# The Average Order of Dominating Sets of a Graph

Iain Beaton

Department of Mathematics & Statistics Dalhousie University Halifax, CA

ibeaton@dal.ca

Jason I. Brown<sup>\*</sup>

Department of Mathematics & Statistics Dalhousie University Halifax, CA Jason.Brown@dal.ca

August 18, 2020

#### Abstract

This papers focuses on the average order of dominating sets of a graph. We find the extremal graphs for the maximum and minimum value over all graphs on n vertices, while for trees we prove that the star minimizes the average order of dominating sets. We prove the average order of dominating sets in graphs without isolated vertices is at most 3n/4, but provide evidence that the actual upper bound is 2n/3. Finally, we show that the normalized average, while dense in [1/2, 1], tends to  $\frac{1}{2}$  for almost all graphs.

### 1 Introduction

For a graph G containing vertex v,  $N_G(v) = \{u | uv \in E(G)\}$  denotes the open neighbourhood of v while  $N_G[v] = N(v) \bigcup \{v\}$  denotes the closed neighbourhood of v (we will omit the subscript G when only referring to one graph). For  $S \subseteq V(G)$ , the closed neighbourhood N[S] of S is simply the union of the closed neighbourhoods for each vertex in S. A subset of vertices S is a dominating set of G if N[S] = V(G), that is, every vertex is either in S or adjacent to a vertex in S. The domination number of G, denoted  $\gamma(G)$ , is the order of the smallest dominating set of G. The study of dominating sets in graphs is extensive (see, for example, [14]).

Let  $\mathcal{D}(G)$  denote the collection of dominating sets of G. Furthermore let  $d_k(G) = |\{S \in \mathcal{D}(G) : |S| = k\}|$ . Then the average order of dominating sets in G, denoted  $\operatorname{avd}(G)$ , is

$$\operatorname{avd}(G) = \frac{\sum_{k=\gamma(G)}^{|V(G)|} k d_k(G)}{\sum_{k=\gamma(G)}^{|V(G)|} d_k(G)},$$

that is, the average cardinality of a dominating set of G.

For graphs with few dominating sets  $\operatorname{avd}(G)$  is relatively easy to compute using the above formula. For example the empty graph  $\overline{K_n}$  has exactly one dominating set of order n, hence  $\operatorname{avd}(\overline{K_n}) = n$ . However if G has many dominating sets other techniques may be more appropriate to compute  $\operatorname{avd}(G)$ . The *domination polynomial* of Gis defined by

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

<sup>\*</sup>Supported by NSERC grant RGPIN 2018-05227

(see [2], for example, for for a discussion of domination polynomials). The average order of dominating sets in G can be regarded as the logarithmic derivative of D(G, x) evaluated at 1, that is,

$$\operatorname{avd}(G) = \frac{d}{dx} \ln(D(G, x)) \Big|_{x=1} = \frac{D'(G, 1)}{D(G, 1)}.$$
 (1)

This allows us to compute  $\operatorname{avd}(G)$  quickly when D(G, x) readily available. For example,

$$D(K_n, x) = (1+x)^n - 1$$
 and  $D(K_{1,n-1}, x) = x(1+x)^{n-1} + x^{n-1}$ ,

 $\mathbf{SO}$ 

$$\operatorname{avd}(K_n) = \frac{n2^{n-1}}{2^n - 1}$$
 and  $\operatorname{avd}(K_{1,n-1}) = \frac{(n+1)2^{n-2} + n - 1}{2^{n-1} + 1}$ .

It is trivial to observe that the domination polynomial is multiplicative over components, that is, for graphs G and H,  $D(G \cup H, x) = D(G, x)D(H, x)$  where  $G \cup H$  denotes the disjoint union of G and H. From this we can obtain a fundamental result which states that the average order of dominating sets is additive over components.

**Lemma 1.1.** Let G and H be graphs. Then  $\operatorname{avd}(G \cup H) = \operatorname{avd}(G) + \operatorname{avd}(H)$ .

Proof. As  $D(G \cup H, x) = D(G, x)D(H, x)$ , it follows that  $D'(G \cup H, x) = D'(G, x)D(H, x) + D(G, x)D'(H, x)$ . Therefore

$$\operatorname{avd}(G \cup H) = \frac{D'(G,1)D(H,1) + D(G,1)D'(H,1)}{D(G,1)D(H,1)} = \frac{D'(G,1)}{D(G,1)} + \frac{D'(H,1)}{D(H,1)} = \operatorname{avd}(G) + \operatorname{avd}(H).$$

Although the average order of dominating sets is a novel area of research, there has been work done on averages of serveral other graphs invariants:

- Closely related to the Weiner Index of a graph [21], the mean distance (between vertices) in a graph was introduced in 1977 by Doyle and Graver [11]. Doyle and Graver showed for connected graphs of order n (that is, with n vertices) the mean distance in a graph was maximized by a path, with mean distance (n + 1)/3, and minimized by the complete graph, with mean distance 1.
- The mean subtree order of a graph was introduced in 1983 by Jamison [15]. Jamison showed for any tree T on n vertices, the average number of vertices in a subtree of T is at least (n+2)/3, with that minimum achieved if and only if T is a path. As the mean subtree order of T is at most n, Jamison naturally defined the mean subtree order of T divided by n to be the density of T and showed there were trees whose density approached 1 as  $n \to \infty$ . Jamison conjectured the tree with maximum density was some caterpillar graph. Additionally, the mean subtree order has been subject to a fair amount of recent work [13, 16, 19, 20].
- The average size of an independent set in a graph was introduced in 2019 by Andriantiana et. el. [4]. Andriantiana et. el. showed the average number of vertices of an independent set in a graph was maximized by the empty graph and minimized by the complete graph. They also showed the average number of vertices of an independent set in a tree was maximized by  $P_n$  and minimized by  $K_{1,n-1}$ .

• In 2020, Andriantiana et. el. [5] introduced the average size of a matching in a graph was introduced. Andriantiana et. el. showed the average number of edges in a matching of a graph was now minimized by the empty graph and maximized by the complete graph. They also showed the average number of edges in a matching in a tree was maximized by  $P_n$  and minimized by  $K_{1,n-1}$ .

We remark that when the domination polynomial has all real roots, the average order of domination sets of graph G can also determine the mode of the coefficients of D(G, x); Darroch [10] showed in general that a positive sequence  $(a_0, a_1, \dots, a_n)$  has its mode at either  $\left\lfloor \frac{f'(1)}{f(1)} \right\rfloor$  or  $\left\lceil \frac{f'(1)}{f(1)} \right\rceil$ , where  $f(x) = a_0 + a_1x + \dots + a_nx^n$  is the associated generating polynomial. Therefore by (1), if D(G, x) has all real roots then it mode is at  $\lfloor \operatorname{avd}(G) \rfloor$  or  $\lceil \operatorname{avd}(G) \rceil$ .

This paper is structured as follows. In Section 2 we determine the extremal graphs for the average order of dominating sets of graphs of order n. In Section 3 we develop bounds for the average order of domination sets for connected graphs, as well as for trees. Section 4 introduces a normalized version of the parameter, describes the distribution of these parameters, and considers the values for Erdös-Renyi random graphs. Finally we conclude with some remarks.

### 2 Extremal Graphs

For a graph on n vertices, it is clear that  $\operatorname{avd}(G) \leq n$  as every dominating set has cardinality at most n. This bound is achieved by  $\overline{K_n}$ , and this graph is the unique extremal graph, as any other graph of order n has a dominating set of size smaller than n. On the other hand, what about the *minimum* value of  $\operatorname{avd}(G)$  over all graphs of order n? As you might expect, the complete graph  $K_n$  is the unique extremal graph in this case, but the argument will be more subtle, and that is what we shall pursue now.

We shall first need some technical results about the average cardinality of sets in collection of sets. Let X be a nonempty finite set and  $\mathcal{P}(X)$  its powerset. For any nonempty subset  $\mathcal{A} \subseteq \mathcal{P}(X)$  we define the average order of  $\mathcal{A}$ , denoted  $\operatorname{av}(\mathcal{A})$  to be

$$\operatorname{av}(\mathcal{A}) = \frac{1}{|\mathcal{A}|} \sum_{A \in \mathcal{A}} |A|.$$

**Lemma 2.1.** For a nonempty finite set X, let  $\mathcal{A} \subset \mathcal{B} \subseteq \mathcal{P}(X)$ . Then

 $\operatorname{av}(\mathcal{B}) \leq \operatorname{av}(\mathcal{A})$  if and only if  $\operatorname{av}(\mathcal{B} - \mathcal{A}) \leq \operatorname{av}(\mathcal{A})$ .

