Communications in Combinatorics and Optimization
Vol. xx, No. x (xxxx), pp. 1-17
https://doi.org/10.22049/cco.2026.31202.2779

CCO
Commun. Comb. Optim.

Research Article

Average size of 1-nearly independent vertex sets

Audace A.V. Dossou-Olory[1], Eric O.D. Andriantiana[2,3,*]

[1]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
audace@aims.ac.za

[2]Department of Mathematics (Pure and Applied), Rhodes University, Makhanda, 6140 South Africa

[3]National Institute for Theoretical and Computational Sciences (NITheCS), Stellenbosch, South Africa
*e.andriantiana@ru.ac.za

Received: 9 November 2025; Accepted: 16 September 2026
Published Online: 23 September 2026

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 attains the upper bound with an additive error tending to zero, as n tends to infinity.

Keywords: 1-nearly independent vertex sets, extremal graph structures, average size.

AMS Subject Classification: Primary 05C30, 05C69; secondary 05C35, 05C75

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 generalizations–plays a crucial role in understanding graph complexity and structure. Key questions include which n-vertex graphs minimize or maximize the number or average size of independent sets, under various structural constraints.

* *Corresponding Author*

An independent vertex set in a graph G is a set of vertices, no two of which are adjacent in G. 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 [10]).

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 among d-regular bipartite graphs G on n vertices such that $2d$ divides n, the maximum value of $i(G)$ is achieved by the disjoint union of $n/2d$ copies of $K_{d,d}$, confirming a special case of a conjecture by Alon and Kahn. In general graphs, Zhao [11] extended Kahn's entropy method to show, as a special case of a broader result, that among all n-vertex, d-regular graphs (not necessarily bipartite), the number of independent vertex sets is maximized by the disjoint union of $n/2d$ copies of $K_{d,d}$, when n is a multiple of $2d$. Among regular graphs, the minimum is less understood, but lower bounds have been given in terms of vertex degrees.

There are many other works on $i(G)$ when G is restricted to various classes of n-vertex graphs such as graphs with a given minimum degree [6, 7], or a given degree sequence [1].

A *1-nearly independent vertex set* allows exactly one pair of adjacent vertices in the collection. This concept was recently introduced by Andriantiana and Shozi in [3], where various properties are discussed and extremal n-vertex graphs and n-vertex trees are characterized.

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

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

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

Andriantiana et al. [2] showed that among all n-vertex graphs, $\bar{s}$ is maximized by the edgeless graph E_n and minimized by the complete graph K_n, while among trees, the path P_n minimizes $\bar{s}$ and the star S_n maximizes it. See also the old short paper [9] which studies the asymptotic behaviour of $\bar{s}(G)$ in terms of both order and chromatic number of G, and the recent results [5] giving a tight lower bound on the expected size of a uniformly chosen independent set in triangle-free graphs of bounded degree.

In this manuscript, we investigate the average size av_1 of 1-nearly independent vertex 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 the bounds, except for the case of the upper bound for n-vertex trees.

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 the 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 characterize 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; see Section 5.

The contributions of the different sections are not of the same nature. Most of the extremal results for general graphs established in Section 3 are consequences of the averaging identity of Lemma 4, which relates the average size of 1-nearly independent vertex sets to the contributions of the edges of the graph. These results provide useful general bounds and, in the relevant cases, characterize the graphs attaining the corresponding extrema. The main new contribution of Section 4 is more specific to trees. We determine the exact minimum of av_1 among n-vertex trees, attained by the star, and establish that the maximum is asymptotically localized around $n/2$: for $n > 8$, it is strictly less than $(n+1)/2$. We further construct an explicit family of n-vertex trees R_n for which $av_1(R_n) = n/2 + \delta_n$ with $\lim_{n\to\infty} \delta_n = 1/2$, showing that the upper bound is essentially sharp. Thus, although the exact tree attaining the maximum is not identified here, the location of the maximum is determined up to a bounded error, and the family R_n provides explicit near-extremal trees approaching the upper bound. Accordingly, Section 4 is the central part of the paper.

