UNIVERSITY OF CALIFORNIA, SAN DIEGO

Random Graphs with Attribute Affinity

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Mary L Radcliffe

Committee in charge:

Professor Fan Chung Graham, Chair
Professor Sanjoy Dasgupta
Professor Ron Graham
Professor Todd Kemp
Professor Jacques Verstraete

2012

The dissertation of Mary L Radcliffe is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____

____

____

____

____

Chair

University of California, San Diego

2012

DEDICATION

To John, whose love and support keeps me grounded,
and to Calvin and Beatrice, whose daily joyfulness
reminds me I can fly.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

There are many people whose support and guidance have contributed to my development as a mathematician, and whose influence I have carried with me. Most notably, I would like to thank Fan Chung for her patience, dedication, and many words of wisdom throughout these last few years. Without such a supportive and understanding advisor, I would most certainly not have been able to achieve all that I have. I would also like to thank my other mathematical collaborators and coauthors, especially Stephen Young and Paul Horn, for their roles not only in working with me, but educating me in ways of which they may not even be aware.

Thanks goes to Ping Zhang, who started my interest in graph theory and helped me to understand and begin to develop a research identity. I would also like to thank Jacques Verstraete, Allen Schwenk, John Martino, Jeff Strom, Jeff Remmel, and Todd Kemp for their fantastic education and support. Very heartfelt thanks go to the late Tom Cotner, who gave me confidence in my abilities at a young age and will be deeply missed.

Special thanks also go to my friends and officemates over the past years. Franklin, Jake, Andy, Adrian, Alex, Bob: where would we be without one another to bounce ideas off and check each other's algebra? To my friends, Joel, Patrick, Chris, Karen, John, Rebecca, Jenn, Val, Kelly, and the entire choir at UCC: thank you for being there for a laugh, a game, a meal, a song, a pie, or even just a coffee. To the UCSD AWM Chapter: thanks for lending an ear and providing advice when it was needed, and for working tirelessly to improve the community of mathematicians at our university.

My family has been a mainstay of strength since my first days. My parents and siblings have never doubted me and their faith and love has kept me going when even I thought I could not. For tolerating all my oddities and caring for me despite them, a thousand thanks to David, Linda, Annie, and Doug.

Through thick and thin, good times and bad, I cannot even begin to express my gratitude to my wonderful loving husband John. For watching the kids while I was working day and night, for making so many sacrifices for me, and for just being there, thank you, thank you, thank you. And to Calvin and Beatrice, who have no idea what's going on, your simple pleasures make my life possible. Every day that I have struggled, your laughter reminds me I can go on. Thank you for just being you.

Chapter 2 is based on the paper "On the spectra of general random graphs", Electronic Journal of Combinatorics, volume 18(1), P215, 2011; joint with Fan Chung. The dissertation author was the primary author of this work.

Chapter 3 is based in part on the paper "The spectra of multiplicative attribute graphs," which has been submitted; joint with Stephen Young. The dissertation author was the primary author of this work.

Section 3.4 is based on unpublished joint work with Stephen Young. The dissertation author was the primary author of this work.

Section 4.2 is based on the paper "Giant components in Kronecker graphs," Random Structures and Algorithms, volume 40(3), pp. 385-397, 2012; joint with Paul Horn, and appears with kind permission of Wiley-Blackwell. The dissertation author was the primary author of this work.

Section 4.3 is based on unpublished joint work with Stephen Young. The dissertation author was the primary author of this work.

VITA

2007	Bachelor of Science in Mathematics *summa cum laude*, Western Michigan University
2007	Bachelor of Arts in Spanish *summa cum laude*, Western Michigan University
2010	Master of Arts in Mathematics, University of California, San Diego
2012	Doctor of Philosophy in Mathematics, University of California, San Diego

PUBLICATIONS

"On the irregular chromatic number of a graph" (with F. Okamoto and P. Zhang), *Congressus Numerantium*, 314, pp 129-150, 2006.

"On irregular colorings of graphs" (with P. Zhang), *AKCE International Journal of Graphs and Combinatorics*, 3, pp. 175-191, 2006.

"Irregular colorings of graphs" (with P. Zhang), *Bulletin of the Institute of Combinatorics and its Applications*, 49, pp. 41-59, 2007.

"Giant components in Kronecker graphs" (with P. Horn), *Random Structures and Algorithms* 40(3), pp. 385-397, 2012.

"On the spectra of general random graphs" (with F. Chung), *Electronic Journal of Combinatorics*, 18(1), P215, 2011.

"The spectra of multiplicative attribute graphs" (with S. Young), submitted.

ABSTRACT OF THE DISSERTATION

Random Graphs with Attribute Affinity

by

Mary L Radcliffe

Doctor of Philosophy in Mathematics

University of California, San Diego, 2012

Professor Fan Chung Graham, Chair

In this thesis, we study problems related to random graphs generated via attribute affinity. A random graph with attribute affinity is a graph for which we associate to each vertex an attribute vector from an alphabet Γ, and generate edges randomly, where the probability of an edge is determined by comparing the attributes of the associated vectors. In particular, we shall do the following:

- For general random graphs in which the probability that $v_i \sim v_j$ is p_{ij}, we develop a technique for obtaining concentration of both the adjacency and normalized Laplacian eigenvalues. This technique can be used to asymptotically establish the spectra of a stochastic Kronecker graph, an affinity graph with vertex set fixed at Γ^t.

- We then generalize these results to a multiplicative attribute graph, an attribute affinity model in which the vertex set is chosen randomly from Γ^t. Specifically, we can determine asymptotically the normalized Laplacian eigenvalues in this regime. This allows us to determine asymptotic bounds on the diameter of the graph, which was shown to be asymptotically constant by Kim and Leskovec [43].

- We establish a necessary and sufficient condition for the emergence of a giant component of a stochastic Kronecker graph with $\Gamma = \{0, 1\}$. Moreover, the proof is adapted to establish the uniqueness and asymptotic size of such a component. This extends previous work of Mahdian and Xu [55], in which necessary and sufficient conditions were established in certain cases.

- Using techniques similar to those for the spectrum, we can determine the uniqueness and asymptotic size of a giant component in a multiplicative attribute graph, when it exists. Conditions on the emergence of the component were established by Kim and Leskovec [43].

Chapter 1

Mathematical Preliminaries and Overview

1.1 Notation

Throughout this paper, we use standard graph theory notation and terminology. A graph $G = (V, E)$ is a set V of vertices and a set $E \subset \binom{V}{2}$ of edges. The number of vertices in G is denoted by n. Given a vertex $v \in V$, we will write $\deg(v)$ to denote the degree of the vertex. Here we will consider graphs that may have edges from a vertex v to itself; such an edge will contribute 1 to the degree of the vertex and will be referred to as a self-loop. We will write $d_{\max}$ and $d_{\min}$ to denote the largest and smallest degrees, respectively. If every vertex of G has degree d, then G is called d-regular.

If two vertices are adjacent in G, we will often write $u \sim v$ or $uv \in E(G)$. If H is a subgraph of G, we will write $u \sim_H v$ to denote that $uv \in E(H)$.

The distance between two vertices u and v, denoted $\mathrm{dist}(u, v)$, is the number of edges in the shortest path between u and v. The diameter of the graph is defined to be $\mathrm{diam}(G) = \max_{u,v \in V} \mathrm{dist}(u, v)$, so that between any two vertices $u, v \in V$ there is a path of length at most $\mathrm{diam}(G)$ between u and v.

Given a set $S \subset V$, we define $\mathrm{vol}\,(S) = \sum_{v \in S} \deg(v)$. We will write vol $(V) = \text{vol}(G)$. For two subsets $S, T \subset V$, define

$$e(S,T) = |\{\{u,v\} \in E(G) : u \in S, v \in T\}|.$$

If $S = \{v\}$ is a single vertex, we will write $e(S,T) = e(v,T)$.

A weighted graph is a graph G together with a function $w : E \to \mathbb{R}^+$ that assigns a weight $w(e)$ to each edge $e \in E$. In this context, the degree of a vertex v is the sum of the weights of the edges incident to v.

We shall also use standard matrix notation throughout. The $n \times 1$ vector with every entry equal to 1 will be denoted by $\mathbf{1_n}$, or simply $\mathbf{1}$ if the dimension is understood. The $n \times n$ matrix with every entry equal to 1 will be denoted by J_n, or J if the dimension is understood.

A matrix M is Hermitian if for all i, j, we have $M_{ij} = \bar{M}_{ji}$. For a Hermitian matrix M, all eigenvalues are real [41]; we write the eigenvalues of M to be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, unless otherwise specified. We write $\|M\| = \max_{i=1,2,\ldots,n} |\lambda_i|$. Note that although in the case of general matrices this is a semi-norm, by restricting to Hermition matrices, $\| \cdot \|$ is in fact a norm, called the spectral norm. We define a matrix to be positive semidefinite if all of its eigenvalues are nonnegative. We can therefore define a partial ordering on the set of Hermitian matrices, called the semidefinite ordering, such that

$$A \preceq B \text{ if and only if } B - A \text{ is positive semidefinite.}$$

That is to say, $A \preceq B$ if and only if $\|B - A\| \geq 0$.

We will make frequent use of the following well-known theorem, known as Weyl's Theorem (see, for example, [41]).

Theorem 1.1 (Weyl's Theorem)**.** *If M and N are Hermitian matrices, then*

$$\max_{1 \leq i \leq n} |\lambda_i(M) - \lambda_i(N)| \leq \|M - N\|.$$

We shall use several applications of functions to matrices. In general, if f is a function with Taylor expansion $f(x) = \sum_{n=0}^{\infty} a_n x^n$, we take $f(A) = \sum_{n=0}^{\infty} a_n A^n$. We note that notions of convergence are as in [41]. In particular, we will often use the matrix exponential, $\exp(A) = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$. We note that $\exp(A)$ is always positive definite when A is Hermitian, and that $\exp(A)$ converges for all choices of A. Moreover, we shall require brief use of the matrix logarithm. In general, if $B = \exp(A)$, we say that A is a logarithm of B. As our matrices will be Hermitian, it is sufficient for uniqueness of this function to require that the logarithm also be Hermitian (see, for example, [64]).

We shall also make use of the Kronecker product of matrices. Given two matrices A and B, we define the Kronecker product $A \otimes B$ as

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{pmatrix}.$$

The Kronecker product satisfies the following properties (see, for example, [53], [49]).

Theorem 1.2. *Let A, B, C and D be $n \times n$ matrices. Then*

- $(A \otimes B)(C \otimes D) = AC \otimes BD$
- $A \otimes (B \otimes C) = (A \otimes B) \otimes C$
- *if $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A and $\mu_1, \mu_2, \dots, \mu_n$ are the eigenvalues of B, then the eigenvalues of $A \otimes B$ are $\{\lambda_i \mu_j : 1 \leq i \leq n, 1 \leq j \leq n\}$.*

Throughout the remainder of this paper, log will refer to the natural logarithm.

1.2 Spectral Graph Theory

Spectral graph theory is the study of eigenvalues of matrices associated to graphs, and the connections between these eigenvalues and properties of the associated graphs. Here we will focus on two such matrices.

The adjacency matrix of a graph is an $n \times n$ matrix, denoted by A_G, or simply A if the graph is understood, indexed by V, in which

$$A_{ij} = \begin{cases} 1 & v_i \sim v_j \\ 0 & \text{otherwise} \end{cases} . \tag{1.1}$$

Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ denote the eigenvalues of A. The largest eigenvalue μ_1 has the property that $d_{\max} \geq \mu_1 \geq \sqrt{d_{\max}}$. Moreover, it is easy to see that if A is a d-regular graph, then $\mu_1 = d$, with corresponding eigenvector $\mathbf{1}$.

The second largest eigenvalue of the adjacency matrix is also of great importance. For regular graphs, we have the following bound.

Theorem 1.3 (Alon-Boppana Bound [3]). *Let G be a d-regular graph with $d \geq 3$. Then*

$$\mu_2 \geq 2\sqrt{d-1}\left(1 - O\left(\frac{\log(d-1)}{\log n}\right)\right).$$

Graphs for which $\mu_2 \leq 2\sqrt{d-1}$ are called Ramanujan graphs. For regular graphs, the second largest eigenvalue of the adjacency matrix is related to expansion in the graph. The following result, known as the Expander Mixing Lemma (see, for example, [47]) illustrates that relationship. In particular, we see that the difference between the number of edges between any two sets and the average number of edges is controlled by μ_2.

Theorem 1.4 (Expander Mixing Lemma). *Suppose G is a d-regular graph. Let $S, T \subseteq V(G)$. Then*

$$\left| e(S,T) - \frac{d|S||T|}{n} \right| \leq \mu_2 \sqrt{|S||T|}$$

Moreover, the second largest eigenvalue is related to connectivity and diameter properties. In particular, we have the following (see, for example, [17]).

Theorem 1.5. *Let G be a d-regular graph. Then the diameter of G satisfies*

$$\operatorname{diam}(G) \leq \frac{\log(n-1)}{\log\left(\frac{2d-\mu_2-\mu_n}{\mu_2-\mu_n}\right)} + 1$$

Notice that in the preceding three theorems, we require G to be a regular graph. In fact, the spectrum of the adjacency matrix can be overwhelmed by degree information, as shown by Chung, Lu, and Vu in [22]. These authors show that the eigenvalues of the adjacency matrix of a random power law graph are themselves power law. Thus, any structural information about the graph is subsumed by the degree sequence. Therefore, in order to obtain analogous structural results for irregular graphs, we turn to the normalized Laplacian matrix $\mathcal{L}$.

For a graph G, define D_G (or simply D if the graph is understood) to be the $n \times n$ diagonal matrix with $D_{ii} = \deg(v_i)$, called the degree matrix. The normalized Laplacian is defined to be $\mathcal{L}_G = I - D^{-1/2}AD^{-1/2}$, with the convention that $D_{ii}^{-1/2} = 0$ if $\deg(v_i) = 0$, and written as simply $\mathcal{L}$ if the graph is understood. It is easy to check that the eigenvalues of $\mathcal{L}$ are between 0 and 2, and 0 is an eigenvalue with eigenvector $D^{1/2}\mathbf{1}$. Let $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$ be the eigenvalues of $\mathcal{L}$.

As with the adjacency matrix, the eigenvalues of $\mathcal{L}$ provide significant structural information about the graph. The following theorem gives some basic facts about the spectrum of $\mathcal{L}$. We will repeatedly use item 4 regarding connectivity of a graph based on λ_1, and thus we here provide a short proof, adapted from [17].

Theorem 1.6. *Let G be a graph with $\mathcal{L}$ and λ_i as above. Then*

1. $\lambda_{n-1} = 2$ *if and only if G is bipartite.*

2. *if G is bipartite and λ_i is an eigenvalue of $\mathcal{L}$, then $2 - \lambda_i$ is an eigenvalue of $\mathcal{L}$ with the same multiplicity.*

3. *the spectrum of G is the union of the spectra of its connected components.*

4. *G is connected if and only if $\lambda_1 > 0$.*

Proof of 4. By the Rayleigh-Ritz Theorem (see, for example, [41]), we have that

$$\begin{aligned}
\lambda_1 &= \inf_{f \perp D^{1/2}\mathbf{1}} \frac{\langle f, \mathcal{L} f\rangle}{\langle f, f\rangle} \\
&= \inf_{f \perp D^{1/2}\mathbf{1}} \frac{\langle f, D^{-1/2}(D-A)D^{-1/2}\rangle}{\langle f, f\rangle} \\
&= \inf_{\substack{f \perp D^{1/2}\mathbf{1} \\ g = D^{-1/2} f}} \frac{\langle g, (D-A)g\rangle}{\langle Dg, g\rangle} \\
&= \inf_{g \perp D\mathbf{1}} \frac{\sum_{u\sim v}(g(u)-g(v))^2}{\sum_{v\in V} g(v)^2 d_v}
\end{aligned}$$

Thus, $\lambda_1 = 0$ if and only if there exists a $g \perp D\mathbf{1}$ such that $\sum_{u\sim v}(g(u) - g(v))^2 = 0$. Suppose G is connected. Then in order to achieve the sum above, we must have that $g(u) = g(v)$ for any two vertices that are adjacent, and thus as there is a path between any two vertices, $g(u) = g(v)$ for any $u, v \in G$ and g is a constant vector. However, no nonzero constant vector is perpendicular to $D\mathbf{1}$, and thus $\lambda_1 > 0$.

If G is disconnected, then the spectrum of $\mathcal{L}$ is the union of the spectra of the connected components. As each component has an eigenvalue 0 and there are at least two components, $\lambda_1 = 0$. □

Let $\bar{\lambda} = \max\{1-\lambda_1, \lambda_{n-1}-1\}$. This is often referred to as the spectral gap of G. As with μ_2 for the adjacency matrix, $\bar{\lambda}$ is related to expansion and connectivity properties of G. In particular, we obtain the following analog of Theorem 1.4 for irregular graphs. As with the Expander Mixing Lemma for regular graphs, this is essentially a bound on the difference between the number of edges between two sets and the average number of edges between sets with the same volume.

Theorem 1.7 (Expander Mixing Lemma for general graphs)**.** *Let G be a graph and $S, T \subset V$. Then*

$$\left| e(S,T) - \frac{\mathrm{vol}\,(S)\,\mathrm{vol}\,(T)}{\mathrm{vol}\,(G)} \right| \le \bar{\lambda}\sqrt{\mathrm{vol}\,(S)\,\mathrm{vol}\,(T)}.$$

Moreover, λ_1 is also related to the diameter of the graph, as illustrated in the following theorem (see, for example, [17]).

Theorem 1.8. *The diameter of G satisfies*

$$\frac{1}{\operatorname{diam}(G)\operatorname{vol}(G)} \leq \lambda_1 \leq 1 - 2\frac{\sqrt{d_{\max}-1}}{d_{\max}}\left(1 - \frac{2}{\operatorname{diam}(G)}\right) + \frac{2}{\operatorname{diam}(G)}. \tag{1.2}$$

Another way in which λ_1 is related to expansion properties is by way of the Cheeger constant. We define the Cheeger constant h_G to be

$$h_G = \min_{S \subset V} \frac{e(S, \bar{S})}{\min\{\operatorname{vol}(S), \operatorname{vol}(\bar{S})\}}. \tag{1.3}$$

In essence, the Cheeger constant is measuring the smallest cut-size in a graph. A graph with good expansion properties should have a relatively large Cheeger constant, as any set that can be disconnected from the graph with relatively few edges does not expand. In fact, this intuition is confirmed by a relationship between λ_1 and h_G, known as the Cheeger inequality. Various proofs of the Cheeger Inequality can be found in [18] and [17].

Theorem 1.9 (Cheeger Inequality)**.** *For a graph G, we have*

$$\frac{h_G^2}{2} \leq \lambda_1 \leq 2h_G. \tag{1.4}$$

An isomorphism of a graph G is a bijective function $\phi : V(G) \to V(G)$ such that $u \sim v$ if and only if $\phi(u) \sim \phi(v)$. In this case, we can induce a bijection from ϕ on the edges of G, also written as ϕ by abuse of notation, such that $\phi(uv) = \phi(u)\phi(v)$. A graph is called *edge transitive* if for any two edges $uv, wx \in E$, there exists an isomorphism of G such that $\phi(uv) = wx$. We shall use the following theorem regarding h_G for edge transitive graphs, found in [17].

Theorem 1.10. *If G is an edge-transitive graph with diameter* $\operatorname{diam}(G)$*, then*

$$h_G \geq \frac{1}{2\operatorname{diam}(G)}. \tag{1.5}$$

1.3 Probability

The main focus of this thesis will be to explore spectral and connectivity properties of several random graph families. We will thus require some probabilistic techniques. In particular, we will often use results related to how far a random variable differs from its mean. Such theorems are referred to as concentration inequalities. The first and perhaps simplest such inequality is Markov's Inequality.

Theorem 1.11 (Markov's Inequality)**.** *Let X be a nonnegative, real-valued random variable. Then*

$$\mathbb{P}(X > a) \leq \frac{\mathbb{E}[X]}{a}.$$

Markov's Inequality can be viewed as a concentration inequality by taking $a = c\mathbb{E}[X]$, so that we obtain $\mathbb{P}(X > c\mathbb{E}[X]) \leq \frac{1}{c}$. As a corollary to Markov's Inequality, one can obtain Chebyshev's Inequality:

Theorem 1.12 (Chebyshev's Inequality)**.** *Let X be a random variable with $|\mathbb{E}[X]| < \infty$. Then*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

In particular, we shall often be concerned with random variables that are sums of independent random indicator functions. A random indicator function that takes the value 1 with probability p and 0 otherwise is referred to as a Bernoulli random variable with parameter p. In this instance, we can obtain further concentration by way of Chernoff bounds. The first and simplest such bound was proven by Chernoff in 1952 [16].

Theorem 1.13 (Chernoff bounds, version 1)**.** *Let $X_1, X_2, \ldots, X_m$ be i.i.d. random variables with $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$. Let $X = X_1 + X_2 + \cdots + X_m$. Then for any $a > 0$,*

$$\mathbb{P}(X > a) < e^{-\frac{a^2}{2n}}.$$

There are many different versions of Chernoff bounds, several of which can be found in the Appendix of [5]. We provide here those versions that are useful for the present purposes.

Theorem 1.14 (Chernoff bounds, version 2)**.** *Let $X = X_1 + X_2 + \cdots + X_m$, where the X_i are i.i.d. random variables that take the value 1 with probability p and 0 otherwise. Then*

$$\mathbb{P}(|X - pm| > \epsilon pm) < 2e^{-c_\epsilon pm}, \tag{1.6}$$

where $c_\epsilon = \min\{-\log(\epsilon^\epsilon(1+\epsilon)^{-(1+\epsilon)}), \epsilon^2/2\}$.

Theorem 1.15 (Chernoff bounds, version 3)**.** *Let* $X = X_1 + X_2 + \cdots + X_m$*, where each* X_i *takes the value* 1 *with probability* p_i *and* 0 *otherwise. Let* $p = \sum_{i=1}^m p_i$*, so that* $\mathbb{E}[X] = pm$*. Then*

$$\mathbb{P}(X - \mathbb{E}[X] < -a) < e^{-\frac{a^2}{2pm}} \tag{1.7}$$

Another useful concentration inequality is Bernstein's inequality, originally proven by Bernstein in 1924 [10]. We state here a version somewhat simpler than Bernstein's original result (see, for example, [21]). In Chapter 2, we will generalize this result to random matrices.

Theorem 1.16 (Bernstein's Inequality)**.** *Let* $X = X_1 + X_2 + \cdots + X_m$*, where the* X_i *are independent random variables with* $|X_i| \le M$ *for each* i*. Then*

$$\mathbb{P}(|X - \mathbb{E}[X]| > a) \le \exp\left(\frac{-a^2}{2\operatorname{Var}(X) + 2Ma/3}\right). \tag{1.8}$$

In addition to concentration inequalities, we shall take advantage of other results in probability theory. One such useful result is Jensen's Inequality (see, for example, [28]).

Theorem 1.17 (Jensen's Inequality)**.** *If* f *is a convex function and* X *is any random variable, then* $f(\mathbb{E}[X]) \le \mathbb{E}[f(X)]$.

If $\{X_n\}_{n\in\mathbb{N}}$ are a family of random variables, we say that X_n has property $\mathcal{A}$ asymptotically almost surely if

$$\lim_{n\to\infty} \mathbb{P}(X_n \text{ has property } \mathcal{A}) = 1.$$

As we shall often concern ourselves with asymptotic results of this nature, we will abbreviate "asymptotically almost surely" to "a.a.s.".

As we will examine the spectra of random graphs, we will require an understanding of random matrices for several of our main results. A random matrix M is a matrix in which each entry is a random variable. We will consider expectation to be taken coordinatewise, so $\mathbb{E}[M]_{ij} = \mathbb{E}[M_{ij}]$. For square matrices, we will take variance in an analogous way to one-dimensional random variables, so $\mathrm{Var}(M) = \mathbb{E}\left[(M - \mathbb{E}[M])^2\right]$.

We note that Jensen's Inequality can be generalized to random matrices (see, for example, [62]).

Theorem 1.18 (Operator Jensen's Inequality). *If f is convex with respect to the semidefinite order and X is a random Hermitian matrix, then $f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$.*

1.4 Random Graphs

A random graph is a graph in which each edge occurs randomly, often independent from the other edges. One of the first such graphs studied is $G_{n,p}$, known as the Erdős-Rényi graph [30], although in fact the model appears in an earlier paper of Gilbert's [36]. In $G_{n,p}$, every edge occurs independently and at random with probability p. Thus, the adjacency matrix of $G_{n,p}$ is a random Hermitian matrix in which every entry above the diagonal is a Bernoulli random variable with probability parameter p, and every diagonal entry is 0.

In general, a random graph has a random Hermitian adjacency matrix, where each entry above the diagonal is a Bernoulli random variable. In the case that we allow self-loops, every diagonal entry is also a Bernoulli random variable. We will write the expectation of A as $\bar{A}$. Similarly, D will be a diagonal matrix whose diagonal entries are sums of Bernoulli random variables, and we write the expectation of D as $\bar{D}$.

