Hitting subgraphs in $P_4$-tidy graphs

Boštjan Brešar$^{a,b,*}$, Tim Kos$^{b,*}$, Rastislav Krivoš-Belluš$^{c†}$
Gabriel Semanišin$^{c‡}$

March 1, 2019

$a$ Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia
$b$ Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia
$c$ Institute of Computer Science, P. J. Šafárik University, Jesenná 5, 041 54 Košice, Slovakia

Abstract

Vertex cover number, which is one of the most basic graph invariants, can be viewed as the smallest number of vertices that hit (or that belong to) every subgraph $K_2$ in a graph $G$. In this paper, we consider the next two smallest cases of connected graphs, which are the path $P_3$ and the cycle $C_3$; the problem is to minimize the set of vertices that hit every subgraph $H$ in $G$, where $H$ is one of the two graphs. We focus on computational aspects of these two variations of the vertex cover number. We show that both problems are NP-complete even in the class of graphs that have no induced cycles of length $t$, where $t \in \{4, \ldots, 8\}$ (in fact, in the $P_3$-hitting case, the problem is NP-complete in even smaller family of graphs, where also triangles are forbidden). On the positive side, we prove a complementary result that in graphs with no induced $P_4$ (the so-called cographs) both problems become linear. Moreover, we consider a generalization of cographs — $P_4$-tidy graphs, for which we establish an efficient algorithm for both problems.

Keywords: path vertex cover number, decycling number, $P_4$-tidy graph

AMS subject classification (2000): 05C05, 05C38, 05C85, 68R10.

*Supported in part by the Slovenian Research Agency (ARRS) under the grant P1-0297.
†The research was partially supported by Slovak grant VEGA 1/0142/15.
‡The research was partially supported by Slovak grants VEGA 1/0142/15, APVV-15-0091 and Technicom ITMS 313011D232.
1 Introduction

Given a property $\mathcal{P}$ of graphs, the variety of problems of determining the minimum subset $S$ of vertices of a graph $G$ such that $G - S$ possesses the property $\mathcal{P}$ is known as the Vertex Deletion Problem (with respect to $\mathcal{P}$). One of the most famous such problems is obtained when $\mathcal{P}$ is the property of having no edges, which yields the Vertex Cover Problem (and in the complement independence number $\alpha(G)$ arises). In this paper, we deal with another two small cases of graphs, namely $C_3$ and $P_3$, which yield two vertex deletion problems when we forbid any one of these graphs as a subgraph in $G - S$. We call the corresponding problems $C_3$-Hitting Problem and $P_3$-Hitting problem, to be consistent with the initial approach. Anyway, closely related concepts have been considered earlier, in particular, two of them are Feedback Vertex Set Problem and $k$-Path Vertex Cover Problem. In the first case the required property is to be acyclic, and in the second case to have no $P_k$. The first problem is related to the invariant decycling number introduced in \cite{2} and the second one to the $k$-path vertex cover number introduced in \cite{5}.

Given a graph $G$, the decycling number $\nabla(G)$ is the cardinality of the smallest set $S$ of vertices in $G$ such that $G - S$ contains no cycle. The invariant was introduced in \cite{2}, while this concept was studied much earlier while searching for the so-called feedback vertex set (i.e., the set of vertices that hit all the cycles of a graph) \cite{12}, whose cardinality is bounded by a given integer. The feedback vertex set appeared already in the context of the well-known Erdős-Pósa theorem \cite{11}, and the corresponding problem (which is exactly the decision version of the decycling number) appeared on the famous Karp’s list of NP-complete problems \cite{20}.

In another related concept we are interested in sets of vertices that hit all paths of a prescribed given size. A set $D$ of vertices in a graph $G$ is called a dissociation set if the subgraph induced by vertices of $D$ has maximum degree at most 1. The cardinality of a maximum dissociation set $D$ in a graph $G$ is called the dissociation number of $G$, and is denoted by $\text{diss}(G)$. The dissociation number was introduced by Papadimitriou and Yannakakis \cite{26} in relation with the complexity of the so-called restricted spanning tree problem. A dual (and more general) concept is the $k$-path vertex cover, which was introduced in \cite{5} and studied in several papers \cite{4, 15}; it is defined as a set $S$ of vertices in $G$ such that $G - S$ does not contain any path $P_k$. The corresponding invariant, the $k$-path vertex cover number of an arbitrary graph $G$, is denoted by $\psi_k(G)$. If $k$-path vertex cover $S$ of $G$ is of cardinality $\psi_k(G)$, then we call $S$ minimum $k$-path vertex cover of
Note that dissociation sets are complements of 3-path vertex covers of $G$, and so $\text{diss}(G) = |V(G)| - \psi_3(G)$. The decision version of the $k$-path vertex cover number is NP-complete [5], moreover, in the case $k = 3$ it is NP-complete even in bipartite graphs which are $C_4$-free and have maximum degree 3 [3]; cf. also [25] for further strengthening of this result and [19] for an approximation algorithm. Some variations of the problem were already studied as well (see e.g. [22]).

The third concept is in a sense related to feedback vertex sets and also to 3-path vertex covers. A set $S$ of vertices in a graph $G$ is called $C_3$-hitting set if a graph induced by vertices $V(G) \setminus S$ contains no triangles. The $C_3$-hitting number of a graph $G$, denoted by $\nabla_3(G)$, is the smallest cardinality of a $C_3$-hitting set of a graph $G$. A set is minimum $C_3$-hitting set if it is a $C_3$-hitting set of size $\nabla_3(G)$.

In this paper, we will focus on a class of graphs that widely generalizes the class of cographs, and are called $P_4$-tidy graphs. This class of graphs contains several other graph families having few $P_4$’s, such as $P_4$-sparse [13], $P_4$-lite [16], $P_4$-extendible [18] and $P_4$-reducible graphs [17]. A graph class generalizing all of them is the class of $P_4$-tidy graphs [14]. They are also high in the hierarchy of the classes of graphs recognizable by modular decomposition (cf. [9] for modular decomposition and [14] for representation of modular graphs with respect to $P_4$-tidy graphs). To determine $C_3$-hitting and $P_3$-hitting sets with respect to modular decomposition and in $P_4$-tidy graphs, will thus yield a solution for these problems in several other interesting classes of graphs.

