CUT AND PENDANT VERTICES AND THE NUMBER OF CONNECTED INDUCED SUBGRAPHS OF A GRAPH

AUDACE A. V. DOSSOU-OLORY

ABSTRACT. A vertex whose removal in a graph G increases the number of components of G is called a cut vertex. For all n, c, we determine the maximum number of connected induced subgraphs in a connected graph with order n and c cut vertices, and also characterise those graphs attaining the bound. Moreover, we show that the cycle has the smallest number of connected induced subgraphs among all cut vertex-free connected graphs. The general case $c > 0$ remains an open task. We also characterise the extremal graph structures given both order and number of pendant vertices, and establish the corresponding formulas for the number of connected induced subgraphs. The 'minimal' graph in this case is a tree, thus coincides with the structure that was given by Li and Wang [Further analysis on the total number of subtrees of trees. *Electron. J. Comb.* 19(4), #P48, 2012].

1. Introduction and Preliminaries

Let G be a simple graph with vertex set $V(G)$ and edge set $E(G)$. The graph G is said to be connected if for all $u, v \in V(G)$, there is a $u - v$ path in G. An induced subgraph H of G is a graph such that $\emptyset \neq V(H) \subseteq V(G)$ and $E(H)$ consists of all those edges of G whose endvertices both belong to $V(H)$. The order of G is the cardinality $|V(G)|$, i.e. the number of vertices of G; the girth of G is the smallest order of a cycle (if any) in G; a pendant vertex (or leaf) of G is a vertex of degree 1 in G.

A general question in extremal/structural graph theory [2, 25, 29] is to find the minimum or maximum value of a prescribed graph parameter in a specified class of graphs. Turán's theorem [29] dating back to 1941, characterises the n-vertex graphs with greatest number of edges that contain no complete graph as a subgraph; this is probably the most classical result in extremal graph theory. This question has been studied quite thoroughly for several other parameters including the popular invariant *number of subtrees* of a tree (a connected graph with no cycle). Substantial work has been reported in the literature on the number of subtrees, see for example [1, 11, 15, 16, 17, 18, 26, 27, 32]. In recent works [5, 6], our main purpose was to extend some extremal results on the number of subtrees of a tree to more general classes of graphs such as connected graphs or unicylic graphs (connected graphs with only one cycle). In [5], order is prescribed for the class of all connected graphs and the class of all unicyclic graphs. Specifically, paper [5] characterises those graphs

2010 *Mathematics Subject Classification.* Primary 05C30; secondary 05C35, 05C05.

Key words and phrases. cut vertex, pendant vertex, induced subgraph, connected graph, extremal graph structure, tree.

The author was partially supported by the National Research Foundation of South Africa, grant 118521.

(or unicyclic graphs) with n vertices that minimise or maximise the number of connected induced subgraphs, thus extending some results of Székely and Wang [27]. In [6], further classes of graphs are considered, namely the class of all unicyclic graphs of order n and with girth g, and the class of all unicyclic graphs of girth g, with n vertices of which p are pendant. For each of the aforementioned classes of graphs, the extreme numbers of connected induced subgraphs were found in [6], and the extremal graph structures were also characterised. Extremal results on the total number of connected subgraphs (not necessary induced subgraphs) appeared recently in [20]. In general, there is no monotone relationship between the number of connected subgraphs and the number of connected induced subgraphs. In other words, if graph G has more connected subgraphs than graph H, it is not necessary true that G also contains more connected induced subgraphs than H.

In this note, we continue our systematic investigation on the number of connected induced subgraphs by considering two further classes of connected graphs. A component of G is a maximal (with respect to the number of vertices) connected induced subgraph of G. By $G - u$, we mean the graph that results from deleting vertex u and all edges incident with u in G. A *cut vertex* of G is a vertex $u \in \mathrm{V}(G)$ with the property that $G-u$ has more components than G. In the present paper, which complements [5, 6], we concentrate on two new classes of connected graphs for which we determine the extreme values and characterise the extremal graphs with respect to the number of connected induced subgraphs. Section 2 deals with the class of all connected graphs of order n with c cut vertices, while in Section 3 the focus is placed on the class of all connected graphs with n vertices of which p are pendant.

The n-vertex path and the n-vertex star are denoted by P_n and S_n, respectively. By $\mathbb{T}^1(n,p)$, we mean the tree obtained from the vertex disjoint graphs $S_{1+\lfloor p/2 \rfloor}$ and $S_{1+\lceil p/2 \rceil}$ by identifying their central vertices with the two leaves of P_{n-p}, respectively. Set $m := \lfloor (n-1)/p \rfloor$, $l := n - 1 - p \cdot m$ and denote by $\mathbb{T}^2(n,p)$ the rooted tree whose branches are l copies of P_{m+2} and $p - l$ copies of P_{m+1}. The extremal tree structures that minimise or maximise the number of subtrees of a tree with prescribed order and number of pendant vertices were characterised by Li and Wang [18], and Andriantiana et al. [1], respectively. Li and Wang's result [18, Theorem 1] states that precisely the tree $\mathbb{T}^1(n,p)$ has the smallest number of subtrees, while Andriantiana et al.'s result [1, Corollary 4] states that the maximum number of subtrees is achieved by the tree $\mathbb{T}^2(n,p)$. We shall prove (see Theorem 17 in Section 3) that $\mathbb{T}^1(n,p)$ is the unique graph of order n and with p pendant vertices that minimises the number of connected induced subgraphs.

The Wiener index of a connected graph G is defined as the sum of distances between all unordered pairs of vertices of G. The first results on this distance-based invariant date back to 1947 and are due to the chemist H. Wiener [31] who observed its strong correlation to the boiling point of certain chemical compounds. Subsequently, several authors have obtained sharp bounds on the Wiener index under various restrictions. A lower bound on the Wiener index, in terms of order and size, was given by Entringer et al. [10]. An upper bound, depending on order, also appeared in [10] by Entringer et al., and in [8] by Doyle and Graver. The maximum Wiener index among all cut vertex-free graphs was obtained by Plesník [22]. The Wiener index has been shown to correlate well with other chemical indices in applications [28, 30]. The tree $\mathbb{T}^1(n,p)$ was previously to Li and Wang's result [18, Theorem 1], shown by Shi [24] to have the maximum Wiener index among all n-vertex trees with p pendant vertices, while Entringer [9], and Entringer and Burns [3] proved that $\mathbb{T}^2(n,p)$ is the tree of order n with p pendant vertices having the smallest Wiener index. The same is observed in our current context: for each of the graph classes in consideration, the graphs that are found to maximise the number of connected induced subgraphs were also recently reported in [22, 21] to minimise the Wiener index, and vice versa.

For a connected graph G, we denote by $n(G)$, $c(G)$, $p(G)$ (or simply n, c, p if there is no danger of confusion) the order, number of cut vertices, and number of pendant vertices of G, respectively. It is well-known that if G is a non-trivial connected graph (i.e. a graph of order at least two), then $c(G) \leq n(G) - 2$ since a leaf of a spanning tree of G cannot be a cut vertex of G. This bound is achieved by paths only (the cut vertices of a path are its vertices of degree 2). From here onwards, we then assume that $n(G) > 2$ and $c(G) < n(G) - 2$. Clearly, if T is a tree, then every vertex of T is either a leaf or a cut vertex. Therefore, the identity $p(T) + c(T) = n(T)$ holds. Hence, the problem of finding the minimum (resp. maximum) number of connected induced subgraphs of an n-vertex tree having c cut vertices is equivalent to the problem of finding the minimum (resp. maximum) number of connected induced subgraphs of an n-vertex tree having $n-c$ pendant vertices. However, as mentioned earlier, the extremal trees for the latter problem were already characterised by Li and Wang [18], and Andriantiana et al. [1]. This is a motivation for us to consider more general classes of connected graphs.

The complete graph of order n and the cycle of order n are denoted by K_n and C_n, respectively. By $\deg_G(u)$, we mean the degree of vertex u in the graph G. We denote by $\mathrm{N}(G)$ the number of connected induced subgraphs of G. By $\mathrm{N}(G)_u$, we mean those connected induced subgraphs of G that contain vertex u, and $\mathrm{N}(G)_{u,v}$ stands for those connected induced subgraphs of G that contain vertices u and v. We simply write $G-u-v$ instead of $(G-u)-v$.

We shall frequently employ the following three lemmas without further reference.

Lemma 1 ([27]). *We have* $\mathrm{N}(P_n) = n(n+1)/2$ *for all* n. *Moreover, if* $u \in V(P_n)$, *then* $\mathrm{N}(P_n)_u \geq n$ *with equality holding if and only if* u *is a leaf.*

Lemma 2 ([5]). *We have* $\mathrm{N}(C_n) = n^2 - n + 1$ *for all* n. *Moreover, if* $u \in V(C_n)$, *then we have* $\mathrm{N}(C_n)_u = 1 + \binom{n}{2}$.

Lemma 3. *We have* $\mathrm{N}(K_n) = 2^n - 1$ *for all* n. *Moreover, if* $u \in V(K_n)$, *then* $\mathrm{N}(K_n)_u = 2^{n-1}$ *for all* n.

Proof. Every induced subgraph of K_n is a complete graph. Thus $\mathrm{N}(K_n) = 2^n - 1$. If $u \in V(K_n)$, then $\mathrm{N}(K_n)_u = \mathrm{N}(K_n) - \mathrm{N}(K_{n-1}) = 2^{n-1}$. □

Let G be a connected graph. A *block* of G is a maximal (with respect to the number of vertices) cut vertex-free connected induced subgraph of G [12]. In particular, if G is a non-trivial connected graph, then so are all blocks of G. Moreover, every block of G is either P_2 or a cyclic graph since every tree of order three or more contains at least one cut vertex. As a first consequence of this definition, one deduces that the intersection of the vertex sets of any two distinct blocks of G consists of at most one vertex [13].

The proof techniques in this work build on several graph transformations, some of which are known to have a counterpart for the Wiener index. The rest of the paper is organised as follows: Section 2 contains extremal results on the number of connected induced subgraphs with c cut vertices. Define $G(n_1; \dots; n_q)$ to be the graph constructed as follows: we consider $q+1 > 3$ pairwise vertex disjoint graphs $K_q, P_{n_1}, \dots, P_{n_q}$ such that $V(K_q) = \{v_1, \dots, v_q\}$; for every $j \in \{1, \dots, q\}$, we let u_j be a leaf of P_{n_j} and identify u_j with v_j. We prove (see Theorem 11) that $G(s; \dots; s; s+1; \dots; s+1)$ ($n-c-t$ copies of s followed by t copies of $s+1$) is the unique connected graph of order n and with c cut vertices that has the greatest number of connected induced subgraphs. A formula in terms of n and c is also provided for $\mathrm{N}(G(s; \dots; s; s+1; \dots; s+1))$. We demonstrate in Theorem 13 that the cycle C_n has the smallest number of connected induced subgraphs among all cut vertex-free connected graphs of order n. The general case $c > 0$ seems to be hard and we leave this as an open problem. Section 3 considers the class of all connected graphs with n vertices of which p are pendant. The 'maximal' graph in this case is already known; see [7]. We summarise this result in Theorem 14 and then prove its minimisation counterpart in Theorems 17 and 24. Specifically, we show that for $p \neq 1$, the tree $\mathbb{T}^1(n,p)$ remains the unique graph of order n and with p pendant vertices that has the smallest number of connected induced subgraphs. For $p = 0$ and $n > 5$, we prove that the minimum number of connected induced subgraphs is realised by the so-called double tadpole graph, and that it is unique with this property. By the n-vertex double tadpole graph, we mean the graph constructed from the path of order $n-4$ and two vertex disjoint triangles by identifying bijectively the two leaves of the path with two other vertices, one from each triangle.

