Rank and Number of Nodal Domains of Cographs

Türker Bıyıkoğlu

Department for Applied Statistics and Data Processing,

University of Economics and Business Administration, Augasse 2-6,

A-1090 Vienna, Austria

E-mail: `tuerker@statistik.wu-wien.ac.at`

Phone: **43 1 31336-4180 Fax : **43 1 31336-738

Institute for Theoretical Chemistry and Structural Biology,

University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria

November 29, 2002

Abstract

A nodal domain of a real vector associated with a graph is the maximal induced subgraph of a graph on which the vector does not change sign. We characterized the (maximum and minimum) number of nodal domains of eigenvectors of Laplacian matrix of a cograph. We also showed that the rank of the adjacency matrix of a cograph is equal to the number of distinct nonzero columns of the adjacency matrix.

Keywords: discrete nodal domain theorem; cograph; threshold graph; graph Laplacian; Sign graph; rank of cograph.

1 Introduction

A graph G is called *cograph* if G has no induced subgraph P_4. Cographs arise in many disparate areas of mathematics and computer science. In this article we consider the Laplacian and adjacency matrices of cographs. We characterize the number of nodal domains of cographs with respect to the Laplacian matrix, and the rank of the adjacency matrix of a cograph.

Let $G = (V, E)$ be a graph with vertex set $V = \{1, \dots, n\}$ and edge set E and let $x = (x_1, \dots, x_n)$ be a real vector. We associate the real numbers x_i with the vertices i of G, for $i = 1, \dots, n$. A *positive (negative) nodal domain* is a maximal connected induced subgraph of G on vertices $i \in V$ with $x_i > 0$ $(x_i < 0)$. We denote by $\eta(x)$ the number of nodal domains of the vector x. For example, let G be the path P_6 and consider the vector $x = (1, 2, -1, 0, -1, 3)$. The vector x has two positive nodal domains, two negative nodal domains, and hence $\eta(x) = 4$.

Let G be a simple, undirected, loop-free graph with n vertices. We call a symmetric real $n \times n$ matrix M a *generalized Laplacian* of G if $m_{uv} < 0$ when u and v are adjacent vertices of G and $m_{uv} = 0$ when u and v are distinct and not adjacent. There are no constraints on the diagonal entries of M. An important example is the *Laplacian matrix* $L(G) = D(G) - A(G)$, where $A(G)$ is the adjacency matrix of G and $D(G)$ is the diagonal matrix of vertex degrees.

Let $\lambda_1 \leq \cdots \leq \lambda_n$ be the eigenvalues of a generalized Laplacian of G. Then any eigenvector corresponding to eigenvalue λ_k with multiplicity r has at most $k + r - 1$ nodal domains [6]. This theorem is called the *discrete nodal domain theorem* and it is the discrete analogue of Courant's nodal domain theorem for elliptic operators on Riemannian manifolds, see e.g. [3].

The number of nodal domains can be much smaller than the bound obtained from the discrete nodal domain theorem. It is also not easy to find the maximum or minimum number of nodal domains. For example, for a tree with n vertices: we can find for each eigenvalue the maximum number of nodal domains is in $O(n^2)$ time [1]. On the other hand to find the minimum number of nodal domains is NP-complete [1]. For the hypercubes the complete spectral information is available for Laplacian matrix but it is an open problem for most eigenvalues to find the minimum or maximum number of nodal domains [2]. As we see this problem is much easier for cographs than that for trees.

