Finding all the extremal trees corresponding to a given degree sequence for certain functions on adjacent vertex degrees

Ivan Damnjanović*
Faculty of Electronic Engineering, University of Niš, Niš, Serbia
ivan.damnjanovic@elfak.ni.ac.rs
Diffine LLC, San Diego, California, U.S.A.
ivan@diffine.com

Abstract

Wang [Cent. Eur. J. Math. 12:1656–1663, 2014] has considered any discrete symmetric function \( F : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R} \) such that
\[
F(x, a) + F(y, b) \geq F(y, a) + F(x, b)
\]
for any \( x \geq y \) and \( a \geq b \), and showed that, among all the trees \( T_D \) corresponding to a given degree sequence \( D \), the greedy tree must attain the maximum value of the graph invariant given by the expression
\[
\sum_{u \sim v} F(\deg(u), \deg(v)),
\]
where \( \deg(u) \) and \( \deg(v) \) represent the degrees of the vertices \( u \) and \( v \), respectively, and the summing is performed across all the unordered pairs of adjacent vertices \( u \) and \( v \). In this paper, we extend Wang’s result by determining all the possible trees from \( T_D \) that attain the maximum value of the aforementioned graph invariant, thereby fully resolving the corresponding extremal problem.

MSC: 05C35, 05C05, 05C07, 05C92.

*Supported by Diffine LLC.
Table 1: Some topological indices $R_F$ with their corresponding functions $F(x, y)$.

| $R_F$                                      | $F(x, y)$                      |
|--------------------------------------------|--------------------------------|
| Randić index [1]                           | $\frac{1}{\sqrt{xy}}$        |
| second Zagreb index [2]                    | $xy$                           |
| second modified Zagreb index [3]           | $\frac{1}{xy}$                |
| sum–connectivity index [4]                 | $\frac{1}{\sqrt{\frac{x^2+y^2}{2}}}$ |
| harmonic index [5]                         | $x+y$                          |
| Sombor index [6]                           | $\sqrt{x^2+y^2}$               |

In this paper we will deal exclusively with standard trees. For this reason, we will consider all graphs to be undirected, finite and non-null, in addition to being connected and acyclic. Thus, each graph will be connected and have at least one vertex, and there shall be no loops, multiple edges or cycles in general.

For some $n \in \mathbb{N}$, let $D \in \mathbb{N}_0^n$ be any non-increasing sequence of $n$ non-negative integers. We shall use $T_D$ to denote the set of all the trees that correspond to the degree sequence $D$. In other words, this set will contain precisely the trees whose vertices can be ordered in such a way that their corresponding degrees yield the sequence $D$. Throughout the paper, we will also rely on $\text{deg}_T(u)$ in order to signify the degree of some vertex $u$ from a given tree of interest $T$.

Now, let $F: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ be an arbitrary discrete symmetric real function. By using

$$R_F(T) = \sum_{u \sim v} F(\text{deg}_T(u), \text{deg}_T(v)),$$

where the summing is performed across all the unordered pairs of adjacent vertices $u$ and $v$ from some tree $T$, it becomes possible to define various adjacent vertex degree based topological indices capable of describing the structure of the given tree in some manner. A brief preview of some of these indices is given in Table 1.

Wang [7] has considered the aforementioned topological indices that correspond to the functions $F$ which satisfy the condition

$$F(x, a) + F(y, b) \geq F(y, a) + F(x, b) \quad \text{for any } x \geq y \text{ and } a \geq b.$$
The condition Eq. (2) arises very naturally in practical usage given the fact that it is satisfied by various functions $F$ that appear throughout chemical graph theory. For example, it is straightforward to check that Eq. (2) holds for each function $F$ that appears in Table 1, except for $F(x, y) = \sqrt{x^2 + y^2}$, in which case this condition is satisfied for $(-F)(x, y) = -\sqrt{x^2 + y^2}$ instead.

