ON THE DISTANCE SPECTRAL RADIUS OF TREES WITH GIVEN DEGREE SEQUENCE

KENNETH DADEDZI

VALISOA RAZANAJATOVO MISANANTENAINA

AND

STEPHAN WAGNER

Department of Mathematical Sciences
Stellenbosch University
South Africa

e-mail: dadedzi@sun.ac.za
valisoa@sun.ac.za
swagner@sun.ac.za

This paper is dedicated to the memory of Slobodan Simić.

Abstract

We consider the problem of maximizing the distance spectral radius and a slight generalization thereof among all trees with some prescribed degree sequence. We prove in particular that the maximum of the distance spectral radius has to be attained by a caterpillar for any given degree sequence. The same holds true for the terminal distance matrix. Moreover, we consider a generalized version of the reverse distance matrix and also study its spectral radius for trees with given degree sequence. We prove that the spectral radius is always maximized by a greedy tree. This implies several corollaries, among them a "reversed" version of a conjecture of Stevanović and Ilić. Our results parallel similar theorems for the Wiener index and other invariants.

Keywords: distance matrix, spectral radius, tree, degree sequence.

2010 Mathematics Subject Classification: 05C05, 05C50, 05C12.

1. INTRODUCTION AND PRELIMINARIES

Let G be a connected graph whose vertices are $v_1, v_2, \ldots, v_n$. The *distance matrix* of G, denoted by $D(G)$, is the symmetric matrix whose ij-th entry is the length

of the shortest path from vertex v_i to vertex v_j; it is one of the classical matrices associated with a graph. Its study in the context of spectral graph theory traces back to the work of Graham and Pollak [13]. Among other things, they proved a very elegant formula for the determinant of the distance matrix of a tree: we have

$$\det(D(T)) = (n-1)(-1)^{n-1}2^{n-2},$$

where n is the number of vertices. Notably, the determinant depends only on the number of vertices of the tree, but not its structure. This was generalized in several different directions. Collins [10] and Bapat [2] (the latter for matrix weights) extended the result to a weighted tree T on n vertices with edge weights α_i $(i = 1, \ldots, n-1)$. They proved that the determinant of the distance matrix (where distance is now defined as the sum of the edge weights along the shortest path) is given by

$$\det(D(T)) = (-1)^{n-1}2^{n-2} \left(\prod_{i=1}^{n-1} \alpha_i \right) \left(\sum_{i=1}^{n-1} \alpha_i \right),$$

which is still independent of the tree structure and the assignment of the weights to the edges, but depends on the edge weights and the number of vertices.

Furthermore, Bapat, Kirkland and Neumann proved in [4] that for a weighted tree T on n vertices with edge weights α_i $(i = 1, \ldots, n-1)$ and any real constant c, the determinant of the matrix $H_c(T) = D(T) + cJ$, where J is a matrix with all entries equal to 1, is

$$(1) \qquad \det(H_c(T)) = (-1)^{n-1}2^{n-2} \left(\prod_{i=1}^{n-1} \alpha_i \right) \left(2c + \sum_{i=1}^{n-1} \alpha_i \right).$$

In particular, if all weights are equal to 1, then we have

$$\det(H_c(T)) = (-1)^{n-1}2^{n-2}(2c + n - 1),$$

which again is independent of the structure of the tree.

In analogy to the matrix $H_c(T)$, let us also consider the matrix $M_c(T) = cJ - D(T)$. By a simple substitution, one finds that

$$\begin{aligned} \det(M_c(T)) &= (-1)^n \det(D(T) - cJ) \\ &= (-1)^n \left((-1)^{n-1}2^{n-2}(-2c + (n-1)) \right) \\ &= 2^{n-2}(2c - n + 1) \end{aligned}$$

for every tree T with n vertices, which is also independent of the structure of the tree, but only depends on the constant c and the number of vertices n. The

reverse distance matrix is usually defined as the matrix $\text{diam}(T)J - D(T)$, where $\text{diam}(T)$ is the diameter of T [1]. This is clearly a special case of the matrix M_c, where $c = \text{diam}(T)$. Note that $\text{diam}(T)$ is also the smallest possible choice of c for which we obtain a matrix with non-negative entries.

The spectrum of the distance matrix only depends on the graph G, not the ordering of its vertices, as it is the case for other classical matrices (e.g. adjacency matrix, Laplacian matrix). Since $D(G)$ is symmetric, all its eigenvalues are real, and they sum to 0 since the trace of $D(G)$ is 0. Moreover, all entries except for those on the diagonal are positive, so by the Perron-Frobenius theorem there is a unique eigenvalue with an eigenvector whose entries are all positive, and this eigenvalue is also equal to the spectral radius of $D(G)$, which is also called the *distance spectral radius* of G. When G is a tree, the spectral radius is in fact the only positive eigenvalue, all the others are negative [3, p. 104].

The Perron-Frobenius theorem also applies to $H_c(G)$ if c is non-negative, and to M_c as long as c is greater than or equal to the diameter, so that all entries are positive. In this paper, we will be interested in the spectral radius of matrices of the form $H_c(G)$ or $M_c(G)$, specifically its maximum value for trees with a given degree sequence. The spectral radius of $H_c(G)$ will be denoted by $\eta_c(G)$, and the spectral radius of $M_c(G)$ by $\mu_c(G)$. Clearly, the distance spectral radius is precisely the special case $\eta_0(G)$.

The distance spectral radius is closely related to other distance-based graph invariants, the Wiener index being a notable example. The *Wiener index* of a graph is defined as the sum of the distances between all (unordered) pairs of vertices, which is exactly half of the sum of all the entries in the distance matrix. The following inequality is a consequence of this fact.

Theorem 1 [17]. *Let $\rho(G)$ be the distance spectral radius of a graph G of order n. Then*

$$\rho(G) \geq \frac{2W(G)}{n}.$$

The proof of this theorem is based on the concept of a Rayleigh quotient, which will also play a major role in this paper.

Recall that the *Rayleigh quotient* of a vector $\boldsymbol{f}$ with respect to a matrix A is the quotient

$$\frac{\boldsymbol{f}^T A \boldsymbol{f}}{\boldsymbol{f}^T \boldsymbol{f}}.$$

If $\boldsymbol{f}$ is a unit vector, this simply reduces to $\boldsymbol{f}^T A \boldsymbol{f}$. It is well known that the largest eigenvalue of a symmetric matrix is the maximum of the Rayleigh quotient, taken either over all nonzero vectors or just all unit vectors. Thus the spectral radius of $D(G)$ is bounded below by the Rayleigh quotient associated with the vector $\boldsymbol{1}$

whose entries are all equal to 1:

$$\rho(G) \geq \frac{\mathbf{1}^T D(G) \mathbf{1}}{\mathbf{1}^T \mathbf{1}} = \frac{2W(G)}{n}.$$

Equality in Theorem 1 holds if and only if the *total distance* $\sum_w d(v, w)$ has the same value for every vertex v of G; for example, this is trivially the case for all vertex-transitive graphs. With this connection in mind, it is not surprising that the results we obtain in this paper also have known analogues for the Wiener index.

As mentioned before, the matrices we are dealing with here are non-negative matrices to which the Perron-Frobenius theorem applies: each such matrix has a unique non-negative unit eigenvector, which is associated with its spectral radius. This unit eigenvector is called the *Perron vector*.

There is already a large amount of literature on the problem of maximizing or minimizing the distance spectral radius in some family of graphs. Trees are a fundamental and natural family of graphs to consider, and indeed there are many results on various types of trees, such as trees with given matching number [16,22], domination number [29], number of leaves [22, 23], number of odd vertices [18], maximum degree [19–21,25], or diameter [31] and some particular classes of trees [30]. See also [6–8, 28].

As mentioned earlier, our main focus in this paper will be on trees with a prescribed degree sequence. There is a simple characterization of finite sequences that are degree sequences of trees: the trivial identity that stems from the handshake lemma is both necessary and sufficient.

Proposition 2. *A sequence $(d_1, d_2, \ldots, d_n)$ of positive integers is a degree sequence of a tree of order n if and only if*

$$\sum_{i=1}^n d_i = 2(n - 1).$$

We will call a sequence that satisfies the condition of Proposition 2 a *tree degree sequence*. For such a tree degree sequence, we are interested in the trees that maximize the spectral radii η_c and μ_c. Greedy trees and caterpillars will play a major role in our study. Let us first present their formal definitions.

Definition [27]. The greedy tree $G(\alpha)$, also known as BFD-tree [5], is a rooted tree obtained from a degree sequence α by the following "greedy algorithm":

(i) assign the largest degree to the root r;

(ii) label the neighbors of r as $v_1, v_2, \ldots, v_{d(r)}$, from left to right, and assign the largest degrees available to them such that $d(v_1) \geq d(v_2) \geq \cdots \geq d(v_{d(r)})$;

(iii) label the neighbors of v_1 (except r) as $v_{11}, v_{12}, \ldots$, and assign degrees as in (ii);

(iv) repeat (iii) for $v_2, v_3, \ldots, v_{d(r)}$, then the vertices at distance 2 from the root, and so on, from top to bottom and left to right; at each step, the next vertex is assigned the largest degree in the sequence that has not been assigned yet.

Example 3. We construct a greedy tree for the degree sequence:

$$\alpha = (4, 3, 3, 2, 2, 2, 1, 1, 1, 1, 1, 1).$$

The degree of the root r is 4, and it has the neighbors v_1, v_2, v_3, v_4. We therefore assign the next largest available degree, which is 3, to v_1 and repeat the procedure for the other vertices. Following the algorithm yields Figure 1.
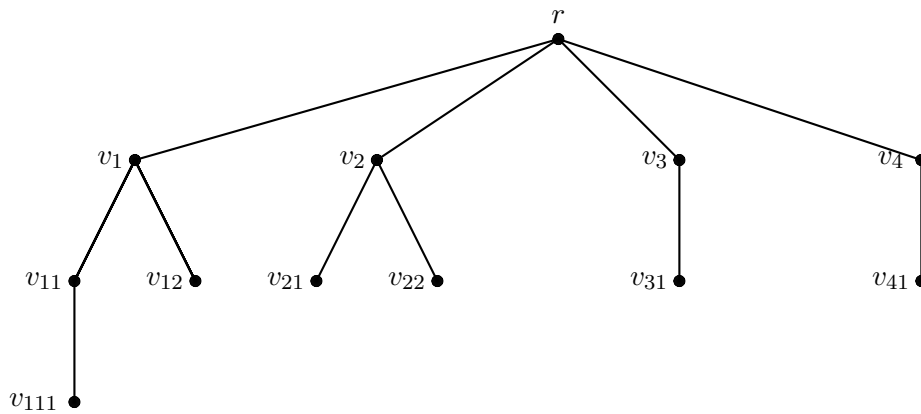

Figure 1. A greedy tree $G(\alpha)$.

The greedy tree is known to maximize, among others, the spectral radius [5] and Laplacian spectral radius [32]. It is also known that the minimum of the Wiener index among all trees with degree sequence α is the greedy tree $G(\alpha)$, see [27]. An important special case is the so-called Volkmann tree. In [11], it is shown that the Volkmann tree minimizes the Wiener index among trees with n vertices and maximum degree Δ.

