

BIPARTITE SUBGRAPHS AND THE SIGNLESS LAPLACIAN MATRIX

Steve Kirkland and Debdas Paul

For a connected graph G , we derive tight inequalities relating the smallest signless Laplacian eigenvalue to the largest normalised Laplacian eigenvalue. We investigate how vectors yielding small values of the Rayleigh quotient for the signless Laplacian matrix can be used to identify bipartite subgraphs. Our results are applied to some graphs with degree sequences approximately following a power law distribution with exponent value 2.1 (scale-free networks), and to a scale-free network arising from protein-protein interaction.

1. Introduction

Let G be a graph with adjacency matrix A , and denote the diagonal matrix of vertex degrees for G by D . The matrix $Q = D + A$ is known as the *signless Laplacian matrix* for G , and has been the subject of a flurry of recent papers. The surveys [5], [6] and [7] give an overview of the research on the signless Laplacian matrix from the perspective of spectral graph theory. In particular, it is known that the signless Laplacian matrix Q for a graph G is positive semi-definite, and that the multiplicity of 0 as an eigenvalue of Q coincides with the number of connected components of G that are bipartite [4]; thus, for a connected graph G , the smallest eigenvalue of Q is positive if and only if G is not bipartite.

In a similar vein, if G is a graph with no isolated vertices, the matrix $\mathcal{L} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ is known as the *normalised Laplacian matrix* for G (here, A and D are as above). The spectral properties of the normalised Laplacian matrix are also well-studied, and [3] gives an extensive discussion of how the spectral properties of \mathcal{L} reflect the structure of G . In particular, it is known that the eigenvalues of the normalised Laplacian matrix fall in the interval $[0, 2]$, and that the multiplicity of 2 as an eigenvalue of \mathcal{L} coincides with the number of connected components of G that are bipartite. Thus, as above, for a connected graph G , the largest eigenvalue of \mathcal{L} is less than 2 if and only if G is not bipartite.

Suppose now that we have a graph G with signless Laplacian matrix Q and normalised Laplacian matrix \mathcal{L} . In view of the observations above, we might interpret the smallest eigenvalue of Q , say μ , as a measure of how bipartite our graph G is; alternatively, we might interpret the largest eigenvalue of \mathcal{L} , say λ , as a measure of how bipartite our graph G is. (We note in passing that both interpretations are

2000 Mathematics Subject Classification: 05C50
Keywords and Phrases: Bipartite subgraph, Signless Laplacian matrix, Normalised Laplacian matrix.

philosophically related to the work of Fielder [8], who proposed that the second smallest eigenvalue of the Laplacian matrix $D - A$ be used as a measure of the connectivity of G .) Since both μ and λ provide notions of how bipartite G is, it is natural to investigate the relationship between these two quantities. We do so in section 2, providing upper and lower bounds on λ in terms of μ ; both bounds are tight, and we give examples of infinite families of nonbipartite, nonregular graphs for which the upper and lower bounds are attained, respectively.

The discussion above gives rise to the following scenario. Suppose that our connected graph G is not bipartite, but its smallest signless Laplacian eigenvalue is close to zero. One might then have the intuition that G contains a bipartite subgraph that is not very well connected with the rest of the graph. How might we identify such a subgraph? In section 3, we provide a condition (based on a Rayleigh quotient) that is sufficient to identify a bipartite subgraph H . We also give a condition under which we can bound the number of vertices of the subgraph H having at least one neighbour in $G \setminus H$. Both results serve to reinforce the intuition noted above.

Part of our interest in identifying bipartite subgraphs that are only weakly connected to the rest of the graph stems from the study of complex networks such as protein-protein interaction networks, the world wide web, and certain social networks (see [14] for a survey of work on this topic). Within such networks, bipartite structures are important in verifying some useful properties [9], [15]. For example, in the case of protein-protein interaction networks, these bipartite subgraphs represent biologically relevant interaction motifs [13]. In section 4, we use eigenvectors of the signless Laplacian matrix to generate certain Rayleigh quotients that enable us to identify bipartite subgraphs using the results of section 3. That approach is applied to randomly generated graphs that approximate scale-free networks, and to a particular protein-protein interaction network.

Throughout the paper we will make use of basic results and techniques from matrix theory and from graph theory. The reader is referred to [10] and [12] for the necessary background material on these topics.

2. Extreme eigenvalues for the signless and normalised Laplacian matrices

As noted in section 1, for a connected graph G , both the smallest signless Laplacian eigenvalue and the largest normalised Laplacian eigenvalue for G serve as indicators as to whether or not G is bipartite. Since both of these eigenvalues identify a common feature of G (i.e. bipartiteness, or the lack thereof) it is natural

to seek a quantifiable relationship between these two eigenvalues. The following result does precisely that.

THEOREM 2.1. *Let G be a connected graph on n vertices with signless Laplacian matrix Q and normalised Laplacian \mathcal{L} . Let the smallest eigenvalue of Q be μ and the largest eigenvalue of \mathcal{L} be λ . Denote the maximum and minimum degrees of G by Δ and δ , respectively. Then*

$$(1) \quad 2 - \frac{\mu}{\delta} \leq \lambda \leq 2 - \frac{\mu}{\Delta}.$$

Proof: Let A denote the adjacency matrix of G , and D denote the diagonal matrix of vertex degrees, so that $Q = D + A$ and $\mathcal{L} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$.