Our approach sometimes follows [21], adapted to our current setting. Throughout this note, all graphs are simply connected. We assume $n \geq 3$ and $p \leq n-2$ since the case $p = n-1 \geq 2$ corresponds to the n-vertex star, i.e. a vertex and $n-1$ leaves attached to it.

2. Connected graphs with c cut vertices

We define $\mathcal{H}(n,c)$ to be the set of all connected graphs with order n and c cut vertices.

2.1. **The maximisation problem.** In order to state the main result of this subsection, we need to go through some preparation. It is obvious that the complete graph K_n uniquely realises the maximum number of connected induced subgraphs among all graphs in $\mathcal{H}(n,0)$.

Let G be a non-trivial connected graph. The following properties about G are elementary; see for instance [13, 14].

(i) Every cut vertex of G belongs to at least two distinct blocks of G;

(ii) Every two distinct blocks of G have at most one vertex in common. Whenever they have a vertex in common, it must be a cut vertex of G.
(iii) If G has at least one cut vertex, then G also has at least one block that contains exactly one cut vertex of G.

We shall make frequent use of these properties without further reference. We begin with a series of important lemmas. The next two lemmas are straightforward.

Lemma 4. *If G' is obtained from a non-trivial connected graph G by adding an edge between two nonadjacent vertices of the same block of G, then*

$$c(G') = c(G).$$

Note that the above graph transformation (Lemma 4) increases the number of edges in a block of G while preserving the number of cut vertices of G. Our next transformation reduces the number of blocks of G by one while preserving its number of cut vertices.

Lemma 5. *Let B_1, B_2, B_3 be three distinct blocks of a non-trivial connected graph G such that $V(B_1) \cap V(B_2) \cap V(B_3) = \{w\}$. Assume that G' is constructed from G by adding an edge between a neighbour v_1 of w in B_1 and a neighbour v_2 of w in B_2. Then we have*

$$c(G) = c(G').$$

Proof. Clearly, every cut vertex of G' is a cut vertex of G by construction. Let z be a cut vertex of G. If $z \notin V(B_1) \cup V(B_2) \cup V(B_3)$, then all vertices in $V(B_1) \cup V(B_2) \cup V(B_3)$ are entirely contained in only one component of $G - z$. Thus z is a cut vertex of G'. Otherwise, let $j \in \{1, 2, 3\}$ such that $z \in V(B_j)$. If $z \in \{v_1, v_2\}$, then $G - z$ and $G' - z$ are isomorphic graphs by definition of G'; otherwise $z \notin \{v_1, v_2\}$. If $z \neq w$, then v_1 and v_2 belong to the same component of $G - z$. Thus z is a cut vertex of G'. Otherwise $z = w$ and so $V(B_1 - z), V(B_2 - z), V(B_3 - z)$ are all contained entirely in distinct components of $G - z$. Since an edge is only added between v_1 and v_2 in G to obtain G', we deduce that the component of $G' - z$ that contains $B_3 - z$ as a subgraph remains isolated in $G' - z$. Hence, z is a cut vertex of G'. □

Consider $q + 1 > 3$ pairwise vertex disjoint graphs $K_q, P_{n_1}, \ldots, P_{n_q}$ such that $1 \leq n_1 \leq n_2 - 1$ and $V(K_q) = \{v_1, \ldots, v_q\}$. For every $j \in \{1, \ldots, q\}$, let u_j be a leaf of P_{n_j} and identify u_j with v_j. We denote by $G(n_1; \ldots; n_q)$ the resulting graph.

Lemma 6. *Let H be a connected graph of order greater than two, and u, v two distinct vertices of H such that $\mathrm{N}(H)_{u,v} > 1$ and $\mathrm{N}(H - u)_v \leq \mathrm{N}(H - v)_u$. Let $H(n_1; n_2)$ be the graph obtained from H by identifying u with a leaf of P_{n_1}, and v with a leaf of P_{n_2} for some $1 \leq n_1 \leq n_2 - 1$. We have*

$$\mathrm{N}(H(n_1; n_2)) \leq \mathrm{N}(H(n_1 + 1; n_2 - 1)).$$

The inequality is strict if and only if $\mathrm{N}(H - u)_v < \mathrm{N}(H - v)_u$ or $n_1 < n_2 - 1$. In particular, we get

$$\mathrm{N}(G(n_1; n_2; \cdots; n_q)) \leq \mathrm{N}(G(n_1 + 1; n_2 - 1; n_3; \cdots; n_q))$$

if and only if $|n_1 - n_2| \geq 1$. Equality holds if and only if $|n_1 - n_2| = 1$.

Proof. We categorise subgraphs of $H(n_1;n_2)$ according to whether they contain an element of $\{u,v\}$ or not. Removing vertices u and v from $H(n_1;n_2)$ yields the graphs (possibly empty) P_{n_1-1}, P_{n_2-1} and $H-u-v$. Thus $\mathrm{N}(P_{n_1-1})+\mathrm{N}(P_{n_2-1})+\mathrm{N}(H-u-v)$ counts the number of connected induced subgraphs of $H(n_1;n_2)$ that contain none of the vertices u,v. On the other hand, $n_1\cdot \mathrm{N}(H-v)_u+n_2\cdot\mathrm{N}(H-u)_v$ counts the number of connected induced subgraphs of $H(n_1;n_2)$ that contain u or v but not both. The number of connected induced subgraphs of $H(n_1;n_2)$ that contain both u and v is given by $n_1\cdot n_2\cdot \mathrm{N}(H)_{u,v}$. Hence, we get

$$\begin{aligned}\mathrm{N}(H(n_1;n_2)) &= n_1\cdot n_2\cdot\mathrm{N}(H)_{u,v}+n_1\cdot\mathrm{N}(H-v)_u+n_2\cdot\mathrm{N}(H-u)_v\\ &\quad+\mathrm{N}(P_{n_1-1})+\mathrm{N}(P_{n_2-1})+\mathrm{N}(H-u-v)\,.\end{aligned}$$

This implies that

$$\begin{aligned}\mathrm{N}(H(n_1;n_2))-\mathrm{N}(H(n_1+1;n_2-1)) &= (n_1\cdot n_2-(n_1+1)(n_2-1))\,\mathrm{N}(H)_{u,v}\\ &\quad+(n_1-(n_1+1))\,\mathrm{N}(H-v)_u+(n_2-(n_2-1))\,\mathrm{N}(H-u)_v\\ &\quad+\binom{n_1}{2}-\binom{n_1+1}{2}+\binom{n_2}{2}-\binom{n_2-1}{2}\\ &=(n_1-n_2+1)(\mathrm{N}(H)_{u,v}-1)+\mathrm{N}(H-u)_v-\mathrm{N}(H-v)_u\le 0\,.\end{aligned}$$

Moreover, this inequality becomes an equality if and only if $\mathrm{N}(H-u)_v=\mathrm{N}(H-v)_u$ and $n_1=n_2-1$. This proves the lemma. □

Lemma 7. *Let $H(n;l)$ be the graph constructed from the two vertex disjoint complete graphs K_l and K_{n+1-l} by identifying $u\in V(K_l)$ with $v\in V(K_{n+1-l})$ for some $n\ge 3$ and $2\le l\le (n+1)/2$. Then we have*

$$\mathrm{N}(H(n;2))>\mathrm{N}(H(n;3))>\cdots>\mathrm{N}(H(n;\lfloor (n+1)/2\rfloor))\,.$$

Proof. We have

$$\begin{aligned}\mathrm{N}(H(n;l)) &= \mathrm{N}(H(n;l))_u+\mathrm{N}(H(n;l)-u)\\ &=\mathrm{N}(K_l)_u\cdot\mathrm{N}(K_{n+1-l})_v+\mathrm{N}(K_{l-1})+\mathrm{N}(K_{n-l})=2^{n-1}+2^{l-1}+2^{n-l}-2\end{aligned}$$

which implies that

$$\mathrm{N}(H(n;l))-\mathrm{N}(H(n;l+1))=2^{n-l-1}-2^{l-1}>0$$

for all $2\le l\le (n-1)/2$. The statement of the lemma follows. □

Lemma 8. *Let $l\ge 3,\ r\ge 2$ be two positive integers and K_l, K_r two vertex disjoint complete graphs such that $V(K_r)=\{w_1,\dots,w_r\}$. Consider $r-1$ vertex disjoint connected graphs $R_2,\dots,R_r$ such that $v_j\in V(R_j)$ for every $j\in\{2,\dots,r\}$ and $|V(R_2)|>1$. Identify w_1 with a fixed vertex $u\in V(K_l)$. Further, identify w_j with v_j for all $j>1$. Call the resulting graph G_1. Fix $v\ne u\in V(K_l)$ and let G_2 be constructed from G_1 by deleting the edges joining v to a neighbour, except u, of v in K_l. Finally, let G_3 be constructed from G_2 by making the graph induced by $V(K_l-u-v)\cup V(K_r-w_1)$ in G_2 a complete graph. We have*

$$\mathrm{N}(G_3)>\mathrm{N}(G_1)\quad\textit{and}\quad c(G_1)=c(G_3)\,.$$

Proof. Denote by H the graph induced by $V(G_1) - V(K_l - u)$ in G_1. We have

$$\mathrm{N}(G_1) = \mathrm{N}(G_1)_u + \mathrm{N}(G_1 - u) = \mathrm{N}(K_l)_u \cdot \mathrm{N}(H)_{w_1} + \mathrm{N}(K_{l-1}) + \mathrm{N}(H - w_1)$$
$$\mathrm{N}(G_2) = \mathrm{N}(G_2)_u + \mathrm{N}(G_2 - u) = 2\,\mathrm{N}(H)_{w_1} \cdot \mathrm{N}(K_l - v)_u + 1 + \mathrm{N}(K_{l-2}) + \mathrm{N}(H - w_1)\,.$$

In particular, we get $\mathrm{N}(G_1) - \mathrm{N}(G_2) = 2^{l-2} - 1$. Clearly, G_2 is a subgraph of G_3 by construction. Let $S_1 \subseteq V(K_l - u - v)$ and $S_2 \subseteq V(K_r - w_1)$ be two nonempty subsets of vertices of $V(G_2) = V(G_3)$. These choices of S_1 and S_2 are possible since $l \geq 3$ and $r \geq 2$. The graph induced by $S_1 \cup S_2$ in G_2 is disconnected while the graph induced by $S_1 \cup S_2$ in G_3 is connected. The total number of these connected induced subgraphs in G_3 is given by $(2^{l-2} - 1)(2^{r-1} - 1)$. Let z be a vertex adjacent to v_2 in R_2. Vertex z exists since $|V(R_2)| > 1$. The graph induced by $S_1 \cup \{v_2, z\}$ in G_3 is connected and different from all the subgraphs induced by $S_1 \cup S_2$ as $z \notin S_1 \cup S_2$. Moreover, $S_1 \cup \{v_2, z\}$ induces a disconnected graph in G_2. The total number of such connected induced subgraphs in G_3 is $2^{l-2} - 1$. Therefore, we deduce that

$$\mathrm{N}(G_3) - \mathrm{N}(G_2) \geq (2^{l-2} - 1)2^{r-1} \geq 2(2^{l-2} - 1)\,.$$

It follows that $\mathrm{N}(G_3) - \mathrm{N}(G_1) \geq 2^{l-2} - 1 > 0$.

