



# On the average order of a dominating set of a forest

Aysel Erey

Department of Mathematics, Gebze Technical University, Kocaeli, Turkey



## ARTICLE INFO

### Article history:

Received 1 April 2021  
 Received in revised form 31 July 2022  
 Accepted 6 August 2022  
 Available online xxxx

### Keywords:

Dominating set  
 Average order  
 Forest

## ABSTRACT

We show that the average order of a dominating set of a forest graph  $G$  on  $n$  vertices with no isolated vertices is at most  $2n/3$ . Moreover, the equality is achieved if and only if every non-leaf vertex of  $G$  is a support vertex with one or two leaf neighbors. Our result answers an open question of Beaton and Brown in [The average order of dominating sets of a graph, Discrete Math. 344(12) (2021)].

© 2022 Elsevier B.V. All rights reserved.

## 1. Introduction

In this article all graphs are finite, simple, loopless and undirected. Given a graph  $G$ , let  $V(G)$  and  $E(G)$  be the vertex set and the edge set of  $G$ , respectively. The *order* of  $G$  is  $|V(G)|$  and the *size* of  $G$  is  $|E(G)|$ . A vertex  $u$  is a *neighbor* of vertex  $v$  in  $G$  if  $u$  and  $v$  are adjacent in  $G$ . The *open neighborhood* of  $v$ ,  $N_G(v)$ , consists of all neighbors of  $v$  in  $G$ , and the *closed neighborhood* of  $v$ ,  $N_G[v]$ , is equal to  $N_G(v) \cup \{v\}$ . A vertex  $v$  of  $G$  is called an *isolated vertex* of  $G$  if  $v$  has no neighbors in  $G$ . For a subset  $S$  of vertices of  $G$ , let  $G \setminus S$  denote the subgraph induced by the vertices in  $V(G) \setminus S$  (if  $S = \{v\}$  is a singleton, we simply write  $G \setminus v$ ). If  $H$  is a subgraph of  $G$ , we write  $G \setminus H$  for the subgraph induced by  $V(G) \setminus V(H)$  in  $G$ . Also, let  $G/u$  be the graph obtained from  $G$  by deleting the vertex  $u$  and adding edges between all pairs of nonadjacent neighbors of  $u$ . The complete graph, empty graph and star graph on  $n$  vertices are denoted by  $K_n$ ,  $\bar{K}_n$  and  $K_{1,n-1}$ , respectively. A vertex of degree one is called a *leaf* and an edge containing a leaf vertex is called a *pendant edge*. Let  $L_G(u)$  denote the set of all leaf neighbors of a vertex  $u$  in  $G$  and  $L_G[u] = L_G(u) \cup \{u\}$ . We say that  $u$  is a *support vertex* of  $G$  if  $u$  is adjacent to a leaf vertex  $v$  of  $G$ , and  $u$  is called the *support of  $v$*  in  $G$ . An acyclic graph is called a *forest* and an acyclic connected graph is called a *tree*.

Average values of various graph invariants have been studied in the literature. In 1971, Doyle and Graver [12] initiated the study of mean (average) distance in a graph which received a considerable attention. They gave a formula for computing the mean distance of trees, and determined extremal graphs with maximum or minimum mean distance among connected graphs of a given order [12]. Extremal problems for the mean distance were also examined in certain other classes of graphs [18], [13]. Chung [7] showed that the independence number is an upper bound on the mean distance and Dankelmann [9] proved upper bounds for it in terms of the so called  $k$ -packing number. Average eccentricity (a distance related notion), was recently studied in [10]. Average order of a subtree of a graph was introduced by Jamison [14] in 1983 and this invariant has been investigated by a number of researchers, see, for example, [6,16,19,20]. Moreover, the average connectivity of a graph was considered in [5]. Most recently, Andriantiana et al. [2,3] studied the average sizes of independent sets and matchings.

E-mail address: aysel.erey@gtu.edu.tr.

Lastly, the average distance [8] and the average size of independent sets [11] were also studied in the context of random graphs.