We first consider the left hand inequality in (1). Let x be an eigenvector of Q corresponding to μ . Without loss of generality, we write x as $\begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$, where both x_1 and x_2 are nonnegative vectors. Partition D and A conformally with x as $\left[\begin{array}{c|c} D_1 & 0 \\ \hline 0 & D_2 \end{array} \right]$ and $\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right]$.

Since $x^T Q x = \mu x^T x$, we find that

$$(2) \quad x_1^T D_1 x_1 + x_1^T A_{11} x_1 - 2x_1^T A_{12} x_2 + x_2^T D_2 x_2 + x_2^T A_{22} x_2 = \mu(x_1^T x_1 + x_2^T x_2).$$

Next, let $z = \begin{bmatrix} D_1^{\frac{1}{2}} x_1 \\ -D_2^{\frac{1}{2}} x_2 \end{bmatrix}$. Then

$$z^T \mathcal{L} z = x_1^T D_1 x_1 - x_1^T A_{11} x_1 + 2x_1^T A_{12} x_2 + x_2^T D_2 x_2 - x_2^T A_{22} x_2,$$

and using (2) we thus find that $z^T \mathcal{L} z = 2x^T D x - \mu x^T x$. Thus we find that $\frac{z^T \mathcal{L} z}{z^T z} = 2 - \mu \frac{x^T x}{z^T z} \geq 2 - \frac{\mu}{\delta}$. Finally, using the fact that $\lambda = \max\{\frac{u^T \mathcal{L} u}{u^T u} | u \neq 0\}$, we find that

$$(3) \quad \lambda \geq 2 - \frac{\mu}{\delta}.$$

Next, we consider the right hand inequality in (1). Now let y be an eigenvector of \mathcal{L} corresponding to λ , and partition y as $\begin{bmatrix} y_1 \\ -y_2 \end{bmatrix}$, where y_1 and y_2 are nonnegative vectors. Partition D and A conformally with y as $\left[\begin{array}{c|c} \hat{D}_1 & 0 \\ \hline 0 & \hat{D}_2 \end{array} \right]$ and $\left[\begin{array}{c|c} \hat{A}_{11} & \hat{A}_{12} \\ \hline \hat{A}_{21} & \hat{A}_{22} \end{array} \right]$ (observe that the partitioning for A and D here may be different than the partitioning arising from x).

Since $y^T \mathcal{L}y = \lambda y^T y$, we have

$$(4) \quad \begin{aligned} y_1^T y_1 - y_1^T \hat{D}_1^{-\frac{1}{2}} \hat{A}_{11} \hat{D}_1^{-\frac{1}{2}} y_1 + 2y_1^T \hat{D}_1^{-\frac{1}{2}} \hat{A}_{12} \hat{D}_2^{-\frac{1}{2}} y_2 + y_2^T y_2 - y_2^T \hat{D}_2^{-\frac{1}{2}} \hat{A}_{22} \hat{D}_2^{-\frac{1}{2}} y_2 \\ = \lambda(y_1^T y_1 + y_2^T y_2). \end{aligned}$$

Now let $w = \begin{bmatrix} \hat{D}_1^{\frac{1}{2}} y_1 \\ -\hat{D}_2^{\frac{1}{2}} y_2 \end{bmatrix}$, and note that $w^T Q w = y_1^T y_1 + y_1^T \hat{D}_1^{-\frac{1}{2}} \hat{A}_{11} \hat{D}_1^{-\frac{1}{2}} y_1 - 2y_1^T \hat{D}_1^{-\frac{1}{2}} \hat{A}_{12} \hat{D}_2^{-\frac{1}{2}} y_2 + y_2^T y_2 + y_2^T \hat{D}_2^{-\frac{1}{2}} \hat{A}_{22} \hat{D}_2^{-\frac{1}{2}} y_2$. From (4) we find that $w^T Q w = (2 - \lambda)y^T y$.

Since $\mu = \min\{\frac{u^T Q u}{u^T u} \mid u \neq 0\}$, we find that $\mu \leq (2 - \lambda) \frac{y^T y}{w^T w}$. Observing that $\frac{y^T y}{w^T w} \leq \Delta$, we thus find that $\mu \leq (2 - \lambda)\Delta$. Rearranging this yields

$$(5) \quad \lambda \leq 2 - \frac{\mu}{\Delta}.$$

□

Evidently if G is regular (so that $\delta = \Delta$) or bipartite (so that $\mu = 0$), then equality holds throughout (1). The next two examples show that for nonbipartite, nonregular graphs, equality can still hold in either of the inequalities in (1). Throughout, we use $\mathbf{1}_p$ to denote an all ones vector of order p ; the subscript will be suppressed when the order is clear from the context.