Now since v is a leaf of G_3 and u is adjacent to v in G_3, we conclude that u remains a cut-vertex of G_3 while v remains a non cut-vertex of G_3. Moreover, all other vertices of G_1 preserve their status (cut vertex or not) in G_3. This proves that $c(G_1) = c(G_3)$, completing the proof. □

Next, we describe another graph transformation that will also be useful for our analysis. It is a result that is similar in nature to but different from Lemma 8. It does, however, complement Lemma 8.

Lemma 9. *Let K_l, K_r be two complete graphs with (disjoint) vertex sets*

$$V(K_l) = \{u_1, \ldots, u_l\},\ V(K_r) = \{w_1, \ldots, w_r\}$$

for some $l, r \geq 3$. Consider $l+r-1$ vertex disjoint connected graphs $M, L_2, \ldots, L_l, R_2, \ldots, R_r$ such that $x_j \in V(L_j)$ and $z_j \in V(R_j)$ for all $j \neq 1$. Let v_1, v_2 be two distinct vertices of M. Identify u_1 with v_1, and w_1 with v_2. Further, identify u_j with x_j, and w_j with z_j for all $j \neq 1$. Denote by G_1 the resulting graph. Let G_2 be obtained from G_1 by removing the edges joining u_1 to a neighbour, except u_2, of u_1 in K_l; see Figure 1. Let w' be a fixed neighbour of v_1 in M such that w' lies on a shortest $v_1 - v_2$ path P in G_2. A new graph G_3 is constructed from G_2 by adding an edge between w' and all vertices $u_3, \ldots, u_l$. We have

$$c(G_1) = c(G_3)\,.$$

Furthermore, let L be the graph induced by $\{u_1\} \cup V(L_2) \cup \cdots \cup V(L_l)$ in G_1, and R the graph induced by $\{w_1\} \cup V(R_2) \cup \cdots \cup V(R_r)$ in G_1. Assume that $\mathrm{N}(R)_{w_1} \geq \mathrm{N}(L)_{u_1}$. Then we have

$$N(G_3) > N(G_1)\,.$$

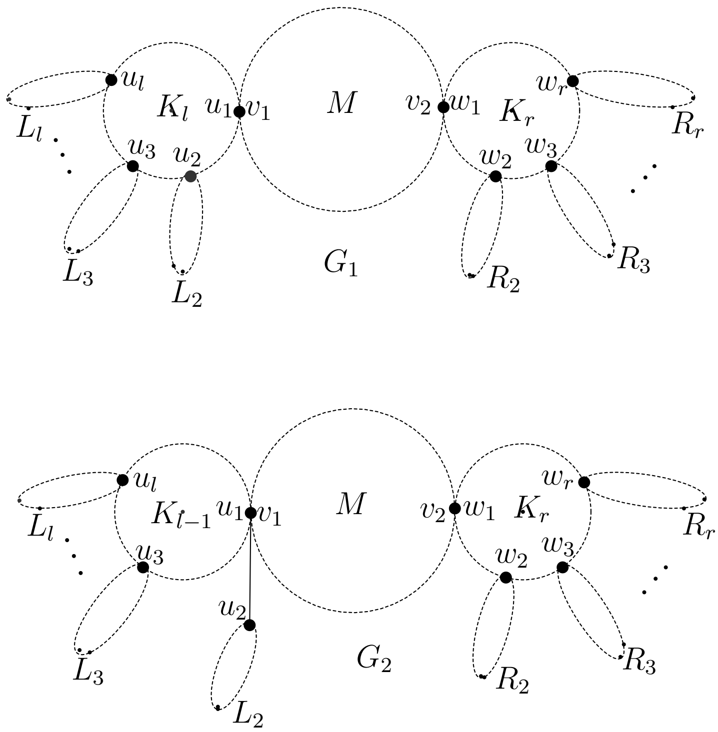

FIGURE 1. The graphs G_1 and G_2 constructed in Lemma 9.

Proof. It is clear by construction that $u_1 \in V(G_1)$ remains a cut vertex of G_2. This is because u_1 is adjacent to u_2, and u_2 is adjacent to no vertex of G_2 outside $V(L_2) \cup \{u_1\}$ in G_2. Thus, all cut (resp. non cut) vertices of G_1 remain cut (resp. non cut) vertices of G_2. Therefore, we have $c(G_1) = c(G_2)$. On the other hand, since edges are only added between w' and the vertices $u_3, \ldots, u_l$ in G_2 to obtain G_3, it is clear that the following hold:

- All non cut vertices of G_2 remain non cut vertices of G_3;
- All cut vertices of G_2 that do not belong to $V(M) - \{v_1, v_2\}$ remain cut vertices of G_3.

Let θ be a cut vertex of G_2 such that $\theta \in V(M) - \{v_1, v_2\}$. We show that θ is also a cut vertex of G_3. If $\theta = w'$, then $G_3 - \theta$ and $G_2 - \theta$ are isomorphic graphs. So assume that $\theta \neq w'$. Then w' must belong to the component, say C of $G_2 - \theta$ that contains v_1, since otherwise, every $v_1 - w'$ path must pass through θ. In particular, we get $\theta \in \{v_1, w'\}$ as v_1w' is an edge of G_2: this is a contradiction to the choice of θ. Hence, $w' \in V(C)$.

Note that C also contains all of $u_3, \ldots, u_l$ since $\theta \in V(M) - \{v_1, v_2\}$ and C is a component of $G_2 - \theta$ that contains $u_1 (= v_1)$. Since $w' \in V(C)$, we then deduce that all other (different from C) components of $G_2 - \theta$ remain components of $G_3 - \theta$. This proves that θ is indeed a cut vertex of G_3. In particular, we get $c(G_3) = c(G_2) = c(G_1)$.

Let $x_1 \in \{u_1, v_1\}$ and denote by L_1 the graph induced by $V(M) \cup V(R)$ in G_1. Then the vertex set of G_1 can be partitioned into $V(L_1), \ldots, V(L_l)$. Thus, for a subset $S \subseteq V(G_1)$

containing a vertex of $V(L_i)$ and a vertex of $V(L_j)$, where $i \neq j$ to induce a connected graph in G_1, it is necessary to have $x_i, x_j \in S$. Therefore, we get

$$\mathrm{N}(G_1) = \sum_{j=1}^{l} \mathrm{N}(L_j - x_j) + \prod_{j=1}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1$$

as a formula for the number of connected induced subgraphs of G_1. Likewise, denote by L' the graph induced by $V(L_1) \cup V(L_2)$ in G_2. The set $V(G_2)$ can also be partitioned into $V(L'), V(L_3), \ldots, V(L_l)$. Thus, we get

$$\mathrm{N}(G_2) = \mathrm{N}(L' - x_1) + \sum_{j=3}^{l} \mathrm{N}(L_j - x_j) + (1 + \mathrm{N}(L')_{x_1}) \prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1$$

in the same way as for G_1. On the other hand, we have

$$\mathrm{N}(L' - x_1) = \mathrm{N}(L_2) + \mathrm{N}(L_1 - x_1) \text{ and } \mathrm{N}(L')_{x_1} = \mathrm{N}(L_1)_{x_1}(1 + \mathrm{N}(L_2)_{x_2})\,.$$

Therefore, we obtain

$$\mathrm{N}(G_1) - \mathrm{N}(G_2) = \mathrm{N}(L_2)_{x_2} \Big(\prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1 \Big)$$

after simplification. By construction, G_3 contains G_2 as a subgraph. We now find a lower bound on $\mathrm{N}(G_3) - \mathrm{N}(G_2)$ by solely counting certain subsets of $V(G_3) = V(G_2)$ that induce a connected graph in G_3 and a disconnected graph in G_2. Let $S_1 \neq \emptyset$ be a subset of $V(L_3) \cup \cdots \cup V(L_l)$ such that S_1 contains x_j whenever S_1 contains an element of $V(L_j)$. Recall that P is a fixed shortest $v_1 - v_2$ path in G_2 that contains w'. Denote by R' the graph induced by $V(R) \cup V(P - v_1)$ in G_2. Let $S_2 \neq \emptyset$ a subset of $V(R')$ that contains w'. Since $w' \neq v_1$ is adjacent to all of $x_3, \ldots, x_l$ in G_3, we deduce that $S_1 \cup S_2$ always induces a connected graph in G_3. However, the graph induced by $S_1 \cup S_2$ in G_2 is always disconnected as there is no edge from an element of S_1 to an element of S_2 in G_2. Therefore, we obtain a total of

$$\mathrm{N}(R')_{w'} \Big(\prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1 \Big)$$

such sets $S_1 \cup S_2$ inducing a connected graph in G_3 and a disconnected graph in G_2. Since $\mathrm{N}(L)_{u_1} > 1 + \mathrm{N}(L_2)_{u_2}$, we use the trivial inequality $\mathrm{N}(R')_{w'} \geq \mathrm{N}(R)_{w_1}$ alongside the assumption $\mathrm{N}(R)_{w_1} \geq \mathrm{N}(L)_{u_1}$ to derive that

$$\mathrm{N}(G_3) - \mathrm{N}(G_2) \geq \mathrm{N}(R)_{w_1} \Big(\prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1 \Big) > (1 + \mathrm{N}(L_2)_{u_2}) \Big(\prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1 \Big)\,.$$

This implies that

$$\mathrm{N}(G_3) - \mathrm{N}(G_1) > \prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1 \geq \mathrm{N}(L_3)_{x_3} > 0\,,$$

completing the proof. □

It is required in Lemma 9 that $|V(M)| > 1$. Lemma 10 below covers the special case where $|V(M)| = 1$.

