

## *Combinatorial Perron values of trees and bottleneck matrices*

Enide Andrade<sup>a†</sup> and Geir Dahl<sup>b\*</sup>

<sup>a</sup> *Center for Research and Development in Mathematics and Applications,  
Department of Mathematics, University of Aveiro, Portugal;*

<sup>b</sup> *Department of Mathematics, University of Oslo, Norway*

*(Received 00 Month 20XX; final version received 00 Month 20XX)*

**Published in *Linear and Multilinear Algebra* in 2017, see  
<http://dx.doi.org/10.1080/03081087.2016.1274363>**

The algebraic connectivity  $a(G)$  of a graph  $G$  is an important parameter, defined as the second smallest eigenvalue of the Laplacian matrix of  $G$ . If  $T$  is a tree,  $a(T)$  is closely related to the Perron values (spectral radius) of so-called bottleneck matrices of subtrees of  $T$ . In this setting we introduce a new parameter called the *combinatorial Perron value*  $\rho_c$ . This value is a lower bound on the Perron value of such subtrees; typically  $\rho_c$  is a good approximation to  $\rho$ . We compute exact values of  $\rho_c$  for certain special subtrees. Moreover, some results concerning  $\rho_c$  when the tree is modified are established, and it is shown that, among trees with given distance vector (from the root),  $\rho_c$  is maximized for caterpillars.

**Keywords:** Perron value; bottleneck matrix; tree; Laplacian matrix; majorization

**AMS Subject Classification:** 15A18; 05C05

### 1. Introduction

Let  $G = (V, E)$  be a graph and  $L(G)$  its Laplacian matrix, i.e.,  $L(G) = D - A$  where  $A$  is the adjacency matrix and  $D$  the diagonal matrix of vertex degrees.  $L(G)$  is positive semidefinite and singular. The algebraic connectivity  $a(G)$  of  $G$  is the second smallest eigenvalue of  $L(G)$ , and it is denoted by  $a(G)$ . This is an important parameter, and its properties have been investigated intensively, see a brief discussion of some literature below.

In this paper we focus on trees. Let  $T$  be a given tree. We study the so-called *bottleneck matrices* associated with the Laplacian matrix  $L(T)$ . These arise by choosing a vertex  $v \in T$  and deleting the corresponding row and column of  $L(T)$ . The resulting matrix  $L$  is invertible, and its inverse  $L^{-1}$  is a direct sum of smaller matrices  $M_i$  associated with the subtrees of  $T \setminus \{v\}$  (the components). These matrices

---

<sup>†</sup> Enide Andrade was supported in part by the Portuguese Foundation for Science and Technology (FCT-Fundação para a Ciência e a Tecnologia), through CIDMA - Center for Research and Development in Mathematics and Applications, within project UID/MAT/04106/2013. Email: enide@ua.pt

\*Corresponding author. Email: geird@math.uio.no

$M_i$  are called bottleneck matrices. The spectral radius of such bottleneck matrices determine  $a(G)$ , see [10, 16] and below. The main goal of this paper is to investigate properties of bottleneck matrices, and introduce and study some combinatorial parameters that give estimates for the spectral radius of bottleneck matrices.

For a general graph  $G$  the algebraic connectivity  $a(G)$  reflects the connectivity of  $G$ , and  $a(G) > 0$  if and only if  $G$  is connected [7]. The importance of this invariant parameter stems from inequalities that relates the vertex connectivity (the minimum number of vertices whose removal results in a disconnected graph) and edge connectivity (the minimum number of edges whose removal results in a disconnected graph) [9]. Some graphs for which the algebraic connectivity attains the vertex connectivity are characterized for instance in [12]. Some studies concerning expanders and isoperimetric numbers are related with graphs with large algebraic connectivity [3, 13]. Moreover, some research on how the algebraic connectivity of a weighted tree behaves when the tree is perturbed by removing one of its branches and replacing it by another can be found in [11]. For more references concerning algebraic connectivity, see [1, 4–6, 16].

In [7] Fiedler introduced a remarkable result that describes some of the structure of the eigenvectors associated with  $a(G)$  and allows to classify the trees into two types, denoted simply Type I and II. In fact, if  $z$  is an eigenvector associated with  $a(T)$ , now called a *Fiedler vector*, then exactly one of the following two cases can occur:

- (i) No entry of  $z$  is zero. In this case there is a unique pair of vertices  $v_i$  and  $v_j$  such that they are adjacent and  $z_i > 0$ ,  $z_j < 0$ . Furthermore, the entries of  $z$  are increasing along any path in  $T$  which starts at  $v_i$  and does not contain  $v_j$ , and the entries of  $z$  are decreasing along any path in  $T$  which starts at  $v_j$  and does not contain  $v_i$ . Such a tree is called a *Type II tree*, and the vertices  $v_i$  and  $v_j$  are called *characteristic vertices*.
- (ii) At least one entry of  $z$  is zero. Then the subgraph of  $T$  induced by the set of vertices corresponding to zeros in  $z$  is connected. Moreover, there is a unique vertex  $v_r$  such that  $z_r = 0$  and  $v_r$  is adjacent to at least one vertex  $v_s$  with  $z_s \neq 0$ . The entries are either increasing, decreasing or identically zero along any path  $T$  starting at  $v_r$ . Such a tree is called a *Type I tree* and the vertex  $v_r$  is called *characteristic vertex*.

The *characteristic set* of a tree is the set of its characteristic vertices. One can show that the previous classification is independent of the choice of eigenvector (in case the eigenspace of  $a(G)$  has higher dimension than 1). There is a close connection between bottleneck matrices and the characterization of Type I and II trees and the algebraic connectivity ([10]). In fact,  $T$  is a Type I tree with characteristic vertex  $v_r$  if and only if  $T$  has two or more branches at  $v_r$  for which the spectral radius is maximum (among the branches); such branches are called *Perron branches* at  $v_r$ . In this case, the algebraic connectivity is the reciprocal of the spectral radius of the bottleneck matrix of any of its Perron branches. Moreover,  $T$  is a Type II tree with (adjacent) characteristic vertices  $v_i$  and  $v_j$  if and only if there exists  $\gamma$ ,  $0 < \gamma < 1$ , such that  $\rho(M_1 - \gamma J) = \rho(M_2 - (1 - \gamma)J)$ , where  $M_1$  is the bottleneck matrix for the branch at  $v_j$  containing  $v_i$  and  $M_2$  is the bottleneck matrix for the branch at  $v_i$  containing  $v_j$ . Let  $J$  be the all ones matrix. Then the algebraic connectivity can be

obtained by

$$a(G) = \frac{1}{\rho(M_1 - \gamma J)} = \frac{1}{\rho(M_2 - (1 - \gamma)J)}.$$

These results, and the close connections to bottleneck matrices, motivated the present study of properties of such matrices. For more on algebraic connectivity of trees and related results, see [1, 2, 8, 10–12].