By taking into consideration the set $T_D$ for some degree sequence $D$ such that $T_D \neq \emptyset$, Wang [7] has found a way to construct a single tree $T$ that maximizes the value $R_F$ on this set. Moreover, it can be shown that the maximum value of $R_F$ must always be attained by the so-called greedy tree, whose construction is elaborated in [7]. In this paper, we consider the problem of determining not one, but all the possible trees that attain the maximum value of $R_F$ on the set $T_D$. We offer the full solution to the said problem for all the possible discrete symmetric real functions $F: \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ that satisfy the condition

$$F(x, a) + F(y, b) > F(y, a) + F(x, b) \quad \text{for any } x > y \text{ and } a > b. \quad (3)$$

Here, it is important to notice that Eq. (3) is slightly stricter than Eq. (2). In other words, it is not difficult to see that each function that satisfies Eq. (3) must also satisfy Eq. (2), but not vice versa. Of course, this observation creates no issues at all given the fact that Eq. (3) does hold for many functions $F$ having practical usage in chemical graph theory. For example, each function $F$ from Table 1 satisfies Eq. (3), except for $F(x, y) = \sqrt{x^2 + y^2}$, in which case $(-F)(x, y) = -\sqrt{x^2 + y^2}$ does.

We disclose the main result of the paper in the form of the next theorem.

**Theorem 1.** For some $n \in \mathbb{N}$, let $D = (d_1, d_2, \ldots, d_n) \in \mathbb{N}_0^n$ be an arbitrarily chosen non-increasing sequence of $n$ non-negative integers such that $T_D \neq \emptyset$, and let $F: \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ be any discrete symmetric real function that satisfies Eq. (3). Now, let $T_D^*$ denote the set of all the trees that can possibly be obtained via the following non-deterministic construction algorithm:

1. Add the vertex 1 and label it with $d_1$.

2. For each integer $j = 2, n$, add the vertex $j$, label it with $d_j$, and connect it to a single pre-existing vertex $h$, $1 \leq h < j$ such that

   (i) the current degree of $h$ is lower than the label assigned to $h$;

   (ii) there is no pre-existing vertex satisfying (i) which has a label greater than $h$.

3}
Any given tree $T \in \mathcal{T}_D$ attains the maximum value of $R_{\mathcal{F}}$ on $\mathcal{T}_D$ if and only if it is isomorphic to at least one tree from $\mathcal{T}_D^\ast$.

The remainder of the paper will focus on providing a full self-contained proof of Theorem 1. We will accomplish this by relying on three auxiliary lemmas before concentrating on finalizing the theorem proof itself. In the end, we will give a brief concrete example of the usage of Theorem 1.

Throughout the rest of the paper, we will consider $D \in \mathbb{N}_0^n$ to be a given non-increasing sequence of $n \in \mathbb{N}$ non-negative integers such that $\mathcal{T}_D \neq \emptyset$. Also, we will consider $\mathcal{F}: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ to be any discrete symmetric real function that satisfies Eq. (3).

To begin with, we shall refer to the ordered pair $((v_0, v_1, v_2, \ldots, v_{n-1}), (f_1, f_2, \ldots, f_{n-1}))$ as a construction scheme of some tree $T$ of order $n \in \mathbb{N}$ provided that this tree can be obtained via the following simple algorithm:

1. Add the vertex $v_0$.

2. For each integer $j = 1, n-1$, add the new vertex $v_j$ and then connect it to the previously added vertex $f_j$.

It is clear that each tree surely has at least one construction scheme. However, it becomes convenient to notice that the trees that attain the maximum value of $R_{\mathcal{F}}$ on $\mathcal{T}_D$ always possess very specific construction schemes. Our immediate goal shall be to elaborate on this fact and provide a result that will later be used while finishing off the proof of Theorem 1. We shall start with the following auxiliary lemma regarding the degrees of vertices that lie on an arbitrary path.

**Lemma 2.** Let $T \in \mathcal{T}_D$ be a tree that attains the maximum value of $R_{\mathcal{F}}$ on $\mathcal{T}_D$ and let $u$ and $v$ be two of its arbitrarily chosen vertices. For any vertex $w$ that lies on the path from $u$ to $v$, we necessarily have

$$\deg_T(w) \geq \min(\deg_T(u), \deg_T(v)).$$

**Proof.** We shall prove the lemma by contradiction. Let $P$ be the $(u, v)$-path in $T$ and suppose that there does lie a vertex on $P$ whose degree is below $\min(\deg_T(u), \deg_T(v))$. It is straightforward to see that $\deg_T(u), \deg_T(v) \geq 2$ must hold. For this reason, we can construct a non-trivial path $Q$ from $v$ to some leaf $t$ so that this path is entirely disjoint with $P$, except for the vertex $v$.  

