# A new characterization of $P_k$-free graphs

Eglantine Camby[1] and Oliver Schaudt[2]

[1] Université Libre de Bruxelles, Département de Mathématique, Boulevard du Triomphe, 1050 Brussels, Belgium
ecamby@ulb.ac.be
[2] Universität zu Köln, Institut für Informatik, Weyertal 80, 50931 Köln, Germany
schaudto@uni-koeln.de

**Abstract.** The class of graphs that do not contain an induced path on $k$ vertices, $P_k$-free graphs, plays a prominent role in algorithmic graph theory. This motivates the search for special structural properties of $P_k$-free graphs, including alternative characterizations. Let $G$ be a connected $P_k$-free graph, $k \geq 4$. We show that $G$ admits a connected dominating set that induces either a $P_{k-2}$-free graph or a graph isomorphic to $P_{k-2}$. Surprisingly, it turns out that every minimum connected dominating set of $G$ has this property. This yields a new characterization for $P_k$-free graphs: a graph $G$ is $P_k$-free if and only if each connected induced subgraph of $G$ has a connected dominating set that induces either a $P_{k-2}$-free graph, or a graph isomorphic to $C_k$. This improves and generalizes several previous results; the particular case of $k = 7$ solves a problem posed by van 't Hof and Paulusma. In the second part of the paper, we present an efficient algorithm that, given a connected graph $G$, computes a connected dominating set $X$ of $G$ with the following property: for the minimum $k$ such that $G$ is $P_k$-free, the subgraph induced by $X$ is $P_{k-2}$-free or isomorphic to $P_{k-2}$. As an application our results, we prove that Hypergraph 2-Colorability, an NP-complete problem in general, can be solved in polynomial time for hypergraphs whose vertex-hyperedge incidence graphs is $P_7$-free.
**keywords:** $P_k$-free graph, connected domination, computational complexity.
**MSC:** 05C69, 05C75, 05C38.

## 1 Introduction

A *dominating set* of a graph $G$ is a vertex subset $X$ such that every vertex not in $X$ has a neighbor in $X$. Dominating sets have been intensively studied in the literature. The main interest in dominating sets is due to their relevance on both theoretical and practical side. Moreover, there are interesting variants of domination and many of them are well-studied.

A *connected dominating set* of a graph $G$ is a dominating set $X$ whose induced subgraph, henceforth denoted $G[X]$, is connected. As usual, a connected dominating set such that every proper subset is not a connected dominating set is called a *minimal connected dominating set.* A connected dominating set of minimum size is called a *minimum connected dominating set.*

We use the following standard notation. Let $P_k$ be the induced path on $k$ vertices and let $C_k$ be the induced cycle on $k$ vertices. If $G$ and $H$ are two graphs, we say that $G$ is $H$*-free* if $H$ does not appear as an induced subgraph of $G$. Furthermore, if $G$ is $H_1$-free and $H_2$-free for some graphs $H_1$ and $H_2$, we say that $G$ is $(H_1, H_2)$*-free*. If two graphs $G$ and $H$ are isomorphic, we write $G \cong H$.

The class of $P_k$-free graphs has received a fair amount of attention in the theory of graph algorithms. Given an NP-hard optimization problem, it is often fruitful to study its complexity when the instances are restricted to $P_k$-free graphs.

Let us mention two recent results in this direction: the polynomial time algorithm to compute a stable set of maximum weight in $P_5$-free graphs, given by Lokshtanov *et al.* [10], and the result of Hoang *et al.* [6] showing that $k$-Colorability is efficiently solvable on $P_5$-free graphs. The proof of the latter result relies on the fact that a connected $P_5$-free graph has a dominating clique or a dominating $P_3$.

**Theorem 1 (Bácso and Tuza [1]).** *Let $G$ be a connected $P_5$-free graph. Then $G$ has a dominating clique or a dominating induced $P_3$.*

An immediate implication of this result is the following.

**Theorem 2 (Bácso and Tuza [1], Cozzens and Kelleher [4]).** *Let $G$ be a graph. The following assertions are equivalent.*

*(i) $G$ is $P_5$-free.*
*(ii) Every connected induced subgraph $H$ of $G$ admits a connected dominating set $X$ such that $H[X]$ is a clique or $H[X] \cong C_5$.*

Later, van 't Hof and Paulusma [13] obtained a characterization for the class of $P_6$-free graphs in the flavour of Theorem 2. An earlier, slightly weaker result was given by Liu *et al.* [8], and the particular case of $(C_3, P_6)$-free graphs was discussed before by Liu and Zhou [9].

**Theorem 3 (van 't Hof and Paulusma [13]).** *Let $G$ be a graph. The following assertions are equivalent.*

*(i) $G$ is $P_6$-free.*
*(ii) Every connected induced subgraph $H$ of $G$ admits a connected dominating set $X$ such that $H[X]$ has a complete bipartite spanning subgraph or $H[X] \cong C_6$.*

Complementing Theorem 3, van 't Hof and Paulusma give a polynomial time algorithm that, given a connected $P_6$-free graph, computes a connected dominating set $X$ such that $G[X]$ has a complete bipartite spanning subgraph or $G[X] \cong C_6$.

In view of Theorems 2 and 3, two questions arise. The first one is whether condition (ii) of Theorem 3 can be tightened, such that $H[X]$ is a $P_4$-free graph or $G[X] \cong C_6$. Note that if $H[X]$ is $P_4$-free, it is a connected cograph, and in particular has a complete bipartite spanning subgraph. This condition is the direct analogue of condition (ii) of Theorem 2 for $P_6$-free graphs. The advantage of the strenghtened version is of course that the structure of cographs is well understood and more restricted compared to the class of graphs having a spanning complete bipartite graph.