A main goal of this paper is to introduce the notion of combinatorial Perron value, and investigate its properties. The motivation for this new concept lies in the structure of Fiedler vectors. As mentioned above, and loosely speaking, the components of a Fiedler vector are increasing (in absolute value) when one moves away from the characteristic vertex, or vertices. We think of the characteristic vertices as a kind of “center” of the tree. This seems to suggest that the components of a Fiedler vector give some kind of measure of distance from the center, and therefore one may ask:

- *Can the (combinatorial) distances in the tree be used to approximate the Fiedler vector, or to approximate the Perron value?*

This idea and question lead to the main concept of this paper. For a rooted tree  $T$ , with root vertex  $r$ , we define the *combinatorial Perron value* of  $T$  by

$$\rho_c(T) = \frac{\sum_i \sigma_i^2}{\sum_i d_i^2}$$

where  $d_i$  is the number of vertices in the unique path from  $r$  to vertex  $i$  in  $T$ , and  $\sigma_i$  is the sum of those  $d_j$  for which the vertex  $i$  is contained in the path between  $r$  and vertex  $j$ . We show how  $\rho_c(T)$  relates to the bottleneck matrix of the tree and the so-called path matrix, and establish several of its properties. In particular, we show that  $\rho_c(T)$ , and a related value, both provide good approximations of the Perron value of the tree (branch). Moreover, we show how  $\rho_c(T)$  is affected by certain modifications of the tree, and prove that  $\rho_c(T)$  is maximized for caterpillars, when distances are fixed. An interesting (future) application of this new concept is to find approximate characteristic sets of trees (see type (i) and (ii) above) by replacing the Perron values of branches by the corresponding combinatorial Perron values. This, and other possible further work is mentioned as concluding remarks in the final section.

The remaining paper is organized as follows. Section 2 treats bottleneck and path matrices, shows relationships, and prepares for the introduction of  $\rho_c(T)$ . Section 3 defines  $\rho_c(T)$ , gives alternative expressions, and proves the exact value of  $\rho_c(T)$  for certain classes of trees. Section 4 shows that  $\rho_c(T)$  is a lower bound of the corresponding Perron value, and also establishes other interesting bounds, along with numerical examples illustrating the quality of these bounds. A connection to the discrete Poisson problem is also discussed. Section 5 presents a result on majorization and Section 6 contains results on how the combinatorial Perron value is affected by certain modifications of the tree. The extremal property of caterpillars, mentioned above, is also shown in that section. Finally, some suggestions for further work are indicated.

*Notation.* The incidence vector of a set  $S \subseteq U$  is a vector  $v$  whose entries are labeled with the members of  $U$ , that is,  $v_u = 1$  if  $u \in S$ , and  $v_u = 0$  otherwise. For a matrix  $A = [a_{ij}]$  define  $\sigma(A) = \sum_{i,j} a_{ij}$ , and let the  $i$ 'th row sum of  $A$  be denoted

by  $R_i(A)$ .

## 2. Bottleneck and path matrices

Let  $\bar{T}$  be a given tree. Following ([4, 11]) a *branch*  $T$  in  $\bar{T}$  is a component subtree of  $\bar{T} \setminus \{\bar{r}\}$  where  $\bar{r}$  is some vertex in  $\bar{T}$ . Let  $r$  be the (unique) vertex in  $T$  which is adjacent to  $\bar{r}$  in  $\bar{T}$ . We call  $\bar{r}$  and  $r$  the *external root* and the (*internal*) *root* of  $T$ , respectively. Note that  $\bar{r} \notin T$ . Note that in [11] one calls  $\bar{r}$  the root of the branch  $T$ . Let  $n$  be the number of vertices of a branch  $T$  with internal root  $r$ . For each vertex  $j$  in  $T$  let  $P_j$  denote the (unique)  $rj$ -path in  $T$ . The *bottleneck matrix*  $M = [m_{ij}]$  of  $T$  is the  $n \times n$  matrix where  $m_{ij}$  is the number of common vertices in the  $ri$ -path  $P_i$  and the  $rj$ -path  $P_j$ . The *Perron value* of  $T$  is defined as the Perron value of  $M$ , i.e., the spectral radius  $\rho(M)$  of  $M$ .

Another view on the bottleneck matrix  $M$  of  $T$  is presented next. For simplicity we denote its vertices by  $1, 2, \dots, n$ . We define the *path matrix*  $N$ , or  $N_T$ , of  $T$  as the  $n \times n$   $(0, 1)$ -matrix whose rows and columns correspond to the vertices in  $T$ , and where column  $j$  of  $N$  is the incidence vector of the path  $P_j$  for  $j \leq n$ , i.e., it contains ones in rows corresponding to vertices in  $P_j$ , and zeros otherwise. The path matrix  $N$  depends on the ordering of the vertices (but the same ordering is used for rows and columns). We use a breadth-first ordering of the vertices: first the root  $r$ , then its neighbors, then each of their neighbors etc. The number of vertices in  $P_j$  is denoted by  $d_j$  we call it the *distance* of vertex  $j$ ; this is the number of edges in the unique path between the external root  $\bar{r}$  and  $j$  in the tree  $\bar{T}$ . If the underlying tree  $T$  needs to be indicated, we write  $d_j = d(T, j)$ . The vector  $d = (d_1, d_2, \dots, d_n)$  is the distance vector of  $T$ . Then we have the following result.

LEMMA 2.1 *The path matrix  $N$  is upper triangular with ones in the first row and on the diagonal. The column sum vector of  $N$  is the distance vector  $d = (d_1, d_2, \dots, d_n)$  and it satisfies  $1 = d_1 < d_2 \leq d_3 \leq \dots \leq d_n$ . Finally,  $N$  and the bottleneck matrix  $M$  are related by*

$$M = N^T N. \tag{1}$$

*In particular,  $N$  is invertible and  $M$  is positive definite.*

**Proof.** The first statement follows from the discussion above and the fact that we use breadth-first ordering of vertices. Moreover, the  $(i, j)$ 'th entry of  $N^T N$  is the inner product of the incidence vector of  $P_i$  and  $P_j$  and it therefore equals that number of common vertices, i.e.,  $m_{ij}$ . As  $M = N^T N$ , and the columns of  $N$  are linearly independent,  $M$  is positive semidefinite.  $\square$

We see from (1) that the bottleneck matrix  $M$  has the Cholesky factorization  $M = VV^T$  where  $V = N^T$ . We remark (although this will not be used here) that  $M$  is a completely positive matrix (as  $N$  is nonnegative) and, moreover, that the path matrix  $N$  is totally unimodular, i.e., that each square submatrix has determinant 0 or  $\pm 1$ . The distance vector  $d = (d_1, d_2, \dots, d_n)$  of  $T$  will be central in this paper. Vertices in  $T$  are denoted by  $1, 2, \dots, n$  or  $v_1, v_2, \dots, v_n$ .

*Example 1* Consider the tree (branch)  $T$  with  $n = 6$  in Figure 1 where  $r$  is the (internal) root; the external root  $\bar{r}$  is also indicated. The vertices are numbered using

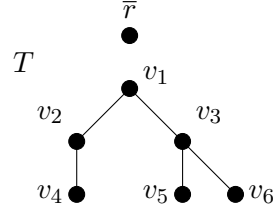


Figure 1. The tree  $T$  in Example 1.

breadth-first-ordering, and  $d = (d_1, d_2, \dots, d_6) = (1, 2, 2, 3, 3, 3)$ . The path matrix is

$$N = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

For instance, column 4 is the incidence vector of the path  $P_4 = v_1, v_2, v_4$ . □

Two special classes of trees are of particular interest: paths and stars. The star  $S_n$ , with root  $r = v_1$ , has edges  $v_1v_i$  ( $2 \leq i \leq n$ ) and its path matrix  $N = N_{S_n}$  is

$$N_{S_n} = \begin{bmatrix} 1 & e^T \\ 0 & I_{n-1} \end{bmatrix},$$

where  $e$  is the all ones vector and  $I_{n-1}$  the identity matrix of order  $n - 1$ . The distance vector is  $d = (1, 2, 2, \dots, 2)$ . Similarly, when  $T$  is a path  $P_n = v_1, v_2, \dots, v_n$  with root  $r = v_1$ , the path matrix  $N = N_{P_n}$  is

$$N_{P_n} = U_n,$$

where  $U_n$  is the  $n \times n$   $(0, 1)$ -matrix with ones in its upper triangular part (including the diagonal). The distance vector is  $d = (1, 2, 3, \dots, n)$ . These two special trees are extreme in the sense that the path matrix  $N$  has the minimum number of ones in the case of stars, and the maximum number of ones in the case of paths.