We can view $\bar{A}$ and $\bar{D}$ as the adjacency matrix and degree matrix for the expectation of G, that is, the weighted complete graph $\bar{G}$ in which every edge is weighted by its expectation in G. We thus will make use of $\bar{\mathcal{L}} = I - \bar{D}^{-1/2}\bar{A}\bar{D}^{-1/2}$, the Laplacian matrix for $\bar{G}$, instead of the expectation of $\mathcal{L}$. We note that $\bar{\mathcal{L}}$ has many computational advantages over $\mathbb{E}[\mathcal{L}]$, as D and A are not independent and thus $\mathbb{E}[\mathcal{L}] = \mathbb{E}\left[I - D^{-1/2}AD^{-1/2}\right]$ cannot be factored into products of $\bar{D}$ and $\bar{A}$.

Returning to $G_{n,p}$, we see that $\bar{A} = p(J - I)$, $\bar{D} = (n-1)pI$, and $\bar{\mathcal{L}} = I - \frac{1}{n-1}(J - I)$.

In general, if G is a random graph, we will use $\bar{\mu}_1 \geq \bar{\mu}_2 \geq \cdots \geq \bar{\mu}_n$ to refer to the spectrum of $\bar{A}$ and $0 = \bar{\lambda}_0 \leq \bar{\lambda}_1 \leq \ldots \bar{\lambda}_{n-1}$ to refer to the spectrum of $\bar{\mathcal{L}}$. We will write $\bar{d}_{\max}$ and $\bar{d}_{\min}$ to refer the the maximum and minimum expected degrees of G, respectively. As with A, D, and $\mathcal{L}$, we will subscript the expected matrices as $\bar{A}_G$, $\bar{D}_G$, and $\bar{\mathcal{L}}_G$ to clarify as needed.

One common class of random graphs that we will refer to in Chapter 4 are percolated random graphs. For any deterministic graph G, we can define G_p to be the random graph where two vertices are adjacent with probability p if they are adjacent in G, and otherwise are not adjacent. That is to say, we choose each edge from G independently and with probability p. This type of percolation is often referred to as bond percolation. Although we will use this notion only in passing, these types of graphs have been widely studied [39], [13], [34], [4], [45].

We will focus here on random graphs defined via attribute affinity. Let $\Gamma_1, \Gamma_2, \ldots, \Gamma_t$ be finite alphabets, and for each $i \in [t]$, let $\Theta^{(i)}$ be a symmetric matrix indexed by Γ_i. These matrices are referred to as *affinity matrices*, or sometimes *generating matrices* for the graph. A random graph G is defined via attribute affinity on the set $\{\Theta^{(i)}\}$ if there exists a function $a : V(G) \to \Gamma_1 \times \Gamma_2 \times \ldots \Gamma_t$, called the attribute function, and a function f, called the probability function, taking values in $[0, 1]$, such that for all $u, v \in V(G)$ with $a(u) = (\gamma_1, \gamma_2, \ldots, \gamma_t)$ and $a(v) = (\delta_1, \delta_2, \ldots, \delta_t)$, we have

$$\mathbb{P}(u \sim v) = f(\Theta^{(1)}_{\gamma_1,\delta_1}, \Theta^{(2)}_{\gamma_2,\delta_2}, \ldots, \Theta^{(t)}_{\gamma_t,\delta_t}).$$

We call $a(v)$ the attribute vector for v.

We can think of the function a as assigning a list of attributes to the vertices, and each pair of attributes has some affinity to one another. Thus, the probability that two vertices are adjacent is a function of their list of affinities. We will typically restrict to the case that f is t-fold multiplication, and the $\Theta^{(i)}$ are $[0, 1]$-valued matrices.

1.4.1 Random graph models

Often, one produces and studies random graph models as a means of understanding the structure of large, realistic graphs. Large graphs often display predictable properties, and the goal of modeling is to generate random graphs that also obey these properties. Specifically, we shall focus on three of the most well-studied properties in large graphs.

First, many large graphs follow a power law degree distribution. Let N_k be the proportion of vertices in a graph that are degree k. The power law property states that there is a constant β such that $N_k \propto k^{-\beta}$ for all k a.a.s.. In other words, a graph is power law if $\log N_k$ is linear in $\log k$, with slope $-\beta$. Such graphs are often called scale-free. Citation graphs have been observed to be scale-free [26], as has the Internet and web-graph [9], [6], [32], [46], [14], collaboration graphs [59], and large social networks [15]. We note that this is a significant deviation from the structure seen in $G_{n,p}$, in which the expected degree of every vertex is uniform at $(n-1)p$.

The second property we shall discuss is the small-world phenomenon. A graph has the small-world property if the diameter of the graph is asymptotically constant. That is to say, even as the number of nodes increases, the diameter of the graph does not. The small-world phenomenon has been observed in the Internet and web-graph, as well as large social networks [2], [56].

Finally, we expect realistic graphs to be relatively sparse. In particular, $|E| \ll \binom{n}{2}$. This has been observed, for example, in the Internet graph [32], the web graph [48], and large social networks [27].

The main focus of this dissertation is to determine the spectra and find conditions for the emergence of the giant component in random graphs defined via attribute affinity. In particular, we focus on two graph models, the stochastic Kronecker graph model (SKG) and the multiplicative attribute graph model (MAG). As seen in Section 1.2, the spectrum of a graph has close ties to its diameter and related connectivity properties. In this way, we seek to study these graphs as large network models via their spectral properties. We will discuss the giant component and its applications more fully in Section 4.1. We begin by defining these two models, as well as some of their basic properties.

1.4.2 Stochastic Kronecker Graphs

We first define a stochastic Kronecker graph. This model was proposed by Leskovec et al. [49] as a means of generating graphs with self-similar structure.

Let Γ be a finite set with $|\Gamma| = k$. Let Θ be a $k \times k$ symmetric matrix indexed by Γ whose entries are in $[0, 1]$. The stochastic Kronecker graph $G = \mathrm{SKG}(\Theta, t)$ is a random graph defined via attribute affinity, such that $V(G) = \Gamma^t$, and the probability function is t-fold multiplication. That is, given $a = (a_1, a_2, \ldots, a_t), b = (b_1, b_2, \ldots, b_t) \in V(G)$, we have

$$\mathbb{P}(a \sim b) = \Theta_{a_1,b_1}\Theta_{a_2,b_2} \ldots \Theta_{a_t,b_t}.$$

Clearly the only necessary parameters to define G are Θ and t, as Θ will imply k, and the underlying set Γ is irrelevant to the resulting random graph. The graph G will have $n = k^t$ vertices. Notice that the SKG is defined precisely so that $\bar{A} = \Theta^{\otimes t}$. We will define $\tilde{D}$ to be the diagonal matrix of column sums of Θ.

Proposition 1.19 (see, for example, [55], [49])**.** *Let $G = \mathrm{SKG}(\Theta, t)$ over the alphabet Γ. Then $\bar{D} = \tilde{D}^{\otimes t}$.*

Proof. Let $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_t) \in V(G)$. It suffices to prove that $\bar{d_\sigma} = \prod_{i=1}^{t} \tilde{D}_{\sigma_i}$. Now,

$$\begin{aligned}
\bar{d_\sigma} &= \sum_{\tau \in \Gamma^t} \mathbb{P}(\sigma \sim \tau) \\
&= \sum_{\tau \in \Gamma^t} \prod_{i=1}^{t} \Theta_{\sigma_i,\tau_i} \\
&= \prod_{i=1}^{t} \sum_{\gamma \in \Gamma} \Theta_{\sigma_i,\gamma} \\
&= \prod_{i=1}^{t} d_{\sigma_i}.
\end{aligned}$$

Therefore, $\bar{D} = \tilde{D}^{\otimes t}$ as desired. □

By Theorem 1.2, we thus have that for $\mathrm{SKG}(\Theta, t)$, $\bar{A} = \Theta^{\otimes t}$ and

$$\bar{\mathcal{L}} = I - (\tilde{D}^{\otimes t})^{-1/2}\Theta^{\otimes t}(\tilde{D}^{\otimes t})^{-1/2} = I - (\tilde{D}^{-1/2}\Theta\tilde{D}^{-1/2})^{\otimes t}. \tag{1.9}$$

In Chapter 4, we will restrict our attention to the case where $\Gamma = \{0, 1\}$, and $n = 2^t$, so that the SKG is generated by a 2×2 affinity matrix

$$\Theta = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}.$$

In this case, the graph is generated by the parameters α, β, γ and t. We shall denote the SKG with 2×2 affinity matrix by $K(t, \alpha, \beta, \gamma)$. Moreover, we typically assume without loss of generality that $\alpha > \gamma$. In this case, the attribute vector of each vertex is a length t binary vector. Let $\omega(v)$ denote the number of 1s in the attribute vector for v, otherwise known as the *weight* of v. In this case, Proposition 1.19 yields the following.

Corollary 1.20. *Let $G = K(t, \alpha, \beta, \gamma)$. For $v \in V(G)$, we have*

$$\mathbb{E}[\deg(v)] = (\alpha + \beta)^{\omega(v)}(\beta + \gamma)^{t-\omega(v)}.$$

This property gives the expected degree distribution for an SKG with 2×2 generating matrix. We notice here that this distribution is not power law, so the SKG is not a scale-free network. However, as the following theorem shows, these graphs do obey the small-world phenomenon whenever they are connected.

Theorem 1.21. *[55] Let $G = K(t, \alpha, \beta, \gamma)$, where wolog $\alpha > \gamma$. Then G is connected a.a.s. if and only if $\beta + \gamma > 1$. Moreover, if $\beta + \gamma > 1$, then the diameter of G is constant a.a.s..*

In fact, in Chapter 3, we are able to provide asymptotic bounds on the diameter of the graph based on the parameters α, β, and γ.

Finally, we have that the expected number of edges in $G = K(t, \alpha, \beta, \gamma)$ is

$$\sum_{v \in V} \mathbb{E}[\deg(v)] = \sum_{i=1}^{t} \binom{t}{i} (\alpha + \beta)^i (\beta + \gamma)^{t-i} = (\alpha + 2\beta + \gamma)^t.$$

As $\binom{t}{2} \sim t^2$, this graph is relatively dense, rather than sparse. Furthermore, the degree distribution is binomial, rather than power law.

Despite these drawbacks, the SKG has been studied as a model for realistic graphs [50], [51]. In [50], Leskovec et. al. show that the SKG with a 2×2

initiator matrix captures some relevant properties of realistic graphs. Specifically, the authors give an algorithm to choose the parameters α, β, γ to model a real network. Using this algorithm, SKGs are produced to model several real networks that roughly capture the eigenvalue distribution for the adjacency matrix, diameter, degree distribution, and hop-plot. The hop-plot is a plot of the number of pairs of vertices connected by paths of length at most k against k, and can be viewed as a way to measure expansion in the graph.

1.4.3 Multiplicative Attribute Graphs

Although the stochastic Kronecker graph model has been proposed as a model for a variety of different complex networks, it has the significant drawback that with a $k \times k$ affinity matrix, it is only possible to generate a network whose size is a power of k. Moreover, as we have seen, this network has a less than satisfactory degree distribution and density. There are several different generalizations of the stochastic Kronecker graph to a network of different degree distribution or arbitrary size. One such method, proposed by Kim and Leskovec, is the multiplicative attribute graph (MAG) [43].

In the most general setting, an MAG can be defined as follows. For $1 \leq i \leq t$, let Γ_i be an alphabet of size k_i. For each i, let $\Theta^{(i)}$ be a $k_i \times k_i$ affinity matrix with entries in $[0, 1]$. Let P be a probability distribution over $\Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_t$. For $n > 0$, we define $G = \mathrm{MAG}(n, \{\Gamma_i\}, \{\Theta^{(i)}\}, P)$ to be a random graph with attribute affinity, with $|V(G)| = n$, and $a(v) \in \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_t$, is chosen from the probability distribution P independently for each v. We take the probability function to be t-fold multiplication, so that for two vertices u, v with attribute vectors $a(u) = (\sigma_1, \sigma_2, \ldots, \sigma_t)$ and $a(v) = (\tau_1, \tau_2, \ldots, \tau_t)$, we have

$$\mathbb{P}(u \sim v) = \Theta^{(1)}_{\sigma_1,\tau_1} \Theta^{(2)}_{\sigma_2,\tau_2} \ldots \Theta^{(t)}_{\sigma_t,\tau_t}.$$

If each alphabet above is $\{0, 1\}$, we have $\Theta^{(i)} = \begin{pmatrix} \alpha_i & \beta_i \\ \beta_i & \gamma_i \end{pmatrix}$. We may take Q to be a product distribution, where the attribute γ_i in $a(v)$ is equal to 1 with probability μ_i and 0 with probability $1 - \mu_i$, and each attribute is independent of each other attribute. In this setting, Leskovec and Kim [43] show that under certain conditions, the MAG can be forced to take a power law degree distribution:

Theorem 1.22. *[43] If $\Theta^{(i)}$, μ_i are as above, and for each i,*

$$\frac{\mu_i}{1-\mu_i} = \left(\frac{\mu_i\alpha_i + (1-\mu_i)\beta_i}{\mu_i\beta_i + (1-\mu_i)\gamma_i}\right)^{-\delta}$$

for some $\delta > 0$, then the degree distribution of G satisfies a power law with coefficient $-\delta - \frac{1}{2}$.

We will focus on a less general version of the MAG. In particular, we will take $\Gamma_1 = \Gamma_2 = \cdots = \Gamma_t = \Gamma$, so that every attribute comes from the same alphabet, and $k = |\Gamma|$. Moreover, if Q is a probability distribution over Γ, we will take P to be the product distribution of Q over Γ^t. If we think of Q as a $k \times k$ diagonal matrix with $Q_{\sigma,\sigma} = Q(\sigma)$, we then have that a vertex v takes attribute vector $a(v) = \sigma$ with probability $Q^{\otimes t}_{\sigma,\sigma}$. Without loss of generality, we may assume that for all $\gamma \in \Gamma$, $Q_{\gamma,\gamma} > 0$ as otherwise we may consider the smaller alphabet $\Gamma' = \Gamma - \{\gamma\}$. We denote this version of the MAG by MAG(n, t, Θ, Q). For a particular word $\sigma \in \Gamma^t$, we will use n_σ to denote the number of vertices of G with attribute vector σ. We define the *signature* of the graph to be the collection $\{n_\sigma\}_{\sigma\in\Gamma^t}$.

Notice that in the multiplicative attribute graphs, edges of the graph do not appear independently. In fact, the probability that each edge will appear depends upon the attribute vector of the vertices incident to that edge. However, if we fix the signature of a MAG, we do have that edges appear independently, a fact that we will take advantage of in Chapters 3 and 4.

In Section 4.3, we will moreover focus on MAG(n, t, Θ, Q) in the case that $\Gamma = \{0, 1\}$. In this case, as with the SKG, we have that the affinity matrix

$$\Theta = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix},$$

and Q is a binary distribution with $\mathbb{P}(1) = \mu$ and $\mathbb{P}(0) = 1 - \mu$. In this setting, we will use notation similar to that in the SKG, so $\omega(v)$ will denote the weight of a vertex v; that is, $\omega(v)$ is the number of 1s in $a(v)$. In general, we will assume that $t \sim \rho \log n$. We denote this more specialized MAG by $K(n, t, \Theta, \mu)$.

In this case, we have the following expected degrees in G.

Theorem 1.23. *[43] Let $G = K(n, t, \Theta, \mu)$. Then for $v \in V$ with weight $\omega(v)$, we have*

$$\mathbb{E}\left[\deg(v)\right] = n(\mu\alpha + (1-\mu)\beta)^{\omega(v)}(\mu\beta + (1-\mu)\gamma)^{t-\omega(v)}.$$

Moreover, Kim and Leskovec [43] have shown the following.

Theorem 1.24. *Let $G = K(n, t, \Theta, \mu)$ where $t \sim \rho \log n$. Let*

$$F(G) = \begin{cases} (\mu\beta + (1-\mu)\gamma)^{\rho} & \text{when } (1-\mu)^{\rho} \geq \frac{1}{e} \\ (\mu\alpha + (1-\mu)\beta)^{\nu\rho}(\mu\beta + (1-\mu)\gamma)^{(1-\nu)\rho} & \text{otherwise} \end{cases},$$

where ν is a solution to $\left(\left(\frac{\mu}{\nu}\right)^{\nu}\left(\frac{1-\mu}{1-\nu}\right)^{1-\nu}\right)^{\rho} = \frac{1}{e}$. Then G is connected a.a.s. if $F(G) > \frac{1}{e}$ and disconnected a.a.s. if $F(G) < \frac{1}{e}$.

Moreover, G has constant diameter a.a.s..

In Chapter 3, we will also provide asymptotic bounds on the diameter of $K(n, t, \Theta, \mu)$ based on the parameters μ, α, β, and γ.

Thus we see that the multiplicative attribute graph has several benefits over the stochastic Kronecker graph. While we are still able to define adjacencies via commonalities between vertices, we can control the degree sequence of the graph and maintain constant diameter, demonstrating many useful realistic properties, such as power law and the small world phenomenon.

As with the SKG, the MAG has been studied as a model for realistic networks. In [44], Kim and Leskovec provide an algorithm for estimating parameters μ_i, $\Theta^{(i)}$ in the event that all attributes are from the alphabet $\{0, 1\}$. Using this algorithm, several real networks are fitted with multiplicative attribute graph models, which accurately approximate the spectrum of the adjacency matrix, hop-plot, degree distribution, and diameter. Moreover, the authors compare these estimates to those of the SKG using the algorithm given in [50], and find that the MAG provides a much more realistic approximation of the graph. This is unsurprising, given the degree distribution and diameter analysis of the MAG compared to the SKG.

We instead will focus on properties of the simplified graphs MAG(n, t, Θ, Q) and $K(n, t, \Theta, \mu)$. It should be mentioned that many of our results may be generalized to the MAG in full generality by taking maxima and minima of the chosen parameter settings appropriately.

1.5 Overview

The remainder of this thesis is organized as follows. In Chapter 2, we discuss some general techniques for concentrating the spectra of the adjacency matrix A and Laplacian matrix $\mathcal{L}$ of a random graph on the spectrum of $\bar{A}$ and $\bar{\mathcal{L}}$. In Chapter 3, we apply these techniques to SKG(Θ, t) and MAG(n, t, Θ, Q) as described above. This will allow us to derive some information about connectivity and diameter of these graphs. As a side discussion, we will also show how these spectral techniques can be applied to a random graph with vertices chosen randomly from a set of possible vertices. In Chapter 4, we will define the giant component in a graph, and derive necessary and sufficient conditions for the emergence of the giant component in $K(t, \alpha, \beta, \gamma)$. Necessary and sufficient conditions for the emergence of the giant component have been established for $K(n, t, \Theta, \mu)$ in [43]. We will further develop these conditions by establishing both the uniqueness and asymptotic size of the giant components in $K(t, \alpha, \beta, \gamma)$ and $K(n, t, \Theta, \mu)$.

Chapter 2

Concentration of Spectra of A and $\mathcal{L}$

2.1 Introduction

In this chapter, we provide general results on the spectra of random graphs which will become useful in determining structure of SKGs and MAGs. In particular, we will use techniques from random matrix theory to generalize Bernstein's Inequality (see Theorem 1.16) to random matrices, and then apply this concentration inequality to matrices arising from the study of random graphs. Such techniques have appeared in [60] to produce slightly weaker results. In this case, we will concentrate the spectral norm of a random matrix on that of its expectation.

To write analogues of Bernstein's Inequality, we must consider a random matrix associated to a graph as a sum of independent random matrices. Here, we will assume the edges of the graph appear independently, and then write the adjacency matrix as a sum of random matrices, each one corresponding to a single edge in the graph. For $\mathcal{L}$, we use more nuanced techniques to concentrate both the degrees of vertices as well as the norm of the graph.

2.2 Matrix Concentration Inequalities

We begin by examining the very rich field of matrix concentration inequalities. Previously, various matrix concentration inequalities have been derived by many authors including Ahlswede-Winter [1], Cristofides-Markström [25], Oliveira

[60], Gross [38], Recht [62], and Tropp [64]. Here we give a short proof for a simple version that is particularly suitable for random graphs.

Theorem 2.1. *Let $X_1, X_2, \ldots, X_m$ be independent random $n \times n$ Hermitian matrices. Moreover, assume that $\|X_i - \mathbb{E}[X_i]\| \leq M$ for all i, and put $v^2 = \|\sum \operatorname{Var}(X_i)\|$. Let $X = \sum X_i$. Then for any $a > 0$,*

$$\mathbb{P}(\|X - \mathbb{E}[X]\| > a) \leq 2n \exp\left(-\frac{a^2}{2v^2 + 2Ma/3}\right).$$

For the proof, we will rely on the following results:

Lemma 2.2 (see, for example, [64])**.** *Let $f, g : \mathbb{R} \to \mathbb{R}$, and suppose there is a subset $S \subseteq \mathbb{R}$ with $f(a) \leq g(a)$ for all $a \in S$. If A is a Hermitian matrix with all eigenvalues contained in S, then $f(A) \preceq g(A)$.*

Lemma 2.3 ([52])**.** *Given a fixed Hermitian matrix A, the function*

$$X \mapsto \operatorname{Tr}(\exp(A + \log X))$$

is concave with respect to the semidefinite order on the set of positive definite X.

By applying Theorem 1.18 together with Lemma 2.3, we obtain the following.

Lemma 2.4. *If A is a fixed matrix and X is a random Hermitian matrix, then*

$$\mathbb{E}[\operatorname{Tr}(\exp(A + X))] \leq \operatorname{Tr}(\exp[A + \log(\mathbb{E}[\exp X])]). \tag{2.1}$$

Proof. By Lemma 2.3, $f : e^X \mapsto \operatorname{Tr}\left(\exp(A + \log e^X)\right)$ is concave on the set of Hermitian matrices X, as e^X is positive definite for all Hermitian X. Thus, by Theorem 1.18, $f(\mathbb{E}[X]) \geq \mathbb{E}[f(X)]$. Therefore, we have

$$\operatorname{Tr}\left(\exp(A + \log(e^{\mathbb{E}[X]}))\right) \geq \mathbb{E}\left[\operatorname{Tr}\left(\exp(A + \log(e^X))\right)\right] = \mathbb{E}[\operatorname{Tr}(\exp(A + X))].$$

□

We shall use this result to overcome the difficulties presented by working with the semidefinite order, as opposed to real numbers. The primary problem that must be overcome is that unlike real numbers, the semidefinite order does not respect products. That is to say, if $A \preceq B$ and $C \preceq D$, it is not necessarily true that $AC \preceq BD$.

Lemma 2.5. *For independent random $n \times n$ Hermitian matrices $X_1, X_2, \ldots, X_k$ with $\mathbb{E}[X_k] = 0$ for all k and $\|X_k\| \leq M$ for all k,*

$$\mathbb{E}\left[\operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m} \theta X_k\right)\right)\right] \leq \operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m} \frac{1}{2} g(\theta M)\theta^2 \mathbb{E}\left[X_k^2\right]\right)\right), \tag{2.2}$$

where $g(x) = \frac{2}{x^2}(e^x - x - 1)$.

Proof. Let

$$g(x) = \frac{2}{x^2}(e^x - x - 1) = 2\sum_{k=2}^{\infty} \frac{x^{k-2}}{k!}.$$

Notice that g is increasing. Given $\theta > 0$, we have that $\|\theta X_k\| \leq \theta M$, and therefore $g(\theta X_k) \preceq g(\theta M)I$ by Lemma 2.2. Therefore, as $\mathbb{E}[X_k] = 0$,

$$\begin{aligned}
\mathbb{E}\left[e^{\theta X_k}\right] &= \mathbb{E}\left[I + \theta X_k + \frac{1}{2}\theta^2 X_k^2 g(\theta X_k)\right] && (2.3)\\
&= I + \frac{1}{2}\theta^2 g(\theta X_k)\mathbb{E}\left[X_k^2\right] && (2.4)\\
&\preceq I + \frac{1}{2} g(\theta M)\theta^2 \mathbb{E}\left[X_k^2\right] && (2.5)\\
&\preceq e^{\frac{1}{2} g(\theta M)\theta^2 \mathbb{E}\left[X_k^2\right]}. && (2.6)
\end{aligned}$$

For a given k, let $\mathbb{E}_k[\cdot] := \mathbb{E}[\cdot | X_1, X_2, \ldots, X_k]$. Then by the tower property of conditional expectation (see, for example, [28]), we have

$$\mathbb{E}\left[\operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m} \theta X_k\right)\right)\right] = \mathbb{E}\mathbb{E}_1\mathbb{E}_2 \ldots \mathbb{E}_{m-1}\left[\operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m-1} \theta X_k + \theta X_m\right)\right)\right]$$

As the X_i are independent, each X_k is fixed with respect to $\mathbb{E}_{m-1}$ except X_m, and $\mathbb{E}_{m-1}[\exp X_m] = \mathbb{E}[\exp(X_m)]$. Applying inequality (2.1) from Lemma 2.4, we have

$$\mathbb{E}\mathbb{E}_1\mathbb{E}_2 \ldots \mathbb{E}_{m-1}\left[\operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m-1} \theta X_k + \theta X_m\right)\right)\right] \leq$$
$$\mathbb{E}\mathbb{E}_1\mathbb{E}_2 \ldots \mathbb{E}_{m-2}\left[\operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m-1} \theta X_k + \log \mathbb{E}\left[\exp(\theta X_m)\right]\right)\right)\right].$$

Iteratively applying this process, we obtain

$$\mathbb{E}\left[\operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m} \theta X_k\right)\right)\right] \leq \operatorname{Tr}\left(\exp\left(\sum_{k=1}^{m} \log \mathbb{E}\left[\exp \theta X_m\right]\right)\right).$$

As both $\log(\cdot)$ and $\mathrm{Tr}\left(\exp(\cdot)\right)$ are monotone with respect to the semidefinite order (these facts can be proven with basic manipulations [64]), inequality (2.6) implies that

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Tr}\left(\exp\left(\sum_{k=1}^{m}\theta X_k\right)\right)\right] &\leq \mathrm{Tr}\left(\exp\left(\sum_{k=1}^{m}\log e^{\frac{1}{2}g(\theta M)\theta^2\mathbb{E}\left[X_k^2\right]}\right)\right) \\
&\leq \mathrm{Tr}\left(\exp\left(\sum_{k=1}^{m}\frac{1}{2}g(\theta M)\theta^2\mathbb{E}\left[X_k^2\right]\right)\right),
\end{aligned}$$

as desired. □

Proof of Theorem 2.1. We assume for the sake of the proof that $\mathbb{E}\left[X_k\right] = 0$ for all k. Clearly this yields the general case by simply replacing each X_k by $X_k - \mathbb{E}\left[X_k\right]$.

