ON THE AVERAGE SIZE OF 1-NEARLY INDEPENDENT VERTEX SETS IN GRAPHS

AUDACE A. V. DOSSOU-OLORY AND ERIC O. ANDRIANTIANA

ABSTRACT. A k-nearly independent vertex subset of a graph G is a set of vertices that induces a subgraph containing exactly k edges. For $k = 0$, this coincides with the classical notion of independent subsets. This paper investigates the average size, $av_1(G)$ of the 1-nearly independent vertex subsets of both graphs and trees of a given order n.

Let E_n denote the n-vertex edgeless graph, so that $av_1(E_n) = 0$. We determine all n-vertex graphs $G \neq E_n$ that minimize or maximize av_1. Similarly, we identify the trees of order n that achieve the minimum value of av_1, and prove that the maximum value lies between $n/2$ and $(n + 1)/2$ if $n > 8$. Finally, we construct a family of n-vertex trees which shows that the bounds are asymptotically sharp.

1. INTRODUCTION

Counting substructures in graphs is a central topic in combinatorics and graph theory. Among these, the enumeration of subgraphs—particularly induced subgraphs, spanning subgraphs, independent sets and their generalisations—plays a crucial role in understanding graph complexity and structure. Key questions include which n-vertex graphs minimise or maximise the number or average size of independent sets, under various structural constraints?

An independent vertex set in a graph G is a set of vertices, no two of which are adjacent. The number of independent vertex sets of G is also known as the Fibonacci number or Merrifield- Simmons index of G in the mathematical literature. Their enumeration has deep ties to certain models in statistical physics via the hard-core model, where each independent set is assigned a weight based on its size (see [9]).

Let $i(G)$ denote the number of independent vertex sets (including the empty set) in graph G. Among n-vertex graphs, the minimum is achieved by the complete graph K_n with $i(K_n) = n + 1$. For graphs with a fixed number of vertices and edges, a classic result of Kahn [8] shows that the number of independent sets is maximized by a disjoint union of complete bipartite graphs. Specifically, among d-regular bipartite graphs G on n vertices such that $2d|n$, the maximum value of $i(G)$ is achieved by the disjoint union of $n/2d$ copies of $K_{d,d}$, confirming a conjecture of Alon and Kahn. In general graphs, Zhao [11] extended Kahn's entropy method to show that among all graphs with maximum degree Δ, the number of independent vertex sets is asymptotically maximized by a disjoint union of $K_{\Delta,\Delta}$. Among regular graphs, the minimum is less understood, but lower bounds

2020 *Mathematics Subject Classification.* Primary 05C30, 05C69; secondary 05C35, 05C75.

Key words and phrases. 1-nearly independent vertex sets, extremal graph structures, average size.

have been given in terms of vertex degrees. There are many other work on extremising $i(G)$ when G is restricted to various classes of n-vertex graphs such as graphs with a given minimum degree [6, 7], or given degree sequence [2].

A *1-nearly independent vertex set* allows exactly one pair of adjacent vertices in the collection. This concept was introduced recently by Andriantiana and Shozi in [1], where various of its properties are discussed and extremal n-vertex graphs and n-vertex trees are characterised.

The analogous edge version [4] counts edge-subsets with one adjacent pair; the star minimises and the path maximises this number among n-vertex trees. These new generalisations open further directions for extremal and enumerative analyses.

Let $\mathcal{S}_0(G)$ denote the set of all independent vertex subsets of graph $G = (V, E)$. Define the average size of independent vertex sets of G by

$$\bar{s}(G) := \frac{1}{i(G)} \sum_{I \in \mathcal{S}_0(G)} |I|.$$

Andriantiana et al. [3] showed that among all n-vertex graphs, $\bar{s}$ is maximised by the edgeless graph and minimised by K_n, while among trees, the path minimises $\bar{s}$ and the star maximises it. See also the old short paper [10] which studies the asymptotic behaviour of $\bar{s}(G)$ in terms of order and the chromatic number of G, and the recent results [5] giving a tight lower bound on the expected size of a uniformly choosen independent set in triangle-free graphs of bounded degree.

In this manuscript, we investigate the average size av_1 of 1-nearly independent subsets in the class of graphs of fixed order n, and subsequently in the class of trees of order n. Both upper and lower bounds are established, along with characterizations of the graphs that attain these bounds.

The paper is structured as follows. In Section 2, we formally define av_1, compute explicit expressions of av_1 for graphs that have less complicated structures, and establish some technical recursive formulas. These formulas will later be needed in proving our main theorems. Section 3 is devoted to the study of average among n-vertex graphs. In particular, graphs with the smallest and largest av_1 are characterized. In Section 4 we show that the star is the n-vertex tree with the smallest av_1. We also prove that the maximum value of av_1 is less than $(n+1)/2$ for $n > 8$. Moreover, we characterise an explicit family of n-vertex trees R_n with $av_1(R_n) = n/2 + \delta_n$ for some $0 < \delta_n < 1/2$, thus matching our bounds. Finally, we provide further relations between 0- and 1-nearly independent vertex sets in graphs in Section 5.