### 3. Combinatorial Perron values

Let again  $T$  be a rooted tree. We use the notation introduced in Section 2. Moreover, let  $\rho(M)$  or  $\rho(T)$  denote the Perron value (maximum eigenvalue) of the bottleneck matrix  $M$  of  $T$ .

We now introduce a new concept which is central in this work. Define

$$\rho_c(N) = \frac{\|Nd\|^2}{\|d\|^2} = \frac{d^T N^T N d}{d^T d} \tag{2}$$

which we call the *combinatorial Perron value* of  $T$ . Our discussion below will show that the combinatorial Perron value is a good approximation to the Perron value  $\rho(M)$  of  $M$ . This also means that the distance vector  $d$  is a good starting point for finding the Perron vector (the eigenvector corresponding to the eigenvalue  $\rho(M)$ ). Thus the distance vector captures much of the information carried by the Perron vector.

The tree  $T$  and its path matrix  $N_T$  are identified, so we sometimes write  $\rho_c(T)$  for  $\rho_c(N)$ , and if different trees are considered we write  $N_T$  for the path matrix of  $T$ . This is a graph invariant, i.e., reordering of vertices does not affect the value of  $\rho_c(N)$ , and, moreover,  $\rho_c(N)$  *only* depends on  $N$  (or  $T$  and  $r$ ), because the distance vector  $d$  is the column sum vector of  $N$ , i.e.,  $d = N^T e$  where  $e$  is the all ones vector. Thus,  $\rho_c(N)$  is the Rayleigh quotient of the bottleneck matrix  $M$  evaluated in the distance vector  $d$ . This immediately implies that  $\rho_c(T)$  is a lower bound on the Perron value  $\rho(M)$ ; more on this below.

For vertices  $i$  and  $j$ , if  $P_j$  contains  $P_i$ , then we write  $j \preceq i$ ; this means that  $j$  is “below” or after  $i$ , seen from the root. This defines a partial order on the vertex set  $V$ . As  $M = N^T N = [m_{ij}]$ , we obtain some alternative expressions for  $\rho_c(T)$  (where the summation is from 1 to  $n$ ):

$$\rho_c(N) = \frac{\sum_i d_i^3 + 2 \sum_{i < j} m_{ij} d_i d_j}{\sum_i d_i^2} = \frac{\sum_i \sigma_i^2}{\sum_i d_i^2} \tag{3}$$

where  $\sigma_i = \sum_{j \preceq i} d_j$ . In fact,  $d^T N^T N d = d^T M d = \sum_i d_i^3 + 2 \sum_{i < j} m_{ij} d_i d_j$  as  $M$  is symmetric and  $m_{ii} = d_i$  ( $i \leq n$ ). Moreover, the  $i$ th row of the path matrix  $N$  is the incidence vector of the set  $\{j : j \preceq i\}$ , so  $\|Nd\|^2 = \sum_i (\sum_{j \preceq i} d_j)^2 = \sum_i \sigma_i^2$ .

For the extreme cases when the tree  $T$  is a star or a path it is possible to obtain an explicit expression for the combinatorial Perron value. For a path this expression is a quadratic polynomial in  $n$ .

**PROPOSITION 3.1** *Let  $P_n$  and  $S_n$  denote the path and the star with  $n$  vertices, respectively. Then*

$$\begin{aligned} \rho_c(P_n) &= (2n^2 + 2n + 1)/5, \\ \rho_c(S_n) &= n + \frac{3n-3}{4n-3}. \end{aligned} \tag{4}$$

**Proof.** We obtain for the path  $P_n$

$$\begin{aligned} \|Nd\|^2 &= \sum_{i=1}^n \sigma_{v_i}^2 \\ &= \sum_{i=1}^n (\sum_{k=i}^n k)^2 \\ &= (4n^5 + 10n^4 + 10n^3 + 5n^2 + n)/30. \end{aligned}$$

Moreover,  $d = (1, 2, \dots, n)$ ,  $\|d\|^2 = n(2n + 1)(n + 1)/6$  and by dividing these expressions we obtain

$$\rho_c(P_n) = \frac{\|Nd\|^2}{\|d\|^2} = (2n^2 + 2n + 1)/5.$$

For the star  $S_n$ ,  $d = (1, 2, 2, \dots, 2)$ , so  $\|d\|^2 = 4n - 3$  and  $\|Nd\|^2 = 4n^2 - 3$  and

therefore  $\rho_c(S_n) = n + (3n - 3)/(4n - 3)$ . □

The Perron value of a path  $P_n$  (i.e., the Perron value of its bottleneck matrix) is known, see [2], and it is given by

$$\rho(P_n) = \frac{1}{2} \left( 1 - \cos\left(\frac{\pi}{2n+1}\right) \right)^{-1}.$$

This is derived from the general connection between the Perron values of branches of a tree and the algebraic connectivity of the tree, see [10]. In fact, an odd path is a tree of Type I, where the middle vertex has two Perron branches, and their bottleneck matrix has spectral radius equal to the inverse of the algebraic connectivity. Using the same approach one may compute the Perron value of the bottleneck matrix of a star  $S_n$ , see [2], and it is

$$\rho(S_n) = (1/2)(n + 1 + \sqrt{n^2 + 2n - 3}).$$

*Example 2* In Figure 2 we compare the Perron value  $\rho(T)$  and the combinatorial Perron value  $\rho_c(T)$ , when  $T$  is a star or a path for some values of  $n = |V|$ . Note that  $\rho_c$  is very close to the Perron value. This is also our general experience from extensive computational experiments, we return to this later. □

|               | $n = 4$ | $n = 5$ | $n = 6$ | $n = 7$ | $n = 9$ |
|---------------|---------|---------|---------|---------|---------|
| $\rho(S_n)$   | 4.7913  | 5.8284  | 6.8541  | 7.873   | 9.8990  |
| $\rho_c(S_n)$ | 4.6923  | 5.7059  | 6.7143  | 7.72    | 9.7272  |
| $\rho(P_n)$   | 8.2909  | 12.3435 | 17.2069 | 22.8808 | 36.6604 |
| $\rho_c(P_n)$ | 8.2     | 12.2    | 17      | 22.6    | 36.2    |

Figure 2. Perron value and  $\rho_c$  for stars and paths

Figure 3 shows the relative difference of Perron and combinatorial Perron values for paths, i.e.,  $100(\rho(P_n) - \rho_c(P_n))/\rho(P_n)\%$ , for different values of  $n$ . So, we see that this relative difference is about 1.3% for  $n$  at least 10.

A tree  $T$  having at most one vertex of degree greater than two is called a *generalized star*, or a *star-like tree*. We call this vertex the central vertex. Both stars and paths are special cases of a generalized star. We restrict the attention to the generalized star with  $p$  paths, each of length  $k$ , and where the internal root is the central vertex. This tree is denoted by  $S_{p,k}$ . So,  $k = 1$  corresponds to a star, and  $p = 1$  corresponds to a single path. The next result computes  $\rho_c(S_{p,k})$ . Similarly, one may compute the combinatorial Perron value of all generalized stars, but the expressions become more complicated, so we omit this here.