*Proof.* Set  $S(\mathcal{C}) = \sum_{C \in \mathcal{C}} |C|$  for any  $\mathcal{C} \subseteq \mathcal{P}(X)$ . Then

$$\begin{aligned} \operatorname{av}(\mathcal{B}) \leq \operatorname{av}(\mathcal{A}) \\ \Leftrightarrow & \frac{S(\mathcal{B})}{|\mathcal{B}|} \leq \frac{S(\mathcal{A})}{|\mathcal{A}|} \\ \Leftrightarrow & \frac{S(\mathcal{A}) + S(\mathcal{B} - \mathcal{A})}{|\mathcal{A}| + |\mathcal{B} - \mathcal{A}|} \leq \frac{S(\mathcal{A})}{|\mathcal{A}|} \\ \Leftrightarrow & S(\mathcal{A})|\mathcal{A}| + S(\mathcal{B} - \mathcal{A})|\mathcal{A}| \leq S(\mathcal{A})|\mathcal{A}| + S(\mathcal{A})|\mathcal{B} - \mathcal{A}| \\ \Leftrightarrow & S(\mathcal{B} - \mathcal{A})|\mathcal{A}| \leq S(\mathcal{A})|\mathcal{B} - \mathcal{A}| \\ \Leftrightarrow & \frac{S(\mathcal{B} - \mathcal{A})}{|\mathcal{B} - \mathcal{A}|} \leq \frac{S(\mathcal{A})}{|\mathcal{A}|} \\ \Leftrightarrow & \operatorname{av}(\mathcal{B} - \mathcal{A}) \leq \operatorname{av}(\mathcal{A}) \end{aligned}$$

**Lemma 2.2.** For a nonempty finite set X, let  $\mathcal{A} \subseteq \mathcal{P}(X)$ . If there exists  $r_1, r_2 \in \mathbb{R}$  and partition  $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$  of  $\mathcal{A}$  such that  $r_1 \leq \operatorname{av}(\mathcal{A}_i) \leq r_2$ , then  $r_1 \leq \operatorname{av}(\mathcal{A}) \leq r_2$ .

Proof. Now

and

$$av(\mathcal{A}) = \frac{S(\mathcal{A})}{|\mathcal{A}|} = \frac{\sum_{i=1}^{k} S(\mathcal{A}_{i})}{|\mathcal{A}|} = \frac{\sum_{i=1}^{k} |\mathcal{A}_{i}| av(\mathcal{A}_{i})}{|\mathcal{A}|} \ge \frac{\sum_{i=1}^{k} |\mathcal{A}_{i}| r_{1}}{|\mathcal{A}|} = \frac{r_{1} \sum_{i=1}^{k} |\mathcal{A}_{i}|}{|\mathcal{A}|} = r_{1}$$
$$av(\mathcal{A}) = \frac{S(\mathcal{A})}{|\mathcal{A}|} = \frac{\sum_{i=1}^{k} S(\mathcal{A}_{i})}{|\mathcal{A}|} = \frac{\sum_{i=1}^{k} |\mathcal{A}_{i}| av(\mathcal{A}_{i})}{|\mathcal{A}|} \le \frac{\sum_{i=1}^{k} |\mathcal{A}_{i}| r_{2}}{|\mathcal{A}|} = \frac{r_{2} \sum_{i=1}^{k} |\mathcal{A}_{i}|}{|\mathcal{A}|} = r_{2}.$$

A simplicial complex  $\mathcal{A}$  is a subset of  $\mathcal{P}(X)$  such that  $\emptyset \in \mathcal{A}$  and  $A \in \mathcal{A}$  implies  $\mathcal{P}(A) \subseteq \mathcal{A}$ . Simplicial complexes have numerous applications in combinatorics (and algebraic topology); here we will need a result on the average size of a set in a complex.

**Proposition 2.3.** Let  $\mathcal{A}$  be a simplicial complex on a nonempty finite set X with n elements. Then for all  $k \leq \frac{n}{2}$ 

$$|\mathcal{A}_k| \ge |\mathcal{A}_{n-k}|,$$

where  $\mathcal{A}_k = \{A \in \mathcal{A} : |A| = k\}$ . Hence  $\operatorname{av}(\mathcal{A}) \leq \frac{n}{2}$ .

Proof. We will use Hall's Theorem (see, for example, [9]) to show  $|\mathcal{A}_k| \geq |\mathcal{A}_{n-k}|$ . Consider the bipartite graph with bipartition  $(\mathcal{A}_{n-k}, \mathcal{A}_k)$  where  $A \in \mathcal{A}_{n-k}$  and  $B \in \mathcal{A}_k$  are adjacent if and only if  $B \subseteq A$ . As  $\mathcal{A}$  is a simplicial complex, the degree of each  $A \in \mathcal{A}_{n-k}$  is  $\binom{n-k}{k}$  and the degree each  $B \in \mathcal{A}_k$  is at most  $\binom{n-k}{n-2k} = \binom{n-k}{k}$ . Furthermore, for any subset  $S \subseteq \mathcal{A}_{n-k}$  there are exactly  $|S|\binom{n-k}{k}$  edges incident with the vertices of S and at most  $|N(S)|\binom{n-k}{k}$  edges incident to the vertices of N(S). Therefore  $|S| \leq |N(S)|$  and by Hall's Theorem you can match  $\mathcal{A}_{n-k}$  into  $\mathcal{A}_k$ , so  $|\mathcal{A}_k| \geq |\mathcal{A}_{n-k}|$ .

Now let  $\mathcal{B}_k = \mathcal{A}_{n-k} \cup \mathcal{A}_k$ . Note that  $\operatorname{av}(\mathcal{B}_k) \leq \frac{n}{2}$  and  $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_{\frac{n}{2}}$  is a partition of  $\mathcal{A}$ . It follows from Lemma 2.2 that  $\operatorname{av}(\mathcal{A}) \leq \frac{n}{2}$ .

On our path to proving that  $K_n$  is the unique graph, among all graphs of order n, with the least average order of dominating sets, we shall need the following that states that every subset of vertices that does not omit the closed neighbourhood of *some* vertex must be dominating.

**Lemma 2.4.** [1] Let G be a graph of order n then  $d_{n-k}(G) = \binom{n}{k}$  for all  $k \leq \delta(G)$ , where  $\delta(G)$  is the minimum degree of G.

**Theorem 2.5.** Let G be a graph of order n then  $\operatorname{avd}(G) \geq \frac{n2^{n-1}}{2^n-1}$  with equality if and only if  $G \cong K_n$ .

*Proof.* Let  $\overline{\mathcal{D}(G)}$  be the collection of subsets  $S \subseteq V(G)$  such that V(G) - S is a dominating set of G. Note  $\overline{\mathcal{D}(G)}$  is a simplicial complex. Therefore by Theorem 2.3 for all  $k \leq \frac{n}{2}$ ,

$$d_{n-k} = |\{S \in \overline{\mathcal{D}(G)} : |S| = k\}| \ge |\{S \in \overline{\mathcal{D}(G)} : |S| = n-k\}| = d_k$$

It follows that  $\operatorname{avd}(G)$  is minimized if  $d_{n-k} = d_k$  for all  $k \leq \frac{n}{2}$ . As  $d_n = 1$  and  $d_0 = 0$  this cannot happen for k = 0, but  $\operatorname{avd}(G)$  will be minimized if and only if  $d_{n-k} = d_k$  for all  $1 \leq k \leq \frac{n}{2}$ , and this does occur for  $K_n$ . Thus  $\operatorname{avd}(G) \geq \frac{n2^{n-1}}{2^n-1} > \frac{1}{2}$ . To find all extremal graphs, suppose  $d_{n-k} = d_k$  for all  $1 \leq k \leq \frac{n}{2}$ .

First we assume that  $\delta(G) \geq 1$ . By Lemma 2.4,  $d_{n-1} = n$  and therefore  $d_1 = n$ . A dominating set of order one must be a vertex of degree n-1. Therefore G has n vertices of degree n-1 and is hence  $K_n$ . Now suppose  $\delta(G) = 0$  and G has  $r \geq 1$  isolated vertices. Then  $D(G, x) = x^r D(H, x)$  for some isolate-free graph on n-r vertices. Furthermore

$$\operatorname{avd}(G) = \frac{D'(G,1)}{D(G,1)} = \frac{rD(H,1) + D'(H,1)}{D(H,1)} = r + \operatorname{avd}(H) \ge r + \frac{n-r}{2} \ge \frac{n+1}{2}.$$

If n = 1, then  $G \cong K_1$  and  $\frac{n+1}{2} = \frac{n2^{n-1}}{2^n-1}$ . An easy induction shows that  $\frac{n+1}{2} > \frac{n2^{n-1}}{2^n-1}$  for  $n \ge 2$ , completing the proof.

### 3 Bounds

#### 3.1 General graphs

For a graph on *n* vertices, we have seen that  $\operatorname{avd}(G) \leq n$ , with the bound achieved uniquely by  $\overline{K_n}$ . However, can we say more if we insist on the graph being connected? Or even just having no isolated vertices? The lower bound occurs for complete graphs, so no improvement is possible there, but the upper bound leaves some room for improvement. We shall do so first in terms of  $\delta$ , the minimum degree.

For a dominating set S of a graph G let

$$a(S) = \{ v \in S : S - v \notin \mathcal{D}(G) \},\$$

the set of *critical* vertices of S with respect to domination (in that their removal makes the set no longer dominating). This parameter is key to improving the upper bound. we will need first an expression for the sum of a(S) over all dominating sets.

