Further analysis on the total number of subtrees of trees

Shuchao Li[1,*] and Shujing Wang[2]

Faculty of Mathematics and Statistics
Central China Normal University
Wuhan 430079, P.R. China

[1]lscmath@mail.ccnu.edu.cn, [2]wang06021@126.com

Submitted: Mar 29, 2012; Accepted: Dec 5, 2012; Published: XX
Mathematics Subject Classifications: 05C05, 05C07, 05C35,

Abstract

When considering the total number of subtrees of trees, the extremal structures which maximize this number among binary trees and trees with a given maximum degree lead to some interesting facts that correlate to some other graphical indices in applications. Along this line, it is interesting to study that over some types of trees with a given order, which trees minimize or maximize this number. Here are our main results: (1) The extremal tree which minimizes the total number of subtrees among n-vertex trees with k pendants is characterized. (2) The extremal tree which maximizes (resp. minimizes) the total number of subtrees among n-vertex trees with a given bipartition is characterized. (3) The extremal tree which minimizes the total number of subtrees among the set of all q-ary trees with n non-leaf vertices is identified. (4) The extremal n-vertex tree with given domination number maximizing the total number of subtrees is characterized.

Keywords: subtrees; Wiener index; leaves; bipartition; q-ary tree; domination number

1 Introduction

We consider only connected simple graphs (i.e., finite, undirected graphs without loops or multiple edges). We follow the notations and terminology in [1] except otherwise stated.

Let $G = (V_G, E_G)$ be a graph with $u, v \in V_G$, $d_G(u)$ (or $d(u)$ for short) denotes the degree of u. Throughout we denote by P_n, $K_{1,n-1}$ the path and star on n vertices,

*Financially supported by the National Natural Science Foundation of China (Grant Nos. 11071096, 11271149) and the Special Fund for Basic Scientific Research of Central Colleges (CCNU11A02015).

respectively. $G-v$, $G-uv$ denote the graph obtained from G by deleting vertex $v \in V_G$, or edge $uv \in E_G$, respectively (this notation is naturally extended if more than one vertex or edge is deleted). Similarly, $G+uv$ is obtained from G by adding edge $uv \notin E_G$. For $v \in V_G$, let $N_G(v)$ (or $N(v)$ for short) denote the set of all the adjacent vertices of v in G. We refer to vertices of degree 1 of G as *leaves* (or *pendant vertices*), and the edges incident to leaves are called *pendant edges*. Let $PV(G)$ denote the set of all leaves of G. For convenience, let $\mathscr{T}_n^k$ be the set of all n-vertex trees with k leaves.

Let G be a connected bipartite graph with n vertices. Hence its vertex set can be partitioned into two subsets V_1 and V_2, such that each edge joins a vertex in V_1 with a vertex in V_2. Suppose that V_1 has p vertices and V_2 has q vertices, where $p+q=n$. Then we say that G has a (p,q)-*bipartition* $(p \leqslant q)$. For convenience, let $\mathscr{P}_n^{p,q}$ be the set of all n-vertex trees, each of which has a (p,q)-bipartition. Given positive integers n, q with $q \geqslant 2$, we call T a *complete q-ary tree* (or *q-ary tree* for short) if any non-pendant vertex v in T has exactly q neighbours. Denote by $\mathscr{A}_n^q$ the class of q-ary trees with n non-leaf vertices ($(q-2)n+2$ leaves).

A subset S of V_G is called a *dominating set* of G if for every vertex $v \in V_G \setminus S$, there exists a vertex $u \in S$ such that v is adjacent to u. For a dominating set S of graph G with $v \in S$, $u \in V_G \setminus S$, if $vu \in E_G$, then u is said to be dominated by v. The *domination number* of G, denoted by $\gamma(G)$, is defined as the minimum cardinality of dominating sets of G. Denote by $\mathscr{D}_n^\gamma$ the set of all n-vertex trees with domination number γ.

Given a tree T, a *subtree* of T is just a connected induced subgraph of T. The number of subtrees $F(T)$ has received much attention. Let T denote a tree with n vertices each of whose non-leaves has degree at least three, Andrew and Wang [12] showed that the average orders in the subtrees of T is at least $\frac{n}{2}$ and strictly less than $\frac{3n}{4}$. Székely and Wang [7] characterized the binary trees with n leaves that have the greatest number of subtrees. Kirk and Wang [5] identified the tree, for a given size and such that the vertex degree is bounded, having the greatest number of subtrees. Székely and Wang [10] gave a formula for the maximal number of subtrees a binary tree can possess over a given number of vertices. They also show that caterpillar trees (trees containing a path such that each vertex not belonging to the path is adjacent to a vertex on the path) have the smallest number of subtrees among binary trees. Yan and Ye [18] characterized the tree with the diameter at least d having the maximum number of subtrees.

In 2012, Zhang, Zhang, Gray, and Wang [23, 24] determined the extremal tree with given degree sequence having the largest number of subtrees. As a consequence, they obtained the extremal trees with maximum number of subtrees among the set of trees with given number of independence number, leaves, matching number, respectively. Zhang and Zhang [22] investigated the structures of an extremal tree which has the minimal number of subtrees in the set of all trees with a very special degree sequence. For some related results on the enumeration of subtrees of trees, one may also see Székely and Wang [8, 9], Wang [15] and Song [6].

An interesting fact that among the n-vertex trees of given degree sequence, the extremal one that maximizes the total number of subtrees is exactly the one that minimizes some chemical indices such as the well known Wiener index, and vice versa; see [13, 19-23].

