2021-07-15

The Non-Backtracking Spectrum of a Graph and Non-Bactracking PageRank

Cory Glover
Brigham Young University

The Non-Backtracking Spectrum of a Graph and Non-Backtracking PageRank

Cory Glover

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Mark Kempton, Chair
Emily Evans
Benjamin Webb

Department of Mathematics

Brigham Young University

ABSTRACT

The Non-Backtracking Spectrum of a Graph and Non-Backtracking PageRank

Cory Glover
Department of Mathematics, BYU
Master of Science

This thesis studies two problems centered around non-backtracking walks on graphs. First, we analyze the spectrum of the non-backtracking matrix of a graph. We show how to obtain the eigenvectors of the non-backtracking matrix using a smaller matrix and in doing so, create a block diagonal decomposition which more clearly expresses the non-backtracking matrix eigenvalues. Additionally, we develop upper and lower bounds on the matrix spectrum and use the spectrum to investigate properties of the graph. Second, we investigate the difference between PageRank and non-backtracking PageRank. We show some instances where there is no difference and develop an algorithm to compare PageRank and non-backtracking PageRank under certain conditions using μ-PageRank.

Keywords: graphs, networks, random walk, non-backtracking random walk, spectral graph theory, PageRank

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. Mark Kempton, for introducing me to spectral graph theory and teaching me the power of persistence in research. Thank you for countless hours of discussion and many more hours reading drafts of my thesis along the way. Thank you to my committee members, Dr. Emily Evans and Dr. Benjamin Webb, for their mentorship and counsel. Many thanks to my other BYU mathematics professors, in particular Dr. Mark Hughes, Dr. Tyler Jarvis, Dr. Blake Barker, and Dr. Jared Whitehead, for their teaching, guidance and patience. And thank you to my high school calculus teacher, Sandy Friedman, for kindling a love of mathematics.

I would also like to thank Alice Oveson, Tyler Jones, Adam Knudson, and Nolan Faught for their many helpful discoveries and insights which were pivotal to this thesis. Also infinite thanks to all in my ACME cohort for their friendship and support. And thank you to my current cohort, in particular Marja Crawford, Rebecca Flores, Jason Gardiner and Michael Nelson, who convinced me I could do graduate level mathematics.

A special thank you to my parent in-laws, Kimberly and David Frandsen, who always ask questions about my research and who gifted a ski pass that helped my stress levels more than they know. And words cannot express my gratitude to my parents, Diana and Kevin Glover, for encouraging me, counseling me, and never giving up on me in every aspect of my life.

Most importantly, thank you to my wife Summer for always listening to me talk about my research, being patient when I interrupt hers', and teaching me to believe in myself.

Contents

List of Tables

List of Figures

Chapter 1. Introduction

Random walks on graphs are frequently used to model behaviors of processes across networks [28]. A natural restriction to a random walk is not allowing backtracks, such as modeling a road network where one wants to avoid u-turns. That is, at a given step, the random walk cannot immediately return to where it just came from. This restricted walk is called a non-backtracking random walk. Non-backtracking random walks have been an area of study as early as 1966 [20]. However, they became a more prominent topic later after Ortner and Wosses' [37] use of non-backtracking random walks to study the cogrowth of graphs and Alon et. al. proved non-backtracking random walks mix faster than simple random walks on regular graphs [2]. Now non-backtracking random walks are frequently studied in graph focused research problems. For example, non-backtracking random walks are often used in conjunction with random graphs (due to their efficacy in modeling real world networks) [7, 8, 11, 30, 41, 45]. Bordenave et. al. [9], Newman [34], and Behnaz et. al. [32] have all used non-backtracking random walks for community detection and Pan, Jiang and Xu used them to maximize influence on social networks [38]. Additionally, centrality measures using non-backtracking random walks have been discussed by a variety of authors [4, 6, 25, 31, 33, 47]. These are just a few examples from the growing catalog of research into non-backtracking random walks.

This heightened interest in non-backtracking random walks is the motivation behind this thesis. The thesis will be broken into two primary topics: 1) the non-backtracking spectrum of a graph and 2) non-backtracking PageRank. We will begin by establishing necessary definitions and notation in Chapter 2. In Chapter 3, we will study the non-backtracking spectrum of a graph. Specifically, we will create upper and lower bounds on the non-backtracking spectrum as well as classify the spectrum for specific classes of graphs (e.g. bipartite graphs, trees, cycles, etc.). Lastly, Chapter 4 will examine the similarities and differences between non-backtracking PageRank and standard PageRank. In particular,

we will emphasize when these two centrality measures are equivalent and when they differ.

Chapter 2. Definitions and Notation

2.1 Graphs and Their Matrices

Let $G = (V, E)$ be a *graph* where V is a set of vertices and E is a set of edges. In general, V is a set of objects and E is a set of connections between those objects. We say $V(G)$ is the vertex set of G and $E(G)$ is the edge set of G. These edges may be undirected, directed, weighted and/or unweighted. An edge is directed if it connected vertex i to j, but not vice versa. An edge is undirected if it is not directed. An edge weight is a numerical values associated with the edge. The greater the value, the stronger the connection is. If the edges of a graph do not have weights, all edges are assumed to have weight 1 and the graph is considered unweighted. If vertices i and j are connected by an edge, we will say $i \sim j$. Alternatively, we say i and j are neighbors in the graph G. Unless otherwise stated, we will assume a graph G is undirected, unweighted, connected and simple. A connected graph is a graph where each vertex v_i can reach every other vertex v_j through some collection edges in $E(G)$. A strongly connected graph is a directed graph where each vertex v_i can reach every other vertex v_j through some collection of edges in $E(G)$. Alternatively, a weakly connected graph is a directed graph where for each pair of vertices v_i and v_j, there exists a collection of edges whereby either v_i can reach v_j or v_j can reach v_i (but not necessarily both). A simple graph is a graph with no more than one edge between two vertices and no loops from a vertex to itself. Throughout the paper we will define $n = |V(G)|$ to be the size of the vertex set of G and $m = |E(G)|$ the size of the edge set of G. Assume n, m are both finite.

To study graphs, we often study the matrices related to the graph (in particular the spectrum of the matrix). Some of the most common matrices used are the *adjacency matrix* (A), the *Laplacian matrix* (L), and the *degree matrix* (D). The content of this paper will primarily use the adjacency matrix and the degree matrix (as well as other matrices to be

defined later).

The adjacency matrix encodes the edge set of a graph in an $n \times n$ matrix where n is the size of the vertex set. The (i, j)-th entry of A is 1 if $i \sim j$ and 0 otherwise:

$$A_{ij} = \begin{cases} 1 & i \sim j \\ 0 & i \not\sim j \end{cases}.$$

Given a vertex $v_i \in V(G)$, the *degree* of v_i (denoted d_i) is the number of vertices v_i is connected to. The degree matrix D encodes the degree of each vertex in a diagonal matrix where $D_{ii} = d_i$. We use the same labeling in A as in D (i.e., the i^{th} row of A represents the same vertex as the i^{th} diagonal entry of D). Additionally, we will say $d_{\max}$ is the maximal entry of D and similarly for $d_{\min}$.

Primarily, we will study the spectrum of the matrices of a graph to understand properties of the graph. Additionally, we will also use the eigenvectors of a graph to understand its properties (e.g. PageRank). The spectrum of the adjacency matrix of a graph has been thoroughly studied [12] and can be used to described how connected a graph is and bipartiteness, among other things. Throughout the paper, we will denote $\sigma(M)$ to be the spectrum of matrix M.

2.2 Random Walks on Graphs

A *walk* along a graph G is a collection of vertices $\{v_1, v_2, ...\}$ where $v_i \in V(G)$ and $v_i \sim v_{i-1}$. A *random walk* along a graph G is a walk where v_i is chosen uniformly at random from the neighbors of v_{i-1}. A random walk is an example of a Markov chain. To encode a random walk in matrix form, we define the *transition probability matrix* $P = D^{-1}A$. The (i, j)-th entry of P is the probability of traveling from i to j in a random walk.

The *stationary distribution* of a random walk describes the probability of being at a vertex in an infinite random walk across G. We can solve for the stationary distribution of

a random walk by finding the left eigenvector of P such that $\mathbf{x}^T P = \mathbf{x}$. By the Gershgorin-Disk theorem, we know the spectral radius of P is 1 and as such, this eigenvector exists [19]. Further, by the Perron-Frobenious theorem we know the stationary distribution is unique for a simple, connected graph G [19].

2.3 Non-backtracking Random Walks

In this section we will define a non-backtracking random walk and some of the associated matrices used when studying non-backtracking random walks.

Definition 2.1 (Non-backtracking Random Walk)**.** A *non-backtracking random walk* is a walk $\{v_1, v_2, ...\}$ where v_i is chosen uniformly from the neighbors of v_i, excluding v_{i-1}.

The idea of a non-backtracking random walk is we perform a modified random walk, where at each vertex v_i we cannot travel to the immediately previous vertex. Note this does not mean we never travel to the immediate previous vertex again, just we will not travel to it immediately after leaving it. This also means that when arriving at a degree 1 vertex, our non-backtracking random walk will enter a sink and the walk will end.

Since at each step the non-backtracking random walk must "remember" where it previously was, a non-backtracking random walk is NOT a Markov chain. To gain back the Markov properties of a random walk, we create a new representation of the non-backtracking random walk. Given a graph G, create a new set of directed edges by taking each undirected edge (i, j) and creating two directed edges (i, j) and (j, i) (see Figure 2.1). We define a new graph $\hat{G}$ where the vertices are the directed edges formed and we say two directed edges $(i, j), (k, l)$ are connected if $j = k$. We define the edge adjacency matrix C in the natural way:

$$C((i,j),(k,l)) = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}.$$

Note that performing a random walk along $\hat{G}$ is equivalent to a random walk on G.

In order to represent a non-backtracking random walk we create a new graph with the same vertex set as $\hat{G}$. However the edge set will be defined as

$$\{(i,j) \sim (k,l) \colon j = k \text{ and } i \neq l\}.$$

Thus two directed edges are connected if they are connected as in $\hat{G}$ AND they do not backtrack each other. With this condition we can define the non-backtracking matrix of a graph G:

$$B((i,j),(k,l)) = \begin{cases} 1 & j = k, i \neq l \\ 0 & \text{otherwise} \end{cases}.$$

Considering B as the adjacency matrix of a directed graph, we can consider a random walk across this new graph which will be equivalent to a non-backtracking random walk across G. As a result, if we consider the non-backtracking random walk as a walk across the graph generated by B, we regain all the Markov chain properties lost by adding the non-backtracking condition.

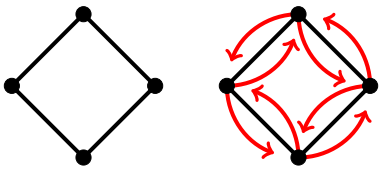


Figure 2.1: A NBRW along the 4-cycle can be considered a Markov chain by performing a simple random walk along the directed edges of the right graph with a non-backtracking condition.

To relate the vertex and edge spaces, we also employ the following matrices. Define

$S \in M_{2m\times n}$ and $T \in M_{n\times 2m}$ where

$$S((u,v),x) = \begin{cases} 1 & v = x \\ 0 & \text{otherwise} \end{cases} \qquad T(x,(u,v)) = \begin{cases} 1 & x = u \\ 0 & \text{otherwise} \end{cases}. \tag{2.1}$$

S is considered the vertex to edge matrix and similarly T is the edge to vertex matrix. Define $\tau \in M_{2m\times 2m}$ to be the backtracking operator

$$\tau((u,v),(x,y)) = \begin{cases} 1 & v = x \text{ and } u = y \\ 0 & \text{otherwise} \end{cases}. \tag{2.2}$$

Using the above matrices, we can build many of the matrices needed to study non-backtracking random walks:

$$C = ST, \qquad B = ST - \tau, \qquad D = T\tau S, \qquad A = TS. \tag{2.3}$$

Chapter 3. Non-backtracking Spectrum

Spectral graph theory is pivotal in the study of graphs. Often it is employed through analyzing the spectrum of the adjacency matrix or Laplacian matrix in order to understand the structure of a graph [12]. Our question is whether we can use the spectrum of the non-backtracking matrix G in the same way. We will call this spectrum the *non-backtracking spectrum*. Many have noted the relationship between the non-backtracking spectrum and zeta functions through Ihara's theorem (see Theorem 3.1) [2, 22, 30]. More specifically, Angel et. al. have studied the non-backtracking spectrum of the universal cover of a graph [3] and Bordenave et. al., Gulikers et. al., Wang and Wood, and others have studied the non-backtracking spectrum of random graphs [9, 11, 17, 44, 48].

In this chapter we will further study the non-backtracking spectrum. We will relate the

non-backtracking spectrum to graph structures with few edges (e.g. trees and cycles) and bipartite graphs. We will also define a new matrix K (see Section 3.3) which relates to B through an invariant subspace. Using K, we will decompose B to more descriptively display its spectrum for a general graph. Lastly, we will create upper and lower bounds on the non-backtracking spectrum.

3.1 Previous Results and Irreducibility

It has previously been proven that the spectrum of B can be found using Ihara's Theorem (see [23]).

Theorem 3.1 (Ihara's Theorem [20])**.** *Given a graph G with n vertices and m edges, let B be the non-backtracking matrix of G as defined above. Let A denote the adjacency matrix of G and D the degree matrix. Then*

$$det(I - uB) = (1 - u^2)^{m-n} det(u^2(D - I) - uA + I). \tag{3.1}$$

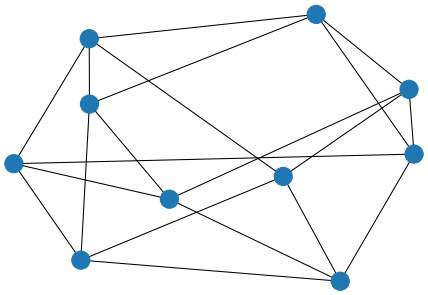
Thus the eigenvalues of B are either ± 1 or solutions to $\det(\mu^2 I - \mu A + (D - I)) = 0$. Our goal is identify properties of the spectrum of B for a given graph G. When G is d-regular (each vertex has degree d), then Ihara's Theorem works out in a straightforward way to give the spectrum of B (see [2, 22, 30]). Similar results have been obtained for bipartite biregular graphs [22].

Theorem 3.2. *Let G be a d-regular graph and A the adjacency matrix of G. Then the eigenvalues of B are either*

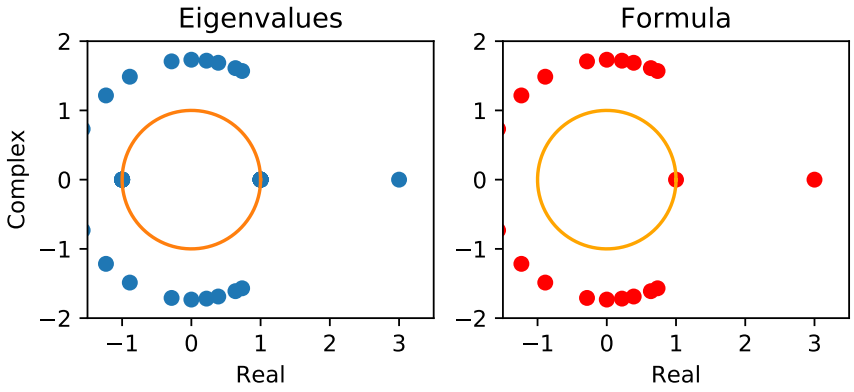
$$\frac{\lambda_i \pm \sqrt{\lambda_i^2 - 4(d-1)}}{2}, (i = 1, ..., n) \tag{3.2}$$

where $\lambda_i \in \sigma(A)$ or ± 1, with the extra ± 1 appearing $m - n$ times each. (Note that the

multiplicity of 1 *is* $m - n + 1$ *and* -1 *has multiplicity* $m - n + 1$ *if* G *is bipartite as an additional* 1 *and* -1 *come from Equation 3.2).*



(a) 4-regular graph



(b) Spectrum Plots

Figure 3.1: The graph above is a 4-regular graph with 10 vertices. The spectrum plot on the left is the spectrum of the non-backtracking matrix of said graph. The spectrum plot on the right shows the values computed using the spectrum of the adjacency matrix of said graph and equation 3.2. As can be seen, -1 is missing from the right plot as it is not computed from equation 3.2 but is in the left plot as it comes from the extra -1 with multiplicity $m - n$.

In general it is useful to know when B is irreducible. Throughout the paper, we will use this condition to employ the Perron-Frobenius theorem.

Proposition 3.3. *Let G be a connected graph that is not a cycle and $d_{\min} \geq 2$. Then B is irreducible.*

Proof. Since B represents a directed graph $\hat{G}$, it suffices to show a random walk on $\hat{G}$ is irreducible. A random walk along a graph is irreducible if for every pair of vertices i and j, there exists a path from i to j. Since $\hat{G}$ is a directed graph, this is equivalent to showing $\hat{G}$ is strongly connected.

Denote i as the directed edge from $a \to b$ and $-i$ as the directed edge from $b \to a$ where a and b are vertices of the original graph G. Note if G is connected, a simple random walk is irreducible on G. Assume a path exists on $\hat{G}$ between i and $-i$ for every i in $\hat{G}$. Then a simple random walk across $\hat{G}$ will have the ability to backtrack after a finite number of steps. Combining this with the connectedness of G implies $\hat{G}$ must be strongly connected. Hence it is sufficient to show there is a path from i to $-i$ for every i in $\hat{G}$.