The paper is organized as follows. In the next section we establish the basic notation used throughout the paper. In Section 3, NP-completeness results are proven. More precisely, we prove that for any fixed integer $k \geq 3$, $C_k$-HITTING PROBLEM is NP-complete for $C_t$-free graphs, where $3 \leq t \leq 8$ and $t \neq k$, and a similar result is proven for the $P_k$-hitting problem. In Section 4, we consider modular decomposition and $P_4$-tidy graphs, and present linear time algorithms that find a minimum $C_3$-hitting set and a minimum $P_3$-hitting set in such graphs.

2 Notation and preliminaries

This section is devoted to notational and preliminary issues. A non-trivial graph is a graph on at least two vertices. For each positive integer $n$, $K_n$ and $P_n$ are, respectively, the complete graph and the path with $n$ vertices. For $n \geq 3$, $C_n$ denotes the cycle with $n$ vertices. Given a graph $G$, $\overline{G}$ denotes its
complement. The subgraph induced by a set \( S \) of vertices of \( G \) is denoted by \( G[S] \). We write \( G - S \) for the subgraph induced by \( V(G) \setminus S \). If \( S = \{v\} \) we simply write \( G - v \). Let \( H \) be a graph. If a graph \( G \) contains no induced subgraph isomorphic to \( H \), we say that \( G \) is \( H \)-free.

The neighborhood of a vertex \( v \in V(G) \) is the set \( N_G(v) = \{ u \in V(G) : uv \in E(G) \} \), while neighborhood of a set \( X \subseteq V(G) \) is defined as \( N_G(X) = \bigcup_{v \in X} N_G(v) \). The closed neighborhood of a vertex \( v \in V(G) \) is the set \( N_G[v] = N(v) \cup \{v\} \), while closed neighborhood of a set \( X \subseteq V(G) \) is defined as \( N_G[X] = \bigcup_{v \in X} N_G[v] \).

Two vertices \( u \) and \( v \) in \( G \) are called true (resp. false) twin vertices if \( N[u] = N[v] \) (resp. \( N(u) = N(v) \)). A vertex \( v \in V(G) \) is called a true (resp. false) twin vertex if there exists \( u \in V(G) \setminus \{v\} \), such that \( u \) and \( v \) are true (resp. false) twin vertices.

The union of graphs \( G_1 \) and \( G_2 \), denoted by \( G_1 \cup G_2 \), is the graph \( G \) with \( V(G) = V(G_1) \cup V(G_2) \) and \( E(G) = E(G_1) \cup E(G_2) \). The join of graphs \( G_1 \) and \( G_2 \), denoted by \( G_1 + G_2 \), is the union of the graphs \( G_1 \) and \( G_2 \) together with all edges joining \( V(G_1) \) and \( V(G_2) \).

A set \( I \subseteq V(G) \) is a independent set (or stable set), if no two vertices in \( I \) are adjacent. The independence number of a graph \( G \), denoted by \( \alpha(G) \), is the cardinality of the largest independent set of \( G \). A set is maximum independent set of \( G \) if is an independent set of size \( \alpha(G) \).

We follow with some preliminary observations about the considered invariants. It is clear that \( \nabla_3(G) \leq \nabla(G) \) if \( G \) is an arbitrary graph, and it is easy to see that \( \nabla_3(G) = \nabla(G) \) in chordal graphs \( G \). Obviously, the difference \( \nabla(G) - \nabla_3(G) \) can be arbitrary large (for example take the complete bipartite graphs). It is known that dissociation number and thus also \( \psi_3(G) \) can be determined in polynomial time for arbitrary chordal graph \( G \) [8]. A simple modification of this approach also results in a polynomial algorithm for the weighted version. Hence the following question is also interesting.

**Question 1** What is the computational complexity of determining \( \nabla(G) \) (resp. \( \nabla_3(G) \)) for arbitrary chordal graph \( G \)? Similar questions can be posed for the weighted versions.

Such questions are very interesting in the context of graph colorings. There are some papers dealing with the relationship between cycle lengths and chromatic number of graphs (cf. [24]). On the other hand, different families of graphs have been studied in the context of their \( k \)-path vertex cover. In particular, several upper bounds were established for some interesting families, like outerplanar graphs, trees [5], some graph products
[15, 23] and cacti [28]. It would be interesting to consider in a similar way the invariants $\nabla$ and $\nabla_3$.

**Observation 2.1** If $G$ is $k$-colorable, then $\nabla_3(G) \leq \frac{k-2}{k}|V(G)|$. In particular, for 3-chromatic graphs $G$ we have $\nabla_3(G) \leq \frac{1}{3}|V(G)|$, and for planar graphs $G$, $\nabla_3(G) \leq \frac{1}{2}|V(G)|$.

Note that the bounds can be very far from exact values. For example take the graph $G$ obtained from the set of $n$ triangles, and identify $n$ vertices, one from each triangle. Clearly, $G$ is 3-chromatic and planar, and $\nabla_3(G) = 1$.

### 3 NP-completeness

The 2-subdivision of a graph $G$ is obtained from $G$ by inserting two new vertices in every edge, that is, replacing each edge $uv \in E(G)$ with $P_4 : uxyv$. A graph is a 2-subdivision if it is the 2-subdivision of some graph.

**Vertex Cover Problem**

| Input: A connected graph $G$ and a positive integer $k$. |
| Question: Does $G$ have a vertex cover of size $\leq k$? |

The following result was (implicitly) proven by Poljak [27]. For completeness reasons we give (sketch of the) proof.

**Theorem 3.1** Vertex Cover Problem is NP-complete for 2-subdivision graphs.

**Proof.** It is known that Vertex Cover Problem is in NP-complete. To prove that Vertex Cover Problem is NP-complete for 2-subdivision graphs we use a polynomial reduction from Vertex Cover Problem for any graph. Let $G'$ be the 2-subdivision of $G$. Let $k$ be an integer, $k \geq 0$. It not hard to see that $G$ has a vertex cover number of size at most $k$, if and only if $G'$ has a vertex cover number of size at most $k + m$, where $m$ is the number of edges of $G$. $\square$

Since 2-subdivision graphs are $C_t$-free, where $3 \leq t \leq 8$, we get the next result:

**Corollary 3.2** Vertex Cover Problem is NP-complete for $C_t$-free graphs, where $3 \leq t \leq 8$. 
Now, we consider the $k$-path vertex cover problem.