Definition. The Volkmann tree $V_{n,\Delta}$ is a special case of a greedy tree; it is characterized by its order n and maximum degree Δ. Its degree sequence is $\alpha = (\Delta, \ldots, \Delta, r, 1, \ldots, 1)$, where $r \in \{1, 2, \ldots, \Delta - 1\}$ satisfies $r \equiv n - 1 \bmod \Delta - 1$.

Figure 2 shows the Volkmann tree $V_{20,3}$, whose degree sequence is

$$\alpha = (3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1).$$

A conjecture of Stevanović and Ilić in [25] states that among all trees with order n and maximum degree Δ, the Volkmann tree minimizes the spectral radius of the distance matrix. A special case was proven by Luo and Zhou [21] and Lin *et al.* [20]. See also [12]. While we will not be able to prove this conjecture, we will add evidence in its favour by proving a "reversed" version (see Corollary 20).
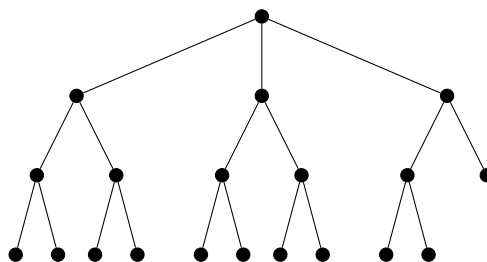

Figure 2. The Volkmann tree $V_{20,3}$.

It is harder to characterize the trees for which the maximum is attained, but one can show that they are always so-called caterpillars.

Definition. A caterpillar is a tree with the property that removing its leaves yields a path. This path will be called the backbone of the caterpillar. In particular, a path is also a caterpillar (and we will even consider a path with only one or two vertices a caterpillar).

Figure 3 shows an example of a caterpillar whose degree sequence is $\alpha = (5, 4, 3, 1, 1, 1, 1, 1, 1, 1, 1)$.
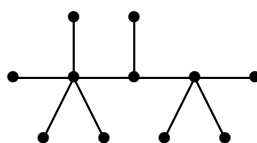

Figure 3. A caterpillar.

In the following section, we prove that the maximum distance spectral radius among all trees with a given degree sequence is always attained by a caterpillar, paralleling similar results e.g. for the Wiener index. In fact, we prove a slightly more general statement on the spectral radius of matrices of the form $H_c(T) = cJ + D(T)$. While we cannot describe the precise shape of this maximizing caterpillar in general, a partial characterization can be given.

Thereafter, we look at a "reversed" problem. For a fixed real number c, we consider matrices of the form $M_c(T) = cJ - D(T)$. For sufficiently large c,

this is still a non-negative matrix, and we will show that the maximum spectral radius of $M_c(T)$, given the degree sequence of T, is always attained by a greedy tree. One can also deduce results on trees with a given number of leaves or given maximum degree from this general theorem by a technique based on majorization of sequences.

2. The Spectral Radius of a Generalized Distance Matrix

In this section, we consider the problem of maximizing the distance spectral radius (and more generally the spectral radius of $H_c(T)$) of trees whose degree sequence is given. Our main result reads as follows.

Theorem 4. *Let $c \geq 0$ be a constant, and let α be a fixed tree degree sequence. Let $T_{\max}$ be a tree that maximizes the spectral radius of $H_c(T) = cJ + D(T)$ among all trees with degree sequence α. Then $T_{\max}$ is a caterpillar.*

The proof will follow the lines of [24] (which dealt with the Wiener index and other distance-based graph invariants) in many ways; however, there are some subtleties that require us to modify the argument. Before we begin the proof, let us state a few definitions.

Definition. A *caterpillar branch* of a tree T is a maximal induced subgraph B of T that is a non-trivial caterpillar (i.e., a caterpillar with more than one vertex) and has the property that $T \setminus B$ is still a tree.

The proof of Theorem 4 is by contradiction: for every non-caterpillar, we construct a tree with the same degree sequence for which the spectral radius is greater. To this end, we consider the Perron vector $\boldsymbol{f}$ and show that the Rayleigh quotient of the same vector with respect to the newly constructed tree is greater.

Let us consider c as fixed in the following. The Rayleigh quotient of the matrix $H_c(T) = cJ + D(T)$ of a tree T on a unit vector $\boldsymbol{f}$, which we denote by $R_T(\boldsymbol{f})$, is given by

$$R_T(\boldsymbol{f}) = \sum_{u \in V(T)} \sum_{v \in V(T)} (c + d_T(u, v)) f(u) f(v)$$

$$= c \left(\sum_{u \in V(T)} f(u) \right)^2 + 2 \sum_{\{u,v\} \subseteq V(T)} d_T(u, v) f(u) f(v),$$

where $f(u)$ is the entry of $\boldsymbol{f}$ associated with the vertex u. If $\boldsymbol{f}$ is the Perron vector of $H_c(T)$ (which is what we will generally assume in the following), then we have $\eta_c(T) = R_T(\boldsymbol{f})$. For a subset S of the vertices of T, we set

$$|S|_{\boldsymbol{f}} = \sum_{u \in S} f(u)$$

and call it the *weight* of S (with respect to $\boldsymbol{f}$). Since $\boldsymbol{f}$ has only positive entries, it is clear that $|S|_{\boldsymbol{f}} > 0$ for every nonempty vertex set S.

***Proof of Theorem* 4.** Let $T_{\max}$ be an optimal tree (i.e., a tree that maximizes the spectral radius of $H_c(T)$ for some fixed c) among all trees with some prescribed degree sequence, and let $\boldsymbol{f}$ be the associated Perron vector. Suppose that $T_{\max}$ is not a caterpillar. We consider pairs B_1, B_2 of caterpillar branches in $T_{\max}$ that are attached to the same vertex v (see Figure 4); such pairs must always exist if the tree is not a caterpillar. To see why this is the case, consider the tree T_{red} that results when all leaves are removed. Since the original tree is not a caterpillar, the result is not a path. Now fix any leaf and consider the branching vertex (vertex of degree at least 3) in T_{red} whose distance is greatest. At least two of the branches going out from this vertex are paths in T_{red} and thus caterpillar branches in $T_{\max}$, since there would otherwise be a branching vertex at a greater distance.
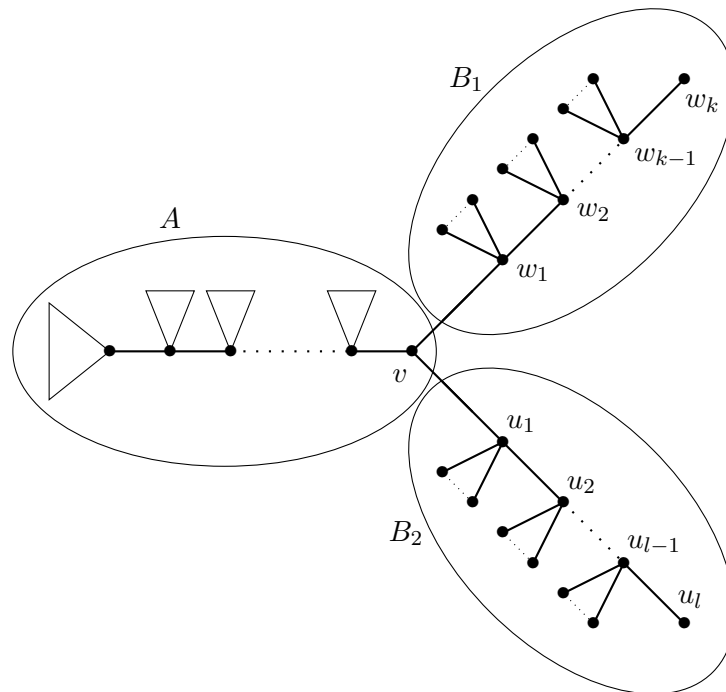

Figure 4. B_1, B_2 and A in an optimal tree $T_{\max}$ that is not a caterpillar.

Among all such pairs of two caterpillar branches attached to the same vertex, we choose the pair B_1, B_2 for which $|B_1|_{\boldsymbol{f}} + |B_2|_{\boldsymbol{f}}$ is minimal. Let $P = vw_1w_2 \cdots w_k$ and $P' = vu_1u_2 \cdots u_l$ be the longest paths (backbones) of the caterpillar branches B_1 and B_2 respectively, as indicated in the figure. Furthermore, we denote $T_{\max} \setminus (B_1 \cup B_2)$ by A. Without loss of generality, we can assume that

$|B_2|_{\boldsymbol{f}} \geq |B_1|_{\boldsymbol{f}}$. Note that $|A|_{\boldsymbol{f}} > |B_2|_{\boldsymbol{f}}$ by our choice of the pair B_1, B_2. If this was not the case, then we could reach a contradiction in one of two ways:

- If $A \setminus v$ is a caterpillar branch, then consider the pair $B_1, A \setminus v$ instead of B_1, B_2; since $|B_1|_{\boldsymbol{f}} + |A \setminus v|_{\boldsymbol{f}} < |B_1|_{\boldsymbol{f}} + |A|_{\boldsymbol{f}} \leq |B_1|_{\boldsymbol{f}} + |B_2|_{\boldsymbol{f}}$, we would have a contradiction to the choice of B_1, B_2.
- Otherwise, we can find a pair B_1', B_2' of caterpillar branches inside of A, and their combined weight would be $|B_1'|_{\boldsymbol{f}} + |B_2'|_{\boldsymbol{f}} \leq |A|_{\boldsymbol{f}} < |B_1|_{\boldsymbol{f}} + |B_2|_{\boldsymbol{f}}$, again with a contradiction.

We form a new tree T' from $T_{\max}$ by exchanging B_2 and w_k. That is, we remove the edges $w_{k-1}w_k$ and vu_1, and add the edges $w_{k-1}u_1$ and vw_k. Clearly, this does not change the degree sequence. Note that this operation only changes the distance between two vertices x and y if $x \in A$ or $x \in B_1 \setminus w_k$ and $y \in B_2$ or $y = w_k$ (or vice versa). Let us now compute

$$R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f}) = 2 \left(\sum_{\{x,y\} \subseteq V(T_{\max})} \left(d_{T'}(x,y) - d_{T_{\max}}(x,y) \right) f(x) f(y) \right).$$

There are four different cases to consider for x and y.