Lemma 10. *Let K_l, K_r be two complete graphs with (disjoint) vertex sets*

$$V(K_l) = \{u_1, \dots, u_l\},\ V(K_r) = \{w_1, \dots, w_r\}$$

for some $l, r \geq 3$. Consider $l+r-2$ vertex disjoint connected graphs $L_2, \dots, L_l, R_2, \dots, R_r$ such that $x_j \in V(L_j)$ and $z_j \in V(R_j)$ for all $j \in \{2, \dots, r\}$. Identify u_1 with w_1, u_j with x_j, and w_j with z_j for all $j \neq 1$. Denote by G_1 the resulting graph. Let G_2 be obtained from G_1 by removing the edges joining u_1 to a neighbour, except u_2, of u_1 in K_l. Let G_3 be constructed from G_2 by making the graph induced by the set $V(K_l - u_1 - u_2) \cup V(K_r - w_1)$ a complete graph. We have $c(G_1) = c(G_3)$. Furthermore, assume that $\mathrm{N}(R_2)_{z_2} \geq \mathrm{N}(L_2)_{x_2}$. Then we have

$$N(G_3) > N(G_1)\,.$$

Proof. The proof is done in analogy to Lemma 9 with the following simple modification. Let R be the graph induced by $\{w_1\} \cup V(R_2) \cup \cdots \cup V(R_r)$ in G_1. Denote by R' the graph induced by $V(R) \cup V(L_2)$ in G_2. We have

$$\mathrm{N}(G_1) = \mathrm{N}(R - w_1) + \sum_{j=2}^{l} \mathrm{N}(L_j - x_j) + (1 + \mathrm{N}(R)_{w_1}) \prod_{j=2}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1\,,$$

$$\mathrm{N}(G_2) = \mathrm{N}(R' - w_1) + \sum_{j=3}^{l} \mathrm{N}(L_j - x_j) + (1 + \mathrm{N}(R')_{w_1}) \prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1\,,$$

and

$$\mathrm{N}(R' - w_1) = \mathrm{N}(L_2) + \mathrm{N}(R - w_1),\ \mathrm{N}(R')_{w_1} = \mathrm{N}(R)_{w_1}(1 + \mathrm{N}(L_2)_{x_2})\,.$$

It follows that

$$\mathrm{N}(G_1) - \mathrm{N}(G_2) = \mathrm{N}(L_2)_{x_2} \Big(\prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1 \Big)\,.$$

Clearly, every subgraph of G_2 is also a subgraph of G_3. Let $S_1 \neq \emptyset$ be a subset of $V(L_3) \cup \cdots \cup V(L_l)$ such that S_1 contains x_j whenever S_1 contains an element of $V(L_j)$. Likewise, let $S_2 \neq \emptyset$ be a subset of $V(R - w_1)$ such that S_2 contains z_j whenever S_2 contains an element of $V(R_j)$. The set $S_1 \cup S_2$ always induces a disconnected graph in G_2, and a connected graph in G_3. Therefore, we get

$$\mathrm{N}(G_3) - \mathrm{N}(G_2) \geq \Big(\prod_{j=2}^{r} (1 + \mathrm{N}(R_j)_{z_j}) - 1 \Big) \Big(\prod_{j=3}^{l} (1 + \mathrm{N}(L_j)_{x_j}) - 1 \Big)\,.$$

On the other hand, we have

$$\prod_{j=2}^{r}(1+\mathrm{N}(R_j)_{z_j})-1 \geq (1+\mathrm{N}(R_2)_{z_2})(1+\mathrm{N}(R_3)_{z_3})-1 \geq 1+2\,\mathrm{N}(R_2)_{z_2}\,.$$

Hence, using the assumption $\mathrm{N}(R_2)_{z_2} \geq \mathrm{N}(L_2)_{x_2}$, we deduce that

$$\mathrm{N}(G_3)-\mathrm{N}(G_2) \geq (1+2\,\mathrm{N}(L_2)_{x_2})\Big(\prod_{j=3}^{l}(1+\mathrm{N}(L_j)_{x_j})-1\Big)\,,$$

which implies that $\mathrm{N}(G_3)-\mathrm{N}(G_1)>0$. This completes the proof of the lemma. □

We are now ready to formulate a characterisation of all graphs maximising the number of connected induced subgraphs in the set $\mathcal{H}(n,c)$. At this point, it can be recalled that $G(n_1;\ldots;n_q)$ is the graph constructed as follows: we consider $q+1>3$ pairwise vertex disjoint graphs $K_q, P_{n_1},\ldots,P_{n_q}$ such that $n_1 \leq n_2$, $n_2>1$ and $V(K_q)=\{v_1,\ldots,v_q\}$. For every $j\in\{1,\ldots,q\}$, we let u_j be a leaf of P_{n_j} and identify u_j with v_j.

Theorem 11. *Let $n>1$ and $0\leq c\leq n-2$. Denote by t the residue of n modulo $n-c$, and set $s=\lfloor n/(n-c)\rfloor$. We have*

$$\mathrm{N}(H) \leq (n-c-t)\binom{s}{2}+t\binom{s+1}{2}+(s+1)^{n-c-t}(s+2)^t-1$$

for all $H\in\mathcal{H}(n,c)$. Equality holds if and only if H is isomorphic to the graph $G(s;\ldots;s;s+1;\ldots;s+1)$ ($n-c-t$ copies of s followed by t copies of $s+1$).

Proof. First off, note that if B is a block of a non-trivial connected graph G, then B is necessarily 'surrounded' by $|V(B)|$ (possibly trivial) connected induced subgraphs of G whose vertex sets are pairwise disjoint. In other words, the removal of all edges of B in G must leave $|V(B)|$ connected graphs.

Let $\mathbb{H}\in\mathcal{H}(n,c)$ be a graph with order n and c cut vertices that maximises the number of connected induced subgraphs. We know, by repeatedly applying Lemma 4, that all blocks of $\mathbb{H}$ are non-trivial complete graphs. We are going to prove that all blocks of $\mathbb{H}$, except possibly only one, are in fact of order 2. The statement is obvious for $c=0$ since $\mathbb{H}=K_n$ in this case. So we assume that $c\geq 1$. By repeatedly invoking Lemma 5, we can further assume that every cut vertex of G belongs to precisely two distinct blocks of $\mathbb{H}$. If $c=1$, then $\mathbb{H}$ has precisely two blocks, say K_l and K_{n+1-l} for some $2\leq l\leq (n+1)/2$. Thus, in this case, the statement holds true by Lemma 7. So we assume that $c\geq 2$. Consider a block K_l of $\mathbb{H}$ such that $l\geq 3$. We consider two separate cases depending on whether K_l contains one or more cut vertices of $\mathbb{H}$.

Assume that K_l contains precisely one cut vertex, say w_1 of $\mathbb{H}$. Let $w_2\neq w_1$ be another cut vertex of $\mathbb{H}$ such that both w_1 and w_2 belong the the same block K_r of $\mathbb{H}$. Thus w_2 also belongs to a further block B of $\mathbb{H}$ different from K_r. This kind of description for $\mathbb{H}$ yields exactly the graph G_1 constructed in Lemma 8, where the graph R_2 in Lemma 8 contains B as a subgraph and $w_2\in V(R_2)$. Note that the graph transformation described in Lemma 8 preserves the number of cut vertices when passing from G_1 to G_3 but creates a new block of order 2 in G_3. It is shown in Lemma 8 that $\mathrm{N}(G_3) > \mathrm{N}(G_1) = \mathrm{N}(\mathbb{H})$. However, this is impossible from the choice of $\mathbb{H}$. Hence, we must have $l = 2$.

Assume that K_l contains two or more cut vertices, say u_1, u_2 of $\mathbb{H}$. If there is no other block that contains two or more cut vertices of $\mathbb{H}$, then we are done immediately by Lemma 8. This is because Lemma 8 states that in a 'maximal' graph, all blocks that contain only one cut vertex must be of order 2.

Otherwise, let $K_r \neq K_l$ be another block containing two or more cut vertices, say w_1, w_2 of $\mathbb{H}$. We can assume that $r \geq 3$ since otherwise, there is nothing more to prove. We observe two possible situations:

Case 1: $V(K_l) \cap V(K_r) = \emptyset$. In this case, there exists a non-trivial connected graph M that contains both u_1, w_1 and no other vertex of $V(K_l) \cup V(K_r)$. In particular, $\mathbb{H}$ can be described in the same way as the graph G_1 defined in Lemma 9 (see Figure 1), where L and R are the two components of $\mathbb{H} - V(M - u_1 - w_1)$ that contain u_1 and w_1, respectively. Without loss of generality, say $\mathrm{N}(R)_{w_1} \geq \mathrm{N}(L)_{u_1}$. Then Lemma 9 shows the existence of another graph G_3 with order n and c cut vertices satisfying $\mathrm{N}(G_3) > \mathrm{N}(G_1) = \mathrm{N}(\mathbb{H})$, which is indeed a contradiction.

Case 2: $V(K_l) \cap V(K_r) \neq \emptyset$. Since K_l and K_r are both blocks of $\mathbb{H}$, they can only have one common vertex, which is therefore a cut vertex of $\mathbb{H}$. Thus, without loss of generality, say $V(K_l) \cap V(K_r) = \{u_1 = w_1\}$. The graph $\mathbb{H}$ can then be given the same description as the graph G_1 defined in Lemma 10, where $x_2 = u_2 \in V(L_2)$ and $z_2 = w_2 \in V(R_2)$. Without loss of generality, say $\mathrm{N}(R_2)_{z_2} \geq \mathrm{N}(L_2)_{x_2}$. Then Lemma 10 applied to $\mathbb{H} = G_1$, which contradicts the choice of $\mathbb{H}$.

Summing up, we have proved that all blocks of $\mathbb{H}$, except possibly only one, are of order 2. Moreover, every cut vertex of $\mathbb{H}$ belongs to precisely two distinct blocks of $\mathbb{H}$. This then makes it simple to derive the full structure of $\mathbb{H}$. It is easy to see that all blocks of $\mathbb{H}$ are of order 2 if and only if $c = n - 2$ ($\mathbb{H}$ is a path in this case). Assume that $c \leq n - 3$ and let K_q be the unique block of $\mathbb{H}$ such that $q > 2$. One immediately deduces that $\mathbb{H}$ consists of K_q to which q paths (possibly trivial) $P_{n_1}, \ldots, P_{n_q}$ are attached to the vertices $u_1, \ldots, u_q$ of K_q, respectively, by identifying u_j with a leaf of P_{n_j} for all j. Therefore, we have $(n_1 - 1) + \cdots + (n_q - 1) = c$ and $n_1 + \cdots + n_q = n$, i.e. $q = n - c$. To complete the proof of the theorem, we need to find the values of all n_j. Lemma 6 yields that $n_1, \ldots, n_q$ must all be as equal as possible, i.e.

$$n_1 = \cdots = n_{n-c-t} = \lfloor n/(n-c) \rfloor = s \text{ and } n_{n-c-t+1} = \cdots = n_{n-c} = s + 1\,,$$

where t is the residue of n modulo $n - c$. Hence, we get

$$\mathrm{N}(\mathbb{H}) = (n-c-t)\binom{s}{2} + t\binom{s+1}{2} + (s+1)^{n-c-t}(s+2)^t - 1$$

as a special case in the proof of Lemma 6. This completes the proof of the theorem. □

2.2. **The minimisation problem.** In this subsection, we consider the special case $c = 0$ of the problem of finding those graphs that minimise the number of connected induced subgraphs among all graphs in the set $\mathcal{H}(n, c)$.