The correlation between these and other indices has been the subject of investigation in the recent work of Székley, Wang and Wu [11]; see also the work of Wagner [13]. Along this line, it is interesting and nature to study the extremal problems on $F(T)$ in some other types of trees. In particular, it is interesting to characterize the extremal trees with minimal number of subtrees among some types of trees. Motivated by these facts, in this paper we focus on characterizing the structure of trees maximizing/minimizing $F(T)$ among $\mathscr{T}_n^k$, $\mathscr{P}_n^{p,q}$, $\mathscr{A}_n^q$ and $\mathscr{D}_n^\gamma$, respectively.

Let $P_k(a,b)$ be a tree obtained by attaching a and b pendant vertices to the two pendant vertices of P_k, respectively. In particular, if $k=1$, then $P_k(a,b)=K_{1,a+b}$. It is straightforward to check that $P_{n-k}(\lfloor\frac{k}{2}\rfloor,\lceil\frac{k}{2}\rceil)\in\mathscr{T}_n^k$. It is known that the Wiener index among n-vertex trees with k pendant vertices is maximized by $P_{n-k}(\lfloor\frac{k}{2}\rfloor,\lceil\frac{k}{2}\rceil)$; see Dobrynin, Entringer and Gutman [2]. We are to show the counterpart of this result for the number of subtrees.

Theorem 1. *Precisely the graph $P_{n-k}(\lfloor\frac{k}{2}\rfloor,\lceil\frac{k}{2}\rceil)$ minimizes the total number of subtrees among $\mathscr{T}_n^k$.*

Consider a star $K_{1,p}$ with $p+1$ vertices and attach $q-1$ pendant edges to a non-central vertex of the star $K_{1,p}$. The resulting tree with $p+q$ vertices has a (p,q)-bipartition. Denote the resulting tree by $D(p,q)$; see Fig. 1. Obviously, $D(p,q)\in\mathscr{P}_n^{p,q}$. We call $D(p,q)$ a *double star*. Ye and Chen [19] characterized the tree with given bipartition having minimal energy and Hosoya index. Then it is natural to consider the extremal problem on the total number of subtrees of trees with given bipartition.

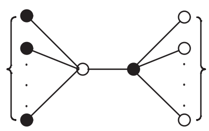

Figure 1: Tree $D(p,q)$.

Theorem 2. *Precisely the graph $D(p,q)$ (resp. $P_{2p-1}(\lfloor\frac{n-2p+1}{2}\rfloor,\lceil\frac{n-2p+1}{2}\rceil)$) maximizes (resp. minimizes) the total number of subtrees among $\mathscr{P}_n^{p,q}$.*

Consider the path P_{n+2} and attach $q-2$ pendant edges to each of the non-leaf vertices of P_{n+2}. Denote the resulting tree by $\hat{T}_n^q$ (see Fig. 2). It is easy to see that $\hat{T}_n^q\in\mathscr{A}_n^q$. In view of Theorem 2.3 in [5], it is easy to determine the tree in $\mathscr{A}_n^q$ which maximizes the total number of subtrees. It is natural and interesting to characterize the tree with minimum number of subtrees of trees among $\mathscr{A}_n^q$.

Theorem 3. *Precisely the graph $\hat{T}_n^q$ (see Fig. 2) minimizes the total number of subtrees among $\mathscr{A}_n^q$.*

Let $A(n,\gamma)$ be the tree that is obtained by attaching $\gamma-1$ pendant edges to $\gamma-1$ pendant vertices of the star $K_{1,n-\gamma}$. It is routine to check that $A(n,\gamma)\in\mathscr{D}_n^\gamma$. He, Wu and

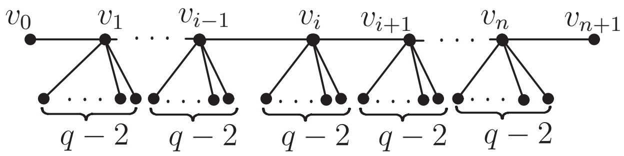

Figure 2: Tree $\hat{T}_n^q$.

Yu [4] characterized the tree with given domination number having minimal energy. It is natural to consider the extremal problem on the total number of subtrees of trees with given bipartition.

Theorem 4. *Precisely the graph* $A(n,\gamma)$ *maximizes the total number of subtrees in* $\mathscr{D}_n^\gamma$.

Theorem 5. *Precisely the graph* $P_{\frac{n}{2}} \circ K_1$ *minimizes the total number of subtrees in* $\mathscr{D}_n^{\frac{n}{2}}$*, where* $P_{\frac{n}{2}} \circ K_1$ *is obtained by attaching a leaf to each vertex of the path* $P_{\frac{n}{2}}$.

2 Some Lemmas

In this section, we give some necessary results which will be used to prove our main results. For a set S, let $|S|$ denote its cardinality. For two graphs G_1, G_2, if G_1 is a *connected subgraph* of G_2, then we denote it by $G_1 \subseteq G_2$. Given a tree T with $u, v \in V_T$, let

$$f_T(u) = |\{T' : \ T' \subseteq T,\, u \in V_{T'}\}|, \qquad f_T(u*v) = |\{T' : \ T' \subseteq T,\, u,\, v \in V_{T'}\}|,$$
$$f_T(u/v) = |\{T' : \ T' \subseteq T,\, u \in V_{T'}, v \notin V_{T'}\}|, \qquad F(T) = |\{T' : T' \subseteq T,\, |V_{T'}| \geqslant 1\}|.$$

Then one has

$$f_T(u) = f_T(u*v) + f_T(u/v), \quad f_T(v) = f_T(u*v) + f_T(v/u). \tag{1}$$

Consider the tree W in Fig. 3 with vertices x and y, and

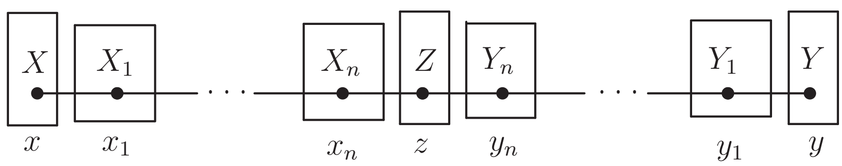

Figure 3: Path $P_W(x,y)$ connecting vertices x and y.