**PROPOSITION 3.2** *The combinatorial Perron value of the generalized star  $S_{p,k}$  is given by*

$$\rho_c(S_{p,k}) = \frac{p(\eta_k - \xi_k^2) + (p\xi_k - p + 1)^2}{pk(2k+1)(k+1)/6 - p + 1} \tag{5}$$

where  $\eta_k = (4k^5 + 10k^4 + 10k^3 + 5k^2 + k)/30$  and  $\xi_k = k(k + 1)/2$ .

**Proof.** The distance vector of  $S_{p,k}$  is  $d = (1, 2, \dots, 2, 3, \dots, 3, \dots, k, \dots, k)$  where

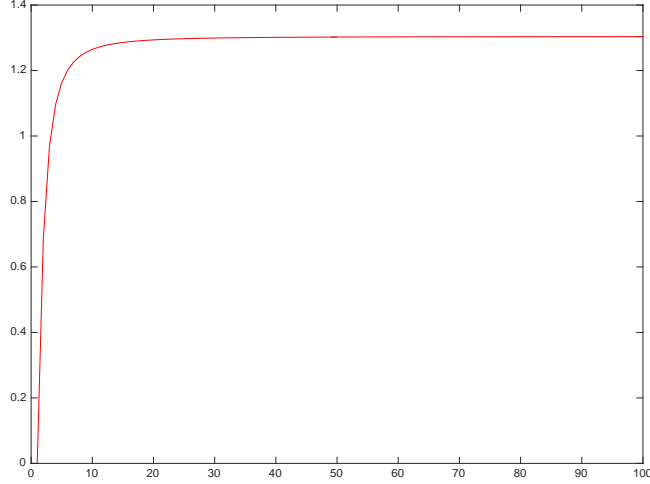


Figure 3. Relative difference,  $\rho$  and  $\rho_c$  for paths.

the number  $i$  appears  $p$  times ( $2 \leq i \leq k$ ). So

$$\begin{aligned} \|d\|^2 &= p \sum_{j=1}^k j^2 - (p-1) \\ &= pk(2k+1)(k+1)/6 - p + 1. \end{aligned}$$

Next, let  $N_{P_k}$  and  $d$  be the path matrix and the distance vector in  $P_k$ , so by the calculations in Proposition 3.1,  $\|N_{P_k}d(P_k, \cdot)\|^2 = (4k^5 + 10k^4 + 10k^3 + 5k^2 + k)/30$ . Then, by comparing  $S_{p,k}$  to  $p$  paths  $P_k$  we get

$$\begin{aligned} \|N_{S_{p,k}}d\|^2 &= p\eta_k - p\xi_k^2 + \sigma_1^2 \\ &= p(\eta_k - \xi_k^2) + (p\xi_k - p + 1)^2. \end{aligned}$$

The expression for  $\rho_c(S_{p,k})$  follows from this. □

#### 4. Perron value bounds

Let  $S_i = \{j : j \preceq i\}$  and  $R_i(N) = |S_i|$  ( $i \leq n$ ). So,  $R_i(N)$  is the number of vertices below  $v_i$ , and also the  $i$ th row sum in the path matrix  $N$ . Define  $\beta_i = \sum_{j:v_i \preceq v_j} r_j$ .



Let  $K = NN^T$ . Then the  $(i, j)$ th entry of  $K$  is  $|S_i \cap S_j|$ . Recall that  $\sigma(A)$  denotes the sum of all entries in a matrix  $A$ . Then

$$\rho_c(N) = \frac{\sigma(K^2)}{\sigma(K)}. \tag{6}$$

In fact, as  $d = N^T e$ , one gets  $\|Nd\|^2 = d^T N^T N d = e^T N N^T N N^T e = e^T K^2 e = \sigma(K^2)$ . Similarly,  $d^T d = e^T K e = \sigma(K)$ , and (6) follows.

A very simple and rough upper bound on  $\rho_c(N)$  is  $\sigma(N)$ , the number of ones in  $N$ , as shown next.

**PROPOSITION 4.1**

$$\rho_c(N) \leq \sum_{j=1}^n d_j = \sigma(N). \tag{7}$$

**Proof.** From the Cauchy-Schwarz inequality we obtain

$$\|Nd\|^2 = \sum_{i=1}^n (\text{row}_i(N) \cdot d)^2 \leq \sum_{i=1}^n \|\text{row}_i(N)\|^2 \|d\|^2$$

where  $\text{row}_i(N)$  denotes the  $i$ th row of  $N$ . But as  $N = [a_{ij}]$  is a  $(0, 1)$ -matrix

$$\sum_{i=1}^n \|\text{row}_i(N)\|^2 = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 = \sum_{i=1}^n \sum_{j=1}^n a_{ij} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} = \sigma(N) = \sum_{j=1}^n d_j.$$

Therefore

$$\rho_c(N) = \frac{\|Nd\|^2}{\|d\|^2} \leq \frac{\sigma(N)\|d\|^2}{\|d\|^2} = \sigma(N).$$

□

Using similar techniques one may find a better upper bound on  $\rho_c(N)$ , as a convex combination of the numbers  $\beta_j$  ( $j \leq n$ ):  $\rho_c(N) \leq (\sum_j \beta_j d_j^2) / (\sum_j d_j^2)$ . This is done using the Cauchy-Schwarz inequality for the subvector  $d_{S_i}$  of  $d$  corresponding to the support  $S_i$  in each row of  $N$ . The simple bound  $\sigma(N) = \sum_i d_i$  in (7) contains combinatorial information, the sum of the distances. We should say that bounds on  $\rho_c(N)$  are not of computational interest, since it is easy to compute  $\rho_c(N)$  anyway.

We now introduce an interesting bound for the Perron value  $\rho(T)$ . This bound is easily computed from the parameters, and has an interpretation in terms of the discrete Poisson problem, as briefly discussed next. The bottleneck matrix  $M$  satisfies (see [10, 16])  $M = L^{-1}$  where  $L$  is obtained from the Laplacian matrix of  $T$  by adding 1 in position  $(1, 1)$ ; this corresponds to the root  $\bar{r}$  (which is not a part of  $T$ ). Associated with  $L$  we have the discrete Poisson problem  $Lu = f$  where  $f$  is the given data (parameter),  $f \in \mathbb{R}^n$ . The Poisson problem is one of basic equations in partial differential equations ([17]). By using the standard approach of discretizing the domain of the (unknown) function  $u$ , the continuous Laplace operator is replaced by the discrete Laplace operator, and this corresponds to the Laplacian matrix, with modifications based on the boundary of the region. In our setting the

“boundary” of the tree  $T$  consists only of the external root  $\bar{r}$ . Thus, the matrix  $L$  also equals the Laplacian matrix of the larger tree  $\bar{T}$  after deleting the row and column corresponding to  $\bar{r}$ . This modification makes the matrix  $L$  invertible, so the solution  $u$  of the discrete Poisson problem  $Lu = f$  is therefore unique. Now, let  $d$  as usual be the distance vector of  $T$ . Consider the discrete Poisson problem with data  $d$ , i.e.,

$$Lu = d. \quad (8)$$

The unique solution will be denoted by  $u^*(d)$ , and may be computed from the bottleneck matrix as  $u^*(d) = L^{-1}d = Md = N^tNd$ . Define

$$\pi(N) = \frac{\|u^*(d)\|}{\|d\|} = \frac{\|Md\|}{\|d\|}. \quad (9)$$

Thus,  $\pi(N)$  measures the size of the *solution* of the Poisson problem relative to the data, where the data is the distance vector  $d$ . The component  $u^*(d)_v$  of  $u^*(d)$  corresponding to vertex  $v$  is  $\sum_{u \in P_v} \sigma_u$  (where  $\sigma_u = \sum_{w \prec u} d_w$ ); this may be regarded as the length/weight of the path  $P_v$ , using vertex weights  $\sigma_u$ . In fact, the measure  $\pi(N)$  has a close connection to the combinatorial Perron value and also to the Perron value.