$k$-path Vertex Cover Problem

**Input:** A connected graph $G$ and positive integers $k \geq 2$ and $t$.

**Question:** Does $G$ have a $k$-path vertex cover of size $\leq t$?

**Theorem 3.3** For any fixed integer $k \geq 2$, $k$-path Vertex Cover Problem is NP-complete for $C_t$-free graphs, where $3 \leq t \leq 8$.

**Proof.** It can be proven the same way as [5, Theorem 1], by using a polynomial reduction from Vertex Cover Problem. Let $G$ be a $C_t$-free graph. Let $G'$ be a graph obtained from $G$ by attaching a path with $\left\lfloor \frac{k-1}{2} \right\rfloor$ new vertices to every vertex of $G$. Clearly, $G'$ is $C_t$-free, and $G$ has a vertex cover of size at most $t$, if and only if $G'$ has a $k$-path vertex cover of size at most $t$. \hfill \Box

$C_k$-hitting Problem

**Input:** A connected graph $G$ and positive integers $k \geq 3$ and $t$.

**Question:** Does $G$ have a $C_k$-hitting set of size $\leq t$?

**Theorem 3.4** For any fixed integer $k \geq 3$, $C_k$-hitting Problem is NP-complete.

**Proof.** First we will prove that $C_k$-hitting Problem is in NP. Let $D$ be a set of $t$ vertices from a graph $G$. To check if $D$ is a $C_k$-hitting set of $G$, we have to check if $G - D$ has no $C_k$ as a subgraph. Or equivalently, for each $k$-tuple of vertices in $G - D$ we can check if it induces a Hamiltonian graph. There are $O(n^k)$ such $k$-tuples. To check if a graph induced by $k$-tuple is Hamiltonian, can be done in $O(2^k)$, which is constant as we are looking at a function of $n$ (note that $k$ is fixed and $n$ can be arbitrarily large; in fact, we are particularly interested in small values of $k$). Hence, to check if $D$ is a $C_k$-hitting set of $G$ can be done in $O(n^k)$ time. It follows that $C_k$-hitting Problem is in NP.

To prove that $C_k$-hitting Problem is NP-complete we use a polynomial reduction from Vertex Cover Problem. Let $G = (V(G), E(G))$ be any graph and let $G' = (V(G'), E(G'))$ be the graph constructed from $G$ as follows. For each $e = v_1v_2 \in E(G)$ we add vertices $V_e = \{v_e^1, \ldots, v_e^{k-2}\}$ and
create a cycle \( \{v_1, v_2, v_{e_1}^1, \ldots, v_{e_{k-2}}^k, v_1\} \). (On each edge in \( G \) we attach a path and form a cycle of size \( k \).)

Next, we prove, that \( G \) has a vertex cover of size at most \( t \), if and only if \( G' \) has a \( C_k \)-hitting set of size at most \( t \).

Let \( S \) be a vertex cover of \( G \) with \( |S| = k \). Clearly, \( S \) is also a \( C_k \)-hitting set of \( G' \). Now, let \( S' \) be \( C_k \)-hitting set of \( G' \) with \( |S'| = k \). Let \( S \subseteq V(G) \) be defined as follows:

- if \( v \in S' \) and \( v \in V(G) \), then add \( v \) to \( S \).
- if \( v \in S' \) and \( v \notin V(G) \), then \( v = v_e^i \) where \( e = v_1v_2 \in E(G) \) and \( 1 \leq i \leq k - 2 \). We add to \( S \) \( v_1 \) or \( v_2 \).

Clearly, \( S \) is a vertex cover of \( G \) with \( |S| \leq k \). \( \square \)

The same way we can prove following:

**Corollary 3.5** For any fixed integer \( k \geq 3 \), \( C_k \)-hitting Problem is NP-complete for \( C_t \)-free graphs, where \( 3 \leq t \leq 8 \) and \( t \neq k \).

Note, that if \( G \) is \( C_k \)-free graph, then \( \nabla_k(G) = 0 \).

## 4 \( P_4 \)-tidy graphs

We start this section by formally defining \( P_4 \)-tidy graphs. Let \( U \) be a subset of vertices inducing a \( P_4 \) in \( G \). A partner of \( U \) is a vertex \( v \in G - U \) such that \( U \cup \{v\} \) induces at least two \( P_4 \)s in \( G \). A graph \( G \) is \( P_4 \)-tidy if any \( P_4 \) has at most one partner. It is known that the class of \( P_4 \)-tidy graphs is self-complementary and hereditary [14].

Modular decomposition is a decomposition of a graph into subsets of vertices by using two operations - join and union (see e.g. [6]). Clearly, if a graph \( G \) (resp. \( \bar{G} \)) is not connected, it can be obtained by disjoint union (resp. join) of two non-empty graphs. If \( G \) and its complement are connected, we say that \( G \) is modular. Given a graph family \( \mathcal{F} \), we denote by \( M(\mathcal{F}) \) the family of modular graphs in \( \mathcal{F} \). In particular, if \( \mathcal{F} \) is the family of cographs, i.e., the graphs that do not contain \( P_4 \) as an induced subgraph, it is known that \( M(\mathcal{F}) \) only contain trivial graphs.

Now, given a graph class \( \mathcal{F} \), if the problem of \( G_1 \cup G_2 \) and \( G_1 + G_2 \) can be solved in polynomial (linear) time from known results of \( G_1 \) and \( G_2 \), and if the problem is solvable in polynomial (linear) time for graphs in \( M(\mathcal{F}) \), then the problem is solvable in polynomial (linear) time for graph class \( \mathcal{F} \).
Non-trivial modular $P_4$-tidy graphs are the graphs $C_5$, $P_3$ and $\overline{P}_5$ and the spider and quasi-spider graphs with $P_4$-tidy heads defined below [14]. (Trivial modular $P_4$-tidy graphs are $K_1$ and the empty graph.)

A graph $G$ is a spider graph if its vertex set can be partitioned into three sets $S$, $C$ and $H'$ ($H'$ possible empty), where $S$ is a stable set, $G[C]$ is a clique, $|S| = |C| \geq 2$, $G[H']$ is completely joined to $G[C]$, and no vertex of $G[H']$ is adjacent to a vertex in $G[S]$. Moreover, if $S = \{s_1, \ldots, s_r\}$ and $C = \{c_1, \ldots, c_r\}$ one of the following conditions must holds:

1. **thin spider**: $s_i$ is adjacent to $c_j$ if and only if $i = j$.
2. **thick spider**: $s_i$ is adjacent to $c_j$ if and only if $i \neq j$.

