

# On the mean connected induced subgraph order of cographs

MATTHEW E. KROEKER\*    LUCAS MOL

ORTRUD R. OELLERMANN†

*University of Winnipeg  
Winnipeg, MB, R3B 2E9  
CANADA*

mattkroeker@shaw.ca    l.mol@uwinnipeg.ca  
o.oellermann@uwinnipeg.ca

## Abstract

In this article, the extremal structures for the mean order of connected induced subgraphs of cographs are determined. It is shown that among all connected cographs of order  $n \geq 7$ , the star  $K_{1,n-1}$  has maximum mean connected induced subgraph order, and for  $n \geq 3$ , the  $n$ -skillet,  $K_1 + (K_1 \cup K_{n-2})$ , has minimum mean connected induced subgraph order. It is deduced that the *density* for connected cographs (i.e. the ratio of the mean to the order of the graph) is asymptotically  $1/2$ . The mean order of all connected induced subgraphs containing a given vertex  $v$  of a cograph  $G$ , called the *local mean* of  $G$  at  $v$ , is shown to be at least as large as the mean order of all connected induced subgraphs of  $G$ , called the *global mean* of  $G$ .

## 1 Introduction

The study of the mean order of subtrees of a tree was initiated by Jamison in [6, 7]. Several open problems posed in [6] were subsequently studied in [9, 11, 12, 13]. In this article we extend the concept of the mean subtree order of trees to other graphs. Since the subtrees of a tree are precisely the connected induced subgraphs of a tree, it is natural to study the problem of finding the mean order of the connected induced subgraphs of a (connected) graph. This extension is of interest since it also relates to the study of node reliability, as we will discuss. One could also extend Jamison's

---

\* Supported by a Natural Sciences and Engineering Research Council of Canada USRA, number 510336-2017.

† Supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant, number RGPIN-2016-05237.

work [6] by studying the mean order of connected subgraphs (or trees) induced by sets of *edges*. However, these measures do not have an immediate connection to node reliability, and also appear to be more difficult to compute in general.

If  $G$  is a graph of order  $n$  and  $a_k(G)$  denotes the number of connected induced subgraphs of order  $k$  in  $G$ , for each  $k \in \{1, \dots, n\}$ , then the *connected induced subgraph polynomial* of  $G$  is given by

$$\Phi_G(x) = \sum_{k=1}^n a_k(G)x^k.$$

The mean order of the connected induced subgraphs of  $G$ , called the *global mean* of  $G$ , is given by the logarithmic derivative

$$M_G = \frac{\Phi'_G(1)}{\Phi_G(1)}.$$

We note that the polynomial  $\Phi_G(x)$  is related to the *residual node connectedness reliability* (or *node reliability*) of  $G$ , first studied in [10]. For  $p \in (0, 1)$  the node reliability of  $G$  is given by

$$R_G(p) = \sum_{k=1}^n a_k(G)p^k(1-p)^{n-k}.$$

It is readily seen that

$$\Phi_G(x) = (1+x)^n R_G\left(\frac{x}{1+x}\right).$$

Given that each vertex of  $G$  operates independently with probability  $p$ , the node reliability  $R_G(p)$  is the probability that the operational vertices induce a connected subgraph. In particular, it was shown in [10] that the problem of counting the total number of connected induced subgraphs of  $G$  (that is, finding  $\Phi_G(1)$ ) is #P-complete, even for split graphs, and for planar bipartite graphs. It follows that computing either  $R_G(p)$  or  $\Phi_G(x)$  is NP-hard in general. However, polynomial algorithms have been found for certain restricted families [3], including trees, series-parallel graphs, and permutation graphs. Several authors have considered optimality questions for node reliability [5, 8, 14], while others have investigated the (complex) roots of the polynomial [1] and the shape on the interval  $(0, 1)$  [2].

Let  $v$  be a vertex of  $G$  and, for each  $k \in \{1, \dots, n\}$ , let  $a_k(G; v)$  denote the number of connected induced subgraphs of  $G$  containing  $v$  and having order  $k$ . Then the generating polynomial for the connected induced subgraphs of  $G$  containing  $v$ , called the *local connected induced subgraph polynomial* of  $G$  at  $v$ , is given by

$$\Phi_{G,v}(x) = \sum_{k=1}^n a_k(G; v)x^k.$$

The mean order of the connected induced subgraphs containing  $v$ , called the *local mean* of  $G$  at  $v$ , is given by the logarithmic derivative  $M_{G,v} = \Phi'_{G,v}(1)/\Phi_{G,v}(1)$ .

The *density* of a graph  $G$  of order  $n$  is defined as  $M_G/n$ . The density of a graph  $G$  is the probability that a randomly chosen vertex belongs to a randomly chosen connected induced subgraph of  $G$ . This notion of density is not to be confused with the edge-density of a graph.

This paper focuses on the mean order of connected induced subgraphs of ‘cographs’. A graph is a *cograph* if it contains no induced path on four vertices. Thus a graph is a cograph if and only if its complement is a cograph. The following useful characterization of cographs can be gleaned from Theorem 2 in [4] and the discussion preceding this result. For graphs  $G_1$  and  $G_2$ , the union of  $G_1$  and  $G_2$  is denoted by  $G_1 \cup G_2$  while the join is denoted by  $G_1 + G_2$ . The complement of a graph  $G$  is denoted by  $\overline{G}$ .

**Result 1.1.** *Let  $G$  be a graph. The following are equivalent:*

- (a)  $G$  is a cograph.
- (b)  $G \cong K_1$ , or there exist two cographs  $G_1$  and  $G_2$  such that either  $G = G_1 \cup G_2$ , or  $G = G_1 + G_2$ .
- (c) For each  $S \subseteq V(G)$  with  $|S| \geq 2$ ,  $S$  induces a connected subgraph of  $G$  if and only if  $S$  induces a disconnected subgraph of  $\overline{G}$  (i.e. each nontrivial vertex set induces a connected graph in exactly one of  $G$  or  $\overline{G}$ ).

This result implies that  $G$  is a connected cograph if and only if  $G \cong K_1$  or  $G$  is the join of two cographs. Moreover,  $G$  is a nontrivial connected cograph if and only if its complement is a disconnected cograph.

We also make frequent use of the following straightforward result from [6] on logarithmic derivatives. Essentially, if we express  $\Phi_G(x)$  as a sum of polynomials with nonnegative coefficients, then  $M_G$  is a weighted average of the logarithmic derivatives (means) of the polynomials in the sum. Here and throughout this article, we use the convention that the logarithmic derivative of the zero polynomial is 0.

**Result 1.2.** *Let  $\Phi_1, \dots, \Phi_k$  be polynomials with nonnegative coefficients. Then the logarithmic derivative of  $\Phi_1 + \dots + \Phi_k$  is a convex combination of the logarithmic derivatives of the  $\Phi_i$ , i.e.*

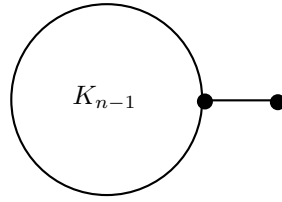
$$\frac{\Phi'_1(1) + \dots + \Phi'_k(1)}{\Phi_1(1) + \dots + \Phi_k(1)} = \frac{\Phi'_1(1)}{\Phi_1(1)}c_1 + \dots + \frac{\Phi'_k(1)}{\Phi_k(1)}c_k$$

where  $c_1 + \dots + c_k = 1$ , and  $c_i \geq 0$  for all  $i \in \{1, \dots, k\}$ .

In several proofs throughout the article, we make use of a variant of the global mean. Let  $G$  be a graph of order  $n$ , and let

$$\Phi_G^*(x) = \Phi_G(x) - nx.$$

The logarithmic derivative of  $\Phi_G^*$ , denoted  $M_G^*$ , is the mean order of the nontrivial (that is, order at least 2) connected induced subgraphs of  $G$ . By our convention that

Figure 1: An  $n$ -skillet

the logarithmic derivative of a zero polynomial is 0, if  $G$  has no edges, then  $M_G^* = 0$ . The value  $M_G^*$  is called the  $M^*$  mean of  $G$ .

The layout of this article is as follows. In Section 2, we show that among all connected cographs of order  $n \geq 7$ , the star  $K_{1,n-1}$  has maximum mean connected induced subgraph order. In Section 3, we show that the local mean at any vertex of a connected cograph is at least as large as the global mean. In Section 4, we show that among all connected cographs of order  $n \geq 3$ , the  $n$ -skillet  $S_n = K_1 + (K_1 \cup K_{n-2})$  (see Fig. 1) has minimum mean connected induced subgraph order. Finally, in Section 5, we find the extremal structures for the mean connected induced subgraph order of disconnected cographs of order  $n$ . We conclude by observing that the density of connected cographs is asymptotically  $1/2$ .

