Discrete Applied Mathematics 373 (2025) 179–195

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Partitioning vertices of graphs into paths of the same length

Oleg Duginov [a,1], Dmitriy Malyshev [b,c,d,*,2], Dmitriy Mokeev [d,b,1]

[a] Belarusian State University of Informatics and Radioelectronics, Gikalo str. 6, 220005 Minsk, Belarus
[b] National Research University Higher School of Economics, Bolshaja Pecherskaja str. 25/12, 603155 Nizhny Novgorod, Russia
[c] Moscow Institute of Physics and Technology, Institutsky lane 9, 141700 Dolgoprudny, Moscow region, Russia
[d] Lobachevsky State University of Nizhny Novgorod, Gagarina ave 23, 603950 Nizhny Novgorod, Russia

ARTICLE INFO

Article history:
Received 3 September 2024
Received in revised form 8 April 2025
Accepted 22 April 2025

Keywords:
Path packing
Path partition
Special graph classes
Computational complexity

ABSTRACT

Given a graph, the (*induced*) P_k*-partition problem* is to decide whether its vertex set can be partitioned into subsets, each of which induces (the k-path) a k-vertex subgraph with a Hamiltonian path. We show that these problems are NP-complete for planar subcubic bipartite $(H_1, H_2, \ldots, H_\ell)$-free graphs of girth g, for any $k, g \geq 3, l \geq 1$, where H_i is obtained by joining central vertices in two copies of P_3 with P_{i+1}. We show that the P_k-partition (induced P_k-partition) problem is NP-complete for split graphs and any $k \geq 5$, chordal graphs and any $k \geq 4$ (any $k \geq 3$), line graphs of planar bipartite graphs and any $k \geq 5$ (any $k \geq 3$). We show that the P_4-partition and, for any $k \geq 5$, induced P_k-partition problems, restricted to split graphs, are polynomial. Additionally, we prove NP-completeness for the optimization version of the induced P_4-partition problem and split graphs.

1. Introduction

1.1. Some packing and partitioning problems in graphs

Let H be a simple graph. The H-Packing problem is to find in a given graph the maximum number of pairwise vertex-disjoint, not necessarily induced, subgraphs, each of which is isomorphic to H. The H-Partition problem is a special decision version of the H-Packing problem, which is to decide whether the vertex set of a given graph can be partitioned into, not necessarily induced, subgraphs, each isomorphic to some fixed graph H. If all the subgraphs are required to be induced, then the last problem is said to be the Induced H-Partition problem. The H-Partition problem arises in many application domains, examples of applications for specific graphs H can be found in [7,10,31]. The Induced H-Partition problem has applications in computer aided electronic circuit board design [23].

The H-Partition and Induced H-Partition problems are widely studied in discrete optimization from structural and complexity perspectives, see, for example, the articles [2,4,6,21,22,24,25,30,38,41]. In this paper, we focus on the complexity issues of the H-Partition problem and its induced version for the case, when H is the k-path P_k. Structural

* Corresponding author at: National Research University Higher School of Economics, Bolshaja Pecherskaja str. 25/12, 603155 Nizhny Novgorod, Russia.

E-mail addresses: oduginov@mail.ru (O. Duginov), dsmalyshev@rambler.ru (D. Malyshev), MokeevDB@gmail.com (D. Mokeev).

1 Results of Sections 1, 2, 4, 5 Supported by Russian Foundation for Basic Research (grant number 20-51-04001) and Belarusian Republican Foundation for Fundamental Research, Republic of Belarus (grant number F21RM-001).

2 Results of Section 3 were supported by the Ministry of Science and Higher Education of the Russian Federation (Goszadaniye) No. 075-03-2024-117, project No. FSMG-2024-0025.

https://doi.org/10.1016/j.dam.2025.04.054
0166-218X/

and complexity aspects of the P_k-Partition problem for $k \leq 3$ are also well-studied [1,4,10,31,33,35]. In this paper, we primarily focus on the case of $k \geq 4$ and continue research, related to the computational complexity of the P_k-Partition problem and its induced version in special graph classes [35,36]. More details are presented in the next subsection.

1.2. Notation, preliminaries, and statements of our results

For any natural n, by $[n]$ we will denote the set $\{1, 2, \ldots, n\}$. We use the standard graph-theoretic terminology, see, for example, [15,39] for any terminology, non-defined in this paper. Unless otherwise stated, we consider only simple graphs $G = (V, E)$, each with vertex set $V = V(G)$ and edge set $E = E(G)$. For any vertex v of a graph G, by $N(v)$ and $\deg(v)$ we denote the neighbourhood and degree of v. The *order* of a graph G is its number of vertices $|V(G)|$. We use the following notation:

(a) P_k is the path of order k,
(b) K_n is the complete graph of order n,
(c) $2K_2$ is the graph consisting of two vertex-disjoint copies of K_2,
(d) $G(U \cup W, E)$ is a bipartite graph with parts U and W,
(e) $K_{p,q}$ is the complete bipartite graph $G = (U \cup W, E)$, such that $|U| = p$ and $|W| = q$,
(f) H_i is the graph, obtained from a $P_i = (x_1, \ldots, x_{i+1})$ by adding vertices y_1, y_2, z_1, z_2 and the edges $\{y_1, x_1\}$, $\{y_2, x_1\}$, $\{z_1, x_k\}$, $\{z_2, x_k\}$,
(g) $L(G)$ is the line graph of a graph G.

Given vertices v_0 and v_k of a graph G, a *walk* (or a v_0, v_k-*walk*) of G is an alternating sequence $v_0, e_1, v_1, e_2, \ldots, v_{k-1}, e_k, v_k$ of vertices v_p, $p \in \{0\} \cup [k]$ and edges e_q, $q \in [k]$, such that $e_i = \{v_{i-1}, v_i\} \in E(G)$, for each $i \in [k]$. The vertices v_0 and v_k are *end-vertices* of the v_0, v_k-walk. The length of the v_0, v_k-walk is the number k. A *trail* is a walk without repeated edges. A *path* is a trail without repeated vertices.

A *subgraph* of a graph G is a graph H, such that $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. Given a non-empty subset $\mathcal{E} \subseteq E(G)$, the subgraph H of a graph G is *induced by the edge set* $\mathcal{E}$ if $E(H) = \mathcal{E}$ and any vertex $v \in V(G)$ belongs to H if and only if v is incident to at least one edge from $\mathcal{E}$. Given a non-empty subset $U \subseteq V(G)$, the subgraph H of a graph G is *induced by the vertex set* U if $V(H) = U$ and any two vertices u and v are adjacent in H if and only if u and v adjacent in G. The subgraph of G, induced by a vertex set U, is denoted by $G[U]$. A subgraph H of a graph G is *induced*, if there exists a non-empty subset $U \subseteq V(G)$, such that $H = G[U]$.

A *clique* of a graph is a set of its pairwise adjacent vertices. An *independent set* of a graph is a set of its pairwise non-adjacent vertices. The *girth* of a graph is the length of its shortest induced cycle.

A *matching* of a graph G is a subset $M \subseteq E(G)$, such that M does not contain two adjacent edges. Given a vertex $v \in V(G)$, we say that the matching M *covers the vertex* v, if v is incident to an edge of M. Given a subset $U \subseteq V(G)$, we say that the matching M *covers the set* U if M covers each vertex from U. The matching M is *perfect* if it covers the vertex set $V(G)$.

Definition 1. Given a graph $G = (V, E)$ and an integer $k \geq 3$, a P_k-*packing* in G is a collection $\mathcal{S}$ of mutually vertex-disjoint, not necessarily induced, subgraphs of G, which are all isomorphic to P_k. A P_k-packing is *induced* if all the subgraphs are induced. The number $|\mathcal{S}|$ of subgraphs in $\mathcal{S}$ is the *size* of $\mathcal{S}$.

For a graph G, we denote by $p_k(G)$ (respectively, by $ip_k(G)$) the maximum size of a P_k-packing (respectively, an induced P_k-packing) in G. The decision problems, associated with these parameters, are defined as follows:

The P_k-Packing problem
Instance: A graph G and an integer b.
Question: Does G have a P_k-packing of size at least b? In other words, is it $p_k(G) \geq b$ or not?

The Induced P_k-Packing problem
Instance: A graph G and an integer b.
Question: Does G have an induced P_k-packing of size at least b? In other words, is it $ip_k(G) \geq b$ or not?

Definition 2. Given a positive integer $k \geq 3$ and a graph $G = (V, E)$, such that the number of vertices $|V|$ is divisible by k, an (induced) P_k-packing $\mathcal{S}$ in G with $|\mathcal{S}| = \frac{|V|}{k}$ is called a (respectively, *induced*) P_k-*factor* of G.

In other words, a P_k-factor of a graph G is a collection of subgraphs of G, such that each subgraph is isomorphic to P_k and each vertex of G is exactly contained in one of the subgraphs. An induced P_k-factor of a graph G is defined similarly with the additional condition that each subgraph is induced. Let us observe that there exist graphs of order $t \cdot k$ with $k \geq 3$ that have a P_k-factor (e.g., Hamiltonian graphs) or an induced P_k-factor (e.g., paths and cycles) and graphs that do not have any P_k-factors and induced P_k-factors (e.g., stars $K_{1,t \cdot k-1}$). Hence, it is natural and reasonable to study the following decision problems:

The P_k-Partition problem
Instance: A graph G.
Question: Does G have a P_k-factor or not?

The Induced P_k-Partition problem
Instance: A graph G.
Question: Does G have an induced P_k-factor or not?

Table 1
The computational complexity of the P_k-Partition/Induced P_k-Partition problems for various graph classes.

Graph classes	$k = 3$	$k = 4$	$k \geq 5$
Cographs	P [10]/?	?	?
Chordal graphs	NPc [10]/**NPc** [Theorem 12]	**NPc** [Theorem 11]/**NPc** [Theorem 12]	**NPc** [Theorem 11]/**NPc** [Theorem 12]
Split graphs	P [10]/P [16]	**P** [Theorem 3]/?	**NPc** [Theorem 8]/**P** [Theorem 9]
Bipartite graphs of maximum degree 3	NPc [35]/NPc [35]	NPc [35]/NPc [35]	NPc [35]/NPc [35]
Planar bipartite graphs of maximum degree 3	NPc [35]/NPc [35]	**NPc** [Theorem 1]/**NPc** [Theorem 2]	**NPc** [Theorem 1]/**NPc** [Theorem 2]
Line graphs of planar bipartite graphs	?/**NPc** [Theorem 13]	?/**NPc** [Theorem 13]	**NPc**/**NPc** [Theorem 13]
Subcubic grids	NP-c [10]/NP-c [10]	?/?	?/?
Interval graphs	P [10]/?	?/?	?/?

The P_k-Partition and P_k-Packing problems are closely related to each other in the following sense [35]: if the P_k-Partition problem is NP-complete for some graph class, then the P_k-Packing problem, restricted to this graph class, is NP-complete as well.

For each fixed $k \geq 3$, the P_k-Partition problem is NP-complete [22]. It remains NP-complete for chordal graphs and $k = 3$ [10], bipartite graphs of maximum degree 3 (for each fixed $k \geq 3$) [35], planar graphs (for each fixed $k \geq 3$) [8], planar bipartite graphs (for each fixed $k \geq 3$) [17,35], planar bipartite graphs of maximum degree 3 and $k = 3$ [35], subcubic grids and $k = 3$ [10]. At the same time, this problem is polynomial-time solvable for interval graphs and $k = 3$ [10], split graphs and $k = 3$ [10], square graphs and $k = 3$ [3], and some graphs G, such that, for any their induced subgraph H, $p_k(H)$ equals the minimum size of vertex subsets, intersecting all k-paths of H [34]. The Induced P_k-Partition problem is NP-complete for bipartite graphs of maximum degree 3, for each fixed $k \geq 3$, planar bipartite graphs of maximum degree 3 and $k = 3$ [35,36], and polynomial-time solvable for split graphs and $k = 3$ [16].

For each fixed $k \geq 3$, the P_k-Packing problem is NP-complete [22,33] and remains NP-complete in any graph class, for which the P_k-Partition problem is NP-complete. On the other hand, for any fixed $k \geq 3$, this problem is polynomial-time solvable for trees [36] and for trees, when k is a part of the input [33]. The Induced P_k-Packing problem is known to be NP-complete for bipartite graphs of maximum degree 3, for each fixed $k \geq 3$, for $k = 3$ and planar bipartite graphs of maximum degree 3 [35]. The P_3-Packing problem has recently been considered in the parameterized complexity framework [14,19,29,40].

