of $\Theta(n)$. The only remaining contribution is from paths of length two which we did in our graph count example. If H is a path of length two, then $\mathcal{M}_4(H)$ equals 4. Hence,

$$M_4 = (1 \pm o(1)) \left(4 \cdot \frac{m(m+1) \log n}{2} \right).$$

For the general case we split the proof into the odd case and the even case.

Even moments: We aim to compute M_{2k}. By equation (9):

$$M_{2k} = \frac{1}{n} \sum_{H \in \mathcal{F}(\mathcal{K}_{m,n})} \mathcal{M}_{2k}(H) \mathbf{E}[X_{m,n}(H)].$$

By equation (5.2) we see that the contribution of an ordered graph H is $O(n^{f(H)/2-1} \log^{g(H)} n)$. Since H comes from a closed walk of length $2k$, we must have $f(H) \leq k$ (to reach each leaf we must walk their corresponding edge back and forth), with equality if and only if we have a star on k edges. This yields the upper bound of $O(n^{k/2-1})$. Now note that this bound is indeed achievable by Theorem 5.2 if we let H be the following ordered graph:
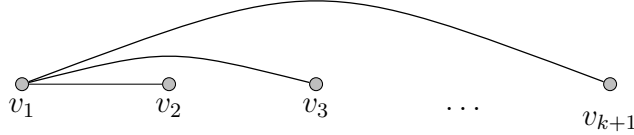
Odd moments: We want to compute M_{2k-1}. Just as above, note that for any H which yields a closed walk of length $2k-1$ we must have $f(H) \leq k-1$. To see this, note that any closed walk must have a cycle, which uses at least one edge (in the case of a loop), and then we can have at most $k-1$ vertices of degree one. Note that we can only have equality when we have a start and a loop around the high degree vertex. This yields an upper bound of $O(n^{(k-1)/2-1})$. In the case when H is the following ordered graph:
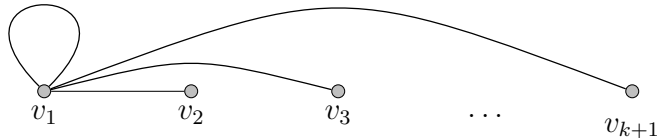
we obtain a matching lower bound, which implies the result. $\qquad \square$

As we pointed out during the introduction, in order to further study the spectrum of the PA graph, we need to break our analysis into two: The bulk of the spectrum and the edge eigenvalues.

7. Spectrum of the truncated graphs and Proof of Theorem 2.2

In this section we will prove Theorem 2.2. As we will see in a later section, the largest eigenvalues will arise from the stars with large degree (which correspond to vertices appearing early in our graph). Hence, to exclude these large eigenvalues from our study, we will consider the graph obtained by deleting the first vertices:

Definition 7.1. *Let $\varepsilon > 0$ be fixed. An ε-truncated PA random graph is a random graph obtained by sampling $G_{m,n} \in \mathcal{G}_{m,n}$, and then deleting the first εn vertices to obtain $G_{\varepsilon,m,n}$. We will denote this space of random graphs by $\mathcal{G}_{\varepsilon,m,n}$. Note that these will be a graphs on $n - \varepsilon n$ vertices and we will refer to its vertices by its original label. That is, $V(G_{\varepsilon,m,n}) = [\varepsilon n, n]$.*

We want to study the limiting behavior of the random measures $\mu_{G_{\varepsilon,m,n}}$ which henceforth we will write as $\mu_{\varepsilon,m,n}$ to shorten the notation. Note that the first part of Theorem 2.2 will be given by Cauchy's interlacing theorem (Theorem 4.1). Namely,

$$dist(\mu_{m,n}, \mu_{\varepsilon,m,n}) \leq \varepsilon.$$

For the second part of the theorem, we define $\mu_{\infty,\varepsilon,m}$ via its moments, which will be the sequence given by the sequence $\{C(k, \varepsilon, m)\}$ which we define below.

Before we define the sequence, we present some notation we will be using:

Definition 7.2. *Let $\mathcal{T}_k$ denote the set of ordered trees with at most k edges.*

Definition 7.3. *Let $D = (d_1, \ldots d_t)$ be a sequence of positive integers with $\sum_i d_i = 2(t-1)$. Let $\varepsilon > 0$ be a fixed constant between zero and one. Then define:*

$$(24) \qquad \psi(D, \varepsilon) = \frac{1}{(2m)^{t-1}} \lim_{n \to \infty} \frac{1}{n} \sum_{\varepsilon n \leq x_1 < \ldots < x_t \leq n} \prod_{i=1}^{t} \frac{1}{x_i^{d_i/2}}.$$

Note the above limit agrees with the earlier definition of $\psi(D, \varepsilon)$:

$$\psi(D, \varepsilon) = \frac{1}{(2m)^{t-1}} \lim_{n \to \infty} \frac{1}{n} \sum_{\varepsilon n \leq x_1 < \ldots < x_t \leq n} \prod_{i=1}^{t} \frac{1}{x_i^{d_i/2}}$$

$$= \frac{1}{(2m)^{t-1}} \lim_{n \to \infty} \frac{1}{n} (1 \pm o(1)) \int_{\varepsilon n}^{n} \int_{\varepsilon n}^{x_t} \ldots \int_{\varepsilon n}^{x_2} \prod_{i=1}^{t} \frac{1}{x_i^{d_i/2}} dx_1 \ldots dx_t \qquad \text{substitute } x_i = ny_i:$$

$$= \frac{1}{(2m)^{t-1}} \left(\int_{\varepsilon}^{1} \int_{\varepsilon}^{y_t} \ldots \int_{\varepsilon}^{y_2} \prod_{i=1}^{t} \frac{1}{y_i^{d_i/2}} dy_1 \ldots dy_t \right) \left(\lim_{n \to \infty} \frac{(1 \pm o(1)) n^t}{n \cdot n^{\sum_i d_i/2}} \right)$$

$$= \frac{1}{(2m)^{t-1}} \int_\varepsilon^1 \int_\varepsilon^{y_t} \ldots \int_\varepsilon^{y_2} \prod_{i=1}^t \frac{1}{y_i^{d_i/2}} dy_1 \ldots dy_t.$$

It is clear that the last formula is a constant depending on m, ε, and $D(T)$. Recall the critical constants $C(k, \varepsilon, m)$, discussed in Section 2.4:

$$C(2k, \varepsilon, m) = \frac{1}{1-\varepsilon} \sum_{T \in \mathcal{T}_k} \varphi(T, m) \mathcal{M}_{2k}(T) \psi(D(T), \varepsilon)$$

$$C(2k+1, \varepsilon, m) = 0.$$

Apriori, the above might be a sequence of constants that do not correspond to the moments of any measure. Moreover, if they do correspond to a measure, they might not uniquely define it. However, this is not the case as shown by the following lemma:

Lemma 7.4. *Let $\{C(k, \varepsilon, m)\}_{k=0}^\infty$ be defined as above. Then:*

(1) *The sequence $\{C(k, \varepsilon, m)\}_{k \geq 0}$ is solvable.*
(2) *The sequence is also M-determinate.*

Proof of Lemma 7.4. We postpone the proof of part 1. For part 2, note that Carleman's condition requires for the moments to not "grow too fast". Write $C_{2k} = C(2k, \varepsilon, m)$ for shorthand. We will show:

$$\limsup_{k \to \infty} \frac{C_{2k}^{1/2k}}{2k} \leq C$$

for some constant C. We will assume that $\mathcal{T}_k$ is the set of labeled trees with exactly k edges (as opposed to the set of trees with *at most* k edges, to see why this is enough see Appendix). To bound $\psi(D(T), \varepsilon)$ let $D(T) = (d_1, \ldots, d_{k+1})$. Then:

$$\psi(D(T), \varepsilon) = \frac{1}{2^k} \lim_{n \to \infty} \frac{1}{n} \sum_{\varepsilon n \leq x_1 < \ldots \leq x_{k+1} \leq n} \prod_{i=1}^{k+1} \frac{1}{x_i^{d_i/2}}$$

$$\leq \frac{1}{2^k} \lim_{n \to \infty} \frac{1}{n} \prod_{i=1}^{k+1} \sum_{x=\varepsilon n}^n \frac{1}{x^{d_i/2}}$$

Note that

$$(25) \qquad \sum_{x=\varepsilon n}^n \frac{1}{x^{d_i/2}} \leq \begin{cases} 2n^{1-d_i/2} & \text{if } d_i = 1 \\ \log(1/\varepsilon)n^{1-d_i/2} & \text{if } d_i = 2 \\ 2n^{1-d_i/2}\varepsilon^{-d_i/2} & \text{if } d_i \geq 3 \end{cases}$$

Thus,

$$\psi(D(T), \varepsilon) \leq \frac{1}{2^k} \lim_{n \to \infty} \frac{1}{n} \left(\frac{2\log(1/\varepsilon)}{\sqrt{\varepsilon}} \right)^{k+1} n^{k+1-\sum_i d_i/2}$$

$$= \frac{1}{2^k} \left(\frac{2\log(1/\varepsilon)}{\sqrt{\varepsilon}} \right)^{k+1} \lim_{n\to\infty} \frac{n^{k+1-k}}{n}$$

$$= \frac{1}{2^k} \left(\frac{2\log(1/\varepsilon)}{\sqrt{\varepsilon}} \right)^{k+1}.$$

Hence, $\psi(D(T), \varepsilon) \le \tilde{C}^k$ for an appropiate constant $\tilde{C}$. Hence, we focus on bounding:

$$\sum_{T \in T_k} \varphi(T, m) \mathcal{M}_{2k}(T)$$

In the Appendix, we show the following combinatorial identity:

$$\tag{26} \mathcal{M}_{2k}(T) = (2k) \prod_{v \in V(T)} (d(v,T) - 1)!$$

Also note that

$$\varphi(T, m) \le \prod_{v \in V(T)} [m]^{d(v,T)} \le (m+1)^{2k} \prod_{v \in V(T)} (d(v,T) - 1)!$$

Hence, we can bound the terms by functions that only depend on the degree sequence of T. Let $\mathcal{D}$ denote the set of degree sequences of trees on $k+1$ vertices. Then,

$$\sum_{T \in T_k} \varphi(T, m) \mathcal{M}_{2k}(T) \le 2k(m+1)^{2k} \sum_{D \in \mathcal{D}} |\{T \in \mathcal{T}_k : D(T) = D\}| \left(\prod_{d_i \in D} (d_i - 1)! \right)^2$$

By Lemma 3.10,

$$|\{T \in \mathcal{T}_k : D(T) = D\}| = \binom{k-1}{d_1 - 1, d_2 - 1, \ldots, d_{k+1} - 1}$$

Thus,

$$\sum_{T \in \mathcal{T}_k} \mathcal{M}_{2k}(T) \varphi(T, m) \le 2k(m+1)^{2k} (k-1)! \sum_{D \in \mathcal{D}} \prod_{d_i \in D} (d_i - 1)!$$

$$= 2k(m+1)^{2k} ((k-1)!)^2 \sum_{D \in \mathcal{D}} \frac{1}{\binom{k-1}{d_1-1, d_2-1, \ldots, d_{k+1}-1}}$$

$$\le 2k(m+1)^{2k} ((k-1)!)^2 |\mathcal{D}|$$

$$= 2k(m+1)^{2k} ((k-1)!)^2 \binom{2k}{k+1}$$

$$\le 2k(2(m+1))^{2k} ((k-1)!)^2$$

$$= 2k(m+1)^{2k} ((k-1)!)^2$$

$$\tag{27} \le 2k(2k(m+1))^{2k}.$$

Hence, we can bound the desired limit by $\sqrt{\tilde{C}}(m+1)$ which proves the lemma.

$\square$

Theorem 7.5. *Let $0 < \varepsilon < 1$ be fixed. Let $G_{\varepsilon,m,n}$ be the ε-truncation of a random PA-graph. Let $\{C(k,\varepsilon,m)\}$ be the sequence of constants as defined above. Then for any fixed k we have:*

$$\mathbf{E}\left[\bar{\mu}_{\varepsilon,m,n}^k\right] \longrightarrow C(k,\varepsilon,m).$$

Proof. We need to compute

$$\mathbf{E}[\mathrm{Tr}(A_{G_{\varepsilon,m,n}}^k)]$$

From equation (9), we see that this is equivalent to

(28) $$\mathbf{E}[\mathrm{Tr}(A_{G_{\varepsilon,m,n}}^k)] = \frac{1}{n - \varepsilon n} \sum_{H \in \mathcal{F}(K_{m,n-\varepsilon n})} \mathcal{M}_k(H)\mathbf{E}[X(H,G_{\varepsilon,m,n})]$$

Let $H \in \mathcal{F}(K_{m,n-\varepsilon n})$ be a graph on t vertices. Let $V = \{x_1 < \ldots < x_t\}$ be a subset of $[\varepsilon n, n]$. Then by equation (19) we have:

$$\mathbf{E}[X(H,V)] = \left(1 \pm O\left(\frac{1}{\varepsilon n}\right)\right) \frac{\varphi(H,m)}{(2m)^{|E(H)|}} \prod_{i=1}^t \frac{1}{x_i^{d(v_i,H)/2}}$$