**Lemma 3.1.** For a graph G with n vertices.

$$\sum_{S \in \mathcal{D}(G)} |a(S)| = 2D'(G, 1) - nD(G, 1).$$

Proof. For a vertex  $v \in V(G)$  let  $a_v(G) = \{S \in \mathcal{D}(G) : S - v \notin \mathcal{D}(G)\}$ , and let  $\mathcal{D}_{+v}(G)$  denote the collection of dominating sets which contain v. Let  $\mathcal{D}_{-v}(G)$  denote the collection of dominating sets which does not contain v. Clearly  $a_v(G) \subseteq \mathcal{D}_{+v}(G)$ . We will now show is a one-to-one correspondence between  $\mathcal{D}_{+v}(G) - a_v(G)$  and  $\mathcal{D}_{-v}(G)$ . For any  $S \in \mathcal{D}_{+v}(G) - a_v(G)$ ,  $S - v \in \mathcal{D}(G)$  so clearly  $S - v \in \mathcal{D}_{-v}(G)$ . Furthermore if  $S \in \mathcal{D}_{-v}(G)$  then  $S \cup \{v\} \in \mathcal{D}_{+v}(G)$  and  $S \cup \{v\} \notin a_v(G)$ . As the maps are injective, it follows that  $|\mathcal{D}_{+v}(G) - a_v(G)| = |\mathcal{D}_{-v}(G)|$ and as  $a_v(G) \subseteq \mathcal{D}_{+v}(G)$  we have  $|a_v(G)| = |\mathcal{D}_{+v}(G)| - |\mathcal{D}_{-v}(G)|$ . Furthermore

$$\sum_{v \in V(G)} |\mathcal{D}_{+v}(G)| = \sum_{i=1}^{n} i \cdot d(G, i) = D'(G, 1)$$
(2)

and

$$\sum_{v \in V(G)} |\mathcal{D}_{-v}(G)| = \sum_{i=1}^{n} (n-i) \cdot d(G,i) = nD(G,1) - D'(G,1).$$
(3)

Therefore

$$\sum_{S \in \mathcal{D}(G)} |a(S)| = \sum_{v \in V(G)} a_v(G) = \sum_{v \in V(G)} (|\mathcal{D}_{+v}(G)| - |\mathcal{D}_{-v}(G)|) = 2D'(G, 1) - nD(G, 1).$$

In order to get to our upper bound, we need to partition a(S). Let S be a dominating set of G containing the vertex v. By definition  $v \in a(S)$  if and only if S - v is not a dominating in G. Therefore  $v \in a(S)$  if and only if there exists  $u \in N[v]$  such that among the vertices of S, u is only dominated by v (u could very well be v). We will call such a vertex u a private neighbour of v with respect to S. Let  $\operatorname{Priv}_S(v)$  denote the collection of all private neighbours of v with respect to S, that is,

$$\operatorname{Priv}_{S}(v) = \{ u \in N[v] : N[u] \cap S = \{v\} \}.$$

Note  $v \in a(S)$  if and only if  $\operatorname{Priv}_S(v) \neq \emptyset$ . Moreover, for  $v \in a(S)$ , note that  $\operatorname{Priv}_S(v) \cap S \subseteq \{v\}$ . We now partition  $a(S) = a_1(S) \cup a_2(S)$ , where

$$a_1(S) = \{ v \in a(S) : \operatorname{Priv}_S(v) \cap (V - S) \neq \emptyset \}$$
$$a_2(S) = \{ v \in a(S) : \operatorname{Priv}_S(v) = \{v\} \}.$$

(We allow either to be empty.) Note that if  $v \in a_2(S)$  then  $N(v) \subseteq V - S$ . We can partition  $V - S = N_1(S) \cup N_2(S)$ , where

$$N_1(S) = \{ v \in V - S : |N[v] \cap S| = 1 \}$$
$$N_2(S) = \{ v \in V - S : |N[v] \cap S| \ge 2 \}.$$

That is,  $N_1(v)$  is the set of those vertices outside of S that have a single neighbour in S, and  $N_2(S)$  are those that have more than one neighbour in S. (Again, we allow either to be empty.)

As an example consider the labelled  $P_5$  in Figure 3.1. Now  $S = \{v_2, v_3, v_5\}$  is a dominating set. Furthermore  $a(S) = \{v_2, v_5\}$  with  $a_1(S) = \{v_2\}$ ,  $a_2(S) = \{v_5\}$ ,  $N_1(S) = \{v_1\}$  and  $N_2(S) = \{v_4\}$ . Alternatively, let  $S' = \{v_1, v_3, v_5\}$ . Now  $a(S') = \{v_1, v_3, v_5\}$  with  $a_1(S') = \emptyset$ ,  $a_2(S') = \{v_1, v_3, v_5\}$ ,  $N_1(S') = \emptyset$  and  $N_2(S') = \{v_2, v_4\}$ .

•				-•
$v_1$	$v_2$	$v_3$	$v_4$	$v_5$

Figure 3.1: A vertex labelled  $P_5$ 

**Lemma 3.2.** Let G be a graph. For any  $S \in \mathcal{D}(G)$ ,  $|a_1(S)| \leq |N_1(S)|$ .

Proof. For any  $u \in N_1(S)$ ,  $N[u] \cap S \in a_1(S)$ . Therefore the map  $f : N_1(S) \to a_1(S)$  where  $f(v) = N[v] \cap S$  is surjective, so  $|N_1(S)| \ge |a_1(S)|$ .

For graph G containing a vertex v let  $p_v(G)$  denote the collection of subsets of V - N[v] which dominate G - v(and hence they dominate G - N[v] as well). We are now ready to improve our upper bound for graph with no isolated vertices.

**Theorem 3.3.** Let G be a graph with  $n \ge 2$  vertices and minimum degree  $\delta \ge 1$ . Then

$$\operatorname{avd}(G) \le \frac{2n(2^{\delta} - 1) + n}{3(2^{\delta} - 1) + 1},$$

and so  $\operatorname{avd}(G) \leq \frac{3n}{4}$ .

*Proof.* We begin by showing

$$\sum_{S \in \mathcal{D}(G)} |a_2(S)| = \sum_{v \in V(G)} |p_v(G)|.$$
 (4)

It suffices to show for any  $S \in \mathcal{D}(G)$  containing  $v, v \in a_2(G)$  if and only if  $S - v \in p_v(G)$ . Suppose  $v \in a_2(S)$ . By definition of  $a_2(G), v \in a_2(G)$  if and only if  $\operatorname{Priv}_S(v) = \{v\}$ . Therefore  $N[v] \cap S = \{v\}$  and  $N(v) \subseteq N_2(S)$ . Furthermore  $S - v \subseteq V - N[v]$  and dominates G - v and thus  $S - v \in p_v(G)$ . Conversely, suppose  $S - v \in p_v(G)$ . By definition of  $p_v(G), S - v \notin \mathcal{D}(G)$  but  $S \in \mathcal{D}(G)$ . Therefore  $v \in a(S)$ . However every neighbour of v is already dominated by S - v; therefore,  $N(v) \subseteq N_2(S)$  and  $v \in a_2(G)$ .

For now fix  $v \in V(G)$ . By definition every  $S \in p_v(G)$  dominates G - v but does not contain any vertices of N[G]. Therefore for any non-empty  $T \subseteq N(v), S \cup T \in \mathcal{D}_{-v}(G)$ . Furthermore for  $S_1, S_2 \in p_v(G)$  and  $T_1, T_2 \in N(v)$ 

if  $S_1 \cup T_1 = S_2 \cup T_2$  then  $S_1 = S_2$  and  $T_1 = T_2$ . Let  $\mathcal{P}(N(v))$  denote the power set of N(v). Then there is an injective map from  $p_v(G) \times (\mathcal{P}(N(v)) - \emptyset)$  to  $\mathcal{D}_{-v}(G)$  and hence  $(2^{\deg(v)} - 1)|p_v(G)| \leq |\mathcal{D}_{-v}(G)|$ . So together with (3) and (4) we obtain

$$\sum_{S \in \mathcal{D}(G)} |a_2(S)| = \sum_{v \in V(G)} |p_v(G)| \le \sum_{v \in V(G)} \frac{|\mathcal{D}_{-v}(G)|}{2^{\deg(v)} - 1} \le \sum_{v \in V(G)} \frac{|\mathcal{D}_{-v}(G)|}{2^{\delta} - 1} = \frac{nD(G, 1) - D'(G, 1)}{2^{\delta} - 1}.$$

By Lemma 3.2,  $|a_1(S)| \leq |N_1(S)|$ . So together with (3) we obtain

$$\sum_{S \in \mathcal{D}(G)} |a_1(S)| \le \sum_{S \in \mathcal{D}(G)} |N_1(S)| \le \sum_{S \in \mathcal{D}(G)} |V - S| = \sum_{v \in V(G)} |\mathcal{D}_{-v}(G)| = nD(G, 1) - D'(G, 1).$$

By Lemma 3.1,  $\sum_{S \in \mathcal{D}(G)} |a(S)| = 2D'(G, 1) - nD(G, 1)$ , and hence from

$$\sum_{S \in \mathcal{D}(G)} |a(S)| = \sum_{S \in \mathcal{D}(G)} |a_1(S)| + \sum_{S \in \mathcal{D}(G)} |a_2(S)|$$

we have that

$$2D'(G,1) - nD(G,1) \le nD(G,1) - D'(G,1) + \frac{nD(G,1) - D'(G,1)}{2^{\delta} - 1}.$$

From this it follows that

$$\frac{D'(G,1)}{D(G,1)} \le \frac{2n(2^{\delta}-1)+n}{3(2^{\delta}-1)+1}.$$

Finally, one can verify that as  $\delta \geq 1$ ,

$$\frac{2n(2^{\delta}-1)+n}{3(2^{\delta}-1)+1} \le \frac{3n}{4}$$

and we are done.