Since G is not a cycle, there must exist at least one vertex with degree greater than or equal to 3. We examine a walk across $\hat{G}$ beginning at directed edge i. Assume i is pointing towards a vertex a of degree at least 3. Assume we take the shortest path from i to some directed edge pointing towards a which is not the trivial path of length 0. If this path arrives at an edge k pointing towards a such that $k \neq i$, then we can take a step from k to $-i$ since G is connected. If this path arrives at i, then take a step onto a directed edge j which is not in the current path. We are guaranteed such j exists since a has degree at least 3. Take the shortest path from j to another directed edge pointing towards a which is not j itself. If this path arrives at some directed edge $l \neq i$, then we can take a step from l to $-i$ since G is connected. If this path arrives at i, then at some point along the path from j to i we intersected at some edge r from the first shortest path found. When arriving at r, rather than continue to i via r, step to $-r$. Then by the existence of the path from i to r, there must exist a path from $-r$ to $-i$. Hence a path always exist between i and $-i$.

Now assume i is pointing towards a vertex b of degree 2. Then a random walk on $\hat{G}$ beginning at i can only travel to one edge. Since G is connected and G is not a cycle, there

exists a path from i to some edge k which points to a vertex of degree at least 3. Thus, by the previous statement, a path exists between k and $-k$. Then there is clearly a path from $-k$ to $-i$. Hence there is a path from i to $-i$. Hence, G is strongly connected and B is irreducible. □

3.2 Examples with few edges

We begin studying the non-backtracking spectrum by looking at graphs where $m \leq n$ (graphs which are trees or cycles). The non-backtracking spectrum of both trees and cycles have been previously analyzed by Torres [46] (see Theorems 3.5 and 3.8). In this section, we further this study by providing alternate proofs to theorems regarding the non-backtracking spectrum of trees (Theorem 3.5 and Corollary 3.6). Additionally, we study the non-backtracking spectrum of graphs with pendant cycles in Theorem 3.9 and Corollary 3.10.

3.2.1 Trees. A tree is a connected graph which contains no cycles. Trees have been found to be difficult to distinguish using the spectrum of A, since Schwenk showed many large trees have the same spectrum [40]. In the case of the non-backtracking matrix, all trees have the same spectrum. To show this, we first find the characteristic polynomial of the edge adjacency matrix C where the G is a directed tree, with all edges pointing towards a root vertex and use this expression to find the non-backtracking spectrum of a tree.

Remark. The non-backtracking spectrum of a tree is straightforward and can be found in Torres' paper [46]. His proof method uses properties of non-backtracking random walks. We will use cofactor expansion to show the same result. Additionally we will give an alternate proof to one of his corollaries using cofactor expansion.

Lemma 3.4. *Let G be a directed tree with n vertices where all edges eventually point to one root vertex. Let C be the edge adjacency matrix of G. Then $det(\lambda I - C) = \lambda^{n-1}$.*

Proof. Note a row representing any edge pointing directly to the root vertex of G will have all 0 entries. Choose an edge j pointing to the root and perform cofactor expansion across

this row of $\lambda I - C$ corresponding to j. This gives $\det(\lambda I - C) = \lambda \det(\lambda I - C_{\neq\{j\}})$, where $C_{\neq\{j\}}$ is C with the row and columns for j deleted. By induction, continue on every edge pointing to the root. The new $\hat{C}$ will then consist of a number of directed trees with all edges pointing to a root. Repeat by induction. Thus, $\det(\lambda I - C) = \lambda^{n-1}$. □

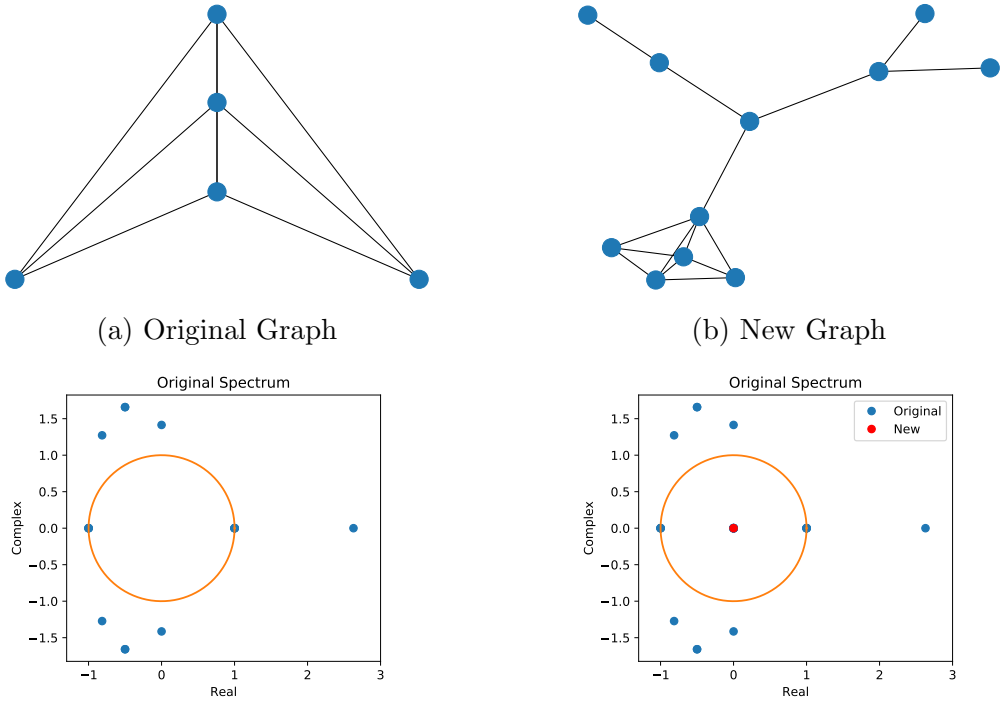
Theorem 3.5 ([46]). *Let B be the non-backtracking matrix of a tree G. Then det$(\lambda I - B) = \lambda^{2(n-1)}$.*

Proof. Let G be a graph with n vertices and k leaves. Let the i^{th} row of B represent an edge pointing towards a leaf in G. Every entry of this row will be 0. We write B such that the first k rows represent the k edges pointing towards the k leaves of the graphs. We perform cofactor expansion, choosing the first row of $\lambda I - B$. Then $\det(\lambda I - B) = \lambda \det(\lambda I - B_{\neq 1})$ where $B_{\neq 1}$ is B without the 1^{st} row or column. Continuing we get $\det(\lambda I - B) = \lambda^k \det(\lambda I - B_{\neq\{1,\ldots,k\}})$.

Let the i^{th} row of $B_{\neq\{1,\ldots,k\}}$ represent the edge pointing to the parent of a leaf vertex. Since all rows representing edges pointing to leaf vertices have been removed, the i^{th} row will only have one nonzero entry λ. Thus, $\det(\lambda I - B) = \lambda^{k+1} \det(\lambda I - B_{\neq\{1,\ldots,k,i\}})$. Continue this process for the parents of all leaf vertices and then for parents of parents. This continues until $B_{\neq\{1,\ldots,k,i_1,i_2,\ldots,i_j\}}$ represents an adjacency matrix of a directed tree of $n-1$ edges. By Lemma 3.4, we get $\det(\lambda I - B) = \lambda^{2(n-1)}$. □

Corollary 3.6 ([46]). *Let G' be a graph with m edges and T be a tree with n vertices. Let B' be the non-backtracking matrix of G'. Define G as the graph constructed by joining G' and T at one vertex (i.e., one vertex in G' will also be a vertex in T). Define B as the non-backtracking matrix of G. Then $\sigma(B)$ is $\sigma(B')$ along with eigenvalue 0 with algebraic multiplicity $2(n-1)$.*

Proof. Order the entries of B such that the last $2(n-1)$ entries represent edges in the tree T. Consider the matrix $\lambda I - B$. We can use the same method in Theorem 3.5 to find $\det(\lambda I - B) = \lambda^{2(n-1)} \det(\lambda I - B_{\neq\{2m+1,2m+2,\ldots,2m+2(n-1)\}}) = \lambda^{2(n-1)} \det(\lambda I - B')$. The result follows. □



(a) Original Graph

(b) New Graph

(c) Original Spectrum

(d) New Spectrum

Figure 3.2: The graph shown in (a) is our original graph and (b) shows the same graph with a tree concatenated at one vertex. Below each respective graph is their spectrum plot. As can be seen, the only difference between the two is the addition of 0 to the spectrum of the original graph in new spectrum.

3.2.2 Cycles. In the case $m = n$, we have G is a cycle C_n. The spectrum of the adjacency matrix of C_n is known to be the $2\cos(2\pi j/n)$ for $j = \{0, ..., n-1\}$ (see [12]). Similarly, the non-backtracking spectrum of C_n can be explicitly calculated. First, we note the following fact about the spectrum of directed cycles:

Lemma 3.7 ([12]). *Let D_n be the directed n-cycle. Then A if is the adjacency matrix of D_n, $\sigma(A)$ consists of the n^{th} roots of unity.*

In Figure 2.1, we see changing each of the edges of C_n to be two directed edges creates two directed cycles D_n. Given the non-backtracking condition, these two directed cycles can be considered disjoint. Using this fact, [46] calculates the non-backtracking spectrum of C_n explicity.

Theorem 3.8 ([46]). *Let C_n be an undirected cycle with n vertices. Let B be the non-*

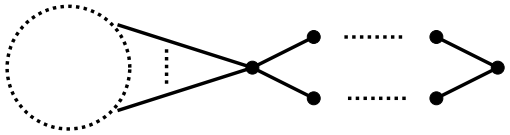


Figure 3.3: A graph with a pendant cycle takes any graph and attaches a cycle of length n to one vertex of the graph.

backtracking matrix of C_n. Then the eigenvalues of B are $e^{2\pi j/n}$ for $j = 0, ..., n-1$ and each eigenvalue has multiplicity 2.

With an understanding of the non-backtracking spectrum of cycles, we can expand to graphs with "pendant" cycles. By this we mean a graph G made from connecting a graph G' at one vertex with a cycle C_n, as in Figure 3.3. We can guarantee the non-backtracking spectra of graphs with pendant cycles contain eigenvalues from the non-backtracking spectrum of C_n.

Theorem 3.9. *Let G' be a graph with m edges and let C_n be a cycle. Let G be a graph created by joining G' and C_n at exactly one vertex. Let B be the non-backtracking matrix of G. Then $e^{2\pi j/n} \in \sigma(B)$ for $j = 0, ..., n-1$.*

Proof. We construct B such that the first $2m$ rows and columns correspond to the edges in the graph G' and the last $2n$ rows and columns correspond to the edges in C_n. We know the non-backtracking matrix of a cycle can be written as a block diagonal matrix. So

$$B = \begin{pmatrix} B' & Q & Q \\ * & D_n & 0 \\ * & 0 & D_n \end{pmatrix}$$

where B' is the non-backtracking matrix of G' and the bottom-right 2×2 block matrix is the non-backtracking matrix of C_n. Let r be the vertex joining G' and C_n, where d_r is the degree of r in the graph G'. Thus there are d_r rows in the first $2m$ rows of B which have nonzero entries in Q. In fact, we can define Q as a block of zeros with one column containing non-zero entries. This column will be identical in the block above each D_n block,

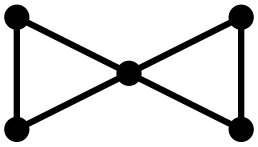


Figure 3.4: An example of a pinwheel graph with two 3-cycles connected at one vertex.

as each edge pointing towards the cycle points to each of the two directed cycles formed on the directed edges. We define a vector

$$\mathbf{x}_j = \begin{pmatrix} 0 & \cdots & 0 & \mathbf{y}_j & -\mathbf{y}_j \end{pmatrix}^T$$

where the first $2m$ entries are 0, and $\mathbf{y}_j = \begin{pmatrix} 1 & e^{2\pi ji/n} & e^{2(2)\pi ji/n} & \cdots & e^{2(n-1)\pi ji/n} \end{pmatrix}^T$. We then we recall $D_n\mathbf{y}_j = e^{2\pi ij/n}\mathbf{y}_j$. Thus we see

$$B\mathbf{x}_j = \begin{pmatrix} B' & Q & Q \\ * & D_n & 0 \\ * & 0 & D_n \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{y}_j \\ -\mathbf{y}_j \end{pmatrix} = \begin{pmatrix} Q\mathbf{y}_j - Q\mathbf{y}_j \\ D_n\mathbf{y}_j \\ -D_n\mathbf{y}_j \end{pmatrix} = \begin{pmatrix} 0 \\ e^{2\pi ij/n}\mathbf{y}_j \\ -e^{2\pi ij/n}\mathbf{y}_j \end{pmatrix} = e^{2\pi ij/n}\mathbf{x}_j.$$

So $e^{2\pi ij/n} \in \sigma(B)$. □

It is often interesting to examine the non-backtracking spectrum of a particular family of graphs. We define a specific family of graphs called pinwheel graphs. A pinwheel graph is made by connected p cycles of the same length onto one vertex (see Figure 3.4). We classify the non-backtracking spectrum of pinwheel graphs.

Corollary 3.10. *Let G be a pinwheel graph made by connecting p cycles of length k at one vertex. Let B be its non-backtracking matrix. Then $e^{2\pi ij/k} \in \sigma(B)$ with multiplicity p for $j \in \{0, ..., k-1\}$, $e^{2\pi ij/(2k)}$ with multiplicity $p-1$ for all $j \in \{1, 3, 5, ..., 2k-1\}$, and the k^{th} complex roots of $2p-1$ are in $\sigma(B)$ with multiplicity 1. These capture the entirety of $\sigma(B)$ and the spectral radius of B is $|(2p-1)^{1/k}|$.*

Proof. We can write B as a block diagonal matrix

$$B = \begin{pmatrix} C_k & 0 & R & R & \cdots & R & R \\ 0 & C_k & R & R & \cdots & R & R \\ R & R & C_k & 0 & \cdots & R & R \\ R & R & 0 & C_k & \cdots & R & R \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ R & R & R & R & \cdots & C_k & 0 \\ R & R & R & R & \cdots & 0 & C_k \end{pmatrix}$$

where each block is $p \times p$ and R is a block of all zeros, with the exception of the last row, which has a 1 in the first entry of the row and zeroes everywhere else. By straightforward computation, we can see

$$\begin{pmatrix} \mathbf{y} & -\mathbf{y} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}^T$$

$$\begin{pmatrix} 0 & 0 & \mathbf{y} & -\mathbf{y} & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}^T$$

$$\vdots$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \mathbf{y} & -\mathbf{y} \end{pmatrix}^T$$

are all eigenvectors with eigenvalue $e^{2\pi ij/k}$ where $C_k\mathbf{y} = e^{2\pi ij/k}\mathbf{y}$. Thus, each root of unity $e^{2\pi ij/k}$ has at least algebraic multiplicity p.

Let $C_{2k}\mathbf{z} = e^{2\pi ij/2k}\mathbf{z}$ for $j \in \{1, 3, 5, ..., 2k-1\}$ and $\mathbf{z} = \begin{pmatrix} z_1 & z_2 \end{pmatrix}^T$ where z_i is the i^{th} half of the eigenvector $\mathbf{z}$. Then $z_1 = \begin{pmatrix} \alpha_1 & \cdots & \alpha_k \end{pmatrix}^T$ and $z_2 = \begin{pmatrix} \beta_1 & \cdots & \beta_k \end{pmatrix}^T$. Since C_{2k} is a circulant matrix, then $\alpha_l = e^{2\pi ij/2k}\alpha_{l+1}$ for all $l \in \{1, ..., k-1\}$, $\beta_l = e^{2\pi ij/2k}\beta_{l+1}$ for all

$l \in \{1, ..., k-1\}$, $\alpha_k = e^{2\pi ij/2k}\beta_1$, and $\beta_k = e^{2\pi ij/2k}\alpha_1$. Lastly define the vector

$$\mathbf{v}_1 = \begin{pmatrix} (p-1)z_1 & -(p-1)z_2 & -z_1 & z_2 & -z_1 & z_2 & \cdots & -z_1 & z_2 \end{pmatrix}^T.$$

Then we see through direct calculation

$$B\mathbf{v}_1 = \begin{pmatrix} -(p-1)\mathbf{r} & \mathbf{r} & \cdots & \mathbf{r} \end{pmatrix} = e^{2\pi ij/2k}\mathbf{v}_1$$

where $\mathbf{r} = \begin{pmatrix} \mathbf{r}_1 & \mathbf{r}_2 \end{pmatrix}^T$, $\mathbf{r}_1 = \begin{pmatrix} -\alpha_2 & \cdots & -\alpha_k & -\beta_1 \end{pmatrix}^T$ and $\mathbf{r}_2 = \begin{pmatrix} \beta_2 & \cdots & \beta_k & \alpha_1 \end{pmatrix}^T$. So $\mathbf{v}_1$ is an eigenvector of B with associated eigenvalue $e^{2\pi ij/2k}$. Similarly, we define

$$\mathbf{v}_2 = \begin{pmatrix} -z_1 & z_2 & (p-1)z_1 & -(p-1)z_2 & -z_1 & z_2 & \cdots & -z_1 & z_2 \end{pmatrix}$$

and so on for $\mathbf{v}_i$ where $i \in \{1, ..., p\}$. By a similar calculation, we see $B\mathbf{v}_i = e^{2\pi ij/2k}\mathbf{v}_i$ for all i. We now show the set $\{\mathbf{v}_1, ..., \mathbf{v}_{p-1}\}$ is linearly independent. Let $0 = \sum_{i=1}^{n-1} \gamma_i \mathbf{v}_i$. Note the last two entries of this summation give $\sum_{i=1}^{p-1} \gamma_i z_1 = 0$ and $\sum_{i=1}^{p-1} \gamma_i z_2 = 0$. Then the first entry gives $\gamma_1(p-1)z_1 = \sum_{i=2}^{p-1} \gamma_i z_1$. From the second to last entry, this must mean $\gamma_1(p-1)z_1 = -\gamma_1 z_1$. Since $z_1 \neq 0$ and $p-1$ is fixed, then $\gamma_1 = 0$. The third entry then gives $\gamma_2(p-1)z_1 = \sum_{i=3}^{p-1} \gamma_i z_1$. Again, by the second to last entry, we get $\gamma_2(p-1)z_1 = -\gamma_2 z_1$. So $\gamma_2 = 0$. Continuing through every odd entry in order of $\sum_{i=1}^{p-1} \gamma_i \mathbf{v}_i$, we get $\gamma_i = 0$ for all $i = \{1, ..., p-1\}$. Thus, the set $\{\mathbf{v}_1, \cdots, \mathbf{v}_{p-1}\}$ must be linearly independent. Thus, $e^{2\pi ij/2k}$ must have algebraic multiplicity at least $p-1$.