2. Preliminary results

Throughout this paper, all graphs are simple. The number of elements of a finite set S is denoted by $|S|$. The empty sum is treated as 0.

Let $G = (V, E)$ be a simple graph and l a non-negative integer. As defined in [1], a set $S \subseteq V$ is called a l-nearly independent vertex set of G if the graph induced by S in G contains exactly l edges. Thus 0-nearly independent vertex sets are precisely the classical independent vertex sets. Denote by $\mathcal{S}_l(G)$ the set of all l-nearly independent vertex sets of G, and set

$$\sigma_l(G) := |\mathcal{S}_l(G)| \quad \text{and} \quad S_l(G) := \sum_{S \in \mathcal{S}_l(G)} |S| \,,$$

that is $\sigma_l(G)$ is the number of l-nearly independent vertex sets of G, while $S_l(G)$ is the sum of the sizes of the l-nearly independent vertex sets of G. We are specifically interested in the average size of a l-nearly independent vertex set of G, which is given by

$$av_l(G) := \frac{S_l(G)}{\sigma_l(G)} \,,$$

provided that $\sigma_l(G) \neq 0$. If $\mathcal{S}_l(G) = \emptyset$, then it is consistent to write $av_l(G) = 0$.

We denote by $\sigma_l(G,k)$ the number of l-nearly independent vertex sets of G that have cardinality k. Clearly, $\sigma_0(G,0) = 1$, while $\sigma_l(G,0) = 0$ for all $l > 0$ and $\sigma_1(G,k) = 0$ for all $k < 2$. We define the l-nearly series $I_l(G;x)$ by

$$I_l(G;x) = \sum_{k \geq 0} \sigma_l(G,k) x^k \,,$$

which has a finite number of terms since $\sigma_l(G,k) = 0$ for all $k > |V(G)|$. We refer to $I_l(G;x)$ as the l-nearly independent vertex polynomial of G. It follows that

$$\begin{aligned} \sigma_l(G) &= \sum_{k \geq 0} \sigma_l(G,k) = I_l(G;1) \quad \text{and} \\ S_l(G) &= \sum_{k \geq 0} k \cdot \sigma_l(G,k) = \frac{d}{dx} I_l(G;x)\Big|_{x=1} := I_l'(G;1) \,. \end{aligned}$$

An edge e with end vertices u, v in a graph G will be denoted by $e = uv$. Let G_1 and G_2 be two graphs such that their vertex sets are disjoints. Then $G_1 \cup G_2$ means the graph with vertex set $V(G_1) \cup V(G_2)$ and edge set $E(G_1) \cup E(G_2)$. In this case, if S is a 1-nearly independent vertex set of $G_1 \cup G_2$, then the unique edge $e = uv$ in the subgraph induced by S necessarily comes from G_1 or G_2, and $S \backslash \{u, v\}$ is an independent vertex set of $G_1 \cup G_2$. Conversely, $S \cap V(G_j)$ is either an independent or a 1-nearly independent vertex set of G_j for $j \in \{1, 2\}$, depending on the location of edge $e = uv$.

We therefore obtain the following identities.

Lemma 1. *Let G_1 and G_2 be two vertex disjoint graphs. We have*

$$\begin{aligned} \sigma_1(G_1 \cup G_2, k) &= \sum_{l=2}^{k} \sigma_1(G_1, l)\sigma_0(G_2, k-l) + \sigma_1(G_2, l)\sigma_0(G_1, k-l) \quad \textit{and} \\ \sigma_1(G_1 \cup G_2) &= \sigma_1(G_1)\sigma_0(G_2) + \sigma_1(G_2)\sigma_0(G_1) \,. \end{aligned}$$

We also obtain the following.

Lemma 2. *Let G_1 and G_2 be two vertex disjoint graphs. We have*

$$I_1(G_1 \cup G_2; x) = I_1(G_1; x) I_0(G_2; x) + I_1(G_2; x) I_0(G_1; x) \quad \text{and} \tag{1}$$

$$S_1(G_1 \cup G_2) = S_1(G_1)\sigma_0(G_2) + \sigma_1(G_1)S_0(G_2) + S_1(G_2)\sigma_0(G_1) + \sigma_1(G_2)S_0(G_1)\,. \tag{2}$$

Proof. Formula (1) follows immediately from Lemma 1. For (2), simply note that

$$\begin{aligned} I_1'(G_1 \cup G_2; x) = I_1'(G_1; x) I_0(G_2; x) + I_1(G_1; x) I_0'(G_2; x) \\ + I_1'(G_2; x) I_0(G_1; x) + I_1(G_2; x) I_0'(G_1; x)\,. \end{aligned}$$

□

Let $G = (V, E)$ be a graph. For $S \subseteq V$, we denote by $G - S$ the graph induced by $V \backslash S$ in G. For simplicity, we write $G - v$ instead of $G - \{v\}$. We use the standard notation that $N(v)$ is the set of all neighbours of $v \in V$ in G and $N[v] = N(v) \cup \{v\}$.