Case 1. One of the two lies in A, the other in B_2. The contribution of all such pairs to $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$ is

$$2 \sum_{\substack{x \in A \\ y \in B_2}} \left(d_{T'}(x,y) - d_{T_{\max}}(x,y) \right) f(x) f(y)$$

$$= 2 \sum_{\substack{x \in A \\ y \in B_2}} \Big(d_{T_{\max}}(x,v) + d_{T_{\max}}(v,w_k) + d_{T_{\max}}(u_1,y)$$

$$- d_{T_{\max}}(x,v) - d_{T_{\max}}(v,u_1) - d_{T_{\max}}(u_1,y) \Big) f(x) f(y)$$

$$= 2 \sum_{\substack{x \in A \\ y \in B_2}} \left(d_{T_{\max}}(v,w_k) - 1 \right) f(x) f(y) = 2(k-1)|A|_{\boldsymbol{f}}|B_2|_{\boldsymbol{f}}.$$

Case 2. One of the two lies in A, the other is w_k. The contribution of all such pairs to $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$ is

$$2 \sum_{x \in A} \left(d_{T'}(x,w_k) - d_{T_{\max}}(x,w_k) \right) f(x) f(w_k)$$

$$= 2 \sum_{x \in A} \left(d_{T_{\max}}(x,v) + 1 - d_{T_{\max}}(x,v) - d_{T_{\max}}(v,w_k) \right) f(x) f(w_k)$$

$$= 2(1-k)|A|_{\boldsymbol{f}} f(w_k).$$

Case 3. One of the two lies in $B_1 \setminus \{w_k\}$, the other in B_2. The contribution of all such pairs to $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$ is

$$2 \sum_{\substack{x \in B_1 \setminus \{w_k\} \\ y \in B_2}} \big(d_{T'}(x,y) - d_{T_{\max}}(x,y)\big) f(x) f(y)$$

$$= 2 \sum_{\substack{x \in B_1 \setminus \{w_{k-1}, w_k\} \\ y \in B_2}} \big(d_{T_{\max}}(x, w_k) + d_{T_{\max}}(u_1, y) - d_{T_{\max}}(x, v) - 1 - d_{T_{\max}}(u_1, y)\big) f(x) f(y)$$

$$+ 2 \sum_{y \in B_2} \big(1 + d_{T_{\max}}(u_1, y) - d_{T_{\max}}(w_{k-1}, v) - 1 - d_{T_{\max}}(u_1, y)\big) f(w_{k-1}) f(y)$$

$$= 2 \sum_{x \in B_1 \setminus \{w_{k-1}, w_k\}} \big(d_{T_{\max}}(x, w_k) - d_{T_{\max}}(x, v) - 1\big) f(x) |B_2|_{\boldsymbol{f}} + 2(1 - k) f(w_{k-1}) |B_2|_{\boldsymbol{f}}.$$

Case 4. One of the two lies in $B_1 \setminus \{w_k\}$, the other is w_k. The contribution of all such pairs to $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$ is

$$2 \sum_{x \in B_1 \setminus \{w_k\}} \big(d_{T'}(x, w_k) - d_{T_{\max}}(x, w_k)\big) f(x) f(w_k)$$

$$= 2 \sum_{x \in B_1 \setminus \{w_{k-1}, w_k\}} \big(d_{T_{\max}}(x, v) + 1 - d_{T_{\max}}(x, w_k)\big) f(x) f(w_k)$$

$$+ 2 \big(d_{T_{\max}}(w_{k-1}, v) + 1 - 1\big) f(w_{k-1}) f(w_k)$$

$$= 2 \sum_{x \in B_1 \setminus \{w_{k-1}, w_k\}} \big(d_{T_{\max}}(x, v) + 1 - d_{T_{\max}}(x, w_k)\big) f(x) f(w_k)$$

$$+ 2(k - 1) f(w_{k-1}) f(w_k).$$

Combining the last two cases, we obtain

$$2 \sum_{x \in B_1 \setminus \{w_{k-1}, w_k\}} \big(d_{T_{\max}}(x, w_k) - d_{T_{\max}}(x, v) - 1\big) f(x) (|B_2|_{\boldsymbol{f}} - f(w_k))$$

$$+ 2 f(w_{k-1})(1 - k)(|B_2|_{\boldsymbol{f}} - f(w_k)).$$

Observe that $|B_2|_{\boldsymbol{f}} - f(w_k) > |B_2|_{\boldsymbol{f}} - |B_1|_{\boldsymbol{f}} \geq 0$ by the assumption that $|B_2|_{\boldsymbol{f}} \geq |B_1|_{\boldsymbol{f}}$. Note also that $d_{T_{\max}}(x, w_k) \geq 2$ for all $x \in B_1 \setminus \{w_{k-1}, w_k\}$. Moreover, $d_{T_{\max}}(x, v) \leq d_{T_{\max}}(w_k, v) = k$ for all $x \in B_1 \setminus \{w_k\}$. Using these simple estimates, we find that the contribution of Cases 3 and 4 to $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$ is greater than or equal to

$$2 \sum_{x \in B_1 \setminus \{w_{k-1}, w_k\}} f(x)(2 - 1 - k)(|B_2|_{\boldsymbol{f}} - f(w_k)) + 2 f(w_{k-1})(1 - k)(|B_2|_{\boldsymbol{f}} - f(w_k))$$

$$= 2(1 - k) |B_1 \setminus \{w_k\}|_{\boldsymbol{f}} (|B_2|_{\boldsymbol{f}} - f(w_k)).$$

Together with Cases 1 and 2, we obtain

$$R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$$

$$\geq 2(k-1)\big(|A|_{\boldsymbol{f}}(|B_2|_{\boldsymbol{f}} - f(w_k)) - |B_1 \setminus \{w_k\}|_{\boldsymbol{f}}(|B_2|_{\boldsymbol{f}} - f(w_k))\big)$$

$$= 2(k-1)\big(|A|_{\boldsymbol{f}} - |B_1 \setminus \{w_k\}|_{\boldsymbol{f}}\big)\big(|B_2|_{\boldsymbol{f}} - f(w_k)\big),$$

which is positive since $|A|_{\boldsymbol{f}} > |B_2|_{\boldsymbol{f}} \geq |B_1|_{\boldsymbol{f}} > f(w_k)$. Hence we have

$$\eta_c(T') \geq R_{T'}(\boldsymbol{f}) > R_{T_{\max}}(\boldsymbol{f}) = \eta_c(T_{\max}).$$

This contradicts the optimality of $T_{\max}$. ∎

Example 5. Figure 5 shows the respective extremal trees for the degree sequences

$$\alpha_1 = (5, 4, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1),$$
$$\alpha_2 = (8, 5, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1),$$
$$\alpha_3 = (10, 4, 3, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1),$$
$$\alpha_4 = (8, 4, 4, 4, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1).$$
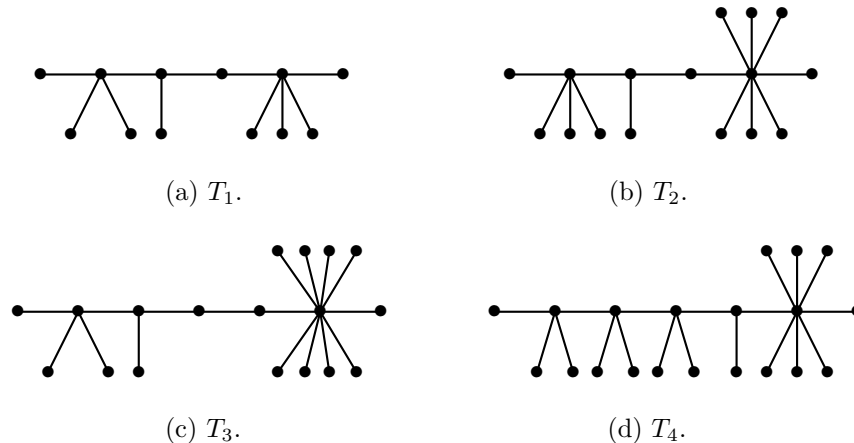

(a) T_1.

(b) T_2.

(c) T_3.

(d) T_4.

Figure 5. Extremal trees T_1, T_2, T_3 and T_4.

Note that Theorem 4 does not uniquely determine $T_{\max}$ except for some very special cases (e.g. trees whose vertices only have degree 1 or d for some fixed d). Characterizing $T_{\max}$ in general appears to be a difficult problem, and even the analogous problem for the Wiener index does not have a simple solution (an efficient algorithm is available, however—see [9]).

However, we can prove the following general statement about the shape of the extremal caterpillar: its vertex degrees form a "V-pattern".

Theorem 6. *Let $c \geq 0$ be a constant, and let α be a fixed tree degree sequence. Let $T_{\max}$ be a caterpillar that maximizes the spectral radius of $H_c(T) = cJ + D(T)$ among all trees with degree sequence α, and let the degrees of the internal vertices along its backbone be $d_1, d_2, \ldots, d_k$, in this order. There exists an index m such that*

$$(2) \qquad\qquad d_1 \geq d_2 \geq \cdots \geq d_m \leq \cdots \leq d_{k-1} \leq d_k.$$

Proof. Let $\boldsymbol{f}$ denote the Perron vector as in the proof of Theorem 4, and let S_i be the subtree induced by the i-th vertex on the backbone and its leaf neighbors ($1 \leq i \leq k$). Clearly, the S_i are disjoint, and their union is the entire tree $T_{\max}$. Choose j to be the largest index for which

$$\sum_{i<j} |S_i|_{\boldsymbol{f}} < \sum_{i>j} |S_i|_{\boldsymbol{f}}.$$

Since this inequality is trivially satisfied for $j = 1$, and trivially not satisfied for $j = k$ (unless $k = 1$, in which case there is nothing to prove), such an index must exist.

We claim that the degrees satisfy the inequality (2) either for $m = j$ or for $m = j + 1$. Let us assume for contradiction that $d_i < d_{i+1}$ for some $i < j$, and consider the tree T' obtained by moving $d_{i+1} - d_i$ leaves from S_{i+1} to S_i. This set of leaves will be denoted by L. Observe that T' and $T_{\max}$ have the same degree sequence. We also note that the only distances that change are those between the leaves in L and the other vertices, and that all these distances change by precisely 1. Now consider the difference

$$R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f}) = 2 \left(\sum_{\{x,y\} \subseteq V(T_{\max})} \big(d_{T'}(x,y) - d_{T_{\max}}(x,y)\big) f(x)f(y) \right)$$

of the Rayleigh quotients. This difference simplifies to

$$R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f}) = 2 \sum_{x \in L} \sum_{y \in (S_{i+1} \setminus L) \cup S_{i+2} \cup \cdots \cup S_k} f(x)f(y) - 2 \sum_{x \in L} \sum_{y \in S_1 \cup \cdots \cup S_i} f(x)f(y)$$

$$= 2|L|_{\boldsymbol{f}} \left(|S_{i+1}|_{\boldsymbol{f}} - |L|_{\boldsymbol{f}} + \sum_{h=i+2}^{k} |S_h|_{\boldsymbol{f}} - \sum_{h=1}^{i} |S_h|_{\boldsymbol{f}} \right)$$

$$> 2|L|_{\boldsymbol{f}} \left(\sum_{h=j+1}^{k} |S_h|_{\boldsymbol{f}} - \sum_{h=1}^{j-1} |S_h|_{\boldsymbol{f}} \right),$$

which is positive by our choice of j. Thus we have

$$\eta_c(T') \geq R_{T'}(\boldsymbol{f}) > R_{T_{\max}}(\boldsymbol{f}) = \eta_c(T_{\max}),$$

contradicting the choice of $T_{\max}$. So we can conclude that

$$d_1 \geq d_2 \geq \cdots \geq d_j.$$

By maximality of j, we have

$$\sum_{i<j+1} |S_i|_{\boldsymbol{f}} \geq \sum_{i>j+1} |S_i|_{\boldsymbol{f}}.$$

So repeating the same argument, we find that

$$d_{j+1} \leq d_{j+2} \leq \cdots \leq d_k.$$

Combining the two chains of inequalities, we obtain (2), either with $m = j$ or with $m = j + 1$. ∎

It is worth pointing out that the results of this section also apply to a slightly different matrix known as the terminal distance matrix (see [15] for some recent results on the spectral radius of the terminal distance matrix of trees). Its formal definition reads as follows.