Theorem 3.3 shows all graphs with no isolated vertices have  $\operatorname{avd}(G) \leq \frac{3n}{4}$ . However for  $\delta \geq 4$  the bound can be improved again, if we are even more careful with our counting. Again, we shall need a couple of technical lemmas first.

Lemma 3.4. For any graph G,

$$\sum_{S \in \mathcal{D}(G)} |N_1(S)| = \sum_{e \in E(G)} |\mathcal{D}(G) - \mathcal{D}(G - e)|.$$

Proof. It suffices to show for every dominating set  $S \in \mathcal{D}(G)$  there are exactly  $N_1(S)$  edges  $e = \{u, v\}$  in G such that  $S \notin \mathcal{D}(G - e)$ . For every  $S \in \mathcal{D}(G)$  consider the edge e in G. If e goes from a vertex  $v \in N_1(S)$  to some vertex in S then v is not dominated by S in G - e, so  $S \in \mathcal{D}(G) - \mathcal{D}(G - e)$ .

Conversely suppose e does not go from a vertex in  $N_1(S)$  to some vertex in S; we need to show that  $S \notin \mathcal{D}(G) - \mathcal{D}(G - e)$ . Note that in G - e, S necessarily dominates every vertex other than possibly u and v. Therefore  $S \in \mathcal{D}(G - e)$  if and only if S dominates both u and v in G - e. Consider the following 3 cases:

Case 1:  $u, v \in S$ . Then both u and v dominate themselves in S so S is a dominating set in G - e. Therefore  $S \notin \mathcal{D}(G) - \mathcal{D}(G - e)$ .

Case 2:  $u, v \notin S$ . As S is a dominating set of G, there exists vertices  $x, y \in S$  (possibly x = y) such that x and y are adjacent to u and v respectively in G. Note x and y are still adjacent to u and v respectively in G - e. Therefore S is a dominating set in G - e and  $S \notin \mathcal{D}(G) - \mathcal{D}(G - e)$ .

Case 3: Either  $u \in S$  and  $v \notin S$ , or  $u \notin S$  and  $v \in S$ . Without loss of generality suppose  $u \in S$  and  $v \notin S$ . As e does not go from a vertex in  $N_1(S)$  to some vertex in S then  $v \notin N_1(S)$  and therefore  $v \in N_2(S)$ . By definition of  $N_2(S)$ , there exists at least one other vertex  $x \in S$  adjacent to v. Therefore x is still adjacent to v in G - e and S is a dominating set in G - e. Therefore  $S \notin \mathcal{D}(G) - \mathcal{D}(G - e)$ .

Therefore for every dominating set  $S \in \mathcal{D}(G)$ , the number of edges e in G which have  $S \in \mathcal{D}(G) - \mathcal{D}(G - e)$  is exactly the number of edges from  $N_1(S)$  to S. By definition of  $N_1(S)$ , each vertex in  $N_1(S)$  is adjacent to exactly one vertex in S. Therefore, the number of edges e in G which have  $S \in \mathcal{D}(G) - \mathcal{D}(G - e)$  is exactly  $|N_1(S)|$ .

**Lemma 3.5.** [17] Let G be a graph. For every edge  $e = \{u, v\}$  of G,

$$|\mathcal{D}(G) - \mathcal{D}(G - e)| = |p_u(G - e)| + |p_v(G - e)| - |p_u(G)| - |p_v(G)|.$$

_	_
_	_

We are now ready to prove another upper bound for  $\operatorname{avd}(G)$ .

**Theorem 3.6.** For any graph G with no isolated vertices,

$$\operatorname{avd}(G) \le \frac{n}{2} + \sum_{v \in V(G)} \frac{\deg(v)}{2^{\deg(v)+1} - 2}.$$

Proof. By Lemma 3.2, Lemma 3.4, Lemma 3.5 we obtain.

$$\sum_{S \in \mathcal{D}(G)} |a_1(S)| \le \sum_{e \in E(G)} (|p_u(G-e)| + |p_v(G-e)| - |p_u(G)| - |p_v(G)|) = \sum_{v \in V(G)} \sum_{u \in N(v)} (|p_v(G-uv)| - |p_v(G)|) \le \sum_{v \in V(G)} |a_1(S)| \le \sum_{v$$

Together with (4) we obtain

$$\sum_{S \in \mathcal{D}(G)} |a(S)| = \sum_{S \in \mathcal{D}(G)} (|a_1(S)| + |a_2(S)|) \le \sum_{v \in V(G)} \sum_{u \in N(v)} |p_v(G - uv)| - \sum_{v \in V(G)} (\deg(v) - 1)|p_v(G)|.$$

Furthermore as G has no isolated vertices we obtain

$$\sum_{S \in \mathcal{D}(G)} |a(S)| \le \sum_{v \in V(G)} \sum_{u \in N(v)} |p_v(G - uv)|.$$
(5)

For each  $v \in V(G)$  and  $e = \{u, v\} \in E(G)$  consider  $S \in p_v(G - e)$ . For any nonempty  $T \subseteq N_G[v] - \{u\}$ ,  $S \cup T \in \mathcal{D}(G - e) \subseteq \mathcal{D}(G)$  and all such sets are distinct. Therefore  $(2^{\deg(v)} - 1)|p_v(G - e)| \leq |\mathcal{D}(G)|$  (where the degree is in the graph G) and together with Lemma 3.1 and (5) we obtain

$$2D'(G,1) - nD(G,1) = \sum_{S \in \mathcal{D}(G)} |a(S)| \le \sum_{v \in V(G)} \frac{\deg(v) \cdot D(G,1)}{2^{\deg(v)} - 1}$$

from which is follows that

$$\frac{D'(G,1)}{D(G,1)} \le \frac{n}{2} + \sum_{v \in V(G)} \frac{\deg(v)}{2^{\deg(v)+1} - 2}.$$

**Corollary 3.7.** For a graph G with minimum degree  $\delta \geq 1$ .

$$\operatorname{avd}(G) \le \frac{n}{2} \left( 1 + \frac{\delta}{2^{\delta} - 1} \right).$$

In particular if  $\delta \ge 2\log_2(n)$  then  $\operatorname{avd}(G) \le \frac{n+1}{2}$ .

*Proof.* Let  $f(x) = \frac{x}{2^{x+1}-2}$ . It is not hard to verify that for  $x \ge 1$ , f(x) is a decreasing function. Therefore for all  $v \in V(G)$ ,  $f(\deg(v)) \le f(\delta)$ , and by Theorem 3.6

$$\operatorname{avd}(G) \le \frac{n}{2} + \sum_{v \in V(G)} \frac{\deg(v)}{2^{\deg(v)+1} - 2} \le \frac{n}{2} + \frac{n \cdot \delta}{2^{\delta+1} - 2} = \frac{n}{2} \left( 1 + \frac{\delta}{2^{\delta} - 1} \right).$$

Now suppose  $\delta \ge 2\log_2(n)$ . As  $\delta \le n-1$ , we know that  $2\log_2(n) \le n-1$ . Again, one can verify that  $g(n) = \delta/(2^{\delta}-1)$  is decreasing for  $\delta \ge 1$ , so

$$\operatorname{avd}(G) \leq \frac{n}{2} \left( 1 + \frac{\delta}{2^{\delta} - 1} \right)$$
$$\leq \frac{n}{2} \left( 1 + \frac{2 \log_2(n)}{2^{2 \log_2(n)} - 1} \right)$$
$$\leq \frac{n}{2} \left( 1 + \frac{n - 1}{n^2 - 1} \right)$$
$$= \frac{n}{2} \left( 1 + \frac{1}{n + 1} \right)$$
$$\leq \frac{n}{2} \left( 1 + \frac{1}{n} \right)$$
$$= \frac{n + 1}{2}.$$

	-	-	-	-

Theorem 3.3 and Corollary 3.7 give two different upper bounds for  $\operatorname{avd}(G)$  based on  $\delta(G)$ . Figure 3.2 plots  $\operatorname{avd}(G)$  sorted by minimum degree for all graphs of order n = 8 and n = 9, respectively. The curve in Figure 3.2 is the minimum of the two bounds of Theorem 3.3 and Corollary 3.7 evaluated for each integer  $0 \le \delta \le n$  and linearly interpolated between each point.

Our best upper bound for all isolate-free graphs remains  $\operatorname{avd}(G) \leq \frac{3n}{4}$ . However by Corollary 3.7 if  $\delta(G) \geq 4$ then  $\operatorname{avd}(G) \leq \frac{19n}{30} \leq \frac{2n}{3}$ . In fact, all graphs up to order 9 with no isolated vertices have  $\operatorname{avd}(G) \leq \frac{2n}{3}$ . This leads



Figure 3.2: The bounds from Theorem 3.3 and Corollary 3.7 compared to avd(G) for n = 8 and n = 9.

us to the following conjecture.

**Conjecture 3.8.** Let G be a graph with  $n \ge 2$  vertices. If G has no isolated vertices (so, in particular, if G is connected) then  $\operatorname{avd}(G) \le \frac{2n}{3}$ .