4
Now, let \( p_1 \) be the first vertex on \( P \) whose degree is below \( \min(\deg_T(u), \deg_T(v)) \), and let \( p_0 \) be the vertex on this path before it. Similarly, let \( q_1 \) be the first vertex on \( Q \) whose degree is below \( \min(\deg_T(u), \deg_T(v)) \), and let \( q_0 \) be the vertex on this path before it. Taking everything into consideration, we obtain a \((p_0, q_1)\)-path as depicted in Figure 1 that will be of further interest.

![Figure 1](image)

Figure 1: The obtained \((p_0, q_1)\)-path in \( T \), alongside the vertices \( u \) and \( t \).

It is clear that the tree \( T \) satisfies

\[
\begin{align*}
\ p_0 & \sim p_1, \\
\ q_0 & \sim q_1, \\
\ q_0 & \not\sim q_1, \\
\ q_1 & \not\sim q_0.
\end{align*}
\]

Bearing this in mind, we can remove the edges \( p_0p_1 \) and \( q_0q_1 \) from \( T \) and add the edges \( p_0q_0 \) and \( p_1q_1 \) in order to obtain another tree \( T_1 \) whose vertices have the same degrees as in \( T \). Hence, \( T_1 \in \mathcal{T}_D \). Furthermore, \( R_\mathcal{F}(T) \) and \( R_\mathcal{F}(T_1) \) will have the same summands in Eq. (1) except for those that correspond to the deleted and newly added edges. This immediately implies

\[
R_\mathcal{F}(T_1) - R_\mathcal{F}(T) = \mathcal{F}(\deg_T(p_0), \deg_T(q_0)) + \mathcal{F}(\deg_T(p_1), \deg_T(q_1)) \\
- \mathcal{F}(\deg_T(p_0), \deg_T(p_1)) - \mathcal{F}(\deg_T(q_0), \deg_T(q_1)).
\] (4)

However, we know that

\[
\deg_T(p_0), \deg_T(q_0) \geq \min(\deg_T(u), \deg_T(v)), \\
\deg_T(p_1), \deg_T(q_1) < \min(\deg_T(u), \deg_T(v)),
\]

which swiftly leads us to

\[
\mathcal{F}(\deg_T(p_0), \deg_T(q_0)) + \mathcal{F}(\deg_T(q_1), \deg_T(p_1)) > \\
> \mathcal{F}(\deg_T(q_1), \deg_T(q_0)) + \mathcal{F}(\deg_T(p_0), \deg_T(p_1))
\]

by virtue of Eq. (3). Now, Eq. (4) tells us that \( R_\mathcal{F}(T_1) - R_\mathcal{F}(T) > 0 \) must hold, which is impossible since the tree \( T \) attains the maximum value of \( R_\mathcal{F} \) on \( \mathcal{T}_D \). Hence, we obtain a contradiction. \( \Box \)

Now, by taking into consideration Lemma 2, we are able to formulate and prove the next lemma regarding the constructibility of trees that attain the maximum \( R_\mathcal{F} \) value.
Lemma 3. If $T \in \mathcal{T}_D$ is some tree that attains the maximum value of $R_F$ on $\mathcal{T}_D$, then this tree surely has a construction scheme $((v_0, v_1, v_2, \ldots, v_{n-1}), (f_1, f_2, \ldots, f_{n-1}))$ such that

- for all the $0 \leq j < h \leq n - 1$, we have $\deg_T(v_j) \geq \deg_T(v_h)$;

- for all the $1 \leq j < h \leq n - 1$ such that $\deg_T(v_j) = \deg_T(v_h)$, the condition $\deg_T(f_j) \geq \deg_T(f_h)$ must hold.

Proof. Let $T \in \mathcal{T}_D$ be a given tree that attains the maximum value of $R_F$ on $\mathcal{T}_D$. Lemma 2 tells us that, for each $j = \min D, \max D$, the subgraph of $T$ induced by the set of vertices whose degree is at least $j$ must be a tree. From here, we quickly conclude that we can construct $T$ by simply constructing its subtree induced by the vertices of degree $\max D$, then extending this subtree to the subtree induced by the vertices of degree at least $\max D - 1$, and so on, until we obtain $T$ itself. Thus, the tree $T$ necessarily has a construction scheme $C' = ((v'_0, v'_1, v'_2, \ldots, v'_{n-1}), (f'_1, f'_2, \ldots, f'_{n-1}))$ such that the degrees of $v'_0, v'_1, v'_2, \ldots, v'_{n-1}$ appear in non-increasing order.