Our focus in this paper will be on the average order of a dominating set of a graph. A subset of vertices  $S$  is called a *dominating set* of  $G$  if every vertex in  $V(G) \setminus S$  is adjacent to some vertex in  $S$ . Let  $\mathcal{D}(G)$  denote the family of all dominating sets of  $G$ . Recently Beaton and Brown [4] introduced the *average order of a dominating set* of  $G$ , denoted  $\text{avd}(G)$ , which is given by

$$\text{avd}(G) = \frac{\sum_{S \in \mathcal{D}(G)} |S|}{|\mathcal{D}(G)|}.$$

They showed that the complete graph  $K_n$  uniquely minimizes the average order of a dominating set among all graphs on  $n$  vertices [4]. It is trivial that  $\text{avd}(\overline{K}_n) = n \geq \text{avd}(G)$  for every graph  $G$  on  $n$  vertices. So, the empty graph  $\overline{K}_n$  has the largest average order of a dominating set among graphs on  $n$  vertices. What if we do not allow isolated vertices? Which graphs have the largest average order of a dominating set among all graphs of order  $n$  without isolated vertices? While this question was examined in [4], the question remained unanswered in general. It was shown that  $\text{avd}(G) \leq \frac{3}{4}n$  for every graph  $G$  of order  $n$  without isolated vertices [4]. However the factor  $3/4$  in the upper bound is not best possible, and the evidence provided in [4] suggests that the constant  $3/4$  can be improved to a smaller number. Indeed, the following conjecture was proposed.

**Conjecture 1.1.** [4] *If  $G$  is a graph of order  $n$  with no isolated vertices, then  $\text{avd}(G) \leq \frac{2n}{3}$ .*

Conjecture 1.1 was verified for all graphs up to 9 vertices, all graphs with minimum degree at least 4, and all quasi-regularizable graphs (which include all graphs containing a perfect matching or a hamiltonian cycle) [4]. Beaton and Brown [4] also studied such extremal problems within the family of trees. They showed that for every tree graph  $G$  of order  $n$  with  $G \not\cong K_{1,n-1}$ ,

$$\text{avd}(G) > \frac{n - 1 + 2^{n-2}(n + 1)}{2^{n-1} + 1} = \text{avd}(K_{1,n-1})$$

and hence the star graph  $K_{1,n-1}$  is the unique extremal graph with minimum average order of a dominating set. On the other hand, the problem of determining extremal graphs maximizing this parameter among trees remained as an open problem. In this article, we settle this problem by proving the following:

**Theorem 1.2.** *If  $G$  is a tree of order  $n \geq 2$ , then  $\text{avd}(G) \leq \frac{2n}{3}$ . Moreover, the equality holds if and only if every non-leaf vertex of  $G$  is a support vertex with one or two leaf neighbors.*

We actually prove Conjecture 1.1 for forests, see our Theorem 3.4 which immediately implies Theorem 1.2.

## 2. Preliminaries

The *domination number*  $\gamma(G)$  of a graph  $G$  is the cardinality of a minimum dominating set of  $G$ . Let  $d_k(G)$  be the number of dominating sets of  $G$  with cardinality  $k$ . The *domination polynomial* of  $G$ , denoted by  $D_G(x)$ , is given by

$$D_G(x) = \sum_{k=\gamma(G)}^{|V(G)|} d_k(G) x^k.$$

Observe that  $\text{avd}(G)$  is equal to the logarithmic derivative of  $D_G(x)$  evaluated at  $x = 1$ , that is,

$$\text{avd}(G) = \frac{D'_G(1)}{D_G(1)}.$$

Let  $H_1, \dots, H_c$  be the connected components of  $G$ . It is well known that

$$D_G(x) = \prod_{i=1}^c D_{H_i}(x),$$

and moreover, it was observed in [4] that

$$\text{avd}(G) = \sum_{i=1}^c \text{avd}(H_i).$$

Let  $V(G) = \{v_1, \dots, v_n\}$  and let  $G(v_1^{k_1}, v_2^{k_2}, \dots, v_n^{k_n})$  be the graph obtained from  $G$  by adding  $k_i$  leaves to each vertex  $v_i$  for  $i \in \{1, \dots, n\}$ . For  $k_1, \dots, k_n > 0$ , it was observed in [1] and [17] that

$$D_{G(v_1^{k_1}, v_2^{k_2}, \dots, v_n^{k_n})}(x) = \prod_{i=1}^n D_{K_{1,k_i}}(x).$$