## 2 Connected cographs with maximum mean

We show in this section that among all connected cographs of order  $n \geq 7$ , the star  $K_{1,n-1}$  has maximum global mean. First note that

$$\Phi_{K_{1,n-1}}(x) = x(1+x)^{n-1} + (n-1)x,$$

so it is easy to verify that

$$M_{K_{1,n-1}} = \frac{(n+1) \cdot 2^{n-2} + n-1}{2^{n-1} + n-1} = \frac{n+1}{2} - \frac{(n-1)^2}{2(2^{n-1} + n-1)}.$$

The fact that this mean is maximum among all connected cographs of order  $n \geq 7$  follows from two intermediate results which we prove in this section:

- (a) Among all complete bipartite graphs  $K_{s,n-s}$ , where  $1 \leq s \leq n-1$ , the star  $K_{1,n-1}$  has largest mean whenever  $n \geq 7$ .
- (b) If  $G = G_1 + G_2$ , where  $|V(G_1)| = s$  and  $|V(G_2)| = n-s$  and  $G_1$  and  $G_2$  are cographs, then  $M_G \leq M_{K_{s,n-s}}$ .

While both results rely on fairly technical arguments, we briefly provide some intuition for (b). The idea is that we can obtain the connected cograph  $G = G_1 + G_2$  from the complete bipartite graph  $K_{s,n-s}$  simply by adding edges in the bipartition sets. The only new connected induced subgraphs that are introduced by this process

lie completely inside of either  $G_1$  or  $G_2$ . Since  $G_1$  and  $G_2$  are cographs, we can show that the average order of these new connected induced subgraphs is smaller than the average order of the connected induced subgraphs of  $K_{s,n-s}$  by an inductive argument. So intuitively, the connected cograph  $G_1 + G_2$  will have largest mean if  $G_1$  and  $G_2$  have no edges.

We begin with some useful observations. Let  $G = G_1 + G_2$ , where  $G_1$  and  $G_2$  are graphs of orders  $s$  and  $n - s$ , respectively. Then the connected induced subgraph polynomial of  $G$  is given by

$$\Phi_G(x) = \Phi_{G_1}(x) + \Phi_{G_2}(x) + \Phi_{K_{s,n-s}}^*(x),$$

as the connected induced subgraphs of  $G$  lie completely in either  $G_1$  or  $G_2$ , or have at least one vertex from each graph (and every vertex of  $G_1$  is joined to every vertex of  $G_2$ ). The polynomial  $\Phi_{K_{s,n-s}}^*(x)$  plays an important role in much of our work. By a basic counting argument, we have

$$\begin{aligned} \Phi_{K_{s,n-s}}^*(x) &= ((1+x)^s - 1)((1+x)^{n-s} - 1) \\ &= (1+x)^n - (1+x)^s - (1+x)^{n-s} + 1. \end{aligned}$$

For ease of notation, we let  $\psi_{s,n-s}(x) = \Phi_{K_{s,n-s}}^*(x)$ , so that

$$\psi_{s,n-s}(1) = 2^n - 2^{n-s} - 2^s + 1, \tag{1}$$

$$\psi'_{s,n-s}(1) = n2^{n-1} - s2^{s-1} - (n-s)2^{n-s-1}, \tag{2}$$

and

$$M_{K_{s,n-s}}^* = \frac{\psi'_{s,n-s}(1)}{\psi_{s,n-s}(1)} = \frac{n2^{n-1} - s2^{s-1} - (n-s)2^{n-s-1}}{2^n - 2^{n-s} - 2^s + 1}.$$

In order to show that  $K_{1,n-1}$  has largest global mean among all complete bipartite graphs of order  $n \geq 7$ , we first show that  $K_{1,n-1}$  has largest  $M^*$  mean among all complete bipartite graphs of order  $n \geq 1$ . The result is trivial for  $n \leq 3$ , and for  $n \geq 4$  we prove the stronger result that  $M_{K_{s,n-s}}^*$  is decreasing in  $s$  for  $s \in \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$ . This stronger result is used to prove several of our later results, so we include it in the Lemma statement below.

**Lemma 2.1.** *For  $n \geq 4$  and  $s \in \{1, \dots, \lfloor \frac{n}{2} \rfloor - 1\}$ ,*

$$M_{K_{s,n-s}}^* > M_{K_{s+1,n-s-1}}^*.$$

*In particular,  $K_{1,n-1}$  uniquely maximizes the  $M^*$  mean among all complete bipartite graphs of order  $n$  for all  $n \geq 1$ .*

*Proof.* First note that  $s \leq \lfloor \frac{n}{2} \rfloor - 1$  is equivalent to  $n \geq 2s + 2$ . We use this inequality throughout the proof. We have

$$\begin{aligned} M_{K_{s,n-s}}^* - M_{K_{s+1,n-s-1}}^* &= \frac{\psi'_{s,n-s}(1)}{\psi_{s,n-s}(1)} - \frac{\psi'_{s+1,n-s-1}(1)}{\psi_{s+1,n-s-1}(1)} \\ &= \frac{\psi'_{s,n-s}(1)\psi_{s+1,n-s-1}(1) - \psi'_{s+1,n-s-1}(1)\psi_{s,n-s}(1)}{\psi_{s,n-s}(1)\psi_{s+1,n-s-1}(1)}. \end{aligned} \tag{3}$$

It suffices to show that the numerator of the expression in (3) is positive for  $1 \leq s \leq \lfloor \frac{n}{2} \rfloor - 1$ . Substituting the appropriate expressions from (1) and (2) into the numerator of (3), and then simplifying, yields

$$2^{2n-2s-2} - (n - s + 1) \cdot 2^{n-s-2} + (s - 1) \cdot 2^{2n-s-2} - (n - s - 2) \cdot 2^{n+s-1} + 3(n - 2s - 1) \cdot 2^{n-2} - 2^{2s} + (s + 2) \cdot 2^{s-1}. \tag{4}$$

The proof that (4) is strictly positive requires two cases:

**Case 1:**  $s = 1$ . Evaluating (4) at  $s = 1$  and simplifying gives

$$2^{2n-4} - (3n - 6)2^{n-3} - 1,$$

which is easily verified to be positive for  $n \geq 4$  by induction.

**Case 2:**  $2 \leq s \leq \lfloor \frac{n}{2} \rfloor - 1$ . Since  $n \geq 2s + 2$ , we may assume that  $n \geq 6$ . Grouping the terms of (4), we claim that

$$2^{2n-2s-2} - (n - s + 1) \cdot 2^{n-s-2} > 0, \tag{5}$$

$$(s - 1) \cdot 2^{2n-s-2} - (n - s - 2) \cdot 2^{n+s-1} \geq 0, \text{ and} \tag{6}$$

$$3(n - 2s - 1) \cdot 2^{n-2} - 2^{2s} \geq 2^{n-1}, \tag{7}$$

which together imply that (4) is strictly positive. While we only need to show that  $3(n - 2s - 1) \cdot 2^{n-2} - 2^{2s}$  is nonnegative here, the stronger inequality (7) is used in the proof of our next result.

For (5), we factor the left-hand side to obtain

$$2^{n-s-2} [2^{n-s} - (n - s + 1)],$$

so it suffices to show that  $2^{n-s} - (n - s + 1) > 0$ . Let  $k = n - s$ , and note that  $k \geq s + 2 \geq 4$ . The inequality  $2^k - (k + 1) > 0$  is easily proven by induction, completing the proof of (5).

For (6), factoring the left-hand side gives

$$2^{n+s-1} [(s - 1) \cdot 2^{n-2s-1} - (n - s - 2)].$$

So it suffices to show that  $(s - 1) \cdot 2^{n-2s-1} - (n - s - 2) \geq 0$ . For each fixed value of  $s \geq 2$ , we prove this inequality by induction on  $n$ . For the base case, we substitute  $n = 2s + 2$ , and we find

$$(s - 1) \cdot 2 - s = s - 2 \geq 0.$$

Now suppose that the inequality holds for some natural number  $n \geq 2s + 2$ . Then

$$\begin{aligned} (s - 1) \cdot 2^{n-2s} - (n - s - 1) &= (s - 1) \cdot 2^{n-2s-1} - (n - s - 2) + (s - 1) \cdot 2^{n-2s-1} - 1 \\ &\geq (s - 1) \cdot 2^{n-2s-1} - 1 \\ &> 0, \end{aligned}$$

completing the proof of (6).

Finally, for (7), we use the inequality  $n \geq 2s + 2$  twice:

$$3(n - 2s - 1) \cdot 2^{n-2} - 2^{2s} \geq 3 \cdot 2^{n-2} - 2^{2s} \geq 3 \cdot 2^{n-2} - 2^{n-2} = 2^{n-1}. \quad \square$$

We now prove the first of the two statements from which our main result will follow.

**Theorem 2.2.** For  $n \geq 7$ ,

$$M_{K_{1,n-1}} > M_{K_{2,n-2}},$$

and for  $n \geq 6$  and  $s \in \{2, \dots, \lfloor \frac{n}{2} \rfloor - 1\}$ ,

$$M_{K_{s,n-s}} > M_{K_{s+1,n-s-1}}.$$

In particular, among all complete bipartite graphs of order  $n \geq 7$ , the star  $K_{1,n-1}$  uniquely maximizes the global mean.