Define, for the bottleneck matrix  $M = [m_{ij}]$ , its  $\ell_1$  matrix norm  $\|M\|_1 = \max_i \sum_j m_{ij}$ . The following result relates several of the parameters discussed, and provide lower and upper bounds on the Perron value  $\rho(M)$  of the bottleneck matrix  $M$ .

**THEOREM 4.2** *Let  $T$  be a tree and  $M = [m_{ij}]$  its bottleneck matrix. Then*

$$\rho_c(N) \leq \pi(N) \leq \rho(M) \leq \|M\|_1. \quad (10)$$

**Proof.** Since  $u^*(d) = N^tNd$  we obtain from the Cauchy-Schwarz inequality

$$\begin{aligned} \rho_c(N) &= \|Nd\|^2 / \|d\|^2 \\ &= d^t N^t Nd / \|d\|^2 \\ &= d^t u^*(d) / \|d\|^2 \\ &\leq \|d\| \|u^*(d)\| / \|d\|^2 \\ &= \|u^*(d)\| / \|d\| \\ &= \pi(N). \end{aligned}$$

Next

$$\pi(N) = \|u^*(d)\| / \|d\| = \|Md\| / \|d\| \leq \sup_{x \neq 0} \|Mx\| / \|x\| = \rho(M).$$

Finally, any matrix norm gives an upper bound for the spectral radius, and therefore

$$\rho(M) \leq \|M\|_1 = \max_i \sum_j m_{ij}.$$

□

| $n$ | $\rho_c(N)$ | $\pi(N)$ | $\rho(M)$ | Err. | $\ M\ _1$ | $\sum_v d_v$ |
|-----|-------------|----------|-----------|------|-----------|--------------|
| 10  | 21.4        | 21.6     | 21.9      | 1.4  | 26        | 32           |
| 20  | 36.4        | 37.0     | 37.8      | 2.1  | 47        | 73           |
| 30  | 54.0        | 55.0     | 56.2      | 2.1  | 69        | 110          |
| 50  | 88.6        | 90.5     | 93.4      | 3.1  | 118       | 234          |
| 100 | 125.6       | 129.2    | 133.2     | 3.0  | 174       | 410          |
| 150 | 287.7       | 294.6    | 304.0     | 3.1  | 397       | 898          |

Figure 4. Random trees, comparison of bounds in Theorem 4.2.

The importance of Theorem 4.2 is that one obtains easily computable bounds on the Perron value, and, moreover, that these bounds are obtained from the distances in the tree  $T$ . Moreover, one may consider the distance vector  $d$  as an approximation to the Perron vector. The quality of the bounds of Theorem 4.2 are suggested by the result of generating some random trees, see the table in Figure 4. Each row corresponds to a random tree, and  $n$  is the number of vertices. The bounds  $\rho_c(N)$  and  $\pi(N)$  and the Perron value  $\rho(M)$  are given (with one decimal). The relative error,  $\text{Err} = 100 \cdot (\rho(M) - \pi(N))/\rho(M)\%$ , is also indicated, as well as the rough upper bound  $\sum_v d_v$  on  $\rho_c(N)$ . Extensive testing confirms that these lower bounds are very good.

**COROLLARY 4.3** *For any tree  $T$  the following inequalities hold*

$$\rho_c(N) \leq \pi(N) \leq \rho(M) \leq n(n+1)/2. \tag{11}$$

**Proof.** Theorem 4.2 gives  $\rho_c(N) \leq \pi(N) \leq \rho(M) \leq \|M\|_1$  so we only need to show

$$\|M\|_1 \leq n(n+1)/2.$$

Let  $i$  be such that  $\sum_j m_{ij} = \|M\|_1$ , where  $m_{ij} = |P_i \cap P_j|$ . Then vertex  $i$  must be a pendant vertex, otherwise there is an adjacent vertex  $l$  with  $d_l = d_i + 1$ , and then  $|P_l \cap P_j| \geq |P_i \cap P_j| = m_{ij}$  for each  $j$ , and with strict inequality holds for  $l$  (and other vertices below  $l$ , not that when  $i = j = l$ ,  $m_{ll} = m_{ii} + 1$ ). This contradicts the choice of  $i$ , so we conclude that  $i$  is a pendant vertex. Let  $k = d_i$ . Therefore

$$\begin{aligned} \|M\|_1 &= \sum_j m_{ij} \\ &= \sum_{s \in P_i} m_{is} + \sum_{s \notin P_i} m_{is} \\ &\leq (1 + 2 + \dots + k) + (n - k)(k - 1) \\ &= (1/2)k(k + 1) + (n - k)(k - 1) \\ &= nk - n - (1/2)k^2 + (3/2)k. \end{aligned}$$

Simple analysis shows that final expression, viewed as a quadratic polynomial in  $k$  is increasing in  $k$ , and it is maximized for  $k = n$  (clearly,  $k \leq n$  must hold). We conclude that  $\|M\|_1 \leq (1/2)n(n+1)$  as desired. □

## 5. Caterpillars and a majorization result

A *caterpillar*  $C$  is a tree  $C = C(n_1, n_2, \dots, n_k)$  consisting of a path  $P = v_1, v_2, \dots, v_k$ , for some  $k$ , and, for each  $i \leq k$ ,  $n_i \geq 0$  additional vertices attached to  $v_i$ . We require  $n_k = 0$  (otherwise we could redefine). The path  $P_k$  is called the *central path*, and  $v_1$  is the internal root of this tree. So  $C$  has  $n = k + \sum_{i=1}^k n_i$  vertices. Stars and paths are special cases of a caterpillar. Let  $N_C$  be the  $n \times n$  path matrix of a caterpillar  $C$ , and let  $d$  be its distance vector, so  $1 = d_1 < d_2 \leq \dots \leq d_n$ . Then the  $j$ th column of  $N_C$  consists of  $d_j - 1$  leading ones in addition to a 1 in row  $j$  (on the diagonal of  $N_C$ ).

For any tree  $T$  with distance vector  $d$ , it is easy to see that there is a unique (up to relabeling) caterpillar  $C$  with the same distance vector  $d$ , and we let  $C_d$  denote this caterpillar.

The following theorem gives a connection to majorization and shows that the path matrices of caterpillars are extreme in a certain sense. We use  $\preceq^{maj}$  to denote the majorization order ([14]), so we write  $x \preceq^{maj} y$ , for vectors  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_n)$  whenever  $\sum_{j=1}^k x_{[j]} \leq \sum_{j=1}^k y_{[j]}$  ( $k < n$ ) and  $\sum_{j=1}^n x_j = \sum_{j=1}^n y_j$ . Here  $x_{[j]}$  denotes the  $j$ th largest component in  $x$ . Recall that the row sum corresponding to vertex  $i$  in the path matrix  $N$  is  $|\{j : j \preceq i\}|$ . Recall that the  $i$ 'th row sum of a matrix  $A$  is denoted by  $R_i(A)$ , and we let  $R(A)$  denote the vector with these numbers as components.