We can show that the upper bound in Conjecture 3.8 is achieved for all  $n \ge 2$ : For n = 2 and n = 3,  $\operatorname{avd}(K_2) = \frac{4}{3}$ and  $\operatorname{avd}(K_{1,2}) = 2$ . For any  $n \ge 4$  there exists non-negative integers k and  $\ell$  such that  $n = 2k+3\ell$ . Then by Lemma 1.1 any graph of the form  $H = kK_2 \cup \ell K_{1,2}$  will have  $\operatorname{avd}(H) = \frac{2n}{3}$ . These graphs are not connected, but one can insist on connectivity as follows. Let G be any graph on  $k + \ell$  vertices, and let G' be the graph obtained by adding one leaf to k vertices of G and two leaves to the other  $\ell$  vertices of G. Oboudi showed [18] that D(G', x) = D(H, x). Therefore  $\operatorname{avd}(G') = \frac{2n}{3}$ , and by choosing G to be connected, the graph G' will be as well.

While we are unable to prove Conjecture 3.8, we can provide some evidence for it. A graph G is called *quasi-regularizable* if one can replace each edge of G with a non-negative number of parallel copies, so as to obtain a regular multigraph of minimum degree at least one. Any graph which contains a spanning subgraph which is both regular and nonempty is quasi-regularizable; in particular, any graph which contain either a perfect matching or a hamiltonian cycle is quasi-regularizable. Berge [6] characterized quasi-regularizable graphs as those for which  $|S| \leq |N(S)|$  holds for any independent set S of G. We will now show that for quasi-regularizable graphs, Conjecture 3.8 holds.

**Theorem 3.9.** If G is a quasi-regularizable graph then  $\operatorname{avd}(G) \leq \frac{2n}{3}$ .

Proof. We begin by showing  $|a(S)| \leq n - |S|$  for every  $S \in \mathcal{D}(G)$ . By Lemma 3.2,  $|a_1(S)| \leq |N_1(S)|$ . Therefore it suffices to show  $|a_2(S)| \leq |N_2(S)|$ . For every  $v \in a_2(S)$ ,  $N(v) \subseteq V - S$  as otherwise  $\operatorname{Priv}_S(v) \neq \{v\}$ . Furthermore,  $N(v) \subseteq N_2(S)$  as otherwise  $v \in a_1(S)$ . Therefore  $a_2(S)$  is an independent set with  $N(a_2(S)) \subseteq N_2(S)$ . As G is a

quasi-regularizable graph then  $|a_2(S)| \leq |N(a_2(S))| \leq |N_2(S)|$ , so

$$|a(S)| = |a_1(S)| + |a_2(S)| \le |N_1(S)| + |N_2(S)| = n - |S|.$$

Finally, as  $|a(S)| \le n - |S|$  then  $\sum_{S \in \mathcal{D}(G)} |a(S)| \le nD(G, 1) - D'(G, 1)$ . Thus together with Lemma 3.1 we obtain

$$2D'(G,1) - nD(G,1) \le nD(G,1) - D'(G,1) \Rightarrow \operatorname{avd}(G) = \frac{D'(G,1)}{D(G,1)} \le \frac{2n}{3}.$$

#### 3.2 Trees

In this section we turn to trees (which are connected and, if are nontrivial, have  $\delta \ge 1$ ). For every  $n \ge 2$  there is a tree T of order n with  $\operatorname{avd}(T) = \frac{2n}{3}$ , satisfying the upper bound from Conjecture 3.8 for isolate-free graphs. Examples of trees which achieve the upper bound given in Conjecture 3.8 are described in the paragraph following Conjecture 3.8 (if one chooses the base graph G there to be a tree as well). It also remains an open question whether this is actual the upper bound amongst trees.

However, what about the lower bound? For graphs the lower bound was achieved by complete graphs, but these are far from being trees. We show now that  $\operatorname{avd}(T) \ge \operatorname{avd}(K_{1,n-1})$ , and the argument is even more involved than for the lower bound for general graphs. For this we require a result similar to that of Proposition 2.3. However the proof of this is considerably more involved. For a tree T of order n, recall  $\mathcal{D}(T)$  denotes the collection of all dominating sets in T. For now fix  $S \in \mathcal{D}(T)$ . Recall in the proof of Proposition 2.3, it was important to bound the number of subsets  $S' \subseteq S$  where S' is also a dominating set and |S'| = k. Let

$$\operatorname{dom}_k(S) = |\{S' \subseteq S : S' \in \mathcal{D}(T) \text{ and } |S'| = k\}|.$$

The trivial upper bound, which was used in the proof of Proposition 2.3, is simply  $\operatorname{dom}_k(S) \leq \binom{|S|}{k}$ , but we need something stronger for trees. Recall  $a(S) = \{v \in S : S - v \notin \mathcal{D}(T)\}$ . Therefore for any  $S' \subseteq S$ , if  $S' \in \mathcal{D}(T)$ then  $a(S) \subseteq S'$ . Therefore  $\operatorname{dom}_k(S) \leq \binom{|S|-|a(S)|}{k-|a(S)|}$ . However this is only useful if  $a(S) \neq \emptyset$ . On the other hand, when  $a(S) = \emptyset$ , S is double dominating set [12], that is, a subset  $S \subseteq V(G)$  such that for every vertex  $v \in V(G)$ ,  $|N[v] \cap S| \geq 2$ . The order of the smallest double dominating set is denoted  $\gamma_{\times 2}(G)$ . Note that for a dominating set S of a tree T, if  $|S| < \gamma_{\times 2}(T)$  then  $a(S) \geq 1$ . Moreover, suppose  $\gamma(T) + \gamma_{\times 2}(T) \geq n + 1$ . Then  $|S| > \gamma_{\times 2}(T)$ implies  $n + 1 - |S| < \gamma(T)$  and hence  $\operatorname{dom}_k(S) = 0$  for  $k \leq n + 1 - |S|$ .

**Theorem 3.10.** If T is a nontrivial tree then  $\gamma_{\times 2}(T) + \gamma(T) \ge n + 1$ .

Proof. We can assume that  $n \ge 3$ , as if n = 2, then  $T = K_2$  and so  $\gamma_{\times 2}(T) = 2$ ,  $\gamma(T) = 1$  and the result holds. Set V(T) = V. It is sufficient to show for any double dominating set S,  $\gamma(T) \ge n - |S| + 1$ . Note that if the m vertices  $v_1, \ldots, v_m$  had pairwise disjoint closed neighbourhoods, then  $\gamma(T) \ge m$  as any dominating set would need to contain at least one vertex from each closed neighbourhood. Therefore it is sufficient to show for any double dominating set S, there exists a collection of |V - S| + 1 vertices with pairwise disjoint closed neighbourhoods. We will induct on the number of vertices in V - S. For  $v \in V$  and  $u \in N(v)$  let B(T, v, u) denote the set of vertices in the same component as u in T - v (See Figure 3.3).



Figure 3.3: An example of B(T, v, u) and B(T, v, w).

Let S be a double dominating set, so that S contains every leaf and stem of T (as stem in a tree is a vertex adjacent of a leaf). The case where |V - S| = 0 is vacuously true for any nontrivial tree. Assume for any nontrivial tree and some  $k \ge 0$  if  $|V - S| \le k$  then there exists a collection of |V - S| + 1 vertices with pairwise disjoint closed neighbourhoods. Now suppose |V - S| = k + 1. Note that for any leaf in T, both it and its stem (i.e. the leaf's only neighbour) must both be in S, otherwise S is not a double dominating set. Fix a leaf  $\ell \in V$ . Now choose a vertex  $v \notin S$  which is of maximum distance to  $\ell$ . Note v is not a stem, as otherwise  $v \in S$  which contradicts  $v \notin S$ . Furthermore,  $\deg(v) \ge 2$ , as otherwise either  $v \in S$  or S is not a double dominating set; in particular,  $v \ne \ell$ .

Let  $w \in N(v)$  be the only neighbour of v which is closer  $\ell$  than v (See Figure 3.3). Note that  $w \neq \ell$ , as otherwise v would be stem and hence belong to S. As  $\deg(v) \geq 2$ , choose  $u \in N(v) - \{w\}$ . Note every vertex in B(T, v, u) is further from  $\ell$  than v and therefore  $B(T, v, u) \subseteq S$ . Moreover, as v is not a stem,  $\deg(u) \geq 2$ . Therefore choose  $u' \in N(u) - \{v\}$ . Note that  $N[u'] \subseteq B(T, v, u)$  (as in Figure 3.3). Now set T' = B(T, v, w) and  $S' = S \cap B(T, v, w)$ . T' is a nontrivial tree as  $w, \ell \in T'$ . For each  $x \in V(T')$ ,  $N_T[x] = N_{T'}[x]$ , except for w where  $N_T[x] = N_{T'}[x] \cup \{v\}$ . As  $v \notin S$ ,  $|N_T[x] \cap S| = |N_{T'}[x] \cap S'| \geq 2$  for all  $x \in V(T')$ . Therefore S' is a double dominating set of T'. Finally the only vertex in V(T) - V(T') which was not in S was v, as v was the furthest vertex from  $\ell$  which was not in S. Therefore |V(T') - S'| = |V(T) - S| - 1 = k and by our induction hypothesis there exists a collection of k+1 vertices with disjoint closed neighbourhoods in T'. Let P denote this collection. As  $v \notin N_T[u']$  then  $N_T[x] \cap N_T[u'] = \emptyset$  for all  $x \in V(T')$ . Therefore  $P \cup \{u'\}$  is a collection of k+2 = |V - S| + 1 vertices with pairwise disjoint closed neighbourhoods in T.