Both the P_k-Partition problem and Induced P_k-Partition problem can be modelled, using the framework from [12]. The P_3-Partition problem has been widely studied from the computational complexity perspective. In contrast to this, the case of $k \geq 4$ is little studied. Monnot and Toulouse in [35,36] posed a question about the computational complexity of the P_k-Partition and the Induced P_k-Partition problems, restricted to planar bipartite graphs of maximum degree 3, for any fixed $k \geq 4$. In Section 2, we answer this question. Namely, we show that, for any fixed integers $k \geq 3$, $\ell \geq 1$, and $g \geq 3$, both problems are NP-complete for planar bipartite $(H_1, H_2, \ldots, H_\ell)$-free graphs of maximum degree 3 and girth at least g. In Section 3, we study the computational complexity of the considered problems in the class of split graphs. We show there that the P_k-Partition problem, restricted to split graphs, is polynomial-time solvable for $k = 4$ and NP-complete, for each fixed $k \geq 5$, the Induced P_k-Partition problem is polynomial-time solvable, for each fixed $k \geq 5$. Besides, we also show that the Induced P_4-Packing problem for split graphs is NP-complete. Thus, for split graphs and the P_k-Partition problem, we have the following complete complexity dichotomy in terms of k: This problem is polynomial-time solvable for them if $k \leq 4$, otherwise, it is NP-complete. In Section 4, we study the computational complexity of the problems in the class of chordal graphs. We show that both problems, restricted to chordal graphs, are NP-complete, for each fixed $k \geq 4$. In addition, we note that the Induced P_3-Partition problem for chordal graphs is NP-complete as well. Finally, in Section 5, we study the computational complexity of both problems for line graphs of planar bipartite graphs. We show NP-completeness of the P_k-Partition and Induced P_ℓ-Partition problems, restricted to line graphs of planar bipartite graphs, for each fixed $k \geq 5$ and for each fixed $\ell \geq 3$, respectively.

Table 1 summarizes previously known and our complexity results for the P_k-Partition and Induced P_k-Partition problems. We abbreviate "NP-complete" by "NPc" and "polynomial time solvable" by "P". Bold entries correspond to new results in this paper.

2. The computational complexity of the P_k-Partition / Induced P_k-Partition problems, restricted to planar bipartite graphs of maximum degree 3

The first result is NP-completeness of the considered problems in a subclass of planar bipartite graphs. A proof of this result is based on some well-known local graph transformation technique [5,11,26,32].

Theorem 1. *For any fixed integers $k \geq 3$, $g \geq 3$, and $\ell \geq 1$, the P_k-Partition problem is NP-complete for planar bipartite $(H_1, H_2, \ldots, H_\ell)$-free graphs of girth at least g and maximum degree 3.*

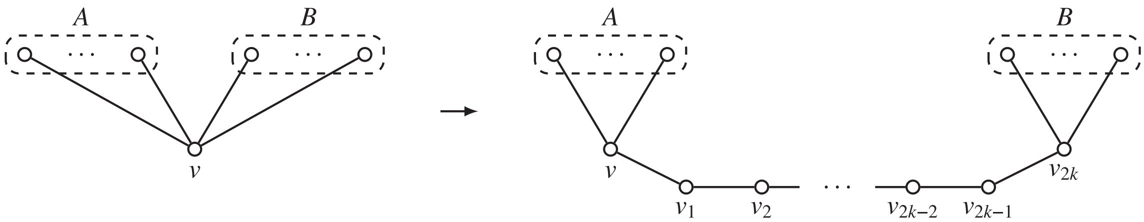

Fig. 1. The vertex $2k$-splitting operation.

Proof. Let $k \geq 3$, $g \geq 3$, and $\ell \geq 1$ be integers. We construct a polynomial-time reduction from the P_k-Partition problem for planar bipartite graphs, which is known to be NP-complete [17,35].

Let $G = (V, E)$ be an arbitrary planar bipartite graph. Let us consider an arbitrary vertex $v \in V$, such that $\deg(v) \geq 2$. The *vertex $2k$-splitting with respect to v* is defined as follows, also see Fig. 1:

(a) arbitrarily split the neighbourhood $N(v)$ of v into two non-empty disjoint subsets A and B,
(b) add a path with vertices $v_1, v_2, \ldots, v_{2k}$,
(c) add the edge $\{v, v_1\}$ and all the edges between v_{2k} and vertices from B,
(d) remove all the edges between v and vertices from B.

We denote by G^v_{2k} the graph, obtained from G by the vertex $2k$-splitting operation with respect to v.

Claim 1. *The graph G has a P_k-factor if and only if G^v_{2k} has a P_k-factor.*

Proof. Assume that G has a P_k-factor $\mathcal{S}$. Our aim is to construct a P_k-factor $\mathcal{S}'$ of G^v_{2k}. Let P_v be the path of $\mathcal{S}$, containing v. We denote the end vertices of P_v by u and w, i.e., the path P_v is a u, w-path of G. We will separately consider two possible cases.

Case 1. Suppose that v is not an end vertex of P_v, i.e., $v \neq u$ and $v \neq w$. In this case, v divides P_v into two subpaths: the u, v-path P^1_v of order t and the v, w-path P^2_v of order $k - t + 1$. Without loss of generality, one may assume that the vertex of P^1_v, adjacent to v, belongs to A. The following subcases are only possible:

Subcase A1. The vertex of P^2_v, adjacent to v, belongs to A. Let us form two paths of order k of G^v_{2k}: one path P, consisting of $v_1, v_2, \ldots, v_k$, and another path Q, consisting of $v_{k+1}, v_{k+2}, \ldots, v_{2k}$. It is not hard to see that $\mathcal{S}' = \mathcal{S} \cup \{P, Q\}$ is a P_k-factor of G^v_{2k}.

Subcase B1. The vertex of P^2_v, adjacent to v, belongs to B. Let us form 3 paths of order k of G^v_{2k}: the path P, consisting of $V(P^1_v)$ and $v_1, v_2, \ldots, v_{k-t}$, the path Q, consisting of $v_{k-t+1}, v_{k-t+2}, \ldots, v_{2k-t}$, and the path R, consisting of $v_{2k-t+1}, v_{2k-t+2}, \ldots, v_{2k}$ and $V(P^2_v) \setminus \{v\}$. It is not hard to see that $\mathcal{S}' = (\mathcal{S} \setminus \{P_v\}) \cup \{P, Q, R\}$ is a P_k-factor of G^v_{2k}.

We obtained a P_k-factor of G^v_{2k} in both subcases. We proceed to the second case.

Case 2. Suppose that v is an end vertex of P_v. Let us form two paths of order k of G^v_{2k}: one path P, consisting of $v_1, v_2, \ldots, v_k$, and another path Q, consisting of $v_{k+1}, v_{k+2}, \ldots, v_{2k}$. It is not hard to see that $\mathcal{S}' = \mathcal{S} \cup \{P, Q\}$ is a P_k-factor of G^v_{2k}. We obtained a P_k-factor of G^v_{2k} in both cases.

Conversely, assume that $\mathcal{S}'$ be a P_k-factor of G^v_{2k}. Our aim is to construct a P_k-factor $\mathcal{S}$ of G. Let P_v be the path of $\mathcal{S}'$, containing v. The only possible cases are the following:

Case 1. The path P_v does not contain v_1. We show that P_v does not contain any vertex from $\{v_1, v_2, \ldots, v_{2k}\}$. Suppose to the contrary that at least one vertex from $\{v_1, v_2, \ldots, v_{2k}\}$ is contained in P_v. Let v_j, $j \in \{2, 3, \ldots, 2k\}$ be the vertex of P_v with the minimum index j. Clearly, v_j is an end vertex of P_v and $j \in \{k+3, k+4, \ldots, 2k\}$. In this case, we have that $v_1, v_2, \ldots, v_{j-1}$ induce one or two paths of order k, belonging to the P_k-factor $\mathcal{S}'$. On the other hand, the quantity $j - 1$ of these vertices is not divisible by k and we obtain a contradiction.

Thus, P_v does not contain any of the vertices $v_1, v_2, \ldots, v_{2k}$. It implies that $\mathcal{S}'$ contains the following two paths: one path P, consisting of $v_1, v_2, \ldots, v_k$, and another path Q, consisting of $v_{k+1}, v_{k+2}, \ldots, v_{2k}$. It is not hard to see that the set $\mathcal{S} = \mathcal{S}' \setminus \{P, Q\}$ is a P_k-factor of G.

Case 2. The path P_v contains v_1. We immediately obtain that at least one vertex of $\{v_1, v_2, \ldots, v_{2k}\}$ is an end vertex of P_v. Let v_{j+1} be the vertex with the minimum index that does not belong to P_v. The vertex v_j, $j \in [k-1]$ is an end vertex of P_v. It implies that the P_k-factor $\mathcal{S}'$ contains the path P of order k, consisting of $v_{j+1}, v_{j+2}, \ldots, v_{j+k}$. We show that v_{j+k+1} is not contained in P_v. Suppose to the contrary that P_v contains v_{j+k+1}. It is not hard to see that P_v contains the vertices $v_{j+k+2}, v_{j+k+3}, \ldots, v_{2k}$ as well. Hence, P_v contains at least $j + 1 + (2k - j - k - 1 + 1) = k + 1$ vertices, which is a contradiction with the assumption that order of P_v is k.

Thus, v_{j+k+1} is not contained in P_v and, hence, v_{j+k+1} is an end vertex of another path $Q \in \mathcal{S}'$. The following subcases are only possible:

Subcase A2. The vertex v is an end vertex of P_v. In this situation, P_v consists of $v, v_1, v_2, \ldots, v_{k-1}$, P consists of $v_k, v_{k+1}, \ldots, v_{2k-1}$, and Q contains v_{2k}. The set $\mathcal{S}$, obtained from $\mathcal{S}'$ by removing P_v, P and replacing v_{2k} in Q by v, is a P_k-factor of G.

Subcase B2. The vertex v is a non-end vertex of P_v. Let us form the path R of order k of G from P_v and Q as follows:

(a) remove $v_1, v_2, \ldots, v_j$ from P_v,
(b) remove $v_{j+k+1}, v_{j+k+2}, \ldots, v_{2k-1}$ from Q,
(c) identify v and v_{2k} of the obtained paths.

The resulting path R is a path of order k of G. It is not hard to see that the set $\mathcal{S} = (\mathcal{S}' \setminus \{P_v, P, Q\}) \cup \{R\}$ is a P_k-factor of G. We obtained a P_k-factor of G in both cases. ⧫

Claim 2. *By means of the vertex 2k-splitting operation, G can be transformed into a planar bipartite $(H_1, H_2, \ldots, H_\ell)$-free graph of girth at least g and maximum degree 3.*

Proof. Assume that G has a vertex v with $\deg(v) \geq 4$. For the vertex $2k$-splitting with respect to v, we split $N(v)$ into two disjoint subsets A and B, such that $|A| \geq 2$ and $|B| = 2$. After it, we obtain a graph with $2k$ new vertices. Clearly, all the degrees of new vertices are at most 3 and degree of v is decreased by one. It is not hard to see that the operation preserves the planarity and bipartiteness. Applying this procedure to every vertex of degree at least 4, we obtain a graph of maximum vertex degree 3.

Let us consider the variant of the vertex $2k$-splitting operation, when $|B| = 1$. It is equivalent to replacing an edge by a path of length $2k+1$. Now, we replace each edge of the planar bipartite graph by a path of length $2k+1$. The length of each induced cycle is increased in $2k+1$ times. Consequently, the resulting graph contains no cycles of odd length, i.e., it is bipartite. As replacing an edge by a path preserves planarity, the resulting graph is planar as well. Applying the operation a necessary number of times, we will avoid induced cycles of orders at most g and induced subgraphs isomorphic to $H_1, H_2, \ldots, H_\ell$.⧫

Applying the handshaking lemma, it is not hard to check that after the transformation of G, described above, we will obtain a graph with at most $2k|E(G)|$ vertices. Therefore, those transformations can be performed in polynomial time. This finishes the proof of this theorem. □

We showed NP-completeness of the P_k-Partition problem for planar bipartite graphs with bridge-like forbidden induced subgraphs and bounded maximum degree and girth. A proof is completely the same for the "induced" version of the P_k-Partition problem. Hence, we have the following result:

Theorem 2. *For any fixed integers $k \geq 3$, $g \geq 3$, and $\ell \geq 1$, the* Induced P_k-Partition *problem is NP-complete for planar bipartite $(H_1, H_2, \ldots, H_\ell)$-free graphs of girth at least g and maximum degree 3.*

3. The computational complexity of the P_k-Partition / Induced P_k-Partition problems, restricted to split graphs

Further, we study the computational complexity of the P_k-Partition problem for split graphs. We will show that it can be solved in polynomial time for $k = 4$ and split graphs (see Theorem 3), but remains NP-complete, for any $k \geq 5$ (see Theorem 8). Theorem 8 will be proven by induction, where the induction base will be NP-completeness of the P_k-Partition problem for split graphs (see Theorems 4,5,6,7) and $k \in \{5, 6, 7, 8\}$ and the induction step will be NP-completeness of the P_{k-4}-Partition problem for split graphs.