**THEOREM 5.1** *Let  $T$  be a tree with  $n$  vertices. Then the following majorization holds*

$$R(N_T) \preceq^{maj} R(N_{C_d}).$$

*In particular,*

$$\sum_{i=1}^n R_i(N_T)^2 \leq \sum_{i=1}^n R_i(N_{C_d})^2.$$

**Proof.** Let the distance vector of  $T$  be  $d = (d_1, d_2, \dots, d_n)$ , and let  $R(N_T) = (r_1^T, r_2^T, \dots, r_n^T)$ ,  $R(N_{C_d}) = (r_1^C, r_2^C, \dots, r_n^C)$ . Clearly

$$\sum_{i=1}^n r_i^C = \sum_{i=1}^n d_i = \sum_{i=1}^n r_i^T.$$

Let  $k < n$  and choose  $i_1, \dots, i_k$  such that  $r_{i_t}^T = r_{[t]}^T$ ; the  $t$ 'th largest component in

$R(N_T)$ , for  $t \leq k$ . Let  $N = [a_{ij}]$  be the path matrix of  $T$ . Then

$$\begin{aligned} \sum_{t=1}^k r_{[t]}^T &= \sum_{t=1}^k r_{i_t}^T \\ &= \sum_{t=1}^k \sum_{j=1}^n a_{i_t j} \\ &= \sum_{t=1}^k (1 + \sum_{j=i_t+1}^n a_{i_t j}) \\ &= k + \sum_{t=1}^k \sum_{j=i_t+1}^n a_{i_t j} \\ &= k + \sum_{j=1}^n \sum_{t:t \leq k, i_t < j} a_{i_t j} \\ &\leq k + \sum_{j=1}^n \min\{k, d_j - 1\} \\ &= \sum_{i=1}^k r_i^C. \end{aligned}$$

This shows the desired majorization. Moreover, this implies that  $\sum_{i=1}^n g(r_i^T) \leq \sum_{i=1}^n g(r_i^C)$  for each convex function  $g : \mathbb{R} \rightarrow \mathbb{R}$ . In fact, by a well-known theorem due to Hardy, Littlewood and Pólya (see Theorem B.2 in Chapter 2 of [14]) the majorization implies that there exists a doubly stochastic matrix  $A = [a_{ij}]$  (that is, a square matrix with nonnegative entries and all rows and column sums equal to one) such that  $R(N_T) = A R(N_{C_d})$ . Combining this with the convexity of  $g$  (Jensen's inequality) gives

$$\sum_i g(r_i^T) = \sum_i g\left(\sum_j a_{ij} r_j^C\right) \leq \sum_i \sum_j a_{ij} g(r_j^C) = \sum_j g(r_j^C) \sum_i a_{ij} = \sum_j g(r_j^C).$$

as  $\sum_i a_{ij} = 1$  for each  $j$ . Applying this for  $g(x) = x^2$  we get the final inequality.  $\square$

## 6. Tree modifications and the combinatorial Perron value

We study how certain different operations on trees influence the combinatorial Perron value.

A natural operation is to extend the tree by adding a vertex and connecting to some existing vertex. In general, this operation does not decrease the Perron value  $\rho(T)$ , as the maximal eigenvalue of a principal submatrix of a nonnegative matrix  $A$  cannot exceed the maximal eigenvalue of  $A$  ([15]). Surprisingly, for the combinatorial Perron value  $\rho_c(T)$ , this is not true: there are examples where adding an edge leads to smaller value on  $\rho_c(T)$ . This has been observed in our computational experiments for randomly generated trees, although this seems to occur very seldom, and indicate a kind of “irregularity” of the combinatorial Perron value.

However, under a certain, rather weak, condition we can show that  $\rho_c$  increases when an edge is added.

**THEOREM 6.1** *Let  $T$  be a tree and  $w \in V(T)$ , and let  $T'$  be the tree obtained from  $T$  by adding a vertex  $w'$  and the edge  $ww'$ . Let  $k = d(T, w) + 1$  and assume that*

$$(2/k) \sum_{v \in P_w} \sigma_v + k > \|M_T\|_1 \tag{12}$$

where  $P_w$  is the path between the root and  $w$  in  $T$ . Then  $\rho_c(T') > \rho_c(T)$ .

**Proof.** Let  $\sigma_v = \sigma(T, v)$  for each vertex  $v \in V(T)$ , and let  $n = |V|$ . The new

tree  $T'$  has  $n + 1$  vertices,  $V(T') = V(T) \cup \{w'\}$  and  $d(T', w') = k$ . Note that  $\sigma(T', v) = \sigma(T, v) + k$  for  $v \in P_w$  and  $\sigma(T', v) = \sigma(T, v)$  for  $v \in V(T) \setminus P_w$ . So

$$\rho_c(T') = \frac{\sum_{v \in V} \sigma_v^2 + \Delta}{\sum_{v \in V} d_v^2 + k^2} \tag{13}$$

where

$$\begin{aligned} \Delta &= \sum_{v \in V(T')} (\sigma(T', v))^2 - \sum_{v \in V} \sigma_v^2 \\ &= k^2 + \sum_{v \in P_w} ((\sigma(T', v))^2 - \sigma_v^2) \\ &= k^2 + \sum_{v \in P_w} ((\sigma_v + k)^2 - \sigma_v^2) \\ &= k^2 + \sum_{v \in P_w} k(2\sigma_v + k) \\ &= k^2 + 2k \sum_{v \in P_w} \sigma_v + \sum_{v \in P_w} k^2 \\ &= 2k \sum_{v \in P_w} \sigma_v + k^3 \end{aligned}$$

as  $P_w$  contains  $k - 1$  vertices. Moreover

$$\frac{\sum_{v \in V} \sigma_v^2}{\sum_{v \in V} d_v^2} = \rho_c(T) \leq \|M_T\|_1 < (2/k) \sum_{v \in P_w} \sigma_v + k$$

where the first inequality is due to Theorem 4.2 and the second is by assumption. Observe the following simple fact, for positive real numbers  $a_1, a_2, b_1, b_2$ , that if  $a_1/b_1 < a_2/b_2$ , then  $a_1/b_1 < (a_1 + a_2)/(b_1 + b_2)$ . We apply this observation with  $a_1 = \sum_{v \in V} \sigma_v^2$ ,  $b_1 = \sum_{v \in V} d_v^2$ ,  $a_2 = \Delta$  and  $b_2 = k^2$ . Then  $\rho_c(T) = a_1/b_1 < a_2/b_2 = (2/k) \sum_{v \in P_w} \sigma_v + k$  as just shown. Therefore

$$\rho_c(T) < (a_1 + a_2)/(b_1 + b_2) = \rho_c(T')$$

as desired. □

It seems complicated to obtain closed expressions for  $\rho_c(T)$  for trees other than generalized stars. However, the rational function of  $d$  in the definition of  $\rho_c(T)$  opens up for another kind of analysis, which may be called a ‘‘marginal asymptotic analysis’’. We give one such result, and it should be clear how to produce similar results.

Let  $T$  be an arbitrary tree with internal root  $r$  and  $n$  vertices. Let  $v \in V(T)$  and define  $T(v; m)$  as the tree obtained from  $T$  by adding  $m$  new vertices and attaching each of these to  $v$ . Define

$$\rho_c^\infty(T, v) = \lim_{m \rightarrow \infty} \frac{\rho_c(T(v; m))}{n + m}.$$

Since  $n + m$  is the number of vertices of  $T(v; m)$ ,  $\rho_c^\infty(T, v)$  gives the asymptotic value of  $\rho_c(T(v; m))$ , scaled by the number of vertices, when  $m$  tends to infinity. This number gives information on the effect of appending several vertices at vertex  $v$ .