$$P_W(x,y) = x_0(x)x_1 \dots x_n z y_n \dots y_1 y_0(y) \ (x_0(x)x_1 \dots x_n y_n \dots y_1 y_0(y))$$

if $d_W(x,y)$ is even (odd) for any $n \geqslant 0$. After the deletion of all the edges of $P_W(x,y)$ from W, some connected components will be remained. Let $X_i\,(X_0)$ denote the component that contains $x_i\,(x_0 = x)$, let $Y_i\,(Y_0)$ denote the component that contains $y_i\,(y_0 = y)$, for $i = 1, 2, \dots, n$, and let Z denote the component that contains z.

Lemma 6 ([7])**.** *In the above situation, if $f_{X_i}(x_i) \geqslant f_{Y_i}(y_i)$ for $i = 0, 1, \ldots, n$, then $f_W(x) \geqslant f_W(y)$. Furthermore, $f_W(x) = f_W(y)$ if and only if $f_{X_i}(x_i) = f_{Y_i}(y_i)$ for all i.*

If we have a tree T with vertices x and y, and two rooted trees X and Y, then we can build two new trees, first T', by identifying the root of X with x and the root of Y with y, second T'', by identifying the root of X with y and the root of Y with x (as shown in Fig. 4).

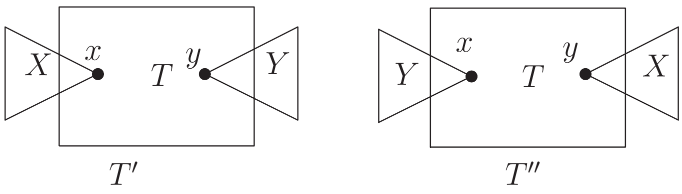

Figure 4: Switching subtrees rooted at x and y.

Lemma 7 ([7])**.** *In the above situation, if $f_T(x) > f_T(y), f_X(x) < f_Y(y)$, then we have $F(T'') > F(T')$.*

Two distinct edges in a graph G are *independent* if they do not have a common end vertex in G. A set of pairwise independent edges of G is called a *matching* of G, while a matching of maximum cardinality is a maximum matching of G. The *matching number* β of G is the cardinality of a maximum matching of G.

Lemma 8 ([23])**.** *Precisely the graph T^n_β maximizes the total number of subtrees of n-vertex trees with matching number β, where T^n_β is obtained from the star $K_{1,n-\beta}$ by adding $\beta - 1$ pendant edges to $\beta - 1$ leaves of $K_{1,n-\beta}$.*

Lemma 9. *Consider a longest path $P_r = v_1v_2 \ldots v_r$ in a tree T with $r \geqslant 3$, there exists a vertex $v_i \in V_{P_r} \setminus \{v_1, v_r\}$ such that*

$$f_T(v_1) < \cdots < f_T(v_{i-1}) < f_T(v_i) \geqslant f_T(v_{i+1}) > \cdots > f_T(v_r). \tag{2}$$

Proof. Choose three vertices x, y, z such that $xy, yz \in E_T$. Let X, Y, Z, respectively, denote the components containing x, y, z after the removal of the edges xy and yz from T. Observe the identities:

$$\begin{aligned}
f_T(x) &= f_X(x) + f_X(x)f_Y(y) + f_X(x)f_Y(y)f_Z(z),\\
f_T(z) &= f_Z(z) + f_Z(z)f_Y(y) + f_Z(z)f_Y(y)f_X(x),\\
f_T(y) &= f_Y(y) + f_X(x)f_Y(y) + f_Z(z)f_Y(y) + f_X(x)f_Y(y)f_Z(z).
\end{aligned}$$

This gives

$$2f_T(y) - f_T(x) - f_T(z) = 2f_Y(y) + (f_X(x) + f_Z(z))(f_Y(y) - 1) > 0. \tag{3}$$

Let

$$i := \min\{j :\ 1 \leqslant j \leqslant r,\ f_T(v_j) \geqslant f_T(u),\ u \in V_{P_r}\}. \tag{4}$$

In view of (1) and $f_T(v_2/v_1) > f_T(v_1/v_2) = 1$, $f_T(v_{r-1}/v_r) > f_T(v_r/v_{r-1}) = 1$, we can see that $i \neq 1, r$. By (3) and (4), we have

$$2f_T(v_{i+1}) - f_T(v_i) - f_T(v_{i+2}) > 0, \quad f_T(v_i) \geqslant f_T(v_{i+1}).$$

Hence, $f_T(v_i) \geqslant f_T(v_{i+1}) > f_T(v_{i+2})$. Repeat the procedure as above to obtain

$$f_T(v_i) \geqslant f_T(v_{i+1}) > f_T(v_{i+2}) > \cdots > f_T(v_r). \tag{5}$$

Similarly, we obtain

$$f_T(v_1) < \cdots < f_T(v_{i-1}) < f_T(v_i). \tag{6}$$

Hence, (5) and (6) imply (2) immediately. □

Let T_1 be the graph as depicted in Fig. 5, where T' (resp. T'') is a tree with at least two vertices. Attaching a pendant edge to u and contracting the edge uv of T_1 yields the graph T_2; see Fig. 5. We call the procedure constructing T_2 from T_1 the *A-transformation* of T_1.

Figure 5: Trees T_1 and T_2.