Thus,  $\text{avd}(G(v_1^{k_1}, v_2^{k_2}, \dots, v_n^{k_n})) = \sum_{i=1}^n \text{avd}(K_{1,k_i})$ . In particular, if  $G'$  is obtained from  $G$  by adding one leaf to each of  $p$  vertices of  $G$  and two leaves to each of  $q$  vertices of  $G$ , where  $|V(G)| = p + q$ , then

$$\text{avd}(G') = p \text{avd}(K_{1,1}) + q \text{avd}(K_{1,2}) = \frac{4}{3}p + 2q = \frac{2}{3}(2p + 3q) = \frac{2}{3}|V(G')|.$$

Thus, we see that the extremal graphs mentioned in Theorem 1.2 indeed achieve the given upper bound.

We will also make use of the following recursive formula for the domination polynomials of graphs containing vertices with nested closed neighborhoods.

**Lemma 2.1.** [15] *Let  $u$  and  $v$  be two vertices of  $G$  such that  $N_G[v] \subseteq N_G[u]$ . Then,*

$$D_G(x) = xD_{G/u}(x) + D_{G \setminus u}(x) + xD_{G \setminus N_G[u]}(x).$$

*In particular, if  $v$  is a leaf vertex and  $u$  is its neighbor, then*

$$D_G(x) = x[D_{G/u}(x) + D_{G \setminus \{u,v\}}(x) + D_{G \setminus N_G[u]}(x)].$$

Lastly, observe that if  $H$  is a subgraph of a graph  $G$ , then  $D_H(1) \leq D_G(1)$  because the function  $f : \mathcal{D}(H) \rightarrow \mathcal{D}(G)$  defined by  $f(S) = S \cup (V(G) \setminus V(H))$  is clearly one-to-one.

### 3. Main result

In the proof of the following result, let  $d_k(G, u)$  (respectively  $d_k(G, \bar{u})$ ) denote the number of dominating sets of  $G$  of order  $k$  which contain the vertex  $u$  (respectively do not contain the vertex  $u$ ). Clearly,  $d_k(G) = d_k(G, u) + d_k(G, \bar{u})$  for every vertex  $u$ .

**Lemma 3.1.** *Let  $G$  be a graph of order  $n$  and  $w$  be a support vertex of  $G$  with  $L_G(w) = \{v_1, \dots, v_t\}$  for some integer  $t \geq 1$ . Let also  $H = G \setminus L_G[w]$ . Suppose that every vertex  $u$  in  $N_G(w) \setminus L_G(w)$  is a support vertex in  $G$ , and  $3D'_H(1) \leq 2(n - t - 1)D_H(1)$ . Then,*

$$3D'_G(1) \leq 2nD_G(1)$$

*with equality if and only if  $t \in \{1, 2\}$  and  $3D'_H(1) = 2(n - t - 1)D_H(1)$ .*

**Proof.** Let  $S$  be a dominating set of  $G$  of order  $k$ . If  $v_t \notin S$ , then  $w$  must be in  $S$ , as  $v_t$  is a leaf. Since every vertex in  $N_G(w) \setminus L_G(w)$  is a support vertex of  $G$ , the vertex subset  $S \cap V(H)$  must be a dominating set of  $H$  on  $k - 1 - i$  vertices where  $i$  is the number of vertices in  $S \cap \{v_1, \dots, v_{t-1}\}$ . Therefore,

$$d_k(G, \bar{v}_t) = \sum_{i=0}^{t-1} \binom{t-1}{i} d_{k-1-i}(H).$$

Similarly, one can check that the number of dominating sets of  $G$  of order  $k$  which contain  $v_t$  but not  $w$  is  $d_{k-t}(H)$ , and the number of dominating sets which contain both  $v_t$  and  $w$  is  $\sum_{i=0}^{t-1} \binom{t-1}{i} d_{k-2-i}(H)$ . Hence,

$$d_k(G, v_t) = d_{k-t}(H) + \sum_{i=0}^{t-1} \binom{t-1}{i} d_{k-2-i}(H).$$

Since  $d_k(G) = d_k(G, v_t) + d_k(G, \bar{v}_t)$ , we have

$$d_k(G) = d_{k-t}(H) + \sum_{i=0}^{t-1} \binom{t-1}{i} d_{k-2-i}(H) + \sum_{i=0}^{t-1} \binom{t-1}{i} d_{k-1-i}(H).$$