The second question is whether similar characterizations can be given for the class of $P_k$-free graphs, for $k > 6$. In their paper, van 't Hof and Paulusma [13] explicitly ask for such a characterization in the case of $k = 7$.

### 1.1 Our contribution

In this paper, we give an affirmative answer to these two questions. We show that every connected $P_k$-free graph, $k \geq 4$, admits a connected dominating set that induces either a $P_{k-2}$-free graph, or a graph isomorphic to $P_{k-2}$. Surprisingly, it turns out that every minimum connected dominating set has this property.

**Theorem 4.** *Let $G$ be a connected $P_k$-free graph, $k \geq 4$, and $X$ be any minimum connected dominating set of $G$. Then $G[X]$ is $P_{k-2}$-free, or $G[X] \cong P_{k-2}$.*

From this result we derive the following characterization of $P_k$-free graphs.

**Theorem 5.** *Let $G$ be a graph and $k \geq 4$. The following assertions are equivalent.*

*(i) $G$ is $P_k$-free.*
*(ii) Every connected induced subgraph $H$ of $G$ admits a connected dominating set $X$ such that $H[X]$ is $P_{k-2}$-free or $H[X] \cong C_k$.*

We now come to an algorithm for the connected domination with a certain structure. The proof of Theorem 4 is constructive in the sense that it yields an algorithm to compute, given a connected $P_k$-free graph, a connected dominating set that induces either a $P_{k-2}$-free graph, or a graph isomorphic to $P_{k-2}$. This algorithm is polynomial only when $k$ is fixed, and hence not useful for arbitrary input graphs. However, recall that the computation of a longest induced path in a graph is an NP-hard problem, as shown in Garey and Johnson [5, p. 196]. In other words, there is little hope of computing in polynomial time the minimum $k$ for which the input graph is $P_k$-free. To overcome this obstacle, our algorithm can only make implicit use of the absence of an induced $P_k$, which is the main difficulty here.

**Theorem 6.** *Given a connected graph $G$ on $n$ vertices and $m$ edges, one can compute in time $\mathcal{O}(n^5(n+m))$ a connected dominating set $X$ with the following property: for the minimum $k \geq 4$ such that $G$ is $P_k$-free, $G[X]$ is $P_{k-2}$-free or $G[X] \cong P_{k-2}$.*

Our last result is an application of the previous theorems. A *2-coloring* of a hypergraph assigns to each vertex one of two colors, such that each hyperedge contains vertices of both colors. The problem Hypergraph 2-Colorability is to decide whether a given hypergraph admits a 2-coloring. Garey and Johnson [5, p. 221] explain that it is NP-complete in general. One successful approach to deal with this hardness is to put restrictions on the bipartite vertex-hyperedge incidence graphs[3] of the input hypergraph.

As an application of Theorem 3, van 't Hof and Paulusma [13] show that Hypergraph 2-Colorability is solvable in polynomial time for hypergraphs with $P_6$-free incidence graphs. Using our results, we settle the case of hypergraphs with $P_7$-free incidence graphs.

**Theorem 7.** Hypergraph 2-Colorability *can be solved in polynomial time for hypergraphs with $P_7$-free incidence graphs. If it exists, a 2-coloring can be computed in polynomial time.*

The proof of our results we give in the subsequent sections. We close the paper with a short discussion of our contribution.

## 2 Proofs

### 2.1 Proof of Theorems 4 and 5

We need the following lemma from an earlier paper of ours [3].

**Lemma 1 (Camby and Schaudt [3]).** *Let $G$ be a connected graph that is $(P_k, C_k)$-free, for some $k \geq 4$, and let $X$ be a minimal connected dominating set of $G$. Then $G[X]$ is $P_{k-2}$-free.*

When applied to connected $P_k$-free graphs, which are in particular ($P_{k+1}$, $C_{k+1}$)-free, the above lemma implies that any minimal connected dominating set induces a $P_{k-1}$-free graph, for $k \geq 3$. We next prove a simple but useful lemma, which plays a key role also in the proof of Theorem 6. Let $X$ be a connected dominating set of a connected graph $G$, and $x \in X$. Assuming that $X$ is a minimal connected dominating set and $|X| \geq 2$, $x$ is a cut-vertex of $G[X]$ or $x$ has a *private neighbor*: a vertex $y \in V(G) \setminus X$ with $N_G(y) \cap X = \{x\}$.

**Lemma 2.** *Let $G$ be a connected $P_k$-free graph, for some $k \geq 4$, and let $X$ be a minimal connected dominating set of $G$. Assume that there is an induced $P_{k-2}$ in $G[X]$, say on the vertices $x_1, x_2, \ldots, x_{k-2}$. Then any private neighbor $y$ of $x_1$ is such that $(X \cup \{y\}) \setminus \{x_{k-2}\}$ is a connected dominating set of $G$.*

*Proof.* Let $X' := \{x_1, x_2, \ldots, x_{k-2}\}$. Suppose $x_1$ has a private neighbor $y$, and let $Y := (X \cup \{y\}) \setminus \{x_{k-2}\}$. We have to prove that $Y$ is a connected dominating set of $G$.