2 The Number of Nodal Domains of Cographs

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be graphs on disjoint sets of r and s vertices, respectively. Their *disjoint union* $G_1 + G_2$ is the graph $G_1 + G_2 =$

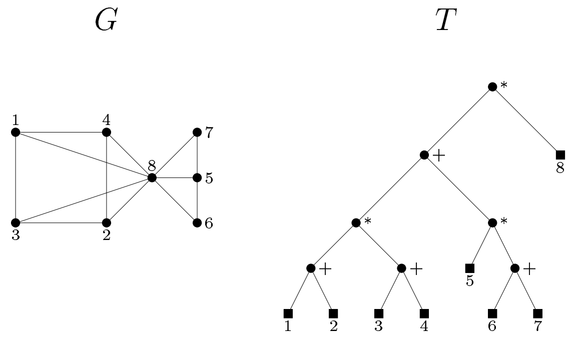

Figure 1: The cograph G and the cotree T of G

$(V_1 \cup V_2, E_1 \cup E_2)$, and their *join* $G_1 * G_2$ is the graph on $n = r + s$ vertices obtained from $G_1 + G_2$ by inserting new edges from each vertex of G_1 to each vertex of G_2.

Cographs have several characterizations. The following tree representation with join and disjoint union operations is more suitable for our purpose.

Lemma 1 ([5]) *To each cograph $G = (V, E)$, one can associate a unique rooted tree T, called the cotree of G. Each leaf node of T corresponds to a (unique) vertex of V. Each internal node is labeled with a $*$ or a $+$. Children of nodes labeled with $+$ are labeled with $*$, and vice versa. It is possible to associate a cograph with each node of the cotree T. Leaf nodes correspond to the cograph with the one vertex they represent. Internal nodes labeled with $*(+)$ correspond to the join (disjoint union) of the cographs, corresponding to the children of the node (see Figure 1). G equals the cograph corresponding with the root of T. Cographs can be in $O(|V| + |E|)$ time recognized, and in the same time the corresponding cotree can be built.*

It means that each cograph G is the disjoint union of two disjoint cographs G_1 and G_2, $G = G_1 + G_2$ or G is the join of two disjoint cographs G_1 and G_2, $G = G_1 * G_2$.

It is well-known that the eigenvalues of Laplacian matrix $L(G)$ of a graph G with n vertices are $0 = \lambda_1 \leq \cdots \leq \lambda_n$ and the trivial eigenvalue $\lambda_1 = 0$

has the eigenvector $e_n = (1, \ldots, 1)$. Each eigenvector of $L(G)$, which is orthogonal to e_n, has at least two entries with opposite sign.

The following lemma invites to look for the nodal domains of cographs.

Lemma 2 ([8]) *Let G_1 and G_2 be graphs on disjoint sets of r and s vertices, respectively. If $\mu_1 \leq \cdots \leq \mu_r$ and $\nu_1 \leq \cdots \leq \nu_s$ are eigenvalues of Laplace matrix of G_1 and G_2, respectively. Then the eigenvalues of $G_1 * G_2$ are $n = r + s$; $\mu_2 + s, \ldots, \mu_r + s$; $\nu_2 + r, \ldots, \nu_s + r$; and 0. Suppose y is an eigenvector of G_1 that is orthogonal to e_r. Extend y to $G_1 * G_2$ by defining it to be zero on $V(G_2)$. If y affords the eigenvalue μ, the extension of y is an eigenvector of $G_1 * G_2$ affording $\mu + s$. Similarly an eigenvector of G_2 affording ν extends to an eigenvector of $G_1 * G_2$ affording $\nu + r$. The eigenvalue $\lambda = r + s$ corresponds to an eigenvector whose value is $-s$ on each of the r vertices of G_1 and r on each of the s vertices of G_2. Finally, the trivial eigenvalue is afforded by e_{r+s}.*

Obviously, the eigenvalues of the Laplace matrix of $G_1 + G_2$ are the union of eigenvalues of G_1 and G_2 (respecting multiplicity). It follows from Lemmas 1 and 2 that the Laplacian eigenvalues of a cograph are integers and easy to compute from its cotree.

Let T be a rooted tree and let v be a node of T. A *subtree at* v is the induced tree by v and all descendants of v. Similarly, a *subtree of* v is the subtree at one of the children of v.