Lemma 10. *Let T_1 and T_2 be the trees defined as above, we have $F(T_1) < F(T_2)$.*

Proof. Let $\bar{T}$ be the component containing v in $T_1 - N_{T''}(v)$. Note that

$$f_{\bar{T}}(u) - f_{\bar{T}}(v) = f_{\bar{T}}(u/v) - f_{\bar{T}}(v/u) = f_{\bar{T}}(u) - 1 > 0,$$

i.e., $f_{\bar{T}}(u) > f_{\bar{T}}(v)$. Hence, by Lemma 7 our result holds. □

3 Proofs of Theorems 1 and 2

In this section, we first determine the extremal tree which minimizes the total number of subtrees among $\mathscr{T}_n^k$. Next we determine the extremal tree which maximizes (resp. minimizes) the total number of subtrees among $\mathscr{P}_n^{p,q}$.

Proof of Theorem 1. Choose $T \in \mathscr{T}_n^k$ such that the total number of its subtrees is as small as possible. If $k = 2$ or, $k = n - 1$, it is easy to see that $\mathscr{T}_n^k = \{P_{n-k-1}(\lfloor \frac{k}{2} \rfloor, \lceil \frac{k}{2} \rceil)\}$, our result follows immediately. Hence, in what follows we consider $2 < k < n - 1$. In order to complete the proof, it suffices to show the following claims.

Claim 11. *If T minimizes the total number of subtrees in $\mathscr{T}_n^k$, then $T \cong P_{n-k}(a,b)$, where $a \geqslant b \geqslant 1$ and $a+b=k$.*

Proof of Claim 11 If $\operatorname{diam}(T)=3$, the claim follows immediately. Hence we consider the trees whose diameter is larger than 3. Suppose that $P_r = v_1 \ldots v_r$ $(r \geqslant 5)$ is one of the longest paths in T, then we have $d_T(v_3) = d_T(v_4) = \cdots = d_T(v_{r-2}) = 2$; otherwise set

$$i := \min\{j :\ d_T(v_j) \geqslant 3,\ v_j \in \{v_3, v_4, \ldots, v_{r-2}\}\}.$$

and denote $N_T(v_i) = \{v_{i-1}, v_{i+1}, z_1, z_2, \ldots, z_s\}$, $\ s \geqslant 1$. Delete all the vertices $z_1, z_2, \ldots, z_s$ from T and let T_0 be the component containing v_i. By Lemma 9, there exists $v_t \in V_{P_r}$ such that

$$f_{T_0}(v_1) < \cdots < f_{T_0}(v_{t-1}) < f_{T_0}(v_t) \geqslant f_{T_0}(v_{t+1}) > \cdots > f_{T_0}(v_r).$$

If $t < i$, then we have $f_{T_0}(v_i) > f_{T_0}(v_{r-1})$. By Lemma 7, we have

$$F(T) > F(T'), \tag{7}$$

where $T' = T - v_i z_1 - \cdots - v_i z_s + v_{r-1} z_1 + \cdots + v_{r-1} z_s$.

If $t \geqslant i$, then we have $f_{T_0}(v_i) > f_{T_0}(v_2)$. By Lemma 7, we have

$$F(T) > F(T''), \tag{8}$$

where $T'' = T - v_i z_1 - \cdots - v_i z_s + v_2 z_1 + \cdots + v_2 z_s$.

Note that $T', T'' \in \mathscr{T}_n^k$, hence (7) (resp. (8)) is a contradiction to the choice of T. This completes the proof of Claim 11. $\square$

Claim 12. *Given positive integers a, b with $a \geqslant b$ and $a+b=k$, one has*

$$F(P_{n-k}(a,b)) = (2^a + 2^b)(n-k-1) + 2^k + k + \binom{n-k-1}{2}. \tag{9}$$

Furthermore, if $a-b \geqslant 2$ then

$$F(P_{n-k}(a,b)) > F(P_{n-k}(a-1,b+1)). \tag{10}$$

Proof of Claim 12 For convenience, assume that $d_{P_{n-k}(a,b)}(v_1) = a+1$ and $d_{P_{n-k}(a,b)}(v_{n-k}) = b+1$. Then we have

$$\begin{aligned} F(P_{n-k}(a,b)) = & f_{P_{n-k}(a,b)}(v_1/v_{n-k}) + f_{P_{n-k}(a,b)}(v_1 * v_{n-k}) + f_{P_{n-k}(a,b)}(v_{n-k}/v_1) \\ & + F(P_{n-k}(a,b) - v_1 - v_{n-k}). \end{aligned} \tag{11}$$

By direct calculation, we have

$$f_{P_{n-k}(a,b)}(v_1/v_{n-k}) = 2^a(n-k-1) \quad \text{and} \quad f_{P_{n-k}(a,b)}(v_{n-k}/v_1) = 2^b(n-k-1). \tag{12}$$

It is straightforward to check that the total number of subtrees of $P_{n-k}(a, b)$ containing both v_1 and v_{n-k} is equal to the total number of subtrees of $K_{1,a+b}$ each contains the center of $K_{1,a+b}$. Hence, we have

$$f_{P_{n-k}(a,b)}(v_1 * v_{n-k}) = 2^{a+b} = 2^k. \tag{13}$$

On the other hand,

$$F(P_{n-k}(a, b) - v_1 - v_{n-k}) = F((a + b)P_1 \cup P_{n-k-2}) = k + \binom{n-k-1}{2}. \tag{14}$$

In view of (11)-(14), (9) holds. Hence,

$$F(P_{n-k}(a, b)) - F(P_{n-k}(a-1, b+1)) = (2^a+2^b-2^{a-1}-2^{b+1})(n-k-1) = (2^{a-1}-2^b)(n-k-1).$$

Note that $a - b \geqslant 2$, hence $(2^{a-1} - 2^b)(n - k - 1) > 0$, i.e., $F(P_{n-k}(a, b)) > F(P_{n-k}(a - 1, b + 1))$, as desired. □

By Claims 11 and 12, Theorem 1 follows immediately. □

Proof of Theorem 2 For convenience, denote by $\iota(T)$ the number of non-pendant vertices of T. For any $T \in \mathscr{P}_n^{p,q}$. If $p = 1$, $\mathscr{P}_n^{p,q} = \{K_{1,n-1}\} = \{D(1, n - 1)\} = \{P_1(\lfloor \frac{n-1}{2})\rfloor, \lceil \frac{n-1}{2} \rceil\}$. Our result holds in this case. Hence, in what follows, we consider $p \geqslant 2$.