The size of $C$ (and $S$) is called the weight of $G$ and the set $H'$ in the partition is called the head of the spider. A thin (resp. thick) spider graph $G$ with the partition $S,C,H'$ will be denoted $G = (S,C,H)_{thin}$ (resp. $G = (S,C,H)_{thick}$), where $H = G[H']$ (induced subgraph of $G$ with vertices in $H'$).

Notice that if $|S| = |C| = 2$ thin and thick spider graphs are isomorphic. In what follows we consider thick spider graphs with $|S| = |C| \geq 3$.

Now, if $G = (S,C,H)_{thin}$ is a thin spider graph, the graph obtained by adding a false or true twin to one vertex $v \in S \cup C$ is called thin quasi-spider graph. Without loss of generality we assume that the vertex that get the twin is $s_r \in S$ (resp. $c_r \in C$). A false (true) twin vertex will be denoted by $s'_r$ or $c'_r$, respectively. We denote by $G = (S',C,H)_{thin}$ (resp. $G = (S,C',H)_{thin}$) the thin quasi-spider graph obtaining from a thin spider graph $(S,C,H)_{thin}$ by adding a false twin to one vertex in $S$ (resp. $C$). The same way, we denote by $G = (S',C,H)_{thin}$ (resp. $G = (S,C',H)_{thin}$) the thin quasi-spider graph obtaining from a thin spider graph $(S,C,H)_{thin}$ by adding a true twin to one vertex in $S$ (resp. $C$). Analogically, thick quasi-spider graphs are defined. We will write $G = (S,C,H)$ if $G$ is any spider or any quasi-spider.

It is known that the partition for spider and quasi-spider graphs is unique and their recognition as well as modular partition can be performed in linear time (see [14]).

### 4.1 $P_3$-hitting set

For computing $\psi_3(G)$ of a $P_4$-tidy graph $G$ we will use the modular decomposition.

For the non-trivial modular $P_4$-tidy graphs $P_3$, $\overline{P}_3$ and $C_5$ it holds:
Observation 4.1 $\psi_3(P_3) = 1, \psi_3(C_3) = \psi_3(\overline{P_3}) = 2$.

The following theorem determines the 3-path vertex cover number of the disjoint union and the join of two graphs.

**Theorem 4.2** Let $G_1$ and $G_2$ be arbitrary graphs. Then

(i) $\psi_3(G_1 \cup G_2) = \psi_3(G_1) + \psi_3(G_2)$;

(ii) $\psi_3(G_1 + G_2) = \min\{|G_1| + |G_2| - 2, \psi_3(G_1) + |G_2|, |G_1| + \psi_3(G_2)\}$.

**Proof.** The first assertion is obvious. For the second one, either all vertices from one of graphs $G_1, G_2$ are in the $P_3$-hitting set for 3-path vertex cover (and the other graph is 3-path vertex covered by hitting set) or at both graphs $G_1, G_2$ at most one vertex is not in the hitting set $H$ (otherwise there will exist a $P_3$, which is not hit). □

Observation 4.3 Let $G = (S, C, H)$ be a thin (quasi-)spider. Then there exists $c \in C$ with $|S \cap N(c)| = 1$.

**Lemma 4.4** If $H$ is a $P_3$-hitting set of a (quasi-)spider $G = (S, C, H)$, then $|C \setminus H| \leq 2$.

**Proof.** Because $C$ is a clique, it is clear that at least $|C| - 2$ vertices have to be in hitting set $H$. This condition holds also for the case when $G$ is a quasi-spider with false twins in $C$ (clique without one edge). □

**Lemma 4.5** Let $G = (S, C, H)$ be a thin (quasi-)spider with $V(H) = \emptyset$. Then $\psi_3(G) = |C| - 1$.

**Proof.** Let $c \in C$ be a vertex from Observation 4.3 and let $H = C \setminus \{c\}$. Clearly $H$ is a $P_3$-hitting set of $G$. Hence, $\psi_3(G) \leq |H| = |C| - 1$.

Suppose, there exists $P_3$-hitting set $H'$ of $G$ with $|H'| \leq |C| - 2$. By Lemma 4.4, $|C \setminus H'| \leq 2$ and $|H'| \geq |C| - 2$. Hence, $|H'| = |C| - 2$ and let $H' = C \setminus \{c', c''\}$ for some $c', c'' \in C, c' \neq c''$. If $c'$ and $c''$ are false twins, then there exists $s \in S \cap (N(c') \cap N(c''))$. Then $\{c', s, c''\}$ induces $P_3$ in $G$, a contradiction. If $c'$ and $c''$ are not false twins, then $c'c'' \in E(G)$. Hence, $\{c', c'', s\}$, where $s \in S \cap N(c'')$, induces $P_3$ in $G$, a contradiction. This implies $\psi_3(G) = |C| - 1$. □

9
**Lemma 4.6** Let $G = (S, C, H)$ be a (quasi-)spider, which is not a thin (quasi-)spider with $V(H) = \emptyset$. Then $\psi_3(G) = |C| + \psi_3(H)$.

**Proof.** Let $\mathcal{H}_P(H)$ be a minimum $P_3$-hitting set of $H$ and $\mathcal{H} = C \cup \mathcal{H}_P(H)$. Clearly $\mathcal{H}$ is a $P_3$-hitting set of $G$. Hence, $\psi_3(G) \leq |H| = |C| + \psi_3(H)$.