We now let $\lambda = (2p-1)^{1/k}$ be one of the k^{th} complex roots of $2p-1$. Then we define $\mathbf{w} = \begin{pmatrix} \lambda & \lambda^2 & \cdots \lambda^k & \lambda & \cdots & \lambda^k & \cdots & \lambda & \cdots & \lambda^k \end{pmatrix}$ where $\mathbf{w} \in \mathbb{C}^{2pk}$. Then we see

$$B\mathbf{w} = \begin{pmatrix} \lambda^2 & \cdots & \lambda^{k+1} & \lambda^2 & \cdots & \lambda^{k+1} & \cdots & \lambda^2 & \cdots & \lambda^{k+1} \end{pmatrix} = \lambda\mathbf{w}.$$

Since there are k complex roots of $2p-1$, there are k unique eigenvalues with at least

algebraic multiplicity 1. Further since $\mathbf{w}$ is positive and B is nonnegative and irreducible by Propostion 3.3, then $|(2p-1)^{1/k}|$ must be the spectral radius of B.

To summarize, we have the following eigenvalues: $e^{2\pi ij/k}$ for $j \in \{0, ..., k-1\}$ with multiplicity at least p, $e^{2\pi il/2k}$ for $l \in \{1, 3, ..., 2k-1\}$ with multiplicity at least $p-1$, and the k roots $(2p-1)^{1/k}$ with multiplicity at least 1. We then see $pk + (p-1)k + k = 2pk$. Since $B \in M_{2pk}$, then all the multiplicities mentioned must be exact. □

Remark. The non-backtracking spectrum of other families of graphs have previously been studied. See [11, 24, 42] for some examples.

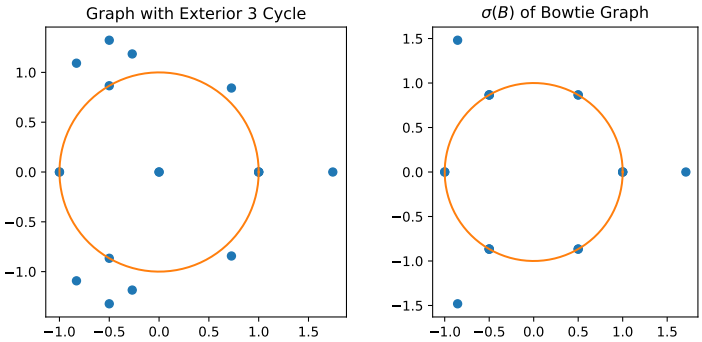


Figure 3.5: The graph on the left shows the spectrum of a random GNP graph with 5 vertices with an exterior 3 cycle. The graph on the right shows the spectrum of a pinwheel graph made from 3 cycles of length 3.

3.3 The K Matrix

The eigenvalues of B coming from Equation 3.1 in Ihara's Theorem can be found independent of B. We define

$$K = \begin{pmatrix} A & D-I \\ -I & 0 \end{pmatrix}$$

which has characteristic polynomial $\det(\mu^2 I - \mu A + (D-I))$. Using K, we can create a decomposition of B which clearly organizes the spectrum of B and gives an alternate proof of Ihara's Theorem.

Remark. Lubetzky and Peres showed the matrix B is unitarily similar to a block diagonal matrix which clearly shows the spectrum of B as well as the eigenvectors.

Theorem 3.11 ([29]). *Let G be a connected d-regular graph ($d \geq 3$) on n vertices. Let $N = dn$ and $\lambda_i \in \sigma(A)$, with $\lambda_1 = d$. Then the operator B is unitarily similar to*

$$\Lambda = diag\left(d-1, \begin{pmatrix} \theta_2 & \alpha_2 \\ 0 & \theta_2' \end{pmatrix}, ..., \begin{pmatrix} \theta_n & \alpha_n \\ 0 & \theta_n' \end{pmatrix}, -1, ..., -1, 1, ..., 1\right)$$

where $|\alpha_i| < 2(d-1)$ for i, θ_i and θ_i' are defined as the solutions of

$$\theta^2 - \lambda_i \theta + d - 1 = 0$$

and -1 has multiplicity $N/2 - n$ and 1 has multiplicity $N/2 - n + 1$.

While we are unable to make such a similarity for a general graph G, we will decompose B in a similar manner into a block diagonal matrix to see the eigenvalues more clearly.

Remark. Torres succeeded in diagonalizing B if the eigenvalues $\mu \in \sigma(B)$ such that $|\mu| > 1$ are simple. We will not use this requirement in our proof.

Before creating our decomposition, we also need to understand the eigenvectors of B for eigenvalues ± 1. Lubetzky and Peres [29] show these eigenvectors come from $\mathscr{E}_{-1} \cap \text{Null(ST)}$ and $\mathscr{E}_1 \cap \text{Null}(ST)$ respectively, where $\mathscr{E}_i$ is the eigenspace of τ corresponding to eigenvalue i. They further show $\dim(\mathscr{E}_{-1} \cap \text{Null(ST)}) = m - n + 1$ and $\dim(\mathscr{E}_1 \cap \text{Null(ST)}) = m - n$ or $m - n + 1$ if G is bipartite. Additionally, we need to relate B and K. Using Equation 2.3 we get

$$B \begin{pmatrix} S & T^T \end{pmatrix} = \begin{pmatrix} S & T^T \end{pmatrix} K. \tag{3.3}$$

With this we create our decomposition.

Theorem 3.12. *Let G be a connected graph and B its non-backtracking matrix. Let $R \in M_{2m \times 2(m-n)}$ where the columns of R are linearly independent and the first $m - n$ columns of*

R are taken from $\mathscr{E}_{-1} \cap Null(ST)$ and the rest are taken from $\mathscr{E}_1 \cap Null(ST)$. Then

$$BX = X \begin{pmatrix} K & 0 & 0 \\ 0 & I_{m-n} & 0 \\ 0 & 0 & -I_{m-n} \end{pmatrix}$$

and $X = \begin{pmatrix} S & T^T & R \end{pmatrix}$.

Proof. This follows directly from matrix multiplication and the properties of the columns of R. □

Remark. It clearly follows from the previous theorem that $\det(\mu I - B)$ is just the characteristic polynomial of this block diagonal decomposition. Hence, Ihara's Theorem is an immediate corollary of the above theorem. Even more, Theorem 3.12 does not just give the eigenvalues of B but also the eigenvectors. Let $\mathbf{x}$ be an eigenvector of K. Then

$$BX \begin{pmatrix} \mathbf{x} \\ 0 \\ 0 \end{pmatrix} = X \begin{pmatrix} K & 0 & 0 \\ 0 & I_{m-n} & 0 \\ 0 & 0 & -I_{m-n} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 0 \\ 0 \end{pmatrix} = X \begin{pmatrix} K\mathbf{x} \\ 0 \\ 0 \end{pmatrix} = \mu X \begin{pmatrix} \mathbf{x} \\ 0 \\ 0 \end{pmatrix}.$$

Hence we know all the eigenvectors of B associated with eigenvalues of K are of the form $X \begin{pmatrix} \mathbf{x} & 0 & 0 \end{pmatrix}^T$. Additionally, from the construction of R, we see the eigenvectors for ± 1 are $\mathbf{y} \in \mathscr{E}_{-1} \cap \mathrm{Null}(ST)$ and $\mathbf{z} \in \mathscr{E}_1 \cap \mathrm{Null}(ST)$ respectively.

3.4 Properties of $\sigma(B)$ using $\sigma(K)$

With a decomposition of B in terms of K, we now want to better understand the eigenvalues of K. Our goal in this section is to use properties of K in order to place bounds on the eigenvalues of B. Using the Perron-Frobenius theorem, Kotani and Sunada show the following simple bound.

Proposition 3.13 ([23])**.** *Let G be a connected graph with B the non-backtracking matrix and $d_{\max}$ the maximum of degree of G. Then $\rho(B) \leq d_{\max} - 1$ with equality if and only if G is regular.*

We will show a stronger bound using the matrix K and its relationship with B. Immediately from the structure of K, we can learn some information about its eigenvalue-eigenvector pairs.

Proposition 3.14. *Let G be a graph and K as defined above. Then the following are true:*

(i) Every eigenvector of K is of the form $\begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}^T$ where $\mu \in \sigma(K)$,

(ii) $1 \in \sigma(K)$ with geometric multiplicity equal to the number of connected components of G,

(iii) the nullity of K is the number of degree 1 vertices, and

(iv) K is invertible with inverse $K^{-1} = \begin{pmatrix} 0 & -I \\ (D-I)^{-1} & (D-I)^{-1}A \end{pmatrix}$ if and only if $d_{\min} \geq 2$.

Proof. (i) Assume $\mu \in \sigma(K)$ with eigenvector $\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}^T$. Then

$$\mu \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} A & D-I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} A\mathbf{x} + (D-I)\mathbf{y} \\ -\mathbf{x} \end{pmatrix}.$$

Then bottom block gives $\mathbf{x} = -\mu\mathbf{y}$, so the eigenvector must be of the form $\begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}$.

(ii) Let $\mathbf{x} = \begin{pmatrix} \mathbf{1} & -\mathbf{1} \end{pmatrix}$ where $\mathbf{1} \in M_{n\times 1}$ is an all-ones vector. Then

$$K\mathbf{x} = \begin{pmatrix} A\mathbf{1} - (D-I)\mathbf{1} \\ -\mathbf{1} \end{pmatrix} = \begin{pmatrix} \mathbf{k} - \mathbf{k} + \mathbf{1} \\ -\mathbf{1} \end{pmatrix} = \mathbf{x}$$

where $\mathbf{k}$ is the degree vector of G. Thus, $K\mathbf{x} = \mathbf{x}$ and $1 \in \sigma(K)$.

Let $(1, \begin{pmatrix} -\mathbf{y} & \mathbf{y} \end{pmatrix}^T)$ be a general eigenvalue-eigenvector pair for $1 \in \sigma(K)$. Then $\mathbf{y} - A\mathbf{y} + (D-I)\mathbf{y} = 0$. Rearranging we see $(D-A)\mathbf{y} = 0$. So $\mathbf{y} \in \text{Null}(L)$ where L is the Laplacian of G. So $\text{mult}(1) \leq \text{Nullity}(L)$.

Now assume $\mathbf{z} \in \text{Null}(L)$. Then $(D-A)\mathbf{z} = 0$. So $\mathbf{z} - A\mathbf{z} + (D-I)\mathbf{z} = 0$. Then we see

$$K \begin{pmatrix} -\mathbf{z} & \mathbf{z} \end{pmatrix}^T = \begin{pmatrix} -A\mathbf{z} + (D-I)\mathbf{z} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} -\mathbf{z} & \mathbf{z} \end{pmatrix}^T.$$

Thus the geometric multiplicity of $1 \in \sigma(K)$ is the nullity of L. By well-known properties of the Laplacian (see [12]), the geometric multiplicity of $1 \in \sigma(K)$ is the number of connected components in G.

(iii) We have

$$\begin{pmatrix} A & D-I \\ -I & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

if and only if

$$\begin{pmatrix} (D-I)\mathbf{y} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

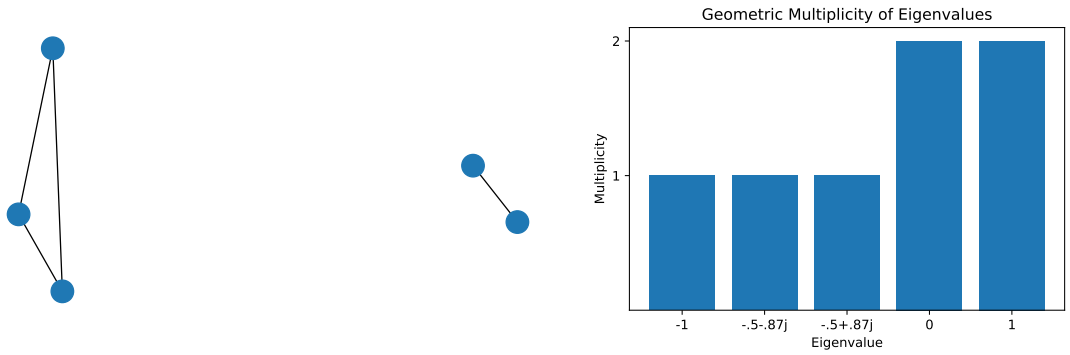
Thus $\begin{pmatrix} 0 & \mathbf{y} \end{pmatrix}^T$ is in the nullspace of K if and only if $(D-I)\mathbf{y} = 0$. Note $D-I$ is diagonal, and so the dimension of its nullspace is equal to the number of diagonal entries which are 0. From this, the result follows.

(iv) Immediately from (iii) we get K is invertible if and only if $d_{\min} \geq 2$. If $d_{\min} \geq 2$, then

$$\begin{pmatrix} A & D-I \\ -I & 0 \end{pmatrix} \begin{pmatrix} 0 & -I \\ (D-I)^{-1} & (D-I)^{-1}A \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$
$$\begin{pmatrix} 0 & -I \\ (D-I)^{-1} & (D-I)^{-1}A \end{pmatrix} \begin{pmatrix} A & D-I \\ -I & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

So $K^{-1} = \begin{pmatrix} 0 & -I \\ (D-I)^{-1} & (D-I)^{-1}A \end{pmatrix}$. □

Remark. Proposition 3.14(iii) can also be proved using the relationship between B and K and results about the invertibility of B found in [46].



(a) Graph (b) Geometric Multiplicity

Figure 3.6: The plot on the right shows the geometric multiplicity of eigenvalues in the spectrum of K of the graph on the left. As can be seen, $1 \in \sigma(K)$ with multiplicity 2 as the graph has two components. Similarly, $0 \in \sigma(K)$ with multiplicity 2 as there are two degree one vertices.

With Proposition 3.14(i), we can create a relationship between the eigenvalues of A and the eigenvalues of B.

Proposition 3.15. *Let $\mu \in \sigma(K)$ with eigenvector $\begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}^T$ and let $\lambda \in \sigma(A)$ such that $A\mathbf{x} = \lambda\mathbf{x}$. If $\mathbf{x}^T\mathbf{y} \neq 0$, then*

$$\mu = \frac{\lambda \pm \sqrt{\lambda^2 - 4\mathbf{x}^T(D-I)\mathbf{y}}}{2}.$$

Proof. Recall $\mu^2\mathbf{y} - \mu A\mathbf{y} + (D-I)\mathbf{y} = 0$ where $\mathbf{y}$ is the second half of the eigenvector $\begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}^T$ of K corresponding to eigenvalue μ. Let $\mathbf{x}^T$ be an eigenvector of A with associated eigenvalue λ, where $\mathbf{x}$ and $\mathbf{y}$ are not orthogonal, and scale $\mathbf{y}$ such that $\mathbf{x}^T\mathbf{y} = 1$. Then left multiplying by $\mathbf{x}^T$ gives $\mu^2 - \mu\lambda + \mathbf{x}^T(D-I)\mathbf{y} = 0$. With the quadratic formula, we know

$$\mu = \frac{\lambda \pm \sqrt{\lambda^2 - 4\mathbf{x}^T(D-I)\mathbf{y}}}{2}.$$ □

Remark. Note this formula does not necessarily give every eigenvalue μ. If K is not diagonalizable (for example, K of any cycle), then there will exist $\mu \in \sigma(K)$ with no corresponding eigenvector $\begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}^T$. Additionally we are not guaranteed that $\mathbf{x}^T\mathbf{y} \neq 0$ for all pairs $(\mathbf{x}, \mathbf{y})$. In fact many $\mathbf{x}$ and $\mathbf{y}$ exist such that $\mathbf{x}^T\mathbf{y} = 0$. However the eigenvectors of A form a basis of $\mathbf{R}^n$, so for each $\mu \in \sigma(K)$ with eigenvector $\begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}^T$ there exists some i such that $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ where $\mathbf{x}_i^T\mathbf{y} \neq 0$.

In the case of regular graphs, Theorem 3.15 simplifies substantially since $D - I = (d-1)I$ where d is the degree of the graph. Exploiting this fact leads to an alternative proof of Theorem 3.2. In doing so, we will also show that K is diagonalizable when G is d-regular.