Given $a > 0$, for all $\theta > 0$, by Lemma 2.5 we have

$$\begin{aligned}
\mathbb{P}(\lambda_{\max}(X) \geq a) &\leq e^{-\theta a}\mathbb{E}\left[\exp(\theta\lambda_{\max}(X))\right] \\
&\leq e^{-\theta a}\mathbb{E}\left[\mathrm{Tr}\left(\exp(\theta X)\right)\right] \\
&= e^{-\theta a}\mathbb{E}\left[\mathrm{Tr}\left(\exp\left(\sum\theta X_k\right)\right)\right] \\
&\leq e^{-\theta a}\,\mathrm{Tr}\left(\exp\left(\sum\frac{1}{2}g(\theta M)\theta^2\mathbb{E}\left[X_k^2\right]\right)\right) \\
&\leq e^{-\theta a}n\lambda_{\max}(\exp\left(\frac{1}{2}g(\theta M)\theta^2\sum\mathbb{E}\left[X_k^2\right]\right)) \\
&\leq n\exp\left(-\theta a + \frac{1}{2}g(\theta M)\theta^2 v^2\right),
\end{aligned}$$

as $v^2 = \|\sum\mathbb{E}\left[X_k^2\right]\| \geq \lambda_{\max}(\sum\mathbb{E}\left[X_k^2\right])$

Notice that if $x < 3$, we have

$$g(x) = 2\sum_{k=2}^{\infty}\frac{x^{k-2}}{k!} \leq \sum_{k=2}^{\infty}\frac{x^{k-2}}{3^{k-2}} = \frac{1}{1-x/3}.$$

Take $\theta = \frac{a}{v^2+Ma/3}$. Then $\theta M = \frac{3aM}{3v^2+aM} \leq 3$, and thus we have

$$\begin{aligned}
\mathbb{P}(\lambda_{\max}(X) \geq a) &\leq n\exp\left(-\theta a + \frac{1}{2}g(\theta M)\theta^2 v^2\right) &\quad (2.7)\\
&\leq n\exp\left(-\frac{a^2}{2v^2+2Ma/3}\right) &\quad (2.8)
\end{aligned}$$

By applying inequality 2.8 to $-X$, we obtain also that

$$\mathbb{P}(\lambda_{\min}(X) \leq -a) \leq n \exp\left(-\frac{a^2}{2v^2 + 2Ma/3}\right).$$

Therefore, $\mathbb{P}(\|X\| \geq a) \leq 2n \exp\left(-\frac{a^2}{2v^2+2Ma/3}\right)$ as desired.

□

2.3 Spectrum of the Adjacency Matrix of a Random Graph

Let G be a random graph, where $\mathbb{P}(v_i \sim v_j) = p_{ij}$ and each edge appears independently of each other edge. Define A_{ij} to be the matrix with a 1 in the ij and ji position and 0 elsewhere. Let h_{ij} be a Bernoulli random variable with parameter p_{ij}. Then the adjacency matrix for G takes the form

$$A = \sum_{1 \leq i \leq j \leq n} h_{ij} A_{ij}.$$

We can thus apply Theorem 2.1 to obtain the following result.

Theorem 2.6. *Let G be a random graph, where $\mathbb{P}(v_i \sim v_j) = p_{ij}$, and each edge is independent of each other edge. Let $\epsilon > 0$, and suppose that for n sufficiently large, $\bar{d}_{\max} > \frac{4}{9}\log(2n/\epsilon)$. Then with probability at least $1 - \epsilon$, for n sufficiently large, the eigenvalues of A and $\bar{A}$ satisfy*

$$|\mu_i - \bar{\mu}_i| \leq \sqrt{4\bar{d}_{\max}\log(2n/\epsilon)}$$

for all $1 \leq i \leq n$.

We note that the bound in this theorem holds simultaneously for all eigenvalues, so that with probability $1 - \epsilon$, we will have concentration on all of the eigenvalues of A, rather than each eigenvalue independently concentrated.

Proof. Let $X_{ij} = h_{ij}A_{ij}$, so $A = \sum_{1 \leq i \leq j \leq n} X_{ij}$. Then $\|X_{ij}\| \leq 1$ for all i, j. In order to apply Theorem 2.1, we must first calculate $v^2 = \|\sum_{1 \leq i \leq j \leq n} \mathrm{Var}\,(X_{ij})\,\|$.

Notice, if $i \neq j$, then

$$\begin{aligned} \mathrm{Var}\,(X_{ij}) &= \mathbb{E}\left[(h_{ij}-p_{ij})^2(A_{ij})^2\right] \\ &= \mathrm{Var}\,(h_{ij})\,(A_{ii}+A_{jj}) \\ &= p_{ij}(1-p_{ij})(A_{ii}+A_{jj}) \end{aligned}$$

Similarly, $\mathrm{Var}\,(X_{ii}) = p_{ii}(1-p_{ii})A_{ii}$. Therefore,

$$\begin{aligned} v_2 = \left\|\sum \mathrm{Var}\,(X_{ij})\right\| &= \left\|\sum_{i=1}^{n}\left(\sum_{j=1}^{n} p_{ij}(1-p_{ij})\right) A_{ii}\right\| \\ &= \max_{i=1,\ldots,n} \sum_{j=1}^{n} p_{ij}(1-p_{ij}) \\ &\leq \max_{i=1,\ldots,n} \sum_{j=1}^{n} p_{ij} = \bar{d}_{\max}. \end{aligned}$$

Take $a = \sqrt{4\bar{d}_{\max}\log(2n/\epsilon)}$. By the assumption on $\bar{d}_{\max}$, we have $a < 3\bar{d}_{\max}$, and thus we obtain

$$\begin{aligned} \mathbb{P}(\|A-\bar{A}\| > a) &\leq 2n\exp\left(-\frac{a^2}{2v^2+2Ma/3}\right) \\ &\leq 2n\exp\left(\frac{-4\bar{d}_{\max}\log(2n/\epsilon)}{4\bar{d}_{\max}})\right) \\ &= \epsilon. \end{aligned}$$

To complete the proof, we apply Weyl's Theorem (see, Theorem 1.1) to obtain that with probability at least $1-\epsilon$, for all $1 \leq i \leq n$,

$$|\mu_i - \bar{\mu}_i| < \sqrt{4\bar{d}_{\max}\log(2n/\epsilon)}.$$

□

As an application, we consider $G_{n,p}$. As mentioned in Section 1.4, for $G_{n,p}$, we have $\bar{A} = p(J-I)$. An application of Theorem 2.6 yields

Theorem 2.7. *For $G_{n,p}$, if $p > \frac{8}{9n}\log(\sqrt{2}n)$, then with probability at least $1 - 1/n = 1-o(1)$, we have*

$$|\mu_i - \bar{\mu}_i| \leq \sqrt{8np\log(\sqrt{2}n)}.$$

As $\bar{A} = p(J-I)$, we have that $\bar{\mu}_1 = p(n-1)$ and $\bar{\mu}_i = -p$ for all $i > 1$. Thus we have proven that $\mu_1 = pn + O(pn)$ and if $i > 1$, $|\mu_i| = O(pn)$. We note that stronger results for the spectrum of the adjacency matrix of $G_{n,p}$ can be found in [35], [33]. Specifically, in [33], it is shown that for $pn \geq c \ln n$, $\mu_1 = pn + O(\sqrt{pn})$, and all other eigenvalues satisfy $|\mu_i| = O(\sqrt{pn})$. However, due to the very strong symmetries in $G_{n,p}$, it seems unlikely that the methods used to investigate this graph in detail will extend to general random graphs.

2.4 Spectrum of the Normalized Laplacian Matrix of a Random Graph

We now turn to concentration of the eigenvalues of $\mathcal{L}$ on those of $\bar{\mathcal{L}}$ for a random graph G.

Theorem 2.8. *Let G be a random graph, where $\mathbb{P}(v_i \sim v_j) = p_{ij}$, and each edge is independent of each other edge. Let $\bar{d}_{\min}$ be the minimum expected degree of G. Choose $\epsilon > 0$. Then there exists a constant $k = k(\epsilon)$ such that if $\bar{d}_{\min} > k \log(n)$, then with probability at least $1 - \epsilon$, the eigenvalues of $\mathcal{L}$ and $\bar{\mathcal{L}}$ satisfy*

$$|\lambda_j - \bar{\lambda}_j| \leq 3\sqrt{\frac{3\log(4n/\epsilon)}{\bar{d}_{\min}}}$$

for all $1 \leq j \leq n$.

As with Theorem 2.6, this bound will hold simultaneously for all eigenvalues with probability $1 - \epsilon$. The value of k above comes out of the proof, and in particular, choosing $k > 3(1 + \log(\frac{\epsilon}{4}))$ is sufficient.

Proof. We will again use Weyl's Theorem (Theorem 1.1), as in the proof of Theorem 2.6, so we need only bound $\|\mathcal{L} - \bar{\mathcal{L}}\|$. For a vertex v_i, we let d_i denote the degree of v_i and $\bar{d}_i$ denote the expected degree of v_i. Let $C = I - \bar{D}^{-1/2} A \bar{D}^{-1/2}$. Then by the triangle inequality, $\|\mathcal{L} - \bar{\mathcal{L}}\| \leq \|\mathcal{L} - C\| + \|C - \bar{\mathcal{L}}\|$. We consider each term separately.

Now, $C - \bar{\mathcal{L}} = \bar{D}^{-1/2}(A - \bar{A})\bar{D}^{-1/2}$. Using notation as in Section 2.3, let

$$\begin{aligned} Y_{ij} &= \bar{D}^{-1/2}((h_{ij} - p_{ij})A_{ij})\bar{D}^{-1/2} \\ &= \frac{h_{ij} - p_{ij}}{\sqrt{\bar{d}_i\bar{d}_j}} A_{ij} \end{aligned}$$

Then $C - \bar{\mathcal{L}} = \sum_{1 \le i \le j \le n} Y_{ij}$, so we can apply Theorem 2.1 to bound $\|C - \bar{\mathcal{L}}\|$. Notice $\|Y_{ij}\| \le (\bar{d}_i\bar{d}_j)^{-1/2} \le \frac{1}{\bar{d}_{\min}}$. Moreover,

$$\mathbb{E}\left[Y_{ij}^2\right] = \begin{cases} \frac{1}{\bar{d}_i\bar{d}_j}(p_{ij}(1 - p_{ij})(A_{ii} + A_{jj}) & i \ne j \\ \frac{1}{\bar{d}_i{}^2}(p_{ii}(1 - p_{ii})A_{ii} & i = j \end{cases}$$

Thus, we obtain

$$\begin{aligned} v^2 = \left\|\sum \mathbb{E}\left[Y_{ij}^2\right]\right\| &= \left\|\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{\bar{d}_i\bar{d}_j}(p_{ij} - p_{ij}^2)A_{ii}\right\| \\ &= \max_{i=1,\ldots,n}\left(\sum_{j=1}^{n} \frac{1}{\bar{d}_i\bar{d}_j}p_{ij} - \sum_{j=1}^{n} \frac{1}{\bar{d}_i\bar{d}_j}p_{ij}^2\right) \\ &\le \max_{i=1,\ldots,n}\left(\frac{1}{\bar{d}_{\min}}\sum_{j=1}^{n} \frac{p_{ij}}{\bar{d}_i}\right) = \frac{1}{\bar{d}_{\min}} \end{aligned}$$

Take $a = \sqrt{\frac{3\log(4n/\epsilon)}{\bar{d}_{\min}}}$. Take k to be large enough so that $\bar{d}_{\min} > k \log n$ implies $a < 1$. Note that $k > 3(1 + \log(\frac{\epsilon}{4}))$ is sufficient. Applying Theorem 2.1, we have

$$\begin{aligned} \mathbb{P}(\|C - \bar{\mathcal{L}}\| > a) &\le 2n \exp\left(\frac{-\frac{3\log(4n/\epsilon)}{\bar{d}_{\min}}}{\frac{2}{\bar{d}_{\min}} + \frac{2a}{3\bar{d}_{\min}}}\right) \\ &\le 2n \exp\left(-\frac{3\log(4n/\epsilon)}{3}\right) \\ &\le \epsilon/2 \end{aligned}$$

For the second term, note that by the Chernoff bound (see Theorem 1.14), for each i,

$$\mathbb{P}(|d_i - \bar{d}_i| > b\bar{d}_i) \le \frac{\epsilon}{2n} \text{ if } b \ge \sqrt{\frac{2\log(4n/\epsilon)}{\bar{d}_i}}$$

Take $b = \sqrt{\frac{\log(4n/\epsilon)}{\bar{d}_{\min}}}$, so that for all i, we have $\mathbb{P}(|d_i - \bar{d}_i| > b\bar{d}_i) \le \frac{\epsilon}{2n}$. Then we obtain

$$\|\bar{D}^{-1/2}D^{1/2} - I\| = \max_{i=1,\ldots,n}\left|\sqrt{\frac{d_i}{\bar{d}_i}} - 1\right|.$$

Note that for $0 < x < 1$, we have $|\sqrt{x} - 1| \leq |x - 1|$. Thus for $x = \frac{d_i}{\bar{d}_i} > 0$, we have that with probability at least $1 - \frac{\epsilon}{2}$,

$$\left|\frac{d_i}{\bar{d}_i} - 1\right| = \bar{d}_i \left|d_i - \bar{d}_i\right| \leq b = \sqrt{\frac{\log(4n/\epsilon)}{\bar{d}_{\min}}} = \frac{1}{\sqrt{3}} a < 1.$$

Thus we obtain

$$\|\bar{D}^{-1/2} D^{1/2} - I\| = \max_{i=1,\ldots,n} \left|\sqrt{\frac{d_i}{\bar{d}_i}} - 1\right| \leq \sqrt{\frac{\log(4n/\epsilon)}{\bar{d}_{\min}}}$$

with probability at least $1 - \frac{\epsilon}{2}$.

We note that as the Laplacian spectrum is contained in $[0, 2]$, we have $\|I - \mathcal{L}\| \leq 1$. Therefore, with probability at least $1 - \frac{\epsilon}{2}$, we have

$$\begin{aligned}
\|\mathcal{L} - C\| &= \|I - D^{-1/2} A D^{-1/2} - I + \bar{D}^{-1/2} A \bar{D}^{-1/2}\| \\
&= \|D^{-1/2} A D^{-1/2} - \bar{D}^{-1/2} D^{1/2} D^{-1/2} A D^{-1/2} D^{1/2} \bar{D}^{-1/2}\| \\
&= \|(I - \mathcal{L}) - (\bar{D}^{-1/2} D^{1/2})(I - \mathcal{L})(D^{1/2} \bar{D}^{-1/2})\| \\
&= \|(\bar{D}^{-1/2} D^{1/2} - I)(I - \mathcal{L}) D^{1/2} \bar{D}^{-1/2} + (I - \mathcal{L})(I - D^{1/2} \bar{D}^{-1/2})\| \\
&\leq \|\bar{D}^{-1/2} D^{1/2} - I\| \|D^{1/2} \bar{D}^{-1/2}\| + \|I - D^{1/2} \bar{D}^{-1/2}\| \\
&\leq b(b+1) + b = b^2 + 2b
\end{aligned}$$

Finally, as $b = \frac{1}{\sqrt{3}} a$ and $a < 1$, we have that with probability at least $1 - \epsilon$,

$$\begin{aligned}
\|\mathcal{L} - \bar{\mathcal{L}}\| &\leq \|C - \bar{\mathcal{L}}\| + \|\mathcal{L} - C\| \\
&\leq a + \frac{1}{3} a^2 + \frac{2a}{\sqrt{3}} \leq 3a,
\end{aligned}$$

completing the proof. □

As a simple application, we again consider $G_{n,p}$. We have that for $G_{n,p}$, $\bar{\mathcal{L}} = I - \frac{1}{n-1}(J - I)$. Applying Theorem 2.8, we obtain the following.

Theorem 2.9. *If* $pn \gg \log n$, *then with probability at least* $1 - 1/n = 1 - o(1)$, *we have*

$$|\lambda_k - \bar{\lambda}_k| \leq 3\sqrt{\frac{6\log(2n)}{pn}} = o(1)$$

for all $1 \leq k \leq n$.

The spectrum of $\bar{\mathcal{L}} = I - \frac{1}{n-1}(J - I)$ is $\left\{\frac{1}{n-1}, 1 + \frac{1}{n-1}\right\}$, where $1 + \frac{1}{n-1}$ has multiplicity $n - 1$. Thus, we see that if $pn \gg \log n$, then with high probability $\mathcal{L}$ has all eigenvalues other than λ_0 close to 1. This result is not new (see [22], [24]), and [24] also considers the case where $pn \leq \log n$.

Acknowledgement: This chapter is based on the paper "On the spectra of general random graphs", written jointly with Fan Chung [23]. This paper appeared in Electronic Journal of Combinatorics, 18(1), P215, 2011.

Chapter 3

Spectra of Stochastic Kronecker Graphs and Multiplicative Attribute Graphs

3.1 Introduction

In this chapter, we apply Theorems 2.6 and 2.8 to the stochastic Kronecker graph and the multiplicative attribute graph. We shall give asymptotic spectral concentration for both these models with arbitrarily sized initiator matrices, subject only to the degree requirements stated in the above two theorems. Moreover, we comment on the implications of these results for other desired graph properties, such as expansion, diameter and connectivity.

In addition, in Section 3.4, we consider the spectra of random graphs with two or more vertices that have the same behavior. We include this analysis here, as the proof is similar to the proof of concentration of spectra for MAGs.

3.2 Spectra of SKGs

We begin by examining the spectrum of the adjacency matrix of an SKG. We note that this is a straightforward application of Theorems 2.6 and 2.8, although we include the details as a prelude to the analysis for the MAG.

Theorem 3.1. *Let* $G = \mathrm{SKG}(\Theta, t)$ *with alphabet* Γ*. Let* δ *and* Δ *be the minimum and maximum diagonal entries of* $\tilde{D}$*, respectively. Let* $\epsilon > 0$*. If* $\Delta > 1$ *and*

$$t \geq \frac{\log(\frac{4}{9}\log(\frac{2k^t}{\epsilon}))}{\log(\Delta)},$$

then with probability at least $1-\epsilon$*, for all* $i \in [k^t]$ *we have*

$$|\mu_i - \bar{\mu}_i| \leq \sqrt{4\Delta^t \log\left(\frac{2k^t}{\epsilon}\right)}.$$

We note that if $\epsilon \sim n^{-c} = k^{-ct}$, we may rewrite the condition on t above so that the bound depends only upon ϵ. That is to say, the condition $t \geq \frac{\log(\frac{4}{9}\log(\frac{2k^t}{\epsilon}))}{\log(\Delta)}$ can be written as $t \geq \frac{\log(\frac{4}{9}\log(\frac{2}{\epsilon}))+\log(1+\frac{1}{2c})}{\log\Delta}$.

Proof. By Proposition 1.19, we have that the maximum expected degree is Δ^t. Moreover, as $\Delta > 1$, if $t \geq \frac{\log(\frac{4}{9}\log(\frac{2k^t}{\epsilon}))}{\log(\Delta)}$, we have

$$\begin{aligned}\Delta^t = \exp(t\log\Delta) &\geq \exp\left(\log\Delta\left(\frac{\log\left(\frac{4}{9}\log\left(\frac{2k^t}{\epsilon}\right)\right)}{\log(\Delta)}\right)\right)\\ &= \frac{4}{9}\log\left(\frac{2k^t}{\epsilon}\right) = \frac{4}{9}\log\left(\frac{2n}{\epsilon}\right).\end{aligned}$$

We thus meet the hypothesis of Theorem 2.6, so we obtain that with probability at least $1-\epsilon$, for all $i \in [k^t]$

$$|\mu_i - \bar{\mu}_i| \leq \sqrt{4\Delta^t \log(2k^t/\epsilon)}$$

as desired. □

Theorem 3.2. *Let* $G, \Gamma, \delta,$ *and* Δ *be as in Theorem 3.1. If* $\delta > 1$ *and* $t \geq \frac{\log\left(\left(3+\log\left(\frac{\epsilon}{4}\right)\right)\log(k^t)\right)}{\log\delta}$*, then with probability at least* $1-\epsilon$*, for all* $i \in [k^t]$ *we have*

$$|\lambda_j - \bar{\lambda}_j| \leq 3\sqrt{\frac{3\log(4k^t/\epsilon)}{\delta^t}}.$$

Proof. As in Theorem 3.1 above, since $\delta > 1$ and $t \geq \frac{\log\left(\left(3+\log\left(\frac{\epsilon}{4}\right)\right)\log(k^t)\right)}{\log\delta}$, we have

$$\begin{aligned}\bar{d}_{\min} = \delta^t &= \exp(t\log\delta)\\ &\geq \exp\left(\log\left(\left(3+\log\left(\frac{\epsilon}{4}\right)\right)\log(k^t)\right)\right)\\ &\geq \left(3+\log\left(\frac{\epsilon}{4}\right)\right)\log(k^t) = \left(3+\log\left(\frac{\epsilon}{4}\right)\right)\log(n).\end{aligned}$$

Thus, we satisfy the hypothesis of Theorem 2.8, so we obtain that with probability at least $1-\epsilon$, for all $i \in [k^t]$

$$|\lambda_j - \bar{\lambda}_j| \leq 3\sqrt{\frac{3\log(4k^t/\epsilon)}{\delta^t}}.$$

□

By way of example, we consider here the application of Theorems 3.1 and 3.2 to $G = K(t,\alpha,\beta,\gamma)$, the stochastic Kronecker graph with a 2×2 affinity matrix. Note that here, if we assume without loss of generality that $\alpha \geq \gamma$, we have $\Delta = \alpha+\beta$ and $\delta = \beta+\gamma$. As we have seen in Theorem 1.21, if $\beta+\gamma < 1$, then G is disconnected a.a.s., so for the sake of studying the spectrum we suppose $\beta+\gamma > 1$. This will imply that $\delta > 1$ and $\Delta > 1$, so both Theorems 3.1 and 3.2 are applicable for sufficiently large t. Note that the eigenvalues of Θ are $x \pm y$, where $x = \frac{\alpha+\gamma}{2}$ and $y = \frac{\sqrt{(\alpha-\gamma)^2+4\beta^2}}{2}$. Thus, by Theorems 1.2 and 3.1, with probability at least $1-2^{-t}$, for each i with $0 \leq i \leq t$ the adjacency matrix has $\binom{t}{i}$ eigenvalues of the form

$$(x+y)^i(x-y)^{t-i} + O(\sqrt{t(\alpha+\beta)^t}).$$

Moreover, we have that

$$\tilde{D}^{-1/2}\Theta\tilde{D}^{-1/2} = \begin{bmatrix} \frac{\alpha}{\alpha+\beta} & \frac{\beta}{\sqrt{(\alpha+\beta)(\beta+\gamma)}} \\ \frac{\beta}{\sqrt{(\alpha+\beta)(\beta+\gamma)}} & \frac{\gamma}{\beta+\gamma} \end{bmatrix},$$

which has eigenvalues 1 and $\lambda = \frac{\alpha\gamma-\beta^2}{(\alpha+\beta)(\beta+\gamma)}$. Thus, as we have shown in equation (1.9) that $\bar{\mathcal{L}} = I - (\tilde{D}^{-1/2}\Theta\tilde{D}^{-1/2})^{\otimes t}$, by Theorems 1.2 and 3.2, with probability at least $1-2^{-t}$, for each i with $0 \leq i \leq t$, the normalized Laplacian $\mathcal{L}(G)$ has $\binom{t}{i}$ eigenvalues of the form

$$1-\lambda^i + c^{-t}$$

for some $c > 1$. In particular, the spectral gap is $|\lambda| + o(1)$. Moreover, $\lambda_1 = 1 - \lambda + o(1)$ if $\alpha\gamma > \beta^2$ and $\lambda_1 = 1 - \lambda^2 + o(1)$ if $\alpha\gamma < \beta^2$.

Although it has been shown by Mahdian and Xu [55] that the diameter of $K(t, \alpha, \beta, \gamma)$ is asymptotically constant, no bound has been given on the diameter of the graph. Using the bounds on the spectrum of the graph, we can use Theorem 1.8 to obtain precise bounds on the diameter of the graph. Specifically, we obtain the following.