*Proof.* For any  $n$  and any  $s \in \{1, \dots, n - 1\}$ , we have

$$\Phi_{K_{s,n-s}} = nx + \psi_{s,n-s}(x),$$

so that

$$M_{K_{s,n-s}} = \frac{n + \psi'_{s,n-s}(1)}{n + \psi_{s,n-s}(1)}$$

Hence, for any  $s \in \{1, \dots, \lfloor \frac{n}{2} \rfloor - 1\}$ ,

$$\begin{aligned} &M_{K_{s,n-s}} - M_{K_{s+1,n-s-1}} \\ &= \frac{[n + \psi'_{s,n-s}(1)][n + \psi_{s+1,n-s-1}(1)] - [n + \psi_{s,n-s}(1)][n + \psi'_{s+1,n-s-1}(1)]}{[n + \psi_{s,n-s}(1)][n + \psi_{s+1,n-s-1}(1)]}. \end{aligned}$$

Expanding the numerator gives

$$\begin{aligned} &\psi'_{s,n-s}(1)\psi_{s+1,n-s-1}(1) - \psi_{s,n-s}(1)\psi'_{s+1,n-s-1}(1) \\ &+ n \left[ \psi'_{s,n-s}(1) + \psi_{s+1,n-s-1}(1) - \psi_{s,n-s}(1) - \psi'_{s+1,n-s-1}(1) \right]. \end{aligned} \tag{8}$$

It suffices to show that this quantity is positive for the given values of  $n$  and  $s$ .

We first demonstrate that (8) is positive when  $s = 1$  and  $n \geq 7$ , proving the first part of the Theorem. When  $s = 1$ , using (1) and (2), we find that (8) simplifies to

$$2^{2n-4} - (n^2 + n - 6) \cdot 2^{n-3} + n - 1,$$

which is easily seen to be positive for  $n \geq 7$  by induction.

Now let  $n \geq 6$  and  $s \in \{2, \dots, \lfloor \frac{n}{2} \rfloor - 1\}$ . An expression for the first line of (8) is given by (4). Applying inequalities (6) and (7), proven in Lemma 2.1, we obtain

$$\begin{aligned} &\psi'_{s,n-s}(1)\psi_{s+1,n-s-1}(1) - \psi_{s,n-s}(1)\psi'_{s+1,n-s-1}(1) \\ &\geq 2^{2n-2s-2} - (n - s + 1) \cdot 2^{n-s-2} + 2^{n-1}. \end{aligned}$$

For the expression on the second line of (8), employing (1) and (2) and then simplifying, we find

$$\begin{aligned} \psi'_{s,n-s}(1) + \psi_{s+1,n-s-1}(1) - \psi_{s,n-s}(1) - \psi'_{s+1,n-s-1}(1) &= s \cdot 2^{s-1} - (n-s-1) \cdot 2^{n-s-2} \\ &> -(n-s-1) \cdot 2^{n-s-2}. \end{aligned}$$

It follows that (8) is greater than

$$\begin{aligned} &2^{2n-2s-2} - (n-s+1) \cdot 2^{n-s-2} + 2^{n-1} - n(n-s-1) \cdot 2^{n-s-2} \\ &= 2^{n-s-2} [2^{n-s} - (n-s+1) + 2^{s+1} - n(n-s-1)] \\ &= 2^{n-s-2} [2^{n-s} + 2^{s+1} - (n^2 - ns - s + 1)]. \end{aligned}$$

So it suffices to show that

$$2^{n-s} + 2^{s+1} - (n^2 - ns - s + 1) > 0. \tag{9}$$

For each fixed value of  $s \geq 2$ , we proceed by induction on  $n$ . Since  $n \geq 2s + 2$ , we begin with the base case  $n = 2s + 2$ . In this case, the left-hand side of (9) evaluates to

$$2^{s+2} + 2^{s+1} - (2s^2 + 5s + 5) = 3 \cdot 2^{s+1} - (2s^2 + 5s + 5),$$

which is positive for  $s \geq 2$  by induction. Now suppose that (9) holds for some  $n \geq 2s + 2$ . Then

$$\begin{aligned} &2^{n+1-s} + 2^{s+1} - [(n+1)^2 - (n+1)s - s + 1] \\ &= 2^{n-s} + [2^{n-s} + 2^{s+1} - (n^2 - ns - s + 1)] - (2n+1) + s \\ &> 2^{n-s} - (2n+1) + s. \end{aligned}$$

So it suffices to show that  $2^{n-s} - (2n+1) + s \geq 0$ . We omit the details, but for each fixed  $s \geq 2$ , the inequality  $2^{n-s} - (2n+1) > 0$  can be shown to hold for  $n \geq 2s + 2$  by induction. This completes the proof.  $\square$

To prove the second of our main two statements the following three lemmas on  $M^*$  means will be useful.

**Lemma 2.3.**  $M^*_{K_{1,n-1}}$  is strictly increasing for  $n \geq 1$ . Moreover,  $\frac{n+1}{2} < M^*_{K_{1,n-1}} \leq \frac{n+2}{2}$  for all  $n \geq 2$ .

*Proof.* First of all, note that  $M^*_{K_{1,1}} = 2 > 0 = M^*_{K_{1,0}}$ . Now, for  $n \geq 2$ , we have

$$M^*_{K_{1,n-1}} = \frac{2^{n-2}(n+1) - 1}{2^{n-1} - 1} = \frac{n+1}{2} + \frac{n-1}{2(2^{n-1} - 1)}.$$

So it suffices to show that  $0 < \frac{n-1}{2(2^{n-1}-1)} \leq \frac{1}{2}$  for  $n \geq 2$ . The first inequality is obvious, while the second holds by a straightforward induction.  $\square$

**Lemma 2.4.** Let  $G$  be a cograph (not necessarily connected) of order  $n \geq 1$ . Then  $M^*_G \leq M^*_{K_{1,n-1}}$ , with equality if and only if  $G \cong K_{1,n-1}$ .



*Proof.* We proceed by induction on  $n$ . The statement is trivial for  $n = 1$ . Suppose now that  $n \geq 2$  and that the statement holds for every integer  $k$ ,  $1 \leq k < n$ . Since  $G$  is a cograph, it is either the join of two cographs or the union of two cographs.

**Case 1:**  $G = G_1 + G_2$ , where  $G_1$  and  $G_2$  are cographs of orders  $s$  and  $n - s$ , respectively.

In this case,

$$\Phi_G(x) = [\Phi_{G_1}(x) - sx] + [\Phi_{G_2}(x) - (n - s)x] + \Phi_{K_{s,n-s}}(x),$$

and hence

$$\Phi_G(x) - nx = [\Phi_{G_1}(x) - sx] + [\Phi_{G_2}(x) - (n - s)x] + [\Phi_{K_{s,n-s}}(x) - nx].$$

Thus, by taking the logarithmic derivative of both sides, we obtain the following expression:

$$M_G^* = c_1 M_{G_1}^* + c_2 M_{G_2}^* + c_3 M_{K_{s,n-s}}^*, \tag{10}$$

for real numbers  $c_1, c_2, c_3 \geq 0$  such that  $c_1 + c_2 + c_3 = 1$ . By the induction hypothesis and then Lemma 2.3,

$$M_{G_1}^* \leq M_{K_{1,s-1}}^* < M_{K_{1,n-1}}^* \quad \text{and} \quad M_{G_2}^* \leq M_{K_{1,n-s-1}}^* < M_{K_{1,n-1}}^*,$$

and by Lemma 2.1,

$$M_{K_{s,n-s}}^* \leq M_{K_{1,n-1}}^*.$$

Moreover, this inequality is strict unless  $s = 1$  or  $s = n - 1$ . Therefore, since each of the means on the right-hand side of (10) is at most  $M_{K_{1,n-1}}^*$ , we must have

$$M_G^* \leq M_{K_{1,n-1}}^*.$$

Since  $M_{G_1}^* < M_{K_{1,n-1}}^*$  and  $M_{G_2}^* < M_{K_{1,n-1}}^*$ , we have equality if and only if  $c_1 = c_2 = 0$  and  $s = 1$  or  $s = n - 1$ , or equivalently,  $G \cong K_{1,n-1}$ .

**Case 2:**  $G = G_1 \cup G_2$ , where  $G_1$  and  $G_2$  are cographs of orders  $s$  and  $n - s$ , respectively.

In this case,

$$\Phi_G(x) - nx = [\Phi_{G_1}(x) - sx] + [\Phi_{G_2}(x) - (n - s)x].$$

Hence

$$M_G^* = c_1 M_{G_1}^* + c_2 M_{G_2}^*,$$

for real numbers  $c_1, c_2 \geq 0$  such that  $c_1 + c_2 = 1$ . By the induction hypothesis and Lemma 2.3,

$$M_G^* \leq c_1 M_{K_{1,s-1}}^* + c_2 M_{K_{1,n-s-1}}^* < c_1 M_{K_{1,n-1}}^* + c_2 M_{K_{1,n-1}}^* = M_{K_{1,n-1}}^*. \quad \square$$

**Lemma 2.5.** *Let  $n \geq 6$  and  $s \in \{1, \dots, n - 1\}$ .*

- (a) If  $G$  is a cograph on  $k \leq n - 2$  vertices, then  $M_G^* < M_{K_{s,n-s}}$ .
- (b) If  $G$  is a cograph on  $n - 1$  vertices, then  $M_G^* < M_{K_{1,n-1}}$ .