Hence,

$$\mathbf{E}[X(H,G_{\varepsilon,m,n})] = \Theta\left(\sum_{\varepsilon n \le x_1 < \ldots < x_t \le n} \prod_{i=1}^t \frac{1}{x_i^{d(v_i,H)/2}}\right) = O\left(\prod_{i=1}^t \sum_{x=\varepsilon n}^n \frac{1}{x^{d(v_i,H)/2}}\right)$$

Using the same bounds as in equation (25), we obtain:

$$\mathbf{E}[X(H,G_{\varepsilon,m,n})] = O\left(n^{t-\sum_i d(v_i,H)/2}\right) = O\left(n^{t-|E(H)|}\right)$$

Since H is a connected graph on t vertices we have $t - |E(H)| \le 1$, with equality if and only if H is isomorphic to a tree. Hence, in equation (28) if we consider the sum over graphs which are not isomorphic to trees we will get a contribution of $O(1/n)$. In particular, note that all the odd moments vanish (since they must contain a cycle). Note that if H is an ordered tree on $t+1$ vertices then we have:

$$\frac{1}{n}\mathbf{E}[X(H,G_{\varepsilon,m,n})] = \frac{1}{n}(1 \pm o(1)) \sum_{\varepsilon n \le x_1 < \ldots < x_{t+1} \le n} \frac{\varphi(H,m)}{(2m)^t} \prod_{i=1}^{t+1} \frac{1}{x_i^{d(v_i,H)/2}}$$

$$\to \varphi(H,m)\psi(D(H),\varepsilon)$$

So to compute the even moments, we obtain:

$$\mathbf{E}[\mathrm{Tr}(A_{G_{\varepsilon,m,n}}^{2k})] = o(1) + \frac{1}{1-\varepsilon} \sum_{H \in \mathcal{T}_k} \mathcal{M}_{2k}(H)\frac{1}{n}\mathbf{E}[X(H,G_{\varepsilon,m,n})]$$

$$\to \frac{1}{1-\varepsilon} \sum_{H \in \mathcal{T}_k} \mathcal{M}_{2k}(H)\varphi(H,m)\psi(D(H),\varepsilon)$$

$$= C(2k,\varepsilon,m)$$

as desired. □

Now that we have shown that our sequence of constants correspond to the limiting moments of the expected ESD, then we have that Hamburguer's criterion must be satisfied: The sequence $\{\bar{\mu}_{\varepsilon,m,n}\}_{n=1}^{\infty}$ satisfies equation (20). In particular, it follows that their limit, $C(k,\varepsilon,m)$, also satisfies (20). Since we already showed that the sequence is M-determinate, we have:

Corollary 7.6. *There exists a unique probability measure with moments equal to $\{C(k,\varepsilon,m)\}_{k\geq 0}$.*

We denote this measure by $\mu_{\varepsilon,m,\infty}$. Hence, to finish the proof of Theorem 2.2, we need to show that $\{\mu_{\varepsilon,m,n}\}$ converge weakly in probability to $\mu_{\varepsilon,m,\infty}$. We do this by showing that the random measures are concentrated around the expected measure.

Lemma 7.7. *Let $G_{\varepsilon,m,n}$ be an ε-truncated PA random graph, and denote its ESD by $\mu_{\varepsilon,m,n}$. Then for any fixed $k,\delta > 0$*

$$\lim_{n\to\infty} \mathbf{P}\left[\left|\mathbf{E}[\mu_{\varepsilon,m,n}^k] - \mathbf{E}[\bar{\mu}_{\varepsilon,m,n}^k]\right| > \delta\right] = 0$$

Proof. We compute the second moment as follows:

$$\mathbf{E}\left[(\mathbf{E}[\mu_{\varepsilon,m,n}^k] - \mathbf{E}[\bar{\mu}_{\varepsilon,m,n}^k])^2\right] = \mathbf{E}\left[\left(\mathbf{E}[\mu_{\varepsilon,m,n}^k]\right)^2\right] - \left(\mathbf{E}[\bar{\mu}_{\varepsilon,m,n}^k]\right)^2$$

$$= \mathbf{E}\left[\frac{1}{(n-\varepsilon n)^2}\left(\mathrm{Tr}(A_{G_{\varepsilon,m,n}}^k)\right)^2\right] - \left(\frac{1}{n-\varepsilon n}\mathbf{E}\left[\mathrm{Tr}(A_{G_{\varepsilon,m,n}}^k)\right]\right)^2$$

$$= \frac{1}{(n-\varepsilon n)^2}\left(\mathbf{E}\left[\left(\sum_{H\in\mathcal{C}(G_{\varepsilon,m,n})}\mathcal{M}_k(H)\mathbb{I}_H\right)^2\right] - \left(\sum_{H\in\mathcal{C}(G_{\varepsilon,m,n})}\mathcal{M}_k(H)\mathbf{E}[\mathbb{I}_H]\right)^2\right)$$

$$= \frac{1}{(n-\varepsilon n)^2}\sum_{H_1,H_2\in\mathcal{C}(G_{\varepsilon,m,n})}\mathcal{M}_k(H_1)\mathcal{M}_k(H_2)\left(\mathbf{E}[\mathbb{I}_{H_1}\mathbb{I}_{H_2}] - \mathbf{E}[\mathbb{I}_{H_1}]\mathbf{E}[\mathbb{I}_{H_2}]\right)$$

If H_1 and H_2 are vertex disjoint, then by Lemma 5.4 we have

$$\mathbf{E}[\mathbb{I}_{H_1}\mathbb{I}_{H_2}] - \mathbf{E}[\mathbb{I}_{H_1}]\mathbf{E}[\mathbb{I}_{H_2}] \leq 0$$

If H_1 and H_2 are not vertex disjoint, then $\mathbb{I}_{H_1}\mathbb{I}_{H_2}$ corresponds to the indicator random variable that a connected labeled graph appears. As we saw in the above proof, connected graphs give a contribution of $O(n)$. Thus,

$$\sum_{H_1\cap H_2\neq\emptyset}\mathcal{M}_k(H_1)\mathcal{M}_k(H_2)\left(\mathbf{E}[\mathbb{I}_{H_1}\mathbb{I}_{H_2}] - \mathbf{E}[\mathbb{I}_{H_1}]\mathbf{E}[\mathbb{I}_{H_2}]\right)$$

$$\leq \sum_{H_1\cap H_2\neq\emptyset}\mathcal{M}_k(H_1)\mathcal{M}_k(H_2)\mathbf{E}[\mathbb{I}_{H_1}\mathbb{I}_{H_2}]$$

$$= O(n).$$

We thus obtain

$$\mathbf{E}\left[(\mathbf{E}[\mu_{\varepsilon,m,n}^k] - \mathbf{E}[\bar{\mu}_{\varepsilon,m,n}^k])^2\right] = O(n^{-1})$$

which implies the lemma via Chebyshev's inequality.   □

We finish off the proof of Theorem 2.2:

Proof. By Lemma 3.9, it is enough to show that the k-th moments of $\{\mu_{\varepsilon,m,n}\}$ converge in probability to the k-th moments of $\mu_{\varepsilon,m,\infty}$. By Theorem 7.5:

$$|\mathbf{E}[\bar{\mu}_{\varepsilon,m,n}^k - C(k,\varepsilon,m)| \to 0$$

deterministically. By Lemma 7.7:

$$|\mathbf{E}[\mu_{\varepsilon,m,n}^k] - \mathbf{E}[\bar{\mu}_{\varepsilon,m,n}^k]| \to 0$$

in probability. We get,

$$|\mathbf{E}[\mu_{\varepsilon,m,n}^k] - C(k,\varepsilon,m)| \to 0$$

in probability as desired.