By rewriting the latter as a polynomial equation and applying the binomial theorem, we get

$$\begin{aligned} D_G(x) &= x^t D_H(x) + \sum_{i=0}^{t-1} \binom{t-1}{i} x^{i+2} D_H(x) + \sum_{i=0}^{t-1} \binom{t-1}{i} x^{i+1} D_H(x) \\ &= x^t D_H(x) + (x^2 + x) D_H(x) \sum_{i=0}^{t-1} \binom{t-1}{i} x^i \\ &= x^t D_H(x) + (x^2 + x) D_H(x) (x + 1)^{t-1} \\ &= [x^t + x(x + 1)^t] D_H(x). \end{aligned}$$

Since  $D_G(x) = [x^t + x(x + 1)^t] D_H(x)$ , we obtain that

$$D'_G(x) = [x^t + x(x + 1)^t] D'_H(x) + [t x^{t-1} + (x + 1)^t + x t (x + 1)^{t-1}] D_H(x).$$

Evaluations of  $D_G(x)$  and  $D'_G(x)$  at  $x = 1$  yield

$$D_G(1) = (1 + 2^t) D_H(1)$$

and

$$D'_G(1) = (1 + 2^t) D'_H(1) + (t + 2^t + t 2^{t-1}) D_H(1).$$

Now, we have

$$\begin{aligned} 2nD_G(1) - 3D'_G(1) &= 2n(1 + 2^t) D_H(1) - 3(1 + 2^t) D'_H(1) - 3(t + 2^t + t 2^{t-1}) D_H(1) \\ &\geq 2n(1 + 2^t) D_H(1) - 2(1 + 2^t)(n - t - 1) D_H(1) - 3(t + 2^t + t 2^{t-1}) D_H(1) \\ &= (2 - t + t 2^{t-1} - 2^t) D_H(1) \\ &= (t - 2)(2^{t-1} - 1) D_H(1) \\ &\geq 0 \end{aligned}$$

where the first inequality holds by the assumption that  $3D'_H(1) \leq 2(n - t - 1)D_H(1)$ , and the last inequality is an equality if and only if  $t \in \{1, 2\}$ . Thus,  $3D'_G(1) \leq 2nD_G(1)$  with equality if and only if  $t \in \{1, 2\}$  and  $3D'_H(1) = 2(n - t - 1)D_H(1)$ .  $\square$

Given a graph  $G$  with a specified vertex  $u \in V(G)$ , we write  $G_{(u,k)}$  to denote the graph obtained by gluing  $G$  and  $K_{k+1}$  at the vertex  $u$ . That is,  $G_{(u,k)} = K_{k+1} \cup G$  and  $K_{k+1} \cap G = \{u\}$ .

**Lemma 3.2.** *Let  $G$  be a graph and  $u \in V(G)$ . Then, for every integer  $k \geq 1$ ,*

$$D_{G_{(u,k)}}(x) = (x + 1)^{k-1} [D_{G_{(u,1)}}(x) + D_{G \setminus u}(x)] - D_{G \setminus u}(x).$$

**Proof.** We proceed by induction on  $k$ . The statement is clear for  $k = 1$ . Suppose that  $k \geq 2$  and let  $V(G_{(u,k)}) \setminus V(G) = \{v_1, \dots, v_k\}$  where the vertices  $v_1, \dots, v_k$  induce a  $k$ -clique in  $G_{(u,k)}$  and  $u$  is adjacent to each  $v_i$  for  $1 \leq i \leq k$ . Since  $v_k$  and  $v_{k-1}$  have the same closed neighborhoods in  $G_{(u,k)}$ , we obtain