We have obtained a construction scheme $C'$ that satisfies the first condition given in the lemma. In order to finalize the proof, we will explain how this construction scheme can be modified so that the second condition surely holds as well. First of all, it is easy to check that the second condition necessarily holds for the vertices of degree $\max D$, hence it becomes sufficient to show that, for any $\beta$, $\min D \leq \beta < \max D$, the addition of vertices of degree $\beta$ within the construction scheme $C'$ can be permuted in some manner so that the second condition becomes satisfied.

The key observation to make is that while $T$ is constructed via the algorithm dictated by $C'$, each vertex of degree $\beta$ is surely connected to a vertex of degree at least $\beta$ upon being added. Moreover, each vertex of degree $\beta$ that is connected to a vertex of degree greater than $\beta$ can certainly freely be reordered among all the vertices of degree $\beta$. In other words, this vertex can be added before or after any other vertex of degree $\beta$, given the fact that its initial neighbor is definitely present to begin with. This directly means that we can reorder the addition of all the vertices of degree $\beta$ so that we first add those whose initial neighbor has the greatest possible degree, then those whose initial neighbor has the second greatest degree, and so on, until we add the vertices of degree $\beta$ whose initial neighbor also has the degree $\beta$, and which cannot freely be reordered. By applying the said transformation on $C'$ for each possible $\beta$, $\min D \leq \beta < \max D$, we obtain a construction scheme...
Lemma 4. Any tree $T \in \mathcal{T}_D$ that attains the maximum value of $R_F$ on $\mathcal{T}_D$ must be isomorphic to at least one tree from $\mathcal{T}^*_D$.

Proof. Let $T$ be any such tree. It is clear that this tree must have a construction scheme $C$ that satisfies the criteria stated in Lemma 3. Now, while $T$ is being constructed via the algorithm dictated by $C$, suppose that there exists a vertex $v$ such that, upon being added, it is not connected to a pre-existing available vertex with the greatest possible degree. Let $p$ be such a pre-existing vertex and let $q$ be the vertex that $v$ gets connected to instead. Due to the criteria imposed on $C$ by virtue of Lemma 3, we see that none of the vertices of degree $\deg_T(v)$ that are added after $v$ can be connected to $p$ either, which means that the vertex $p$ necessarily has a neighbor $u$ in $T$ such that $\deg_T(u) < \deg_T(v)$. Taking everything into consideration, we obtain that the tree $T$ bears a structure as demonstrated in Figure 2.

![Figure 2: The structure of the tree $T$.](image)

It is obvious that the tree $T$ satisfies

$$u \sim p, \qquad v \sim q, \qquad u \not\sim q, \qquad v \not\sim p.$$  

If we remove the edges $up$ and $vq$ from $T$ and add the edges $uq$ and $vp$, we get another tree $T_1$ whose vertices have the same degrees as in $T$. For this reason, we have $T_1 \in \mathcal{T}_D$. Using the same logic as in the proof of Lemma 2, it is easy to show that

$$R_F(T_1) - R_F(T) = F(\deg_T(u), \deg_T(q)) + F(\deg_T(v), \deg_T(p)) - F(\deg_T(u), \deg_T(p)) - F(\deg_T(v), \deg_T(q)).$$  

(5)

Taking into consideration that

$$\deg_T(p) > \deg_T(q) \geq \deg_T(v) > \deg_T(u),$$

the proof is completed. \qed
it becomes straightforward to obtain

\[ F(\text{deg}_T(p), \text{deg}_T(v)) + F(\text{deg}_T(q), \text{deg}_T(u)) > \\
> F(\text{deg}_T(q), \text{deg}_T(v)) + F(\text{deg}_T(p), \text{deg}_T(u)) \]

by directly implementing Eq. (3). Now, by using Eq. (5), this immediately leads us to \( R_{T_1} - R_T > 0 \), which is clearly not possible due to the fact that \( T \) attains the maximum value of \( R_T \) on \( \mathcal{T}_D \).