Theorem 3. *The P_4-Partition problem is polynomial-time solvable for split graphs.*

Proof. Recall that a split graph G is a graph, whose vertex set can be partitioned into a clique C and an independent set I. We will solve the P_4-Partition problem for G by reducing in polynomial time this problem to the problem of finding a matching in G, covering I.

Claim 3. *The graph G admits a P_4-factor if and only if $|I| + |C|$ is a multiple of 4, and there exists a matching M of G, covering I.*

Proof. Assume that there is a P_4-factor $\mathcal{S}$ of G. Clearly, $|C| + |I|$ is a multiple of 4. Let us form a perfect matching M of G as follows. Initially, we put $M = \emptyset$. For each path P from $\mathcal{S}$, we include two non-adjacent edges of P to M. The obtained set M is a perfect matching of G. Finally, we remove from M all the edges that are not incident to any vertices of the set I. The resulting set M is a matching of G, and it covers I.

Conversely, assume that $|I| + |C|$ is a multiple of 4 and M is a matching of G, such that M covers I. Note that any two edges e_1 and e_2 of M are connected with an edge e of $G[C]$, so that these 3 edges e_1, e_2, and e form P_4. Either M (if $|M|$ is even) or M minus any of its elements e' (if $|M|$ is odd) can be divided into pairs. Each such a pair produces P_4. The edge e' together with any edge of C, not adjacent to any element from M, also forms P_4.

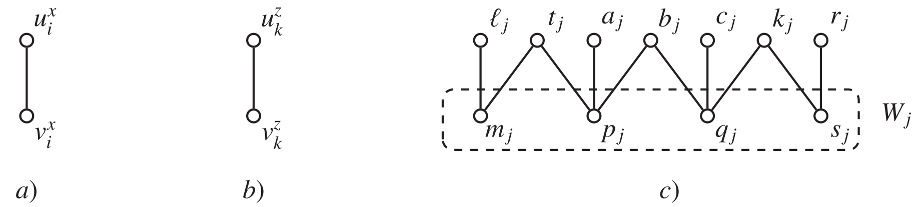

Fig. 2. (a) a component H_i^x, (b) a component H_k^z and (c) a component H_j^y.

Therefore, there is a set $\mathcal{S}$ of pairwise vertex-disjoint 4-paths that covers all the vertices of I. All the remaining non-covered vertices of G belong to C. These vertices can be partitioned into groups of 4 vertices. Each such a group of vertices forms a clique, from which we can pick P_4. We add these paths to $\mathcal{S}$. It is not hard to see that the resulting set $\mathcal{S}$ is a P_4-factor of G. ♦

Let B be the bipartite graph, which is obtained from G by removing all the edges from C. It is not hard to see that there is a matching M in G, covering I, if and only if the maximum size of a matching in B equals $|I|$. As computing a maximum matching in a graph with n vertices and t edges can be done in $O(\sqrt{n}\cdot(n+t))$ time [27, p. 256], we can solve the P_4-Partition problem for G with n vertices and m edges, using Claim 3, in $O(\sqrt{n}\cdot(n+m))$ time. This finishes the proof of this theorem. □

Theorem 4. *The P_5-Partition problem is NP-complete for split graphs.*

Proof. We present a polynomial-time reduction from the 3-Dimensional Matching problem, which is known to be NP-complete [20]. Any instance of this problem consists of a set $\mathcal{C} = \{t_1, t_2, \dots, t_m\} \subseteq X \times Y \times Z$ of triples, such that each element of $X \cup Y \cup Z$ is contained in at most 3 triples from $\mathcal{C}$, where $X = \{x_1, x_2, \dots, x_n\}$, $Y = \{y_1, y_2, \dots, y_n\}$, and $Z = \{z_1, z_2, \dots, z_n\}$ are pairwise disjoint sets. A *3d-matching* is a subset $\mathcal{M}$ of $\mathcal{C}$, such that no two triples in $\mathcal{M}$ agree in at least one coordinate. The 3-Dimensional Matching problem is to decide whether there exists a 3d-matching $\mathcal{M} \subseteq \mathcal{C}$ of size n or not.

Given an instance $I = (X, Y, Z, \mathcal{C})$ of the 3-Dimensional Matching problem, we construct a split graph $G = (V, E)$ as follows:

(a) for each element $x_i \in X$, we create the component H_i^x, illustrated in Fig. 2, (a),
(b) for each element $z_k \in Z$, we create the component H_k^z, illustrated in Fig. 2, (b),
(c) for each element $y_j \in Y$, we create the component H_j^y, illustrated in Fig. 2, (c), and denote the set $\{m_j, p_j, q_j, s_j\}$ by W_j,
(d) next, we perform the following procedure:

for *each triple* $(x_i, y_j, z_k) \in \mathcal{C}$ **do**
  if $\deg a_j = 1$ **then**
    connect v_i^x with a_j and a_j with v_k^z by edges
  if $\deg(a_j) > 1$ *and* $\deg b_j = 2$ **then**
    connect v_i^x with b_j and b_j with v_k^z by edges
  if $\deg(a_j) > 1$ *and* $\deg b_j > 2$ **then**
    connect v_i^x with c_j and c_j with v_k^z by edges

For any $i, j, k \in [n]$, put $N_{ij,k} = N(v_i^x) \cap N(v_k^z) \cap \{a_j, b_j, c_j\}$. By construction of G, we have

$$N_{ijk} \neq \emptyset \Longrightarrow |N_{ijk}| = 1 \Longleftrightarrow (x_i, y_j, z_k) \in \mathcal{C}.$$

(e) finally, we denote $W_1 \cup W_2 \cup \dots W_n \cup \{v_1^x, v_2^x, \dots, v_n^x\} \cup \{v_1^z, v_2^z, \dots, v_n^z\}$ by W and connect any two its distinct vertices by an edge.

We denote the obtained graph by G. Clearly, this graph G is split into the clique W and the independent set $I = V(G) \backslash W$. It is clear that $|W| = 6n$ and $|I| = 9n$.

Let us show that there exists a 3d-matching $\mathcal{M} \subseteq \mathcal{C}$ of size n if and only if G has a P_5-factor. Assume that there is a 3d-matching $\mathcal{M}$ of size n. Let us construct a P_5-factor $\mathcal{S}$ of G as follows. Initially, we put $\mathcal{S} = \emptyset$. For each triple $(x_i, y_j, z_k) \in \mathcal{M}$, we will form paths P, Q, and R in the following way:

(a) if v_i^x, a_j are adjacent and a_j, v_k^z are adjacent, then P, Q, and R are the paths with the vertex sets $\{u_i^x, v_i^x, a_j, v_k^z, u_k^z\}$, $\{\ell_j, m_j, t_j, p_j, b_j\}$, and $\{c_j, q_j, k_j, s_j, r_j\}$, correspondingly,
(b) if v_i^x, b_j are adjacent and b_j, v_k^z are adjacent, then P, Q, and R are the paths with the vertex sets $\{u_i^x, v_i^x, b_j, v_k^z, u_k^z\}$, $\{\ell_j, m_j, t_j, p_j, a_j\}$, and $\{c_j, q_j, k_j, s_j, r_j\}$, correspondingly,

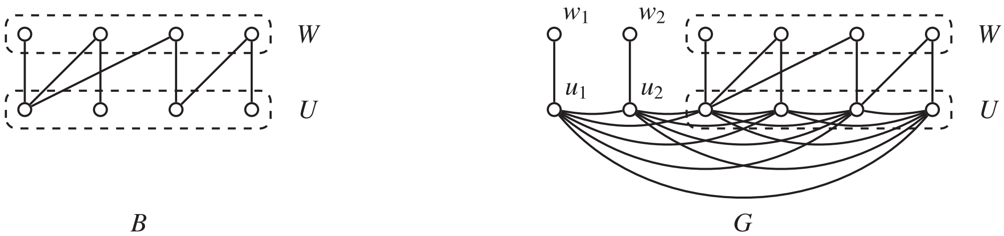

Fig. 3. A balanced bipartite graph B and the graph G, obtained from B.

(c) if v_i^x, c_j are adjacent and c_j, v_k^z are adjacent, then P, Q, and R are the paths with the vertex sets $\{u_i^x, v_i^x, c_j, v_k^z, u_k^z\}$, $\{\ell_j, m_j, t_j, p_j, a_j\}$, and $\{b_j, q_j, k_j, s_j, r_j\}$, correspondingly,

and add these 3 paths P, Q, and R to $\mathcal{S}$. Therefore, the set $\mathcal{M}$ gives $3n$ pairwise vertex-disjoint 5-paths of G, and G has $15n$ vertices. Hence, $\mathcal{S}$ is a 5-factor of G.

Conversely, assume that $\mathcal{S}$ is a P_5-factor of G. We have that $|S| = (|W| + |I|)/5 = 3n$. As I is an independent set, each path from $\mathcal{S}$ contains at most 3 vertices of I. All the vertices of I are covered by paths from $\mathcal{S}$ and $|I| = 9n$. By the pigeonhole principle, each path from $\mathcal{S}$ contains at least 3 vertices of I. Consequently, each path from $\mathcal{S}$ contains exactly 3 vertices of I. It implies that each path from $\mathcal{S}$ does not contain any edge of W.

For any $i \in [n]$, as $\deg(u_i^x) = 1$, there exists a unique path $P_i \in \mathcal{S}$, starting at u_i^x. As u_i^x is adjacent to the single vertex v_i^x, the second vertex of P_i is v_i^x. The path P_i does not contain any edge of W. Hence, the third vertex of P_i is a vertex from $I \setminus \{u_i^x\}$. By construction of G, if v_i^x is adjacent to a vertex w of $I \setminus \{u_i^x\}$, then w can only be either a_j, or b_j, or c_j, for some $j \in [n]$. Let us consider the following three cases, depending on whether $w = a_j$ or $w = b_j$ or $w = c_j$:

Case 1. Suppose that $w = a_j$. By construction of G, we have that $N(a_j) = \{v_i^x, p_j, v_k^z\}$, for some $k \in [n]$. Let us observe that the fourth vertex of P_i cannot be the vertex p_j, otherwise, in $\mathcal{S}$, there is no a path, containing ℓ_j, and, as a consequence, $\mathcal{S}$ is not a P_5-factor of G. Consequently, the fourth vertex is v_k^z. The fifth vertex of P_i is u_k^z, otherwise, there is no a path from $\mathcal{S}$, covering u_k^z. As each path from $\mathcal{S}$ does not contain any edge of W, the P_5-factor $\mathcal{S}$ includes the path with the vertex set $\{\ell_j, m_j, t_j, p_j, b_j\}$ and the path with the vertex set $\{r_j, s_j, k_j, q_j, c_j\}$.

Case 2. Suppose that $w = b_j$. The fourth vertex of P_i is distinct to the vertex p_j, otherwise, in $\mathcal{S}$, there is no a path, containing ℓ_j, and, as a consequence, $\mathcal{S}$ is not a P_5-factor of G. Analogously, the fourth vertex of P_i cannot be the vertex q_j, otherwise, in $\mathcal{S}$, there is no path containing r_j, and, as a consequence, $\mathcal{S}$ is not a P_5-factor of G. By construction of G, $N(b_j) = \{v_i^x, p_j, q_j, v_k^z\}$, for some $k \in [n]$, and the fourth vertex of P_i is v_k^z. As in the previous case, the fifth vertex of P_i is u_k^z. As each path of $\mathcal{S}$ does not contain any edge of W, the P_5-factor $\mathcal{S}$ includes the path with the vertex set $\{\ell_j, m_j, t_j, p_j, a_j\}$ and the path with the vertex set $\{r_j, s_j, k_j, q_j, c_j\}$. The path P_i gives us the triple $\theta_i = (x_i, y_j, z_k) \in \mathcal{C}$, which we include to the set $\mathcal{M}$.