Proposition 3.16. *Let G be a d-regular graph with $d \geq 2$. Then $\sigma(A) = \{f(\mu) \colon \mu \in \sigma(K)\}$ where $f(x) = x + \frac{1}{x}(d-1)$.*

Proof. We first show K is diagonalizable when G is d-regular. Assume $d \geq 2$ for a given graph G. Let $\mathbf{x}$ be an eigenvector of A with associated eigenvalue λ. Define $\mu_1 = \frac{\lambda + \sqrt{\lambda^2 - 4(d-1)}}{2}$ and $\mu_2 = \frac{\lambda - \sqrt{\lambda^2 - 4(d-1)}}{2}$. This implies $\mu_1^2 - \mu_1\lambda + (d-1) = 0$ and $\mu_2^2 - \mu_2\lambda + (d-1) = 0$. Define the vector $\mathbf{v}_i = \begin{pmatrix} -\mu_i x & x \end{pmatrix}$. Then we see

$$K\mathbf{v}_i = \begin{pmatrix} A & (d-1)I \\ -I & 0 \end{pmatrix} \begin{pmatrix} -\mu_i\mathbf{x} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} -\mu_i\lambda\mathbf{x} + (d-1)\mathbf{x} \\ \mu_i\mathbf{x} \end{pmatrix} = \begin{pmatrix} -\mu_i^2\mathbf{x} \\ \mu_i\mathbf{x} \end{pmatrix} = \mu_i\mathbf{v}_i.$$

Thus, both $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors. Since $\mu_1 \neq \mu_2$, we know $\mathbf{v}_1 \neq \mathbf{v}_2$. Further, we know A is diagonalizable. Thus, μ_i has the same algebraic and geometric multiplicity. Lastly, there are $2n$ distinct μ_i since each eigenvector $\mathbf{x}$ of A creates two unique eigenvalue-eigenvector pairs. Since $K \in M_{2n}$, K is diagonalizable.

Now that we know K is diagonalizable, define $f(x) = x + \frac{1}{x}(d-1)$. Let $\mu \in \sigma(K)$ be an eigenvalue of K. Recall all eigenvectors of K can be written as $\begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}^T$, implying $-\mu A\mathbf{y} + (d-1)\mathbf{y} = -\mu^2\mathbf{y}$. Also note $d \geq 2$, so by Proposition 3.14(iii) $\mu \neq 0$ and $A\mathbf{y} = (\mu + \frac{1}{\mu}(d-1))\mathbf{y}$. So there exists some $\lambda \in \sigma(A)$ such that $\mu + \frac{1}{\mu}(d-1) = \lambda$. Thus, $\{f(\mu) \colon \mu \in \sigma(K)\} \subset \sigma(A)$.

We also know there are two solutions to the equation $\mu + \frac{1}{\mu}(d-1) = \lambda$:

$$\mu = \frac{\lambda \pm \sqrt{\lambda^2 - 4(d-1)}}{2}.$$

Let μ_1, μ_2 be the two solutions to this equation where μ_1 and μ_2 are the plus and minus solutions respectively. Assume $\begin{pmatrix} -\mu_1\mathbf{y} & \mathbf{y} \end{pmatrix}^T$ is an eigenvector of K. Then we know $A\mathbf{y} = (\mu_1 + \frac{1}{\mu_1}(d-1))\mathbf{y}$ and hence $A\mathbf{y} = (\mu_2 + \frac{1}{\mu_2}(d-1))\mathbf{y}$. Rearranging we get $-\mu_2^2\mathbf{y} = -\mu_2 A\mathbf{y} + (d-1)\mathbf{y}$. So, $K\begin{pmatrix} -\mu_2\mathbf{y} & \mathbf{y} \end{pmatrix}^T = \mu_2\begin{pmatrix} -\mu_2\mathbf{y} & \mathbf{y} \end{pmatrix}^T$. Similarly, if $\begin{pmatrix} -\mu_2\mathbf{y} & \mathbf{y} \end{pmatrix}^T$ is an eigenvector of K with eigenvalue μ_2, then $\begin{pmatrix} -\mu_1\mathbf{y} & \mathbf{y} \end{pmatrix}^T$ is an eigenvector of K. Since K is diagonalizable, then μ_1 and μ_2 have the same multiplicity. Let $\text{mult}(\mu_i)$ be the algebraic multiplicity of μ_i. Order the eigenvalues of K such that if μ and $\hat{\mu}$ are both solutions of $\lambda = \mu + \frac{1}{\mu}(d-1)$, then $\mu = \mu_i$ and $\hat{\mu} = \mu_{n+i}$. Then we know $2n = \sum_{i=1}^{2n} \text{mult}(\mu_i) = \sum_{i=1}^{n} 2(\text{mult}(\mu_i)) = 2\sum_{i=1}^{n} \text{mult}(\mu_i)$. So $n = \sum_{i=1}^{n} \text{mult}(\mu_i)$. Then we know $|\{f(\mu) \colon \mu \in \sigma(K)\}| = \sum_{i=1}^{n} \text{mult}(\mu_i) = n$. Since $|\sigma(A)| = n$ and $\{f(\mu) \colon \mu \in \sigma(K)\} \subseteq \sigma(A)$, $\sigma(A) = \{f(\mu) \colon \mu \in \sigma(K)\}$. □

Our goal now is to use Proposition 3.15 to bound the spectrum of B. We begin by bounding the "bottom" of the spectrum using a formula similar to that of Proposition 3.15 but not quite equal.

Theorem 3.17. *Let G be a connected graph where $d_{\min} \geq 2$. Then $|\mu| \geq 1$ for all $\mu \in \sigma(B)$.*

Proof. Recall any $\mu \neq \pm 1$ satisfies the equation $\mu^2 \mathbf{y} - \mu A\mathbf{y} + (D-I)\mathbf{y} = 0$ for some vector $\mathbf{y} \neq 0$. Assume $\|\mathbf{y}\|_2 = 1$. Then left multiplying by $\mathbf{y}^T$ gives $\mu^2 - \mu \mathbf{y}^T A\mathbf{y} + \mathbf{y}^T(D-I)\mathbf{y} = 0$. The quadratic equation then gives

$$\mu = \frac{\mathbf{y}^T A\mathbf{y} \pm \sqrt{(\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T(D-I)\mathbf{y}}}{2}.$$

We work by cases:

(i) Assume $(\mathbf{y}^T A\mathbf{y})^2 < 4\mathbf{y}^T(D-I)\mathbf{y}$. Then

$$|\mu| = \frac{(\mathbf{y}^T A\mathbf{y})^2}{4} + \mathbf{y}^T(D-I)\mathbf{y} - \frac{(\mathbf{y}^T A\mathbf{y})^2}{4} = \mathbf{y}^T D\mathbf{y} - 1 \geq 2 - 1 = 1.$$

(ii) Assume $(\mathbf{y}^T A\mathbf{y})^2 = 4\mathbf{y}^T(D-I)\mathbf{y}$. Then

$$|\mu| = \left|\frac{\mathbf{y}^T A\mathbf{y}}{2}\right| = \left|\sqrt{\mathbf{y}^T(D-I)\mathbf{y}}\right| \geq \sqrt{2} \geq 1.$$

(iii) Assume $(\mathbf{y}^T A\mathbf{y})^2 > 4\mathbf{y}^T(D-I)\mathbf{y}$. If $\mu = \frac{\mathbf{y}^T A\mathbf{y} + \sqrt{(\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T(D-I)\mathbf{y}}}{2}$, then

$$|\mu| = \left|\frac{\mathbf{y}^T A\mathbf{y} + \sqrt{(\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T(D-I)\mathbf{y}}}{2}\right| \geq \left|\frac{\mathbf{y}^T A\mathbf{y}}{2}\right| > |\sqrt{\mathbf{y}^T(D-I)\mathbf{y}}| \geq 1.$$

Assume $\mu = \frac{\mathbf{y}^T A\mathbf{y} - \sqrt{(\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T(D-I)\mathbf{y}}}{2}$. Define the function

$$g(x) = \left|\frac{x - \sqrt{x^2 - 4r}}{2}\right|$$

on the domain $(-\infty, -2\sqrt{r}] \cup [2\sqrt{r}, \infty)$ where $r = \mathbf{y}^T(D-I)\mathbf{y}$. This function is clearly non-increasing. We also know $(\mathbf{y}^T A\mathbf{y})^2 > 4\mathbf{y}^T(D-I)\mathbf{y}$. So either $-\mathbf{y}^T A\mathbf{y} > 2\sqrt{r}$ or $\mathbf{y}^T A\mathbf{y} > 2\sqrt{r}$.

First assume $-\mathbf{y}^T A\mathbf{y} > 2\sqrt{r}$. Thus, $\mathbf{y}^T A\mathbf{y} < 0$. Since $r \geq 1$, we know $\mathbf{y}^T A\mathbf{y} < 2\sqrt{r}$. Since g is non-increasing, then $g(\mathbf{y}^T A\mathbf{y}) > g(2\sqrt{r}) \geq 1$. Thus, $|\mu| \geq 1$.

Now assume $\mathbf{y}^T A\mathbf{y} > 2\sqrt{r}$. Note if $g(\mathbf{y}^T A\mathbf{y}) = 1$, then

$$\begin{aligned}
1 &= \frac{\mathbf{y}^T A\mathbf{y} - \sqrt{(\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T(D-I)\mathbf{y}}}{2} \\
\sqrt{(\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T(D-I)\mathbf{y}} &= \mathbf{y}^T A\mathbf{y} - 2 \\
(\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T(D-I)\mathbf{y} &= (\mathbf{y}^T A\mathbf{y})^2 - 4\mathbf{y}^T A\mathbf{y} + 4 \\
\mathbf{y}^T(A - (D-I))\mathbf{y} &= 1 \\
\mathbf{y}^T(D-A)\mathbf{y} &= 0.
\end{aligned}$$

Thus, if $g(\mathbf{y}^T A\mathbf{y}) = 1$, then $\mathbf{y}^T A\mathbf{y} = \mathbf{y}^T D\mathbf{y}$. Further, we know the Laplacian is positive semidefinite [13], so $\mathbf{y}^T D\mathbf{y} \geq \mathbf{y}^T A\mathbf{y}$. Since g is non-increasing, then $g(\mathbf{y}^T A\mathbf{y}) \geq 1$ for all $\mathbf{y}$. So $|\mu| \geq 1$.

Thus, $|\mu| \geq 1$ for all $\mu \in \sigma(B)$. □

Remark. A similar proof of Theorem 3.17 in terms of Zeta functions can be found in [23].

With the fact that the modulus of all $\mu \in \sigma(B)$ are bounded below by one, we turn our attention to the spectral radius. Our first goal is to show the spectral radius of B is strictly greater than 1 when G is not a cycle and has no pendant vertices.

Corollary 3.18. *Let G be a connected graph such that G is not a tree or cycle and $d_{\min} \geq 2$. Then $\rho(K) > 1$.*

Proof. Let B be the non-backtracking matrix. Note $\rho(B) = \rho(K)$ and $\rho(K) \geq 1$ since $\sigma(K) \subset \sigma(B)$ and $1 \in \sigma(K)$. Assume G is not d-regular. Since B is nonnegative and irreducible by Proposition 3.3 and $d_{\min} \geq 2$, then [27] tells us $\rho(B) > \min_i \sum_j b_{ij} \geq 1$. Assume G is d-regular. Since G is not a cycle, $d \geq 3$. Then by Theorem 3.2,

$$\rho(B) = \frac{d + \sqrt{d^2 - 4(d-1)}}{2} \geq \frac{3+1}{2} = 2 > 1.$$

So $\rho(B) > 1$. Since $\rho(B) = \rho(K)$, then $\rho(K) > 1$. $\square$

We now turn our attention to upper bounds on the spectrum of B. To do this, we employ Proposition 3.15. Recall the necessary condition for the proposition was $\mathbf{x}^T\mathbf{y} \neq 0$ for some eigenvalue-eigenvector pairs $(\lambda, \mathbf{x})$ and $(\mu, \begin{pmatrix} -\mu\mathbf{y} & \mathbf{y} \end{pmatrix}^T)$ of A and K respectively. To obtain this condition, we show for the eigenvalue-eigenvector pair $(\rho(K), \begin{pmatrix} -\rho(K)\mathbf{y} & \mathbf{y} \end{pmatrix}^T)$ the vector $\mathbf{y}$ can always be chosen to be positive.

Lemma 3.19. *Let G be a connected graph such that G is not a tree or cycle and $d_{\min} \geq 2$. Then $\mathbf{y}$ can be strictly positive for the eigenvector $\begin{pmatrix} -\rho(K)\mathbf{y} & \mathbf{y} \end{pmatrix}^T$.*

Proof. Let $K \begin{pmatrix} -\rho(K)\mathbf{y} & \mathbf{y} \end{pmatrix}^T = \rho(K) \begin{pmatrix} -\rho(K)\mathbf{y} & \mathbf{y} \end{pmatrix}^T$. Now scale $\mathbf{y}$ such that there exists some $y_k > 0$. By the Perron-Frobenius Theorem and Theorem 3.12, $T^T\mathbf{y} \succ \rho(K)S\mathbf{y}$ or $\rho(K)S\mathbf{y} \succ T^T\mathbf{y}$, where S and T are defined as in equation 2.1. First assume $T^T\mathbf{y} \succ \rho(K)S\mathbf{y}$. From the definitions of T and S, we get $y_i \geq \rho(K)y_j$ for all $i \sim j$. Choose $y_k > 0$. Thus for all i such that $i \sim k$, $\frac{y_i}{y_k} \geq \rho(K)$. So $y_i \neq 0$ for all $i \sim k$ since $\rho(K) > 1$ by Proposition 3.18. Then by similar argument, $\frac{y_k}{y_i} \geq \rho(K)$. This implies $\frac{y_k}{y_i} = 1$ which is a contradiction since $\rho(K) > 1$.

Assume $\rho(K)S\mathbf{y} \succ T^T\mathbf{y}$. Again, choose $y_k > 0$. So for all i such that $i \sim k$, $\rho(K)\frac{y_i}{y_k} > 1$. Since $\rho(K) > 1$, then y_i must be positive. Since G is connected, by induction we get $y_j > 0$ for all vertices j. Thus y is positive. $\square$

Theorem 3.20. *Let G be a connected graph with A the adjacency matrix and B the non-backtracking matrix. If $\rho(A) \geq 2\sqrt{\mathbf{x}^T(D-I)\mathbf{y}}$, then*

$$\rho(B) \leq \frac{\rho(A) + \sqrt{\rho(A)^2 - 4(d_{\min} - 1)}}{2}.$$

Proof. Let G be a tree. Then $\rho(B) = 0$ by Theorem 3.5 and $\rho(B) \leq \frac{\rho(A)+\sqrt{\rho(A)^2-4(d_{\min}-1)}}{2}$.

Now assume G is a cycle. Then $\rho(B) = 1$ by Theorem 3.8 and $\frac{2+\sqrt{4-4}}{2} = 1 \geq \rho(B)$. Now assume G is a cycle with dangling vertices. By Corollary 3.6 $\rho(B)$ is the spectral radius of $G\backslash S$ where S is the set of dangling vertices. Then by the work above the result holds.

Assume G is not a tree or cycle and $d_{\min} \geq 2$. Note $\rho(B) \geq \rho(K)$. By the Perron-Frobenius theorem and Lemma 3.19, $\mathbf{x}^T\mathbf{y} \neq 0$. Thus by Proposition 3.15,

$$\rho(B) \leq \rho(K) = \frac{\rho(A) \pm \sqrt{\rho(A)^2 - 4\mathbf{x}^T(D-I)\mathbf{y}}}{2} \leq \frac{\rho(A) + \sqrt{\rho(A)^2 - 4(d_{\min}-1)}}{2}.$$

Now assume G is not a tree or cycle and has at least one dangling vertex. By Corollary 3.6 $\rho(B)$ is the spectral radius of $G\backslash S$ where S is the set of all dangling vertices. Then by the same argument as above the result holds. □

There are many bounds shown for the spectral radius of the adjacency matrix of a graph [16, 26, 36, 43]. Using a bound provided by Das and Kumar [16], we get a simple bound on the spectral radius of B dependent on the minimum degree $d_{\min}$, number of vertices n, and number of edges m.

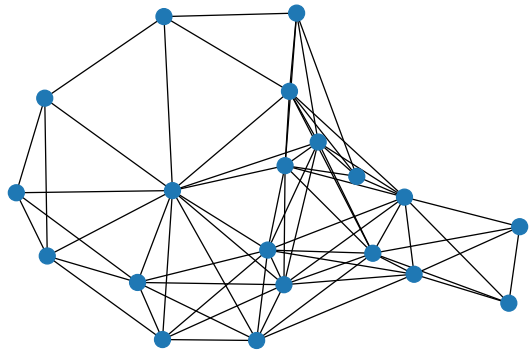
Corollary 3.21. *Let G be a connected graph with A the adjacency matrix and B the non-backtracking matrix. If $\rho(A) \geq 2\sqrt{\mathbf{x}^T(D-I)\mathbf{y}}$, then*

$$\rho(B) \leq \frac{\sqrt{2m-n-1} + \sqrt{2m-n-4d_{\min}+3}}{2}.$$

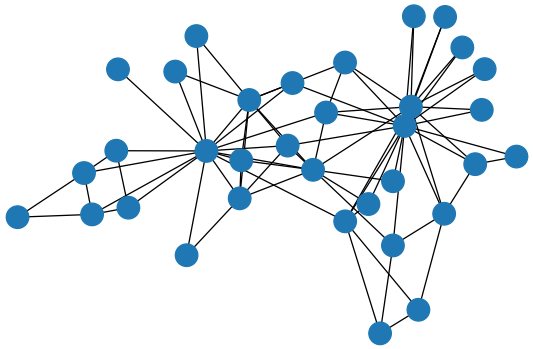
3.4.1 Bipartite Graphs.

A bipartite graph is a graph where all vertices can be divided into two subsets, V_1 and V_2, where vertices in V_1 only connect to vertices in V_2 and vice versa. It is widely known the spectrum of the adjacency matrix can determine the bipartiteness of a graph.