Theorem 3.3. *Let $G = K(t, \alpha, \beta, \gamma)$, and suppose $\beta + \gamma > 1$. Let $\hat{\lambda} = 1 - \lambda_1$, so $\hat{\lambda} = \lambda + o(1)$ if $\lambda > 0$ and $\lambda^2 + o(1)$ if $\lambda < 0$. Then the diameter of G satisfies*

$$(1 + o(1))\frac{1}{(1 - \hat{\lambda})(\alpha + 2\beta + \gamma)^t} \leq \mathrm{diam}(G) \leq (1 + o(1))\frac{2}{\hat{\lambda}} \tag{3.1}$$

a.a.s..

Proof. For the lower bound, we note that by Theorem 1.8, $\mathrm{diam}(G) \geq \frac{1}{\lambda_1 \mathrm{vol}(G)}$. As degrees in G are binomial random variables, the degrees are concentrated by Chernoff bounds (Theorem 1.14), so we obtain

$$\begin{aligned}
\mathrm{diam}(G) &\geq \frac{1}{\lambda_1 \mathrm{vol}\,(G)} \\
&\geq \frac{1}{(1 - \hat{\lambda} + o(1)) \sum_{k=0}^{t} \binom{t}{k} (\alpha + \beta)^k (\beta + \gamma)^k (1 + o(1))} \\
&\geq (1 + o(1))\frac{1}{(1 - \hat{\lambda})(\alpha + 2\beta + \gamma)^t}.
\end{aligned}$$

For the upper bound, we note that Theorem 1.8 gives

$$\mathrm{diam}(G) \leq \frac{4\sqrt{d_{\max} - 1} + 2d_{\max}}{d_{\max}(\lambda_1 - 1) + 2\sqrt{d_{\max} - 1}}.$$

As with the lower bound, concentration in the degrees of vertices gives that $d_{\max} = (1 + o(1))(\alpha + \beta)^t$. Thus we obtain

$$\begin{aligned}
\mathrm{diam}(G) &\leq \frac{4\sqrt{d_{\max} - 1} + 2d_{\max}}{d_{\max}(\lambda_1 - 1) + 2\sqrt{d_{\max} - 1}} \\
&\leq (1 + o(1))\frac{4(\alpha + \beta)^{t/2} + 2(\alpha + \beta)^t}{(\alpha + \beta)^t \hat{\lambda} + 2(\alpha + \beta)^{t/2}} \\
&= (1 + o(1))\frac{4 + 2(\alpha + \beta)^{t/2}}{(\alpha + \beta)^{t/2} \hat{\lambda} + 2} \\
&= (1 + o(1))\frac{2}{\hat{\lambda}}.
\end{aligned}$$

□

Remark 3.4. As mentioned above, the spectral gap of $K(t,\alpha,\beta,\gamma)$ is $\frac{|\alpha\gamma-\beta^2|}{(\alpha+\beta)(\beta+\gamma)} + o(1)$. As seen in Theorem 1.7, the smaller the spectral gap, the better expansion we experience. Thus in order for the Kronecker graph to be a reasonable expander, we would desire that the spectral gap be close to 0. However, this can only be attained for very specific parameter settings, namely, for β very close to $\sqrt{\alpha\gamma}$, which may not be the desired settings to produce realistic graphs [50]. Thus we see that in general, the Kronecker graph can not be assumed to be a very good expander.

3.3 Spectra of MAGs

We now turn to considering the spectrum of $\mathcal{L}$ for an MAG of the form $G = \mathrm{MAG}(n,t,\Theta,Q)$. In the following theorem we show that the qualitative behavior of the spectral properties of G are essentially that of an appropriate SKG, plus some nearly trivial eigenvalues (that is, eigenvalues very close to 1). In particular, if the distribution Q on Γ is uniform then (up to trivial eigenvalues) the spectrum of a t^{th}-order MAG is the essentially the same as that of the t^{th}-order SKG with the same affiliation matrix.

Theorem 3.5. *Let $G = \mathrm{MAG}(n,t,\Theta,Q)$. Let $\hat{D}$ be the diagonal matrix of column sums of $Q\Theta$, and let δ be the minimum diagonal entry of $\hat{D}$. Let $q_{\min}$ be the minimum diagonal entry of Q. Fix $\epsilon > 0$. If*

$$t \le \min\left\{ \frac{\log(n) - \log\left(6\log\left(\frac{8n}{\epsilon}\right)\right)}{\log\left(\frac{1}{\delta}\right)}, \frac{\log(n) - \log\left(12\log\left(\frac{2n}{\epsilon}\right)\right)}{\log\left(\frac{1}{q_{min}}\right)} \right\}, \tag{3.2}$$

then with probability at least $1-\epsilon$ there is a set $A \subset [n]$, where $A = \{a_1, \ldots, a_{k^t}\}$, such that for $i \in [k^t]$,

$$\left|\lambda_{a_i}(\mathcal{L}(G)) - \left(1 - \lambda_{k^t+1-i}\left(\left(\hat{D}^{-1/2}Q^{1/2}\Theta Q^{1/2}\hat{D}^{-1/2}\right)^{\otimes t}\right)\right)\right|$$
$$\le 3\sqrt{\frac{6\log\left(\frac{8n}{\epsilon}\right)}{\delta^t n}} + 4\sqrt{\frac{3\log\left(\frac{2n}{\epsilon}\right)}{q_{\min}^t n}}$$

$$\leq \sqrt{\frac{108 \log\left(\frac{8n}{\epsilon}\right)}{n \min\left\{q_{\min}^t, \delta^t\right\}}}.$$

Furthermore, for all $j \notin A$, $|\lambda_j(\mathcal{L}(G)) - 1| \leq 3\sqrt{\frac{6\log\left(\frac{8n}{\epsilon}\right)}{\delta^t n}}$.

Proof. In order to analyze the spectrum of G we consider the random assignment of vertices to words in Γ^t separately from the random generation of edges.

Fix the signature $\{n_\sigma\}_{\sigma\in\Gamma^t}$ of the graph G, and define for each $\sigma \in \Gamma^t$, $d_\sigma = \sum_{\tau\in\Gamma^t} n_\tau \Theta^{\otimes t}_{\sigma,\tau}$. We note that for any vertex v, $\mathbb{E}\left[\deg(v)\right] = d_{a(v)}$ and thus the minimum expected degree is $\bar{d}_{\min} = \min_{\sigma\in\Gamma^t} d_\sigma$ for graphs with a fixed signature. Now using Theorem 2.8, if $\bar{d}_{\min} \geq 3\log\left(\frac{8n}{\epsilon}\right)$, then with probability at least $1 - \frac{\epsilon}{2}$,

$$|\lambda_i(\mathcal{L}(G)) - (1 - \lambda_{n-i+1}(M))| \leq 3\sqrt{\frac{3\log\left(\frac{8n}{\epsilon}\right)}{\bar{d}_{\min}}}$$

for all i, where

$$M_{u,v} = \frac{\Theta^{\otimes t}_{a(u),a(v)}}{\sqrt{d_{a(u)}d_{a(v)}}}.$$

In order to understand the spectrum of M, we consider the case where every element of the signature is at least 1, that is, $n_\sigma \geq 1$ for all $\sigma \in \Gamma^t$. Thus for every $\sigma \in \Gamma^t$, there exists a vertex $v \in V$ such that $a(v) = \sigma$. We will abuse notation and refer to any such vertex as $a^{-1}(\sigma)$. Thus we may define the $k^t \times k^t$ matrix by $\tilde{M}_{\sigma,\tau} = \sqrt{n_\sigma n_\tau} M_{a^{-1}(\sigma),a^{-1}(\tau)}$. We claim that $\tilde{M}$ captures the non-trivial portion of the spectrum of M.

Observe that for any two vertices u and v with $a(u) = a(v)$, the corresponding rows of M are identical, and thus for each $\sigma \in \Gamma^t$, M has an eigenvalue of 0 of multiplicity $n_\sigma - 1$. Hence, the multiplicity of 0 as an eigenvalue of M is at least $n - k^t$. In order to show that the remaining eigenvalues of M are the spectrum of $\tilde{M}$, let φ be an eigenvector for M with corresponding eigenvalue $\lambda \neq 0$. Define $\psi \in \mathbb{R}^{k^t}$ by

$$\psi(\sigma) = \frac{1}{n_\sigma} \sum_{a(v)=\sigma} \varphi(v).$$

Now for any $v \in V(G)$ with $a(v) = \sigma$,

$$\lambda\varphi(v) = M\varphi(v) = \sum_{u\in V} M_{v,u}\varphi(u) = \sum_{\tau\in\Gamma^t} \tilde{M}_{\sigma,\tau} n_\tau \psi(\tau).$$

As this quantity is independent of the choice of v (up to choice of representative) this implies that the eigenvector φ has $\varphi(u) = \varphi(v)$ as long as $a(u) = a(v)$. Thus we define $\tilde{\varphi} \in \mathbb{R}^{k^t}$ by

$$\tilde{\varphi}(\sigma) = \sqrt{n_\sigma}\psi(\sigma)$$

and we have for $a(v) = \sigma$

$$\begin{aligned}
\tilde{M}\tilde{\varphi}(\sigma) &= \sum_{\tau\in\Gamma^t} \tilde{M}_{\sigma,\tau}\tilde{\varphi}(\tau) \\
&= \sum_{\tau\in\Gamma^t} M_{a^{-1}(\sigma),a^{-1}(\tau)}\sqrt{n_\sigma n_\tau}\sqrt{n_\tau}\psi(\tau) \\
&= \sqrt{n_\sigma}\sum_{\tau\in\Gamma^t} M_{a^{-1}(\sigma),a^{-1}(\tau)} n_\tau \psi(\tau) \\
&= \sqrt{n_\sigma} M\varphi(v) \\
&= \sqrt{n_\sigma}\lambda\psi(\sigma) \\
&= \lambda\tilde{\varphi}(\sigma).
\end{aligned}$$

Therefore, the nonzero eigenvalues of M are the same as the nonzero eigenvalues of $\tilde{M}$, and hence it suffices to consider the spectrum of $\tilde{M}$.

For each $\sigma \in \Gamma^t$ define $q_\sigma = Q^{\otimes t}_{\sigma,\sigma}$, that is, q_σ is the probability that an arbitrary vertex v has $a(v) = \sigma$. Now the expected value of n_σ is $q_\sigma n \geq q^t_{\min} n$, and thus by Chernoff bounds (see Theorem 1.14), $(1-\epsilon^*)\, q_\sigma n \leq n_\sigma \leq (1+\epsilon^*) q_\sigma n$ with probability at least $1 - \frac{\epsilon}{2}$, where

$$\epsilon^* = \sqrt{\frac{3\log\left(\frac{2k^t}{\epsilon}\right)}{q^t_{\min} n}} \leq \sqrt{\frac{3\log\left(\frac{2n}{\epsilon}\right)}{q^t_{\min} n}} \leq \frac{1}{2}.$$

We note that $\delta = \min_{j\in[k]}\left(\sum_{i=1}^k q_i p_{ij}\right) \leq \sum_{i=1}^k q_i = 1$. Moreover, equality holds only in the case that Θ is the all ones matrix, in which case the graph is complete and the theorem is trivial. Thus, we may assume $\delta < 1$, and we obtain

$$\bar{d}_{\min} \geq \frac{1}{2}\delta^t n \geq \frac{1}{2} n \delta^{\frac{\log\left(6\log\left(\frac{8n}{\epsilon}\right)\right) - \log(n)}{\log(\delta)}} = 3\log\left(\frac{8n}{\epsilon}\right),$$

and thus with probability at least $1-\epsilon$, the spectrum of $\mathcal{L}(G)$ is controlled by the spectrum of $\tilde{M}$ with a signature near the expected signature. Thus we define

$$\begin{aligned}
\mathcal{M}_{\sigma,\tau} &= \sqrt{q_\sigma n q_\tau n}\frac{\Theta^{\otimes t}_{\sigma,\tau}}{\sqrt{\sum_{\eta\in\Gamma^t} nq_\eta \Theta^{\otimes t}_{\sigma,\eta} \sum_{\nu\in\Gamma^t} nq_\nu \Theta^{\otimes t}_{\tau,\nu}}} \\
&= \sqrt{q_\sigma q_\tau}\frac{\Theta^{\otimes t}_{\sigma,\tau}}{\sqrt{\sum_{\eta\in\Gamma^t} q_\eta \Theta^{\otimes t}_{\sigma,\eta} \sum_{\nu\in\Gamma^t} q_\nu \Theta^{\otimes t}_{\tau,\nu}}} \\
&= \sqrt{Q^{\otimes t}_{\sigma,\sigma} Q^{\otimes t}_{\tau,\tau}}\frac{\Theta^{\otimes t}_{\sigma,\tau}}{\sqrt{\sum_{\eta\in\Gamma^t} Q^{\otimes t}_{\eta,\eta} \Theta^{\otimes t}_{\sigma,\eta} \sum_{\nu\in\Gamma^t} Q^{\otimes t}_{\nu,\nu} \Theta^{\otimes t}_{\tau,\nu}}} \\
&= \sqrt{Q^{\otimes t}_{\sigma,\sigma} Q^{\otimes t}_{\tau,\tau}}\frac{\Theta^{\otimes t}_{\sigma,\tau}}{\sqrt{\hat{D}^{\otimes t}_{\sigma,\sigma} \hat{D}^{\otimes t}_{\tau,\tau}}} \\
&= \left(\hat{D}^{-1/2} Q^{1/2} \Theta Q^{1/2} \hat{D}^{-1/2}\right)^{\otimes t}_{\sigma,\tau},
\end{aligned}$$

where the last two equalities come from the fact that both Q and $\hat{D}$ are diagonal matrices.

We make the standard observation that for any matrix A and invertible matrix S, the spectrum of $S^{-1}AS^{-1}$ is the same as the spectrum of $S^{-2}A$, as the eigenpairs (λ, v) for $S^{-1}AS^{-1}$ correspond to the eigenpairs $(\lambda, S^{-1}v)$ for $S^{-2}A$. In particular

$$\|\mathcal{M}\| = \left\|\hat{D}^{-1/2} Q^{1/2} \Theta Q^{1/2} \hat{D}^{-1/2}\right\| = \left\|\hat{D}^{-1} Q \Theta\right\| = 1,$$

as $\hat{D}$ is formed from the column sums of $Q\Theta$. We thus obtain

$$\begin{aligned}
\left\|\tilde{M} - \mathcal{M}\right\| &= \max_{\|f\|=1} \left| f^T \left(\tilde{M} - \mathcal{M}\right) f \right| \\
&= \max_{\|f\|=1} \left| \sum_{\sigma\in\Gamma^t} \sum_{\tau\in\Gamma^t} f_\sigma \left(\tilde{M} - \mathcal{M}\right)_{\sigma,\tau} f_\tau \right| \\
&\leq \max_{\|f\|=1} \sum_{\sigma\in\Gamma^t} \sum_{\tau\in\Gamma^t} |f_\sigma| \left|\tilde{M} - \mathcal{M}\right|_{\sigma,\tau} |f_\tau| \\
&\leq \max_{\|f\|=1} \sum_{\sigma\in\Gamma^t} \sum_{\tau\in\Gamma^t} |f_\sigma| \frac{2\epsilon^*}{1-\epsilon^*} |\mathcal{M}_{\sigma,\tau}| |f_\tau| \\
&= \frac{2\epsilon^*}{1-\epsilon^*} \max_{\|f\|=1} \sum_{\sigma\in\Gamma^t} \sum_{\tau\in\Gamma^t} |f_\sigma| \mathcal{M}_{\sigma,\tau} |f_\tau| \\
&\leq \frac{2\epsilon^*}{1-\epsilon^*} \|\mathcal{M}\| \\
&= \frac{2\epsilon^*}{1-\epsilon^*}
\end{aligned}$$

Thus, by Weyl's theorem (Theorem 1.1), for any i, $\left|\lambda_i(\tilde{M}) - \lambda_i(\mathcal{M})\right| \leq \frac{2\epsilon^*}{1-\epsilon^*} \leq 4\epsilon^*$. Therefore, for k^t eigenvalues λ of $\mathcal{L}$, we obtain

$$\begin{aligned} |\lambda - \lambda_i(\mathcal{M})| &\leq \left|\lambda - \lambda_i(\tilde{M})\right| + \left|\lambda_i(\tilde{M}) - \lambda_i(\mathcal{M})\right| \\ &\leq 3\sqrt{\frac{3\log\left(\frac{8n}{\epsilon}\right)}{\bar{d}_{\min}}} + 4\epsilon^* \\ &\leq 3\sqrt{\frac{3\log\left(\frac{8n}{\epsilon}\right)}{\bar{d}_{\min}}} + 4\sqrt{\frac{3\log\left(\frac{2k^t}{\epsilon}\right)}{q_{\min}^t n}} \end{aligned}$$

as desired. Moreover, by the observation above about the remaining eigenvalues, we have that all other eigenvalues satisfy

$$|\lambda - 1| \leq 3\sqrt{\frac{3\log\left(\frac{8n}{\epsilon}\right)}{\bar{d}_{\min}}}.$$

□

We now consider the results of this theorem applied to the special case $K(n, t, \Theta, \mu)$, where $\Gamma = \{0, 1\}$, $\Theta = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$, and $Q = \begin{pmatrix} \mu & 0 \\ 0 & 1-\mu \end{pmatrix}$. We will assume that $t \sim \rho \log n$. In this case, we have

$$\hat{D} = \begin{pmatrix} \mu\alpha + (1-\mu)\beta & 0 \\ 0 & \mu\beta + (1-\mu)\gamma \end{pmatrix},$$

so $\delta = \min\{\mu\alpha + (1-\mu)\beta, \mu\beta + (1-\mu)\gamma\}$. We shall make the standard assumption that $\mu\alpha + (1-\mu)\beta > \mu\beta + (1-\mu)\gamma$; this is without loss of generality as the graph is symmetric in 0s and 1s. Thus, $\delta = \mu\beta + (1-\mu)\gamma$.

Let $c_1 = \mu\alpha + (1-\mu)\beta$ and $c_2 = \mu\beta + (1-\mu)\gamma$. We thus obtain that

$$\hat{D}^{-1/2}Q^{1/2}\Theta Q^{1/2}\hat{D}^{-1/2} = \begin{pmatrix} \frac{\mu\alpha}{c_1} & \beta\sqrt{\frac{\mu(1-\mu)}{c_1c_2}} \\ \beta\sqrt{\frac{\mu(1-\mu)}{c_1c_2}} & \frac{(1-\mu)\gamma}{c_2} \end{pmatrix},$$

which has eigenvalues 1 and $\lambda = \frac{\mu(1-\mu)(\alpha\gamma-\beta^2)}{(\mu\alpha+(1-\mu)\beta)(\mu\beta+(1-\mu)\gamma)}$. Therefore, we have that if t satisfies condition (3.2), then with probability at least $1 - \frac{1}{n}$, G has $\binom{t}{i}$ eigenvalues of the form $1 - \lambda^i + o(1)$, and all remaining eigenvalues are of the form

$1 + o(1)$. Note that since $\rho > 0$, it suffices to satisfy the bound (3.2) that $\rho \leq \min\{\frac{1}{\log(1/c_2)}, \frac{1}{\log(1/\mu)}, \frac{1}{\log(1/(1-\mu))}\}$.

We thus have that the spectral gap is $|\lambda| + o(1)$. Moreover, if $\alpha\gamma > \beta^2$, we have $\lambda_1 = 1 - \lambda$ and if $\alpha\gamma < \beta^2$, then $\lambda_1 = 1 - \lambda^2$. We note that if G satisfies the bound (3.2), we must have that $\rho \leq \frac{1}{\log(\frac{1}{c_2})}$. This is precisely the condition required for G to be connected, as noted in Theorem 1.24.

As with the stochastic Kronecker graph, this concentration of eigenvalues allows us to obtain precise bounds on the diameter of G. That the diameter is constant was established in [43], but no precise bounds were provided.

Theorem 3.6. *For $G = K(n, t, \Theta, \mu)$ as above, satisfying (3.2), put $\hat{\lambda} = 1 - \lambda_1$, so $\hat{\lambda} = \lambda + o(1)$ if $\lambda > 0$ and $\lambda^2 + o(1)$ if $\lambda < 0$. Then we have*

$$(1 + o(1))\frac{1}{(1 + \hat{\lambda})n^{1+c\log(\mu\alpha+\beta+(1-\mu)\gamma)}} \leq \operatorname{diam}(G) \leq (1 + o(1))\frac{2}{\hat{\lambda}}.$$

Proof. The proof is similar to that of Theorem 3.3. For the lower bound, we use Theorem 1.8 and concentration on both the number of vertices of a particular weight and degree of vertices to obtain

$$\begin{aligned}
\operatorname{diam}(G) &\geq \frac{1}{\lambda_1 \operatorname{vol}(G)} \\
&\geq \frac{1}{(1 + o(1))(1 - \hat{\lambda}) \sum_{k=0}^{t} \binom{t}{k} n c_1^k c_2^{t-k} (1 + o(1))} \\
&\geq (1 + o(1))\frac{1}{(1 - \hat{\lambda})n(c_1 + c_2)^{\rho \log n}} \\
&= (1 + o(1))\frac{1}{(1 - \hat{\lambda})n^{1+\rho\log(c_1+c_2)}},
\end{aligned}$$

the desired result.

For the upper bound, we note that the expected maximum degree is $(1 + o(1))nc_1^t$ by concentration of degrees, and the assumption that $c_1 > c_2$. Using

Theorem 1.8, we thus obtain

$$\begin{aligned}
\operatorname{diam}(G) &\leq \frac{4\sqrt{d_{\max}-1}+2d_{\max}}{d_{\max}(1-\lambda_1)+2\sqrt{d_{\max}-1}} \\
&\leq (1+o(1))\frac{4n^{1/2}c_1^{t/2}+2nc_1^t}{nc_1^t\hat{\lambda}+2n^{1/2}c_1^{t/2}} \\
&\leq (1+o(1))\frac{4+2\sqrt{n}c_1^{t/2}}{\hat{\lambda}\sqrt{n}c_1^{t/2}+2} \\
&\leq (1+o(1))\frac{4+2n^{\frac{1}{2}+\frac{\rho}{2}\log(c_1)}}{\hat{\lambda}n^{\frac{1}{2}+\frac{\rho}{2}\log(c_1)}+2}
\end{aligned}$$

We note that as $c_1^\rho \geq c_2^\rho > \frac{1}{e}$, we must have $\rho\log(c_1) > -1$, and thus $\frac{1}{2}+\frac{\rho}{2}\log(c_1) > 0$. Thus, we obtain

$$\operatorname{diam}(G) \leq (1+o(1))\frac{2}{\hat{\lambda}},$$

as desired. □

Remark 3.7. As noted above, the spectral gap in $K(n,t,\Theta,\mu)$ is

$$\frac{\mu(1-\mu)\,|\alpha\gamma-\beta^2|}{c_1c_2}+o(1) = \left|1-\frac{\mu^2\alpha\beta+(1-\mu)^2\beta\gamma}{(\mu\alpha+(1-\mu)\beta)(\mu\beta+(1-\mu)\gamma)}\right|+o(1).$$

It is unclear what settings of parameters will make this relatively small, to ensure good expansion. Certainly, as noted with $K(t,\alpha,\beta,\gamma)$ in Remark 3.4, it would be sufficient to require β to be close to $\sqrt{\alpha\gamma}$, however this choice of parameters does not seem to necessarily generate realistic graphs [44].

3.4 Spectra of Graphs with Repeated Vertices

In this section we consider random graphs with repeated vertices. By "repeated vertices," we mean a collection of vertices $v_1, v_2, \ldots, v_k$ such that for all $u \in V(G)$, we have $\mathbb{P}(v_i \sim u) = \mathbb{P}(v_j \sim u)$ for all $1 \leq i,j \leq k$. Similar problems were considered for deterministic graphs in [7], [8] and [11]. Here we consider the problem of determining eigenvalues of a random graph with repeated vertices, and in fact we find that if a graph has a large enough minimum expected degree, we may choose vertices to repeat uniformly at random essentially without changing the eigenvalues of the graph.

Let $\Sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_k\}$, where for each pair (σ_i, σ_j) we associate a probability p_{ij}, and let P be the $k \times k$ matrix indexed by Σ with $P_{\sigma_i,\sigma_j} = p_{ij}$. Let Q be a probability distribution on Σ, which we shall represent as a diagonal $k \times k$ matrix with $Q_{ii} = Q(\sigma_i) = q_i = q_{\sigma_i}$. Define a graph G as follows. Let n be the number of vertices of G. To each vertex in G we associate an element of Σ, called the *type* of v and denoted by $t(v)$. We choose $t(v)$ independently for each v according to the distribution Q. For two vertices u and v of type $t(u) = \sigma_i$ and $t(v) = \sigma_j$, we define $\mathbb{P}(u \sim v) = p_{ij}$. We denote this graph by $G(n, \Sigma, P, Q)$. For each i, let n_{σ_i} denote the number of vertices of type σ_i. We call the collection $\{n_\sigma\}_{\sigma\in\Sigma}$ the *signature* of the graph G.

This is in fact a generalization of $\mathrm{MAG}(n, t, \Theta, Q)$, where $\Sigma = \Gamma^t$, the distribution Q in $G(n, \Sigma, P, Q)$ is given by the product distribution $Q^{\otimes t}$ in $\mathrm{MAG}(n, t, \Theta, Q)$, and P is defined via attribute affinity in Θ.