THEOREM 6.2 *Let  $T(v; m)$  be as above. Then*

$$\rho_c^\infty(T, v) = d_v. \quad (14)$$

**Proof.** Let  $k = d_v + 1$ . Also, let  $V$  be the vertex set of  $T$ , and  $d = (d_1, d_2, \dots, d_n)$  its distance vector. Then  $\rho_c(T(v; m))$  may be written as follows

$$\frac{mk^2 + \sum_{w \in P_v} (\sigma_w + mk)^2 + \sum_{w \notin P_v} \sigma_w^2}{mk^2 + \sum_v d_v^2} = \frac{k^2 + \sum_{w \in P_v} \frac{(\sigma_w + mk)^2}{m} + \sum_{w \notin P_v} \frac{\sigma_w^2}{m}}{k^2 + \sum_v \frac{d_v^2}{m}}$$

Note that the denominator here tends to  $k^2$  as  $m \rightarrow \infty$ . Moreover,

$$\begin{aligned} & \lim_{m \rightarrow \infty} (1/(m+n)) \left( k^2 + \sum_{w \in P_v} \frac{(\sigma_w + mk)^2}{m} + \sum_{w \notin P_v} \frac{\sigma_w^2}{m} \right) \\ &= \sum_{w \in P_v} \lim_{m \rightarrow \infty} \frac{(\sigma_w + mk)^2}{m(m+n)} \\ &= \sum_{w \in P_v} \lim_{m \rightarrow \infty} \frac{(\sigma_w/m + k)^2}{1+n/m} \\ &= \sum_{w \in P_v} k^2 \\ &= |P_v| \cdot k^2. \end{aligned}$$

Therefore we get the limit

$$\rho_c^\infty(T, v) = \lim_{m \rightarrow \infty} \frac{\rho_c(T(v; m))}{n+m} = \frac{|P_v| \cdot k^2}{k^2} = |P_v| = d_v.$$

□

From this theorem, for  $m$  large, we obtain the approximation

$$\rho_c(T(v; m)) \approx (n+m)d_v$$

where  $n+m = |V(T(v; m))|$ .

We now turn to the coalescence of two trees. Let  $T_0$  and  $T_1$  be rooted trees with disjoint vertex sets. Let  $v$  be a vertex in  $T_0$ , and let  $v_1$  denote the internal root in  $T_1$ . Let  $T_0 \oplus_v T_1$  be the *coalescence* of  $T_0$  and  $T_1$  with respect to  $v$ , i.e.,  $T_0 \oplus_v T_1$  is the tree obtained by the union of  $T_0$  and  $T_1$  where the internal root  $v_1$  of  $T_1$  is identified with  $v$ . Note that  $|V(T_0 \oplus_v T_1)| = |V(T_0)| + |V(T_1)| - 1$ . An example of this construction is shown in Figure 5. We obtain different such trees depending on the choice of the connecting vertex  $v$ . Recall that the distance in a tree  $T$  from its root to a vertex  $v$  is sometimes indicated by  $d(T, v)$ . Similarly, we write  $\sigma(T, v) = \sum_{w \preceq v} d(T, w)$  where the partial order refers to  $T$ . Let  $k = d(T_0, v)$  and define

$$\delta_{T_1} = \sum_{w \in V(T_1) \setminus \{v_1\}} d(T_1, w) + (k-1)(|V(T_1)| - 1)$$

which only depends on  $T_1$  and  $k$ . We let  $V = V(T_0 \oplus_v T_1)$ .

THEOREM 6.3 *Let  $T_0$  and  $T_1$  be as above and let  $p$  and  $q$  be distinct vertices in  $T_0$  with  $d(T_0, p) = d(T_0, q) = k$  for some  $k$ , i.e.,  $p$  and  $q$  have the same distance from*

the root. Then

$$\rho_c(T_0 \oplus_p T_1) = \rho_c(T_0 \oplus_q T_1) + \frac{\Delta}{\sum_{v \in V} d_v^2} \quad (15)$$

where  $\Delta = 2\delta_{T_1} \left( \sum_{v \in P_p \setminus P_q} \sigma(T_0, v) - \sum_{v \in P_q \setminus P_p} \sigma(T_0, v) \right)$ . In particular,

$$\rho_c(T_0 \oplus_p T_1) > \rho_c(T_0 \oplus_q T_1)$$

if and only if  $\sum_{v \in P_p \setminus P_q} \sigma(T_0, v) > \sum_{v \in P_q \setminus P_p} \sigma(T_0, v)$ .

**Proof.** To simplify notation, define  $T_p = T_0 \oplus_p T_1$  and  $T_q = T_0 \oplus_q T_1$ , and let  $V$  be the vertex set of  $T_p$  and  $T_q$  with the natural identification of vertices. Since  $d(T_0, p) = d(T_0, q)$ , it follows that the distances (from the internal root) are the same in  $T_p$  and  $T_q$ , i.e., the distance functions  $d(T_p, \cdot)$  and  $d(T_q, \cdot)$  coincide. In particular, for each  $v \in V(T_1)$ ,

$$d(T_p, v) = d(T_q, v) = k + d(T_1, v) - 1$$

as  $d(T_1, v_1) = 1$  where  $v_1$  is the internal root in  $T_1$ . Therefore

$$\sum_{v \in V(T_1) \setminus \{v_1\}} d(T_p, v) = \sum_{v \in V(T_1) \setminus \{v_1\}} (k + d(T_1, v) - 1) = \delta_{T_1}.$$

Now,  $\sigma(T_p, v) = \sigma(T_q, v)$  for all  $v \notin P_p \Delta P_q := (P_p \setminus P_q) \cup (P_q \setminus P_p)$ , as for such  $v$ , the subtrees  $\{w : w \preceq v\}$  are the same in  $T_p$  and  $T_q$ . Moreover,  $\sigma(T_p, v) = \sigma(T_0, v) + \delta_{T_1} = \sigma(T_q, v) + \delta_{T_1}$  for all  $v \in P_p \setminus P_q$ , and  $\sigma(T_p, v) = \sigma(T_0, v) = \sigma(T_q, v) - \delta_{T_1}$  for all  $v \in P_q \setminus P_p$ .

By summing over all  $v \in V$  we get

$$\begin{aligned} & \sum_{v \in V} \sigma(T_p, v)^2 - \sum_{v \in V} \sigma(T_q, v)^2 \\ &= \sum_{v \in P_p \Delta P_q} (\sigma(T_p, v)^2 - \sigma(T_q, v)^2) \\ &= \sum_{v \in P_p \Delta P_q} (\sigma(T_p, v) - \sigma(T_q, v))(\sigma(T_p, v) + \sigma(T_q, v)) \\ &= \sum_{v \in P_p \setminus P_q} \delta_{T_1} (2\sigma(T_0, v) + \delta_{T_1}) + \sum_{v \in P_q \setminus P_p} (-\delta_{T_1}) (2\sigma(T_0, v) + \delta_{T_1}) \\ &= 2\delta_{T_1} \sum_{v \in P_p \setminus P_q} \sigma(T_0, v) - 2\delta_{T_1} \sum_{v \in P_q \setminus P_p} \sigma(T_0, v) \\ &\quad + \delta_{T_1}^2 (|P_p \setminus P_q| - |P_q \setminus P_p|) \\ &= 2\delta_{T_1} \left( \sum_{v \in P_p \setminus P_q} \sigma(T_0, v) - \sum_{v \in P_q \setminus P_p} \sigma(T_0, v) \right) \\ &= \Delta \end{aligned}$$

as  $|P_p \setminus P_q| = |P_q \setminus P_p|$  because  $d(T_0, p) = d(T_0, q)$ . The theorem now follows.  $\square$

Based on the previous theorem we may define a useful operation on a tree (using the same notation as in the theorem). A *switching* is to replace the tree  $T_0 \oplus_p T_1$  by the tree  $T_0 \oplus_q T_1$ . Loosely speaking, this corresponds to “moving” the subtree  $T_1$  from a vertex  $p$  to a vertex  $q$ . In particular, a switching preserves all distances from the root, and the change in  $\rho_c$  is described in Theorem 6.3.