Case 3. Suppose that $w = c_j$. The vertex c_j can actually not have a degree larger than 3, since y_j cannot belong to more than three triples from C. Let us observe that the fourth vertex of P_i cannot be the vertex q_j, otherwise, in $\mathcal{S}$, there is no a path, containing r_j, and, as a consequence, $\mathcal{S}$ is not a P_5-factor of G. Hence, the P_5-factor $\mathcal{S}$ contains 3 paths with the following vertex sets: $\{u_i^x, v_i^x, c_j, v_k^z, u_k^z\}$, $\{\ell_j, m_j, t_j, p_j, a_j\}$, and $\{r_j, s_j, k_j, q_j, b_j\}$, for some $j, k \in [n]$.

Now, we construct a 3d-matching $\mathcal{M} \subseteq \mathcal{C}$ of size n. Initially, we put $\mathcal{M} = \emptyset$. Our idea is to add a triple $\theta_i \in \mathcal{C}$ to $\mathcal{M}$, for each $i \in [n]$. For any $i \in [n]$, the path P_i contains a vertex v_i^x, a unique vertex in $\{a_j, b_j, c_j\}$, and a vertex v_k^z. By construction of G, the triple $\theta_i = (x_i, y_j, z_k)$ is in $\mathcal{C}$, and we include this triple to $\mathcal{M}$. The resulting set $\mathcal{M}$ consists of n triples $\theta_i = (x_i, y_j, z_k)$, $i \in [n]$, which is a 3d-matching, as $\mathcal{S}$ is a P_5-factor. Taking into account that G can be constructed in polynomial time, we obtain the desired polynomial-time reduction. This finishes the proof of this theorem. □

Theorem 5. *The P_6-Partition problem for split graphs is NP-complete.*

Proof. In order to show that it is NP-complete, we construct a polynomial-time reduction from the P_4-Partition problem, restricted to balanced bipartite graphs, which is known to be NP-complete [35, Theorem 3.1].

Let us consider an arbitrary bipartite graph $B = (U \cup W, E)$, such that its order $|U| + |W|$ is a multiple of 4 and $|U| = |W|$. We transform this graph B into a split graph G as follows, also see Fig. 3:

(a) for each $i \in [\frac{|U|+|W|}{4}]$, add new vertices u_i and w_i to the graph, and connect them with an edge,
(b) connect any two distinct vertices of $U \cup \{u_1, u_2, \ldots, u_{\frac{|U|+|W|}{4}}\}$ with an edge.

The obtained graph G is split, in which $I = W \cup \{w_1, w_2, \ldots, w_{\frac{|U|+|W|}{4}}\}$ is an independent set and $C = U \cup \{u_1, u_2, \ldots, u_{\frac{|U|+|W|}{4}}\}$ is a clique.

Claim 4. *The graph B admits a P_4-factor if and only if the graph G admits a P_6-factor.*

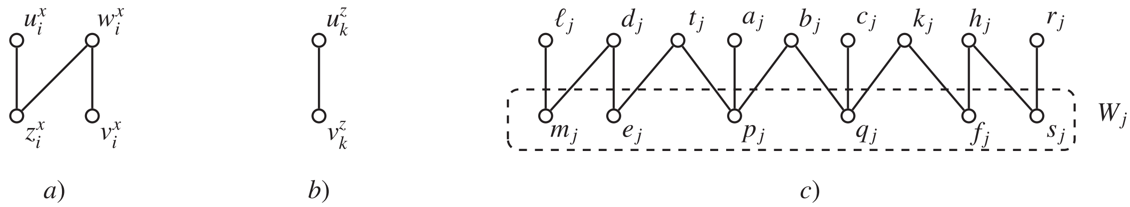

Fig. 4. (a) a component H_i^x, (b) a component H_k^z and (c) a component H_j^y.

Proof. Let $\mathcal{S}$ be a P_4-factor of B. The number of paths from $\mathcal{S}$ equals $\frac{|U|+|W|}{4}$. Append to each path from $\mathcal{S}$ an edge of the form $\{u_i, w_i\}$, $i \in [\frac{|U|+|W|}{4}]$, so that each edge $\{u_i, w_i\}$, $i \in [\frac{|U|+|W|}{4}]$ is appended to exactly one path from $\mathcal{S}$. We obtain a collection of 6-paths of G. This collection is a P_6-factor of G.

Conversely, let us consider a P_6-factor $\mathcal{S}$ of G. By construction of G, we have that $|\mathcal{S}| = \frac{|U|+|W|}{4}$. The following observations are easy to see:

- (a) the graph G does not contain a 6-path P, such that P has more than 3 vertices of I,
- (b) as $|I| = |C|$, each path from $\mathcal{S}$ contains exactly 3 vertices of C and 3 vertices of I,
- (c) there is no a path from $\mathcal{S}$, containing, simultaneously, a vertex of $\{u_i, w_i\}$ and a vertex of $\{u_j, w_j\}$, where $i \neq j$,
- (d) each path from $\mathcal{S}$ starts at a vertex w_i, $i \in [\frac{|U|+|W|}{4}]$ and does not contain edges of $G[U]$.

Consequently, in order to obtain a P_4-factor of B, it suffices to delete vertices w_i and u_i, $i \in [\frac{|U|+|W|}{4}]$, from each path of $\mathcal{S}$.♦

It is not hard to see that the split graph G can be constructed in polynomial time. It finishes the proof of this theorem. □

Theorem 6. *The P_7-Partition problem is NP-complete for split graphs.*

Proof. This proof is similar to the proof of Theorem 4. We present a polynomial-time reduction from the NP-complete 3-Dimensional Matching problem [20].

Let $(X, Y, Z, \mathcal{C})$ be an instance of the 3-Dimensional Matching problem. We construct the split graph G as in the proof of Theorem 4 with the following modifications: instead of the components H_i^x, H_k^z, and H_j^y, illustrated in Fig. 2, we use the components, illustrated in Fig. 4; in (c), we denote by W_j the set $\{m_j, e_j, p_j, q_j, f_j, s_j\}$; we put

$$W = W_1 \cup W_2 \cup \cdots \cup W_n \cup \{z_1^x, z_2^x, \ldots, z_n^x\} \cup \{v_1^x, v_2^x, \ldots, v_n^x\} \cup \{v_1^z, v_2^z, \ldots, v_n^z\}$$

and connect with an edge any two distinct vertices of W.

The obtained graph G is split with a clique W and an independent set $I = V(G) \setminus W$. Let us observe that $|W| = 9n$ and $|I| = 12n$.

We will show that there exists a 3d-matching $\mathcal{M} \subseteq \mathcal{C}$ of size n if and only if G has a P_7-factor. Assume that $\mathcal{M} \subseteq \mathcal{C}$ is a 3d-matching of size n. We construct a P_7-factor $\mathcal{S}$ of G as follows. Initially, we put $\mathcal{S} = \emptyset$. For each triple $(x_i, y_j, z_k) \in \mathcal{M}$, we form 3 paths P, Q, and R in the following way:

- (a) if v_i^x, a_j are adjacent and a_j, v_k^z are adjacent, then P, Q, and R are the paths with the vertex sets $\{u_i^x, z_i^x, w_i^x, v_i^x, a_j, v_k^z, u_k^z\}$, $\{\ell_j, m_j, d_j, e_j, t_j, p_j, b_j\}$, and $\{r_j, s_j, h_j, f_j, k_j, q_j, c_j\}$, respectively,
- (b) if v_i^x, b_j are adjacent and b_j, v_k^z are adjacent, then P, Q, and R are the paths with the vertex sets $\{u_i^x, z_i^x, w_i^x, v_i^x, b_j, v_k^z, u_k^z\}$, $\{\ell_j, m_j, d_j, e_j, t_j, p_j, a_j\}$, and $\{r_j, s_j, h_j, f_j, k_j, q_j, c_j\}$, respectively,
- (c) if v_i^x, c_j are adjacent and c_j, v_k^z are adjacent, then P, Q, and R are the paths with the vertex sets $\{u_i^x, z_i^x, w_i^x, v_i^x, c_j, v_k^z, u_k^z\}$, $\{\ell_j, m_j, d_j, e_j, t_j, p_j, a_j\}$, and $\{r_j, s_j, h_j, f_j, k_j, q_j, b_j\}$, respectively,

and add these paths P, Q, and R to $\mathcal{S}$. It is not hard to see that $\mathcal{S}$ is a P_7-factor of G.

Conversely, assume that $\mathcal{S}$ is a P_7-factor of G. Let us observe that $|\mathcal{S}| = (|W| + |I|)/7 = 3n$. As I is an independent set, each path from $\mathcal{S}$ contains at most 4 vertices of I. All the vertices of I are contained in paths from $\mathcal{S}$ and $|I| = 12n$. By the pigeonhole principle, we have that each path of $\mathcal{S}$ contains at least 4 vertices of I. Hence, each path from $\mathcal{S}$ contains exactly 4 vertices of I. It is not hard to see that each path from $\mathcal{S}$ does not contain any edge of W. We may recover a 3d-matching $\mathcal{M}$ of size n in a similar way as in the proof of Theorem 4. The proof of this theorem is complete. □

Theorem 7. *The P_8-Partition problem is NP-complete for split graphs.*

Proof. In order to show that it is NP-complete, we present a polynomial-time reduction from the P_6-Partition problem, restricted to balanced bipartite graphs, which is known to be NP-complete [35, Theorem 3.1].

Let us consider an arbitrary bipartite graph $B = (U \cup W, E)$, such that its order $|U| + |W|$ is a multiple of 6 and $|U| = |W|$. We transform this graph B into a split graph G as follows:

(a) for each $i \in [\frac{|U|+|W|}{6}]$, add to the graph two additional new vertices u_i and w_i, connected by an edge,
(b) connect any two distinct vertices of $U \cup \{u_1, u_2, \ldots, u_{\frac{|U|+|W|}{6}}\}$ by an edge.

The obtained graph G is split, in which $I = W \cup \{w_1, w_2, \ldots, w_{\frac{|U|+|W|}{6}}\}$ is an independent set and $C = U \cup \{u_1, u_2, \ldots, u_{\frac{|U|+|W|}{6}}\}$ is a clique.

Claim 5. *The graph B admits a P_6-factor if and only if G admits a P_8-factor.*

Proof. Let $\mathcal{S}$ be a P_6-factor of B. We have that $|\mathcal{S}| = \frac{|U|+|W|}{6}$. Append an edge $\{u_i, w_i\}$, $i \in [\frac{|U|+|W|}{6}]$ to each path from $\mathcal{S}$, so that each edge $\{u_i, w_i\}$, $i \in [\frac{|U|+|W|}{6}]$ is appended to exactly one path from $\mathcal{S}$. We obtain a collection of 8-paths of G, which is a P_8-factor of G. Conversely, let us consider a P_8-factor $\mathcal{S}$ of G. By construction of G, we have that $|\mathcal{S}| = \frac{|U|+|W|}{6}$. The following observations are easy to see:

(a) the graph G contains no a path P of order 8, such that P contains more than 4 vertices of I,
(b) as $|I| = |C|$, every path from $\mathcal{S}$ contains exactly 4 vertices of C and 4 vertices of I,
(c) there is no a path from $\mathcal{S}$, containing, simultaneously, a vertex of $\{u_i, w_i\}$ and a vertex of $\{u_j, w_j\}$, where $i \neq j$,
(d) every path from $\mathcal{S}$ starts at a vertex w_i, $i \in [\frac{|U|+|W|}{6}]$ and contains no edges of $G[U]$.

Consequently, in order to obtain a P_6-factor of B, it suffices to delete vertices w_i and u_i, $i \in [\frac{|U|+|W|}{6}]$ from each path from $\mathcal{S}$.♦

It is not hard to see that the split graph G can be constructed in polynomial time. The proof of this theorem is complete. □

Theorem 8. *For each fixed integer $k \geq 5$, the P_k-PARTITION problem is NP-complete for split graphs.*

Proof. We use the mathematical induction method. The claim for $k = 5$, $k = 6$, $k = 7$, and $k = 8$ is the basis step, given by Theorems 4, 5, 6, and 7.

Let $k \geq 9$ be an arbitrary integer. By the inductive hypothesis, the P_{k-4}-PARTITION problem for split graphs is NP-complete. We give a polynomial-time reduction from the P_{k-4}-PARTITION problem, restricted to split graphs.

Given a split graph G with a clique C and an independent set I, such that $|C| + |I|$ is a multiple of $(k - 4)$, we construct a split graph G^* as follows: add $2t$ paths of order 2 with the vertex sets $\{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_{2t}, y_{2t}\}$, such that $x_1, x_2, \ldots, x_{2t}$ are adjacent to all the vertices of G, where $t = (|C| + |I|)/(k - 4)$. It is not hard to see that the obtained graph G^* can be split into the clique $C \cup \{x_1, x_2, \ldots, x_{2t}\}$ and the independent set $I \cup \{y_1, y_2, \ldots, y_{2t}\}$. The graph G^* can be constructed in polynomial time.

