

Full Length Article



Invitation to the subpath number

Martin Knor^a, Jelena Sedlar^{b,d, }, Riste Škrekovski^{c,d,e, },* , Yu Yang^f

^a Slovak University of Technology, Bratislava, Slovakia

^b University of Split, FGAG, Split, Croatia

^c University of Ljubljana, FMF, Ljubljana, Slovenia

^d Faculty of Information Studies, Novo Mesto, Slovenia

^e Rudolfovo - Science and Technology Centre, Novo mesto, Slovenia

^f School of Software, Pingdingshan University, Pingdingshan, China

A B S T R A C T

In this paper we count all the subpaths of a given graph G , including the subpaths of length zero, and we call this quantity the subpath number of G . The subpath number is related to the extensively studied number of subtrees, as it can be considered as counting subtrees with the additional requirement of maximum degree being two. We first give the explicit formula for the subpath number of trees and unicyclic graphs. We show that among connected graphs on the same number of vertices, the minimum of the subpath number is attained for any tree and the maximum for the complete graph. Further, we show that the complete bipartite graph with partite sets of almost equal size maximizes the subpath number among all bipartite graphs. The explicit formula for cycle chains, i.e. graphs in which two consecutive cycles share a single edge, is also given. This family of graphs includes the unbranched catacondensed benzenoids which implies a possible application of the result in chemistry. The paper is concluded with several directions for possible further research where several conjectures are provided.

1. Introduction

The number $N(G)$ of non-empty subtrees in a graph G is a concept that was first studied for trees [10]. Various properties of $N(G)$ have been studied for several subclasses of trees, such as trees with given maximum degree [5], binary trees [11], trees with given degree sequence [19]. Recently, the study of the number of subtrees has been extended to some classes of general graphs, such as graphs with given number of cut edges [14] and cacti and block graphs [6]. The inverse problem for the number of subtrees was studied in [3]. For some other interesting results on the topic we refer the reader to [16–18].

Motivated by this research, in this paper we study the number of subpaths in a graph G , we call this quantity the subpath number. The number of subtrees has an enormous magnitude due to which it is hard to evaluate, but we believe that the number of subpaths should be more manageable. Nevertheless, the complexity of counting subpaths of a graph has been studied in [15], where it is established that this problem is $\#P$ -hard.

Most research regarding the subtree number is done on the subtrees of trees. In the case of counting subpaths, every pair of vertices in a tree is connected by only one path, so the subpath number is the same for all trees on a given number of vertices. Hence, it is interesting to consider the subpath number only for classes of general graphs. We establish the exact value of the subpath number for unicyclic graphs, which yields extremal graphs in that class. By showing that the subpath number is strictly increasing with respect to the edge addition, we show that the complete graph maximizes the subpath number, and that trees minimize it among all connected graphs on a given number of vertices. In the class of bipartite graphs, we show that the complete bipartite graph with partite sets of (almost) equal size maximizes the subpath number. We also consider graphs which are a sequence of cycles such that two consecutive

* Corresponding author.

E-mail address: skrekovski@gmail.com (R. Škrekovski).

<https://doi.org/10.1016/j.amc.2025.129646>

Received 27 February 2025; Received in revised form 2 July 2025; Accepted 17 July 2025

Available online 25 July 2025

0096-3003/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

cycles share a single edge. For such graphs we establish the exact formula for the subpath number. Since this class of graphs include ladder graphs, but also the graphs of the unbranched catacondensed benzenoids, our results might have a potential application to chemistry [1].

The Wiener index $W(G)$ of a graph G is defined as the sum of distances over all pairs of vertices in G . It was introduced in the seminal paper by Wiener [12], where a correlation of the Wiener index with the chemical properties of some molecular compounds is established. Since then, the Wiener index has become one of the most researched indices in chemical graph theory, for an overview of the results we refer the reader to two surveys [7,8]. An interesting “negative” correlation of the number of subtrees and the Wiener index is also observed, in the sense that on many graph classes the graph which maximizes the number of subtrees is the same graph which minimizes the Wiener index, and vice versa. This is observed on several graph classes such as trees, unicyclic graphs and cacti, for more details see [13]. One possible direction for further research is to establish whether the subpath number correlates with one of these quantities in this sense, or a class of graphs on which it does not correlate to either of them.

This paper is concluded with some other considerations regarding further work. Given the result for bipartite graphs, we conjecture that the complete bipartite graph with partite sets of (almost) equal size maximizes the subpath number among all triangle-free graphs also. Moreover, since the addition of an edge strictly increases the subpath number, an interesting question is whether a regular graph maximizes (resp. minimizes) the subpath number among all graphs with prescribed maximum degree (resp. minimum degree). This focuses one’s attention to regular graphs and the question which graphs are maximal and which are minimal with respect to the subpath number in the class of r -regular graphs. Since a 2-regular graph is a collection of cycles for which the subpath number is easily established, the first non-trivial case is $r = 3$, i.e. cubic graphs. For this case, we conjecture extremal graphs.

2. Preliminary results

For a graph G , the *subpath number* is defined as the number of paths in G , including the paths of length 0. The subpath number of G is denoted by $pn(G)$. Let us first observe some interesting properties of such a number. Assume that the vertices of G are denoted by v_1, \dots, v_n , where d_i denotes the degree of the vertex v_i . We can now partition the paths of the graph G into classes of given length, i.e., we define $pn_l(G)$ to be the number of all paths of length l in G . Obviously, it holds that

$$pn(G) = \sum_{l=0}^{n-1} pn_l(G).$$

Before we proceed to the proposition which gives the value of $pn_l(G)$ for small l , let us first introduce the first and the second Zagreb index of a graph G , denoted by $M_1(G)$ and $M_2(G)$. These two indices are defined by

$$M_1(G) = \sum_{i=1}^n d_i^2 \text{ and } M_2(G) = \sum_{v_i v_j \in E(G)} d_i d_j,$$

respectively. Also, a cycle of length 3 will be called a *triangle*. For the paths of a small length the following proposition holds.

Proposition 1. *Let G be a graph on n vertices denoted by v_i , for $i = 1, \dots, n$, where d_i denotes the degree of the vertex v_i . Then the following holds:*

- $pn_0(G) = |V(G)|$;
- $pn_1(G) = |E(G)|$;
- $pn_2(G) = \sum_{i=1}^n \binom{d_i}{2}$;
- $pn_3(G) = M_2(G) - M_1(G) + |E(G)| - 3t$, where t denotes the number of triangles in G .

Proof. Since every vertex of G represents one path of length 0 in G , it follows that $pn_0(G) = |V(G)|$. Similarly, since each edge of G represents one path of length 1 in G , it follows that $pn_1(G) = |E(G)|$. Let us next consider $pn_2(G)$. Denote by \mathcal{P}_i the set of all paths of length two which have v_i for the middle vertex. Both edges of such a path are incident to v_i , hence $|\mathcal{P}_i| = \binom{d_i}{2}$. The formula for $pn_2(G)$ now follows from the fact that the sets \mathcal{P}_i and \mathcal{P}_j are disjoint for $i \neq j$.