EXAMPLE 2.2. Suppose that $m \in \mathbb{N}$ with $m \geq 2$, and let G be the graph given by $(K_m \cup K_m) \vee K_1$, where \cup denotes the union and \vee denotes the join operation (see [12]). Observe that G is not bipartite, not regular, and has minimum degree m . The signless Laplacian matrix for G can be written as

$$Q = \left[\begin{array}{c|c|c} (m-1)I + J & 0 & \mathbf{1} \\ \hline 0 & (m-1)I + J & \mathbf{1} \\ \hline \mathbf{1}^T & \mathbf{1}^T & 2m \end{array} \right];$$

(here J denotes an all ones matrix). From this we find that the signless Laplacian spectrum consists of $2m - 1$, $\frac{4m-1 \pm \sqrt{8m+1}}{2}$, and $m - 1$, the latter with multiplicity $2m - 2$. Similarly, the normalised Laplacian matrix for G is

$$\mathcal{L} = \left[\begin{array}{c|c|c} \frac{m+1}{m}I - \frac{1}{m}J & 0 & -\frac{1}{m\sqrt{2}}\mathbf{1} \\ \hline 0 & \frac{m+1}{m}I - \frac{1}{m}J & -\frac{1}{m\sqrt{2}}\mathbf{1} \\ \hline -\frac{1}{m\sqrt{2}}\mathbf{1}^T & -\frac{1}{m\sqrt{2}}\mathbf{1}^T & 1 \end{array} \right];$$

we then find that the normalised Laplacian spectrum is given by 0 , $\frac{1}{m}$ and $\frac{m+1}{m}$, the latter with multiplicity $2m - 1$.

Consequently we find that the largest normalised Laplacian eigenvalue is $\lambda = \frac{m+1}{m}$, the smallest normalised Laplacian eigenvalue is $\mu = m - 1$, and that equality holds in (3).

EXAMPLE 2.3. Suppose that $m \in \mathbb{N}$ with $m \geq 3$, and consider the graph G on $2m + 1$ vertices constructed as follows: start with $K_{m,m}$, and delete an edge from it, say between vertices i and j ; then take an isolated vertex k , and add the edges $k \sim i$ and $k \sim j$. Observe that G is not regular, is not bipartite, and has maximum degree m .

The signless Laplacian matrix for G can be written as

$$Q = \begin{bmatrix} m & 0^T & 0 & \mathbf{1}^T & 1 \\ 0 & mI & \mathbf{1} & J & 0 \\ 0 & \mathbf{1}^T & m & 0^T & 1 \\ \mathbf{1} & J & 0 & mI & 0 \\ 1 & 0^T & 1 & 0^T & 2 \end{bmatrix}.$$

It now follows that the signless Laplacian spectrum of G consists of m (with multiplicity $2m - 4$), $\frac{m+1 \pm \sqrt{m^2+2m-3}}{2}$, and the roots of the cubic $z^3 - (3m + 1)z^2 + (2m^2 + 4m - 3)z - (4m^2 - 8m + 4)$. An uninteresting computation shows that the smallest signless Laplacian eigenvalue for G is $\mu = \frac{m+1 - \sqrt{m^2+2m-3}}{2}$.

The normalised Laplacian matrix for G is given by

$$\mathcal{L} = \begin{bmatrix} 1 & 0^T & 0 & -\frac{1}{m}\mathbf{1}^T & -\frac{1}{\sqrt{2m}} \\ 0 & I & -\frac{1}{m}\mathbf{1} & -\frac{1}{m}J & 0 \\ 0 & -\frac{1}{m}\mathbf{1}^T & 1 & 0^T & -\frac{1}{\sqrt{2m}} \\ -\frac{1}{m}\mathbf{1} & -\frac{1}{m}J & 0 & I & 0 \\ -\frac{1}{\sqrt{2m}} & 0^T & -\frac{1}{\sqrt{2m}} & 0^T & 1 \end{bmatrix}.$$

The eigenvalues of \mathcal{L} are $0, 1$ (with multiplicity $2m - 4$), $2 - \frac{1}{m}(\frac{m+1 \pm \sqrt{m^2+2m-3}}{2})$, and $\frac{2 + \frac{1}{m} \pm \sqrt{\frac{4}{m} - \frac{3}{m^2}}}{2}$. It follows that the largest normalised Laplacian eigenvalue is $\lambda = 2 - \frac{1}{m}(\frac{m+1 - \sqrt{m^2+2m-3}}{2})$, and that equality holds in (5).

3. Small Rayleigh quotients for the signless Laplacian matrix

From the results of section 2, we find that both the largest normalised Laplacian eigenvalue, and the smallest signless Laplacian eigenvalue can be thought of as providing a measure of how close a connected graph is to being bipartite. In this section, we focus on the signless Laplacian matrix, as the analysis for that matrix is somewhat more tractable than for the normalised Laplacian matrix.

Recall that the *algebraic connectivity* of a graph is the second smallest eigenvalue of its Laplacian matrix; see [1] and [11] for surveys on this remarkable quantity. The algebraic connectivity plays a role in the following result, which provides a sufficient condition for a subgraph to be bipartite.