Definition. Let $\mathcal{P}(T) = \{w_1, w_2, \ldots, w_k\}$ be the set of leaves of a tree T. The terminal distance matrix (or reduced distance matrix) $TD(T)$ is a symmetric matrix whose ij-th entry is $d(w_i, w_j)$.

Theorem 7. *Let $c \geq 0$ be a constant, and let α be a fixed tree degree sequence. Let $T_{\max}$ be a tree that maximizes the spectral radius of $cJ + TD(T)$ among all trees with degree sequence α. Then $T_{\max}$ is a caterpillar, and its vertex degrees follow the pattern described in Theorem 6.*

Proof. The proof is analogous to the proofs of Theorems 4 and 6. ∎

3. The Spectral Radius of a Generalized Reverse Distance Matrix

It is natural to also consider the minimization problem for the distance spectral radius or generally the spectral radius of $H_c(T)$. Indeed, as pointed out in the introduction, there are concrete conjectures in this regard, and since greedy trees are often extremal with respect to distance-based graph invariants (such as the Wiener index), one would expect that this is the case for the distance spectral radius as well. While this appears to be the case, we have been unable to prove it, but we provide evidence in this direction. Minimizing the spectral radius of $D(T)$ is similar to maximizing the spectral radius of a "reversed" matrix of the form $M_c(T) = cJ - D(T)$, as defined earlier. For the latter problem, we succeed in proving that the greedy trees are extremal.

The matrix $M_c(T)$ has only non-negative entries if c is greater than or equal to the diameter of T. If we let ν denote the number of internal vertices (non-leaves) of T, then the diameter is at most equal to $\nu+1$, since any diametral path starts and ends at a leaf and otherwise only contains internal vertices, thus at most $\nu+2$ vertices in total. For a caterpillar, this value is actually attained. If the entries of $M_c(T)$ are non-negative, then the Perron-Frobenius theorem applies, which will be important for us in our analysis of the spectral radius. Let us now formulate the main theorem of this section.

Theorem 8. *Let α be a tree degree sequence, and let ν be the number of entries greater than 1 (number of non-leaves) in α. For every constant $c \geq \nu + 1$, the greedy tree $G(\alpha)$ is the unique tree with degree sequence α that maximizes the spectral radius of the matrix $M_c(T) = cJ - D(T)$.*

The proof of Theorem 8 is based on similar results for the Wiener index as obtained by Wang in [27]. The basic idea is to analyze the behaviour of the spectral radius of $M_c(T)$ under certain operations on the tree T. We first need some definitions and lemmas. The following lemma gives a characterization of a greedy tree.

Lemma 9 [27]**.** *A rooted tree T is a greedy tree if*

 (i) *the root r has the largest degree;*

(ii) *the heights of any two leaves differ by at most 1; here, the height $h_T(v)$ of a vertex v is the distance from v to the root r;*

(iii) *for any two vertices v and w, if $h_T(v) < h_T(w)$, then $d(v) \geq d(w)$;*

(iv) *for any two vertices v and w of the same height, if $d(v) > d(w)$, then $d(v') \geq d(w')$ for any successors v' of v and w' of w of the same height;*

 (v) *for any two vertices v and w of the same height, if $d(v) > d(w)$, then $d(v') \geq d(w')$ and $d(v'') \geq d(w'')$ for any siblings v' of v and w' of w or successors v'' of v' and w'' of w' of the same height.*

Now let T be a tree of order n with vertex set $V(T)$, and let $\boldsymbol{f}$ be an $n \times 1$ unit vector. The Rayleigh quotient of $M_c(T)$ on the vector $\boldsymbol{f}$ is given by

$$R_T(\boldsymbol{f}) = \frac{\boldsymbol{f}^T M \boldsymbol{f}}{\boldsymbol{f}^T \boldsymbol{f}} = \boldsymbol{f}^T M \boldsymbol{f} = \sum_{u \in V(T)} \sum_{v \in V(T)} (c - d_T(u,v)) f(u) f(v),$$

where $f(u)$ is the component of $\boldsymbol{f}$ associated with the vertex u. We know that the spectral radius $\mu_c(T)$ of $M_c(T)$ is the maximum of $R_T(\boldsymbol{f})$ over all unit vectors, and this maximum is attained when $\boldsymbol{f}$ is the Perron vector, which is the unit eigenvector corresponding to $\mu_c(T)$. In the following, we assume that $T = T_{\max}$

is a tree with degree sequence α that maximizes the spectral radius of $M_c(T)$, and we will always let $\boldsymbol{f}$ denote the Perron vector of $M_c(T_{\max})$.

Next, we describe a decomposition of a tree T taken from [27] which will aid us in the proof of some technical lemmas that will lead us to Theorem 8. Consider two vertices x and y in the tree T. Let the path $P_T(x,y)$ from the vertex x to the vertex y be given by

$$P_T(x,y) = x_k x_{k-1} \cdots x_2 x_1 y_1 y_2 y_3 \cdots y_{k-1} y_k$$

when its length is odd, or

$$P_T(x,y) = x_k x_{k-1} \cdots x_2 x_1 z y_1 y_2 \cdots y_{k-1} y_k$$

when its length is even, where $x_k = x$, $y_k = y$. Furthermore, we let X_i, Y_i and Z denote the components that contain x_i, y_i and z respectively when the edges of the path $P_T(x,y)$ are removed. Also, we let $X_{>k}$ (respectively, $Y_{>k}$) be the subtree of the tree induced by the vertices in $X_{k+1} \cup X_{k+2} \cup \cdots$ (respectively, $Y_{k+1} \cup Y_{k+2} \cup \cdots$). This decomposition is shown in Figure 6 for odd path length. As in the previous section, we set

$$|S|_{\boldsymbol{f}} = \sum_{u \in S} f(u)$$

for every set S of vertices. Since $\boldsymbol{f}$ has only positive entries, we have $|S|_{\boldsymbol{f}} > 0$ for all nonempty sets S again.
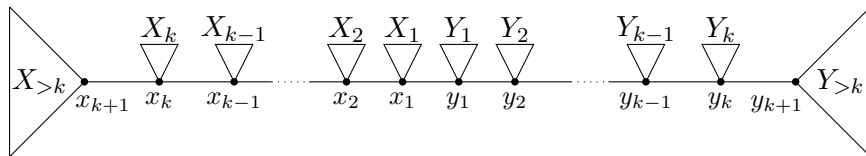

Figure 6. Labelling of a path and the components.

Lemma 10. *Let $T_{\max}$ be decomposed as depicted in Figure 6. If $|X_i|_{\boldsymbol{f}} \geq |Y_i|_{\boldsymbol{f}}$ for $i = 1, 2, \ldots, k-1$ and $|X_{>k-1}|_{\boldsymbol{f}} \geq |Y_{>k-1}|_{\boldsymbol{f}}$, then $d(x_k) \geq d(y_k)$.*

Proof. We assume (for contradiction) that $d(x_k) < d(y_k)$, so that $d(y_k) - d(x_k) = a > 0$. We let u_i ($i = 1, \ldots, a$) be neighbors of y_k other than y_{k-1} and y_{k+1} (chosen arbitrarily). We increase the degree of x_k by removing all the edges $y_k u_i$ ($i = 1, \ldots, a$) and adding all edges $x_k u_i$ ($i = 1, \ldots, a$) instead. Note that this operation does not change the degree sequence; we want to show that it increases the Rayleigh quotient $R_{T_{\max}}(\boldsymbol{f})$. Let T' be the new tree obtained after performing this operation and A be the set of vertices in the components of $T_{\max} \setminus \{y_k u_i : i = 1, \ldots, a\}$ that contain $u_1, u_2, \ldots, u_a$.

We compute the difference $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$:

$$
\begin{aligned}
R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f}) &= 2 \sum_{\{u,v\} \subseteq V(T_{\max})} (c - d_{T'}(u,v)) f(u) f(v) \\
&\quad - 2 \sum_{\{u,v\} \subseteq V(T_{\max})} (c - d_{T_{\max}}(u,v)) f(u) f(v) \\
&= 2 \sum_{\{u,v\} \subseteq V(T_{\max})} (d_{T_{\max}}(u,v) - d_{T'}(u,v)) f(u) f(v).
\end{aligned}
$$

We observe that the distance between two vertices u and v only changes if $u \in A$ and $v \in T_{\max} \backslash A$ (or vice versa). We will consider the case when the path shown in Figure 6 has odd length, the other case being similar. There are four possibilities for the vertices u and v.

Case 1. One of the two lies in X_i for some $i \le k-1$, and the other lies in A. In this case, the distance between the vertices decreases by $2i - 1$ in T'. Thus we get

$$
\begin{aligned}
2 \sum_{\substack{u \in A \\ v \in X_i}} \left(d_{T_{\max}}(u,v) - d_{T'}(u,v)\right) f(u) f(v) &= 2 \sum_{u \in A} f(u) \sum_{i=1}^{k-1} (2i-1) |X_i|_{\boldsymbol{f}} \\
&= 2 \sum_{i=1}^{k-1} (2i-1) |X_i|_{\boldsymbol{f}} |A|_{\boldsymbol{f}}.
\end{aligned}
$$

Case 2. If one of the two lies in Y_i for some i and the other lies in A, the distance increases by $2i - 1$. This gives us

$$
\begin{aligned}
2 \sum_{\substack{u \in A \\ v \in Y_i}} \left(d_{T_{\max}}(u,v) - d_{T'}(u,v)\right) f(u) f(v) &= -2 \sum_{u \in A} f(u) \sum_{i=1}^{k-1} (2i-1) |Y_i|_{\boldsymbol{f}} \\
&= -2 \sum_{i=1}^{k-1} (2i-1) |Y_i|_{\boldsymbol{f}} |A|_{\boldsymbol{f}}.
\end{aligned}
$$

Case 3. One of the two lies in $X_{>k-1}$, and the other lies in A. In this case, the distance between the vertices decreases by $2k - 1$. Thus we get

$$
\begin{aligned}
2 \sum_{\substack{u \in A \\ v \in X_{>k-1}}} \left(d_{T_{\max}}(u,v) - d_{T'}(u,v)\right) f(u) f(v) &= 2 \sum_{u \in A} f(u) (2k-1) |X_{>k-1}|_{\boldsymbol{f}} \\
&= 2(2k-1) |X_{>k-1}|_{\boldsymbol{f}} |A|_{\boldsymbol{f}}.
\end{aligned}
$$