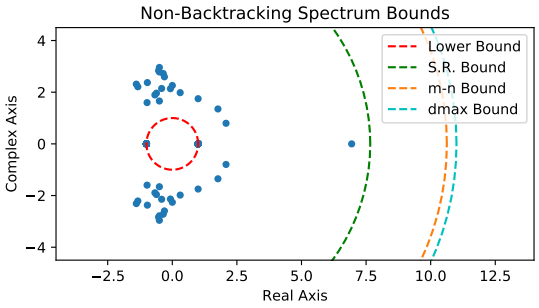
Theorem 3.22 ([12])**.** *Let G be a graph and A its associated adjacency matrix.*



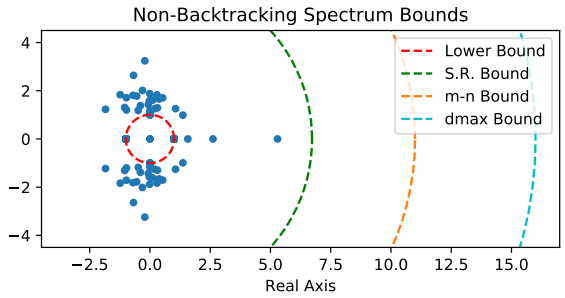
(a) Random Geometric Graph



(b) Karate Club



(c) Upper and Lower Bounds



(d) Upper and Lower Bounds

Figure 3.7: The plots in (c) and (d) describe the non-backtracking spectrum of the graphs in (a) and (b) respectively with the upper and lower bounds described above (red - Theorem 3.17, green - Theorem 3.20, orange - Proposition 3.13, blue - Corollary 3.21). The graph in (a) is a random geometric graph with 20 vertices and radius 0.4. The graph in (b) is the Zachary Karate Club network [39].

(i) G is bipartite if and only if, for each eigenvalue $\lambda \in \sigma(A)$, $-\lambda \in \sigma(A)$ with the same multiplicity ($\sigma(A)$ is symmetric).

(ii) If G is connected and λ_1 is the largest eigenvalue of A, then G is bipartite if and only if $-\lambda_1$ is an eigenvalue of A.

The spectrum of the non-backtracking matrix B can also indicate whether a graph is bipartite in the same way. Additionally, the same properties hold for the spectrum of K when G is bipartite.

Theorem 3.23. *Let G be a connected graph, B its associated non-backtracking matrix, and K defined as above. Then the following are equivalent:*

(i) G is a bipartite graph,

(ii) $\sigma(K)$ is symmetric,

(iii) $\sigma(B)$ is symmetric,

(iv) $\lambda_n = -\lambda_1$ where $\lambda_i \in \sigma(K)$, and

(v) $\mu_n = -\mu_1$ where $\mu_i \in \sigma(B)$,

(vi) $-1 \in \sigma(K)$.

Proof. $(i) \to (ii)$: Assume G is bipartite. If G is bipartite, then the adjacency matrix A can be written as $\begin{pmatrix} 0 & A_2 \\ A_1 & 0 \end{pmatrix}$ (see [12]). We know $A = TS$ by equation 2.3. Thus for a bipartite graph, we define the following matrices:

$$T_1 \colon = \begin{cases} 1 & i_1 \mapsto (i_1, i_2) \\ 0 & \text{otherwise} \end{cases}, \qquad T_2 \colon = \begin{cases} 1 & i_2 \mapsto (i_2, i_1) \\ 0 & \text{otherwise} \end{cases},$$

$$S_1 \colon = \begin{cases} 1 & (i_2, i_1) \mapsto i_1 \\ 0 & \text{otherwise} \end{cases}, \qquad S_2 \colon = \begin{cases} 1 & (i_1, i_2) \mapsto i_2 \\ 0 & \text{otherwise} \end{cases}.$$

In these matrices, i_j represents a vertex in partition j and (i_j, i_k) represents an edge from partition j to partition k. Thus by simple computation, we see $A = \begin{pmatrix} 0 & T_1S_2 \\ T_2S_1 & 0 \end{pmatrix}$ and the matrix $C = \begin{pmatrix} 0 & S_2T_2 \\ S_1T_1 & 0 \end{pmatrix}$. Hence, the edges are also divided into two edge partitions.

In order to compute B, we also define a matrix

$$\tau_1 \colon = \begin{cases} 1 & (w, x)_1 \mapsto (y, z)_2 \\ 0 & \text{otherwise} \end{cases} \text{ and } \tau_2 \colon = \begin{cases} 1 & (w, x)_2 \mapsto (y, z)_1 \\ 0 & \text{otherwise} \end{cases},$$

where $(w,x)_j$ represents an edge in edge partition j. We then see $\tau = \begin{pmatrix} 0 & \tau_2 \\ \tau_1 & 0 \end{pmatrix}$. So the matrix $B = \begin{pmatrix} 0 & S_2T_2 - \tau_2 \\ S_1T_1 - \tau_1 & 0 \end{pmatrix}$. Defining $B_j = S_jT_j - \tau_j$, we get $B = \begin{pmatrix} 0 & B_2 \\ B_1 & 0 \end{pmatrix}$.

Let $\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}^T$ be an eigenvector of B with corresponding eigenvalue μ. Then

$$B \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}^T = \mu \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}^T$$
$$\begin{pmatrix} B_2\mathbf{y} & B_1\mathbf{x} \end{pmatrix}^T = \mu \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}^T.$$

Consider the vector $\begin{pmatrix} \mathbf{x} & -\mathbf{y} \end{pmatrix}^T$. We see

$$B \begin{pmatrix} \mathbf{x} & -\mathbf{y} \end{pmatrix}^T = \begin{pmatrix} -B_2\mathbf{y} & B_1\mathbf{x} \end{pmatrix}^T = \begin{pmatrix} -\mu\mathbf{x} & \mu\mathbf{y} \end{pmatrix}^T = -\mu \begin{pmatrix} \mathbf{x} & -\mathbf{y} \end{pmatrix}^T.$$

So $-\mu$ is an eigenvalue of B with eigenvector $\begin{pmatrix} \mathbf{x} & -\mathbf{y} \end{pmatrix}^T$. Hence the spectrum of B is symmetric around 0.

$(ii) \to (iii)$: Recall $\sigma(B) = \sigma(K) \cup \{\pm 1\}$. Since the set of ± 1 is symmetric by Ihara's theorem, $\sigma(K)$ is symmetric.

$(iii) \to (iv)$: Assume $\sigma(K)$ is symmetric. Then $\lambda_n = -\lambda_1$ where λ_1 is the spectral radius of $\sigma(K)$.

$(iv) \to (v)$: If G is a tree, then the result follows trivially from Theorem 3.5.

If G is not a tree then $\sigma(K) \subseteq \sigma(B)$. Since $1 \in \sigma(K)$, we know $\rho(K) \geq 1$. Thus, $\rho(B) \geq 1$. If $\rho(B) = 1$, then the dominant eigenvalue of B must be $\mu_1 = \pm 1$. From Ihara's Theorem, $\pm 1 \in \sigma(B)$.

If $\rho(B) \neq 1$, then $\mu_1 \in \sigma(K)$ by Ihara's Theorem and $\mu_1 = \lambda_1$. Thus, $\lambda_n = -\mu_1$. Then, we know $\mu_n = -\mu_1$.

$(v) \to (i)$: If G is tree or cycle, the result holds from Theorems 3.5 and 3.8 respectively.

Assume G is not a tree or cycle and $\mu_n = -\mu_1$. Then $|\mu_n| = |\mu_1|$. By Proposition 3.3 tells us B is irreducible. Then from the Perron-Frobenius theorem, we know the period d of B must be at least 2. Further, since $\mu_n = -\mu_1$, $d = 2k$ for some $k \in \{1, 2, 3, ...\}$. Then every cycle must be of even length, so B is the adjacency matrix of a directed bipartite graph. So $B = \begin{pmatrix} 0 & B_2 \\ B_1 & 0 \end{pmatrix}$, implying $A = \begin{pmatrix} 0 & A_2 \\ A_1 & 0 \end{pmatrix}$. Hence G is bipartite.

(iii) → (vi): Recall $1 \in \sigma(K)$ from Theorem 3.14(ii). Since $\sigma(K)$ is symmetric, $-1 \in \sigma(K)$.

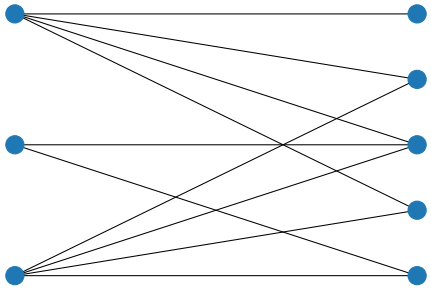
(vi) → (i): Since $-1 \in \sigma(K)$, we know $\begin{pmatrix} \mathbf{y} & \mathbf{y} \end{pmatrix}$ is the associated eigenvector by Theorem 3.14(i). Thus, $A\mathbf{y}+(D-I)\mathbf{y} = -\mathbf{y}$ implying $D^{-1}A\mathbf{y} = -\mathbf{y}$. So $-1 \in \sigma(D^{-1}A)$. Note $D^{-1}A$ is the transition probability matrix of G with spectral radius 1. Thus, G must be bipartite. □

Remark. An alternate proof of Theorem 3.23(vi) can be derived directly as a corollary of Proposition 4.15 of [46].

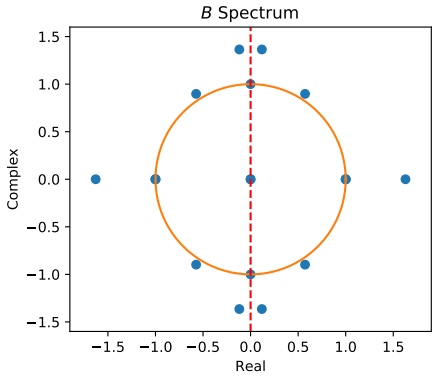
Recall in Theorem 3.5 we calculated $\sigma(B)$ when G is a tree. Since $\sigma(K) \supset \sigma(B)$ in the case of a tree, we can use the properties of bipartite graphs to classify $\sigma(K)$ for trees as well.

Corollary 3.24. *Let G be a tree. Let K be defined as above. Then $1, -1, 0 \in \sigma(K)$. The eigenvalue 0 has multiplicity at least r, where r is the number of leaves in G.*

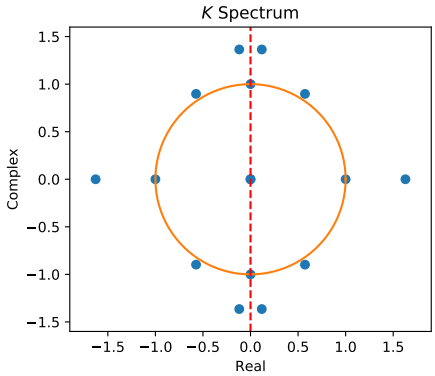
Proof. Let i be a leaf in G. Then $d_i - 1 = 0$. Thus e_{n+i} is an eigenvector of K with eigenvalue 0. This is true for all leaves. Thus, 0 has multiplicity at least r. From Theorem 3.14 and Corollary 3.23, we also know $1 \in \sigma(K)$ and $-1 \in \sigma(K)$ since every tree is a bipartite graph. □



(a) Bipartite Graph



(b) Spectrum of B



(c) Spectrum of K

Figure 3.8: The graph is (a) is bipartite and the plots in (b) and (c) are spectrum plots for B and K respectively. As can be seen, both spectrum plots are symmetric along the real axis. Additionally, -1 is in the spectrum of K.

Chapter 4. Non-backtracking PageRank

One question often asked when examining a graph is which vertices are the most "important" in the graph and defining a ranking. This ranking is often called a centrality measure. Various centrality measures exists for graphs (e.g., degree centrality, eigenvector centrality, betweenness centrality, etc.) each of which define the "importance" of a vertex differently [35]. The centrality measure in question, PageRank, determines the importance of a vertex by a weighted average of the degree of the vertex and the importance of it's neighbors. Calculating PageRank is traditionally done using simple random walks. In this chapter, we will calculate non-backtracking PageRank in the same way as traditional PageRank but with non-backtracking random walks rather than simple random walks.

The idea of calculating PageRank using non-backtracking random walks has been an area of recent study due to its modeling advantages. Arrigo et. al. developed a solver to efficiently solve for non-backtracking PageRank [5] and applied it to road networks. Aleja et. al. and Criado et. al. used non-backtracking PageRank to analyze Madrid's public transportation system [1, 15]. While each of these show the efficacy of using non-backtracking PageRank through experimental results, none tackle when and why PageRank and non-backtracking PageRank provide different results. The goal of this chapter will be the explore when PageRank and non-backtracking PageRank coincide, when they do not, and what causes them to differ.

4.1 Calculating PageRank, Non-Backtracking PageRank, and μ-PageRank

In order to compare PageRank and non-backtracking PageRank, we give explicit definitions of each as well as other tools useful in comparing them (i.e., edge PageRank and μ-PageRank).

4.1.1 Standard PageRank. To calculate PageRank, we consider a modified random walk $\{v_1, v_2, ...\}$ across G. With probability $\epsilon \in (0, 1)$, v_i is chosen uniformly at random from the neighbors of v_{i-1}. With probability $1 - \epsilon$, v_i is chosen from the set of vertices V with probability distribution u. Then the stationary distribution of the modified random walk is the PageRank vector π of G, where the i^{th} entry of π is the PageRank value of a given vertex.

Definition 4.1 (PageRank)**.** Let G be a graph with adjacency matrix A and diagonal matrix D. Let $\epsilon \in (0, 1)$ and v be a initial distribution vector such that $\|v\|_1 = 1$. Then the stationary distribution of

$$\epsilon D^{-1}A + (1 - \epsilon)\mathbf{1}v^T$$

is the PageRank vector π of G.

4.1.2 Edge PageRank. Non-backtracking PageRank is the same as PageRank but the modified random walk is a non-backtracking random walk with probability ϵ. In order to compare standard and non-backtracking PageRank, we look at an alternative representation of standard PageRank. Recall the edge adjacency matrix

$$C((i,j),(k,l)) = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{4.1}$$

and the graph $\hat{G}$ it generates. Let $\hat{D}$ be the diagonal degree matrix of $\hat{G}$. Additionally we recall the matrices

$$S((i,j),x) = \begin{cases} 1 & j = x \\ 0 & j \neq x \end{cases} \qquad T(x,(i,j)) = \begin{cases} 1 & i = x \\ 0 & i \neq x \end{cases} \tag{4.2}$$

which are used to project from the vertices of G to its lifted edges and vice versa. Arrigo et. al. use the matrices above to calculate PageRank of G using the graph $\hat{G}$ [5].

Definition 4.2 (Edge PageRank)**.** Let G be a graph with lifted graph $\hat{G}$. Let C be the adjacency matrix of $\hat{G}$ and $\hat{D}$ the (edge) degree matrix. Let $u = T^T(TT^T)^{-1}v$ be the initial distribution vector from standard PageRank projected onto $\hat{G}$. Then the stationary distribution of

$$H_1 = \epsilon\hat{D}^{-1}C + (1-\epsilon)\mathbf{1}u^T$$

is the edge PageRank $\hat{\pi}$ of G. The PageRank of G is then $\pi = T\hat{\pi}$.

4.1.3 Non-backtracking PageRank. Using the lift of G onto its directed edges, we can also calculate non-backtracking PageRank. To do this we consider the graph created by the non-backtracking matrix. We then can consider a modified random walk along this

new graph (i.e., along the edges of G without backtracking) where e_i is chosen uniformly at random from the neighbors of e_{i-1} with probability ϵ and e_i is chosen from the probability distribution u as defined in Definition 4.2. The stationary distribution of this random walk on the edges of G is the non-backtracking PageRank on the edges of G (which can then be projected using T). More precisely, we can define the non-backtracking PageRank as follows:

Definition 4.3 (Non-backtracking PageRank [5])**.** Let G be a graph with let B be its non-backtracking matrix and $\hat{D}$ its edge degree matrix. Let $u = T^T(TT^T)^{-1}v$ be the initial distribution vector from standard PageRank projected onto the directed edges of G. Then the stationary distribution of

$$H_0 = \epsilon(\hat{D} - I)^{-1}B + (1-\epsilon)\mathbf{1}u^T$$

is the edge non-backtracking PageRank $\hat{\pi}_0$ of G. The non-backtracking PageRank of G is $\pi_0 = T\hat{\pi}_0$.

4.1.4 μ-PageRank. The notion of μ-PageRank (or μ-centrality) has been studied to give an alternative ranking to vertices [14]. The main idea of μ-PageRank is to again perform a modified random walk $\{v_1, v_2, ...\}$ where v_i is chosen at random with probability ϵ from the neighbors of v_{i-1} and the edge from $v_{i-2} \sim v_{i-1}$ has weight μ and the edges to all other vertices have weight 1 and with probability $1-\epsilon$, v_i is chosen from the probability distribution v. Again to encapsulate this modified random walk into a Markov chain we lift the graph G to a graph of directed edges. We weight every backtracking connection with the probability μ (see Figure 4.1). To do this we recall the backtracking operator τ:

$$\tau((i,j),(k,l)) = \begin{cases} 1 & j = k, i = l \\ 0 & \text{otherwise} \end{cases}.$$

Intuitively, we are calculating PageRank as it becomes more and less likely to backtrack in the PageRank modified random walk. Precisely we define μ-PageRank as follows:

Definition 4.4 (μ-PageRank [14])**.** Let G be a graph with C its edge adjacency matrix, $\hat{D}$ its edge degree matrix and τ the backtracking operator. Let $u = T^T(TT^T)^{-1}v$ be the initial distribution vector from standard PageRank projected onto the directed edges of G. Then the stationary distribution of

$$H_\mu = \epsilon(\hat{D} - (1-\mu)I)^{-1}(C - (1-\mu)\tau) + (1-\epsilon)\mathbf{1}u^T$$

is the edge μ-PageRank $\hat{\pi}_\mu$ of G. The *μ-PageRank* of G is $\pi_\mu = T\hat{\pi}_\mu$.