                                                                                       □

8. EDGE EIGENVALUES

In this section we prove Theorem 2.9. As we saw in previous sections, the main contribution towards the even moments come from stars where the high degree appears first. It turns out, that these stars will also dictate the behavior of the largest eigenvalues.

In order to show Theorem 2.9, we will need control on the degrees, in particular we need to show they do not deviate too much from their mean. One could prove these results using the tools developed above, but fortunately some of the bounds needed were already done in [12], so rather than reprove them we will simply state them.

For a graph G, we will write $\|G\|$ for the operator norm of the adjacency matrix. Note that this quantity is always equal to $\lambda_1(G)$ by the Perron-Frobenius Theorem.

Theorem 8.1 (Theorem 1.1 in [12]). *Let $t \to \infty$. Then whp we have $\Delta_1(G_{m,t}) \leq \sqrt{t}\log t$.*

Theorem 8.2 (Theorem 1.2 in [12]). *Let $t \to \infty$. Then whp we have $\|G_{m,t}\| \leq t^{1/4}\sqrt{\log t}$.*

Theorem 8.3 (Claim 2.6 in [12]). *Let $\delta > 0$ be fixed. Then for any $s := s(n)$ such that $s \to \infty$ as $n \to \infty$ we have that for all $v \in [s,n]$ and $r \in [s,v]$ we have $d(r,G_{m,v}) \leq n^\delta\sqrt{v/r}$ with probability at least $1 - O(s^{-C})$ for any arbitrary constant C.*

Remark 8.4. *The above theorem is stated slightly differently in [12]. To obtain the form stated above, change ℓ (as in their proof) to be an integer satisfying $2 - \delta\ell \leq -C$, and then keep track of the error probability.*

Lastly, we will be using a result from [19] which gives us explicit rates of convergence from the degrees to their limiting distribution. To describe this limiting distribution, we first define

$$\kappa_s(x) = \Gamma(s)\sqrt{\frac{2}{s\pi}} \exp\left(\frac{-x^2}{2s}\right) U\left(s-1, \frac{1}{2}, \frac{x^2}{2s}\right)$$

where $\Gamma(s)$ denotes the gamma function, and $U(a, b, z)$ denotes the confluent hypergeometric function of the second kind. For more information about the above distribution and its properties, see [19] and references therein. Denote by $d_K(\mu_1, \mu_2)$ to be the Kolmogorov distance between two probability measures defined as:

$$d_K(\mu_1, \mu_2) = \sup_{x \in \mathbb{R}} |\mu_1(-\infty, x], -\mu_2(-\infty, x]|$$

Theorem 8.5 (Theorem 1.1 in [19]). *Let $b_{N,i} = \sqrt{\mathbf{E}[d(i, G_{1,N})^2]}$. For $1 \leq i \leq N$ and some constants $c, C > 0$ independent of N, we have:*

$$\frac{c}{\sqrt{N}} \leq d_K\left(\frac{d(i, G_{1,N})}{b_{N,i}}, \kappa_{i-1/2}\right) \leq \frac{C}{\sqrt{N}}$$

In particular, the only property we will need from $\kappa_{i-1/2}(x)$ is that they are uniformly bounded: For all positive integers i, we have $\kappa_{i-1/2}(x) \leq 2$ for all $x \in [0, \infty)$. We present our final tool,

Theorem 8.6. *Let $\omega(n) \to \infty$. Then, for all $i \in [1, k]$ we have*

$$d(i, G_{m,n}) \in \left[\frac{\mathbf{E}d(i, G_{m,n})}{\omega(n)}, \omega(n)\mathbf{E}d(i, G_{m,n})\right]$$

with probability at least $1 - O(k/\omega(n) + k/\sqrt{n})$.

Proof. A calculation in the Appendix shows

$$\mathbf{E}[d(i, G_{1,n})] = \Theta\left(\sqrt{\frac{n}{i}}\right) \tag{29}$$

$$b_{n,i} = \Theta\left(\sqrt{\frac{n}{i}}\right). \tag{30}$$

Note that $d(i, G_{m,n}) = \sum_{j=(i-1)m+1}^{im} d(j, G_{1,mn})$, so we obtain $\mathbf{E}d(i, G_{m,n}) = \Theta(\sqrt{n/i})$. Let $\tilde{C}$ be an absolute constant such that for any $i \in [n]$ and any $j \in [(i-1)m+1, im]$ we have:

$$\frac{\mathbf{E}d(i, G_{m,n})}{b_{mn,j}} \leq \tilde{C}.$$

For any fixed function $\omega(n)$, we have:

$$\mathbf{P}\left[d(i, G_{m,n}) \leq \frac{\mathbf{E}d(i, G_{m,n})}{\omega(n)}\right] \leq \sum_{j=(i-1)m+1}^{im} \mathbf{P}\left[d(j, G_{1,mn}) \leq \frac{\mathbf{E}d(i, G_{m,n})}{\omega(n)}\right]$$

$$= \sum_{j=(i-1)m+1}^{im} \mathbf{P}\left[\frac{d(j, G_{1,mn})}{b_{mn,j}} \leq \frac{\mathbf{E}d(i, G_{m,n})}{b_{mn,j}\omega(n)}\right]$$

$$\leq \sum_{j=(i-1)m+1}^{im} \mathbf{P}\left[\frac{d(j, G_{1,mn})}{b_{mn,j}} \leq \frac{\tilde{C}}{\omega(n)}\right]$$

$$\leq \sum_{j=(i-1)m+1}^{im} \frac{C}{\sqrt{mn}} + \int_0^{\frac{\tilde{C}}{\omega(n)}} \kappa_{j-1/2}(x)dx$$

$$= O\left(\frac{1}{\sqrt{n}} + \frac{1}{\omega(n)}\right)$$

Hence, with probability at least $1 - O\left(\frac{k}{\sqrt{n}} + \frac{k}{\omega(n)}\right)$ we have that for all $i \leq k$:

$$d(i, G_{m,n}) \geq \frac{\mathbf{E}d(i, G_{m,n})}{\omega(n)}.$$

For the upper bound one can just apply Markov's to obtain:

$$\mathbf{P}\left[d(i, G_{m,n}) \geq \omega(n)\mathbf{E}d(i, G_{m,n})\right] \leq \frac{1}{\omega(n)}.$$