[3] Recall that for a hypergraph $H = (V, E)$ we define the bipartite vertex-hyperedge incidence graph as the bipartite graph on the set of vertices $V \cup E$ with the edges $vY$ such that $v \in V$, $Y \in E$ and $v \in Y$. In the following, we just say the *incidence graph*.

Suppose for a contradiction that $G[Y]$ is not connected. Hence, $x_{k-2}$ is a cut-vertex of $G[X]$. In particular, there is some vertex $x' \in X$ such that $N_G(x') \cap X' = \{x_{k-2}\}$. But then $G[X' \cup \{y, x'\}] \cong P_k$, a contradiction.

It remains to show that $Y$ is a dominating set. Suppose the contrary, that is, there is some vertex $z$ with $N_G[z] \cap Y = \emptyset$. As $X$ is a dominating set, $N_G[z] \cap X = \{x_{k-2}\}$. Because $x_{k-2}$ is adjacent to $Y$ and $z$ is not adjacent to $Y$, $z \neq x_{k-2}$. But this means that $G[X' \cup \{y, z\}] \cong P_k$, a contradiction. □

Now we can state the proof of Theorem 4.

*Proof (Proof of Theorem 4.).* Let $X$ be a minimum connected dominating set of $G$. As $G$ is in particular $(P_{k+1}, C_{k+1})$-free, $G[X]$ is $P_{k-1}$-free, by Lemma 1. We have to show that $G[X]$ is $P_{k-2}$-free or isomorphic to $P_{k-2}$.

To see this, assume there is an induced $P_{k-2}$ in $G[X]$, say on the vertices $x_1, x_2, \dots, x_{k-2}$. Let $X' := \{x_1, x_2, \dots, x_{k-2}\}$. We will prove that $X \setminus X' = \emptyset$. Note that $x_1$ is not a cut-vertex of $G[X]$: otherwise there is some vertex $x' \in X$ such that $N_G(x') \cap X' = \{x_1\}$, and hence $G[X' \cup \{x'\}] \cong P_{k-1}$. This is a contradiction. Thus, $x_1$ is not a cut-vertex of $G[X]$ and therefore has a private neighbor w.r.t. $X$, say $y_1$. By Lemma 2, $Y_1 := (X \cup \{y_1\}) \setminus \{x_{k-2}\}$ is a connected dominating set of $G$. As $X$ is a minimum connected dominating set, $Y_1$ is a minimum connected dominating set, too. Moreover, $y_1$ has no neighbor in $X \setminus \{x_1\}$, in particular in $X \setminus X'$.

By reapplying the argumentation to $Y_1$ and the induced $P_{k-2}$ on $y_1, x_1, x_2, \dots, x_{k-3}$, we obtain a vertex $y_2 \in V(G) \setminus Y_1$ such that $Y_2 := (Y_1 \cup \{y_2\}) \setminus \{x_{k-3}\}$ is a minimum connected dominating set of $G$ and $G[Y_2]$ contains an induced $P_{k-2}$ on the vertices $y_2, y_1, x_1, x_2, \dots, x_{k-4}$. Moreover, $y_2$ has no neighbor in $Y_1 \setminus \{y_1\}$, in particular in $X \setminus X'$.

Iteratively, we end up with a minimum connected dominating set $Y_{k-2}$, which is exactly $(X \setminus X') \cup \{y_1, \dots, y_{k-2}\}$. Since, for $i = 1, 2, \dots, k-2$, $y_i$ is not adjacent to $X \setminus X'$ and $G[Y_{k-2}]$ is connected, $X \setminus X'$ must be empty, hence $X = X'$. Thus, $G[X] = G[X'] \cong P_{k-2}$. This completes the proof. □

*Proof (Proof of Theorem 5.).* Clearly $P_k$ does not have a connected dominating set which satisfies (ii). Hence, (ii) implies (i).

Conversely, let $H$ be any connected induced subgraph of $G$, and let $X$ be a minimum connected dominating set of $H$. By Theorem 4, $H[X]$ is $P_{k-2}$-free or $H[X] \cong P_{k-2}$. If $H[X]$ is $P_{k-2}$-free, the assertion of (ii) is satisfied. Otherwise, let $x_1, x_2, \dots, x_{k-2}$ be a consecutive ordering of the induced path $H[X]$. In particular, $x_1$ and $x_{k-2}$ are not cut-vertices of $H[X]$. As $X$ is minimum, there exists a private neighbor $y_i$ of $x_i$, for $i \in \{1, k-2\}$. It must be that $y_1 y_{k-2} \in E(H)$, since otherwise $H[X \cup \{y_1, y_{k-2}\}] \cong P_k$. Hence, $H[X \cup \{y_1, y_{k-2}\}] \cong C_k$, as desired. So, (i) implies (ii). □

### 2.2 Proof of Theorem 6

Before we state our algorithm, we need to introduce some notation and definitions. For this, let us assume we are given a connected input graph $G$ on $n$

vertices and $m$ edges. Let $X$ be an arbitrary connected dominating set of $G$ with at least two vertices.

By $NC(X)$ we denote the set of vertices in $X$ that are non-cutting in $G[X]$, i.e. for every $x \in NC(X)$, $G[X \setminus \{x\}]$ is connected. Let $x$ be a degree-1 vertex of $G[X]$. We define the *half-path* starting in $x$ to be the maximal path $(x, x_1, x_2, \ldots, x_s)$ in $X$ such that $|N_{G[X]}(x_i)| = 2$ for each $i \in \{1, 2, \ldots, s-1\}$. The *length* of the half-path is then $s$. For example, if the neighbor $y \in X$ of $x$ has degree at least 3, the half-path is simply $(x, y)$.