Theorem 1 *For each eigenvalue of the Laplace matrix of a cograph $G = (V, E)$ we can find an eigenvector with maximum or minimum number of nodal domains in $O(|V| + |E|)$ time.*

Proof: By Lemma 1 a cograph G has a unique cotree T. Let v be a node of the cotree T with subtrees $T_1, \ldots, T_k$ and $G_1, \ldots, G_k$ the respective cographs. Let G_v be the cograph corresponding with v as root. Now we show that the number of nodal domains of G_v can be expressed in terms of the number of nodal domains of $G_1, \ldots, G_k$. Let $\mathrm{MaxND}(\lambda)$ and $\mathrm{MinND}(\lambda)$ be eigenvectors of the eigenvalue λ with maximum and minimum number of nodal domains, respectively.

If v has the label $+$ (disjoint union), then the eigenvalues of G_v are the union of eigenvalues of $G_1, \ldots, G_k$. Let $x^1, \ldots, x^k$ be the eigenvectors of λ with maximum number of nodal domains. Then $x = (x^1, \ldots, x^k)$ is the eigenvector of λ of cograph G_v with maximum number of nodal domains and $\eta(x) = \sum_{i=1}^{n} \eta(x^i)$. Similarly let $y^1, \dots, y^k$ be the eigenvectors of λ with minimum number of nodal domains. Then $y = (0, \dots, 0, y^i, 0, \dots, 0)$ is the eigenvector of λ and $\mathrm{MinND}(\lambda) = \min\{\eta(y^1), \dots, \eta(y^k)\}$.

If v has the label $*$ (join operation), then an easy induction gives that the eigenvalues of G_v are $|V(G_v)|$ and $\lambda_{G_i} + \sum_{j \neq i} |V(G_j)|$, where $\lambda_{G_i} > 0$ is an eigenvalue of G_i for $i = 1, \dots, k$. By Lemma 2 the extension $(0, \dots, 0, x^i, 0, \dots, 0)$ of the eigenvector x^i of λ_{G_i} is an eigenvector of $\mu = \lambda_{G_i} + \sum_{j \neq i} |V(G_j)|$. The eigenvectors $\{(0, \dots, 0, x^{i_1}, 0, \dots, 0), \dots, (0, \dots, 0, x^{i_p}, 0, \dots, 0)\}$ span the eigenspace of $\mu \neq |V(G_v)|$ with respect to the choice of the basis of $\lambda_{G_{i_1}}, \dots, \lambda_{G_{i_p}}$, where $\mu = \lambda_{G_{i_1}} + \sum_{j \neq i_1} |V(G_j)| = \cdots = \lambda_{G_{i_p}} + \sum_{j \neq i_p} |V(G_j)|$. The eigenvectors $x^{i_1}, \dots, x^{i_p}$ have at least two vertices with opposite sign. By join operation all linear combinations of $(0, \dots, 0, x^{i_1}, 0, \dots, 0), \dots, (0, \dots, 0, x^{i_p}, 0, \dots, 0)$ have two nodal domains. Therefore $\mathrm{MaxND}(\mu) = \max\{\mathrm{MaxND}(\lambda_{G_{i_1}}), \dots, \mathrm{MaxND}(\lambda_{G_{i_p}})\}$. Similarly, $\mathrm{MinND}(\mu) = 2$ for $p \geq 2$ and $\mathrm{MinND}(\mu) = \mathrm{MinND}(\lambda_{G_{i_1}})$ for $p = 1$. For the eigenvalue $\mu = |V(G_v)|$ by Lemma 1 the children of the node v are labeled with $+$. Therefore each of the graphs $G_1, \dots, G_k$ is either not connected or a single vertex. Let $c_1, \dots, c_k$ be the number of connected components of $G_1, \dots, G_k$. By Lemma 2 it is easy to see that $\mathrm{MaxND}(\mu) = \max\{c_1, \dots, c_k\} + 1$ when the node v has more than two children and $\mathrm{MaxND}(\mu) = c_1 + c_2$ when v has two children.