Suppose, there exists $P_3$-hitting set $\mathcal{H}'$ of $G$ with $|\mathcal{H}'| < |C| + \psi_3(H)$. Clearly, $|\mathcal{H}' \cap V(H)| \geq \psi_3(H)$. By Lemma 4.4, $|\mathcal{H}' \cap C| \geq |C| - 2$. Hence, $|\mathcal{H}'| \geq |C| - 2 + \psi_3(H)$. Let’s distinguish 2 cases:

- $|C \setminus \mathcal{H}'| = 1$
  
i.e. all but one vertex $c \in C$ from $C$ are in the hitting set $\mathcal{H}'$. Hence, $|\mathcal{H}'| = |C| - 1 + \psi_3(H)$, $|\mathcal{H}' \cap V(H)| = \psi_3(H)$ and $|\mathcal{H}' \cap S| = 0$. In order to not have an unhit $P_3$ with endvertices from $V(H) \cup S$ through $c$ it is easy to see that all but one vertex from $V(H) \cup (S \cap N(c))$ have to be in the hitting set $\mathcal{H}'$. This implies, $|V(H) \cup (S \cap N(c)) \setminus \mathcal{H}'| \leq 1$. On the other hand, $|V(H) \cup (S \cap N(c)) \setminus \mathcal{H}'| = |V(H) \setminus \mathcal{H}'| + |(S \cap N(c)) \setminus \mathcal{H}'| = |V(H)| - \psi_3(H) + |S \cap N(c)|$. However, $|V(H)| - \psi_3(H) + |S \cap N(c)| \geq 1$, a contradiction (contrary happens just in the case when $V(H) = \emptyset$ and $G$ is a thin (quasi-)spider - note that thick (quasi-)spiders have $r > 2$).

- $|C \setminus \mathcal{H}'| = 2$
  
i.e. two vertices $c', c'' (c' \neq c'')$ from $C$ are not in the hitting set $\mathcal{H}'$. So all vertices from $V(H) \cup (S \cap (N(c') \cup N(c''))) \setminus \mathcal{H}'$ (otherwise it would form $P_3$ with vertices $c', c''$ - either $c'c''v$ when $c', c''$ are not false twins, or $c'vc''$ when $c', c''$ are false twins). Hence, $|\mathcal{H}'| \geq |C| - 2 + |V(H)| + |S \cap (N(c') \cup N(c''))|$. Note, that if $|V(H)| \neq \emptyset$, then $|V(H)| > \psi_3(H)$. Since $G$ is not a thin spider with $V(H) = \emptyset$, $|C| - 2 + |V(H)| + |S \cap (N(c') \cup N(c''))| > |C| - 2 + \psi_3(H) + 1$. Hence, $|\mathcal{H}'| > |C| - 1 + \psi_3(H)$, a contradiction.

This implies $\psi_3(G) = |C| + \psi_3(H)$. \hfill \Box

By a combination of the previous two results we have:

**Theorem 4.7** Let $G = (S, C, H)$ be a (quasi-)spider. Then

$$\psi_3(G) = \begin{cases} |C| - 1 & \text{for thin (quasi-)spider} \\ |C| + \psi_3(H) & \text{with } V(H) = \emptyset \\ |C| - \psi_3(H) & \text{otherwise} \end{cases}.$$ 

**Theorem 4.8** The 3-path vertex cover number and the smallest 3-path vertex cover can be computed in linear time for $P_4$-tidy graphs.
Proof. By Observation 4.1 and Theorem 4.7, the 3-path vertex cover number can be computed in linear time for non-trivial modular $P_4$-tidy graphs. Combining this result and Theorem 4.2, we derive that 3-path vertex cover number can be computed in linear time for $P_4$-tidy graphs. □

4.2 $C_3$-hitting set

Let us start with two easy observations concerning non-trivial $P_4$-tidy graphs.

Observation 4.9 $\nabla_3(P_5) = 0, \nabla_3(C_5) = 0$ and $\nabla_3(P_5) = 1$.

Proposition 4.10

(i) If $H$ is a $C_3$-hitting set of a spider $G = (S, C, H)$, then $|C \setminus H| \leq 2$.

(ii) If $H$ is a $C_3$-hitting set of a quasi-spider $G = (S, C, H)$ then $|C \setminus H| \leq 3$. Moreover if $G$ does not contain false twins in $C$, then $|C \setminus H| \leq 2$.

Proof. Let $G = (S, C, H)$ be a (quasi-)spider. If $|C \setminus H| \geq 3$ then, since $C$ is complete, $G \setminus H$ contains a $C_3$, a contradiction. The only exception is when $G$ is a quasi-spider graph $(S, C^f, H)_{thick}$ or $(S, C^f, H)_{thin}$. In such a case $C \setminus H$ can contain false twins and one additional vertex of $C$. □

The next lemma describes the relationship between the independence number and $C_3$-hitting number of graphs with at least one edge.

Proposition 4.11

(i) For an arbitrary graph $G$, $\nabla_3(G) \leq |V(G)| - \text{diss}(G)$.

(ii) If $G$ has at least one edge, then $\nabla_3(G) \leq |V(G)| - \alpha(G) - 1$.

Proof. (i) If $D$ is the dissociation set, then it is triangle-free and therefore a minimum set hitting all triangles has cardinality at most $|V(G)| - \text{diss}(G)$.

(ii) Let $I$ be a maximum independent set of $G$. Let $u$ be an arbitrary vertex of $G$ not belonging to $I$. Then $I \cup \{u\}$ induces a triangle-free subgraph of $G$ and therefore $\nabla_3(G) \leq |V(G)| - \alpha(G) - 1$. □

The following theorem determines the $C_3$-hitting number of the disjoint union and the join of two graphs.

Theorem 4.12 Let $G_1$ and $G_2$ be arbitrary graphs. Then
\( \nabla_3(G_1 \cup G_2) = \nabla_3(G_1) + \nabla_3(G_2) ; \)

\( \nabla_3(G_1 + G_2) = \min \{|V(G_1)| - \alpha(G_1) + |V(G_2)| - \alpha(G_2), \nabla_3(G_1) + |V(G_2)|, |V(G_1)| + \nabla_3(G_2)\} . \)

**Proof.** The first assertions is obvious and therefore we shall concentrate to the second one. If one of the graphs is edgeless, say \( G_1 \), then is obvious that \( \nabla_3(G_1 + G_2) = \min \{|V(G_2)| - \alpha(G_2), |V(G_1)| + \nabla_3(G_2)\} \). Let us denote by \( \mathcal{H}_1, \mathcal{H}_2 \) and \( \mathcal{H} \) the \( C_3 \)-hitting sets of \( G_1, G_2 \) and \( G_1 + G_2 \), respectively. We claim that \( G_i - \mathcal{H}_i \) contains an edge. If \( \mathcal{H}_i \) is empty then this assertion follows immediately from the assumption. If \( \mathcal{H}_i \) is non-empty and the opposite is true then all the vertices of triangles of \( G_i \) belongs to \( \mathcal{H}_i \). However in such a case \( \mathcal{H}_i \) is not minimal because after removing an arbitrary vertex \( u \) from \( \mathcal{H}_i \) the resulting set is still \( C_3 \)-hitting set of \( G_i \). Therefore we have only three possibilities for the structure of \( \mathcal{H} \):