$$\begin{aligned} D_{G_{(u,k)}}(x) &= x D_{G_{(u,k)}/v_k}(x) + D_{G_{(u,k)} \setminus v_k}(x) + x D_{G_{(u,k)} \setminus N_{G_{(u,k)}}[v_k]}(x) \\ &= x D_{G_{(u,k-1)}}(x) + D_{G_{(u,k-1)}}(x) + x D_{G \setminus u}(x) \\ &= (x + 1) D_{G_{(u,k-1)}}(x) + x D_{G \setminus u}(x) \\ &= (x + 1) \left[ (x + 1)^{k-2} (D_{G_{(u,1)}}(x) + D_{G \setminus u}(x)) - D_{G \setminus u}(x) \right] + x D_{G \setminus u}(x) \\ &= (x + 1)^{k-1} [D_{G_{(u,1)}}(x) + D_{G \setminus u}(x)] - D_{G \setminus u}(x). \quad \square \end{aligned}$$

**Lemma 3.3.** *Let  $T$  be a tree with  $|V(T)| \geq 3$  and  $u$  be a vertex of  $T$ . Suppose that  $u$  is not a support vertex of  $T$  and  $u$  has at most one neighbor in  $T$  which is not a support vertex in  $T$ . Let  $G_1$  be the graph obtained from  $T$  by attaching a new leaf vertex  $v$  at  $u$ . Then,*

- (i)  $D_{G_1}(1) \leq D_T(1) + 3D_{T \setminus u}(1)$ ,
- (ii)  $D_{G_1}(1) \leq 5D_{T \setminus u}(1)$  and
- (iii)  $D_T(1) \leq 3D_{T \setminus u}(1)$ .

**Proof.** Given a graph  $G$  containing  $T \setminus u$  as a subgraph, let us first define  $\mathcal{A}_G$  and  $\mathcal{B}_G$  as follows:  $\mathcal{A}_G = \{S : S \in \mathcal{D}(G) \text{ and } S \cap V(T \setminus u) \in \mathcal{D}(T \setminus u)\}$  and  $\mathcal{B}_G = \mathcal{D}(G) \setminus \mathcal{A}_G$ . It is clear that  $D_G(1) = |\mathcal{A}_G| + |\mathcal{B}_G|$ . By adding  $v, u$  or both, one can extend every dominating set of  $T \setminus u$  to a dominating set of  $G_1$  in three different ways. So, we have  $|\mathcal{A}_{G_1}| = 3D_{T \setminus u}(1)$ . If all neighbors of  $u$  in  $T$  are support vertices in  $T$ , then  $\mathcal{B}_{G_1} = \emptyset$ , and the results in (i) and (ii) are clear as  $D_{G_1}(1) = 3D_{T \setminus u}(1)$  in this case. So, we may assume that  $u$  is adjacent to exactly one non-support vertex in  $T$ , say  $u'$ . Let  $T'$  be the component of  $T \setminus u$  which contains  $u'$  and let  $T^* = (T \setminus u) \setminus T'$ . Note that if  $S$  belongs to  $\mathcal{B}_{G_1}$ , then  $S \cap V(T^*)$  is a dominating set of  $T^*$  since every neighbor of  $u$  in  $T^*$  is a support vertex of  $T$ . Also,  $u$  must be in  $S$  and  $S$  contains no vertex of  $N_{T'}[u']$ . Now, the function  $f : \mathcal{B}_{G_1} \rightarrow \mathcal{D}(T)$  defined by

$$f(S) = \begin{cases} S & \text{if } v \notin S \\ (S \setminus \{u, v\}) \cup \{u'\} & \text{if } v \in S \end{cases}$$

is one-to-one. Hence,  $|\mathcal{B}_{G_1}| \leq D_T(1)$  and the result in (i) follows. Let  $\lambda(T' \setminus u')$  be the number of dominating sets of  $T' \setminus u'$  which do not contain any vertex from  $N_{T'}(u')$ . Observe that

$$|\mathcal{B}_{G_1}| = 2D_{T^*}(1)\lambda(T' \setminus u') \leq 2D_{T^*}(1)D_{T' \setminus u'}(1) \leq 2D_{T^*}(1)D_{T'}(1) = 2D_{T \setminus u}(1)$$