We can count the contribution of v to $\sigma_1(G, k)$ as follows. If v is an element of $S \in \mathcal{S}_1(G)$, then S can only contain at most one neighbour of v in G. If S contains both v and $u \in N(v)$, then no other element of $N(v) \cup N(u)$ can belong to S and $S\backslash\{u, v\}$ is an independent vertex set of G. Thus the contribution of v to $\sigma_1(G, k)$ is given by

$$\sigma_1(G - N[v], k - 1) + \sum_{u \in N(v)} \sigma_0(G - (N[v] \cup N[u]), k - 2)$$

and thus

$$\sigma_1(G, k) = \sigma_1(G - v, k) + \sigma_1(G - N[v], k - 1) + \sum_{u \in N(v)} \sigma_0(G - (N[v] \cup N[u]), k - 2)\,.$$

Note that $G - (N[v] \cup N[u])$ and $G - (N(v) \cup N(u))$ are isomorphic graphs since $uv \in E$.

The next lemma follows immediately, and its proof is omitted. Note that the recursive formula for S_1 is obtained by differentiating the expression for I_1 and evaluating it at $x = 1$.

Lemma 3. *Let G be a graph and v a vertex of G. We have*

$$I_1(G; x) = I_1(G - v; x) + x I_1(G - N[v]; x) + x^2 \sum_{u \in N(v)} I_0\big(G - (N(v) \cup N(u)); x\big)\,,$$

$$\sigma_1(G) = \sigma_1(G - v) + \sigma_1(G - N[v]) + \sum_{u \in N(v)} \sigma_0\big(G - (N(v) \cup N(u))\big)\,, \qquad \text{and}$$

$$\begin{aligned} S_1(G) &= S_1(G - v) + \sigma_1(G - N[v]) + S_1(G - N[v]) \\ &\quad + \sum_{u \in N(v)} \Big(2\sigma_0\big(G - (N(v) \cup N(u))\big) + S_0\big(G - (N(v) \cup N(u))\big)\Big)\,. \end{aligned}$$

In what follows, we compute $\sigma_1(G), S_1(G)$ for a few classes of graphs. We use the following usual notation for n-vertex graphs:

- E_n: graph with no edges (edgeless graph),
- S_n: star (a central vertex adjacent to $n - 1$ pendent edges),

- K_n: complete graph (every two vertices are adjacent),
- P_n: path (a tree with maximum degree less than or equal to 2).

Proposition 4.

$$\begin{aligned}
&\sigma_1(E_n) = S_1(E_n) = 0\,,\\
&\sigma_1(S_n) = n-1,\ S_1(S_n) = 2(n-1)\,, \qquad \textit{and}\\
&\sigma_1(K_n) = n(n-1)/2,\ S_1(K_n) = n(n-1).
\end{aligned}$$

Proof. All these expressions are easily obtained by direct counting arguments. Nevertheless, let us use the above decomposition formulas. E_n has no edge, so $\mathcal{S}_1(E_n) = \emptyset$. Let v be the central vertex of S_n, we have $S_n - v = E_{n-1}$, $S_n - N(v) = \emptyset$, and

$$\sigma_1(S_n) = \sigma_1(E_{n-1}) + \sigma_1(\emptyset) + \sum_{u\in N(v)} \sigma_0(\emptyset) = |N(v)| = n-1\,,$$

$$S_1(S_n) = S_1(E_{n-1}) + \sigma_1(\emptyset) + S_1(\emptyset) + \sum_{u\in N(v)} (2\sigma_0(\emptyset) + S_0(\emptyset)) = 2|N(v)| = 2(n-1)\,.$$

For the complete graph,

$$\sigma_1(K_n) = \sigma_1(K_{n-1}) + \sigma_1(\emptyset) + \sum_{u\in N(v)} \sigma_0(\emptyset) = \sigma_1(K_{n-1}) + |N(v)| = \sigma_1(K_{n-1}) + n-1\,,$$

$$\begin{aligned}
S_1(K_n) &= S_1(K_{n-1}) + \sigma_1(\emptyset) + S_1(\emptyset) + \sum_{u\in N(v)} \big(2\sigma_0(\emptyset) + S_0(\emptyset)\big)\\
&= S_1(K_{n-1}) + 2|N(v)| = S_1(K_{n-1}) + 2(n-1)\,.
\end{aligned}$$

Solving these recursions yield $\sigma_1(K_n) = n(n-1)/2$ and $S_1(K_n) = n(n-1)$. □