THEOREM 3.1. *Let G be a graph on k vertices with signless Laplacian matrix Q , and let $x \in \mathbb{R}^k$ be a vector with at least one positive entry and at least one negative entry. By permuting the entries of x and simultaneously permuting the rows and columns of Q , we assume without loss of generality that $x_j > 0, j = 1, \dots, l, x_j < 0, j = l + 1, \dots, n$, and $x_j = 0, j = n + 1, \dots, k$. Let y denote the subvector of x on its first n entries, let $s = \begin{bmatrix} \mathbf{1}_l \\ -\mathbf{1}_{n-l} \end{bmatrix}$, let $z = y - \frac{y^T s}{n} s$, and let H_0 denote the subgraph of G induced by the edges in G of the form $i \sim j$ where $1 \leq i \leq l < j \leq n$. Set $\nu = \frac{x^T Q x}{x^T x}$ and $\theta = \frac{n^2 y^T y}{(y^T s)^2} - n$; denote the algebraic connectivity of H_0 by α , and let $\epsilon = \min\{(y_i + y_j)^2 | i, j = 1, \dots, l, i \neq j\} \cup \{(y_i + y_j)^2 | i, j = l + 1, \dots, n, i \neq j\}$. If*

$$(6) \quad \nu < \frac{\frac{n^2 \epsilon}{(y^T s)^2} + \alpha \theta}{n + \theta},$$

then the nonzero entries of x induce a bipartite subgraph of G .

Proof. We begin by remarking that θ provides a measure of how close y is to s , since $\theta = 0$ if and only if $y = s$, by the Cauchy-Schwarz inequality. Observing that z is orthogonal to s , we find readily that $z^T z = \frac{(y^T s)^2}{n^2} \theta$. We partition z conformally with s as $z = \begin{bmatrix} u \\ -v \end{bmatrix}$, and let $\tilde{z} = \begin{bmatrix} u \\ v \end{bmatrix}$; observe that $\tilde{z}^T \mathbf{1} = 0$. For each $j = 1, \dots, n$, let $d_0(j) = |\{r | j \sim r, n + 1 \leq r \leq k\}|$. Since $x^T Q x = \nu x^T x$, we find that

$$\nu y^T y = \sum_{i \sim j, 1 \leq i, j \leq l} (y_i + y_j)^2 + \sum_{i \sim j, l+1 \leq i, j \leq n} (y_i + y_j)^2 + \sum_{i \sim j, 1 \leq i \leq l, l+1 \leq j \leq n} (y_i + y_j)^2 + \sum_{j=1}^n d_0(j) y_j^2.$$

Hence,

$$\begin{aligned} \nu y^T y \geq & \sum_{i \sim j, 1 \leq i, j \leq l} (y_i + y_j)^2 + \sum_{i \sim j, l+1 \leq i, j \leq n} (y_i + y_j)^2 + \sum_{i \sim j, 1 \leq i \leq l, l+1 \leq j \leq n} (y_i + y_j)^2 = \\ & \sum_{i \sim j, 1 \leq i, j \leq l} (y_i + y_j)^2 + \sum_{i \sim j, l+1 \leq i, j \leq n} (y_i + y_j)^2 + \tilde{z}^T L(H_0) \tilde{z}, \end{aligned}$$

where $L(H_0)$ is the Laplacian matrix of H_0 . As $\tilde{z}^T \mathbf{1} = 0$, we have $\tilde{z}^T L(H_0) \tilde{z} \geq \alpha \tilde{z}^T \tilde{z} = \alpha z^T z$.

We proceed by contraposition, so suppose that the induced subgraph on vertices $1, \dots, n$, which we denote by H , is not bipartite. Then without loss of generality we may assume that $l \geq 2$ and that H contains the edge $1 \sim 2$. From the above, we have $\nu y^T y \geq (y_1 + y_2)^2 + \alpha z^T z \geq \epsilon + \alpha z^T z$. Since $z^T z = \frac{(y^T s)^2}{n^2} \theta$ and $y^T y = \frac{(y^T s)^2}{n} + z^T z$, we find that

$$\nu \geq \frac{\epsilon + \alpha \theta \frac{(y^T s)^2}{n^2}}{(n + \theta) \frac{(y^T s)^2}{n^2}}.$$

The conclusion now follows. \square

EXAMPLE 3.2. In this example we illustrate the fact that, in the context of Theorem 3.1, some constraint on ν is needed in order to conclude that the subgraph is bipartite.

Here we consider K_3 , whose corresponding signless Laplacian matrix is $Q = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. Let $x = \begin{bmatrix} \frac{3}{4} \\ \frac{3}{4} \\ \frac{-3}{2} \end{bmatrix}$. We find readily that $\nu = \frac{x^T Q x}{x^T x} = 1$, $y^T s = 3 = n$, $\theta = \frac{3}{8}$ and $\epsilon = \frac{9}{4}$; the subgraph H_0 induced by the edges $1 \sim 3, 2 \sim 3$ is $K_{1,2}$, so $\alpha = 1$. All of this yields $\frac{\epsilon + \alpha \theta}{n + \theta} = \frac{7}{9}$, which is of course less than ν , as K_3 is not bipartite.

REMARK 3.3. Suppose that in the context of Theorem 3.1, the subgraph of G on vertices $1, \dots, n - H$ say - is bipartite. Letting R denote the principal submatrix of Q on rows and columns $1, \dots, n$, we find readily that $s^T R s$ is the number of edges between vertices in H and vertices in $G \setminus H$. It is not difficult to determine that $s^T R s \geq \mu n$, where μ denotes the smallest eigenvalue of Q .

It is natural to anticipate that if ν is small, then the number of edges between H and $G \setminus H$ will also be small; the results below are an attempt to reinforce that intuition.