and therefore (ii) follows. Lastly,  $|\mathcal{A}_T| \leq 2D_{T \setminus u}(1)$  because adding the vertex  $u$  to a dominating set of  $T \setminus u$  may or may not yield a dominating set for  $T$ . Moreover, if  $S \in \mathcal{B}_T$ , then  $u$  must be in  $S$  and  $S$  contains no vertex of  $N_{T'}[u']$ . Now the function  $g : \mathcal{B}_T \rightarrow \mathcal{D}(T \setminus u)$  defined by  $g(S) = (S \setminus u) \cup \{u'\}$  is one-to-one and therefore we obtain that  $|\mathcal{B}_T| \leq D_{T \setminus u}(1)$ . Thus,  $D_T(1) = |\mathcal{A}_T| + |\mathcal{B}_T| \leq 3D_{T \setminus u}(1)$  and (iii) is established.  $\square$

We are now ready to prove our main result.

**Theorem 3.4.** *Let  $G$  be a forest on  $n$  vertices with no isolated vertices. Then,  $\text{avd}(G) \leq \frac{2n}{3}$  and, moreover, equality holds if and only if every non-leaf vertex of  $G$  is a support vertex with one or two leaf neighbors.*

**Proof.** We proceed by strong induction on the number of vertices. If  $n = 2$ , then  $G \cong K_2$  and it is clear that  $\text{avd}(K_2) = 4/3$ . We may assume that  $n \geq 3$ . First suppose that  $G$  is connected, that is  $G$  is a tree. Let  $G$  be a rooted tree and let  $v$  be a support vertex of maximum distance from the root of  $G$ . Also let  $L_G(v) = \{v_1, \dots, v_k\}$  for some  $k \geq 1$ . If  $V(G) = L_G[v]$ , then  $G \cong K_{1,k}$  and it is easy to see that  $\text{avd}(K_{1,k}) \leq 2|V(G)|/3$  with equality iff  $1 \leq k \leq 2$ . So we may assume that  $V(G) \neq L_G[v]$ . Let  $T = G \setminus L_G[v]$  and  $u \in V(T)$  be the parent of  $v$  in  $G$ . Note that  $N_G(v) = L_G(v) \cup \{u\}$ , as the chosen support vertex  $v$  is of maximum distance from the root of  $G$ . If  $u$  is a support vertex in  $G$ , then the result follows from Lemma 3.1 and the induction hypothesis. So we may assume that  $u$  has no leaf neighbors in  $G$ . Since  $u$  is neither a leaf nor a support vertex of  $G$ , we have  $|V(T)| \geq 3$ . Now we shall show that the strict inequality  $3D'_G(1) < 2nD_G(1)$  holds.

$$D_G(x) = x [D_{G/v}(x) + D_{G \setminus \{v, v_k\}}(x) + D_{G \setminus N_G[v]}(x)] \tag{1}$$

$$= x [D_{T(u,k)}(x) + x^{k-1}D_T(x) + D_{T \setminus u}(x)] \tag{2}$$

$$= x [(x+1)^{k-1}[D_{T(u,1)}(x) + D_{T \setminus u}(x)] - D_{T \setminus u}(x) + x^{k-1}D_T(x) + D_{T \setminus u}(x)] \tag{3}$$

$$= x(x+1)^{k-1} [D_{G \setminus L_G(v)}(x) + D_{T \setminus u}(x)] + x^k D_T(x) \tag{4}$$

where (1) follows from Lemma 2.1; (2) holds as  $G/v \cong T_{(u,k)}$ ,  $G \setminus \{v, v_k\} \cong \bar{K}_{k-1} \cup T$  and  $G \setminus N_G[v] \cong T \setminus u$ ; (3) follows from Lemma 3.2; and (4) follows since  $T_{(u,1)} \cong G \setminus L_G(v)$ . Let us write  $G_1 = G \setminus L_G(v)$ , then

$$D'_G(x) = [(x+1)^{k-1} + (k-1)x(x+1)^{k-2}] [D_{G_1}(x) + D_{T \setminus u}(x)] + x(x+1)^{k-1} [D'_{G_1}(x) + D'_{T \setminus u}(x)] + kx^{k-1}D_T(x) + x^k D'_T(x).$$

