Study of some graph theoretic problems via vertex orderings

Thesis submitted in fulfillment of the requirements for the degree

of

Doctor of Philosophy in Computer Science

by

Dalu Jacob

Under the supervision of

Dr. Mathew C Francis

COMPUTER SCIENCE UNIT

INDIAN STATISTICAL INSTITUTE CHENNAI CENTRE

JUNE 2022

Dedicated to

My mother (Philomina Babu) and my sister Unni (Sharlet Jacob)

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Dr. Mathew Francis, for his unfailing support, patience and invaluable guidance throughout my PhD journey. The realization of this thesis would have been impossible without him. His profound knowledge in the subject and critical thinking have deeply inspired me. My heartfelt thanks to him for the trust he placed in me, when I was struggling to be confident about myself.

I would like to offer my special thanks to Prof. Pavol Hell. It was indeed a great experience working with him. I am grateful for his quick responses to the mails and constant encouragement. I would like to extend my sincere gratitude to my other co-authors Dr. Satyabarata Jana, Dr. Suresh Dara, and Dr. N. Narayanan for collaborating with me.

I would like to express my sincere thanks to my M. Phil thesis advisor Prof. A. Vijayakumar (CUSAT), who motivated me to pursue a PhD and was instrumental in the development of my interest in graph theory. I am thankful to all the faculties at ISI Chennai: in particular, Dr. Sudheesh Kattumannil, Dr. Sujata Ghosh, Dr. Karthick, Dr. Venkateswarlu, Dr. Sampangi Raman, Dr. Amit Biswas for their love, support and for creating such a friendly atmosphere at the institute. My special thanks to our librarian Dr. Kalpana for her love and affection. Also, I extend my gratitude to all other office staffs at ISI Chennai for being so helpful.

One of the stimulating part of my PhD life at ISI Chennai was the friendship with the students and scholars there. I will always remember the colorful days I spent with Sreelakshmi chechi – our long walks, laughter, and late night conversations at IIT Madras. Also the interactions with Anjana chechi, Ranjith, Navneet, Nijjwal, Aishwarya, Suparna, Litty, Divya, Pramit have substantially contributed to my happiness. I miss the after-lunch Table Tennis sessions which sometimes used to extend till evening. Particularly eventful were the short intermittent stays of Daphna chechi that involved several interesting discussions. I am also thankful to Muthukumar and Arnab Char for the fruitful academic discussions. During my PhD course work at ISI Kolkata, I was bestowed with many wonderful friendships. Avishek, Koushik, Soumya, Pritam, Nayana, Ramgovind, Snehalika, Sameer, Mostaf are a few of those people who I am thankful for their invaluable company and friendship. Along the same line, a very special thanks to Satya for always helping me out.

I can never thank my Amma and my sister Unni enough for the incredible support, love and care they have given in ups and downs of my life. Also, the loving memory of my late father will last forever in my life. Whatever I have achieved in my life so far, are the result of their sacrifices and selfless love for me. A special thanks to my nephew Evan for the immense joy that he has brought in my life. I would also like to thank Aji chettan, Baba, Maa and Didi for their love, support and encouragement. I am indebted to all my beloved teachers for their invaluable contribution in my life.

Words fall short to thank Arnab. He is indeed one of the greatest blessings in my PhD life. I am grateful for his love, care, deep understanding and most of all for being always there for me.

ABSTRACT

Graph traversal algorithms like breadth-first search and depth-first search typically produce an ordering of the vertices of the input graph. Properties of these vertex orderings often provide new insights about the structure of the graphs under consideration. The existence of vertex orderings that satisfy some special properties characterizes some well-known graph classes like chordal graphs, comparability graphs and cocomparability graphs. Moreover, the availability of such a characterization for a class of graphs often helps us obtain efficient recognition algorithms for the class and also efficient algorithms for a multitude of optimization problems on graphs belonging to the class. As a contribution to this line of research, we show how the vertex ordering approach is useful in the study of some structural and algorithmic questions about graphs and digraphs.

Threshold graphs are a class of graphs that have many equivalent definitions and have applications in integer programming and set packing problems. A graph is said to have a threshold cover of size k if its edges can be covered using k threshold graphs. Let $\mathrm{th}(G)$ denote the least integer k such that G has a threshold cover of size k. In 1977, Chvátal and Hammer observed that $\mathrm{th}(G) \geq \chi(G^*)$, where G^* is a suitably constructed auxiliary graph. They also asked the question of whether there is any graph G such that $\mathrm{th}(G) > \chi(G^*)$. Cozzens and Leibowitz showed that for every $k \geq 4$, there exists a graph G such that $\chi(G^*) = k$ but $\mathrm{th}(G) > k$. Later, Raschle and Simon settled this question for the case $k = 2$, by proving that for any graph G such that $\chi(G^*) = 2$, we have $\mathrm{th}(G) = \chi(G^*)$. In the first part of this thesis, we show how the lexicographic method of Hell and Huang can be used to obtain a completely new and, we believe, simpler proof for this result. For the case when G is a split graph, our method yields a proof that is much shorter than the ones known in the literature.

The problem of computing a minimum cardinality dominating set or absorbing set or kernel (an independent and absorbing set of a digraph), and the problems of computing a maximum cardinality independent set or kernel are all known to be NP-hard for general digraphs. In the second part of the thesis, we explore the algorithmic complexity of these problems in the well known class of *interval digraphs*. A digraph G is an interval digraph if a pair of intervals (S_u, T_u) can be assigned to each vertex u of G such that $(u, v) \in E(G)$ if and only if $S_u \cap T_v \neq \emptyset$. Many different subclasses of interval digraphs have been defined and studied in the literature by restricting the kinds of pairs of intervals that can be assigned to the vertices. We observe that several of these classes, like interval catch digraphs, interval nest digraphs, adjusted interval

digraphs and chronological interval digraphs, are subclasses of the more general class of *reflexive interval digraphs* – which arise when we require that the two intervals assigned to a vertex have to intersect. Here we identify the class of reflexive interval digraphs as an important class of digraphs. We show that while the problems mentioned above are NP-complete, and even hard to approximate, on interval digraphs (even on some very restricted subclasses of interval digraphs called *point-point digraphs*, where the two intervals assigned to each vertex are required to be degenerate), they are all efficiently solvable, in most of the cases linear-time solvable, in the class of reflexive interval digraphs. We also provide a vertex ordering characterization for the class of reflexive interval digraphs and two structural characterizations for the class of point-point digraphs. The results we obtain improve and generalize several existing algorithms and structural results for subclasses of reflexive interval digraphs. Along the way, we also obtain some new results for undirected graphs as well.

A vertex in a directed graph is said to have a *large second neighborhood* if it has at least as many second out-neighbors as out-neighbors. The Second Neighborhood Conjecture, first stated by Seymour, asserts that there is a vertex having a large second neighborhood in every oriented graph (a directed graph without loops or digons). In the third part of the thesis, we extend some results on this conjecture. It is straightforward to see that the conjecture is true for any oriented graph whose underlying undirected graph is bipartite. We extend this by showing that the conjecture holds for oriented graphs whose vertex set can be partitioned into an independent set and a 2-degenerate graph. Fisher proved the conjecture for tournaments and later Havet and Thomassé provided a different proof for the same using median orders of tournaments. Havet and Thomassé in fact showed the stronger statement that if a tournament contains no sink, then it contains at least two vertices with large second neighborhoods. Using their techniques, Fidler and Yuster showed that the conjecture remains true for tournaments from which either a matching or a star has been removed. Using the same tool of median orders, we extend this result to show that the conjecture holds even for tournaments from which both a matching and a star have been removed. This implies that a tournament from which a matching has been removed contains either a sink or two vertices with large second neighborhoods.

Contents

List of Figures

LIST OF PUBLICATIONS

Published/Accepted papers

- *Uniquely Restricted Matchings in interval graphs*, Mathew Francis, Dalu Jacob, and Satyabrata Jana, *SIAM Journal on Discrete Mathematics.* 32(1): 148 – 172 (2018).

- *The Lexicographic Method for the Threshold Cover Problem*, Mathew Francis and Dalu Jacob, Proceedings of the *Conference on Algorithms and Discrete Applied Mathematics (CALDAM)*, 2020, 41 – 52.

- *On the Kernel and Related Problems in Interval Digraphs*, Mathew Francis, Pavol Hell, and Dalu Jacob, Proceedings of the *International Symposium on Algorithms and Computation (ISAAC)*, 2021, 17:1 – 17.

- *Extending some results on the Second Neighborhood Conjecture*, Suresh Dara, Mathew Francis, Dalu Jacob, and N. Narayanan, *Discrete Applied Mathematics.* 311:1 – 17 (2022).

Submitted papers

- The full journal version of the paper, *The Lexicographic Method for the Threshold Cover Problem*, Mathew Francis and Dalu Jacob, submitted to *Discrete Mathematics* (Manuscript No: DM-31394) on 14th March 2022, is currently under review.

- The full journal version of the paper, *On the Kernel and Related Problems in Interval Digraphs*, Mathew Francis, Pavol Hell, and Dalu Jacob, submitted to *Algorithmica* (special issue ISAAC 2021, Manuscript No: ALGO-D-21-00200) on 7th November 2021, is currently undergoing a minor revision.

Chapter 1

Introduction

The origin of graph theory dates back to 18^{th} century, when Euler, one of the greatest mathematicians of all time, proposed an elegant solution to the famous *Königsberg bridge* problem. It is quite interesting to see, how an answer to this simple puzzle has flourished into *'graph theory'*, an important field of study in both mathematics and computer science. Its magical power to embrace the connections between a family of objects paved the way for its use in several disciplines and real-world problems. In this work, we study a few problems in graph theory, that can be approached via vertex orderings. Through the solution to these problems, we will try to illustrate a couple of instances that shows how certain orderings of vertices help us to reveal some interesting facts about some special classes of graphs.

1.1 Basic definitions and notations

First we define some basic concepts concerning undirected graphs. Let $G = (V, E)$ (or simply G) be an undirected graph, where $V(G)$ denotes the vertex set of G and $E(G)$ denotes the edge set of G. An edge between two vertices u and v in G is denoted as uv. An edge that is not present in G is called a *missing edge (or non-edge)*. Any two vertices u and v are said to be *adjacent* (or *neighbors*) in G if $uv \in E(G)$; otherwise they are *non-adjacent* (or *non-neighbors*). The *neighborhood* of a vertex u in G, denoted as $N_G(u)$ is defined as $N_G(u) = \{v \in V(G) : uv \in E(G)\}$. Sometimes we omit the subscript in this notation, if the graph under consideration is clear from the context. The *degree* of a vertex u in G is defined as $|N_G(u)|$. A vertex in G with zero degree is referred to as an *isolated vertex* in G. The *complement* $\overline{G}$ of G is the graph with vertex set $V(\overline{G}) = V(G)$ and edge set $E(\overline{G}) = \{uv : u, v \in V(G), u \neq v$ and $uv \notin E(G)\}$. A graph H is said to be an *induced subgraph* of G if $V(H) \subseteq V(G)$ and

$E(H) = \{uv \in E(G) : u, v \in V(H)\}$; we may also say that H is a subgraph *induced by* $V(H)$. For any set $S \subseteq V(G)$, we denote by $G[S]$ the graph induced by the set S in G. A *path* P_n on n vertices is defined as the graph with $V(P_n) = \{v_1, v_2, \ldots, v_n\}$ and $E(P_n) = \{v_j v_{j+1} : 1 \leq j \leq n-1\}$. A *cycle* C_n on n vertices is defined as the graph with $V(C_n) = \{v_1, v_2, \ldots, v_n\}$ and $E(C_n) = \{v_j v_{j+1} : 1 \leq j \leq n-1\} \cup \{v_n v_1\}$. Given a graph H, a graph G is said to be *H-free* if G contains no induced subgraph isomorphic to H. For example, a graph is said to be *triangle-free*, if it does not contain C_3 as an induced subgraph. A set $S \subseteq V(G)$ is said to be a *clique* if all the vertices in S are pair-wise adjacent in G. In particular, if $V(G)$ itself is a clique, then G is called a *complete graph*. A graph is said to be *connected* if any pair of vertices in it is connected by a path. A maximal connected subgraph of a graph G is referred to as a *component* of G. A graph G with $V(G) = \{u\} \cup \{v_1, v_2, \ldots, v_n\}$ is said to be a *star* if $E(G) = \{uv_j : 1 \leq j \leq n\}$. A graph G is called a *bipartite graph* if $V(G)$ can be partitioned into two subsets A and B such that no two vertices that belong to the same partition are adjacent; i.e. any edge in G has one end-point in A and other end-point in B. Thus a bipartite graph is denoted as $G = (A, B, E)$. Given a bipartite graph $G = (A, B, E)$, the *bipartite complement* of G is defined as the bipartite graph $\overline{G^b}$ with $V(\overline{G^b}) = V(G)$ and $E(\overline{G^b}) = \{ab : a \in A, b \in B \text{ and } ab \notin E(G)\}$. A bipartite graph $G = (A, B, E)$ is said to be a *complete bipartite graph* if all the vertices in A are adjacent to all the vertices in B; i.e. $E(G) = \{ab : a \in A, b \in B\}$.

A set $S \subseteq V(G)$ is said to be an *independent set* in G if all the vertices in S are pair-wise non-adjacent in G. A *proper coloring* of $V(G)$ is an assignment of colors to the vertices of G, such that the set of vertices in each color class form an independent set in G. Note that any graph has a trivial proper coloring using $|V(G)|$ different colors. Therefore the interesting parameter here is, the minimum number of colors needed for a proper coloring of $V(G)$ known as the *chromatic number* of G and is denoted as $\chi(G)$. It follows from the definition of a bipartite graph that, graphs with chromatic number two are precisely the class of bipartite graphs. A set $S \subseteq V(G)$ is said to be a *dominating set* in G if for any $v \in V(G) \setminus S$, there exists $u \in S$ such that $uv \in E(G)$. Given a graph G, a set $M \subseteq E(G)$ is said to be a *matching* if no two edges in M have a common vertex as an end-point.

Directed graphs

Directed graphs (or digraphs) are graphs in which each edge has a direction on it. A digraph is also denoted as $G = (V, E)$ (or G) where $V(G)$ (or V) is the vertex set of G and $E(G)$ (or E) is the edge set (or arc set) of G. Analogous to the previous definitions for undirected graphs, we have the following: An *edge (or arc)* in a digraph, connecting two vertices u and v (possibly $u = v$), and directed from u to v is denoted as (u, v). A *loop* in a digraph G is an arc of the form (u, u) for some $u \in V(G)$. A digraph G is said to be *reflexive* if all the vertices in G have loops on them. For two distinct vertices u and v in a digraph G, if both the arcs (u, v) and (v, u) are present in $E(G)$ then we refer to them as a pair of *symmetric arcs*. A digraph G is said to be an *oriented graph* if it does not contain any loops or symmetric arcs. Let H be an undirected graph. The digraph obtained from H by replacing each edge of H by a pair of symmetric arcs is called the *symmetric digraph* of H. On the other hand, given a digraph G, the *underlying undirected graph* of G is defined to be the undirected graph H, with $V(H) = V(G)$ and $E(H) = \{uv : u \neq v,$ either $(u, v) \in E(G)$ or $(v, u) \in E(G)\}$. Refer to Figure 1.1 for an example of an undirected graph and its symmetric digraph and an example of a digraph and its underlying undirected graph. An oriented graph is called a *tournament* if its underlying undirected graph is a complete graph. The *complement* of a digraph G is defined to be the digraph $\overline{G}$ with $V(G) = V(\overline{G})$ and $E(\overline{G}) = \{(u, v) : u, v \in V(G)$ and $(u, v) \notin E(G)\}$. Let G be a digraph. A vertex v in G is said to be an *out-neighbor (resp. in-neighbor)* of a vertex u if $(u, v) \in E(G)$ (resp. $(v, u) \in E(G)$). Two vertices u and v are said to be adjacent in G if either (u, v) or (v, u) is in $E(G)$. The *out-neighborhood (resp. in-neighborhood)* of a vertex u in G denoted as $N_G^+(u)$ (resp. $N_G^-(u)$) is defined as $N_G^+(u) = \{v : (u, v) \in E(G)\}$ (resp. $N_G^-(u) = \{v : (v, u) \in E(G)\}$). The *out-degree* of a vertex u in G is defined to be $|N_G^+(u)|$. Let G be a digraph with $V(G) = \{v_1, v_2, \ldots, v_n\}$. Then G is said to be a *directed path* if $E(G) = \{(v_j, v_{j+1}) : 1 \leq j < n\}$ and G is said to be a *directed cycle* if $E(G) = \{(v_j, v_{j+1}) : 1 \leq j < n\} \cup \{(v_n, v_1)\}$. A digraph G is said to be a *acyclic* if it does not contain any directed cycle in it. Such digraphs are commonly referred to as *directed acyclic graphs (DAG)*.

Now here we consider the directed analogues of some concepts that we defined earlier for undirected graphs. Let $G = (V, E)$ be a directed graph. A set $S \subseteq V(G)$ is said to be an *independent set* of G, if for any two vertices $u, v \in S$, $(u, v), (v, u) \notin E(G)$. Therefore the independent sets of a digraph G are exactly the independent sets of its underlying undirected graph. On the other hand, we define a set $S \subseteq V(G)$ to be a *weak independent set* of G, if for any two vertices $u, v \in S$, either $(u, v) \notin E(G)$ or $(v, u) \notin E(G)$. It follows that, the weak independent sets of a digraph G are precisely the independent sets of the undirected graph S_G with $V(S_G) = V(G)$ and $E(S_G) = \{uv : u, v \in V(G)$ and $(u, v), (v, u) \in E(G)\}$. A set $S \subseteq V(G)$ is said to be an *absorbing (resp. dominating) set* of G, if for any $v \in V(G) \setminus S$, there exists $u \in S$ such that $(v, u) \in E(G)$ (resp. $(u, v) \in E(G)$). Note that any absorbing set or dominating set

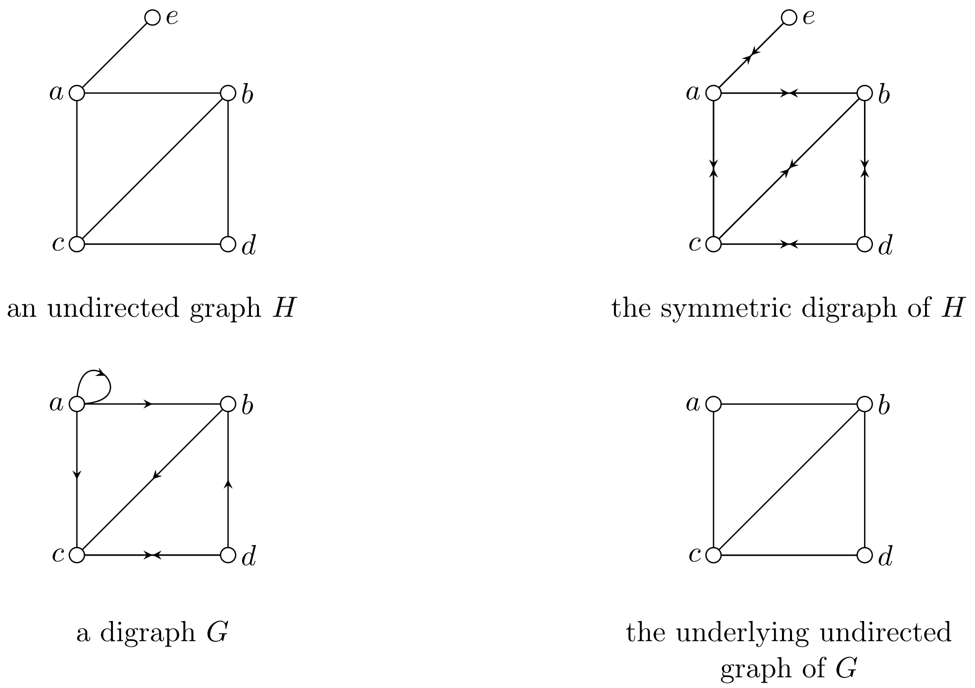


Figure 1.1: The symmetric digraph and the underlying undirected graph

of a digraph G is a dominating set of its underlying undirected graph, but the converse is not necessarily true. Given a digraph G, a set $S \subseteq V(G)$ is said to be a *directed feedback vertex set* of G if the digraph induced by the vertices in $V(G) \setminus S$ is a DAG (where loops are allowed to be present). In other words, the removal of a directed feedback vertex set of a digraph G destroys all the directed cycles (except for the loops) in G.

1.2 Some important classes of graphs

Interval graphs

Interval graphs are exactly the intersection graphs of a family of intervals on a real line. Formally defining,

Definition 1 (Interval graphs)**.** *An undirected graph G is said to be an interval graph if there exists a collection $\{I_u\}_{u \in V(G)}$, of closed intervals on the real line such that for any $u, v \in V(G)$ we have $uv \in E(G)$ if and only if $I_u \cap I_v \neq \emptyset$. The collection $\{I_u\}_{u \in V(G)}$ is called the interval representation of G.*

See Figure 1.2 for an example of an interval graph and a corresponding interval representation of it (note that the intervals are drawn on different horizontal lines just for ease of understanding).

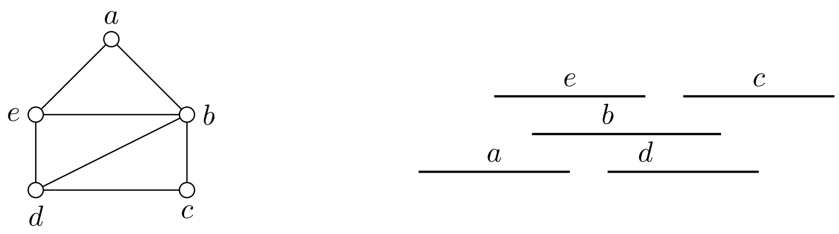


Figure 1.2: An interval graph and a corresponding interval representation

Häjos [62] initiated the study of interval graphs, by posing the question of characterizing graphs that are exactly intersection graphs of intervals on a real line. He proposed this problem from a purely mathematical perspective. On the other hand, independently a well known molecular biologist, Seymour Benzer [6] also asked a related question from an entirely different point of view. During his investigations on the structure of genes, he wondered whether the internal subelements of a gene are linked together in a linear order. A solution to this problem involved finding whether the internal structure of a gene can be modeled as an interval graph. Later interval graphs has became an extensively studied topic in the literature [11, 108, 55, 42, 116] and have proven to be a very useful mathematical structure in modeling many real-world problems. The following theorem gives a vertex ordering characterization for interval graphs.

Theorem 1 ([101])**.** *A graph $G = (V, E)$ is an interval graph if and only if $V(G)$ has a linear ordering $<$ such that for any $u, v, w \in V(G)$ where $u < v < w$, if $uw \in E(G)$ then $uv \in E(G)$ (i.e. the configuration in Figure 1.3 is forbidden).*

Proof. Suppose that G is an interval graph with an interval representation $\{I_u\}_{u \in V(G)}$. It is then easy to verify that the ordering of the vertices with respect to the increasing order of the left endpoints of the intervals representing the vertices in G, satisfies the condition for $<$. On the other hand, suppose that $V(G)$ has a linear ordering $<$ such that for any $u, v, w \in V(G)$ where $u < v < w$, we have $uw \in E(G)$ implies that $uv \in E(G)$. Let $V(G) = \{1, 2, \ldots, n\}$. Let us assume without loss of generality that $<: (1, 2, \ldots, n)$. Note that for each vertex $j \in \{1, 2 \ldots, n\}$, the vertices in $N(j) \cap \{j+1, \ldots, n\}$ appear consecutively in $<$. Now for each vertex $j \in \{1, 2 \ldots, n\}$, define $I(j) = [j, k]$ where $k = \max\{\{j\} \cup N(j) \cap \{j+1, \ldots, n\}\}$. Then $\{I_j\}_{j \in V(G)}$ can be easily verified to be an interval representation of G. $\square$

Note that throughout the thesis, for the figures containing dashed lines in it, we follow the following terminology: a dashed line connecting u and v (resp. from u to v) indicates the absence of the edge uv (resp. (u, v)) in the graph (resp. digraph). The edges of the graphs (resp.

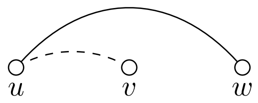


Figure 1.3: The forbidden configuration for a vertex ordering of interval graphs

digraphs) that are not drawn in such figures, may or may not be present in the corresponding graph (resp. digraph).

Chordal graphs

Definition 2 (Chordal graphs). *A graph is said to be chordal (triangulated) if G does not contain any C_k where $k \geq 4$, as an induced subgraph.*

Gavril [51] proved that chordal graphs are exactly the intersection graphs of subtrees in a tree. Chordal graphs can also be characterized using a special ordering of vertices called *perfect elimination ordering* as defined below.

Definition 3 (Perfect elimination ordering). *Let G be an undirected graph. A vertex v in G is said to be a simplicial vertex in G if $N(v)$ is a clique in G. A linear ordering $(v_1, v_2, \dots, v_n)$ of vertices of G is said to be a perfect elimination ordering of G if for each $i \in \{1, 2, \dots, n\}$, the vertex v_i is a simplicial vertex in $G[\{v_i, v_{i+1}, \dots, v_n\}]$.*

First Dirac [36] and later Lekkerkerker and Boland [80] proved that any chordal graph has a simplicial vertex. Using this, Fulkerson and Gross [45] defined the perfect elimination ordering, which yields an iterative procedure to recognize chordal graphs (note that chordal graphs are closed under taking induced subgraphs). In fact they proved that chordal graphs are precisely the graphs whose vertex set admits a perfect elimination ordering [45].

An undirected graph G is said to be *weakly chordal (or weakly triangulated)* if G does not contain C_k or $\overline{C_k}$ for $k \geq 5$, as an induced subgraph.

Cocomparability graphs:

The class of comparability and cocomparability graphs can be defined as follows.

Definition 4 (Comparability, Cocomparability graphs). *An undirected graph G is said to be a comparability graph if the edges in G admit a transitive orientation. i.e. edges of G can be oriented to obtain an oriented graph H, such that for any three distinct vertices $u, v, w \in V(H) = V(G)$,*

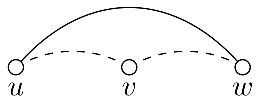


Figure 1.4: The forbidden configuration for an umbrella-free ordering

$(u, v), (v, w) \in E(H)$ implies that $(u, w) \in E(H)$. The complement of a comparability graph is called a cocomparability graph.

An *umbrella-free ordering* of a graph as defined below was first proposed by Damaschke [29] and is a classical way to characterize cocomparability graphs.

Definition 5 (Umbrella-free ordering)**.** *Let G be an undirected graph. An ordering $<$ of $V(G)$ is said to be an umbrella-free ordering of G if for any three distinct vertices $u, v, w \in V(G)$ such that $u < v < w$, $uw \in E(G)$ implies that either $uv \in E(G)$ or $vw \in E(G)$ (i.e. the structure in Figure 1.4 is forbidden).*

We then have the following due to Damaschke [29].

Theorem 2 ([29])**.** *An undirected graph $G = (V, E)$ is a cocomparability graph if and only if G has an umbrella-free ordering.*

Proof. Suppose that G is a cocomparability graph. Then as $\overline{G}$ is a comparability graph, we have that the edges in $\overline{G}$ admit a transitive orientation. Let $\overline{G}_T$ be an oriented graph obtained by transitively orienting the edges of $\overline{G}$. Since any transitively oriented graph is a DAG, consider $<$ to be an ordering of the $V(\overline{G}_T) = V(\overline{G}) = V(G)$ with respect to a topological ordering of $\overline{G}_T$. Then we can verify that $<$ satisfies the conditions for an umbrella-free ordering of $V(G)$. Otherwise, suppose that there exist $u < v < w$ that forms an umbrella. Then $uw \in E(G)$, $uv, vw \notin E(G) \implies uw \notin E(\overline{G}), uv, vw \in E(G)$. Since $<$ is a topological ordering, $(u, v), (v, w) \in E(\overline{G}_T)$. As $(u, w) \notin E(\overline{G}_T)$, we have a contradiction to the fact that $\overline{G}_T$ is transitively oriented.

On the other hand, assume that $<$ is an umbrella-free ordering of G. Then $<$ also has the property that for $u, v, w \in V(G) = V(\overline{G})$, if $(u, v), (v, w) \in E(\overline{G})$ then $(u, w) \in E(\overline{G})$. Now for each edge $xy \in E(\overline{G})$, orient xy from x to y if and only if $x < y$. In this way, we obtain an orientation of edges of $\overline{G}$, which can be easily verified to be a transitive orientation. This implies that $\overline{G}$ is a comparability graph and therefore, G is a cocomparability graph. □

Gilmore and Hoffman [54] characterized interval graphs to be exactly the graphs which are both chordal and cocomparability. Interval graphs also have several other equivalent characterizations [80, 45, 56].

Tolerance graphs: The class of tolerance graphs was introduced by Golumbic and Monma [58] as a generalization of interval graphs. Like interval graphs, tolerance graphs also arise from the intersection of intervals on a real line but in a special way. The edges between the vertices are determined by a measure of intersection region of their corresponding intervals. Informally speaking, if the intervals corresponding to a pair of vertices can 'tolerate' the intersection between them, then they are not connected by an edge. For an interval $I = [x, y]$ on the real line (here $x, y \in \mathbb{R}$ and $x \leq y$), the length of the interval denoted by $|I|$ is defined to be the value $y - x$. The class of tolerance graphs can be formally defined as follows.

Definition 6 (Tolerance graphs)**.** *An undirected graph G is said to be a tolerance graph if each vertex u in G can be assigned an interval I_u and a tolerance $t_u \in \mathbb{R}^+$ in such a way that for any two vertices u and v in G, $uv \in E(G)$ if and only if $|I_u \cap I_v| \geq \min\{t_u, t_v\}$. In addition, if $t_u \leq |I_u|$ for each vertex $u \in V(G)$, then G is called a bounded tolerance graph.*

See the book by Golumbic and Trenk [59] on tolerance graphs for a detailed study of tolerance graphs and their variants.

Split graphs

The class of split graphs is defined as follows.

Definition 7 (Split graphs)**.** *An undirected graph G is said to be a split graph if $V(G)$ can be partitioned into two sets A and B such that A is a clique and B is an independent set.*

Threshold graphs

The class of threshold graphs was introduced by Chvátal and Hammer [22]. Threshold graphs have several equivalent characterizations and one way to define this class of graphs is as follows:

Definition 8 (Threshold graphs)**.** *Two edges ab, cd in an undirected graph G are said to form an* alternating 4-cycle *if $ad, bc \notin E(G)$. Threshold graphs are precisely the class of graphs that does not contain any pair of edges that form an alternating 4-cycle.*

We will study more about threshold graphs in Chapter 2.

Chain graphs: Chain graphs can be considered as the bipartite analogues of threshold graphs.

Definition 9 (Chain graphs)**.** *A bipartite graph $G = (A, B, E)$ is called a* chain graph *if for any pair of vertices u and v that belong to A, either $N(u) \subseteq N(v)$ or $N(v) \subseteq N(u)$.*

Observe that if the vertices in A satisfy this property, then the vertices in B also automatically satisfy this property. Therefore, chain graphs are precisely the class of bipartite graphs for which the neighborhoods of the vertices in either partite set of G have a linear order with respect to inclusion.

Interval bigraphs

A bipartite analogue of interval graphs, called *interval bigraph* was introduced by Harary, Kabell, and McMorris [64] and can be defined as follows:

Definition 10 (Interval bigraph). *A bipartite graph $G = (A, B, E)$ is said to be an interval bigraph if there exists a collection $\{I_u\}_{u \in V(G)}$, of closed intervals on a real line such that $ab \in E(G)$ if and only if $a \in A$, $b \in B$ and $I_a \cap I_b \neq \emptyset$.*

See Figure 1.5 for an example of an interval bigraph and a corresponding interval representation of it.

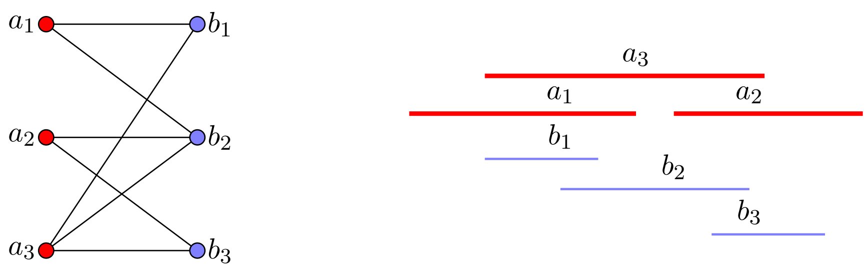


Figure 1.5: An interval bigraph and a corresponding interval representation

Interval digraphs

Interval digraphs can be considered as a directed analogue of interval graphs. The following definition of interval digraphs is proposed by Das, Roy, Sen and West [32].

Definition 11 (Interval digraph). *A digraph G is said to be an interval digraph if there exists a collection $\{(S_u, T_u)\}_{u \in V(G)}$, of pairs of closed intervals on a real line such that for any $u, v \in V(G)$, we have $(u, v) \in E(G)$ if and only if $S_u \cap T_v \neq \emptyset$. The collection $\{(S_u, T_u)\}_{u \in V(G)}$ is called the interval representation of G.*

See Figure 1.6 for an example of an interval digraph and a corresponding interval representation of it (source and destination intervals of each vertex are shown in thin green and bold red respectively).

We will study more about interval digraphs and interval bigraphs in the upcoming chapters.

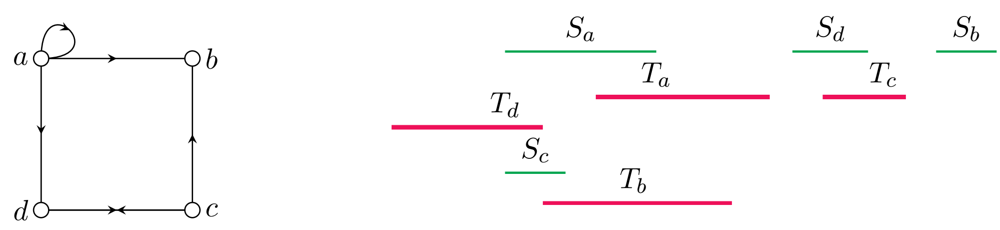


Figure 1.6: An interval digraph and a corresponding interval representation

1.3 Graph covering and graph recognition problems

Definition 12 (Graph cover)**.** *A graph G is said to be covered by graphs, $H_1, H_2, \ldots, H_k$ if for each $i = \{1, 2, \ldots, k\}$, $V(H_i) = V(G)$ and $E(G) = \bigcup_{1 \leq i \leq k} E(H_i)$.*

The problem of covering an arbitrary graph using graphs that belong to special graph classes, is a well known problem in graph theory [77]. As a part of this work, we study the following special graph covering problems.

Definition 13 (Threshold cover problem)**.** *The k-threshold cover problem asks whether an input graph G can be covered by k threshold graphs.*

Given a graph G, for each $e \in E(G)$, let G_e denotes the graph with $V(G_e) = V(G)$ and $E(G_e) = \{e\}$. Then note that for each $e \in E(G)$, G_e is trivially a threshold graph and $\bigcup_{e \in E(G)} E(G_e) = E(G)$. This implies that any graph G can be covered by $|E(G)|$ threshold graphs. Therefore the interesting parameter here is, $\mathrm{th}(G) = \min\{k : G$ can be covered by k threshold graphs$\}$.

Definition 14 (Chain cover problem)**.** *The k-chain subgraph cover problem (in short k-CSC) asks whether an input bipartite graph G can be covered by k chain graphs.*

As in the case of threshold cover, since the graph induced by each edge is trivially a chain graph, we have that any bipartite graph can be covered by using $|E(G)|$ chain graphs. Therefore the interesting parameter here is, ch(G) $= \min\{k : G$ has a chain subgraph cover of size $k\}$.

The above mentioned covering problems are well studied in the literature and we will learn more about them and their algorithmic complexities in Chapter 2 and Chapter 3.

The *graph recognition problem* is an important algorithmic problem in graph theory [49].

Definition 15 (Graph recognition problem)**.** *Given a class of graphs $\mathcal{C}$, the graph recognition problem asks whether an input graph belong to the class $\mathcal{C}$.*

Note that for any integer $k \geq 1$, the k-threshold cover problem (resp. k-CSC problem) is exactly the problem of recognizing graphs (resp. bipartite graphs) that can be covered by k threshold graphs (resp. k chain graphs).

Given two problems P_1 and P_2, we say that the problem P_1 is *reducible* to the problem P_2 in polynomial time if there exists a polynomially computable function f such that for any instance, say x of problem P_1, we have that $f(x)$ is an instance of problem P_2 and x is a YES instance of problem P_1 if and only if $f(x)$ is a YES instance of problem P_2.

For example, we will see in Chapter 2 that for any integer $k \geq 1$, the k-CSC problem can be reduced to the k-threshold cover problem in polynomial time.

1.4 Some well-known vertex orderings and their applications

Let us begin with the well-known graph traversal algorithms, *Breadth First Search (BFS)* and *Depth First Search (DFS)* which are fundamentally important ingredients in various graph algorithms. Roughly speaking, as their name indicates, BFS visits the vertices in the non-decreasing order of their distances from the starting vertex, whereas DFS starts at a vertex and explores as far as possible from the current vertex before it backtracks. Both of these were introduced in the context of maze traversal algorithms but later became an inevitable part of various graph algorithms. For example, BFS has applications in shortest path problems, recognition problems of several graph classes, network flows etc. and DFS plays a very important role in the algorithms for connectivity, planarity, finding cycles in graph etc. [115]. Both BFS and DFS can be implemented in time linear to the size of the input graph.

Lex-BFS ordering: Rose, Lueker and Tarjan [109] introduced a variant of BFS called *Lexicographic Breadth First search (Lex-BFS)* to construct a linear-time algorithm for recognizing chordal graphs. Later, Lex-BFS based algorithms were discovered for the recognition of many different graph classes (see [24] for a survey). A Lex-BFS ordering is also a BFS ordering – i.e., a breadth-first search algorithm can also visit the vertices in that order – but it has some additional properties.

Let $a = (a_1, a_2, \ldots, a_k)$ and $b = (b_1, b_2, \ldots, b_k)$ are two vectors of same length defined on the set of positive integers. We say that a is *lexicographically smaller* than b, if there exists an integer $i \leq k$ such that $a_i < b_i$ and $a_j = b_j$ for each $j \in \{1, 2, \ldots, i-1\}$.

The procedure [14] described below is an easy way to produce a Lex-BFS ordering starting

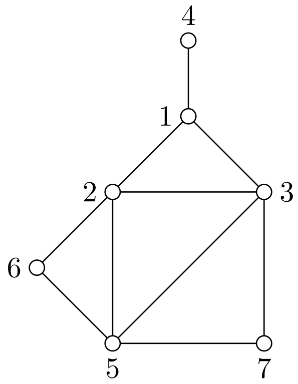


Figure 1.7: An example of a graph labeled with respect to a Lex-BFS ordering of its vertices

at a vertex x of an input graph G.

Procedure: *Generate a Lex-BFS ordering σ starting at a vertex x of an input graph G.*

$\sigma(1) = x$, mark x as **visited**;

for each k from 2 to n **do**

for each **unvisited vertex** u, let $b(u) = (b_1, b_2, \ldots, b_{k-1})$, where $b_i = 1$, if $\sigma(i) \in N(u)$ or $b_i = 0$, otherwise;

Let v be a vertex with lexicographically largest vector b (i.e. $b(v)$ is a lexicographically largest vector in $\{b(u) : u \text{ is unvisited}\}$);

assign $\sigma(k) = v$;

mark v as **visited**;

See Figure 1.7, for an example of a graph that has been labeled with respect to a Lex-BFS ordering of its vertices. Note that even though the above procedure does not run in linear-time, there exist linear-time algorithms [61] to produce a Lex-BFS ordering of an input graph.

In one part of this work we exploit an important property of Lex-BFS ordering, which we will come to see in an upcoming chapter. As we have mentioned before, Lex-BFS was introduced in the context of recognizing chordal graphs. We have already noted that chordal graphs are precisely the undirected graphs that admits a *perfect elimination ordering* (see Definition 3). In fact, Rose, Tarjan and Lueker [109] proved that an undirected graph is chordal if and only if any

Lex-BFS ordering is the reverse of a perfect elimination ordering. Consequently, the recognition problem of chordal graphs can be solved in linear time. Perfect elimination orderings of chordal graphs are extremely useful in efficiently solving some of the classic graph problems like maximum independent set, maximum clique, minimum coloring, minimum clique cover [50] etc on chordal graphs.

Revisiting the vertex ordering characterizations for some of the graph classes that we have defined before, we can see similar algorithmic applications. For example, the vertex ordering characterization for interval graphs (see Theorem 1) has applications in solving several algorithmic problems for interval graphs including the minimum coloring problem [96], the domination problem and its variants like independent domination, connected domination and total domination problems [101], optimal path cover problem [4], the longest path problem [71] and many more. We have also seen a characterization of cocomparability graphs in terms of an *umbrella-free ordering* (see Theorem 13) of the vertex set. This ordering is also highly useful in solving many algorithmic problems for the class of cocomparability graphs including the domination problem and its variants like independent domination, connected domination and total domination problems [79], Hamiltonian path problem [30], longest path problem [92] etc. Note that over the years, several researchers have come up with various special vertex orderings with interesting properties. Many other classes of graphs that admits vertex/edge ordering characterizations have also been studied in the literature. See Chapter 5 of the book [15] for a detailed survey of vertex and edge orderings.

Lexicographic method: The technique called *lexicographic method* is introduced by Hell and Huang [68]. This method has a significant importance in our work. Hell and Huang [68], demonstrated how this method can lead to shorter proofs and simpler recognition algorithms for certain problems that involve constructing a specific 2-coloring of an auxiliary bipartite graph whose vertices correspond to the edges of the graph. This method involves fixing an ordering $<$ of the vertices of the graph, and then processing the edges in the "lexicographic order" implied by the ordering $<$. They showed how this technique can lead to simpler characterization proofs and recognition algorithms for comparability graphs, proper interval graphs and proper circular-arc graphs. We adapt this technique in our work, to construct a 2-threshold cover (if it exists) for an input graph.

Now let us review a few special vertex orderings concerning the class of directed graphs.

Topological ordering: It is one of the most well-known vertex ordering algorithms for digraphs. A linear ordering $<$ of the vertices of a digraph G is called a *topological ordering* if for any edge $(u, v) \in E(G)$ we have $u < v$. It is easy to see that directed acyclic graphs are precisely the digraphs that have a topological ordering. This fact can be used to develop a linear time recognition algorithm for the class of DAG [23]. Other than its applications to solve algorithmic problems for DAG, topological ordering also has several applications to real world problems including scheduling, operation systems etc. and it plays an important role in problems related to network analysis [73].

There are other classes of digraphs that can be characterized using vertex orderings. As a part of this work, we also propose a vertex ordering characterization for a subclass of interval digraphs and derive some interesting consequences.

Median order: This is yet another interesting notion of ordering in directed graphs. A *median order* of a digraph is a linear ordering of the vertex set that maximizes the number of forward arcs. The problem of finding a median order of a digraph is NP-hard [5] for general digraphs and the complexity of the median order problem even in the class of tournaments seems to be unknown [17]. The concept of median orders arises in the context of voting theory [5] and is well studied. Havet and Thomasse [65] used median order as a tool to provide a short and constructive proof of the Second Neighborhood Conjecture for the class of tournaments. This conjecture was shown to be true for the class of tournaments initially by Fisher [43] using some probabilistic arguments. A part of our work involves solving the Seymour Second Neighborhood Conjecture for some special classes of digraphs. We also explore some properties of median orders in our approach towards this work, which we will come across later.

1.5 Scope of the thesis and a brief review of related works

In the previous section, we have witnessed the importance of special vertex orderings in solving various kinds of problems in graph theory. As a contribution to this line of research, we follow an ordering approach towards most of the problems that we encounter in this thesis. The work mainly consists of the following three problems that incorporate the flavors of both structural and algorithmic graph theory.

1.5.1 Lexicographic method for the threshold cover problem

An important approach towards the threshold cover problem of an input graph G that we can see in the literature is, by studying a related parameter in the auxiliary graph G^* that has vertex set $V(G^*) = E(G)$ and edge set $E(G^*) = \{\{ab, cd\}: ab, cd \in E(G)$ such that ab, cd form an alternating 4-cycle in $G\}$. Chvátal and Hammer observed that since for any subgraph H of G that is a threshold graph, $E(H)$ is an independent set in G^*, we have that $\mathrm{th}(G) \geq \chi(G^*)$. This gave rise to the question of whether there is any graph G such that $\mathrm{th}(G) > \chi(G^*)$. Cozzens and Leibowitz [27] showed the existence of such graphs. In particular, they showed that for every $k \geq 4$, there exists a graph G such that $\chi(G^*) = k$ but $\mathrm{th}(G) > k$. The question of whether such graphs exist for $k = 2$ remained open for a decade (see [84]). Ibaraki and Peled [70] showed, by means of some very involved proofs, that if G is a split graph or if G^* contains at most two non-trivial components, then $\chi(G^*) = 2$ if and only if $\mathrm{th}(G) = 2$. They further conjectured that for any graph G, $\chi(G^*) = 2 \Leftrightarrow \mathrm{th}(G) = 2$. If the conjecture held, it would show immediately that graphs having a threshold cover of size 2 can be recognized in polynomial time, since the auxiliary graph G^* can be constructed and its bipartiteness checked in polynomial time. Consequently, we then also have polynomial-time algorithms for 2-CSC problem. In contrast to the case $k \leq 2$, Yannakakis [118] showed that 3-CSC is NP-complete. This again implies by a suitable reduction that the problem of deciding whether $\mathrm{th}(G) \leq 3$ is also NP-complete, even if the input graph G is a split graph. In fact, all these covering problems become NP-complete for any arbitrary $k \geq 3$ [118, 22]. Thus Ibaraki and Peled's conjecture was significant to completely settle the complexity of threshold cover problem and related covering problems. Cozzens and Halsey [26] studied some properties of graphs having a threshold cover of size 2 and showed that it can be decided in polynomial time whether the complement of a bipartite graph has a threshold cover of size 2. Finally, in 1995, Raschle and Simon [102] proved the conjecture of Ibaraki and Peled by extending the methods in [70]. To be precise, they proved that for any graph G, $\chi(G^*) = 2$ if and only if $\mathrm{th}(G) = 2$.