Let us finally consider $pn_3(G)$. Here, we focus on the middle edge e of a path in G of length 3. Denote by $\mathcal{P}_{i,j}$ the set of all paths of G which have $v_i v_j$ as the middle edge. Notice that the other edge of such a path incident to v_i can be chosen in $d_i - 1$ ways and the other edge incident to v_j in $d_j - 1$ ways. Thus, we obtain quantity $(d_i - 1)(d_j - 1)$, but here we also counted the triangles of G which contain e as an edge. Since two paths of length three with distinct middle edge must be distinct, we obtain

$$\begin{aligned} pn_3(G) &= \sum_{v_i v_j \in E(G)} (d_i - 1)(d_j - 1) - 3t \\ &= \sum_{v_i v_j \in E(G)} d_i d_j - \sum_{v_i v_j \in E(G)} (d_i + d_j) + |E(G)| - 3t. \end{aligned}$$

The first member of the above sum equals $M_2(G)$ by definition, so it remains to show that the second member of the sum equals $M_1(G)$. To see this, notice that for each vertex v_i of G , the degree d_i is added in the sum d_i times, hence we have

$$\sum_{v_i, v_j \in E(G)} (d_i + d_j) = \sum_{v_i \in V(G)} d_i^2 = M_1(G),$$

and we are done. ■

Let us next consider trees on n vertices. From the well known fact that every pair of vertices in a tree is connected by the unique path, the following observation follows.

Observation 2. If T is a tree on n vertices, then $\text{pn}(T) = \binom{n+1}{2}$.

Since unicyclic graphs are obtained from trees by introducing a single edge, the natural next step is to consider unicyclic graphs on n vertices. In a unicyclic graph G with the cycle of length g , removing all edges of the cycle results in precisely g connected components.

Proposition 3. Let G be a unicyclic graph on n vertices with the cycle C of length g . Denote by n_1, n_2, \dots, n_g the number of vertices in the components which occur after deleting all edges of C . Then, $\text{pn}(G) = n + 2\binom{n}{2} - \binom{n_1}{2} - \binom{n_2}{2} - \dots - \binom{n_g}{2}$.

Proof. Denote by $C = v_1 v_2 \dots v_g v_1$ the only cycle in G . Let G_i be the connected component of $G \setminus E(C)$ which contains v_i . We may assume that G_i has n_i vertices. Notice the following, if a pair of vertices belongs to the same connected component G_i , then there exists a unique path connecting them. On the other hand, if two vertices belong to two distinct connected components of $G \setminus E(C)$, say v_i and v_j , then there are precisely two distinct paths in G connecting such a pair of vertices. So, we count the paths of G as follows. First, there are n paths of length zero, as each vertex of G represents one such paths. Second, for each pair of vertices of G we may count that there are two paths which connect them, this gives $2\binom{n}{2}$ paths. From this quantity we must subtract the number of pairs of vertices in G which are connected by only one path. For each connected component G_i there are $\binom{n_i}{2}$ such pairs. Hence, we obtain

$$\text{pn}(G) = n + 2\binom{n}{2} - \binom{n_1}{2} - \binom{n_2}{2} - \dots - \binom{n_g}{2},$$

as claimed. ■

The above result easily yields the value of the subpath number for the cycle on n vertices.

Corollary 4. For the cycle on n vertices we have $\text{pn}(C_n) = n^2$.

Also, from the formula for the subpath number of a unicyclic graph given in Proposition 3, it is easy to deduce extremal unicyclic graphs with respect to the subpath number.

Corollary 5. Among unicyclic graphs on n vertices, $\text{pn}(G)$ attains maximum value if and only if $G = C_n$, and it attains the minimum value if and only if the only cycle of G is a triangle with two of its vertices being of degree two.

Let us next consider a random graph $G(n, p)$ on n vertices where the probability of each edge is p , for $0 \leq p \leq 1$. For a set of outcomes Ω , a random variable X is a mapping from Ω to the real numbers. Given a random variable X , the expected value $\mathbb{E}(X)$ is defined as the average value of the random variable where each value is weighted according to its probability, i.e. $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \cdot \text{pr}(\omega)$ where $\text{pr}(\omega)$ denotes the probability of the outcome ω . Notice that the family of graphs $G(n, p)$ can be considered as the set of outcomes, and $\text{pn}(G(n, p))$ as the random variable on that set. The following proposition gives the expected value of the subpath number for $G(n, p)$.

Proposition 6. For a random graph $G(n, p)$, where $0 \leq p \leq 1$, it holds that

$$\mathbb{E}(\text{pn}(G(n, p))) = \frac{1}{2} \sum_{k=1}^n \frac{n!}{(n-k)!} p^{k-1} + \frac{n}{2}.$$

Proof. A sequence of k vertices induces a path $P = v_1 v_2 \dots v_k$ ($k \geq 2$) in $G(n, p)$ with probability p^{k-1} . There are $n!/(n-k)!$ such sequences, so the expectation of the number of k -paths X_k in $G(n, p)$ is $\mathbb{E}(X_k) = 1/2 n!/(n-k)! p^{k-1}$, wherein factor $1/2$ appears since each path corresponds to two such sequences. As there are always n paths of length 0, we obtain

$$\mathbb{E}(\text{pn}(G(n, p))) = \sum_{k=1}^n \mathbb{E}(X_k) = \frac{1}{2} \sum_{k=1}^n \frac{n!}{(n-k)!} p^{k-1} + \frac{n}{2},$$

as claimed. ■

The above proposition yields the value of the subpath number for the complete graph K_n , since the probability p of each edge in K_n equals 1. It is easily verified that

$$\sum_{k=1}^n \binom{n}{k} k! = n! \sum_{k=0}^{n-1} \frac{1}{k!},$$

so we have the following result.

Corollary 7. For the complete graph K_n , we have

$$\text{pn}(K_n) = \frac{n!}{2} \sum_{k=0}^{n-1} \frac{1}{k!} + \frac{n}{2}.$$

Approximating further the expression for $\text{pn}(K_n)$, we obtain

$$\text{pn}(K_n) \sim \frac{n!}{2} e + \frac{n}{2}$$

when $n \rightarrow +\infty$. Also, we establish the value of the subpath number for the complete bipartite graph.