Taking a union bound gives the desired result. $\qquad\square$

8.1. **Proof of Theorem 2.9.** The proof will be similar in nature to the one done in [8]. We are going to split the graph into multiple subgraphs, and we are going to show that the eigenvalues can be deduced from a very well structured subgraph, and the rest can be viewed as "noise". For ease of notation, we will simply write G for $G_{m,n}$. Throughout the proof, we will have the following notation

$$\delta = 1/10000 \qquad k = n^{1/25} \qquad b = n^{1/20} \qquad s = n^{1/7} \qquad t = n^{13/25}.$$

Let $S = [1, s]$ and $T = [t, n]$. Consider the following subgraphs of G defined as follows:

- $G_1 = G[1, t]$.
- $G_2 = G[s, n]$.
- $G(S, T)$ is the bipartite graph spanned by the vertex partition $[1, s]$ and $[t, n]$.
- G_3 is the subgraph of $G(S, T)$ spanned by the vertices in T which have more than one neighbor in S. If we have a vertex $v \in T$ which connects to a $u \in S$ through two distinct edges, we consider t having two neighbors in S, so G_3 will contain all the parallel edges of $G(S, T)$.
- G_4 is the complement of G_3 in $G(S, T)$.

The advantage of the decomposition above is that we obtain G_4 which is very well structured: It is a union of s disjoint stars, where we have one star for every $u \in S$. Hence, for $i \leq k$ we

will have:

$$\lambda_i(G_4) = \sqrt{\Delta_i(G_4)}.$$

We will show that the eigenvalues for G can be deduced from the eigenvalues of G_4. This consists of two parts:

(1) The subgraphs G_1, G_2 and G_3 have negligible norm.
(2) For a vertex $u \in S$, the degree in G_4 is close to the degree in G. That is, $d(u, G_4) \approx d(u, G)$.

Thus, we define the "loss" of a vertex $u \in S$ as:

$$L(u) = d(u, G_1) + d(u, G_3).$$

Note that we obtain the following relation:

$$d(u, G) = d(u, G_4) + L(u).$$

Once we have shown that the norms of G_1, G_2, G_3 are negligible, we can apply Weyl's inequality: how the required bounds through the following five steps:

Step 1: Whp we have that for all $i \leq k$: $d(i, G) \geq \mathbf{E}[d(i, G)]/b$.

Step 2: Whp we have $\|G_1\| \leq t^{1/4} \log t$.

Step 3: Whp we have $\|G_2\| = O(n^\delta (n/s)^{1/4})$.

Step 4: Whp we have that for all $u \in S$: $L(u) \leq s^{1.5} \log^2 n + t^{1/2} \log t$.

Step 5: Whp we have $\|G_3\| \leq s^{1.5} \log^2 n$.

Assuming Steps 1 - 5, the proof of Theorem 2.9 follows readily:

Proof of Theorem 2.9. Recall that

$$\mathbf{E}[d(i, G)] = \Theta\left(\sqrt{\frac{n}{i}}\right)$$

which implies by Step 1 that whp we have:

$$\Delta_i(G) \geq \frac{C}{b}\sqrt{\frac{n}{i}} \geq \frac{C}{b}\sqrt{\frac{n}{k}}$$

for some appropriate constant C. Note that by Step 4 and the above lower bound, we have that whp for all $i \leq k$:

$$
\begin{aligned}
\Delta_i(G_4) &\geq \Delta_i(G) - \max_{u \in S} L(u) \\
&\geq \Delta_i(G) - (s^{1.5} \log^2 n + t^{1/2} \log t) \\
&= (1 - o(1))\Delta_i(G).
\end{aligned}
$$

(31)

As $\Delta_i(G) \geq \Delta_i(G_4)$, we have $\Delta_i(G_4) = (1 - o(1))\Delta_i(G)$. Since G_4 is a union of disjoint stars, we have that $\lambda_i(G_4) = \sqrt{\Delta_i(G_4)}$, which means:

$$\lambda_i(G_4) = (1 - o(1))\sqrt{\Delta_i(G)}.$$

By using our lower bound on $\Delta_i(G)$ and the fact that $i \leq k$, we have:

$$\lambda_i(G_4) = (1 - o(1))\sqrt{\Delta_i(G)} \geq (1 - o(1))\sqrt{\Delta_k(G)} \geq \frac{C}{2}\left(\frac{n^{1/4}}{b^{1/2}k^{1/4}}\right) = n^{43/200}.$$

By Weyl's inequality we have:

$$|\lambda_i(G_4) - \lambda_i(G)| \leq \|G_1\| + \|G_2\| + \|G_3\|$$

which using Steps 2, 3 and 5, we have that whp:

$$\begin{aligned}
\|G_1\| + \|G_2\| + \|G_3\| &\leq t^{1/4}\log t + O(n^\delta (n/s)^{1/4}) + s^{1.5}\log^2 n \\
&= O(n^\delta n^{6/28}) \\
(32) \qquad\qquad &= o\left(\left(\frac{1}{b}\sqrt{\frac{n}{k}}\right)^{1/2}\right) \\
&= o(\lambda_k(G_4)).
\end{aligned}$$

Thus,

$$|\lambda_i(G_4) - \lambda_i(G)| = o(\lambda_i(G_4))$$

Combining all of the above, we obtain that whp for all $i \leq k$:

$$\begin{aligned}
\lambda_i(G) &= (1 \pm o(1))\lambda_i(G_4) \\
&= (1 \pm o(1))\sqrt{\Delta_i(G_4)} \\
&= (1 \pm o(1))\sqrt{\Delta_i(G)}
\end{aligned}$$

just as desired.   □

We now prove each one of the required Steps.

Step 1:   Follows from Theorem 8.6 by letting $\omega(n) = b$.

Step 2:   Follows from Theorem 8.2.