Case 4. If one of the two lies in $Y_{>k-1} \backslash A$ and the other lies in A, the distance increases by $2k - 1$. This gives us

$$2 \sum_{\substack{u \in A \\ v \in Y_{>k-1} \backslash A}} \big(d_{T_{\max}}(u,v) - d_{T'}(u,v)\big) f(u) f(v)$$
$$= -2 \sum_{u \in A} f(u)(2k-1)\big(|Y_{>k-1}|_{\boldsymbol{f}} - |A|_{\boldsymbol{f}}\big)$$
$$= -2(2k-1)|A|_{\boldsymbol{f}}\big(|Y_{>k-1}|_{\boldsymbol{f}} - |A|_{\boldsymbol{f}}\big).$$

Combining the four cases, we find that $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f})$ is equal to

$$2|A|_{\boldsymbol{f}} \left(\sum_{i=1}^{k-1}(2i-1)(|X_i|_{\boldsymbol{f}} - |Y_i|_{\boldsymbol{f}}) + (2k-1)(|X_{>k-1}|_{\boldsymbol{f}} - |Y_{>k-1}|_{\boldsymbol{f}} + |A|_{\boldsymbol{f}}) \right)$$

if the length of the path $P_T(x,y)$ is odd, and similarly equal to

$$2|A|_{\boldsymbol{f}} \left(\sum_{i=1}^{k-1}(2i)(|X_i|_{\boldsymbol{f}} - |Y_i|_{\boldsymbol{f}}) + (2k)(|X_{>k-1}|_{\boldsymbol{f}} - |Y_{>k-1}|_{\boldsymbol{f}} + |A|_{\boldsymbol{f}}) \right)$$

if the length of the path is even.

We are assuming that $|X_i|_{\boldsymbol{f}} \geq |Y_i|_{\boldsymbol{f}}$ for $1 \leq i \leq k-1$ and $|X_{>k-1}|_{\boldsymbol{f}} \geq |Y_{>k-1}|_{\boldsymbol{f}}$, so $R_{T'}(\boldsymbol{f}) - R_{T_{\max}}(\boldsymbol{f}) > 0$. It follows that

$$\mu_c(T') \geq R_{T'}(\boldsymbol{f}) > R_{T_{\max}}(\boldsymbol{f}) = \mu_c(T_{\max}),$$

which contradicts the assumption that $T_{\max}$ maximizes the spectral radius of $M_c(T)$. $\blacksquare$

Lemma 11. *Let P be a path of an optimal tree $T_{\max}$ whose end vertices are leaves. If the length of P is odd $(2l-1)$, we label the vertices of P as $u_l u_{l-1} \cdots u_1 w_1 w_2 \cdots w_l$ and denote by U_i, W_i the components that contain u_i, w_i respectively $(i \in \{1,2,\ldots,l\})$, when the edges of P are removed. Assume (without loss of generality, reversing the labelling if necessary) that $|U_j|_{\boldsymbol{f}} > |W_j|_{\boldsymbol{f}}$ for the first index j for which $|U_j|_{\boldsymbol{f}} \neq |W_j|_{\boldsymbol{f}}$. Then we have*

$$|U_1|_{\boldsymbol{f}} \geq |W_1|_{\boldsymbol{f}} \geq |U_2|_{\boldsymbol{f}} \geq |W_2|_{\boldsymbol{f}} \geq \cdots \geq |U_l|_{\boldsymbol{f}} \geq |W_l|_{\boldsymbol{f}}.$$

If the length of P is even $(2l)$, we label the vertices as $u_{l+1} u_l u_{l-1} \cdots u_1 w_1 w_2 \cdots w_l$. Again, we have

$$|U_1|_{\boldsymbol{f}} \geq |W_1|_{\boldsymbol{f}} \geq |U_2|_{\boldsymbol{f}} \geq |W_2|_{\boldsymbol{f}} \geq \cdots \geq |W_l|_{\boldsymbol{f}} \geq |U_{l+1}|_{\boldsymbol{f}},$$

reversing the labelling if necessary.

In order to prove Lemma 11, we first find an equivalent representation for the Rayleigh quotient of $M_c(T)$ on its Perron vector $\boldsymbol{f}$.

Proposition 12. *For every tree T, we can write*

$$(3) \qquad \sum_{\{u,v\} \subseteq V(T)} d_T(u,v)f(u)f(v) = \sum_{uv \in E(T)} \sum_{u' \in C(u,v)} f(u') \sum_{v' \in C(v,u)} f(v'),$$

where $E(T)$ is the set of edges of T and $C(u,v)$ is the set of all vertices closer to the vertex u than to the vertex v.

Therefore the Rayleigh quotient of $M_c(T)$ on its Perron vector $\boldsymbol{f}$ can be written as

$$R_T(\boldsymbol{f}) = \sum_{u \in V(T)} \sum_{v \in V(T)} cf(u)f(v) - 2 \sum_{\{u,v\} \subseteq V(T)} d_T(u,v)f(u)f(v)$$

$$= c\left(\sum_{u \in V(T)} f(u)\right)^2 - 2 \sum_{uv \in E(T)} \sum_{u' \in C(u,v)} f(u') \sum_{v' \in C(v,u)} f(v').$$

Proof. We only have to prove equation (3); the second statement follows immediately. Firstly, the right hand side of the equation, when multiplied out, is equal to the sum over unordered pairs of vertices $\{u',v'\}$, each with its weight $f(u')f(v')$, such that the unique path between u' and v' contains the edge uv, where u' is closer to u and v' is closer to v (u' might coincide with u, and v' might coincide with v). Therefore, for each pair $\{u',v'\}$ the weight $f(u')f(v')$ occurs precisely $d(u',v')$ times in the sum, which means that the sums coincide. ■

For our next step, we need a general inequality on rearrangements, see the book of Hardy, Littlewood and Pólya [14]. For a sequence s of the form

$$s_{-l}, s_{-l+1}, \ldots, s_{-1}, s_0, s_1, \ldots, s_{l-1}, s_l,$$

we obtain the sequence s^+ by rearranging the elements so that

$$s_0^+ \geq s_1^+ \geq s_{-1}^+ \geq s_2^+ \geq s_{-2}^+ \geq \cdots \geq s_l^+ \geq s_{-l}^+.$$

Theorem 13 (see [14], Theorem 371). *Let $p, q,$ and r be non-negative sequences. Suppose the sequence r is symmetrically increasing, that is*

$$r_0 \leq r_1 = r_{-1} \leq r_2 = r_{-2} \leq \cdots \leq r_{2k} = r_{-2k},$$

and the sequences p and q have no prescribed order. Then the bilinear form

$$\sum_{a=-k}^{k} \sum_{b=-k}^{k} r_{a-b}p_a q_b$$

attains its minimum among all possible orders of p and q when p is p^+ and q is q^+.

Proof of Lemma **11.** We will only prove the statement for odd path length; the proof for even path length is similar. We set

$$S_{-l+1} = |U_l|_{\boldsymbol{f}}, S_{-l+2} = |U_{l-1}|_{\boldsymbol{f}}, \ldots, S_0 = |U_1|_{\boldsymbol{f}},$$
$$S_1 = |W_1|_{\boldsymbol{f}}, S_2 = |W_2|_{\boldsymbol{f}}, \ldots, S_l = |W_l|_{\boldsymbol{f}}.$$

We look at all possible permutations of the branches $U_1, \ldots, U_l, W_1, \ldots, W_l$ and show that a permutation for which the maximum value of the Rayleigh quotient $R_T(\boldsymbol{f})$ is attained must have the property stated in the lemma. We now consider the total contribution C of all edges of the path P to the Rayleigh quotient $R_T(\boldsymbol{f})$ as described in our Proposition 12, because the contribution of all other edges remains the same when the order of the branches $U_1, \ldots, U_l, W_1, \ldots, W_l$ is changed. We get

$$\begin{aligned}
C = {}& S_{-l+1}(S_{-l+2} + \cdots + S_l) \\
& + (S_{-l+1} + S_{-l+2})(S_{-l+3} + \cdots + S_l) \\
& \vdots \\
& + (S_{-l+1} + S_{-l+2} + \cdots + S_{l-1})S_l.
\end{aligned}$$

It can be seen that the summands of C are all of the form $S_a S_b$, it remains to count how often each such term occurs. To obtain a specific product $S_a S_b$, we must have S_a and S_b occurring in different parentheses, which happens $|a - b|$ times. Thus we define coefficients as follows:

$$r_{a-b} = r_{b-a} = \frac{|a - b|}{2}.$$

Moreover, we add $S_{-l} = 0$ as a dummy variable to achieve symmetry (this is not needed in the case where the path length is even). Then we have

$$C = \sum_{a=-l+1}^{l} \sum_{b=-l+1}^{l} r_{a-b} S_a S_b = \sum_{a=-l}^{l} \sum_{b=-l}^{l} r_{a-b} S_a S_b,$$

where r is symmetrically increasing, namely

$$r_0 = 0 \le r_1 = r_{-1} = \frac{1}{2} \le \cdots \le r_{2l} = r_{-2l} = l.$$

Therefore, by Theorem 13, C attains its minimum if S is S^+, so

$$S_0^+ \ge S_1^+ \ge S_{-1}^+ \ge S_2^+ \ge S_{-2}^+ \ge \cdots \ge S_{-l+1}^+ \ge S_l^+ \ge S_{-l}^+ = 0,$$

or in other words

$$(4) \qquad |U_1|_{\boldsymbol{f}} \ge |W_1|_{\boldsymbol{f}} \ge |U_2|_{\boldsymbol{f}} \ge |W_2|_{\boldsymbol{f}} \ge \cdots \ge |U_l|_{\boldsymbol{f}} \ge |W_l|_{\boldsymbol{f}}.$$

Note that the right hand side of (3) attains its minimum under permutations of the branches $U_1, \ldots, U_l, W_1, \ldots, W_l$ when C does. Equivalently, the Rayleigh quotient $R_T(\boldsymbol{f})$ attains its maximum value by Proposition 12. Therefore, the optimal tree has to satisfy (4). ∎

Lemma 14. *Let $T_{\max}$ be an optimal tree. For a path P with labelling as described in Lemma* 11, *we have*

$$d(u_1) \geq d(w_1) \geq d(u_2) \geq d(w_2) \geq \cdots \geq d(u_l) = d(w_l) = 1$$

if the path length is odd and

$$d(u_1) \geq d(w_1) \geq d(u_2) \geq d(w_2) \geq \cdots \geq d(u_l) \geq d(w_l) = d(u_{l+1}) = 1$$

if the path length is even.

Proof. Again we will show the proof for odd path length, the other case is treated in a similar manner. From Lemma 11 we have

$$|U_1|_{\boldsymbol{f}} \geq |W_1|_{\boldsymbol{f}} \geq |U_2|_{\boldsymbol{f}} \geq |W_2|_{\boldsymbol{f}} \geq \cdots \geq |U_l|_{\boldsymbol{f}} \geq |W_l|_{\boldsymbol{f}}.$$

We will now apply Lemma 10, each time with an appropriate choice of the vertices y_i and x_i.