Essentially, we can think of Σ as a random graph in which the probability that $\sigma_i \sim \sigma_j$ is p_{ij}. We then produce a new graph G by repeating vertices in Σ. In this way we obtain clusters of vertices that behave identically with respect to the rest of the graph, and the graph induced on each cluster is G_{n_i,p_i}.

In the $\mathrm{MAG}(n, t, \Theta, Q)$, the underlying graph is in fact an SKG, namely $\mathrm{SKG}(\Theta, t)$. Each word in Γ^t forms a cluster of vertices as described, which we saw in Theorem 3.5 essentially have no effect on the spectrum of $\mathcal{L}$ for the graph. In fact, this is generally true for these graphs $G(n, \Sigma, P, Q)$.

Theorem 3.8. *Let $G = G(n, \Sigma, P, Q)$. Let $\hat{D}$ be the diagonal matrix of column sums of QP. Let δ be the minimum diagonal entry of $\hat{D}$. Then if $\delta > C\frac{\log n}{n}$ for C sufficiently large, there exists a set $A = \{a_1, a_2, \ldots, a_k\} \subset [n]$ such that for $i \in [k]$,*

$$\left|\lambda_{a_i} - \left(1 - \lambda_i(\hat{D}^{-1/2}Q^{1/2}PQ^{1/2}D^{-1/2})\right)\right| \leq \sqrt{\frac{118\log\left(\frac{4n}{\epsilon}\right)}{\min\{\bar{d}_{\min}, nq_{\min}\}}}.$$

Moreover, for $i \notin A$, we have

$$|\lambda_i - 1| \leq 3\sqrt{\frac{3\log\left(\frac{4n}{\epsilon}\right)}{\bar{d}_{\min}}}.$$

The precise value of C will come out of the proof.

Proof. We first fix the signature $\{n_\sigma\}_{\sigma\in\Sigma}$. For each i, define $d_{\sigma_i} = d_i = \sum_{1\leq j\leq k} n_{\sigma_j} p_{ij}$, so that for any vertex v, $\mathbb{E}[\deg(v)] = d_{t(v)}$. Moreover, the minimum expected degree is $\bar{d}_{\min} = \min_{\sigma\in\Sigma} d_\sigma$. Define M to be the $n \times n$ matrix with

$$M_{u,v} = \frac{P_{t(v),t(u)}}{\sqrt{d_{t(v)} d_{t(u)}}}.$$

Note that if $\bar{d}_{\min} \geq k(\epsilon)\log(n)$, then by Theorem 2.8, with probability at least $1 - \frac{\epsilon}{2}$, we have

$$|\lambda_i(\mathcal{L}) - (1 - \lambda_{n-i+1}(M))| \leq 3\sqrt{\frac{3\log\left(\frac{4n}{\epsilon}\right)}{\bar{d}_{\min}}}$$

for all i.

Suppose $n_\sigma \geq 1$ for all $\sigma \in \Sigma$. Then for every $\sigma \in \Sigma$, there exists a vertex $v \in V$ such that $t(v) = \sigma$. We shall abuse notation and refer to any such vertex as $t^{-1}(\sigma)$. We thus define the $k \times k$ matrix $\tilde{M}$ by $\tilde{M}_{\sigma,\tau} = \sqrt{n_\sigma n_\tau} M_{t^{-1}(\sigma),t^{-1}(\tau)}$.

Note that for any two vertices u, v with $t(u) = t(v)$, the corresponding rows of M are identical, and thus for each $\sigma \in \Sigma$, M has an eigenvalue of 0 of multiplicity $n_\sigma - 1$. Thus, the multiplicity of 0 as an eigenvalue of M is at least $\sum_{\sigma\in\Sigma}(n_\sigma - 1) = n - k$.

Moreover, if φ is an eigenvector for M with eigenvalue $\lambda \neq 0$, notice that for any $v \in t^{-1}(\sigma)$, we have

$$\begin{aligned}\lambda\varphi(v) = M\varphi(v) &= \sum_{u\in V} M_{v,u}\varphi(u) \\ &= \sum_{\tau\in\Sigma} \tilde{M}_{\sigma,\tau}\left(\sum_{v\in t^{-1}(\tau)} \varphi(v)\right).\end{aligned}$$

Note that this is independent of the choice of representative v, so we must have that each $v \in t^{-1}(\sigma)$ has the same value for φ. Put $\tilde{\varphi} \in \mathbb{R}^k$ to be the vector with $\tilde{\varphi}(\sigma) = \sqrt{n_\sigma}\varphi(t^{-1}(\sigma))$. Then we obtain

$$\begin{aligned}
\tilde{M}\tilde{\varphi}(\sigma) &= \sum_{\tau\in\Sigma} \tilde{M}_{\sigma,\tau}\tilde{\varphi}(\tau) \\
&= \sum_{\tau\in\Sigma} M_{t^{-1}(\sigma),t^{-1}(\tau)}\sqrt{n_\sigma n_\tau}\sqrt{n_\tau}\varphi(t^{-1}(\sigma)) \\
&= \sqrt{n_\sigma}\sum_{\tau\in\Sigma} n_\tau M_{t^{-1}(\sigma),t^{-1}(\tau)}\varphi(t^{-1}(\sigma)) \\
&= \sqrt{n_\sigma}M\varphi(t^{-1}(\sigma)) \\
&= \lambda\sqrt{n_\sigma}\varphi(t^{-1}(\sigma)) = \lambda\tilde{\varphi}(\sigma).
\end{aligned}$$

Therefore, the nonzero eigenvalues of M are the same as the nonzero eigenvalues of $\tilde{M}$. Hence, it suffices to consider the spectrum of $\tilde{M}$.

Note that for each $\sigma \in \Sigma$, the expected value of n_σ is $q_\sigma n \geq q_{\min} n$, where $q_{\min} = \min_{\sigma\in\Sigma} q_\sigma$. Therefore, by Chernoff bounds (Theorem 1.14), we have that $(1-\epsilon^*)q_\sigma n \leq n_\sigma \leq (1+\epsilon^*)q_\sigma n$ for all σ with probability at least $1 - \frac{\epsilon}{2}$, where

$$\epsilon^* = \sqrt{\frac{2\log\left(\frac{4k}{\epsilon}\right)}{q_{\min} n}} \leq \sqrt{\frac{2\log\left(\frac{4n}{\epsilon}\right)}{q_{\min} n}} \leq \frac{1}{2}.$$

Therefore, we have that with probability at least $1 - \frac{\epsilon}{2}$ the minimum expected degree in G satisfies

$$\begin{aligned}
d_{\min} &= \min_{1\leq i\leq k}\sum_{j=1}^{k} n_j p_{ij} \\
&\geq \min_{1\leq i\leq k}\sum_{j=1}^{n} (1-\epsilon^*)n q_j p_{ij} \\
&= (1-\epsilon^*)n\delta \geq (1-\epsilon^*)C\log n
\end{aligned}$$

We choose C so that $(1-\epsilon^*)C > k(\epsilon)$ as in Theorem 2.8. Thus, we obtain that with probability at least $1-\epsilon$, we have concentration of the Laplacian eigenvalues of G on the eigenvalues of $I - M$, and G has a signature within a factor of ϵ^* of the expected signature. It remains only to consider the eigenvalues of $\tilde{M}$. Define

$$\begin{aligned}
\mathcal{M}_{\sigma,\tau} &= \sqrt{q_\sigma n q_\tau n}\frac{P_{\sigma,\tau}}{\sqrt{\sum_{\eta\in\Sigma} nq_\eta P_{\sigma,\eta}\sum_{\nu\in\Sigma} nq_\nu P_{\tau,\nu}}} \\
&= \sqrt{q_\sigma q_\tau}\frac{P_{\sigma,\tau}}{\sqrt{\sum_{\eta\in\Sigma} q_\eta P_{\sigma,\eta}\sum_{\nu\in\Sigma} q_\nu P_{\tau,\nu}}} \\
&= \sqrt{Q_{\sigma,\sigma} Q_{\tau,\tau}}\frac{P_{\sigma,\tau}}{\sqrt{\sum_{\eta\in\Sigma} Q_{\eta,\eta} P_{\sigma,\eta}\sum_{\nu\in\Sigma} Q_{\nu,\nu} P_{\tau,\nu}}} \\
&= \sqrt{Q_{\sigma,\sigma} Q_{\tau,\tau}}\frac{P_{\sigma,\tau}}{\sqrt{\hat{D}_{\sigma,\sigma}\hat{D}_{\tau,\tau}}} \\
&= \left(\hat{D}^{-1/2} Q^{1/2} P Q^{1/2} D^{-1/2}\right)_{\sigma,\tau},
\end{aligned}$$

as both Q and $\hat{D}$ are diagonal matrices.

As in the proof of Theorem 3.5, we note that for any matrix A and invertible matrix S, the spectrum of $S^{-1}AS^{-1}$ is the same as that of $S^{-2}A$. Therefore,

$$\|\mathcal{M}\| = \|\hat{D}^{-1/2} Q^{1/2} P Q^{1/2} D^{-1/2}\| = \|\hat{D}^{-1} Q P\| = 1,$$

as $\hat{D}$ is precisely the column sums of QP. Moreover,

$$\begin{aligned}
\|\tilde{M} - \mathcal{M}\| &= \max_{\|f\|=1} \left| f^T(\tilde{M} - \mathcal{M}) f \right| \\
&= \max_{\|f\|=1} \left| \sum_{\sigma\in\Sigma}\sum_{\tau\in\Sigma} f_\sigma (\tilde{M} - \mathcal{M})_{\sigma,\tau} f_\tau \right| \\
&\leq \max_{\|f\|=1} |f_\sigma| \left| \tilde{M} - \mathcal{M}_{\sigma,\tau} \right| |f_\tau| \\
&\leq \max_{\|f\|=1} |f_\sigma| \left(\frac{2\epsilon^*}{1-\epsilon^*} \right) \mathcal{M}_{\sigma,\tau} |f_\tau| \\
&= \frac{2\epsilon^*}{1-\epsilon^*} \|\mathcal{M}\| = \frac{2\epsilon^*}{1-\epsilon^*}.
\end{aligned}$$

Therefore, by Wey's Theorem (Theorem 1.1),

$$\left| \lambda_i(\tilde{M}) - \lambda_i(\mathcal{M}) \right| \leq \frac{2\epsilon^*}{1-\epsilon^*} \leq 4\epsilon^*.$$

We thus obtain that G has k eigenvalues λ with

$$\begin{aligned}\left|\lambda - \left(1 - \lambda_i(\hat{D}^{-1/2}Q^{1/2}PQ^{1/2}D^{-1/2})\right)\right| &\leq 3\sqrt{\frac{3\log\left(\frac{4n}{\epsilon}\right)}{\bar{d}_{\min}}} + 4\sqrt{\frac{2\log\left(\frac{4n}{\epsilon}\right)}{q_{\min}n}} \\ &\leq \sqrt{\frac{118\log\left(\frac{4n}{\epsilon}\right)}{\min\{\bar{d}_{\min}, nq_{\min}\}}}.\end{aligned}$$

Moreover, all remaining eigenvalues satisfy

$$|\lambda - 1| \leq 3\sqrt{\frac{3\log\left(\frac{4n}{\epsilon}\right)}{\bar{d}_{\min}}}.$$

□

Corollary 3.9. *Let H be a random graph with h vertices and minimum expected degree $\bar{d}_{\min} > r\log h$. Let $\epsilon > 0$, and let G be obtained from H by choosing $n \leq h^{r/C}$ vertices uniformly at random from the vertex set of H, where C is the constant defined in Theorem 3.8, and joining vertices with the same probability as in H. Then with probability at least $1-\epsilon$, the graph G has h eigenvalues λ that satisfy*

$$|\lambda - \lambda_i(H)| \leq \sqrt{\frac{118\log\left(\frac{4n}{\epsilon}\right)}{\min\{\bar{d}_{\min}, \frac{n}{h}\}}},$$

and all remaining eignvalues of G satisfy

$$|\lambda - 1| \leq 3\sqrt{\frac{3\log\left(\frac{4n}{\epsilon}\right)}{\bar{d}_{\min}}}.$$

Proof. Let $P = \bar{A}_H$. As we have chosen the vertices of G uniformly at random from H, we have $Q = \frac{1}{h}I$, so $QP = \frac{1}{h}P$ and $\hat{D} = \frac{1}{h}\bar{D}_H$. Thus, we have

$$\delta = \frac{1}{h}r\log h \geq \frac{1}{n}C\log h^{r/C} = C\frac{\log n}{n},$$

and we may apply Theorem 3.8. To complete the result, notice that

$$\begin{aligned}\hat{D}^{-1/2}Q^{1/2}PQ^{1/2}D^{-1/2} &= (\frac{1}{h}\bar{D})^{-1/2}(\frac{1}{h}I)^{1/2}P(\frac{1}{h}I)^{1/2}(\frac{1}{h}\bar{D})^{-1/2} \\ &= \sqrt{h}\bar{D}_H^{-1/2}\frac{1}{\sqrt{h}}\bar{A}_H\frac{1}{\sqrt{h}}\sqrt{h}\bar{D}_H^{-1/2} \\ &= I - \bar{\mathcal{L}}(H).\end{aligned}$$

□

Acknowledgement: Part of this chapter is based on the paper "The spectra of multiplicative attribute graphs", written jointly with Stephen Young [61]. This paper has been submitted.

Part of this chapter is based on unpublished joint work with Stephen Young.

Chapter 4

Giant components in Stochastic Kronecker and Multiplicative Attribute Graphs

4.1 Introduction

A giant component in a graph G on n vertices is a connected component of size $\Theta(n)$. In [31], Erdős and Rényi analyzed the random graph $G_{n,p}$ as p changes from 0 to 1. Parametrizing p as $p = \frac{c}{n}$, they observed a sharp change in behavior, occurring at $c = 1$. In particular, if $c < 1$, the graph is composed of components with size at most $O(\log n)$, but if $c > 1$, the graph contains one giant component of size $\Theta(n)$, and all other components of size $O(\log n)$. The value $c = 1$ is called the threshold for the emergence of the giant component in $G_{n,p}$.

As the study of random graphs has grown, this type of behavior has been observed in many different random graph models [58], [57], [20], [39], [13], [34], [37]. That is to say, in many different random graph models, one can find a threshold for the parameters that define the model so that if one chooses parameters below the threshold, all components are of size $o(n)$, but if one chooses parameters above the threshold, there exists a giant component of size $\Theta(n)$.

The study of giant components fundamentally consists of three aspects.

First, and most importantly, is the determination of the giant component threshold. Once such a threshold has been established, two questions naturally arise: is the giant component unique in the graph, and what is its asymptotic size? As the giant component has size Cn for some constant C, it is natural to ask if we can bound C, and if it is possible to have two components with size $\Theta(n)$.

Finally, it is of interest to study what happens at the point of transition. This has been widely studied for various graphs [12], [19], [29], [42], [54]. In $G_{n,p}$, it is known that if $p = \frac{1}{n} + \frac{\lambda}{n^{2/3}}$, the value of λ (which may itself be a function of n) determines the component structure of $G_{n,p}$.

In this chapter, we consider the questions of emergence, uniqueness, and size of the giant component for $K(t, \alpha, \beta, \gamma)$ and $K(n, t, \Theta, \mu)$. For both of these random graphs, we provide a necessary and sufficient condition on parameters for the emergence of the giant component, as well as establish its uniqueness and asymptotic size.

Throughout this chapter, as we will be studying random graphs with attribute affinity where the attribute vectors are binary, we will make extensive use of bounds on binomial coefficients. The following theorem, known as the entropy bounds, will be particularly useful.

Theorem 4.1 (Entropy Bounds). *[63] For any n and $0 < p < 1$,*

$$\binom{n}{pn} > \frac{p^{-pn}(1-p)^{-(1-p)n}}{e\sqrt{2\pi np(1-p)}} \tag{4.1}$$

and

$$\binom{n}{pn} < \frac{e^{nH(p)}}{\sqrt{2\pi np(1-p)}}. \tag{4.2}$$

Moreover, in this chapter we will often talk about binomial coefficients of the form $\binom{n}{pn}$. We will typically treat pn as an integer for the purpose of bounding the coefficients. Of course, it may not be the case that pn is in fact integral; in that case we should be taking $\lfloor pn \rfloor$ in its place. However, the computations we shall use (such as Theorem 4.1) will differ by at most a $1 + o(1)$ factor from the true computations using $\lfloor pn \rfloor$, and as such we will omit the floors.

4.2 Giant Components in SKGs

In this section we examine conditions for the emergence of the giant component in a stochastic Kronecker graph with a 2×2 affinity matrix and $n = 2^t$ vertices. We note that Mahdian and Xu proved a necessary and sufficient condition for the emergence of the giant component in $K(t, \alpha, \beta, \gamma)$ in the case that $\alpha > \beta > \gamma$ in [55]. Here, we generalize that result to any $K(t, \alpha, \beta, \gamma)$ via a substantially different proof. Moreover, we establish the asymptotic size of the giant component in the SKG with 2×2 affinity matrix.

4.2.1 Emergence of the Giant Component

We will first prove the following theorem.

Theorem 4.2. *Let $G = K(t, \alpha, \beta, \gamma)$. A necessary and sufficient condition for G to have a giant component of size $\Theta(n)$ a.a.s. is that $(\alpha + \beta)(\beta + \gamma) > 1$.*

We shall assume throughout, without loss of generality, that $\alpha > \gamma$. As we have seen in Remark 3.4, G does not necessarily have good expansion properties. The approach to establishing Theorem 4.2 will be to find a section of G with good expansion, and show that this is contained in a giant component. In so doing, we gain structural information about the giant component itself, which will allow us to establish the asymptotic size of the giant component in Theorem 4.9, as well as establish the uniqueness of such a component.

Given $s \in [t]$, let V_s denote the set of vertices with weight s and G_s the subgraph of G induced on V_s. For a vertex v with weight k, we say a neighbor u of v is of type (a, b) if u shares $k - a$ 1s with v and $(t - k) - b$ 0s with v. That is to say, we may think of u being obtained from v by changing a 1s to 0s and b 0s to 1s.

Let

$$k = \frac{\alpha + \beta}{\alpha + \gamma + 2\beta} t, \qquad \ell = \frac{\beta}{\alpha + 2\beta + \gamma} t = \frac{\beta}{\alpha + \beta} k, \tag{4.3}$$

and let H denote the subgraph of G with $V(H) = V_k$ and $E(H) = \{uv \in G : u \cdot v = k - \ell\}$. That is, H is the subgraph of vertices with weight k and edges of type (ℓ, ℓ).

Notice that if $\omega(v) = k$, then the expected number of neighbors of v of type (a, b) is

$$\binom{k}{a}\binom{t-k}{b}\beta^a\alpha^{k-a}\beta^b\gamma^{t-k-b}.$$

To select the above parameters k and ℓ, we first find a and b that maximize this expression in terms of k, and then find k so that $a = b$, resulting in the above parameters.

Lemma 4.3. *For a vertex $v \in V_k$, we have $\mathbb{E}\left[\deg_H(v)\right]$ is*

$$\binom{k}{\ell}\binom{t-k}{\ell}\beta^{2\ell}\alpha^{k-\ell}\gamma^{t-k-\ell} > (1+o(1))c^t.$$

for some $c > 1$.

Proof. Now, $\frac{\ell}{k} = \frac{\beta}{\alpha+\beta}$ and $\frac{\ell}{t-k} = \frac{\beta}{\beta+\gamma}$, so the entropy bound (4.1) gives us

$$\begin{aligned}
&\binom{k}{\ell}\binom{t-k}{\ell}\beta^{\ell}\alpha^{k-\ell}\beta^{\ell}\gamma^{t-k-\ell} \\
&> \quad \frac{(\alpha+\beta)(\beta+\gamma)\left(\frac{\alpha+\beta}{\beta}\right)^{\ell}\left(\frac{\alpha+\beta}{\alpha}\right)^{k-\ell}\left(\frac{\beta+\gamma}{\beta}\right)^{\ell}\left(\frac{\beta+\gamma}{\gamma}\right)^{t-k-\ell}}{2e^2\pi k\beta\sqrt{\alpha\gamma}}\beta^{2\ell}\alpha^{k-\ell}\gamma^{t-k-\ell} \\
&= \quad \frac{(\alpha+\beta)(\beta+\gamma)}{2e^2\pi k\beta\sqrt{\alpha\gamma}}(\alpha+\beta)^k(\beta+\gamma)^{t-k} \\
&= \quad \frac{1}{2e^2\pi k\beta\sqrt{\alpha\gamma}}((\alpha+\beta)(\beta+\gamma))^{t-k+1}(\alpha+\beta)^{2k-t}. \qquad (4.4)
\end{aligned}$$

Notice that $\alpha > \gamma$ and $(\alpha+\beta)(\beta+\gamma) > 1$ implies that $\alpha + \beta > 1$ and also that $k > \frac{t}{2}$. Thus (4.4) is exponential in t as desired. Note that, while for our purposes it will not be required, the c we obtain is

$$c = ((\alpha+\beta)(\beta+\alpha))^{\frac{\alpha+\beta}{\alpha+\gamma+2\beta}}(\alpha+\beta)^{\frac{\alpha-\gamma}{\alpha+\gamma+2\beta}} > 1.$$

□

Let $G(t, k, \ell)$ denote the graph with vertex set $V(G(t, k, \ell)) = \{S \subset [t] : |S| = k\}$, where two vertices S and T are adjacent if $|S \cap T| = k - \ell$. Note that our graph H is a percolated version of this graph where each edge is taken independently with probability $\beta^{2\ell}\alpha^{k-\ell}\gamma^{t-k-\ell}$. In order to show that H is a.a.s. connected, we will first derive some information on $G(t, k, \ell)$, and then take advantage of the fact that the expected degree of each vertex in H is exponential in t, as shown above in Lemma 4.3.

Lemma 4.4. *Suppose $t \geq k + \ell$ and $\ell < k$. For any two distinct vertices U and V in $G(t, k, \ell)$, let $s = |U \cap V|$. Then*

$$\operatorname{dist}(U, V) \leq \begin{cases} 1 & s = k - \ell \\ 2 & s > k - \ell \text{ and } s \geq 2\ell + k - t \\ 2\left\lceil \frac{2\ell + k - t - s}{t - 2\ell} \right\rceil + 2 & s > k - \ell \text{ and } s < 2\ell + k - t \\ \left\lceil \frac{k - \ell - s}{\ell} \right\rceil + 2 & s < k - \ell \text{ and } s > 2\ell + k - t \\ \left\lceil \frac{k - \ell - s}{\ell} \right\rceil + 2\left\lceil \frac{3\ell - t}{t - 2\ell} \right\rceil + 2 & s < k - \ell \text{ and } s < 2\ell + k - t \end{cases}$$

Proof. Let $\zeta = t - k - \ell$, so $2\ell + k - t = \ell - \zeta$. For ease of notation we let $M = U \triangle V$ and $Z = [t] - (U \cup V)$. Note that $|M| = |U \cap M| + |V \cap M| = 2(k - s)$, and $|Z| = t - s - 2(k - s) = t - 2k + s = s - k + \ell + \zeta$. Thus we have that $s \geq k - \ell - \zeta$. Now if $s = k - \ell$, then $U \sim V$ and $\operatorname{dist}(U, V) = 1$, hence we may assume that $k - \ell - \zeta \leq s < k - \ell$ or $k - \ell < s < k$. We consider several cases.

Case 1: $s > k - \ell$ and $s \geq \ell$.

As $s > \ell$, there is a set $A \subseteq U \cap V$ where $|A| = s - \ell$, and as $|Z| = s - k + \ell + \zeta$, there is a set $B \subseteq Z$ where $|B| = s - k + \ell$. Define $W = A \cup M \cup B$. Now $|W| = |A| + |M| + |B| = (s - \ell) + (2k - 2s) + (s - k + \ell) = k$ so W is a vertex in $G(t, k, \ell)$. Furthermore, $|U \cap W| = |A| + |U \cap M| = (s - \ell) + (k - s) = k - \ell$ and thus $U \sim W$. Similarly, $V \sim W$. Hence, if $s > k - \ell$ and $s \geq \ell$, then $\operatorname{dist}(U, V) = 2$.

Case 2: $s > k - \ell$ and $\ell > s \geq \ell - \zeta$.

Then $|M \cap U| = |M \cap V| = k - s > k - \ell$, and thus there are sets $A \subseteq M \cap U$ and $B \subseteq M \cap V$ such that $|A| = k - \ell = |B|$. Now, as

$$|Z| = s - k + \ell + \zeta \geq (\ell - \zeta - k) + \ell + \zeta = 2\ell - k$$

and $2\ell - k = \ell + (\ell - k) > \ell - s > 0$, there is a set $C \subseteq Z$ such that $|C| = 2\ell - k$. Let $W = A \cup B \cup C$. Then $|W| = 2(k - \ell) + 2\ell - k = k$, and $|W \cap U| = |W \cap V| = k - \ell$. Thus W is a vertex in $G(t, k, \ell)$ and $U \sim W \sim V$ and $\operatorname{dist}(U, V) = 2$.

At this point, in order to construct a path between U and V, it suffices to exhibit a vertex W such that $\mathrm{dist}(U,W) < \infty$ and $|W \cap V| > |U \cap V|$, as then repeated applications will terminate in some vertex W' where $|W' \cap V| > k-\ell$ and $|W' \cap V| \geq \ell - \zeta$.