Proposition 8. For the complete bipartite graph $K_{a,b}$, where $a + b = n$ and $a \leq b$, we have

$$\begin{aligned} \text{pn}(K_{a,b}) &= \binom{a}{2} \sum_{k=1}^{a-1} \binom{b}{k} k! \cdot \binom{a-2}{k-1} (k-1)! \\ &\quad + \binom{b}{2} \sum_{k=1}^a \binom{a}{k} k! \cdot \binom{b-2}{k-1} (k-1)! \\ &\quad + a \cdot b \sum_{k=1}^a \binom{a-1}{k-1} (k-1)! \cdot \binom{b-1}{k-1} (k-1)! + n. \end{aligned}$$

Proof. Denote the two bipartite sets of $K_{a,b}$ by A and B , so that $|A| = a$ and $|B| = b$. Assume first that $x, y \in A$ and $x \neq y$. Notice that there are $\binom{a}{2}$ such pairs. Since $K_{a,b}$ is bipartite, x and y can be connected only by a path of an even length. The number of subpaths of G of length $2k$ which connect x and y is $\binom{b}{k} k! \cdot \binom{a-2}{k-1} (k-1)!$. Since a longest subpath of G connecting x and y is of length at most $2(a-1)$, it follows that $1 \leq k \leq a-1$. So, there are

$$\binom{a}{2} \sum_{k=1}^{a-1} \binom{b}{k} k! \cdot \binom{a-2}{k-1} (k-1)!$$

paths which connect such pairs x and y . Notice that this is the expression from the first line of the formula for $\text{pn}(K_{a,b})$ from the statement. The same reasoning for $x, y \in B$ and $x \neq y$, yields the expression from the second line of the formula. Notice that due to $a \leq b$, here we have $1 \leq k \leq a$ and not $1 \leq k \leq b-1$.

Assume next that $x \in A$ and $y \in B$. There are $a \cdot b$ such pairs and obviously $x \neq y$. Vertices x and y can be connected only by a path of an odd length $2k-1$ for $1 \leq k \leq a$. For such a pair x and y , there are $\binom{a-1}{k-1} (k-1)! \cdot \binom{b-1}{k-1} (k-1)!$ subpaths of $K_{a,b}$ connecting them. Finally, the n subpaths of zero length must be added to $\text{pn}(K_n)$ and we are done. ■

Next, we consider all connected graphs where we show that the minimum value of the subpath number is obtained for any tree and the maximum value for the complete graph. For that purpose we need the following statement.

Lemma 9. Let G be a connected graph on n vertices and let e be an edge of G . Let G' be a graph obtained from G by removing the edge e . Then $\text{pn}(G') < \text{pn}(G)$.

Proof. The claim follows from the fact that every path in G' is also a path in G . On the other hand, the edge e is a path of length 1 in G , which is not contained in G' . ■

We now have the following result.

Theorem 10. Let G be a connected graph on n vertices. Then

$$\binom{n+1}{2} \leq \text{pn}(G) \leq \frac{n!}{2} \sum_{i=0}^{n-1} \frac{1}{i!} + \frac{n}{2},$$

where the lower bound is attained if and only if G is a tree, and the upper bound if and only if $G = K_n$.

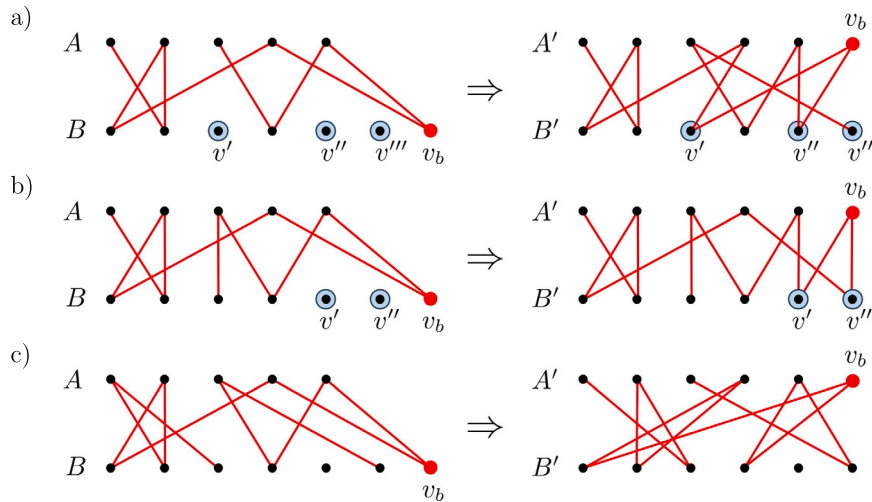


Fig. 1. All figures on the left show a path P in the complete bipartite graph $K_{5,7}$ which contains $v_b = v_7$. The figures on the right show the corresponding path $f(P) = P'$ in $K_{6,6}$. Edges of $K_{5,7}$ and $K_{6,6}$ not contained on the path are not shown for the simplicity sake. Notice that: a) $P = u_1 v_2 u_2 v_1 u_4 v_3 v_5 u_5 v_4 u_3 \in \mathcal{P}_A$ so $P' = u_1 v_2 u_2 v_1 u_4 v_3 v_5 u_5 v_4 u_3 \in \mathcal{P}'_{A,B}$, b) $P = u_1 v_2 u_2 v_1 u_4 v_3 v_5 v_4 u_3 v_3 \in \mathcal{P}_{A,B}$ so $P' = u_1 v_2 u_2 v_1 u_4 v_3 v_5 u_5 v_4 u_3 v_3 \in \mathcal{P}'_{A,B}$, c) $P = v_3 u_1 v_2 u_2 v_1 u_4 v_3 v_5 v_4 u_3 v_6 \in \mathcal{P}_B$ so $P' = u_1 v_3 u_2 v_2 u_4 v_1 v_3 v_4 u_5 v_6 u_3 \in \mathcal{P}_A$.

Proof. The equality in lower and upper bound follows from Observation 2 and Corollary 7, respectively. Let us now prove the inequalities. Let G be a connected graph on n vertices which is neither a tree, nor a complete graph. Then, there exists a sequence of edges of G by removing which a tree is obtained. Also, there exists a sequence of edges which can be added to G to obtain the complete graph K_n . Then the claim follows from Lemma 9. ■

3. Bipartite graphs

In the previous section, we established the subpath number of the complete bipartite graphs. In this section we consider bipartite graphs more generally, namely we characterize extremal bipartite graphs with respect to the subpath number. Since every tree is a bipartite graph, Theorem 10 implies that a bipartite graph G minimizes the subpath number if and only if G is a tree. It remains to consider the bipartite graphs which maximize the subpath number.

Theorem 11. Among bipartite graphs on n vertices the maximum value of the subpath number is attained only by the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lfloor n/2 \rfloor}$.

Proof. Let G be a bipartite graph with partition sets of sizes a and b , where we assume $a \leq b$. If G is not a complete bipartite graph, then Lemma 9 implies $\text{pn}(G) < \text{pn}(K_{a,b})$. Thus, assume that G is a complete bipartite graph, i.e., $G = K_{a,b}$. If $b - a \leq 1$, then $G = K_{\lfloor n/2 \rfloor, \lfloor n/2 \rfloor}$, and the claim follows. Otherwise, assume $b - a \geq 2$. Let $A = \{u_1, \dots, u_a\}$ be the smaller partite set of G and $B = \{v_1, \dots, v_b\}$ the larger one. We denote $u_i < u_j$ (resp. $v_i < v_j$) if $i < j$. Define G' as the graph obtained from G by removing the edges $v_b u_i$ for $1 \leq i \leq a$ and introducing the edges $v_b v_i$ for $1 \leq i \leq b - 1$. Notice that $G' = K_{a+1, b-1}$, where the two partite sets are $A' = A \cup \{v_b\}$ and $B' = B \setminus \{v_b\}$. As for the labels and the corresponding ordering of vertices of A' and B' , they are inherited from G , where we additionally have to define only $u_i < v_b$ for every $i \leq a$.