*Proof.* Part (a): Without loss of generality we may assume that  $s \leq \lfloor \frac{n}{2} \rfloor$ , or equivalently  $n \geq 2s$ . By Lemma 2.4,  $M_G^* \leq M_{K_{1,k-1}}^*$ . So it suffices to show that  $M_{K_{1,k-1}}^* < M_{K_{s,n-s}}$ . From Lemma 2.3, we have

$$M_{K_{1,k-1}}^* \leq \frac{k+2}{2} \leq \frac{n}{2}.$$

So it suffices to show that  $M_{K_{s,n-s}} > \frac{n}{2}$ . Indeed, we have

$$\begin{aligned} M_{K_{s,n-s}} - \frac{n}{2} &= \frac{n2^{n-1} - s2^{s-1} - (n-s)2^{n-s-1} + n}{2^n - 2^s - 2^{n-s} + 1 + n} - \frac{n}{2} \\ &= \frac{(n-s)2^s + s2^{n-s} + n - n^2}{2[2^n - 2^s - 2^{n-s} + 1 + n]}. \end{aligned} \tag{11}$$

One can show, for fixed  $s \geq 1$  and by induction on  $n \geq 2s$ , that the numerator in (11) is positive.

Part (b): By Lemma 2.4,  $M_G^* \leq M_{K_{1,n-2}}^*$ . So it suffices to show that  $M_{K_{1,n-2}}^* < M_{K_{1,n-1}}$ . We have

$$M_{K_{1,n-1}} - M_{K_{1,n-2}}^* = \frac{(n+1)2^{n-2} + (n-1)}{2^{n-1} + (n-1)} - \frac{n2^{n-3} - 1}{2^{n-2} - 1}.$$

The numerator of this expression over common denominator  $[2^{n-1} + (n-1)][2^{n-2} - 1]$  is given by

$$\begin{aligned} &(n+1)2^{2n-4} + (n-1)2^{n-2} - (n+1)2^{n-2} - (n-1) \\ &\quad - [n2^{2n-4} + n(n-1)2^{n-3} - 2^{n-1} - (n-1)] \\ &= 2^{2n-4} + (n-1)2^{n-2} + 2^{n-1} - (n+1)2^{n-2} - n(n-1)2^{n-3} \\ &= 2^{n-3} [2^{n-1} - n(n-1)], \end{aligned}$$

and  $2^{n-1} > n(n-1)$  holds for  $n \geq 6$  by induction. □

We now prove the second of the two main intermediate results of this section.

**Theorem 2.6.** *Let  $G$  be a connected cograph of order  $n \geq 6$ , obtained from the join of two cographs  $G_1$  and  $G_2$  of orders  $s$  and  $n - s$ , respectively. Then  $M_G \leq M_{K_{s,n-s}}$ , with equality if and only if  $G \cong K_{s,n-s}$ .*

*Proof.* Since  $G = G_1 + G_2$ , the generating polynomial for the connected induced subgraphs of  $G$  is given by

$$\Phi_G(x) = [\Phi_{G_1}(x) - sx] + [\Phi_{G_2}(x) - (n-s)x] + \Phi_{K_{s,n-s}}(x).$$

Thus,

$$M_G = c_1 M_{G_1}^* + c_2 M_{G_2}^* + c_3 M_{K_{s,n-s}}$$

for some real numbers  $c_1, c_2, c_3 \geq 0$  such that  $c_1 + c_2 + c_3 = 1$ . By Lemma 2.5,  $M_{G_1}^* < M_{K_{s,n-s}}$  and  $M_{G_2}^* < M_{K_{s,n-s}}$ . Therefore,  $M_G$  is a weighted average of means of which  $M_{K_{s,n-s}}$  is the largest, from which we conclude that  $M_G \leq M_{K_{s,n-s}}$ . The inequality is strict unless  $c_1 = c_2 = 0$ , i.e.  $G \cong K_{s,n-s}$ .  $\square$

Our main result now follows easily.

**Corollary 2.7.** *If  $G$  is a connected cograph of order  $n \geq 7$ , then*

$$M_G \leq \frac{n+1}{2} - \frac{(n-1)^2}{2(2^{n-1} + n - 1)},$$

*with equality if and only if  $G \cong K_{1,n-1}$ . Generally, if  $G$  is a connected cograph of order  $n \geq 1$ , then  $M_G \leq \frac{n+1}{2}$ , with equality if and only if  $n = 1$ .*

*Proof.* The first statement follows immediately from Theorems 2.2 and 2.6. For  $n \geq 7$ , the second statement follows easily from the first. For  $n \leq 6$ , we check the second statement exhaustively using a computer algebra system. The cographs of maximum global mean for  $n \leq 6$  and their means are given in Table 1. Decimal expansions are provided for ease of verifying the second statement.  $\square$

Order	Cograph	Global Mean	Decimal Expansion
1	$K_1$	1	1
2	$K_2$	4/3	1.33...
3	$K_3$	12/7	1.71...
4	$K_{2,2}$	28/13	2.15...
5	$K_{2,3}$	69/26	2.65...
6	$K_{2,4}$	54/17	3.18...

Table 1: The cographs of maximum global mean among all connected cographs of order  $n \leq 6$ .

Finally, we observe that our upper bound on the mean connected induced subgraph order of connected cographs generalizes nicely to all cographs, as described below.

**Corollary 2.8.** *Let  $G$  be a cograph and let  $s$  be the order of a largest component of  $G$ . Then  $M_G \leq \frac{s+1}{2}$  with equality if and only if  $s = 1$ .*

*Proof.* Suppose  $G$  has components  $G_1, \dots, G_k$ . Since  $\Phi_G(x) = \sum_{i=1}^k \Phi_{G_i}(x)$ , the mean  $M_G$  is a convex combination of the  $M_{G_i}$ . So the result follows immediately from Corollary 2.7.  $\square$

### 3 Comparing local and global means

Recall that if  $v$  is a vertex of a graph  $G$ , then  $M_{G,v}$  denotes the mean order of the connected induced subgraphs of  $G$  that contain  $v$ . It was shown in [6], for any tree  $T$  and any vertex  $v$  of  $T$ , that

$$M_{T,v} \geq M_T,$$

with equality if and only if  $T \cong K_1$ . It is natural to ask whether this property extends to the mean connected induced subgraph order of graphs in general. We provide an infinite family of graphs with a vertex for which this inequality does not hold.

For each natural number  $n$ , let  $\mathcal{H}_n$  be the family of graphs of order  $n$  with mean greater than  $\frac{n+2}{2}$ . From [9], there is a tree (in fact, a caterpillar)  $T_n$  of order  $n$  such that  $M_{T_n} \geq n - \lceil 2 \log_2 n \rceil - 1$  for all  $n$ . It follows that  $\mathcal{H}_n$  is nonempty for all  $n \geq 25$ , and further, there is a tree in  $\mathcal{H}_n$  for all  $13 \leq n \leq 24$  by computational work. Define

$$\mathcal{G}_n = \{H + K_1 : H \in \mathcal{H}_{n-1}\},$$

and note that  $\mathcal{G}_n$  is nonempty for all  $n \geq 14$ .

Let  $G \in \mathcal{G}_n$  for some  $n \geq 14$ . In particular, let  $G = H + v$ , where  $H \in \mathcal{H}_{n-1}$ . Then

$$\Phi_G(x) = \Phi_H(x) + \Phi_{G,v}(x). \tag{12}$$

Since  $v$  is universal in  $G$ , we see that  $\Phi_G(x) = x(1+x)^{n-1}$ , and thus  $M_{G,v} = \frac{n+1}{2}$ . Since  $H \in \mathcal{H}_{n-1}$ , we have  $M_H > \frac{(n-1)+2}{2} = \frac{n+1}{2} = M_{G,v}$ . We conclude that  $M_H > M_G > M_{G,v}$ , since  $M_G$  is a convex combination of  $M_H$  and  $M_{G,v}$  by (12). Therefore, for each  $n \geq 14$ ,  $\mathcal{G}_n$  is nonempty and every member of  $\mathcal{G}_n$  has a vertex at which the local mean is less than the global mean.

Using results from the previous section, we now demonstrate that  $M_{G,v} \geq M_G$  for any connected *cograph*  $G$ , and any vertex  $v \in V(G)$ . Aside from being interesting in its own right, this result will be used in the next section when we determine the structure of connected cographs with minimum global mean.

**Theorem 3.1.** *Let  $G$  be a connected cograph of order  $n \geq 1$  and  $v$  any vertex of  $G$ . Then*

$$M_{G,v} \geq \frac{n+1}{2}.$$

*Proof.* We proceed by induction on  $n$ . The result holds for  $n = 1$ . Suppose  $n \geq 2$  and that the result holds for every connected cograph of order less than  $n$ , and let  $G$  be a connected cograph of order  $n$ . Since  $G$  is a connected cograph,  $G = G_1 + G_2$ , where  $G_1$  is a cograph of order  $s$  (say), and  $G_2$  is a cograph of order  $n - s$ . Without loss of generality, suppose that  $v \in V(G_1)$ . More specifically, since  $G_1$  is not necessarily connected,  $v$  belongs to a component  $H$  of  $G_1$ . Let  $k$  be the order of  $H$  (note that  $1 \leq k \leq s$ ).

Now the connected induced subgraphs of  $G$  containing  $v$  can be partitioned into those containing no vertex of  $G_2$ , and those containing at least one vertex of  $G_2$ .