Case 3: $s < k - \ell$ and $s \geq \ell - \zeta$.

As $k - \ell - s > 0$, there exists a set $B \subset M \cap V$ with $|B| = \ell$ and a set $A \subset M \cap U$ with $|A| = k - \ell - s$. Let $W = (U \cap V) \cup A \cup B$, so $|W| = s + k - \ell + s + \ell = k$, and $U \sim W$. Moreover, $|W \cap B| = s + \ell$. Thus, after $m = \lceil \frac{k-\ell-s}{\ell} \rceil$ applications, we obtain a vetex W' with $|W' \cap B| = s + m\ell \geq s + k - \ell - s = k - \ell$, and we may apply Case 2 to obtain

$$\mathrm{dist}(U,V) \leq \left\lceil \frac{k-\ell-s}{\ell} \right\rceil + 2.$$

Case 4: $s > k - \ell$ and $s < \ell - \zeta$. As $\ell - s - \zeta > 0$ we may partition $U \cap M$ into disjoint sets M_U^ℓ, M_U^ζ, M_U, where $|M_U^\ell| = k - \ell$, $|M_U^\zeta| = \zeta$ and $|M_U| = \ell - s - \zeta$. Define M_V^ℓ, M_V^ζ and M_V analogously. Let $W_1 = M_U^\ell \cup M_V^\ell \cup M_V \cup Z$ and define $W_2 = M_U \cup (U \cap V) \cup M_V^\ell \cup M_V^\zeta$. Then we have

$$\begin{aligned} |W_1| &= 2(k-\ell) + (\ell - s - \zeta) + (s - k + \ell + \zeta) = k \\ |W_2| &= (\ell - s - \zeta) + s + (k - \ell) + \zeta = k \\ |U \cap W_1| &= |M_U^\ell| = k - \ell \\ |W_1 \cap W_2| &= |M_V^\ell| = k - \ell \\ |W_2 \cap V| &= s + k - \ell + \zeta > s \end{aligned}$$

Therefore, $U \sim W_1 \sim W_2$, and $|W_2 \cap V| = s + (k-\ell) + \zeta > s$. Thus, after $\lceil \frac{2\ell+k-t-s}{t-2\ell} \rceil$ steps, we have vertices W_1' and W_2' such that $|W_2' \cap V| = s + m(k - \ell + \zeta) = s + m(t - 2\ell) \geq s + 2\ell + k - t - s = \ell - \zeta$. Thus we may apply Case 2 or Case 3 to obtain

$$\mathrm{dist}(U,V) \leq 2 \left\lceil \frac{2\ell + k - t - s}{t - 2\ell} \right\rceil + 2.$$

Case 5: $s < k - \ell$ and $s < \ell - \zeta$.

As $|M \cap U| = |M \cap V| = k - s$, and $s < k - \ell$, there exists $A \subseteq M \cap U$ with $|A| = k - \ell - s$. As $k - s > k - (k - \ell) = \ell$, there exists $B \subseteq M \cap V$ with $|B| = \ell$. Let $W = (U \cap V) \cup A \cup B$. Then we have $|W| = s + (k - \ell - s) + \ell = k$ and $|W \cap U| = |U \cap V| + |A| = s + (k - \ell - s) = k - \ell$. Thus W is a vertex in $G(t,k,\ell)$

and $U \sim W$. Furthermore, $|V \cap W| = |U \cap V| + |B| = s + \ell$. Thus after at most $m = \lceil \frac{k-\ell-s}{\ell} \rceil$ steps, we have a vertex W' with $|W' \cap V| = s + m\ell \geq s + k - \ell - s = k - \ell$. We may then apply Case 4 to obtain

$$\begin{aligned} \operatorname{dist}(U,V) &\leq \left\lceil \frac{k-\ell-s}{\ell} \right\rceil + 2 \left\lceil \frac{2\ell + k - t - (s + \ell \lceil \frac{k-\ell-s}{\ell} \rceil)}{t-2\ell} \right\rceil + 2 \\ &\leq \left\lceil \frac{k-\ell-s}{\ell} \right\rceil + 2 \left\lceil \frac{2\ell + k - t - (k-\ell)}{t-2\ell} \right\rceil + 2 \\ &\leq \left\lceil \frac{k-\ell-s}{\ell} \right\rceil + 2 \left\lceil \frac{3\ell - t}{t-2\ell} \right\rceil + 2. \end{aligned}$$

□

In the case of k and ℓ as defined in equation (4.3), we have

$$k + \ell = \frac{\alpha+\beta}{\alpha+\gamma+2\beta} t + \frac{\beta}{\alpha+2\beta+\gamma} t = \frac{\alpha+2\beta}{\alpha+2\beta+\gamma} t \leq t$$

and

$$\ell = \frac{\beta}{\alpha+\beta} k < k.$$

Therefore, applying Lemma 4.4, we have

$$\begin{aligned} \operatorname{diam}(G(t,k,\ell)) &\leq \max \Bigg\{ 2, 2 \left\lceil \frac{2\ell + k - t - (k-\ell)}{t-2\ell} \right\rceil + 2, \\ &\qquad \left\lceil \frac{k - \ell - (2\ell + k - t)}{\ell} \right\rceil + 2, \left\lceil \frac{k-\ell}{\ell} \right\rceil + 2 \left\lceil \frac{3\ell - t}{t - 2\ell} \right\rceil + 2 \Bigg\} \\ &\leq \max \left\{ 2, \frac{2\beta}{\alpha+\gamma} + 2, \frac{\alpha+\gamma}{\beta} + 2, \frac{\alpha}{\beta} + \frac{2\beta}{\alpha+\gamma} + 3 \right\} = \Theta(1) \end{aligned}$$

As any permutation of $[t]$ is an automorphism of $G(t,k,\ell)$ and we can easily construct a permutation that maps any edge to any other, we have that $G(t,k,\ell)$ is edge transitive. Thus, we may apply Theorem 1.10 together with the fact that $G(t,k,\ell)$ has constant diameter to obtain the following.

Lemma 4.5. *For $K = G(t,k,\ell)$ as above, $h_K \geq C$ for some constant C. In particular, for any set $S \subseteq V(K)$ with $|S| = s \leq |V(K)|/2$, we have*

$$e(S,\bar{S}) \geq C \operatorname{vol}(S) = Cs\binom{k}{\ell}\binom{t-k}{\ell}.$$

As observed above, the graph H is a percolated version of $G(t,k,\ell)$ where each edge is chosen independently with probability $\beta^{2\ell}\alpha^{k-\ell}\gamma^{t-k-\ell}$. We are now prepared to prove that H is connected with high probability.

Theorem 4.6. *Let H be as described above. Then H is connected a.a.s..*

Proof. Let $S \subseteq V(H)$, with $|S| = s \leq |V(H)|/2$. By Lemma 4.5, in $G(t,k,\ell)$, we have $e(S,\bar{S}) \geq Cs\binom{k}{\ell}\binom{t-k}{\ell}$. Thus, in H, we have at least $Cs\binom{k}{\ell}\binom{t-k}{\ell}$ potential edges from S to $\bar{S}$. Therefore,

$$\begin{aligned}\mathbb{E}\left[e(S,\bar{S})\right] &\geq Cs\binom{k}{\ell}\binom{t-k}{\ell}\beta^{2\ell}\alpha^{k-\ell}\gamma^{t-k-\ell}\\ &\geq s\Theta(c^t)\end{aligned}$$

by Lemma 4.3. Moreover, $e(S,\bar{S})$ is binomially distributed, so by Chernoff bounds (see Theorem 1.15), we have

$$\begin{aligned}\mathbb{P}\left(e(S,\bar{S}) \leq \frac{1}{2}\mathbb{E}\left[e(S,\bar{S})\right]\right) &= \mathbb{P}\left(e(S,\bar{S}) - \mathbb{E}\left[e(S,\bar{S})\right] \leq -\frac{1}{2}\mathbb{E}\left[e(S,\bar{S})\right]\right)\\ &\leq e^{-\frac{1}{8}\mathbb{E}\left[e(S,\bar{S})\right]}\\ &\leq e^{-s\Theta(c^t)}.\end{aligned}$$

Let $\mathcal{A}$ denote the event that there exists a set $S \subseteq V(H)$ with $e(S,\bar{S}) = 0$. Then $\mathcal{A}$ is precisely the event that H is disconnected. By the union bound

$$\begin{aligned}\mathbb{P}(\mathcal{A}) &\leq \sum_{s=1}^{|V(H)|/2}\sum_{|S|=s}\mathbb{P}(e(S,\bar{S})=0)\\ &\leq \sum_{s=1}^{|V(H)|/2}\binom{\binom{t}{k}}{s}\exp(-s\cdot\Theta(c^t))\\ &\leq \sum_{s=1}^{|V(H)|/2}\left(\binom{t}{k}\exp(-\Theta(c^t))\right)^s\\ &\leq \sum_{s=1}^{|V(H)|/2}(o(1))^s = o(1).\end{aligned}$$

Note that we used the fact that $\binom{t}{k}\exp(-\Theta(c^t)) = o(1)$. This is easily verified by taking logs: $\log\binom{t}{k} = o(t\log t)$ while $\log(\exp(\Theta(c^t))) = \Theta(c^t)$. Therefore, $\mathbb{P}(\mathcal{A}) = o(1)$ and thus H is connected a.a.s.. □

We are now ready to complete the proof of Theorem 4.2, and to establish the upper bound in Theorem 4.9.

We first establish the necessity of the condition in Theorem 4.2 with the following theorem. We note that a similar technique is used in [55], although we include a proof here for the sake of completeness.

Theorem 4.7. *Suppose $G = K(t,\alpha,\beta,\gamma)$, and $(\alpha+\beta)(\beta+\gamma) < 1$. Then G contains $n - o(n)$ isolated vertices a.a.s..*

Proof. Let $\epsilon = 1-(\alpha+\beta)(\beta+\gamma) > 0$. Let $v \in V$ have weight $\omega(v) = w \le t/2+t^{2/3}$. Then the expected degree of v is

$$\begin{aligned}
(\alpha+\beta)^w(\beta+\gamma)^{t-w} &\le (\alpha+\beta)^{t/2+t^{2/3}}(\beta+\gamma)^{t/2-t^{2/3}} \\
&\le ((\alpha+\beta)(\beta+\gamma))^{t/2}\left(\frac{\alpha+\beta}{\beta+\gamma}\right)^{t^{2/3}} \\
&\le (1-\epsilon)^{\frac{1}{2}\log n}\left(\frac{\alpha+\beta}{\beta+\gamma}\right)^{(\log n)^{2/3}} \quad = o(1)
\end{aligned}$$

Thus, v is isolated a.a.s.. Moreover, the proportion of vertices with weight at most $t/2+t^{2/3}$ is at least $1-e^{-(t^{2/3})^2/2t} = 1-o(1)$ by the Chernoff bound (see Theorem 1.13). Therefore, there are at least $n-o(n)$ isolated vertices in G a.a.s., and thus G has no giant component a.a.s.. □

We now turn our attention to establishing the sufficiency of the condition in Theorem 4.2. For the remainder of this section, we assume $K(t,\alpha,\beta,\gamma)$ has $(\alpha+\beta)(\beta+\gamma) > 1$, and $\alpha \ge \gamma$. Also, we set k and ℓ as in equation (4.3).

Theorem 4.8. *Suppose $s \ne k$, with*

$$s \ge \frac{2\log(t) - t\log(\beta+\gamma) - \log\left(\frac{(\alpha+\beta)(\beta+\gamma)}{2e^2\pi\beta t\sqrt{\alpha\gamma}}\right)}{\log(\frac{\alpha+\beta}{\beta+\gamma})}. \tag{4.5}$$

Then for every vertex $v \in V_s$, a.a.s. there exists r with $|r-k| < |s-k|$ such that there is a vertex $u \in V_r$ with $v \sim u$.

The precise bound on s in the statement of Theorem 4.8 is quite technical, and falls out from the proof. Note that

$$s \geq \frac{-\log(\beta+\gamma)}{-\log(\beta+\gamma)+\log(\alpha+\beta)}t + \Theta(\log t).$$

As $(\alpha+\beta)(\beta+\gamma) > 1$, we have that

$$m = \frac{-\log(\beta+\gamma)}{-\log(\beta+\gamma)+\log(\alpha+\beta)} < \frac{1}{2}. \tag{4.6}$$

In particular, Theorem 4.8 holds for all vertices with weight at least $\frac{t}{2}$. Thus, a.a.s. every vertex of weight at least $\frac{t}{2}$ can be connected by a path to a vertex of weight k. Moreover, Lemma 4.6 shows that a.a.s. G_k is connected, and thus we have a giant component of size at least $\frac{n}{2}$. Thus, Theorem 4.8 is sufficient to complete the proof of Theorem 4.2.

If $\beta+\gamma > 1$, then note that all non-negative s satisfy equation (4.5). That is, the graph is connected, as shown in [55] and mentioned above in Theorem 1.21.

Proof of Theorem 4.8. Suppose $v \in V_s$. The expected number of neighbors of v of type (a,b) is

$$\binom{s}{a}\binom{t-s}{b}\beta^a\alpha^{s-a}\beta^b\gamma^{t-s-b}.$$

Let $a = \frac{\beta}{\alpha+\beta}s$ and $b = \frac{\beta}{\beta+\gamma}(t-s)$. Then the weight of a neighbor of v of type (a,b) is

$$f(s) = \frac{\alpha}{\alpha+\beta}s + \frac{\beta}{\beta+\gamma}(t-s).$$

Note that as f is linear, we obtain

$$\begin{aligned} k \;>\; f(s) \;&>\; s \text{ when } s < k.\\ k \;<\; f(s) \;&<\; s \text{ when } s > k.\\ f(s) \;&=\; s \text{ when } s = k. \end{aligned}$$

Therefore, a neighbor of v of type (a,b) has weight r, with $|r-k| < |s-k|$. Using the entropy bound (equation (4.1)), the expected number of neighbors of v of type

(a,b) is

$$
\begin{aligned}
&\binom{s}{a}\binom{t-s}{b}\beta^a\alpha^{s-a}\beta^b\gamma^{t-s-b}\\
&> \frac{(\alpha+\beta)(\gamma+\beta)\left(\frac{\alpha+\beta}{\beta}\right)^a\left(\frac{\alpha+\beta}{\alpha}\right)^{s-a}\left(\frac{\beta+\gamma}{\beta}\right)^b\left(\frac{\beta+\gamma}{\gamma}\right)^{t-s-b}}{2e^2\pi\beta\sqrt{s(t-s)\alpha\gamma}}\beta^a\alpha^{s-a}\beta^b\gamma^{t-s-b}\\
&= \frac{(\alpha+\beta)(\gamma+\beta)}{2e^2\pi\beta\sqrt{s(t-s)\alpha\gamma}}(\alpha+\beta)^s(\gamma+\beta)^{t-s} \geq t^2. \qquad (4.7)
\end{aligned}
$$

The lower bound on s in the statement of the theorem is chosen precisely so that inequality (4.7) holds.

The number of neighbors of v with weight r is binomially distributed, so the Chernoff bounds (see Theorem 1.13) imply that the probability that v has no neighbors with weight r is bounded as

$$
\begin{aligned}
\mathbb{P}(e(v,V_r)=0) &\leq \frac{1}{2}\mathbb{E}\left[e(v,V_r)\right]\\
&\leq \exp\left(-\frac{1}{8}\mathbb{E}\left[e(v,V_r)\right]\right)\\
&\leq \exp\left(-\frac{t^2}{8}\right).
\end{aligned}
$$

As there are 2^t vertices in $K(t,\alpha,\beta,\gamma)$, by the union bound

$$\mathbb{P}(\exists v : \omega(v) > s \text{ and } e(v,V_r)=0) \leq 2^t \exp\left(-\frac{t^2}{8}\right) = o(1).$$

Therefore, a.a.s. no such v exists, and the Theorem is proven. □

4.2.2 Size of the Giant Component

In this section we investigate the asymptotic size of the giant component in $K(t,\alpha,\beta,\gamma)$. By definition, the giant component has size at least Cn for some constant C; in Theorem 4.2, we establish that the giant component is of size at least $n-O\left(\binom{t}{mt+\log t}\right)$, where m is as defined below in Theorem 4.9. Here we show that in fact that the number of vertices that are not in the giant component is indeed $\Theta\left(\binom{t}{mt}\right)$, so the bound determined above on the size of the giant component is in fact accurate. Note moreover that this result will also imply the uniqueness of the giant component, as we will show that in fact most vertices that are not in the giant component are isolated.

Theorem 4.9. *Suppose $\alpha, \beta, \gamma \in (0,1)$ with $(\alpha+\beta)(\beta+\gamma) > 1$ and $\alpha \geq \gamma$. By Theorem 4.2, $K(t,\alpha,\beta,\gamma)$ has a giant component a.a.s.. Suppose, moreover, that $\beta+\gamma < 1$. Let X denote the set of vertices of $K(t,\alpha,\beta,\gamma)$ that are not in the giant component. Then a.a.s.*

$$|X| = \Theta\left(\binom{t}{mt}\right)$$

where

$$m = \frac{-\log(\beta+\gamma)}{-\log(\beta+\gamma)+\log(\alpha+\beta)} \tag{4.8}$$

We note that in Theorem 4.9, the assumption that $\beta+\gamma < 1$ is not very restrictive; if $\beta+\gamma \geq 1$ then the graph is connected a.a.s. (see Theorem 1.21). Furthermore, if $\beta+\gamma \geq 1$ then the constant m is negative, so one could omit the restriction by replacing m with the maximum of the stated value or 0. In order to derive the precise result, we must examine the vertices with weight $mn+o(n)$ more closely.

Again, we consider the set V_s of vertices with weight s. For a vertex $v \in V_s$, its expected degree is

$$\begin{aligned}\mathbb{E}[\deg(v)] &= \sum_{a=0}^{s}\sum_{b=0}^{t-s}\binom{s}{a}\binom{t-s}{b}\beta^a\alpha^{s-a}\beta^b\gamma^{t-s-b} - \alpha^s\gamma^{t-s} &(4.9)\\ &= (\alpha+\beta)^s(\beta+\gamma)^{t-s} - \alpha^s\gamma^{t-s}. &(4.10)\end{aligned}$$

The $\alpha^s\gamma^{t-s}$ term corresponds to a 'self loop' at a vertex. For the purposes of this proof, we wish to consider edges that leave the vertex, so we will omit this term.

In the proof of Theorem 4.2, we chose $a = \frac{\beta}{\alpha+\beta}s$ and $b = \frac{\beta}{\beta+\gamma}s$ to approximately maximize Equation (4.10). In order to establish Theorem 4.9 we need a more precise understanding of the summation.

Lemma 4.10. *Let $\epsilon > 0$ be such that $(1+\epsilon)\frac{\beta}{\alpha+\beta} < 1$. Moreover, suppose that*

$s = s(t)$, and there exists r with $0 < r < 1$, and $\frac{s}{t} \to r$. Then

$$\sum_{a=(1+\epsilon)\frac{\beta}{\alpha+\beta}s}^{s} \sum_{b=0}^{(1-\epsilon)\frac{\beta}{\beta+\gamma}(t-s)} \binom{s}{a}\binom{t-s}{b}\beta^a\alpha^{s-a}\beta^b\gamma^{t-s-b} = o((\alpha+\beta)^s(\beta+\gamma)^{t-s}).$$

Proof. Let

$$g(a,b) = \binom{s}{a}\binom{t-s}{b}\beta^a\alpha^{s-a}\beta^b\gamma^{t-s-b}.$$

Then

$$\frac{g(a+1,b)}{g(a,b)} = \frac{s-a}{a+1}\frac{\beta}{\alpha}.$$

For $a \geq (1+\epsilon)\frac{\beta}{\alpha+\beta}s$,

$$\begin{aligned}
\frac{g(a+1,b)}{g(a,b)} &= \frac{s-(1+\epsilon)\frac{\beta}{\alpha+\beta}s}{(1+\epsilon)\frac{\beta}{\alpha+\beta}s+1}\cdot\frac{\beta}{\alpha} \\
&\leq \frac{\alpha+\beta}{\alpha(1+\epsilon)} - \frac{\beta}{\alpha} \\
&= \frac{\alpha-\epsilon\beta}{\alpha(1+\epsilon)} \\
&\leq \frac{1}{1+\epsilon}.
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{a=(1+\epsilon)\frac{\beta}{\alpha+\beta}s}^{s} g(a,b) &\leq \sum_{a=(1+\epsilon)\frac{\beta}{\alpha+\beta}s}^{s} (1+\epsilon)^{a-(1+\epsilon)\frac{\beta}{\alpha+\beta}s} g\left((1+\epsilon)\frac{\beta}{\alpha+\beta}s, b\right) \\
&\leq Cg\left((1+\epsilon)\frac{\beta}{\alpha+\beta}s, b\right),
\end{aligned}$$

where C is obtained by summing the geometric series.

A similar bound on $\frac{g(a,b-1)}{g(a,b)}$ allows us to derive that

$$\sum_{l=(1+\epsilon)\frac{\beta}{\alpha+\beta}s}^{s} \sum_{b=0}^{(1-\epsilon)\frac{\beta}{\beta+\gamma}s} g(a,b) \leq C'g\left((1+\epsilon)\frac{\beta}{\alpha+\beta}s, (1-\epsilon)\frac{\beta}{\beta+\gamma}s\right). \tag{4.11}$$

Note that

$$\sum_{(1+\epsilon)\frac{\beta}{\alpha+\beta}s-\log(s)}^{(1+\epsilon)\frac{\beta}{\beta+\alpha}s} \sum_{b=(1-\epsilon)\frac{\beta}{\beta+\gamma}(t-s)}^{(1-\epsilon)\frac{\beta}{\beta+\gamma}(t-s)+\log s} g(a,b) = \omega\left(g\left((1+\epsilon)\frac{\beta}{\alpha+\beta}s, (1-\epsilon)\frac{\beta}{\beta+\gamma}\right)\right).$$

as the sums are bounded below by a geometric series with ratio greater than one. Together with equation (4.11), this completes the proof. □

We next use the following lemma in order to establish the lower bound in Theorem 4.9.

Lemma 4.11. *Let X and m be as in Theorem 4.9. Then a.a.s.*

$$|X| = \Omega\left(\binom{t}{mt}\right).$$

Proof. Let $l = mt - 1$. Then the expected degree of a vertex v with weight l is

$$\begin{aligned}
\mathbb{E}[\deg(v)] &= (\alpha+\beta)^{mt-1}(\beta+\gamma)^{t-mt+1} \\
&= \frac{\beta+\gamma}{\alpha+\beta}\left((\alpha+\beta)^m(\beta+\gamma)^{1-m}\right)^t \\
&= \frac{\beta+\gamma}{\alpha+\beta} =: q < 1
\end{aligned}$$

by the definition of m. Take $p = \mathbb{P}(\deg(v) = 0) \geq 1 - q > 0$.

Let Y denote the set of isolated vertices with weight l. Then $\mathbb{E}[|Y|] = p\binom{t}{l}$. We write $|Y| = \sum_{v \in V_l} z_v$ where z_v is the indicator that v is isolated. The z_v are not independent, however we have

$$\begin{aligned}
\mathbb{E}[z_v z_u] &= \prod_{x \in V}(1 - \mathbb{P}(x \sim v)) \prod_{\substack{y \in V \\ y \neq v}}(1 - \mathbb{P}(y \sim u)) \\
&= \frac{\mathbb{E}[z_v]\,\mathbb{E}[z_u]}{1 - \mathbb{P}(u \sim v)} \\
&= \mathbb{E}[z_v]\,\mathbb{E}[z_u] + \frac{\mathbb{P}(u \sim v)}{1 - \mathbb{P}(u \sim v)}\mathbb{E}[z_v]\,\mathbb{E}[z_u]
\end{aligned}$$

Thus, as $\mathbb{E}[z_v] = \mathbb{E}[z_u] = p$ and $\mathbb{P}(u \sim v) \to 0$, we obtain

$$\begin{aligned}
\mathrm{Cov}(z_v, z_u) &= \mathbb{E}[z_v z_u] - \mathbb{E}[z_v]\,\mathbb{E}[z_u] \\
&= \frac{\mathbb{P}(u \sim v)}{1 - \mathbb{P}(u \sim v)}\mathbb{E}[z_v]\,\mathbb{E}[z_u] \\
&= o(p^2).
\end{aligned}$$

Moreover,

$$\mathrm{Var}(z_v) = p(1-p).$$

Thus we have

$$\begin{aligned}\operatorname{Var}\left(|Y|^2\right) &= \sum_v \operatorname{Var}(z_v) + \sum_{v\in V_l}\sum_{\substack{u\in V_l\\ u\neq v}} \operatorname{Cov}(z_v, z_u)\\ &\le p(1-p)\binom{t}{l} + o\left(p^2\binom{t}{l}^2\right) = o\left(\binom{t}{l}^2\right).\end{aligned}$$

Since $\operatorname{Var}(|Y|^2) = o(\binom{t}{l}^2) = o(\mathbb{E}[|Y|]^2)$, Chebyshev's inequality (see Theorem 1.12) implies that for any $c > 0$,

$$\mathbb{P}\left(\left||Y| - p\binom{t}{l}\right| \ge c\binom{t}{l}\right) \le \frac{o\left(\binom{t}{l}^2\right)}{c^2\binom{t}{l}^2} = o(1).$$

Therefore, a.a.s. we have that $|Y| = p\binom{t}{l} + o\left(\binom{t}{l}\right)$, and thus $|X| \ge c\binom{t}{l} \ge c'\binom{t}{mt}$ a.a.s., establishing the result. □

Using Lemma 4.11, we can derive from Theorem 4.8 that

$$\binom{t}{mt} \ll |X| \ll \binom{t}{mt + C\log(t)}$$

for some absolute constant C, but these differ by a factor polynomial in t (and hence by a poly-logarithmic factor in n.) Here, the $\ll$ symbol is in the traditional number theoretic sense, that is, $f(x) \ll g(x)$ if $f(x) = O(g(x))$.