Let $n \geq 4$ and G be a graph consisting of the cycle C_{n-1} together with a vertex $z \notin V(C_{n-1})$ which is adjacent to precisely two vertices $x, v \in V(C_{n-1})$. In the sequel, we shall refer to every such graph as *special*.

Lemma 12. *If G is a special graph of order n, then we have* $\mathrm{N}(G) > n^2 - n + 1 = \mathrm{N}(C_n)$.

Proof. Let G be a special graph of order n. A simple lower bound on $\mathrm{N}(G)$ can be obtained as follows: a z-containing connected induced subgraph of G is either the single vertex z, or consists of z and at least a neighbour of z in G. Thus, we get

$$\mathrm{N}(G)_z = 1 + \mathrm{N}(G - v)_{x,z} + \mathrm{N}(G - x)_{v,z} + \mathrm{N}(G)_{x,v,z} .$$

Since $G - z$ is a cycle and $G - v - z$ as well as $G - x - z$ are paths, we deduce that

$$\mathrm{N}(G)_z \geq 1 + (n-2) + (n-2) + 2 = 2n - 1$$

from which the inequality

$$\mathrm{N}(G) = \mathrm{N}(G - z) + \mathrm{N}(G)_z \geq (n-1)(n-2) + 1 + 2n - 1 > n^2 - n + 1$$

follows. □

A cut vertex-free connected graph with at least three vertices is also referred to as a 2-connected graph.

Theorem 13. *For all $n \geq 3$, the cycle C_n has the smallest number of connected induced subgraphs among all graphs in the set $\mathcal{H}(n, 0)$.*

Proof. Throughout the proof, it is assumed that $n \geq 3$. Let $G \in \mathcal{H}(n, 0)$ be a graph of order n that minimises the number of connected induced subgraphs. Then G must necessary be minimally 2-connected. In other words, G must have the property that removing an edge in G destroys 2-connectivity. Moreover, in view of Lemma 12, G cannot be a special graph. Suppose that G is not a cycle. Clearly, we have $n \geq 5$. Let us prove that we can always identify $n^2 + n + 1 > n^2 - n + 1 = \mathrm{N}(C_n)$ connected induced subgraphs in G.

Let u, v be two non-adjacent vertices of G. By the vertex version of Menger's theorem [19], there must exist two internally vertex disjoint paths between u and v. Among all such $u - v$ paths, we choose two of them that are of smallest lengths. The vertex sets of these chosen paths must necessarily induce paths in G since otherwise, the property of these paths being shortest is violated. Let m denote the number of edges of G. Then the number of unordered pairs of non-adjacent vertices of G is $\binom{n}{2} - m$, and therefore G has at least

$$2\left(\binom{n}{2} - m\right) = n^2 - n - 2m$$

connected induced subgraphs, each of them is a path of order three or more.

Let x, y be two adjacent vertices of G. We claim that the graph $G - x - y$ is connected and moreover it is not a path. For the claim, suppose that $G - x - y$ is not connected and let G_1, G_2 be two (connected) components of $G - x - y$. Both x and y must have a neighbour in G_1 and G_2 because neither x nor y is a cut vertex of G. This implies that G contains a cycle that passes through x, y and never uses the edge xy. This cycle can be obtained as follows: let x_1 (resp. x_2) be a neighbour of x in G_1 (resp. G_2), and y_1 (resp. y_2) be a neighbour of y in G_1 (resp. G_2). Then this cycle is made of xx_1, a shortest $x_1 - y_1$ path in G_1, y_1y, yy_2, a shortest $y_2 - x_2$ path in G_2, and x_2x, in this order. However, by a result of Dirac [4, Theorem 3], this cannot happen in a minimally 2-connected graph. Hence, $G - x - y$ is connected. It remains to show that $G - x - y$ is not a path. Suppose to the contrary that $G - x - y$ is a path and let u_1, u_2 be the endvertices of $G - x - y$. Since G is cut vertex-free, both u_1 and u_2 must have x or y as a neighbour. In a minimally 2-connected graph, this gives rise to essentially two possibilities (up to exchanging the role of x and y, or u_1 and u_2) for G: either G itself is C_n, or G consists of a cycle C_{n-1} together with a vertex $z \notin V(C_{n-1})$ which is adjacent to precisely two vertices of C_{n-1}. The former situation is avoided by assumption while the latter defines G as a special graph. Hence, $G - x - y$ is not a path.

Let $x \in V(G)$. We further claim that the connected graph $G - x$ is not a path. To see this, note that if $G - x$ was a path, then its two endvertices would both be adjacent to x since G is 2-connected. In particular G would be a cycle since G is minimally 2-connected. Hence, $G - x$ is not a path.

Now we note that the following are all distinct connected induced subgraphs of G and none of them is a path of order at least three:

- all single vertices of G;
- all 2-vertex connected subgraphs of G;
- all subgraphs obtained by removing two adjacent vertices of G;
- all subgraphs obtained by removing one vertex of G;
- the whole graph G.

By combining all the above cases, we obtain $n + m + m + n + 1 = 2n + 2m + 1$ additional connected induced subgraphs of G that are not paths of order three or more. Together with the induced paths enumerated earlier, we conclude that

$$\mathrm{N}(G) \geq n^2 - n - 2m + 2n + 2m + 1 = n^2 + n + 1 > n^2 - n + 1 = \mathrm{N}(C_n)\,.$$

This completes the proof of the theorem. □

We observe that all blocks of a graph that minimises the number of connected induced subgraphs in the set $\mathcal{H}(n, c)$ must be minimally 2-connected. However, there are usually many minimally 2-connected graphs having the same order n. For $n \geq 3$, the sequence starts

$$1, 1, 2, 3, 6, 12, 28, 68, 184, 526, 1602, 5075, 16711, 56428, 195003, 685649, \dots,$$

see A003317 in [23]. It is then natural to formulate this intriguing problem for further investigation:

Problem 1. Find a constructive characterisation of those graphs with order n and $c > 0$ cut vertices that have the smallest number of connected induced subgraphs.

3. Connected graphs with p pendant vertices

We define $\mathcal{G}(n,p)$ to be the set of all connected graphs with n vertices of which p are pendant. In [7] Andriantiana and the author of the present paper investigated inequalities which relate the number of connected induced subgraphs of a graph to that of its complement. They also arrived at the following result which settles the extremal graph structure for the maximum number of connected induced subgraphs among all graphs in $\mathcal{G}(n,p)$.

Theorem 14 ([7])**.** *Let $G \in \mathcal{G}(n,p)$ with $n \geq 5$ and $0 \leq p \leq n-2$.*

- *If $p < n-2$, then we have*
$$\mathrm{N}(G) \leq 2^{n-1} + 2^{n-p-1} + p - 1\,.$$
Equality happens if and only if G can be obtained by identifying one vertex of K_{n-p} with the central vertex of S_{p+1}.
- *If $p = n-2$, then we have*
$$\mathrm{N}(G) \leq n + 3 \cdot 2^{n-3}\,.$$
Equality happens if and only if G can be obtained by inserting one vertex into an edge of S_{n-1}.

In order to obtain the minimisation counterpart of Theorem 14, we need to state two intermediate results.

3.1. **The case $p \neq 0$.** Sharp bounds on the number of connected induced subgraphs in terms of order were obtained in [5]. One of the results in [5] will be needed for our purpose.

Theorem 15 ([5]–Theorem 9)**.** *If G is a unicylic graph of order n, then*
$$N(G) \geq (n^2 + 3n - 4)/2.$$
The bound is attained if and only if G can be obtained by identifying a vertex of K_3 with a leaf of P_{n-2}.

At this stage, recall that $\mathbb{T}^1(n,p)$ is the tree obtained from the vertex disjoint graphs $S_{1+\lfloor p/2 \rfloor}$ and $S_{1+\lceil p/2 \rceil}$ by identifying their central vertices with the two leaves of P_{n-p}, respectively. We recall Li and Wang's result as stated in the introduction.

Theorem 16 ([18]–Theorem 1)**.** *If $n \geq 4$ and $2 \leq p \leq n-2$, then $\mathbb{T}^1(n,p)$ is the unique tree with order n and p leaves that attains the minimum number of subtrees.*

Our next theorem, which is essentially extracted from Theorem 16, reads as follows:

Theorem 17. *Let $G \in \mathcal{G}(n,p)$ with $n \geq 4$ and $1 \leq p \leq n-2$.*

- *If $p = 1$, then we have*
$$\mathrm{N}(G) \geq (n^2 + 3n - 4)/2\,.$$
Equality happens if and only if G can be obtained by identifying a vertex of K_3 with a leaf of P_{n-2}.

- *If $p \neq 1$, then we have*
$$\mathrm{N}(G) \geq 2^p + (n-p-1)(2^{\lfloor p/2 \rfloor} + 2^{\lceil p/2 \rceil}) + p + (n-p-1)(n-p-2)/2\,.$$
Equality happens if and only if G is isomorphic to $\mathbb{T}^1(n,p)$.

Proof. Suppose that $p = 1$. Then G is not a tree. One can then remove edges (possibly none) from G to get a new connected graph G' of order n that contains exactly one cycle. It follows from Theorem 15 that $\mathrm{N}(G) \geq \mathrm{N}(G') \geq (n^2+3n-4)/2$. Moreover, the unique graph attaining this bound also has exactly one pendant vertex as it can be obtained by identifying a vertex of K_3 with a leaf of P_{n-2}. Thus the result follows in this case.