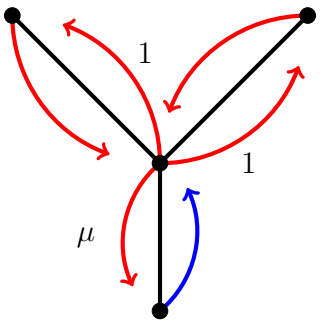


Figure 4.1: In the modified random walk for μ-PageRank, the blue directed edge connects to each edge with weight 1 with the exception of the backtracking edge. The backtracking edge has weight μ.

Remark. Setting $\mu = 1$ weights all the edges the same giving the standard PageRank of G. If $\mu = 0$, this makes it impossible to choose v_{i-2} as v_i with probability ϵ and is therefore non-backtracking PageRank. This motivates the notation H_1 and H_0 for standard and non-backtracking PageRank respectively. This has been rigorously proven by Criado et. al. [14].

4.2 When PageRanks are Equal

To begin our discussion on when standard and non-backtracking PageRank differ, we will begin by looking at when they are the same. Arrigo et. al. found they are equivalent

when our graph in question is regular.

Theorem 4.5 (Arrigo et. al.)**.** *Let A be the adjacency matrix of an k-regular graph with $k \geq 2$. Then for $\epsilon \in [0,1]$, define the matrices $\hat{P} = \epsilon(\hat{D} - I)^{-1}B + \frac{1-\epsilon}{n}\mathbf{1}\nu^T$ and $P = \epsilon D^{-1}A + \frac{1-\epsilon}{n}\mathbf{1}\mathbf{1}^T$ where $\nu = T^T D^{-1}$. If $P^T\mathbf{x} = \mathbf{x}$ and $\hat{P}^T\hat{\mathbf{y}} = \hat{\mathbf{y}}$, and $\|\mathbf{x}\| = \|\hat{\mathbf{y}}\| = 1$, then $T\hat{\mathbf{y}} = \mathbf{x}$.*

Using a similar method, we prove the same result holds for bipartite biregular graphs. A bipartite biregular graph is a bipartite graph where the vertices in V_1 have degree d_1 and vertices in V_2 have degree d_2. The PageRank value of a given vertex v_i solely depends on d_1 and d_2. Specifically, vertices in the vertex set with greater degree (d_1 or d_2) have greater rank than those in the opposite set. The result is stated explicitly as follows:

Theorem 4.6. *Let A be the adjacency matrix of an bipartite biregular graph with $d_1, d_2 \geq 2$ where d_i is the degree of each vertex in sets V_1 and V_2 respectively. Then for $\epsilon \in [0,1]$, define the matrices $\hat{P} = \epsilon(\hat{D} - I)^{-1}B + \frac{1-\epsilon}{n}\mathbf{1}\nu^T$ and $P = \epsilon D^{-1}A + \frac{1-\epsilon}{n}\mathbf{1}\mathbf{1}^T$ where $\nu = T^T D^{-1}$. If $P^T\mathbf{x} = \mathbf{x}$ and $\hat{P}^T\hat{\mathbf{y}} = \hat{\mathbf{y}}$, and $\|\mathbf{x}\| = \|\hat{\mathbf{y}}\| = 1$, then $T\hat{\mathbf{y}} = \mathbf{x}$. Specifically, vertices in V_1 have PageRank value $\frac{1+\epsilon\frac{d_1}{d_2}}{n(1+\epsilon)}$ and vertices in V_2 have PageRank value $\frac{1+\epsilon\frac{d_2}{d_1}}{n(1+\epsilon)}$ where $\mathbf{x}$ is the PageRank vector.*

Proof. Since G is bipartite, we can write B and T as

$$B = \begin{pmatrix} 0 & B_2 \\ B_1 & 0 \end{pmatrix} \qquad\qquad T = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}.$$

We can solve for $\hat{\mathbf{y}}$ by solving

$$\left(\begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix} - \epsilon \begin{pmatrix} 0 & B_1^T \\ B_2^T & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{d_2-1} & 0 \\ 0 & \frac{1}{d_1-1} \end{pmatrix}\right)\hat{\mathbf{y}} = \frac{1-\epsilon}{n} \begin{pmatrix} T_1^T & 0 \\ 0 & T_2^T \end{pmatrix} \begin{pmatrix} \frac{1}{d_1} & 0 \\ 0 & \frac{1}{d_2} \end{pmatrix} \mathbf{1}.$$

Recall T is the out degree matrix. Thus $T^T\mathbf{1}$ counts the number of vertices which point to a certain edge (the edge in-degree). Note edges can only have one vertex pointing to them.

So $T^T\mathbf{1} = \mathbf{1}$. We can then simplify the right side of the equation to

$$\frac{1-\epsilon}{n}\begin{pmatrix} T_1^T & 0 \\ 0 & T_2^T \end{pmatrix}\begin{pmatrix} \frac{1}{d_1} & 0 \\ 0 & \frac{1}{d_2} \end{pmatrix}\mathbf{1} = \frac{1-\epsilon}{n}\begin{pmatrix} \frac{1}{d_1}T_1^T & 0 \\ 0 & \frac{1}{d_2}T_2^T \end{pmatrix}\mathbf{1} = \frac{1-\epsilon}{n}\begin{pmatrix} \frac{1}{d_1}\mathbf{1} \\ \frac{1}{d_2}\mathbf{1} \end{pmatrix}.$$

We use the inverse of the matrix on the left side of the equation to get

$$\hat{\mathbf{y}} = \frac{1-\epsilon}{n}\left(\begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix} - \epsilon\begin{pmatrix} 0 & \frac{1}{d_1-1}B_1^T \\ \frac{1}{d_2-1}B_2^T & \end{pmatrix}\right)^{-1}\begin{pmatrix} \frac{1}{d_1}\mathbf{1} \\ \frac{1}{d_2}\mathbf{1} \end{pmatrix}.$$

We note the solution to a geometric series on the right side of the equation and replace the inverse matrix with a geometric series.

$$\hat{\mathbf{y}} = \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^r\begin{pmatrix} 0 & \frac{1}{d_1-1}B_1^T \\ \frac{1}{d_2-1}B_2^T & \end{pmatrix}^r\begin{pmatrix} \frac{1}{d_1}\mathbf{1} \\ \frac{1}{d_2}\mathbf{1} \end{pmatrix}.$$

By induction, we can rewrite this sum as

$$\sum_{r=0}^{\infty}\epsilon^r\begin{pmatrix} 0 & \frac{1}{d_1-1}B_1^T \\ \frac{1}{d_2-1}B_2^T & \end{pmatrix}^r =$$

$$\sum_{r=0}^{\infty}\epsilon^{2r}\begin{pmatrix} \left(\frac{1}{(d_1-1)(d_2-1)}B_1^TB_2^T\right)^r & 0 \\ 0 & \left(\frac{1}{(d_1-1)(d_2-1)}B_2^TB_1^T\right)^r \end{pmatrix}$$

$$+\sum_{r=0}^{\infty}\epsilon^{2r+1}\begin{pmatrix} 0 & \frac{1}{d_1-1}B_1^T\left(\frac{1}{(d_1-1)(d_2-1)}B_2^TB_1^T\right)^r \\ \frac{1}{d_2-1}B_2^T\left(\frac{1}{(d_1-1)(d_2-1)}B_1^TB_2^T\right)^r & 0 \end{pmatrix}.$$

So we can rewrite the non-backtracking PageRank equation as

$$\begin{aligned}
\hat{\mathbf{y}} =& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r}\begin{pmatrix}\left(\frac{1}{(d_1-1)(d_2-1)}B_1^TB_2^T\right)^r & 0\\ 0 & \left(\frac{1}{(d_1-1)(d_2-1)}B_2^TB_1^T\right)^r\end{pmatrix}\begin{pmatrix}\frac{1}{d_1}\mathbf{1}\\ \frac{1}{d_2}\mathbf{1}\end{pmatrix}\\
&+\frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1}\begin{pmatrix}0 & \frac{1}{d_1-1}B_1^T\left(\frac{1}{(d_1-1)(d_2-1)}B_2^TB_1^T\right)^r\\ \frac{1}{d_2-1}B_2^T\left(\frac{1}{(d_1-1)(d_2-1)}B_1^TB_2^T\right)^r & 0\end{pmatrix}\begin{pmatrix}\frac{1}{d_1}\mathbf{1}\\ \frac{1}{d_2}\mathbf{1}\end{pmatrix}\\
=&\frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r}\begin{pmatrix}\frac{1}{d_1}\left(\frac{1}{(d_1-1)(d_2-1)}B_1^TB_2^T\right)^r\mathbf{1}\\ \frac{1}{d_2}\left(\frac{1}{(d_1-1)(d_2-1)}B_2^TB_1^T\right)^r\mathbf{1}\end{pmatrix}\\
&+\frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1}\begin{pmatrix}\frac{1}{d_2(d_1-1)}B_1^T\left(\frac{1}{(d_1-1)(d_2-1)}B_2^TB_1^T\right)^r\mathbf{1}\\ \frac{1}{d_1(d_2-1)}B_2^T\left(\frac{1}{(d_1-1)(d_2-1)}B_1^TB_2^T\right)^r\mathbf{1}\end{pmatrix}.
\end{aligned}$$

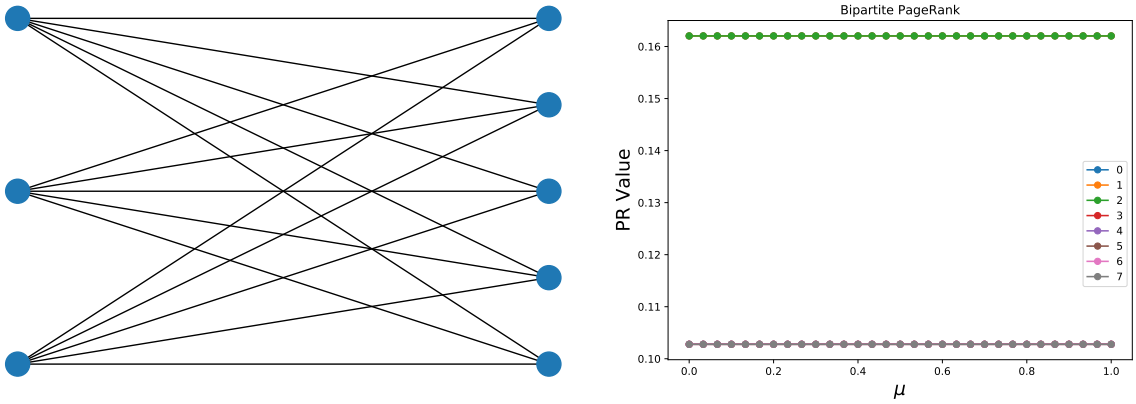
Note B is biregular and B_1 is an adjacency matrix mapping edges from partition 1 to edges from partition 2 (similarly B_2 maps from 2 to 1). Thus, $B_1^T\mathbf{1}$ will count the number of incoming edges to partition 1 from partition 2. This is the same as counting the number of edges leaving a vertex from partition 1 to parition 2 (the degree of vertices in parition 1).

Thus, $B_1^T\mathbf{1} = (d_1-1)\mathbf{1}$. Similarly $B_2^T\mathbf{1} = (d_2-1)\mathbf{1}$. Hence,

$$
\begin{aligned}
\hat{\mathbf{y}} &= \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r}\begin{pmatrix}\frac{1}{d_1}\left(\frac{(d_1-1)(d_2-1)}{(d_1-1)(d_2-1)}\mathbf{1}\right)^r\\ \frac{1}{d_2}\left(\frac{(d_1-1)(d_2-1)}{(d_1-1)(d_2-1)}\mathbf{1}\right)^r\end{pmatrix} + \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1}\begin{pmatrix}\frac{1}{d_2(d_1-1)}B_1^T\left(\frac{(d_1-1)(d_2-1)}{(d_1-1)(d_2-1)}\mathbf{1}\right)^r\\ \frac{1}{d_1(d_2-1)}B_2^T\left(\frac{(d_1-1)(d_2-1)}{(d_1-1)(d_2-1)}\mathbf{1}\right)^r\end{pmatrix}\\
&= \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r}\begin{pmatrix}\frac{1}{d_1}\mathbf{1}\\ \frac{1}{d_2}\mathbf{1}\end{pmatrix} + \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1}\begin{pmatrix}\frac{1}{d_2(d_1-1)}B_1^T\mathbf{1}\\ \frac{1}{d_1(d_2-1)}B_2^T\mathbf{1}\end{pmatrix}\\
&= \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r}\begin{pmatrix}\frac{1}{d_1}\mathbf{1}\\ \frac{1}{d_2}\mathbf{1}\end{pmatrix} + \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1}\begin{pmatrix}\frac{1}{d_2}\mathbf{1}\\ \frac{1}{d_1}\mathbf{1}\end{pmatrix}\\
&= \frac{1-\epsilon}{n}\begin{pmatrix}\left(\sum_{r=0}^{\infty}\frac{(\epsilon^2)^r}{d_1}\right)\mathbf{1}\\ \left(\sum_{r=0}^{\infty}\frac{(\epsilon^2)^r}{d_2}\right)\mathbf{1}\end{pmatrix} + \frac{1-\epsilon}{n}\begin{pmatrix}\left(\sum_{r=0}^{\infty}\frac{\epsilon^{2r+1}}{d_2}\right)\mathbf{1}\\ \left(\sum_{r=0}^{\infty}\frac{\epsilon^{2r+1}}{d_1}\right)\mathbf{1}\end{pmatrix}\\
&= \frac{1-\epsilon}{n}\begin{pmatrix}\frac{1}{d_1}\left(\sum_{r=0}^{\infty}(\epsilon^2)^r\right)\mathbf{1}\\ \frac{1}{d_2}\left(\sum_{r=0}^{\infty}(\epsilon^2)^r\right)\mathbf{1}\end{pmatrix} + \frac{1-\epsilon}{n}\begin{pmatrix}\frac{1}{d_2}\left(\sum_{r=0}^{\infty}\epsilon^{2r+1}\right)\mathbf{1}\\ \frac{1}{d_1}\left(\sum_{r=0}^{\infty}\epsilon^{2r+1}\right)\mathbf{1}\end{pmatrix}\\
&= \frac{1-\epsilon}{n}\begin{pmatrix}\frac{1}{d_1}\left(\sum_{r=0}^{\infty}(\epsilon^2)^r\right)\mathbf{1}\\ \frac{1}{d_2}\left(\sum_{r=0}^{\infty}(\epsilon^2)^r\right)\mathbf{1}\end{pmatrix} + \frac{1-\epsilon}{n}\begin{pmatrix}\frac{\epsilon}{d_2}\left(\sum_{r=0}^{\infty}\epsilon^{2r}\right)\mathbf{1}\\ \frac{\epsilon}{d_1}\left(\sum_{r=0}^{\infty}\epsilon^{2r}\right)\mathbf{1}\end{pmatrix}\\
&= \frac{1-\epsilon}{n}\begin{pmatrix}\left(\frac{1}{d_1}+\frac{\epsilon}{d_2}\right)\left(\sum_{r=0}^{\infty}(\epsilon^2)^r\right)\mathbf{1}\\ \left(\frac{1}{d_2}+\frac{\epsilon}{d_1}\right)\left(\sum_{r=0}^{\infty}(\epsilon^2)^r\right)\mathbf{1}\end{pmatrix}.
\end{aligned}
$$

Since $\epsilon^2 < 1$, the summation converges to $\frac{1}{1-\epsilon^2}$. Further, we simplify the fractions and get

$$
\begin{aligned}
\hat{\mathbf{y}} &= \frac{1-\epsilon}{n}\begin{pmatrix}\left(\frac{d_2+\epsilon d_1}{d_1 d_2}\right)\left(\frac{1}{1-\epsilon^2}\right)\mathbf{1}\\ \left(\frac{d_1+\epsilon d_2}{d_1 d_2}\right)\left(\frac{1}{1-\epsilon^2}\right)\mathbf{1}\end{pmatrix}\\
&= \frac{1-\epsilon}{nd_1d_2(1-\epsilon^2)}\begin{pmatrix}(d_2+\epsilon d_1)\mathbf{1}\\ (d_1+\epsilon d_2)\mathbf{1}\end{pmatrix}\\
&= \frac{1}{nd_1d_2(1+\epsilon)}\begin{pmatrix}(d_2+\epsilon d_1)\mathbf{1}\\ (d_1+\epsilon d_2)\mathbf{1}\end{pmatrix}.
\end{aligned}
$$



(a) (b)

Figure 4.2: The graph in (a) is a bipartite biregular graph with 3 and 5 nodes in each respective set, where the degree of the first set is 5 and the degree of the second set is 3. The plot in (b) shows that μ-PageRank is constant on the interval $\mu \in [0, 1]$, implying that standard and non-backtracking PageRank are equivalent. Only two nodes are shown as the plots for each node in the same set are layered.