In our next result, we continue with the notation of Theorem 3.1.

THEOREM 3.4. *Suppose that $\nu < \alpha$, and let $q = |\{j | d_0(j) \geq 1, j = 1, \dots, n\}|$. Then $q \leq \nu n \left(\frac{1 + \alpha - \nu}{\alpha - \nu} \right)$.*

Proof: Let H denote the subgraph of G induced by vertices $1, \dots, n$. We begin by remarking that q denotes the number of vertices in the subgraph H that are adjacent to at least one vertex in $G \setminus H$. Suppose for concreteness that $d_0(j_p) \geq 1$ for $p = 1, \dots, q$. Arguing as in the proof of Theorem 3.1, we have $\nu n - (\alpha - \nu) z^T z \geq \sum_{p=1}^q d_0(j_p) (1 + z_{j_p})^2 \geq \sum_{p=1}^q (1 + z_{j_p})^2$.

Set $w^T = [z_{j_1} \ \dots \ z_{j_q}]$. Then $\nu n - (\alpha - \nu)w^T w \geq \sum_{i=1}^q (1 + w_i)^2$. Applying the Cauchy-Schwarz inequality (twice), and letting $\beta = w^T \mathbf{1}$, we find that $\nu n - (\alpha - \nu)\frac{\beta^2}{q} \geq \frac{(q + \beta)^2}{q}$. Consequently, we find that $(1 + \alpha - \nu)\beta^2 + 2q\beta + q^2 - \nu qn \leq 0$, and by considering the left hand side as a quadratic in β , and then minimising that quadratic, we find that $-\frac{q^2}{1 + \alpha - \nu} + q^2 - \nu qn \leq 0$. It now follows that $q \leq \nu n \left(\frac{1 + \alpha - \nu}{\alpha - \nu} \right)$, as desired. \square

REMARK 3.5. In the context of Theorem 3.4, we always (trivially) have $q \leq n$, so Theorem 3.4 only yields useful information in the case that $\frac{\nu(1 + \alpha - \nu)}{\alpha - \nu} < 1$. It is straightforward to determine that this last holds only when either $\nu < \frac{2 + \alpha - \sqrt{4 + \alpha^2}}{2}$ or $\nu > \frac{2 + \alpha + \sqrt{4 + \alpha^2}}{2}$. Evidently the latter condition violates the hypothesis that $\nu < \alpha$ in Theorem 3.4, so we conclude that Theorem 3.4 yields a nontrivial bound on q when $\nu < \frac{2 + \alpha - \sqrt{4 + \alpha^2}}{2}$.

As above, we continue with the notation of Theorem 3.1. In the special case that the vector x of Theorem 3.1 is an eigenvector corresponding to the smallest signless Laplacian eigenvalue, we can derive an upper bound on the number of edges between H and $G \setminus H$.

THEOREM 3.6. *Let G be a graph on k vertices, and let x be an eigenvector corresponding to the smallest eigenvalue μ of Q . Suppose that $x_j > 0, j = 1, \dots, l, x_j < 0, j = l + 1, \dots, n$, and $x_j = 0, j = n + 1, \dots, k$. Let y denote the subvector of x on its first n entries, let $s = \begin{bmatrix} \mathbf{1}_l \\ -\mathbf{1}_{n-l} \end{bmatrix}$. Let $\theta = \frac{n^2 y^T y}{(y^T s)^2} - n$, let R denote the principal submatrix of Q on vertices $1, \dots, n$, and denote the largest eigenvalue of R by ρ .*

$$\text{Then } s^T R s \leq \mu \frac{n^2}{n + \theta} + \rho \frac{n\theta}{n + \theta}.$$

Proof: Note that y is an eigenvalue of R corresponding to its smallest eigenvalue, which is necessarily μ . Denote the remaining eigenvalues of R by $\lambda_2 \leq \dots \leq \lambda_n \equiv \rho$, and let $u_j, j = 2, \dots, n$ denote the corresponding eigenvectors, which we take (without loss of generality) to be pairwise orthogonal, and orthogonal to y . We find that $s = \frac{y^T s}{y^T y} y + \sum_{j=2}^n \frac{u_j^T s}{u_j^T u_j} u_j$. In particular, we find that $n = s^T s = \frac{(y^T s)^2}{y^T y} + \sum_{j=2}^n \frac{(u_j^T s)^2}{u_j^T u_j}$.

Next, we note that $s^T R s = \mu \frac{(y^T s)^2}{y^T y} + \sum_{j=2}^n \lambda_j \frac{(u_j^T s)^2}{u_j^T u_j} \leq \mu \frac{(y^T s)^2}{y^T y} + \rho \sum_{j=2}^n \frac{(u_j^T s)^2}{u_j^T u_j} = \mu \frac{(y^T s)^2}{y^T y} + \rho \left(n - \frac{(y^T s)^2}{y^T y} \right)$. Using the fact that $\theta = \frac{n^2 y^T y}{(y^T s)^2} - n$, and rewriting the inequality slightly now yields the desired conclusion. \square