We have shown that it is enough to build the cotree of a cograph to find $\mathrm{MaxND}(\lambda)$ or $\mathrm{MinND}(\lambda)$. By Lemma 1 we can build the cotree of $G = (V, E)$ in $O(|V| + |E|)$ time. □

Corollary 1 *The Laplacian eigenvalues of a complete k-partite graph $K_{n_1,\dots,n_k}$ with $n_1 \geq \cdots \geq n_k$ are 0; $n = n_1 + \cdots + n_k$; and $n - n_i$, for $i = 1, \dots, k$. The maximum number of nodal domains of eigenvalues n and $n - n_i$ are equal to $n_1 + n_2$ and n_i, respectively. The minimum number of nodal domains of all eigenvalues are equal to two.*

For an important subclass of cographs, namely threshold graphs, we can directly compute the number of nodal domains without using Theorem 1. A graph $G = (V, E)$ is called as a *threshold graph*, if G does not contain one of the three forbidden induced subgraph graphs, $K_2 + K_2$, C_4, or P_4. Another useful characterization of threshold graph is the following.

Lemma 3 ([4]) *G is a connected threshold graph if and only if $G = (K, U)$, where K is a complete graph with a partition of non empty cliques $K_1, \dots, K_s$*

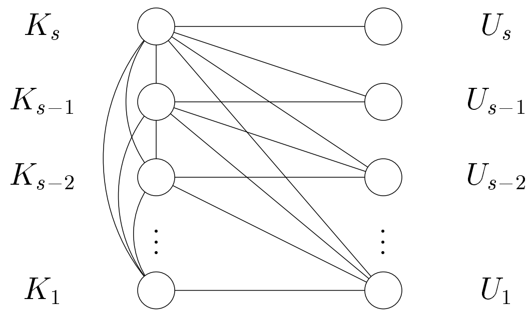

Figure 2: The typical structure of a threshold graph. A line between cells K_i and U_j indicates that each vertex in K_i is adjacent to each vertex of U_j.

and U is an independent set with a partition of non empty independent sets $U_1, \ldots, U_s$. All vertices of K_i are adjacent with all vertices of U_h, for $1 \leq h \leq i$ and for $i = 1, \ldots, s$. (see Figure 2)

In the next section we use this lemma to characterize the rank of a cograph.

By Lemmas 2 and 3 the Laplacian eigenvalues of a threshold graph are obtained easily by induction; for a similar procedure see [7].

Corollary 2 *Let $G = (K, U)$ be a connected threshold graph with the partitions K_i and U_i, for $i = 1, \ldots, s$. The eigenvalues of the Laplacian matrix of G are 0; $\sum_{i=1}^{h} |U_i| + \sum_{j=1}^{s} |K_j|$ for $h = 1, \ldots, s$; $\sum_{j=h}^{s} |K_j|$ for $h = 2, \ldots, s$; $\sum_{j=1}^{s} |K_j|$ when $|U_1| \geq 2$. The bounds for the number of nodal domains are:*

(i) If $\lambda = \sum_{i=1}^{h} |U_i| + \sum_{j=1}^{s} |K_j|$, then

1. $2 \leq \eta(x) \leq |U_h| + 2$ *when* $h \geq 2$,
2. $2 \leq \eta(x) \leq |U_1| + 1$ *when* $h = 1$.

(ii) If $\lambda = \sum_{j=h}^{s} |K_j|$, then

1. $2 \leq \eta(x) \leq |U_h| + 1$ *when* $h \geq 2$,
2. $\eta(x) \leq |U_1|$ *when* $h = 1$ *and* $|U_1| \geq 2$.

These bounds on $\eta(x)$ are sharp. The special case $G = K_n$ is trivial.