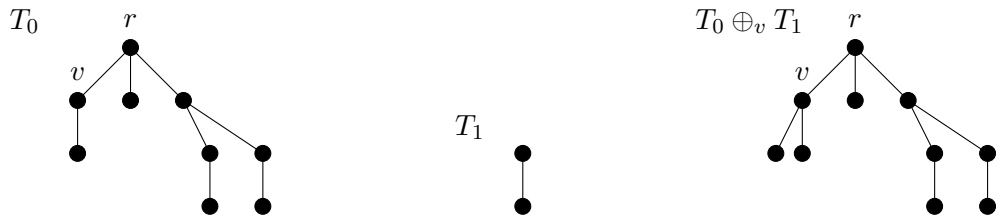


Figure 5. The tree  $T_0 \oplus_v T_1$

The next result uses such switchings to maximize  $\rho_c$  for a given distance vector. Let  $\mathcal{D}_n$  denote the set of all distance vectors of rooted trees with  $n$  vertices, i.e., all vectors  $d = (d_1, d_2, \dots, d_n)$  with

$$1 = d_1, 2 = d_2 = \dots = d_{i_2}, 3 = d_{i_2+1} = \dots = d_{i_3} = \dots$$

where  $1 = i_1 < i_2 < i_3 < \dots$  and  $i_k = n$  for some  $k$ . For a given  $d \in \mathcal{D}_n$ , let  $C_d$  denote the (unique) caterpillar having distance vector  $d$ .

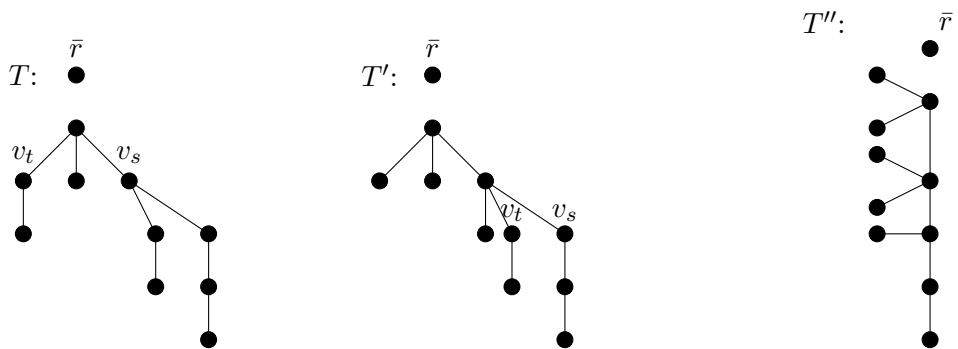


Figure 6. Proof of Theorem 6.4 and  $C_d$  with  $d = (1, 2, 2, 2, 3, 3, 3, 4, 4, 5)$ .

**THEOREM 6.4** *Let  $d \in \mathcal{D}_n$ . The maximum of  $\rho_c(T)$  over all rooted trees with distance vector  $d$  is attained for the caterpillar  $C_d$ , and no other tree attains this maximum.*

**Proof.** Suppose that  $T$  is not a caterpillar. We organize the tree  $T$  in  $k \geq 1$  levels, where level  $s$  consists of vertices  $j$  with  $d_j = s$ . The only vertex at level 1 is  $v_1$ , the internal root. Let  $v_s$  and  $v_t$  be two vertices at level 2 both being adjacent to at least one vertex at level 3. Then two possible operations can be done:

- (1) “moving” all vertices at level 3 from  $v_s$  to  $v_t$ , and
- (2) the opposite, that is, “moving” all vertices from  $v_t$  over to  $v_s$ .

Both possibilities satisfy the condition of Theorem 6.3 and therefore must increase  $\rho_c(T)$ . Then, choose the switching to vertex  $v_s$ . As long as there are at least 2 vertices at level 2 as described above we continue this procedure. After that, only one vertex at level 2 has adjacent vertices at level 3. Therefore we have a caterpillar from the root  $r$  down to the level 2 considering all vertices except those that are adjacent to vertices at level 4, if any. Then, if there are vertices adjacent to vertices at level 4, consider again two vertices at level 3 and do the same procedure until we get a caterpillar  $C_d$  with the same distance vector  $d$  as  $T$ .  $\square$

The construction in the proof of Theorem 6.4 is illustrated in Figure 6 for a particular tree.

One may also consider the problem of minimizing  $\rho_c$  for a given distance vector. This seems more difficult: the reverse of the operation used for maximization may be used to reduce  $\rho_c$ , but this does not directly lead to a global minimizer.

**Concluding remarks.** We mention some interesting further questions to study in this area; some work along these lines is in progress. As mentioned in the Introduction, and given that  $\rho_c(T)$  approximates the Perron value of the branch well, it seems natural to use  $\rho_c(T)$  for studying algebraic connectivity and approximation of the characteristic set of a tree. Another interesting topic is if upper bounds on the Perron value, better than  $\|M\|_1$ , may be found by distance-based techniques similar to what we used here. Finally, it would be good to establish further results on how the combinatorial Perron value reflects certain properties of trees.

**Acknowledgment.** The authors wish to thank two referees for very useful comments that improved the paper significantly, and for suggestions for further work.

## References

- [1] Abreu NMM. Old and new results on algebraic connectivity of graphs. *Linear Algebra Appl.* 2007; 423: 53–73.
- [2] Abreu NMM, Markenzon L, Lee, Rojo O. On trees with maximum algebraic connectivity. *Appl. Anal. and Discrete Math.* 2016; 10: 88–101.
- [3] Alon N. Eigenvalues and expanders. *Combinatorics.* 1986; 6: 85–98.
- [4] Bapat RB. *Graphs and Matrices.* London: Springer; 2010.
- [5] Brualdi RA, Ryser HJ. *Combinatorial Matrix Theory.* Cambridge: Cambridge University Press; 1991.
- [6] Cvetković D, Rowlinson D, Simić S. *An Introduction to the Theory of Graph Spectra.* Cambridge: Cambridge University Press; 2010.
- [7] Fiedler M. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal.* 1973; 23: 298–305.
- [8] Grone R, Merris R. The Laplacian spectrum of a graph II. *SIAM J. Discrete Math.* 1994; 7 (2): 221–229.
- [9] Harary F. *Graph Theory.* Reading: Addison-Wesley; 1969.
- [10] Kirkland S, Neumann M, Shader S. Characteristic Vertices of weighted trees via Perron values. *Linear Multilinear Algebra.* 1996; 40: 311–325.
- [11] Kirkland S, Neumann M. Algebraic connectivity of weighted trees under perturbation. *Linear and Multilinear Algebra.* 1997; 42: 187–203.
- [12] Kirkland S, Moliterno JJ, Neumann M, Shader B. On graphs with equal algebraic and vertex connectivity. *Linear Algebra Appl.* 2002; 341: 45–56.
- [13] Mohar B. Isoperimetric numbers of graphs. *J. Comb. Theory (B).* 1989; 47: 274–291.
- [14] Marshall AW, Olkin I, Arnold BC. *Inequalities: Theory of Majorization and Its Applications.* New York: Springer, Second edition; 2011.
- [15] Minc H. *Nonnegative Matrices.* John Wiley & Sons, Inc.; 1988.
- [16] Moliterno JJ. *Applications of Combinatorial Matrix Theory to Laplacian Matrices of Graphs.* Boca Raton: CRC Press; 2012.
- [17] Tveito A, Winther R. *Introduction to Partial Differential Equations: A Computational Approach,* Berlin: Springer, Second edition; 2005.