We claim that G has a P_{k-4}-factor if and only if G^* has a P_k-factor. Let $\mathcal{S} = \{S_1, S_2, \ldots, S_t\}$ be a P_{k-4}-factor of G and $t = (|C|+|I|)/(k-4)$. We transform each path S_i, $i \in [t]$ of order $k-4$ into a path S_i^* of order k by adding 4 vertices x_i, y_i and x_{t+i}, y_{t+i}. Clearly, $S_1^*, S_2^*, \ldots, S_t^*$ form a P_k-factor of G^*. Conversely, assume that there is a P_k-factor $\mathcal{S}^* = \{S_1^*, S_2^*, \ldots, S_t^*\}$ of G^*. As $y_1, y_2, \ldots, y_{2t}$ are leaves, both end vertices of each path S_i^*, $i \in [t]$ are elements of the set $\{y_1, y_2, \ldots, y_{2t}\}$. It is not hard to see that each path S_i^*, $i \in [t]$ contains exactly two vertices of $\{x_1, x_2, \ldots, x_{2t}\}$. Now, we transform the paths $S_1^*, S_2^*, \ldots, S_{t-1}^*$, and S_t^* of order k into paths $S_1, S_2, \ldots, S_{t-1}$, and S_t of order $k-4$ by removing vertices of $\{x_1, x_2, \ldots, x_{2t}\}$, and $\{y_1, y_2, \ldots, y_{2t}\}$. The paths $S_1, S_2, \ldots, S_{t-1}$, and S_t form a P_{k-4}-factor of G. The proof of this theorem is complete. □

Next, we will study the computational complexity of the "induced" version of the P_k-PARTITION problem for split graphs.

Theorem 9. *For each fixed integer $k \geq 5$, the* INDUCED P_k-PARTITION *problem can be solved in polynomial time for split graphs.*

Proof. Let $k \geq 5$ be an integer. Any split graph does not contain $2K_2$ as an induced subgraph [13, p. 106]. On the other hand, P_k has $2K_2$ as an induced subgraph. Consequently, split graphs do not admit induced P_k-factors and the INDUCED P_k-PARTITION problem for split graphs can be solved in a constant time. □

The computational status of the INDUCED P_4-PARTITION problem, restricted to split graphs, remains open. We suspect that it is NP-complete. Further, we will prove a weaker claim by showing that the C_4-PACKING problem is NP-complete for bipartite tournaments.

Let $G = (V, E)$ be a split graph with a clique C and an independent set I, such that $|C| = |I|$. Given an induced P_4-packing $\mathcal{S}$ of G, we have that each path from $\mathcal{S}$ has 4 vertices, say a, b, c, d, where b, c belong to C and a, d belong to I. Let us transform G into a bipartite graph $H = (C \cup I, \mathcal{E})$ by removing all the edges between distinct vertices of C. Therefore, all the induced subgraphs of G, each isomorphic to P_4, bijectively correspond to all the induced subgraphs of H, each isomorphic to $2K_2$. Hence, all the induced P_4-packings of G bijectively correspond all the induced $2K_2$-packings of H. Consequently,

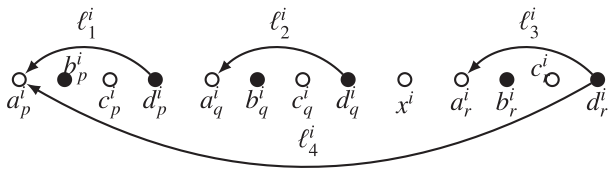

Fig. 5. A bipartite tournament L_i, associated with a variable x_i, $i \in [n]$.

the Induced P_4-Packing problem, restricted to split graphs, is polynomial-time equivalent to the Induced $2K_2$-Packing problem for bipartite graphs. We carry over the latter problem to bipartite tournaments. Recall that a *bipartite tournament* is an oriented complete bipartite graph.

Suppose that $H = (C \cup I, \mathcal{E})$ is an arbitrary bipartite graph with parts C and I, where $|C| = |I|$. Let us define a bipartite tournament $T = (C \cup I, A)$ with parts C and I, such that, for each pair of vertices $u \in C$ and $v \in I$, the tournament T contains the arc (u, v), if $\{u, v\} \in \mathcal{E}$, or the arc (v, u), if $\{u, v\} \notin \mathcal{E}$. It is not hard to see that all the induced $2K_2$-packings of H bijectively correspond to collections of mutually vertex-disjoint (oriented) 4-cycles in T. Thus, the Induced $2K_2$-Packing problem for bipartite graphs is polynomial-time equivalent to the C_4-Packing problem for balanced bipartite tournaments: Given a bipartite tournament $T = (C \cup I, A)$, such that $|C| = |I|$, and an integer k, does the tournament T have at least k mutually vertex-disjoint 4-cycles? A close problem, in which it is asked whether there exist at least k mutually vertex-disjoint 3-cycles in a given *tournament*, i.e., an oriented complete graph, is known to be NP-complete [9].

Theorem 10. *The C_4-Packing problem is NP-complete for bipartite tournaments.*

Proof. We construct a polynomial-time reduction from a special case of the well-known 1-In-3 Satisfiability problem, which is a variant of the 3-Satisfiability problem, where it is asked whether there is an assignment of Boolean values to variables, such that exactly one of the literals in each clause is true. Moore and Robson in [37] showed NP-completeness of the 1-In-3 Satisfiability problem in the case, when

(a) a conjunctive normal form Φ is defined over variables $x_1, x_2, \dots, x_n$ and consists of clauses $C_1, C_2, \dots, C_n$, n is divided by 3;
(b) each clause consists of three distinct variables without negations;
(c) each variable is included in exactly three clauses;
(d) the graph, whose vertices are clauses and variables, edges join clauses and the variables that they contain, is planar.

This special case of the 1-In-3 Satisfiability problem is called the Cubic Planar Monotone 1-In-3 Satisfiability problem.

Let Φ be a conjunctive normal form from an instance of the Cubic Planar Monotone 1-In-3 Satisfiability problem. We construct a bipartite tournament T as follows. For each variable x_i, $i \in [n]$, which is included in three clauses, say C_p, C_q, and C_r, where $p < q < r$, we create a bipartite tournament L_i, see Fig. 5, with parts $U_i = \{a_p^i, c_p^i, a_q^i, c_q^i, x^i, a_r^i, c_r^i\}$ and $W_i = \{b_p^i, d_p^i, b_q^i, d_q^i, b_r^i, d_r^i\}$. Vertices of this tournament are sequentially arranged one after another as depicted in Fig. 5. The tournament L_i includes exactly 4 backward arcs, i.e., arcs from right to left, which are $\ell_1^i = (d_p^i, a_p^i)$, $\ell_2^i = (d_q^i, a_q^i)$, $\ell_3^i = (d_r^i, a_r^i)$, $\ell_4^i = (d_r^i, a_p^i)$. All the remaining undepicted arcs between vertices of U_i and vertices of W_i are arcs from left to right.

We put together the bipartite tournaments $L_1, L_2, \dots, L_n$ into one bipartite tournament L with parts $X = U_1 \cup U_2 \cup \dots \cup U_n$ and $Y = W_1 \cup W_2 \cup \dots \cup W_n$. To this end, we arrange $L_1, L_2, \dots, L_n$ one after another in the natural order. Let us consider two arbitrary tournaments L_i and L_j, $i < j$. Arcs between these tournaments are defined from left to right, i.e., from L_i to L_j. In other words, L is the concatenation of $L_1, L_2, \dots, L_n$ with arcs from L_i to L_j, $i < j$.

For each clause C_p, $p \in [n]$, we create two new vertices C_p^1 and C_p^2. We add new vertices $C_1^1, C_1^2, C_2^1, C_2^2, C_3^1, C_3^2, \dots, C_n^1, C_n^2$ to the end of L. After this, we add arcs in order to obtain a bipartite tournament T with parts $A = X \cup \{C_1^2, C_2^2, \dots, C_n^2\}$ and $B = Y \cup \{C_1^1, C_2^1, \dots, C_n^1\}$. Add the arc (C_p^1, C_p^2), for each $p \in [n]$. For each pair of indices $p, q \in [n]$, such that $p < q$, we add the arcs (C_p^1, C_q^2) and (C_p^2, C_q^1). For each clause $C_p = (x_i \vee x_j \vee x_k)$, $p \in [n]$, and $i < j < k$, add the three backward arcs $h_p^i = (C_p^2, b_p^i)$, $h_p^j = (C_p^2, b_p^j)$, $h_p^k = (C_p^2, b_p^k)$, see Fig. 6, and all the remaining possible arcs from vertices of $L_1, L_2, \dots, L_n$ to the vertices C_p^1 and C_p^2.

Claim 6. *There is an assignment of Boolean values to $x_1, x_2, \dots, x_n$, such that in each clause C_p, $p \in [n]$ exactly one of its variables is true if and only if there are at least $\frac{10n}{3}$ mutually vertex-disjoint 4-cycles of T.*

Proof. Let us consider an assignment of Boolean values to $x_1, x_2, \dots, x_n$, such that exactly one of its variables in each clause C_p, $p \in [n]$ is true. Let us construct a set $\mathcal{S}$ of mutually vertex-disjoint 4-cycles of T. Initially, we put $S = \emptyset$. Let x_i be an arbitrary variable, $i \in [n]$. This variable is exactly included in three clauses, say C_p, C_q, and C_r, where $p < q < r$. The considered variable x_i corresponds to a sub-tournament L_i of T, see Fig. 7.

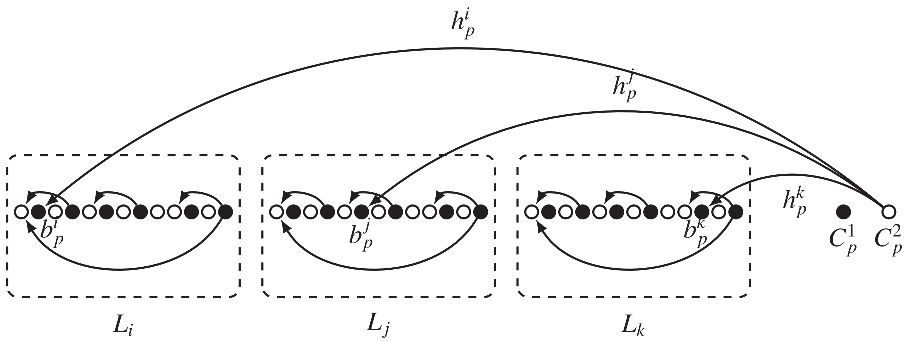

Fig. 6. A fragment of the tournament T.

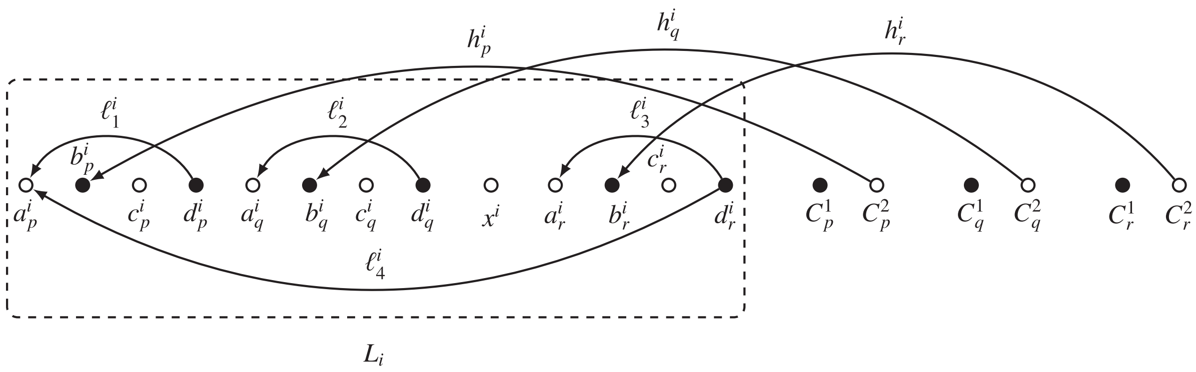

Fig. 7. A fragment of the tournament T.