Let $G = (V, E)$ be a graph. Since every element of $\mathcal{S}_1(G)$ induces a graph that contains only one edge, the set $\mathcal{S}_1(G)$ can be partitioned into

$$\mathcal{S}_1(G) = \bigcup_{e\in E} \bigcup_{k\geq 2} \mathcal{S}^e_{1,k}(G)\,,$$

where $\mathcal{S}^e_{1,k}(G)$ comprises all those k-element 1-nearly independent vertex sets involving the end vertices of edge e. One can therefore determine $|\mathcal{S}^e_{1,k}(G)|$, the contribution of edge $e = uv$ to $\sigma_1(G, k)$. Following the same argument given above, we get

$$\begin{aligned}
|\mathcal{S}^e_{1,k}(G)| &= \sigma_0\big(G - (N(v)\cup N(u)), k-2\big)\,,\\
\sigma_1(G,k) &= \sum_{uv\in E} \sigma_0\big(G - (N(v)\cup N(u)), k-2\big)\,.
\end{aligned}$$

As an immediate consequence, we obtain the following formulas:

Lemma 5. *For every graph G, it holds that*

$$I_1(G;x) = x^2 \sum_{uv\in E} I_0(G - (N(v)\cup N(u)); x)\,,$$

$$\begin{aligned}
\sigma_1(G) &= \sum_{uv\in E} \sigma_0\big(G-(N(v)\cup N(u))\big)\,,\\
S_1(G) &= \sum_{uv\in E} \Big(2\sigma_0\big(G-(N(v)\cup N(u))\big)+S_0\big(G-(N(v)\cup N(u))\big)\Big)\\
&= \sum_{uv\in E} \sigma_0\big(G-(N(v)\cup N(u))\big)\Big(2+av_0\big(G-(N(v)\cup N(u))\big)\Big)\,.
\end{aligned}$$

For example, setting $E(P_n) := \{v_0v_1, v_1v_2, \dots, v_{n-2}v_{n-1}\}$ and using this lemma we can compute $\sigma_1(P_n)$ and $S_1(P_n)$ as

$$\sigma_1(P_n) = \sum_{j=1}^{n-1} \sigma_0(P_{j-2}\cup P_{n-2-j}) = \sum_{j=1}^{n-1} \sigma_0(P_{j-2})\sigma_0(P_{n-2-j})\,,$$

where the second equality comes from Lemma 6 below.

Lemma 6 ([3])**.** *Let G_1 and G_2 be two vertex disjoint graphs. We have*

$$\begin{aligned}
\sigma_0(G_1\cup G_2) &= \sigma_0(G_1)\sigma_0(G_2)\,,\\
av_0(G_1\cup G_2) &= av_0(G_1)+av_0(G_2)\,.
\end{aligned}$$

The next lemma provides an alternative formula for $av_1(G)$ that will be useful in some cases.

Lemma 7. *If G is not an edgeless graph, then*

$$av_1(G) = \frac{S_1(G)}{\sigma_1(G)} = 2 + \frac{\sum_{uv\in E} S_0\big(G-(N(v)\cup N(u))\big)}{\sum_{uv\in E} \sigma_0\big(G-(N(v)\cup N(u))\big)}\,.$$

Furthermore, there exists edges u_1v_1, u_2v_2 of G such that

$$2+av_0\big(G-(N(v_1)\cup N(u_1))\big) \le av_1(G) \le 2+av_0\big(G-(N(v_1)\cup N(u_1))\big)\,.$$

Proof. According to Lemma 5,

$$av_1(G) = 2 + \frac{\sum_{uv\in E} S_0\big(G-(N(v)\cup N(u))\big)}{\sum_{uv\in E} \sigma_0\big(G-(N(v)\cup N(u))\big)}\,.$$

Now set $H_{uv} := G-(N(v)\cup N(u))$ and note that

$$av_1(G) = 2 + \frac{\sum_{uvE} S_0(H_{uv})}{\sum_{uv\in E} \sigma_0(H_{uv})} = 2 + \sum_{uv\in E} \alpha_{uv}\frac{S_0(H_{uv})}{\sigma_0(H_{uv})} = 2+\sum_{uv\in E}\alpha_{uv}av_0(H_{uv})\,,$$

where $\alpha_{uv} = \sigma_0(H_{uv})/\sum_{uv\in E}\sigma_0(H_{uv})$. Since $\alpha_{uv} > 0$ and $\sum_{uv\in E}\alpha_{uv} = 1$, we can interpret $\sum_{uv\in E}\alpha_{uv}\frac{S_0(H_{uv})}{\sigma_0(H_{uv})}$ as an average. Hence, there is a case of uv that corresponds to a term not less than the average, and there is a case that corresponds to a term not larger than the average. The lemma follows. □

In the next section, we consider graphs with a given order n.

3. Extremal n-vertex graphs

In this section, we characterise the structure of graphs that minimise or maximise av_1. Trivially, $av_1(G) \geq 0$ with equality if and only if G is an edgeless graph. Here and in the following, we assume $G \neq E_n$ for all n.