To each $x \in X$ we assign a weight $w_X(x)$ as follows:

1. if $|N_{G[X]}(x)| \geq 2$, put $w_X(x) = 0$, and
2. if $|N_{G[X]}(x)| = 1$, put $w_X(x) = s$, where $s$ is the length of the half-path starting in $x$.

Finally, the weight $w(X)$ of the set $X$ is given by

$$w(X) = \sum_{x \in X} (w_X(x))^2.$$

The squares in this sum are needed for technical reasons. See Fig. 1 for an illustration of the above definitions.

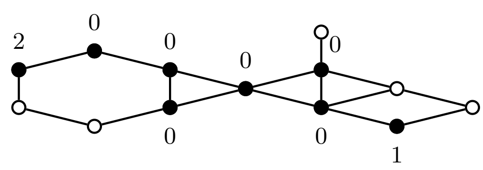


**Fig. 1.** A graph $G$. The black vertices form a connected dominating set $X$ of $G$, with weights $w_X$ as shown. We have $w(X) = 5$.

Let $\mathcal{X}$ be the family of all connected dominating sets of $G$. We next define a strict partial order $\prec$ on $\mathcal{X}$ as follows. For any two sets $X, Y \in \mathcal{X}$, we put $X \prec Y$ if

1. $|X| > |Y|$, or
2. $|X| = |Y|$ and $w(X) < w(Y)$.

We now consider the strict poset $(\mathcal{X}, \prec)$. The *height* of $(\mathcal{X}, \prec)$ is the maximum order of a set of mutually comparable elements of $\mathcal{X}$.

**Lemma 3.** *For a connected $n$-vertex graph $G$, the height of $(\mathcal{X}, \prec)$ is in $\mathcal{O}(n^3)$.*

*Proof.* Let $X$ be a connected dominating set of a connected graph $G$. If $G[X]$ is not an induced path, every vertex in $X$ of degree at most 2 in $G[X]$ is contained in at most one half-path. Hence, $\sum_{x \in X} w_X(x) \leq |X|$. If $G[X]$ is an induced path, every vertex appears in exactly two half-paths, implying $\sum_{x \in X} w_X(x) \leq 2|X|$. Thus

$$w(X) = \sum_{x \in X} (w_X(x))^2 \leq (\sum_{x \in X} w_X(x))^2 \leq 4|X|^2,$$

and so the weight of a connected dominating set is in $\mathcal{O}(n^2)$. Since there are at most $n$ different possible sizes of connected dominating sets of $G$, the height of $(\mathcal{X}, \prec)$ is in $\mathcal{O}(n^3)$. □

*Proof (Proof of Theorem 6).* Assume we are given a connected graph $G$ on $n$ vertices and $m$ edges as input and let $k \in \mathbb{N}$ be the smallest integer such that $G$ is $P_k$-free. Our algorithm works as follows, starting with the connected dominating set $Y := V(G)$. Its output is a connected dominating set $X$ with the properties stated in Theorem 6.

1. Compute a minimal connected dominating set $X \subseteq Y$.
2. If $G[X]$ is an induced path, return $X$ and terminate the algorithm.
3. Compute the set $NC(X)$ and the weight $w_X(x)$ for every $x \in NC(X)$.
4. Order the vertices of $NC(X)$ with non-increasing weight $w_X$, breaking ties arbitrarily. Let that order be $v_1, v_2, \ldots, v_{|NC(X)|}$.
5. For $i$ from 1 to $|NC(X)|$ do the following:
   (a) Compute a private neighbor $y_i$ of $v_i$ w.r.t. $X$.
   (b) For $j$ from $i + 1$ to $|NC(X)|$ do the following:
      i. Check whether $Y_{ij} := (X \cup \{y_i\}) \setminus \{v_j\}$ is a connected dominating set.
      ii. If yes, put $Y \leftarrow Y_{ij}$ and go to Step 1.
6. Return $X$ and terminate the algorithm.

We remark that the computation of $y_i$ in Step 5a is always possible, since $x_i$ is non-cutting in $G[X]$ and $X$ is a minimal connected dominating set.

See Fig. 2 for an illustration of Step 5(b)ii.

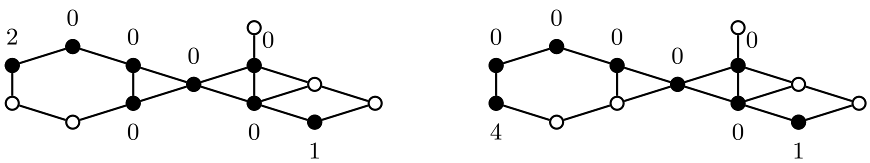


**Fig. 2.** Before (left) and after (right) an application of Step 5(b)ii. In the next iteration, the algorithm terminates with the right connected dominating set as output.

The proof is completed by the following sequence of claims.

*Claim.* When the algorithm terminates, the output $X$ is a connected dominating set and $G[X]$ is $P_{k-2}$-free or $G[X] \cong P_{k-2}$.

Since Step 1 is applied before the return is called, $X$ is a minimal connected dominating set. Since $G$ is $P_k$-free and hence $(P_{k+1}, C_{k+1})$-free, $G[X]$ is $P_{k-1}$-free, by Lemma 1. If the algorithm terminates with Step 2, it holds that $G[X]$ is isomorphic to $P_\ell$ for some $\ell < k - 2$. Hence, either $G[X] \cong P_{k-2}$ or $G[X]$ is $P_{k-2}$-free.