This proof of Raschle and Simon is very technical and involves the use of a number of complicated reductions and previously known results. In this part of the work, we propose a completely different and self-contained proof for the theorem of Raschle and Simon that a graph G can be covered by two threshold graphs if and only if G^* is bipartite. We show how the lexicographic method of Hell and Huang can be used to obtain a completely new and, we believe, simpler proof for this result. This shows that the applicability of the lexicographic method may not be limited to only problems involving orientation of edges. However, it should be noted that unlike in the work of Hell and Huang, we start with a *Lex-BFS ordering* of the vertices of the graph instead of an arbitrary ordering. We demonstrate by an example, why in the proof for general graphs it is important to start with a Lex-BFS ordering, instead of an arbitrary ordering. This strengthens the relevance of Lex-BFS orderings, as a useful tool in solving graph theoretic problems. Now for the case when the input graph G is a split graph, we can use the lexicographic method starting with any arbitrary ordering and our method yields a proof that is much shorter than the ones known in the literature. Moreover, our proof can be used to obtain a simple algorithm to find a 2-threshold cover (if exists) for an input graph. Further, using the relation between threshold cover for split graphs and chain cover for bipartite graphs we have a very simple algorithm to find a 2-chain cover (if exists) for an input bipartite graph. Both the algorithms, for finding a 2-threshold cover and 2-chain cover can be implemented in $O(m^2)$ time. Note that faster algorithms for finding a threshold cover of size 2 for an input graph (that runs in $O(n^3)$ time) [112] and finding a chain cover of size 2 for an input bipartite graph (that runs in $O(n^2)$ time) [84] are known. But these algorithms are quite involved and has several reductions. Whereas, the strategy of our corresponding algorithms are basic, straight forward and much simpler to implement.

1.5.2 Kernel and related problems in interval digraphs

In this part of the work, we illustrate the fact that the reflexivity of an interval digraph has a huge impact on the algorithmic complexity of several problems related to domination and independent sets in digraphs. In particular, here we study the following problems in the class of interval digraphs and its subclasses (the formal definitions of these problems can be found in Chapter 4 and Chapter 6).

(a) Independent-Set

(b) Absorbing-Set (resp. Dominating-Set)

(c) Kernel

(d) Min-Kernel and Max-Kernel

(e) Weak Independent-Set

(f) Feedback Vertex-Set

We will also see along the way that, how the solution to some of the problems above, can be used to obtain few interesting results in some special classes of undirected graphs. Note that the following subclasses of interval digraphs are important to us. The class of *reflexive* *interval digraphs* is a subclass of interval digraphs which arise when we require that the two intervals assigned to each vertex (in an interval representation of it) have to intersect and the class of interval nest digraphs forms a subclass of reflexive interval digraphs that has an interval representation in which for each vertex, the destination interval is completely contained inside the source interval. On the other hand, the class of *point-point digraphs* is a subclass of interval digraphs where the two intervals assigned to each vertex are required to be degenerate (i.e. pair of intervals corresponding to each vertex are two points). Let us briefly have a look at our contributions in this part of the work.

Prisner [100] proved that the underlying undirected graphs of interval nest digraphs (a subclass of reflexive interval digraphs) are weakly triangulated graphs and notes that this means that any algorithm that solves the maximum independent set problem on weakly triangulated graphs can be used to solve the INDEPENDENT-SET problem on interval nest digraphs and their reversals. Since the problem of computing a maximum independent set can be solved in $O(nm)$ time in weakly triangulated graphs [67], it follows that there is an $O(nm)$-time algorithm for the INDEPENDENT-SET problem in interval nest digraphs and their reversals, even when only the adjacency list of the input graph is given.

We provide a vertex-ordering characterization for the class of reflexive interval digraphs and two simple characterizations for the class of point-point digraphs including a forbidden structure characterization. Our characterization of point-point digraphs directly yields a linear time recognition algorithm for that class of digraphs. From our vertex-ordering characterization of reflexive interval digraphs, it follows that the underlying undirected graphs of every reflexive interval digraph is a cocomparability graph. Also a natural question that arises here is whether the underlying graphs of reflexive interval digraphs is the same as the class of cocomparability graphs. We show that this is not the case by demonstrating that the underlying graphs of reflexive interval digraphs cannot contain an induced $K_{3,3}$. Also, as the maximum independent set problem is linear time solvable on cocomparability graphs [89] we now have that the INDEPENDENT SET problem is also linear time solvable on reflexive interval digraphs. This improves and generalizes the $O(nm)$-time algorithm for the same problem on interval nest digraphs. In contrast, we prove that the INDEPENDENT SET problem is APX-hard even for the class of point-point digraphs.

Domination in digraphs is a topic that has been explored less when compared to its undirected counterpart. Even though bounds on the minimum dominating sets in digraphs have been obtained by several authors (see the book [66] for a survey), not much is known about the computational complexity of finding a minimum cardinality absorbing set (or dominating set) in directed graphs. Even for tournaments, the best known algorithm for Dominating-Set does not run in polynomial time [91, 103]. In [91], the authors give an $n^{O(\log n)}$ time algorithm for the Dominating-Set problem in tournaments and they also note that Sat can be solved in $2^{O(\sqrt{v})}n^K$ time (where v is the number of variables, n is the length of the formula and K is a constant) if and only if the Dominating-Set in a tournament can be solved in polynomial time. Thus, determining the algorithmic complexity of the Dominating-Set problem even in special classes of digraphs seems to be much more challenging than the algorithmic question of finding a minimum cardinality dominating set in undirected graphs.

We observe that the problem of solving Absorbing-Set (resp. Dominating-Set) on a reflexive interval digraph G can be reduced to the problem of solving Red-Blue Dominating Set on an interval bigraph whose interval representation can be constructed from an interval representation of G in linear time. Further, we show that Red-Blue Dominating Set is linear time solvable on interval bigraphs (given an interval representation). Thus the problem Absorbing-Set (resp. Dominating-Set) is linear-time solvable on reflexive interval digraphs, given an interval representation of the digraph as input. If no interval representation is given, an algorithm in [95] can be used to construct one in polynomial time, and therefore these problems are polynomial-time solvable on reflexive interval digraphs even when no interval representation of the input graph is known. In contrast, we prove that the Absorbing-Set and Dominating-Set problems remain APX-hard even for point-point digraphs.

An independent absorbing set of a directed graph is more well-known as a *kernel* of the digraph. This term is introduced by Von Neumann and Morgenstern [94] in the context of game theory. They showed that for digraphs associated with certain combinatorial games, the existence of a kernel implies the existence of a winning strategy. Most of the work related to domination in digraphs has been mainly focused on kernels. We follow the terminology in [100] and call an independent dominating set in a directed graph a *solution* of the graph. Unlike its undirected counterpart, a kernel (resp. solution) of a digraph does not necessarily exist. For example, a directed triangle with edges $(a, b), (b, c)$ and (c, a) does not have a kernel (resp. solution). Therefore, besides the computational problem of finding a minimum or maximum sized kernel, called Min-Kernel and Max-Kernel respectively, the comparatively easier problem of determining whether a given digraph has a kernel in the first place, called Kernel, is itself a non-trivial one. In fact, the Kernel problem is known to be NP-hard for general digraphs [21]. Later, Fraenkel [44] proved that the Kernel problem remains NP-complete even for planar digraphs of degree at most 3 having in- and out-degrees at most 2. It can be easily seen that the

MIN-KERNEL and MAX-KERNEL problems are NP-complete for those classes of graphs for which the KERNEL problem is NP-complete. A digraph is said to be *kernel perfect* if every induced subgraph of it has a kernel. Several sufficient conditions for digraphs to be kernel perfect has been explored [106, 38, 94]. The KERNEL problem is trivially solvable in polynomial time on any kernel perfect family of digraphs. But the algorithmic complexity status of the problem of computing a kernel in a kernel perfect digraph also seems to be unknown [98]. Prisner [100] proved that interval nest digraphs and their reversals are kernel-perfect, and a kernel can be found in these graphs in time $O(n^2)$ if a representation of the graph is given. We remark that the MIN-KERNEL problem can be shown to be NP-complete even in some kernel perfect families of digraphs that have a polynomial-time computable kernel.

We show that reflexive interval digraphs are kernel perfect and hence the KERNEL problem is trivial on this class of digraphs. We construct a linear-time algorithm that computes a kernel in a reflexive interval digraph, given an interval representation of digraph as an input. This improves and generalizes Prisner's similar results about interval nest digraphs mentioned above. Moreover, we give an $O((n+m)n)$ time algorithm for the MIN-KERNEL and MAX-KERNEL problems for a superclass of reflexive interval digraphs (here m denotes the number of edges in the digraph other than the self-loops at each vertex). As a consequence, we obtain an improvement over the $O(n^3)$ time algorithm for finding a minimum independent dominating set in cocomparability graphs that was given by Kratsch and Stewart [79]. Our algorithm for MIN-KERNEL and MAX-KERNEL problems has a better running time of $O(n^2)$ for adjusted interval digraphs. On the other hand, we show that the problem KERNEL is NP-complete for point-point digraphs and MIN-KERNEL and MAX-KERNEL problems are APX-hard for point-point digraphs.

The directed FEEDBACK VERTEX-SET problem is one of the classic NP-complete problems in the literature and it has various real world applications. As a consequence of the vertex ordering characterization for reflexive interval digraphs, we can show that the problems WEAK INDEPENDENT-SET and FEEDBACK VERTEX-SET are reducible to each other in linear time for reflexive interval digraphs. We study these problems in certain subclasses of interval digraphs. In particular we show that, the problem WEAK INDEPENDENT-SET and therefore, the problem FEEDBACK VERTEX-SET are polynomial-time solvable for the class of interval nest digraphs. Further we use this solution of the WEAK INDEPENDENT-SET problem for interval nest digraphs, to give a polynomial-time algorithm for the problem of finding a maximum cardinality uniquely restricted matching for the class of interval graphs. This result settles one of the open problem posed by Golumbic in [57].

1.5.3 The Seymour Second Neighborhood Conjecture for some special graph classes

Given a digraph G, we denote by $N_G^{++}(v)$ the *second out-neighborhood* of v in G, which is defined to be the set of vertices whose distance from v is exactly 2, i.e. $N_G^{++}(v) = \{u \in V(G) \colon N_G^{-}(u) \cap N_G^{+}(v) \neq \emptyset$ and $u \notin N_G^{+}(v) \cup \{v\}\}$. A vertex v in a digraph G is said to have a *large second neighborhood* if $|N_G^{++}(v)| \geq |N_G^{+}(v)|$. Paul Seymour conjectured the following in 1990 (see [35]):

Conjecture (Paul Seymour): *Every oriented graph contains a vertex with a large second neighborhood.*

The above conjecture, if true, implies a special case of another open question concerning digraphs called the Caccetta-Häggkvist Conjecture [16]. The Seymour Second Neighborhood Conjecture for the special case of tournaments, was known as Dean's Conjecture [35] and was later solved by Fisher [43] in 1996 using some basic linear algebraic and probabilistic arguments. Later in 2000, Havet and Thomasse [65] gave a short combinatorial proof of Dean's Conjecture using median orders of tournaments. They could in fact prove something stronger: in a tournament without a sink, there exist two vertices with large second neighborhoods. Using the approach of Havet and Thomasse, Fidler and Yuster [41] in 2007 proved that the Second Neighborhood Conjecture is true for oriented graphs that can be obtained from tournaments by removing edges in some specific ways. In particular, they showed that a tournament missing a matching (an oriented graph whose missing edges form a matching), a tournament missing a star and a tournament missing a complete graph all satisfy the conjecture. As these results hold even if the missing matching (or star, or complete graph) is empty, they extend the proof of Dean's Conjecture by Havet and Thomasse. Using techniques from this paper, Salman Ghazal [52] proved that the Second Neighborhood Conjecture is true for tournaments missing a "generalized star" (or threshold graphs), thereby extending the result of Fidler and Yuster for tournaments missing a star and tournaments missing a complete graph. It has to be noted that among these results that all use the median order approach, the case of the tournament missing a matching is by far the most difficult one, requiring a complicated proof. In this work, we introduce new ideas to refine and extend this proof, allowing us to prove the conjecture for a superclass of tournaments missing a matching: we show that oriented graphs whose missing edges can be partitioned into a (possibly empty) matching and a (possibly empty) star also satisfy the Second Neighborhood Conjecture. In fact, we prove the stronger statement that in such a graph that does not contain a sink, there exists a vertex that has a large second neighborhood and is not the center of the missing star.

Ghazal [53] attempts to generalize the theorem of Havet and Thomasse by trying to prove that there exist two vertices with large second neighborhoods in every tournament missing a matching that does not contain a sink. He shows that if a tournament missing a matching satisfies certain additional technical conditions, then such a result can be obtained. Our result mentioned above directly yields a proof that shows that every tournament missing a matching that does not contain a sink has at least two vertices with large second neighborhoods.

Other researchers have tried to attack special cases of the Second Neighborhood Conjecture without using the median order approach. Lladó [82] proved the conjecture in regular oriented graphs with high connectivity. Kaneko and Locke [74] verified the conjecture for oriented graphs with minimum out-degree at most 6.

It is easy to verify that in any oriented graph, a minimum out-degree vertex whose out-neighborhood is an independent set is a vertex with a large second neighborhood. Therefore, the conjecture is true for bipartite graphs (in fact, it is true if the underlying undirected graph is triangle-free). It appears difficult to prove the conjecture even for oriented graphs whose underlying undirected graph is 3-colorable. We show that the conjecture is true for every oriented graph whose vertices can be partitioned into two sets such that one is an independent set and the other induces a 2-degenerate graph in the underlying undirected graph.

1.6 Outline of the thesis

The thesis is organized into three parts. The first part is primarily based on the threshold cover problem and this consists of two chapters. In Chapter 2, we study the threshold cover problem using the lexicographic method. This is followed by Chapter 3, that reviews the connections established in the literature between some subclasses of bipartite graphs and digraphs. The second part of the work, which is based on the kernel and related problems in interval digraphs, consists of three chapters. In Chapter 4, along with an ordering characterization for the class of reflexive interval digraphs, we propose efficient algorithms for several problems like Independent Set, Absorbing Set, Kernel, Min-Kernel and Max-Kernel for reflexive interval digraphs. On the other hand, in Chapter 5, we show that all these problems turn out to be NP-complete and/or APX-hard for even a restricted subclass of interval digraphs called point-point digraphs. In the same chapter, we also characterize point-point digraphs in two ways, and one of the characterizations leads to a linear-time recognition algorithm for the class of point-point digraphs. In Chapter 6, we study the problems Weak Independent-Set and Feedback Vertex-Set for some particular subclasses of interval digraphs. The third part is a study of the Seymour Second Neighborhood Conjecture (SSNC) for some special graph classes. In particular, in Chapter 7 we show that SSNC is true for oriented graphs whose missing edges can be partitioned into a (possibly empty) matching and a (possibly empty) star. We also observe some interesting consequences of this result. In Chapter 8, we show that SSNC is true for oriented graphs in which their vertex set can be partitioned into two sets such that one is an independent set and the other induces a 2-degenerate graph in the underlying undirected graph. Finally, we conclude the thesis in Chapter 9 by proposing some interesting open problems related to the work.

Part I

The Threshold Cover Problem

Chapter 2

The Lexicographic Method for the Threshold Cover Problem

Let us begin with a brief study of the well-known class of threshold graphs.

2.1 Threshold graphs

The class of threshold graphs was introduced by Chvátal and Hammer [22]. It is a well studied graph class in the literature, which has numerous applications in set packing problems [63], interger programming [22] and many more. Mahadev and Peled have given a comprehensive survey on threshold graphs and their applications [88]. Before defining the class of threshold graphs, first let us define a related concept.

Threshold dimension: Let $G = (V, E)$ be an undirected graph with $V(G) = \{v_1, v_2, \ldots, v_n\}$. Note that any set $S \subseteq V(G)$ can be represented as a characteristic vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ such that

$$x_i = \begin{cases} 1, & \text{if } v_i \in S \\ 0, & \text{otherwise} \end{cases}$$

This implies that, any set $S \subseteq V(G)$ corresponds to a corner of an n-dimensional hypercube. Then the *threshold dimension* $\theta(G)$ of a graph G can be defined as follows.

Definition 16 (Threshold dimension)**.** *Given a graph G, its threshold dimension $\theta(G)$ is defined*

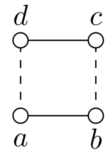


Figure 2.1: Alternating 4-cycle

as the minimum number k of linear inequalities

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq t_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &\leq t_2 \\ &\vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &\leq t_k \end{aligned}$$

in $\mathbb{R}$ such that S is an independent set in G if and only if the characteristic vector corresponding to S satisfies all the linear inequalities above.

Speaking geometrically, the threshold dimension of a graph G can be interpreted as the minimum number of halfspaces whose intersection contains exactly those corners of the n-dimensional hypercube that correspond to independent sets in G. The *threshold graphs* are precisely the class of graphs G for which $\theta(G) \leq 1$. Formally, we define this class of graphs as follows.

Definition 17 (Threshold graphs)**.** *A graph G is a threshold graph if there exist real number t and real vertex weights a_i for each vertex v_i in G, such that the linear inequality $a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq t$, separates independent sets and non-independent sets of G.*

Threshold graphs have several equivalent characterizations. But in this work we use the following forbidden structure characterization of threshold graphs, given by Chvátal and Hammer [22]. Two edges ab, cd in an undirected graph G are said to form an *alternating 4-cycle* if $ad, bc \notin E(G)$ (see Figure 2.1). We then have the following theorem.

Theorem 3 ([22])**.** *An undirected graph G is a threshold graph if and only if G does not contain any pair of edges that form an alternating 4-cycle (or equivalently, G is a $\{2K_2, P_4, C_4\}$-free graph).*

Since the complement of each forbidden structure is also forbidden (as $\overline{2K_2} = C_4$, $\overline{P_4} = P_4$ and $\overline{C_4} = 2K_2$), threshold graphs are closed under taking complements.

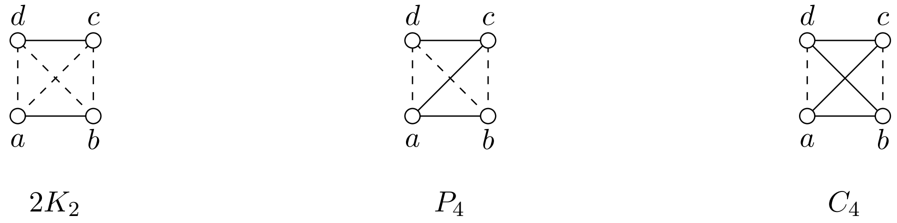


Figure 2.2: Forbidden structures for threshold graphs

Threshold cover: A graph G is said to be covered by graphs, $H_1, H_2, \dots, H_k$ if for each $i = \{1, 2, \dots, k\}$, $V(H_i) = V(G)$ and $E(G) = \bigcup_{1 \le i \le k} E(H_i)$.

Definition 18 (Threshold cover problem)**.** *The k-threshold cover problem asks whether an input graph G can be covered by k threshold graphs.*

Given a graph G, for each $e \in E(G)$, let G_e denotes the graph with $V(G_e) = V(G)$ and $E(G_e) = \{e\}$. Then note that for each $e \in E(G)$, G_e is trivially a threshold graph and $\bigcup_{e \in E(G)} E(G_e) = E(G)$. This implies that any graph G can be covered by $|E(G)|$ threshold graphs. Therefore the interesting parameter here is, $\mathrm{th}(G) = \min\{k : G$ can be covered by k threshold graphs$\}$.

The threshold cover number and the threshold dimension are equal

Suppose that G can be covered by k threshold graphs, i.e. there exist k threshold graphs $H_1, H_2, \dots, H_k$ such that for each $i = \{1, 2, \dots, k\}$, $V(H_i) = V(G)$ and $E(G) = \bigcup_{1 \le i \le k} E(H_i)$. Since H_i is a threshold graph for each i, by Definition 17 we have that for each $i \in \{1, 2, \dots, k\}$, there exist real number t_i and real vertex weights a_{ij} for each vertex v_j in H_i, such that the characteristic vector of a set $S \subseteq V(H_i)$ satisfies the linear inequality, say $L_i : a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \le t_i$ if and only if S is an independent set of H_i. Note that as $E(G) = \bigcup_{1 \le i \le k} E(H_i)$, the independent sets of G remain as independent sets in each of the subgraphs H_i. Therefore, the characteristic vectors of the independent sets of G satisfies each of the k linear inequalities L_i. On the other hand, if S is not an independent set in G, then there exist vertices $u, v \in S$ such that $uv \in E(G)$. Then by the definition of cover, there exists some $i \in \{1, 2, \dots, k\}$ such that $uv \in E(H_i)$. This implies that S is not an independent set in H_i and therefore, the characteristic vector of S does not satisfy the linear inequality L_i. Since the linear inequalities $L_1, L_2, \dots, L_k$ satisfies the required conditions in Definition 16, we can conclude that the threshold dimension, $\theta(G) \le \mathrm{th}(G)$.

On the other hand, assume that the threshold dimension $\theta(G) = k$. Therefore by Defini-

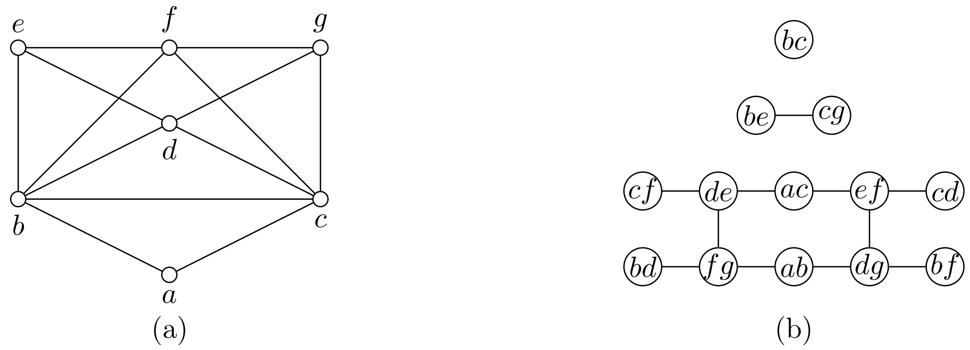


Figure 2.3: (a) a graph G, and (b) the auxiliary graph G^* of G.

tion 16, we have that there exist k linear inequalities, say $L_1, L_2, \ldots, L_k$ (as in the definition) such that a set $S \subseteq V(G)$ is an independent set in G if and only if the characteristic vector of S satisfies all the k linear inequalities. Now for each $i \in \{1, 2, \ldots, k\}$ we define H_i to be the graph with $V(H_i) = V(G)$ and $E(H_i) = \{uv : u, v \in V(G), u \neq v$, and the characteristic vector of the set $\{u, v\}$ does not satisfies the inequality $L_i\}$. It can be then verified that the graph H_i as defined above is a threshold graph for each $i \in \{1, 2, \ldots, k\}$ and $E(G) = \bigcup_{1 \leq i \leq k} E(H_i)$. This implies that G can be covered by k threshold graphs. Therefore we have $\mathrm{th}(G) \leq \theta(G)$.

Thus we have the following theorem.

Theorem 4. *For any graph G, the threshold dimension $\theta(G) = \mathrm{th}(G) = \min\{k : G$ can be covered by k threshold graphs$\}$.*

An important approach in the literature toward the study of threshold cover problem, is by defining an auxiliary graph and studying a related parameter on it.

2.2 The auxiliary graph G^*

Chvátal and Hammer [22] defined the auxiliary graph G^* corresponding to a graph G as follows.

Definition 19 (The auxiliary graph G^*)**.** *Given a graph G, the graph G^* has vertex set $V(G^*) = E(G)$ and edge set $E(G^*) = \{\{ab, cd\} : ab, cd \in E(G)$ such that ab, cd form an alternating 4-cycle in $G\}$.* (See Figure 2.3 for an illustration).

Chvátal and Hammer observed that the following lower bound on $\mathrm{th}(G)$ holds.

Lemma 1. $\mathrm{th}(G) \geq \chi(G^*)$.

Proof. Let G be covered by two threshold graphs H_1 and H_2. By the definition of G^*, if $\{ab, cd\} \in E(G^*)$ then $ad, bc \in E(\overline{G})$. The fact that H_1 and H_2 are threshold subgraphs of G then implies that neither H_1 nor H_2 can contain both the edges ab and cd. We therefore conclude that the sets $E(H_1)$ and $E(H_2)$ are both independent sets in G^*. Since G is covered by H_1 and H_2, we have that $V(G^*) = E(H_1) \cup E(H_2)$. Thus, $\{E(H_1), E(H_2) \setminus E(H_1)\}$ forms a bipartition of G^* into two independent sets. This completes the proof. □

This gave rise to the question of whether there is any graph G such that $\mathrm{th}(G) > \chi(G^*)$. Cozzens and Leibowitz [27] showed that for every $k \geq 4$, there exists a graph G such that $\chi(G^*) = k$ but $\mathrm{th}(G) > k$. The question of whether such graphs exist for $k = 2$ remained open for a decade (see [84]). Ibaraki and Peled [70] showed, by means of some very involved proofs, that if G is a split graph or if G^* contains at most two non-trivial components, then $\chi(G^*) = 2$ if and only if $\mathrm{th}(G) = 2$. They further conjectured that for any graph G, $\chi(G^*) = 2 \Leftrightarrow \mathrm{th}(G) = 2$. Cozzens and Halsey [26] studied some properties of graphs having a threshold cover of size 2 and showed that it can be decided in polynomial time whether the complement of a bipartite graph has a threshold cover of size 2. Finally, in 1995, Raschle and Simon [102] proved the conjecture of Ibaraki and Peled by extending the methods in [70].

Theorem 5. *For any graph G, $\chi(G^*) = 2$ if and only if* $\mathrm{th}(G) = 2$.

2.2.1 An overview of Raschle and Simon's proof

By proving Theorem 5, Raschle and Simon not only settled Ibraki and Peled's conjecture, they also gave the first polynomial-time algorithm for recognizing graphs that can be covered by two threshold graphs. But their proof for Theorem 5 is very technical and involves the use of previously known results and a number of complicated reductions. Their approach towards the proof was using *threshold completions*. Given a graph $G = (V, E)$ and a set $S \subseteq E(G)$, S is said to have a threshold completion in G if there exists a subset of edges E' such that $S \subseteq E' \subseteq E(G)$ and $G' = (V, E')$ is a threshold graph. Now the problem of finding two threshold graphs H_1 and H_2 such that $E(G) = E(H_1) \cup E(H_2)$ can be seen to be equivalent to the problem of finding two partitions S_1 and S_2 of $E(G)$ for which threshold completions exist. A sequence of vertices $v_0, v_1, \ldots, v_{2k-1}$ of a graph G is said to form an AC_{2k} in G if $v_{2i}v_{2i+1} \in E(G)$ and $v_{2i-1}v_{2i} \notin E(G)$ (indices modulo 2k). An AC_{2k} *with respect to a set* $S \subseteq E(G)$ in G, has all its present edges in S. Some fundamental results in the literature regarding threshold completions are the following [63, 102]:

(a) If a set $S \subseteq E(G)$ has a threshold completion in G, then such a threshold completion can be found in linear time.

(b) A set $S \subseteq E(G)$ has a threshold completion if and only if G does not contain any AC_{2l} with respect to S for $l \geq 2$.

Using the above results, it is now enough to prove that, if G^* is bipartite then the vertices in G^* can be partitioned into two sets where both the sets are AC_{2l}-free with respect to G for each $l \geq 2$. i.e. their goal was to find a suitable 2-coloring for the vertices in G^*, in such a way that there does not exist an AC_{2l} for $l \geq 2$ in G with respect to each of the color classes (such a coloring is called an AC_{2l}-free coloring). In order to reduce this problem further, they make use of the following result in [63].

Let G be a graph with $\chi(G^*) = 2$. Then there exists an AC_{2l} for $l \geq 3$ in G with respect to one of the color classes of G^* if and only if there exists an AC_6 with respect to one of the color classes of G^*.

Note that by the definition of G^*, the color classes corresponding to any 2-coloring of G^* are AC_4-free with respect to G. Therefore by combining the above observation with the previous results, if G^* is bipartite then the problem of finding a 2-threshold cover for G boils down to the problem of finding an AC_6-free 2-coloring of G^*. Now given a 2-coloring of G^* with color classes E_1 and E_2, they define an AP_6 in G to be a sequence $v_0, v_1, \ldots, v_5, v_0$ of distinct vertices of G such that $v_0v_1, v_2v_3, v_4v_5 \in E_i$ for some $i \in \{1, 2\}$ and $v_1v_2, v_3v_4, v_5v_0 \in E(\overline{G})$. A 2-coloring of G^* is said to be AP_6-free if there is no AP_6 in G with respect to that coloring. They proved that, given an AP_6-free 2-coloring of G^*, an AC_6-free 2-coloring of G^* can be computed in $O(m^2)$ time. They further define a *double* AP_6 to be an AP_6 with an additional constraint that $v_0v_2, v_1v_5 \in E_i$ and proved that given a double AP_6-free 2-coloring of G^*, an AP_6-free 2-coloring of G^* can be computed in $O(m^2)$ time. The most intricate part is the proof of correctness of an algorithm that computes a double AP_6-free 2-coloring of G^*. In fact, they proved that a double AP_6-free coloring of G^* can be computed in $O(m^2)$ time. Consequently, the algorithm for finding a 2-threshold cover has to go through each of these reductions backwards, in order to find a valid 2-threshold cover for an input graph G. Figure 2.4 briefly summarizes major steps of the algorithm (as well as their proof for Theorem 5) for finding a 2-threshold cover given by Raschle and Simon [102].

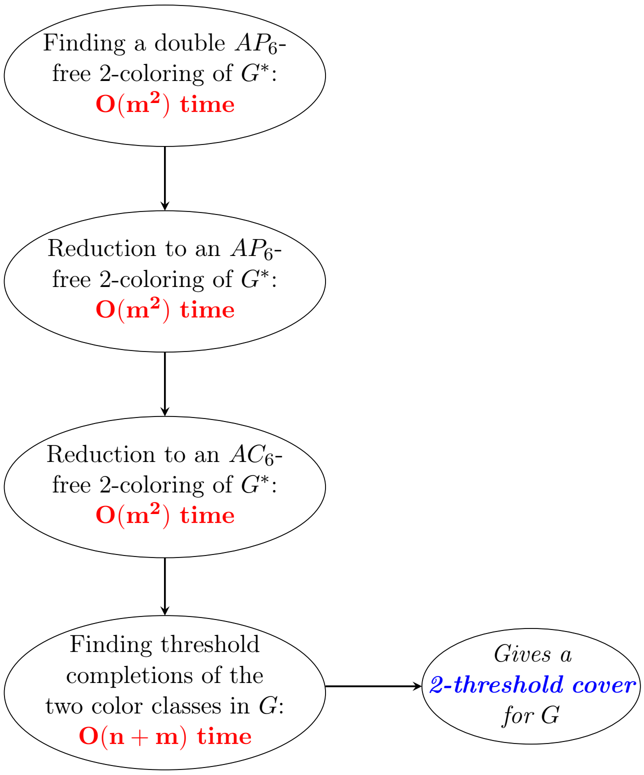
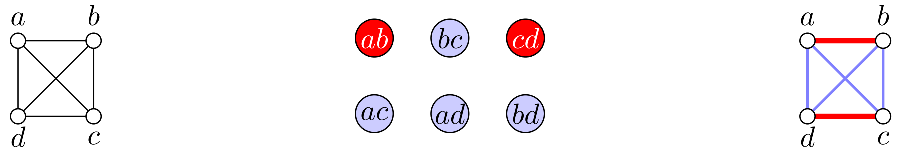


Figure 2.4: Major steps in Raschle and Simon's proof and algorithm

2.3 Our contribution to the threshold cover problem

In this part of the work, we propose a completely different and self-contained proof for the theorem of Raschle and Simon (Theorem 5) using the *lexicographic method* introduced by Hell and Huang [68]. In particular, we adapt this technique to construct a specific 2-coloring of G^* that can be used to generate a 2-threshold cover of G (without using the notion of threshold completions). Our proof is direct, and also gives rise to a simpler recognition algorithm for graphs having a threshold cover of size 2. Note that faster algorithms for determining if a graph has a threshold cover of size 2 are known. After the algorithm of Raschle and Simon [102], Sterbini and Raschle [112] used some observations of Ma [83] to construct an $\mathcal{O}(|V(G)|^3)$ algorithm for the problem. But this algorithm also involves several reductions in order to construct a 2-threshold cover (if exists).

Let G be any graph such that G^* is bipartite. We would like to prove that G has a threshold cover of size 2. Note that if we take an arbitrary 2-coloring of G^* having color classes X_1, X_2, the subgraph G_i of G formed by the edges in X_i (where $i \in \{1, 2\}$) need not necessarily be a threshold graph. Figure 2.5 demonstrates such an example. The graph G in (a) is a complete graph in which no pair of edges form an alternating C_4. Thus G^* has only isolated vertices and



(a) G (b) G^* (c) corresponding cover

Figure 2.5: An arbitrary 2-coloring of G^* that does not directly correspond to a 2-threshold cover for G

(b) shows an arbitrary 2-coloring of G^* using the colors dark red and light blue. If you look at the corresponding edges in the respective color classes, we can observe that none of the color classes directly provides a threshold graph (as shown in (c) the bold red edges that corresponds to the dark red color class of G^* form a $2K_2$ and the thin blue edges that corresponds to the light blue color class of G^* form a C_4). Now as there can be an exponential number of 2-colorings possible for G^* (since G^* can have many connected components), the crux of the problem is to find a "special 2-coloring of G^*" (where some isolated vertices in G^* may have both the colors in it). This was the motivation for using the *lexicographic method* here.

2.4 The lexicographic method

The technique called the *lexicographic method* was introduced by Hell and Huang [68]. They demonstrated how this method can lead to shorter proofs and simpler recognition algorithms for certain problems that involve constructing a specific 2-coloring of an auxiliary bipartite graph whose vertices correspond to the edges of the graph. The method involves fixing an arbitrary ordering $<$ of the vertices of the graph, and then processing the edges in the "lexicographic order" implied by the ordering $<$. They showed how this method can lead to simpler characterization proofs and recognition algorithms for comparability graphs, proper interval graphs and proper circular-arc graphs. The method starts by taking an arbitrary ordering of the vertices of the graph. It then prescribes choosing the lexicographically smallest (with respect to the given vertex ordering) edge to orient and then orienting it in one way or the other, along with all the edges whose orientations are forced by it. Hell and Huang showed that the lexicographic approach makes it easy to ensure that the orientation so produced satisfies the necessary conditions, if such an orientation exists. We adapt this technique to construct a 2-coloring of G^* that can be used to generate a 2-threshold cover of G. This shows that the applicability of the lexicographic method may not be limited to only problems involving orientation of edges. However, it should be noted that unlike in the work of Hell and Huang, we start with a *Lex-BFS ordering* of the vertices of the graph instead of an arbitrary ordering. It is an ordering of the vertices with the property that it is possible for a *Lexicographic Breadth First Search (Lex-BFS)* algorithm to visit the vertices of the graph in that order. A Lex-BFS ordering is also a BFS ordering – i.e., a breadth-first search algorithm can also visit the vertices in that order – but it has some additional properties. Lex-BFS can be implemented to run in time linear in the size of the input graph and was introduced by Rose, Tarjan and Lueker [109] to construct a linear-time algorithm for recognizing chordal graphs. Later, Lex-BFS based algorithms were discovered for the recognition of many different graph classes (see [24] for a survey).

2.5 Preliminaries

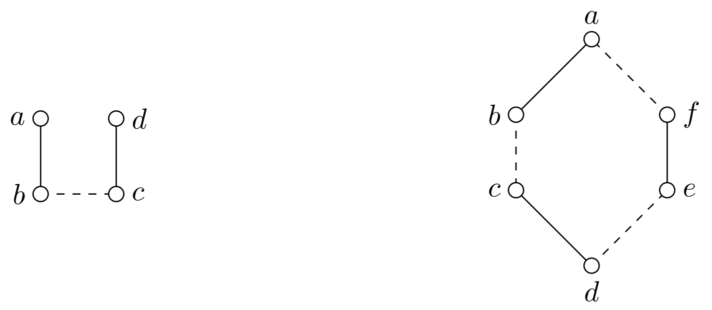
Let $G = (V, E)$ be any graph. Recall that edges $ab, cd \in E(G)$ form an alternating 4-cycle if $bc, da \in E(\overline{G})$. In this case, we also say that a, b, c, d, a is an alternating 4-cycle in G (alternating 4-cycles are called AC_4s in [102]). The edges ab and cd are said to be the *opposite edges* of the alternating 4-cycle a, b, c, d, a. Thus for a graph G, the auxiliary graph G^* is the graph with $V(G^*) = E(G)$ and $E(G^*) = \{\{ab, cd\} \colon ab, cd \in E(G)$ are the opposite edges of an alternating 4-cycle in $G\}$. Note that it follows from the definition of an alternating 4-cycle that if a, b, c, d, a is an alternating 4-cycle, then the vertices a, b, c, d are pairwise distinct. We shall refer to the vertex of G^* corresponding to an edge $ab \in E(G)$ alternatively as $\{a, b\}$ or ab, depending upon the context.

Our goal is to provide a new proof for Theorem 5.

It is easy to see that $\chi(G^*) = 1$ if and only if $\mathrm{th}(G) = 1$. Therefore, by Lemma 1, it is enough to prove that if G^* is bipartite, then G can be covered by two threshold graphs. In order to prove this, we find a specific 2-coloring of the non-trivial components of G^* (components of size at least 2) using the lexicographic method of Hell and Huang [68].

We say that (A_0, A_1, A_2) is a *valid 3-partition* of $E(G)$ if $\{A_0, A_1, A_2\}$ is a partition of $E(G)$ with the property that in any alternating 4-cycle in G, one of the opposite edges belongs to A_1 and the other to A_2. In other words, for any edge $\{ab, cd\} \in E(G^*)$, one of ab, cd is in A_1 and the other in A_2.

Given a valid 3-partition (A_0, A_1, A_2) of $E(G)$ and $A \in \{A_1, A_2\}$, we say that a, b, c, d is an *alternating A-path* if $a \neq d$, $ab, cd \in A \cup A_0$, and $bc \in E(\overline{G})$. Further, we say that a, b, c, d, e, f, a is an *alternating A-circuit* if $a \neq d$, $ab, cd, ef \in A \cup A_0$, and $bc, de, fa \in E(\overline{G})$. See Figure 2.6 for the illustration of an A-alternating path and an A-alternating circuit.



(i) an A-alternating path (ii) an A-alternating circuit

Figure 2.6: For (i) $a \neq d$, $ab, cd \in A \cup A_0$ and for (ii) $a \neq d$ and possibly, $b = e$ or $c = f$, $ab, cd, ef \in A \cup A_0$, $bc, de, fa \in E(\overline{G})$

Observation 1. *Let (A_0, A_1, A_2) be a valid 3-partition of $E(G)$ and let $\{A, \overline{A}\} = \{A_1, A_2\}$.*

(a) If a, b, c, d is an alternating A-path, then $ad \in E(G)$.

(b) If a, b, c, d, e, f, a is an alternating A-circuit, then $ef \in A$ and $ad \in \overline{A}$.

Proof. To prove (a), it just needs to be observed that if $ad \in E(\overline{G})$, then a, b, c, d, a would be an alternating 4-cycle in G whose opposite edges both belong to $A \cup A_0$, which contradicts the fact that (A_0, A_1, A_2) is a valid 3-partition of $E(G)$. To prove (b), suppose that a, b, c, d, e, f, a is an alternating A-circuit. Since a, b, c, d is an alternating A-path, we have by (a) that $ad \in E(G)$. Then since a, d, e, f, a is an alternating 4-cycle in G and $ef \in A \cup A_0$, it follows that $ef \in A$ and $ad \in \overline{A}$. $\square$

We shall use the above observation throughout this chapter without referring to it explicitly.

Let (A_0, A_1, A_2) be a valid 3-partition of $E(G)$ and let $\{A, \overline{A}\} = \{A_1, A_2\}$. We say that (a, b, c, d, e) is an *A-pentagon in G with respect to* (A_0, A_1, A_2) if $a, b, c, d, e \in V(G)$, $ac, ad, be \in E(\overline{G})$, $ab, ae \in A$, $bc, bd, ec, ed \in \overline{A}$ and $cd \in A \cup A_0$. We abbreviate this to just "A-pentagon" when the graph G and the 3-partition (A_0, A_1, A_2) of G are clear from the context. We say that an A-pentagon (a, b, c, d, e) is a *strict A-pentagon* if $cd \in A$. We say that (a, b, c, d, e) is a *pentagon* (resp. *strict pentagon*) if it is an A-pentagon (resp. strict A-pentagon) for some $A \in \{A_1, A_2\}$. (Pentagons are similar to the "AP_5-s" in [102]). Figure 2.7(i) and Figure 2.7(ii) illustrate an A-pentagon and a strict A-pentagon.

We say that (x, y, z, w) is an *A-switching path in G with respect to* (A_0, A_1, A_2) if $x, y, z, w \in V(G)$, $xw \in E(\overline{G})$, $xy, zw \in A \cup A_0$, and $yz \in \overline{A}$. When the graph G and the 3-partition (A_0, A_1, A_2) of G are clear from the context, we abbreviate this to just "A-switching path". We

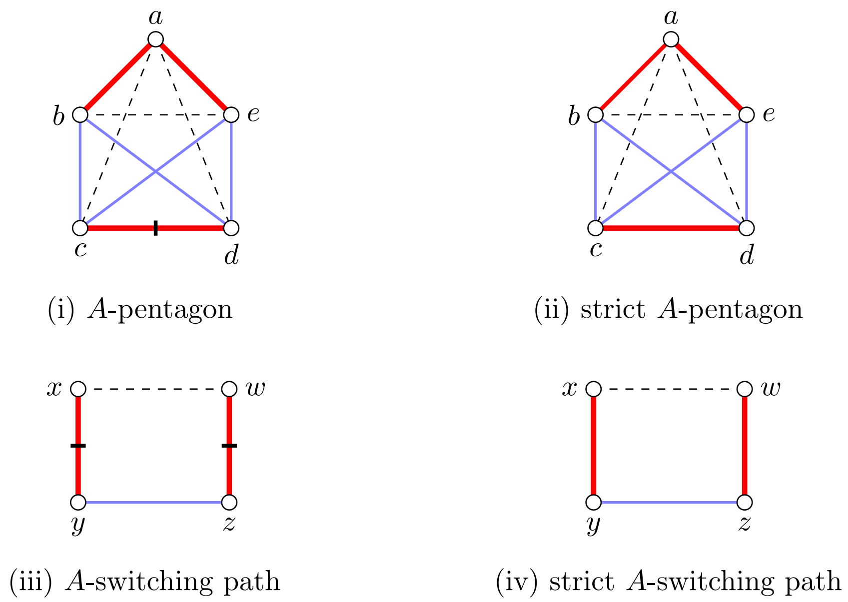


Figure 2.7: Edges belonging to the set A are shown in bold red and that belonging to the set $\overline{A}$ are shown in thin blue. A black dash on an edge in (i) and (iii) indicates the possibility of the edge being in A_0 (as for A-pentagons, $cd \in A \cup A_0$ and for A-switching paths, $xy, zw \in A \cup A_0$).

say that (x, y, z, w) is a *strict A-switching path* if it is an A-switching path and in addition, $xy, zw \in A$. We say that (x, y, z, w) is a *switching path* (resp. *strict switching path*) if it is an A-switching path (resp. strict A-switching path) for some $A \in \{A_1, A_2\}$. Figure 2.7(iii) and Figure 2.7(iv) illustrate an A-switching path and a strict A-switching path.

Note that from the definitions of pentagons and switching paths, it follows that if (a, b, c, d, e) is a pentagon, then the vertices a, b, c, d, e are pairwise distinct, and if (a, b, c, d) is a switching path, then the vertices a, b, c, d are pairwise distinct.

Lemma 2. *Let (A_0, A_1, A_2) be a valid 3-partition of $E(G)$. Let $\{A, \overline{A}\} = \{A_1, A_2\}$. Let (x, y, z, w) be an A-switching path in G and let $y'z' \in E(G)$ be such that $yz', zy' \in E(\overline{G})$. Then,*

(a) if $x = y'$, then $(x = y', y, z, w, z')$ is an A-pentagon and

(b) if $w = z'$, then $(w = z', z, y, x, y')$ is an A-pentagon.

Proof. Since $yz \in \overline{A}$ and $\{yz, y'z'\} \in E(G^*)$ we have that $y'z' \in A$. Suppose that $x = y'$. Then $y, (x = y'), z, w, (x = y'), z', y$ is an alternating A-circuit (note that $y \neq w$ as $x \in N(y) \setminus N(w)$), implying that $yw \in \overline{A}$. This further implies that $z' \neq w$. Then we also have alternating A-circuits z', y', z, w, x, y, z' and $z', (y' = x), w, z, (y' = x), y, z'$, implying that $xy \in A$ and $z'w, z'z \in \overline{A}$. Consequently, $(x = y', y, z, w, z')$ is an A-pentagon. Since (w, z, y, x) is also an A-switching path, we can similarly conclude that if $w = z'$, then $(w = z', z, y, x, y')$ is an A-pentagon. $\square$

Let $<$ be an ordering of the vertices of G. Given two k-element subsets $S = \{s_1, s_2, \ldots, s_k\}$ and $T = \{t_1, t_2, \ldots, t_k\}$ of $V(G)$, where $s_1 < s_2 < \cdots < s_k$ and $t_1 < t_2 < \cdots < t_k$, S is said to be *lexicographically smaller* than T, denoted by $S < T$, if $s_j < t_j$ for some $j \in \{1, 2, \ldots, k\}$, and $s_i = t_i$ for all $1 \leq i < j$. In the usual way, we let $S \leq T$ denote the fact that either $S < T$ or $S = T$. For a set $S \subseteq V(G)$, we abbreviate $\min_< S$ to just $\min S$. Note that the relation $<$ ("is lexicographically smaller than") that we have defined on k-element subsets of $V(G)$ is a total order. Therefore, given a collection of k-element subsets of $V(G)$, the lexicographically smallest one among them is well-defined.