If the value of x_i is true, then add to $\mathcal{S}$ the 4-cycle on $b_p^i, c_p^i, C_p^1, C_p^2$, the 4-cycle on $b_q^i, c_q^i, C_q^1, C_q^2$, the 4-cycle on $b_r^i, c_r^i, C_r^1, C_r^2$, and the 4-cycle on a_p^i, d_q^i, x^i, d_r^i. Otherwise, add to $\mathcal{S}$ the 4-cycles on

$$a_p^i, b_p^i, c_p^i, d_p^i;\ \ a_q^i, b_q^i, c_q^i, d_q^i;\ \ a_r^i, b_r^i, c_r^i, d_r^i.$$

Having this done for each variable $x_i, i \in [n]$, one after another, we will obtain the set $\mathcal{S}$ of mutually vertex-disjoint 4-cycles of T, such that $|\mathcal{S}| = 4 \cdot \frac{n}{3} + 3 \cdot \frac{2n}{3} = \frac{10n}{3}$.

If $a = (u, v)$ is an arc of T, then u is the *initial vertex* and v is the *terminal vertex* of a. Let $\mathcal{S}$ be a set of mutually vertex-disjoint 4-cycles of T, such that $|\mathcal{S}| \geq \frac{10n}{3}$. Without loss of generality, we may assume that $\mathcal{S}$ is a set of maximum cardinality. Let X be the set of arcs h_p^i, which are in $\mathcal{S}$, where $i, p \in [n]$. Our aim is to transform $\mathcal{S}$ without decreasing its size, so that each 4-cycle from the set, having $h_p^i \in X$, satisfies the following condition (C1): it contains c_p^i and C_p^1. We consider all the arcs from X in the increasing lexicographical order of (i, p). The terminal vertices of arcs from X form an increasing sequence.

Let us consider arcs from X in the aforementioned lexicographical order, starting from the first arc. Let $h_p^i \in X$ and $C \in \mathcal{S}$ be the 4-cycle with the arc h_p^i. If it contains c_p^i and C_p^1, then we do nothing. Otherwise, we correct 4-cycles from $\mathcal{S}$, such that C will satisfy the condition (C1).

Assume that c_p^i is included in $C' \in \mathcal{S}$ and $C' \neq C$. As c_p^i is not contained in any backward arc of T and there are no two backward arcs of T, such that the initial vertex of one arc is the terminal vertex of another arc, the 4-cycle C' contains exactly one backward arc of T, namely the backward arc ℓ_4^i. Let us focus on C' and h_p^i. Fig. 8 illustrates the only four possible cases of their mutual location:

In any case, x^i is either not included in any 4-cycle from $\mathcal{S}$ or included in C. Indeed, let us consider the situation, in which x^i belongs to a 4-cycle from $\mathcal{S}$. As x^i is not contained in any backward arcs of T, the 4-cycle from $\mathcal{S}$ with x^i contains exactly one backward arc, namely an arc $h_q^j \in X$, such that its terminal vertex is at most the terminal vertex of h_p^i. It means that $j \leq i$ and $q \leq p$, if $j = i$. The cases $j < i$ and $j = i, q < p$ are not possible, since the 4-cycle from $\mathcal{S}$ with h_q^j contains c_q^j and C_q^1, not the vertex x^i. Hence, $i = j$ and $q = p$. Consequently, $h_p^i = h_q^j$ and x^i is contained in C.

If x^i does not belong to C, then we replace c_p^i by x^i in C'. After this, c_p^i is not involved in any 4-cycle from $\mathcal{S}$. Assume that x^i is in C. Then, we exchange x^i and c_p^i in C and C'. After this, c_p^i belongs to C. So, we may assume that either c_p^i belongs to C or it does not belong to any 4-cycle from $\mathcal{S}$.

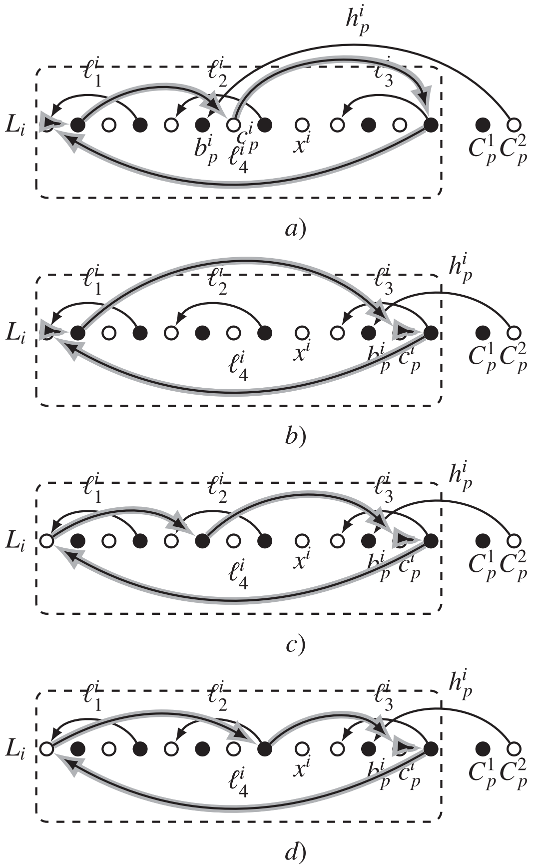

Fig. 8. The four possible cases for mutual location of C', depicted by bold lines, and h_p^i.

Assume that the vertex C_p^1 is contained in a 4-cycle $C' \in \mathcal{S}$ and $C' \neq C$. As C_p^1 is not contained in any backward arc of T, the cycle C' contains exactly one backward arc of T, which is an arc $h_{p_1}^{i_1} \in X$, $p_1 > p$. Similarly, if $C_{p_1}^1$ is contained in a 4-cycle from $\mathcal{S}$, then this 4-cycle contains a single backward arc $h_{p_2}^{i_2} \in X$, $p_2 > p_1$. If $C_{p_2}^1$ is contained in a 4-cycle from $\mathcal{S}$, then this 4-cycle contains a single backward arc $h_{p_3}^{i_3} \in X$, $p_3 > p_2$ and so on. Thus, we obtain a sequence of indices $p = p_0 < p_1 < p_2 < \cdots < p_m$, such that

(a) for each $k \in [m]$, there is a 4-cycle from $\mathcal{S}$, which contains the vertex $C_{p_{k-1}}^1$, the arc $h_{p_k}^{i_k} = (C_{p_k}^2, b_{p_k}^{i_k})$ and, as a consequence, the arc $(C_{p_{k-1}}^1, C_{p_k}^2)$;

(b) the vertex $C_{p_m}^1$ is not contained in any 4-cycle from $\mathcal{S}$.

The latter condition follows from the fact that any 4-cycle of T with an arc $h_p^i \in X$ contains at most one vertex C_q^1, $q \in [n]$. Finally, we reconstruct all the 4-cycles from $\mathcal{S}$ as follows. For each $k \in [m]$, we replace $C_{p_{k-1}}^1$ by $C_{p_k}^1$ in the 4-cycle from $\mathcal{S}$, which contains $C_{p_{k-1}}^1$. After this, C_p^1 is not included in any 4-cycle from $\mathcal{S}$. So, we may assume that either C_p^1 belongs to C or it does not belong to any 4-cycle from $\mathcal{S}$

Now, replace the two vertices of C, different from b_p^i and C_p^2, by c_p^i and C_p^1 (if necessary). Having done this transformation for each arc $h_p^i \in X$, one after another, we obtain the set $\mathcal{S}$, such that each its 4-cycle, containing h_p^i, satisfies the condition (C1). Thus, each 4-cycle from $\mathcal{S}$ with an arc $h_p^i \in X$ consists of $b_p^i, c_p^i, C_p^1, C_p^2$. Note that any other 4-cycle from $\mathcal{S}$ is located inside its L_i, $i \in [n]$.

We will transform $\mathcal{S}$ of mutually vertex-disjoint 4-cycles into a set $\mathcal{S}'$ of mutually vertex-disjoint 4-cycles without decreasing its size. For each variable x_i, $i \in [n]$, that is included in 3 clauses, say C_p, C_q, and C_r, the set $\mathcal{S}'$ will contain

(a) the three 4-cycles of a sub-tournament L_i: on $a_p^i, b_p^i, c_p^i, d_p^i$, on $a_q^i, b_q^i, c_q^i, d_q^i$, and on $a_r^i, b_r^i, c_r^i, d_r^i$;

or

(*b*) the four 4-cycles: on $b_p^i, c_p^i, C_p^1, C_p^2$, on $b_q^i, c_q^i, C_q^1, C_q^2$, on $b_r^i, c_r^i, C_r^1, C_r^2$, and on a_p^i, d_q^i, x^i, d_r^i.

Put $\mathcal{S}' = \emptyset$. Let us consider an arbitrary variable x_i, $i \in [n]$, which is simultaneously included in clauses C_p, C_q, and C_r. This variable corresponds to a sub-tournament L_i of T, see Fig. 7. The backward arcs h_p^i, h_q^i, and h_r^i enter the vertices b_p^i, b_q^i, and b_r^i, respectively. The only following five cases are possible:

Case 1. None of the arcs h_p^i, h_q^i, and h_r^i is contained in 4-cycles from $\mathcal{S}$. In this case, the set $\mathcal{S}$ contains all the three 4-cycles, pointed out in *a*), otherwise, $\mathcal{S}$ is not a maximum-cardinality set of mutually vertex-disjoint 4-cycles of T. Add these three cycles to $\mathcal{S}'$.

Case 2. All the arcs h_p^i, h_q^i, and h_r^i are contained in 4-cycles from $\mathcal{S}$. In this case, $\mathcal{S}$ contains all the 4 cycles, pointed out in (*b*). Add these four cycles to $\mathcal{S}'$.

Case 3. Exactly one of the arcs h_p^i, h_q^i, and h_r^i is contained in 4-cycles from $\mathcal{S}$. It is easy to check that $\mathcal{S}$ contains at most three 4-cycles, whose vertex sets are subsets of $V(L_i) \cup \{C_p^1, C_p^2, C_q^1, C_q^2, C_r^1, C_r^2\}$. Add to $\mathcal{S}'$ the three 4-cycles, pointed out in *a*).

Case 4. There are a 4-cycle $C_1 \in \mathcal{S}$, containing h_q^i, and a 4-cycle $C_2 \in \mathcal{S}$, containing h_p^i or h_r^i. It is easy to check that $\mathcal{S}$ contains at most one 4-cycle, located in L_i. Add to $\mathcal{S}'$ the three 4-cycles, pointed out in *a*).

Case 5. There are a 4-cycle $C_1 \in \mathcal{S}$, containing h_p^i, and a 4-cycle $C_2 \in \mathcal{S}$, containing h_r^i. It is easy to check that $\mathcal{S}$ contains at most one 4-cycle, located in L_i. Add to $\mathcal{S}'$ the three 4-cycles, pointed out in *a*).

Having this transformation done for each variable x_i, $i \in [n]$, we obtain the set $\mathcal{S}'$ of mutually vertex-disjoint 4-cycles of T, such that $|\mathcal{S}'| \geq |\mathcal{S}|$ and, for each variable x_i, $i \in [n]$, which is included in three clauses, say C_p, C_q, and C_r, the set $\mathcal{S}'$ contains either three 4-cycles from *a*) or four 4-cycles from *b*).

For any $i \in [n]$, assign $x_i = true$ if $\mathcal{S}'$ contains four 4-cycles from *b*) and $x_i = false$ if $\mathcal{S}'$ contains three 4-cycles from *a*). Let us denote the number of variables with the true value by k. We obtain that

$$4k + 3(n-k) = |\mathcal{S}'| \geq |\mathcal{S}| \geq \frac{10n}{3}$$

and, as a consequence of this, $k \geq \frac{n}{3}$. As $k \geq \frac{n}{3}$ and $\mathcal{S}'$ consists of mutually vertex-disjoint 4-cycles, each pair of vertices C_p^1 and C_p^2, $p \in [n]$, is contained in a 4-cycle from $\mathcal{S}'$ and, as a consequence of this, exactly one of the variables from each clause C_p, $p \in [n]$, is true.

The bipartite tournament T can be constructed in polynomial time. This completes the proof. □

Corollary 1. *The* Induced P_4-Packing *problem is NP-complete for split graphs.*

4. The computational complexity of the P_k-Partition / Induced P_k-Partition problems, restricted to chordal graphs

Now, we study the computational complexity of the P_k-Partition / Induced P_k-Partition problems, restricted to chordal graphs. NP-completeness of the P_3-Partition problem for chordal graphs is proved in [10]. We extend this result to arbitrary k.