First choose $T \in \mathscr{P}_n^{p,q}$ such that the total number of its subtrees is as large as possible. In order to characterize the structure of T, it suffices to show that $\iota(T) = 2$.

Note that when $p \geqslant 2$, hence $\iota(T) \neq 1$. So we assume to the contrary that $\iota(T) \geqslant 3$. Choose three vertices, say u, v, w, such that each of them is of degree at least 2. Let $V_T = V_1 \cup V_2$. It is straightforward to check that $\{u, v, w\}$ contains two elements in V_1 or V_2. We assume, without loss of generality, that $u, v \in V_1$ with $N_T(u) = \{u_1, z_1, \ldots, z_t\}$, $N_T(v) = \{u_{2k-1}, r_1, \ldots, r_s\}$, $t \geqslant 1$, $s \geqslant 1$ and the unique path joining u and v is $P = uu_1 \ldots u_{2k-1}v$. Consider the component in $T - uz_1 - \cdots - uz_t - vr_1 - \cdots - vr_s$, say T', which contains both u and v.

Without loss of generality, we assume $f_{T'}(u) \leqslant f_{T'}(v)$, then in view of (1), we can see that $f_{T'}(u/v) \leqslant f_{T'}(v/u)$. Let T'' be the component containing both u and v in the graph $T - uz_1 - \cdots - uz_t$, it is straightforward to see that

$$f_{T''}(u) - f_{T''}(v) = f_{T''}(u/v) - f_{T''}(v/u) = f_{T'}(u/v) - f_{T''}(v/u) < f_{T'}(u/v) - f_{T'}(v/u) \leqslant 0.$$

i.e.,

$$f_{T''}(u) < f_{T''}(v). \tag{15}$$

Let

$$T^* = T - uz_1 - uz_2 - \cdots - uz_t + vz_1 + vz_2 + \cdots + vz_t.$$

Note that u and v are in V_1, hence we have $T^* \in \mathscr{P}_n^{p,q}$. On the other hand, by (15) and Lemma 7, we have $F(T) < F(T^*)$, a contradiction to the choice of T. Hence, we get that $\iota(T) = 2$, i.e., $T \cong D(p, q)$, as desired.

Now choose $T \in \mathscr{P}_n^{p,q}$ such that the total number of its subtrees is as small as possible. If $p = q$ or $p = q-1$, it is easy to see that $P_n \in \mathscr{P}_n^{p,q}$, as P_n minimizes the total number of subtrees of n-vertex tree, we can see that $P_n = P_{2p-1}(1,1)$ or $P_n = P_{2p-1}(0,1)$, minimizes the total number of subtrees among $\mathscr{P}_n^{p,q}$. Hence, in what follows we consider $1 < p < \lfloor \frac{n}{2} \rfloor$. In order to complete the proof, it suffices to show the following claim.

Claim 13. *If T minimizes the total number of subtrees of trees in $\mathscr{P}_n^{p,q}$, then $T \cong P_{2p-1}(a,b)$, where $a+b = n-2p+1$ with $a \geqslant b \geqslant 1$.*

Proof of Claim 13 If $1 < p < \lfloor \frac{n}{2} \rfloor$, it is easy to see that $T \ncong D(p,q)$, so $\mathrm{diam}(T) \geqslant 3$. If $\mathrm{diam}(T) = 3$, the claim follows immediately. Hence in what follows we consider the trees whose diameter is larger than 3. Suppose that $P_r = v_1 \dots v_r$ $(r \geqslant 5)$ is one of the longest paths in T, we are to show that $d_T(v_3) = d_T(v_4) = \cdots = d_T(v_{r-2}) = 2$ and $r = 2p+1$. First assume to the contrary that there exists a vertex $v \in \{v_3, v_4, \dots, v_{r-2}\}$ such that $d_T(v) \geqslant 3$. Let

$$i = \min\{j : \ d_T(v_j) \geqslant 3, \quad 3 \leqslant j \leqslant r-2\}, \qquad N_T(v_i) = \{v_{i-1}, v_{i+1}, z_1, z_2, \dots, z_s\}, \quad s \geqslant 1.$$

Let T_0 be the component that contains v_i in $T - \{z_1, z_2, \dots, z_s\}$. By Lemma 9, there exists $v_t \in V_{P_r}$ such that

$$f_{T_0}(v_1) < \cdots < f_{T_0}(v_{t-1}) < f_{T_0}(v_t) \geqslant f_{T_0}(v_{t+1}) > \cdots > f_{T_0}(v_r).$$

If $t < i$, then we have $f_{T_0}(v_i) > f_{T_0}(v_{r-1}) > f_{T_0}(v_r)$. If v_i and v_{r-1} are in the same part, by Lemma 7, we have

$$F(T) > F(T'), \tag{16}$$

where

$$T' = T - v_i z_1 - \cdots - v_i z_s + v_{r-1} z_1 + \cdots + v_{r-1} z_s, \qquad T' \in \mathscr{P}_n^{p,q}$$

otherwise, v_i and v_r are in the same part, we have

$$F(T) > F(T''), \tag{17}$$

where

$$T'' = T - v_i z_1 - \cdots - v_i z_s + v_r z_1 + \cdots + v_r z_s, \qquad T'' \in \mathscr{P}_n^{p,q}.$$

If $t \geqslant i$, repeat the procedure as above, we have a $T''' \in \mathscr{P}_n^{p,q}$ such that

$$F(T) > F(T'''), \qquad T''' \in \mathscr{P}_n^{p,q}. \tag{18}$$

Hence, (16)-(18) are contradictions to the choice of T. So we have $T \cong P_{r-2}(a,b)$.