Now to project $\hat{\mathbf{y}}$ to the vertices, we get

$$\begin{aligned} T\hat{\mathbf{y}} &= \frac{1}{nd_1d_2(1+\epsilon)} \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix} \begin{pmatrix} (d_2+\epsilon d_1)\mathbf{1} \\ (d_1+\epsilon d_2)\mathbf{1} \end{pmatrix} \\ &= \frac{1}{nd_1d_2(1+\epsilon)} \begin{pmatrix} (d_2+\epsilon d_1)(T_1\mathbf{1}) \\ (d_1+\epsilon d_2)(T_2\mathbf{1}) \end{pmatrix}. \end{aligned}$$

Note $T\mathbf{1}$ counts the number of outgoing edges from a given vertex. Thus,

$$\begin{aligned} T\hat{\mathbf{y}} &= \frac{1}{nd_1d_2(1+\epsilon)} \begin{pmatrix} (d_2+\epsilon d_1)d_1\mathbf{1} \\ (d_1+\epsilon d_2)d_2\mathbf{1} \end{pmatrix} \\ &= \frac{1}{n(1+\epsilon)} \begin{pmatrix} 1+\epsilon\frac{d_1}{d_2} & 0 \\ 0 & 1+\epsilon\frac{d_2}{d_1} \end{pmatrix} \mathbf{1}. \end{aligned}$$

Recall in a bipartite, biregular graph, the adjacency matrix is $A = \begin{pmatrix} 0 & A_1 \\ A_2 & 0 \end{pmatrix}$. Thus

$$\begin{aligned}
\mathbf{x} =& \frac{1-\epsilon}{n}(I - \epsilon A^T D^{-1})^{-1}\mathbf{1} \\
=& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^r (A^T)^r (D^{-1})^r \mathbf{1} \\
=& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^r \begin{pmatrix} 0 & A_2^T \\ A_1^T & 0 \end{pmatrix}^r \begin{pmatrix} \frac{1}{d_1} & 0 \\ 0 & \frac{1}{d_2} \end{pmatrix}^r \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \end{pmatrix} \\
=& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^r \begin{pmatrix} 0 & \frac{A_2^T}{d_2} \\ \frac{A_1^T}{d_1} & 0 \end{pmatrix}^r \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \end{pmatrix} \\
=& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r} \begin{pmatrix} \left(\frac{A_1^T A_2^T}{d_1 d_2}\right)^r & 0 \\ 0 & \left(\frac{A_2^T A_1^T}{d_1 d_2}\right)^r \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \end{pmatrix} \\
&+ \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1} \begin{pmatrix} 0 & \frac{A_2^T}{d_2}\left(\frac{A_2^T A_1^T}{d_1 d_2}\right)^r \\ \frac{A_1^T}{d_1}\left(\frac{A_1^T A_2^T}{d_1 d_2}\right)^r & 0 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \end{pmatrix} \\
=& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r} \begin{pmatrix} \left(\frac{A_1^T A_2^T}{d_1 d_2}\right)^r \mathbf{1} \\ \left(\frac{A_2^T A_1^T}{d_2 d_1}\right)^r \mathbf{1} \end{pmatrix} + \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1} \begin{pmatrix} \frac{A_2^T}{d_2}\left(\frac{A_2^T A_1^T}{d_2 d_1}\right)^r \mathbf{1} \\ \frac{A_1^T}{d_1}\left(\frac{A_1^T A_2^T}{d_1 d_2}\right)^r \mathbf{1} \end{pmatrix} \\
=& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r}\mathbf{1} + \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r+1} \begin{pmatrix} \frac{d_1}{d_2} \\ \frac{d_2}{d_1} \end{pmatrix} \mathbf{1} \\
=& \frac{1-\epsilon}{n}\sum_{r=0}^{\infty}\epsilon^{2r}\mathbf{1} + \frac{\epsilon(1-\epsilon)}{n}\sum_{r=0}^{\infty}\epsilon^{2r} \begin{pmatrix} \frac{d_1}{d_2} \\ \frac{d_2}{d_1} \end{pmatrix} \mathbf{1} \\
=& \frac{1-\epsilon}{n(1-\epsilon^2)} \begin{pmatrix} 1+\epsilon\frac{d_1}{d_2} \\ 1+\epsilon\frac{d_2}{d_1} \end{pmatrix} \mathbf{1} \\
=& \frac{1}{n(1+\epsilon)} \begin{pmatrix} 1+\epsilon\frac{d_1}{d_2} & 0 \\ 0 & 1+\epsilon\frac{d_2}{d_1} \end{pmatrix} \mathbf{1}.
\end{aligned}$$

Thus, $T\hat{\mathbf{y}} = \mathbf{x}$. □

4.3 Analysis of μ-PageRank

To compare standard and non-backtracking PageRank for a general graph, we will use μ-PageRank. As mentioned in Remark 4.1.4, if $\mu = 1$ we have standard PageRank and if $\mu = 0$ we have non-backtracking PageRank. Then for a given vertex, we can consider its PageRank value as a function of μ and consider the behavior of the function. We will denote the function in terms of the PageRank vector as π_μ.

Criado et. al. [14] originally considered π_μ on the interval $\mu \in [0, 1]$. We will consider π_μ on the interval $\mu \in [0, \infty)$. Intuitively, when μ gets larger than 1, the modified random walk used to calculated PageRank is increasingly likely to backtrack. On this interval, we show $(\pi_\mu)_i$ has a limit.

Proposition 4.7 (∞-PageRank)**.** *Let G be a graph with initial distribution vector v and μ-PageRank as in Definition 4.4. Let $\pi_\infty = \lim_{\mu\to\infty} \pi_\mu$ be the ∞-PageRank of G. Then*

$$\pi_\infty = (1+\epsilon)^{-1}v + \epsilon(1+\epsilon)^{-1}AD^{-1}v.$$

Therefore the ∞-Pagerank of vertex i is $\frac{v_i}{1+\epsilon} + \frac{\epsilon}{1+\epsilon}\sum_{j\sim i}\frac{v_j}{d_j}$ where d_j is the degree of vertex j and v_j is the j^{th} entry of v.

Proof. For $\epsilon \in (0, 1)$ we have

$$H_\infty = \lim_{\mu\to\infty} \epsilon(\hat{D} - (1-\mu)I)^{-1}(C - (1-\mu)\tau) + (1-\epsilon)\mathbf{1}u^T \tag{4.3}$$

$$= \epsilon\tau + (1-\epsilon)\mathbf{1}u^T. \tag{4.4}$$

Then we have

$$\hat{\pi}_\infty^T = \epsilon\hat{\pi}_\infty^T\tau + (1-\epsilon)\hat{\pi}_\infty^T\mathbf{1}u^T \tag{4.5}$$

$$\hat{\pi}_\infty^T(I - \epsilon\tau) = (1-\epsilon)v^T(TT^T)^{-1}T \tag{4.6}$$

$$\hat{\pi}_\infty^T = (1-\epsilon)v^T(TT^T)^{-1}T(I-\epsilon\tau)^{-1}. \tag{4.7}$$

Direct computation shows $(I - \epsilon\tau)^{-1} = (1-\epsilon^2)^{-1}I + \frac{\epsilon}{1-\epsilon^2}\tau$. So

$$\hat{\pi}_\infty^T = v^T(TT^T)^{-1}T\Big((1+\epsilon)^{-1}I + \epsilon(1+\epsilon)^{-1}\tau\Big).$$

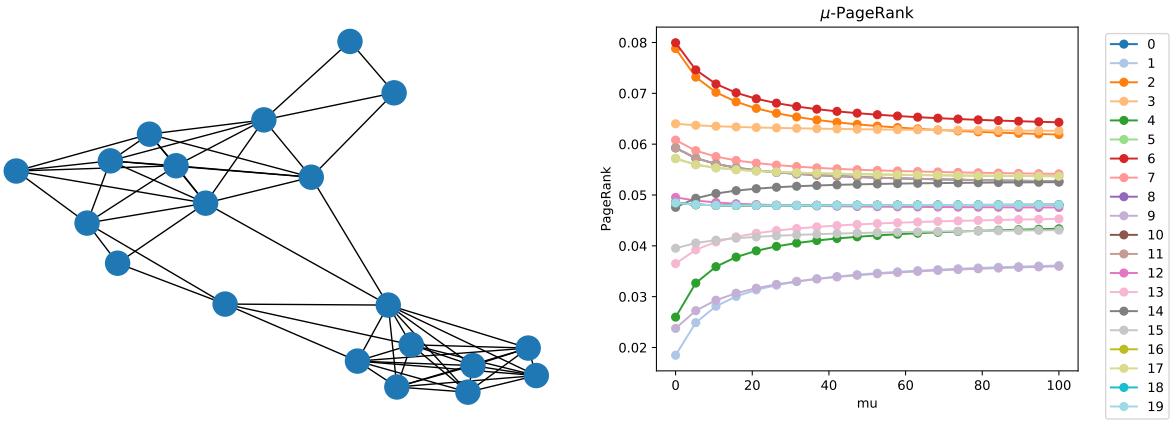
To project onto the vertices of G, we right multiply by T^T. To simplify this computation, we use the facts $TT^T = D$, $T\tau = S^T$ and $TS = A$. This gives

$$\begin{aligned}\pi_\infty^T &= (1+\epsilon)^{-1}v^T(TT^T)^{-1}TT^T + \epsilon(1+\epsilon)^{-1}v^T(TT^T)^{-1}T\tau T^T \qquad (4.8)\\ &= (1+\epsilon)^{-1}v^T + \epsilon(1+\epsilon)^{-1}v^T D^{-1}A. \qquad \square\end{aligned}$$

Additionally we conjecture $(\pi_\mu)_i$ is monotonic as a function of μ. To test this conjecture, we developed a Python script which creates 1000 random geometric graphs with 20 vertices and radius 0.4 and 1000 random GNP graphs with 20 vertices and probability 0.2 of two vertices being connected. We ensured each graph was connected. The script then computes the μ-PageRank for 10 values of μ between 0 and 20 (with 20 being chosen as the end point as most graphs μ-PageRank empirically converge to their limit by this value) for each graph. We also only do 10 values of μ to account for machine error and ensure significant change is happening between the PageRank values. For computing PageRank, $\epsilon \in (0,1)$ is chosen uniformly at random. The differences between μ-PageRank values for each vertex in a given graph is then found and signed as μ increases. Lastly, the sum of the signs is taken for each graph and analyzed to determine whether μ-PageRank is monotonic for each vertex. This script can be found in A.2.

We note because some vertices' PageRank does not change as μ increases, machine error implies $(\pi_\mu)_i$ is not monotonic, but fluctuates at infinitesimal margins of error. To more closely examine these cases, we save a plot of each experiment to analyze in the case of a PageRank value not changing monotonically. After this analysis, experimentation showed monotonicity in μ-PageRank for each vertex. These results lead us to the following conjecture:

Conjecture 4.8. *The function $(\pi_\mu)_i$ is monotonic in μ for all i.*



(a) Random Geometric Graph

(b) μ-PageRank

Figure 4.3: The graph in (a) is a random geometric graph with 20 vertices and radius 0.4. Letting $\epsilon = 0.8$, we calculated π_μ. The plot in (b) shows $(\pi_\mu)_i$ for each vertex i on the interval $\mu \in [0, 100)$. It is clear each function is monotonic and converges as $\mu \to \infty$.

4.4 Comparing standard and Non-Backtracking PageRank

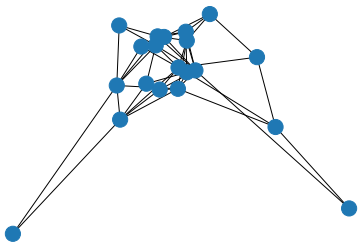
Using the properties in Section 4.3, we can analyze how standard and non-backtracking PageRank compare with relative ease. Given a graph G, we calculate the standard PageRank. This can be done with a variety of computationally efficient tools [10, 18, 21]. We will use the NetworkX implementation. Then the PageRank limit of a given vertex can be calculated using Proposition 4.7. Assuming we have monotonicity in μ-PageRank, we can compare standard PageRank with the limit to better understand non-backtracking PageRank using Algorithm 1.

It is worth noting what this Algorithm does not do. It does not show whether the ranking of a given vertex has changed from standard to non-backtracking PageRank. Rather, Algorithm 1 only shows whether the PageRank value changes with respect to μ. As seen in the next example, the PageRank value changing does not necessarily imply the ranking of a given vertex changes.

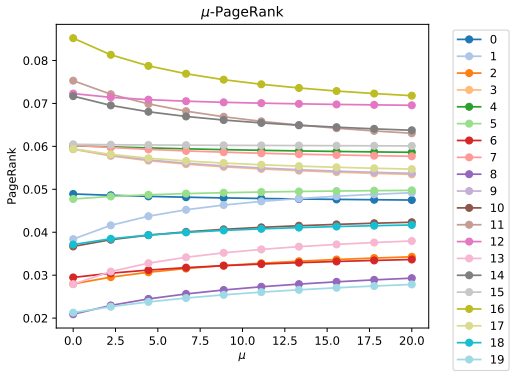
Algorithm 1: Comparing standard and Non-Backtracking PageRank

Result: Determine if non-backtracking PageRank is greater or smaller
$G \leftarrow$ Graph;
$\epsilon \leftarrow$ jumping factor;
$\pi_1 \leftarrow$PageRank(G,ϵ);
$\pi_\infty \leftarrow$PageRankLimit(G,ϵ);
def *compare*(i,π_1,π_∞)**:**
  if $(\pi_1)_i > (\pi_\infty)_i$ **then**
    return -1;
  end
  if $(\pi_1)_i < (\pi_\infty)_i$ **then**
    return 1;
  else
    return 0
  end

4.4.1 Example. We create a random GNP graph with 20 vertices and probability 0.2 and take the 2-core of the largest component of said graph. In Figure 4.4, the graph can be seen as well as a plot describing how μ-PageRank changes for each vertex on the interval $\mu \in [0, 20]$. It is clearly seen these functions are monotonic with respect to μ up to machine error. They are also converging to the limit found in Proposition 4.7. Given the conditions of Algorithm 1 are satisfied, we may use the algorithm to compare standard and non-backtracking PageRank.



(a) GNP Graph



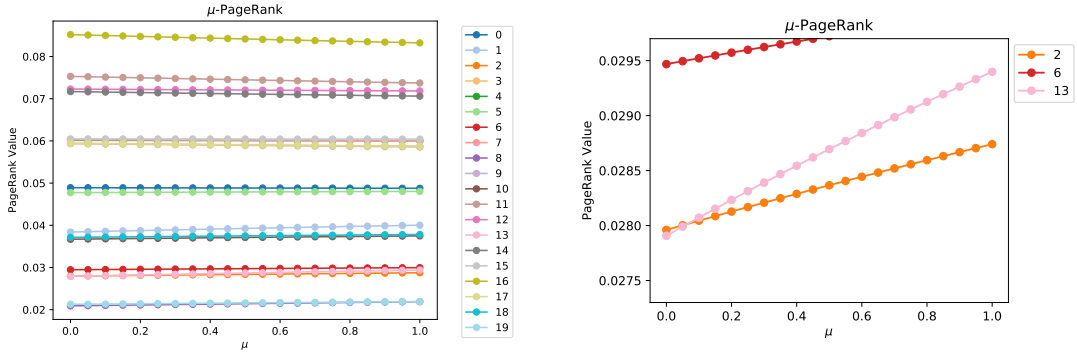
(b) μ-PageRank

Figure 4.4: The graph generated is the 2-core of the largest connected component of a GNP graph with 20 vertices and probability 0.2. The plot on the right shows the PageRank value of each vertex in the graph on the interval $\mu \in [0, 20]$.

Vertex	Output	Vertex	Output
0	-1	10	1
1	1	11	-1
2	1	12	-1
3	-1	13	1
4	-1	14	-1
5	1	15	-1
6	1	16	-1
7	-1	17	-1
8	1	18	1
9	-1	19	1

Table 4.1: Output of Algorithm 1 on the graph in Figure 4.4.

In Table 4.1 the results of Algorithm 1 for the graph generated are displayed. Comparing with Figure 4.4(b) we see the algorithm accurately tracks when non-backtracking PageRank is greater than, lesser than, or equal to standard PageRank (-1 indicates greater, 1 indicates lesser, and 0 indicates equality). However these outputs do not indicate change in rank. For example, Figure 4.4 shows the PageRank value of vertex 16 decreases as μ increases however it remains the vertex with highest rank for all $\mu \in [0, 20]$ (however $(\pi_\infty)_{12} > (\pi_\infty)_{16}$).



(a) μ-PageRank on $[0, 1]$

(b) μ-PageRank of vertices 2 and 13

Figure 4.5: The plot on the left show each function $(\pi_\mu)_i$ on the interval $\mu \in [0, 1]$. The plot on the left zooms in on the functions $(\pi_\mu)_2$ and $(\pi_\mu)_{13}$.

To compare standard and non-backtracking ranking changes, we must examine the behavior of the set of functions on the interval $\mu \in [0, 1]$. Figure 4.5(a) shows each of these functions on this smaller interval. Each continue to appear monotonic. To identify rank changes, we look for intersections between functions. In Figure 4.5(b), we can more clearly

see vertices 2 and 13 switch rank between standard and non-backtracking PageRank.

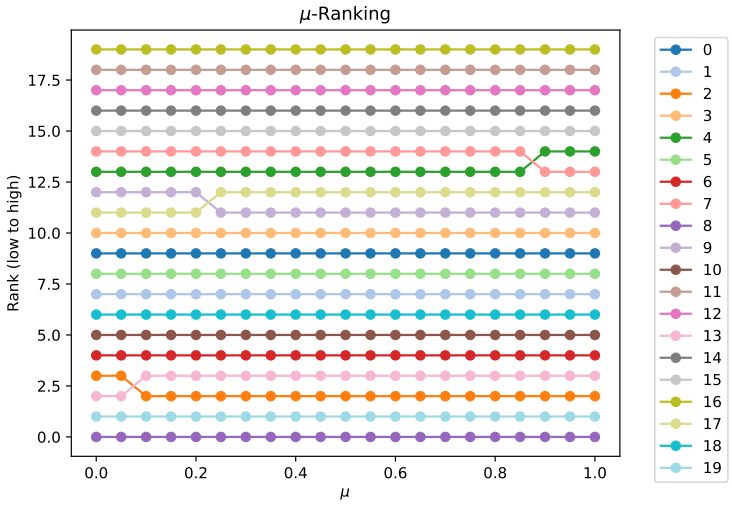


Figure 4.6: The plot shows the ranking of each vertex in the graph from Figure 4.4 for $\mu \in [0, 1]$. Higher rankings imply greater PageRank values (e.g., the vertex with greatest PageRank value is vertex 16 which has rank 19).