Step 3:   We will be using equation (10). Our matrix A will be the adjacency matrix for G_2, where for ease of notation the rows and columns are labeled $[s, n]$. Let $r = \sqrt{ns}$, $c_i = \frac{n^{1/4}}{i^{1/4}}$ for $i \in [s, r]$, and $c_i = 1$ for $i \in (r, n]$. Define,

$$R_i = \frac{1}{c_i}\sum_{j=s}^{n} c_j a_{ij}.$$

It suffices to show that for all $i \in [s, n]$ we have $R_i = O(n^\delta (n/s)^{1/4})$ with probability $1 - o(n^{-1})$, so the result would follow from taking a union bound. Let C be a large enough constant such that $s^{-C} = o(n^{-1})$. Then, using Theorem 8.3 we have:

Case $i \in [s, r]$: We have

$$
\begin{aligned}
R_i &= \frac{1}{c_i} \left(\sum_{j=s}^{i} c_j a_{ij} + \sum_{j=i+1}^{r} c_j a_{ij} + \sum_{j=r+1}^{n} c_j a_{ij} \right) \\
&\leq \frac{mc_s}{c_i} + d(i, G_{m,r}) + \frac{d(i, G_{m,n})}{c_i} \\
&\leq \frac{mc_s}{c_i} + \sqrt{\frac{r}{i}} r^\delta + \frac{1}{c_i} \sqrt{\frac{n}{i}} n^\delta \\
&= O \left(\left(\frac{n}{s} \right)^{1/4} n^\delta \right)
\end{aligned}
$$

occurs with probability $1 - O(s^{-C}) = 1 - o(n^{-1})$.

Case $i \in (r, n]$: In this case one has:

$$
\begin{aligned}
R_i &= \sum_{j=s}^{i} c_j a_{ij} + \sum_{j=i+1}^{n} a_{ij} \\
&\leq mc_s + d(i, G_{m,n}) \\
&= m \sqrt[4]{\frac{n}{s}} + \sqrt{\frac{n}{r}} n^\delta \\
&= O \left(\left(\frac{n}{s} \right)^{1/4} n^\delta \right)
\end{aligned}
$$

occurs with probability $1 - O(r^{-C}) = 1 - o(n^{-1})$. Hence, whp we have $R_i = O(n^\delta (n/s)^{1/4})$ for all $i \in [s, n]$. Thus, whp we have:

$$
\|G_2\| = O \left(\left(\frac{n}{s} \right)^{1/4} n^\delta \right).
$$

Step 4: For $u \in S$, we have:

$$
(33) \qquad\qquad d(u, G_1) \leq \Delta_1(G_1) \leq \sqrt{t} \log t
$$

whp by Theorem 8.1. Let H_1 be the following ordered graph $\{(v_1, v_3), (v_2, v_3)\}$, and let $H_2 = \{(v_1, v_2), (v_1, v_2)\}$. Then,

$$d(u, G_3) \leq \sum_{x_2 \in S, x_3 \in T} X_{m,n}(H_1, \{u, x_2, x_3\}) + 2 \sum_{x_2 \in T} X_{m,n}(H_2, \{u, x_2\}).$$

By computing the above graph counts (see Appendix) we obtain:

(34) $$\mathbf{E}[d(u, G_3)] = O(\sqrt{s} \log n)$$

Hence, for an appropriate constant C (independent of n), we have:

$$\mathbf{P}[d(u, G_3) \geq s^{1.5} \log^2 n] \leq \mathbf{P}[d(u, G_3) \geq Cs \log n \mathbf{E}[d(u, G_3)]]$$

$$\leq \frac{1}{Cs \log n}.$$

By taking a union bound over all $u \in S$, we obtain that with high probability:

$$d(u, G_3) \leq s^{1.5} \log^2 n.$$

Combining with equation (33) we obtain:

$$L(u) \leq s^{1.5} \log^2 n + \sqrt{t} \log t.$$

Step 5: We will use the fact that the norm of a graph is always bounded by the maximum degree. Recall, G_3 is a bipartite graph where for every $v \in T$ we have that $d(v, G_3) \leq m$ (by definition of our model), and for $u \in S$ we have $d(u, G_3) \leq s^{1.5} \log^2 n$ (by above). Hence, whp:

$$\|G_3\| \leq s^{1.5} \log^2 n.$$

This concludes the proof of Theorem 2.9. Now we proceed to study the largest eigenvectors.

9. Localization of Eigenvectors

In this section, we prove Theorem 2.8. The strategy is as follows: We will use the decomposition and all the notation we used in the previous section (e.g., $k, t, G_1, G_2, \ldots$). Let $G_{m,n}$ be a random PA graph, and let A_n denote its adjacency matrix. Let B_n denote the adjacency matrix of G_4, and let $C_n := A_n - B_n$. From the Davis-Kahan Theorem (Theorem 4.4), we see that the two quantities one needs to control when studying eigenvectors of perturbations are the eigenvalue gap and the size of the noise. We present the following bounds on the two quantities:

Lemma 9.1. *Let C_n be defined as above. Then whp:*

$$\|C_n\| = O(n^{43/200}).$$

Proof. C_n corresponds to the adjacency matrix of the complement of G_4 in $G_{m,n}$. As we saw in the proof of Theorem 2.9, we see that this complement is contained in $G_1 \cup G_2 \cup G_3$. Hence, by equation (32) we obtain whp

$$\|C_n\| \leq \|G_1\| + \|C_3\| + \|G_3\|$$

$$= O(n^{43/200}).$$

$\square$

Recall that the eigenvalues of B_n arise from the (square root of) max degrees. Hence, to control the eigenvalue gap, we need to control the maximum degree gap. We use the following theorem:

Theorem 9.2 (Theorem 1 in [12])**.** *Let K be some fixed constant. We have whp:*

$$\frac{\sqrt{n}}{\log n} \leq \Delta_i(G) \leq \Delta_{i-1}(G) - \frac{\sqrt{n}}{\log n}$$

$$\frac{\sqrt{n}}{\log n} \leq \Delta_i(G).$$

We can now show a lower bound on the eigenvalue gap:

Lemma 9.3. *Using the notation from above, for $i \leq K$ we have whp:*

$$\lambda_i(B_n) - \lambda_{i+1}(B_n) \geq \frac{n^{1/4}}{\log^3 n}.$$

For a proof of the above lemma, see Appendix. With these two bounds, we can present the proof.

Proof of Theorem 2.8. Since B_n is the union of disjoint stars we have that for $i \leq K$ whp:

$$\|w_i\|_\infty = \frac{1}{\sqrt{2}}$$

and all coordinates which do not realize the infinity norm are equal to $1/\sqrt{2\Delta_i(G_4)} = o(1)$. Hence, if we show:

$$\|v_i - w_i\|_2 = o(1)$$

then the proof would follow. First of all note that for two unit vectors u and v, whose angle θ is in $[0, \pi/2]$ we have:

$$\|u - v\|_2^2 \leq 2\sin^2(\theta).$$

Hence, it suffices to show that the (sine of the) angle between w_i and v_i is $o(1)$. By Lemmas 9.1 and 9.3:

$$\|C_n\| = O(n^{43/200})$$

$$\lambda_i(B_n) - \lambda_{i+1}(B_n) \geq \frac{n^{1/4}}{\log^3 n}$$

occur whp for all $i \leq K$. Hence,

$$|\lambda_{i-1}(A_n) - \lambda_i(B_n)| = |\lambda_{i-1}(A_n) - \lambda_{i-1}(B_n) + \lambda_{i-1}(B_n) - \lambda_i(B_n)|$$

$$\geq \frac{n^{1/4}}{2\log^3 n}.$$

Similarly, we can bound $|\lambda_i(B_n) - \lambda_{i+1}(A_n)|$ from below by the same quantity. Thus, by Theorem 4.4:

$$\sin \angle(v_i, w_i) = O\left(\frac{\log^3 n}{n^{7/200}}\right)$$

which implies the desired result.