Now assume that the algorithm terminates in Step 6. In particular, $G[X]$ is not an induced path. Suppose for a contradiction that $G[X]$ contains an induced

$P_{k-2}$, say on the vertices $x_1, x_2, \ldots, x_{k-2}$. Like in the proof of Lemma 2, both $x_1$ and $x_{k-2}$ cannot be cut-vertices of $G[X]$. Thus, $x_1, x_{k-2} \in NC(X)$.

After Step 4, the vertices of $NC(X)$ are ordered $v_1, v_2, \ldots, v_{|NC(X)|}$ with non-increasing weight. W.l.o.g. $x_1 = v_i$, $x_{k-2} = v_j$, and $i < j$ and let $y_i$ be the private neighbor of $v_i$ that is computed by the algorithm. As $X$ is returned, the set $Y_{ij} := (X \cup \{y_i\}) \setminus \{v_j\}$ is not a connected dominating set, in contradiction to Lemma 2. This proves our claim.

*Claim.* Let $X$ be a minimal connected dominating set considered in some iteration of the algorithm. Assume that the 'go to' is called in Step 5(b)ii because $Y_{ij} := (X \cup \{y_i\}) \setminus \{v_j\}$ is a connected dominating set. Let $X'$ be the minimal connected dominating set computed in the subsequent Step 1. Then $X \prec X'$.

Note that $|X| \geq 2$ since the 'go to' is called in Step 5(b)ii and $G[X]$ is not an induced path. So we can assume that every vertex in $G[X]$ has degree at least 1. Clearly $|X'| \leq |X|$. If $|X'| < |X|$, $X \prec X'$ by definition. So we may assume that $|X'| = |X|$, and hence $X' = Y_{ij}$. It remains to show that $w(X) < w(X')$.

Let $z \in X \setminus \{v_i, v_j\}$. If $z$ has degree at least 2, then $w_X(z) = 0$ and hence $w_{X'}(z) \geq w_X(z)$. Now suppose $z$ has degree 1. Let $(z, x_1, x_2, \ldots, x_s)$ be a half-path starting in $z$. As $G[X]$ is not a path, $x_s$ is a cut-vertex of $G[X]$. In particular, $x_s \neq v_j$. Hence, in $G[X']$, $(z, x_1, x_2, \ldots, x_s)$ is the initial segment of a half-path starting in $z$. In particular, $w_{X'}(z) \geq w_X(z)$.

If $v_i$ is not a degree-1 vertex of $G[X]$, $w_{X'}(v_i) = w_X(v_i) = 0$, and $(y_i, v_i)$ is the initial segment of a half-path starting in $y_i$. Hence, $w_{X'}(y_i) \geq 1$, and thus

$$w_{X'}(v_i) = 0 \text{ and } w_{X'}(y_i) \geq w_X(v_i) + 1. \tag{1}$$

If the degree of $v_i$ in $G[X]$ is 1, let $(v_i, x_1, x_2, \ldots, x_s)$ be a half-path starting in $v_i$. Again, $x_s$ is a cut-vertex of $G[X]$, and so $x_s \neq v_j$. Hence, in $G[X']$, $(y_i, v_i, x_1, x_2, \ldots, x_s)$ is the initial segment of a half-path starting in $y_i$. Again (1) holds.

Summing up, we see that (1) holds, and

$$w_{X'}(z) \geq w_X(z) \text{ for every vertex } z \in X' \setminus \{y_i, v_i\}. \tag{2}$$

We now turn to the vertex $v_j$. First assume that the degree of $v_j$ in $G[X]$ is at least 2, and thus $w_X(v_j) = 0$. Combining this with (1) and (2) yields

$$w(X') - w(X) \geq w_{X'}(y_i)^2 - w_X(v_j)^2 > 0,$$

and so $w(X') - w(X) > 0$.

Now assume that $v_j$ is a vertex of degree 1 in $G[X]$, and so $w_X(v_j) \geq 1$. Let $N_{G[X]}(v_j) = \{x\}$. As $G[X]$ is not a path, $|N_{G[X]}(x)| \geq 2$, and so $w_X(x) = 0$. Thus $w_{X'}(x) = w_X(v_j) - 1$. Recall that (1) and (2) hold. We obtain the following inequality.

$$\begin{aligned} w(X') - w(X) &\geq w_{X'}(y_i)^2 + w_{X'}(x)^2 - w_X(v_i)^2 - w_X(v_j)^2 \\ &= (w_{X'}(y_i)^2 - w_X(v_i)^2) - (w_X(v_j)^2 - w_{X'}(x)^2) \\ &\geq [(w_X(v_i) + 1)^2 - w_X(v_i)^2] - [w_X(v_j)^2 - (w_X(v_j) - 1)^2] \end{aligned}$$

But $w_X(v_i) \geq w_X(v_j)$ implies

$$(w_X(v_i) + 1)^2 - w_X(v_i)^2 > w_X(v_j)^2 - (w_X(v_j) - 1)^2,$$

and thus $w(X') - w(X) > 0$ holds as in the previous case.

Hence, $X \prec X'$, proving our claim.

*Claim.* The algorithm terminates in $\mathcal{O}(n^5(n+m))$ time.

By Claim 2.2, each call of the 'go to'-step and the subsequent application of Step 1 result in a connected dominating set that is properly larger in the order $\prec$. By Lemma 3, the height of the poset $(\mathcal{X}, \prec)$, and hence the number of iterations the whole algorithm performs, is in $\mathcal{O}(n^3)$.