The following observation states a well-known property of Lex-BFS orderings [24].

Observation 2 ([24]). *Let $<$ denote a Lex-BFS ordering of the vertices of a graph G. For $a, b, c \in V(G)$, if $a < b < c$, $ab \notin E(G)$ and $ac \in E(G)$, then there exists $x \in V(G)$ such that $x < a < b < c$, $xb \in E(G)$ and $xc \notin E(G)$ (See Figure 2.8).*

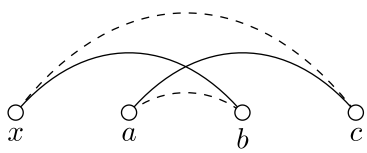


Figure 2.8: Property of Lex-BFS ordering

2.6 Proof of Theorem 5

Assume that G^* is bipartite.

We shall now construct a partial 2-coloring of the vertices of G^* using the colors $\{1, 2\}$ by means of an algorithm that consists of three phases. We shall describe the first two phases here, after which a partial 2-coloring of G^* is obtained. The third phase, which will be described later, modifies this coloring so as to obtain a 2-threshold cover of G.

Phase I. Construct a Lex-BFS ordering $<$ of G.

Recall that every vertex of G^* is a two-element subset of $V(G)$.

Phase II. For every non-trivial component C of G^*, perform the following operation:

> Choose the lexicographically smallest vertex in C (with respect to the ordering $<$) and assign the color 1 to it. Extend this to a proper coloring of C using the colors $\{1, 2\}$.

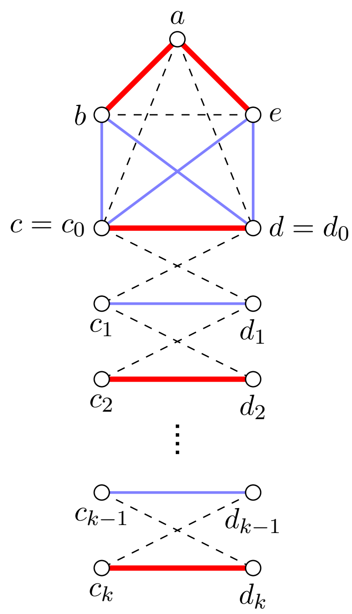


Figure 2.9: A strict F-pentagon together with the path $c_0d_0, c_1d_1, \ldots, c_kd_k$ (possibly, some of the vertices in the figure may coincide)

Note that after Phase II, every vertex of G^* that is in a non-trivial component has been colored either 1 or 2. For $i \in \{1, 2\}$, let $F_i = \{e \in V(G^*)\colon e \text{ is colored } i\}$. Further, let F_0 denote the set of all isolated vertices (trivial components) in G^*. Clearly, F_0 is exactly the set of uncolored vertices of G^* and we have $V(G^*) = F_0 \cup F_1 \cup F_2$. Note that since the opposite edges of any alternating 4-cycle in G correspond to adjacent vertices in G^*, one of them receives color 1 and the other color 2 in the partial 2-coloring of G^* constructed in Phase II. It follows that (F_0, F_1, F_2) is a valid 3-partition of $E(G)$.

2.6.1 No strict pentagons

In this section we shall prove that there are no strict pentagons in G with respect to (F_0, F_1, F_2).

Let $\{F, \overline{F}\} = \{F_1, F_2\}$. Let (a, b, c, d, e) be a strict F-pentagon and $c_0d_0, c_1d_1, \ldots, c_kd_k$ be a path in G^*, where $c_0 = c$, $d_0 = d$, $k \geq 0$, and for each $i \in \{0, 1, \ldots, k-1\}$, $c_id_{i+1}, d_ic_{i+1} \in E(\overline{G})$. Since $cd = c_0d_0 \in F$, it follows that $c_id_i \in F$ for all even i and $c_id_i \in \overline{F}$ for all odd i (See Figure 2.9).

Observation 3. *For each $i \in \{0, 1, \ldots, k\}$, the edges c_ib, c_ie, d_ib, d_ie exist and they belong to F when i is odd and to $\overline{F}$ when i is even.*

Proof. We prove this by induction on i. This is easily seen to be true when $i = 0$. Suppose that $i > 0$. We shall assume without loss of generality that i is odd as the other case is symmetric.

Then by the induction hypothesis, $c_{i-1}b, d_{i-1}b, c_{i-1}e, d_{i-1}e \in \overline{F}$. Then $c_i, d_i, c_{i-1}, b, e, d_{i-1}, c_i$ is an alternating $\overline{F}$-circuit (note that $c_i \neq b$ as $d_{i-1} \in N(b) \setminus N(c_i)$), implying that $c_i b \in F$. By symmetric arguments, we get $c_i e, d_i b, d_i e \in F$. □

Remark 1. *By the above observation, we have that:*

(a) for each $i \in \{0, 1, \ldots, k\}$, $c_i, d_i \notin \{b, e\}$,

(b) if $\{c_i, d_i\} \cap \{c_j, d_j\} \neq \emptyset$ for some $0 \leq i, j \leq k$, then $i \equiv j \mod 2$, and

(c) for each even $i \in \{0, 1, \ldots, k\}$, we have $a \notin \{c_i, d_i\}$.

Observation 4. *If $c_1 \neq a$, then (d, b, c_1, a, e) is a strict $\overline{F}$-pentagon. Similarly, if $d_1 \neq a$, then (c, b, d_1, a, e) is a strict $\overline{F}$-pentagon.*

Proof. By Observation 3, we have $c_1 b, c_1 e, d_1 b, d_1 e \in F$. Suppose that $c_1 \neq a$. Then c_1, b, e, a, c, d, c_1 is an alternating F-circuit, and therefore we have that $ac_1 \in \overline{F}$. It now follows that (d, b, c_1, a, e) is a strict $\overline{F}$-pentagon. By similar arguments, it can be seen that if $d_1 \neq a$, then $ad_1 \in \overline{F}$ and therefore (c, b, d_1, a, e) is a strict $\overline{F}$-pentagon. □

Observation 5. *Let $S_0 = \{a, c_0, d_0\}$ and for $1 \leq i \leq k$, let $S_i = S_{i-1} \cup \{c_i, d_i\}$. Let $i \in \{0, 1, \ldots, k\}$. For each $z \in \{c_i, d_i\}$, there exist $x_z, y_z \in S_i$ such that (x_z, b, y_z, z, e) is a strict F-pentagon when i is even and a strict $\overline{F}$-pentagon when i is odd.*

Proof. We are given an $i \in \{0, 1, \ldots, k\}$ and a vertex z that is either c_i or d_i. First let us consider the case when $z = a$. Since $z \in \{c_i, d_i\}$, we have by Remark 1(c), that i is odd, which implies that $i \geq 1$. Note that we have either $c_1 \neq a$ or $d_1 \neq a$. If $c_1 \neq a$, we define $x_z = d$, $y_z = c_1$ and if $d_1 \neq a$, we define $x_z = c$, $y_z = d_1$. Clearly, $x_z, y_z \in S_1 \subseteq S_i$, since $i \geq 1$. By Observation 4, we get that (x_z, b, y_z, z, e) is a strict $\overline{F}$-pentagon, and so we are done. Therefore, we shall now assume that $z \neq a$.

We shall now prove the statement of the observation by induction on i. Clearly, when $i = 0$, $z \in \{c_0, d_0\}$, so we can choose $x_z = a$, $y_z \in \{c, d\} \setminus \{z\}$ such that (x_z, b, y_z, z, e) is a strict F-pentagon (note that $x_z, y_z \in S_0$ as required). So let us assume that $i \geq 1$. If $z \in \{c_j, d_j\}$ for some $j < i$, then by Remark 1(b) we have that $j \equiv i \mod 2$ and by the induction hypothesis applied to j and z, there exist $x_z, y_z \in S_j \subseteq S_i$ (as $j < i$) such that (x_z, b, y_z, z, e) is a strict F-pentagon if i is even and a strict $\overline{F}$-pentagon if i is odd, completing the proof. Therefore, we assume that there is no $j < i$ such that $z \in \{c_j, d_j\}$. Since we have already assumed that $z \neq a$, we now have $z \notin S_{i-1}$.

Observe that there exists $z' \in \{c_{i-1}, d_{i-1}\}$ such that $z'z \in E(\overline{G})$. Then by the induction hypothesis, there exist $x_{z'}, y_{z'} \in S_{i-1}$ such that $(x_{z'}, b, y_{z'}, z', e)$ is a strict F-pentagon if $i-1$ is even and a strict $\overline{F}$-pentagon if $i-1$ is odd. Define $x_z = z'$ and $y_z = x_{z'}$. Then we have $x_z, y_z \in S_{i-1} \subseteq S_i$. Since $y_z \in S_{i-1}$ and $z \notin S_{i-1}$, we also have that $y_z \neq z$. Using Observation 3 and the fact that $(x_{z'}, b, y_{z'}, z', e)$ is a strict F-pentagon (resp. $\overline{F}$-pentagon) if i is odd (resp. even), we now have that $(y_z = x_{z'}), b, e, z, z', y_{z'}, (x_{z'} = y_z)$ is an alternating F-circuit (resp. $\overline{F}$-circuit). Therefore, $y_z z \in \overline{F}$ if i is odd and $y_z z \in F$ if i is even. Consequently we get that (x_z, b, y_z, z, e) is a strict F-pentagon when i is even and a strict $\overline{F}$-pentagon when i is odd. □

It is easy to see that Observation 5 implies the following.

Remark 2. *Let $\{F, \overline{F}\} = \{F_1, F_2\}$ and let (a, b, c, d, e) be any strict F-pentagon in G with respect to (F_0, F_1, F_2). Let $c'd'$ be a vertex in the same component as cd in G^*. Then for each $z \in \{c', d'\}$, there exist $x_z, y_z \in V(G)$ such that (x_z, b, y_z, z, e) is a strict F-pentagon if $c'd' \in F$ and a strict $\overline{F}$-pentagon if $c'd' \in \overline{F}$.*

Suppose that there is at least one strict pentagon in G with respect to (F_0, F_1, F_2). We say that a pentagon (a, b, c, d, e) is lexicographically smaller than a pentagon (a', b', c', d', e') if $\{a, b, c, d, e\} < \{a', b', c', d', e'\}$. Consider the lexicographically smallest strict pentagon (a, b, c, d, e) in G. Let $\{F, \overline{F}\} = \{F_1, F_2\}$ such that (a, b, c, d, e) is a strict F-pentagon. Since $cd \in F$, it belongs to a non-trivial component C of G^*. Therefore, there exists $uv \in E(G)$ such that $cv, du \in E(\overline{G})$ (so that $\{cd, uv\} \in E(G^*)$). Clearly, at least one of u, v is distinct from a. We assume without loss of generality that $u \neq a$ (by interchanging the labels of c and d if necessary). By applying Observation 4 to the path $(c_0 d_0 = cd), (c_1 d_1 = uv)$ in G^*, we get that (d, b, u, a, e) is a strict $\overline{F}$-pentagon, which implies that $au \in \overline{F}$. By Observation 3 applied to the same path, we get that $ub, ue \in F$.

Observation 6. $a > \min\{c, d\}$.

Proof. Suppose for the sake of contradiction that $a < \min\{c, d\}$. If $u < c$, then (d, b, u, a, e) is a strict $\overline{F}$-pentagon that is lexicographically smaller than (a, b, c, d, e), which is a contradiction. So we can assume that $c < u$, which gives us $a < c < u$. As $ac \in E(\overline{G})$ and $au \in E(G)$, by Observation 2, there exists a vertex x such that $x < a < c < u$, $xc \in E(G)$ and $xu \in E(\overline{G})$. Since a, u, x, c, a is an alternating 4-cycle in which $au \in \overline{F}$, we have that $xc \in F$. Then b, a, c, x, u, e, b is an alternating F-circuit (note that $b \neq x$ as $u \in N(b) \setminus N(x)$), and therefore $xb \in \overline{F}$. Symmetrically, we also get that $xe \in \overline{F}$. Then d, b, e, x, u, a, d is an alternating $\overline{F}$-circuit (note that $x \neq d$ as $x < a < \min\{c, d\}$), and therefore we have $xd \in F$. Now (u, b, x, d, e) is a strict F-pentagon that is lexicographically smaller than (a, b, c, d, e), which is a contradiction. $\square$

Let $c'd'$ be the lexicographically smallest vertex in C.

Observation 7. $\min\{c', d'\} = \min\{c, d\}$.

Proof. We know that $c'd' \leq cd$, and therefore $\min\{c', d'\} \leq \min\{c, d\}$. Suppose that $z = \min\{c', d'\} < \min\{c, d\}$. From Remark 2, we have that for each $z \in \{c', d'\}$, there exist vertices $x_z, y_z \in V(G)$ such that (x_z, b, y_z, z, e) is a strict pentagon. Since $a > \min\{c, d\}$ by Observation 6, we have $a > z$. Then (x_z, b, y_z, z, e) is a lexicographically smaller strict pentagon than (a, b, c, d, e) which is a contradiction. $\square$

Observation 8. $a > \max\{c, d\}$.

Proof. Let $\{y, \overline{y}\} = \{c, d\}$ such that $y < \overline{y}$. By Observation 6 it is now enough to show that $y < a < \overline{y}$ is not possible. Since $ya \in E(\overline{G})$ and $y\overline{y} \in E(G)$, $y < a < \overline{y}$ implies by Observation 2 that there exists $x < y$ such that $xa \in E(G)$ but $x\overline{y} \in E(\overline{G})$. Then $x, a, y, \overline{y}, x$ is an alternating 4-cycle, and therefore xa and $y\overline{y} = cd$ belong to the same component C of G^*. Thus $c'd' \leq xa$, which implies that $\min\{c', d'\} \leq \min\{x, a\}$. Since $\min\{x, a\} = x < y = \min\{c, d\}$, we now have $\min\{c', d'\} \leq \min\{x, a\} < \min\{c, d\}$. This contradicts Observation 7. $\square$

Since $c'd'$ is the lexicographically smallest vertex in C, our algorithm would have colored it with the color 1. Therefore, we have $c'd' \in F_1$. Consider a path $c_0d_0, c_1d_1, \ldots, c_kd_k$ in G^*, where $c_0 = c$, $d_0 = d$, $c_k = c'$ and $d_k = d'$, in which for each $i \in \{0, 1, \ldots, k-1\}$, $c_id_{i+1}, d_ic_{i+1} \in E(\overline{G})$. Suppose that $cd \in F_2$. Then since $c_kd_k = c'd' \in F_1$, we have that k is odd. Now by Remark 1(b), we have that $\{c_0, d_0\} \cap \{c_k, d_k\} = \emptyset$. But this contradicts Observation 7. Thus we have that $cd \in F_1$. Therefore, (a, b, c, d, e) is a strict F_1-pentagon, or in other words, $F = F_1$. Then, our earlier observations imply that $ub, ue \in F_1$ and $au \in F_2$.

Since ec, ab, ed and bc, ae, bd are paths in G^*, it follows that ec, ed lie in one component of G^* and bc, bd also lie in one component of G^*. Let D be the component containing bc, bd and D' the component containing ec, ed in G^*. Consider the lexicographically smallest vertex in $D \cup D'$. Let us assume without loss of generality that this vertex is in D (we can interchange the labels of b and e if required). Define $p_0 = b$, $q_0 = c$. Then in G^*, there exists a path $p_0q_0, p_1q_1, \ldots, p_tq_t$ between bc and the lexicographically smallest vertex p_tq_t in D. As before, for $0 \leq i \leq t-1$, we have $p_iq_{i+1}, q_ip_{i+1} \in E(\overline{G})$ and for $0 \leq i \leq t$, we have $p_iq_i \in F_1$ when i is odd and $p_iq_i \in F_2$

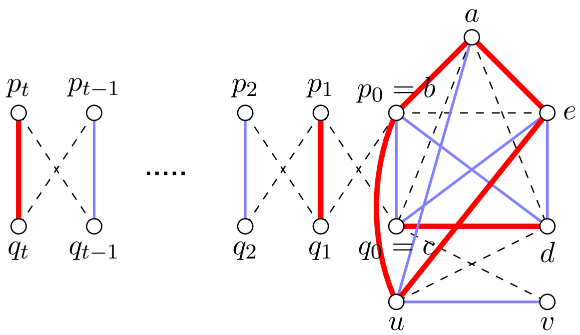


Figure 2.10: A strict F-pentagon together with the paths $p_0q_0, p_1q_1, \ldots, p_tq_t$ and cd, uv where $a \neq u$ (possibly, some other vertices in the figure may coincide) – (d, b, u, a, e) is an $\overline{F}$-pentagon

when i is even. Also, since p_tq_t is the lexicographically smallest vertex in its component in G^*, we know that $p_tq_t \in F_1$, which implies that t is odd (See Figure 2.10).

Observation 9. *Let $i \in \{0, 1, \ldots, t\}$. Then if i is odd, we have*

(a) $p_i \notin \{b, e\}$,

(b) $q_i \notin \{a, c, d\}$,

(c) $p_ib, p_ie \in F_1$,

(d) Either $p_i = a$ *or* $p_ia \in F_2$, *and*

(e) Either $q_ic \in F_2$ *or* $q_id \in F_2$.

and if i is even, we have

(a) $q_i \notin \{b, e\}$,

(b) $p_i \notin \{a, c, d\}$,

(c) $q_ib, q_ie \in F_2$,

(d) Either $q_i = d$ *or* $q_id \in F_1$, *and*

(e) Either $p_iu \in F_1$ *or* $p_ia \in F_1$.

Proof. We shall prove this by induction on i. If $i = 0$, then the statement of the lemma can be easily seen to be true. Suppose that $i > 0$. We give a proof for the case when i is odd (the case when i is even is symmetric and can be proved using similar arguments). By the induction hypothesis, $q_{i-1}b, q_{i-1}e \in F_2$, and therefore since $p_iq_{i-1} \in E(\overline{G})$, we have $p_i \notin \{b, e\}$. We now prove the following claim.

Claim 1. For $x \in \{a, u\}$, if $p_i = x$ or $p_i x \in F_2$, then $p_i b, p_i e \in F_1$.

If $p_i = x$ then there is nothing to prove as we already know that $ab, ae, ub, ue \in F_1$. So assume that $p_i x \in F_2$. Let $\{z, \bar{z}\} = \{b, e\}$. Then $p_i, x, d, z, \bar{z}, q_{i-1}, p_i$ is an alternating F_2-circuit (recall that $p_i \notin \{b, e\}$), which implies that $p_i z \in F_1$. We thus get that $p_i b, p_i e \in F_1$. This proves the claim.

By the induction hypothesis we know that either $p_{i-1} a \in F_1$ or $p_{i-1} u \in F_1$, and also that $p_{i-1} \notin \{c, d\}$. First suppose that $p_{i-1} a \in F_1$. This implies that $q_i \neq a$. Let $\{y, \bar{y}\} = \{c, d\}$. Then we have that $p_{i-1}, a, \bar{y}, y$ is an alternating F_1-path implying that $p_{i-1} y \in E(G)$. Thus, $p_{i-1} c, p_{i-1} d \in E(G)$. This implies that $q_i \notin \{c, d\}$. By the induction hypothesis we also have that $q_{i-1} y \in F_1$ for some $y \in \{c, d\}$. Then $q_i, p_i, q_{i-1}, y, a, p_{i-1}, q_i$ is an alternating F_1-circuit, which implies that $q_i y \in F_2$. If $p_i \neq a$, then $p_i, q_i, p_{i-1}, a, y, q_{i-1}, p_i$ is an alternating F_1-circuit, implying that $p_i a \in F_2$. Since we have either $p_i = a$ or $p_i a \in F_2$ we are done by Claim 1.

Therefore we can assume that $p_{i-1} a \notin F_1$. If $i = 1$, then we know that $p_{i-1} a = ba \in F_1$, so we can assume that $i \geq 2$. By the induction hypothesis, we have that for some $y \in \{c, d\}$, $q_{i-2} y \in F_2$. Therefore if $p_{i-1} a \in E(G)$, then we have that $p_{i-1}, a, y, q_{i-2}, p_{i-1}$ is an alternating 4-cycle in which $q_{i-2} y \in F_2$, implying that $p_{i-1} a \in F_1$ which is a contradiction. Since we know that $p_{i-1} \neq a$ by the induction hypothesis, we can assume that $p_{i-1} a \in E(\overline{G})$. Note that since $p_{i-1} a \notin F_1$, we have by the induction hypothesis that $p_{i-1} u \in F_1$. If $q_{i-1} = d$, then $p_{i-1}, (q_{i-1} = d), u, a, p_{i-1}$ is an alternating 4-cycle whose opposite edges both belong to F_2, which is a contradiction. Therefore by the induction hypothesis we have $q_{i-1} d \in F_1$. If $q_i = a$ (resp. $q_i = c$) then $p_i, (q_i = a), d, q_{i-1}, p_i$ (resp. $p_{i-1}, u, d, (c = q_i), p_{i-1}$) is an alternating 4-cycle whose opposite edges are both in F_1, which is a contradiction. Therefore, $q_i \notin \{a, c\}$. If $p_i a \in F_2$ then we have that $a, p_i, q_{i-1}, p_{i-1}, a$ is an alternating 4-cycle whose opposite edges are both in F_2, which is a contradiction. This implies that $p_i a \notin F_2$ and therefore $p_i \neq u$. Then $p_i, q_i, p_{i-1}, u, d, q_{i-1}, p_i$ is an alternating F_1-circuit, implying that $p_i u \in F_2$. Therefore by Claim 1, we have that $p_i b, p_i e \in F_1$. Now if $a \neq p_i$, then $p_i, b, e, a, d, q_{i-1}, p_i$ is an alternating F_1-circuit, which implies that $p_i a \in F_2$ which is a contradiction. This implies that $a = p_i$, which further implies that $q_i \neq d$. Then $q_i, p_i, q_{i-1}, d, u, p_{i-1}, q_i$ is an alternating F_1-circuit, which implies that $q_i d \in F_2$ and we are done. $\square$

Observation 10. *For each even $i \in \{0, 1, 2, \ldots, t\}$, either $ap_i \in E(G)$ or both $dq_{i-1}, dq_{i+1} \in E(G)$.*

Proof. Suppose that there exists an even $i \in \{0, 1, 2, \ldots, t\}$ and $j \in \{i-1, i+1\}$ such that

$ap_i, dq_j \notin E(G)$. By Observation 9, we know that $p_i \neq a$ and $q_j \neq d$. So we have $ap_i, dq_j \in E(\overline{G})$. Now if $d \neq q_i$, then we have by Observation 9 that $q_i d \in F_1$. Then p_j, q_j, d, q_i, p_j is an alternating 4-cycle whose both opposite edges belong to F_1, which is a contradiction. Therefore we can assume that $d = q_i$. Then $(d = q_i), p_i, a, u, (d = q_i)$ is an alternating 4-cycle whose opposite edges both belong to F_2, which is again a contradiction. □

Recall that D' is the component containing ec in G^*.

Observation 11. *For any odd $i \in \{0, 1, \ldots, t\}$, if $ap_{i-1} \in E(G)$, then for each $y \in \{c, d\}$ for which $yq_i \in E(G)$, we have $yq_i \in D'$. On the other hand, if $ap_{i-1} \notin E(G)$, then $dq_i \in D'$.*

Proof. We prove this by induction on i. When $i = 1$, we have $ap_0 = ab \in E(G)$ and for each $y \in \{c, d\}$ such that $yq_1 \in E(G)$, we have that $ec, (ab = ap_0), yq_1$ is a path in G^*. We thus have the base case. We shall now prove the claim for $i \geq 3$ assuming that the claim is true for $i - 2$. Suppose that $ap_{i-1} \in E(G)$. By Observation 9, there exists $y'' \in \{c, d\}$ such that $y''q_{i-2} \in E(G)$. By the induction hypothesis, either $y''q_{i-2} \in D'$ or $dq_{i-2} \in D'$ (depending upon whether ap_{i-3} is an edge or not). Thus in any case, we have that there exists $y' \in \{c, d\}$ such that $y'q_{i-2} \in D'$. Now for each $y \in \{c, d\}$ such that $yq_i \in E(G)$, since $y'q_{i-2}, ap_{i-1}, yq_i$ is a path in G^*, we get that $yq_i \in D'$, so we are done. Next, suppose that $ap_{i-1} \notin E(G)$. Then by Observation 9, we have $up_{i-1} \in E(G)$ and by Observation 10, we have $dq_{i-2}, dq_i \in E(G)$. We then have by the induction hypothesis that $dq_{i-2} \in D'$. Since dq_{i-2}, up_{i-1}, dq_i is a path in G^*, we have $dq_i \in D'$. □

Recall that C is the component of G^* containing the vertex cd.

Observation 12. *For each odd $i \in \{0, 1, \ldots, t\}$, if $a \neq p_i$ then $ap_i \in C$.*

Proof. We prove this by induction on i. The base case when $i = 1$ is true since if $a \neq p_1$ then by Observation 9, $ap_1 \in E(G)$, and since $\{ap_1, (q_0 = c)d\} \in E(G^*)$, we have $ap_1 \in C$. Assume that $i \geq 3$ and the claim is true for $i - 2$. Suppose that $a \neq p_i$. Then we have $ap_i \in E(G)$ by Observation 9. If $d = q_{i-1}$ then we have $\{ap_i, c(q_{i-1} = d)\} \in E(G^*)$, so we have $ap_i \in C$. So we assume that $d \neq q_{i-1}$. Then by Observation 9, we have that $dq_{i-1} \in E(G)$. By the induction hypothesis, we have that either $ap_{i-2} \in C$ or $a = p_{i-2}$. If $ap_{i-2} \in C$, then since ap_i, dq_{i-1}, ap_{i-2} is a path in G^*, we have $ap_i \in C$. On the other hand, if $a = p_{i-2}$ then we again have $ap_i \in C$ as $ap_i, dq_{i-1}, u(p_{i-2} = a), cd$ is a path in G^*. □

Recall that t is odd, $p_t q_t \in D$, and $p_t q_t$ is the lexicographically smallest vertex in $D \cup D'$.

Observation 13. $p_t < \min\{c, d\}$

Proof. Let $\{y, \bar{y}\} = \{c, d\}$, where $y < \bar{y}$. Note that $p_t \notin \{c, d\}$, since by Observation 9, $p_t b \in F_1$, but we know that $cb, db \in F_2$. By the same lemma, we also have that $q_t \notin \{c, d\}$. Therefore as $\min\{p_t, q_t\} \leq \min\{c, d\}$ (since $p_t q_t < bc, bd$), we have that $\min\{p_t, q_t\} < \min\{c, d\} = y$. Now if $p_t = \min\{p_t, q_t\}$ then we are done. Therefore let us assume that $q_t = \min\{p_t, q_t\}$, and so $q_t < y$.

Suppose that $yq_t \in E(G)$. If $yq_t \notin D'$, then by Observation 11, we have that $ap_{t-1} \notin E(G)$ and $\overline{y}q_t \in D'$. By Observation 9, we know that $p_{t-1} \neq a$, which implies that $ap_{t-1} \in E(\overline{G})$. By our choice of $p_t q_t$, we now have that $p_t q_t < \overline{y}q_t$, which implies that $p_t < \overline{y}$. Now by Observation 8, $p_t \neq a$, which implies by Observation 9 that $p_t a \in F_2$. Then $a, p_t, q_{t-1}, p_{t-1}, a$ is an alternating 4-cycle in which both opposite edges belong to F_2, which is a contradiction. We can thus conclude that $yq_t \in D'$. Then by our choice of $p_t q_t$, we have that $p_t < y$, and we are done. So we assume that $yq_t \notin E(G)$.

Recall that $q_t < y$ (and therefore $yq_t \in E(\overline{G})$). Now if $y < p_t$ then we have $q_t < y < p_t$ where $q_t y \notin E(G)$ and $q_t p_t \in E(G)$. By Observation 2, this implies that there exists $x < q_t$ such that $xy \in E(G)$ and $xp_t \notin E(G)$ (which means that $xp_t \in E(\overline{G})$ since $x < p_t$). Then $\{xy, p_t q_t\} \in E(G^*)$, which implies that $xy \in D$. But $xy < p_t q_t$, which contradicts our choice of $p_t q_t$. We can thus conclude that $p_t < y$ (recall that $p_t \neq y$ as $p_t \notin \{c, d\}$) and we are done. □

Note that by Observation 13 and Observation 8 we have that $a \neq p_t$. Then by Observation 12, we have $ap_t \in C$. By Observation 13 and Observation 7, $p_t < \min\{c', d'\}$, which implies that $ap_t < c'd'$. This is a contradiction to our choice of $c'd'$. Therefore we have the following lemma.

Lemma 3. *There are no strict pentagons in G (with respect to (F_0, F_1, F_2)).*

2.6.2 No strict switching paths

In this section, we show that there are no strict switching paths either in G with respect to (F_0, F_1, F_2). First we note the following observation.

Observation 14. *Let (x, y, z, w) be a strict switching path with respect to (F_0, F_1, F_2). Let $y'z' \in E(G)$ be such that $yz', zy' \in E(\overline{G})$. Then, $y' \neq x$ and $z' \neq w$.*

Proof. Let $\{F, \overline{F}\} = \{F_1, F_2\}$. Suppose that (x, y, z, w) is a strict F-switching path. Then we have that $xy, zw \in F$, $yz \in \overline{F}$, and $xw \in E(\overline{G})$. By Lemma 2 and the fact that $zw, xy \in F$, we know that if $y' = x$ then $(x = y', y, z, w, z')$ is a strict F-pentagon, and if $z' = w$ then $(w = z', z, y, x, y')$ is a strict F-pentagon. Since we know by Lemma 3 that there are no strict pentagons in G, we can conclude that $y' \neq x$ and $z' \neq w$. □

Now we show that there are no strict switching paths in G. Suppose not. Then let (a, b, c, d) be the lexicographically smallest strict switching path in G.

Observation 15. *(a, b, c, d) is not a strict F_1-switching path.*

Proof. Suppose for the sake of contradiction that (a, b, c, d) is a strict F_1-switching path. Let C be the component of G^* containing bc. Let $b_0c_0, b_1c_1, \ldots, b_kc_k$, where $b_0 = b$ and $c_0 = c$, be a path in C between bc and the lexicographically smallest vertex b_kc_k in C. We assume that for each $i \in \{0, 1, \ldots, k-1\}$, $b_ic_{i+1}, c_ib_{i+1} \in E(\overline{G})$. As $b_0c_0 \in F_2$, it follows that $b_ic_i \in F_2$ for each even i and $b_ic_i \in F_1$ for each odd i. Since b_kc_k is the lexicographically smallest vertex in its component in G^*, we know that $b_kc_k \in F_1$, which implies that k is odd.

We claim that $b_ia, c_id \in F_1$ for each even i and $b_ia, c_id \in F_2$ for each odd i, where $0 \leq i \leq k$. We prove this by induction on i. The case where $i = 0$ is trivial as $b_0 = b$ and $c_0 = c$. So let us assume that $i > 0$. Consider the case where i is odd. As $i - 1$ is even, by the induction hypothesis we have $b_{i-1}a, c_{i-1}d \in F_1$. Since $b_{i-1}c_{i-1} \in F_2$, we can observe that, (a, b_{i-1}, c_{i-1}, d) is a strict F_1-switching path. Then by Observation 14, we have that $a \neq b_i$ and $d \neq c_i$. Now the alternating F_1-circuits $b_i, c_i, b_{i-1}, a, d, c_{i-1}, b_i$ and $c_i, b_i, c_{i-1}, d, a, b_{i-1}, c_i$ imply that $b_ia, c_id \in F_2$. The case where i is even is symmetric and hence the claim.

By the above claim, $b_ka, c_kd \in F_2$. Since $b_kc_k \in F_1$, we now have that (a, b_k, c_k, d) is a strict F_2-switching path. Since $b_kc_k < bc$, we have that $\{a, b_k, c_k, d\} < \{a, b, c, d\}$, which is a contradiction to our assumption that (a, b, c, d) is the lexicographically smallest strict switching path in G. □

Observation 16. *(a, b, c, d) is not a strict F_2-switching path.*

Proof. Suppose for the sake of contradiction that (a, b, c, d) is a strict F_2-switching path. By the symmetry between a and d, we can assume without loss of generality that $a < d$.

As $bc \in F_1$, the vertex bc belongs to a non-trivial component of G^*. Then there exists a neighbor uv of bc in G^* such that $bv, uc \in E(\overline{G})$. As $bc \in F_1$, we have $uv \in F_2$. By Observation 14, we have that $u \neq a$. Then a, b, v, u, c, d, a is an alternating F_2-circuit, implying that $au \in F_1$. As $ab \in F_2$, we know that ab is not the lexicographically smallest vertex in its component. Let $a_0b_0, a_1b_1, \ldots, a_kb_k$ be a path in G^* between ab and the lexicographically smallest vertex a_kb_k in its component, where $a_0 = a$, $b_0 = b$, and for $0 \leq i < k$, $a_ib_{i+1}, a_{i+1}b_i \in E(\overline{G})$. Note that for $0 \leq i \leq k$, $a_ib_i \in F_2$ if i is even and $a_ib_i \in F_1$ if i is odd. Since $a_kb_k \in F_1$ (as it is the lexicographically smallest vertex in its component in G^*), this implies that k is odd.

We claim that for $0 \leq i \leq k$, $a_iu, b_ic \in F_1$ if i is even and $a_iu, b_ic \in F_2$ if i is odd. We prove this by induction on i. The base case when $i = 0$ is trivial, since $au, bc \in F_1$. Let $i > 0$ be odd. By the induction hypothesis we have that $a_{i-1}u, b_{i-1}c \in F_1$. Since $a_{i-1}b_{i-1} \in F_2$ we can observe that (u, a_{i-1}, b_{i-1}, c) is a strict F_1-switching path. Therefore by Observation 14, we have that $a_i \neq u$ and $b_i \neq c$. Then we have alternating F_1-circuits $a_i, b_i, a_{i-1}, u, c, b_{i-1}, a_i$ and $b_i, a_i, b_{i-1}, c, u, a_{i-1}, b_i$, implying that $a_iu, b_ic \in F_2$. The case when i is even is symmetric. This proves our claim. Since k is odd, we now have that $a_ku, b_kc \in F_2$. Note that now (c, b_k, a_k, u) is a strict F_2-switching path.

Suppose that $d < b$. Then we have that $a < d < b$, where $ad \in E(\overline{G})$ and $ab \in E(G)$. Therefore by Observation 2, there exists $x < a$ such that $xd \in E(G)$ and $xb \in E(\overline{G})$. Then x, d, a, b, x is an alternating 4-cycle in which $ab \in F_2$, implying that $xd \in F_1$. Then we have a strict F_1-switching path (x, d, c, b) such that $\{x, d, c, b\} < \{a, b, c, d\}$, which is a contradiction to the choice of (a, b, c, d). Therefore we can assume that $b < d$. Since $a_kb_k < ab$ and $a, b < d$, we have that $\{c, b_k, a_k, u\} < \{a, b, c, d\}$. As (c, b_k, a_k, u) is a strict switching path, this contradicts the choice of (a, b, c, d). □

From Observation 15 and Observation 16, we have the following lemma.

Lemma 4. *There are no strict switching paths in G (with respect to (F_0, F_1, F_2)).*

Remark 3. *Recall that, given a 2-coloring of G^* in which the color classes are denoted by E_1 and E_2, Raschle and Simon [102] define an AP_6 in G to be a sequence $v_0, v_1, \ldots, v_5, v_0$ of distinct vertices of G such that $v_0v_1, v_2v_3, v_4v_5 \in E_i$ for some $i \in \{1, 2\}$ and $v_1v_2, v_3v_4, v_5v_0 \in E(\overline{G})$. Raschle and Simon observed that if G^* has an AP_6-free 2-coloring, then G has a 2-threshold cover and it can be computed in time $O(|E(G)|^2)$ (using Theorem 3.1, Theorem 2.5, Fact 2 and Fact 1 in [102]). The major part of the work of Raschle and Simon is to show that an AP_6-free 2-coloring of G^* always exists if G^* is bipartite and that it can be computed in time $O(|E(G)|^2)$ (Sections 3.2 and 3.3 of [102]). It can be seen that any 2-coloring of G^* obtained by extending the partial 2-coloring of G^* computed after Phases I and II of our algorithm is in fact an AP_6-free 2-coloring of G^* as follows. Let E_1 and E_2 be the color classes of such a 2-coloring of G^*. We can assume without loss of generality that $F_1 \subseteq E_1$ and $F_2 \subseteq E_2$. Note that $F_0 \subseteq E_1 \cup E_2$. Suppose that there is an AP_6 $v_0, v_1, \ldots, v_5, v_0$ in G with respect to this coloring where the edges $v_0v_1, v_2v_3, v_4v_5 \in E_i$, where $i \in \{1, 2\}$. Note that $(\emptyset, E_1, E_2)$ is a valid 3-partition of $E(G)$. For each even $j \in \{0, 1, \ldots, 5\}$, since $v_j, v_{j+1}, v_{j+2}, v_{j+3}$ (subscripts modulo 6) is an alternating E_i-path, we have that $v_jv_{j+3} \in E(G)$. This implies that for each even*

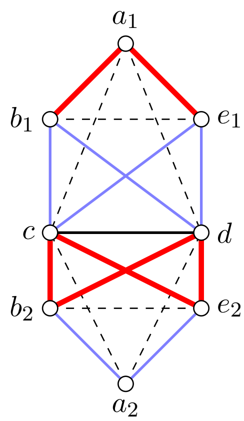


Figure 2.11: An F-pentagon and an $\overline{F}$-pentagon having same edge cd (possibly, $a_1 = a_2$)

$j \in \{0, 1, \ldots, 5\}$, $v_j, v_{j+1}, v_{j+2}, (v_{j+5} = v_{j-1}), v_j$ is an alternating 4-cycle in G (note that from the previous observation, we have $v_{j+2}v_{j+5} \in E(G)$), from which it follows that v_jv_{j+1} is in a non-trivial component of G^*. Therefore, $v_0v_1, v_2v_3, v_4v_5 \notin F_0$. Since these edges belong to E_i, it follows that $v_0v_1, v_2v_3, v_4v_5 \in F_i$. Then $v_0, v_1, \ldots, v_5, v_0$ is an alternating F_i-circuit, and therefore $v_0v_3 \in F_{3-i}$. This implies that (v_2, v_3, v_0, v_1) is a strict F_i-switching path in G, which contradicts Lemma 4. Thus the proof of Theorem 5 can already be completed using the observations in [102]. In the next section, we nevertheless give a self-contained proof that shows that G has a 2-threshold cover without using the "threshold completion" method used in [70, 102]. Also note that since it is clear that Phases I and II of the algorithm, and also the initial construction of G^*, can be done in time $O(|E(G)|^2)$, we obtain a simple algorithm with the same time complexity that computes the 2-threshold cover of a graph G whose auxiliary graph G^* is bipartite (note however that there is a faster algorithm for computing a 2-threshold cover due to Sterbini and Raschle [112]).

2.6.3 Constructing the 2-threshold cover of G

Observation 17. *There does not exist $a_1, a_2, b_1, b_2, e_1, e_2, c, d \in V(G)$ such that (a_1, b_1, c, d, e_1) is an F_1-pentagon and (a_2, b_2, c, d, e_2) is an F_2-pentagon.*

Proof. Suppose not (see Figure 2.11). Then as $b_1c, e_1c \in F_2$ and $b_2c, e_2c \in F_1$, we have $\{b_1, e_1\} \cap \{b_2, e_2\} = \emptyset$. Then b_1, a_1, c, b_2 and e_1, a_1, c, e_2 are alternating F_1-paths, implying that $b_1b_2, e_1e_2 \in E(G)$. As b_1, b_2, e_2, e_1, b_1 is an alternating 4-cycle, we have $\{b_1b_2, e_1e_2\} \in E(G^*)$. Thus, $b_1b_2 \notin F_0$, or in other words, $b_1b_2 \in F_1 \cup F_2$. If $b_1b_2 \in F_1$, then (c, b_1, b_2, a_2) is a strict F_2-switching path, which contradicts Lemma 4. On the other hand, if $b_1b_2 \in F_2$, then (c, b_2, b_1, a_1) is a strict F_1-switching path, which again gives a contradiction to Lemma 4. $\square$

We shall now describe Phase III of the algorithm that yields a partial 2-coloring of G^* that can be directly converted into a 2-threshold cover of G.

Phase III. For each $i \in \{1, 2\}$, let

$$S_i = \{cd \in F_0 \colon \exists a, b, e \in V(G) \text{ such that } (a, b, c, d, e) \text{ is an } F_i\text{-pentagon in } G \text{ with respect to } (F_0, F_1, F_2)\}.$$

Color every vertex in S_1 with 2 and every vertex in S_2 with 1.

Let F_0' be the set of vertices of G^* that are uncolored after Phase III, and for $i \in \{1, 2\}$, let F_i' be the set of vertices of G^* that are colored i. Clearly, $F_0' = F_0 \setminus (S_1 \cup S_2)$, $F_1' = F_1 \cup S_2$ and $F_2' = F_2 \cup S_1$. Note that $S_1, S_2 \subseteq F_0$ and that $S_1 \cap S_2 = \emptyset$ by Observation 17. It is easy to see that $\{F_0', F_1', F_2'\}$ is a partition of $E(G)$. Further, since $F_1 \subseteq F_1'$, $F_2 \subseteq F_2'$ and (F_0, F_1, F_2) is a valid 3-partition of $E(G)$, it follows that (F_0', F_1', F_2') is also a valid 3-partition of $E(G)$. We shall show that $(V(G), F_0' \cup F_1')$ and $(V(G), F_0' \cup F_2')$ are both threshold graphs, thereby completing the proof of Theorem 5. From here onwards, we use the terms "pentagons" and "switching paths" with respect to (F_0', F_1', F_2') unless otherwise mentioned.

Lemma 5. *There are no pentagons in G with respect to (F_0', F_1', F_2').*

Proof. Suppose for the sake of contradiction that (a, b, c, d, e) is a pentagon in G with respect to (F_0', F_1', F_2'). Let $i \in \{1, 2\}$ such that (a, b, c, d, e) is an F_i'-pentagon. Recall that ec, ab, ed and bc, ae, bd are paths in G^* and hence each of ab, ae, bc, bd, ec, ed is in a non-trivial component of G^*. Thus none of them is in F_0. Since $ab, ae \in F_i'$ and $bc, bd, ec, ed \in F_{3-i}'$, this implies that $ab, ae \in F_i$ and $bc, bd, ec, ed \in F_{3-i}$. Since (a, b, c, d, e) is an F_i'-pentagon, we have $cd \in F_0' \cup F_i'$. This implies that $cd \notin F_{3-i}'$ and that $cd \in F_0 \cup F_i$. If $cd \in F_0$, then (a, b, c, d, e) is an F_i-pentagon in G with respect to (F_0, F_1, F_2), which implies that $cd \in S_i$, and therefore $cd \in F_{3-i}'$. Since this is a contradiction, we can assume that $cd \in F_i$. Then (a, b, c, d, e) is a strict F_i-pentagon in G with respect to (F_0, F_1, F_2), contradicting Lemma 3. □

Lemma 6. *There are no switching paths in G with respect to (F_0', F_1', F_2').*

Proof. Suppose not. Let (a, b, c, d) be a switching path in G with respect to (F_0', F_1', F_2'). Let $i \in \{1, 2\}$ such that (a, b, c, d) is an F_i'-switching path. Then we have $ad \in E(\overline{G})$, $ab, cd \in F_i' \cup F_0'$, and $bc \in F_{3-i}'$. Suppose that bc belongs to a non-trivial component of G^*. Then there exists $uv \in E(G)$ such that $bv, cu \in E(\overline{G})$. By Lemma 2 and Lemma 5, we have that $a \neq u$ and $d \neq v$. Notice that since $bc \in F_{3-i}'$ and b, c, u, v, b is an alternating 4-cycle, we have $uv \in F_i'$.

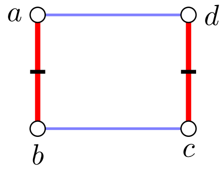


Figure 2.12: F-switching cycle – $ab, cd \in F \cup F'_0$ and $bc, ad \in \overline{F}$

Then d, c, u, v, b, a, d and a, b, v, u, c, d, a are alternating F'_i-circuits, implying that $dv, au \in F'_{3-i}$ and $ab, cd \in F'_i$. This further implies that (a, b, c, d) is a strict F'_i-switching path with respect to (F'_0, F'_1, F'_2). Since b, a, d, v, b and c, d, a, u, c and b, c, u, v, b are alternating 4-cycles, we also have that $ab, cd, bc \notin F_0$, which further implies that $ab, cd \in F_i$ and $bc \in F_{3-i}$. Then (a, b, c, d) is also a strict F_i-switching path with respect to (F_0, F_1, F_2), which is a contradiction to Lemma 4.