We now show that $\text{pn}(G) < \text{pn}(G')$. Let \mathcal{P} and \mathcal{P}' denote the sets of all subpaths of G and G' , respectively. Our goal is to construct an injective but not surjective mapping $f : \mathcal{P} \rightarrow \mathcal{P}'$, as the existence of such f implies $\text{pn}(G) < \text{pn}(G')$.

First, we define partitions of \mathcal{P} and \mathcal{P}' which will be useful to us. A path $P \in \mathcal{P}$ belongs to \mathcal{P}_A (resp. \mathcal{P}_B) if both its end-vertices belong to A (resp. B). Similarly, a path $P' \in \mathcal{P}'$ belongs to \mathcal{P}'_A (resp. \mathcal{P}'_B) if both its end-vertices belong to A' (resp. B'). We also define $\mathcal{P}_{AB} = \mathcal{P} \setminus (\mathcal{P}_A \cup \mathcal{P}_B)$ and we obtain the partition of \mathcal{P} into \mathcal{P}_A , \mathcal{P}_B and \mathcal{P}_{AB} . Similarly we define \mathcal{P}'_{AB} to obtain the analogous partition of \mathcal{P}' .

The mapping f is defined as follows. If a path $P \in \mathcal{P}$ does not contain v_b , set $f(P) = P$. Now, consider a subpath P of G that contains v_b . Denote vertices of P by $x_1 x_2 \dots x_k$ so that $x_i x_{i+1}$ is an edge of P . If $P \in \mathcal{P}_A \cup \mathcal{P}_{AB}$, we assume $x_1 \in A$. Additionally, if $P \in \mathcal{P}_A \cup \mathcal{P}_B$, we assume $x_1 < x_k$. Denote by i^* the index for which $x_{i^*} = v_b$. We construct the corresponding subpath P' in the following way illustrated by Fig. 1:

- If $P \in \mathcal{P}_A$, then $b - a \geq 2$ implies $B \setminus V(P)$ contains at least three vertices. Denote them by v', v'' and v''' so that $v' < v'' < v'''$. Let P' be the path obtained from P by inserting v' right before x_{i^*} , v'' right after x_{i^*} and v''' after x_k in the vertex sequence of P . Notice that P' is a subpath of G' and $P' \in \mathcal{P}'_{AB}$.
- If $P \in \mathcal{P}_{AB}$, then $b - a \geq 2$ implies $B \setminus V(P)$ contains at least two vertices. Denote them by v' and v'' so that $v' < v''$. Let P' be a path obtained from P by inserting v'' right before x_{i^*} and v' right after x_{i^*} . Then again, P' is a subpath of G' with $P' \in \mathcal{P}'_{AB}$.

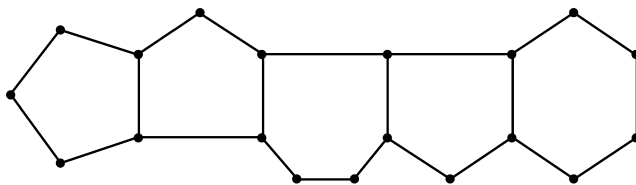


Fig. 2. Graph $G(S)$ with $S = (3, 2, 1, 1, 4; 3, 1, 3, 2, 4)$.

- If $P \in \mathcal{P}_B$, then i^* is odd and k is odd. Let P' be the path obtained from P by swapping x_{2i-1} and x_{2i} for $2i < i^*$, and also swapping x_{2i} and x_{2i+1} for $2i > i^*$. Again, P' is a subpath of G' and $P' \in \mathcal{P}'_A$.

We define $f(P) = P'$, so the mapping f is completely defined. Notice that $f(P)$ contains v_b if and only if P contains v_b .

Let us first show that f is injective, i.e. that $f(P) = f(Q)$ implies $P = Q$. For that purpose, let $P, Q \in \mathcal{P}$ be a pair of paths with $f(P) = f(Q)$. We denote $P' = f(P)$ and $Q' = f(Q)$. If $P' = Q'$ does not contain v_b , then $P = P' = Q' = Q$. So, let us assume that $P' = Q'$ contains v_b . Let $P = x_1 \cdots x_k$ and $Q = y_1 \cdots y_q$ be the vertex labelings of P and Q . Similarly, assume $P' = x'_1 \cdots x'_{k'}$ and $P' = Q'$ implies this is also a labeling of Q' . Analogously as in \mathcal{P} , we assume that if $P' \in \mathcal{P}'_A \cup \mathcal{P}'_{AB}$, then $x'_1 \in A$. Additionally, if $P' \in \mathcal{P}'_A \cup \mathcal{P}'_B$, then we assume $x'_{i^*} < x'_{k'}$. Let i^{**} satisfy $x'_{i^{**}} = v_b$. We distinguish three cases.

Case 1: $P' \in \mathcal{P}'_{A,B}$ and $x'_{i^{**}-1} < x'_{i^{**}+1}$. Then both P' and Q' were obtained from paths in \mathcal{P}_A , i.e. $P, Q \in \mathcal{P}_A$. Hence, both P and Q are obtained from P' by deleting $x'_{i^{**}-1}, x'_{i^{**}+1}, x'_{k'}$ which implies $P = Q$.

Case 2: $P' \in \mathcal{P}'_{A,B}$ and $x'_{i^{**}+1} < x'_{i^{**}-1}$. Then both P' and Q' were obtained from paths in $\mathcal{P}_{A,B}$. Hence, P as well as Q are obtained from P' by deleting $x'_{i^{**}-1}, x'_{i^{**}+1}$, so $P = Q$.

Case 3: $P' \in \mathcal{P}'_A$. Then both P' and Q' were obtained from paths in \mathcal{P}_B . Hence, both P and Q are obtained from P' by swapping x'_{2i-1} and x'_{2i} for $2i < i^{**}$, and also by swapping x'_{2i} and x'_{2i+1} for $2i > i^{**}$. So again $P = Q$, although we have to reverse both P and Q if $x'_{k'-1} < x'_2$.

Therefore, f is an injection. It is easily seen that f is not a surjection since the path $v_1 v_b$ of G' is not $f(P)$ for any path P of G . ■

From the above theorem and Proposition 8, the sharp upper bound on $\text{pn}(G)$ for a bipartite graph G can be derived.