Notice that $T \in \mathscr{P}_n^{p,q}$ with $1 < p < \lfloor \frac{n}{2} \rfloor$, it is easy to see that $r \leqslant 2p+1$. If $r < 2p+1$, it means that v_1 and v_r are in different parts (otherwise we have $p < \lceil \frac{r-2}{2} \rceil$ or $q < \lceil \frac{r-2}{2} \rceil$). As $a \geqslant b$, we have $v_1 \in V_2$ and $v_r \in V_1$, where V_1 and V_2 are two parts of V_T with $|V_1| = p, |V_2| = q$. Assume that $N_T(v_2) = \{v_3, v_1, w_2, \dots, w_a\}, N_T(v_{r-1}) = \{v_{r-2}, v_r, u_2, \dots, u_b\}$.

Let $\hat{T} = T - \{v_r, u_2, \dots, u_b\}$. By Lemma 6, we have $f_{\hat{T}}(v_{r-1}) < f_{\hat{T}}(v_1)$. Hence, in view of Lemma 7 we have

$$F(T) > F(\tilde{T}), \tag{19}$$

where $\tilde{T} = T - v_{r-1}v_r - v_{r-1}u_2 - \cdots - v_r u_b + v_1 v_r + v_1 u_2 + \cdots + v_1 u_b$. Note that $v_{r-1}, v_1 \in V_2$, $\tilde{T} \in \mathscr{P}_n^{p,q}$, hence (19) is a contradiction to the choice of T. So we have $T \cong P_{2p-1}(a, b)$. $\square$

Combining with Claims 12 and 13, we have $T \cong P_{2p-1}(\lfloor \frac{n-2p+1}{2} \rfloor, \lceil \frac{n-2p+1}{2} \rceil)$, as desired. $\square$

Remark 14. By direct calculation, we have

$$F(D(p, q)) = 2^{n-2} + 2^{p-1} + 2^{q-1} + n - 2.$$

This gives

$$F(D(p, q)) - F(D(p - 1, q + 1)) = 2^{p-2} - 2^{q-1} < 0$$

for $1 < p \leqslant q$. Hence, we have

$$F(D(p, q)) < F(D(p - 1, q + 1)) < \cdots < F(D(1, n - 1)) = F(K_{1,n-1}), \tag{20}$$

for $1 < p \leqslant q$. Note that $D(p, q)$ maximizes the total number of subtrees among $\mathscr{P}_n^{p,q}$, hence in view of (20) and first part of Theorem 2, the following corollary holds immediately.

Corollary 15 ([10]). *The star $K_{1,n-1}$ has $2^{n-1} + n - 1$ subtrees, more than any other tree on n vertices.*

4 Proof of Theorem 3

In this section we determine the extremal tree which minimizes the the total number of subtrees among $\mathscr{A}_n^q$.

Proof of Theorem 3 In order to characterize the structure of the tree, say T, minimizing the total number of subtrees in $\mathscr{A}_n^q$, it suffices to show that the diameter of T is $n + 1$. Without loss of generality, we assume one of the longest paths in T is $P_{r+1} = v_0 v_1 \dots v_r$. If $r = n + 1$, our result holds obviously. So in what follows, we assume that $r \leqslant n$.

For convenience, let T_i be the component that contains v_i in $T - E_{P_{r+1}}$ for $i = 1, 2, \dots r - 1$. Set

$$l := \min\{i : \ 1 \leqslant i \leqslant r - 1, T_i \ncong K_{1,q-2}\}.$$

We can see that there exists $j \in \{1, 2, \dots, q - 2\}$ such that $d_T(v_{lj}) > 1$, i.e., $d_T(v_{lj}) = q$. Note that P_{r+1} is a longest path in T, hence

$$1 < l < r - 1.$$

Thus, we can partition T into two subtrees, say S and T_0, such that $E_T = E_S \cup E_{T_0}, V_T = V_S \cup V_{T_0}$ and $V_S \cap V_{T_0} = \{v_{l_j}\}$; see Fig. 6. For convenience, let $N_{T_0}(v_{l_j}) = \{w_1, w_2, \dots, w_{q-1}\}$.

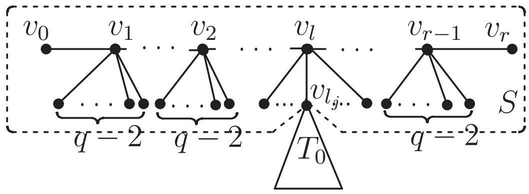

Figure 6: A q-arc tree T.