2. Preliminary results

Throughout this paper, all graphs are simple. The number of elements in 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 [3], a set $S \subseteq V$ is called a l-nearly independent vertex set in 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 $\mathbb{S}_l(G)$ the set of all l-nearly independent vertex sets in G, and set

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

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

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

provided that $\sigma_l(G) \neq 0$. If $\mathbb{S}_l(G) = \emptyset$, 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 in 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 disjoint. 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 in $G_1 \cup G_2$, 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 in $G_1 \cup G_2$. Conversely, $S \cap V(G_j)$ is either an independent or a 1-nearly independent vertex set in 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*

$$I_1(G_1 \cup G_2; x) = I_1(G_1; x)I_0(G_2; x) + I_1(G_2; x)I_0(G_1; x) \quad \textit{and} \tag{2.1}$$

$$S_1(G_1 \cup G_2) = S_1(G_1)\sigma_0(G_2) + \sigma_1(G_1)S_0(G_2) + S_1(G_2)\sigma_0(G_1) + \sigma_1(G_2)S_0(G_1) \,. \tag{2.2}$$

Proof. Formula (2.1) follows immediately from Lemma 1. For (2.2), 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 neighbors 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 $S \in \mathbb{S}_1(G)$, then S can only contain at most one neighbor 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 in G. Thus the contribution of v to $\sigma_1(G, k)$ is given by

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

Therefore,

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

holds. 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*

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

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

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

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 incident to $n-1$ pendant 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 1.

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

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 $\mathbb{S}_1(E_n) = \emptyset$. 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 yields $\sigma_1(K_n) = n(n-1)/2$ and $S_1(K_n) = n(n-1)$. □

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

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

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

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

As an immediate consequence, we obtain the following formulas.

Lemma 4. *For every graph G, it holds that*

$$\begin{aligned} I_1(G;x) &= x^2 \sum_{uv\in E} I_0(G-(N(v)\cup N(u));x)\,,\\ \sigma_1(G) &= \sum_{uv\in E} \sigma_0\big(G-(N(v)\cup N(u))\big)\,,\\ S_1(G) &= \sum_{uv\in E} \Big(2\sigma_0\big(G-(N(v)\cup N(u))\big)+S_0\big(G-(N(v)\cup N(u))\big)\Big)\\ &= \sum_{uv\in E} \sigma_0\big(G-(N(v)\cup N(u))\big)\Big(2+av_0\big(G-(N(v)\cup N(u))\big)\Big)\,. \end{aligned}$$

For example, by defining both P_{-1} and P_0 as the graph with empty vertex set, setting $E(P_n) := \{v_0v_1, v_1v_2, \ldots, v_{n-2}v_{n-1}\}$, 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 5 below.

Lemma 5 ([2]). *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 6. *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 exist edges u_1v_1, u_2v_2 of G such that

$$2+av_0\big(G-(N(v_1)\cup N(u_1))\big) \le av_1(G) \le 2+av_0\big(G-(N(v_2)\cup N(u_2))\big)\,.$$

Proof. According to Lemma 4,

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

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

$$av_1(G) = 2 + \frac{\sum\limits_{uvE} S_0(H_{uv})}{\sum\limits_{uv\in E} \sigma_0(H_{uv})} = 2 + \sum_{uv\in E} \alpha_{uv} \frac{S_0(H_{uv})}{\sigma_0(H_{uv})}$$
$$= 2 + \sum_{uv\in E} \alpha_{uv} \cdot av_0(H_{uv})\,, \tag{2.3}$$

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 characterize the structure of graphs that minimize or maximize 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 [2] it was proved that the star maximizes av_0 among all n-vertex trees, while [3] shows that the star minimizes 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 [3], 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 characterization 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 1. *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 minimizes av_1 among n-vertex trees, and it is one of the n-vertex graphs that minimizes av_1.