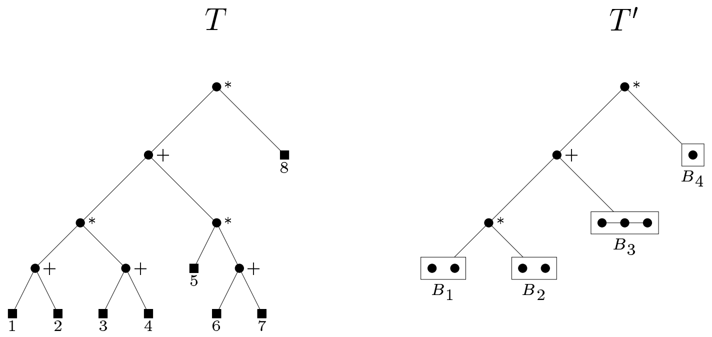

Figure 3: The cotree T and the threshold cotree T' from T with leaves $B_1, \ldots, B_4$.

3 The Rank of Cographs

In this section we prove the following conjecture of T. Sillke about the rank of cographs [10]: The rank of the adjacency matrix $A(G)$ of a cograph G is equal to the number of distinct nonzero columns of $A(G)$.

First we characterize cographs G, where all columns of $A(G)$ are different. We prove that adjacency matrices of such cographs have also full rank.

By Lemma 1 we know that each cograph G has an associated rooted cotree T. It is easy to see that we can get a new tree T' from T such that the leaf nodes of T' correspond to the set of threshold graphs. We look the *threshold maximal* tree T' with leaf nodes $B_1, \ldots, B_s$, that means T' has no internal node with children B_i and B_j such that the corresponding cograph of $B_i * B_j$ $(B_i + B_j)$ is a set of threshold graphs (see Figure 3). We call T' as *threshold cotree* of G.

Proposition 1 *Let $G = (V, E)$ be a cograph and T its (threshold) cotree T. Let v be a node of T and L the subtree at v. Let F be vertices of G which correspond to leaf nodes of L. Then the vertices of F have same neighbors in $V - F$ (we say also outside of F or outside of the subtree L).*

Proof: It is easy to see by induction from leaves to the root. $\square$

Lemma 4 *Let G be cograph and $A(G)$ its adjacency matrix. All columns of $A(G)$ are distinct and nonzero if and only if G has a threshold cotree T with leaves $B_1, \ldots, B_s$, where B_j are a set of threshold graphs and it holds:*

(i) Each independent set U_i of a threshold graph of B_j has at most one vertex.

(ii) All internal nodes of T with label $+$ have at most one subtree L such that the corresponding cograph of L has at most one isolated vertex.

Proof: Let all columns of $A(G)$ be distinct and nonzero. Let T be a threshold cotree of G with leaves $B_1, \ldots, B_s$. B_j are set of threshold graphs. We assume that B_j has a threshold graph with $|U_i| \geq 2$. By Proposition 1 all vertices of U_i have the same neighbors. Therefore the columns of vertices of U_i are equal, a contradiction. Let v be an internal node of T and v has at least two subtrees with corresponding cographs with isolated vertices. By Proposition 1, the corresponding vertices of these isolated vertices have the same neighbors outside of these subtrees. Therefore their columns are equal. The sufficiency part is easy to see by induction from leaves to the root. $\square$

Theorem 2 *Let $G = (V, E)$ be a cograph and let $A(G)$ be its adjacency matrix. The rank of $A(G)$ is equal to the number of distinct nonzero columns of $A(G)$.*

Proof: We show by induction on the number of vertices of G. The case $|V(G)| \leq 2$ is trivial. We assume that the assertion holds for $|V(G)| \leq n-1$. We first consider the case that $A(G)$ has at least two equal columns. Without loss of generality we may assume that $A(G) = [a_1, ..., a_{n-2}, a_{n-1}, a_n = a_{n-1}]$, where a_i are the columns of the $A(G)$ and the last two columns are equal. Then $\text{rank}(A(G)) = \text{rank}(A(G - v_n))$. $G - v_n$ is a cograph and by induction hypothesis, $\text{rank}(A(G - v_n))$ is equal to the number of distinct nonzero columns of the $A(G - v_n)$.