In [3] it was proved that the star maximises av_0 among all n-vertex trees, while [1] shows that the star minimises σ_1 among all n-vertex connected graphs. It will be shown that the star also uniquely achieves the minimum value of av_1 among connected graphs.

In paper [1], a *good graph* $G = (V, E)$ is defined as a graph in which every edge uv satisfies $N(v) \cup N(u) = V$. A full characterisation of the set $\mathcal{H}$ of all good graphs is provided there. For any $H \in \mathcal{H} \setminus \{K_1\}$, we have $av_1(H) = 2$. Hence we have the following first main theorem of this paper.

Theorem 8. *For every n-vertex graph G that is not edgeless, we have*

$$av_1(G) \geq 2\,.$$

Equality holds if and only if $G \in \mathcal{H}\backslash\{K_1\}$.

In particular, S_n uniquely minimises av_1 among n-vertex trees, and it is one of the n-vertex graphs that minimises av_1.

Proof. By Lemma 7, $av_1(G) \geq 2$ for all $G = (V, E)$ that are not edgeless, and equality holds if and only if $S_0\big(G-(N(v)\cup N(u))\big) = 0$ for all $uv \in E$, i.e. every $G-(N(v)\cup N(u))$ is the empty graph, or equivalently $N(v) \cup N(u) = V$.

The only edgeless graph from $\mathcal{H}$ is K_1. □

We use Lemma 7 to obtain some bounds on av_1. For a graph $G = (V, E)$, define

$$\delta_1(G) = \min_{uv\in E}|N(v)\cup N(u)| \quad \text{and} \quad \delta_2(G) = \max_{uv\in E}|N(v)\cup N(u)|\,.$$

It is clear that $2 \leq \delta_1(G) \leq \delta_2(G) \leq |V|$ for all G that are not edgless.

Proposition 9. *For every n-vertex graph $G \neq E_n$,*

$$2 \leq \frac{3n+2-3\delta_2(G)}{n+1-\delta_2(G)} \leq av_1(G) \leq \frac{n+4-\delta_1(G)}{2} \leq \frac{n+2}{2}\,.$$

Proof. Using u_1, v_1, u_2 and v_2 as in Lemma 7 and the result

$$\frac{n}{n+1} = av_0(K_n) \leq av_0(H) \leq av_0(E_n) = \frac{n}{2}$$

that holds for all n-vertex graph H (see [3]), we have

$$2 + \frac{n - |N(v_1) \cup N(u_1)|}{n + 1 - |N(v_1) \cup N(u_1)|} \leq av_1(G) \leq 2 + \frac{n - |N(v_2) \cup N(u_2)|}{2} .$$

The function $x \mapsto (a - x)/(b - x)$ is decreasing in x, provided that $a < b$. It follows that

$$2 + \frac{n - \delta_2(G)}{n + 1 - \delta_2(G)} \leq av_1(G) \leq 2 + \frac{n - \delta_1(G)}{2} \leq 2 + \frac{n - 2}{2} .$$

□

We now aim to obtain a sharp upper bound on av_1.

Proposition 10 ([3]). *For every n-vertex graph G that is not edgeless, we have*

$$av_0(G) < av_0(E_n) = n/2.$$

Let $av_1(G, e)$ denote the average of all 1-nearly independent subsets of G that contain the endvertices of edge e. For $e = uv$, we have

$$av_1(G, e) = 2 + av_0\big(G - (N(u) \cup N(v))\big) .$$

Moreover,

$$av_1(G) = 2 + \sum_{uv \in E} \alpha_{uv} \cdot av_0\big(G - (N(v) \cup N(u))\big)$$

for certain $\alpha_{uv} > 0$ such that $\sum_{uv \in E} \alpha_{uv} = 1$, see the proof of Lemma 7. Since $av_0(E_m)$ increases with m, the above identities imply that

$$av_1(G) \leq \max_{e \in E(G)} av_1(G, e) \leq 2 + av_0(E_{n-2})$$

holds for every graph G with n vertices.

Define $G_n := K_2 \cup (n - 2)K_1$ for $n \geq 6$. This graph G_n contains only one edge, thus

$$av_1(G_n) = 2 + av_0(E_{n-2}) .$$

Hence, our next theorem follows:

Theorem 11. *Let $n \geq 6$. For every n-vertex graph G,*

$$av_1(G) \leq \frac{n}{2} + 1 .$$

Equality holds if and only if $G = G_n$.

We now shift our focus to the family of n-vertex trees.

4. Extremal n-vertex trees

As a particular case of Theorem 8, the star S_n uniquely minimises av_1 among n-vertex trees. The next task is to search for n-vertex trees with the maximum average size of 1-nearly independent vertex sets.

We call an internal vertex in a tree T and vertex with degre greater than 1.