Therefore we can assume that bc belongs to a trivial component in G^*, i.e. $bc \in F_0$. Since $bc \in F'_{3-i}$, it should be the case that $bc \in S_i$, which implies that there exists an F_i-pentagon (x, y, b, c, z) in G with respect to (F_0, F_1, F_2). Since $ab, cd \in F'_i \cup F'_0 \subseteq F_i \cup F_0$, we know that $a, d \notin \{x, y, z\}$. Since a, b, x, y and d, c, x, z are alternating F_i-paths, we have that $ay, dz \in E(G)$. Since a, y, z, d, a is an alternating 4-cycle, we know that one of ay, dz is in F_i and the other in F_{3-i}. Because of symmetry, we can assume without loss of generality that $ay \in F_i$ and $dz \in F_{3-i}$ (by renaming (a, b, c, d) as (d, c, b, a) and interchanging the labels of y and z if necessary). Then a, y, z, x, c, d, a is an alternating F_i-circuit, implying that $ax \in F_{3-i}$. Then (a, x, z, d) is a strict F_{3-i}-switching path in G with respect to (F_0, F_1, F_2), which again contradicts Lemma 4. $\square$

Let $\{F, \overline{F}\} = \{F'_1, F'_2\}$. We say that (a, b, c, d) is an *F-switching cycle* in G with respect to (F'_0, F'_1, F'_2) if $ab, cd \in F \cup F'_0$ and $bc, ad \in \overline{F}$. As before, we say that (a, b, c, d) is a *switching cycle* in G with respect to (F'_0, F'_1, F'_2) if there exists $F \in \{F'_1, F'_2\}$ such that (a, b, c, d) is an F-switching cycle. See Figure 2.12 for an illustration.

Lemma 7. *There are no switching cycles in G with respect to (F'_0, F'_1, F'_2).*

Proof. Suppose not. Let (a, b, c, d) be a switching cycle in G with respect to (F'_0, F'_1, F'_2). Let $i \in \{1, 2\}$ such that (a, b, c, d) is an F'_i-switching cycle. Then we have $ab, cd \in F'_i \cup F'_0$ and $ad, bc \in F'_{3-i}$. Suppose that bc belongs to a non-trivial component of G^*. Then there exists $uv \in E(G)$ such that $bv, cu \in E(\overline{G})$. Since b, c, u, v, b is an alternating 4-cycle and $bc \in F'_{3-i}$, we have that $uv \in F'_i$. If $u = a$ and $v = d$, then $b, (a = u), c, (d = v), b$ is an alternating 4-cycle in which both the opposite edges belong to $F'_i \cup F'_0$, which is a contradiction. Therefore, either $u \neq a$ or $v \neq d$. Because of symmetry, we can assume without loss of generality that $u \neq a$ (by renaming (a, b, c, d) as (d, c, b, a) and interchanging the labels of u and v if necessary). Then

a, b, v, u is an alternating F'_i-path, implying that $au \in E(G)$. If $au \in F'_i \cup F'_0$ then (c, d, a, u) is an F'_i-switching path, and if not, then $au \in F'_{3-i}$, in which case (b, a, u, v) is an F'_i-switching path. In both cases, we have a contradiction to Lemma 6.

Therefore we can assume that bc belongs to a trivial component of G^*, i.e. $bc \in F_0$. Since $bc \in F'_{3-i}$, it should be the case that $bc \in S_i$, which implies that there exists an F_i-pentagon (x, y, b, c, z) in G with respect to (F_0, F_1, F_2). Since $ab, cd \in F'_i \cup F'_0 \subseteq F_i \cup F_0$, $a, d \notin \{x, y, z\}$. As y, x, b, a and z, x, c, d are alternating F_i-paths, we have that $ya, zd \in E(G)$. Now if both $ya, zd \in F'_i \cup F'_0$ we have that (y, a, d, z) is an F'_i-switching path, which is a contradiction to Lemma 6. On the other hand, if $ya \in F'_{3-i}$ or $zd \in F'_{3-i}$, then since $xy, xz \in F_i \subseteq F'_i$, we have that either (x, y, a, b) or (x, z, d, c) is an F'_i-switching path, which again contradicts Lemma 6. □

We are now ready to complete the proof of Theorem 5. Consider the graphs H_1, H_2, having $V(H_1) = V(H_2) = V(G)$, $E(H_1) = F'_1 \cup F'_0$ and $E(H_2) = F'_2 \cup F'_0$. We claim that H_1 and H_2 are both threshold graphs. Suppose for the sake of contradiction that H_i is not a threshold graph for some $i \in \{1, 2\}$. Then there exist edges $ab, cd \in E(H_i)$ such that $bc, ad \in E(\overline{H_i})$. If $bc, ad \in E(\overline{G})$, then a, b, c, d, a is an alternating 4-cycle in G whose opposite edges both belong to $F'_i \cup F'_0$, which contradicts the fact that (F'_0, F'_1, F'_2) is a valid 3-partition. So we can assume by symmetry that $bc \in E(G)$. Since $bc \in E(\overline{H_i})$, $bc \notin F'_i \cup F'_0$, which implies that $bc \in F'_{3-i}$. Now if $ad \in E(\overline{G})$, then (a, b, c, d) is an F'_i-switching path in G with respect to (F'_0, F'_1, F'_2), which is a contradiction to Lemma 6. On the other hand, if $ad \in E(G)$, then $ad \in F'_{3-i}$ (since $ad \in E(\overline{H_i})$), which implies that (a, b, c, d) is an F'_i-switching cycle in G with respect to (F'_0, F'_1, F'_2), which contradicts Lemma 7. Thus we can conclude that both H_1 and H_2 are threshold graphs. Since $E(G) = E(H_1) \cup E(H_2)$, we further get that $\{H_1, H_2\}$ is a 2-threshold cover of G.

2.7 Simpler proofs for paraglider-free graphs and split graphs

We now show that our proof of Theorem 5, as well as the algorithm to construct a 2-threshold cover of a graph G whose auxiliary graph G^* is bipartite, becomes considerably simpler if G is a "paraglider-free graph" or a "split graph". A *paraglider* is the graph $\overline{P_3 \cup K_2}$ (See Figure 2.13). Note that the subgraph formed by the edges of a pentagon in a graph is a paraglider. A graph is said to be *paraglider-free* if it contains no induced subgraph isomorphic to a paraglider. Thus, paraglider-free graphs cannot contain any pentagons with respect to any valid 3-partition of $E(G)$.

A graph $G = (X, Y, E)$ is said to be a *split graph* if X is a clique in G, Y is an independent

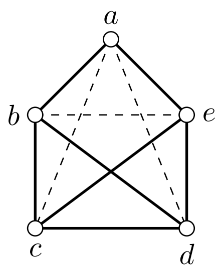


Figure 2.13: Paraglider graph

set in G and $V(G) = X \cup Y$. It is also known that split graphs are precisely $\{2K_2, C_4, C_5\}$-free graphs. As the paraglider contains an induced C_4, split graphs are paraglider-free.

Let G be a graph such that G^* is bipartite. Suppose that G is paraglider-free. Then the proof of Theorem 5 can be simplified as follows. We skip Phase III of our algorithm. Thus, once we finish running Phases I and II of our algorithm and obtain the valid 3-partition (F_0, F_1, F_2) of $E(G)$, we output $H_1 = (V(G), F_1 \cup F_0)$ and $H_2 = (V(G), F_2 \cup F_0)$ as the two threshold graphs that form a 2-threshold cover of G. We can do this because, the fact that G is paraglider-free implies that G does not contain any pentagons after Phase II. Thus we can conclude that Lemma 3 holds without any proof (the whole of Section 2.6.1 can be omitted). Using Lemma 3, we can prove Observations 14, 15 and 16 as before without any modification. We now simply set $F'_0 = F_0$, $F'_1 = F_1$ and $F'_2 = F_2$ without running Phase III, as the fact that there are no pentagons in G implies that $S_1 = S_2 = \emptyset$. The statement of Lemma 5 can be directly seen to be true without any proof. Lemmas 6 and 7 can be proved as before; actually, the second paragraphs of both these proofs can be omitted as these cases only arise when $bc \in S_i$ for some $i \in \{1, 2\}$. It now follows as before that H_1 and H_2 form a 2-threshold cover of G.

For the case of split graphs (which are a special kind of paraglider-free graphs), we can additionally also skip Phase I of our algorithm. Suppose that $G = (X, Y, E)$ is a split graph. We start with an arbitrary ordering $<$ of the vertices of G, and once we get the valid 3-partition (F_0, F_1, F_2) after running Phase II of the algorithm, we can output $H_1 = (V(G), F_1 \cup F_0)$ and $H_2 = (V(G), F_2 \cup F_0)$ as the two threshold graphs that form a 2-threshold cover of G. We follow the same proof as the one for paraglider-free graphs, with the only change being made to the last paragraph of the proof of Observation 16, where Observation 2 is used (note that Observation 2 no longer holds as $<$ is not necessarily a Lex-BFS ordering). We replace this paragraph with the following:

Recall that $a_0b_0, a_1b_1, \ldots, a_kb_k$ is a path in G^*, such that for any $i \in \{0, 1, \ldots, k-$

$1\}$, $a_ib_{i+1} \in E(\overline{G})$ and $b_ia_{i+1} \in E(\overline{G})$. Let $i \in \{0, 1, \ldots, k-1\}$. If a_i and b_{i+1} both belong to one of X or Y, then it should be the case that $a_i, b_{i+1} \in Y$ (recall that X is a clique in G). Since $a_ib_i, a_{i+1}b_{i+1} \in E(G)$ and Y is an independent set in G, we then have $b_i, a_{i+1} \in X$. Since X is a clique, this contradicts the fact that $b_ia_{i+1} \in E(\overline{G})$. Therefore we can conclude that for each $i \in \{0, 1, \ldots, k-1\}$, one of a_i, b_{i+1} belongs to X and the other to Y. By the same argument, we can also show that for each $i \in \{0, 1, \ldots, k-1\}$, one of b_i, a_{i+1} belongs to X and the other to Y. Since k is odd, it now follows that one of $(a = a_0), b_k$ belongs to X and the other to Y, and similarly, one of $(b = b_0), a_k$ belongs to X and the other to Y. We can therefore conclude that $a \neq b_k$ and $b \neq a_k$. Recall that $a_kb_k < ab$, $a < d$, (c, b_k, a_k, u) is a strict F_2-switching path, and $a_iu, b_ic \in F_1$ (resp. $a_iu, b_ic \in F_2$) for each even i (resp. odd i). Then we have $a_ku, b_kc \in F_2$ and $au, bc \in F_1$, which implies that $a_k \neq a$ and $b_k \neq b$. We now have $\{a, b\} \cap \{a_k, b_k\} = \emptyset$, and therefore $\min\{a_k, b_k\} < \min\{a, b\}$. But then as $a < d$, we have $\{c, b_k, a_k, u\} < \{a, b, c, d\}$, which is a contradiction to the choice of (a, b, c, d).

Ibaraki and Peled [70] were the first to show that if G is a split graph, then G has a 2-threshold cover if and only if G^* is bipartite. Our proof, simplified as described above, yields a different proof for this fact which we believe is much simpler than the proofs in [70] or [102].

2.7.1 The chain subgraph cover problem

A bipartite graph $G = (A, B, E)$ is called a *chain graph* if for any pair of vertices u and v that belong to A, either $N(u) \subseteq N(v)$ or $N(v) \subseteq N(u)$. Equivalently, chain graphs are the class of bipartite graphs that does not contain a pair of edges whose end-points induce a $2K_2$ in G. A collection of chain graphs $\{H_1, H_2, \ldots, H_k\}$ is said to be a *k-chain subgraph cover* of a bipartite graph G if it is covered by $H_1, H_2, \ldots, H_k$. The problem of deciding whether a bipartite graph G can be covered by k chain graphs, i.e. whether G has a k-chain subgraph cover, is known as the *k-chain subgraph cover (k-CSC)* problem.

Yannakakis [118] credits Golumbic for the following observation. Given a bipartite graph $G = (A, B, E)$, let $\hat{G}$ be the split graph obtained from G by adding edges between every pair of vertices in one of the partite sets, say A: i.e. $V(\hat{G}) = V(G)$ and $E(\hat{G}) = E(G) \cup \{uv : u, v \in A\}$. Suppose that there exist vertices $a, c \in A$ and $b, d \in B$ such that the edges $ab, cd \in E(G)$ form a $2K_2$ in G. Then as $ba, ac, cd \in E(\hat{G})$ and $ad, bc, bd \notin E(\hat{G})$ the vertices a, b, c, and d induces a P_4 in $\hat{G}$. This implies that, if G is not a chain graph then $\hat{G}$ is not a threshold graph.

On the other hand, we know that the split graph $\hat{G}$ does not contain $2K_2$ or C_4 (as they are forbidden for split graphs). Suppose that the vertices x, y, u, and v induces a P_4 in $\hat{G}$ with edges say, $xy, yu, uv \in E(\hat{G})$ and $xu, yv, xv \notin E(\hat{G})$. Then, as $\hat{G}[A]$ is a complete graph and B is an independent set, it should be the case that both the edges xy, uv has one end-point in A and other end-point in B. This implies that $x, v \in B$ and $y, u \in A$, and therefore the edges $xy, uv \in E(G)$ form a $2K_2$ in G. Thus if $\hat{G}$ is not a threshold graph then G is not a chain graph. Therefore we can conclude that G is a chain graph if and only if $\hat{G}$ is a threshold graph. Further we can see that G has a k-chain subgraph cover if and only if the split graph $\hat{G}$ has a k-threshold cover. This implies that the k-chain subgraph cover problem for bipartite graphs can be reduced to the k-threshold cover problem for split graphs in polynomial time. In fact, the reverse reduction is also possible in polynomial time. Let G be a split graph with vertex partitions A and B, where A is a clique and B is an independent set. By removing all the edges between the vertices in A, we obtain a bipartite graph $G' = (A, B, E)$. As in the former case, it can be easily verified that G has a k-threshold cover if and only if the bipartite graph G' has a k-chain subgraph cover.

Thus we have the following theorem.

Theorem 6 ([118]). *For $k \geq 1$, the following problems are reducible to each other in polynomial-time.*

(*a*) *Recognizing whether a bipartite graph can be covered by k chain graphs.*

(*b*) *Recognizing whether a split graph can be covered by k threshold graphs.*

Yannakakis [118] showed that k-CSC is NP-complete for each fixed $k \geq 3$, which implies by the above theorem that the problem of deciding whether $\mathrm{th}(G) \leq k$ for an input graph G is also NP-complete for each fixed $k \geq 3$. He also pointed out that using the results of Ibaraki and Peled [70], the 2-CSC problem can be solved in polynomial time (since by the above theorem, 2-CSC can be reduced to the problem of determining whether a split graph can be covered by two threshold graphs). Thus our algorithm for split graphs described in Section 2.7 can also be used to compute a 2-chain subgraph cover, if one exists, for an input bipartite graph G in time $O(|E(G)|^2)$ (note that even though $|E(\hat{G})| > |E(G)|$, the vertices in $\hat{G}^*$ corresponding to the edges in $E(\hat{G}) \setminus E(G)$ are all isolated vertices and hence can be ignored while computing the partial 2-coloring of $\hat{G}^*$). Note that Ma and Spinrad [84] proposed a more involved $O(|V(G)|^2)$ algorithm for the problem. However, our algorithm for split graphs, and hence the algorithm for computing a 2-chain subgraph cover that it yields, is considerably simpler to implement than the algorithms of [70, 84, 102, 112].

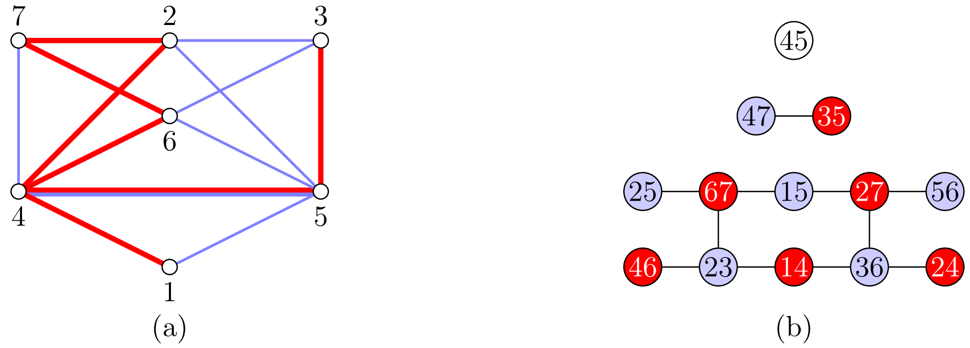
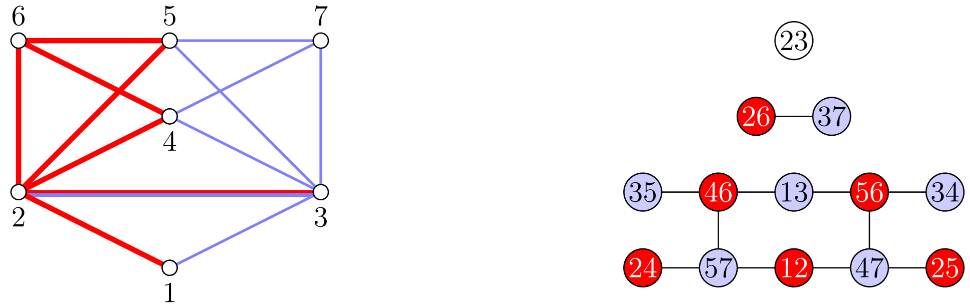


Figure 2.14: (a) the graph G from Figure 2.3, with its vertices numbered according to a non-Lex-BFS ordering, and (b) the graph G^* and its partial 2-coloring after Phases II and III.

2.8 Significance of Lex-BFS ordering

Would running just Phases II and III of our algorithm always produce a valid 2-threshold cover of G for any graph G? That is, could we have started with an arbitrary ordering of the vertices of G instead of a Lex-BFS ordering? We show that the algorithm may fail to produce a 2-threshold cover of the graph G shown in Figure 2.3 if the algorithm starts by taking an arbitrary ordering of vertices in Phase I. Suppose that the vertices of the graph are ordered according to their labels as shown in Figure 2.14(a). Clearly, it is not a Lex-BFS ordering of the vertices, as since the vertex in the second position is not a neighbor of the vertex in the first position, it is not even a BFS ordering. The sets F_0', F_1', F_2' computed by our algorithm after Phases II and III will be as shown in Figure 2.14(b) – the vertices of G^* in the set F_1' are shown as dark red, the ones in F_2' as light blue and the ones in F_0' as white. In Figure 2.14(a), the bold red edges form the graph H_1 and the thin blue edges form the graph H_2. Clearly, neither is a threshold graph (for example, both contain a C_4). On the other hand, Figure 2.15 shows the 2-threshold cover of G computed by our algorithm if it starts with the Lex-BFS ordering of the vertices of G as indicated by the labels of the vertices in Figure 2.15(a). Note that starting with a BFS ordering instead of a Lex-BFS ordering will also not work, since we can always add a universal vertex to the graph G shown in Figure 2.14(a) and number it 0, so that the vertex ordering is now a BFS ordering. It is not difficult to see that the graphs H_1 and H_2 computed in this case also fail to be threshold graphs (in fact, the edges incident on the vertex labeled 0 are all isolated vertices in the auxiliary graph, and none of them belong to any pentagons; hence they all belong to F_0', and the sets F_1' and F_2' will be exactly the same as before).

Thus the graph G shown in Figure 2.3 demonstrates that even though Phase I is optional for split graphs, for general graphs, our algorithm may not produce a 2-threshold cover of the



(a) (b)

Figure 2.15: (a) the graph G from Figure 2.3, with its vertices numbered according to a Lex-BFS ordering, and (b) the graph G^* and its partial 2-coloring after Phases II and III.

input graph if Phase I is skipped. Note that the graph G is not a paraglider-free graph. We have not found an example of a paraglider-free graph for which our algorithm will fail if Phase I is skipped. We believe that our result demonstrates once again the power of the lexicographic method in yielding elegant proofs for certain kinds of problems that otherwise seem to need more complicated proofs. Further research could establish the applicability of the method to a wider range of problems.

Chapter 3

Bigraphs and Digraphs

In this chapter, we review some connections established in the literature between some special subclasses of bipartite graphs and digraphs. In particular, we are interested in the directed analogue of chain cover problem and the close relationship between the class of interval bigraphs and the class of interval digraphs. We also discuss a few characterizations that are known in the literature for both of these classes of graphs.

3.1 Chain graphs and Ferrers digraphs

In this section, we revisit the class of chain graphs and chain cover problem and review some results concerning the directed analogue of chain cover problem. Here we also study a method by which the classes of bigraphs and digraphs can be transformed to each other.

Recall that, *chain graphs* are precisely the class of bipartite graphs for which the neighborhoods of the vertices in either partite set of G have a linear order with respect to inclusion. Equivalently, they are $2K_2$-free bipartite graphs. Since the bipartite complement of a $2K_2$ is also a $2K_2$, we can observe that the class of chain graphs are closed under taking bipartite complements. Analogous to the threshold cover problem for general undirected graphs, we came across the *chain cover problem* for bipartite graphs: the k-CSC problem asks whether an input bipartite graph G can be covered by k chain graphs.

3.1.1 Ferrers digraphs

The class of digraphs called Ferrers digraphs were independently introduced by Guttman [60] and Riguet [107]. A digraph G is said to be a *Ferrers digraph* if the out-neighborhoods (or in-

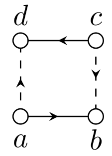


Figure 3.1: Alternating 4-anticircuit

neighborhoods) of vertices in G have a linear order with respect to inclusion. It is evident from this definition that the class of Ferrers digraphs is very closely related to the classes of threshold graphs and chain graphs. A directed notion of alternating 4-cycle called *alternating 4-anticircuit* can be defined as follows. Given a digraph G, two edges $(a, b), (c, d)$ in G are said to form an *alternating 4-anticircuit* in G if $(a, d), (c, b) \notin E(G)$ (refer to Figure 3.1). Thus equivalently, a Ferrers digraph can be defined as follows:

Definition 20 (Ferrers digraphs)**.** *A digraph G is a Ferrers digraph if and only if no pair of edges in G form an alternating 4-anticircuit.*

From the above definitions it is not difficult to see that the Ferrers digraphs can also be characterized to be the digraphs whose adjacency matrix does not contain a 2×2 permutation matrix, i.e. one of the matrices $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, as a submatrix.

Given a digraph G, we say that G has a *Ferrers cover of size k* if the input digraph G can be covered by k Ferrers digraphs.

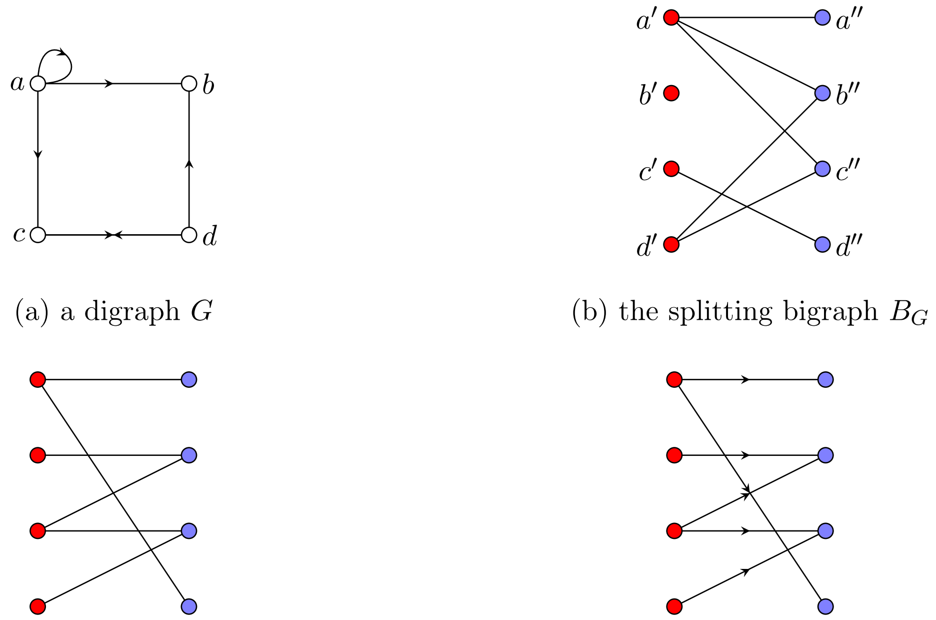
Definition 21 (Ferrers cover problem)**.** *The k-Ferrers cover problem asks whether an input digraph G can be covered by k Ferrers digraphs.*

As in the case of threshold cover and chain cover problems, since each edge of a digraph is trivially a Ferrers digraph, it is easy to see that any digraph G can be covered by using $|E(G)|$ Ferrers digraphs.

3.1.2 Transformations between bigraphs and digraphs

Given a digraph $G = (V, E)$, we can define the *splitting bigraph* of G as follows. This concept appears in the work of Müller [95].

Definition 22 (Splitting bigraph)**.** *Let G be a digraph. The splitting bigraph B_G of G is defined as the bipartite graph with two partite sets, $V' = \{u' : u \in V(G)\}$ and $V'' = \{u'' : u \in V(G)\}$, and $E(B_G) = \{u'v'' : (u, v) \in E(G)\}$.* (Refer to (a) and (b) of Figure: 3.2 for an example)



(a) a digraph G (b) the splitting bigraph B_G

(c) a bipartite graph G (d) the digraph $\overrightarrow{G}$

Figure 3.2: Transformations from a digraph to a bipartite graph and vice-versa

Note that for a digraph G, two edges $(a, b), (c, d) \in E(G)$ form an alternating 4-anticircuit in G if and only if the edges $a'b'', c'd'' \in E(B_G)$ form a $2K_2$ in B_G. This implies that G is a Ferrers digraph if and only if B_G is a chain graph.

Connection between threshold cover, chain cover, and Ferrers cover: It is easy to verify that a digraph G has a Ferrers cover of size k if and only if the splitting bigraph B_G has a chain cover of size k. This implies that the k-Ferrers cover problem for digraphs can be reduced to the k-chain subgraph cover problem for bipartite graphs in polynomial time (since given any digraph G, the splitting bigraph B_G can be constructed in polynomial time). On the other hand, given a bipartite graph $G = (A, B, E)$, consider the digraph $\overrightarrow{G}$ obtained from G by orienting all the edges from A to B. Then we have that, for vertices $a, c \in A$ and $b, d \in B$, two edges $ab, cd \in E(G)$ form a $2K_2$ in G if and only if the edges $(a, b), (c, d) \in E(\overrightarrow{G})$ form an alternating 4-anticircuit in $\overrightarrow{G}$. This implies that G is a chain graph if and only if $\overrightarrow{G}$ is a Ferrers digraph. Further this implies that the k-chain subgraph cover problem for bipartite graphs can be reduced to the k-Ferrers cover problem for digraphs in polynomial time. Figure 3.2 provides an example that illustrates the above transformations – a digraph transformed to a bipartite graph and vice-versa.

The above observations together with Theorem 6, gives the following theorem.

Theorem 7 ([118, 32]). *For $k \geq 1$, the following problems are reducible to each other in polynomial time.*

(a) *Recognizing whether a digraph can be covered by k Ferrers digraphs.*

(b) *Recognizing whether a bipartite graph can be covered by k chain graphs.*

(c) *Recognizing whether a split graph can be covered by k threshold graphs.*

Recall that, the k-CSC problem is NP-complete for each fixed $k \geq 3$ where as it is polynomial-time solvable for $k \leq 2$. Thus, above theorem implies that k-Ferrers cover problem is also NP-complete for each fixed $k \geq 3$ and it is polynomial-time solvable for $k \leq 2$.

Graph partitioning problem: We define a variant of cover called partition as follows.

Definition 23 (Graph partition). *A graph G is said to be partitioned into k graphs, $H_1, H_2, \ldots, H_k$ if for each $i \in \{1, 2, \ldots, k\}$ we have $V(H_i) = V(G)$ and $E(G) = \bigcup_{1 \leq i \leq k}, E(H_i)$, where for any $i, j \in \{1, 2, \ldots, k\}$ such that $i \neq j$, $E(H_i) \cap E(H_j) = \emptyset$; i.e. G can be covered by k pair-wise (edge) disjoint subgraphs.*

As in the case of graph covering problems, it is interesting to study the problem of partitioning an input graph into graphs that belong to special graph classes.

In fact, using the same transformations that we have discussed above and in Theorem 6, we have the following corollary of Theorem 7.

Corollary 1. *For $k \geq 1$, the following problems are reducible to each other in polynomial time.*

(a) *Recognizing whether a digraph can be partitioned into k Ferrers digraphs.*

(b) *Recognizing whether a bipartite graph can be partitioned into k chain graphs.*

(c) *Recognizing whether a split graph can be partitioned into k threshold graphs.*

We will see the complexity status of the above problems later in this chapter.

3.2 Interval bigraphs and interval digraphs

First let us define one of the well-known classes of undirected graphs called intersection graphs.

Definition 24 (Intersection graph)**.** *Let $\mathcal{F}$ be a family of subsets of a universal set U. Given an undirected graph G, a collection $\{S_u\}_{u\in V(G)}$ where $S_u \in \mathcal{F}$ is said to be an intersection representation of G with respect to $\mathcal{F}$ if for any $u, v \in V(G)$ we have $uv \in E(G)$ if and only if $S_u \cap S_v \neq \emptyset$. A graph that has an intersection representation with respect to $\mathcal{F}$ is called an intersection graph of sets from $\mathcal{F}$.*

It was Marczewski [113] who first noted that any undirected graph G is an intersection graph. This can be seen by defining $\mathcal{F} = \{S_v : v \in V\}$, where S_v denotes the set of all edges incident to v in G. Note that for different choices of geometric objects as the universal set, we obtain different graph classes. Graphs like interval graphs (intersection graphs of intervals on the real line), circular-arc graphs (intersection graph of arcs on a circle), rectangle graphs (intersection graphs of rectangles in $\mathbb{R}^2$), string graphs (intersection graphs of curves on the plane) etc... are a few of them [111]. McKee and McMorris [90] have presented a detailed overview of intersection graphs and their several variants. The class of interval graphs is one of the most prominent variants of intersection graphs.

Intersection bigraphs and interval bigraphs: A bipartite analogue of intersection graphs, called *intersection bigraph* was introduced by Harary, Kabell, and McMorris [64] and can be defined as follows:

Definition 25 (Intersection bigraph)**.** *Let $\mathcal{F}$ be a family of subsets of a universal set U. Given a bipartite graph $G = (A, B, E)$, a collection $\{S_u\}_{u\in V(G)}$, where $S_u \in \mathcal{F}$ is said to be an intersection representation of G w.r.t. $\mathcal{F}$ if $ab \in E(G)$ if and only if $a \in A$, $b \in B$ and $S_a \cap S_b \neq \emptyset$. A bipartite graph that has such an intersection representation with respect to $\mathcal{F}$ is called an intersection bigraph of sets from $\mathcal{F}$.*

In particular, the class of bipartite graphs called ***interval bigraphs*** are precisely the intersection bigraphs of a family of intervals on the real line: i.e. a bipartite graph $G = (A, B, E)$ is said to be an *interval bigraph* if there exists a collection $\{I_u\}_{u\in V(G)}$, of closed intervals on the real line such that $ab \in E(G)$ if and only if $a \in A$, $b \in B$ and $I_a \cap I_b \neq \emptyset$ (refer to Figure 1.5 for an example).

Intersection digraphs and interval digraphs The class of intersection digraphs can be considered as a directed analogue of intersection graphs. Note that in an intersection representation of a digraph, along with the connection between a pair of vertices we also have to incorporate the direction on the edge connecting them. This gives us an indication that, unlike in the case of undirected graphs, a collection of single sets may not be sufficient for an intersection representation for digraphs. The study of intersection digraphs was initiated by Maehara [85]. He introduced the notion called *pointed-set*, which is defined as a set S together with a base point $b \in S$ and is denoted by the pair (S, b). He then defined a family of digraphs called *catch digraphs*, which are exactly the digraphs G for which there exist a family $\{(S_u, b_u)\}_{u \in V(G)}$ of pointed-sets such that $(u, v) \in E(G)$ if and only if $b_v \in S_u$. He proved that every digraph can be represented as a catch digraph of pointed convex sets in $\mathbb{R}^2$ as follows: Let G be a digraph. For each vertex u in G, let b_u be a distinct point on the circumference of a circle in $\mathbb{R}^2$. Now by defining S_u to be the convex hull of the points $\{b_u\} \cup \{b_v : (u, v) \in E(G)\}$, it can be seen that the catch digraph of the pointed convex sets $\{(S_u, b_u)\}_{u \in V(G)}$ is isomorphic to G.

Later Das, Roy, Sen and West [32] generalized the concept of catch digraphs to intersection digraphs by replacing the pointed set (S, b) with a pair of sets (S, T).

Definition 26 (Intersection digraph)**.** *Let $\mathcal{F}$ be a family of subsets of a universal set U. Given a digraph G, a collection $\{(S_u, T_u)\}_{u \in V(G)}$ of $\mathcal{F} \times \mathcal{F}$ is said to be an intersection representation of G w.r.t. $\mathcal{F}$ if for any $u, v \in V(G)$, we have $(u, v) \in E(G)$ if and only if $S_u \cap T_v \neq \emptyset$. A digraph that has an intersection representation with respect to $\mathcal{F}$ is called an intersection digraph of sets from $\mathcal{F} \times \mathcal{F}$.*

Similar to the observation made for the class of intersection graphs, it can be shown that any digraph is an intersection digraph – given a digraph G, define for each $u \in V(G)$, $S_u = \{(u, v) \in E(G) : v \in V(G)\}$ and $T_u = \{(v, u) \in E(G) : v \in V(G)\}$. Then it can be verified that $\{(S_u, T_u)\}_{u \in V(G)}$ is a valid intersection representation of G. As in the case of undirected graphs, there are several variants for intersection digraphs, of which ***interval digraphs*** is the most popular one. Interval digraphs are exactly the intersection digraphs of intervals on a real line. Das, Roy, Sen and West [32] defined interval digraphs as follows: a digraph G is said to be an *interval digraph* if there exists a collection $\{(S_u, T_u)\}_{u \in V(G)}$, of pairs of closed intervals on a real line such that for any $u, v \in V(G)$, we have $(u, v) \in E(G)$ if and only if $S_u \cap T_v \neq \emptyset$. The collection $\{(S_u, T_u)\}_{u \in V(G)}$ is called the interval representation of G (refer to Figure 1.6 for an example).

Müller [95] has noted the following connection between interval digraphs and interval bigraphs, using transformations similar to the ones we encountered in Section 3.1.2.

Proposition 1 ([95])**.** *The following statements are true.*

(a) *A digraph G is an interval digraph if and only if the splitting bigraph B_G is an interval bigraph.*

(b) *A bipartite graph $G = (A, B, E)$ is an interval bigraph if and only if the digraph $\overrightarrow{G}$ obtained from G by orienting all the edges from one partite sets to other is an interval digraph.*

Moreover, the recognition problems for the classes of interval digraphs and interval bigraphs are reducible to each other in linear time.

Proof. (a) Suppose that $G = (V, E)$ is an interval digraph with an interval representation $\{(S_u, T_u)\}_{u\in V(G)}$. Consider the splitting bigraph $B_G = (V', V'', E)$ of G where $V' = \{u' : u \in V(G)\}$ and $V'' = \{u'' : u \in V(G)\}$, and $E(B_G) = \{u'v'' : (u, v) \in E(G)\}$. Define $I_{u'} = S_u$ for each $u' \in V'$ and $I_{u''} = T_u$ for each $u'' \in V''$. Then it can be seen that $\{I_{u'} : u' \in V'\} \cup \{I_{u''} : u'' \in V''\}$ is an interval representation of B_G, implying that B_G is an interval bigraph. On the other hand, suppose that B_G is an interval bigraph with an interval representation $\{I_{u'} : u' \in V'\} \cup \{I_{u''} : u'' \in V''\}$. Define $S_u = I_{u'}$ and $T_u = I_{u''}$. It can be then verified that G is an interval digraph with an interval representation $\{(S_u, T_u)\}_{u\in V(G)}$.

(b) Suppose that $G = (A, B, E)$ is an interval bigraph with an interval representation $\{I_u\}_{u\in V(G)}$. Let $\overrightarrow{G}$ denote the digraph obtained from G by orienting all the edges from one partite sets to other, say A to B. For each $u \in V(\overrightarrow{G})$, define $S_u = I_u$, $T_u = \emptyset$, if $u \in A$ and $S_u = \emptyset$, $T_u = I_u$, if $u \in B$. Then it can be verified that $\{(S_u, T_u)\}_{u\in V(G)}$ is an interval representation of $\overrightarrow{G}$, which implies that $\overrightarrow{G}$ is an interval digraph. On the other hand, if the digraph $\overrightarrow{G}$ obtained from G by orienting all the edges from A to B is an interval digraph, with an interval representation $\{(S_u, T_u)\}_{u\in V(\overrightarrow{G})}$, then by defining $I_u = S_u$ for each $u \in A$ and $I_u = T_u$ for each $u \in B$ we obtain an interval representation of G, implying that G is an interval bigraph. □

Now let us review some of the important structural results in the literature, concerning the classes of interval digraphs and interval bigraphs.

The first characterization for the class of interval digraphs was proposed by Das, Roy, Sen and West [32]. They proved the following theorem.

Theorem 8 ([32])**.** *Let G be a digraph. Then the following conditions are equivalent.*

(a) *G is an interval digraph.*

(b) *The rows and columns of the adjacency matrix of G can be independently permuted to obtain a matrix in which each 0 can be replaced by either R or C in such a way that every R has only R's to its right and every C has only C's below it.*

(*c*) $\overline{G}$ *can be partitioned into two Ferrers digraphs. i.e. there exist Ferrers digraphs* H_1 *and* H_2 *such that* $E(\overline{G}) = E(H_1) \cup E(H_2)$ *and* $E(H_1) \cap E(H_2) = \emptyset$.

Later Sanyal and Sen [110] used a notion called *consistent ordering* of edges to give an edge ordering characterization for interval digraphs. Given a digraph $G = (V, E)$, a linear ordering $<$ of $E(G)$ is said to be a consistent ordering if, for $(p, q), (t, s), (p, u), (t, q) \in E(G)$ we have: $(p, q) < (t, s) < (p, u)$ implies that $(p, s) \in E(G)$ $(q \neq u)$ and $(p, q) < (t, s) < (t, q)$ implies that $(t, s) \in E(G)$ $(p \neq t)$. They proved that G is an interval digraph if and only if $E(G)$ has a consistent ordering.

The class of interval bigraphs are also studied by several authors. First we note the following characterization for interval bigraphs.

Theorem 9 ([32]). *An undirected bipartite graph* G *is an interval bigraph if and only if the bipartite complement* $\overline{G^b}$ *of* G *can be partitioned into two chain graphs.*

Proof. Let G be an undirected bipartite graph. Note that by a transformation that we have used to prove Corollary 1, we have that the bipartite complement $\overline{G^b}$ of G can be partitioned into two chain graphs if and only if the digraph, $H = \overrightarrow{\overline{G^b}}$ (obtained from $\overline{G^b}$ by orienting all the edges, say from A to B) can be partitioned to two Ferrers digraphs. As the bipartite complement $\overline{G^b}$ of G is also a bipartite graph, by the definition of H we have that $V(H)$ also has the same partite sets A and B and all the edges in H are directed from A to B. Let H' be the digraph obtained by adding all possible symmetric arcs between the vertices belonging to the same partition of $V(H)$. I.e. $V(H') = V(H)$ and $E(H') = E(H) \cup \{(a, a'), (a', a) : a, a' \in A\} \cup \{(b, b'), (b', b) : b, b' \in B\}$. Then we can prove that the digraph H has a partition into two Ferrers digraphs if and only if the digraph H' has a partition into two Ferrers digraphs. In fact, it can be shown as follows: If there exist two Ferrers digraphs H_1 and H_2 such that $E(H) = E(H_1) \cup E(H_2)$ and $E(H_1) \cap E(H_2) = \emptyset$ then define the digraphs, H'_1 and H'_2 with $V(H'_1) = V(H'_2) = V(H') = V(H)$, $E(H'_1) = E(H_1) \cup \{(a, a'), (a', a) : a, a' \in A\}$ and $E(H'_2) = E(H_2) \cup \{(b, b'), (b', b) : b, b' \in B\}$. Note that the edges which belong to the same partite sets are added to the subgraphs H'_1 and H'_2 in such a way that, these edges cannot be a part of any alternating 4-anticircuit in H'_1 or H'_2. Now the fact that H_1 and H_2 are Ferrers digraphs gives us that H'_1 and H'_2 are also both Ferrers digraphs. Clearly H'_1 and H'_2 form a partition of H' as well. On the other hand, if there exist two Ferrers digraphs H'_1 and H'_2 such that $E(H') = E(H'_1) \cup E(H'_2)$ and $E(H'_1) \cap E(H'_2) = \emptyset$ then define the digraphs, H_1 and H_2 with $V(H_1) = V(H_2) = V(H) = V(H')$, $E(H_1) = E(H'_1) \cap E(H)$ and $E(H_2) = E(H'_2) \cap E(H)$. Since the edges in H_1 and H_2 are all oriented from A to B, by the definitions of $E(H_1)$ and $E(H_2)$, we can observe that any pair of edges that forms an alternating 4-anticircuit in H_1 (resp. H_2) remain as an alternating 4-anticircuit in H_1' (resp. H_2'). Therefore the fact that H_1' and H_2' are Ferrers digraphs imply that both the digraphs H_1 and H_2 are Ferrers digraphs that form a partition of H. Applying Theorem 8(c), we then have that the digraph H' can be partitioned into two Ferrers digraphs if and only if the complement $\overline{H'}$ of H' is an interval digraph. Recall that the digraph $H = \overrightarrow{\overline{G^b}}$ is the digraph obtained from $\overline{G^b}$ by orienting all its edges from A to B and H' is the digraph obtained from H by adding all possible symmetric arcs between the vertices in the same partite sets. We then have that the complement $\overline{H'}$ of H' is exactly the digraph, say $\overleftarrow{G}$ whose underlying undirected graph is G and all its edges are oriented from the partite set B to A. Therefore by Proposition 1, we have that the digraph $\overleftarrow{G} = \overline{H'}$ is an interval digraph if and only if G is an interval bigraph. Combining with the previous observations, we then have the theorem. $\square$

Note that, the above theorem tells us that the recognition problem of interval bigraphs can be reduced to the *2-chain partition problem* – whether there exists two chain graphs that form an (edge) partition of the input graph

Now the following characterization for interval bigraphs proposed by Hell and Huang [69] connects interval bigraphs to a subclass of circular-arc graphs.

Theorem 10 ([69]). *A bipartite graph G is an interval bigraph if and only if the complement $\overline{G}$ of G is a circular-arc graph with clique number two (vertices can be partitioned into two cliques) that has a circular-arc representation in which no two arcs together cover the whole circle.*

Hell and Huang [69] also proposed the following two vertex ordering characterizations for the class of interval bigraphs.

Theorem 11 ([69]). *Let G be a bipartite graph with partite sets X and Y. Then the following statements are equivalent.*

(a) *G is an interval bigraph.*

(b) *The vertex set of G has an ordering $<$ such that for any $u, v, w \in V(G) = X \cup Y$ with $u < v < w$, the configuration in Figure 3.3 is forbidden (dark red vertices are in X and light blue vertices are in Y, or conversely).*

(c) *The vertex set of G has an ordering $<$ such that for any $u, v, w, x \in V(G) = X \cup Y$ with $u < v < w < x$, the configurations in Figure 3.4 is forbidden (dark red vertices are in X and light blue vertices are in Y, or conversely).*

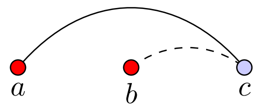


Figure 3.3: Forbidden configuration

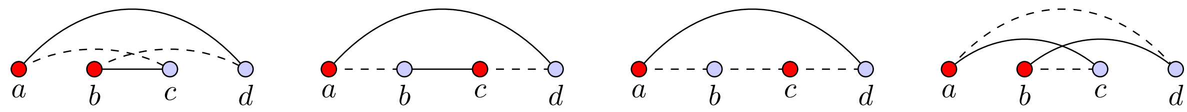


Figure 3.4: Forbidden configurations

Unfortunately, neither of the above characterizations directly yields a polynomial-time recognition algorithm for the class of interval bigraphs (resp. interval digraphs). The only known polynomial-time algorithm known for recognizing the class of interval bigraphs (resp. interval digraphs) is the one proposed by Müller [95], which can be implemented in $O(n.m^6(n+m)\log n)$ time. He used a dynamic programming approach for the same. He also notes that, any interval bigraph is chordal bipartite, and this observation is very crucial in the development of this algorithm. Applying Theorem 9 and Theorem 10, Mülller's algorithm can also be used to solve the recognition problems for the class of 2-chain partitionable graphs and the class of circular-arc graphs with clique number two and having a circular-arc representation in which no two arcs cover the whole circle. By Corollary 1, this algorithm can be further used to solve the recognition problems for the class of digraphs that can be partitioned into two Ferrers digraphs and the split graphs that can be partitioned into two threshold graphs.

However the problem of finding a simpler recognition algorithm for the class of interval bigraphs or interval digraphs is a long standing open problem. In the same paper [95], Müller also provides some structures that are necessarily forbidden in an interval bigraph, and conjectured that they are sufficient as well. But later Hell and Huang [69] disproved this conjecture. Thus the question of a forbidden structure characterization for interval bigraphs or interval digraphs still remains as an interesting open problem. Moreover, the complexities of all the recognition problems listed in Corollary 1 also remain open for $k > 2$.

Part II

On the Kernel and Related Problems in Interval Digraphs

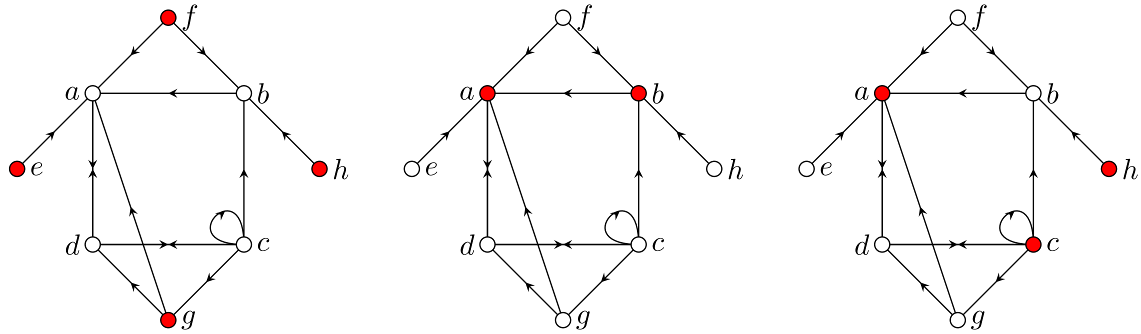
Chapter 4

Algorithms for Reflexive Interval Digraphs

4.1 Introduction

In this chapter, we study some of the classic computation problems on digraphs in a subclass of interval digraphs, called reflexive interval digraphs. Our basic goal here is to illustrate the fact that the reflexivity of an interval digraph has a huge impact on the algorithmic complexity of several problems related to domination and independent sets in digraphs. First let us recall the definitions of the undirected counterparts of those problems.