□

References

[1] R. Albert, and A. Barabási. *Statistical mechanics of complex networks*, Reviews of Modern Physics, Volume 74, January 2002.

[2] R. Albert, H. Jeong and A.-L. Barabási. *Diameter of the world-wide web*, Nature, 401 (1999), 130131.

[3] R. Albert. *Scale-free networks in cell biology.* Journal of Cell Science 118, 4947-4957 Published by The Company of Biologists 2005.

[4] A. Barabási, R. Albert, and H. Jeong*Scale-free characteristics of random networks: the topology of the world-wide web*, Physica A 281 (2000), 6977.

[5] B. Bollobás. *Mathematical results on scale-free random graphs.*

[6] B. Bollobás, and O. Riordan. *The Diameter of a Scale-Free random graph.* Journal of Combinatorial Theory, Series B 97 (2007) 269290 Combinatorica 24(1)(2004), 5-34.

[7] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnady. *The degree sequence of a scale-free random graph process.* Random Structures and Algorithms (2001), 18, 79-90.

[8] F. Chung, L. Lu, and V. Vu. *Eigenvalues of Random Power Law Graphs.* Annals of Combinatorics 7 (2003), 21-33.

[9] F. Chung, L. Lu, and V. Vu *Spectra of random graphs with given expected degrees. PNAS*(2003), 100, 6313-6318.

[10] N. Eikmeier, and D. Gleich. *Revisiting Power-law Distributions in Spectra of RealWorld Networks.* KDD17, August 1317, 2017, Halifax, NS, Canada.

[11] I. Farkas, I. Derényi, A. Barabási and T. Vicsek. *Spectra of "real-world" graphs: Beyond the semicircle law.* Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, Volume 64.

[12] A. Flaxman, A. Frieze and T. Fenner. *High Degree Vertices and Eigenvalues in the Preferential Attachment Graph.* Internet Mathematics. Volume 2, 2005.

[13] A. Frieze, and M. Karoński. *Introduction to Random graphs.*

[14] J. Holland, and M. Handcock. *An assessment of preferential attachment as a mechanism for human sexual network formation.* Proc. R. Soc. Lond. B (2003) 270, 11231128.

[15] J. Moon. *On the maximum degree in a random tree.* Michigan Math. J. Volume 15, Issue 4 (1968), 429-432.

[16] T. Martin, X. Zhang, and M.E.J. Newman. *Localization and centrality in networks.*

[17] W. Powell, K. Koput, D. White, and J. Owen-Smith. *Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences.* AJS Volume 110 Number 4 (January 2005): 11321205.

[18] P. Pollner, G. Palla, and T. Vicsek. *Preferential attachment of communities: The same principle, but a higher level.* Europhys. Lett., 73 (3), pp. 478484 (2006).

[19] E. Pekös, A. Röllin, and N. Ross. *Degree asymptotics with rates for preferential attachment random graphs.* The Annals of Applied Probability, 2013, Vol. 23, No. 3, 1188-1218.

[20] V. M. Preciado and A. Jadbabaie, *Spectral analysis of dynamically evolving networks with linear preferential attachment,* 2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL, 2009, pp. 1293-1299.

[21] E. Wigner. *On the Distribution of the Roots of Certain Symmetric Matrices.* Ann. of Math. 67, 325-328, 1958.

[22] W. Xiao, M. Li, and G. Chen. *Small-World Features of Real-World Networks.* Journal of Communications and Networks, Vol. 19, No. 3, June 2017.

[23] Y. Yu, T. Wang, and R. Samworth. *A useful variant of the DavisKahan theorem for statisticians.* Biometrika, Volume 102, Issue 2, June 2015.

10. Appendix

Proof of lemma 3.6: Note that we have that the sequence $\{\mu_n\}$ is tight. Hence, by Prokhorov's theorem there exists a subsequence $\{\mu_{n_j}\}_{j=1}^\infty$ such that it converges weakly to a measure, call it $\tilde{\mu}$. Note that the moments of $\tilde{\mu}$ are equal to the moments of μ, but since they uniquely determine the measure, we have $\tilde{\mu} = \mu$. By Prokhorov's theorem, since every weakly convergent subsequence has the same limit μ, then the sequence $\{\mu_n\}$ converges weakly to μ.

Proof of lemma 3.9: Fix f to be a bounded continuous function. Let $X_n = \int f d\mu_n$ and $X = \int f d\mu$. We want to show that X_n converges in probability to X. This is equivalent to showing that for any subsequence $\{n_m\}$ one can find a further subsequence $\{n_{m_j}\}$ such that $X_{n_{m_j}}$ converges almost surely to X.

Given $\{n_m\}$, we extract a subsequence $\{n_{m_j}\}$ such that $\int x^k d\mu_{n_{m_j}}$ converges almost surely to $\int x^k d\mu$ for all k: This can be done by first choosing a subsequence in which the first moments converges a.s., then pick a subsequence where the second moment converges a.s., and continue inductively. Since μ is uniquely determined by its moments, we have that $\mu_{n_{m_j}}$ converges weakly to μ with probability equal to one. This implies that $X_{n_{m_j}}$ converges almost surely to X, just as desired.

Moreover part of Theorem 5.2: We wish to lower bound:

$$\sum_{1 \leq x_1 < ... < x_t \leq n} \Theta\left(\prod_{i=1}^t \frac{1}{x_i^{d(v_i,H)/2}}\right)$$

where we have $d(v_i, H) \geq d(v_{i+1}, H)$. Let I_1 and I_2 be such that: $d(v_i, H) \geq 3$ for $i \in [1, I_1]$, $d(v_i, H) = 2$ for $i \in [I_1 + 1, I_2]$ and $d(v_i, H) = 1$ for $i \in [I_2 + 1, t]$. Then,

$$\sum_{1 \leq x_1 < ... < x_t \leq n} \Theta\left(\prod_{i=1}^t \frac{1}{x_i^{d(v_i,H)/2}}\right) \geq \sum_{I_1+1 \leq x_{I_1+1} < ... < x_t \leq n} \Theta\left(\prod_{i=I_1+1}^t \frac{1}{x_i^{d(v_i,H)/2}}\right)$$

$$\geq \Theta \left(\left(\sum_{I_1+1\leq x_{I_1+1}<...<x_{I_2}\leq\sqrt{n}} \prod_{i=I_1+1}^{I_2} \frac{1}{x_i} \right) \left(\sum_{\sqrt{n}\leq x_{I_2+1}<...<x_t\leq n} \prod_{i=I_2+1}^{t} \frac{1}{\sqrt{x_i}} \right) \right)$$