EXAMPLE 3.7. In this example, we revisit the graph G of Example 2.3, and use it to illustrate the results of Theorems 3.1, 3.4 and 3.6. As we saw earlier, the signless Laplacian matrix for G can be written as

$$Q = \left[\begin{array}{c|c|c|c|c} m & 0^T & 0 & \mathbf{1}^T & 1 \\ \hline 0 & mI & \mathbf{1} & J & 0 \\ \hline 0 & \mathbf{1}^T & m & 0^T & 1 \\ \hline \mathbf{1} & J & 0 & mI & 0 \\ \hline 1 & 0^T & 1 & 0^T & 2 \end{array} \right],$$

and the smallest signless Laplacian eigenvalue is $\mu = \frac{m+1-\sqrt{m^2+2m-3}}{2}$. It is straight-

forward to determine that the vector $x = \begin{bmatrix} 1 - \mu \\ \mathbf{1}_{m-1} \\ -1 + \mu \\ -\mathbf{1}_{m-1} \\ 0 \end{bmatrix}$ serves as a μ -eigenvector.

Applying Theorem 3.1 with the vector x , we note that the subgraph H identified by the nonzero entries of x is the subgraph of G formed by deleting the vertex of degree 2. Computing the relevant quantities, we find that $\nu = \mu$, $\epsilon = (2 - \mu)^2$, $y^T s = 2(m - \mu)$, $n = 2m$, $\alpha = \frac{3m-2-\sqrt{m^2+4m-2}}{2}$ and $\theta = \frac{2m(m-1)\mu^2}{(m-\mu)^2}$. It is straightforward to show that as $m \rightarrow \infty$, μ is asymptotic to $\frac{1}{m}$; on the other hand, the right-hand side of (6) is asymptotic to $\frac{2}{m}$ as $m \rightarrow \infty$, so that for all sufficiently large values of m , we find that (6) is satisfied.

Referring to Theorem 3.4, we find that for the subgraph H of G , the value of q is 2. Since α is readily seen to be asymptotic to $m - 2$ as $m \rightarrow \infty$, it follows that as $m \rightarrow \infty$, the expression $\nu n \left(\frac{1+\alpha-\nu}{\alpha-\nu} \right)$ converges to 2 as $m \rightarrow \infty$. Thus we find that for all sufficiently large values of m , the upper bound on q furnished by Theorem 3.4 is accurate.

Turning to Theorem 3.6, we find that for the subgraph H , we have $s^T R s = 2$. Note that the largest eigenvalue ρ of R is readily seen to lie between $2m - 1$ and $2m$. It now follows that the expression $\mu \frac{n^2}{n+\theta} + \rho \frac{n\theta}{n+\theta}$ converges to 2 as $m \rightarrow \infty$, so that Theorem 3.6 provides an accurate estimate of $s^T R s$ for this example.

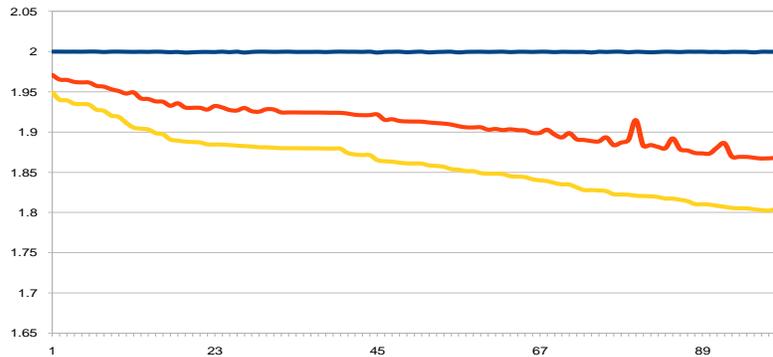


FIGURE 1. $2 - \frac{\mu}{\delta} \leq \lambda \leq 2 - \frac{\mu}{\Delta}$

4. Computations and applications

4.1. Inequality (1) for large graphs. Using *NetworkX* we randomly generated 100 multigraphs, each on 1000 vertices, whose degree sequences followed a power law distribution with exponent 2.1. Loops and multiple edges were then removed to generate simple loop-free graphs whose degree sequences were close to following a power law distribution. For each such graph, we considered the connected component with the largest number of vertices, and for that component, we computed the largest normalised Laplacian eigenvalue λ , as well as the expressions $2 - \frac{\mu}{\delta}$ and $2 - \frac{\mu}{\Delta}$, where μ is the smallest signless Laplacian eigenvalue, and where δ, Δ represent the minimum and maximum degrees, respectively. The results are depicted in Figure 1, which plots $2 - \frac{\mu}{\delta}$ (blue), λ (red), and $2 - \frac{\mu}{\Delta}$ (yellow), for each graph; here the graphs were sorted according to increasing values of μ . The values of μ for these examples ranged between approximately 0.05055 and 0.23191. Since the graphs that were generated have degree sequences that are roughly distributed according to a power law, their maximum degrees are typically quite large, while the minimum degrees are quite small. These observations are reflected in Figure 1:

the blue graph corresponding to $2 - \frac{\mu}{\Delta}$ is approximately constant, while the yellow graph corresponding to $2 - \frac{\mu}{\delta}$ is more sensitive to the value of μ . Note that for all of the graphs, the quantities $2 - \frac{\mu}{\delta}$ and $2 - \frac{\mu}{\Delta}$ provide a fairly small interval in which λ is contained.