Step 1. Let $x_i = u_i$ and $y_i = w_i$ for $i = 1, 2, \ldots, l$; then we obtain $d(u_i) \geq d(w_i)$ by Lemma 10.

Step 2. Let $y_1 = u_{i+1}, y_2 = u_{i+2}, \ldots$ and $x_1 = u_i, x_2 = u_{i-1}, \ldots, x_i = u_1, x_{i+1} = w_1, \ldots$ Then we get

$$|Y_{>1}|_{\boldsymbol{f}} = \sum_{k=i+2}^{l} |U_k|_{\boldsymbol{f}} \quad \text{and} \quad |X_{>1}|_{\boldsymbol{f}} = \sum_{k=1}^{l} |W_k|_{\boldsymbol{f}} + \sum_{k=1}^{i-1} |U_k|_{\boldsymbol{f}},$$

implying $|X_{>1}|_{\boldsymbol{f}} > |Y_{>1}|_{\boldsymbol{f}}$, so by Lemma 10 we get $d(x_1) = d(u_i) \geq d(y_1) = d(u_{i+1})$. That is, we have

$$d(u_1) \geq d(u_2) \geq d(u_3) \geq \cdots \geq d(u_l).$$

Step 3. Let $y_1 = w_i, y_2 = w_{i-1}, \ldots, y_{i+1} = u_1, \ldots$ and $x_1 = w_{i+1}, x_2 = w_{i+2}, \ldots$ Then we get

$$|X_{>1}|_{\boldsymbol{f}} = \sum_{k=i+2}^{l} |W_k|_{\boldsymbol{f}} \quad \text{and} \quad |Y_{>1}|_{\boldsymbol{f}} = \sum_{k=1}^{l} |U_k|_{\boldsymbol{f}} + \sum_{k=1}^{i-1} |W_k|_{\boldsymbol{f}},$$

implying $|X_{>1}|_{\boldsymbol{f}} < |Y_{>1}|_{\boldsymbol{f}}$, so by Lemma 10 we get $d(y_1) = d(w_i) \geq d(x_1) = d(w_{i+1})$. That is, we have

$$d(w_1) \geq d(w_2) \geq d(w_3) \geq \cdots \geq d(w_l).$$

Step 4. Let $z = u_1$ and $y_i = u_{i+1}, x_i = w_i$ for $i = 1, 2, \ldots, l$. Since

$$|X_{>i}|_{\boldsymbol{f}} = \sum_{k=i+1}^{l} |W_k|_{\boldsymbol{f}} \quad \text{and} \quad |Y_{>i}|_{\boldsymbol{f}} = \sum_{k=i+2}^{l} |U_k|_{\boldsymbol{f}},$$

we have $|X_{>i}|_{\boldsymbol{f}} > |Y_{>i}|_{\boldsymbol{f}}$ and thus $d(x_i) = d(w_i) \geq d(u_{i+1}) = d(y_i)$ by Lemma 10. Combining all these results, we get

$$d(u_1) \geq d(w_1) \geq d(u_2) \geq d(w_2) \geq \cdots \geq d(u_l) = d(w_l) = 1. \qquad \blacksquare$$

Definition. We say that a function g defined on the vertices of a path $x_1 x_2 \cdots x_l$ is concave if $g(x_{k+1}) - g(x_k) < g(x_k) - g(x_{k-1})$ for all k.

Lemma 15. *Let T be a tree, and let $\boldsymbol{f}$ be the Perron vector corresponding to $M_c(T)$. For a vertex u, we set*

$$N_{\boldsymbol{f}}(u) = \sum_{v \in V(T)} (c - d_T(u, v)) f(v),$$

which is the coordinate corresponding to u in $M_c(T)\boldsymbol{f}$, thus equal to $\mu_c(M)f(u)$. This expression is concave along paths. It follows that the maximum of $N_{\boldsymbol{f}}$ is attained either at one or two adjacent points (vertices) in the tree T.

Proof. Let $P_1 = x_1 x_2 \cdots x_l$ be a path in a tree T and let $A_{\leq i}$ and $B_{\geq i+1}$ be the components of $T \setminus x_i x_{i+1}$ that contain the vertices $x_1 \cdots x_i$ and $x_{i+1} \cdots x_l$, respectively. We have

$$N_{\boldsymbol{f}}(x_{i+1}) - N_{\boldsymbol{f}}(x_i) = \sum_{v \in V(T)} \big(d_T(x_i, v) - d_T(x_{i+1}, v) \big) f(v)$$

$$= \sum_{v \in A_{\leq i}} (-1) f(v) + \sum_{v \in B_{\geq i+1}} f(v) = -|A_{\leq i}|_{\boldsymbol{f}} + |B_{\geq i+1}|_{\boldsymbol{f}}.$$

It is easy to see that when i increases, $|A_{\leq i}|_{\boldsymbol{f}}$ increases and $|B_{\geq i+1}|_{\boldsymbol{f}}$ decreases and hence $N_{\boldsymbol{f}}(x_{i+1}) - N_{\boldsymbol{f}}(x_i)$ decreases. So $N_{\boldsymbol{f}}$ is concave, thus reaches its maximum either at one vertex x_k, or two adjacent vertices x_k, x_{k+1} along the path P_1.

Now, let y, z be any two distinct vertices in the tree T for which $N_{\boldsymbol{f}}(y)$ and $N_{\boldsymbol{f}}(z)$ are both equal to the maximum of $N_{\boldsymbol{f}}$ in the tree. We know that $N_{\boldsymbol{f}}$ is concave along the path that passes through the two vertices y and z, so they have to be adjacent. It is important to note that there cannot be more than two vertices where the maximum of $N_{\boldsymbol{f}}$ is attained since they would form a triangle, which is impossible in a tree. Therefore we obtain that the maximum of $N_{\boldsymbol{f}}$ is attained either at one or two adjacent vertices in the tree T. $\qquad \blacksquare$

Now, using all the auxiliary results, we prove Theorem 8 by showing that the optimal tree satisfies all the properties of the greedy tree as in Lemma 9. Again, the steps are essentially following [27].

***Proof of Theorem* 8.** Consider a path as described in Lemma 11. We have

$$
N_{\boldsymbol{f}}(u_1) - N_{\boldsymbol{f}}(u_2) = \sum_{v \in V(T)} \big(d_T(u_2, v) - d_T(u_1, v)\big)f(v)
$$
$$
= \sum_{v \in U_{>1}} (-1)f(v) + \sum_{v \in W_{\geq 1} \cup U_1} f(v) = -|U_{>1}|_{\boldsymbol{f}} + |W_{\geq 1} \cup U_1|_{\boldsymbol{f}},
$$

which is greater than 0 by Lemma 11. It follows that $N_{\boldsymbol{f}}(u_1) > N_{\boldsymbol{f}}(u_2)$. Also,

$$
N_{\boldsymbol{f}}(u_1) - N_{\boldsymbol{f}}(w_1) = \sum_{v \in V(T)} \big(d_T(w_1, v) - d_T(u_1, v)\big)f(v)
$$
$$
= \sum_{v \in U_{\geq 1}} f(v) + \sum_{v \in W_{\geq 1}} (-1)f(v) = |U_{\geq 1}|_{\boldsymbol{f}} - |W_{\geq 1}|_{\boldsymbol{f}},
$$

which is greater than or equal to 0 by Lemma 11. Thus $N_{\boldsymbol{f}}(u_1) \geq N_{\boldsymbol{f}}(w_1)$.

From Lemmas 11 and 14, we know that for an optimal tree labelled as in those lemmas, $d(u_1)$ and $|U_1|_{\boldsymbol{f}}$ are the respective maxima among all vertices along the path. We also know that the function $N_{\boldsymbol{f}}$ is concave along the path, $N_{\boldsymbol{f}}(u_1) > N_{\boldsymbol{f}}(u_2)$ and $N_{\boldsymbol{f}}(u_1) \geq N_{\boldsymbol{f}}(w_1)$, so $N_{\boldsymbol{f}}(u_1)$ is the maximum along the path.

We recall from Lemma 15 that $N_{\boldsymbol{f}}$ is maximal at either only one vertex, which we then label r (this vertex will become the root), or two adjacent vertices in the tree $T_{\max}$, which we then label r and v_1. We will present the proof for the case when $N_{\boldsymbol{f}}$ is maximal at only one vertex r, and otherwise it follows in a similar manner.

Now we consider the optimal tree $T_{\max}$ to be a rooted tree with root r, so since $N_{\boldsymbol{f}}(r) > N_{\boldsymbol{f}}(z)$ for every vertex z in $T_{\max}$ by assumption, r must be the vertex u_1 in any path containing it, and thus have the largest degree by Lemma 14. Thus (i) in Lemma 9 is satisfied.

Suppose we have a path from a leaf u passing through r to another leaf v in $T_{\max}$ such that the only ancestor of u and v is r. By Lemma 11 we have $|d_T(u, r) - d_T(v, r)| = 0$ if the path length is even and $|d_T(u, r) - d_T(v, r)| = 1$ if the path length is odd. This remains true if the first common ancestor of u and v is a vertex s other than the root: since N_f is concave along paths with its maximum at r, it must be decreasing along the paths from r to u through s and from r to v through s. Thus the maximum of N_f on the path from u to v occurs at s, and Lemma 11 applies again. Hence the heights of any two leaves differ by at most 1, which yields (ii) in Lemma 9.

Let $h_{T_{\max}}(y)$ and $h_{T_{\max}}(z)$ be the heights of two vertices y and z for which $h_{T_{\max}}(y) < h_{T_{\max}}(z)$. These heights are with respect to the root vertex r of the optimal tree. If z is a successor of y and we consider the path from a leaf u to a leaf v passing through y, z, then we get $d(y) \geq d(z)$ by Lemma 14. Now, suppose z is not a successor of y. We consider a path from a leaf y' to a leaf z' passing through the vertices y, s, z, where s is the first common ancestor of y, z.

Since $N_{\boldsymbol{f}}$ is concave along paths and attains its maximum at r by our choice of r its values are decreasing on the paths from r to y' and to z', respectively. In particular, it decreases on the paths from s to y' and z', which means that the maximum of $N_{\boldsymbol{f}}$ on the path from y' to z' is attained at s. Thus we can set $s = u_1$ in Lemma 11 when the path from y' to z' is considered. Let $i = h_{T_{\max}}(y) - h_{T_{\max}}(s)$ and $j = h_{T_{\max}}(z) - h_{T_{\max}}(s)$ be the heights of y and z with respect to the vertex s. We get two possible labellings of the path, namely either $y = u_{i+1}, z = w_j$ or $y = w_i, z = u_{j+1}$. If we consider the first labelling then we get $i + 1 \leq j$, which by Lemma 14 implies that $d(y) \geq d(z)$. Note that this is also true for the second labelling. Thus (iii) in Lemma 9 is also satisfied.