4. Cycle chains

As the next class of graphs we consider the so called cycle chains where two consecutive cycles share a single edge and no vertex is shared by more than 2 cycles. More precisely, let $k \geq 2$ and let $S = (a_1, a_2, \dots, a_k; b_1, b_2, \dots, b_k)$ be a sequence of positive integers such that $a_1 = b_1$ and $a_k = b_k$. Take a collection of $k - 1$ independent edges $u_1 v_1, u_2 v_2, \dots, u_{k-1} v_{k-1}$. Now join u_1 with v_1 by a path of length $a_1 + 1$ and join u_{k-1} with v_{k-1} by a path of length $a_k + 1$. Further, for $2 \leq i \leq k - 1$, join u_{i-1} with v_i by a path of length a_i and join v_{i-1} with v_i by a path of length b_i . Finally, denote the resulting graph by $G(S)$ and called a cycle chain. A graph $G(S)$ is illustrated by Fig. 2.

Notice that $G(S)$ is a graph which consists of the sequence of k cycles C_i , such that two consecutive cycles C_i and C_{i+1} share a single edge, this edge being $u_i v_i$ for $1 \leq i \leq k - 1$. Further, the length of cycles C_1 and C_k is $a_1 + 2$ and $a_k + 2$, respectively. The length of a cycle C_i , for $2 \leq i \leq k - 1$, is $a_i + b_i + 2$. Also, observe that $G(S)$ has

$$a_1 + a_k + 2(k - 1) + \sum_{i=2}^{k-1} ((a_i - 1) + (b_i - 1)) = \sum_{i=1}^{k-1} (a_i + b_{i+1}) + 2$$

vertices.

It is worthy to note that this family of graphs includes the so called unbranched catacondensed benzenoids, this is the case when all cycles in the sequence are of the length six. Hence, our results have a possible application in chemistry, as they can be used to derive chemical properties of the mentioned chemical compounds, see for example [1] and all the references within.

We have the following statement.

Theorem 12. Let $k \geq 2$ and let $S = (a_1, a_2, \dots, a_k; b_1, b_2, \dots, b_k)$ be a sequence of positive integers such that $a_1 = b_1$ and $a_k = b_k$. Denote $n = \sum_{i=1}^{k-1} (a_i + b_{i+1}) + 2$. Then

$$\begin{aligned} \text{pn}(G(S)) &= \sum_{1 \leq i < j \leq k} (a_i b_j + b_i a_j) 2^{j-i-1} (i+1)(k-j+2) - a_1 a_k 2^k \\ &+ \sum_{2 \leq i < j \leq k-1} (a_i a_j + b_i b_j) 2^{j-i-1} (i+1)(k-j+2) \end{aligned}$$

$$\begin{aligned}
 & + \sum_{i=2}^{k-1} \left(\binom{a_i+1}{2} + \binom{b_i+1}{2} \right) (1+i(k-i+1)) \\
 & + \left(\binom{a_1+2}{2} + \binom{a_k+2}{2} \right) (k+1) \\
 & + \sum_{i=2}^{k-1} (a_i+1)(b_i+1)(k+1) - (k-1)(k+1) + n.
 \end{aligned}$$

Proof. First, denote the set of a_1 (resp. a_k) interior vertices of a path connecting u_1 with v_1 (resp. u_{k-1} with v_{k-1}) by A_1 (resp. A_k) and let $B_1 = A_1$ and $B_k = A_k$. Further, denote the set of $a_i - 1$ (resp. $b_i - 1$) interior vertices of a path connecting u_{i-1} with u_i (resp. v_{i-1} with v_i) by A_i (resp. B_i), where $2 \leq i \leq k - 1$. Also, recall that C_i denotes the cycle consisting of $u_{i-1}, v_{i-1}, u_i, v_i, A_i, B_i$, for $2 \leq i \leq k - 1$, and that C_1 (resp. C_k) denotes the cycle consisting of $u_1 v_1, A_1$ (resp. u_{k-1}, v_{k-1}, A_k). To simplify the argumentation, we assume that edges $u_1 v_1, u_2 v_2, \dots, u_{k-1} v_{k-1}$ are drawn in a vertical position and the cycles C_1, C_2, \dots, C_k are convex polygons. For each pair of distinct vertices x and y of $G(S)$ we count the number of paths between x and y .

Case 1: $x \in \{u_{i-1}\} \cup A_i$ and $y \in B_j \cup \{y_j\}$, where $1 \leq i < j \leq k$. Consider a path P between x and y . If we cut $G(S)$ vertically between x and y , the cut must intersect P odd number of times, hence once. On the other hand if we cut $G(S)$ vertically on the left-hand side of x or on the right-hand side of y , then the cut intersects P twice or 0 times. Thus, if P contains $u_\ell v_\ell$ for $\ell < i$, this edge determines P on the left-hand side from x . Analogously, if P contains $u_\ell v_\ell$ for $\ell \geq j$, this edge determines P on the right-hand side from y . But, between x and y any choice of edges $u_i v_i, u_{i+1} v_{i+1}, \dots, u_{j-1} v_{j-1}$ is possible since each such choice can be completed to a path in $G(S)$ (in fact to 2 distinct paths in $G(S)$) in the vertical strip of plane between x and y .

Denote $E_{i,j} = \{u_i v_i, u_{i+1} v_{i+1}, \dots, u_{j-1} v_{j-1}\}$ and let $T \subseteq E_{i,j}$. Let us count the paths P connecting x and y such that $E(P) \cap E_{i,j} = T$. Assume first that T contains an even number of edges of $E_{i,j}$. If the path P starts by going right from x , then there are $k - j + 1$ such paths. On the other hand, if the path P starts from x by going left or down, then there are i such paths. Since there are $2^{j-i-1} = 2^{(j-1)-(i-1)}$ subsets T of $E_{i,j}$, half of which are even sized, there are

$$2^{j-i-1}((k-j+1) + i)$$

paths between x and y which contain an even sized $T \subseteq E_{i,j}$.

Assume next that T contains an odd number of edges of $E_{i,j}$. If the path P starts from x by going right, there is only one path between x and y containing T . On the other hand, if P starts from x by going left or down, there are $i(k-j+1)$ such paths between x and y . Again, since there are 2^{j-i-1} odd sized subsets T of $E_{i,j}$, we conclude that there are

$$2^{j-i-1}(1 + i(k-j+1))$$

paths between x and y containing odd sized $T \subseteq E_{i,j}$.

Summing the above two numbers yields

$$2^{j-i-1}((k-j+1) + i + 1 + i(k-j+1)) = 2^{j-i-1}(i+1)(k-j+2),$$

which is the total number of subpaths of $G(S)$ connecting x and y .

Case 2: $x \in \{v_{i-1}\} \cup B_i$ and $y \in A_j \cup \{u_j\}$, where $1 \leq i < j \leq k$. By the analogous reasoning as in Case 1, we conclude that there are $2^{j-i-1}(i+1)(k-j+2)$ subpaths of $G(S)$ connecting x and y .