Let $H = (V, E)$ be an undirected graph. A set $S \subseteq V(H)$ is said to be an *independent set* in H if for any two vertices $u, v \in S$, $uv \notin E(H)$. A set $S \subseteq V(H)$ is said to be a *dominating set* in H if for any $v \in V(H) \setminus S$, there exists $u \in S$ such that $uv \in E(H)$. A set $S \subseteq V(H)$ is said to be an *independent dominating set* in H if S is dominating as well as independent. Note that any maximal independent set in H is an independent dominating set in H, and therefore every undirected graph contains an independent dominating set, which implies that the problem of deciding whether an input undirected graph contains an independent dominating set is trivial. On the other hand, finding an independent dominating set of *maximum cardinality* is NP-complete for general graphs, since independent dominating sets of maximum cardinality are exactly the independent sets of maximum cardinality in the graph. The problem of finding a minimum cardinality independent dominating set is also NP-complete for general graphs [49] and also in many special graph classes (refer to [81] for a survey). Now we study the directed analogues of these problems, which are also well-studied in the literature.



(a) an independent set (b) an absorbing set (c) a kernel

Figure 4.1: Examples: the vertices that constitute the corresponding sets are shown in dark red.

Some computational problems in digraphs: Let $G = (V, E)$ be a directed graph. A set $S \subseteq V(G)$ is said to be an *independent set* in G, if for any two vertices $u, v \in S$, $(u, v), (v, u) \notin E(G)$. A set $S \subseteq V(G)$ is said to be an *absorbing (resp. dominating) set* in G, if for any $v \in V(G) \setminus S$, there exists $u \in S$ such that $(v, u) \in E(G)$ (resp. $(u, v) \in E(G)$). As any set of vertices that consists of a single vertex is independent and the whole set $V(G)$ is absorbing as well as dominating, the interesting computational problems that arise here are that of finding a maximum independent set, called INDEPENDENT-SET, and that of finding a minimum absorbing (resp. dominating) set in G, called ABSORBING-SET (resp. DOMINATING-SET). A set $S \subseteq V(G)$ is said to be an *independent dominating (resp. absorbing) set* if S is both independent and dominating (resp. absorbing). Note that unlike undirected graphs, the problem of finding a maximum cardinality independent dominating (resp. absorbing) set is different from the problem of finding a maximum cardinality independent set for directed graphs. An independent absorbing set in a directed graph is more well-known as a *kernel* of the digraph. We follow the terminology in [100] and call an independent dominating set in a directed graph a *solution* of the graph. It is easy to see that a kernel in a directed graph G is a solution in the directed graph obtained by reversing every arc of G and vice versa. Note that unlike in the case of undirected graphs, a kernel need not always exist in a directed graph. For example, it is easy to see that a directed triangle with arcs say, $(a, b), (b, c)$ and (c, a) does not have a kernel. Therefore, besides the computational problems of finding a minimum or maximum sized kernel, called MIN-KERNEL and MAX-KERNEL respectively, the comparatively easier problem of determining whether a given directed graph has a kernel in the first place, called KERNEL, is itself a non-trivial one. Figure 4.1 illustrates through an example the notions of independent sets, absorbing sets and kernels for digraphs.

Recall that given a digraph G, a collection $\{(S_u, T_u)\}_{u \in V(G)}$ of pairs of intervals is said to be an *interval representation* of G if $(u, v) \in E(G)$ if and only if $S_u \cap T_v \neq \emptyset$. A digraph G that has an interval representation is called an *interval digraph* [32]. We consider a loop to be present on a vertex u of an interval digraph if and only if $S_u \cap T_u \neq \emptyset$. An interval digraph is a *reflexive interval digraph* if there is a loop on every vertex.

By imposing restrictions on the source and destination intervals for each vertex in the representation of interval digraphs, we obtain several subclasses of interval digraphs.

4.1.1 Subclasses of interval digraphs

Many subclasses of interval digraphs have attracted the interest of researchers over the years. The authors of [32] studied a special subclass of interval digraphs called *interval point digraphs*, which are those interval digraphs G with an interval representation $\{(S_u, T_u)\}_{u \in V(G)}$ such that T_u is a degenerate interval, i.e. a point. When, in addition, the point T_u lies inside the interval S_u for each vertex $u \in V(G)$, the graph G is said to be an *interval catch digraph.* Even more restrictively, if the point T_u is the left end-point of the interval S_u for each vertex u, then the graph is said to be a *chronological interval digraph*, which were introduced and characterized in [31]. Note here that interval catch digraphs were defined and studied in the work of Maehara [85] that predates the introduction of interval digraphs (the term "interval digraph" was used with a different meaning in this work). Prisner [100] generalized interval catch digraphs to *interval nest digraphs* – they are the interval digraphs with an interval representation $\{(S_u, T_u)\}_{u \in V(G)}$ such that $T_u \subseteq S_u$ for each $u \in V(G)$. When the intervals S_u and T_u for each vertex $u \in V(G)$ are required to have a common left end-point, the interval digraphs that arise are called *adjusted interval digraphs*, which were introduced by Feder, Hell, Huang, and Rafiey [40]. The class of adjusted interval digraphs has an association to the list homomorphism problem.

For digraphs G and H, a function $f : V(G) \to V(H)$ is said to be a *homomorphism* from G to H if $(f(u), f(v)) \in E(H)$ whenever $(u, v) \in E(G)$. Given a list $L = \{L(v) \subseteq V(H) : v \in V(G)\}$, a *list homomorphism* (with respect to L) is a homomorphism f from G to H, with an additional property that $f(v) \in L(v)$ for each $v \in V(G)$. The *list homomorphism problem L-HOM(H)* asks whether an input digraph G with a list L, admits a list homomorphism to a target digraph H with respect to L. Feder, Hell, Huang, and Rafiey [40] showed that the list homomorphism problem for a target digraph H is polynomial-time solvable if H is an adjusted interval digraph and conjectured that if H is not an adjusted interval digraph, then the problem is NP-complete (see [40]). This conjecture if true, is analogous to the dichotomy result of list homomorphism

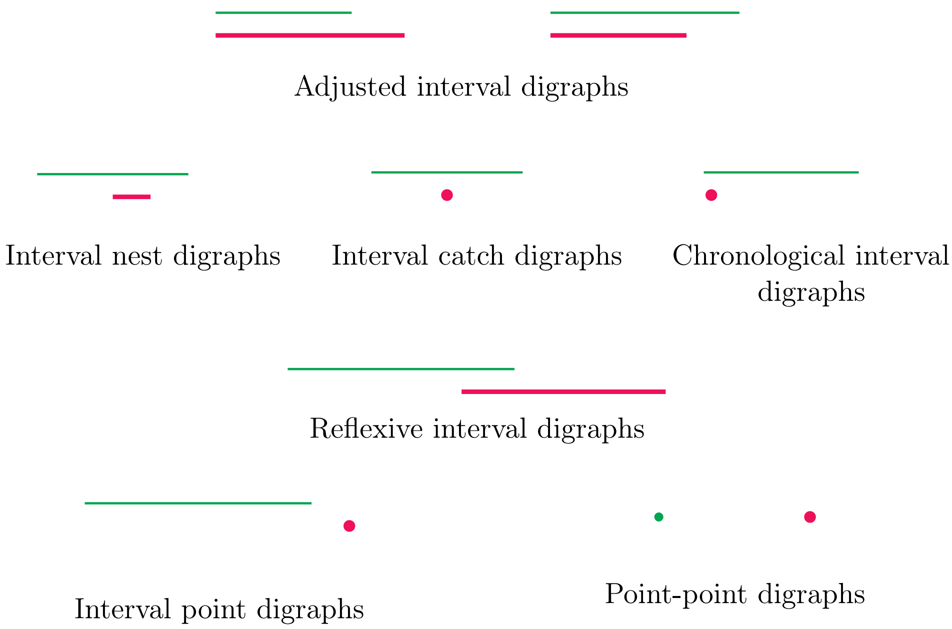


Figure 4.2: Subclasses of interval digraphs (source intervals are shown in thin green and destination intervals are shown in bold red)

problem for reflexive undirected graphs which states that, the list homomorphism problem for a target reflexive undirected graph H is polynomial-time solvable if H is an interval graph, and NP-complete otherwise.

We observe that several of these classes, like interval catch digraphs, interval nest digraphs, adjusted interval digraphs and chronological interval digraphs, are subclasses of the more general class of reflexive interval digraphs – which arise when we require that the two intervals assigned to a vertex have to intersect. On the other hand, interval point digraphs and even the restricted class of interval point digraphs called point-point digraphs, where the two intervals assigned to each vertex are required to be degenerate, (i.e. they consist of a single point each) need not be reflexive. Figure 4.2 depicts the properties of the intervals representing a vertex that belongs to the corresponding subclass of interval digraphs. Note that the interval representation of a vertex in an adjusted interval digraph can be either of two types as shown in the figure.

Tolerance digraphs The class of tolerance digraphs naturally evolve from undirected tolerance graphs. Note that when two vertices u and v of a tolerance graph are adjacent, then it should be the case that either $|I_u \cap I_v| \geq t_u$ or $|I_u \cap I_v| \geq t_v$ or both. This observation motivates Bogart and Trenk [10] to think of a directed analogue of tolerance graphs called tolerance digraphs which is defined as follows: A graph G is said to be a *tolerance digraph* if each vertex u in G can be assigned an interval I_u on a real line and a tolerance $t_u \in \mathbb{R}^+$ in such a way that for two vertices

u and v in G, $(u,v) \in E(G)$ if and only if $|I_u \cap I_v| \geq t_v$. In addition, if $t_u \leq |I_u|$ for each $u \in V(G)$, then we call G as a *bounded tolerance digraph.* Bogart and Trenk generalized the notion of bounded tolerance digraphs to *bounded bitolerance digraphs.* In bounded bitolerance digraphs, each vertex is assigned with a pair of tolerances instead of one.

Definition 27 (Totally bounded bitolerance digraphs)**.** *A digraph G is said to be a bounded bitolerance digraph if each vertex $u \in V(G)$ is assigned a real interval $I_u = [l(u), r(u)]$, a left tolerant point $lt(u) \in I_u$, and a right tolerant point $rt(u) \in I_u$ such that for any two vertices u and v, $(u,v) \in E(G)$ if and only if $I_u \cap I_v \nsubseteq [l(v), lt(v))$ and $I_u \cap I_v \nsubseteq (rt(v), r(v)]$. In addition, if $lt(u) \leq rt(u)$ for each vertex $u \in V(G)$ then G is called a totally bounded bitolerance digraph.*

For more on these aspects of tolerance in graphs, see Golumbic and Trenk [59].

For an interval $I = [x, y]$ of the real line (here $x, y \in \mathbb{R}$ and $x \leq y$), we denote by $l(I)$ the left end-point x of I and by $r(I)$ the right end-point y of I.

It is interesting to note that the class of totally bounded bitolerance digraphs coincides with the class of interval nest digraphs [10]. Suppose that G is a totally bounded bitolerance digraph with a collection of intervals $\{I_u\}_{u \in V(G)}$, left tolerant points $\{lt(u)\}_{u \in V(G)}$ and right tolerant points $\{rt(u)\}_{u \in V(G)}$. Note that by the definition of totally bounded bitolerance digraphs, we know that for each vertex $u \in V(G)$, we have $lt(u), rt(u) \in I_u$ and $lt(u) \leq rt(u)$. It is then easy to verify that, G has an interval nest representation $\{(S_u, T_u)\}_{u \in V(G)}$ by defining for each vertex $u \in V(G)$, $S_u = I_u$ and $T_u = [lt(u), rt(u)]$. On the other hand, suppose that G has an interval nest representation $\{(S_u, T_u)\}_{u \in V(G)}$. Thus for each vertex $u \in V(G)$, we have that $T_u \subseteq S_u$. Then for each vertex $u \in V(G)$, by defining the interval representing the vertex u as, $I_u = S_u$, the left tolerant point of the vertex u, $lt(u) = l(T_u)$ and the right tolerant point of the vertex u, $rt(u) = r(T_u)$, it can be proved that G is a totally bounded bitolerance digraph.

In this chapter, we see as our main contribution the identification of the class of reflexive interval digraphs as an important class of digraphs. In particular, we show that the problems Kernel, Min-Kernel, Max-Kernel, Absorbing-Set, Dominating-Set, and Independent-Set are efficiently solvable in the class of reflexive interval digraphs. The significance of reflexivity of interval digraphs will be more evident when we show in an upcoming chapter that, all these problems are NP-complete and/or APX-hard even on point-point digraphs. Along the way we obtain new characterizations of both these graph classes, which reveal some of the properties of these digraphs.

For a bipartite graph having two specified partite sets A and B, a set $S \subseteq B$ such that

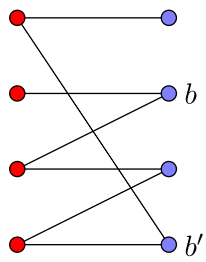


Figure 4.3: The set $\{b, b'\}$ forms an A-dominating set

$\bigcup_{u \in B} N(u) = A$ is called an A-*dominating set*. For example, for the bipartite graph given in Figure 4.3, $\{b, b'\}$ is an A-dominating set. Note that the bipartite graph $G = (A, B, E)$ does not contain an A-dominating set if and only if there are isolated vertices in A. The problem of finding an A-dominating set of minimum cardinality in a bipartite graph with partite sets A and B is more well-known as the Red-Blue Dominating Set problem, which was introduced for the first time in the context of the European railroad network [117] and plays an important role in the theory of fixed parameter tractable algorithms [37]. This problem is equivalent to the well known Set Cover and Hitting Set problems [49] and therefore, it is NP-complete for general bipartite graphs. The problem remains NP-complete even for planar bipartite graphs [2]. Recall from Chapter 3 that the class of *interval bigraphs* are closely related to the class of interval digraphs. These are undirected bipartite graphs with partite sets A and B such that there exists a collection of intervals $\{S_u\}_{u \in V(G)}$ such that $uv \in E(G)$ if and only if $u \in A$, $v \in B$, and $S_u \cap S_v \neq \emptyset$. We also study the Red-Blue Dominating Set problem in the class of interval bigraphs and we use this to solve the problem Absorbing-Set for reflexive interval digraphs.

4.1.2 Literature survey

The problems of computing a maximum independent set and minimum dominating set in undirected graphs are two classic optimization problems in graph theory. As we have noted before, the Independent-Set problem in a directed graph coincides with the problem of finding a maximum cardinality independent set of its underlying undirected graph. Also, in order to find a maximum independent set in an undirected graph, one could just orient the edges of the graph in an arbitrary fashion and solve the Independent-Set problem on the resulting digraph. Therefore, there is an easy reduction from the problem of computing a maximum independent set in undirected graphs to the Independent-Set problem on digraphs and vice versa, implying that these two problems have the same algorithmic complexity. On the other hand, it seems that the directed analogue of the domination problem is harder than the undirected version, since even though one can find a minimum dominating set in an undirected graph by replacing every edge with symmetric arcs and then using an algorithm for Dominating-Set on digraphs, a reduction in the other direction is not known. In particular, a minimum dominating set in the underlying undirected graph of a digraph need not even be a dominating set of the digraph. For example, any vertex of a complete graph is a dominating set of size 1, implying that the problem of finding a minimum cardinality dominating set in a complete graph is trivial, while no polynomial time algorithm is known to solve the Dominating-Set problem for the class of tournaments, which are precisely orientations of complete graphs. Domination in tournaments is well studied in the literature [91, 3, 20], but still the best known algorithm for Dominating-Set does not run in polynomial time [91, 103]. In [91], the authors give an $n^{O(\log n)}$ time algorithm for the Dominating-Set problem in tournaments and they also note that Sat can be solved in $2^{O(\sqrt{v})}n^K$ time (where v is the number of variables, n is the length of the formula and K is a constant) if and only if the Dominating-Set in a tournament can be solved in polynomial time. Thus, determining the algorithmic complexity of the Dominating-Set problem even in special classes of digraphs seems to be much more challenging than the algorithmic question of finding a minimum cardinality dominating set in undirected graphs. Even though bounds on the minimum dominating sets in digraphs have been obtained by several authors (see the book [66] for a survey), not much is known about the computational complexity of finding a minimum cardinality absorbing set (or dominating set) in directed graphs. Nevertheless, Kernel is a variant of Dominating-Set that has gained the attention of researchers over the years.

The term *kernel* of a digraph, is introduced by Von Neumann and Morgenstern [94] in the context of game theory. They showed that for digraphs associated with certain combinatorial games, the existence of a kernel implies the existence of a winning strategy. Most of the work related to domination in digraphs has been mainly focused on kernels. In fact, the Kernel problem was shown to be NP-complete in general digraphs by Chvátal [21]. Later, Fraenkel [44] proved that the Kernel problem remains NP-complete even for planar digraphs of degree at most 3 having in- and out-degrees at most 2. It can be easily seen that the Min-Kernel and Max-Kernel problems are NP-complete for those classes of graphs for which the Kernel problem is NP-complete. A digraph is said to be *kernel perfect* if every induced subgraph of it has a kernel. Several sufficient conditions for digraphs to be kernel perfect has been explored [106, 38, 94]. The Kernel problem is trivially solvable in polynomial-time on any kernel perfect family of digraphs. But the algorithmic complexity status of the problem of computing a kernel in a kernel perfect digraph also seems to be unknown [98]. Prisner [100] proved that interval nest digraphs and their reversals are kernel-perfect, and a kernel can be found in these graphs in time $O(n^2)$ if a representation of the graph is given. Note that the MIN-KERNEL problem can be shown to be NP-complete even in some kernel perfect families of digraphs that have a polynomial-time computable kernel (see Remark 4). Apart from game theory, the notion of kernel historically played an important role as an approach towards the proof of the celebrated 'Strong perfect graph conjecture' (now Strong Perfect Graph Theorem). A digraph G is called *normal* if every clique in G has a kernel (that is, every clique contains a vertex that is an out-neighbor of every other vertex of the clique). Berge and Duchet (see [12]) introduced a notion called *kernel-solvable* graphs, which are undirected graphs for which every normal orientation (symmetric arcs are allowed) of it has a kernel. They conjectured that kernel solvable graphs are exactly the perfect graphs. This conjecture was shown to be true for various special graph classes [9, 86, 87]. In general graphs, it was proved by Boros and Gurvich [12] that perfect graphs are kernel solvable and the converse direction follows from the Strong Perfect Graph Theorem. Kernels are also closely related to *Grundy functions* in digraphs (for a digraph $G = (V, E)$, a non-negative function $f : V \to \mathbb{N}_{>0}$ is called a Grundy function, if for each vertex $v \in V$, $f(v)$ is the smallest non-negative integer that does not belong to the set $\{f(u) : u \in N^+(v)\}$). Berge [7] showed that if a digraph has a Grundy function then it has a kernel. Even though the converse is not necessarily true for general digraphs, Berge [7] proved that every kernel-perfect graph has a Grundy function. It is known that almost every random digraph has a kernel [34]. Kernels, its variants and kernel-perfect graphs are topics that have been extensively studied in the literature, including in the works by Richardson [105], Galeana-Sánchez and Neumann-Lara [47], Berge and Duchet [8], and many more. See [13] for a detailed survey of results related to kernels.

Though every normal orientation of a perfect graph has a kernel, the question of finding a kernel has been noted as a challenging problem even in such digraphs. Polynomial-time algorithms for the KERNEL problem, that also compute a kernel in case one exists, have been obtained for some special graph classes. König (see [66]), who was one of the earliest to study domination in digraphs (he called an independent dominating set a 'basis of second kind'), proves that every minimal absorbing set of a transitive digraph is a kernel and every kernel in a transitive digraph has the same cardinality. Thus the KERNEL problem is trivial for transitive digraphs and there is a simple linear time algorithm for the MIN-KERNEL problem in such digraphs. The problem of computing a kernel, if one exists, can be solved in polynomial time for digraphs that do not contain odd directed cycles using Richardson's Theorem [104]. This implies that this problem is also polynomial-time solvable in directed acyclic graphs. Polynomial-time algorithms for finding a kernel, if one exists, is also known for digraphs that are normal orientations of permutation graphs [1], Meyniel orientations (an orientation D of G for which every triangle in D has at least two symmetric arcs) of comparability graphs [1], normal orientations (without symmetric arcs) of claw-free graphs [98], normal orientations of chordal graphs [98] and normal orientations of directed edge graphs (intersection graphs of directed paths in a directed tree) [98, 33]. For the class of normal orientations of line graphs of bipartite graphs, Maffray [87] observed that kernels in such graphs coincide with the stable matchings in the corresponding bipartite graphs. Thus in this graph class, a kernel can be computed in polynomial time using the famous Gale and Shapley algorithm [46] for stable matchings in bipartite graphs. It is shown in [98] that for any orientation (without symmetric arcs) of circular-arc graphs, KERNEL can be solved in polynomial time and a kernel, if one exists, can also be computed in polynomial time. The problem was also solved for the class of interval nest digraphs by Prisner [100].

4.1.3 Notation

For a closed interval $I = [x, y]$ of the real line (here $x, y \in \mathbb{R}$ and $x \leq y$), we denote by $l(I)$ the left end-point x of I and by $r(I)$ the right end-point y of I. We use the following observation throughout the paper: if I and J are two intervals, then $I \cap J = \emptyset \Leftrightarrow (r(I) < l(J)) \vee (r(J) < l(I))$. Given an interval representation of a graph, we can always perturb the end-points of the intervals slightly to obtain an interval representation of the same graph which has the property that no end-point of an interval coincides with any other end-point of an interval. We assume that every interval representation considered in this paper has this property.

Let $G = (V, E)$ be a directed graph. For $u, v \in V(G)$, we say that u is an *in-neighbor* (resp. *out-neighbor*) of v if $(u, v) \in E(G)$ (resp. $(v, u) \in E(G)$). For a vertex v in G, we denote by $N_G^+(v)$ and $N_G^-(v)$ the set of out-neighbors and the set of in-neighbors of the vertex v in G respectively. When the graph G under consideration is clear from the context, we abbreviate $N_G^+(v)$ and $N_G^-(v)$ to just $N^+(v)$ and $N^-(v)$ respectively. We denote by n the number of vertices in the digraph under consideration, and by m the number of edges in it not including any self-loops.

For $i, j \in \mathbb{N}$ such that $i \leq j$, let $[i, j]$ denote the set $\{i, i+1, \ldots, j\}$. Let G be a digraph with vertex set $[1, n]$. Then for $i, j \in [1, n]$, we define $N_{>j}^+(i) = N^+(i) \cap [j+1, n]$, $N_{>j}^-(i) = N^-(i) \cap [j+1, n]$, $N_{<j}^+(i) = N^+(i) \cap [1, j-1]$, and $N_{<j}^-(i) = N^-(i) \cap [1, j-1]$. We denote by $\overline{N_{>j}^+(i)}$ and $\overline{N_{>j}^-(i)}$ the sets $[j+1, n] \setminus N_{>j}^+(i)$ and $[j+1, n] \setminus N_{>j}^-(i)$ respectively.

4.2 Ordering characterization

We first show that a digraph is a reflexive interval digraph if and only if there is a linear ordering of its vertex set such that none of the structures shown in Figure 4.4 are present.

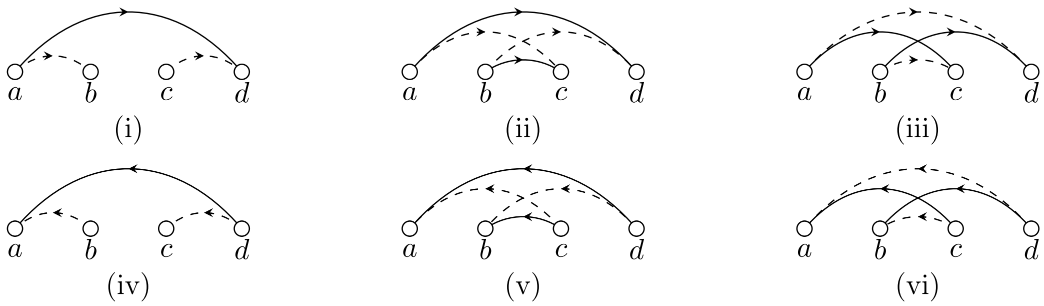


Figure 4.4: Forbidden structures for reflexive interval digraphs (possibly $b = c$ in (i), (ii), (iv) and (v)). Note that the vertices are assumed to have self-loops since a vertex without a self-loop is itself forbidden in a reflexive interval digraph.

Theorem 12. *A digraph G is a reflexive interval digraph if and only if $V(G)$ has an ordering $<$ in which for any $a, b, c, d \in V(G)$ such that $a < b < c < d$, none of the structures in Figure 4.4 occur (b and c can be the same vertex in (i), (ii), (iv), (v) of Figure 4.4).*

Proof. Let G be a reflexive interval digraph with an interval representation $\{(S_v, T_v) : v \in V(G)\}$. For any vertex $v \in V(G)$, let x_v be the left-most end-point of the interval $S_v \cap T_v$(which is well defined as G is a reflexive interval digraph). Let $<$ be an ordering of $V(G)$ with respect to the increasing order of the points x_v. Now we can verify that structures in Figure 4.4 are forbidden with respect to the order $<$.

Suppose not. Let $a < b < c < d$ be such that of Figure 4.4(i). $a < b$, $c < d$ and $(a, b), (c, d) \notin E(G)$ implies that $r(S_a) < l(T_b)$ and $r(S_c) < l(T_d)$. Since $b \leq c$, we also have that $l(T_b) \leq r(S_c)$. Combining these observations we then have that $r(S_a) < l(T_d)$, which further implies that $(a, d) \notin E(G)$, which is a contradiction to Figure 4.4(i). Let $a < b < c < d$ be such that of Figure 4.4(ii). Then $a < c$, $b < d$ and $(a, c), (b, d) \notin E(G)$ implies that $r(S_a) < l(T_c)$ and $r(S_b) < l(T_d)$. Since $(a, d) \in E(G)$, we also have that $l(T_d) < r(S_a)$. Combining these observations we then have, $r(S_b) < l(T_c)$ implying that $(b, c) \notin E(G)$, which is a contradiction to Figure 4.4(ii). Suppose that $a < b < c < d$ be such that of Figure 4.4(iii). Then $(a, c), (b, d) \in E(G)$ implies that $l(T_c) < r(S_a)$ and $l(T_d) < r(S_b)$. Since $a < d$, $(a, d) \notin E(G)$, we also have that $r(S_a) < l(T_d)$. Combining these observations we then have, $l(T_c) < r(S_b)$ implying that $(b, c) \in E(G)$, which is a contradiction to Figure 4.4(iii). Since we arrive at a contradiction in every case, we can conclude that none of the structures in Figures 4.4(i), (ii) or (iii) can be present. Similarly, by interchanging the roles of source and destination intervals in the above proof, we can also prove that none of the structures in Figures 4.4(iv), (v) or (vi) can be present with respect to the ordering $<$.

Conversely, assume that $<$ is an ordering of $V(G)$ for which the structures in Figure 4.4 are absent. Let $n = |V(G)|$. We can assume that $V(G) = [1, n]$ and that $<$ is the ordering $(1, 2, \ldots, n)$. First, we note the following observation.

Observation 18. *For any two vertices i, j such that $i < j$, we have the following:*

(a) *either $N^+_{>j}(i) \subseteq N^+_{>j}(j)$ or $N^+_{>j}(j) \subseteq N^+_{>j}(i)$ and*

(b) *either $N^-_{>j}(i) \subseteq N^-_{>j}(j)$ or $N^-_{>j}(j) \subseteq N^-_{>j}(i)$.*

Proof. Suppose not. Due to the symmetry between (a) and (b), we prove only the case where (a) is not true. Then there exists two distinct vertices $x_i, x_j \in \{j + 1, \ldots, n\}$ such that $x_i \in N^+_{>j}(i) \setminus N^+_{>j}(j)$ and $x_j \in N^+_{>j}(j) \setminus N^+_{>j}(i)$. Now if $x_i < x_j$, then the vertices $i < j < x_i < x_j$ form Figure 4.4(iii) which is forbidden and if $x_j < x_i$, then the vertices $i < j < x_j < x_i$ form Figure 4.4(ii) which is also forbidden. As we have a contradiction in both the cases, we are done. □

We now define for each $i \in \{1, 2, \ldots, n\}$, a pair of intervals (S_i, T_i) as follows. For each $i \in \{1, 2, \ldots, n\}$, let

$$y_i = \begin{cases} \min \overline{N^+_{>i}(i)}, & \text{if } \overline{N^+_{>i}(i)} \neq \emptyset \\ n + 1, & \text{otherwise} \end{cases} \qquad \text{and} \qquad z_i = |N^+_{>y_i}(i)|.$$

.

Define, $r(S_i) = y_i - 1 + \frac{z_i}{n+1}$ and $l(T_i) = \min\big(\{i\} \cup \{r(S_j) : j \in N^-_{<i}(i)\}\big)$.

Similarly let,

$$y'_i = \begin{cases} \min \overline{N^-_{>i}(i)}, & \text{if } \overline{N^-_{>i}(i)} \neq \emptyset \\ n + 1, & \text{otherwise} \end{cases} \qquad \text{and} \qquad z'_i = |N^-_{>y'_i}(i)|.$$

.

Define, $r(T_i) = y'_i - 1 + \frac{z'_i}{n+1}$ and $l(S_i) = \min\big(\{i\} \cup \{r(T_j) : j \in N^+_{<i}(i)\}\big)$.

Note that for each vertex $i \in V(G)$, by the above definition of intervals corresponding to i, we have that the point $i \in S_i \cap T_i$, $y_i - 1 \leq r(S_i) < y_i$ and $y'_i - 1 \leq r(T_i) < y'_i$. Also for any two vertices i, j such that $y_i = y_j = p$, we have by Observation 18 that $r(S_i) \leq r(S_j)$ if and only if

$z_i \leq z_j$ if and only if $N^+_{>p}(i) \subseteq N^+_{>p}(j)$. Similarly, for any two vertices i, j such that $y'_i = y'_j = q$, we have that $r(T_i) \leq r(T_j)$ if and only if $z'_i \leq z'_j$ if and only if $N^-_{>q}(i) \subseteq N^-_{>q}(j)$.

Now we have to prove that $E(G) = \{(i,j) : S_i \cap T_j \neq \emptyset\}$. Let $(i,j) \in E(G)$ be such that $i < j$. If $j < y_i$, then we have $l(S_i) \leq i < j \leq r(S_i)$, implying that $S_i \cap T_j \neq \emptyset$ (recall that $j \in S_j \cap T_j$). Suppose that $y_i < j$. Then we have $l(T_j) \leq r(S_i) < y_i < j < r(T_j)$ implying that $S_i \cap T_j \neq \emptyset$. In a similar way, by interchanging the roles of source and destination intervals and that of i and j, we can also prove that: if $(i,j) \in E(G)$ be such that $j < i$, then $S_i \cap T_j \neq \emptyset$. On the other hand, suppose that $(i,j) \notin E(G)$, where $i < j$. Clearly, then $y_i \leq j$. For the sake of contradiction assume that $S_i \cap T_j \neq \emptyset$. Since $r(S_i) < y_i$, this is possible only if $l(T_j) \leq r(S_i) < y_i \leq j$. Thus, $l(T_j) < j$, which implies by the definition of intervals that $N^-_{<j}(j) \neq \emptyset$. Let $k \in N^-_{<j}(j)$ such that $r(S_k) = \min\{r(S_l) : l \in N^-_{<j}(j)\}$. Since $(k,j) \in E(G)$ and $r(S_k) = l(T_j) < j$, we can conclude by the definition of $r(S_k)$ that $y_k < j$. Suppose that $y_i = y_k = p$. Then, since $r(S_k) = l(T_j) \leq r(S_i)$, we can conclude by our earlier observation that $N^+_{>p}(k) \subseteq N^+_{>p}(i)$, which contradicts the fact that $j \in N^+_{>p}(k) \setminus N^+_{>p}(i)$. We can thus infer that $y_i \neq y_k$. This together with the fact that $y_k - 1 \leq r(S_k) = l(T_j) < y_i$, implies that $y_k < y_i$. Suppose that $y_k \leq i$, then $k < y_k \leq i < j$, $(k,j) \in E(G)$, and $(k, y_k), (i,j) \notin E(G)$, which gives us Figure 4.4(i), which is a contradiction. Therefore we can assume that $i < y_k$, which further implies that $(i, y_k) \in E(G)$ (recall that $y_k < y_i$). Now we have $y_k \in N^+_{>\max\{i,k\}}(i) \setminus N^+_{>\max\{i,k\}}(k)$ and $j \in N^+_{>\max\{i,k\}}(k) \setminus N^+_{>\max\{i,k\}}(i)$, which contradicts Observation 18. As we arrive at a contradiction in every case, we can conclude that $S_i \cap T_j = \emptyset$. The case where $(i,j) \notin E(G)$ such that $j < i$ is symmetric. $\square$

Now we define the following.

Definition 28 (DUF-ordering). *A* directed umbrella-free ordering *(or in short a* DUF-ordering*) of a digraph G is an ordering of $V(G)$ satisfying the following properties for any three distinct vertices $i < j < k$:*

(*a*) *if $(i,k) \in E(G)$, then either $(i,j) \in E(G)$ or $(j,k) \in E(G)$, and*

(*b*) *if $(k,i) \in E(G)$, then either $(k,j) \in E(G)$ or $(j,i) \in E(G)$.*

Definition 29 (DUF-digraph). *A digraph G is a* directed umbrella-free digraph *(or in short a* DUF-digraph*) if it has a DUF-ordering.*

Then the following corollary is an immediate consequence of Theorem 12.

Corollary 2. *Every reflexive interval digraph is a DUF-digraph.*

Let G be an undirected graph. We define the *symmetric digraph of* G to be the digraph obtained by replacing each edge of G by symmetric arcs.

The following characterization of cocomparability graphs was first given by Damaschke [29].

Theorem 13 ([29]). *An undirected graph G is a cocomparability graph if and only if there is an ordering $<$ of $V(G)$ such that for any three vertices $i < j < k$, if $ik \in E(G)$, then either $ij \in E(G)$ or $jk \in E(G)$.*

Then we have the following corollary.

Corollary 3. *The underlying undirected graph of every DUF-digraph is a cocomparability graph.*

Note that there exist digraphs which are not DUF-digraphs but their underlying undirected graphs are cocomparability (for example, a directed triangle with edges $(a, b), (b, c)$ and (c, a)). But we can observe that the class of underlying undirected graphs of DUF-digraphs is precisely the class of cocomparability graphs, since it follows from Theorem 13 that the symmetric digraph of any cocomparability graph is a DUF-digraph. In contrast, the class of underlying undirected graphs of reflexive interval digraphs forms a strict subclass of cocomparability graphs. We prove this by showing that no directed graph that has $K_{3,3}$ as its underlying undirected graph can be a reflexive interval digraph ($K_{3,3}$ can easily be seen to be a cocomparability graph). This would also imply by Corollary 2 that the class of reflexive interval digraphs forms a strict subclass of DUF-digraphs.

Theorem 14. *The underlying undirected graph of a reflexive interval digraph cannot contain $K_{3,3}$ as an induced subgraph.*

Proof. Since the class of reflexive interval digraphs is closed under taking induced subgraphs, it is enough to prove that the underlying undirected graph of a reflexive interval digraph cannot be $K_{3,3}$. Let H be an undirected graph. An ordering $<$ of $V(H)$ is said to be a *special umbrella-free* ordering of H, if for any four distinct vertices $a, b, c, d \in V(G)$ such that $a < b < c < d$, $ad \in E(H)$ implies that either $ab \in E(H)$ or $cd \in E(H)$. Let G be any reflexive interval digraph. By Theorem 12, we have that $V(G)$ has an ordering such that none of the structures in Figure 4.4 are present. It follows that this ordering is also a special umbrella-free ordering of the underlying undirected graph of G. Therefore we can conclude that the underlying undirected graph of any reflexive interval digraph has a special umbrella-free ordering. We claim that $K_{3,3}$ does not have a special umbrella-free ordering, which then implies the theorem.

Let A and B denote the two partite sets of the bipartite graph $K_{3,3}$. Suppose for the sake of contradiction that $K_{3,3}$ has a special umbrella-free ordering $<= (v_1, v_2, \ldots, v_6)$. Suppose that v_1 and v_6 belong to different partite sets of $K_{3,3}$. Without loss of generality, we can assume that $v_1 \in A$ and $v_6 \in B$. This implies that there cannot exist vertices $v_i, v_j \in \{v_2, v_3, v_4, v_5\}$ such that $v_i < v_j$, $v_i \in A$ and $v_j \in B$, as otherwise we have $v_1 < v_i < v_j < v_6$, $v_1v_6 \in E(K_{3,3})$, and $v_1v_i, v_jv_6 \notin E(K_{3,3})$, which contradicts the fact that $<$ is a special umbrella-free ordering. This further implies that $v_2, v_3 \in B$ and $v_4, v_5 \in A$. Then we have $v_2 < v_3 < v_4 < v_5$, $v_2v_5 \in E(K_{3,3})$, and $v_2v_3, v_4v_5 \notin E(K_{3,3})$, which is again a contradiction. Therefore we can assume that v_1 and v_6 belong to the same partite set of $K_{3,3}$. Without loss of generality, we can assume that $v_1, v_6 \in A$. Now if $v_2 \in A$, then we have $v_3, v_4, v_5 \in B$. Then we have $v_1 < v_2 < v_3 < v_4$, $v_1v_4 \in E(K_{3,3})$, and $v_1v_2, v_3v_4 \notin E(K_{3,3})$, which is again a contradiction. This implies that $v_2 \in B$. Now if there exists a vertex $x \in \{v_4, v_5\} \cap A$, then we have $v_3 \in B$, in which case we have $v_2 < v_3 < x < v_6$, $v_2v_6 \in E(K_{3,3})$, and $v_2v_3, xv_6 \notin E(K_{3,3})$, which is again a contradiction. Therefore we can assume that $v_4, v_5 \in B$, implying that $v_3 \in A$. Then we have $v_1 < v_3 < v_4 < v_5$, $v_1v_5 \in E(K_{3,3})$, and $v_1v_3, v_4v_5 \notin E(K_{3,3})$, which is again a contradiction. This shows that $K_{3,3}$ has no special umbrella-free ordering, thereby proving the theorem. □

Prisner [100] proved the following.

Theorem 15 ([100]). *The underlying undirected graphs of interval nest digraphs are weakly triangulated graphs.*

By Corollaries 2, 3 and Theorem 14, we can conclude that the underlying undirected graph of reflexive interval digraphs are $K_{3,3}$-free cocomparability graphs. This strengthens Theorem 15, since now we have that the underlying undirected graphs of interval nest digraphs are $K_{3,3}$-free weakly triangulated cocomparability graphs.

4.3 Algorithms for reflexive interval digraphs

Here we explore the three different problems defined in Section 4.1 in the class of reflexive interval digraphs.

Let G be a reflexive interval digraph. Note that any induced subdigraph of G is also a reflexive interval digraph and that the "reversal" of G – the digraph obtained by replacing each edge (u, v) of G by (v, u)) – is also a reflexive interval digraph. Since in any digraph, a set S is an absorbing set (resp. kernel) if and only if it is a dominating set (resp. solution) in its reversal, this means that any algorithm that solves ABSORBING-SET (resp. KERNEL) problem for the class of reflexive interval digraphs can also be used to solve the DOMINATING-SET (resp. SOLUTION) problem on an input reflexive interval digraph. Therefore, in the sequel, we only study the ABSORBING-SET and KERNEL problems on reflexive interval digraphs.

4.3.1 Kernel

We use the following result of Prisner that is implied by Theorem 4.2 of [100].

Theorem 16 ([100])**.** *Let $\mathcal{C}$ be a class of digraphs that is closed under taking induced subgraphs. If in every graph $G \in \mathcal{C}$, there exists a vertex z such that for every $y \in N^-(z)$, $N^+(z) \setminus N^-(z) \subseteq N^+(y)$, then the class $\mathcal{C}$ is kernel-perfect.*

In fact, the following lemma guarantees the existence of such a vertex z (as specified in Theorem 16) in reflexive interval digraphs. Since the class of reflexive interval digraphs are closed under taking induced subgraphs, further we will see that how such a vertex can be iteratively used to find a kernel for any digraph belonging to this class.

Lemma 8. *Let G be a reflexive interval digraph G with interval representation $\{(S_u, T_u)\}_{u \in V(G)}$. Let z be the vertex such that $r(S_z) = \min\{r(S_v) : v \in V(G)\}$. Then for every $y \in N^-(z)$, $N^+(z) \setminus N^-(z) \subseteq N^+(y)$.*

Proof. Let $x \in N^+(z) \setminus N^-(z)$ and $y \in N^-(z)$. We have to prove that $x \in N^+(y)$. By the choice of z, we have that $r(S_x), r(S_y) > r(S_z)$. As $S_z \cap T_z \neq \emptyset$ (since G is reflexive interval digraph), we have $l(T_z) < r(S_z)$. Combining with the previous inequality, we have $l(T_z) < r(S_x)$. As $x \notin N^-(z)$, it then follows that $l(S_x) > r(T_z)$. Since $y \in N^-(z)$, we have that $l(S_y) < r(T_z)$. We now have that $l(S_y) < l(S_x)$. As $l(S_x) < r(T_x)$ this further implies that $l(S_y) < r(T_x)$. Now if $x \notin N^+(y)$, it should be the case that $l(T_x) > r(S_y) > r(S_z)$ which is a contradiction to the fact that $x \in N^+(z)$. □

Since reflexive interval digraphs are closed under taking induced subgraphs, by Theorem 16 and Lemma 8, we have the following.

Theorem 17. *Reflexive interval digraphs are kernel-perfect.*

It follows from the above theorem that the decision problem KERNEL is trivial on reflexive interval digraphs. As explained below, we can also compute a kernel in a reflexive interval digraph efficiently, if an interval representation of the digraph is known.

Let G be a reflexive interval digraph with an interval representation $\{(S_u, T_u)\}_{u \in V(G)}$. Let $G_0 = G$ and z_0 be the vertex in G such that $r(S_{z_0}) = \min\{r(S_v) : v \in V(G)\}$. For $i \geq 1$, recursively define G_i to be the induced subdigraph of G with $V(G_i) = V(G_{i-1}) \setminus (\{z_{i-1}\} \cup N^-(z_{i-1}))$ and if $V(G_i) \neq \emptyset$, define z_i to be the vertex such that $r(S_{z_i}) = \min\{r(S_v) : v \in V(G_i)\}$. Let t be smallest integer such that $V(G_{t+1}) = \emptyset$. Note that this implies that $V(G_t) = \{z_t\} \cup N^-_{G_t}(z_t)$. Clearly $t \leq n$ and $r(S_{z_0}) < r(S_{z_1}) < \cdots < r(S_{z_t})$. By Lemma 8, we have that for each $i \in \{1, 2, \ldots, t\}$, z_i has the following property: for any $y \in N^-_{G_i}(z_i)$ we have $N^+_{G_i}(z_i) \setminus N^-_{G_i}(z_i) \subseteq N^+_{G_i}(y)$.

We now recursively define a set $K_i \subseteq V(G_i)$ as follows: Define $K_t = \{z_t\}$. For each $i \in \{t-1, t-2, \ldots, 0\}$,

$$K_i = \begin{cases} \{z_i\} \cup K_{i+1} & \text{if } (z_i, z_j) \notin E(G), \text{ where } j = \min\{l : z_l \in K_{i+1}\} \\ K_{i+1} & \text{otherwise.} \end{cases}$$

Lemma 9. *For each $i \in \{1, 2, \ldots, t\}$, K_i is a kernel of G_i.*

Proof. We prove this by reverse induction on i. The base case where $K_t = \{z_t\}$ is trivial since $V(G_t) = \{z_t\} \cup N^-_{G_t}(z_t)$. Assume that the hypothesis is true for all j such that $j > i$. If $K_i = K_{i+1}$ then it implies that there exists $z_j \in K_{i+1}$ such that $z_j \in N^+(z_i)$. Further as $z_j \in V(G_{i+1}) = V(G_i) \setminus (\{z_i\} \cup N^-_{G_i}(z_i))$, we have that $z_j \in N^+_{G_i}(z_i) \setminus N^-_{G_i}(z_i)$. Let $y \in N^-_{G_i}(z_i)$. Since $N^+_{G_i}(z_i) \setminus N^-_{G_i}(z_i) \subseteq N^+_{G_i}(y)$, we then have that $y \in N^-_{G_i}(z_j)$. Thus $N^-_{G_i}(z_i) \subseteq N^-_{G_i}(z_j)$. As $z_i \in N^-(z_j)$, it follows that every vertex in $V(G_i) \setminus V(G_{i+1}) = \{z_i\} \cup N^-_{G_i}(z_i)$ is an in-neighbor of z_j. We can now use the induction hypothesis to conclude that $K_i = K_{i+1}$ is a kernel of G_i. On the other hand, if $K_i = \{z_i\} \cup K_{i+1}$, then it should be the case that $(z_i, z_j) \notin E(G)$ where $j = \min\{l : z_l \in K_{i+1}\}$. Now consider any $z_l \in K_{i+1}$ where $z_l \neq z_j$. By definition of j, we have $l > j$. If $(z_i, z_l) \in E(G)$, then as $r(S_{z_i}) < r(S_{z_j}) < r(S_{z_l})$, it should be the case that $l(T_{z_l}) < r(S_{z_i}) < r(S_{z_j}) < r(S_{z_l})$. We also have $l(S_{z_j}) < l(S_{z_l})$ as otherwise $S_{z_j} \subseteq S_{z_l}$, implying that $S_{z_l} \cap T_{z_j} \neq \emptyset$, contradicting the fact that $(z_l, z_j) \notin E(G)$ (as z_l and z_j both belong to K_{i+1}, which by the induction hypothesis is a kernel of G_{i+1}). Since $r(T_{z_l}) > l(S_{z_l}) > l(S_{z_j})$ and $r(S_{z_j}) > l(T_{z_l})$, we now have that $S_{z_j} \cap Tz_l \neq \emptyset$, which is a contradiction to the fact that $(z_j, z_l) \notin E(G)$ (as $z_j, z_l \in K_{i+1}$, which by the induction hypothesis is a kernel of G_{i+1}). Thus no vertex in K_{i+1} can be an out-neighbor of z_i. By definition of G_{i+1}, no vertex in G_{i+1}, and hence no vertex in K_{i+1}, can be an in-neighbor of z_i. Then we have by the induction hypothesis that $K_i = \{z_i\} \cup K_{i+1}$ is an independent set. Since the only vertices in $V(G_i) \setminus V(G_{i+1})$ are

$\{z_i\} \cup N^-_{G_i}(z_i)$, and K_{i+1} is an absorbing set of G_{i+1} by the induction hypothesis, we can conclude that $K_i = \{z_i\} \cup K_{i+1}$ is an absorbing set of G_i. Therefore K_i is a kernel of G_i. □

By the above lemma, we have that K_0 is a kernel of G. We can now construct an algorithm that computes a kernel in a reflexive interval digraph G, given an interval representation of it. We assume that the interval representation of G is given in the form of a list of left and right end-points of intervals corresponding to the vertices. We can process this list from left to right in a single pass to compute the list of vertices $z_0, z_1, \dots, z_t$ in $O(n+m)$ time. We then process this new list from right to left in a single pass to generate a set K as follows: initialize $K = \{z_t\}$ and for each $i \in \{t-1, t-2, \dots, 0\}$, add z_i to K if it is not an in-neighbor of the last vertex that was added to K. Clearly, the set K can be generated in $O(n+m)$ time. It is easy to see that $K = K_0$ and therefore by Lemma 9, K is a kernel of G. Thus, we have the following theorem.