Now we are in the position to apply Lemma 6 in the following setting:

$$x \leftarrow v_0,\ x_1 \leftarrow v_1, \ldots, y_1 \leftarrow v_l, y \leftarrow v_{l_j}.$$

It is easy to see that

$$f_{X_i}(x_i) = f_{Y_i}(y_i) = 2^{q-2}, \qquad i = 2, \ldots, \lfloor l/2 \rfloor$$

and $f_{Y_1}(y_1) > 2^{q-2} = f_{X_1}(x_1)$. Therefore, by Lemma 6, we have $f_S(x) < f_S(y)$, where S is defined as above. By Lemma 7, we have

$$F(T') < F(T), \tag{21}$$

where $T' = T - \{v_{l_j}w_1, v_{l_j}w_2, \ldots, v_{l_j}w_{q-1}\} + \{v_0w_1, v_0w_2, \ldots, v_0w_{q-1}\}$. Inequality (21) is a contradiction to the choice of T. Hence, we obtain $r = n+1$, i.e., $T \cong \hat{T}_n^q$, as desired. □

Remark 16. In particular, let $q = 3$ in Theorem 3, we can obtain that just the n-leaf binary caterpillar tree minimizes the total number of subtrees among n-leaf binary trees, which is obtained by Székely and Wang [10].

Corollary 17 ([10])**.** *For any $n \geqslant 2$, precisely the n-leaf binary caterpillar tree $\hat{T}^3_{n-2}$ minimizes the number of subtrees among n-leaf binary trees.*

5 Proof of Theorems 4 and 5

In this section we determine the extremal n-vertex tree with domination number γ maximizing the total number of subtrees. Furthermore, the extremal n-vertex tree with domination number $\frac{n}{2}$ minimizing the total number of subtrees is also characterized.

Proof of Theorem 4 It is known from [17] that $\gamma(G) \leqslant \beta(G)$, where $\beta(G)$ is the matching number of G. In what follows we are to show: If $T_0 \in \mathscr{D}_n^\gamma$ maximizes the total number of subtrees, then $\gamma(T_0) = \beta(T_0)$.

In fact, it suffices to show that $\gamma(T_0) \geqslant \beta(T_0)$. Otherwise, by the definition of the set $\mathscr{D}_n^\gamma$, we have $\beta(T_0) > \gamma(T_0) = \gamma$. Assume that $S = \{v_1, v_2, \ldots, v_\gamma\}$ is a dominating set of cardinality γ. Then there exist γ independent edges $v_1v_1', v_2v_2', \ldots, v_\gamma v_\gamma'$ in T_0. Note that $\beta(T_0) > \gamma(T_0) = \gamma$, there must exist another edge, say w_1w_2, which is independent of each of edges $v_iv_i',\ i = 1, 2, \ldots, \gamma$.

Figure 7: The structures of T_0 and T'_0 in the proof of Theorem 4.

If the two vertices w_1, w_2 are dominated by the same vertex $v_i \in S$, then a triangle $C_3 = w_1w_2v_i$ occurs. This is impossible because of the fact that T_0 is a tree. Therefore w_1, w_2 are dominated by two different vertices from S. Without loss of generality, assume that w_i is dominated by the vertex v_i for $i = 1, 2$ (see Fig. 7). Now we construct a new tree $T'_0 \in \mathscr{D}_n^\gamma$ by A-transformation of T_0 on the edges v_1w_1 and v_2w_2, respectively. By Lemma 10, we have $F(T_0) < F(T'_0)$, a contradiction. Thus, Theorem 4 follows immediately from Lemma 8.

This completes the proof. $\square$

Proof of Theorem 5 It is known [3, 16] that if $n = 2\gamma$, then a tree T belongs to $\mathscr{D}_n^\gamma$ if and only if there exists a tree H on $\gamma = \frac{n}{2}$ vertices such that $T = H \circ K_1$, where $H \circ K_1$ is obtained by attaching a leaf to each vertex of H. Hence in what follows, we are to show: For any tree T, one has $F(T \circ K_1) \geqslant F(P_{|V_T|} \circ K_1)$ with equality if and only if $T \cong P_{|V_T|}$.

In fact, let $\mathscr{R}(T)$ be the set of all the subtrees of tree T. For any u in V_T and $1 \leqslant m \leqslant |V_T|$, let $\mathscr{R}^m(T;u)$ denote the set of all m-vertex subtrees of a tree T each of which contains u. It is routine to check that

$$\begin{aligned} F(T \circ K_1) &= \sum_{T_1 \in \mathscr{R}(T)} 2^{|V_{T_1}|} + |V_T| \qquad (22) \\ &= \sum_{T_1 \in \mathscr{R}(T-u)} 2^{|V_{T_1}|} + \sum_{T_1 \in \mathscr{R}(T;u)} 2^{|V_{T_1}|} + |V_T| \\ &= \sum_{T_1 \in \mathscr{R}(T-u)} 2^{|V_{T_1}|} + \sum_{m=1}^{|V_T|} |\mathscr{R}^m(T;u)| 2^m + |V_T|. \qquad (23) \end{aligned}$$

Assume that $T \ncong P_{|V_T|}$. If $|V_T| = 2$ or 3, our result is clearly true. If $|V_T| = 4$, there exist only two trees, i.e., P_4 and $K_{1,3}$, hence $T = K_{1,3}$. In this case, for any $u \in PV(T)$ we have

$$|\mathscr{R}^1(T;u)| = |\mathscr{R}^2(T;u)| = |\mathscr{R}^4(T;u)| = 1, |\mathscr{R}^3(T;u)| = 2 \qquad (24)$$

And for any $v \in PV(P_4)$, we have

$$|\mathscr{R}^1(P_4;v)| = |\mathscr{R}^2(P_4;v)| = |\mathscr{R}^3(P_4;v)| = |\mathscr{R}^4(P_4;u)| = 1. \qquad (25)$$

Note that $P_4 - u = K_{1,3} - v$, hence by (23)-(25) we have $F(K_{1,3} \circ K_1) > F(P_4 \circ K_1)$.