Proof. By Lemma 6, $av_1(G) \geq 2$ 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 6 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 $2 \le \delta_1(G) \le \delta_2(G) \le |V|$ for all G that are not edgeless.

Proposition 2. *For every n-vertex graph $G \neq E_n$,*

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

Proof. Using u_1, v_1, u_2 and v_2 as in Lemma 6 and the result

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

that holds for all n-vertex graph H (see [2]), we have

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

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

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

□

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

Proposition 3 ([2]). *For every n-vertex graph G, we have*

$$av_0(G) \le av_0(E_n) = n/2 \,.$$

Equality holds if and only if G is isomorphic to E_n.

Let $av_1(G, e)$ denote the average of all 1-nearly independent vertex subsets of G that contain the two end vertices 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 $\sum_{uv \in E} \alpha_{uv} = 1$, provided G is not an edgeless graph; see the proof of Lemma 6.
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 $n \geq 2$. 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 2. *Let $n \geq 2$. For every n-vertex graph G,*

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

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

Proof. It is only left to prove that if $av_1(G) = \frac{n}{2} + 1$, with $n = |V(G)|$, then $G = G_n$. Clearly, G cannot be edgeless. From (2.3), we have

$$av_1(G) = 2 + \sum_{uv \in E(G)} \alpha_{uv} \cdot av_0(H_{uv}),$$

where $H_{uv} = G - (N(v) \cup N(u))$, $\alpha_{uv} > 0$ and $\sum_{uv \in E(G)} \alpha_{uv} = 1$. Moreover, by Proposition 3, we know that for any $uv \in E(G)$,

$$\mathrm{av}_0(H_{uv}) \leq \frac{|V(H_{uv})|}{2} \leq \frac{n-2}{2}$$

holds, with equality if and only if H_{uv} is edgeless and $|V(H_{uv})| = n-2$. In other words, $\mathrm{av}_0(H_{uv}) = \frac{n-2}{2}$ holds if and only if $H_{uv} = E_{n-2}$, and thus $N(v) \cup N(u) = \{u, v\}$. So, $av_1(G) = \frac{n}{2} + 1 = 2 + \frac{n-2}{2}$ holds if and only if both $G - (N(v) \cup N(u)) = E_{n-2}$ and $N(v) \cup N(u) = \{u, v\}$ hold for every edge uv in G.
Assume that $av_1(G) = 2 + \frac{n-2}{2}$ and let $u_0v_0 \in E(G)$. Since $N(v_0) \cup N(u_0) = \{u_0, v_0\}$, vertices u_0 and v_0 span a K_2 connected component in G. Since $H_{u_0v_0} = E_{n-2}$, the graph G contains no edge other than u_0v_0. Hence $G = K_2 \cup (n-2)K_1 = G_n$. $\square$

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

4. Extremal n-vertex trees

As a particular case of Theorem 1, the star S_n uniquely minimizes 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.
By an *internal vertex* in a tree T, we mean a vertex with degree greater than 1 in T. Any tree with order at least 3 contains an internal vertex.

Proposition 4. *Let $n \geq 3$ be an integer and T a tree of order n. Denote by δ' the minimum of the internal vertex degrees in T. Then we have*

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

Equality holds if and only if $n = 1 + \delta'$.

Proof. Let T be chosen as stated in the proposition. As already seen in the proof of Lemma 6,

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

for some positive numbers α_{uv} that sum to 1. Given that $n \geq 3$, if $uv \in E(T)$ then at least u or v is necessarily an internal vertex of T. Thus, we have

$$|N(u) \cup N(v)| \geq 1 + \delta' \tag{4.1}$$

with equality if and only if u is a leaf (a vertex of degree 1) in T and v is of degree δ' in T, or vice versa. Now we invoke Proposition 3 to obtain

$$\begin{aligned} av_0\big(T - (N(u) \cup N(v))\big) &\leq \frac{\big|V\big(T - (N(u) \cup N(v))\big)\big|}{2} = \frac{|V(T)| - |N(u) \cup N(v)|}{2} \\ &\leq \frac{n - (1 + \delta')}{2}\,. \end{aligned} \tag{4.2}$$