It remains to consider the case that all columns of $A(G)$ are different and nonzero. This is the main part of the proof. We show that if all columns of $A(G)$ are different, then all columns of $A(G)$ are linearly independent. Let $\alpha_1, \ldots, \alpha_n$ be the coefficient of columns $a_1, \ldots, a_n$ such that $\sum_{i=1}^{n} \alpha_i a_i = 0$. We have to show $\alpha_1 = \cdots = \alpha_n = 0$. Let T be the threshold cotree of G with leaves $B_1, \ldots, B_s$. The B_j are the set of threshold graphs. By Lemma, 4 each independent set U_i of a threshold graph of B_j has at most one vertex, i.e. the adjacency matrix of such a threshold graph has the form

$$\begin{array}{ccccccccc}
 & K_1 & K_2 & \cdots & K_s & U_1 & U_2 & \cdots & U_s \\
K_1 & A(K_1) & 1 & \cdots & 1 & 1 & 0 & \cdots & 0 \\
K_2 & 1 & A(K_2) & 1 & 1 & 1 & 1 & 0 & 0 \\
\vdots & 1 & 1 & \ddots & 1 & 1 & 1 & \ddots & 0 \\
K_s & 1 & \cdots & 1 & A(K_s) & 1 & \cdots & \cdots & 1 \\
U_1 & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
U_2 & 0 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
\vdots & 0 & 0 & \ddots & 1 & \vdots & \vdots & \ddots & \vdots \\
U_s & 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0
\end{array}$$

Claim: Let v be the node of the threshold cotree T and L be the subtree at v. Let $a_1, \ldots, a_k$ be the columns of vertices of corresponding cograph of the subtree L. It exists a coefficient α_h, where $1 \leq h \leq k$ such that α_j are either $\alpha_j = c_j \alpha_h$, $c_j > 0$ or $\alpha_j = 0$ for $j = 1, \ldots, k$.

Before we prove the claim, let us apply the claim to the root of the threshold cotree T. Then $\alpha_j = c_j \alpha_h$, $c_j > 0$ or $\alpha_j = 0$ for $j = 1, \ldots, n$. From the row of an arbitrary vertex x of the cograph G, we have

$$\sum_{xj \in E(G)} \alpha_j = \alpha_h \sum_{xj \in E(G)} c_j = 0, \text{ then } \alpha_h = 0.$$

Hence all coefficients are equal to zero. Therefore all columns of $A(G)$ are linearly independent.

We prove the claim by induction from leaves to the root of T. Let H be one of the threshold graphs of the leaf B_j (for $A(H)$ see above). By Lemma 3, $H = (K, U)$ and K_i are cliques and U_i are independent sets for $i = 1, \ldots, s$. It is easy to show that the coefficients of the vertices of the clique K_i are equal for $i = 1, \ldots, s$. By Proposition 1 each vertex of threshold graph H has the same neighbors outside of H. Let R_H be the sum of the coefficient of these neighbors. By using the row belonging to U_s the coefficients of $K_1, \ldots, K_{s-1}$ are zero. By the rows belonging to $U_{s-1}, \ldots, U_1$, the coefficients $\alpha_{U_1} = \cdots = \alpha_{U_{s-1}} = 0$. By one of the rows of K_s and U_s we get $\alpha_{U_s} = \alpha_{K_s}$ where α_{K_s} is the coefficient of each vertex of K_s, since