Theorem 18. *A kernel of a reflexive interval digraph can be computed in linear-time, given an interval representation of the digraph as input.*

The linear-time algorithm described above is an improvement and generalization of the Prisner's result that, interval nest digraphs and their reversals are kernel-perfect, and a kernel can be found in these graphs in time $O(n^2)$ if a representation of the graph is given [100].

Now it is interesting to note that even for some kernel perfect digraphs with a polynomial-time computable kernel, the problems Min-Kernel and Max-Kernel turn out to be NP-complete. The following remark provides an example of such a class of digraphs.

Remark 4. *Let $\mathcal{C}$ be the class of symmetric digraphs of undirected graphs. Note that the class $\mathcal{C}$ is kernel-perfect, as for any $G \in \mathcal{C}$ the kernels of the digraph G are exactly the independent dominating sets of its underlying undirected graph. Note that any maximal independent set of an undirected graph is also an independent dominating set of it. Therefore, as a maximal independent set of any undirected graph can be found in linear-time, the problem* Kernel *is linear-time solvable for the class $\mathcal{C}$. On the other hand, note that the problems* Min-Kernel *and* Max-Kernel *for the class $\mathcal{C}$ is equivalent to the problems of finding a minimum cardinality independent dominating set and a maximum cardinality independent set for the class of undirected graphs, respectively. Since the latter problems are NP-complete for the class of undirected graphs, we have that the problems* Min-Kernel *and* Max-Kernel *are NP-complete in $\mathcal{C}$.*

Note that unlike the class of reflexive interval digraphs, the class of DUF-digraphs are not kernel perfect. Figure 4.5 provides an example for a DUF-digraph that has no kernel. Since that

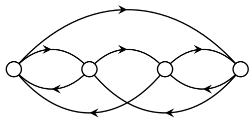


Figure 4.5: An example of a DUF-digraph that has no kernel.

graph is a semi-complete digraph (i.e. each pair of vertices is adjacent), and every vertex has an out-neighbor which is not its in-neighbor, it cannot have a kernel. The ordering of the vertices of the graph that is shown in the figure can easily be verified to be a DUF-ordering.

In contrast to Remark 4, even though DUF-digraphs may not have kernels, we show in the next section that the problems KERNEL and MIN-KERNEL can be solved in polynomial time in the class of DUF-digraphs. In fact we give a polynomial-time algorithm that, given a DUF-digraph G with a DUF-ordering as input, either finds a minimum sized kernel in G or correctly concludes that G does not have a kernel.

4.3.2 Minimum sized kernel

Let G be a DUF-digraph with vertex set $[1, n]$. We assume without loss of generality that $<: (1, 2, \dots, n)$ is a DUF-ordering of G. For ease of notation, in this section, we shorten $N^+_{>i}(i)$ and $N^-_{>i}(i)$ to $N^+_>(i)$ and $N^-_>(i)$ respectively. We further define $N_>(i) = N^+_>(i) \cup N^-_>(i)$ and for each $A \in \{N^+_>(i), N^-_>(i), N_>(i)\}$, we define $\overline{A} = [i+1, n] \setminus A$.

For any vertex $i \in \{1, 2, \dots, n\}$, let $P_i = \{j : j \in \overline{N_>(i)}$ such that $[i+1, j-1] \subseteq N^-(i) \cup N^-(j)\}$ and let $G[i, n]$ denote the subgraph induced in G by the set $[i, n]$. Note that we consider $[i+1, j-1] = \emptyset$, if $j = i+1$. For a collection of sets $\mathcal{S}$, we denote by $\mathrm{Min}(\mathcal{S})$ an arbitrarily chosen set in $\mathcal{S}$ of the smallest cardinality. For each $i \in \{1, 2, \dots, n\}$, we define a set $K(i)$ as follows. Here, when we write $K(i) = \infty$, we mean that the set $K(i)$ is undefined.

$$K(i) = \begin{cases} \{i\}, & \text{if } N^-_>(i) = \{i+1, \dots, n\} \\ \{i\} \cup \mathrm{Min}\{K(j) \neq \infty : j \in P_i\}, & \text{if } P_i \neq \emptyset \text{ and } \exists j \in P_i \text{ such that } K(j) \neq \infty \\ \infty, & \text{otherwise} \end{cases}$$

Note that it follows from the above definition that $K(n) = \{n\}$. For each $i \in \{1, 2, \dots, n\}$, let $OPT(i)$ denote a minimum sized kernel of $G[i, n]$ that also contains i. If $G[i, n]$ has no kernel that contains i, then we say that $OPT(i) = \infty$. We then have the following lemma.

Lemma 10. *The following hold.*

(a) *If $K(i) \neq \infty$, then $K(i)$ is a kernel of $G[i,n]$ that contains i, and*

(b) *if $OPT(i) \neq \infty$, then $K(i) \neq \infty$ and $|K(i)| = |OPT(i)|$.*

Proof. (a) We prove this by the reverse induction on i. Suppose that $K(i) \neq \infty$. The base case where $i = n$ is trivially true. Assume that the hypothesis is true for every $j > i$. It is clear from the definition of $K(i)$ that $i \in K(i)$. If $K(i) = \{i\}$, then it should be the case that $N^-_>(i) = \{i+1, \ldots, n\}$, implying that the set $K(i) = \{i\}$, is both an independent set and an absorbing set in $G[i,n]$, and we are done. Otherwise, $K(i) = \{i\} \cup K(j)$ for some $j \in P_i$ such that $K(j) \neq \infty$. By the definition of P_i, we have that $j \in \overline{N_>(i)}$ and $[i+1, j-1] \subseteq N^-(i) \cup N^-(j)$. Since $j > i$, we have by the induction hypothesis that $K(j)$ is an independent and absorbing set in $G[j,n]$. Suppose that there exists $k \in K(j)$, such that $k \in N(i)$. Since $j \in \overline{N_>(i)}$ we have that $j \neq k$, which implies that $k > j$. We then have vertices $i < j < k$ such that $k \in N(i)$, $j \notin N(i)$ and $k \notin N(j)$, which is a contradiction to the fact that $<$ is a DUF-ordering. Therefore we can conclude that $K(i) = \{i\} \cup K(j)$ is an independent set in $G[i,n]$. Since $j \in P_i$, we have by the definition of P_i that $[i+1, j-1] \subseteq N^-(i) \cup N^-(j)$. It then follows from the fact that $K(j)$ is an absorbing set of $G[j,n]$ containing j that $K(i) = \{i\} \cup K(j)$ is an absorbing set of $G[i,n]$. Thus $K(i)$ is a kernel of $G[i,n]$ that contains i.

(b) Suppose that $OPT(i) \neq \infty$. The proof is again by reverse induction on i. The base case where $i = n$ is trivially true. Assume that the hypothesis is true for any $j > i$. If $|OPT(i)| = 1$, then it should be the case that $OPT(i) = \{i\}$ and $j \in N^-(i)$ for each $j \in \{i+1, \ldots, n\}$, i.e. $N^-_>(i) = \{i+1, \ldots, n\}$. By the definition of $K(i)$, we then have $K(i) = \{i\}$, and we are done. Therefore we can assume that $|OPT(i)| > 1$. Let $j = \min(OPT(i) \setminus \{i\})$. Clearly, $j > i$. As $OPT(i)$ is an independent set, we have that $j \in \overline{N_>(i)}$. We claim that $j \in P_i$. Suppose that there exists a vertex $y \in [i+1, j-1]$ such that $y \notin N^-(i) \cup N^-(j)$. Since $OPT(i)$ is an absorbing set in $G[i,n]$, there exists a vertex $k \in OPT(i) \setminus \{i,j\}$ such that $y \in N^-(k)$. By the choice of j and the definition of k, we have that $j < k$ and $(j,k) \notin E(G)$. Then we have $y < j < k$, $(y,k) \in E(G)$, and $(y,j), (j,k) \notin E(G)$, which is a contradiction to the fact that $<$ is a DUF-ordering. Therefore we can conclude that $[i+1, j-1] \subseteq N^-(i) \cup N^-(j)$, which implies by the definition of P_i that $j \in P_i$. This proves our claim. Note that if there exists a vertex $z \in (N^-(i) \setminus N^-(j)) \cap [j,n]$, then we have vertices $i < j < z$ such that $(z,i) \in E(G)$ and $(z,j), (j,i) \notin E(G)$, which is a contradiction to the fact that $<$ is a DUF-ordering. Therefore we can assume that $N^-(i) \cap [j,n] \subseteq N^-(j) \cap [j,n]$. This implies that $OPT(i) \setminus \{i\}$ is a kernel of $G[j,n]$ that contains j. Thus $OPT(j) \neq \infty$, which implies by the induction hypothesis that $K(j) \neq \infty$ and $|K(j)| = |OPT(j)| \leq |OPT(i) \setminus \{i\}|$. Since $j \in P_i$ and $K(j) \neq \infty$, we have $K(i) \neq \infty$, and further we have $|K(i)| \leq |\{i\} \cup K(j)| \leq 1 + |OPT(i) \setminus \{i\}| = |OPT(i)|$. By (a), $K(i)$ is a kernel of $G[i,n]$ that contains i, and hence we have $|K(i)| = |OPT(i)|$. $\square$

Suppose that G has a kernel. Now let OPT denote a minimum sized kernel in G. Let $\mathcal{K} = \{K(j) \neq \infty : [1, j-1] \subseteq N^-(j)\}$. Note that we consider $[1, j-1] = \emptyset$ if $j = 1$. By Lemma 10(a), it follows that every member of $\mathcal{K}$ is a kernel of G. So if G does not have a kernel, then $\mathcal{K} = \emptyset$. The following lemma shows that the converse is also true.

Lemma 11. *If G has a kernel, then $\mathcal{K} \neq \emptyset$ and $|OPT| = \big|\mathrm{Min}(\mathcal{K})\big|$.*

Proof. Suppose that G has a kernel. Then clearly, OPT exists. Let $j = \min\{i : i \in OPT\}$. Then it should be the case that $[1, j-1] \subseteq N^-(j)$. As otherwise, there exist vertices $j' \in [1, j-1]$ and $k \in OPT$ such that $j' \in N^-(k) \setminus N^-(j)$. Since OPT is an independent set, this implies that we have vertices, $j' < j < k$ such that $(j', k) \in E(G)$ and $(j', j), (j, k) \notin E(G)$ which is a contradiction to the fact that $<$ is a DUF-ordering. Also by the choice of j, we have that $OPT \subseteq [j, n]$. Then as OPT is a kernel of G, OPT is a kernel of $G[j, n]$ that contains j. This implies that $OPT(j) \neq \infty$ and $|OPT(j)| \leq |OPT|$. Therefore by Lemma 10, we have that $K(j) \neq \infty$ and $|K(j)| = |OPT(j)|$. Thus $K(j) \in \mathcal{K}$, which implies that $\mathcal{K} \neq \emptyset$. Further, $\big|\mathrm{Min}(\mathcal{K})\big| \leq |K(j)| = |OPT(j)| \leq |OPT|$. Since every member of $\mathcal{K}$ is a kernel of G, it now follows that $\big|\mathrm{Min}(\mathcal{K})\big| = |OPT|$. $\square$

We thus have the following theorem.

Theorem 19. *The DUF-digraph G has a kernel if and only if $K(j) \neq \infty$ for some j such that $[1, j-1] \subseteq N^-(j)$. Further, if G has a kernel, then the set $\{K(j) \neq \infty : [1, j-1] \subseteq N^-(j)\}$ contains a kernel of G of minimum possible size.*

Let G be a DUF-digraph with vertex set $[1, n]$. For each $i \in [1, n]$, we can compute the set P_i in $O(n + m)$ time as follows. We mark the in-neighbors of i in $[i, n]$ and then scan the vertices from i to n in a single pass in order to collect the vertices which are not in-neighbors of i in an ordered list L. Initialize $P_i = \emptyset$. We mark every out-neighbor of i in L. Now for each unmarked vertex j in L (processed from left to right), we add j to P_i if and only if every vertex of L before j is an in-neighbor of j. Note that this computation of P_i can be done in $O(n + m)$ time. This implies that we can precompute the set $\{P_i : i \in [1, n]\}$ in $O((n+m)n)$ time. Now since $|P_i| \leq n$, it is easy to see from the recursive definition for $K(i)$ that $\{K(i) : i \in [1, n]\}$ can be computed in

$O(n^2)$ time. For $j \in [1, n]$, we can check in $O(n+m)$ time whether $[1, j-1] \subseteq N^-(j)$. Thus in $O((n+m)n)$ time, we can compute the minimum sized set in $\{K(j) \neq \infty : [1, j-1] \subseteq N^-(j)\}$. Therefore by Theorem 19, we have the following corollary.

Corollary 4. *The* MIN-KERNEL *problem can be solved for DUF-digraphs in $O((n+m)n)$ time if the DUF-ordering is known. Consequently, for a reflexive interval digraph, the* MIN-KERNEL *problem can be solved in $O((n+m)n)$ time if the interval representation is given as input.*

Let G be a cocomparability graph. Let H be the symmetric digraph of G. Now it is easy to see that a set $K \subseteq V(H) = V(G)$, is a kernel of H if and only if K is an independent dominating set of G. Therefore a kernel of minimum possible size in H will be a minimum independent dominating set in G. Note that a vertex ordering of a cocomparability graph that satisfies the properties in Theorem 13 can be found in linear time [89]. Let $<$ be such a vertex ordering of G. As noted before, H is a DUF-digraph with DUF-ordering $<$. Thus an algorithm that computes a minimum sized kernel in H also computes a minimum independent dominating set in G. From Corollary 4, we now have the following.

Corollary 5. *An independent dominating set of minimum possible size can be found in $O((n+m)n)$ time in cocomparability graphs.*

The above corollary is an improvement over the result by Kratsch and Stewart [79] that an independent dominating set of minimum possible size problem can be computed in $O(n^3)$ time for cocomparability graphs.

We now show that a minimum sized kernel of an adjusted interval digraph, whose interval representation is known, can be computed more efficiently than in the case of DUF-digraphs. Let G be an adjusted interval digraph with an interval representation $\{(S_u, T_u)\}_{u \in V(G)}$. Note that by the definition of adjusted interval digraphs, we have that $l(S_u) = l(T_u)$ for each $u \in V(G)$. Let $<$ be an ordering of vertices in G with respect to the common left end-points of intervals corresponding to each vertex. Then $<$ has the following property: for any three distinct vertices $u < v < w$, if $(u, w) \in E(G)$ then $(u, v) \in E(G)$ and if $(w, u) \in E(G)$ then $(v, u) \in E(G)$. Then note that $<$ is also a DUF-ordering of $V(G)$. Further, the out-neighbors (resp. in-neighbors) of every vertex occur consecutively in $<$. This implies that the neighbors of every vertex occur consecutively in $<$. Further we have

$$\text{if } [x, y] \subseteq N^-(y) \text{ (resp. } N^+(y)\text{) then for any } z \in [x, y], \text{ we have } [x, z] \subseteq N^-(z) \text{ (resp. } N^+(z)\text{).} \tag{4.1}$$

Let $V(G) = [1, n]$ so that $<= (1, 2, \ldots, n)$. We can compute the sets $\{\max N^+(i) \colon i \in [1, n]\}$ and $\{\max N^-(i) \colon i \in [1, n]\}$ in $O(n + m)$ time by just preprocessing the adjacency list of G. Since the vertices in $N^+_>(i)$ (resp. $N^-_>(i)$) occur consecutively in $<$, we can also compute the set $\{\min \overline{N_>(i)} = \max N(i) + 1 \colon i \in [1, n]\}$ in $O(n + m)$ time (if $\max N(i) = n$, then we set $\min \overline{N_>(i)} = n + 1$). Let $i \in [1, n]$. We can construct P_i as follows. We compute $x = \min\{\max N^+(j) \colon j \in [\min \overline{N^-_>(i)} = \max N^-(i)+1, n]\}$ in $O(n)$ time (note that if $\max N^-(i) = n$, then $\min \overline{N_>(i)} = n+1$, in which case we can just set $P_i = \emptyset$). We claim that $P_i = [\min \overline{N_>(i)}, x]$. To see this, first note that for every vertex i, $N^+_>(i) = [i + 1, \max N^+(i)]$. Therefore, since for any vertex $z \in [\min \overline{N^-_>(i)}, x]$, we have $\max N^+(z) \geq x$, we can conclude that $(z, x) \in E(G)$. Thus $[\min \overline{N^-_>(i)}, x] \subseteq N^-(x)$. Therefore by property (4.1), for each $z \in [\min \overline{N_>(i)}, x] \subseteq [\min \overline{N^-_>(i)}, x]$, we have that $[\min \overline{N^-_>(i)}, z] \subseteq N^-(z)$. Since the vertices of $N^-_>(i)$ are consecutive in $<$, this means that $[i, z] \subseteq N^-(i) \cup N^-(z)$. Therefore we have that $[\min \overline{N_>(i)}, x] \subseteq P_i$. Now consider any $z > x$. By the definition of x, there exists $j \in [\min \overline{N^-_>(i)}, x]$ such that $\max N^+(j) = x$. Then as $x < z$, we have $(j, z) \notin E(G)$, which implies that $j \notin N^-(i) \cup N^-(z)$. Thus $z \notin P_i$. Therefore we can conclude that $P_i = [\min \overline{N_>(i)}, x]$. Note that if $\min \overline{N_>(i)} > x$, then $P_i = \emptyset$. It is clear that the set P_i can be computed in this way in $O(n)$ time for an $i \in [1, n]$. So the set $\{P_i : i \in [1, n]\}$ can be computed in $O(n^2)$ time. The sets $\{K(i) \colon i \in [1, n]\}$ can then be computed in $O(n^2)$ time as before. Now we compute $y = \min\{\max N^+(j) \colon j \in [1, n]\}$ in $O(n)$ time. Then $[1, y - 1] \subseteq N^-(y)$. Therefore by property (4.1), for each $z \in [1, y]$ we have $[1, z - 1] \subseteq N^-(z)$. Now consider any $z > y$. By the definition of y, there exists $j \in [1, y]$ such that $y = \max N^+(j)$. Then as $y < z$, we have that $(j, z) \notin E(G)$, which implies that $[1, z - 1] \nsubseteq N^-(z)$. Therefore we can conclude that $[1, y] = \{j \colon [1, j - 1] \subseteq N^-(j)\}$. By Theorem 19, we can just output in $O(n)$ time a set of minimum size in $\{K(i) \colon i \in [1, y]$ and $K(i) \neq \infty\}$ as a minimum sized kernel of G. Thus we have the following corollary.

Corollary 6. *The* Min-Kernel *problem can be solved in $O(n^2)$ time in adjusted interval digraphs, given an adjusted interval representation of the input graph.*

Remark 5. *Note that the* Max-Kernel *problem can also be solved in $O((n + m)n)$ time for the class of DUF-digraphs, by a minor modification of our algorithm that solves* Min-Kernel *problem (replace* $\mathrm{Min}\{K(j) \neq \infty : j \in P_i\}$ *in the recursive definition of $K(i)$ by* $\mathrm{Max}\{K(j) \neq \infty : j \in P_i\}$ *and follow the same procedure. Then we have that if kernel exists, then a maximum sized kernel is given by* $\mathrm{Max}(\mathcal{K})$*). Further, the recursive definition can also be easily adapted to the weighted version of the problems* Min-Kernel *and* Max-Kernel *in $O((n + m)n)$ time.*

4.3.3 Minimum absorbing set

Recall that given any digraph G, the splitting bigraph B_G is defined as follows: $V(B_G)$ is partitioned into two sets $V' = \{u' : u \in V(G)\}$ and $V'' = \{u'' \colon u \in V(G)\}$, and $E(B_G) = \{u'v'' \colon (u,v) \in E(G)\}$ (refer to (a) and (b) of Figure 3.2 for an illustration). Müller [95] observed that G is an interval digraph if and only if B_G is an interval bigraph (Proposition 1 (b)).

Recall that for a bipartite graph having two specified partite sets A and B, a set $S \subseteq B$ such that $\bigcup_{u \in S} N(u) = A$ is called an A-dominating set (or a red-blue dominating set). If G is a reflexive interval digraph, then every V'-dominating set of B_G corresponds to an absorbing set of G and vice versa. To be precise, if $S \subseteq V''$ is a V'-dominating set of B_G, then $\{u \colon u'' \in S\}$ is an absorbing set of G and if $S \subseteq V(G)$ is an absorbing set of G, then $\{u'' \colon u \in S\}$ is a V'-dominating set of B_G (note that this is not true for general interval digraphs). Thus finding a minimum cardinality absorbing set in G is equivalent to finding a minimum cardinality V'-dominating set in the bipartite graph B_G. We show in this section that the problem of computing a minimum cardinality A-dominating set is linear time solvable for interval bigraphs. This implies that the Absorbing-Set problem can be solved in linear time on reflexive interval digraphs.

Consider an interval bigraph H with partite sets A and B. Let $\{I_u\}_{u \in V(H)}$ be an interval representation for H; i.e. $uv \in E(H)$ if and only if $u \in A$, $v \in B$ and $I_u \cap I_v \neq \emptyset$. Let $|A| = t$. We assume without loss of generality that $A = \{1, 2, \ldots, t\}$, where $r(I_i) < r(I_j) \Leftrightarrow i < j$. We also assume that there are no isolated vertices in A, as otherwise H does not have any A-dominating set. For each $i \in \{1, 2, \ldots, t\}$, we compute a minimum cardinality subset $DS(i)$ of B that dominates $\{i, i+1, \ldots, t\}$, i.e. $\{i, i+1, \ldots, t\} \subseteq \bigcup_{u \in DS(i)} N(u)$. Then $DS(1)$ will be a minimum cardinality A-dominating set of H. We first define some parameters that will be used to define $DS(i)$.

Let $i \in \{1, 2, \ldots, t\}$. We define $\rho(i) = \max_{u \in N(i)} r(I_u)$ and let $R(i)$ be a vertex in $N(i)$ such that $r(I_{R(i)}) = \rho(i)$. Since A does not contain any isolated vertices, $\rho(i)$ and $R(i)$ exist for each $i \in \{1, 2, \ldots, t\}$. Let $\lambda(i) = \min\{j \colon \rho(i) < l(I_j)\}$. Note that $\lambda(i)$ may not exist. It can be seen that if $\lambda(i)$ exists, then $\lambda(i) > i$ in the following way. Let $j = \lambda(i)$. Clearly, $\rho(i) < l(I_j)$. As $R(i) \in N(i)$, we have $l(I_i) < \rho(i)$, which implies that $i \neq j$. If $j < i$, then it should be the case that $l(I_i) < \rho(i) < l(I_j) < r(I_j) < r(I_i)$, which implies that any interval I_x, where $x \in B$, that intersects I_j also intersects I_i, and $r(I_x) > \rho(i)$. But this contradicts our choice of $\rho(i)$ and $R(i)$. Thus $N(j) = \emptyset$, implying that j is an isolated vertex in A, which is a contradiction. Therefore, we can conclude that for any $i \in A$, $\lambda(i) > i$.

Lemma 12. *Let $i \in \{1, 2, \dots, t\}$. If $\lambda(i)$ exists, then $R(i)$ dominates every vertex in $\{i, i+1, \dots, \lambda(i)-1\}$ and otherwise, $R(i)$ dominates every vertex in $\{i, i+1, \dots, t\}$.*

Proof. We first note that as $R(i) \in N(i)$, we have $l(I_{R(i)}) \leq r(I_i)$, as otherwise the intervals $I_{R(i)}$ and I_i will be disjoint.

Suppose that $\lambda(i)$ exists. Then consider any $j \in \{i, i+1, \dots, \lambda(i)-1\}$. Suppose for the sake of contradiction that $R(i) \notin N(j)$. Clearly, $j \neq i$ as $R(i) \in N(i)$. So we have $i < j < \lambda(i)$. Since $I_{R(i)}$ and I_j are disjoint, we have either $\rho(i) = r(I_{R(i)}) < l(I_j)$ or $r(I_j) < l(I_{R(i)})$. In the former case, since $i < j < \lambda(i)$, we have a contradiction to the choice of $\lambda(i)$. So we can assume that $r(I_j) < l(I_{R(i)})$. Recalling that $l(I_{R(i)}) \leq r(I_i)$, we now have that $r(I_j) < r(I_i)$, which contradicts the fact that $j > i$. Thus, $R(i)$ dominates every vertex in $\{i, i+1, \dots, \lambda(i)-1\}$. Next, suppose that $\lambda(i)$ does not exist. Then consider any vertex $j > i$. Since $\lambda(i)$ does not exist, we have $l(I_j) \leq \rho(i) = r(I_{R(i)})$. Since $l(I_{R(i)}) \leq r(I_i)$ and $r(I_i) < r(I_j)$, we have $l(I_{R(i)}) < r(I_j)$. Thus, the intervals I_j and $I_{R(i)}$ intersect for every $j > i$, implying that $R(i)$ dominates every vertex in $\{i, i+1, \dots, t\}$. $\square$

We now explain how to compute $DS(i)$ for each $i \in \{1, 2, \dots, t\}$. We recursively define $DS(i)$ as follows:

$$DS(i) = \begin{cases} \{R(i)\} \cup DS(\lambda(i)) & \text{if } \lambda(i) \text{ exists} \\ \{R(i)\} & \text{otherwise} \end{cases}$$

Lemma 13. *For each $i \in \{1, 2, \dots, t\}$, the set $DS(i)$ as defined above is a minimum cardinality subset of B that dominates $\{i, i+1, \dots, t\}$.*

Proof. We prove this by reverse induction on i. The base case where $i = t$ is trivial, by the definition of $R(t)$. Let $i < t$. Assume that the hypothesis holds for any $j > i$. If $\lambda(i)$ does not exist, then by Lemma 12, $R(i)$ dominates every vertex in $\{i, i+1, \dots, t\}$. This implies that $DS(i) = \{R(i)\}$ is a minimum cardinality subset of B that dominates $\{i, i+1, \dots, t\}$ and we are done. Therefore let us assume that $\lambda(i)$ exists. Then by the recursive definition of $DS(i)$, we have that $DS(i) = \{R(i)\} \cup DS(\lambda(i))$. Since $\lambda(i) > i$, we have by the inductive hypothesis that $DS(\lambda(i))$ is a minimum cardinality subset of B that dominates every vertex in $\{\lambda(i), \lambda(i)+1, \dots, t\}$. Since by Lemma 12, we have that $R(i)$ dominates every vertex in $\{i, i+1, \dots, \lambda(i)-1\}$, we then have that $DS(i) = \{R(i)\} \cup DS(\lambda(i))$ dominates every vertex in $\{i, i+1, \dots, t\}$. Consider any set $OPT \subseteq B$ that dominates $\{i, i+1, \dots, t\}$. Clearly, there exists a $u \in OPT$ such that $i \in N(u)$. By the definition of $R(i)$, we know that $r(I_u) \leq r(I_{R(i)}) = \rho(i)$.

Since $\rho(i) < l(I_{\lambda(i)})$, this implies that $\lambda(i) \notin N(u)$. Then, since $\lambda(i) \in \{i, i+1, \ldots, t\}$, there must exist a vertex $v \in OPT \setminus \{u\}$ such that $\lambda(i) \in N(v)$. Now consider any vertex $j \in N(u) \cap \{\lambda(i), \lambda(i)+1, \ldots, t\}$. We have $r(I_j) \geq r(I_{\lambda(i)}) \geq l(I_{\lambda(i)}) > \rho(i) \geq r(I_u)$. Since $j \in N(u)$, we have $l(I_j) \leq r(I_u)$, which implies that $l(I_j) < l(I_{\lambda(i)}) \leq r(I_{\lambda(i)}) < r(I_j)$. This implies that every interval that intersects $I_{\lambda(i)}$ also intersects I_j, in particular $j \in N(v)$. Applying the argument for every $j \in N(u) \cap \{\lambda(i), \lambda(i)+1, \ldots, t\}$, we can conclude $N(u) \cap \{\lambda(i), \lambda(i)+1, \ldots, t\} \subseteq N(v)$. Since OPT dominates every vertex in $\{\lambda(i), \lambda(i)+1, \ldots, t\}$, this implies that $OPT \setminus \{u\}$ dominates every vertex in $\{\lambda(i), \lambda(i)+1, \ldots, t\}$. Since by the inductive hypothesis, $DS(\lambda(i))$ is a minimum cardinality subset of B that dominates every vertex in the same set, we have that $|OPT \setminus \{u\}| \geq |DS(\lambda(i))|$. Then $|OPT| \geq |DS(\lambda(i)) \cup \{R(i)\}| = |DS(i)|$. This proves that $DS(i)$ is a minimum cardinality subset of B that dominates every vertex in $\{i, i+1, \ldots, t\}$. □

It is not difficult to verify that given an interval representation of the interval bigraph H with partite sets A and B, the parameters $R(i)$ and $\lambda(i)$ can be computed for each $i \in A$ in $O(n+m)$ time. Also, given a reflexive interval digraph G, the interval bigraph B_G can be constructed in linear time. Thus we have the following corollary.

Corollary 7. *The* Red-Blue Dominating Set *problem can be solved in interval bigraphs in linear time, given an interval representation of the bigraph as input. Consequently, the* Absorbing-Set *(resp.* Dominating-Set*) problem can be solved in linear time in reflexive interval digraphs, given an interval representation of the input digraph.*

Note that even if an interval representation of the interval bigraph is not known, it can be computed in polynomial time using Müller's algorithm [95]. Thus given just the adjacency list of the graph as input, the Red-Blue Dominating Set problem is polynomial-time solvable on interval bigraphs and the Absorbing-Set (resp. Dominating-Set) problem is polynomial-time solvable on reflexive interval digraphs.

4.3.4 Maximum independent set

We have the following theorem due to McConnell and Spinrad [89].

Theorem 20. *An independent set of maximum possible size can be computed for cocomparability graphs in $O(n+m)$ time.*

Let G be a DUF-digraph. Let H be the underlying undirected graph of G. Then by Corollary 3, we have that H is a cocomparability graph. Note that the independent sets of G and H are exactly the same. Therefore any algorithm that finds a maximum cardinality independent set in cocomparability graphs can be used to solve the INDEPENDENT-SET problem in DUF-digraphs. Thus by the above theorem, we have the following corollary.

Corollary 8. *The* INDEPENDENT-SET *problem can be solved for DUF-digraphs in $O(n+m)$ time. Consequently, the* INDEPENDENT-SET *problem can be solved for reflexive interval digraphs in $O(n+m)$ time.*

The above corollary generalizes and improves the $O(mn)$ time algorithm due to Prisner's [100] observation that underlying undirected graph of interval nest digraphs are weakly triangulated (Theorem 15) and the fact that maximum cardinality independent set problem can be solved for weakly triangulated graphs in $O(mn)$ time [67]. Note that the weighted INDEPENDENT-SET problem can also be solved for DUF-digraphs in $O(n+m)$ time, as the problem of finding a maximum weighted independent set in a cocomparability graphs can be solved in linear time [76].

Chapter 5

Hardness Results for Point-Point Digraphs

In this chapter, we prove the hardness results for the problems KERNEL, MIN-KERNEL, MAX-KERNEL, ABSORBING-SET, DOMINATING-SET, and INDEPENDENT-SET (studied in the last chapter) for the class of point-point digraphs. Here we also provide two characterizations for point-point digraphs. One of the characterization helps us to study the hardness of the before mentioned problems for point-point digraphs and the other characterization yields a linear-time recognition algorithm for the class of point-point digraphs.

First let us give a brief introduction to the approximation hardness of the problems.

5.1 Approximation hardness

Refer to [28] for a detailed overview of the following concepts.

NPO: An *NP-optimization problem* Q can be defined as a quadruple $(I_Q, sol_Q, c_Q, type)$ satisfying the following conditions.

- (a) I_Q denotes the set of all instances of Q which are polynomial-time recognizable.
- (b) For an instance $x \in I_Q$, $sol_Q(x)$ denotes the set of all feasible solutions of x. There exists a polynomial p such that for any $y \in sol_Q(x)$, $|y| \leq p(|x|)|$. Moreover for any $x \in I_Q$ and y with $|y| \leq p(|x|)$, whether $y \in sol_Q(x)$ can be determined in polynomial time.
- (c) Given an instance $x \in I_Q$ and a feasible solution $y \in sol_Q(x)$, c_Q is a polynomial-time computable function that measures the value of the solution y, and is denoted as $c_Q(y)$.
- (d) $type \in \{\min, \max\}$.

The objective of an NP-optimization problem Q is to find an *optimum solution*, OPT_Q for a given instance $x \in I_Q$, i.e. a feasible solution OPT_Q of x such that $c_Q(OPT_Q) = type\{c_Q(y) : y \in sol_Q(x)\}$. *NPO* stands for the class of all NP-optimization problems.

Let Q be an NP- optimization problem. Given an instance $x \in I_Q$ and a feasible solution $y \in sol_Q(x)$, we can define the ratio, $R_Q(x, y) = \max\left\{\frac{c_Q(y)}{c_Q(OPT_Q)}, \frac{c_Q(OPT_Q)}{c_Q(y)}\right\}$.

APX: Let Q be an NP-optimization problem. Let A be an algorithm such that $A \in sol_Q$. For a rational number $r > 1$, A is said to be an r*-approximation algorithm* for Q if for any instance x of Q, $A(x)$ can be computed in polynomial time and the ratio, $R_Q(x, A(x)) \leq r$. Moreover, an NPO-problem Q is said to belongs to the class APX if there exists an r-approximation algorithm for Q for some rational $r > 1$.

PTAS: An NP-optimization problem Q is said to belongs to the class $PTAS$ if for each rational $r > 1$, there exists an algorithm A such that A is an r-approximation algorithm for Q.

Clearly, $PTAS \subseteq APX \subseteq NPO$. Note that the inclusions are strict if and only if $P \neq NP$ [28]. The reducibility $\leqslant$ is said to *preserves membership* in a class $\mathcal{C}$, if $A \leqslant B$ and $B \in \mathcal{C}$ implies that $A \in \mathcal{C}$. Moreover, the reducibility $\leqslant$ is said to be *approximation preserving* if it preserves membership in either APX, PTAS or both. There are various types of approximation preserving reductions available in the literature (see [28]).

An approximation preserving reduction, called *L-reduction* [97] can be defined as follows.

Definition 30 (L-reduction)**.** *Let P_1 and P_2 be two NP-optimization problems with cost functions c_{P_1} and c_{P_2} respectively. Let f be a polynomially computable function that maps the instances of problem P_1 to the instances of problem P_2. Then f is said to be an L-reduction from P_1 to P_2 if there exist a polynomially computable function g and constants $\alpha, \beta \in (0, \infty)$ such that the following conditions hold:*

(*a*) *If $y \in sol_{P_2}(f(x))$ then $g(y) \in sol_{P_1}(x)$, where $x \in I_{P_1}$.*

(*b*) *$OPT_{P_2}(f(x)) \leq \alpha OPT_{P_1}(x)$, where $OPT_{P_2}(f(x))$ and $OPT_{P_1}(x)$ denote the optimum value of respective instances for the problems P_2 and P_1 respectively.*

(*c*) *$|OPT_{P_1}(x) - c_{P_1}(g(y))| \leq \beta |OPT_{P_2}(f(x)) - c_{P_2}(y)|$, for every instance x of P_1 and $y \in sol_{P_2}(f(x))$.*

Note that for any instance x of P_1, y denote a solution of $f(x)$ that is produced by the approximation algorithm for the problem P_2 when run on the instance $f(x)$.

Let Q be an NP-optimization problem. Then Q is said to have a polynomial-time approximation algorithm with *worst case error* ϵ if for any instance $x \in I_Q$ and a feasible solution $y \in sol_Q(x)$, we have $\frac{|c_Q(y)-OPT_Q(x)|}{OPT_Q(x)} \leq \epsilon$.

The following proposition was proved in [97].

Proposition 2 ([97]). *Let P_1 and P_2 be two minimization (resp. maximization) problems such that P_1 is L-reducible to P_2. Let α and β be the constants as given in Definition 30. If there is a polynomial-time approximation algorithm for the problem P_2 with worst case error ϵ then there is a polynomial-time approximation algorithm for the problem P_1 with worst case error $\alpha\beta\epsilon$. Consequently, an L-reduction preserves membership in PTAS.*

Proof. Suppose that P_1 and P_2 are two minimization problems. Since P_1 is L-reducible to P_2, by Definition 30 we have that there exist polynomially computable functions f, g and constants $\alpha, \beta \in (0, \infty)$ such that the following conditions (a) and (b) hold for every instance x of P_1 and $y \in sol_{P_2}(f(x))$:

(a) $OPT_{P_2}(f(x)) \leq \alpha OPT_{P_1}(x)$,

(b) $c_{P_1}(g(y)) \leq OPT_{P_1}(x) + \beta[c_{P_2}(y) - OPT_{P_2}(f(x))]$.

Since P_2 is a minimization problem that has an approximation algorithm with worst case error ϵ, we have $\frac{c_{P_2}(y)-OPT_{P_2}(f(x))}{OPT_{P_2}(f(x))} \leq \epsilon$, i.e.

(c) $c_{P_2}(y) \leq (1+\epsilon)OPT_{P_2}(f(x))$, for any $y \in sol_{P_2}(f(x))$.

As in Definition 30, also note that if $y \in sol_{P_2}(f(x))$ then $g(y) \in sol_{P_1}(x)$, where $x \in I_{P_1}$.

Therefore, for any $x \in I_{P_1}$ and $g(y) \in sol_{P_1}(x)$, we have,

$$\begin{aligned}
c_{P_1}(g(y)) &\leq OPT_{P_1}(x) + \beta[c_{P_2}(y) - OPT_{P_2}(f(x))] && (\text{ by } (b)) \\
&\leq OPT_{P_1}(x) + \beta[(1+\epsilon)OPT_{P_2}(f(x)) - OPT_{P_2}(f(x))] && (\text{ applying } (c)) \\
&\leq OPT_{P_1}(x) + \beta\epsilon\alpha OPT_{P_1}(x) && (\text{ applying } (a)) \\
&= (1+\alpha\beta\epsilon)OPT_{P_1}(x).
\end{aligned}$$

This implies that P_1 has a $(1+\alpha\beta\epsilon)$ approximation algorithm.

A similar proof holds for the case in which P_1 and P_2 are two maximization problems. This proves the proposition. □

A problem Q is said to be *APX-hard* if there is a PTAS preserving reduction from every other problem in APX to Q. If a problem Q is APX-hard as well as belongs to the class APX, then we say that the problem Q is *APX-complete*. Note that $PTAS \neq APX$, if $P \neq NP$ [28]. Thus we have that no APX-hard problem has a PTAS, if $P \neq NP$. Therefore, in order to prove that an NPO-problem Q is APX-hard, it is enough to show that the problem Q has a PTAS preserving reduction from an APX-hard problem. In particular, in this work, in order to prove that an NPO-problem Q is APX-hard, we show that the problem Q has an L-reduction from an APX-hard problem (this is valid by Proposition 2).

5.2 Characterizations for point-point digraphs

Let $G = (V, E)$ be a digraph. We say that a, b, c, d is an *anti-directed walk* of length 3 if $a, b, c, d \in V(G)$, $(a, b), (c, b), (c, d) \in E(G)$ and $(a, d) \notin E(G)$ (the vertices a, b, c, d need not be pairwise distinct, but it follows from the definition that $a \neq c$ and $b \neq d$). Refer to Figure 5.1.

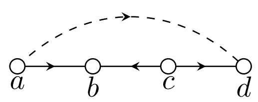


Figure 5.1: An anti-directed walk of length 3 (possibly, some of the vertices may coincide)

Recall that given a digraph G, $B_G = (X, Y, E)$ is a splitting bigraph of G where $X = \{x_u : u \in V(G)\}$ and $Y = \{y_u : u \in V(G)\}$ and $x_u y_v \in E(B_G)$ if and only if $(u, v) \in E(G)$. We then have the following theorem.

Theorem 21. *Let G be a digraph. Then the following conditions are equivalent:*

(a) *G is a point-point digraph.*

(b) *G does not contain any anti-directed walk of length 3.*

(c) *The splitting bigraph of G is a disjoint union of complete bipartite graphs.*

Proof. $(a) \Rightarrow (b)$: Let G be a point-point digraph with a point-point representation $\{(S_u, T_u)\}_{u \in V(G)}$. Suppose that there exist vertices a, b, c, d in G such that $(a, b), (c, b), (c, d) \in E(G)$. By the definition of point-point representation, we then have $S_a = T_b = S_c = T_d$. This implies that $(a, d) \in E(G)$. Therefore we can conclude that G does not contain any anti-directed walk of length 3.

$(b) \Rightarrow (c)$: Suppose that G does not contain any anti-directed walk of length 3. Let $B_G = (X, Y, E)$ be the splitting bigraph of G. Let $x_u y_v$ be any edge in B_G, where $u, v \in V(G)$. Clearly, by the definition of B_G, $(u, v) \in E(G)$. We claim that the graph induced in B_G by the vertices

$N(x_u) \cup N(y_v)$ is a complete bipartite graph. Suppose not. Then it should be the case that there exist two vertices $x_a \in N(y_v)$ and $y_b \in N(x_u)$ such that $x_a y_b \notin E(B_G)$, where $a, b \in V(G)$. By the definition of B_G, we then have that $(a, v), (u, v), (u, b) \in E(G)$ and $(a, b) \notin E(G)$. So a, v, u, b is an anti-directed walk of length 3 in G, which is a contradiction to (b). This proves that for every $p \in X$ and $q \in Y$ such that $pq \in E(B_G)$, the set $N(p) \cup N(q)$ induces a complete bipartite subgraph in B_G. Therefore, each connected component of B_G is a complete bipartite graph. (This can be seen as follows: Suppose that there is a connected component C of B_G that is not complete bipartite. Choose $p \in X \cap C$ and $q \in Y \cap C$ such that $pq \notin E(B_G)$ and the distance between p and q in B_G is as small as possible. Let t be the distance between p and q in B_G. Clearly, t is odd and $t \geq 3$. Consider a shortest path $p = z_0, z_1, z_2, \ldots, z_t = q$ from p to q in B_G. By our choice of p and q, we have that $z_1 z_{t-1} \in E(B_G)$. But then $p \in N(z_1)$, $q \in N(z_{t-1})$ and $pq \notin E(B_G)$, contradicting our observation that $N(z_1) \cup N(z_{t-1})$ induces a complete bipartite graph in B_G.)

$(c) \Rightarrow (a)$: Suppose that G is a digraph such that the splitting bigraph B_G is a disjoint union of complete bipartite graphs, say $H_1, H_2, \ldots, H_k$. Now we can obtain a point-point representation $\{(S_u, T_u)\}_{u \in V(G)}$ of the digraph G as follows: For each $i \in \{1, 2, \ldots, k\}$, define $S_u = i$ if $x_u \in V(H_i)$ and $T_v = i$ if $y_v \in V(H_i)$. Note that $(u, v) \in E(G)$ if and only if $x_u y_v \in E(B_G)$ if and only if $x_u, y_v \in V(H_i)$ for some $i \in \{1, 2, \ldots, k\}$. Therefore we can conclude that $(u, v) \in E(G)$ if and only if $S_u = T_v = i$ for some $i \in \{1, 2, \ldots, k\}$. Thus the digraph G is a point-point digraph. □

Using the equivalences of (a) and (c) we have the following corollary.

Corollary 9. *Point-point digraphs can be recognized in linear time.*

5.3 Hardness results for point-point digraphs

5.3.1 Subdivision of an irreflexive digraph

For an undirected graph G, the *k-subdivision* of G, where $k \geq 1$, is defined as the graph H having vertex set $V(H) = V(G) \cup \bigcup_{ij \in E(G)} \{u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^k\}$, obtained from G by replacing each edge $ij \in E(G)$ by a path $i, u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^k, j$.

The following theorem is adapted from Theorem 5 of Chlebík and Chlebíková [19].