Thus, we conclude that while \( T \) is being constructed in accordance with the construction scheme \( C \), the vertices must be added in such a way their degrees yield a non-increasing sequence, with each vertex after the first being connected to a pre-existing available vertex with the greatest possible degree. However, this is precisely how the construction algorithm described in Theorem 1 works. Hence, it promptly follows that \( T \) must indeed be isomorphic to at least one tree from \( \mathcal{T}_D^* \).

We are now finally in position to put all the pieces of the puzzle together and complete the proof of Theorem 1.

**Proof of Theorem 1**

If a tree \( T \in \mathcal{T}_D \) attains the maximum value of \( R_T \) on \( \mathcal{T}_D \), then it is surely isomorphic to at least one tree from \( \mathcal{T}_D^* \), by virtue of Lemma 4. Thus, in order to finish the theorem proof, we need to show that each tree \( T \in \mathcal{T}_D \) isomorphic to at least one tree from \( \mathcal{T}_D^* \) must also attain the maximum value of \( R_T \) on \( \mathcal{T}_D \). In other words, it is sufficient to prove that all the trees from \( \mathcal{T}_D^* \) attain the maximum value of \( R_T \).

It is obvious that all the trees from \( \mathcal{T}_D^* \) necessarily attain the same value of \( R_T \). This follows by simply inspecting the construction algorithm given in Theorem 1 and noticing that, regardless of which vertex we choose as \( h \) in each step, the contribution to Eq. (1) from the newly added edge will be the same. Now, it is clear that there exists a tree \( T_0 \in \mathcal{T}_D \) attaining the maximum value of \( R_T \) on \( \mathcal{T}_D \). Due to Lemma 4 we know that \( T_0 \) is certainly isomorphic to at least one tree from \( \mathcal{T}_D^* \). However, this means that there exists a tree in \( \mathcal{T}_D^* \) that attains the maximum value of \( R_T \) on \( \mathcal{T}_D \), which directly implies that all of them must attain the maximum value of \( R_T \), as desired. \( \Box \)
We end the paper by giving a brief example of how Theorem 1 can be used on a concrete degree sequence. Let $D = (4, 4, 3, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1)$. If we implement the result obtained by Wang [7], we get that the greedy tree depicted in Figure 3 surely attains the maximum value of $R_F$ on $T_D$.

Now, if we apply Theorem 1 we are able to obtain a much stronger result. More precisely, we conclude that some tree $T$ attains the maximum value of $R_F$ on $T_D$ if and only if it belongs to one of the three isomorphism classes shown in Figure 4. This observation is straightforward to notice — the trees in $T_D^*$ are precisely such that the two vertices of degree four are adjacent, while each vertex of degree three or two must have a neighbor of degree four. Thus, there essentially exist exactly three different trees that attain the maximum of $R_F$ on $T_D$, with the greedy tree obtained by Wang corresponding to the tree given in Figure 4a.

![Figure 3: The greedy tree for $D = (4, 4, 3, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1)$](image_url)
Figure 4: All three isomorphism classes corresponding to the trees from the set $T_D^*$ for $D = (4, 4, 3, 3, 2, 1, 1, 1, 1, 1, 1, 1)$. The vertices of degree four, three and two are colored in black, red and yellow, respectively.

References

[1] M. Randić. On characterization of molecular branching. *J. Amer. Chem. Soc.* 97(23):6609–6615, 1975.

[2] I. Gutman and N. Trinajstić. Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons. *Chem. Phys. Lett.* 17(4):535–538, 1972.

[3] A. Miličević, S. Nikolić and N. Trinajstić, On reformulated Zagreb indices. *Mol. Divers.* 8(4):393-399, 2004.
[4] B. Zhou and N. Trinajstić. On a novel connectivity index. *J. Math. Chem.* 46:1252–1270, 2009.

[5] S. Fajtlowicz. On conjectures of Graffiti–II. *Congr. Numer.* 60:187–197, 1987.

[6] I. Gutman. Geometric approach to degree-based topological indices: Sombor indices. *MATCH Commun. Math. Comput. Chem.* 86:11–16, 2021.

[7] H. Wang. Functions on adjacent vertex degrees of trees with given degree sequence. *Cent. Eur. J. Math.* 12:1656–1663, 2014.