To compare all vertices, we create a ranking plot found in Figure 4.6. This indicates the rank of each vertex in the graph as μ increases from 0 to 1. The ranking change observed in Figure 4.5(b) can be seen in the bottom left of the plot. Overall, there are three ranking changes occur as μ increases from 0 to 1 in this example.

4.5 Future Work

Continuing this project, we would like to prove Conjecture 4.8. Proving Conjecture 4.8 would provide certainty that Algorithm 1 may always be used to identify differences in PageRank values in standard and non-backtracking PageRank. Additionally, we would like to bound the difference between standard and non-backtracking PageRank values. That is, if the non-backtracking PageRank value of a vertex is found to be greater than the standard PageRank value, we want to know how much greater can it be.

As seen in Figure 4.4, whether a function $(\pi_\mu)_i$ is monotonically increasing or decreasing depends on the vertex i. We would like to identify the condition determining whether the

function increases or decreases. Through data exploration and analysis of Proposition 4.7, we believe the condition depends on the degree of the vertex i, the degree of the neighbors of i, and the jumping factor ϵ.

Lastly, and perhaps most importantly, we would like to develop an algorithm similar to Algorithm 1 which determines whether a given vertex differs in ranking for standard and non-backtracking PageRank. A development of this type of algorithm could allow for analysis of non-backtracking PageRank without the need to compute non-backtracking PageRank (which proves to be extremely computationally expensive for large graphs).

Chapter 5. Conclusion

In this thesis we have centered our focus on non-backtracking random walks. Chapter 3 focused on the spectrum of the non-backtracking matrix. We showed the non-backtracking matrix can be decomposed to more explicitly describe its spectrum and developed both upper and lower bounds for the non-backtracking spectrum. In addition, we analyzed the spectrum of various classes of graphs. Chapter 4 focused on the difference between standard and non-backtracking PageRank. We developed the concept of μ-PageRank and used μ-PageRank to develop an algorithm which determines whether non-backtracking PageRank produces a larger, smaller, or equal value than standard PageRank for a given vertex. For bipartite biregular graphs, we showed standard and non-backtracking PageRank are equivalent.

Appendix A. Code

A.1 General Anaylsis

The following is a collection of functions which create the matrices used to understand the non-backtracking matrix of a graph.

```
import numpy as np
```

```
import scipy.linalg as la
import networkx as nx


def create_k(G):
    """
    Creates matrix K.
    Inputs:
        G (networkx graph)
    Output:
        K (ndarray): block matrix K
    """
    # Get necessary matrices
    A = np.array(nx.adjacency_matrix(G).todense())
    D = np.diag(list(dict(G.degree()).values()))
    I = np.eye(D.shape[0])
    # Create K
    K = np.block([[A,D-I],[-I,np.zeros((D.shape[0],D.shape[0]))]])
    return K
def create_s_t(G):
    """
    Creates S and T matrix for graph
    Inputs:
        G (networkx graph)
    Outputs:
        S (ndarray): S matrix
        T (ndarray): T matrix
    """
```

```
    direct = G.to_directed()
    # Find S and T
    S = np.zeros((len(direct.edges),len(G.nodes)))
    T = np.zeros((len(G.nodes),len(direct.edges)))
    # Check edge-node connections
    for i,a in enumerate(direct.edges):
        for j,b in enumerate(G.nodes):
            if a[1] == b:
                S[i,j] = 1
            if a[0] == b:
                T[j,i] = 1
    return S, T
def to_edge_space(G, B=False, graph=True, ret_tau = False):
    """
    Converts a graph into the edge space
    """
    direct = G.to_directed()
    # Find S and T
    S, T = create_s_t(G)
    # Create tau
    tau = np.zeros((len(direct.edges),len(direct.edges)))
    for i,a in enumerate(direct.edges):
        for j,b in enumerate(direct.edges):
            if a[0]==b[1] and a[1]==b[0]:
                tau[i][j] = 1
    # Create non-backtracking graph
    if B:
```

```
    # Create graph generated by non-backtracking matrix
    if graph:
        # Return tau, non-backtracking and edge graph
        if ret_tau:
            return nx.DiGraph(S@T),
                nx.DiGraph(S@T-tau),
                nx.DiGraph(tau)
        # Return edge and non-backtracking graph
        return nx.DiGraph(S@T), nx.DiGraph(S@T-tau)
    # Return tau, non-backtracking and edge matrix
    if ret_tau:
        return S@T, S@T - tau, tau
    return S@T, S@T - tau
# Create graph generated by edge matrix
if graph:
    # Return edge graph and tau graph
    if ret_tau:
        return nx.DiGraph(S@T), nx.DiGraph(tau)
    # Return edge graph
    return nx.DiGraph(S@T)
# Return edge and tau matrix
if ret_tau:
    return  S@T, tau
# Return edge matrix
return S@T
```

A.2 Monotonicity Check

The following code was used to analyze many graphs at once and check $(\pi_\mu)_i$ is monotonic for each vertex i.

```
import numpy as np
import networkx as nx
import scipy.linalg as la
from ipyparallel import Client
import matplotlib.pyplot as plt
import pandas as pd
import nb_general.py as nb


client = Client()
client.ids
dview = client[:]
dview.execute(``import numpy as np")
dview.execute(``import networkx as nx")
dview.execute(``import scipy.linalg as la")
dview.block = True


def mu_pr(G,mu, eps=.8):
    """
    Calculate mu-pagerank
    Input:
        G (networkx graph)
        mu (float)
        eps (float): jumping factor
    Output:
```

```
        mu_pagerank (ndarray): mu-PageRank values
    """

    # Get S, T, C, and tau
    S, T = nb.create_s_t(G)
    C, tau = nb.to_edge_space(G,graph=False,ret_tau=True)
    # Calculate the inverse of degree matrix
    Dinv = np.diag(1/(T@S).sum(axis=1))
    # Calculate uniform distribution
    # across vertices on the edge space
    u = T.T@Dinv@np.ones(len(G.nodes()))
    u = u/la.norm(u,ord=1)
    personalization = dict()
    for i in range(len(u)):
        personalization[i] = u[i]
    # Generate graph for mu-PageRank
    new_graph = nx.from_numpy_array(C-(1-mu)*tau)
    # Calculate mu-PageRank on the edge space
    edge_pr = np.array(list(nx.pagerank(new_graph,alpha=eps,
                    personalization=personalization).values()))
    # Project to vertex space
    return T@edge_pr
def make_graph(A,_):
    """
    Create plot showing change in mu-Pagerank for each vertex
    Input:
        A (ndarray): matrix containing mu values for all nodes
```

```
                    and a range of mu values
        _ (int): counter
    """
    # Set colors and domain
    cm = plt.cm.get_cmap('tab20')
    domain = np.linspace(0,20,10)
    # Create plot
    for i in range(A.shape[1]):
        plt.plot(domain,A[:,i],'-o',color=cm.colors[i],label=i)
    plt.legend(bbox_to_anchor=(1.05,1))
    plt.title(f"Figure {_}")
    plt.xlabel("mu")
    plt.ylabel("PageRank")
    plt.tight_layout()
    # Save plot
    plt.savefig(f"figures/figure_{_}.pdf")
    plt.close()
# Set mu values and initialize
mus = np.linspace(0,20,10)
evaluations = np.zeros((20,2000))
# Create random geometric graphs
for _ in range(1000):
    try:
        # Create graph
        G = list(nx.connected_component_subgraphs(
                nx.random_geometric_graph(20,.4)))[0]
        # Choose jumping factor
```

```
            eps = 0
            while eps == 0:
                eps = np.random.random()
            # Get mu-pagerank matrix
            output = dview.map(mu_pr,[G for i in range(len(mus))],
                               mus,[eps for i in range(len(mus))])
            A = np.vstack(output)
            # Check for monotonicity
            evaluation = np.sign(A-
                             np.roll(A,1,axis=0))[1:,:].sum(axis=0)
            evaluations[:,_] = evaluation
            # Plot
            make_graph(A,_)
        # Continue if graph was not connected
        except:
            print(f"{_} BROKE")
            continue
# Create GNP graphs
for _ in range(1000):
    try:
        # Create graph
        G = list(nx.connected_component_subgraphs(
                nx.gnp_random_graph(20,.2)))[0]
        # Choose jumping factor
        eps = 0
        while eps == 0:
            eps = np.random.random()
```

```
            # Get mu-pagerank matrix
            output = dview.map(mu_pr,[G for i in range(len(mus))],mus)
            A = np.vstack(output)
            # Check for monotonicity
            evaluation = np.sign(A-
                        np.roll(A,1,axis=0))[1:,:].sum(axis=0)
            evaluations[:,int(1000+_)] = evaluation
            # Plot
            make_graph(A,int(1000+_))
        # Continue if graph was not connected
        except:
            print(f''{int(1000+_)} BROKE")
            continue
# Save monotonicity information
evaluations = np.array(evaluations)
data = pd.DataFrame(evaluations)
data.to_csv("monotonicity_check.csv")
```

Bibliography

[1] Aleja, D., Criado, R., del Amo, A. J. G., Pérez, Á., and Romance, M. Non-backtracking PageRank: From the classic model to hashimoto matrices. *Chaos, Solitons & Fractals 126* (2019), 283–291.

[2] Alon, N., Benjamini, I., Lubetzky, E., and Sodin, S. Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics 9*, 04 (2007), 585–603.

[3] Angel, O., Friedman, J., and Hoory, S. The non-backtracking spectrum of the universal cover of a graph. *Transactions of the American Mathematical Society 367* (2007), 4287–4318.

[4] Arrigo, F., Grindrod, P., Higham, D. J., and Noferini, V. Non-backtracking walk centrality for directed networks. *Journal of Complex Networks 6*, 1 (2018), 54–78.

[5] Arrigo, F., Higham, D. J., and Noferini, V. Non-backtracking PageRank. *Journal of Scientific Computing 80*, 3 (2019), 1419–1437.

[6] Arrigo, F., Higham, D. J., and Noferini, V. Beyond non-backtracking: non-cycling network centrality measures. *Proceedings of the Royal Society A 476*, 2235 (2020), 20190653.

[7] Avena, L., Guldacs, H., Hofstad, R., Hollander, F., and Nagy, O. Linking the mixing times of random walks on static and dynamic random graphs.

[8] Ben-Hamou, A., and Salez, J. Cutoff for nonbacktracking random walks on sparse random graphs. *The Annals of Probability 45*, 3 (2017), 1752–1770.

[9] Bordenave, C., Lelarge, M., and Massoulié, L. Non-backtracking spectrum of random graphs: Community detection and non-regular Ramanujan graphs. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (2015), pp. 1347–1357.

[10] Borgs, C., Brautbar, M., Chayes, J., and Teng, S.-H. A sublinear time algorithm for PageRank computations. In *Proceedings of the 9th Workshop on Algorithms and Models for the Web Graph (WAW)* (January 2012), pp. 41–53.

[11] Brito, G., Dumitriu, I., and Harris, K. D. Spectral gap in random bipartite biregular graphs and applications. *arXiv preprint arXiv:1804.07808* (2018).

[12] Brouwer, A. E., and Haemers, W. H. *Spectra of graphs*. Springer Science & Business Media, 2011.

[13] Chen, W., Liu, J., Chen, Y., Khong, S. Z., Wang, D., Başar, T., Qiu, L., and Johansson, K. H. Characterizing the positive semidefiniteness of signed Laplacians via effective resistances. In *2016 IEEE 55th Conference on Decision and Control (CDC)* (2016), IEEE, pp. 985–990.

[14] Criado, R., Flores, J., García, E., del Amo, A. J. G., Ángel Pérez, and Romance, M. On the α-nonbacktracking centrality for complex networks: Existence and limit cases. *Journal of Computational and Applied Mathematics 350* (2019), 35–45.

[15] Criado, R., Moral, S., Pérez, Á., and Romance, M. On the edges' PageRank and line graphs. *Chaos: An Interdisciplinary Journal of Nonlinear Science 28*, 7 (2018), 075503.

[16] Das, K. C., and Kumar, P. Some new bounds on the spectral radius of graphs. *Discrete Mathematics 281*, 1-3 (2004), 149–161.

[17] Gulikers, L., Lelarge, M., and Massoulié, L. Non-backtracking spectrum of degree-corrected stochastic block models. *ArXiv abs/1609.02487* (2017).

[18] Hagberg, A. A., Schult, D. A., and Swart, P. J. Exploring network structure, dynamics, and function using networkx. In *Proceedings of the 7th Python in Science Conference* (Pasadena, CA USA, 2008), G. Varoquaux, T. Vaught, and J. Millman, Eds., pp. 11 – 15.

[19] Horn, R. A., and Johnson, C. R. *Matrix analysis*. Cambridge university press, 2012.

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

[21] Kamvar, S., Haveliwala, T., and Golub, G. Adaptive methods for the computation of pagerank. *Linear Algebra and its Applications 386* (2004), 51–65. Special Issue on the Conference on the Numerical Solution of Markov Chains 2003.

[22] Kempton, M. Non-backtracking random walks and a weighted Ihara's theorem. *Open Journal of Discrete Mathematics 6* (2016), 207–226.

[23] Kotani, M., and Sunada, T. Zeta functions of finite graphs. *Journal of Mathematical Sciences-University of Tokyo 7*, 1 (2000), 7–26.

[24] Kwon, S., and Lee, S. Ihara zeta function of finite graphs with circuit rank two. *The Korean Journal of Mathematics 28*, 1 (2020), 123–136.

[25] Lin, Y., and Zhang, Z. Non-backtracking centrality based random walk on networks. *The Computer Journal 62*, 1 (2019), 63–80.

[26] Liu, B. On an upper bound of the spectral radius of graphs. *Discrete mathematics 308*, 23 (2008), 5317–5324.

[27] Liu, Q., Chen, G., and Zhao, L. Some new bounds on the spectral radius of matrices. *Linear algebra and its applications 432*, 4 (2010), 936–948.

[28] Lovász, L. Random walks on graphs: A survey, 1993.

[29] Lubetzky, E., and Peres, Y. Cutoff on all Ramanujan graphs. *Geometric and Functional Analysis 26*, 4 (2016), 1190–1216.

[30] Lubetzky, E., Sly, A., et al. Cutoff phenomena for random walks on random regular graphs. *Duke Mathematical Journal 153*, 3 (2010), 475–510.

[31] Martin, T., Zhang, X., and Newman, M. E. Localization and centrality in networks. *Physical review E 90*, 5 (2014), 052808.

[32] Moradi-Jamei, B., Shakeri, H., Poggi-Corradini, P., and Higgins, M. J. A new method for quantifying network cyclic structure to improve community detection. *Physica A: Statistical Mechanics and its Applications 561* (2021), 125116.

[33] Morone, F., and Makse, H. A. Influence maximization in complex networks through optimal percolation. *Nature 524*, 7563 (2015), 65–68.

[34] Newman, M. Spectral community detection in sparse networks. *arXiv preprint arXiv:1308.6494* (2013).

[35] Newman, M. *Networks*. Oxford university press, 2018.

[36] Nikiforov, V. Bounds on graph eigenvalues I. *Linear Algebra and Its Applications 420*, 2-3 (2007), 667–671.

[37] Ortner, R., and Woess, W. Non-backtracking random walks and cogrowth of graphs. *Canadian Journal of Mathematics 59* (03 2004).

[38] Pan, J., Jiang, F., and Xu, J. Influence maximization in social networks based on non-backtracking random walk. In *2016 IEEE First International Conference on Data Science in Cyberspace (DSC)* (2016), pp. 260–267.

[39] Rossi, R. A., and Ahmed, N. K. The network data repository with interactive graph analytics and visualization. In *AAAI* (2015).

[40] Schwenk, A. J. Almost all trees are cospectral. *New directions in the theory of graphs* (1973), 275–307.

[41] Sodin, S., and Sackler, B. Random matrices, nonbacktracking walks, and orthogonal polynomials. *Journal of Mathematical Physics 48* (03 2007).

[42] Somodi, M., and Cheuk, D. Ihara zeta functions of coronae. *Linear Algebra and its Applications* (2020).

[43] Stanić, Z. *Inequalities for graph eigenvalues*, vol. 423. Cambridge University Press, 2015.

[44] Stephan, L., and Massouli'e, L. Non-backtracking spectra of weighted inhomogeneous random graphs. *arXiv: Probability* (2020).

[45] Stephan, L., and Massoulié, L. Non-backtracking spectra of weighted inhomogeneous random graphs. 80 pages, Feb. 2021.

[46] Torres, L. Non-backtracking spectrum: Unitary eigenvalues and diagonalizability. *arXiv preprint arXiv:2007.13611* (2020).

[47] Torres, L., Chan, K. S., Tong, H., and Eliassi-Rad, T. Node immunization with non-backtracking eigenvalues. *arXiv preprint arXiv:2002.12309* (2020).

[48] Wang, K., and Wood, P. M. Limiting empirical spectral distribution for the non-backtracking matrix of an erdős-rényi random graph. *arXiv: Probability* (2017).