Hence

$$\Phi_{G,v}(x) = \Phi_{G_1,v}(x) + x \left[ \sum_{i=1}^{n-s} \binom{n-s}{i} x^i \right] \left[ \sum_{j=0}^{s-1} \binom{s-1}{j} x^j \right].$$

Thus,

$$\begin{aligned} \Phi_{G,v}(x) &= \Phi_{G_1,v}(x) + x [(1+x)^{n-s} - 1] (1+x)^{s-1} \\ &= \Phi_{G_1,v}(x) + x [(1+x)^{n-1} - (1+x)^{s-1}], \end{aligned}$$

and

$$\begin{aligned} \Phi'_{G,v}(x) &= \Phi'_{G_1,v}(x) + [(1+x)^{n-1} - (1+x)^{s-1}] \\ &\quad + x [(n-1)(1+x)^{n-2} - (s-1)(1+x)^{s-2}]. \end{aligned}$$

Therefore,

$$M_{G,v} = \frac{\Phi'_{G,v}(1)}{\Phi_{G,v}(1)} = \frac{\Phi'_{G_1,v}(1) + 2^{n-2}(n+1) - 2^{s-2}(s+1)}{\Phi_{G_1,v}(1) + 2^{n-1} - 2^{s-1}}.$$

Straightforward algebra gives

$$M_{G,v} - \frac{n+1}{2} = \frac{2^{s-2}(n-s) + \Phi_{G_1,v}(1) [M_{G_1,v} - (\frac{n+1}{2})]}{\Phi_{G_1,v}(1) + 2^{n-1} - 2^{s-1}}.$$

So it suffices to show that

$$2^{s-2}(n-s) + \Phi_{G_1,v}(1) [M_{G_1,v} - \frac{n+1}{2}] \geq 0.$$

By the induction hypothesis,

$$\begin{aligned} 2^{s-2}(n-s) + \Phi_{G_1,v}(1) [M_{G_1,v} - \frac{n+1}{2}] &\geq 2^{s-2}(n-s) + \Phi_{G_1,v}(1) [\frac{k+1}{2} - \frac{n+1}{2}] \\ &= 2^{s-2}(n-s) - \Phi_{G_1,v}(1) (\frac{n-k}{2}). \end{aligned}$$

Moreover,  $\Phi_{G_1,v}(1) \leq 2^{k-1}$  with equality if and only if  $v$  is a universal vertex of  $H$ , so we have

$$\begin{aligned} 2^{s-2}(n-s) - \Phi_{G_1,v}(1) \left(\frac{n-k}{2}\right) &\geq 2^{s-2}(n-s) - 2^{k-1} \left(\frac{n-k}{2}\right) \\ &= 2^{s-2}(n-s) - 2^{k-2}(n-k) \\ &= 2^{k-2} [(n-s)2^{s-k} - n+k]. \end{aligned}$$

Now it suffices to show that  $(n-s)2^{s-k} - n+k \geq 0$ . Regrouping and applying the inequality  $n-s \geq 1$ , we obtain

$$(n-s)2^{s-k} - n+k = (n-s)(2^{s-k} - 1) + n-s-n+k \geq 2^{s-k} - 1 - s+k = 2^{s-k} - (s-k) - 1.$$

The result follows by letting  $m = s - k$ , and observing that, by induction on  $m \geq 0$ , we have  $2^m - m - 1 \geq 0$ . □

The fact that the local mean at any vertex of a connected cograph is at least as large as the global mean now follows directly from Corollary 2.7 and Theorem 3.1.

**Corollary 3.2.** *If  $G$  is a connected cograph with vertex  $v$ , then  $M_{G,v} \geq M_G$ , with equality if and only if  $G \cong K_1$ .*

*Proof.* Let  $G$  be a connected cograph of order  $n$ , and  $v \in V(G)$ . By Corollary 2.7 and then Theorem 3.1, we have

$$M_G \leq \frac{n+1}{2} \leq M_{G,v},$$

with equality at the first inequality if and only if  $n = 1$ . □

## 4 Connected cographs with minimum mean

In this section we determine the structure of those connected cographs for which the global mean is a minimum. Recall that we can write any connected cograph  $G$  as the join of two cographs  $G_1$  and  $G_2$ . If  $|V(G_1)| = s$  and  $|V(G_2)| = n - s$ , then we can obtain  $G$  from  $K_{s,n-s}$  by adding edges in the bipartition sets. As mentioned previously, we expect that the new connected induced subgraphs created by adding the edges of  $G_1$  and  $G_2$  will have smaller order on average than those of  $K_{s,n-s}$ . Thus, to minimize the mean connected induced subgraph order of  $G_1 + G_2$ , it stands to reason that  $G_1$  and  $G_2$  should have many edges (so that they have many connected induced subgraphs). Surprisingly, we will show that the extremal structures for the minimum global mean are not complete graphs. In order to minimize the global mean of  $G_1 + G_2$ , we want  $G_1$  and  $G_2$  to have many connected induced subgraphs, but we also want them to have small mean order.

Recall that the  $n$ -skillet, denoted by  $S_n$ , is the cograph defined by  $S_n = K_1 + (K_1 \cup K_{n-2})$ , pictured in Fig. 1. We first demonstrate by direct proof that if  $G$  is a connected cograph of order  $n$  with a cut vertex, then  $M_G \geq M_{S_n}$ . Then we consider the general case where  $G$  is any connected cograph of order  $n$ ; the proof is by induction.

We begin by deriving expressions for  $M_{S_n}$  and  $M_{K_n}$  and showing that  $M_{S_n} < M_{K_n}$  for all  $n \geq 3$ .

**Lemma 4.1.** *For any  $n \geq 3$ ,*

- (a)  $M_{S_n} = \frac{n}{2} + \frac{1}{3 \cdot 2^{n-2}}$ ,
- (b)  $M_{K_n} = \frac{n}{2} + \frac{n}{2^{n+1}-2}$ , and
- (c)  $M_{S_n} < M_{K_n}$ .

*Proof.* Parts (a) and (b) are easily obtained from the connected induced subgraph polynomials for  $S_n$  and  $K_n$ , respectively:

$$\begin{aligned}\Phi_{S_n}(x) &= (1+x)^{n-1} - 1 + x + x^2(1+x)^{n-3} \\ \Phi_{K_n}(x) &= (1+x)^n - 1.\end{aligned}$$

For part (c), we show that the difference  $M_{K_n} - M_{S_n}$  is positive for  $n \geq 3$ . We have

$$M_{K_n} - M_{S_n} = \frac{n}{2^{n+1} - 2} - \frac{1}{3 \cdot 2^{n-2}} = \frac{2^{n-2}(3n - 8) + 2}{3 \cdot 2^{n-2} \cdot (2^{n+1} - 2)} > 0$$

where the last inequality follows easily from the fact that  $n \geq 3$ . □

The following three lemmas on  $M^*$  means of cographs will be used to prove the main result of this section.

**Lemma 4.2.**  $M_{K_{s,n-s}}^* > \frac{n}{2}$  for any  $n \geq 2$  and  $1 \leq s \leq n - 1$ .

*Proof.* We have

$$\begin{aligned} M_{K_{s,n-s}}^* - \frac{n}{2} &= \frac{n2^{n-1} - s2^{s-1} - (n-s)2^{n-s-1}}{2^n - 2^s - 2^{n-s} + 1} - \frac{n}{2} \\ &= \frac{(n-s)2^s + s2^{n-s} - n}{2(2^n - 2^s - 2^{n-s} + 1)} \\ &\geq \frac{2(n-s) + 2s - n}{2(2^n - 2^s - 2^{n-s} + 1)} \\ &> 0. \end{aligned}$$

Hence  $M_{K_{s,n-s}}^* > \frac{n}{2}$ . □

**Lemma 4.3.** Let  $G = G_1 + G_2$ , where  $G_1$  and  $G_2$  are cographs of orders  $s$  and  $n - s$  respectively, where  $n \geq 4$  and  $2 \leq s \leq n - 2$ . Then  $M_G^* \leq M_{K_{s,n-s}}^*$ , with equality if and only if  $G \cong K_{s,n-s}$ .

*Proof.* Since  $G$  is a connected cograph, observe that

$$\Phi_G(x) - nx = [\Phi_{G_1}(x) - sx] + [\Phi_{G_2}(x) - (n - s)x] + [\Phi_{K_{s,n-s}}(x) - nx].$$

Hence,

$$M_G^* = c_1M_{G_1}^* + c_2M_{G_2}^* + c_3M_{K_{s,n-s}}^*, \tag{13}$$

for some nonnegative numbers  $c_1, c_2, c_3$ , such that  $c_1 + c_2 + c_3 = 1$ . By Lemma 2.4, and then Lemma 2.3,

$$\begin{aligned} M_{G_1}^* &\leq M_{K_{1,s-1}}^* \leq \frac{s+2}{2} \leq \frac{n}{2}, \text{ and} \\ M_{G_2}^* &\leq M_{K_{1,n-s-1}}^* \leq \frac{n-s+2}{2} \leq \frac{n}{2}. \end{aligned}$$

Finally, by Lemma 4.2,  $M_{K_{s,n-s}}^* > \frac{n}{2}$ , and hence  $M_G^* < M_{K_{s,n-s}}^*$  follows from (13). Moreover, we have equality if and only if  $c_1 = c_2 = 0$ , that is, if and only if  $G \cong K_{s,n-s}$ . □

**Lemma 4.4.**  $M_{K_{2,n-3}}^* \leq M_{K_n}$  for  $n \geq 6$ .

*Proof.* Consider the difference

$$\begin{aligned} M_{K_n} - M_{K_{2,n-3}}^* &= \frac{n2^{n-1}}{2^n - 1} - \frac{2^{n-4}(3n - 1) - 4}{3 \cdot 2^{n-3} - 3} \\ &= \frac{n2^{n-1}[3 \cdot 2^{n-3} - 3] - (2^n - 1)[2^{n-4}(3n - 1) - 4]}{(2^n - 1)(3 \cdot 2^{n-3} - 3)}. \end{aligned}$$

Expanding and regrouping the numerator, we obtain

$$M_{K_n} - M_{K_{2,n-3}}^* = \frac{2^{n-4}(2^n - 21n + 63) - 4}{(2^n - 1)(3 \cdot 2^{n-3} - 3)}.$$

Thus it suffices to show that  $2^n - 21n + 63 \geq 1$  for  $n \geq 6$ , and this follows easily by induction.  $\square$

We now prove the first special case of the main result of this section, namely the case where  $G$  is a connected cograph with a cut vertex. This case is further divided into the subcases where  $G$  does or does not have a leaf. Let  $G$  be a connected cograph of order  $n$  with a cut vertex. Then it is straightforward to show that  $G = K_1 + H$  where  $H$  is a disconnected cograph. We use this structure to show that  $M_G > M_{S_n}$ .