We need two additional lemmas on the way to finding the extremal tree with the least average order of dominating sets.

**Lemma 3.11.** [8] 
$$K_{1,n-1}$$
 has the most dominating sets amongst all trees of order n.

**Lemma 3.12.** [7] For every nontrivial tree T,  $2\gamma(T) \leq \gamma_{\times 2}(T)$ .

**Lemma 3.13.** If T is a graph with n vertices then  $d_{n-k} \ge d_{k+1}$  for all  $k+1 \le \frac{n+1}{2}$ .

*Proof.* Fix  $k \leq \frac{n+1}{2}$ . If  $k + 1 < \gamma(T)$  then clearly  $d_{n-k} \geq d_{k+1}$  holds as  $d_{k+1} = 0$ . So suppose for the remiander of this proof that  $k + 1 \geq \gamma(T)$ . We will now use Hall's Theorem again. As before, let  $\mathcal{D}_k$  denote the collection

of all dominating sets of order k. We now construct a bipartite graph with bipartition  $(\mathcal{D}_{k+1}, \mathcal{D}_{n-k})$ ; two vertices  $A \in \mathcal{D}_{k+1}$  and  $B \in \mathcal{D}_{n-k}$  are adjacent if  $A \subseteq B$ . As every superset of a dominating set remains dominating, the degree of each  $A \in \mathcal{D}_{k+1}$  is  $\binom{n-k-1}{n-2k-1} = \binom{n-k-1}{k}$ . By the same argument used in the proof of Proposition 2.3, it suffices to show for any  $B \in \mathcal{D}_{n-k}$  there are at most  $\binom{n-k-1}{k}$  subsets of B which are in  $\mathcal{D}_{k+1}$ .

By Theorem 3.10,  $\gamma_{\times 2}(T) + \gamma(T) \ge n + 1$  and hence  $k + 1 \ge n + 1 - \gamma_{\times 2}(T)$  We now consider two cases:

Case 1: Suppose  $k + 1 > n + 1 - \gamma_{\times 2}(T)$  then  $\gamma_{\times 2}(T) > n - k$ . For any dominating set  $B \in \mathcal{D}_{n-k}$  there exists a vertex  $v \in T$  such that  $N[v] \cap B$  contains exactly one vertex. Let  $\{u\} = N[v] \cap B$ . Then B - u is no longer a dominating set. Hence there are at most  $\binom{n-k}{k+1} - \binom{n-k-1}{k+1} = \binom{n-k-1}{k}$  subsets of B which are also in  $\mathcal{D}_{k+1}$ .

Case 2: Suppose  $k + 1 = n + 1 - \gamma_{\times 2}(T)$  then  $\gamma_{\times 2}(T) = n - k$ . As  $\gamma_{\times 2}(T) + \gamma(T) \ge n + 1$  then  $\gamma(T) \ge k + 1$ . Furthermore, as  $\gamma(T) \le k + 1$ , it follows that  $k + 1 = \gamma(T)$ . For any dominating set  $B \in \mathcal{D}_{n-k}$ , if B is not a double dominating set then it follows for Case 1 that there are at most  $\binom{n-k-1}{k}$  subsets of B which are also in  $\mathcal{D}_{k+1}$ . So suppose B is a double dominating set. Let m be the number stems in T. If m = 1 then  $T = K_{1,n-1}$ . It is easy to see  $k + 1 = \gamma(K_{1,n-1}) = 1$  so n - k = n. Furthermore more  $d_n(K_{1,n-1}) = d_1(K_{1,n-1}) = 1$  and therefore  $d_n(K_{1,n-1}) \ge d_1(K_{1,n-1})$ . Now suppose  $m \ge 2$ . Choose two stems  $s_1$  and  $s_2$  along with leaves  $\ell_1$  and  $\ell_2$  which are adjacent to  $s_1$  and  $s_2$  respectively. As B is a double dominating set then  $s_1, s_2, \ell_1, \ell_2 \in B$ , otherwise  $\ell_1$  or  $\ell_2$  with not be double dominated. Furthermore if  $A \subseteq B$  such that  $A \in \mathcal{D}_{k+1}$ , then A is a minimum dominating set. Therefore A contains exactly one of  $s_i$  or  $\ell_i$  for each i = 1, 2 and remaining k + 1 vertices of A are chosen from then remaining n - k - 4 vertices in B. Therefore there are at most  $4\binom{n-k-4}{k-1}$  subsets  $A \subseteq B$  such that  $A \in \mathcal{D}_{k+1}$ . If  $n - k - 4 \ge 1$ , then  $\binom{n-k-4}{k-1} = \binom{n-k-5}{k-2} + \binom{n-k-4}{k-1} \le \binom{n-k-4}{k}$  and

$$\begin{aligned} 4\binom{n-k-4}{k-1} &\leq \left(\binom{n-k-4}{k-2} + \binom{n-k-4}{k-1}\right) + \binom{n-k-4}{k-1} + \left(\binom{n-k-4}{k-1} + \binom{n-k-4}{k}\right) \\ &= \binom{n-k-3}{k-2} + \binom{n-k-4}{k-1} + \binom{n-k-3}{k} \\ &\leq \left(\binom{n-k-3}{k-2} + \binom{n-k-3}{k-1}\right) + \left(\binom{n-k-3}{k-1} + \binom{n-k-3}{k}\right) \\ &= \binom{n-k-2}{k-1} + \binom{n-k-2}{k} \\ &= \binom{n-k-1}{k}, \end{aligned}$$

Otherwise, suppose n - k - 4 < 1. As  $\gamma_{\times 2}(T) = n - k$  then  $\gamma_{\times 2}(T) \leq 4$ . By Lemma 3.12,  $2\gamma(T) \leq \gamma_{\times 2}(T)$ . Therefore  $\gamma(T) \leq 2$ . Furthermore as T has two stems,  $\gamma(T) \geq 2$  and therefore  $\gamma(T) = 2$ . Now  $\gamma_{\times 2}(T) + \gamma(T) = n - k + k + 1 = n + 1$ , so  $n \leq 5$ . There are exactly three trees with  $\gamma(T) = 2$  and  $n \leq 5$ . They are shown below.

However  $\gamma_{\times 2}(P_4) = 4$ ,  $\gamma(P_4) = 2$  so  $\gamma_{\times 2}(P_4) + \gamma(P_4) = 6 \neq n+1$  and  $\gamma_{\times 2}(T_5) = 5$ ,  $\gamma(T_5) = 2$  so  $\gamma_{\times 2}(T_5) + \gamma(T_5) = 2$ 

 $7 \neq n+1$ , so we can omit these cases. Finally,  $D(P_5, x) = x^5 + 5x^4 + 8x^3 + 3x^2$  which satisfies  $d_{n-k} \geq d_{k+1}$  for  $k+1 \leq \frac{n+1}{2}$ .

**Theorem 3.14.** If T is a graph with n vertices  $\operatorname{avd}(T) \ge \operatorname{avd}(K_{1,n-1})$  with equality if and only if  $T \cong K_{1,n-1}$ .

Proof. By Lemma 3.13,  $\operatorname{avd}(T) \geq \frac{n+1}{2}$  and  $d_{n-k} \geq d_{k+1}$  for  $k+1 \leq \frac{n+1}{2}$ . Suppose  $T \not\cong K_{1,n-1}$ , so  $\gamma(T) \geq 2$  and d(T,1) = 0. Now consider the mean domination order of all dominating sets except for the dominating set V(T). Let  $\mathcal{D}^*(T) = \mathcal{D}(T) - \{V(T)\}$ . Note that

$$\operatorname{av}(\mathcal{D}^*(T)) = \frac{D'(T,1) - n}{D(T,1) - 1}$$

For  $k+1 \leq \frac{n+1}{2}$  let  $\mathcal{B}_k = \mathcal{D}_{n-k}(T) \cup \mathcal{D}_{k+1}(T)$ . Note that  $\operatorname{av}(\mathcal{B}_k) \geq \frac{n+1}{2}$  as  $d_{n-k} \geq d_{k+1}$ . Furthermore  $\mathcal{B}_2, \ldots, \mathcal{B}_{\frac{n+1}{2}}$  is a partition of  $\mathcal{D}^*(T)$ . It follows from Lemma 2.2 that

$$\frac{D'(T,1) - n}{D(T,1) - 1} = \operatorname{av}(\mathcal{D}^*(T)) \ge \frac{n+1}{2}.$$