Theorem 11. *For each fixed integer $k \geq 4$, the P_k-Partition problem is NP-complete for chordal graphs.*

Proof. Suppose that $k \geq 5$. As the class of split graphs is a subclass of the chordal graphs class and the P_k-Partition problem for split graphs is NP-complete, the P_k-Partition problem remains NP-complete for chordal graphs as well. Hence, it remains to prove the statement for $k = 4$. We slightly modify the reduction from the 3-Dimensional Matching problem to the P_3-Partition problem, restricted to chordal graphs from [10, p. 332]. Let $(X, Y, Z, \mathcal{C})$ be an instance of the 3-Dimensional Matching problem, where X, Y, and Z are pairwise disjoint sets, such that $|X| = |Y| = |Z| = n$ and $\mathcal{C} \subseteq X \times Y \times Z$. Let us construct a graph G as follows (similar to [10, p. 332]):

(*a*) for each element $e \in X \cup Y \cup Z$, add a pair of adjacent vertices u_e and u'_e;
(*b*) for each triple $t = (x, y, z) \in \mathcal{C}$, add three vertices v_x^t, v_y^t, v_z^t and connect by edges v_x^t with v_y^t and u_x, v_z^t with v_y^t and u_z, v_y^t with u_x and u_y, u_z;
(*c*) add an edge between any two distinct vertices of the set $\{u_e : e \in X \cup Y \cup Z\}$.

The graph G is chordal, see [10, p. 332]. In addition, we add to G the following elements:

(*d*) a new vertex v^t and the edge $\{v^t, v_x^t\}$, for each triple $t = (x, y, z) \in \mathcal{C}$;
(*e*) a new vertex u''_e and the edge $\{u''_e, u'_e\}$, for each element $e \in Y \cup Z$.

We denote the obtained graph by H. Clearly, the graph H is chordal. We show that there is a 3d-matching $\mathcal{M} \subseteq \mathcal{C}$ of size n if and only if the graph H admits a P_4-factor.

Suppose that $\mathcal{M} \subseteq \mathcal{C}$ is a 3d-matching of size n. We construct a P_4-factor $\mathcal{S}$ of H. Initially, we put $\mathcal{S} = \emptyset$. For each triple $t \in \mathcal{C} \setminus \mathcal{M}$, we form the 4-path P_t, induced by $\{v^t, v_x^t, v_y^t, v_z^t\}$, and add it to $\mathcal{S}$. For each triple $t = (x, y, z) \in \mathcal{M}$, we form the 4-path S_t, induced by $\{v^t, v_x^t, u_x, u'_x\}$, the 4-path Q_t, induced by $\{v_y^t, u_y, u'_y, u''_y\}$, the 4-path R_t, induced by $\{v_z^t, u_z, u'_z, u''_z\}$, and add them to $\mathcal{S}$. It is easy to see that $\mathcal{S}$ is a P_4-factor of H.

Conversely, now suppose that $\mathcal{S}$ is a P_4-factor of H.

Claim 7. *For each triple $t=(x,y,z)\in\mathcal{C}$, the set $\mathcal{S}$ contains the 4-path Z_1^t on v^t, v_x^t, u_x, u_x' or the 4-path Z_2^t on v^t, v_x^t, v_y^t, v_z^t.*

Proof. Let T be the 4-path from $\mathcal{S}$, containing v^t. As v^t is adjacent to the single vertex v_x^t, the second vertex of T is v_x^t. The third vertex of T can be only u_x or v_y^t. Let us consider the following two cases (note that the fourth vertex of T cannot be u_x):

Case 1. Let the third vertex of T be u_x. The fourth vertex of T is u_x', otherwise, u_x' is not contained in any 4-path from $\mathcal{S}$. In this case, $T=Z_1^t$.

Case 2. Let the third vertex of T be v_y^t. The fourth vertex of T cannot be the vertex u_y, otherwise, neither u_y' nor u_y'' is contained in any 4-path from $\mathcal{S}$, and the vertex u_z, otherwise, neither u_z' nor u_z'' is contained in any 4-path from $\mathcal{S}$. Hence, the fourth vertex of T is v_z^t. The fourth vertex of T cannot be the vertex u_x, otherwise, u_x' is not contained in any 4-path from $\mathcal{S}$. Hence, the fourth vertex of T is v_z^t. In this case, $T=Z_2^t$. ♦

Claim 8. *If a 4-path from $\mathcal{S}$ contains a vertex from $\{u_y'' : y\in Y\}\cup\{u_z'' : z\in Z\}$, then this 4-path contains a vertex from $\{v_y^t : t\in\mathcal{C}, y\in Y\}\cup\{v_z^t : t\in\mathcal{C}, z\in Z\}$ as well.*

Proof. Let $T\in\mathcal{S}$ be the 4-path, containing a vertex u_y'', for some $y\in Y$. For the case of a vertex u_z'', $z\in Z$, a proof is similar. As u_y'' is adjacent to the single vertex u_y', the second vertex of T is u_y'. The third vertex of T is u_y. The fourth vertex of T cannot be a vertex u_x, $x\in X$, otherwise, u_x' is not contained in any 4-path from $\mathcal{S}$, a vertex $u_{y'}$, $y'\in Y$, otherwise, neither $u_{y'}'$ nor $u_{y'}''$ is contained in any 4-path from $\mathcal{S}$, a vertex u_z, $z\in Z$, otherwise, neither u_z' nor u_z'' is contained in any 4-path from $\mathcal{S}$. Hence, the fourth vertex of T is v_y^t, for some $t\in\mathcal{C}$, such that y is included in t. ♦

Let $x\in X$ be an arbitrary element. Let us consider the 4-path $P_x\in\mathcal{S}$, containing u_x'. As u_x' is adjacent to the single vertex u_x, the second vertex of P_x is u_x. The third vertex of P_x cannot be v_y^t, for each triple $t\in\mathcal{C}$, such that x is included in t, otherwise, v_x^t, for some $t\in\mathcal{C}$, is not contained in any 4-path from $\mathcal{S}$. Similarly, any of the vertices u_y, $y\in Y$ and u_z, $z\in Z$ cannot be the third vertex of P_x. The third vertex of P_x cannot be $u_{x'}$, for any $x'\in X$. Indeed, let us suppose that some vertex $u_{x'}$, $x'\in X$ is the third vertex of P_x. In this case, the fourth vertex of P_x is $u_{x'}'$. By Claim 7, we have $Z_1^t\in\mathcal{S}$ or $Z_2^t\in\mathcal{S}$, for each triple $t\in\mathcal{C}$. As $P_x\in\mathcal{S}$, we have

$$|\{t\in\mathcal{C} : Z_1^t\in\mathcal{S}\}|\le|X|-2=n-2. \tag{1}$$

Let us consider the set $F=\{v_y^t : t\in\mathcal{C}, y\in Y\}\cup\{v_z^t : t\in\mathcal{C}, z\in Z\}$. Let $F'\subseteq F$ be the set of all the vertices that are contained in 4-paths from $\mathcal{S}$, which contain at least one of vertices from $\{u_y'' : y\in Y\}\cup\{u_z'' : z\in Z\}$. Let us observe that each 4-path from $\mathcal{S}$ contains at most one element of $\{u_y'' : y\in Y\}\cup\{u_z'' : z\in Z\}$. Claim 8 implies

$$|F'|\ge|Y|+|Z|=2n. \tag{2}$$

On the other hand, if a vertex v_y^t or v_z^t ($t\in\mathcal{C}, y\in Y, z\in Z$) is included in F', then $t\in\mathcal{C}$, such that $Z_1^t\in\mathcal{S}$. Taking into account (1), we obtain

$$|F'|\le 2\cdot|\{t\in\mathcal{C} : Z_1^t\in\mathcal{S}\}|\le 2(n-2),$$

which is a contradiction with (2). Consequently, the third vertex of P_x is v_x^t, for some $t\in\mathcal{C}$, such that x is included in t. Let us consider this triple $t=(x,y,z)\in\mathcal{C}$. As v^t is adjacent to the single vertex v_x^t, the fourth vertex of P_x is v^t. Thus, P_x is the subgraph of H, induced by $\{u_x', u_x, v_x^t, v^t\}$. It is not hard to see that v_y^t can only be an end-vertex of a 4-path $Q_x\in\mathcal{S}$, otherwise, u_y'' or u_z'' is not contained in any 4-path from $\mathcal{S}$. The second vertex of Q_x cannot be the vertex u_z, otherwise, v_z^t is not contained in any 4-path from $\mathcal{S}$ and the vertex v_z^t, otherwise, u_z'' is not contained in any 4-path from $\mathcal{S}$. Hence, the second vertex of Q_x is u_y. The third vertex of Q_x is u_y', otherwise, u_y' is not contained in any 4-path from $\mathcal{S}$. The fourth vertex of Q_x is u_y'', otherwise, u_y'' is not contained in any 4-path from $\mathcal{S}$. The vertex v_z^t is an end-vertex of some 4-path $R_x\in\mathcal{S}$. The second vertex of R_x is u_z. The third vertex of R_x is u_z', otherwise, u_z' is not contained in any 4-path from $\mathcal{S}$. The fourth vertex of R_x is u_z'', otherwise, u_z'' is not contained in any 4-path from $\mathcal{S}$.

Now, let us construct a $3d$-matching $\mathcal{M}\subseteq\mathcal{C}$. Initially, we put $\mathcal{M}=\emptyset$. For each $x\in X$, we add a triple $t\in\mathcal{C}$ to $\mathcal{M}$, such that the corresponding path P_x contains v_x^t. It is not hard to see that $\mathcal{M}$ is a 3d-matching of size n. This finishes the proof of this statement. □

Let us observe that all the paths in the latter proof are induced subgraphs of H. It works for the Induced P_4-Partition problem without any changes. This proof can be slightly modified for the Induced P_k-Partition problem and $k\ge5$ as follows. We add to H the following elements:

(f) the path on new vertices $w_1^t, w_2^t, \dots, w_{k-4}^t$ and the edge $\{v^t, w_1^t\}$, for each triple $t\in\mathcal{C}$;
(g) the path on new vertices $w_y^1, w_y^2, \dots, w_y^{k-4}$ and the edge $\{u_y'', w_y^1\}$, for each element $y\in Y$;
(h) the path on new vertices $w_z^1, w_z^2, \dots, w_z^{k-4}$ and the edge $\{u_z'', w_z^1\}$, for each element $z\in Z$.

We denote the resulting graph by Γ. It is easy to see that Γ is chordal. There is a 3d-matching $\mathcal{M} \subseteq \mathcal{C}$ of size n if and only if Γ admits a P_k-factor. The proof of this fact is similar to the corresponding reasonings from the proof of Theorem 11.

Theorem 9 from [10, p. 332], claiming that the P_3-Partition problem for chordal graphs is NP-complete, uses a polynomial-time reduction from the 3-Dimensional Matching problem with a graph G, constructed by $\mathcal{C}$ as follows:

(a) for each element $e \in X \cup Y \cup Z$, add a pair of adjacent vertices u_e and u'_e;
(b) for each triple $t = (x, y, z) \in \mathcal{C}$, add three vertices v^t_x, v^t_y, v^t_z and connect by edges v^t_x with v^t_y and u_x, v^t_z with v^t_y and u_z, v^t_y with u_x and u_y, u_z;
(c) add an edge between any two distinct vertices of the set $\{u_e : e \in X \cup Y \cup Z\}$.

It was shown in the proof of Theorem 9 from [10, p. 332] that if $\mathcal{M} \subseteq \mathcal{C}$ is a 3d-matching, then

$$\bigcup_{(x,y,z)\notin\mathcal{M}} \{(v^t_x, v^t_y, v^t_z)\} \cup \bigcup_{(x,y,z)\in\mathcal{M}} \{(u'_x, u_x, v^t_x), (u'_y, u_y, v^t_y), (u'_z, u_z, v^t_z)\}$$

is a P_3-factor of G. This is an induced P_3-factor. Otherwise, it was also shown there that any P_3-factor of G has the mentioned form (hence, it is induced), for some $\mathcal{M}' \subseteq \mathcal{C}$, meaning that $\mathcal{M}'$ is a 3d-matching. Hence, the proof of Theorem 9 from [10, p. 332] also works for the Induced P_3-Partition problem, restricted to chordal graphs.