Now suppose that $p \neq 1$. Let us first derive a formula for $\mathrm{N}(\mathbb{T}^1(n,p))$. Denote by u and v the two vertices of $\mathbb{T}^1(n,p)$ such that $\deg_{\mathbb{T}^1(n,p)}(u) = 1 + \lfloor p/2 \rfloor$ and $\deg_{\mathbb{T}^1(n,p)}(v) = 1 + \lceil p/2 \rceil$. We have

$$\begin{aligned}
\mathrm{N}(\mathbb{T}^1(n,p)) &= \mathrm{N}(\mathbb{T}^1(n,p))_{u,v} + \mathrm{N}(\mathbb{T}^1(n,p)-v)_u + \mathrm{N}(\mathbb{T}^1(n,p)-u)_v + \mathrm{N}(\mathbb{T}^1(n,p)-u-v)\\
&= 2^{\lfloor p/2 \rfloor + \lceil p/2 \rceil} + (n-p-1)2^{\lfloor p/2 \rfloor} + (n-p-1)2^{\lceil p/2 \rceil}\\
&\quad + \big(\lfloor p/2 \rfloor + (n-p-1)(n-p-2)/2 + \lceil p/2 \rceil\big)\,.
\end{aligned}$$

We claim that $\mathrm{N}(\mathbb{T}^1(n,p))$ is an increasing function in p. Indeed, we have

$$\begin{aligned}
\mathrm{N}(\mathbb{T}^1(n,p)) &= 2^p + (n-p-1)(2^{p/2} + 2^{p/2}) + p + \binom{n-p-1}{2}\,,\\
\mathrm{N}(\mathbb{T}^1(n,p+1)) &= 2^{p+1} + (n-p-2)(2^{p/2} + 2^{p/2+1}) + p + 1 + \binom{n-p-2}{2}
\end{aligned}$$

if p is even, and

$$\begin{aligned}
\mathrm{N}(\mathbb{T}^1(n,p)) &= 2^p + (n-p-1)(2^{(p-1)/2} + 2^{(p+1)/2}) + p + \binom{n-p-1}{2}\,,\\
\mathrm{N}(\mathbb{T}^1(n,p+1)) &= 2^{p+1} + (n-p-2)(2^{(p+1)/2} + 2^{(p+1)/2}) + p + 1 + \binom{n-p-2}{2}
\end{aligned}$$

if p is odd. In particular, we get

$$\mathrm{N}(\mathbb{T}^1(n,p+1)) - \mathrm{N}(\mathbb{T}^1(n,p)) = 2^p - 1 + (n-p-4)(2^{p/2}-1) \geq 1 + 2^p - 2^{p/2+1} > 0$$

if p is even, and

$$\begin{aligned}
\mathrm{N}(\mathbb{T}^1(n,p+1)) - \mathrm{N}(\mathbb{T}^1(n,p)) &= 2^p - 2 + (n-p-5)(2^{(p-1)/2}-1)\\
&\geq 1 + 2^p - 2^{(p-1)/2} - 2^{(p+1)/2} > 0
\end{aligned}$$

if p is odd. Let T be a spanning tree of G and note that T has at least p leaves. Since $\mathrm{N}(\mathbb{T}^1(n,p))$ is an increasing function in p, we deduce from Theorem 16 that

$$\mathrm{N}(G) \geq \mathrm{N}(T) \geq \mathrm{N}(\mathbb{T}^1(n,p(T))) > \mathrm{N}(\mathbb{T}^1(n,p))$$

if $p(T) \neq p$. If $p(T) = p$, then we have

$$\mathrm{N}(G) \geq \mathrm{N}(T) \geq \mathrm{N}(\mathbb{T}^1(n,p))\,,$$

and the inequality becomes an equality if and only if G is isomorphic to the tree $\mathbb{T}^1(n,p)$. This completes the proof of the theorem. □

3.2. **The case** $p = 0$**.** Let l, n, r be three positive integers such that $l, r \geq 3$ and $n \geq l + r$. We define the *double tadpole* graph $D_n(l;r)$ as the graph constructed from the three pairwise vertex disjoint graphs $C_l, C_r, P_{n+2-l-r}$ by taking $u \in V(C_l)$, $v \in V(C_r)$ and identifying u with one leaf of $P_{n+2-l-r}$ and v with the other leaf of $P_{n+2-l-r}$.

For $n > 5$, we shall prove that the double tadpole graph $D_n(3;3)$ has the smallest number of connected induced subgraphs among all graphs in the set $\mathcal{G}(n,0)$, and that $D_n(3;3)$ is unique with this property.

We first give some important preliminaries, then formally state and prove our result. From here onwards, we shall simply write D_n instead of $D_n(3;3)$.

Proposition 18. *For the double tadpole graph* $D_n(l;r)$*, we have*

$$\mathrm{N}(D_n(3;3)) = \mathrm{N}(D_n) = \frac{(n-1)(n+6)}{2}\,.$$

Furthermore, if $(l,r) \neq (3,3)$*, then we have*

$$\mathrm{N}(D_n(l;r)) > \mathrm{N}(D_n)\,.$$

Proof. Let $u, v \in V(D_n(l;r))$ be the two vertices of $D_n(l;r)$ whose degree is 3. We use our standard decomposition with respect to u, v:

$$\begin{aligned}
\mathrm{N}(D_n(l;r)) &= \mathrm{N}(D_n(l;r))_{u,v} + \mathrm{N}(D_n(l;r)-v)_u + \mathrm{N}(D_n(l;r)-u)_v + \mathrm{N}(D_n(l;r)-u-v)\\
&= \mathrm{N}(C_l)_u \cdot \mathrm{N}(C_r)_v + (n+1-l-r)(\mathrm{N}(C_l)_u + \mathrm{N}(C_r)_v)\\
&\quad + \mathrm{N}(P_{l-1}) + \mathrm{N}(P_{n-l-r}) + \mathrm{N}(P_{r-1})\\
&= \Big(1+\binom{l}{2}\Big)\,\mathrm{N}(C_r)_v + (n+1-l-r)\Big(1+\binom{l}{2} + \mathrm{N}(C_r)_v\Big)\\
&\quad + \binom{l}{2} + \binom{n+1-l-r}{2} + \mathrm{N}(P_{r-1})\,.
\end{aligned}$$

Assume that $r \geq l \geq 4$. Taking the difference $\mathrm{N}(D_n(l;r)) - \mathrm{N}(D_n(l-1;r))$, we get

$$\begin{aligned}
\mathrm{N}(D_n(l;r)) - \mathrm{N}(D_n(l-1;r)) &= (l-1)\,\mathrm{N}(C_r)_v + (l-1)(n+1-l-r)\\
&\quad - \Big(1+\binom{l-1}{2} + \mathrm{N}(C_r)_v\Big) + (l-1) - (n+1-l-r)\\
&= (l-2)\,\mathrm{N}(C_r)_v - \mathrm{N}(C_l)_u + (l-2)(n+1-l-r) + 2(l-1) > 0
\end{aligned}$$

since $\mathrm{N}(C_r)_v \geq \mathrm{N}(C_l)_u$ and $n \geq l + r$. It follows from this inequality that the minimum of $\mathrm{N}(D_n(l;r))$, given r, is attained when $l = 3$. We have

$$\mathrm{N}(D_n(3;3)) = \mathrm{N}(D_n) = 22 + 8(n-5) + \binom{n-5}{2} = \frac{(n-1)(n+6)}{2}\,.$$

Assume that $l = 3$ and $r \geq 4$. Taking the difference $\mathrm{N}(D_n(3;r)) - \mathrm{N}(D_n)$, we get

$$\begin{aligned}\mathrm{N}(D_n(3;r)) - \mathrm{N}(D_n) &= 4 \cdot \mathrm{N}(C_r)_v + (n-r-2)\big(4 + \mathrm{N}(C_r)_v\big) \\ &\quad + 3 + \binom{n-r-2}{2} + \binom{r}{2} - \frac{(n-1)(n+6)}{2} \\ &= \frac{(r-3)(-r^2 + (n+2)r - 2)}{2} > 0\end{aligned}$$

after a simple manipulation (recall that $n \geq 3 + r$). It follows from this inequality that the minimum of $\mathrm{N}(D_n(3;r))$ is attained when $r = 3$. □

Lemma 19. *If G is a graph constructed from two vertex disjoint cycles C_l and C_{n-l+1} by identifying $u \in V(C_l)$ with $v \in V(C_{n-l+1})$, then we have*

$$\mathrm{N}(G) > N(C_n) > \mathrm{N}(D_n)$$

for all $n > 5$.

Proof. Simply note that

$$\mathrm{N}(G) = \mathrm{N}(C_l)_u \cdot \mathrm{N}(C_{n-l+1})_v + \mathrm{N}(P_{l-1}) + \mathrm{N}(P_{n-l})\,,$$

and that the difference $\mathrm{N}(G) - \mathrm{N}(C_n)$ is given by

$$\begin{aligned}\mathrm{N}(G) - \mathrm{N}(C_n) &= \Big(1 + \binom{l}{2}\Big)\Big(1 + \binom{n-l+1}{2}\Big) + \binom{l}{2} + \binom{n-l+1}{2} - (n^2 - n + 1) \\ &= \frac{(l-1)(l-n)(l^2 - (n+1)l + 8)}{4} > 0\end{aligned}$$

as $3 \leq l \leq n-2$. Likewise, the difference $\mathrm{N}(C_n) - \mathrm{N}(D_n)$ is given by

$$\begin{aligned}\mathrm{N}(C_n) - \mathrm{N}(D_n) &= (n^2 - n + 1) - \frac{(n-1)(n+6)}{2} \\ &= \frac{n^2 - 7n + 8}{2} > 0\,.\end{aligned}$$

□

Lemma 20. *Let L, R be two fixed non-trivial vertex disjoint connected graphs such that $u \in V(L)$ and $v \in V(R)$. Consider two vertex disjoint paths P_k, P_q for some $q \geq 2$. Identify u with both a leaf of P_k as well as a leaf of P_q; further, identify v with the other leaf of P_q. Denote by $H(k;q)$ the resulting graph. If $k > 1$, then we have*

$$\mathrm{N}(H(k;q)) > \mathrm{N}(H(1;q+k-1))\,.$$

Proof. We use our standard decomposition again:

$$\begin{aligned}\mathrm{N}(H(k;q)) &= \mathrm{N}(H(k;q))_{u,v} + \mathrm{N}(H(k;q)-v)_u + \mathrm{N}(H(k;q)-u)_v + \mathrm{N}(H(k;q)-u-v) \\ &= k \cdot \mathrm{N}(L)_u \cdot \mathrm{N}(R)_v + (q-1)(k \cdot \mathrm{N}(L)_u + \mathrm{N}(R)_v) \\ &\quad + \mathrm{N}(L-u) + \binom{k}{2} + \binom{q-1}{2} + \mathrm{N}(R-v)\,,\end{aligned}$$

and

$$\mathrm{N}(H(1;q+k-1)) = \mathrm{N}(L)_u \cdot \mathrm{N}(R)_v + (q+k-2)(\mathrm{N}(L)_u + \mathrm{N}(R)_v) \\ + \mathrm{N}(L-u) + \binom{q+k-2}{2} + \mathrm{N}(R-v)\,.$$

It follows that

$$\mathrm{N}(H(k;q)) - \mathrm{N}(H(1;q+k-1)) = (k-1)(\mathrm{N}(L)_u - 1)\,\mathrm{N}(R)_v \\ + (k-1)(q-2)\,\mathrm{N}(L)_u + \binom{k}{2} + \binom{q-1}{2} - \binom{q+k-2}{2}\,.$$

On the other hand, we have

$$2(k-1)(q-2) + \binom{k}{2} + \binom{q-1}{2} - \binom{q+k-2}{2} = (k-1)(q-2) \geq 0\,.$$

Therefore, using the assumption that $\mathrm{N}(L)_u \geq 2$, we derive that

$$\mathrm{N}(H(k;q)) - \mathrm{N}(H(1;q+k-1)) \geq (k-1)(\mathrm{N}(L)_u - 1)\,\mathrm{N}(R)_v + (k-1)(q-2) > 0$$

provided that $k \neq 1$. This completes the proof. □

Our next lemma captures the special case $q = 1$ that is missing in Lemma 20.

Lemma 21. *Let L, R be two fixed non-trivial vertex disjoint connected graphs such that $u \in V(L)$ and $v \in V(R)$. Consider the path $P_k, k \geq 2$ and let w be a leaf of P_k. Identify w with both u and v. Denote by $H(k;1)$ the resulting graph. Then we have*

$$\mathrm{N}(H(k;1)) > \mathrm{N}(H(1;k))\,,$$

where $H(1;k)$ is the graph described in Lemma 20.