4.2. Scale-free networks. Theorem 3.1 suggests a strategy for identifying bipartite subgraphs of a given graph G . The approach is as follows:

- a) compute a unit eigenvector v corresponding to the smallest signless Laplacian eigenvalue for G ;
- b) construct the vector x from v by setting its entries of small absolute value equal to zero;
- c) if it happens that (6) holds, then the nonzero entries of x induce a bipartite subgraph of G .

We implemented that approach on a collection of randomly generated graphs. Specifically, using *NetworkX* we randomly generated 500 multigraphs, each on 600 vertices, whose degree sequences followed a power law distribution with exponent 2.1. Loops and multiple edges were then removed to generate simple loop-free graphs whose degree sequences were close to following a power law distribution. For each graph so generated, we considered the connected component G containing the maximum number of vertices, and computed the unit eigenvector v_G for the smallest eigenvalue of its signless Laplacian matrix.

Next, we formed the vector x_G from v_G by rounding all of its entries to the first decimal place. We then considered the subgraph H_G induced by the nonzero entries of x_G . We remark here that our decision to round the entries of v_G to the first decimal place is motivated by a need to introduce some zeros into the vectors with which we work, since an actual value of 0 is a rare occurrence when computing eigenvectors numerically. Evidently different strategies for approximating v_G will yield different subgraphs; however the results below suggest that our rounding method is reasonably successful in identifying bipartite subgraphs that are weakly connected to the rest of the graph G .

For a total of 403 different graphs G , the subgraphs H_G were connected, and the corresponding vectors x_G satisfied (6), so that bipartiteness was assured by Theorem 3.1. These bipartite subgraphs were all of small order, containing between 3 and 10 vertices, while the original connected graphs had orders ranging from 548 to 598 vertices. For each of these 403 graphs the number of edges between H_G and the rest of G was at most 9; for 208 of these graphs, there was just one edge between H_G and the rest of G .

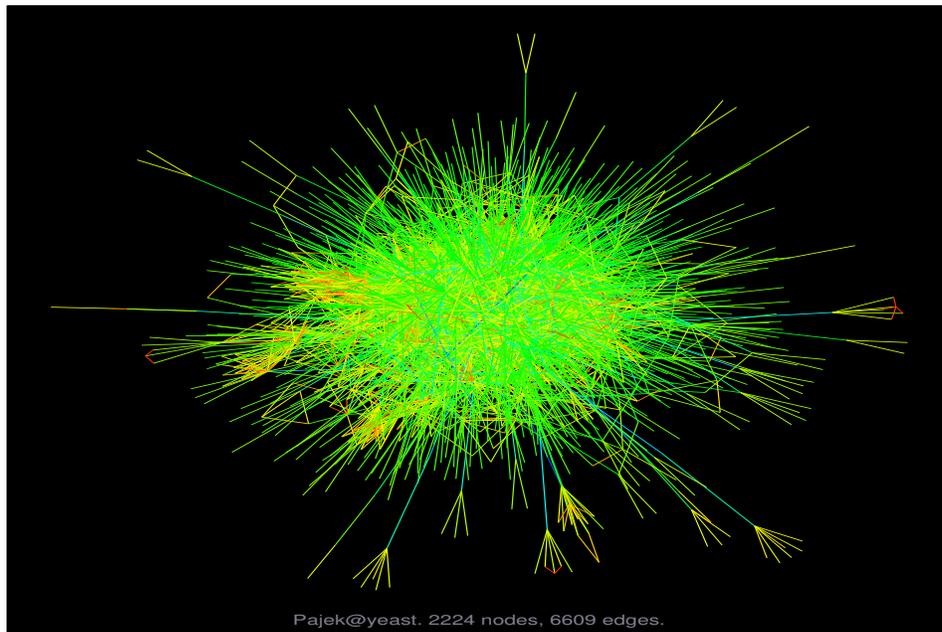


FIGURE 2. The connected component \hat{G}

For a further 63 graphs G , the subgraphs H_G were connected and bipartite, but (6) did not hold. Again, the H_G 's were of small order – between 3 and 16 vertices – while the orders of the connected graphs G ranged from 568 to 592. In each of these cases, there were at most 11 edges connecting H_G to the rest of G . For the remaining 34 graphs, the subgraph H_G was either not connected, or not bipartite, or contained just two vertices.

Based on these computations, it appears that there is some utility in the approach to identifying bipartite subgraphs suggested by Theorem 3.1.

4.3. A protein-protein interaction network. In [2], the authors consider protein-protein interaction networks, and are interested in identifying bipartite (or nearly bipartite) subgraphs within such networks; for, finding such subgraphs leads to an enhanced understanding of the function of the corresponding proteins. The approach to identifying such bipartite subgraphs proposed in [2] is to consider the network as a graph, and then use eigenvectors of the adjacency matrix corresponding to small eigenvalues in order to identify bipartite subgraphs. For the example of a protein-protein interaction network on 2617 vertices (corresponding to budding yeast), it is reported in [2] that six so-called quasi-bipartite subgraphs are identified by this adjacency matrix eigenvector method.

In view of the results in section 3, it is natural to conjecture that eigenvectors of the signless Laplacian matrix corresponding to small eigenvalues may also yield an effective technique for identifying bipartite subgraphs of a protein-protein interaction network. Below we describe the results of an implementation of this idea.