Notice that there are $a_i b_j + b_i a_j$ pairs of vertices x and y which belong to Cases 1 or 2. Also, the case $i = 1$ and $j = k$ is included in both cases, so in the final sum we have to subtract $a_1 b_k 2^{k-2}(1+1)(k-k+2) = a_1 a_k 2^k$ from these two sums. Hence, the total number of subpaths of $G(S)$ which connect pairs of vertices from Cases 1 or 2 equals

$$\sum_{1 \leq i < j \leq k} (a_i b_j + b_i a_j) 2^{j-i-1} (i+1)(k-j+2) - a_1 a_k 2^k$$

which is the expression from the first line of the formula from the statement of Theorem 12.

Case 3: $x \in \{u_{i-1}\} \cup A_i$ and $y \in A_j \cup \{u_j\}$, where $2 \leq i < j \leq k - 1$. Observe that the cases when $x \in A_1$ or $y \in A_k$ were counted already in the previous two cases, so they are not included here. Here, the role of even and odd subsets $T \subseteq E_{i,j}$ is reversed. Namely, for an even sized T , there is 1 path P starting from x to the right and ending in y , and there are $i(k-j+1)$ paths P starting from x to the left or down and ending in y . On the other hand, if T is odd sized, then there is $k-j+1$ paths starting from x to the right, and i paths starting from x to the left or down. Since the number of both even sized and odd sized subsets $T \subseteq E_{i,j}$ is equal to 2^{j-i-1} , we obtain that there are precisely

$$2^{j-i-1}(1 + i(k-j+1) + (k-j+1) + i) = 2^{j-i-1}(i+1)(k-j+2)$$

subpaths of $G(S)$ which connect x and y . Notice that this number is the same as in Cases 1 and 2.

Case 4: $x \in \{v_{i-1}\} \cup B_i$ and $y \in B_j \cup \{v_j\}$, where $2 \leq i < j \leq k - 1$. By the analogous reasoning as in Case 3, we obtain the same number of subpaths between x and y .

Since there are $a_i a_j$ (resp. $b_i b_j$) pairs of vertices x and y which belong to Case 3 (resp. Case 4), these pairs contribute to $pn(G(S))$ by the sum from the second line of the formula from the statement of this theorem.

Case 5: $x, y \in V(C_i)$, where $1 \leq i \leq k$. Here, we distinguish several subcases, the one is when both x and y belong to $A_i \cup \{u_{i-1}, u_i\}$, the other is when both x and y belong to $B_i \cup \{v_{i-1}, v_i\}$, and the last is when x belongs to the former and y to the latter set. The case when $i = 1$ and $i = k$ has to be considered separately.

Let us first assume that $x, y \in A_i \cup \{u_{i-1}, u_i\}$, where $2 \leq i \leq k - 1$ and $x \neq y$. We may assume that x is on the left-hand side from y . If P goes from x to the right towards y , then we have just 1 possibility. But going from x to the left-hand side we have i possibilities at the beginning and $k - j + 1$ possibilities at the end, which gives $1 + i(k - j + 1)$ distinct subpaths of $G(S)$ connecting x and y . Since there are $\binom{a_i+1}{2}$ pairs of vertices x and y with $x, y \in A_i \cup \{u_{i-1}, u_i\}$, such pairs contribute to $pn(G(S))$ with

$$\sum_{i=2}^{k-1} \binom{a_i+1}{2} (1 + i(k - j + 1)).$$

Assume next that $x, y \in B_i \cup \{v_{i-1}, v_i\}$, where $2 \leq i \leq k - 1$ and $x \neq y$. The same reasoning as above yields the analogous sum, hence we obtain the sum in the third line of the formula from the statement of this theorem.

Let us next consider the cases $i = 1$ and $i = k$. Observe that there are two paths connecting x with y which contain none of $u_1 v_1, u_2 v_2, \dots, u_{k-1} v_{k-1}$ and for each $u_i v_i$, $1 \leq i \leq k - 1$, there is a unique path connecting x with y which contains $u_i v_i$. Hence, all pairs $x, y \in V(C_1)$ contribute to $G(S)$ by $\binom{a_1+2}{2}(k + 1)$. The same reasoning yields analogous sum for $x, y \in V(C_k)$, which gives the expression in the fourth line of the formula from the statement of this theorem.

Finally, assume that $x \in A_i \cup \{u_{i-1}, u_i\}$ and $y \in B_i \cup \{v_{i-1}, v_i\}$, where $2 \leq i \leq k - 1$. There are $(a_i + 1)(b_i + 1)$ such pairs x and y , and notice that each path connecting x and y contains at most one vertical edge of $G(S)$. So there are $k + 1$ paths connecting x with y , two of which contain no edge from $u_1 v_1, u_2 v_2, \dots, u_{k-1} v_{k-1}$. The problem is that paths connecting u_j with v_j are counted twice here if $2 \leq j \leq k - 2$, and paths connecting u_1 with v_1 and those connecting u_{k-1} with v_{k-1} are already included in $\binom{a_1+2}{2} + \binom{a_k+2}{2}(k + 1)$. Therefore, we must subtract their number once. Since there are $k - 1$ such pairs each connected by $k + 1$ subpaths, this number is $(k - 1)(k + 1)$. Finally, a number of subpaths of zero length which is equal to the number n of vertices in $G(S)$ must be included into $pn(G(S))$, which gives the last line of the formula from the statement of this theorem. ■

In a cycle chain $G(S)$, let us denote by g_i the length of cycle C_i for $1 \leq i \leq k$. Notice that this implies $g_1 = a_1 + 2$, $g_k = a_k + 2$, and $g_i = a_i + b_i + 2$ for $2 \leq i \leq k - 1$. A cycle C_i of $G(S)$ is called *interior* if $2 \leq i \leq k - 1$. An interior cycle C_i is *linear* (resp. *almost linear*) if $a_i = b_i$ (resp. $|a_i - b_i| = 1$). Notice that an interior cycle of an even length can be linear but not almost linear, and for an odd length cycle it is vice versa. Further, an interior cycle C_i is called a *kink* if $a_i = 1$ or $b_i = 1$. A chain $G(S)$ is a *kink chain* if every interior cycle is a kink. Similarly, $G(S)$ is *linear* if every interior cycle of $G(S)$ is linear. If $G(S)$ contains an odd length cycle, then it cannot be linear, so we additionally introduce the notion of an “almost” linear chain. A chain $G(S)$ is *almost linear* if it is not linear and each of its interior cycles is either linear or almost linear.

Corollary 13. Let (g_1, \dots, g_k) be a sequence of integers with $g_i \geq 4$ for every i . Let $\mathcal{G}(g_1, \dots, g_k)$ be the family of cycle chains with k cycles C_i of length g_i where $k \geq 3$ and $1 \leq i \leq k$. The maximum of the subpath number in the family $\mathcal{G}(g_1, \dots, g_k)$ is attained only for a kink chain, and the minimum only for a linear or an almost linear chain.