1. \( \mathcal{H} = (V(G_1) \setminus \mathcal{H}_\alpha(G_1)) \cup (V(G_1) \setminus \mathcal{H}_\alpha(G_2)) \), where \( \mathcal{H}_\alpha(G_1) \) and \( \mathcal{H}_\alpha(G_2) \) are maximum independent sets;
2. \( \mathcal{H} = \mathcal{H}_1 \cup V(G_2) \);
3. \( \mathcal{H} = \mathcal{H}_2 \cup V(G_1) \)

and \( \nabla_3(G) \) is then equal to the minimum cardinality of the sets described above. \( \square \)

**Observation 4.13** Let \( G = (S, C, H) \) be a (quasi-)spider. Then there exists \( c \in C \) with \( S \cap N(c) \) that does not contain an edge.

**Observation 4.14** Let \( G = (S, C^f, H) \) and let \( \mathcal{H} \) be a \( C_3 \)-hitting set of \( G \). Then \(|C \setminus \mathcal{H}| \leq 3 \) and \(|\mathcal{H}| \geq |C| - 3\).

**Observation 4.15** Let \( G = (S, C, H) \) with no false twins in \( C \) and let \( \mathcal{H} \) be a \( C_3 \)-hitting set of \( G \). Then \(|C \setminus \mathcal{H}| \leq 2 \) and \(|\mathcal{H}| \geq |C| - 2\).

The next lemmas determines the value \( \nabla_3 \) for (quasi-)spider graphs.

**Lemma 4.16** Let \( G = (S, C^f, H) \) (thin or thick), where \( V(H) = \emptyset \). Then

\[
\nabla_3(G) = \begin{cases} 
|C| - 3 & \text{for thin quasi-spider} \\
|C| - 2 & \text{for thick quasi-spider} 
\end{cases} .
\]

**Proof.** Let \( G \) be a thin quasi-spider graph and let \( c_1, c'_1, c_2 \in C \) be three vertices from \( C \), where \( c_1 \) and \( c'_1 \) are false twins. Let \( \mathcal{H} = C \setminus \{c_1, c'_1, c_2\} \). Clearly, \( \mathcal{H} \) is a \( C_3 \)-hitting set of \( G \). Hence, \( \nabla_3(G) \leq |\mathcal{H}| = |C| - 3 \). By...
Observation 4.14, $\nabla_3(G) \geq |C| - 3$. Hence, $\nabla_3(G) = |C| - 3$, if $G$ is a thin quasi-spider.

Let $G$ be a thick quasi-spider graph, let $c, c' \in C$ be false twins and let $H = C \setminus \{c, c'\}$. Clearly, $H$ is a $C_3$-hitting set of $G$. Hence, $\nabla_3(G) \leq |H| = |C| - 2$. By Observation 4.14, $\nabla_3(G) \geq |C| - 3$. Suppose, there exists $C_3$-hitting set $H'$ of $G$ with $|H'| = |C| - 3$. Hence, $H \subset C$ and $C \setminus H' = \{c_1, c_2, c_3\}$. Since $C$ is a clique without one edge, there exists $c_i$ and $c_j$ ($c_i \neq c_j$) that are connected. Because $G$ is a thick quasi-spider with weight $r > 2$, there exists vertex $s \in S \cap (N(c_i) \cap N(c_j))$. Hence, $c_i, c_j$ and $s$ induce a $C_3$, a contradiction. Hence, $\nabla_3(G) = |C| - 2$, if $G$ is a thick quasi-spider. □

Lemma 4.17 Let $G = (S, C, H)$ (thin or thick), where $H = \overline{K}_k, k \in \mathbb{N}$. Then $\nabla_3(G) = |C| - 2$.

**Proof.** Let $c, c' \in C$ be false twins and let $H = C \setminus \{c, c'\}$. Clearly, $H$ is a $C_3$-hitting set of $G$. Hence, $\nabla_3(G) \leq |H| = |C| - 2$. By Observation 4.14, $\nabla_3(G) \geq |C| - 3$. Suppose, there exists $C_3$-hitting set $H'$ of $G$ with $|H'| = |C| - 3$. Hence, $H \subset C$ and $C \setminus H' = \{c_1, c_2, c_3\}$. Since $C$ is a clique without one edge, there exists $c_i$ and $c_j$ ($c_i \neq c_j$) that are connected. Hence, $c_i, c_j$ and any $h \in V(H)$ induce a $C_3$, a contradiction. Hence, $\nabla_3(G) = |C| - 2$. □

Lemma 4.18 Let $G = (S, C, H)$ (thin or thick), where $H$ is not edgeless. Then $\nabla_3(G) = \min\{|C| - 2 + |V(H)| - \alpha(H), |C| + \nabla_3(H)\}$.

**Proof.** Let $H_{C_3}(H)$ be minimum $C_3$-hitting set and let $H_1 = C + H_{C_3}(H)$. Clearly, $H_1$ is a $C_3$-hitting set of $G$. Hence, $\nabla_3(G) \leq |H_1| = |C| + \nabla_3(H)$. Let $I$ be an maximal independent set of $H$, let $c, c' \in C$ be false twins and let $H_2 = (C \setminus \{c, c'\}) + (V(H) \setminus I)$. Clearly, $H_2$ is a $C_3$-hitting set of $G$. Hence, $\nabla_3(G) \leq |H_2| = |C| - 2 + |V(H)| - \alpha(H)$. Follows, $\nabla_3(G) \leq \min\{|C| - 2 + |V(H)| - \alpha(H), |C| + \nabla_3(H)\}$

Let $H'$ be a $C_3$-hitting set of $G$ and suppose $|H'| < \min\{|C| - 2 + |V(H)| - \alpha(H), |C| + \nabla_3(H)\}$. By Observation 4.14, $|C \setminus H'| \leq 3$. Let us distinguish 4 cases:

1. $|C \setminus H'| = 0$,
   i.e., all vertices from $C$ are in the hitting set $H'$. Clearly, $H' - C$ is a $C_3$-hitting set of $H$. Hence, $|H'| \geq |C| + \nabla_3(H)$, a contradiction.