Proposition 12. *For any n-vertex tree T with minimum degree of internal vertices equal to δ', we have*

$$av_1(T) \leq 2 + \frac{n-\delta'-1}{2}\,.$$

Proof. As already seen in the proof of Lemma 7,

$$av_1(T) = 2 + \sum_{uv\in E(T)} \alpha_{uv} \cdot av_0\big(T-(N(u)\cup N(v))\big)$$

for some positive numbers α_{uv} that sum to 1. If T has n vertices and minimum internal degree δ', then by Proposition 10,

$$av_0\big(T-(N(u)\cup N(v))\big) \leq \frac{\max\{n-\delta'-1,0\}}{2}$$

for every $uv \in E(T)$. □

We can then deduce the following upper bound for n-vertex trees, obtained for $\delta' \in \{1,2\}$. The case of equality for $T = P_i$ with $2 \leq i \leq 4$, can be checked with easy calculations.

Corollary 13. *For any n-vertex tree T, we have*

$$av_1(T) \leq 2 + \frac{\max\{n-3,0\}}{2}$$

with equality holding for $T \in \{P_2, P_3, P_4\}$.

We will prove a matching lower bound on the maximum average size of 1-nearly independent vertex sets in trees.

For $n \geq 4$, define R_n to be the tree obtained by subdividing once an edge of the star S_{n-1}, i.e. by adding an edge between a leaf of S_{n-1} and a new vertex.

Let v be a leaf of R_n such that v is adjacent to a vertex u of degree 2. With this choice of vertex v, we apply Lemma 3 alongside Proposition 4 to obtain

$$\begin{aligned} av_1(R_n) &= \frac{S_1(S_{n-1}) + S_1(S_{n-2}) + \sigma_1(S_{n-2}) + S_0(E_{n-3}) + 2\sigma_0(E_{n-3})}{\sigma_1(S_{n-1}) + \sigma_1(S_{n-2}) + \sigma_0(E_{n-3})} \\ &= \frac{2(n-2) + 2(n-3) + (n-3) + (n-3)2^{n-4} + 2\times 2^{n-3}}{(n-2)+(n-3)+2^{n-3}}\,, \end{aligned}$$

where in the last step we used the formulas

$$\sigma_0(E_m) = 2^m\,, \quad S_0(E_m) = m2^{m-1}\,.$$

Thus

$$\begin{aligned} av_1(R_n) &= \frac{5n-13+(n+1)2^{n-4}}{2n-5+2^{n-3}} \\ &= \frac{n}{2} + \frac{2^{n-4}-n^2+15n/2-13}{2n-5+2^{n-3}}\,, \end{aligned}$$

and we also have

$$\begin{aligned} av_1(R_n) &= \frac{5n-13+(n+1)2^{n-4}}{2n-5+2^{n-3}} \\ &= \frac{n+1}{2} - \frac{n^2-13n/2+21/2}{2n-5+2^{n-3}} < \frac{n+1}{2}\,. \end{aligned} \tag{3}$$

Note that

$$\lim_{n\to\infty} av_1(R_n) - \frac{n+1}{2} = 0$$

and that

$$av_1(R_n) = \frac{n}{2} + \frac{2^{n-4}-n^2+15n/2-13}{2n-5+2^{n-3}} > \frac{n}{2}\,,$$

for all n, except $n \in \{6,7,8\}$.

Theorem 14. *Let $n > 8$. Among n-vertex trees, the maximum average size of 1-nearly independent vertex sets is given by $n/2 + \delta_n$ for some function $0 < \delta_n < 1/2$. Moreover, both the upper and lower bounds are asymptotically sharp.*

At this stage, it is reasonable to make the following conjecture.

Conjecture 1. The tree R_n reaches the maximum value of av_1 over the set of all n-vertex trees.

We conclude this work with further bounds relating σ_0 and σ_1.

5. Further bounds

In this short section, we establish further identities and bounds. In particular, we show that it is possible to have $\sigma_0 \leq 3\sigma_1$. We begin with a lemma.

Lemma 15. *For every graph $G = (V, E)$ that is not edgeless and every $uv \in E$, we have*

$$\frac{1}{2^{l_{uv}-2} + 2^{l_{uv}-m_u} + 2^{l_{uv}-m_v} + 1} \leq \frac{\sigma_0(G - (N[v] \cup N[u]))}{\sigma_0(G)} \leq 1 - \frac{\sigma_0(G-v)}{\sigma_0(G)}\,,$$

where

$$l_{uv} = |N(v) \cup N(u)|\,, \quad m_u = |N[u]|\,, \quad m_v = |N[v]|\,.$$

In particular,

$$\frac{\sigma_0(G - (N[v] \cup N[u])}{\sigma_0(G)} \geq \frac{1}{(3 \cdot 2^{l_{uv}-2} + 1)}\,.$$