Proof. We consider the formula for $pn(G(S))$ from Theorem 12. First, notice that a_1 and a_k are the same for all chains in the family $\mathcal{G}(g_1, \dots, g_k)$, hence they do not contribute to the difference of the subpath number for different chains in the family. Next, denote $p_{i,j} = a_i b_j + b_i a_j$ and $q_{i,j} = a_i a_j + b_i b_j$. It is easily verified that $p_{i,j} + q_{i,j} = (a_i + b_i)(a_j + b_j)$. Hence, for $2 \leq i \leq k - 1$ we have $p_{i,j} + q_{i,j} = (g_i - 2)(g_j - 2)$. Also, since $a_1 = b_1$, we have $p_{1,j} = a_1(g_j - 2)$ for $2 \leq j \leq k - 1$. Similarly, $p_{i,k} = (g_i - 2)a_k$ for $2 \leq i \leq k - 1$. Finally, $p_{1,k} = 2a_1 a_k$. This implies that the expressions from the first two lines of the formula are the same for all chains in the considered family.

The expression from the fourth line is also the same for all chains in the family, so let us consider the expressions from the third and fifth line of the formula. Denote $r_i = \binom{a_i+1}{2} + \binom{b_i+1}{2}$ and $s_i = (a_i + 1)(b_i + 1)$. For $2 \leq i \leq k - 1$, it is easily verified that $r_i + s_i = g_i(g_i - 1)/2$. Hence, the difference of the subpath number for a pair of chains in the family $\mathcal{G}(g_1, \dots, g_k)$ is determined by the expression

$$\sum_{i=2}^{k-1} r_i(1 + i(k - i + 1)) - (k + 1) = \sum_{i=2}^{k-1} r_i(-i^2 + i(k + 1) - k).$$

It is easily seen that $-i^2 + i(k + 1) - k$ is strictly positive for $2 \leq i \leq k - 1$. Hence, the subpath number is maximized (resp. minimized) by the chain with maximum (resp. minimum) possible r_i for every $2 \leq i \leq k - 1$. Notice that $g_i = a_i + b_i + 2$ implies

$$r_i = \binom{a_i+1}{2} + \binom{g_i - a_i - 1}{2}.$$

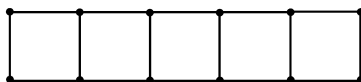


Fig. 3. The figure shows the only 4-chain with $k = 5$.

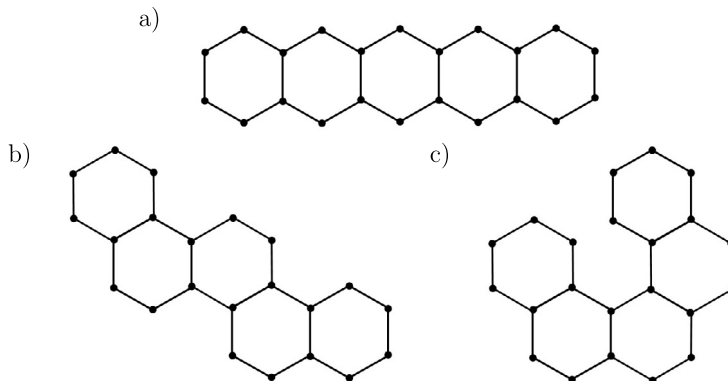


Fig. 4. The figure shows hexagonal chains: a) the linear chain L_5 , b) the zig-zag chain Z_5 and c) the helicene H_5 .

Now, it is easily verified that r_i is minimum for $a_i = \lfloor g_i/2 \rfloor - 1$ or $a_i = \lceil g_i/2 \rceil - 1$. And r_i is maximum for a_i minimum or maximum possible which is equivalent with $a_i = 1$ or $b_i = 1$. ■

Let us next show that the results of Theorem 12 and Corollary 13 can be applied to various families of chains frequently studied in the literature. For an integer $g \geq 4$, a g -chain is a cycle chain in which every cycle is of the length g . For $g = 4$ there exists only one 4-chain on a given number of cycles k . This chain is illustrated by Fig. 3. The 4-chain is also called the ladder graph, and a cycle of length 4 in a ladder graph will be called a square. Using Theorem 12 the subpath number of the ladder graph is easily established.

Corollary 14. For the ladder graph G with k squares we have

$$pn(G) = 36 \cdot 2^k - \frac{1}{3}k^3 - 4k^2 - \frac{56}{3}k - 33.$$

The next interesting value of g is $g = 6$. Cycle chains with $g = 6$ are also called hexagonal chains and the cycles of length 6 are called hexagons. This family of chains has applications in chemistry, since such graphs model unbranched catacondensed benzenoids (see [1] for example). For $k \leq 2$ there is only one hexagonal chain with k hexagons. For $k \geq 3$, there exist multiple such chains, raising the natural question of which chain maximizes or minimizes a particular quantity. This question was investigated for many graph quantities, among them for the Wiener index [4] and the subtree number [18]. In particular, it is established that for a hexagonal chain G with k hexagons, it holds that $W(H_k) < W(G) < W(L_k)$ and $N(L_k) < N(G) < N(H_k)$, provided that $G \notin \{L_k, H_k\}$.

Regarding the same question of extremal hexagonal chains with respect to the subpath number, Corollary 13 provides the answer. Namely, notice that the linear hexagonal chain with k cycles is unique, denote it by L_k . On the other hand, there are many kink hexagonal chains with k hexagons, for example the zig-zag chain Z_k in which $a_{2i} = 1$ and $a_{2i+1} = 3$ for $2 \leq 2i < 2i + 1 \leq k - 1$ or the helicene H_k in which $a_i = 1$ for $2 \leq i \leq k - 1$. These chains are illustrated by Fig. 4. Corollary 13 now yields the following result.

Corollary 15. Let G be a hexagonal chain with k hexagons. Then

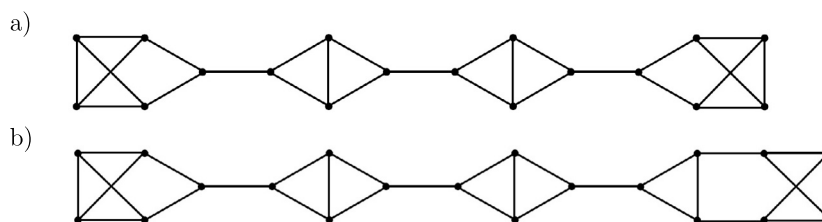
$$144 \cdot 2^k - \frac{5}{3}k^3 - 21k^2 - \frac{265}{3}k - 141 \leq pn(G) \leq 144 \cdot 2^k - \frac{3}{2}k^3 - \frac{43}{2}k^2 - 88k - 141.$$

The lower bound is attained if and only if $G = L_k$, and upper bound is attained only by a kink chain G .

Notice that the subpath number differs from the subtree number in the sense that any kink chain maximizes it, and not only the helicene H_k . For example, the zig-zag chain Z_k also maximizes the subpath number among chains with k hexagons.

5. Concluding remarks and further work

As a direction for further work, it would be interesting to investigate the subpath number for specific graph classes. Lemma 9 implies that removing edges from a graph decreases the subpath number, and that adding edges increases the subpath number. These two processes lead to a graph which contains a vertex of degree 1 or vertex of maximum possible degree, respectively. Hence, the result of Theorem 10 that trees are minimal and the complete graph maximal with respect to the subpath number is implied. An interesting question that arises is what happens when the degree is prescribed?