It follows that  $D_G(1) = 2^{k-1} [D_{G_1}(1) + D_{T \setminus u}(1)] + D_T(1)$  and

$$D'_G(1) = (k+1)2^{k-2} [D_{G_1}(1) + D_{T \setminus u}(1)] + 2^{k-1} [D'_{G_1}(1) + D'_{T \setminus u}(1)] + kD_T(1) + D'_T(1).$$

It is not difficult to calculate that  $2nD_G(1) - 3D'_G(1)$  is equal to

$$2^{k-1} [2(n-k)D_{G_1}(1) - 3D'_{G_1}(1)] + 2^{k-1} [2(n-k-2)D_{T \setminus u}(1) - 3D'_{T \setminus u}(1)] + [2(n-k-1)D_T(1) - 3D'_T(1)] + 2^{k-2} [(k-3)D_{G_1}(1) + (k+5)D_{T \setminus u}(1)] - (k-2)D_T(1).$$

Each of the subgraphs  $T$ ,  $G_1$  and  $T \setminus u$  is a proper subforest of  $G$  without isolated vertices. So, by the induction hypothesis, we have  $3D'_T(1) \leq 2(n-k-1)D_T(1)$ ,  $3D'_{G_1}(1) \leq 2(n-k)D_{G_1}(1)$  and  $3D'_{T \setminus u}(1) \leq 2(n-k-2)D_{T \setminus u}(1)$ . Therefore, in order to show that  $3D'_G(1) \leq 2nD_G(1)$  it suffices to prove only that

$$(k-2)D_T(1) \leq 2^{k-2} [(k-3)D_{G_1}(1) + (k+5)D_{T \setminus u}(1)]. \quad (5)$$

Observe that all but possibly one connected components of  $T \setminus u$  are star graphs because  $v$  is of maximum distance from the root of  $G$ . So, the vertex  $u$  has at most one neighbor in  $T$  which is not a support vertex of  $T$  and Lemma 3.3 applies here. If  $k=1$ , the inequality (5) is  $D_{G_1}(1) \leq D_T(1) + 3D_{T \setminus u}(1)$  and this follows from Lemma 3.3(i). If all neighbors of  $u$  in  $T$  are support vertices in  $T$ , then by the proof of Lemma 3.3(i), we have  $D_{G_1}(1) = 3D_{T \setminus u}(1)$  and therefore strict inequality holds in (5) which implies  $3D'_G(1) < 2nD_G(1)$ . If  $u$  has a neighbor in  $T$  which is not a support vertex of  $T$ , then by the induction hypothesis the strict inequality  $3D'_T(1) < 2(n-k-1)D_T(1)$  holds and again we get  $3D'_G(1) < 2nD_G(1)$ . For  $k \geq 2$ , we shall show that (5) holds strictly.

If  $k=2$ , the strict inequality for (5) is  $D_{G_1}(1) < 7D_{T \setminus u}(1)$  and this follows from Lemma 3.3(ii). If  $k=3$ , it is  $D_T(1) < 16D_{T \setminus u}$  and this is verified by Lemma 3.3(iii). If  $k \geq 4$ , we have  $(k-2) < 2^{k-2}(k-3)$  and  $D_T(1) \leq D_{G_1}(1)$  since  $T$  is a subgraph of  $G_1$ . Thus,  $3D'_G(1) < 2nD_G(1)$  is established for all  $k$ .

Lastly, suppose that  $G$  is a disconnected forest with connected components  $H_1, \dots, H_c$  where  $c \geq 2$ . By the induction hypothesis, for each  $i$  we have  $\text{avd}(H_i) \leq \frac{2|V(H_i)|}{3}$  with equality iff every non-leaf vertex of  $H_i$  is a support vertex with one or two leaf neighbors. Since  $\text{avd}(G) = \sum_{i=1}^c \text{avd}(H_i) \leq \sum_{i=1}^c \frac{2|V(H_i)|}{3} = 2n/3$ , the proof is completed.  $\square$

#### 4. Concluding remarks

Every graph  $G$  without isolated vertices contains a spanning forest  $F$  without isolated vertices, and  $F$  can be obtained from  $G$  by a succession of non-pendant edge removals. So it would be interesting to investigate how  $\text{avd}(G)$  is effected by the removal of a non-pendant edge. In particular we ask the following:

**Question 4.1.** In every graph  $G$  (which is not a disjoint union of stars or empty graphs) does there exist a non-pendant edge  $e$  of  $G$  such that  $\text{avd}(G) < \text{avd}(G \setminus e)$ ?