Proof. We return to the recursion

$$\sigma_0(G) = \sigma_0(G - v) + \sigma_0(G - N[v])\,.$$

Since $G - (N[v] \cup N[u])$ is an induced subgraph of $G - N[v]$, we have

$$\begin{aligned}\sigma_0(G - N[v]) &\geq \sigma_0(G - (N[v] \cup N[u]))\,,\\ 1 &\geq \frac{\sigma_0(G-v)}{\sigma_0(G)} + \frac{\sigma_0(G - (N[v] \cup N[u]))}{\sigma_0(G)}\,,\end{aligned}$$

which proves the second inequality. On the other hand, we iterate the inequality

$$\sigma_0(G) = \sigma_0(G - w) + \sigma_0(G - N[w]) \leq 2\sigma_0(G - w)\,,$$

which holds for all $w \in V$, to obtain

$$\begin{aligned}\sigma_0((G-v)-u) &\leq 2^{|(N[v]\cup N[u])\setminus\{u,v\}|}\sigma_0\big(G - (N[v] \cup N[u])\big)\\ &= 2^{l_{uv}-2}\sigma_0\big(G - (N[v] \cup N[u])\big)\,,\\ \sigma_0((G-v)-N[u]) &\leq 2^{|(N[v]\cup N[u])\setminus(\{v\}\cup N[u])|}\sigma_0\big(G - (N[v] \cup N[u])\big)\\ &= 2^{l_{uv}-m_u}\sigma_0\big(G - (N[v] \cup N[u])\big)\end{aligned}$$

and

$$\begin{aligned}\sigma_0((G-N[v])-u) &\leq 2^{|(N[v]\cup N[u])\setminus(\{u\}\cup N[v])|}\sigma_0\big(G - (N[v] \cup N[u])\big)\\ &= 2^{l_{uv}-m_v}\sigma_0\big(G - (N[v] \cup N[u])\big)\,.\end{aligned}$$

It follows that

$$\begin{aligned}\sigma_0(G) &= \sigma_0(G-v) + \sigma_0(G - N[v])\\ &= \sigma_0((G-v)-u) + \sigma_0((G-v)-N[u])\\ &\quad + \sigma_0((G-N[v])-u) + \sigma_0((G-N[v])-N[u])\\ &\leq (2^{l_{uv}-2} + 2^{l_{uv}-m_u} + 2^{l_{uv}-m_v} + 1)\sigma_0(G - (N[v] \cup N[u])\,,\end{aligned}$$

which proves the first inequality in the lemma. In particular,

$$\begin{aligned}\sigma_0(G) &\leq (2^{l_{uv}-2} + 2^{l_{uv}-2} + 2^{l_{uv}-2} + 1)\sigma_0(G - (N[v] \cup N[u])\\ &= (3 \cdot 2^{l_{uv}-2} + 1)\sigma_0(G - (N[v] \cup N[u])\,,\end{aligned}$$

proving the last inequality. □

In the next proposition, we show that σ_1/σ_0 is a somewhat additive function in the union of graphs, and that $\sigma_0 \leq 3\sigma_1$ can possibly hold in certain graph classes.

Proposition 16. *Let G_1 and G_2 be two vertex disjoint graphs. We have*

$$\frac{\sigma_1(G_1 \cup G_2)}{\sigma_0(G_1 \cup G_2)} = \frac{\sigma_1(G_1)}{\sigma_0(G_1)} + \frac{\sigma_1(G_2)}{\sigma_0(G_2)}\,.$$

Moreover, for every graph $G = (V, E)$ with no isolated vertex and maximum degree 2,

$$\frac{\sigma_1(G)}{\sigma_0(G)} \geq \frac{1}{3}\,.$$

Equality holds for P_2.

Proof. The first identity immediately follows from Lemma 2:

$$\frac{\sigma_1(G_1 \cup G_2)}{\sigma_0(G_1 \cup G_2)} = \frac{\sigma_1(G_1)\sigma_0(G_2) + \sigma_1(G_2)\sigma_0(G_1)}{\sigma_0(G_1)\sigma_0(G_2)}\,.$$

Now let $G = (V, E)$ be a graph with no isolated vertex and maximum degree 2. All the connected components of G are paths and/or cycles. If P_2 is such a component, then set $G_1 := P_2$ and $G_2 := G \backslash V(G_1)$ to obtain

$$\frac{\sigma_1(G)}{\sigma_0(G)} = \frac{\sigma_1(P_2)}{\sigma_0(P_2)} + \frac{\sigma_1(G_2)}{\sigma_0(G_2)} \geq \frac{\sigma_1(P_2)}{\sigma_0(P_2)} = \frac{1}{3}\,.$$

Otherwise,

$$l_{uv} := |N(v) \cup N(u)| \in \{3, 4\}$$

for all $uv \in E$. Here we distinguish two cases:

- Case 1: $|E| \geq 5$.
  We can partition E as $E = E_4 \cup E_3$, where

  $$E_4 = \{uv \in E : \ l_{uv} = 4\} \quad \text{and} \quad E_3 = \{uv \in E : \ l_{uv} = 3\}\,.$$

  We combine (see Lemma 15)

  $$\frac{\sigma_0(G - (N[v] \cup N[u])}{\sigma_0(G)} \geq \frac{1}{(3 \cdot 2^{l_{uv}-2} + 1)}\,.$$

  and (see Lemma 5)

  $$\frac{\sigma_1(G)}{\sigma_0(G)} = \sum_{uv \in E} \frac{\sigma_0\big(G - (N(v) \cup N(u))\big)}{\sigma_0(G)}$$

  to obtain

  $$\begin{aligned}\frac{\sigma_1(G)}{\sigma_0(G)} &\geq \sum_{uv \in E_4} \frac{1}{(3 \cdot 2^{4-2} + 1)} + \sum_{uv \in E_3} \frac{1}{(3 \cdot 2^{3-2} + 1)} = \frac{|E_4|}{13} + \frac{|E_3|}{7} \\ &\geq \frac{5 - |E_3|}{13} + \frac{|E_3|}{7} = \frac{35 + 6|E_3|}{91} > \frac{1}{3}\,.\end{aligned}$$

- Case 2: $1 \leq |E| \leq 4$.
  Here all the connected components of G must belong to the set

  $$Q := \{P_5, C_4, P_4, C_3, P_3\}\,,$$

  while easy calculations yield

  $$\frac{\sigma_1(P_5)}{\sigma_0(P_5)} = \frac{10}{13}\,, \quad \frac{\sigma_1(C_4)}{\sigma_0(C_4)} = \frac{4}{7}\,, \quad \frac{\sigma_1(P_4)}{\sigma_0(P_4)} = \frac{5}{8}\,,$$

$$\frac{\sigma_1(C_3)}{\sigma_0(C_3)} = \frac{3}{4}\,, \quad \frac{\sigma_1(P_3)}{\sigma_0(P_3)} = \frac{2}{5}\,.$$

All these values are greater than $1/3$. Now $G = G_1 \cup G_2$ for some $G_1 \in Q$. We have

$$\frac{\sigma_1(G_1 \cup G_2)}{\sigma_0(G_1 \cup G_2)} = \frac{\sigma_1(G_1)}{\sigma_0(G_1)} + \frac{\sigma_1(G_2)}{\sigma_0(G_2)} > \frac{1}{3}\,,$$

which concludes the proof.

□

References

[1] E. O. Andriantiana and Z. B. Shozi. The number of 1-nearly independent vertex subsets. *Quaestiones Mathematicae*, 47(12):2353–2373, 2024.

[2] E. O. D. Andriantiana. Energy, hosoya index and merrifield–simmons index of trees with prescribed degree sequence. *Discrete Applied Mathematics*, 161(6):724–741, 2013.

[3] E. O. D. Andriantiana, V. R. Misanantenaina, and S. Wagner. The average size of independent sets of graphs. *European Journal of Mathematics*, 6(6):561–576, 2020.

[4] E. O. D. Andriantiana and Z. B. Shozi. The number of 1-nearly independent edge subsets. *Iranian Journal of Mathematical Chemistry*, 16(1):65–84, 2025.

[5] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. On the average size of independent sets in triangle-free graphs. *Proceedings of the American Mathematical Society*, 146(11):4541–4553, 2018.

[6] D. Galvin. Two problems on independent sets in graphs. *Discrete Mathematics*, 311(20):2105–2112, 2011.

[7] W. Gan, P.-S. Loh, and B. Sudakov. Maximizing the number of independent sets of a fixed size. *Combinatorics, Probability and Computing*, 24(3):521–527, 2015.

[8] J. Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combinatorics, Probability and Computing*, 10(3):219–237, 2001.

[9] R. D. Martin. Statistical mechanics of the independent set problem. *Discrete Mathematics*, 309(10):3323–3331, 2009.

[10] A. Nilli. The average size of an independent set in graphs with a given chromatic number. *Discrete Mathematics*, 67(2):147–153, 1987.

[11] Y. Zhao. The number of independent sets in a regular graph. *Combinatorics, Probability and Computing*, 19(2):315–320, 2010.

Audace A. V. Dossou-Olory, Institut National de l'Eau, Centre d'Excellence d'Afrique pour l'Eau et l'Assainissement, and Institut de Mathématiques et de Sciences Physiques, Dangbo, Université d'Abomey-Calavi, Bénin, **ORCID** 0000-0003-2065-117X

Email address: audace@aims.ac.za

Eric O. D. Andriantiana, Department of Mathematics (Pure and Applied), Rhodes University, Makhanda, 6140 South Africa, and National Institute for Theoretical and Computational Sciences (NITheCS), Stellenbosch, South Africa

Email address: e.andriantiana@ru.ac.za