Proof of Theorem 4.9. By Lemma 4.11, it suffices to show that $|X| = O(\binom{t}{mt})$ a.a.s. Let f be as in the proof of Theorem 4.8.

Suppose $s = mt + O(\log(t))$. Choose $\epsilon > 0$ and small enough that

$$\frac{\alpha - \epsilon\beta}{\alpha + \beta}s + \frac{\beta - \epsilon\gamma}{\alpha + \beta}(t - s) > (m + \epsilon)t;$$

such exists by our observation on $f(s)$.

Note that, by Theorem 4.8, for t sufficiently large, if $s' \ge (m + \epsilon)t$, then all vertices in $V_{s'}$ are in the giant component a.a.s.. Thus, a vertex in V_s which is not in the giant component has no edges into $V_{s'}$ for $s' \ge (m + \epsilon)t$.

Consider a vertex v with weight $\omega(v) = mt + O(\log(t))$. We say that an edge from v is *good* if it of type (a, b), where $a \le (1 + \epsilon)\frac{\beta}{\alpha+\beta}s$ and $b \ge (1 - \epsilon)\frac{\beta}{\beta+\gamma}(t - s)$.

(Note that when we say an edge vu incident to v is good, we are assuming the edge type is considered at the vertex v. In this way, an edge may be good when considered from v but not from u). Let Y be the set of vertices with weight $mt + O(\log(t))$ that have no incident good edges. Note that if v has an incident good edge, then it is connected to a vertex $u \in V_{s'}$ with $s' \geq (m+\epsilon)t$, so every vertex not in Y is in the giant component a.a.s., hence $|Y| \geq |X|$.

Let Z_v denote the number of good edges incident to v. By Lemma 4.10

$$\mathbb{E}\left[Z_v\right] = (1+o(1))(\alpha+\beta)^s(\gamma+\beta)^{t-s}.$$

For each v, put W_v to be the set of vertices u for which the edge vu is good (if it exists). Since each edge occurs independently, Z_v is the sum of independent indicator functions, so we can write

$$\begin{aligned} \mathbb{P}(Z_v = 0) &= \prod_{u \in W_v} (1 - \mathbb{P}(v \sim u)) \\ &\leq \exp\left(-\sum_{u \in W_v} \mathbb{P}(v \sim u)\right) = \exp(-\mathbb{E}\left[Z_v\right]). \end{aligned}$$

Thus, for t sufficiently large,

$$\mathbb{P}(v \in Y) = \mathbb{P}(Z_v = 0) \leq \exp\left(-\frac{1}{2}(\alpha+\beta)^s(\gamma+\beta)^{t-s}\right).$$

By Theorem 4.8, there exists a C such that if $s > mt + C\log(t)$, then all vertices in V_s are in the giant component a.a.s.. Thus

$$|X| \leq \sum_{s=0}^{mt+C\log(t)} |Y \cap V_s|.$$

Choose r to be the least integer such that

$$\exp\left(-\frac{1}{2}\left(\frac{\alpha-\gamma}{\beta+\gamma}\right)\left(\frac{\alpha+\beta}{\beta+\gamma}\right)^r\right)\frac{1-m}{m} < \frac{1}{2}.$$

Define

$$g(k) = \exp\left(-\frac{1}{2}(\alpha+\beta)^{mt+r+k}(\gamma+\beta)^{t-mt-r-k}\right)\binom{t}{mt+r+k}.$$

We have chosen r precisely so that for $k \geq 0$,

$$\frac{g(k+1)}{g(k)} < \frac{1}{2}.$$

This implies that $g(k) < \frac{1}{2}g(k-1) < \frac{1}{4}g(k-2) < \cdots < 2^{-k}g(0)$. We therefore obtain

$$\begin{aligned}
\mathbb{E}\left[\sum_{s=mt+r}^{mt+C\log(t)} |Y \cap V_s|\right] &= \sum_{s=mt+r}^{mt+C\log(t)} \mathbb{E}\left[|Y \cap V_s|\right] \\
&\leq \sum_{s=mt+r}^{mt+C\log(t)} \binom{t}{s} \exp\left(-\frac{1}{2}(\alpha+\beta)^s(\gamma+\beta)^{t-s}\right) \\
&= \sum_{k=0}^{C\log(t)-r} g(k) \\
&\leq \sum_{k=0}^{C\log(t)-r} 2^{-k}g(0) \\
&\leq 2g(0).
\end{aligned}$$

As $\sum_{s=mt+r}^{mt+C\log(t)} |Y \cap V_s|$ can be written as the sum of independent indicator functions, it is tightly concentrated by the Chernoff bounds and hence a.a.s.

$$\sum_{s=mt+r}^{mt+C\log(t)} |Y \cap V_s| \leq (1+o(1))2g(0) = \Theta\left(\binom{t}{mt+r}\right) = O\left(\binom{t}{mt}\right).$$

Note that

$$\sum_{s=0}^{mt+r} |Y \cap V_s| \leq \sum_{s=0}^{mt+r} \binom{t}{s} = \Theta\left(\binom{t}{mt+r}\right) = \Theta\left(\binom{t}{mt}\right).$$

Thus

$$|X| \leq \sum_{s=0}^{mt+C\log t} |Y \cap V_s| = \Theta\left(\binom{t}{mt}\right) + O\left(\binom{t}{mt}\right) = \Theta\left(\binom{t}{mt}\right),$$

a.a.s., completing the proof. □

4.3 Giant Components in MAGs

We now turn to the giant component in $G = K(n, t, \Theta, \mu)$. As we have from Theorem 1.23 that the expected degree of a vertex in G is

$$n(\mu\alpha + (1-\mu)\beta)^{\omega(v)}(\mu\beta + (1-\mu)\gamma)^{t-\omega(v)}, \tag{4.12}$$

the conditions for the emergence of the giant component in G will depend upon $\mu\alpha+(1-\mu)\beta$ and $\mu\beta+(1-\mu)\gamma$. For ease of notation, we will write $c_1 = \mu\alpha+(1-\mu)\beta$ and $c_2 = \mu\beta + (1-\mu)\gamma$. Moreover, by symmetry, we may assume that $c_1 \geq c_2$, as otherwise we may swap the roles of 0 and 1 and also interchange μ with $1-\mu$ and γ with α.

The following theorem was proven in [43] for the case that $\alpha > \beta > \gamma$.

Theorem 4.12. *Let $G = K(n, t, \Theta, \mu)$, where $t = \rho \log n$. Then G contains a giant component if and only if*

$$\rho < \frac{1}{\log\left(\frac{c_2^{\mu-1}}{c_1^{\mu}}\right)} \tag{4.13}$$

Here, we provide an alternative proof to this theorem to that given in [43], with a structure similar to the proof of Theorem 4.2 and using spectral concentration techniques developed in Chapters 2 and 3. This will allow us to extend this result for arbitrary α, β, γ, and will also allow us to establish the asymptotic size of the giant component in an analogous manner as Theorem 4.9.

4.3.1 Emergence of the Giant Component

Remark 4.13. We note that condition (4.13) is chosen precisely so that if v is a vertex with $mt = \omega(v) \geq \mu t$ (that is, v has at least the expected number of 1s in its attribute vector), then the expected degree of v satisfies

$$\begin{aligned} \mathbb{E}\left[\deg(v)\right] &= n(\mu\alpha + (1-\mu)\beta)^{\omega}(\mu\beta + (1-\mu)\gamma)^{t-\omega} \\ &= n^{1+\rho\log c_2 + m\rho\log\left(\frac{c_1}{c_2}\right)} = n^{C}, \end{aligned}$$

where C is some positive constant. Thus, any vertex with weight at least as large as the expectation has expected degree polynomial in the number of vertices of G.

As with $K(t, \alpha, \beta, \gamma)$, given $s \in [t]$, we define V_s denote the set of vertices with weight s and G_s the subgraph of G induced on V_s. Moreover, for $s \in [t]$, we will define S_s to be the set of words in $\{0,1\}^t$ with weight s, so that V_s is a multi-subset of S_s. Also, we let V_σ be the set of vertices with attribute vector σ, so $|V_\sigma| = n_\sigma$. For a vertex v with weight k, we say a neighbor u of v is of type

(a, b) if u shares $k - a$ 1s with v and $(t - k) - b$ 0s with v. That is to say, we may think of u being obtained from v by changing a 1s to 0s and b 0s to 1s.

Let
$$k = \frac{\mu c_1}{\mu c_1 + (1-\mu)c_2} t, \qquad \ell = \frac{\mu(1-\mu)\beta}{\mu c_1 + (1-\mu)c_2} t.$$
Notice that, as we have assumed $c_1 \geq c_2$, we have
$$k = \frac{\mu c_1}{\mu c_1 + (1-\mu)c_2} t \geq \frac{\mu c_1}{\mu c_1 + (1-\mu)c_1} t = \mu t.$$
Let H denote the subgraph of G with $V(H) = V_k$ and $E(H) = \{uv \in G : u \cdot v = k - \ell\}$. That is, H is the subgraph of vertices with weight k and edges of type (ℓ, ℓ).

Lemma 4.14. *The graph H described above is connected a.a.s..*

Proof. Let $\mathcal{L}_H$ be the normalized Laplacian for H. We make the standard observation that H is connected if and only if the second smallest eigenvalue λ of $\mathcal{L}_H$ is positive by Theorem 1.6.

We first consider the case where $n_\sigma = 1$ for all $\sigma \in S_k$. Let $\bar{D}_H$ and $\bar{A}_H$ denote the expected adjacency matrix and degree matrix for H.

Let $G(t, k, \ell)$ be the intersection graph with vertex set $\binom{[t]}{k}$ and $u \sim v$ if and only if $|u \cap v| = k - \ell$. By Lemma 4.4, we have
$$\begin{aligned}
\operatorname{diam}(G(t,k,\ell)) &\leq \max\left\{3, \left\lceil \frac{k}{k-\ell} \right\rceil, 2\left\lceil \frac{k-\ell}{t-2k+2\ell} \right\rceil\right\} \\
&\leq \max\left\{3, \left\lceil \frac{\mu\alpha + (1-\mu)\beta}{\mu\alpha + (1-\mu)^2\beta} \right\rceil, \right. \\
&\qquad \left. \left\lceil \frac{\mu^2\alpha + \mu(1-\mu)^2\beta}{2\mu(1-\mu)\beta + (1-\mu)^2\gamma - \mu^2\alpha} \right\rceil \right\} \\
&= \Theta(1).
\end{aligned}$$

Thus, the Cheeger constant for $G(t, k, \ell)$ is $h \geq \frac{1}{2c}$ for some $c > 0$. Using Cheeger's Inequality (Theorem 1.9), we obtain that $\lambda \geq \frac{1}{16c^2}$.

Now suppose we have an arbitrary signature $\{n_\sigma\}_{\sigma \in S_k}$. Note that the expected degree of a vertex $v \in V(H)$ is given by
$$\mathbb{E}\left[\deg_H(v)\right] = \sum_{\substack{\sigma \in S_k \\ v \sim_{G(t,k,\ell)} a^{-1}(\sigma)}} n_\sigma \alpha^\ell \beta^{2\ell} \gamma^{t-k-\ell}$$

Take $\delta = n^{-b}$ for some b. Now, given $\sigma \in S_k$, the probability that a given vertex is in V_σ is $\mu^k(1-\mu)^{t-k} = p$. Thus, the expected value of n_σ is pn for all $\sigma \in S_k$, and thus by Chernoff bounds (see Theorem 1.14), we have that $(1-\epsilon^*)np < n_\sigma < (1+\epsilon^*)np$ for all $\sigma \in S_k$ with probability at least $1-\delta$, where

$$\epsilon^* = \sqrt{\frac{2\log\left(\frac{2n}{\delta}\right)}{np}}$$

for n sufficiently large. Moreover, this yields that the number of vertices in V_k is

$$\begin{aligned} |V_k| &= \sum_{\sigma \in S_k} n_\sigma \\ &= \sum_{\sigma \in S_k} (1+o(1))np \\ &= (1+o(1))np\binom{t}{k} \end{aligned}$$

Thus we obtain

$$\begin{aligned} \mathbb{E}\left[\deg_H(v)\right] &= \sum_{\substack{\sigma \in S_k \\ v \sim_{G(t,k,\ell)} a^{-1}(\sigma)}} n_\sigma \alpha^\ell \beta^{2\ell} \gamma^{t-k-\ell} \\ &= (1+o(1))np\binom{k}{\ell}\binom{t-k}{\ell}\alpha^\ell \beta^{2\ell} \gamma^{t-k-\ell} \\ &\geq (1+o(1))np\binom{t}{k}\frac{\binom{k}{\ell}\binom{t-k}{\ell}\alpha^\ell \beta^{2\ell} \gamma^{t-k-\ell}}{\binom{t}{k}} \\ &= (1+o(1))|V_k|\frac{\binom{k}{\ell}\binom{t-k}{\ell}\alpha^\ell \beta^{2\ell} \gamma^{t-k-\ell}}{\binom{t}{k}}. \end{aligned}$$

We note that here, all edges occur with the same probability, so every vertex has the same expected degree. Thus, in order to apply Theorem 2.8, we need that the expected degree of any vertex in H is at least $\log(|V_k|)$. In fact, we will show that $\frac{\binom{k}{\ell}\binom{t-k}{\ell}\alpha^\ell \beta^{2\ell} \gamma^{t-k-\ell}}{\binom{t}{k}}$ is asymptotically larger than $\frac{1}{|V_k|}$, so that the expected degree of any vertex in H is at least polynomial in $|V_k|$.

Now, using inequality (4.1), we obtain

$$\begin{aligned}\frac{\binom{k}{\ell}\binom{t-k}{\ell}\alpha^{\ell}\beta^{2\ell}\gamma^{t-k-\ell}}{\binom{t}{k}} &\geq \frac{\left(\frac{c_1}{(1-\mu)\beta}\right)^{\ell}\left(\frac{c_1}{\mu\alpha}\right)^{k-\ell}\left(\frac{c_2}{\mu\beta}\right)^{\ell}\left(\frac{c_2}{(1-\mu)\gamma}\right)^{t-k-\ell}}{\binom{t}{k}2e^2\pi\ell\sqrt{\frac{\mu(1-\mu)\alpha\gamma}{c_1c_2}}}\alpha^{k-\ell}\beta^{2\ell}\gamma^{t-k-\ell} \\ &= \frac{1}{O(t)}\frac{c_1^k c_2^{t-k}}{\binom{t}{k}\mu^k(1-\mu)^k} \\ &= \frac{1}{O(t)}\frac{nc_1^k c_2^{t-k}}{|V_k|}\end{aligned}$$

Note that $nc_1^k c_2^{t-k}$ is precisely the degree of v in G, and by Remark 4.13 there exists a positive constant C such that $nc_1^k c_2^{t-k} \geq n^C$. Thus, the expected degree of any vertex in H is at least polynomial in $|V_k|$. We first consider the case where each n_σ is precisely np.

Let L be the graph with vertex set npS_k, that is, the union of np disjoint copies of S_k, where adjacency is defined as in $G(t,k,\ell)$. Notice by the above argument that the diameter in L is asymptotically constant, as if the diameter in $G(t,k,\ell)$ is bounded by c, we have

$$\text{dist}(u,v) \leq \begin{cases} c & a(u) \neq a(v) \\ 2 & a(u) = a(v) \end{cases}.$$

Moreover, clearly L is edge transitive, so we obtain $\lambda_1(L) \geq \frac{1}{16c^2}$ by Theorem 1.9 together with Theorem 1.10. Let $q = \alpha^{k-\ell}\beta^{2\ell}\gamma^{t-k-\ell}$. If we let L_q be the percolated version of L, where each edge is chosen independently and with probability q, then we may use an argument similar to the above argument on degree in H to obtain that the expected degree of a vertex in L_q is at least polynomial in $|L_q|$. Write $\bar{d}$ to denote the expected degree of a vertex in L_q. Then we may apply Theorem 2.8 to obtain that $\lambda_1(L_q)$ satisfies

$$\left|\lambda - \frac{1}{16c^2}\right| \leq \frac{1}{32c^2}$$

with probability at least $1 - \exp\left(-\frac{\bar{d}}{27648c^4} - \log 4\right) = 1 - o(1)$.

Let $\tilde{N}, \tilde{M}$ be $\binom{t}{k} \times \binom{t}{k}$ matrices, with

$$\tilde{N}_{\sigma,\tau} = \begin{cases} \sqrt{np \cdot np}\frac{q}{\sqrt{(npq\binom{k}{\ell}\binom{t-k}{\ell})^2}} & |\sigma \cap \tau| = k - \ell \\ 0 & \text{otherwise} \end{cases},$$

$$\text{and}\quad \tilde{M}_{\sigma,\tau} = \begin{cases} \sqrt{n_\sigma n_\tau}\frac{q}{\sqrt{d_\sigma d_\tau}} & |\sigma\cap\tau| = k-\ell \\ 0 & \text{otherwise} \end{cases}.$$

Note that if $N = \bar{D}_{L_q}^{-1/2}\bar{A}_{L_q}\bar{D}_{L_q}^{-1/2}$, and $M = \bar{D}_H^{-1/2}\bar{A}_H\bar{D}_H^{-1/2}$, then by an argument identical to that in Theorems 3.5 and 3.8, $\tilde{N}$ and $\tilde{M}$ capture the non-trivial eigenvalues of N and M, respectively. Moreover, as $\tilde{N} = I - \bar{\mathcal{L}}_{L_q}$, and the eigenvalues of $\mathcal{L}$ are between 0 and 2 for any graph, we must have $\|\tilde{N}\| \leq 1$. Moreover,

$$\begin{aligned}
\|\tilde{M}-\tilde{N}\| &= \max_{\|f\|=1}\left|f^T\left(\tilde{M}-\tilde{N}\right)f\right| \\
&= \max_{\|f\|=1}\left|\sum_{\sigma\in S_k}\sum_{\tau\in S_k} f_\sigma\left(\tilde{M}-\tilde{N}\right)_{\sigma,\tau} f_\tau\right| \\
&\leq \max_{\|f\|=1}\sum_{\sigma\in S_k}\sum_{\substack{\tau\in S_k\\|\sigma\cap\tau|=k-\ell}} |f_\sigma|\left|\tilde{M}-\tilde{N}\right|_{\sigma,\tau}|f_\tau| \\
&\leq \max_{\|f\|=1}\sum_{\sigma\in S_k}\sum_{\substack{\tau\in S_k\\|\sigma\cap\tau|=k-\ell}} |f_\sigma|\frac{2\epsilon^*}{1-\epsilon^*}|\tilde{N}_{\sigma,\tau}||f_\tau| \\
&= \frac{2\epsilon^*}{1-\epsilon^*}\max_{\|f\|=1}\sum_{\sigma\in S_k}\sum_{\substack{\tau\in S_k\\|\sigma\cap\tau|=k-\ell}} |f_\sigma|\tilde{N}_{\sigma,\tau}|f_\tau| \\
&\leq \frac{2\epsilon^*}{1-\epsilon^*}\|\tilde{N}\| \\
&\leq \frac{2\epsilon^*}{1-\epsilon^*} \leq 4\epsilon^*.
\end{aligned}$$

We therefore obtain by Weyl's Theorem (Theorem 1.1) that $|\lambda_{\max}(\tilde{M}) - \lambda_{\max}(\tilde{N})| \leq 4\epsilon^*$, and thus

$$\lambda_{\max}(\tilde{M}) \leq 1 - \frac{1}{32c^2} + 4\epsilon^* = 1 - \frac{1}{32c^2} + 4\sqrt{\frac{2\log\left(\frac{2n}{\delta}\right)}{np}}.$$

Recall that we chose $\delta = n^{-b}$. In fact, if we set $b = \frac{p}{2^{17}c^4} - 1$, we obtain

$$\begin{aligned}
4\sqrt{\frac{2\log\left(\frac{2n}{\delta}\right)}{np}} &= 4\sqrt{\frac{2(1+b)\log(2n)}{np}} \\
&= 4\sqrt{\frac{2\frac{p}{2^{17}c^4}\log(2n)}{np}} \\
&= \frac{1}{2^6c^2}\sqrt{\frac{\log(2n)}{n}} \leq \frac{1}{64c^2},
\end{aligned}$$

and thus $\lambda_{\max}(\tilde{M}) \leq 1 - \frac{1}{64c^2}$. Therefore, with high probability, the second smallest eigenvalue of $\bar{\mathcal{L}}_H$ satisfies $\bar{\lambda}_1 \geq \frac{1}{64c^2}$. Moreover, by the degree considerations above, we obtain concentration on this eigenvalue by Theorem 2.8, and with probability at least $1-\epsilon$,

$$|\lambda_1 - \bar{\lambda}_1| \leq 3\sqrt{\frac{3\log(4|V_k|/\epsilon)}{\bar{d}_{\min}}}.$$

As $\bar{d}_{\min} = \mathbb{E}\left[\deg_H(v)\right]$ which is at least polynomial in $|V_k|$, we may take

$$\epsilon = \frac{1}{4|V_k|} e^{-\frac{\mathbb{E}[\deg_H(v)]}{27 \cdot 2^{14} c^4}}$$

to obtain $|\lambda_1 - \bar{\lambda}_1| \leq \frac{1}{128c^2}$, and thus with high probability,

$$\lambda_1(H) \geq \frac{1}{128c^2} > 0.$$

Thererfore, H is connected a.a.s.. □

We now establish the necessity of the condition in Theorem 4.12.

Theorem 4.15. *Let $G = K(n,t,\Theta,\mu)$ where $c_1 < c_2$ and $\rho > \frac{1}{\log\left(\frac{c_2^{\mu-1}}{c_1^{\mu}}\right)}$. Then G has $n - o(n)$ isolated vertices a.a.s..*

Proof. Let $v \in V$ with $\omega(v) = w$, where $|w - \mu t| \leq (\mu t)^{2/3}$. Write $\rho = \frac{1}{\log\left(\frac{c_2^{\mu-1}}{c_1^{\mu}}\right)} + \epsilon$. Then the expected degree of v is

$$\begin{aligned}
\mathbb{E}\left[\deg(v)\right] &= n c_1^w c_2^{t-w} \\
&\leq n c_1^{\mu t + (\mu t)^{2/3}} c_2^{(1-\mu)t - (\mu t)^{2/3}} \\
&= \exp\left(\frac{t}{\rho} + t\log\left(\frac{c_1^{\mu}}{c_2^{\mu-1}}\right) + (\mu t)^{2/3}\log\left(\frac{c_1}{c_2}\right)\right) \\
&= \exp\left(\frac{t}{\rho} - \frac{t}{\rho-\epsilon} + Ct^{2/3}\right) \\
&= \exp\left(-\frac{\epsilon}{\rho(\rho-\epsilon)}t + Ct^{2/3}\right).
\end{aligned}$$

As $\rho - \epsilon > 0$, this approaches 0, and thus v is isolated a.a.s.. Moreover, by Chernoff bounds (Theorem 1.14), the proportion of vertices with weight w as described is at least $1 - e^{-\frac{1}{2}(\mu t)^{1/3}} = 1 - o(1)$. Therefore, there are at least $n - o(n)$ isolated vertices in G a.a.s.. □

Thus, by Theorem 4.15, if condition (4.13) is not met, G has no giant component a.a.s..

Now, suppose G meets condition (4.13). Let

$$B = \frac{-\log n + \log\left(\frac{e^2 t\pi\beta\mu(1-\mu)\sqrt{\alpha\gamma}}{2c_1c_2}\right) - t\log c_2 - \log\left(1+\frac{\log n}{\log\log n}\right)}{\log\left(\frac{c_1}{c_2}\right)}.$$

Since $t \sim \rho\log n$, we have

$$B = \left(\frac{-\frac{1}{\rho} - \log(c_2)}{\log(\frac{c_1}{c_2})}\right) t + O(\log t)$$

By (4.13), we have $B < \mu t + O(\log t)$. Let $T_s = \{v \in V | \omega(v) < s\}$. We wish to determine T_B, that is, the number of vertices with weight less than B. Write $\zeta = \frac{-\frac{1}{\rho} - \log(c_2)}{\log(\frac{c_1}{c_2})}$. We will consider $T_{\zeta t} \sim T_B$.