Theorem 22 (Chlebík and Chlebíková)**.** *Let G be an undirected graph having m edges. Let $k \geq 1$.*

(a) *The problem of computing a maximum cardinality independent set is APX-complete when restricted to* $2k$*-subdivisions of 3-regular graphs for any fixed integer* $k \geq 0$.

(b) *The problem of finding a minimum cardinality dominating set (resp. independent dominating set) is APX-complete when restricted to* $3k$*-subdivisions of graphs having degree at most 3 for any fixed integer* $k \geq 0$.

Note that the independent sets, dominating sets and independent dominating sets of an undirected graph G are exactly the independent sets, dominating sets (which are also the absorbing sets), and solutions (which are also the kernels) of the symmetric digraph of G. Clearly the symmetric digraph of G is irreflexive. Since the MAX-KERNEL problem is equivalent to the INDEPENDENT-SET problem in symmetric digraphs, we then have the following corollary of Theorem 22.

Corollary 10. *The problems* INDEPENDENT-SET*,* ABSORBING-SET*,* MIN-KERNEL *and* MAX-KERNEL *problems are APX-complete on irreflexive symmetric digraphs of in- and out-degree at most 3.*

Suppose that $k \geq 0$. Let H be the $2k$-subdivision or $3k$-subdivision of an undirected graph and let G be the symmetric digraph of H. Note that the independent sets, dominating sets, and independent dominating sets of H are exactly the independent sets, dominating sets (which are also the absorbing sets), and solutions (which are also the kernels) of G. Therefore from Theorem 22 we have that the INDEPENDENT-SET problem is APX-hard on irreflexive symmetric digraphs of $2k$-subdivisions of 3-regular graphs, and that the ABSORBING-SET and MIN-KERNEL problems are APX-hard on the symmetric digraphs of $3k$-subdivisions of graphs of degree at most 3 for each $k \geq 0$. But note that for $k \geq 1$, the symmetric digraph of the $2k$-subdivision or $3k$-subdivision of an undirected graph contains an anti-directed walk of length 3 (unless the graph contains no edges), and therefore by Theorem 21, is not a point-point digraph. Thus we cannot directly deduce the APX-hardness of the problems under consideration for point-point digraphs from Theorem 22.

We define the subdivision of an irreflexive digraph, so that the techniques of Chlebík and Chlebíková can be adapted for proving hardness results on point-point digraphs.

Definition 31 (k-subdivision)**.** *Let G be an irreflexive digraph (i.e. G contains no loops). For $k \geq 1$, define the* k-subdivision *of G to be the digraph H having vertex set $V(H) = V(G) \cup \bigcup_{(i,j) \in E(G)} \{u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^k\}$, obtained from G by replacing each edge $(i,j) \in E(G)$ by a directed path $i, u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^k, j$.*

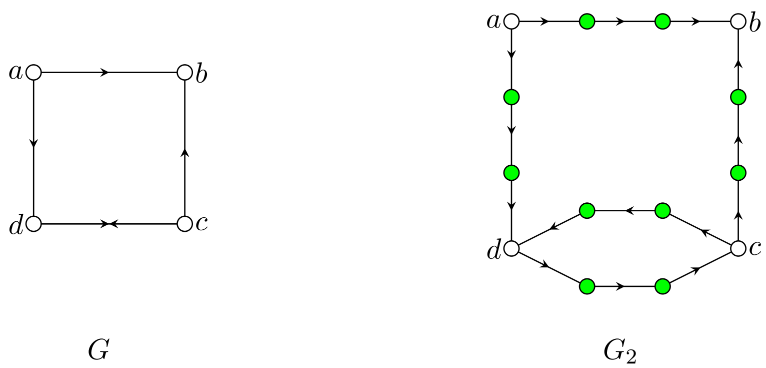


Figure 5.2: An irreflexive digraph and its 2-subdivision

Figure 5.2 provides an example of a 2-subdivision of an irreflexive digraph G. The additional vertices in G_2 that are not present in G are shaded green. Note that the k-subdivision of any irreflexive digraph is also an irreflexive digraph. We then have the following lemma.

Lemma 14. *For any $k \geq 1$, the k-subdivision of any irreflexive digraph is a point-point digraph.*

Proof. Let $k \geq 1$ and let G be any irreflexive digraph. By Theorem 21, it is enough to show that the k-subdivision H of G does not contain any anti-directed walk of length 3. Note that by the definition of k-subdivision, all the vertices in $V(H) \setminus V(G)$ have both in-degree and out-degree exactly equal to one. Further, for every vertex v in H such that $v \in V(G)$, we have that $N^+(v), N^-(v) \subseteq V(H) \setminus V(G)$. Suppose for the sake of contradiction that u, v, w, x is an anti-directed walk of length 3 in H. Recall that we then have $(u, v), (w, v), (w, x) \in E(H)$, $u \neq w$ and $v \neq x$. By the above observations, we can then conclude that $v \in V(G)$ and further that $u, w \in V(H) \setminus V(G)$. Then since $(w, x) \in E(H)$ and $v \neq x$, we have that w has out-degree at least 2, which contradicts our earlier observation that every vertex in $V(H) \setminus V(G)$ has out-degree exactly one. This proves the lemma. $\square$

Theorem 23. *The problem* Independent-Set *is APX-hard for point-point digraphs having maximum degree at most 3.*

Proof. We show a reduction from the Independent-Set problem in 2-subdivisions of 3-regular undirected graphs (which is APX-hard by Theorem 22(a)). Let G be a 3-regular undirected graph and let H be its 2-subdivision. Let D be the digraph obtained by assigning an arbitrary direction to each edge of G. Clearly, D is irreflexive. Let D' be a 2-subdivision of the directed graph D. Note that the underlying undirected graph of D' is H. It is clear that given H, the graph D' can be constructed in polynomial time. By Lemma 14, D' is a point-point digraph. Since the independent sets of H are exactly the independent sets of D', and D' has maximum degree at most 3, we can conclude from Theorem 22(a) that the problem Independent-Set is APX-hard for point-point digraphs having maximum degree at most 3. □

5.3.2 Kernel

Lemma 15. *Let G be an irreflexive digraph and let $k \geq 1$. Then G has a kernel if and only if the $2k$-subdivision of G has a kernel. Moreover, G has a kernel of size q if and only if the $2k$-subdivision of G has a kernel of size $q + km$. Further, given a kernel of size $q + km$ of the $2k$-subdivision of G, we can construct a kernel of size q of G in polynomial time.*

Proof. Let H be the $2k$-subdivision of G and let $\bigcup_{(i,j)\in E(G)}\{u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^{2k}\}$ be the vertices in $V(H) \setminus V(G)$ as defined in Section 5.3.1.

Suppose that G has a kernel $K \subseteq V(G)$. We define the set $K' \subseteq V(H)$ as $K' = K \cup \bigcup_{(i,j)\in E(G)} S(i,j)$, where

$$S(i,j) = \begin{cases} \{u_{ij}^{2l} : l \in \{1,2,\ldots,k\}\}, & \text{if } j \notin K \\ \{u_{ij}^{2l-1} : l \in \{1,2,\ldots,k\}\}, & \text{if } j \in K \end{cases}$$

We claim that K' is a kernel in H. Note that as K is an independent set in G, for any edge $(i,j) \in E(G)$, we have that $i \notin K$ whenever $j \in K$. Thus by the definition of $2k$-subdivision and K', it is easy to see that K' is an independent set in H. Therefore in order to prove our claim, it is enough to show that K' is an absorbing set in H. Consider any $(i,j) \in E(G)$. It is clear from the definition of K' that for each $t \in \{1,2,\ldots,2k-1\}$, either the vertex u_{ij}^t or u_{ij}^{t+1} is in K'. Further, we also have that either the vertex u_{ij}^{2k} or j is in K'. Thus for every vertex $x \in V(H) \setminus V(G)$, either x or one of its out-neighbors is in K'. Now consider a vertex i in $V(H)$ such that $i \in V(G)$. If $i \in K$, then $i \in K'$. On the other hand, if $i \notin K$, then since K is a kernel of G, there exists an out-neighbor j of i such that $j \in K$, in which case we have $u_{ij}^1 \in K'$. Thus in any case, either i or an out-neighbor of i is in K'. This shows that K' is a kernel of H.

Note that by the definition of K', we have $|K' \setminus K| = km$. Therefore if $|K| = q$ then $|K'| = q + km$.

Now suppose that $K' \subseteq V(H)$ is a kernel in H.

Claim 1. Let $(i,j) \in E(G)$ and $t \in \{1,2,\ldots,2k-1\}$. Then $u_{ij}^t \in K'$ if and only if $u_{ij}^{t+1} \notin K'$.

If $u_{ij}^t \in K'$, then since K' is an independent set in H, we have $u_{ij}^{t+1} \notin K'$. On the other hand, if $u_{ij}^t \notin K'$, then since K' is an absorbing set in H, we have $u_{ij}^{t+1} \in K'$. This proves the claim.

We first show that $K' \cap V(G)$ is an independent set of G. Consider any edge $(i, j) \in E(G)$. Suppose that $i \in K'$. Then since K' is an independent set in H, we have $u_{ij}^1 \notin K'$. Applying Claim 1 repeatedly, we have that $u_{ij}^{2k} \in K'$, which implies that $j \notin K'$. Thus, the set $K' \cap V(G)$ is an independent set in G. Next, we note that $K' \cap V(G)$ is also an absorbing set of G. To see this, consider any vertex i of H such that $i \in V(G)$. If $i \notin K'$, then since K' is an absorbing set in H, there exists $(i, j) \in E(G)$ such that $u_{ij}^1 \in K'$. Applying Claim 1 repeatedly, we have that $u_{ij}^{2k} \notin K'$. Then since K' is an absorbing set in H, we have that $j \in K'$. Thus $K' \cap V(G)$ is an absorbing set of G, which implies that $K' \cap V(G)$ is a kernel of G.

Note that by Claim 1, we have that $|K' \setminus V(G)| \leq km$. Let $(i, j) \in E(G)$. Since K' is an absorbing set in H, for each $t \in \{1, 2, \ldots, 2k - 1\}$, either $u_{ij}^t \in K'$ or $u_{ij}^{t+1} \in K'$. This implies that $|K' \cap \{u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^{2k}\}| \geq k$. This further implies that $|K' \setminus V(G)| \geq km$. Therefore we can conclude that $|K' \setminus V(G)| = km$. Thus, if $|K'| = q + km$ then $|K' \cap V(G)| = q$. Clearly, given the kernel K' of H, the kernel $K' \cap V(G)$ of G can be constructed in polynomial time. □

Theorem 24. *The problem* Kernel *is NP-complete for point-point digraphs.*

Proof. We show a reduction from the Kernel problem in general digraphs to the Kernel problem in point-point digraphs. Let G be any digraph. Let G' be the digraph obtained from G by removing all loops in it. Then note that the kernels in G and G' are exactly the same. Let H be a 2-subdivision of G. Since G' is an irreflexive digraph, by Lemma 15 we have that G' has a kernel if and only if H has a kernel. Also, we have by Lemma 14 that H is a point-point digraph. Therefore we can conclude that G has a kernel if and only if the point-point digraph H has a kernel. Thus a polynomial-time algorithm that solves the Kernel problem in point-point digraphs can be used to solve the Kernel problem in general digraphs in polynomial time. This proves the theorem. □

Note that Kernel is known to be NP-complete even on planar digraphs having degree at most 3 and in- and out-degrees at most 2 [44]. The above reduction transforms the input digraph in such a way that every newly introduced vertex has in- and out-degree exactly 1 and the in- and out-degrees of the original vertices remain the same. Moreover, if the input digraph is planar, the digraph produced by the reduction is also planar. Thus we can conclude that the problem Kernel remains NP-complete even for planar point-point digraphs having degree at most 3 and in- and out-degrees at most 2.

As we have noted in Section 5.1, in order to prove that a problem Q is APX-hard, it is enough to show that the problem Q has an L-reduction from an APX-hard problem.

Theorem 25. *For $k \geq 1$, the problems* Min-Kernel *and* Max-Kernel *are APX-hard for $2k$-subdivisions of irreflexive symmetric digraphs having in- and out-degree at most 3. Consequently, the problems* Min-Kernel *and* Max-Kernel *are APX-hard for point-point digraphs having in- and out-degree at most 3.*

Proof. By Corollary 10, we have that the problems Min-Kernel and Max-Kernel are APX-complete for irreflexive symmetric digraphs having in- and out-degree at most 3. Here we give an L-reduction from the Min-Kernel and Max-Kernel problems for irreflexive symmetric digraphs having in- and out-degree at most 3 to the Min-Kernel and Max-Kernel problems for $2k$-subdivisions of irreflexive symmetric digraphs having in- and out-degree at most 3. Let G be an irreflexive symmetric digraph of in- and out-degree at most 3, where $|V(G)| = n$ and $|E(G)| = m$. For $k \geq 1$, let H be the $2k$-subdivision of G. Clearly, H can be constructed in polynomial time. And let $K(G)$ (resp. $K'(G)$) and $K(H)$ (resp. $K'(H)$) denote a minimum (resp. maximum) sized kernel in G and H respectively. Since G is a digraph of in- and out-degree at most 3, we have that $m \leq 3n$. Note that every absorbing set of G has size at least $\frac{n}{4}$, since each vertex has at most 3 in-neighbors. As a minimum (resp. maximum) kernel of G is an absorbing set of G, we have $|K(G)| = q \geq \frac{n}{4}$ (resp. $|K'(G)| = q' \geq \frac{n}{4}$). By Lemma 15, we have that $|K(H)| = q + km$ (resp. $K'(H) = q' + km$). Therefore, $\frac{|K(H)|}{|K(G)|} \leq 1 + 12k$ (resp. $\frac{|K(H')|}{|K(G')|} \leq 1 + 12k$). We can now choose $\alpha = 1 + 12k$ and $\beta = 1$ so that our reduction satisfies the requirements of Definition 30 (Lemma 15 guarantees that condition (c) of Definition 30 holds, and also that the function g in the definition is polynomial-time computable). Thus our reduction is an L-reduction, which implies that the problems Min-Kernel and Max-Kernel are APX-hard for $2k$-subdivisions of irreflexive symmetric digraphs having in- and out-degree at most 3. Now by Lemma 14, we have that the $2k$-subdivision of any irreflexive digraph G is a point-point digraph. Therefore, now we can conclude that the problems Min-Kernel and Max-Kernel are APX-hard for point-point digraphs. □

5.3.3 Minimum absorbing set

Lemma 16. *Let G be an irreflexive digraph and let $k \geq 1$. Then G has an absorbing set of size at most q if and only if the $2k$-subdivision of G has an absorbing set of size at most $q + km$. Further, given an absorbing set of size at most $q + km$ in the $2k$-subdivision of G, we can construct in polynomial time an absorbing set of size at most q in G.*

Proof. Let H be the $2k$-subdivision of G and let $\bigcup_{(i,j)\in E(G)} \{u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^{2k}\}$ be the vertices in $V(H) \setminus V(G)$ as defined in Section 5.3.1.

Suppose that G has an absorbing set $A \subseteq V(G)$ such that $|A| \leq q$. We define the set $A' \subseteq V(H)$ as $A' = A \cup \bigcup_{(i,j)\in E(G)} A(i,j)$, where

$$A(i,j) = \begin{cases} \{u_{ij}^{2l} : l \in \{1,2,\ldots,k\}\}, & \text{if } j \notin A \\ \{u_{ij}^{2l-1} : l \in \{1,2,\ldots,k\}\}, & \text{if } j \in A \end{cases}$$

We claim that A' is an absorbing set in H of size at most $q + km$. Consider any $(i,j) \in E(G)$. It is clear from the definition of A' that for each $t \in \{1,2,\ldots,2k-1\}$, either the vertex u_{ij}^t or u_{ij}^{t+1} is in A'. Further, we also have that either the vertex u_{ij}^{2k} or j is in A'. Thus for every vertex $x \in V(H) \setminus V(G)$, either x or one of its out-neighbors is in A'. Now consider a vertex i in H such that $i \in V(G)$. If $i \in A$, then $i \in A'$. On the other hand, if $i \notin A$, then since A is an absorbing set in G, there exists an out-neighbor j of i such that $j \in A$, in which case we have $u_{ij}^1 \in A'$. Thus in any case, either i or an out-neighbor of i is in A'. This shows that A' is an absorbing set in H. As A' is obtained from A by adding exactly k new vertices corresponding to each of the m edges in G, we also have that $|A'| \leq q + km$. This proves our claim.

For any set $S \subseteq V(H)$ and $(i,j) \in E(G)$, we define $S_{ij} = S \cap \{u_{ij}^1, u_{ij}^2, \ldots, u_{ij}^{2k-1}, u_{ij}^{2k}\}$. Now suppose that H has an absorbing set A' of size at most $q + km$. Let $F = \{(i,j) \in E(G) \colon |A'_{ij}| > k\}$. Now define the set $A'' = (A' \setminus \bigcup_{(i,j)\in F} A'_{ij}) \cup \bigcup_{(i,j)\in F}(\{u_{ij}^{2l-1} : l \in \{1,2,\ldots,k\}\} \cup \{j\})$. Clearly, A'' is also an absorbing set in H, $|A''| \leq |A'| \leq q + km$. Since A'' is an absorbing set in H, for $(i,j) \in E(G)$ and each $t \in \{1,2,\ldots,2k-1\}$, either $u_{ij}^t \in A''$ or $u_{ij}^{t+1} \in A''$. This implies that $|A''_{ij}| \geq k$. From the construction of A'', it is clear that for each $(i,j) \in E(G)$, $|A''_{ij}| \leq k$. Therefore, we can conclude that $|A''_{ij}| = k$ for each $(i,j) \in E(G)$. It then follows that for each $t \in \{1,2,\ldots,2k-1\}$, exactly one of u_{ij}^t, u_{ij}^{t+1} is in A''. We now claim that $A = A'' \cap V(G)$ is an absorbing set in G. Let $i \in V(G)$. Suppose that $i \notin A$, which means that $i \notin A''$. Since A'' is an absorbing set in H, we have that there exists a vertex $j \in N_G^+(i)$ such that $u_{ij}^1 \in A''$. By our earlier observation that exactly one of $u_{ij}^t, u_{ij}^{t+1} \in A''$ for each $t \in \{1,2,\ldots,2k-1\}$, we now have that $u_{ij}^{2k} \notin A''$. This would imply that $j \in A''$. Therefore we can conclude that for any vertex $i \in V(G)$, either $i \in A$ or one of its out-neighbors is in A. This implies that A is an absorbing set in G. Since $|A''_{ij}| = k$ for each $(i,j) \in E(G)$ and $|E(G)| = m$, we now have that $|A| = |A''| - km \leq q$. It is also easy to see that given the absorbing set A' of H, we can construct A'' and then $A'' \cap V(G)$ in polynomial time. This proves the lemma. □

Theorem 26. *For $k \geq 1$, the problem* Absorbing-Set *is APX-hard for $2k$-subdivisions of irreflexive symmetric digraphs having in- and out-degree at most 3. Consequently, the problem*

ABSORBING-SET *is APX-hard for point-point digraphs having in- and out-degree at most 3.*

Proof. This can be proved in a way similar to that of Theorem 25. By Corollary 10, we have that the ABSORBING-SET problem is APX-complete for irreflexive symmetric digraphs having in- and out-degree at most 3. We give an L-reduction from the ABSORBING-SET problem for irreflexive symmetric digraphs having in- and out-degree at most 3 to the ABSORBING-SET problem for $2k$-subdivisions of irreflexive symmetric digraphs having in and out-degree at most 3. Let G be an irreflexive symmetric digraph of in- and out-degree at most 3, where $|V(G)| = n$ and $|E(G)| = m$. For $k \geq 1$, let H be the $2k$-subdivision of G. Clearly, H can be constructed in polynomial time. And let $A(G)$ and $A(H)$ denote a minimum sized absorbing set in G and H respectively. Since G is a digraph of in- and out-degree at most 3, as noted in the proof of Theorem 25, we have that $m \leq 3n$ and $|A(G)| \geq \frac{n}{4}$. By Lemma 16, we have that $|A(H)| \leq |A(G)| + km$. Therefore as $\frac{|A(H)|}{|A(G)|} \leq 1 + 12k$, we can now choose $\alpha = 1 + 12k$ and $\beta = 1$ so that our reduction satisfies the requirements of Definition 30 (Lemma 16 guarantees that condition (c) of Definition 30 holds, and also that the function g in the definition is polynomial-time computable). Thus our reduction is an L-reduction, which implies that ABSORBING-SET is APX-hard for $2k$-subdivisions of irreflexive symmetric digraphs having in- and out-degree at most 3. Since the $2k$-subdivision of any irreflexive digraph G is a point-point digraph by Lemma 14, we can now conclude that the problem ABSORBING-SET is APX-hard for point-point digraphs.

□

5.4 Comparability relations between classes of digraphs

Figure 5.3 shows the inclusion relations between the classes of digraphs that we studied for the problems, KERNEL, MIN-KERNEL, MAX-KERNEL, ABSORBING-SET, DOMINATING-SET, and INDEPENDENT-SET.

Note that the class of interval digraphs and the class of DUF-digraphs are incomparable to each other. This can be shown as follows: a directed triangle with edges $(a, b), (b, c), (c, a)$ is a point-point digraph (refer to Figure 5.4), but it is easy to see that there is no DUF-ordering for this digraph. Thus, the class of point-point digraphs is not contained in the class of DUF-digraphs. On the other hand, consider a symmetric triangle G as shown in Figure 5.5. Then any permutation of the vertices in G is a DUF-ordering of G. Note that the splitting bigraph B_G of G is an induced cycle of length 6. If G is an interval digraph, then B_G is an interval bigraph, which contradicts Müller's observation [95] that interval bigraphs are chordal bipartite

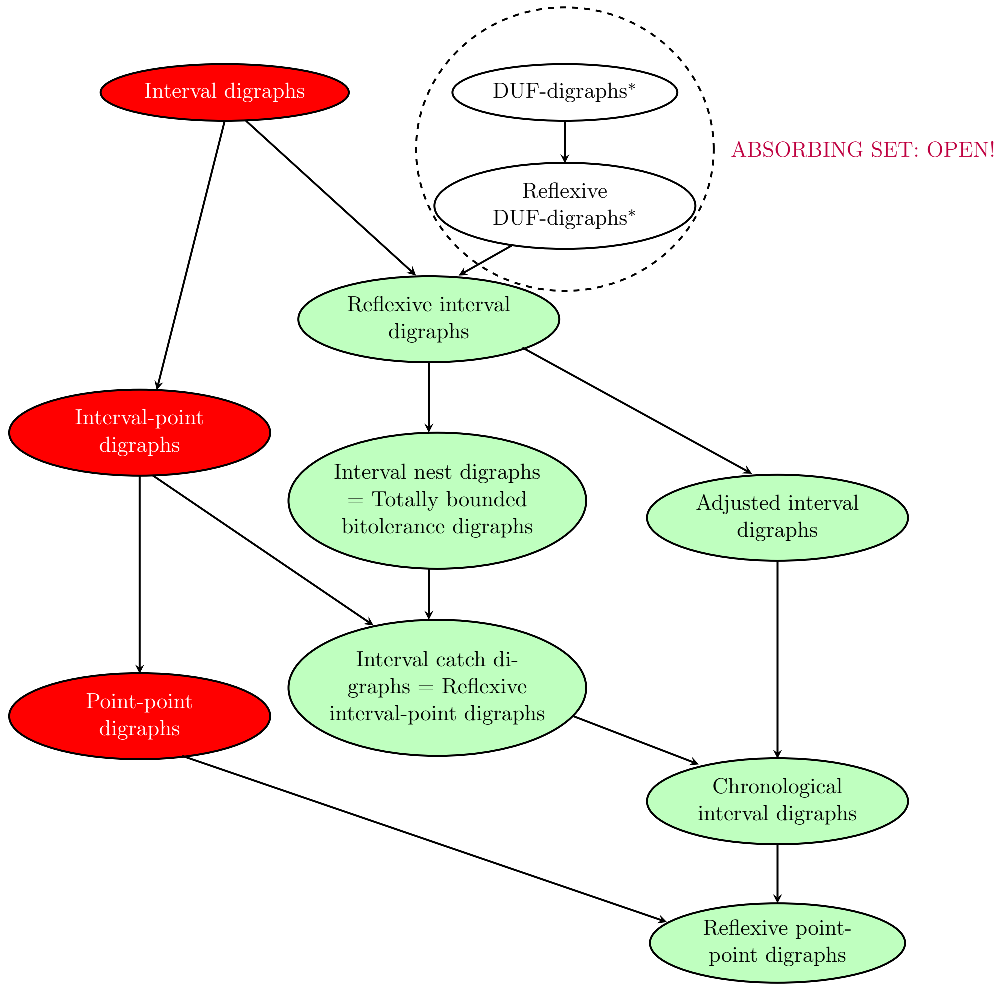


Figure 5.3: Inclusion relations between graph classes. In the diagram, there is an arrow from $\mathcal{A}$ to $\mathcal{B}$ if and only if the class $\mathcal{B}$ is contained in the class $\mathcal{A}$. Moreover, each inclusion is strict. The problems studied are efficiently solvable in the classes shown in light green, while they are NP-hard and/or APX-hard in the classes shown in dark red (* the complexity of the ABSORBING-SET problem on DUF-digraphs and reflexive DUF-digraphs remain open).

graphs (bipartite graphs that do not contain any induced cycle C_k, for $k \geq 6$). Thus G is not an interval digraph, implying that the class of DUF-digraphs is not contained in the class of interval digraphs. Further note that, even the class of reflexive DUF-digraphs is not contained in the class of interval digraphs, as otherwise every reflexive DUF-digraph should have been a reflexive interval digraph, which is not true: by Theorem 14, the underlying undirected graph of a reflexive interval digraph cannot contain $K_{3,3}$ as an induced subgraph, but orienting every edge of a $K_{3,3}$ from one partite set to the other and adding a self-loop at each vertex gives a reflexive DUF-digraph (any ordering of the vertices in which the vertices in one partite set all come before every vertex in the other partite set is a DUF-ordering of this digraph). Clearly, there are DUF-digraphs that are not reflexive, implying that the class of reflexive DUF-digraphs

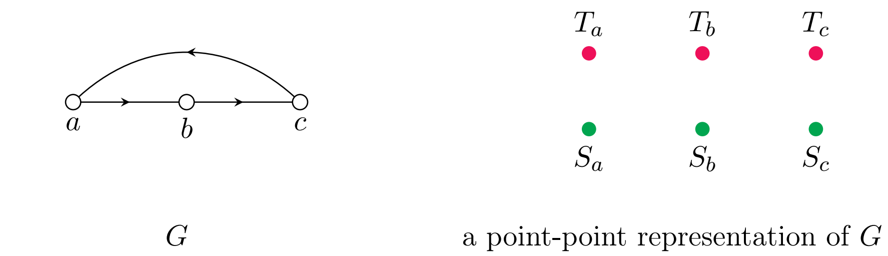


Figure 5.4: A point-point digraph that is not a DUF-digraph

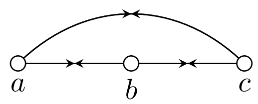


Figure 5.5: A DUF-digraph that is not an interval digraph

forms a strict subclass of DUF-digraphs.

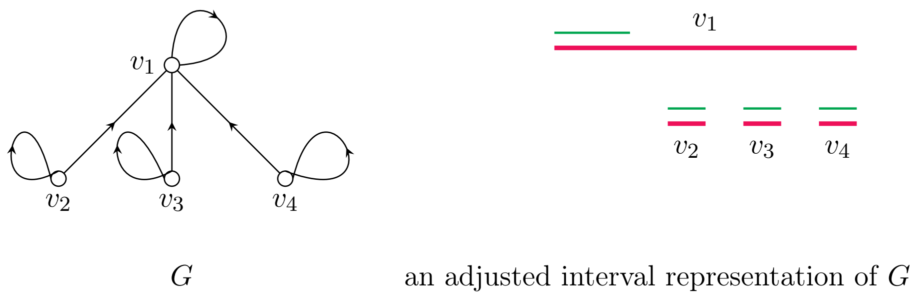
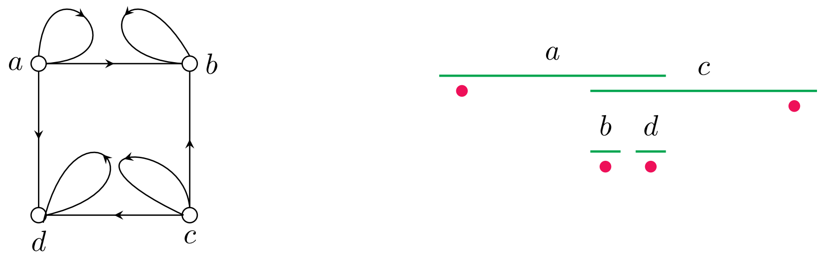


Figure 5.6: An adjusted interval digraph that is not an interval nest digraph

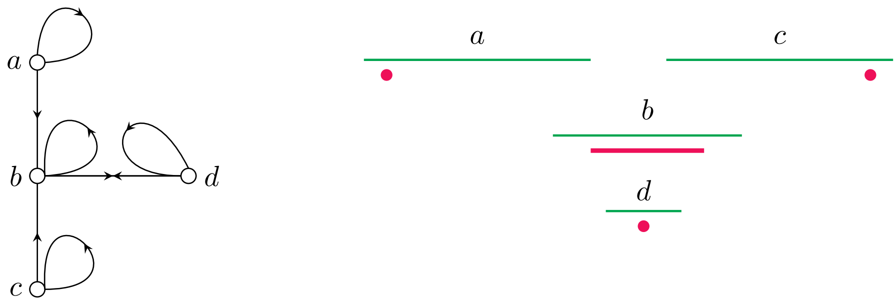
It is easy to see that the class of point-point digraphs is not contained in the class of reflexive interval digraphs as point-point digraphs contain digraphs that are not reflexive. Now in [32], the authors give an example of a digraph which is not an interval point digraph as follows: The digraph has vertex set $\{v_1, v_2, v_3, v_4\}$ and edge set $\{(v_2, v_2), (v_3, v_3), (v_4, v_4), (v_2, v_1), (v_3, v_1), (v_4, v_1)\}$. They observed that this digraph is not an interval point digraph. We slightly modify the above example by adding a loop at v_1 and call the resulting reflexive digraph as G (refer to Figure 5.6). It is then easy to verify that the modified digraph G is not an interval nest digraph (Note that in any interval nest representation of G, there exists $x \in \{v_2, v_3, v_4\}$ such that $S_x \subseteq S_{v_1} \cup \bigcup_{a \in \{v_2, v_3, v_4\} \setminus \{x\}} S_a$. As $T_x \subseteq S_x$, this implies that either $(v_1, x) \in E(G)$ or there exists an $a \in \{v_2, v_3, v_4\} \setminus \{x\}$ such that $(a, x) \in E(G)$, which is a contradiction to the definition of G.) But Figure 5.6 gives an adjusted interval representation of G, where thin green and bold red intervals respectively denote the source and destination intervals corresponding to the vertices in G. This shows that G as defined above is an adjusted interval digraph. Since G is not an interval nest digraph, we can conclude that the class of adjusted interval digraphs (and therefore, the class of reflexive interval digraphs) is not contained in the class of interval nest digraphs (and therefore, not contained in the class of interval catch digraphs). Since interval catch digraphs are exactly reflexive interval-point digraphs, this also means that the class of adjusted interval digraphs (and therefore, the class of reflexive interval digraphs) is not contained in the class of interval-point digraphs.



G an interval catch representation of G

Figure 5.7: An interval catch digraph that is not an adjusted interval digraph

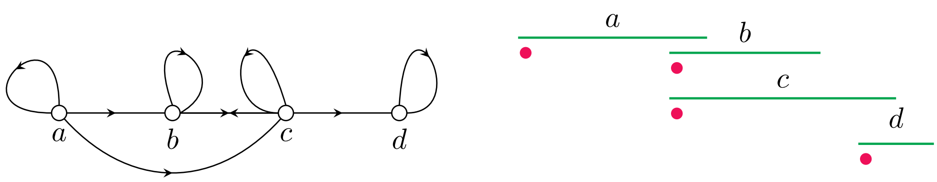
Now consider the digraph G with $V(G) = \{a, b, c, d\}$ and edges $(a, b), (a, d), (c, b), (c, d)$ in addition to loops at each vertex. Figure 5.7 gives an interval catch representation of G, where thin green intervals and dark red points respectively denote the source and destination intervals corresponding to the vertices in G. But note that the underlying undirected graph of G is an induced C_4. This implies that G is not an adjusted interval digraph, as otherwise it contradicts the fact that the underlying undirected graphs of adjusted interval digraphs are interval graphs [40]. This proves that the class of interval catch digraphs (and therefore, the class of reflexive interval digraphs) is not contained in the class of adjusted interval digraphs.



G an interval nest digraph representation of G

Figure 5.8: An interval nest digraph that is not an interval catch digraph

Now let G be a digraph with $V(G) = \{a, b, c, d\}$ and edges $(a, b), (c, b), (b, d), (d, b)$ in addition to loops at each vertex (refer to Figure 5.8). The digraph G is not an interval catch digraph, as in any interval catch representation of G, the point T_b contained in each of the intervals S_a, S_b and S_c. Thus the intervals S_a, S_b, S_c intersect pairwise, which implies that one of the intervals S_a, S_b, S_c is contained in the union of the other two. We have that S_a is not contained in $S_b \cup S_c$, since otherwise the fact that $T_a \in S_a$ implies that either (b, a) or (c, a) is an edge in G, which is a contradiction. For the same reason, we also have that S_c is not contained in $S_a \cup S_b$. We can therefore conclude that $S_b \subseteq S_a \cup S_c$. But as $(b, d) \in E(G)$, we have that $T_d \in S_b$, which implies that either (a, d) or (c, d) is an edge in G – a contradiction. Thus G is not an interval catch digraph. On the other hand, G has an interval nest representation as shown in Figure 5.8, where the thin green and bold red intervals respectively denotes the source and destination intervals corresponding to the vertices in G. This implies that G is an interval nest digraph that is not an interval catch digraph. This further implies that the class of interval nest digraphs is not contained in the class of interval catch digraph (and therefore not contained in the class of interval point digraphs, as we have noted before).



G a chronological interval representation of G

Figure 5.9: A chronological interval digraph that is not a point-point digraph

Consider a digraph G with $V(G) = \{a, b, c, d\}$ and edges $(a, b), (a, c), (b, c), (c, b), (c, d)$ in addition to loops at each vertex. Figure 5.9 gives a chronological interval representation for G, where the thin green and bold red intervals respectively denotes the source and destination intervals corresponding to the vertices in G. But as
$(a, b), (c, b), (c, d) \in E(G)$ and $(a, d) \notin E(G)$, we have that a, b, c, d is an anti-directed walk of length 3. Therefore by Theorem 21, we have that G is not a point-point digraph. Thus we have that the class of chronological interval digraphs is not contained in the class of point-point digraphs. The above observations explains the comparability relations for the classes of digraphs in Figure 5.3.

5.5 Some remarks

After work on this part had been completed, we have been made aware of a recent manuscript of Jaffke, Kwon and Telle [72], in which unified polynomial-time algorithms have been obtained for the problems considered in this part for some classes reflexive intersection digraphs including reflexive interval digraphs. Their algorithms are more general in nature, and consequently have much higher time complexity, while our targeted algorithms are much more efficient; for example, our algorithm finds a minimum dominating (or absorbing) set in a reflexive interval digraph in time $O(m+n)$, while the general algorithm of [72] has complexity $O(n^8)$.

Müller [95] showed the close connection between interval digraphs and interval bigraphs and used this to construct the only known polynomial-time recognition algorithm for both these classes (refer to Chapter 3 for more details). Since this algorithm takes $O(nm^6(n+m)\log n)$ time, the problem of finding a forbidden structure characterization for either of these classes or a faster recognition algorithm are long standing open questions in this field. But many of the subclasses of interval digraphs, like adjusted interval digraphs [114], chronological interval digraphs [31], interval catch digraphs [99], and interval point digraphs [100] have simpler and much more efficient recognition algorithms. It is quite possible that simpler and efficient algorithms for recognition exist also for reflexive interval digraphs. As for the case of interval nest digraphs, no polynomial-time recognition algorithm is known. The complexities of the recognition problem and Absorbing-Set problem for DUF-digraphs also remain as open problems.

Chapter 6

The Weak Independent Set and Directed Feedback Vertex Set Problems

6.1 Introduction

Given a digraph G, we call a set $S \subseteq V(G)$ a *weak independent set* of G, if for any two vertices $u, v \in S$, either $(u, v) \notin E(G)$ or $(v, u) \notin E(G)$. As any set that consists of a single vertex is a weak independent set of G, the interesting computational problem that arises here is that of finding a maximum cardinality weak independent set, called Weak Independent-Set problem. Since independent sets of an undirected graph G are exactly the weak independent sets of the symmetric digraph of G, it can be easily seen that the Weak Independent-Set problem is NP-complete. We will see that in the case of reflexive interval digraphs, the notion of weak independent set has a close connection to another well-known problem in the literature. Given a digraph G, a set $S \subseteq V(G)$ is said to be a *directed feedback vertex set* of G if the digraph induced by the vertices in $V(G) \setminus S$ is a DAG (where loops are allowed to be present). In other words, the removal of a feedback vertex set of a digraph G destroys all the directed cycles (except for the loops) in G. The problem of finding a minimum cardinality directed feedback vertex set, called Feedback Vertex-Set problem is a classic problem that is shown to be NP-complete by Karp [75] along with the first list of NP-complete problems. The Feedback Vertex-Set problem has a significant role in the study of *deadlock recovery* in the field of database systems [48] and is applicable to many other real life problems as well. Several approaches can be seen towards the Feedback Vertex-Set problem including a parameterized approach which was motivated by Chen, Liu and Lu [18].

Figure 6.1 illustrates through an example the notions of weak independent sets and feedback

(a) A weak independent set

(b) a feedback vertex set

Figure 6.1: Examples: the dark red vertices in (a) and (b) respectively denote a weak independent set and a feedback vertex set of the given digraph.

vertex sets for digraphs.

In this chapter we will see that the problems Weak Independent-Set and Feedback Vertex-Set can be reduced to each other for DUF digraphs in linear time. We also study these problems in some particular subclasses of interval digraphs such as interval nest digraphs, point-point digraphs and adjusted interval digraphs. Moreover our solution to the Weak Independent-Set problem for interval nest digraphs has an interesting consequence as well.

6.2 The weak independent set and feedback vertex set problems for DUF digraphs

We have the following lemma.

Lemma 17. *Let G be a DUF digraph. A set $S \subseteq V(G)$ is a weak independent set of G if and only if $V(G) \setminus S$ is a feedback vertex set of G.*

Proof. Suppose that $V(G) \setminus S$ is a feedback vertex set of G. Then we have that the subgraph $G[S]$ is a DAG. Therefore as $G[S]$ does not contain any directed cycles, we have in particular that $G[S]$ does not contain any directed cycles of length 2. In other words, there cannot exist two vertices $u, v \in V(G)$ such that $(u, v), (v, u) \in E(G)$, implying that S is a weak independent set of G.

On the other hand, assume that $S \subseteq V(G)$ is a weak independent set of G. For the sake of contradiction assume that $V(G) \setminus S$ is not a feedback vertex set. Therefore we have that $G[S]$ has at least one directed cycle in it and let C be the directed cycle in $G[S]$ whose length is minimum, say k. Clearly, C is an induced cycle in G. As S is a weak independent set of G, we have that

$k > 2$. Since G is a DUF digraph, we have that the vertices in G have a DUF ordering $<$. Let x be the vertex in C that has the least index in the ordering $<$. I.e. $x = \min_< V(G) \cap V(C)$. Since $x \in V(G) \cap V(C)$ and $k > 2$, there exist two distinct vertices, say $y, z \in V(G) \cap V(C)$ such that $y \in N_G^+(x)$ and $z \in N_G^-(x)$. By symmetry we can assume that $y < z$. Since each vertex in C has exactly one out-neighbor and one in-neighbor in $V(G) \cap V(C)$, and $z \in N_G^-(x)$, we then have that $z \notin N_G^-(y)$. Also, as S is a weak independent set and $y \in N_G^+(x)$ we have that $y \notin N_G^-(x)$. Thus we have vertices $x < y < z$ such that $(z, x) \in E(G)$ and $(z, y), (y, x) \notin E(G)$, which form a directed umbrella that is forbidden in $<$. This is a contradiction and therefore we can conclude that $V(G) \setminus S$ is a feedback vertex set. $\square$

Now the following corollary is an easy consequence of the above theorem.

Corollary 11. *The problems* WEAK INDEPENDENT-SET *and* FEEDBACK VERTEX-SET *are reducible to each other for DUF digraphs and therefore, for reflexive interval digraphs in linear time.*

6.2.1 The weak independent set problem for point-point digraphs and adjusted interval digraphs

Now we evaluate the complexities of the above problems in some subclasses of interval digraphs. First we note the following observation. Let G be a digraph. Let S_G be the undirected graph with $V(S_G) = V(G)$ and $E(S_G) = \{uv : (u, v), (v, u) \in E(G)\}$ (see Figure 6.2 for an example). Note that by the definition of S_G, the weak independent sets of G are exactly the independent sets of S_G. Therefore we have that the WEAK INDEPENDENT-SET problem for G is equivalent to the maximum independent set problem for S_G. We use this equivalence to solve the WEAK INDEPENDENT-SET problem for point-point digraphs and adjusted interval digraphs.

Lemma 18. *The following is true.*

(*a*) *If G is a point-point digraph then S_G is P_4-free.*

(*b*) *If G is an adjusted interval digraph then S_G is an interval graph.*

Proof. Suppose that G is a point-point digraph. By Theorem 21, we have that G does not contain an anti-directed walk of length 3. Recall that a, b, c, d is an anti-directed walk of length 3 if $a, b, c, d \in V(G)$, $(a, b), (c, b), (c, d) \in E(G)$ and $(a, d) \notin E(G)$. We then claim that S_G is P_4-free. Suppose not. Let x, y, z, w be a P_4 in S_G; then $x, y, z, w \in V(G) = V(S_G)$, $xy, yz, zw \in E(S_G)$, and $xw \notin E(S_G)$. Therefore by the definition of S_G, we have that

(a) G (b) S_G

Figure 6.2: An example

$(x,y),(y,x),(y,z),(z,y),(z,w),(w,z) \in E(G)$ and since $xw \notin E(G)$, we also know that either $(x,w) \notin E(G)$ or $(w,x) \notin E(G)$. Now if $(x,w) \notin E(G)$, then x,y,z,w forms an an anti-directed walk of length 3, since $(x,y),(z,y),(z,w) \in E(G)$ and $(x,w) \notin E(G)$. On the other hand, if $(w,x) \notin E(G)$, then w,z,y,x forms an an anti-directed walk of length 3, since $(w,z),(y,z),(y,x) \in E(G)$ and $(w,x) \notin E(G)$. As we have a contradiction in both the cases we can conclude that S_G is P_4-free.

Now suppose that G is an adjusted interval digraph. Let $<$ be an ordering of vertices in G with respect to the common left end-points of intervals corresponding to each vertex. Then $<$ has the following property: for any three distinct vertices $u < v < w$, if $(u,w) \in E(G)$ then $(u,v) \in E(G)$ and if $(w,u) \in E(G)$ then $(v,u) \in E(G)$. Thus for any three distinct vertices $u < v < w$, if the vertices u and w are connected by a symmetric arc in G, then it implies that the vertices u and v are also connected by a symmetric arc in G. By the definition of S_G, this implies that the $V(S_G) = V(G)$ has an ordering $<$ with the property that, for any $u,v,w \in V(S_G) = V(G)$ such that $u < v < w$ we have: $uw \in E(S_G) \implies uv \in E(S_G)$. By Theorem 1, we then have that S_G is an interval graph. $\square$

Since the maximum independent set problem is linear-time solvable for P_4-free graphs [89] and interval graphs [96], we then have the following theorem by a previous observation and Lemma 18.

Theorem 27. *The* Weak Independent-Set *problem can be solved in linear time for point-point digraphs and adjusted interval digraphs.*

As the class of adjusted interval digraphs forms a subclass of reflexive interval digraphs, the following corollary is a consequence of Corollary 11.

Corollary 12. *The* FEEDBACK VERTEX-SET *problem can be solved in linear time for adjusted interval digraphs.*

6.2.2 The weak independent set problem for interval nest digraphs

Recall that the class of *interval nest digraphs* is a subclass of reflexive interval digraphs that has an interval representation $\{(S_u, T_u)\}_{u \in V(G)}$ having the property that $T_u \subseteq S_u$ for every $u \in V(G)$. In this section, we present a polynomial-time dynamic programming algorithm to compute a maximum cardinality weak independent set for an interval nest digraph G, whose interval nest representation is given. Note that we can assume that every interval in the interval nest representation has distinct integer end-points. As there are four end-points corresponding to each vertex, we can assume that each end-point in the representation is a unique integer in $[1, 4|V(G)|]$.

Let the interval nest representation of the input interval nest digraph G be $\{(S_u, T_u)\}_{u \in V(G)}$. For a vertex $u \in V(G)$, let $S_u = [L_u, R_u]$ and $T_u = [l_u, r_u]$. Because of our assumptions about the interval nest representation, we have $L_u < l_u < r_u < R_u$.

For any vertex $x \in V(G)$, let $\eta(x)$ denote the vertex such that $l_x < l_{\eta(x)}$ but there does not exist any vertex $x' \in V(G)$ such that $l_x < l_{x'} < l_{\eta(x)}$.