By Lemma 3.11,  $D(T, 1) \leq D(K_{1,n-1}, 1)$  and

$$\operatorname{avd}(T) = \frac{n + D'(T, 1) - n}{D(T, 1)}$$
$$= \frac{n}{D(T, 1)} + \left(\frac{D(T, 1) - 1}{D(T, 1)}\right) \frac{D'(T, 1) - n}{D(T, 1) - 1}$$
$$\geq \frac{n}{D(T, 1)} + \left(\frac{D(T, 1) - 1}{D(T, 1)}\right) \frac{n + 1}{2}$$
$$\geq \frac{n}{D(K_{1,n-1}, 1)} + \left(\frac{D(K_{1,n-1}, 1) - 1}{D(K_{1,n-1}, 1)}\right) \frac{n + 1}{2}$$
$$> \frac{n - 1}{D(K_{1,n-1}, 1)} + \left(\frac{D(K_{1,n-1}, 1) - 1}{D(K_{1,n-1}, 1)}\right) \frac{n + 1}{2}$$
$$= \frac{n - 1}{2^{n-1} + 1} + \left(\frac{2^{n-1}}{2^{n-1} + 1}\right) \frac{n + 1}{2}$$
$$= \frac{n - 1 + 2^{n-2}(n + 1)}{2^{n-1} + 1} = \operatorname{avd}(K_{1,n-1})$$

# 4 Distribution of Average Order of Dominating Sets

What are the possible values for  $\operatorname{avd}(G)$ ? If G is a graph of order n, we showed in the previous section  $\operatorname{avd}(G) \in (\frac{n}{2}, n]$ , but it seems unlikely that one can say precisely what values in the interval are average orders of dominating sets. A natural variant of  $\operatorname{avd}(G)$  is  $\widehat{\operatorname{avd}}(G) = \frac{\operatorname{avd}(G)}{n}$  which we shall refer to as the normalized average order of dominating sets in G. (Similar kinds of normalized graph parameters have been investigated throughout the literature – for example, [13, 15, 19].)

We start with some examples. We say a graph contains a *simple k-path* if there exists k vertices of degree two which induce a path in G. Two families of graphs which contains simple k-paths are paths  $P_n$  and cycles  $C_n$  (where k = n - 2 and n - 1, respectively). The following holds for graphs which contain simple 3-paths.

**Theorem 4.1.** [17] Suppose G is a graph with vertices u, v, w which form a simple 3-path. Then

$$D(G, x) = x(D(G/u, x) + D(G/u/v, x) + D(G/u/v/w, x))$$

where G/u is the graph formed by joining every neighbour of u and then deleting u.

There is no known closed formula for all coefficients of  $D(P_k, x)$  and  $D(C_k, x)$  respectively. This makes determining mean dominating order of paths and cycles difficult. We will now show that for a family of graphs satisfying a recurrence relation similar to that in Theorem 4.1, we can calculate the limit of the normalized average order of dominating sets as  $n \to \infty$ . First we shall put forward a way to calculate the limits of average values of functions of a certain type (which include those that arise from solving linear polynomial recurrences); the proof is straightforward and omitted.

**Theorem 4.2.** Suppose functions  $f_n(x)$  satisfy

$$f_n(x) = \alpha_1(x)(\lambda_1(x))^n + \alpha_2(x)(\lambda_2(x))^n + \dots + \alpha_k(x)(\lambda_k(x))^n$$

where  $\alpha_i(x)$  and  $\lambda_i(x)$  are fixed non-zero analytic functions, such that  $|\lambda_1(1)| > |\lambda_i(1)|$  for all i > 1. Then

$$\lim_{n \to \infty} \frac{f'_n(1)}{nf_n(1)} = \frac{\lambda'_1(1)}{\lambda_1(1)}.$$

**Theorem 4.3.**  $\lim_{n \to \infty} \widehat{\operatorname{avd}}(P_n) = \lim_{n \to \infty} \widehat{\operatorname{avd}}(C_n) \approx 0.618419922.$ 

*Proof.* For both paths and cycles, we have a sequence of graphs  $(G_n)_{n\geq 1}$  which satisfy

$$D(G_n, x) = x(D(G_{n-1}, x) + D(G_{n-2}, x) + D(G_{n-3}, x))$$

As  $G_n$  follows the homogeneous linear recursive relation  $D(G_n, x) = x(D(G_{n-1}, x) + D(G_{n-2}, x) + D(G_{n-3}, x))$ , then  $D(G_n, x) = \alpha_1(x)\lambda_1(x)^n + \alpha_2(x)\lambda_2(x)^n + \alpha_3(x)\lambda_3(x)^n$  where each  $\lambda_1(x)$  satisfies

$$\lambda_i(x)^3 - x\lambda_i(x)^2 - x\lambda_i(x) - x = 0.$$

We solve this cubic polynomial (see also [3]). The solutions are

$$\lambda_1(x) = \frac{x}{3} + p(x) + q(x),$$
  

$$\lambda_2(x) = \frac{x}{3} - p(x) - q(x) + \frac{\sqrt{3}}{2} (p(x) - q(x)) i,$$
  

$$\lambda_3(x) = \frac{x}{3} - p(x) - q(x) - \frac{\sqrt{3}}{2} (p(x) - q(x)) i,$$

where

$$p(x) = \sqrt[3]{\frac{x^3}{27} + \frac{x^2}{6} + \frac{x}{2} + \sqrt{\frac{x^4}{36} + \frac{7x^3}{54} + \frac{x^2}{4}}}$$
$$q(x) = \sqrt[3]{\frac{x^3}{27} + \frac{x^2}{6} + \frac{x}{2} - \sqrt{\frac{x^4}{36} + \frac{7x^3}{54} + \frac{x^2}{4}}}$$

Note  $|\lambda_1(1)| \approx 1.83929 > |\lambda_2(1)| = |\lambda_3(1)| \approx 0.73735$ . Therefore by Theorem 4.2,  $\lim_{n \to \infty} \frac{\operatorname{avd}(G_n)}{n} = \frac{\lambda_1'(1)}{\lambda_1(1)}$ . It follows that

$$\lim_{n \to \infty} \frac{\operatorname{avd}(G_n)}{n} = \frac{\lambda_1'(1)}{\lambda_1(1)}$$
$$= \frac{\frac{1}{3} + p'(1) + q'(1)}{\frac{1}{3} + p(1) + q(1)}$$
$$= \frac{\frac{1}{3} + \frac{27\sqrt{33} + 187}{66(19 + 3\sqrt{33})^{\frac{2}{3}}} - \frac{27\sqrt{33} - 187}{66(19 - 3\sqrt{33})^{\frac{2}{3}}}}{\frac{1}{3} + \frac{(19 + 3\sqrt{33})^{\frac{1}{3}}}{3} + \frac{(19 - 3\sqrt{33})^{\frac{1}{3}}}{3}}$$
$$= \frac{1}{3} + \frac{(88 - 8\sqrt{33})(19 + 3\sqrt{33})^{\frac{1}{3}}}{1056} + \frac{(55 - 7\sqrt{33})(19 + 3\sqrt{33})^{\frac{2}{3}}}{1056},$$

which we will denote by r. By Theorem 4.1, both  $C_n$  and  $P_n$  satisfy the same recurrence as  $G_n$  and hence  $\lim_{n \to \infty} \frac{\operatorname{avd}(P_n)}{n} = \lim_{n \to \infty} \frac{\operatorname{avd}(C_n)}{n} = r \approx 0.618419922.$ 

For all graphs of order 9 we counted the number of graphs with  $\widehat{\operatorname{avd}}(G) \in [\frac{1}{2} + \frac{k}{20n}, \frac{1}{2} + \frac{k+1}{20n})$  for each integer  $0 \leq k \leq 10n - 1$ . Figure 4.1 shows the linearly interpolated distribution of  $\widehat{\operatorname{avd}}(G)$  for all graphs of order 9. The distribution appears to be skewed towards  $\frac{1}{2}$ . However, our next result shows  $\widehat{\operatorname{avd}}(G)$  can be arbitrarily closed to any value in  $[\frac{1}{2}, 1]$ .



Figure 4.1: Distribution of  $\operatorname{avd}(G)$  for all graphs of order 9

**Proposition 4.4.** The set  $\left\{\widehat{\operatorname{avd}}(G) : G \text{ is a graph}\right\}$  is dense in  $\left[\frac{1}{2}, 1\right]$ .

Proof. It suffices to show for every rational number  $\frac{a}{b} \in [0.5, 1]$  (a and b positive) there exists a sequence of graphs  $(G_k)_{k\geq 1}$ , with orders  $n_k$  respectively, such that  $\lim_{k\to\infty} n_k = \infty$  and  $\lim_{k\to\infty} \frac{\operatorname{avd}(G_k)}{n_k} = \frac{a}{b}$ . Let  $G_k = (2b-2a)K_k \cup (2a-b)\overline{K_k}$ ;  $G_k$  has order (2b-2a)k + (2a-b)k = bk. Note such a graph exists as  $\frac{a}{b} \in [0.5, 1]$  and hence  $a \leq b \leq 2a$ . Recall  $D(K_k, x) = (1+x)^k - 1$  and  $D(\overline{K_k}, x) = x^k$ . Therefore

$$\lim_{k \to \infty} \frac{\operatorname{avd}(K_k)}{k} = \lim_{k \to \infty} \frac{k2^{k-1}}{k(2^k - 1)} = 0.5 \quad \text{and} \quad \lim_{k \to \infty} \frac{\operatorname{avd}(\overline{K_k})}{k} = 1.$$

Therefore by Lemma 1.1

$$\lim_{k \to \infty} \frac{\operatorname{avd}(G_k)}{bk} = \lim_{k \to \infty} \frac{(2b - 2a)\operatorname{avd}(K_k) + (2a - b)\operatorname{avd}(\overline{K_k})}{bk}$$
$$= \lim_{i \to \infty} \frac{(2b - 2a)\operatorname{avd}(K_k)}{bk} + \lim_{k \to \infty} \frac{(2a - b)\operatorname{avd}(\overline{K_k})}{bk}$$
$$= \frac{(2b - 2a) \cdot 0.5}{b} + \frac{2a - b}{b} = \frac{a}{b}$$

While we have shown that the closure of the normalized average order of dominating sets fills the interval [1/2, 1], where do most values lie? Let  $\mathcal{G}(n, p)$  denote the sample space of random graphs on n vertices (each edge exists is independent present with probability p). We will now show with probability tending to 1, the normallized average order of dominating sets of a random graph approaches  $\frac{1}{2}$  (even if the graph is sparse with p close to 0); this explains the "bundling up" of values near n/2 in Figure 4.1.