13
2. \(|C \setminus H'| = 1\),
i.e., all but one vertex \(c \in C\) from \(C\) are in the hitting set \(H'\). Since \(c\) is connected to every vertex in \(H\) and since \(|E(H)| \geq 1\), also some vertices from \(H\) have to be in \(H'\). To be more clear, \(V(H) \setminus H'\) has to be an independent set. Hence, \(|V(H) \setminus H'| \leq \alpha(H)\). Thus, \(|H'| \geq |C \cap H'| + |V(H) \cap H'| \geq |C| - 1 + |V(H)| - \alpha(H)\), a contradiction.

3. \(|C \setminus H'| = 2\),
i.e., two vertices \(c_1, c_2\) from \(C\) are not in the hitting set \(H\). If \(c_1\) and \(c_2\) are not false twins, then all vertices from \(H\) have to be in hitting set \(H\). Note, that \(|V(H)| \geq 2\), hence \(\Delta_3(H) \leq |V(H)| - 2\). We infer \(|H'| \geq |C| - 2 + |V(H)| \geq |C| + \Delta_3(H)\), a contradiction. If \(c_1\) and \(c_2\) are false twins, then \(V(H) \setminus H'\) has to be independent set. Thus, \(|H'| \geq |C \cap H'| + |V(H) \cap H'| \geq |C| - 2 + |V(H)| - \alpha(H)\), a contradiction.

4. \(|C \setminus H'| = 3\),
i.e., three vertices \(c_1, c_2, c_3\) from \(C\) are not in the hitting set \(H'\). Clearly, two of them must be false twins. Hence, all vertices from \(H\) have to be in hitting set \(H\). Note, that \(\alpha(H) \geq 1\). Thus, \(|H'| \geq |C| - 3 + |V(H)| \geq |C| - 2 - \alpha(H) + |V(H)|\), a contradiction.

□

Figure 1: Quasi-spider \(G_1\) in Lemma 4.19.

**Lemma 4.19** Let \(G = (S, C, H)\) (quasi-)spider without false twins in \(C\), where \(V(H) = \emptyset\). Then,

\[
\Delta_3(G) = \begin{cases} 
|C| - 2 & \text{if } G \text{ is thin quasi-spider, which is not } G_1 \\
|C| - 1 & \text{if } G \text{ is thick quasi-spider or } G_1
\end{cases}
\]

**Proof.** Let \(G\) be a thin (quasi-)spider without false twins in \(C\) and let \(c_1, c_2 \in C\). Let \(H = C \setminus \{c_1, c_2\}\). Clearly, if \(G = G_1\), then \(\Delta_3(G) = 1 = |C| - 1\). If \(G \neq G_1\), then \(H\) is a \(C_3\)-hitting set of \(G\). Hence, \(\Delta_3(G) \leq |H| =
Observation 4.15, $\nabla C$ Let $\Box h$ and any of $G$

Proof. Let $\nabla C$ $H$ where $\nabla C$. Hence, $\nabla C = |C| - 2$, if $G$ is a thin (quasi-)spider.

Let $G$ be a thick (quasi-)spider without false twins in $C$, let $c \in C$ and let $\mathcal{H} = C \setminus \{c\}$. Clearly, $\mathcal{H}$ is a $C_3$-hitting set of $G$. Hence, $\nabla_3(G) \leq |\mathcal{H}| = |C| - 1$. By Observation 4.15, $\nabla_3(G) \geq |C| - 2$. Suppose, there exists $C_3$-hitting set $\mathcal{H}'$ of $G$ with $|\mathcal{H}'| = |C| - 2$. Hence, $\mathcal{H} \subset C$ and $C - \mathcal{H}' = \{c_1, c_2\}$. Because $G$ is a thick quasi-spider with weight $r > 2$, there exists vertex $s \in S \cap (N(c_1) \cap N(c_2))$. Hence, $c_1, c_2$ and $s$ induce a $C_3$, a contradiction. Hence, $\nabla_3(G) = |C| - 1$, if $G$ is a thick (quasi-)spider. □

Lemma 4.20 Let $G = (S, C, H)$ (quasi-)spider without false twins in $C$, where $H = K_k, k \in \mathbb{N}$. Then, $\nabla_3(G) = |C| - 1$.

Proof. Let $c \in C$ vertex described in Observation 4.13 and let $\mathcal{H} = C \setminus \{c\}$. Clearly, $\mathcal{H}$ is a $C_3$-hitting set of $G$. Hence, $\nabla_3(G) \leq |\mathcal{H}| = |C| - 1$. By Observation 4.15, $\nabla_3(G) \geq |C| - 2$. Suppose, there exists $C_3$-hitting set $\mathcal{H}'$ of $G$ with $|\mathcal{H}'| = |C| - 2$. Hence, $\mathcal{H} \subset C$ and $C - \mathcal{H}' = \{c_1, c_2\}$. Hence, $c_i, c_j$ and any $h \in V(H)$ induce a $C_3$, a contradiction. Hence, $\nabla_3(G) = |C| - 1$. □

Lemma 4.21 Let $G = (S, C, H)$ be a (quasi-)spider without false twins in $C$, where $H$ is not edgeless. Then, $\nabla_3(G) = |C| + \nabla_3(H)$.

Proof. Let $\mathcal{H}_{C_3}(H)$ be minimum $C_3$-hitting set and let $\mathcal{H} = C + \mathcal{H}_{C_3}(H)$. Clearly, $\mathcal{H}$ is a $C_3$-hitting set of $G$. Hence, $\nabla_3(G) \leq |\mathcal{H}| = |C| + \nabla_3(H)$.

Let $\mathcal{H}'$ be a $C_3$-hitting set of $G$ and suppose $|\mathcal{H}'| < |C| + \nabla_3(H)$. By Observation 4.15, $|C \setminus \mathcal{H}'| \leq 2$. Let’s distinguish 2 cases:

1. $|C \setminus \mathcal{H}'| = 0$
   i.e. all vertices from $C$ are in the $C_3$-hitting set $\mathcal{H}'$. Hence, $|V(H) \cap \mathcal{H}'| < \nabla_3(H)$, a contradiction.