Equality holds in (4.2) if and only if it holds in (4.1) and $T - (N(u) \cup N(v))$ is edgeless. It follows that

$$av_1(T) = 2 + \sum_{uv \in E(T)} \alpha_{uv} \cdot av_0\big(T - (N(u) \cup N(v))\big) \leq 2 + \frac{n - (1 + \delta')}{2}$$

with equality if and only if both $T - (N(u) \cup N(v)) = E_{n-(1+\delta')}$ and

$$|N(u) \cup N(v)| = 1 + \delta' \quad \text{hold} \quad \text{for every} \quad uv \in E(T)\,. \tag{4.3}$$

By (4.3), the tree T cannot contain an induced path of order 4, since otherwise we could find on such a path an edge whose end vertices are both internal in T. Hence, $av_1(T) = 2 + \frac{n-(1+\delta')}{2}$ implies that T is of diameter at most 2, that is T is a star. This means $n = 1 + \delta'$. □

We can then deduce the following upper bound for n-vertex trees, obtained for $\delta' = 2$.

Corollary 1. *Let $n \geq 3$ be an integer. For any n-vertex tree T, we have*

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

with equality holding if and only if $T = P_3$.

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 1 to obtain

$$av_1(R_n) = \frac{S_1(S_{n-1}) + S_1(S_{n-2}) + \sigma_1(S_{n-2}) + S_0(E_{n-3}) + 2\sigma_0(E_{n-3})}{\sigma_1(S_{n-1}) + \sigma_1(S_{n-2}) + \sigma_0(E_{n-3})}$$

and

$$av_1(R_n) = \frac{2(n-2) + 2(n-3) + (n-3) + (n-3)2^{n-4} + 2 \times 2^{n-3}}{(n-2) + (n-3) + 2^{n-3}},$$

respectively, where in the latter 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

$$\begin{aligned} av_1(R_n) &= \frac{5n - 13 + (n+1)2^{n-4}}{2n - 5 + 2^{n-3}} \\ &= \frac{n+1}{2} - \frac{n^2 - 13n/2 + 21/2}{2n - 5 + 2^{n-3}} < \frac{n+1}{2}. \end{aligned} \tag{4.4}$$

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 3. *Let $n > 8$. Then*

$$\frac{n}{2} < \max_{\substack{T \text{ is a tree} \\ |V(T)|=n}} \mathrm{av}_1(T) < \frac{n+1}{2}\,.$$

Moreover, the upper bound is asymptotically sharp.

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

Conjecture 4. 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

This short section investigates further relations between the numbers σ_0 and σ_1 of 0- and 1-nearly independent vertex sets, respectively. The quantity σ_1 is directly involved in the definition of av_1. Thus, σ_1 is the natural counting parameter underlying this average, while σ_0 provides a comparison with the preceding level of near-independence. We establish further identities and bounds relating these two counting parameters. In particular, we show that the ratio σ_1/σ_0 has an additive-type behavior under graph unions. We also prove that $\sigma_0 \leq 3\sigma_1$ for graphs with maximum degree 2 and no isolated vertex. These results complement the study of av_1 by describing the relationship between σ_0 and σ_1; they are presented as further structural observations rather than as part of the principal extremal analysis.

We begin with a lemma.