$$\alpha_{K_s}(|K_s| - 1) + \alpha_{U_s} + R_H = 0 = \alpha_{K_s}|K_s| + R_H.$$

Therefore we are finished for each threshold graph of B_j. By Proposition 1 each threshold graph of B_j has the same neighbors outside of B_j. Hence we are also finished for the leaves B_j. Let us now consider an internal node v of T. Let L_1 and L_2 be the subtrees of v and G_1 and G_2 the corresponding cographs (we argue analogously to more subtrees). Let $a_1, \ldots, a_{k-1}$ and $a_k, \ldots, a_r$ be the columns of corresponding vertices of the cographs of L_1 and L_2, respectively. By induction hypothesis, $\alpha_i = b_i \alpha_h$, $b_i > 0$ or $\alpha_i = 0$ for $i = 1, \ldots, k-1$, where $1 \leq h \leq k-1$ and $\alpha_j = c_j \alpha_m$, $c_j > 0$ or $\alpha_j = 0$ for $j = k, \ldots, r$, where $k \leq p \leq r$. Then $\sum_{i=1}^{k-1} \alpha_i = \alpha_1 \sum_{i=1}^{k-1} b_i$ and $\sum_{j=k}^{r} \alpha_j = \alpha_k \sum_{j=k}^{r} c_j$. By Proposition 1 all vertices of the corresponding cographs of L_1 and L_2 have the same neighbors outside of $L_1 \cup L_2$ and let R be the sum of the coefficients of these neighbors. If v has the label $*$, we look at the rows of the vertices h and p and obtain

$$\sum_{hi \in E(G_1)} b_i \alpha_h + \sum_{j=k}^{r} c_j \alpha_p + R = 0 = \sum_{pj \in E(G_2)} c_j \alpha_p + \sum_{i=1}^{k-1} b_i \alpha_h + R.$$

If v has the label $+$, then G_1 and G_2 are connected. We look at the rows y and z such that $yh \in E(G_1)$ and $zp \in E(G_2)$ (otherwise $\alpha_1 = 0$ or $\alpha_k = 0$) and we have

$$\sum_{yi \in E(G_1)} b_i \alpha_h + R = 0 = \sum_{zj \in E(G_2)} c_j \alpha_p + R.$$

For both labels $*$ and $+$ it follows that $\alpha_1 = c\alpha_k$ and $c > 0$, and hence induction is complete. □

During the final stages of the preparation of this article we became aware that G.F. Royle [9] has recently found a quite different proof for Sillke's conjecture. Royle's proof is based on properties of the characteristic polynomials, while we exploit here the structure of cographs with respect to cotree and threshold graphs.

4 Acknowledgment

I thank J. Leydold and P.F. Stadler for helpful discussions. This work was supported by the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung* Proj. 14094-MAT.

References

[1] T. Bıyıkoğlu, A Discrete Nodal Domain Theorem for Trees, Linear Algebra Appl., to appear.

[2] T. Bıyıkoğlu, W. Hordijk, J. Leydold, T. Pisanski, P.F. Stadler, Graph Laplacians, Nodal Domain, and Hyperplane Arrangements, submitted.

[3] I. Chavel, Eigenvalues in Riemannian Geometry, Academic Press, Orlando Fl., 1984.

[4] V. Chvátal, P.L. Hammer, Univ. Waterloo Res. Report, CORR 73-21 (1973).

[5] D.G. Corneil, Y. Perl, L.K. Stewart, A linear recognition algorithm for cographs, SIAM J. Comput. 14 (1985) 926-934.

[6] E.B. Davies, G.M.L. Gladwell, J. Leydold, P.F. Stadler, Discrete nodal domain theorems, Linear Algebra Appl. 336 (2001) 51-60.

[7] P.L. Hammer, A.K. Kelmans, Laplacian spectra and spanning trees of threshold graphs, Discrete Appl. Math. 65 (1996) 255-273.

[8] R. Merris, Laplacian graph eigenvectors, Linear Algebra Appl. 278 (1998) 221-236.

[9] G.F. Royle, The rank of a cograph, submitted.

[10] T. Sillke, http://www.mathematik.uni-bielefeld.de/~sillke/PROBLEMS/cograph.