**Lemma 4.5.** *Let  $G$  be a connected cograph of order  $n$ . If  $G$  has a cut vertex, then  $M_G \geq M_{S_n}$ , with equality if and only if  $G \cong S_n$ .*

*Proof.* Write  $G = K_1 + H$ , where  $H$  is a disconnected cograph of order  $n - 1$ . We consider two cases, depending on whether or not  $H$  has an isolated vertex.

Suppose first that  $H$  has no isolated vertex. In this case, we must have  $n \geq 5$ . If  $n = 5$ , then  $H \cong K_2 \cup K_2$ , and  $M_G > M_{S_5}$  can be verified directly. Now assume that  $n \geq 6$ . We will prove the stronger result that  $M_G \geq M_{K_n}$ , from which the desired result follows by Lemma 4.1(c). We can write

$$\Phi_{K_n}(x) = \Phi_G(x) + \Phi_{\overline{G}}(x) - nx, \tag{14}$$

since every nonempty subset of vertices induces a connected subgraph of  $K_n$ , while, by Result 1.1(c), every subset of vertices of size at least 2 induces a connected subgraph in either  $G$  or  $\overline{G}$ , but not in both. The singleton subsets are counted by both  $\Phi_G(x)$  and  $\Phi_{\overline{G}}(x)$ , which is why we subtract  $nx$ . Since  $\overline{G} = K_1 \cup \overline{H}$ , we substitute  $\Phi_{\overline{G}}(x) = \Phi_{\overline{H}}(x) + x$  into (14) to obtain

$$\Phi_{K_n}(x) = \Phi_G(x) + [\Phi_{\overline{H}}(x) - (n - 1)x].$$

Therefore, we can write  $M_{K_n}$  as a convex combination as follows:

$$M_{K_n} = c_1 M_G + c_2 M_{\overline{H}}^*. \tag{15}$$

Since  $\overline{H}$  is a connected cograph of order  $n - 1$ , and  $H$  has no isolated vertex, we must have  $\overline{H} = H_1 + H_2$  with  $|V(H_1)| \geq 2$  and  $|V(H_2)| \geq 2$ . Let  $|V(H_1)| = s$ . By Lemma 4.3,  $M_{\overline{H}}^* \leq M_{K_{s,n-1-s}}^*$ . Further, by Lemma 2.1,  $M_{K_{s,n-1-s}}^* \leq M_{K_{2,n-3}}^*$ . Finally,



by Lemma 4.4,  $M_{K_{2,n-3}}^* \leq M_{K_n}$  for  $n \geq 6$ . So  $M_{\overline{H}}^* \leq M_{K_n}$ , and we conclude from (15) that  $M_G \geq M_{K_n}$ .

Suppose otherwise that  $H$  has an isolated vertex. Then  $G = K_1 + H$  contains a leaf. So  $G$  is a spanning subgraph of  $S_n$ . Label the vertices of  $G$  and  $S_n$  with the same labels such that  $v$  denotes a leaf in each graph and  $u$  denotes the cut vertex of each graph. The sets of vertices that induce a connected subgraph in  $S_n$  can be partitioned into two sets, namely (i) those that also induce a connected subgraph of  $G$  and (ii) those that do not induce a connected subgraph of  $G$ . The latter are precisely those sets of vertices that induce a disconnected subgraph of  $\overline{H-v}$  and thus the nontrivial sets of vertices that induce a connected subgraph in  $\overline{H-v}$ . Hence

$$\Phi_{S_n}(x) = \Phi_G(x) + [\Phi_{\overline{H-v}}(x) - (n-2)x].$$

Thus we obtain the following convex combination:

$$M_{S_n} = c_1 M_G + c_2 M_{\overline{H-v}}^*, \tag{16}$$

where  $c_2 > 0$  unless  $G \cong S_n$ . Since  $\overline{H-v}$  is a cograph of order  $n-2$ , Lemma 2.4 and Lemma 2.3 give

$$M_{\overline{H-v}}^* \leq M_{K_{1,n-3}}^* \leq \frac{n}{2}.$$

Finally, note that  $M_{S_n} = \frac{n}{2} + \frac{1}{3 \cdot 2^{n-2}} > \frac{n}{2}$ . Thus  $M_{\overline{H-v}}^* < M_{S_n}$ , and we conclude from (16) that  $M_{S_n} < M_G$ . □

We now focus on the case where  $G$  is a connected cograph without a cut vertex, i.e.  $G$  is a 2-connected cograph. In this case,  $\Phi_G(x)$  can be expressed as the sum of a local generating polynomial  $\Phi_{G,v}(x)$  and the generating polynomial  $\Phi_{G-v}(x)$  for the connected cograph  $G-v$ . We require one more lemma in order to prove the main result of this section.

**Lemma 4.6.** *Let  $G$  be a 2-connected cograph of order  $n \geq 4$ . There is some vertex  $v \in V(G)$  such that  $\Phi_{G-v}(1) < \Phi_{G,v}(1)$ .*

*Proof.* Since  $G$  is 2-connected, there exist non-trivial cographs  $G_1$  and  $G_2$  such that  $G = G_1 + G_2$ . Let  $v$  be a vertex of maximum degree in  $G$ . We may assume that  $v \in V(G_1)$ . If  $v$  is a universal vertex, then the result follows, since, in this case, there are  $2^{n-1}$  connected induced subgraphs that contain  $v$  and at most  $2^{n-1} - 1$  connected induced subgraphs in  $G-v$ .

Suppose now that  $v$  has  $k \geq 1$  non-neighbours. Let  $\mathcal{P}_1$  be the set of connected induced subgraphs of  $G-v$  that contain a neighbour of  $v$  and let  $\mathcal{P}_2$  be the set of connected induced subgraphs of  $G-v$  that do not contain a neighbour of  $v$ . If  $H$  belongs to  $\mathcal{P}_1$ , then the vertices of  $H$  together with  $v$  induce a connected subgraph that contains  $v$ . Let  $\mathcal{P}_{1,v}$  be the family of such connected induced subgraphs. If  $H$  is in  $\mathcal{P}_2$ , then the vertices of  $H$  belong to  $V(G_1) - N[v]$ . So  $\mathcal{P}_2$  has at most  $2^k - 1$  elements. Let  $u \in V(G_2)$ . Since  $\deg(u) \leq \deg(v)$ ,  $u$  has at least  $k$  non-neighbours

that necessarily belong to  $G_2$ . Let  $S$  be the set of non-neighbours of  $u$ . Each non-empty subset  $X$  of  $S$  together with  $u$  and  $v$  yields a connected induced subgraph that contains  $v$  and does not belong to  $\mathcal{P}_1$ , since  $\langle X \cup \{u\} \rangle$  is disconnected. Let  $\mathcal{P}_{2,v}$  be the family of connected induced subgraphs of  $G$  containing  $v$  that are constructed in this manner. Note that  $|\mathcal{P}_{i,v}| \geq |\mathcal{P}_i|$  for  $i = 1, 2$ , and  $\mathcal{P}_{1,v} \cap \mathcal{P}_{2,v} = \emptyset$ , which gives  $\Phi_{G-v}(1) \leq \Phi_{G,v}(1)$ . Finally, the singleton set containing  $v$  induces a connected subgraph of  $G$  containing  $v$  and is not contained in  $\mathcal{P}_{1,v} \cup \mathcal{P}_{2,v}$ , which gives the desired strict inequality.  $\square$

**Theorem 4.7.** *Let  $G$  be a connected cograph of order  $n \geq 3$ . Then  $M_G \geq \frac{n}{2} + \frac{1}{3 \cdot 2^{n-2}}$ , with equality if and only if  $G \cong S_n$ .*

*Proof.* We begin by deriving a useful expression for  $M_{S_n}$ . Let  $u$  be a vertex of degree  $n - 2$  in  $S_n$ . First note that

$$\Phi_{S_n}(x) = \Phi_{S_n,u}(x) + \Phi_{S_{n-1}}(x).$$

Taking the logarithmic derivative of both sides, we obtain

$$M_{S_n} = \left[ \frac{\Phi_{S_n,u}(1)}{\Phi_{S_n,u}(1) + \Phi_{S_{n-1}}(1)} \right] M_{S_n,u} + \left[ \frac{\Phi_{S_{n-1}}(1)}{\Phi_{S_n,u}(1) + \Phi_{S_{n-1}}(1)} \right] M_{S_{n-1}}.$$

One easily verifies that  $\Phi_{S_n,u}(1) = \Phi_{S_{n-1}}(1) = 2^{n-2} + 2^{n-3}$  and  $M_{S_n,u} = \frac{n+1}{2}$ , which gives

$$M_{S_n} = \frac{1}{2} \cdot \frac{n+1}{2} + \frac{1}{2} M_{S_{n-1}}. \tag{17}$$

We now proceed with the proof of the statement by induction on  $n$ . The result is easily verified for  $n = 3$  as  $M_{K_3} > M_{S_3}$  and  $K_3$  and  $S_3$  are the only distinct connected cographs of order 3. Suppose now that  $n > 3$  and that the statement holds for all connected cographs of order  $k$ ,  $3 \leq k < n$ . If  $G$  has a cut vertex, then we are done, by Lemma 4.5. So we may assume that  $G$  is 2-connected.