Next, let y, z be on the same level, thus $h_{T_{\max}}(y) = h_{T_{\max}}(z)$, such that $d(y) > d(z)$, and let y', z' be their respective successors such that $h_{T_{\max}}(y') = h_{T_{\max}}(z')$. Consider the longest path that passes through y', y, s, z, z', where s is the first ancestor of y and z. Again we can set $s = u_1$ in Lemma 11 with $y = w_i, y' = w_j, z = u_{i+1}, z' = u_{j+1}$ where $i = h_{T_{\max}}(y) - h_{T_{\max}}(s), j = h_{T_{\max}}(y') - h_{T_{\max}}(s)$. Then by Lemma 14, we see that $d(y) > d(z)$ implies $d(y') \geq d(z')$. This yields (iv) in Lemma 9.

Finally, let y_1, z_1 be the respective parents of y, z, let y', z' be their respective siblings and y'', z'' successors of y', z' at the same level. We let Y denote the subtree induced by the vertices $v \in V(T_{\max})$ for which $P_{T_{\max}}(v, r)$ contains y and $h_{T_{\max}}(v) \geq h_{T_{\max}}(y); Y_1, Z, Z_1$ are defined in an analogous way. Assuming again that $d(y) > d(z)$, we consider a path from one leaf to another that passes through y, z and their common ancestor s. Applying Lemma 11 and Lemma 14 (where s becomes u_1) to this path now shows that

$$|Y|_{\boldsymbol{f}} \geq |Z|_{\boldsymbol{f}} \quad \text{and} \quad |Y_1 \setminus Y|_{\boldsymbol{f}} \geq |Z_1 \setminus Z|_{\boldsymbol{f}},$$

thus $|Y_1|_{\boldsymbol{f}} \geq |Z_1|_{\boldsymbol{f}}$. Applying Lemmas 11 and 14 once again, now to a path from one leaf to another that passes through $y'', y', y_1, s, z_1, z', z''$ in this order, we find that $d(y') \geq d(z')$ and $d(y'') \geq d(z'')$. Thus (v) in Lemma 9 is also satisfied. In conclusion, we have proven the optimal tree $T_{\max}$ to be the greedy tree. ∎

In the following we compare the greedy trees associated with different degree sequences, paralleling results on other graph invariants (such as the Wiener index, see [33]). We first introduce the concept of majorization for this purpose.

Definition. Let $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$ be sequences of non-negative numbers with $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$. If for all $1 \leq k \leq n$ we have

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i$$

then we say that Y majorizes X. Let S_n be the set of all permutations of $\{1, \ldots, n\}$. If for any $\sigma \in S_n$ the sequence Y still majorizes $(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$, then we write $X \preceq Y$.

The set S_n consists of $n!$ permutations. However, one does not need to check all of them in order to show that $X \preceq Y$. This is because, if we let $\sigma \in S_n$ be a permutation for which $x_{\sigma(1)} \geq \cdots \geq x_{\sigma(n)}$, then we can see that $(x_{\sigma'(1)}, \ldots, x_{\sigma'(n)})$ is majorized by $(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ for all other $\sigma' \in S_n$. So proving that $X \preceq Y$ is equivalent to showing that $(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ is majorized by Y.

Remark 16. It is important to note that majorization is transitive. That is, if $X \preceq Y$ and $Y \preceq Z$, then we also have $X \preceq Z$.

The following lemma on majorizing degree sequences has been used repeatedly in similar contexts, see for instance [33]. We present its proof for completeness.

Lemma 17. *Let $X = (x_1, \ldots, x_n)$ and $X' = (x'_1, \ldots, x'_n)$ be two non-increasing tree degree sequences. If $X \preceq X'$, then there exists a sequence of k tree degree sequences $X_1, \ldots, X_k$ such that $X \preceq X_1 \preceq \cdots \preceq X_k \preceq X'$, where exactly two entries of X_i and X_{i+1} differ by 1. Namely, for two indices $p > q$, the entries x_p, x_q of X_i and the entries x_p^*, x_q^* of X_{i+1} satisfy $x_p^* = x_p - 1, x_q^* = x_q + 1$.*

Proof. Suppose $X \preceq X'$ and $X \neq X'$. Now, we let p be the first index for which $x_p > x'_p$ and let q be the last index less than p for which $x_q < x'_q$, so that $x_r = x'_r$ for all $q < r < p$. Note that by the definition of majorization such positions exist. Now we construct the sequence X_k from X' by replacing x'_q by $x'_q - 1$ and x'_p by $x'_p + 1$. So we get $X_k = (x'_1, \ldots, x'_q - 1, \ldots, x'_p + 1, \ldots, x'_n)$, which is still a valid degree sequence of a tree. We have $X \preceq X_k \preceq X'$. Now, we can apply the same procedure to X and X_k to obtain X_{k-1}. Repeated application of this process yields the result. ∎

Lemma 18. *Let $G(\alpha)$ be the greedy tree for the degree sequence α. Let x and y be two vertices of $G(\alpha)$ satisfying the inequality $d_{G(\alpha)}(x) \geq d_{G(\alpha)}(y) \geq 2$. If $d_{G(\alpha)}(x) = d_{G(\alpha)}(y)$, we assume that y comes after x in the ordering of vertices in the greedy tree. Let s be a successor of y and let T' be the tree obtained from $G(\alpha)$ by deleting the edge ys and adding the edge xs. Then*

$$\mu_c(G(\alpha)) < \mu_c(T').$$

Proof. We consider a longest path that contains the vertices x, y, z such that z is the common ancestor of x and y. We will present the proof for odd path length, the other case is similar. We can set $z = u_1$ in the labelling described in Lemma 11. In view of our assumption that $d(x) \geq d(y)$ (and that x comes first in the ordering in case of equality), we can set $x = u_p$, $y = w_q$, where $p \leq q$, or $x = w_p$, $y = u_q$, where $p < q$ (depending on whether the part of the path from z to a leaf through x or the part from z to a leaf through y is the longer one). We only consider the first possibility, the calculations in the second case are similar.

Let S be the subtree induced by the vertex s and its successors, and let $\boldsymbol{f}$ be the Perron vector of $G(\alpha)$. It can be observed that the distance between two vertices v and w will only change from $G(\alpha)$ to T' if $v \in S$ and $w \in G(\alpha) \setminus S$ (or vice versa). We compute $R_{T'}(\boldsymbol{f}) - R_{G(\alpha)}(\boldsymbol{f})$ by distinguishing the following cases.

Case 1. The contribution to the difference between $R_{G(\alpha)}(\boldsymbol{f})$ and $R_{T'}(\boldsymbol{f})$ of vertex pairs where one vertex is in S and the other in some U_i is

$$
2 \sum_{i=1}^{l} \sum_{\substack{v \in S \\ w \in U_i}} \big((c - d_{T'}(v, w)) - (c - d_{G(\alpha)}(v, w)) \big) f(v) f(w)
$$

$$
= 2 \sum_{i=1}^{l} \sum_{\substack{v \in S \\ w \in U_i}} \big(d_{G(\alpha)}(v, w) - d_{T'}(v, w) \big) f(v) f(w)
$$

$$
= 2 \sum_{i=1}^{l} \sum_{\substack{v \in S \\ w \in U_i}} \Big(\big(d_{G(\alpha)}(s, v) + i + q + d_{G(\alpha)}(u_i, w) \big)
$$

$$
- \big(d_{G(\alpha)}(s, v) + |i - p| + 1 + d_{G(\alpha)}(u_i, w) \big) \Big) f(v) f(w)
$$

$$
= 2 \sum_{i=1}^{l} |S|_{\boldsymbol{f}} \sum_{w \in U_i} \big(i + q - |i - p| - 1 \big) f(w)
$$

$$
= 2 \sum_{i=1}^{p} (2i - 1 + q - p) |S|_{\boldsymbol{f}} \sum_{w \in U_i} f(w) + 2 \sum_{i=p+1}^{l} (p + q - 1) |S|_{\boldsymbol{f}} \sum_{w \in U_i} f(w)
$$

$$
= 2 \sum_{i=1}^{p} (2i - 1 + q - p) |U_i|_{\boldsymbol{f}} |S|_{\boldsymbol{f}} + 2 \sum_{i=p+1}^{l} (p + q - 1) |U_i|_{\boldsymbol{f}} |S|_{\boldsymbol{f}}.
$$

Case 2. The contribution to the difference between $R_{G(\alpha)}(\boldsymbol{f})$ and $R_{T'}(\boldsymbol{f})$ of vertex pairs where one vertex is in S and the other in some W_i (but not in S if $i = q$) is

$$2\sum_{i=1}^{l}\sum_{\substack{v\in S\\ w\in W_i\setminus S}}\big(d_{G(\alpha)}(v,w)-d_{T'}(v,w)\big)f(v)f(w)$$

$$=2\sum_{i=1}^{q-1}(-2i+1+q-p)|W_i|_{\boldsymbol{f}}|S|_{\boldsymbol{f}}$$

$$-2(p+q-1)(|W_q|_{\boldsymbol{f}}-|S|_{\boldsymbol{f}})|S|_{\boldsymbol{f}}-2\sum_{i=q+1}^{l}(p+q-1)|W_i|_{\boldsymbol{f}}|S|_{\boldsymbol{f}}.$$

Summing the two cases, we get

$$R_{T'}(\boldsymbol{f})-R_{G(\alpha)}(\boldsymbol{f})$$

$$=2\sum_{i=1}^{p}(2i-1)(|U_i|_{\boldsymbol{f}}-|W_i)|_{\boldsymbol{f}})|S|_{\boldsymbol{f}}+2\sum_{i=1}^{p}(q-p)(|W_i|_{\boldsymbol{f}}+|U_i|_{\boldsymbol{f}})|S|_{\boldsymbol{f}}$$

$$+2\sum_{i=p+1}^{q-1}(p+q-1)(|U_i|_{\boldsymbol{f}}-|W_i|_{\boldsymbol{f}})|S|_{\boldsymbol{f}}+2\sum_{i=p+1}^{q-1}(2q-2i)|W_i|_{\boldsymbol{f}}|S|_{\boldsymbol{f}}$$

$$+2\sum_{i=q+1}^{l}(p+q-1)(|U_i|_{\boldsymbol{f}}-|W_i|_{\boldsymbol{f}})|S|_{\boldsymbol{f}}$$

$$+2(p+q-1)(|U_q|_{\boldsymbol{f}}-|W_q|_{\boldsymbol{f}}+|S|_{\boldsymbol{f}})|S|_{\boldsymbol{f}}.$$

By Lemma 18, we have $|U_i|_{\boldsymbol{f}}\geq|W_i|_{\boldsymbol{f}}$ for all i, and we also note that $q\geq p$ and $2q\geq 2i$ for $i\leq q-1$. Therefore, each term in the summation is non-negative. Also, we know that $|S|_{\boldsymbol{f}}>0$, so $R_{T'}(\boldsymbol{f})-R_{G(\alpha)}(\boldsymbol{f})>0$ because of the term $2(p+q-1)|S|_{\boldsymbol{f}}^2$. Therefore,

$$\mu_c(G(\alpha))=R_{G(\alpha)}(\boldsymbol{f})<R_{T'}(\boldsymbol{f})\leq\mu_c(T').\qquad\blacksquare$$

Theorem 19. *Let X and Y be two tree degree sequences of the same length. Suppose $X\preceq Y$; then*

$$\mu_c(G(X))\leq\mu_c(G(Y)),$$

with equality only if $X=Y$.