Therefore, we obtain the following result:

Theorem 12. *For each fixed $k \geq 3$, the* Induced P_k-Partition *problem is NP-complete for chordal graphs.*

5. The computational complexity of the P_k-Partition / Induced P_k-Partition problems, restricted to line graphs of planar bipartite graphs

Theorem 13. *For each fixed $k \geq 3$, the* Induced P_k-Partition *problem is NP-complete for line graphs of planar bipartite graphs.*

Proof. To prove this statement, we need the Edge P_{k+1}-Partition problem: Given a graph, the problem is to decide whether its edge set can be partitioned into subsets, each generating P_{k+1}. Namely, we construct a polynomial-time reduction from this problem, restricted to planar bipartite graphs, which is known to be NP-complete [17]. Let G be an arbitrary planar bipartite graph. Clearly, any subgraph of G, isomorphic to P_{k+1}, corresponds to an induced subgraph of $L(G)$, isomorphic to P_k, and vice versa. It establishes the one-to-one correspondence between partitions of $E(G)$ into subsets, each generating P_{k+1}, and partitions of $L(G)$ into its induced subgraphs, each isomorphic to P_k. This fact gives the desired reduction, finishing the proof. □

Theorem 14. *For each fixed $k \geq 5$, the P_k-*Partition *problem is NP-complete for line graphs of planar bipartite graphs.*

Proof. Let $G = (V, E)$ be an arbitrary graph. A u, v-walk of G is a *restricted s-walk*, where $s \geq 3$, if the following conditions are simultaneously true:

(a) the walk contains exactly s distinct edges of G;
(b) each edge of G is included into the walk at most twice;
(c) if e_1 and e_2 are arbitrary edges of the walk with $e_1 = e_2$, then e_1 and e_2 are consecutive in the walk;
(d) the vertices u and v are not necessarily distinct.

It is not hard to see that any restricted k-walk of G corresponds to a subgraph of $L(G)$, isomorphic to P_k, and vice versa. Consequently, the P_k-Partition problem for line graphs is polynomial-time equivalent to the Restricted k-Walk Partition problem: Given a graph, the problem is to decide whether its edge set can be partitioned into restricted k-walks. It remains only to show NP-completeness of the latter problem, restricted to planar bipartite graphs for $k \geq 5$.

The needed polynomial-time reduction (between a version of the 3-Dimensional Matching problem and the Edge P_{k+1}-Partition problem) is presented in [17, p. 150]. Let us describe it in detail. The 3-Dimensional Matching problem, defined on disjoint n-element sets X, Y, Z and a subset $\mathcal{C} \subseteq X \times Y \times Z$, remains NP-complete for instances, for which the bipartite graph $G_{\mathcal{C}}$ with parts $X \cup Y \cup Z$ and $\mathcal{C}$ and edge sets $\bigcup_{t=(x,y,z)\in\mathcal{C}} \{\{x, t\}, \{y, t\}, \{z, t\}\}$ is planar [18]. Fix an instance $\mathcal{C}$ and choose any naturals r, b, l with $r, b, l < \frac{k}{2}$ and $r + b + l = k$. For any $t = (x, y, z) \in \mathcal{C}$, change the 3-star on $\{t, x, y, z\}$ in $G_{\mathcal{C}}$ as it is shown in the following Fig. 9:

Notice that, for any $v \in \{x, y, z\}$, a set of $d_v - 1$ independent P_{l_v+1} has been attached to v, where d_v is the degree of v in $G_{\mathcal{C}}$ and $l_v = r\ (v = x) \vee l_v = b\ (v = y) \vee l_v = l\ (v = z)$. Arguing similarly (e.g., considering restricted k-walks, starting from the gadget pendant vertices) to the proof from [17, p. 150] (which shows that G has an edge P_{k+1}-partition if and only if $\mathcal{C}$ has a 3d-matching), one can verify that any restricted k-walk partition of G must be some its edge P_{k+1}-partition. Thus, the presented reduction works for the Restricted k-Walk Partition problem, showing its NP-completeness for planar bipartite graphs. This finishes the proof of this theorem. □

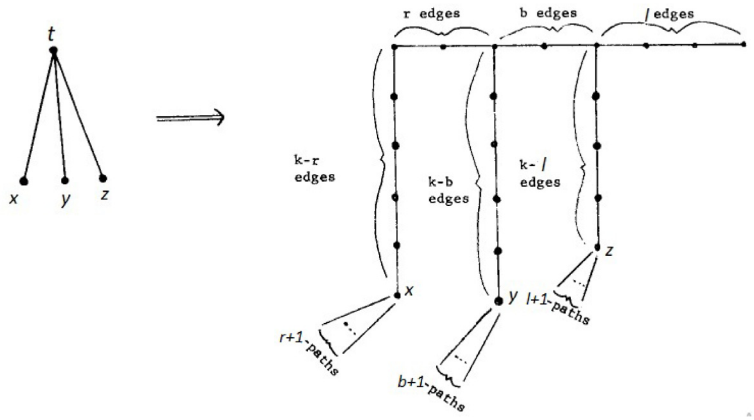

Fig. 9. The local transformation of the 3-star.

The computational status of the P_k-Partition problem for line graphs of planar bipartite graphs is open for $k \in \{3, 4\}$. Let us consider the case of $k = 3$. In this case, restricted 3-walks of a graph induce subgraphs, isomorphic to $K_{1,3}$ or P_4 or K_3. As the problem of partitioning the edge set of a given graph into subgraphs, each isomorphic to $K_{1,3}$ or P_4 or K_3, is NP-complete [28], the P_3-Partition problem is NP-complete for line graphs.

6. Conclusion and further work

In this paper, we considered and clarified the computational complexity of partitioning the vertex set of a given graph into (induced) subgraphs, each isomorphic to a path of order k in the following graph classes: planar bipartite $(H_1, H_2, \ldots, H_\ell)$-free graphs with large girth and maximum degree 3, split graphs, chordal graphs, line graph of planar bipartite graphs, see Table 1. Some of these results resolve known open problems. There are several open questions, which were not answered in this paper, for future research. What is the computational complexity of the Induced P_4-Partition problem for split graphs? What is the computational complexity of the P_k-Partition problem for line graphs of planar bipartite graphs and $k \in \{3, 4\}$? It would be also interesting to clarify the complexity status of both problems for other graph classes, for example, for grid graphs, unit interval graphs, interval graphs, cographs, see Table 1 as well.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable comments.

Results of Sections 1,2,4,5 Supported by Russian Foundation for Basic Research (grant number 20-51-04001) and Belarusian Republican Foundation for Fundamental Research, Republic of Belarus (grant number F21RM-001). Results of Section 3 were supported by the Ministry of Science and Higher Education of the Russian Federation (Goszadaniye) No. 075-03-2024-117, project No. FSMG-2024-0025.

Data availability

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

References

[1] J. Akiyama, V. Chvatal, Packing paths perfectly, Discrete Math. 85 (3) (1990) 247–255.
[2] J. Akiyama, M. Kano, Factors and factorizations of graphs: a survey, J. Graph Theory 9 (1) (1985) 1–42.
[3] J. Akiyama, M. Kano, Graphs and applications, Wiley, 1985, pp. 1–21.
[4] J. Akiyama, M. Kano, Factors and factorizations of graphs: proof techniques in factor theory, Lecture Notes in Mathematics 2031, Springer-Verlag Berlin Heidelberg, 2011.
[5] V. Alekseev, V. Lozin, Local transformations of graphs preserving independence number, Discrete Appl. Math. 135 (1–3) (2004) 17–30.
[6] N. Alon, R. Yuster, H-Factors in dense graphs, J. Combin. Theory Ser. B 66 (1996) 269–282.
[7] B. Bai, W. Chen, K. Letaief, Z. Cao, Diversity-multiplexing tradeoff in OFDMA systems: An H-matching approach, IEEE Trans. Wireless Commun. 10 (11) (2011) 3675–3687.
[8] F. Berman, D. Johnson, T. Leighton, P. Shor, L. Snyder, Generalized planar matching, J. Algorithms 11 (2) (1990) 153–184.
[9] S. Bessy, M. Bougeret, J. Thiebaut, Triangle packing in (sparse) tournaments: Approximation and kernelization, in: 25th Annual European Symposium on Algorithms, ESA 2017, September 4–6, 2017, Vienna, Austria, vol. 87, 2017, pp. 14:1–14:13.
[10] R. van Bevern, R. Bredereck, L. Bulteau, C. Jiehua, F. Vincent, R. Niedermeier, G. Woeginger, Partitioning perfect graphs into stars, J. Graph Theory 85 (2) (2017) 297–335.

[11] R. Boliac, K. Cameron, V. Lozin, On computing the dissociation number and the induced matching number of bipartite graphs, 72, 2004, pp. 241–253,
[12] F. Bonomo-Braberman, C. Gonzalez, A new approach on locally checkable problems, Discrete Appl. Math. 314 (2022) 53–80.
[13] A. Brandstadt, V. Le, J. Spinrad, Graph Classes: a Survey, Society for Industrial and Applied Mathematics, 1999.
[14] M.-S. Chang, L.-H. Chen, L.-J. Hung, A 5k kernel for P_2-packing in net-free graphs, in: 2014 International Computer Science and Engineering Conference, ICSEC, 2014, pp. 12–17.
[15] R. Diestel, Graph Theory, Springer-Verlag Berlin Heidelberg, 2017.
[16] O. Duginov, Partitioning a split graph into induced subgraphs isomorphic to the path of order 3, Proc. Natl. Acad. Sci. Belarus. Phys. Math. Ser. 55 (2019) 32–49, (in Russian).
[17] M. Dyer, A. Frieze, On the complexity of partitioning graphs into connected subgraphs, Discrete Appl. Math. 10 (2) (1985) 139–153.
[18] A. Dyer, A. Frieze, Planar 3DM is NP-complete, J. Algorithms 7 (1998) 174–184.
[19] H. Fernau, D. Raible, A parameterized perspective on packing paths of length two, J. Comb. Optim. 18 (2019) 319–341.
[20] M. Garey, D. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W. H. Freeman & Co., 1979.
[21] J. Han, The complexity of perfect packings in dense graphs, in: T. Gopal, G. Jäger, S. Steila (Eds.), Theory and Applications of Models of Computation, Springer International Publishing, 2017, pp. 289–303.
[22] P. Hell, D. Kirkpatrick, On the complexity of general graph factor problems, SIAM J. Comput. 12 (3) (1983) 601–609.
[23] A. Hope, Component placement through graph partitioning in computer-aided printed-wiring-board design, Electron. Lett. 8 (1972) 87–88.
[24] V. Kann, Maximum bounded H-matching is MAX snp-complete, Inform. Process. Lett. 49 (6) (1994) 309–318.
[25] D. Knop, Partitioning graphs into induced subgraphs, Discrete Appl. Math. 272 (2020) 31–42.
[26] D. Korobitsyn, On the complexity of domination number determination in monogenic classes of graphs, 2, (2) 1994, pp. 191–199,
[27] B. Korte, J. Vygen, Combinatorial Optimization: Theory and Algorithms, Sixth ed., Springer-Verlag Berlin Heidelberg, 2018.
[28] B. Laurent, F. Guillaume, L. Anthony, R. Romeo, R. Irena, Decomposing cubic graphs into connected subgraphs of size three, in: T.N. Dinh, M.T. Thai (Eds.), Computing and Combinatorics, Springer International Publishing, 2016, pp. 393–404.
[29] W. Li, J. Ye, Y. Cao, A 5k-vertex kernel for P_2-packing, Theoret. Comput. Sci. 910 (2022) 1–13.
[30] M. Loebl, S. Poljak, Subgraph packing: A survey, in: R. Bodendiek, R. Henn (Eds.), Topics in Combinatorics and Graph Theory, 1990, pp. 491–503.
[31] L. Lovász, M. Plummer, Matching Theory, North-Holland, 1986.
[32] V. Lozin, M. Milanic, On the maximum independent set problem in subclasses of planar graphs, J. Graph Algorithms Appl. 14 (2) (2010) 269–286.
[33] S. Masuyama, T. Ibaraki, Chain packing in graphs, Algorithmica 6 (1991) 826–839.
[34] D. Mokeev, D. Malyshev, A polynomial-time algorithm of finding a minimum k-path vertex cover and a maximum k-path packing in some graphs, Optim. Lett. 14 (2020) 1317–1322.
[35] J. Monnot, S. Toulouse, The path partition problem and related problems in bipartite graphs, Oper. Res. Lett. 35 (5) (2007) 677–684.
[36] J. Monnot, S. Toulouse, in: V.T. Paschos (Ed.), Combinatorial Optimization and Theoretical Computer Science: Interfaces and Perspectives, Willey, 2008.
[37] C. Moore, J. Robson, Hard tiling problems with simple tiles, Discrete Comput. Geom. 26 (2001) 573–590.
[38] M. Shalu, S. Vijayakumar, T. Sandhya, J. Mondal, Induced star partition of graphs, Discrete Appl. Math. 319 (15) (2021) 81–91.
[39] D. West, Introduction to graph theory, Prentice Hall, 1996.
[40] W. Xi, W. Lin, On maximum P_3-packing in claw-free subcubic graphs, J. Comb. Optim. 41 (2021) 694–709.
[41] R. Yuster, Combinatorial and computational aspects of graph packing and graph decomposition, Comput. Sci. Rev. 1 (2007) 12–26.