Observe that an affirmative answer to Question 4.1 would yield a proof of Conjecture 1.1 in general because of our Theorem 3.4 and the remark above. We also note that Beaton and Brown [4] conjectured that in every non-empty graph  $G$ , there exists an edge  $e$  such that  $\text{avd}(G) < \text{avd}(G \setminus e)$  and they verified this conjecture for graphs on up to 7 vertices. On the other hand, they did not specify any property of such an edge.

#### Declaration of competing interest

There is no conflict of interest.

#### Data availability

No data was used for the research described in the article.

#### Acknowledgements

I would like to thank two anonymous referees for their helpful comments.

#### References

- [1] S. Akbari, S. Alikhani, Y. Peng, Characterization of graphs using domination polynomials, *Eur. J. Comb.* 31 (2010) 1714–1724.
- [2] E.O.D. Andriantiana, V.R. Misanantenaina, S. Wagner, The average size of independent sets of graphs, *Eur. J. Math.* 6 (2020) 561–576.
- [3] E.O.D. Andriantiana, V.R. Misanantenaina, S. Wagner, The average size of matchings in graphs, *Graphs Comb.* 36 (3) (2020) 539–560.
- [4] I. Beaton, J.I. Brown, The average order of dominating sets of a graph, *Discrete Math.* 344 (12) (2021) 112595.
- [5] L.W. Beineke, O.R. Oellermann, R.E. Pippert, The average connectivity of a graph, *Discrete Math.* 252 (2002) 31–45.
- [6] B. Cameron, L. Mol, On the mean subtree order of graphs under edge addition, *J. Graph Theory* 96 (3) (2021) 403–413.
- [7] F. Chung, The average distance and the independence number, *J. Graph Theory* 12 (2) (1988) 229–235.
- [8] F. Chung, L. Lu, The average distance in a random graph with given expected degrees, *Internet Math.* 1 (1) (2003) 91–113.
- [9] P. Dankelmann, Average distance and generalised packing in graphs, *Discrete Math.* 310 (2010) 2334–2344.
- [10] P. Dankelmann, F.J. Osaye, Average eccentricity, k-packing and k-domination in graphs, *Discrete Math.* 342 (5) (2019) 1261–1274.
- [11] E. Davies, M. Jenssen, W. Perkins, B. Roberts, On the average size of independent sets in triangle-free graphs, *Proc. Am. Math. Soc.* 146 (2018) 111–124.
- [12] J.K. Doyle, J.E. Graver, Mean distance in a graph, *Discrete Math.* 17 (1977) 147–154.
- [13] G.R.T. Hendry, On mean distance in certain classes of graphs, *Networks* 19 (4) (1989) 451–457.
- [14] R.E. Jamison, On the average number of nodes in a subtree of a tree, *J. Comb. Theory, Ser. B* 35 (1983) 207–223.

- [15] T. Kotek, J. Preen, F. Simon, P. Tittman, M. Trinks, Recurrence relations and splitting formulas for the domination polynomial, *Electron. J. Comb.* 19 (3) (2012) P47.
- [16] L. Mol, O. Oellermann, Maximizing the mean subtree order, *J. Graph Theory* 91 (4) (2019) 326–352.
- [17] M.R. Oboudi, On the roots of domination polynomial of graphs, *Discrete Appl. Math.* 205 (2016) 126–131.
- [18] D. Rautenbach, Average distance and domination number revisited, *Discrete Appl. Math.* 159 (11) (2011) 1180–1182.
- [19] A. Vince, H. Wang, The average order of a subtree of a tree, *J. Comb. Theory, Ser. B* 100 (2) (2010) 161–170.
- [20] S. Wagner, H. Wang, On the local and global means of subtree orders, *J. Graph Theory* 81 (2) (2016) 154–166.