Proof. Simply note that

$$\mathrm{N}(H(k;1)) = k \cdot \mathrm{N}(L)_u \cdot \mathrm{N}(R)_v + \mathrm{N}(L-u) + \binom{k}{2} + \mathrm{N}(R-v)\,,$$

and that

$$\mathrm{N}(H(1;k)) = \mathrm{N}(L)_u \cdot \mathrm{N}(R)_v + (k-1)(\mathrm{N}(L)_u + \mathrm{N}(R)_v) \\ + \mathrm{N}(L-u) + \binom{k-1}{2} + \mathrm{N}(R-v)\,.$$

In particular, we get

$$\mathrm{N}(H(k;1)) - \mathrm{N}(H(1;k)) = (k-1)(\mathrm{N}(L)_u - 1)(\mathrm{N}(R)_v - 1) > 0\,.$$

□

The following lemma is a variant of the combination of Lemmas 20 and 21.

Lemma 22. *Let L, R be two fixed non-trivial vertex disjoint connected graphs such that $u, w \in V(L)$, $u \neq w$ and $v \in V(R)$. Consider three vertex disjoint paths P_k, P_q, P_{q+k-1} for some $k > 1$. Let G_1, G_2 be the two graphs constructed as follows:*

- *If $q>1$, then identify w with a leaf of P_k, u with a leaf of P_q, and v with the other leaf of P_q to obtain G_1. If $q=1$, then identify w with a leaf of P_k, and u with v to obtain G_1.*
- *Identify u with a leaf of P_{q+k-1}, and v with the other leaf of P_{q+k-1} to obtain G_2.*

We have

$$|V(G_1)| = |V(G_2)| \quad \textit{and} \quad \mathrm{N}(G_1) > \mathrm{N}(G_2)\,.$$

Proof. Denote by J the subgraph of G_1 that consists of L and a leaf of P_k attached to L at vertex w. Assume that $q>1$. Then we have

$$\begin{aligned}\mathrm{N}(G_1) &= \mathrm{N}(G_1)_{u,v} + \mathrm{N}(G_1-v)_u + \mathrm{N}(G_1-u)_v + \mathrm{N}(G_1-u-v)\\ &= \mathrm{N}(J)_u\cdot\mathrm{N}(R)_v + (q-1)(\mathrm{N}(J)_u+\mathrm{N}(R)_v)\\ &+\mathrm{N}(J-u) + \binom{q-1}{2} + \mathrm{N}(R-v)\,,\end{aligned}$$

and

$$\begin{aligned}\mathrm{N}(G_2) = \mathrm{N}(L)_u\cdot\mathrm{N}(R)_v + (q+k-2)(\mathrm{N}(L)_u+\mathrm{N}(R)_v)\\ +\mathrm{N}(L-u) + \binom{q+k-2}{2} + \mathrm{N}(R-v)\,.\end{aligned}$$

In particular, we get

$$\begin{aligned}\mathrm{N}(G_1)-\mathrm{N}(G_2) = \mathrm{N}(R)_v(\mathrm{N}(J)_u - \mathrm{N}(L)_u + 1 - k) + (q-1)\,\mathrm{N}(J)_u - (q+k-2)\,\mathrm{N}(L)_u\\ +\mathrm{N}(J-u) - \mathrm{N}(L-u) + \binom{q-1}{2} - \binom{q+k-2}{2}\,.\end{aligned}$$

On the other hand, we have

$$\mathrm{N}(J)_u \geq \mathrm{N}(L)_u + k\cdot\mathrm{N}(L)_{u,w} \geq \mathrm{N}(L)_u + k$$

and

$$\mathrm{N}(J-u) \geq \mathrm{N}(L-u) + \binom{k}{2} + k\cdot\mathrm{N}(L-u)_w \geq \mathrm{N}(L-u) + \binom{k}{2} + k\,.$$

This implies that

$$\begin{aligned}\mathrm{N}(G_1)-\mathrm{N}(G_2) &\geq \mathrm{N}(R)_v + (q-1)(k+\mathrm{N}(L)_u) - (q+k-2)\,\mathrm{N}(L)_u + k\\ &+\binom{q-1}{2} + \binom{k}{2} - \binom{q+k-2}{2}\\ &= \mathrm{N}(R)_v + k\cdot q + (k-1)(q-2)\,\mathrm{N}(L)_u\\ &+\binom{q-1}{2} + \binom{k}{2} - \binom{q+k-2}{2}\,.\end{aligned}$$

It follows from the identity

$$2(k-1)(q-2) + \binom{k}{2} + \binom{q-1}{2} - \binom{q+k-2}{2} = (k-1)(q-2)$$

that

$$\mathrm{N}(G_1) - \mathrm{N}(G_2) \geq \mathrm{N}(R)_v + k \cdot q + (k-1)(q-2) > 0\,.$$

Assume that $q = 1$. Then we have

$$\begin{aligned}\mathrm{N}(G_1) &= \mathrm{N}(G_1)_u + \mathrm{N}(G_1 - u) = \mathrm{N}(J)_u \cdot \mathrm{N}(R)_v + \mathrm{N}(J-u) + \mathrm{N}(R-v) \\ &\geq \mathrm{N}(R)_v(\mathrm{N}(L)_u + k) + \mathrm{N}(L-u) + \binom{k}{2} + k + \mathrm{N}(R-v)\,,\end{aligned}$$

and

$$\begin{aligned}\mathrm{N}(G_2) &= \mathrm{N}(L)_u \cdot \mathrm{N}(R)_v + (k-1)(\mathrm{N}(L)_u + \mathrm{N}(R)_v) + \mathrm{N}(L-u) + \binom{k-1}{2} + \mathrm{N}(R-v) \\ &\geq \mathrm{N}(L)_u \cdot \mathrm{N}(R)_v + (k-1)(2 + \mathrm{N}(R)_v) + \mathrm{N}(L-u) + \binom{k-1}{2} + \mathrm{N}(R-v)\,.\end{aligned}$$

In particular, we get

$$\mathrm{N}(G_1) - \mathrm{N}(G_2) \geq \mathrm{N}(R)_v + \binom{k}{2} - \binom{k-1}{2} + k - 2(k-1) = 1 + \mathrm{N}(R)_v > 0\,.$$

This completes the proof of the lemma. □

We finish our preliminaries with the following lemma, which is similar in nature but different to Lemma 6 (see Section 2).

Lemma 23. *Let H be a connected graph of order greater than two, and u, v two distinct vertices of H such that $\mathrm{N}(H)_{u,v} > 1$ and $\mathrm{N}(H-v)_u \leq \mathrm{N}(H-u)_v$. Let $H(n_1; n_2)$ be the graph obtained from H by identifying u with a leaf of P_{n_1}, and v with a leaf of P_{n_2} for some $n_1, n_2 \geq 1$. We have*

$$\mathrm{N}(H(n_1; n_2)) \geq \mathrm{N}(H(n_1 + n_2 - 1; 1))\,.$$

Moreover, the inequality is strict if $n_1, n_2 > 1$.

Proof. By the proof of Lemma 6, we have

$$\begin{aligned}\mathrm{N}(H(n_1; n_2)) &= n_1 \cdot n_2 \cdot \mathrm{N}(H)_{u,v} + n_1 \cdot \mathrm{N}(H-v)_u + n_2 \cdot \mathrm{N}(H-u)_v \\ &\quad + \mathrm{N}(P_{n_1-1}) + \mathrm{N}(P_{n_2-1}) + \mathrm{N}(H-u-v)\,.\end{aligned}$$

In particular, we get

$$\begin{aligned}\mathrm{N}(H(n_1+n_2-1; 1)) &= (n_1+n_2-1)\,\mathrm{N}(H)_{u,v} + (n_1+n_2-1)\,\mathrm{N}(H-v)_u + \mathrm{N}(H-u)_v \\ &\quad + \mathrm{N}(P_{n_1+n_2-2}) + \mathrm{N}(H-u-v)\,,\end{aligned}$$

which implies that

$$\begin{aligned}\mathrm{N}(H(n_1{+}n_2-1;1)) - \mathrm{N}(H(n_1;n_2)) &= (n_1+n_2-1-n_1\cdot n_2)\,\mathrm{N}(H)_{u,v}\\ &+ (n_2-1)\,\mathrm{N}(H-v)_u + (1-n_2)\,\mathrm{N}(H-u)_v\\ &+ \binom{n_1+n_2-1}{2} - \binom{n_1}{2} - \binom{n_2}{2}\\ &= (n_2-1)(\mathrm{N}(H-v)_u - \mathrm{N}(H-u)_v) - (n_1-1)(n_2-1)(\mathrm{N}(H)_{u,v}-1) \le 0\,.\end{aligned}$$

Moreover, we have $\mathrm{N}(H(n_1+n_2-1;1)) < \mathrm{N}(H(n_1;n_2))$ if $n_1 > 1$ and $n_2 > 1$. The statement of the lemma follows. □

By rooted path, we mean a path rooted at one of its leaves. Our main result reads as follows:

Theorem 24. *Let $n > 5$ be a positive integer. For every graph $G \in \mathcal{G}(n,0)$, we have*

$$\mathrm{N}(G) \ge \mathrm{N}(D_n) = \frac{(n-1)(n+6)}{2}\,,$$

and $D_n \in \mathcal{G}(n,0)$ is the only graph with this property.

Proof. Let $\mathbb{G} \in \mathcal{G}(n,0)$ be a connected graph with order n and no pendant vertex that minimises the number of connected induced subgraphs. We are going to show that $\mathbb{G}$ can be obtained from certain graphs $H_1 \in \mathcal{G}(n,0)$ through a series of graph transformations that preserve the number of vertices.

First off, note that $\mathbb{G}$ must have at least one cut vertex, since otherwise $\mathrm{N}(\mathbb{G}) \ge \mathrm{N}(C_n)$ by virtue of Theorem 13, while Lemma 19 implies that $\mathrm{N}(C_n) > \mathrm{N}(D_n)$. Fix $H_1 \in \mathcal{G}(n,0)$ such that H_1 has at least one cut vertex. If we remove edges from a graph, the number of connected induced subgraphs decreases. Starting from H_1, we can thus remove certain edges until we reach a connected graph with only two distinct cyclic blocks, say B_1, B_2. More precisely, all blocks, except only two of H_1 are replaced with any generic of their spanning trees. This yields a new graph H_2 which may contain a pendant vertex. Moreover, we have $\mathrm{N}(H_1) \ge \mathrm{N}(H_2)$.

In the graph H_2, we can remove edges from the blocks B_1, B_2 in such a way that the two cyclic blocks of the resulting graph, say H_3 are all cycles, say C_l and C_r. Hence, H_3 consists of two distinct cycles C_l, C_r 'separated' by a (possibly trivial) path P, together with some trees attached to all vertices of $V(C_l) \cup V(C_r) \cup V(P)$ in H_3; see Figure 2 for a picture. Moreover, we have $\mathrm{N}(H_2) \ge \mathrm{N}(H_3)$.