**Theorem 4.5.** Let  $G_n \in \mathcal{G}(n, p)$ . Then

$$\lim_{n \to \infty} \widehat{\operatorname{avd}}(G_n) = \frac{1}{2}$$

*Proof.* It follows from Theorem 2.5 that  $\widehat{\operatorname{avd}}(G_n) \ge \frac{1}{2}$ . Therefore it is sufficient to show  $\lim_{n \to \infty} \widehat{\operatorname{avd}}(G_n) \le \frac{1}{2}$ .

The degree of any vertex v of  $G_n$  has a binomial distribution  $X_v$  with N = n - 1, and hence has mean p(n - 1). From Hoeffding's well known bound on the tail of a binomial distribution, it follows that for any fixed  $\varepsilon > 0$ ,

$$\operatorname{Prob}\left(X_v \le (p-\varepsilon)(n-1)\right) \le e^{-2\varepsilon^2(n-1)}$$

Thus

$$\operatorname{Prob}\left(\cup_{v} X_{v} \leq (p-\varepsilon)(n-1)\right) \leq n e^{-2\varepsilon^{2}(n-1)} \to 0.$$

It follows that  $\delta(G_n) > (p-\varepsilon)(n-1) > 2\log_2(n)$  with probability tending to 1. By Corollary 3.7, if  $\delta(G_n) \ge 2\log_2(n)$  then  $\operatorname{avd}(G_n) \le \frac{n+1}{2}$ . Therefore with probability tending to 1,

$$\lim_{n \to \infty} \widehat{\operatorname{avd}}(G_n) \le \lim_{n \to \infty} \frac{n+1}{2n} = \frac{1}{2},$$

and we are done.

### 5 Conclusion and Open Problems

The most salient open problem is that in Conjecture 3.8, namely, a tight upper bound on the average order of dominating sets among all connected graphs of order n. As mentioned previously, any graph which contains a perfect matching is quasi-regularizable. Let  $\nu(G)$  denote the matching number of G, that is, the largest cardinality of a matching. We alter the proof of the previous theorem to put  $\operatorname{avd}(G)$  in terms of  $\nu(G)$ . This will not improve the bound from Theorem 3.9 for graphs with perfect matchings. However there are graphs which contain near perfect matchings which are not quasi-regularizable and therefore not subject to the bound in Theorem 3.9, for example paths of odd order. However, we can get an upper bound via the matching number.

**Theorem 5.1.** Let G be a graph with n vertices. Then  $\operatorname{avd}(G) \leq n - \frac{2\nu(G)}{3}$ .

Proof. We begin by showing  $|a(S)| \leq 2(n - \nu(G)) - |S|$  for every  $S \in \mathcal{D}(G)$ . By Lemma 3.2,  $|a_1(S)| \leq |N_1(S)|$ . Therefore it suffices to show  $|a_2(S)| \leq |N_2(S)| + n - 2\nu(G)$ . For every  $v \in a_2(S)$ ,  $N(v) \subseteq V - S$  otherwise  $\operatorname{Priv}_S(v) \neq \{v\}$ . Furthermore  $N(v) \subseteq N_2(S)$  otherwise  $v \in a_1(S)$ . Fix a maximum matching in G. Each vertex in  $a_2(S)$  is either unmatched or matched with a vertex in  $N_2(S)$ . Note there are at most  $n - 2\nu(G)$  unmatched vertices in G. Therefore  $|a_2(S)| \leq |N_2(S)| + n - 2\nu(G)$ .

Finally as  $|a(S)| \le 2(n - \nu(G)) - |S|$  then  $\sum_{S \in \mathcal{D}(G)} |a(S)| \le 2(n - \nu(G))D(G, 1) - D'(G, 1)$ . Thus together with Lemma 3.1 we obtain

$$2D'(G,1) - nD(G,1) \le 2(n - \nu(G))D(G,1) - D'(G,1) \Rightarrow \operatorname{avd}(G) = \frac{D'(G,1)}{D(G,1)} \le n - \frac{2\nu(G)}{3}.$$

Another avenue of research is investigating the monotonicity of  $\operatorname{avd}(G)$  with respect to vertex or edge deletion. For example the removal of any edge or vertex in a graph decreases the number of dominating sets. However this is not necessarily the case for  $\operatorname{avd}(G)$ . Let G be the graph pictured in Figure 5.1.



Figure 5.1: A vertex labelled graph

It is not difficult to determine  $D(G, x) = x^6 + 6x^5 + 12x^4 + 10x^3 + 5x^2 + x$  and therefore avd(G) = 25/7. However

- $\operatorname{avd}(G v_1) = \frac{58}{19} < \operatorname{avd}(G) < \frac{13}{3} = \operatorname{avd}(G v_1)$
- $\operatorname{avd}(G v_5 v_6) = \frac{39}{11} < \operatorname{avd}(G) < \frac{78}{19} = \operatorname{avd}(G v_1 v_4)$

Despite this the following conjecture holds for all graphs on up to 7 vertices

**Conjecture 5.2.** For a nonempty graph G there exists a vertex v and edge e such that

 $\operatorname{avd}(G-v) < \operatorname{avd}(G) < \operatorname{avd}(G-e).$ 

# References

- S. Akbari, S. Alikhani, and Y. H. Peng. Characterization of graphs using domination polynomials. *Eur. J. Comb.*, 31:1714–1724, 2010.
- [2] S. Alikhani. Dominating sets and domination polynomials of graphs. Lambert Academic Publishing, first edition, 2012.
- [3] Saeid Alikhani. On the domination polynomials of non P<sub>4</sub>-free graphs. Iran. J. Math. Sci. Inform., 8(2):49–55, 2013.
- [4] Eric O.D. Andriantiana, Valisoa Razanajatovo Misanantenaina, and Stephan Wagner. The average size of independent sets of graphs. *Eur. J. Math.*, 2019.
- [5] Eric O.D. Andriantiana, Valisoa Razanajatovo Misanantenaina, and Stephan Wagner. The Average Size of Matchings in Graphs. *Graphs Comb.*, 36(3):539–560, 2020.
- [6] Claude Berge. Some common properties for regularizable graphs, edge-critical graphs and B-graphs, volume 108. Springer-Verlag, Berlin, 1980.
- [7] Mostafa Blidia, Mustapha Chellali, Teresa W Haynes, and Michael A. Henning. Independent and double domination in trees. Util. Math., 70, 2006.
- [8] Dorota Brod and Zdzis law Skupien. Trees with extremal numbers of dominating sets. Australas. J. Comb., 35:273–290, 2006.
- [9] R.A. Brualdi. Introductory Combinatorics. Pearson/Prentice Hall, fifth edition, 2010.
- [10] J. N. Darroch. On the Distribution of the Number of Successes in Independent Trials. Ann. Math. Stat., 35(3):1317–1321, 1964.
- [11] J.K. Doyle and J.E. Graver. Mean Distance in a Graph. Discrete Math., 17:147–154, 1977.
- [12] Frank Harary and Teresa W Haynes. Double domination in graphs. Ars Comb., 55:201–213, 2000.
- [13] J. Haslegrave. Extremal results on average subtree density of series reduced trees. J. Comb. Theory, Ser. B, 107:26–41, 2014.
- [14] T.W. Haynes, S. Hedetniemi, and P.J. Slater. Fundamentals of domination in graphs. Marcel Dekker, 1998.
- [15] Robert E Jamison. On the Average Number of Nodes in a Subtree of a Tree. J. Comb. Theory, Ser. B, 35:207–223, 1983.

- [16] Robert E Jamison. Monotonicity of the Mean Order of Subtrees. J. Comb. Theory, Ser. B, 37:70-78, 1984.
- [17] T. Kotek, J. Preen, F. Simon, P. Tittmann, and M. Trinks. Recurrence relations and splitting formulas for the domination polynomial. *Electron. J. Comb.*, 19:1–27, 2012.
- [18] Mohammad Reza Oboudi. On the roots of domination polynomial of graphs. Discret. Appl. Math., 205:126–131, 2016.
- [19] Andrew Vince and Hua Wang. The average order of a subtree of a tree. J. Comb. Theory, Ser. B, 100(2):161– 170, 2010.
- [20] Stephan Wagner and Hua Wang. On the Local and Global Means of Subtree Orders. J. Graph Theory, 70560:154–166, 2015.
- [21] Harry Wiener. Structural determination of paraffin boiling points. J. Am. Chem. Soc., 69(1):17–20, 1947.