It remains to discuss the complexity of the particular steps. For this, recall that it can be checked in time $\mathcal{O}(n+m)$ whether a given vertex subset is a connected dominating set. Consequently, Step 1 can be performed in time $\mathcal{O}(n(n+m))$ by the immediate greedy procedure.

Step 2 and the computation of the weights in Step 3 can both be performed in linear time using the degree sequence of $G[X]$. The computation of the set $NC(X)$ in Step 3 can be done straightforwardly in time $\mathcal{O}(n(n+m))$.

It remains to discuss the complexity of the loop of Step 5. The computation of a private neighbor in Step 5a can clearly be done in $\mathcal{O}(n+m)$ time. The inner loop of Step 5b performs $\mathcal{O}(n)$ checks whether some vertex set is a connected dominating set, requiring $\mathcal{O}(n+m)$ time each. Hence, Step 5 can be done in $\mathcal{O}(n^2(n+m))$ time.

The overall running time amounts to $\mathcal{O}(n^5(n+m))$, which completes the proof of both our claim and Theorem 6. □

## 3 Conclusion

In this paper we gave a description of the structure of connected dominating sets in $P_k$-free graphs. We have shown that any connected $P_k$-free graph admits a connected dominating set whose induced subgraph is $P_{k-2}$-free or isomorphic to $P_{k-2}$. In fact, any minimum connected dominating set has this property. Loosely speaking, this means that the restricted structure of connected $P_k$-free graphs results in an even more restricted structure of the induced subgraph of their minimum connected dominating sets.

Although we think that our results are of their own interest, our hope is that they might be useful in other contexts, too. One example we gave is the polynomial time solvability of Hypergraph 2-Colorability for hypergraphs with $P_7$-free incidence graphs. It seems possible that, with more work, one could push this result to hypergraphs with $P_8$-free incidence graphs. However, more interesting would be to know whether there is any $k$ for which Hypergraph 2-Colorability for hypergraphs with $P_k$-free incidence graphs is *not* solvable in polynomial time.

Other possible future applications of our results include the coloring of $P_k$-free graphs. As mentioned earlier, Hoang *et al.* [6] showed that $k$-Colorability is efficiently solvable on $P_5$-free graphs, using the fact that a connected $P_5$-free graph has a dominating clique or a dominating induced $P_3$. To our knowledge, an open problem, conjectured by Huang [7], in this context is whether 4-colorability can be decided in polynomial time for $P_6$-free graphs. From Theorem 6 it follows that, given a $P_6$-free graph, we can efficiently compute a connected dominating set that induces a $P_4$-free graph (that is a cograph) or a $P_4$. Of course cographs are less trivial than cliques, especially when it comes to coloring – but that does not rule out an approach similar to that of Hoáng *et al.* [6]. The fact that each vertex of the graph has some neighbor in this cograph leaves a 3-coloring problem for the rest of the graph, once the coloring of the cograph is fixed. Here, one might use the fact that 3-coloring is polynomial time solvable for $P_6$-free graphs, shown by Randerath and Schiermeyer [11], even in the pre-coloring extension version, proven by Broersma *et al.* [2]. However, one would rather need to use a list coloring version for this particular problem.

## References


1. G. Bácso and Z. Tuza, *Dominating cliques in $P_5$-free graphs*, Period. Math. Hungar. **21** (1990), 303–308.
2. H. Broersma, F.V. Fomin, P.A. Golovach, and D. Paulusma, *Three complexity results on coloring $P_k$-free graphs*, Europ. J. Combin. **34** (2013), 609–619.
3. E. Camby and O. Schaudt, *The price of connectivity for dominating sets: upper bounds and complexity*, preprint (2013).
4. M.B. Cozzens and L.L. Kelleher, *Dominating cliques in graphs*, Disc. Math. **86** (1990), 101–116.
5. M.R. Garey and D.S. Johnson, *Computers and Intractability*, W.H. Freeman and Co., New York, 1979.
6. C.T. Hoáng, M. Kaminski, V.V. Lozin, J. Sawada, and X. Shu, *Deciding k-colorability of $P_5$-free graphs in polynomial time*, Algorithmica **57** (2010), 74–81.
7. S. Huang, *Improved Complexity Results on k-Coloring $P_t$-Free Graphs*, proc. MFCS 2013, 551–558.
8. J. Liu, Y. Peng, and C. Zhao, *Characterization of $P_6$-free graphs*, Disc. App. Math. **155** (2007), 1038–1043.
9. J. Liu and H. Zhou, *Dominating subgraphs in graphs with some forbidden structures*, Disc. Math. **135** (1994), 163–168.
10. D. Lokshtanov, M. Vatshelle, and Y. Villanger, *Independent set in $P_5$-free graphs in polynomial time*, to appear in proc. SODA 2014.
11. B. Randerath and I. Schiermeyer, *3-colorability $\in \mathcal{P}$ for $P_6$-free graphs*, Disc. App. Math. **136** (2004), 299–313.
12. P. van 't Hof and D. Paulusma, *A new characterization of $P_6$-free graphs*, proc. COCOON 2008, 415–424.
13. ———, *A new characterization of $P_6$-free graphs*, Disc. App. Math. **158** (2010), 731–740.

## Appendix: Proof of Theorem 7

*Proof (Proof of Theorem 7).* Let $H = (V, E)$ be a hypergraph whose incidence graph is $P_7$-free. A 2-coloring of $H$ we denote by $(A, B)$, where $A, B \subseteq V$ are two non-empty sets with $A \cup B = V$, each of which intersects every hyperedge.