Fig. 5. The graphs: a) L_{18} , b) L_{20} .

The most simple case of prescribed degrees is a r -regular graph G in which all vertices have the same degree r . If $r = 2$, then G is the cycle for which the subpath number is established. So, the first non-trivial case is $r = 3$, i.e. cubic graphs. Let us define graphs for which we believe are extremal in this family.

Assume $n \geq 10$ is even. If $n = 4q - 2$ for some integer q , then the graph L_n is obtained from $(n - 10)/4$ copies of $K_4 - e$ by adding $(n - 10)/4 - 1$ edges so that the resulting graph is connected and has maximum degree 3. Thus, we obtain a path like structure which is then ended on both sides by a pendant block, each on 5 vertices. Otherwise, if $n = 4q$, then L_n is obtained from $(n - 12)/4$ copies of $K_4 - e$ connected into a path like structure in the same way as before, but now the structure is ended with two pendant blocks, one on 5 vertices and the other on 7 vertices. The graphs L_n are illustrated by Fig. 5, they are an example of graphs with rather small subpath number among cubic graphs on n vertices, but they are not minimal [2]. Hence, we propose the following problem.

Problem 16. In the class of cubic graphs on n vertices, characterize graphs which minimize the subpath number.

The cubic graphs which maximize the subpath number seem to be more elusive. Interestingly, for $n = 8$ it is the Möbius strip and for $n = 10$ it is the Petersen graph. For other small values of $n \geq 12$, the maximal graphs are rather symmetrical, though not always transitive, and they tend to have a relatively small diameter, though not always the smallest possible. Hence, we propose the following problem.

Problem 17. Find the graphs which maximize the subpath number in the class of cubic graphs on n vertices.

Finally, based on our investigation of bipartite graphs, where the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lfloor n/2 \rfloor}$ maximizes the subpath number, we believe that the same graph maximizes the subpath number in a wider family of graphs also, namely the triangle-free graphs. We state it as the following conjecture.

Conjecture 18. Among triangle-free graphs on n vertices the maximum value of the subpath number is attained only by the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lfloor n/2 \rfloor}$.

Final remarks on the above problems. Problem 16 was originally stated (in the draft, see arXiv:2503.02683 [math.CO]) as a conjecture claiming that L_n is the only graph which minimizes the subpath number. However, this was soon disproved by Stijn Cambie [2], nevertheless it may still be of interest to investigate which graphs are indeed optimal, so we state it as a problem. In addition, we were informed that Conjecture 18 has a positive answer by both Stijn Cambie [2] and Sean Longbrake [9].

Acknowledgements

This work is partially supported by Slovak research grants VEGA 1/0069/23, VEGA 1/0011/25, APVV-22-0005 and APVV-23-0076, by Slovenian Research and Innovation Agency ARIS program P1-0383, project J1-3002, by bilateral Croatian-Slovenian project BI-HR/25-27-004 and the annual work program of Rudolfovo, by Project KK.01.1.1.02.0027 co-financed by the Croatian Government and the European Union through the European Regional Development Fund - the Competitiveness and Cohesion Operational Programme, by the Croatian Science Foundation under project number HRZZ-IP-2024-05-2130, by Croatian Ministry of Science, Education and Youth through the bilateral Croatian-Slovenian project 2025-26 and by the Key Scientific and Technological Project of Henan Province, China (grant nos. 252102521077, 252102240118, 242102521023), China Henan International Joint Laboratory for Multidimensional Topology and Carcinogenic Characteristics Analysis of Atmospheric Particulate Matter (PM2.5).

Data availability

No data was used for the research described in the article.

References

- [1] C.S. Anstöter, N. Bašić, P.W. Fowler, T. Pisanski, Catacondensed chemical hexagonal complexes: a natural generalisation of benzenoids, *Croat. Chem. Acta* 93 (4) (2020) 289–301.

- [2] S. Cambie, private communication, 4.3.2025.
- [3] É. Czabarka, L. Székely, S. Wagner, The inverse problem for certain tree parameters, *Discrete Appl. Math.* 157 (15) (2009) 3314–3319.
- [4] I. Gutman, Wiener numbers of benzenoid hydrocarbons: two theorems, *Chem. Phys. Lett.* 136 (1987) 134–136.
- [5] R. Kirk, H. Wang, Largest number of subtrees of trees with a given maximum degree, *SIAM J. Discrete Math.* 22 (2008) 985–995.
- [6] J. Li, K. Xu, T. Zhang, H. Wang, S. Wagner, Maximum number of subtrees in cacti and block graphs, *Aequ. Math.* 96 (2022) 1027–1040.
- [7] M. Knor, R. Škrekovski, A. Tepeh, Mathematical aspects of Wiener index, *Ars Math. Contemp.* 11 (2016) 327–352.
- [8] M. Knor, R. Škrekovski, A. Tepeh, Selected topics on Wiener index, *Ars Math. Contemp.* (2024) #P4.07.
- [9] S. Longbrake, private communication, 5.3.2025.
- [10] L.A. Székely, H. Wang, On subtrees of trees, *Adv. Appl. Math.* 34 (2005) 138–155.
- [11] L.A. Székely, H. Wang, Binary trees with the largest number of subtrees, *Discrete Appl. Math.* 155 (2007) 374–385.
- [12] H. Wiener, Structural determination of paraffin boiling points, *J. Am. Chem. Soc.* 69 (1947) 17–20.
- [13] K. Xu, J. Li, Z. Luo, Comparative results between the number of subtrees and Wiener index of graphs, *RAIRO Oper. Res.* 56 (2022) 2495–2511.
- [14] K. Xu, J. Li, H. Wang, The number of subtrees in graphs with given number of cut edges, *Discrete Appl. Math.* 304 (2021) 283–296.
- [15] M. Yamamoto, Approximately counting paths and cycles in a graph, *Discrete Appl. Math.* 217 (2009) 381–387.
- [16] Y. Yang, X.-X. Li, M.-Y. Jin, L. Li, H. Wang, X.-D. Zhang, Enumeration of subtrees and BC-subtrees with maximum degree no more than k in trees, *Theor. Comput. Sci.* 892 (2021) 258–278.
- [17] Y. Yang, H. Liu, H. Wang, X.-D. Zhang, C.P. Chen, On enumerating algorithms of novel multiple leaf-distance granular regular α -subtrees of trees, *Inf. Comput.* 289 (2022) 104942.
- [18] Y. Yang, H.B. Liu, H. Wang, A.S. Deng, C. Magnant, On algorithms for enumerating subtrees of hexagonal and phenylene chains, *Comput. J.* 60 (2017) 690–710.
- [19] X.M. Zhang, X.D. Zhang, D. Gray, H. Wang, The number of subtrees of trees with given degree sequence, *J. Graph Theory* 73 (2013) 280–295.