By Lemma 4.6 there is a vertex  $v \in V(G)$  such that  $\Phi_{G,v}(1) > \Phi_{G-v}(1)$ . We express  $M_G$  as follows:

$$M_G = \left[ \frac{\Phi_{G,v}(1)}{\Phi_{G,v}(1) + \Phi_{G-v}(1)} \right] M_{G,v} + \left[ \frac{\Phi_{G-v}(1)}{\Phi_{G,v}(1) + \Phi_{G-v}(1)} \right] M_{G-v}.$$

From our choice of  $v$ , it follows that the coefficient of  $M_{G,v}$  is greater than  $\frac{1}{2}$ , while the coefficient of  $M_{G-v}$  is less than  $\frac{1}{2}$ . Since  $M_{G,v} \geq \frac{n+1}{2}$ , by Theorem 3.1, and  $M_{G-v} < \frac{n}{2}$ , by Corollary 2.7, we have  $M_{G,v} > M_{G-v}$ . Thus

$$M_G > \frac{1}{2} M_{G,v} + \frac{1}{2} M_{G-v} \geq \frac{1}{2} \cdot \frac{n+1}{2} + \frac{1}{2} M_{G-v}.$$

Applying the induction hypothesis to  $G - v$  gives

$$M_G > \frac{1}{2} \cdot \frac{n+1}{2} + \frac{1}{2} M_{S_{n-1}} = M_{S_n},$$

where the last equality is due to (17).  $\square$

It follows easily from our results that the global means of connected cographs increase with order.

**Corollary 4.8.** *Let  $G_1$  be a connected cograph of order  $n_1 \geq 2$ , and let  $G_2$  be a cograph (not necessarily connected) of order  $n_2 < n_1$ . Then  $M_{G_1} > M_{G_2}$ .*

*Proof.* The result is easily verified for  $n_1 = 2$ , so we may assume that  $n_1 \geq 3$ . From Theorem 4.7 and Corollary 2.8, we know that  $M_{G_1} \geq M_{S_{n_1}} > \frac{n_1}{2}$ , and  $M_{G_2} \leq \frac{n_2+1}{2}$ , respectively. Since  $n_1 > n_2$ , it follows that  $M_{G_1} > \frac{n_1}{2} \geq \frac{n_2+1}{2} > M_{G_2}$ .  $\square$

### 5 Disconnected cographs

In this section we investigate extremal structures for the mean order of connected induced subgraphs in disconnected cographs. For any order  $n$ , the edgeless cograph  $\overline{K_n}$  has minimum mean. We show that the maximum mean among all disconnected cographs of order  $n \geq 2$  is obtained by the disjoint union of  $K_1$  and the cograph of maximum mean among all connected cographs of order  $n - 1$ . In particular, for  $n \geq 8$ , the cograph  $K_1 \cup K_{1,n-2}$  has maximum mean among all disconnected cographs of order  $n$ .

We will refer to the expressions given below for various means, which are easily verified.

**Lemma 5.1.** *The following formulae hold.*

- (a) For  $n \geq 3$ ,  $M_{K_1 \cup K_{1,n-2}} = \frac{(n-1)+n2^{n-3}}{(n-1)+2^{n-2}} = \frac{n}{2} + \frac{\frac{3n}{2} - \frac{n^2}{2} - 1}{2^{n-2} + (n-1)}$ .
- (b) For  $n \geq 4$ ,  $M_{K_{1,n-3}}^* = \frac{(n-1)2^{n-4}-1}{2^{n-3}-1} = \frac{n}{2} + \frac{\frac{n}{2} - 2^{n-4} - 1}{2^{n-3} - 1}$ .
- (c) For  $n \geq 4$ ,  $M_{K_{2,n-3}} = \frac{3n2^{n-4}-2^{n-4}+n-5}{3 \cdot 2^{n-3} + n - 4} = \frac{n}{2} + \frac{-2^{n-4} + 3n - 5 - \frac{n^2}{2}}{n + 3 \cdot 2^{n-3} - 4}$ .
- (d) For  $n \geq 4$ ,  $M_{K_{1,n-3}} = \frac{(n-3)+(n-1)2^{n-4}}{(n-3)+2^{n-3}} = \frac{n}{2} + \frac{\frac{5n}{2} - \frac{n^2}{2} - 3 - 2^{n-4}}{2^{n-3} + n - 3}$ .

In order to establish our main result we begin by proving some useful inequalities between the means described in Lemma 5.1.

**Lemma 5.2.** *The following inequalities hold.*

- (a)  $M_{K_{1,n-3}}^* < M_{K_1 \cup K_{1,n-2}}$  for  $n \geq 8$ .
- (b)  $M_{K_{2,n-3}} < M_{K_1 \cup K_{1,n-2}}$  for  $n \geq 9$ .
- (c)  $M_{K_{1,n-3}} < M_{K_1 \cup K_{1,n-2}}$  for  $n \geq 4$ .

*Proof.* (a) From Lemma 5.1,  $M_{K_{1,n-3}}^* < M_{K_1 \cup K_{1,n-2}}$  if and only if

$$\frac{\frac{3n}{2} - \frac{n^2}{2} - 1}{2^{n-2} + (n-1)} - \frac{\frac{n}{2} - 2^{n-4} - 1}{2^{n-3} - 1} > 0.$$

The numerator of this expression, given by

$$\left(\frac{3n}{2} - \frac{n^2}{2} - 1\right) (2^{n-3} - 1) - \left(\frac{n}{2} - 2^{n-4} - 1\right) (2^{n-2} + n - 1),$$

simplifies to  $2^{n-4} [2^{n-2} - (n^2 - 2n - 1)]$ . It follows readily by induction, that  $2^{n-2} + 2n + 1 - n^2 > 0$  for  $n \geq 8$ .

(b) From Lemma 5.1,  $M_{K_{2,n-3}} < M_{K_1 \cup K_{1,n-2}}$  if and only if

$$\frac{\frac{3n}{2} - \frac{n^2}{2} - 1}{2^{n-2} + (n - 1)} - \frac{-2^{n-4} + 3n - 5 - \frac{n^2}{2}}{n + 3 \cdot 2^{n-3} - 4} > 0.$$

The numerator of this expression, given by

$$\left(\frac{3n}{2} - \frac{n^2}{2} - 1\right) (n + 3 \cdot 2^{n-3} - 4) - (2^{n-2} + n - 1) \left(-2^{n-4} + 3n - 5 - \frac{n^2}{2}\right),$$

simplifies to  $2^{n-4} [2^{n-2} - 2n - n^2 + 13 + \frac{n-1}{2^{n-4}}]$ . It follows readily by induction on  $n \geq 9$ , that  $2^{n-2} - 2n - n^2 + 13 > 0$ .

(c) From Lemma 5.1,  $M_{K_{1,n-3}} < M_{K_1 \cup K_{1,n-2}}$  if and only if

$$\frac{\frac{3n}{2} - \frac{n^2}{2} - 1}{2^{n-2} + (n - 1)} - \frac{\frac{5n}{2} - \frac{n^2}{2} - 3 - 2^{n-4}}{2^{n-3} + n - 3} > 0.$$

The numerator of this expression, given by

$$\left(\frac{3n}{2} - \frac{n^2}{2} - 1\right) (2^{n-3} + n - 3) - (2^{n-2} + n - 1) \left(\frac{5n}{2} - \frac{n^2}{2} - 3 - 2^{n-4}\right),$$

simplifies to  $2^{n-4} [2^{n-2} + n^2 + 9 - 6n]$ . Once again it is readily shown by induction on  $n \geq 4$ , that  $2^{n-2} + n^2 + 9 - 6n > 0$ .  $\square$

Now we are ready to prove the main result of this section.

**Theorem 5.3.** *For each  $n \geq 1$ , let  $Q_n$  denote the cograph of maximum mean among all connected cographs of order  $n$ . If  $G$  is a disconnected cograph of order  $n \geq 2$ , then  $M_G \leq M_{K_1 \cup Q_{n-1}}$ . Moreover, equality holds for  $n \geq 8$  if and only if  $G \cong K_1 \cup K_{1,n-2}$ .*

*Proof.* We refer the reader to Table 1 and Corollary 2.7 for an explicit description of  $Q_n$  for all  $n$ . Using a computer algebra system, we have verified the statement for  $n \leq 8$ .