Proof. In view of Lemma 17, it is sufficient to prove the statement in the case that the degree sequences differ in exactly two places, with a difference of 1 in each place. The general case follows inductively. So take two non-increasing sequences $X=(x_1,x_2,\ldots,x_i,\ldots,x_j,\ldots,x_n)$ and $Y=(x_1,x_2,\ldots,x_i+1,\ldots,x_j-1,\ldots,x_n)$. Let v,w be two vertices in the greedy tree $G(X)$ such that $d_{G(X)}(v)=x_i$ and

$d_{G(X)}(w) = x_j$. We have $x_i \geq x_j \geq 2$. If $x_i = x_j$, we can assume (without loss of generality) that w comes after v in the ordering of vertices in the greedy tree. If $x_i > x_j$, this follows from the definition of a greedy tree. Let s be a neighbor of w. If we remove the edge ws and add the edge vs to $G(X)$ then we obtain a new tree T' with degree sequence Y. From Lemma 18, we obtain $\mu_c(T') > \mu_c(G(X))$. Also, by Theorem 8, we have $\mu_c(T') \leq \mu_c(G(Y))$. So

$$\mu_c(G(Y)) \geq \mu_c(T') > \mu_c(G(X)).$$

This completes the proof. ∎

We obtain the following corollaries using Theorems 8 and 19.

Corollary 20. *Among trees of order n with maximum degree Δ, the Volkmann tree $V_{n,\Delta}$ maximizes the spectral radius of $M_c(T)$ for every constant $c \geq n - \Delta + 1$.*

Proof. It is easy to see that the degree sequence $(\Delta, \ldots, \Delta, r, 1, \ldots, 1)$ of the Volkmann tree majorizes the degree sequence of every other tree with the same number of vertices whose maximum degree is Δ. Therefore, the statement follows from Theorem 8, combined with Theorem 19. Note here that the maximum number of internal nodes in a tree with n vertices and maximum degree Δ is $n - \Delta$, so the diameter is at most $n - \Delta + 1$. ∎

Corollary 21. *Among all trees of order n with k leaves, the greedy tree corresponding to the degree sequence $\alpha = (k, 2, \ldots, 2, 1, \ldots, 1)$ maximizes the spectral radius of $M_c(T)$ for every constant $c \geq n - k + 1$.*

Proof. This corollary follows in the same way as the previous one, by observing that the sequence $(k, 2, \ldots, 2, 1, \ldots, 1)$ majorizes every other possible degree sequence of a tree with the same number of vertices and the same number of leaves. ∎

One might ask whether a result analogous to Theorem 8 also holds when we replace the distance matrix $D(T)$ by the terminal distance matrix $TD(T)$. For comparison, it is known that the greedy tree has minimum terminal Wiener index (sum of distances between leaves), see [26]. Goubko [12] considered the problem of minimizing the terminal distance spectral radius for trees with given degree sequence and provided evidence that the greedy tree is always extremal. However, this is not always the case for the maximum spectral radius of $M_c(T)$, as the following examples show.

Example 22. Let us consider the matrix $M_4(T) = 4J - TD(T)$. Figure 7 shows the respective extremal trees for the degree sequences

$$\alpha_1 = (3, 2, 2, 1, 1, 1),$$
$$\alpha_2 = (3, 3, 2, 1, 1, 1, 1).$$

(a) T_1.   (b) T_2.
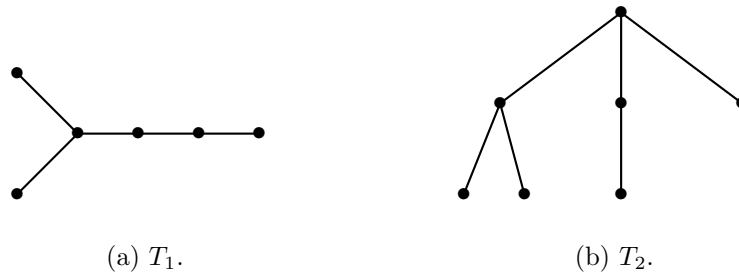
Figure 7. Extremal trees T_1 and T_2.

It can be observed that the extremal tree T_2 is the greedy tree, but T_1 is not. Hence the greedy tree is extremal for some, but not all possible tree degree sequences.

References

[1] A.T. Balaban, D. Mills, O. Ivanciuc and S.C. Basak, *Reverse Wiener indices*, Croat. Chem. Acta **73** (2000) 923–941.

[2] R.B. Bapat, *Determinant of the distance matrix of a tree with matrix weights*, Linear Algebra Appl. **416** (2006) 2–7.
doi:10.1016/j.laa.2005.02.022

[3] R.B. Bapat, Graphs and Matrices (Springer, 2010).
doi:10.1007/978-1-84882-981-7

[4] R.B. Bapat, S.J. Kirkland and M. Neumann, *On distance matrices and Laplacians*, Linear Algebra Appl. **401** (2005) 193–209.
doi:10.1016/j.laa.2004.05.011

[5] T. Bıyıkoğlu and J. Leydold, *Graphs with given degree sequence and maximal spectral radius*, Electron. J. Combin. **15** (2008) #R119.

[6] S.S. Bose, M. Nath and S. Paul, *Distance spectral radius of graphs with r pendent vertices*, Linear Algebra Appl. **435** (2011) 2828–2836.
doi:10.1016/j.laa.2011.04.041

[7] S.S. Bose, M. Nath and S. Paul, *On the maximal distance spectral radius of graphs without a pendent vertex*, Linear Algebra Appl. **438** (2013) 4260–4278.
doi:10.1016/j.laa.2013.01.019

[8] S.S. Bose, M. Nath and D. Sarma, *Maximal distance spectral radius of trees*, Discrete Math. Algorithms Appl. **11** (2019) 1950025.
doi:10.1142/S1793830919500253

[9] E. Çela, N.S. Schmuck, S. Wimer and G.J. Woeginger, *The Wiener maximum quadratic assignment problem*, Discrete Optim. **8** (2011) 411–416.
doi:10.1016/j.disopt.2011.02.002

[10] K.L. Collins, Distance Matrices of Trees (Ph.D. Thesis, Massachusetts Institute of Technology, 1986).

[11] M. Fischermann, A. Hoffmann, D. Rautenbach, L. Székely and L. Volkmann, *Wiener index versus maximum degree in trees*, Discrete Appl. Math. **122** (2002) 127–137.
doi:10.1016/S0166-218X(01)00357-2

[12] M. Goubko, *On minimum terminal distance spectral radius of trees with given degree sequence*, (2015).
arXiv:1507.01733

[13] R.L. Graham and H.O. Pollak, *On the addressing problem for loop switching*, Bell System Tech. J. **50** (1971) 2495–2519.
doi:10.1002/j.1538-7305.1971.tb02618.x

[14] G.H. Hardy, J.E. Littlewood and G. Pólya, Inequalities (Cambridge University Press, 1952).

[15] A. Heydari, *On extremal trees with respect to their terminal distance spectral radius*, Australas. J. Combin. **69** (2017) 159–168.

[16] A. Ilić, *Distance spectral radius of trees with given matching number*, Discrete Appl. Math. **158** (2010) 1799–1806.
doi:10.1016/j.dam.2010.06.018

[17] G. Indulal, *Sharp bounds on the distance spectral radius and the distance energy of graphs*, Linear Algebra Appl. **430** (2009) 106–113.
doi:10.1016/j.laa.2008.07.005

[18] H. Lin and B. Zhou, *The distance spectral radius of graphs with given number of odd vertices*, Electron. J. Linear Algebra **31** (2016) 286–305.
doi:10.13001/1081-3810.2877

[19] H. Lin and B. Zhou, *The distance spectral radius of trees*, Linear Multilinear Algebra **67** (2018) 370–390.
doi:10.1080/03081087.2017.1418830

[20] W. Lin, Y. Zhang, Q. Chen, J. Chen, C. Ma and J. Chen, *Ordering trees by their distance spectral radii*, Discrete Appl. Math. **203** (2016) 106–110.
doi:10.1016/j.dam.2015.09.009

[21] Z. Luo and B. Zhou, *On distance spectral radius of trees and fixed maximum degree*, Filomat **29** (2015) 2021–2026.
doi:10.2298/FIL1509021L

[22] M. Nath and S. Paul, *On the distance spectral radius of trees*, Linear Multilinear Algebra **61** (2013) 847–855.
doi:10.1080/03081087.2012.711324

[23] W. Ning, L. Ouyang and M. Lu, *Distance spectral radius of trees with fixed number of pendent vertices*, Linear Algebra Appl. **439** (2013) 2240–2249.
doi:10.1016/j.laa.2013.06.030

[24] N.S. Schmuck, S.G. Wagner and H. Wang, *Greedy trees, caterpillars, and Wiener-type graph invariants*, MATCH Commun. Math. Comput. Chem. **68** (2012) 273–292.

[25] D. Stevanović and A. Ilić, *Distance spectral radius of trees with fixed maximum degree*, Electron. J. Linear Algebra **20** (2010) 168–179.
doi:10.13001/1081-3810.1366

[26] L.A. Székely, H. Wang and T. Wu, *The sum of the distances between the leaves of a tree and the semi-regular property*, Discrete Math. **311** (2011) 1197–1203.
doi:10.1016/j.disc.2010.06.005

[27] H. Wang, *The extremal values of the Wiener index of a tree with given degree sequence*, Discrete Appl. Math. **156** (2008) 2647–2654.
doi:10.1016/j.dam.2007.11.005

[28] Y. Wang, R. Xing, B. Zhou and F. Dong, *A note on distance spectral radius of trees*, Spec. Matrices **5** (2017) 296–300.
doi:10.1515/spma-2017-0021

[29] Y. Wang and B. Zhou, *On distance spectral radius of graphs*, Linear Algebra Appl. **438** (2013) 3490–3503.
doi:10.1016/j.laa.2012.12.024

[30] R. Xing, B. Zhou and F. Dong, *The effect of a graft transformation on distance spectral radius*, Linear Algebra Appl. **457** (2014) 261–275.
doi:10.1016/j.laa.2014.05.024

[31] G. Yu, S. Guo and M. Zhai, *Distance spectral radius of a tree with given diameter*, Ars Combin. **134** (2017) 351–362.

[32] X.-D. Zhang, *The Laplacian spectral radii of trees with given degree sequences*, Discrete Math. **308** (2008) 3143–3150.
doi:10.1016/j.disc.2007.06.017

[33] X.-D. Zhang, Q.-Y. Xiang, L.-Q. Xu and R.-Y. Pan, *The Wiener index of trees with given degree sequences*, MATCH Commun. Math. Comput. Chem. **60** (2008) 623–644.