A hypergraph for which any two hyperedges are not comparable (w.r.t. inclusion) is called a *clutter*. The following observation was proven by van 't Hof and Paulusma [13]. In order to be self-contained, we give a quick proof of it.

*Claim.* We may assume that $H$ is a clutter.

*Proof.* Assume there are hyperedges $e, f \in E$ with $e \subseteq f$. Such a pair of hyperedges we can detect in polynomial time.

Every 2-coloring of $H$ is a 2-coloring of the hypergraph $H' = (V, E \setminus \{f\})$ in particular. If $(A, B)$ is a 2-coloring of $H'$, it holds that $e \cap A \neq \emptyset$ and $e \cap B \neq \emptyset$. Thus, $f \cap A \neq \emptyset$ and $f \cap B \neq \emptyset$, and so $(A, B)$ is a 2-coloring of $H$.

So we may delete, for every such pair $e, f \in E$ with $e \subseteq f$ the hyperedge $f$ from $H$. It is clear that the resulting hypergraph is a clutter, and its incidence graph is still $P_7$-free. This proves Claim 3. □

Although immediate, Claim 3 considerably simplifies the argumentation of the following proof. We now assume that $H$ is a clutter. Moreover, we may assume that $H$ is connected, that is, its incidence graph is connected. In the following, we prove a sequence of claims that discuss all relevant cases for the 2-coloring problem. We state the polynomial algorithm along the way.

Let $G$ be the incidence graph of $H$. Since we are searching for a 2-coloring, we may assume that $|N_G(f)| \geq 2$ for every $f \in E$. By Theorem 5, there is a connected dominating set $X$ of $G$ such that $G[X]$ is $P_5$-free or $G[X] \cong C_7$. However, the latter case contradicts the fact that $G$ is bipartite. So, $G[X]$ is a connected $P_5$-free graph.

Using Theorem 5 again, we see that $G[X]$ has a dominating $P_3$-free graph. That is, there is a pair of adjacent vertices, say $v \in V$ and $e \in E$, that together dominate $G[X]$. In particular, $e$ intersects every other hyperedge. It is clear that we can compute such hyperedge in polynomial time.

*Claim.* If there is a proper subset $X \subset e$ that dominates $E$, $(X, V \setminus X)$ is a 2-coloring of $H$.

Let $f \in E$ be arbitrary. By assumption, $f \cap X \neq \emptyset$. Since $H$ is a clutter, $f \not\subseteq e$, and thus $f \not\subseteq X$. Hence, $f \setminus X \neq \emptyset$, proving Claim 3.

Indeed, it can be checked in polynomial time whether there is a proper subset $X \subset e$ that dominates $E$. (If so, $X$ is found in polynomial time, too.) In view of Claim 3, we may assume that no proper subset of $e$ dominates $E$.

We now make a distinction of the cases $|e| = 2$ and $|e| \geq 3$. Let us first assume that $|e| = 2$, say $e = \{x, y\}$. Since $H$ is a clutter, every hyperedge $f$ of $H$ contains either $x$ or $y$. Let $X, Y \subseteq E \setminus \{e\}$ such that every $f \in X$ contains $x$, every $g \in Y$ contains $y$, and $X \cup Y = E \setminus \{e\}$.

If $|X| = 0$, every hyperedge contains $y$ and, as $H$ is a clutter, some other vertex. Thus a 2-coloring of $H$ is given by $(\{y\}, V \setminus \{y\})$. By symmetry, we may now assume that $|X|, |Y| \geq 1$. Observe that, if $|X| = 1$, say $X = \{f\}$, $H$ is 2-colorable if and only if there is some vertex $v \in f$ such that $(\{v, y\}, V \setminus \{v, y\})$ is a 2-coloring of $H$. Indeed, if for every vertex $v \in f$, $(\{v, y\}, V \setminus \{v, y\})$ is not a 2-coloring of $H$, there exists a hyperedge $e_v = \{v, y\}$ for each such vertex $v$. Let now $v$ be an arbitrary vertex in $f$. Then $x$ and $v$ must have the same color, and so there is a vertex $v' \in f$ with the second color. Then, the hyperedge $e_{v'} = \{v', y\}$ is monochromatic, a contradiction. This condition can clearly be checked in polynomial time.

Now let $|X|, |Y| \geq 2$. We next show that $H$ admits a 2-coloring. To see this, pick any $f \in X$ and $g \in Y$. Since $H$ is a clutter, $f \setminus e, g \setminus e \neq \emptyset$. Pick any $u \in f \setminus e$ and $v \in g \setminus e$. If $fv, gu \notin E(G)$, $G[\{u, f, x, e, y, g, v\}] \cong P_7$, a contradiction. As $u$ and $v$ were arbitrary, it must be that $f \setminus e \subseteq g \setminus e$ or $g \setminus e \subseteq f \setminus e$.

Now let $f, f' \in X$ and $g \in Y$ be three mutually distinct hyperedges. As shown above, the sets $f \setminus e, f' \setminus e$ are comparable to $g \setminus e$. Since $H$ is a clutter, $f \setminus e$ is not comparable to $f' \setminus e$. Hence, either $f \setminus e, f' \setminus e \subseteq g$, or $g \setminus e \subseteq f, f'$.