We may now assume that  $n \geq 9$ . Suppose  $G$  is a disconnected cograph of order  $n \geq 9$ , and we wish to show that  $M_G \leq M_{K_1 \cup K_{1,n-2}}$ . Let  $H_1, \dots, H_k$  denote the components of  $G$ . Then

$$\Phi_G(x) = \Phi_{H_1}(x) + \dots + \Phi_{H_k}(x).$$

Hence  $M_G$  can be expressed as a convex combination as follows:

$$M_G = c_1 M_{H_1} + \dots + c_k M_{H_k}.$$

Let  $M' = \max\{M_{H_i} : i \in \{1, \dots, k\}\}$ . Then  $M_G \leq M'$ . Let  $H_i$  be such that  $M_{H_i} = M'$ . By Corollary 4.8,  $H_i$  is a component of  $G$  of largest order.

**Case 1:** Suppose  $G$  does not have a component of order  $n - 1$ . Then the largest possible order of a component of  $G$  is  $n - 2$ , and by Corollary 2.7, it follows that  $M_G < M_{K_{1,n-3}}$ . By Lemma 5.2(c),  $M_{K_{1,n-3}} < M_{K_1 \cup K_{1,n-2}}$ . Hence  $M_G < M_{K_1 \cup K_{1,n-2}}$  in this case.

**Case 2:** Suppose  $G$  has a component of order  $n - 1$ . It follows that  $G = K_1 \cup H_2$ , where  $H_2$  is a connected cograph of order  $n - 1$ . In this case

$$\Phi_G(x) = x + \Phi_{H_2}(x).$$

**Subcase 1:** Suppose  $H_2$  has a universal vertex  $v$ . Then  $K_{1,n-2}$  is a spanning subgraph of  $H_2$ . Thus

$$\begin{aligned} \Phi_G(x) &= x + \Phi_{K_{1,n-2}}(x) + \Phi_{H_2-v}(x) - (n - 2)x \\ &= \Phi_{K_1 \cup K_{1,n-2}}(x) + \Phi_{H_2-v}(x) - (n - 2)x. \end{aligned}$$

Since  $H_2 - v$  is a cograph of order  $n - 2$ , it follows that

$$M_G = c_1 M_{K_1 \cup K_{1,n-2}} + c_2 M_{H_2-v}^*,$$

where  $c_2 = 0$  if and only if  $G \cong K_1 \cup K_{1,n-2}$ . By Lemma 2.4,  $M_{H_2-v}^* \leq M_{K_{1,n-3}}^*$ . By Lemma 5.2(a),  $M_{K_{1,n-3}}^* < M_{K_1 \cup K_{1,n-2}}$ . Hence

$$M_G \leq M_{K_1 \cup K_{1,n-2}},$$

with equality in this subcase if and only if  $G \cong K_1 \cup K_{1,n-2}$ .

**Subcase 2:** Suppose  $H_2$  does not have a universal vertex. Since  $H_2$  is connected,  $H_2 = G_1 + G_2$ , where  $G_1$  and  $G_2$  are cographs of orders  $s$  and  $n - s - 1$  respectively, where  $2 \leq s \leq \lfloor \frac{n}{2} \rfloor$ . Thus

$$\Phi_G(x) = (x + \Phi_{K_{s,n-s-1}}(x)) + (\Phi_{G_1}(x) - sx) + (\Phi_{G_2}(x) - (n - s - 1)x).$$

Hence

$$M_G = c_1 M_{K_1 \cup K_{s,n-s-1}} + c_2 M_{G_1}^* + c_3 M_{G_2}^*.$$

Since  $G_1$  and  $G_2$  are cographs of order at most  $n - 3$ , it follows from Lemmas 2.3 and 2.4 that  $M_{G_1}^*, M_{G_2}^* < M_{K_{1,n-4}}^* < M_{K_{1,n-3}}^*$ . By Lemma 5.2(a),  $M_{K_{1,n-3}}^* < M_{K_1 \cup K_{1,n-2}}$ . Further,  $M_{K_1 \cup K_{s,n-s-1}}$  can itself be expressed as a convex combination of  $M_{K_1}$  and  $M_{K_{s,n-s-1}}$ . From this and Theorem 2.2, it follows that  $M_{K_1 \cup K_{s,n-s-1}} < M_{K_{2,n-3}}$ . By Lemma 5.2(b),  $M_{K_{2,n-3}} < M_{K_1 \cup K_{1,n-2}}$  for  $n \geq 9$ . Hence  $M_G < M_{K_1 \cup K_{1,n-2}}$  in this subcase.  $\square$

## 6 Conclusion

In this article we studied the extremal structures for the mean order of the connected induced subgraphs of cographs, in both the connected and disconnected cases. For a connected cograph  $G$  of order  $n \geq 1$ , we proved that

$$\frac{n}{2} < M_G \leq \frac{n+1}{2},$$

with equality on the right if and only if  $n = 1$ . This means that the *density* of  $G$ , defined as  $\frac{M_G}{n}$ , is always close to  $\frac{1}{2}$ ; to be precise, the density of any infinite sequence of distinct connected cographs tends to  $\frac{1}{2}$ . This contrasts the situation for trees, where the density lies between  $\frac{1}{3}$  and 1, and these bounds are asymptotically sharp (see [6] for details).

In Section 3 we showed, for connected cographs, that the local mean at each vertex is at least as large as the global mean, as was the case for trees. We demonstrated that this property does not extend to all connected graphs, providing an infinite family of counterexamples. It remains an interesting open problem to characterize those graphs for which the local mean at each vertex is at least as large as the global mean.

Finally, it remains an interesting open problem to determine the extremal structures for the mean connected induced subgraph order among all connected graphs of order  $n$ . We conjecture that the minimum mean connected induced subgraph order is obtained by the path  $P_n$ , as for trees, and we have verified this statement for  $n \leq 9$ . Determining the structure of graphs with maximum mean seems much more difficult. Using a computer algebra system, we have determined that the maximum is not obtained by a tree for  $3 \leq n \leq 9$ , but instead by a 2-connected graph. The connected graphs with maximum mean are given in Table 2. In the table,  $\Theta_{i,j,k}$  denotes the  $(i, j, k)$ -theta graph, obtained from the multigraph on two vertices with three edges between them by subdividing the first edge  $i$  times, the second edge  $j$  times, and the third edge  $k$  times, and  $G \square H$  denotes the Cartesian product of graphs  $G$  and  $H$ .

Order	Graph	Global Mean	Decimal Expansion
3	$K_3$	$12/7$	1.71...
4	$K_{2,2}$	$28/13$	2.15...
5	$\Theta_{1,1,1}$	$69/26$	2.65...
6	$\Theta_{2,1,1}$	$67/21$	3.19...
7	$\Theta_{2,2,1}$	$83/22$	3.77...
8	$\Theta_{2,2,2}$	$22/5$	4.4
9	$P_3 \square P_3$	$996/197$	5.05...

Table 2: The graphs of maximum global mean among all connected graphs of small order.

## Acknowledgements

The authors wish to thank the anonymous referees for their useful comments that helped to improve this manuscript.

## References

- [1] J.I. Brown and L. Mol, On the roots of the node reliability polynomial, *Networks* **68**(3) (2016), 238–246.
- [2] J.I. Brown and L. Mol, The shape of node reliability, *Discrete Appl. Math.* **238** (2018), 41–55.
- [3] C.J. Colbourn, A. Satyanarayana, C. Suffel and K. Sutner, Computing residual connectedness reliability for restricted networks, *Discrete Appl. Math.* **44** (1993), 221–232.
- [4] D.G. Corneil, H. Lerchs and L. Stewart Burlingham, Complement reducible graphs, *Discrete Appl. Math.* **3** (1981), 163–174.
- [5] O. Goldschmidt, P. Jaillet and R. LaSota, On reliability of graphs with node failures, *Networks* **24** (1994), 251–259.
- [6] R.E. Jamison, On the average number of nodes in a subtree of a tree, *J. Combin. Theory Ser. B* **35** (1983), 207–223.
- [7] R.E. Jamison, Monotonicity of the mean order of subtrees, *J. Combin. Theory Ser. B* **37** (1984), 70–78.
- [8] S. Liu, K.-H. Cheng and X. Liu, Network reliability with node failures, *Networks* **35**(2) (2000), 109–117.
- [9] L. Mol and O.R. Oellermann, Maximizing the mean subtree order, preprint, arXiv:1707.01874[math.CO], 2017.
- [10] K. Sutner, A. Satyanarayana and C. Suffel, The complexity of the residual node connectedness reliability problem, *SIAM J. Comput.* **20**(1) (1991), 149–155.
- [11] A. Vince and H. Wang, The average order of a subtree of a tree, *J. Combin. Theory Ser. B* **100** (2010), 161–170.
- [12] S. Wagner and H. Wang, Indistinguishable trees and graphs, *Graphs Combin.* **30** (2014), 1593–1605.
- [13] S. Wagner and H. Wang, On the local and global means of subtree orders, *J. Graph Theory* **81**(2) (2016), 154–166.
- [14] S. Yu, F.-M. Shao and H. Meng, Uniformly optimal graphs in some classes of graphs with node failures, *Discrete Math.* **310**(1) (2010), 159–166.