$$= \Theta(n^{f(H)/2} \log^{g(H)} n).$$

Proof of equations (29) **and** (30)**:** We have

$$\mathbf{E}[d(i, G_{1,n})] = 1 + \sum_{j=i+1}^{n} \mathbf{P}[(i,j) \in G_{1,n}]$$

$$= 1 + \Theta \left(\sum_{j=i+1}^{n} \frac{1}{\sqrt{ij}} \right)$$

$$= \Theta \left(\sqrt{\frac{n}{i}} \right).$$

Moreover,

$$b_{n,i}^2 = \mathbf{E}[d(i, G_{1,n})^2]$$

$$= \mathbf{E}\left[\left(1 + \sum_{j=i+1}^{n} \mathbb{I}_{(i,j)\in G_{1,n}} \right)^2 \right]$$

$$= 1 + \mathbf{E}\left[3d(i, G_{1,n}) + 2 \sum_{i+1\leq j<j'\leq n} \mathbb{I}_{(i,j)\in G_{1,n}}\mathbb{I}_{(i,j')\in G_{1,n}} \right]$$

$$= 1 + \Theta \left(\sqrt{\frac{n}{i}} \right) + 2 \sum_{i+1<j<j'\leq n} \mathbf{P}[(i,j),(i,j') \in G_{1,n}]$$

$$= 1 + \Theta \left(\sqrt{\frac{n}{i}} \right) + 2 \sum_{i+1<j<j'\leq n} \Theta \left(\frac{1}{i\sqrt{jj'}} \right)$$

$$= \Theta \left(\frac{n}{i} \right).$$

Taking a square root yields equation (30).

Proof of equation (34)**:** Let $V = \{x_1, x_2, x_3\}$. Then from equation (19) we obtain the following upper bound:

$$\mathbf{E}[X_{m,n}(H_1, V)] = O \left(\frac{1}{\sqrt{x_1 x_2 x_3}} \right)$$

Hence,

$$\sum_{x_2\in S, x_3\in T} X_{m,n}(H_1, \{u, x_2, x_3\}) = O \left(\sum_{x_2\in S, x_3\in T} \frac{1}{\sqrt{u x_2 x_3}} \right)$$

$$= O(\sqrt{s}\log n)$$

Similarly for H_2, if $V = \{x_1, x_2\}$ then:

$$\mathbf{E}[X_{m,n}(H_2, V)] = O\left(\frac{1}{x_1 x_2}\right).$$

Hence,

$$\sum_{x_2 \in T} X_{m,n}(H_2, \{u, x_2\}) = O\left(\sum_{x_2 \in T} \frac{1}{ux_2}\right)$$
$$= O(\log n)$$

which imply equation (34).

Proof of Lemma 9.3: First of all note that $\lambda_i(B_n) = \sqrt{\Delta_i(G_4)}$ for $i \leq k$ whp. More-over, we also have:

$$\left(1 - O\left(n^{13/50}\log n\right)\right)\Delta_i(G) \leq \Delta_i(G_4) \leq \Delta_i(G)$$

Thus,

$$\frac{\Delta_{i+1}(G_4)}{\Delta_i(G_4)} \leq \frac{\Delta_{i+1}(G)}{\Delta_i(G) - n^{13/50}\log n}$$
$$= \frac{1}{1 - \frac{n^{13/50}\log n}{\Delta_i(G)}} \frac{\Delta_{i+1}(G_4)}{\Delta_i(G_4)}$$
$$= \left(1 + O\left(\frac{1}{n^{12/50}\log n}\right)\right)\left(1 - O\left(\frac{1}{\log^2 n}\right)\right)$$
$$= \left(1 - O\left(\frac{1}{\log^2 n}\right)\right).$$

Hence,

$$\lambda_i(B_n) - \lambda_{i+1}(B_n) = \sqrt{\Delta_i(G_4)} - \sqrt{\Delta_{i+1}(G_4)}$$
$$= \sqrt{\Delta_i(G_4)}\left(1 - \sqrt{\frac{\Delta_{i+1}(G_4)}{\Delta_i(G_4)}}\right)$$
$$= \sqrt{\Delta_i(G_4)}\left(1 - \sqrt{1 - O\left(\frac{1}{\log^2 n}\right)}\right)$$
$$\geq \frac{n^{1/4}}{\log^{1/2} n} \cdot O\left(\frac{1}{\log^2 n}\right)$$
$$\geq \frac{n^{1/4}}{\log^3 n}.$$

Proof of equation (26)**:** Let T be a tree. We wish to count the number of closed walks in T which use each edge exactly twice (back and forth in the walk). Note that we can rewrite equation (26) as:

$$\sum_{v \in T} d(v, T)! \prod_{w \neq v} (d(w, T) - 1)!$$

We claim that the number of closed walks in T which use each edge exactly twice, starting and ending at v is given by:

$$d(v, T)! \prod_{w \neq v} (d(w, T) - 1)!$$

which follows easily from induction.

Reduction of $\mathcal{T}_k$ on the proof of Lemma 7.4: We want to show that it is enough to consider the case when $\mathcal{T}_k$ is the set of trees with exactly k edges. If T is a tree on $\tilde{k}$ many edges, where $\tilde{k} < k$, then:

$$\mathcal{M}_{2\tilde{k}+2}(T) \leq (\Delta(T)(2\tilde{k})) \mathcal{M}_{2\tilde{k}}(T)$$

to see this, note that given a closed walk of length $2\tilde{k}$, then we can choose a point during the walk to add an extra step. The number of places where the extra step goes is $2\tilde{k}$, and at each vertex we have at most $\Delta(T)$ many choices on which vertex to visit from there. Hence,

$$\mathcal{M}_{2k}(T) \leq (2k\Delta(T))^{k-\tilde{k}} \mathcal{M}_{2\tilde{k}}(T).$$

Note that calculation (27) can be carried the same with k replaced by $\tilde{k}$. Hence, when we sum over the trees with exactly $\tilde{k}$ many edges we get an upper bound of:

$$(2k\Delta(T))^{k-\tilde{k}} 2\tilde{k}(2\tilde{k}(m+1))^{2\tilde{k}} \leq (2k^2)^{k-\tilde{k}} 2k(2k(m+1))^{2\tilde{k}}$$
$$\leq C^k k^{2k}$$

Hence, by summing over all $\tilde{k} < k$ we get an extra factor of k to obtain:

$$C^k k^{2k+1}$$

which satisfies the required limit needed in equation (21).

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY
E-mail address: `daniel.montealegre@yale.edu`

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY
E-mail address: `van.vu@yale.edu`