In the first case, $f \setminus e \subseteq g$ for any $f \in X, g \in Y$. Thus, $(\bigcup_{f \in X} f) \setminus e \subseteq \bigcap_{g \in Y} g$. Since $H$ is a clutter, every $g \in Y$ has a neigbor outside the set $\{y\} \cup \bigcap_{g \in Y} g$. Hence,

$$(\{y\} \cup \bigcap_{g \in Y} g, V \setminus (\{y\} \cup \bigcap_{g \in Y} g))$$

is a 2-coloring of $H$. The second case, $g \setminus e \subseteq f, f'$, is dealt with in a similar fashion.

So we may assume $|e| \geq 3$. Since no proper subset of $e$ dominates $E$ in $G$, the following holds: for every $x \in e$ there is a hyperedge $f_x$ such that $f_x \cap e = \{x\}$.

*Claim.* For all $x, y \in e$, $f_x \setminus e = f_y \setminus e$.

Let $x, y \in e$. The case that $x = y$ is trivial. So we may assume that $x \neq y$.

Suppose that there is a vertex $z \in f_x \setminus (e \cup f_y)$. If there is a vertex $z' \in f_y \setminus (e \cup f_x)$, $G[\{z, f_x, x, e, y, f_y, z'\}] \cong P_7$, a contradiction. Thus, $f_y \setminus (e \cup f_x) = \emptyset$, and so $f_y \setminus e \subseteq f_x \setminus e$.

Since $H$ is a clutter, there is a vertex $u \in f_y \setminus e$. As $f_y \setminus e \subseteq f_x \setminus e$, $u \in (f_x \cap f_y) \setminus e$. Since $|e| \geq 3$, there is a vertex $v \in e \setminus \{x, y\}$. But then $G[\{z, f_x, u, f_y, y, e, v\}] \cong P_7$, a contradiction.

So, $f_x \setminus (e \cup f_y) = \emptyset$ and, for symmetry, $f_y \setminus (e \cup f_x) = \emptyset$. This proves Claim 3. For an illustration, see Fig. 3.

*Claim.* If $|f_x \setminus e| = 1$ for some $x \in e$, $H$ does not admit a 2-coloring.

Assume that $|f_x \setminus e| = 1$ for some $x \in e$. By Claim 3, there is a vertex $v \in V$ such that $f_y \setminus e = \{v\}$ for all $y \in e$.

Suppose that $(A, B)$ is a 2-coloring of $H$. We may assume that $v \in A$. Since for every $z \in e$, $f_z \cap B \neq \emptyset$, $e \subseteq B$ holds, a contradiction. So Claim 3 holds.

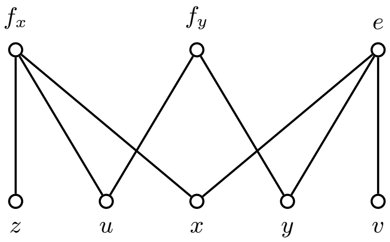


**Fig. 3.** The situation in the proof of Claim 3.

It can be checked in polynomial time whether $|f_x \setminus e| = 1$ for some $x \in e$. In view of Claim 3 and Claim 3, we may now assume that $|f_x \setminus e| \geq 2$ for all $x \in e$.

*Claim.* Let $x, y \in e$ be two arbitrary, distinct vertices and let $z \in f_x \setminus e$. A 2-coloring of $H$ is given by $(\{x, y, z\}, V \setminus \{x, y, z\})$.

Let $x, y, z$ be chosen according to the claim. Suppose that $(\{x, y, z\}, V \setminus \{x, y, z\})$ is not a 2-coloring of $H$. Thus there is an hyperedge $f$ with $f \subseteq \{x, y, z\}$ or $f \cap \{x, y, z\} = \emptyset$.

Let us first assume $f \subseteq \{x, y, z\}$. In particular, $|f \setminus e| \leq 1$. Since $|f_{x'} \setminus e| \geq 2$ for all $x' \in e$, we know that $|f \cap e| \neq 1$. As $N_G(e)$ dominates the set $E$, $|f \cap e| \geq 2$ and so $x, y \in f$. Since $H$ is a clutter, $f \not\subseteq e$, and so $f = \{x, y, z\}$.

By assumption, $|f_x \setminus e| \geq 2$, and so there is a vertex $z' \in f_x \setminus (e \cup \{z\})$. Moreover, since $|e| \geq 3$, there is a vertex $x' \in E \setminus \{x, y\}$. But then $G[\{z', f_x, z, f, y, e, x'\}] \cong P_7$, a contradiction.

So we may assume $f \cap \{x, y, z\} = \emptyset$. As $N_G(e)$ dominates $E$, $e \cap f \neq \emptyset$. Let $x' \in e \cap f$. As $H$ is not a clutter, there is some $z' \in f \setminus e$. This situation is illustrated in Fig. 4.

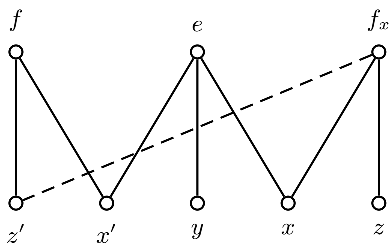


**Fig. 4.** The situation in the proof of Claim 3. The dashed edge is optional.

If $f_x z' \in E(G)$, $G[\{z, f_x, z', f, x', e, y\}] \cong P_7$, a contradiction. Otherwise, $G[\{z, f_x, x, e, x', f, z'\}] \cong P_7$, another contradiction. This proves Claim 3.

Clearly, a 2-coloring as provided by Claim 3 can be constructed efficiently. This completes the proof. □