Lemma 4.16. *With $T_{\zeta t}$ as defined above, a.a.s. we have*

$$|T_{\zeta t}| \leq Cn\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}.$$

Proof. Let $s = \zeta t$. Notice that $\zeta < \mu$ by condition (4.13). Then the probability that a vertex is in T_s is given by

$$\begin{aligned}
p = \sum_{i=0}^{s}\binom{t}{i}\mu^i(1-\mu)^{t-i} &= \binom{t}{s}\mu^s(1-\mu)^{t-s}\sum_{i=0}^{s}\frac{\binom{t}{i}}{\binom{t}{s}}\mu^{i-s}(1-\mu)^{s-i} \\
&= \binom{t}{s}\mu^s(1-\mu)^{t-s}\sum_{i=0}^{s}\prod_{j=0}^{s-i-1}\frac{1-\mu}{\mu}\frac{s-j}{t-s+1+j} \\
&\leq \binom{t}{s}\mu^s(1-\mu)^{t-s}\sum_{i=0}^{s}\prod_{j=0}^{s-i-1}\frac{1-\mu}{\mu}\frac{s}{t-s} \\
&\leq \binom{t}{s}\mu^s(1-\mu)^{t-s}\sum_{i=0}^{s}\left(\frac{1-\mu}{\mu}\frac{\zeta}{1-\zeta}\right)^{s-i} \\
&= \binom{t}{s}\mu^s(1-\mu)^{t-s}\sum_{i=0}^{s}\left(\frac{1-\mu}{1-\zeta}\frac{\zeta}{\mu}\right)^{s} \\
&\leq \binom{t}{s}\mu^s(1-\mu)^{t-s}\sum_{i=0}^{\infty}\left(\left(\frac{1-\mu}{1-\zeta}\right)\left(\frac{\zeta}{\mu}\right)\right)^{s} \\
&= \binom{t}{s}\mu^s(1-\mu)^{t-s}\frac{\mu(1-\zeta)}{\mu-\zeta}
\end{aligned}$$

As μ, ζ are constants, take $C' = \frac{\mu(1-\zeta)}{\mu-\zeta}$. Then $\mathbb{P}(v \in T_s) \leq C'\binom{t}{s}\mu^s(1-\mu)^{t-s}$.

On the other hand, notice that since $s < \mu$, by the entropy bound (4.1)

$$\begin{aligned} p &= \sum_{i=0}^{s} \binom{t}{i}\mu^i(1-\mu)^{t-i} && (4.14)\\ &\geq \binom{t}{s}\mu^s(1-\mu)^{t-s} && (4.15)\\ &\geq \frac{1}{e\sqrt{2\pi\zeta(1-\zeta)t}}\left(\frac{\mu}{\zeta}\right)^s\left(\frac{1-\mu}{1-\zeta}\right)^{t-s} && (4.16)\\ &= \frac{A}{\sqrt{\log n}} n^{\rho\zeta\log\left(\frac{\mu}{\zeta}\right)+\rho(1-\zeta)\log\left(\frac{1-\mu}{1-\zeta}\right)}, && (4.17)\end{aligned}$$

for $A = \frac{1}{e\sqrt{2\pi\zeta(1-\zeta)\rho}}$ a constant. Put $b = \rho\zeta\log\left(\frac{\mu}{\zeta}\right) + \rho(1-\zeta)\log\left(\frac{1-\mu}{1-\zeta}\right)$.

Let Z be the random variable whose value is $|T_s|$. Then we may think of Z as a sum of independent indicator functions $Z = \sum_{i=1}^n x_i$, where x_i takes value 1 if $v_i \in T_s$ and 0 otherwise. We consider two cases, according as whether $b > -1$ or $b \leq -1$.

If $b > -1$, then $np > \frac{A}{\log n}n^{1+b} > n^r$ for some positive constant r. Then by Theorem 1.14, we have that with probability at least $1 - 2e^{-np/8} = 1 - 2e^{-n^r/8} = 1 - o(1)$

$$|Z - pn| \leq \frac{1}{2}pn, \qquad (4.18)$$

and thus $\frac{1}{2}pn \leq Z \leq \frac{3}{2}pn$, so for $C = \frac{3}{2}C'$, we obtain $Z \leq Cn\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}$ as desired.

If $b \leq -1$, then we apply the entropy bound (4.2) to obtain

$$\begin{aligned} p &= \sum_{i=0}^{s} \binom{t}{i}\mu^i(1-\mu)^{t-i} && (4.19)\\ &\leq C'\binom{t}{s}\mu^s(1-\mu)^{t-s} && (4.20)\\ &\leq C'\frac{1}{\sqrt{2\pi\zeta(1-\zeta)t}}\left(\frac{\mu}{\zeta}\right)^s\left(\frac{1-\mu}{1-\zeta}\right)^{t-s} && (4.21)\\ &\leq \frac{C'Ae}{\sqrt{\log n}}n^b. && (4.22)\end{aligned}$$

Therefore, $np \leq \frac{C'Ae}{\sqrt{\log n}}n^{1+b} \leq n^{-r}$ for some $r \geq 0$. Then as the x_i are independent, we obtain $\mathrm{Var}(Z) = \sum_{i=1}^n \mathrm{Var}(x_i) = np(1-p) \leq np \leq n^{-r}$. By

Chebyshev's Inequality (Theorem 1.12), we have that for any $\epsilon > 0$,

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq n^{-r/2+\epsilon}) \leq \frac{\mathrm{Var}(Z)}{n^{-r+2\epsilon}} \leq \frac{1}{n^{2\epsilon}} = o(1).$$

Therefore, with probability $1 - o(1)$, we have that for $C = 2C'$ and ϵ sufficiently small, $Z \leq pn + n^{-r/2+\epsilon} \leq Cn\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}$ as desired. □

Theorem 4.17. *Let $v \in V(G)$ have weight $\omega(v) > B$. Then v is connected by a path to a vertex of weight k a.a.s..*

Proof. Suppose $v \in V_s$. The expected number of neighbors of v of type (a, b) is

$$n\mu^{s-a+b}(1-\mu)^{t-s+a-b}\binom{s}{a}\binom{t-s}{b}\beta^{a+b}\alpha^{s-a}\gamma^{t-s-b}. \tag{4.23}$$

Let

$$a = \frac{(1-\mu)\beta}{c_1}s, \text{ and } b = \frac{\mu\beta}{c_2}(t-s)$$

Then any neighbor of v of type (a, b) has weight

$$f(s) = \frac{\mu\alpha}{c_1}s + \frac{\mu\beta}{c_2}(t-s). \tag{4.24}$$

Notice that

$$\begin{array}{rclll} f(s) & = & s & \text{when} & s = k \\ f(s) & > & s & \text{when} & s < k \\ f(s) & < & s & \text{when} & s > k \end{array}$$

For $v \in V_s$, let $R_v = \{u \in V | u \text{ is a neighbor of } v \text{ of type } (a, b)\}$. Then as f is linear, if $u \in R_v$, then u has weight $f(s)$ with $|f(s) - k| < |s - k|$.

Now, $|R_v|$ may be considered as a sum of i.i.d. random variables $X_1, X_2, \ldots,$ X_{n-1}, where X_i is the indicator of whether $v_i \in R_v$ (where here we have labeled v as v_n for convenience). Using the entropy bound from inequality (4.1), we obtain

$$\begin{aligned}
\mathbb{P}(v_i \in R_v) &= \binom{s}{a}\binom{t-s}{b}\mu^{s-a+b}(1-\mu)^{t-s+a-b}\beta^{a+b}\alpha^{s-a}\gamma^{t-s-b} \\
&= \binom{s}{a}\binom{t-s}{b}(\mu\alpha)^{s-a}((1-\mu)\beta)^{a}(\mu\beta)^{b}((1-\mu)\gamma)^{t-s-b} \\
&> \frac{c_1c_2(\mu\alpha)^{s-a}((1-\mu)\beta)^{a}(\mu\beta)^{b}((1-\mu)\gamma)^{t-s-b}}{e^2 2\pi\beta\mu(1-\mu)\sqrt{s(t-s)\alpha\gamma}} \cdot \\
&\quad \cdot \left(\frac{c_1}{(1-\mu)\beta}\right)^{a}\left(\frac{c_1}{\mu\alpha}\right)^{s-a}\left(\frac{c_2}{\mu\beta}\right)^{b}\left(\frac{c_2}{(1-\mu)\gamma}\right)^{t-s-b} \\
&= \frac{c_1c_2}{e^2 2\pi\beta\mu(1-\mu)\sqrt{s(t-s)\alpha\gamma}}c_1^s c_2^{t-s} \\
&> \frac{2c_1c_2}{e^2 t\pi\beta\mu(1-\mu)\sqrt{\alpha\gamma}}c_1^s c_2^{t-s} =: q.
\end{aligned}$$

By Chernoff bounds (Theorem 1.15), we thus obtain

$$\begin{aligned}
\mathbb{P}(R_v = \emptyset) &= \mathbb{P}(|R_v| = 0) \\
&\leq \mathbb{P}(|R_v| - \mathbb{E}[|R_v|] < -\mathbb{E}[|R_v|]) \\
&\leq e^{\frac{-(nq)^2}{2nq}} \\
&\leq e^{-nq/2}.
\end{aligned}$$

The constant B was chosen precisely so that $q > \frac{2\log(n)}{n} + \frac{2}{n}\log\log n$, and thus $e^{-nq/2} < \frac{1}{n\log n}$. Thus, v is connected to a vertex with weight $f(s)$ with probability at least $1 - \frac{1}{n\log n}$. □

Theorem 4.18. *The giant component in $G = K(n,t,\Theta,\mu)$ is of size at least $n - |T_B| \geq n - C\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}$.*

Proof. By the above result, the probability that there exists a vertex in $V \backslash T_B$ of weight s that is not connected to a vertex with weight $f(s)$ is at most $1 - \frac{1}{\log n}$ by the union bound. Thus, as the graph induced on V_k is connected by Lemma 4.14, this will imply that we have a component of asymptotic size

$$n - |T_B| \geq n - C\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t} \geq C'n$$

for appropriate choices of C and C', by Lemma 4.16. □

Thus, we have established both the sufficiency of the condition in Theorem 4.12, and in addition, have a lower bound on the size of the giant component in G.

4.3.2 Upper Bound on Size of the Giant Component

The goal of this section is to develop an upper bound on the size of the giant component in $K(n,t,\Theta,\mu)$. In fact, we will prove the following.

Theorem 4.19. *Let X be the set of vertices that are not in the giant component of G. Then*

$$|X| = \Theta\left(\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}\right). \tag{4.25}$$

Theorem 4.19 will yield both the size of the giant component and the uniqueness of the giant component, as with Theorem 4.9. We note that by Theorem 4.18, we have that $|X| = O\left(\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}\right)$, so we need only show that

$$|X| = \Omega\left(\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}\right).$$

Moreover, we will show that there are at least $\Omega\left(\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}\right)$ isolated vertices in X, which is sufficient to establish the uniqueness of the giant component in G.

Lemma 4.20. *Let $G = K(n,t,\Theta,\mu)$, and let $\zeta = \frac{-\frac{1}{\rho}-\log(c_2)}{\log\left(\frac{c_1}{c_2}\right)}$ as above. If v is a vertex with $\omega(v) = \zeta t - 1$, then v is isolated with at least constant probability.*

Proof. Let $v \in V$ with $\omega(v) = \zeta t - 1$. Note then that

$$\begin{aligned}
\mathbb{E}[\deg v] &= nc_1^{\omega(v)}c_2^{t-\omega v} \\
&= n\frac{c_2}{c_1}\left(c_1^{\zeta}c_2^{1-\zeta}\right)^t \\
&= n\frac{c_2}{c_1}\left(c_1^{\zeta}c_2^{1-\zeta}\right)^{\rho\log n} \\
&= \frac{c_2}{c_1}n^{1+\rho\zeta\log(\frac{c_1}{c_2})+\rho\log(c_2)} \\
&= \frac{c_2}{c_1}n^{1+\rho(-\frac{1}{\rho}-\log(c_2))+\rho\log(c_2)} \\
&= \frac{c_2}{c_1} = q < 1.
\end{aligned}$$

Then $p = \mathbb{P}(\deg(v) = 0) \geq 1 - q > 0$ is at least constant. ∎

Proof of Theorem 4.19. Let Y be the number of isolated vertices of weight $\zeta t - 1$. Put $s = \zeta t$. Write $Y = \sum_{v \in V_{s-1}} z_v$, where z_v is the indicator that v is isolated. Notice that for $u, v \in V_{s-1}$, we have

$$\begin{aligned}
\mathbb{E}[z_v z_u] &= \prod_{x \in V} (1 - \mathbb{P}(x \sim v)) \prod_{\substack{y \in V \\ y \neq v}} (1 - \mathbb{P}(y \sim u)) \\
&= \mathbb{E}[z_v]\,\mathbb{E}[z_u] + \frac{\mathbb{P}(u \sim v)\mathbb{E}[z_v]\,\mathbb{E}[z_u]}{1 - \mathbb{P}(u \sim v)} \\
&\leq \mathbb{E}[z_v]\,\mathbb{E}[z_u] + \mathbb{P}(u \sim v)\mathbb{E}[z_v]\,\mathbb{E}[z_u]
\end{aligned}$$

Moreover, $\mathbb{E}[z_v] = p$ as shown in Lemma 4.20, and $\mathbb{P}(u \sim v) \to 0$. Therefore, we have

$$\mathrm{Cov}(z_v, z_u) = \mathbb{E}[z_v z_u] - \mathbb{E}[z_v]\,\mathbb{E}[z_u] = o(\mathbb{E}[z_v]\,\mathbb{E}[z_u]) = o(p^2).$$

Moreover, $\mathrm{Var}(z_v) = p(1-p)$, where p. Therefore,

$$\begin{aligned}
\mathrm{Var}\left(Y^2\right) &= \sum_{v \in V_{s-1}} \mathrm{Var}(z_v) + \sum_{\substack{u,v \in V_{s-1} \\ u \neq v}} \mathrm{Cov}(z_v, z_u) \\
&\leq |V_{s-1}|p(1-p) + o\left(p^2 |V_{s-1}|^2\right) = o\left(|V_{s-1}|^2\right).
\end{aligned}$$

Therefore, by Chebyshev's inequality (see Theorem 1.12), we have that for any $c > 0$,

$$\mathbb{P}\left(|Y - p|V_{s-1}|| \geq c|V_{s-1}|\right) \leq \frac{o\left(|V_{s-1}|^2\right)}{c^2 |V_{s-1}|^2} = o(1).$$

Therefore, $Y = p|V_{s-1}| + o(|V_{s-1}|)$.

Moreover, $|V_{s-1}| = \sum_{i=1}^n Z_i$, where Z_i is the indicator of whether $\omega(v_i) = s - 1$. We note that this is a sum of i.i.d. indicator functions with probability $r = \binom{t}{s-1} \mu^{s-1}(1-\mu)^{t-s+1}$. Note then that by inequality (4.17),

$$\begin{aligned}
r &= \binom{t}{s-1} \mu^{s-1}(1-\mu)^{t-s+1} \\
&= \frac{s}{t-s+1} \frac{1-\mu}{\mu} \binom{t}{s} \mu^s (1-\mu)^{t-s} \\
&\geq \frac{A'}{\sqrt{\log n}} n^{\rho\zeta \log\left(\frac{\mu}{\zeta}\right) + \rho(1-\zeta)\log\left(\frac{1-\mu}{1-\zeta}\right)},
\end{aligned}$$

where $A' = \frac{\zeta(1-\mu)}{(2-\zeta)\mu e \sqrt{2\pi\zeta(1-\zeta)\rho}}$ is constant. We take $b = \rho\zeta \log\left(\frac{\mu}{\zeta}\right) + \rho(1-\zeta)\log\left(\frac{1-\mu}{1-\zeta}\right)$. As with the proof of Theorem 4.16, we consider whether $b > -1$ or $b \leq -1$.

If $b > -1$, then by Chernoff bounds (Theorem 1.14) we will obtain that with probability at least $1 - 2e^{-np/8} = 1 - o(1)$, $||V_{s-1}| - rn| \leq \frac{1}{2}rn$. If $b \leq -1$, then by Chebyshev's Inequality (Theorem 1.12), we will obtain that with probability at least $1 - \frac{np}{\log^2(n)} = 1 - o(1)$, $||V_{s-1}| - rn| \leq rn$. These two arguments are identical to those in Theorem 4.16.

Therefore, we obtain that with high probability, $|V_{s-1}| = O(\mathbb{E}[|V_{s-1}|])$, and thus $Y = (1 + o(1))C\mathbb{E}[|V_{s-1}|]$. It remains only to bound $\mathbb{E}[|V_{s-1}|]$. But

$$\begin{aligned} \mathbb{E}[|V_{s-1}|] = rn &= \frac{s}{t-s+1}\frac{1-\mu}{\mu}n\binom{t}{s}\mu^s(1-\mu)^{t-s} \\ &= C'n\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}, \end{aligned}$$

and thus $Y = (1 + o(1))Cn\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}$.

Now, let X be the set of vertices that are not in the giant component. By Theorem 4.18, we have that $|X| \leq C''n\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}$. On the other hand, by the above argument, since $Y \subset X$, we must have $|X| \geq (1+o(1))Cn\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}$. Therefore, $X = \Theta\left(\binom{t}{\zeta t}\mu^{\zeta t}(1-\mu)^{t-\zeta t}\right)$ as desired. $\square$

Acknowledgement: Part of this chapter is based on the paper "Giant components in Kronecker graphs", written jointly with Paul Horn [40]. This paper has appeared in Random Structures and Algorithms 40(3), pp. 385-397, 2012. Material appears with kind permission of Wiley-Blackwell.

Part of this chapter is based on unpublished joint work with Stephen Young.

Bibliography

[1] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Trans. Inform. Theory*, 48(3):569–579, 2002.

[2] R. Albert and A. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2001.

[3] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6:83–96, 1986.

[4] N. Alon, I. Benjamini, and A. Stacey. Percolation on finite graphs and isoperimetric inequalities. *The Annals of Probability*, 32(3):1727–1745, 2004.

[5] N. Alon and J. Spencer. *The Probabilistic Method*. John Wiley & Sons, 3 edition, 2008.

[6] L. Amaral, A. Scala, M. Barthelemy, and H. Stanley. Classes of behavior of small-world networks. *Proc. Natl. Acad. Sci. U. S. A.*, 97(21):11149–11152, 2000.

[7] A. Banerjee. *The spectrum of the graph Laplacian as a tool for analyzing structure and evolution of networks*. PhD thesis, Universität Leipzig, 2008.

[8] A. Banerjee and J. Jost. *Dynamics On and Of Complex Networks; Modeling and Simulation in Science, Engineering, and Technology*, chapter Spectral characterization of network structures and dynamics, pages 117–132. Springer, 2009.

[9] A. Barabási and A. Réka. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[10] S. N. Bernstein. On a modification of chebyshev's inequality and of the error formula of laplace. *Ann. Sci. Inst. Sav. Ukraine*, 4(5), 1924.

[11] J. Bevis, K. Blount, G. Davis, G. Domke, and V. Miller. The rand of a graph after vertex addition. *Linear Algebra and its Applications*, 265:55–69, 1997.

[12] B. Bollobás, C. Borgs, J. Chayes, J. Kim, and D. Wilson. The scaling window of the 2-sat transition. *Random Structures and Algorithms*, pages 182–195, 2002.

[13] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan. Percolation on dense graph sequences. *Annals of Probability*, 38(1):150–183, 2010.

[14] A. Broder, R. Kumar, F. Maghoull, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web: experiments and models. *Proceeings of the Ninth International World-Wide Web Conference*, pages 309–320, 2000.

[15] D. Chakrabarti, Y. Zhan, and C. Faloutsos. R-mat: A recursive model for graph mining. *SIAM International Conference on Data Mining*, 2004.

[16] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.*, 23(4):493–507, 1952.

[17] F. Chung. *Spectral Graph Theory.* AMS Publications, 1997.

[18] F. Chung and R. Graham. Quasi-random tournaments. *Journal of Graph Theory*, 15:173–198, 1991.

[19] F. Chung, P. Horn, and L. Lu. The giant component in a random subgraph of a given graph. In *Proceedings of the 6th Workshop on Algorithms and Models for the Web-graph*, pages 38–49. Lecture Notes in Computer Science, 2009.

[20] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Ann. Combin.*, 6:125–145, 2002.

[21] F. Chung, L. Lu, and V. Vu. Eigenvalues of random power law graphs. *Ann. Combin.*, 7:21–33, 2003.

[22] F. Chung, L. Lu, and V. Vu. The spectra of random graphs with given expected degrees. *Internet Math*, 1:257–275, 2004.

[23] F. Chung and M. Radcliffe. On the spectra of general random graphs. *Electronic Journal of Combinatorics*, 18(1):P215, 2011.

[24] A. Coja-Oghlan. On the laplacian eigenvalues of $g_{n,p}$. *Comb. Prob. Comp.*, 16:923–946, 2007.

[25] D. Cristofides and K. Markström. Expansion properties of random cayley graphs and vertex transitive graphs via matrix martingales. *Random Structures and Algorithms*, 32(8):88–100, 2008.

[26] D. de Solla Price. Networks of scientific papers. *Science*, 149(3683):510–515, 1965.

[27] P. Domingos and M. Richardson. Mining the network value of customers. *Prceedings of the 7th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57–66, 2001.

[28] R. Durrett. *Probability: Theory and Examples.* Duxbury Advanced Series. Brooks/Cole, third edition, 2005.

[29] P. Erdős and T. Luczak. Changes of leadership in a random graph process. *Random Structures and Algorithms*, 243-252, 1994.

[30] P. Erdős and A. Rényi. On random graphs. i. *Publ. Math. Debrecen*, 6:290–297, 1959.

[31] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[32] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Comput. Common. Rev.*, 29(4):251–262, 1999.

[33] U. Feige and E. Ofek. Spectral techniques applied to sparse random graphs. *Random Structures and Algorithms*, 27(2):251–275, 2005.

[34] A. Frieze, M. Krivelevich, and R. Martin. The emergence of a giant component in random subgraphs of pseudo-random graphs. *Random Structures and Algorithms*, 24(1):42–50, 2004.

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

[36] E. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30:1141–1144, 1959.

[37] A. Goerdt, T. Chemnitz, and T. Informatik. The giant component threshold for random regular graphs with edge faults. In *In MFCS*, pages 279–288. Springer, 1997.

[38] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inform. Theory*, 57:1548–1566, 2011.

[39] P. Horn. *Random subgraphs of a given graph.* PhD thesis, University of California, San Diego, 2009.

[40] P. Horn and M. Radcliffe. Giant components in kronecker graphs. *Random Structures and Algorithms*, 40(3):385–397, 2012. Random Structures and Algorithms.

[41] R. Horn and C. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

[42] M. Kang and T. Seierstad. The critical phase for random graphs with a given degree sequence. *Comb. Prob. Comp.*, 17:67–86, 2008.

[43] M. Kim and J. Leskovec. Multiplicative attribute graph model of real-world networks. *Seventh Workshop on Algorithms and Models for the Web Graph*, 2010.

[44] M. Kim and J. Leskovec. Modeling social networks with node attributes using the multiplicative attribute graph model. *Conference on Uncertainty in Artificial Intelligence*, 2011.

[45] W. Kirsch and P. Müller. Spectral properties of the laplacian on bond-percolation graphs. *Math. Z*, 252:899–916, 2006.

[46] J. Kleinberg, S. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. The web as a graph: Measurements, models, and methods. *Proceedings of the 5th International Conference on Combinatorics and Computing*, pages 1–17, 1999.

[47] M. Krivelevich and B. Sudakov. *Pseudo-random graphs*, volume 15 of *Bolyai Soc. Math. Stud.*, pages 199–262. Springer, 2006.

[48] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. Extracting large-scale knowledge bases from the web. *Proceedings of the 25th VLDB Conference*, pages 639–650, 1999.

[49] J. Leskovec, D. Chakrabarti, J. Kleinberg, and C. Faloutsos. Realistic, mathematically tractable graph generation and evolution, using kronecker multiplication. *PKDD*, pages 133–145, 2005.

[50] J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani. Kronecker graphs: an approach to modeling networks. *J. Mach. Learn. Res.*, 11:985–1042, 2010.

[51] J. Leskovec and C. Faloutsos. Scalable modeling of real graphs using kronecker multiplication. *International Conference on Machine Learning*, pages 497–504, 2007.

[52] E. H. Lieb. Convex trace functions and the wigner-yanase-dyson conjecture. *Adv. Math.*, 11:267–288, 1973.

[53] C. V. Loan. The ubiquitous kronecker product. *J. Comput. Appl. Math*, 123:85–100, 2000.

[54] T. Luczak. Component behavior near the critical point of the random graph process. *Random Structures and Algorithms*, 1(3):287–310, 1990.

[55] M. Mahdian and Y. Xu. Stochastic kronecker graphs. *Proceedings of the 5th Workshop on Algorithms and Models for the Web-Graph*, 2007.

[56] S. Milgram. The small-world problem. *Psychology Today*, 2:60–67, 1967.

[57] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6:161–179, 1995.

[58] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, 7(3):295–305, 1998.

[59] M. Newman. The strcture of scientific collaboration networks. *Proc. Natl. Acad. Sci. U. S. A.*, 98(2):404–409, 2001.

[60] R. Oliveira. Concentration of the adjacency matrix and of the laplacian in random graphs with independent edges.

[61] M. Radcliffe and S. Young. The spectra of multiplicative attribute graphs. Submitted.

[62] B. Recht. Simpler approach to matrix completion. to appear in *J. Mach. Learn. Res.*

[63] P. Stanica. Good lower and upper bounds on binomial coefficients. *Journal of Inequalities in Pure and Applied Mathematics*, 2(3), 2001.

[64] J. Tropp. User-friendly tail bounds for sums of random matrices.