In what follows we assume that $F(T \circ K_1) > F(P_{|V_T|} \circ K_1)$ holds for all trees of order less than $|V_T|$. On the one hand, for any $u \in PV(T)$ and $v \in PV(P_{|V_T|})$, we have

$$F((T-u) \circ K_1) \geqslant F((P_{|V_T|} - v) \circ K_1), \tag{26}$$

Each of the equalities in (26) holds if and only if $T - u \cong P_{|V_T|} - v$. Hence by (22), we have

$$\sum_{T_1 \in \mathscr{R}(T-u)} 2^{|V_{T_1}|} \geqslant \sum_{T_1 \in \mathscr{R}(P_{|V_T|}-v)} 2^{|V_{T_1}|}. \tag{27}$$

On the other hand, it is easy to see that for any $w \in PV(T) \setminus \{u\}$, $T - w \in \mathscr{D}^{|V_T|-1}(T; u)$, so we have

$$|\mathscr{R}^{|V_T|-1}(T; u)| > 1 = |\mathscr{R}^{|V_T|-1}(P_{|V_T|}; v)|. \tag{28}$$

Furthermore, for $m = 1, 2, \ldots, |V_T| - 2, |V_T|$,

$$|\mathscr{R}^m(T; u)| \geqslant 1 = |\mathscr{R}^m(P_{|V_T|}; v)|. \tag{29}$$

Hence, $F(T \circ K_1) > F(P_{|V_T|} \circ K_1)$ follows by (23) and (27)-(29) for $T \ncong P_{|V_T|}$.

This completes the proof. □

Acknowledgments

The authors would like to express their sincere gratitude to the referee for his/her careful reading of the paper and for all the insightful comments and valuable suggestions, which considerably improved the presentation of this paper.

References

[1] J.A. Bondy, U.S.R. Murty, Graphs Theroy. Springer, New York (2008)

[2] A.A. Dobrynin, R. Entringer, I. Gutman, Wiener index of trees: theory and applications, Acta Appl. Math. 66 (3) (2001) 211-249.

[3] J.F. Fink, M.S. Jocobson, L.F. Kinch, J. Roberts, On graphs having domination number half their order, Period. Math. Hungar. 16 (1985) 287-293.

[4] C.X. He, B.F. Wu, Z.S. Yu, On the energy of trees with given domination number, MATCH Commun. Math. Comput. Chem. 64 (1) (2010) 169-180.

[5] R. Kirk, H. Wang, Largest number of subtrees of trees with a given maximum degree, SIAM J. Discrete Math. 22 (3) (2008) 985-995.

[6] C.W. Song, Counting special families of labelled trees, Ann. Comb. 10 (2) (2006) 271-283.

[7] L.A. Székely, H. Wang, Binary trees with the largest number of subtrees, Discrete Appl. Math. 155 (3) (2007) 374-385.

[8] L.A. Székely, H. Wang, On subtrees of trees, 2004 Industrial Mathematics Institute Research Reports 04:04, University of South Carolina, http://www.math.sc.edu/~imip/04.html,2004.

[9] L.A. Székely, H.Wang, Binary trees with the largest number of subtrees with at least one leaf, Congr. Numer. 177 (2005) 147-169.

[10] L.A. Székely, H. Wang, On subtrees of trees, Adv. Appl. Math. 34 (1) (2005) 138-155.

[11] L. A. Székely, H. Wang, T.Y. Wu, The sum of the distances between the leaves of a tree and the 'semi-regular' property, Discrete Math. 311 (13) (2011) 1197-1203.

[12] A. Vince, H. Wang, The average order of a subtree of a tree, J. Combin. Theory Ser. B 100 (2) (2010) 161-170.

[13] S. Wagner, Correlation of graph-theoretical indices, SIAM J. Discrete Math. 21 (1) (2007) 33-46.

[14] H. Wang, The extremal values of the Wiener index of a tree with given degree sequence, Discrete Appl. Math. 156 (14) (2008) 2647-2654.

[15] H. Wang, Some results on trees, PhD Thesis, Department of Mathematics, University of South Carolina, 2005.

[16] B. Xu, E.J. Cockayne, T.W. Haynes, S.T. Hedetniemi, S.C. Zhou, Extremal graphs for inequalities involving domination parameters, Discrete Math. 216 (2000) 1-10.

[17] K.X. Xu, L.H. Feng, Extremal energies of trees with a given domination number, Linear Algebra Appl. 435 (2011) 2382-2393.

[18] W.G. Yan, Y.N. Ye, Enumeration of subtrees of trees, Theoret. Comput. Sci. 369 (2006) 256-268.

[19] L.Z. Ye, R.S. Chen, Ordering of trees with a given bipartition by their energies and Hosoya indices, MATCH Commun. Math. Comput. Chem. 52 (2004) 193-208.

[20] X.D. Zhang, Y. Liu, M.X. Han, Maximum Wiener index of trees with given degree sequence, MATCH Commun. Math. Comput. Chem. 64 (3) (2010) 661-682.

[21] X.D. Zhang, Q.Y. Xiang, L.Q. Xu, R.Y. Pan, The Wiener index of trees with given degree sequences, MATCH Commun. Math. Comput. Chem. 60 (2) (2008) 623-644.

[22] X.M. Zhang, X.D. Zhang, Trees with given degree sequences that have minimal subtrees, arXiv:submit/0543782 [math.CO] 3 Sep 2012.

[23] X.M. Zhang, X.D. Zhang, D. Gray, and H. Wang, The number of subtrees of trees with given degree sequence, Journal of Graph Theory (2012): DOI 10.1002/jgt.

[24] X.M. Zhang, X.D. Zhang, D. Gray, and H. Wang, Trees with the most subtrees– an algorithmic approach, J. Comb. 3 (2) (2012) 207-223.