2. $|C \setminus \mathcal{H}'| = 1$
   i.e. all but one vertex $c \in C$ from $C$ are in the hitting set $\mathcal{H}'$. Since $c$ is connected to every vertex in $H$ and since $E(H) \geq 1$, also some vertices from $H$ have to be in $\mathcal{H}'$. To be more clear, $V(H) \setminus \mathcal{H}'$ has to be independent set. Hence, $|V(H) \setminus \mathcal{H}'| \leq \alpha(H)$. Follows, $|\mathcal{H}'| \geq |C \setminus \mathcal{H}'| + |V(H) \cap \mathcal{H}'| \geq |C| - 1 + |V(H)| - \alpha(H)$. By Lemma 4.11 (ii), $|\mathcal{H}'| \geq |C| + |V(H)| - \alpha(H) - 1 \geq |C| + \nabla_3(H)$, a contradiction.
3. \(|C \setminus \mathcal{H}'| = 2\)
   i.e. two vertices \(c_1, c_2\) from \(C\) are not in the hitting set \(\mathcal{H}\), so all vertices from \(H\) have to be in hitting set \(\mathcal{H}\). Note, that \(|V(H)| \geq 2\), hence \(\nabla_3(H) \leq |V(H)| - 2\). Follows, \(|\mathcal{H}'| \geq |C| - 2 + |V(H)| \geq |C| + \nabla_3(H)\), a contradiction.

\[\square\]

By a combination of the previous results we have:

**Theorem 4.22** Let \(G\) be a (quasi-)spider.

- \(\nabla_3(G) = |C| - 3\) if \(G = (S, C, H)_{\text{thin}}\) and \(V(H) = \emptyset\).
- \(\nabla_3(G) = |C| - 2\) if
  - \(G = (S, C, H)_{\text{thick}}\) and \(V(H) = \emptyset\) or \(H = \overline{K}_k, k \in \mathbb{N}\), or
  - \(G = (S, C, H)_{\text{thin}}\) and \(H = \overline{K}_k, k \in \mathbb{N}\), or
  - \(G \in \{(S, C, H)_{\text{thin}}, (S^t, C, H)_{\text{thin}}, (S, C^t, H)_{\text{thin}}\}\) and \(V(H) = \emptyset\), or
  - \(G = (S^t, C, H)_{\text{thin}}\) except \(G_1\).
- \(\nabla_3(G) = |C| - 1\) if
  - \(G \in \{(S, C, H)_{\text{thin}}, (S^t, C, H)_{\text{thin}}, (S^t, C, H)_{\text{thick}}, (S, C^t, H)_{\text{thick}}\}\) and \(V(H) = \emptyset\) or \(H = \overline{K}_k, k \in \mathbb{N}\), or
  - \(G \in \{(S, C, H)_{\text{thin}}, (S^t, C, H)_{\text{thin}}, (S^t, C, H)_{\text{thick}}, (S, C^t, H)_{\text{thick}}\}\) and \(H = \overline{K}_k, k \in \mathbb{N}\), or
  - \(G = G_1\).
- \(\nabla_3(G) = \min\{C| - 2 + |V(H)| - \alpha(H), |C| + \nabla_3(H)\}\) if \(G \in \{(S, C, H)_{\text{thin}}, (S, C^t, H)_{\text{thick}}\}\) and \(|E(H)| \geq 1\).
- \(\nabla_3(G) = |C| + \nabla_3(H)\) if \(G \in \{(S, C, H)_{\text{thin}}, (S^t, C, H)_{\text{thin}}, (S^t, C, H)_{\text{thick}}, (S, C, H)_{\text{thick}}\}\) and \(|E(H)| \geq 1\).

**Proof.** Follows directly from Lemmas 4.16, 4.17, 4.18, 4.19, 4.20 and 4.21.

\[\square\]

The previous results implies the following.
\textbf{Theorem 4.23} $C_3$-hitting number $\nabla_3$ can be computed in linear time for $P_4$-tidy graphs.

\textbf{Proof.} First we show that $C_3$-hitting number $\nabla_3$ can be computed in linear time for non-trivial modular $P_4$-tidy graphs. Clearly, this is true for graphs $C_5$, $P_5$ and $\overline{P_5}$.

By using Theorem 4.22, we will show that $\nabla_3$ can be computed in linear time also for (quasi-)spider graphs. This is not obvious just in the case where $G \in \{(S, C^f, H)_{thin}, (S, C^f, H)_{thick}\}$, $|E(H)| \geq 1$ and $|C| - 2 + |V(H)| - \alpha(H) < |C| + \nabla_3(H)$. In this case $\nabla_3(G)$ depends on $\alpha(H)$. However, the head of a (quasi-)spider is a $P_4$-tidy graph, and it is known the independence number can be computed in linear time for $P_4$-tidy graphs [14]. Hence, also in this case $\nabla_3(H)$ can be computed in linear time.

Combining this result and Lemma 4.12, we derive that $C_3$-hitting number $\nabla_3$ can be computed in linear time for $P_4$-tidy graphs. \hfill \square

As both problems that we consider in this paper have linear computational complexity in $P_4$-free graphs (and several their generalizations), one can be interested in the studied problems in the class of $P_5$-graphs that provide an extension of co-graphs and are very interesting with respect to \textsc{Maximum Independent Set Problem} and \textsc{Graph Colouring Problem}.

\textbf{Problem 4.24} What is the computational complexity of the $C_3$-hitting Problem and the $P_3$-hitting Problem in $P_5$-free graphs?

\textbf{References}


[19] F. Kardoš, J. Katrenič, I. Schiermeyer, On computing the minimum 3-

posium on Complexity of Computer Computations, IBM Thomas J. 
Watson Res. Center, Yorktown Heights, N.Y., New York: Plenum, 85-
103.

[21] M. Kumar, S. Mishra, N. Safina Devi, S. Saurabh, Approximation al-
gorithms for node deletion problems on bipartite graphs with finite for-
90–96.

[22] Y. Li, Z. Yang, W. Wang, Complexity and algorithms for the connected 

[23] Z. Li, L. Zuo, The $k$-path vertex cover in Cartesian product graphs and 

[24] P. Mihók, I. Schiermeyer, Cycle lengths and chromatic number of 

ity of dissociation set problems in graphs, *Discrete Appl. Math.* 159 


[28] J. Tu, Efficient algorithm for the vertex cover $P_k$ problem on cacti, 