Lemma 7. *For every graph $G = (V, E)$ that is not edgeless and every $uv \in E$, we have*

$$\frac{1}{3 \cdot 2^{l_{uv}-2}} \leq \frac{1}{2^{l_{uv}-2} + 2^{l_{uv}-m_u} + 2^{l_{uv}-m_v}} \leq \frac{\sigma_0\big(G - (N(v) \cup N(u))\big)}{\sigma_0(G)} \leq 1 - \frac{\sigma_0(G - v)}{\sigma_0(G)}\,,$$

where

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

In particular,

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

Proof. By partitioning the set of independent vertex subsets of G into those that contain neither u nor v, those that contain u but not v, and those that contain v and not u we obtain

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

This proves the lower bounds.
Now, we prove the upper bound. Note that $G-(N(v)\cup N(u))$ is an induced subgraph of $G - N[v]$, thus $\sigma_0(G - N[v]) \geq \sigma_0\big(G - (N(v) \cup N(u))\big)$ and

$$\sigma_0(G) = \sigma_0(G - v) + \sigma_0(G - N[v]) \geq \sigma_0(G - v) + \sigma_0\big(G - (N(v) \cup N(u))\big)\,.$$

This implies

$$1 \geq \frac{\sigma_0(G - v)}{\sigma_0(G)} + \frac{\sigma_0\big(G - (N(v) \cup N(u))\big)}{\sigma_0(G)}$$

as desired. □

In the next proposition, we show that σ_1/σ_0 is a somewhat additive function in the union of graphs.

Proposition 5. *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 Lemmas 1 and 5:

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

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 $G_2 := G - V(G_1)$ 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 7)

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

and (see Lemma 4)

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

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. □

Acknowledgment: The authors are grateful to one anonymous reviewer for his/her valuable comments and suggestions which have improved the quality of the paper.

Conflict of Interest: The authors declare that they have no conflict of interest.

Data Availability: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] E.O.D. Andriantiana, *Energy, Hosoya index and Merrifield–Simmons index of trees with prescribed degree sequence*, Discrete Appl. Math. **161** (2013), no. 6, 724–741.
https://doi.org/10.1016/j.dam.2012.10.010.

[2] E.O.D. Andriantiana, V. Razanajatovo Misanantenaina, and S. Wagner, *The average size of independent sets of graphs*, Eur. J. Math. **6** (2020), no. 2, 561–576.
https://doi.org/10.1007/s40879-019-00333-8.

[3] E.O.D. Andriantiana and Z.B. Shozi, *The number of 1-nearly independent vertex subsets*, Quaest. Math. **47** (2024), no. 12, 2353–2373.
https://doi.org/10.2989/16073606.2024.2367714.

[4] ———, *The number of 1-nearly independent edge subsets*, Iranian J. Math. Chem. **16** (2025), no. 1, 65–84.
https://doi.org/10.22052/ijmc.2024.254977.1871.

[5] E. Davies, M. Jenssen, W. Perkins, and B. Roberts, *On the average size of independent sets in triangle-free graphs*, Proc. Amer. Math. Soc. **146** (2018), no. 1, 111–124.
http://doi.org/10.1090/proc/13728.

[6] D. Galvin, *Two problems on independent sets in graphs*, Discrete Math. **311** (2011), no. 20, 2105–2112.
https://doi.org/10.1016/j.disc.2011.06.015.

[7] W. Gan, P.S. Loh, and B. Sudakov, *Maximizing the number of independent sets of a fixed size*, Combin. Probab. Comput. **24** (2015), no. 3, 521–527.
https://doi.org/10.1017/S0963548314000546.

[8] J. Kahn, *An entropy approach to the hard-core model on bipartite graphs*, Combin. Probab. Comput. **10** (2001), no. 3, 219–237.
https://doi.org/10.1017/S0963548301004631.

[9] A. Nilli, *The average size of an independent set in graphs with a given chromatic number*, J. Combin. Theory Ser. B **45** (1988), no. 1, 112–114.
https://doi.org/10.1016/0095-8956(88)90060-3.

[10] D. Weitz, *Counting independent sets up to the tree threshold*, Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '06, Association for Computing Machinery, 2006, p. 140–149.

[11] Y. Zhao, *The number of independent sets in a regular graph*, Combin. Probab. Comput. **19** (2010), no. 2, 315–320.
https://doi.org/10.1017/S0963548309990538.