In the graph H_3, replace all components C of $H_3 - (E(C_l) \cup E(C_r) \cup E(P))$ with a rooted path of order $|V(C)|$ rooted at the unique vertex of C that belongs to $V(C_l) \cup V(C_r) \cup V(P)$. This gives us a new graph H_4. We claim that $\mathrm{N}(H_3) \ge \mathrm{N}(H_4)$ with equality if and only if H_3 and H_4 are isomorphic. Indeed, construct from H_3 a new graph H_3' by replacing (without loss of generality) M_1 with the rooted path $P_{|V(M_1)|}$ whose root is v_1. Thus, $H_3 - V(M_1 - v_1)$ and $H_3' - V(M_1 - v_1)$ are isomorphic graphs. On the other hand, if A denotes the number of connected induced subgraphs of $H_3 - V(M_1 - v_1) = H_3' - V(M_1 - v_1)$

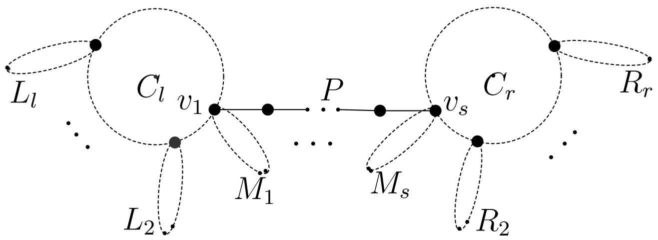

FIGURE 2. The graph H_3 in the proof of Theorem 24: P is a path starting at v_1 and ending at v_s; $L_2, \dots, L_l, M_1, \dots, M_s, R_2, \dots, R_r$ are all trees.

that contain v_1, then

$$\mathrm{N}(M_1 - v_1) + A \cdot \mathrm{N}(M_1)_{v_1} \text{ (resp. } \mathrm{N}(P_{|V(M_1)|-1}) + A \cdot \mathrm{N}(P_{|V(M_1)|})_{v_1})$$

counts precisely the number of connected induced subgraphs of H_3 (resp. H_3') that contain a vertex of $M_1 - v_1$. Since the path P_m (rooted at a leaf) minimises both the total number of subtrees and the number of subtrees containing a specific vertex u among all m-vertex trees (see Székely and Wang [27]), we deduce that $\mathrm{N}(P_{|V(M_1)|-1}) \leq \mathrm{N}(M_1 - v_1)$ and $\mathrm{N}(P_{|V(M_1)|})_{v_1} \leq \mathrm{N}(M_1)_{v_1}$. This implies that

$$\mathrm{N}(H_3') - \mathrm{N}(H_3) = \mathrm{N}(P_{|V(M_1)|-1}) - \mathrm{N}(M_1 - v_1) + A(\mathrm{N}(P_{|V(M_1)|})_{v_1} - \mathrm{N}(M_1)_{v_1}) \leq 0\,.$$

Hence, we have $\mathrm{N}(H_3') \leq \mathrm{N}(H_3)$. Equality holds if and only if M_1 is a rooted path (see Székely and Wang [27]), i.e. H_3 and H_3' are isomorphic graphs. Since H_4 can be obtained from H_3 by a repetitive application of this process of moving from H_3 to H_3', we derive that $\mathrm{N}(H_4) \leq \mathrm{N}(H_3)$ with equality if and only if H_3 and H_4 are isomorphic.

In the graph H_4, fix two distinct vertices $u, v \in V(C_l) \cup V(C_r) \cup V(P)$, and consider H_4 as the graph $H(n_1; n_2)$ described in Lemma 23 where n_1 (resp. n_2) is the order of the path attached at u (resp. v) in H_4. Lemma 23 states that whenever $n_1 > 1$ or $n_2 > 1$, two new graphs $H(n_1 + n_2 - 1; 1)$ and $H(1; n_1 + n_2 - 1)$ can always be constructed from $H(n_1; n_2)$ such that at least one of the inequalities

$$\mathrm{N}(H(n_1; n_2)) > \mathrm{N}(H(n_1 + n_2 - 1; 1)) \quad \text{and} \quad \mathrm{N}(H(n_1; n_2)) > \mathrm{N}(H(1; n_1 + n_2 - 1))$$

holds. In other words, this shows that a graph H_5 with order n and the property that $\mathrm{N}(H_4) \geq \mathrm{N}(H_5)$, can be obtained from H_4 by making all components (paths), except possibly only one of $H_4 - (E(C_l) \cup E(C_r) \cup E(P))$ trivial. This leaves H_5 with two possible shapes if $H_5 - (E(C_l) \cup E(C_r) \cup E(P))$ has a non trivial component (rooted path), say P_k:

- Vertex v_1 or v_s is the root of P_k. In this case, we invoke Lemma 21 or Lemma 20 on H_5 depending on whether the path P 'separating' the cycles C_l and C_r in H_5 is trivial or not;
- Neither v_1 nor v_s is the root of P_k. In this case, we apply Lemma 22.

In either case, the combination of Lemmas 20, 21, and 22 shows the existence of another graph $H_6 \in \mathcal{G}(n, 0)$ with the property that $\mathrm{N}(H_5) \geq \mathrm{N}(H_6)$ with equality if and only if H_5 and H_6 are isomorphic. Moreover, by construction H_6 is either a double tadpole graph, or a graph constructed from two vertex disjoint cycles C_l and C_{n-l+1} by identifying $u \in V(C_l)$ with $v \in V(C_{n-l+1})$. The latter situation corresponds to the graph described in Lemma 19. Consequently, H_6 can only be a double tadpole graph if $\mathrm{N}(H_6)$ is to be the minimum number of connected induced subgraphs that a connected graph with order n and no pendant vertices can have.

Finally, we invoke Proposition 18 on H_6 to obtain the double tadpole graph D_n which satisfies $\mathrm{N}(H_6) > \mathrm{N}(D_n)$ provided that $H_6 \neq D_n$. Summing up, we have proved that $\mathbb{G}$ is indeed the double tadpole graph D_n. □

Acknowledgement

The author is pleased to acknowledge a discussion with Stephan Wagner.

References

[1] E. O. D. Andriantiana, S. Wagner and H. Wang. Greedy trees, subtrees and antichains *The Electronic Journal of Combinatorics*, 20(3), #P28, 2013.

[2] B. Bollobás. Extremal graph theory. Vol. 11 of *London Mathematical Society Monographs*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London - New York, 1978.

[3] K. Burns and R. C. Entringer. A graph–theoretic view of the United States postal service. *Combinatorics, Graph Theory, and Algorithms*, Wiley, New York, 323–334, 1995.

[4] G. A. Dirac. Minimally 2-connected graphs. *Journal für die reine und angewandte Mathematik*, 228:204–216, 1967.

[5] A. A. V. Dossou-Olory. Graphs and unicyclic graphs with extremal connected subgraphs. Preprint, arXiv:1812.02422v2, 2019.

[6] A. A. V. Dossou-Olory. Maximising the number of connected induced subgraphs of unicyclic graphs. Preprint, 2018.

[7] A. A. V. Dossou-Olory and E. O. D. Andriantiana. Nordhaus-Gaddum inequalities for the number of connected induced subgraphs. In preparation, 2019.

[8] J. K. Doyle and J. E. Graver. Mean distance in a graph. *Discrete Mathematics*, 17(2):147-154, 1977.

[9] R. C. Entringer. Bounds for the average distance-inverse degree product in trees. *Combinatorics, Graph Theory, and Algorithms*, New Issues Press, Kalamazoo, 335–352, 1999.

[10] R. C. Entringer, D. E. Jackson and D. A. Snyder. Distance in graphs. *Czechoslovak Mathematical Journal*, 26(2): 283–296, 1976.

[11] F. R. K. Graham, R. L. Chung, and D. Coppersmith. On trees containing all small trees. *The Theory and Applications of Graphs*, John Wiley and Sons, 265–272, 1981.

[12] F. Harary. An elementary theorem on graphs. *The American Mathematical Monthly*, 66(5):405–407, 1959.

[13] F. Harary. A characterization of block-graphs. *Canadian Mathematical Bulletin*, 6(1):1–6, 1963.

[14] F. Harary and G. Prins. The block-cutpoint-tree of a graph. *Publicationes Mathematicae - Debrecen*, 13:103–107, 1966.

[15] R. E. Jamison. Monotonicity of the mean order of subtrees. *Journal of Combinatorial Theory, Series B*, 37(1):70–78, 1984.

[16] R. E. Jamison. On the average number of nodes in a subtree of a tree. *Journal of Combinatorial Theory, Series B*, 35(3):207–223, 1983.

[17] R. Kirk and H. Wang. Largest number of subtrees of trees with a given maximum degree. *SIAM Journal on Discrete Mathematics*, 22(3):985–995, 2008.

[18] S. Li and S. Wang. Further analysis on the total number of subtrees of trees. *The Electronic Journal of Combinatorics*, 19(4), #P48, 2012.

[19] K. Menger. Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10(1):96–115, 1927.
[20] D. Pandey and K. L. Patra. Extremizing the number of connected subgraphs of graphs. Preprint, arXiv:1811.11411, 2018.
[21] D. Pandey and K. L. Patra. Wiener index of graphs with fixed number of pendant or cut vertices. Preprint arXiv:1907.13481, 2019.
[22] J. Plesník. On the sum of all distances in a graph or digraph. *J. Graph Theory*, 8(1):1–21, 1984.
[23] N. J. A. Sloane. The Online Encyclopedia of Integer Sequences. Published electronically at http://oeis.org, 2019.
[24] R. Shi. The average distance of trees. *Systems Science and Mathematical Sciences*, 6(1):18–24, 1993.
[25] M. Simonovits. Extremal graph theory, Selected topics in graph theory, II. *Academic Press, London*, New York, San Francisco, 161–200, 1983.
[26] L. A. Székely and H. Wang. Binary trees with the largest number of subtrees. *Discrete Applied Mathematics*, 155(3):374–385, 2007.
[27] L. A. Székely and H. Wang. On subtrees of trees. *Advances in Applied Mathematics*, 34(1):138–155, 2005.
[28] L. A. Székely, H. Wang and T. Y. Wu. The sum of the distances between the leaves of a tree and the 'semi-regular' property. *Discrete Mathematics*, 311:(13), 1197–1203, 2011.
[29] P. Turán. Eine extremalaufgabe aus der graphentheorie. *Mat. Fiz. Lapok*, 48: 436–452, 1941.
[30] S. Wagner. Correlation of graph-theoretical indices. *SIAM Journal on Discrete Mathematics*, 21(1): 33–46, 2007.
[31] H. Wiener. Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69(1):17–20, 1947.
[32] W. Yan and Y.-N. Yeh. Enumeration of subtrees of trees. *Theoretical Computer Science*, 369(1-3):256–268, 2006.

Audace A. V. Dossou-Olory, Department of Mathematics and Applied Mathematics, University of Johannesburg, P.O. Box 524, Auckland Park, Johannesburg 2006, South Africa

E-mail address: audace@aims.ac.za