We used the dataset describing a protein-protein interaction network for budding yeast available at <http://vlado.fmf.uni-lj.si/pub/networks/data/bio/Yeast/Yeast.htm>. That network has 2361 vertices, and 7182 edges, of which 536 are loops. We removed the loops to yield a loop-free graph G on 2361 vertices with 6646 edges. The graph G consists of 101 connected components. One hundred of these connected components are either trees (and so, necessarily bipartite) or isolated vertices, of orders ranging from one to eight vertices; taken together they contain 137 vertices and 37 edges. The remaining connected component of G , which we denote by \hat{G} , is a graph on 2224 vertices, with 6609 edges. Since the smallest signless Laplacian eigenvalue for \hat{G} is positive (approximately 0.0609), \hat{G} is not bipartite. Figure 2 gives a depiction of \hat{G} ; the source image can be found at <http://www2.research.att.com/~yifanhu/GALLERY/GRAPHS/PDF/Pajek@yeast.pdf>.

We considered unit eigenvectors of the signless Laplacian matrix for \hat{G} corresponding to its 50 smallest eigenvalues. For each such unit eigenvector v , we formed the vector x_v by rounding the entries of v to the first decimal place. We then considered the subgraph of \hat{G} , say H_v , induced by the nonzero entries of x_v .

For a total of 12 such eigenvectors v , the corresponding subgraph H_v was connected and the vector x_v satisfied (6), thus ensuring that H_v was bipartite. These subgraphs were of small order, between 3 and 8 vertices, and in each case were joined to the rest of \hat{G} by at most 15 edges. For 7 of these H_v 's, there was just one edge joining H_v to the rest of \hat{G} . For a further 15 eigenvectors v , the corresponding subgraph H_v was connected and bipartite, but (6) was not satisfied by x_v . Again the subgraphs were of orders between 3 and 8, and were joined to the rest of \hat{G} by at most 12 edges. For the remaining 23 eigenvectors under consideration, the corresponding subgraph was either not connected, or not bipartite, or consisted of just two vertices.

Based on these results, it seems that the technique of using signless Laplacian eigenvectors for small eigenvalues to identify bipartite subgraphs represents an improvement on the adjacency eigenvector approach employed in [2].

Acknowledgement: S.K.'s research is supported in part by the Science Foundation Ireland under Grant No. SFI/07/SK/I1216b. D.P. gratefully acknowledges

Professor Mahantapas Kundu, Department of Computer Science and Engineering,
Jadavpur University, for his constant support and encouragement.

Bibliography

- [1] N. Abreu, Old and new results on algebraic connectivity of graphs, *Linear Algebra and its Applications* 423(1), 53–73, 2007.
- [2] D. Bu, Y. Zhao, L. Cai, H. Xue, X. Zhu, H. Lu, J. Zhang, S. Sun, L. Ling, N. Zhang, G. Li and R. Chen, Topological structure analysis of the protein–protein interaction network in budding yeast, *Nucleic Acids Research*, 31(9): 2443–2450, 2003.
- [3] F. Chung, *Spectral Graph Theory*, American Mathematical Society, Providence, 1997.
- [4] D. Cvtekovic, P. Rowlinson and S. Simic, Signless Laplacians of finite graphs, *Linear Algebra and its Applications* 423(1): 155-171, 2007.
- [5] D. Cvtekovic and S. Simic, Towards a spectral theory of graphs based on signless Laplacian I, *I. Publ. Inst. Math. (Beograd)*, 85(99):19-33, 2009.
- [6] D. Cvtekovic and S. Simic, Towards a spectral theory of graphs based on the signless Laplacian II, *Linear Algebra and its Applications* 432(9), 2257-2272, 2010.
- [7] D. Cvtekovic and S. Simic, Towards a spectral theory of graphs based on the signless Laplacian III, *Applicable Analysis and Discrete Mathematics* 4(1), 156-166, 2010.
- [8] M. Fiedler, Algebraic connectivity of graphs, *Czechoslovak Mathematical Journal* 23(98), 298–305, 1973.
- [9] J.-L. Guillaume and M. Latapy, Bipartite structure of all complex networks, *Information Processing Letters*, 90(5):215-221, 2004.
- [10] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
- [11] S. Kirkland, Algebraic connectivity, in *Handbook of Linear Algebra*, L. Hogben, ed. Chapman Hall/CRC, Boca Raton, 36-1–36-12, 2007.
- [12] R. Merris, *Graph Theory*, Wiley, New York, 2001.
- [13] J. Morrison, R. Breitling, D. Higham and D. Gilbert, A lock-and-key model for protein-protein interactions, *Bioinformatics* 22(16): 2012-2019, 2006.
- [14] M. Newman, The structure and function of complex networks, *SIAM Review* 45(2), 167-256, 2003.
- [15] A. Thomas, R. Cannings, N. Monk, and C. Cannings, On the structure of protein-protein interaction networks, *Biochemical Society Transactions*, 31(6):1491-1496, 2003.

Steve Kirkland, Hamilton Institute, National University of Ireland Maynooth, Ireland, stephen.kirkland@nuim.ie

Debdas Paul, Department of Computer Science and Engineering, Jadavpur University, Kolkata, India, paul.debdas@gmail.com