For vertices $u, v \in V(G)$, we define

$$X(u,v) = \begin{cases} \{y \in V(G) \colon r_u < L_y < R_y < l_v\} & \text{when } r_u < l_v \text{ and } L_v < r_u, \\ \emptyset & \text{otherwise} \end{cases}$$

In addition, for $u, v, x \in V(G)$, define

$$Y(u,v,x) = \begin{cases} \{y \in X(u,v) \colon l_y \geq l_x\} & \text{when } r_u < l_x < l_v \\ \emptyset & \text{otherwise} \end{cases}$$

Note that $X(u,v) = Y(u,v,\eta(u))$. We shall now define our dynamic programming table S in which there is an entry $S(u,v,x) \subseteq Y(u,v,x)$ for every triple of vertices $(u,v,x) \in V(G)^3$ $(= V(G) \times V(G) \times V(G))$. Note that by our definition of $X(u,v)$ and $Y(u,v,x)$, the entry $S(u,v,x)$ corresponding to the triple (u,v,x) will be $\emptyset$ if at least one of the conditions $r_u < l_v$, $L_v < r_u$ or $r_u < l_x < l_v$ is not true. We shall ensure that $S(u,v,x)$ is a weak independent set of maximum possible cardinality among all the weak independent sets that contain only vertices in $Y(u,v,x)$. In other words, $S(u,v,x)$ is a maximum cardinality weak independent set in the subdigraph induced in G by $Y(u, v, x)$.

We give below the pseudocode for a procedure that computes $S(u, v, x)$, given $u, v, x \in V(G)$.

Procedure ComputeS(u, v, x)

---

1. **if** $Y(u, v, x) = \emptyset$ **then**
2.   set $S(u, v, x) = \emptyset$
3.   **return**
4. Set $T = S(u, v, \eta(x))$
5. **if** $x \in X(u, v)$ **then**
6.   Set $T' = \{x\} \cup S(x, v, \eta(x))$
7.   **if** $|T'| > |T|$ **then** set $T = T'$
8.   Set $B = \{y \in Y(u, v, x) \colon L_y < r_x \text{ and } R_x < l_y\}$
9.   **for each** $y \in B$,
10.   Set $T' = \{x\} \cup S(x, y, \eta(x)) \cup S(u, v, y)$
11.   **if** $|T'| > |T|$ **then** set $T = T'$
12. Set $S(u, v, x) = T$

Lemma 19. *For $(u, v, x) \in V(G)^3$, $S(u, v, x)$ is a subset of $Y(u, v, x)$ and is a weak independent set in G.*

Proof. We shall prove this by induction on $l_v - l_x$. If $l_v - l_x \leq 0$, then we have $Y(u, v, x) = \emptyset$ and therefore $S(u, v, x) = \emptyset$. Clearly, the statement of the lemma is true in this case. Now let us assume that the statement has been proved for all $(u', v', x') \in V(G)^3$ such that $l_{v'} - l_{x'} < l_v - l_x$.

If $S(u, v, x) = \emptyset$, then there is nothing to prove. Otherwise, it is the set that got assigned to T in the last step where the value of T was changed. This last step where T's value was changed might be step 4, step 7 or an iteration of step 11. Moreover, $Y(u, v, x) \neq \emptyset$ from which it follows that $l_x > r_u$.

First, let us consider the case when the last time T got assigned was in step 4. In this case, $S(u, v, x) = S(u, v, \eta(x))$. As $l_v - l_x > l_v - l_{\eta(x)}$, we can use the induction hypothesis to conclude that $S(u,v,\eta(x)) \subseteq Y(u,v,\eta(x))$. Since $Y(u,v,\eta(x)) \subseteq Y(u,v,x)$ (recall that $l_x > r_u$), we have $S(u,v,x) = S(u,v,\eta(x)) \subseteq Y(u,v,x)$. It is immediately clear from the induction hypothesis that $S(u,v,\eta(x)) = S(u,v,x)$ is a weak independent set in G.

Next, we consider the case when the last time T got assigned a set was in step 7. Then, we know that $x \in X(u,v)$ which implies that $x \in Y(u,v,x)$. We also have $S(u,v,x) = \{x\} \cup S(x,v,\eta(x))$. Again, by the induction hypothesis, we have $S(x,v,\eta(x)) \subseteq Y(x,v,\eta(x))$ and that $S(x,v,\eta(x))$ is a weak independent set in G. As $x \in X(u,v)$, we have $Y(x,v,\eta(x)) \subseteq Y(u,v,x)$. Since we also have $x \in Y(u,v,x)$, we can conclude that $S(u,v,x) = (\{x\} \cup S(x,v,\eta(x))) \subseteq Y(u,v,x)$. To see that $\{x\} \cup S(x,v,\eta(x))$ is a weak independent set in G, observe that for every vertex $w \in S(x,v,\eta(x)) \subseteq Y(x,v,\eta(x)) = X(x,v)$, we have $r_x < L_w$, implying that $(w,x) \notin E(G)$.

Finally, consider the case when the last time that an assignment to T took place was in an iteration of step 11. Again, it must be the case that $x \in X(u,v)$, which implies that $x \in Y(u,v,x)$. Also, we have $S(u,v,x) = \{x\} \cup S(x,y,\eta(x)) \cup S(u,v,y)$ for some $y \in B \subseteq Y(u,v,x) \subseteq X(u,v)$. By the induction hypothesis, we have $S(x,y,\eta(x)) \subseteq Y(x,y,\eta(x))$ and $S(u,v,y) \subseteq Y(u,v,y)$. As we have $x,y \in X(u,v)$, we can conclude that $Y(x,y,\eta(x)) \subseteq Y(u,v,x)$ and thereby $S(x,y,\eta(x)) \subseteq Y(u,v,x)$. From the definition of B, it is clear that $l_x < l_y$, implying that $Y(u,v,y) \subseteq Y(u,v,x)$, and therefore $S(u,v,y) \subseteq Y(u,v,x)$. Altogether, we now have $S(u,v,x) = (\{x\} \cup S(x,y,\eta(x)) \cup S(u,v,y)) \subseteq Y(u,v,x)$. It only remains to be shown that $S(u,v,x) = \{x\} \cup S(x,y,\eta(x)) \cup S(u,v,y)$ is a weak independent set in G. It is easy to see that for every vertex $w \in S(x,y,\eta(x)) \subseteq Y(x,y,\eta(x)) = X(x,y)$, we have $r_x < L_w$ and therefore, $(w,x) \notin E(G)$. Now consider a vertex $w' \in S(u,v,y) \subseteq Y(u,v,y)$. Clearly, $l_{w'} \geq l_y$. From the definition of B, we have $R_x < l_y$ which now gives us $R_x < l_{w'}$. This means that $(x,w') \notin E(G)$. Finally, let us consider a vertex $w \in S(x,y,\eta(x)) \subseteq Y(x,y,\eta(x)) = X(x,y)$ and a vertex $w' \in S(u,v,y) \subseteq Y(u,v,y)$. Clearly, $R_w < l_y \leq l_{w'}$, implying that $(w,w') \notin E(G)$. □

Lemma 20. *Let $(u,v,x) \in V(G)^3$ and let $S' \subseteq Y(u,v,x)$ be a weak independent set in G. Then $|S'| \leq |S(u,v,x)|$.*

Proof. We shall prove this by induction on $l_v - l_x$. If $l_v - l_x \leq 0$, then we have $Y(u,v,x) = \emptyset$ and therefore $S(u,v,x) = \emptyset$. Clearly, the statement of the lemma is true in this case. Now let us assume that the statement has been proved for all $(u',v',x') \in V(G)^3$ such that $l_{v'} - l_{x'} < l_v - l_x$.

First let us note that the procedure ComputeS(u,v,x) actually computes $S(u,v,x)$ as given by the following expression, where Max$(\mathcal{F})$ denotes a set of maximum cardinality in a family $\mathcal{F}$

of sets.

$$S(u,v,x) = \begin{cases} S(u,v,\eta(x)) & \text{if } x \notin X(u,v) \\ \text{Max}\begin{pmatrix} \{S(u,v,\eta(x)), \{x\} \cup S(x,v,\eta(x))\} \\ \cup \\ \{\{x\} \cup S(x,y,\eta(x)) \cup S(u,v,y) \colon y \in B\} \end{pmatrix} & \text{if } x \in X(u,v) \end{cases} \tag{6.1}$$

Let $S' \subseteq Y(u,v,x)$ be a weak independent set in G. Let us first consider the case in which $x \notin S'$. In this case, it is easy to see that $S' \subseteq Y(u,v,\eta(x))$. From the induction hypothesis, we have $|S'| \le |S(u,v,\eta(x))|$. It follows from equation (6.1) that $|S(u,v,x)| \ge |S(u,v,\eta(x))|$ and therefore we are done.

Now let us consider the case when $x \in S'$. Note that since $S' \subseteq Y(u,v,x) \subseteq X(u,v)$, we now have $x \in X(u,v)$.

Suppose first that there exists some vertex $z \in S' \setminus \{x\}$ such that $L_z < r_x$. Then let z be that vertex in $S' \setminus \{x\}$ with $L_z < r_x$ such that there exists no vertex $z' \in S' \setminus \{x\}$ with $l_{z'} < l_z$ and $L_{z'} < r_x$. Let $S'_1 = S' \cap X(x,z)$ and $S'_2 = S' \setminus (\{x\} \cup S'_1)$. Note that S' is a disjoint union of the sets $\{x\}$, S'_1 and S'_2 and that $z \in S'_2$. We claim that for each vertex $w \in S'_2$, we have $l_w \ge l_z$. Suppose that there exists $w \in S'_2$ such that $l_w < l_z$. As S' is a weak independent set containing both w and z, it must be the case that $r_w < l_z$ (otherwise, $[l_w, r_w] \cap [l_z, r_z] \ne \emptyset$, implying that both $(w,z), (z,w) \in E(G)$). If $L_w < r_x$, then we have a contradiction to our choice of z. Therefore, we have $r_x < L_w$. Recalling that $L_z < r_x$, we now have $L_z < r_x < L_w < r_w < l_z$. Then, the only reason $w \notin X(x,z)$ must be the fact that $l_z < R_w$. But now we have $L_z < r_w < l_z < R_w$, implying that both $(w,z), (z,w) \in E(G)$. But this is impossible as both z and w belong to a weak independent set S' of G. This allows us to conclude that every vertex $w \in S'_2$ has the property that $l_w \ge l_z$. Therefore, recalling that $S' \subseteq X(u,v)$, we can infer that $S'_2 \subseteq Y(u,v,z)$. Clearly, $S'_1 \subseteq X(x,z) = Y(x,z,\eta(x))$. Since $S' \subseteq Y(u,v,x)$ and $z \in S' \setminus \{x\}$, we have $l_x < l_z < R_z < l_v$, implying that $l_v - l_z < l_v - l_x$ and $l_z - l_{\eta(x)} < l_v - l_x$. By the induction hypothesis, we now have $|S'_2| \le |S(u,v,z)|$ and $|S'_1| \le |S(x,z,\eta(x))|$. Therefore, $|S'| = 1 + |S'_1| + |S'_2| \le 1 + |S(x,z,\eta(x))| + |S(u,v,z)|$. Recalling that $l_x < l_z$, $L_z < r_x$ and that both z and x belong to a weak independent set S' of G, we can conclude that $R_x < l_z$. This means that $z \in B$ and from equation (6.1), we now have

$|S(u,v,x)| \geq |\{x\} \cup S(x,z,\eta(x)) \cup S(u,v,z)| = 1 + |S(x,z,\eta(x))| + |S(u,v,z)|$ (as the sets $\{x\}$, $S(x,z,\eta(x))$ and $S(u,v,z)$ are pairwise disjoint). This shows that $|S(u,v,x)| \geq |S'|$.

Next, we shall consider the case when there does not exist any vertex $z \in S' \setminus \{x\}$ such that $L_z < r_x$. Then for every $w \in S' \setminus \{x\}$, we have $r_x < L_w$, which implies that $S' \setminus \{x\} \subseteq X(x,v) = Y(x,v,\eta(x))$. As $S' \setminus \{x\}$ is a weak independent set in G and $l_v - l_{\eta(x)} < l_v - l_x$, we have $|S' \setminus \{x\}| \leq |S(x,v,\eta(x))|$ by our induction hypothesis. Therefore, $|S'| = 1 + |S' \setminus \{x\}| \leq 1 + |S(x,v,\eta(x))| = |\{x\} \cup S(x,v,\eta(x))|$ (note that $x \notin S(x,v,\eta(x))$). From equation (6.1), it is clear that $|S(u,v,x)| \geq |\{x\} \cup S(x,v,\eta(x))|$. We thus have $|S'| \leq |S(u,v,x)|$ as required. $\square$

Theorem 28. *The* Weak Independent-Set *(resp.* Feedback Vertex-Set*) problem for interval nest digraphs can be solved in $O(n^4)$ time, given the interval nest representation of the digraph as input.*

Proof. Add the intervals corresponding to two dummy vertices a and b to the input interval nest representation. Recalling that the left-most end-point in the input representation was 1 and the right-most $4|V(G)|$, let $L_a = -4$, $l_a = -3$, $r_a = -1$, $R_a = 0$, $L_b = -2$, $l_b = 4|V(G)|+1$, $r_b = 4|V(G)|+2$ and $R_b = 4|V(G)|+3$. The sets $X(u,v)$ for all $u,v \in V(G) \cup \{a,b\}$ and $Y(u,v,x)$ for all $(u,v,x) \in (V(G) \cup \{a,b\})^3$ can be computed in $O(n^4)$ time. The algorithm then calls the procedure ComputeS$(a,b,\eta(a))$ and outputs the set $S(a,b,\eta(a))$. Note that this being a dynamic programming algorithm, a call to $S(u,v,x)$ for some $(u,v,x) \in (V(G) \cup \{a,b\})^3$ is made only if $S(u,v,x)$ has not been computed before – or in other words, the algorithm ensures that a call to ComputeS(u,v,x) is made at most once for each triple $(u,v,x) \in (V(G) \cup \{a,b\})^3$. Therefore, the total number of times the procedure ComputeS needs to be called recursively during the execution of ComputeS$(a,b,\eta(a))$ is at most $(n+2)^3$. It is easy to see from the procedure ComputeS(u,v,x) that the time spent in the computation of $S(u,v,x)$ outside the recursive calls to the procedure is $O(n)$. Therefore, the total running time of ComputeS$(a,b,\eta(a))$ is $O(n^4)$, implying that our algorithm has time complexity $O(n^4)$. We only need to show that the output of the algorithm, $S(a,b,\eta(a))$, is a maximum cardinality weak independent set in G. It is clear that $X(a,b) = Y(a,b,\eta(a)) = V(G)$. Therefore, by Lemmas 19 and 20, $S(a,b,\eta(a))$ is a maximum cardinality weak independent set in G. This together with Corollary 11 proves the theorem. $\square$

Now in the next section we will see how the solution for the Weak Independent-Set problem for interval nest digraphs can be used to solve an interesting problem for the class of interval graphs.

6.3 The uniquely restricted matching problem

Given an undirected graph G, a set $M \subseteq E(G)$ is said to be a *matching* if no two edges in M has a common vertex as an end-point. The problem of finding a maximum cardinality matching in a given graph is polynomial-time solvable [39]. A matching M in an undirected graph G is said to be *uniquely restricted* if there is no other matching in G that matches the same set of vertices as M. Unlike the problem of maximum cardinality matching, the problem of finding a maximum cardinality uniquely restricted matching is shown to be NP-hard even for the special graph classes like bipartite graphs and split graphs [57] by Golumbic, Hirst, and Lewenstein. In fact, this problem is shown to be APX-complete even for the bipartite graphs of degree at most 3 by Mishra [93]. In their paper initiating the study of uniquely restricted matchings, Golumbic, Hirst, and Lewenstein [57] proposed linear time algorithms for the problem on threshold graphs, proper interval graphs, cacti, and block graphs while leaving open the question of whether polynomial-time algorithms exist for the problem on interval graphs and permutation graphs. As a consequence of our solution for the WEAK INDEPENDENT-SET problem for interval nest digraphs, here we settle the complexity of the maximum cardinality uniquely restricted matchings for the class of interval graphs.

Before going to our theorem, we state the following definitions and some of the results from [57].

Definition 32 (Alternating cycle with respect to M)**.** *Let G be an undirected graph and $M \subseteq E(G)$ be a matching in M. An even length cycle with edges, say $e_0, e_1, \ldots, e_k$ is said to be an alternating cycle with respect to M in G if $e_i \in M$ and $e_{i+1} \notin M$ (i modulo k) for each $i \in \{0, 1, \ldots, k\}$.*

Golumbic, Hirst, and Lewenstein [57] proved the following theorem:

Theorem 29 ([57])**.** *Let G be an undirected graph and let M be a matching in G. Then M is a uniquely restricted matching in G if and only if there is no alternating cycle in G with respect to M.*

In particular for interval graphs they proved the following theorem.

Theorem 30 ([57])**.** *Let G be an undirected graph and let M be a matching in G. Then the following conditions are equivalent.*

- (a) *M is a uniquely restricted matching.*
- (b) *G does not contain any alternating cycle of length 4 with respect to M.*

(c) *For any pair of edges $e, e' \in E(G)$ we have that e, e' is a uniquely restricted matching in G.*

Now for the alternating cycles of length 4, we note the following observation.

Observation 19. *Let G be a graph and let $\{e, e'\}$ be a matching in G such that $e = uv$ and $e' = u'v'$. Then $\{e, e'\}$ form an alternating cycle of length 4 if and only if for each $w \in \{u, v\}$ we have $N(w) \cap \{u', v'\} \neq \emptyset$ and for each $w' \in \{u', v'\}$ we have $N(w') \cap \{u, v\} \neq \emptyset$.*

Proof. Suppose that $\{e, e'\}$ form an alternating cycle of length 4. Then it should be either of the forms $uvu'v'u$ or $uvv'u'u$. By the definition of alternating cycle, in both the cases it is easy to see that, for each $w \in \{u, v\}$, we have $N(w) \cap \{u', v'\} \neq \emptyset$ and for each $w' \in \{u', v'\}$, we have $N(w') \cap \{u, v\} \neq \emptyset$.

On the other hand, suppose that for each $w \in \{u, v\}$ we have $N(w) \cap \{u', v'\} \neq \emptyset$ and for each $w' \in \{u', v'\}$ we have $N(w') \cap \{u, v\} \neq \emptyset$. If $vu' \notin E(G)$ or $v'u \notin E(G)$, then by our assumption, we have $uu' \in E(G)$ and $vv' \in E(G)$, implying that $uvv'vu$ is an alternating cycle of length 4 in G. Therefore we can assume that $vu' \in E(G)$ and $v'u \in E(G)$, implying that $vu'v'uv$ is an alternating cycle of length 4 in G. □

Here we solve the uniquely restricted matching problem for interval graphs by reducing it from the weak independent set problem for the interval nest digraphs as described below.

6.3.1 The uniquely restricted matching problem for interval graphs

Let G be an interval graph for which we wish to compute a maximum cardinality uniquely restricted matching. Note that we can assume that the interval representation of the input graph G is at our disposal. This is because even if the input graph is provided as an adjacency list, there are well-known algorithms that can generate an interval representation of G in linear-time [80, 25, 61]. Let $\{I_u\}_{u \in V(G)}$ be an interval representation of G. For a vertex $u \in V(G)$, let $I_u = [l_u, r_u]$.

We shall define an interval nest digraph H with $V(H) = E(G)$. The arcs of H are defined by specifying the interval nest representation $\{(S_e, T_e)\}_{e \in V(H)}$ of H as follows. For each $e = uv \in V(H)$, where $u, v \in V(G)$, we define $S_e = I_u \cup I_v$ and $T_e = I_u \cap I_v$. Clearly, for each $e \in V(H)$, we have $T_e \subseteq S_e$ and therefore this is an interval nest representation (note that the union or intersection of any two intervals that have a nonempty intersection is again an interval). Thus, H is an interval nest digraph.

Theorem 31. *Let G and H be as defined above. Let $S \subseteq E(G)$. Then S is a weak independent set in H if and only if S is a uniquely restricted matching in G.*

Proof. Suppose that S is a weak independent set in H. Let $e, e' \in S$ and let $e = uv$ and $e' = u'v'$. We first show that e and e' cannot be incident on a common vertex. Suppose for the sake of contradiction that the edges e and e' of G share a common vertex. We shall assume without loss of generality that $v = v'$. Then clearly, $T_{e'} = (I_{u'} \cap I_{v'}) \subseteq I_{v'} = I_v \subseteq (I_u \cup I_v) = S_e$, implying that $S_e \cap T_{e'} \neq \emptyset$ and therefore, $(e, e') \in E(H)$. Similarly, we have $T_e = (I_u \cap I_v) \subseteq I_v = I_{v'} \subseteq (I_{u'} \cup I_{v'}) = S_{e'}$, leading us to infer that $(e', e) \in E(H)$. But this contradicts the fact that both e and e' belong to a weak independent set S in H. Thus, we can conclude that the edges e and e' in G have no common vertex, or in other words, $\{e, e'\}$ is a matching in G. Next, we show that there is no alternating cycle with respect to $\{e, e'\}$ in G. Suppose for the sake of contradiction that there is such a cycle. Then by Observation 19, we know that in G, each of u, v has at least one neighbor in $\{u', v'\}$ and each of u', v' has at least one neighbor in $\{u, v\}$. This means that each of I_u and I_v intersects $I_{u'} \cup I_{v'}$ and each of $I_{u'}$ and $I_{v'}$ intersects $I_u \cup I_v$. Since $uv, u'v' \in E(G)$, this implies that $I_u \cap I_v$ intersects $I_{u'} \cup I_{v'}$ and $I_{u'} \cap I_{v'}$ intersects $I_u \cup I_v$. We thus have $T_e \cap S_{e'} \neq \emptyset$ and $T_{e'} \cap S_e \neq \emptyset$. By definition of H, it must then be the case that $(e, e'), (e', e) \in E(H)$. But this contradicts the fact that both e and e' belong to a weak independent set S in H. We thus conclude that for any two edges $e, e' \in S$, $\{e, e'\}$ is a matching in G and that there is no alternating cycle with respect to $\{e, e'\}$ in G. As G is an interval graph, this implies, by Theorem 30, that S is a uniquely restricted matching in G.

Now suppose that S is a uniquely restricted matching in G. Again let $e, e' \in S$ and let $e = uv$ and $e' = u'v'$. Suppose for the sake of contradiction that $(e, e'), (e', e) \in E(H)$. As $(e, e') \in E(H)$, we can infer that $S_e \cap T_{e'} \neq \emptyset$, which means that $I_{u'} \cap I_{v'}$ intersects $I_u \cup I_v$. Therefore, both $I_{u'}$ and $I_{v'}$ intersect at least one of I_u or I_v. We can thus conclude that each of u', v' is adjacent to at least one vertex in $\{u, v\}$. Now since $(e', e) \in E(H)$, we can follow the same arguments to reach the conclusion that each of u, v is adjacent to at least one vertex in $\{u', v'\}$. From Observation 19, we now have that there is an alternating cycle with respect to $\{e, e'\}$ in G. But this contradicts the fact that both e and e' belongs to a uniquely restricted matching S in G. Therefore, for any pair of edges $e, e' \in S$, we have either $(e, e') \notin E(H)$ or $(e', e) \notin E(H)$, which allows us to conclude that S is a weak independent set in H. ☐

Theorem 32. *There is a polynomial-time algorithm that computes a maximum cardinality uniquely restricted matching in an interval graph.*

Proof. We can generate an interval representation of the input graph G in $O(n + m)$ time using any of the several well-known algorithms (for example, [78]). The interval nest representation of the digraph H corresponding to the interval representation of G can be computed in $O(m)$ time. The algorithm described in the proof of Theorem 28 can now be used to compute a maximum cardinality weak independent set in H in $O(m^4)$ time. It follows from Theorem 31 that this weak independent set corresponds to a maximum cardinality uniquely restricted matching in G. $\square$

In this chapter we gave polynomial-time algorithms for the problems Weak Independent Set and Feedback Vertex-Set for some particular subclasses of interval digraphs. But the complexity status of these problems for the class of reflexive interval digraphs or for a more general class of interval digraphs remain open.

Part III

Extending Some Results on the Seymour Second Neighborhood Conjecture (SSNC)

Chapter 7

SSNC for Tournaments Missing a Matching and a Star

7.1 Introduction

Let $G = (V, E)$ be a digraph with vertex set $V(G)$ and arc set $E(G)$. As usual, $N_G^+(v)$ (resp. $N_G^-(v)$) denotes the out-neighborhood (resp. in-neighborhood) of a vertex $v \in V(G)$. Let $N_G^{++}(v)$ denote the second out-neighborhood of v, which is the set of vertices whose distance from v is exactly 2, i.e. $N_G^{++}(v) = \{u \in V(G) \colon N_G^-(u) \cap N_G^+(v) \neq \emptyset \text{ and } u \notin N_G^+(v) \cup \{v\}\}$. The *out-degree* of a vertex v is defined to be $|N_G^+(v)|$. The minimum out-degree of G is the minimum value among the out-degrees of all vertices of G. We omit the subscript if the digraph under consideration is clear from the context.

A vertex v in a digraph G is said to have a *large second neighborhood* if $|N^{++}(v)| \geq |N^+(v)|$. Oriented graphs are digraphs without loops or digons: i.e. they can be obtained by orienting the edges of a simple undirected graph. Paul Seymour conjectured the following in 1990 (see [35]):

Conjecture 1 (The Second Neighborhood Conjecture)**.** *Every oriented graph contains a vertex with a large second neighborhood.*

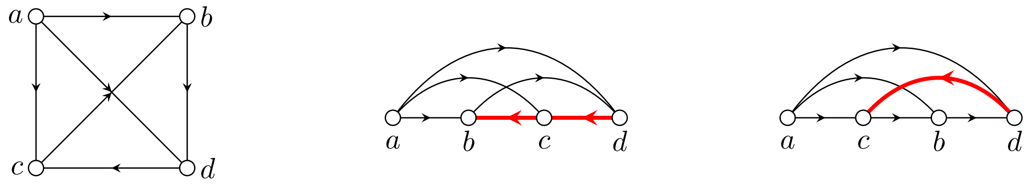
The above conjecture, if true, implies a special case of another open question concerning digraphs called the Caccetta-Häggkvist Conjecture [16], which says that every oriented graph with n vertices and minimum out-degree r contains a directed cycle of length at most $\lceil \frac{n}{r} \rceil$. To be precise, if the Second Neighborhood Conjecture is true then it would imply the Caccetta-Häggkvist Conjecture for the particular case in which both in and out-degree is at least $\frac{n}{3}$. Note that a *sink* trivially has a large second neighborhood and therefore the Second Neighborhood

Conjecture is true for any oriented graph that contains a sink.

Conjecture 1 for the special case of tournaments, was known as Dean's Conjecture [35] and was later solved by Fisher [43] in 1996 using some basic linear algebraic and probabilistic arguments. Later in 2000, Havet and Thomasse [65] gave a short combinatorial proof of Dean's Conjecture using "median orders" of tournaments. They could in fact prove something stronger: in a tournament without a sink, there exist two vertices with large second neighborhoods. Using the approach of Havet and Thomasse, Fidler and Yuster [41] in 2007 proved that the Second Neighborhood Conjecture is true for oriented graphs that can be obtained from tournaments by removing edges in some specific ways. In particular, they showed that a tournament missing a matching (an oriented graph whose missing edges form a matching), a tournament missing a star and a tournament missing a complete graph all satisfy the conjecture. As these results hold even if the missing matching (or star, or complete graph) is empty, they extend the proof of Dean's Conjecture by Havet and Thomasse. Using techniques from this paper, Salman Ghazal [52] proved that the Second Neighborhood Conjecture is true for tournaments missing a "generalized star" – a $\{P_4, C_4, 2K_2\}$-free graph (or equivalently, a threshold graph) – thereby extending the result of Fidler and Yuster for tournaments missing a star and tournaments missing a complete graph. It has to be noted that among these results that all use the median order approach, the case of the tournament missing a matching is by far the most difficult one, requiring a complicated proof. In this chapter, we introduce new ideas to refine and extend this proof, allowing us to prove the conjecture for a superclass of tournaments missing a matching: we show that oriented graphs whose missing edges can be partitioned into a (possibly empty) matching and a (possibly empty) star also satisfy the Second Neighborhood Conjecture. In fact, we prove the stronger statement that in such a graph that does not contain a sink, there exists a vertex that has a large second neighborhood and is not the center of the missing star.

Ghazal [53] attempts to generalize the theorem of Havet and Thomasse by trying to prove that there exist two vertices with large second neighborhoods in every tournament missing a matching that does not contain a sink. He shows that if a tournament missing a matching satisfies certain additional technical conditions, then such a result can be obtained. Our result mentioned above directly yields a proof that shows that every tournament missing a matching that does not contain a sink has at least two vertices with large second neighborhoods.

We also ask whether it is true that if there is exactly one vertex with a large second neighborhood in an oriented graph, then it is a sink. We note that such a result would imply the Second Neighborhood Conjecture.



(i) T (ii) (a,b,c,d) is not a median order (ii) (a,c,b,d) is a median order

Figure 7.1: An example to illustrate median order of a tournament

7.2 Graphs that are almost tournaments

In this section, our main aim will be to show that Conjecture 1 is true for tournaments whose missing edges can be partitioned into a matching and a star. By reviewing median orders of tournaments and their properties, we then study tournaments missing a matching, wherein we introduce the notions and structural results that we need to prove our main result. Along the way, we reprove the result of Fidler and Yuster that the Second Neighborhood Conjecture is true for tournaments missing a matching using these ideas.

7.2.1 Median orders of tournaments

Given an ordering of the vertices of a tournament, an arc of the tournament is said to be a "forward arc" if the starting vertex of the arc occurs earlier than its ending vertex in the ordering. A *median order* of a tournament is an ordering of its vertices with the most number of forward arcs. Formally, an ordering $(x_1, x_2, \dots, x_n)$ of the vertices of a tournament T that maximizes $|\{(x_i, x_j) \in E(T) \colon i < j\}|$ is said to be a median order of T. The *feed vertex* of a median order $(x_1, x_2, \dots, x_n)$ is the last vertex x_n in that ordering of vertices. Figure 7.1 provides an example of a tournament for which (a, b, c, d) is an ordering of the vertices in T that is not a median order (as there are two backward arcs), where as (a, c, b, d) is a median order of T (as the tournament T contains a directed cycle b, c, d in it, any ordering of $V(T)$ would have at least one backward arc). Havet and Thomasse [65] proved the following.

Theorem 33 ([65])**.** *Let T be a tournament and L be a median order of T with feed vertex d. Then $|N_T^+(d)| \leq |N_T^{++}(d)|$.*

The following properties of median orders of tournaments are not difficult to verify (see [65]).

Proposition 3. *Let $(x_1, x_2, \dots, x_n)$ be a median order of a tournament T, and let x_i and x_j be such that $1 \leq i < j \leq n$. If $T' = T[\{x_1, x_2, \dots x_n\}]$, then:*

(*a*) $(x_i, x_{i+1}, \ldots, x_j)$ *is a median order of* T', *and*

(*b*) *if* $(y_1, y_2, \ldots, y_{j-i+1})$ *is a median order of* T', *then* $(x_1, x_2, \ldots, x_{i-1}, y_1, y_2, \ldots, y_{j-i+1}, x_{j+1}, x_{j+2}, \ldots, x_n)$ *is a median order of* T.

Proposition 4 (The feedback property [65])**.** *Let* $(x_1, x_2, \ldots, x_n)$ *be a median order of a tournament* T *and let* x_i *and* x_j *be such that* $1 \le i < j \le n$. *Then:*

(*a*) $|N^+(x_i) \cap \{x_{i+1}, \ldots, x_j\}| \ge |N^-(x_i) \cap \{x_{i+1}, \ldots, x_j\}|$, *and*

(*b*) $|N^+(x_j) \cap \{x_i, \ldots, x_{j-1}\}| \le |N^-(x_j) \cap \{x_i, \ldots, x_{j-1}\}|$.

Proposition 5. *Let* $(x_1, x_2, \ldots, x_n)$ *be a median order of a tournament* T *and let* x_i *and* x_j *be such that* $1 \le i < j \le n$. *Then:*

(*a*) *if* $|N^+(x_i) \cap \{x_{i+1}, \ldots, x_j\}| = |N^-(x_i) \cap \{x_{i+1}, \ldots, x_j\}|$, *then* $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, x_{i+2}, \ldots, x_j, x_i, x_{j+1}, x_{j+2}, \ldots, x_n)$ *is also a median order of* T, *and*

(*b*) *if* $|N^+(x_j) \cap \{x_i, \ldots, x_{j-1}\}| = |N^-(x_j) \cap \{x_i, \ldots, x_{j-1}\}|$, *then* $(x_1, x_2, \ldots, x_{i-1}, x_j, x_i, x_{i+1}, \ldots, x_{j-1}, x_{j+1}, x_{j+2}, \ldots, x_n)$ *is also a median order of* T.

Proposition 6. *Let* $L = (x_1, x_2, \ldots, x_n)$ *be a median order of a tournament* T *and let* $(x_j, x_i) \in E(T)$, *where* $i < j$. *Then* L *is also a median order of the tournament* T' *with* $V(T') = V(T)$ *and* $E(T') = (E(T) \setminus \{(x_j, x_i)\}) \cup \{(x_i, x_j)\}$.

Proof. If L is not a median order of T', then there exists an ordering $\hat{L}$ of $V(T') = V(T)$ such that $(T', \hat{L})$ has at least one more forward arc than (T', L) and therefore at least two more forward arcs than (T, L). But then $(T, \hat{L})$ has at least one more forward arc than (T, L), contradicting the fact that L is a median order of T. Therefore, L is a median order of T' as well. □

Modules Given an oriented graph G, a set $S \subseteq V(G)$ is said to be a *module* in G, if for any two vertices $u, v \in S$, $N^+(u) \setminus S = N^+(v) \setminus S$ and $N^-(u) \setminus S = N^-(v) \setminus S$.

Proposition 7. *Let* G *be an oriented graph and* S *a module in it.*

(*a*) *For* $u \in S$, *let* $G' = G - (S \setminus \{u\})$. *Then,* $N^{++}_{G'}(u) = N^{++}_G(u) \setminus S$.

(*b*) *For* $u, v \in S$, $N^{++}_G(u) \setminus S = N^{++}_G(v) \setminus S$.

Proof. Clearly, $N^{++}_{G'}(u) \subseteq N^{++}_G(u) \setminus S$. Consider any vertex $x \in N^{++}_G(u) \setminus S$. Then $(u, x) \notin E(G)$ and there exists $w \in V(G)$ such that $(u, w), (w, x) \in E(G)$. As we have $x \notin S$, $(w, x) \in E(G)$, $(u, x) \notin E(G)$, and S is a module containing u, we have $w \notin S$. Then since $u, w, x \in V(G')$, we have that $(u, w), (w, x) \in E(G')$ and $(u, x) \notin E(G')$, implying that $x \in N^{++}_{G'}(u)$. Therefore, $N^{++}_G(u) \setminus S \subseteq N^{++}_{G'}(u)$, proving (a).

Note that for proving (b), we only need to prove that $N_G^{++}(u) \setminus S \subseteq N_G^{++}(v) \setminus S$, as u and v are symmetric. Consider any vertex $x \in N_G^{++}(u) \setminus S$. As noted above, $(u,x) \notin E(G)$ and there exists $w \in V(G)$ such that $(u,w),(w,x) \in E(G)$. Since u and v belong to the module S in G, we have that $(v,w) \in E(G)$ and $(v,x) \notin E(G)$, implying that $x \in N_G^{++}(v) \setminus S$. Therefore, $N_G^{++}(u) \setminus S \subseteq N_G^{++}(v) \setminus S$. $\square$

Proposition 8. *Let $(x_1, x_2, \ldots, x_n)$ be a median order of a tournament T. Let $i, j \in \{1, 2, \ldots, n\}$ such that $i < j-1$ and x_i and x_j belong to a module in T and every vertex in $\{x_{i+1}, \ldots, x_{j-1}\}$ is outside this module. Then:*

(a) $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, x_{i+2}, \ldots, x_{j-1}, x_i, x_j, x_{j+1}, \ldots, x_n)$ is a median order of T, and

(b) $(x_1, x_2, \ldots, x_i, x_j, x_{i+1}, x_{i+2}, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ is a median order of T.

Proof. Consider the set of vertices $X = \{x_{i+1}, x_{i+2}, \ldots, x_{j-1}\}$. Suppose that $|N^+(x_i) \cap X| > |N^-(x_i) \cap X|$. As x_i and x_j belong to a module in T and every vertex of X is outside this module, we have $N^+(x_j) \cap X = N^+(x_i) \cap X$ and $N^-(x_j) \cap X = N^-(x_i) \cap X$. This gives us $|N^+(x_j) \cap X| > |N^-(x_j) \cap X|$, which contradicts Proposition 4(b) applied on x_{i+1} and x_j. Therefore, $|N^+(x_i) \cap X| \leq |N^-(x_i) \cap X|$. Then by Proposition 4$(a)$ applied on x_i and x_{j-1}, we have $|N^+(x_i) \cap X| = |N^-(x_i) \cap X|$. Applying Proposition 5$(a)$ on x_i and x_{j-1}, we now get that $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, x_{i+2}, \ldots, x_{j-1}, x_i, x_j, x_{j+1}, \ldots, x_n)$ is a median order of T. This proves (a). It is easy to see, by repeating the same arguments for x_j and X, that (b) is also true. $\square$

Good median orders We now define a special kind of median order of tournaments, along the lines of Ghazal [53]. Given a partition $\mathcal{I} = \{I_1, I_2, \ldots, I_r\}$ of $V(T)$ such that each I_i, $1 \leq i \leq r$, is a module in T, we say that a median order of T is a *good median order with respect to $\mathcal{I}$* if for each $i \in \{1, 2, \ldots, r\}$, the vertices of I_i appear consecutively in it (note that this is slightly different from the "good median orders" defined by Ghazal [53]). Ghazal notes the following property of good median orders (which can be considered as a consequence of Proposition 8(a)).

Lemma 21 ([53])**.** *Let $\mathcal{I} = \{I_1, I_2, \ldots, I_r\}$ be a partition of the vertex set of a tournament T into modules and let L be a median order of T. Then there is a good median order L' of T with respect to $\mathcal{I}$ such that L and L' have the same feed vertex.*

Proof. Given an ordering P of the vertices of T and a module $I \in \mathcal{I}$, a maximal subset of I that is consecutive in P is said to be a "fragment" of I in P. Clearly, the fragments of a module $I \in \mathcal{I}$ are ordered from left to right in P. We define the "weight" of a vertex $v \in I$ with respect to P

to be the number of fragments of I that occur after the fragment of I containing v. The weight of P is defined to be the sum of the weights of all the vertices with respect to P. Note that the median orders of T with zero weight are exactly the good median orders of T with respect to $\mathcal{I}$. Now suppose that P is a median order of T with non-zero weight. Then there exists $I \in \mathcal{I}$ and $u, v \in I$ such that they are not consecutive in P and no vertex between them in P belongs to I. Let P' be the median order of T obtained from P by applying Proposition 8(a) to P, u and v. It can be verified that the weight of P' is strictly less than the weight of P and that P and P' have the same feed vertex. This means that by applying the above procedure repeatedly to the median order L of T, we can obtain a median order L' of T with zero weight (hence, it is a good median order of T with respect to $\mathcal{I}$) having the same feed vertex as L. $\square$

Proposition 9. *Let d be the feed vertex of a median order of a tournament T and let I be a module in T containing d. Then for any vertex $v \in I$, $|N^+(v) \setminus I| \leq |N^{++}(v) \setminus I|$.*

Proof. Let $\mathcal{I} = \{I\} \cup \{\{u\} \colon u \notin I\}$. It is easy to see that $\mathcal{I}$ is a partition of $V(T)$ into modules. By Lemma 21, there exists a good median order $L = (x_1, x_2, \ldots, x_n = d)$ of T with respect to $\mathcal{I}$. Then, there exists $i \in \{1, 2, \ldots, n\}$ such that $I = \{x_i, x_{i+1}, \ldots, x_n\}$. By Proposition 3($a$), $L' = (x_1, x_2, \ldots, x_i)$ is a median order of $T' = T - (I \setminus \{x_i\})$. By Theorem 33, $|N^+_{T'}(x_i)| \leq |N^{++}_{T'}(x_i)|$. Consider any $v \in I$. As I is a module containing x_i and v, $N^+_T(v) \setminus I = N^+_T(x_i) \setminus I = N^+_{T'}(x_i)$. By Proposition 7, we also have that $N^{++}_{T'}(x_i) = N^{++}_T(x_i) \setminus I = N^{++}_T(v) \setminus I$. Combining the above observations, we get $|N^+_T(v) \setminus I| \leq |N^{++}_T(v) \setminus I|$. $\square$

7.2.2 Tournaments missing a matching

In this section, we prove that the Second Neighborhood Conjecture is true for tournaments missing a matching. Throughout this section, we denote by G an oriented graph that can be obtained from a tournament by removing a (possibly empty) matching.

For a vertex $u \in V(G)$, we say that the vertices in $N^+_G(u) \cup N^-_G(u)$ are the *neighbors* of u and that the vertices in $V(G) \setminus (N^+_G(u) \cup N^-_G(u))$ are the *non-neighbors* of u. It is easy to see that every vertex in G has at most one non-neighbor. If there is no edge between two distinct vertices x and y in G, i.e., x is a non-neighbor of y (and vice versa), then we say that $\{x, y\}$ is a *missing edge* in G. We denote this missing edge as $x \dashrightarrow\!\!\!\!\!- y$ (or, equivalently $y \dashrightarrow\!\!\!\!\!- x$). For an arc $(x, y) \in E(G)$, we use the notation $x \to y$ (in other words, $y \in N^+_G(x)$). If $(x, y) \in E(G)$ is an arc with the additional property that $x \notin N^{++}_G(y)$, then we say that (x, y) is a *special arc*, and denote it as $x \twoheadrightarrow y$. Note that there can be no directed triangle in G containing a special arc.

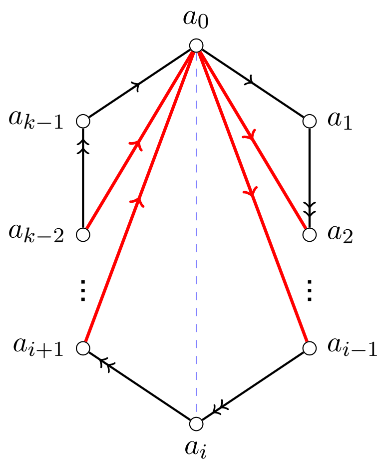


Figure 7.2: Illustration of Lemma 22

Lemma 22. *Let $C = a_0 \to a_1 \twoheadrightarrow a_2 \twoheadrightarrow a_3 \twoheadrightarrow \cdots \twoheadrightarrow a_{k-1} \to a_0$ be a cycle in G. Then:*

(a) *a_0 has a non-neighbor in C, and*

(b) *if $a_0 \dashrightarrow a_i$, then for $j \in \{1, \ldots, i-1\}$, $a_0 \to a_j$ and for $j \in \{i+1, \ldots, k-1\}$, $a_j \to a_0$.*

(See Figure 7.2 for an illustration of the lemma)

Proof. Since G is an oriented graph that has no directed triangle containing a special arc, we have that $k \geq 4$.

(a) Assume to the contrary that a_0 has no non-neighbor in C, i.e., $\forall i \neq 0, a_0 \to a_i$ or $a_i \to a_0$. For some $i \neq 0$, if $a_0 \to a_i$, then $a_0 \to a_{i+1}$, because otherwise, $a_0 \to a_i \twoheadrightarrow a_{i+1} \to a_0$ forms a directed triangle containing a special arc. Now since $a_0 \to a_1$, applying this observation repeatedly gives us $a_0 \to a_2$, $a_0 \to a_3$, $\ldots$, $a_0 \to a_{k-1}$, which is a contradiction to the fact that $a_{k-1} \to a_0$.

(b) Let $a_0 \dashrightarrow a_i$. As a_i is the only non-neighbor of a_0 in G, for each $j \notin \{0, i\}$, we have either $a_0 \to a_j$ or $a_j \to a_0$. Suppose that for some $j \in \{1, \ldots, i-1\}$, we have $a_j \to a_0$, then consider the cycle $C' = a_0 \to a_1 \twoheadrightarrow \cdots \twoheadrightarrow a_j \to a_0$. Then a_0 has no non-neighbor in C', which is a contradiction to (a). Similarly, if there is some $j \in \{i+1, \ldots, k-1\}$ such that $a_0 \to a_j$, then there is no non-neighbor of a_0 in the cycle $a_0 \to a_j \twoheadrightarrow a_{j+1} \twoheadrightarrow \cdots \twoheadrightarrow a_{k-1} \to a_0$, again contradicting (a). $\square$

Special cycles We call a cycle in G a *special cycle* if it consists only of special arcs. It is easy to see that any special cycle contains at least 4 vertices. The following corollary is an immediate consequence of Lemma 22.

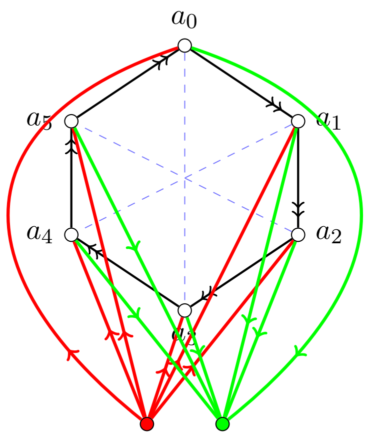


Figure 7.3: Illustration of Lemma 23: dark red and light green vertices belong to $V(G) \setminus C$ and the vertices in C are given in white

Corollary 13. *Let $C = a_0 \twoheadrightarrow a_1 \twoheadrightarrow a_2 \twoheadrightarrow a_3 \twoheadrightarrow \cdots \twoheadrightarrow a_{k-1} \twoheadrightarrow a_0$ be a special cycle in G. Then:*

(a) *Each vertex in C has a non-neighbor in C,*

(b) *if $a_i \dashrightarrow\!\!\!\!- a_j$, then $N^+_G(a_i) \cap V(C) = \{a_{i+1}, a_{i+2}, \ldots, a_{j-1}\}$ and $N^-_G(a_i) \cap V(C) = \{a_{j+1}, a_{j+2}, \ldots, a_{i-1}\}$, where subscripts are modulo k.*

Lemma 23. *Let $C = a_0 \twoheadrightarrow a_1 \twoheadrightarrow a_2 \twoheadrightarrow a_3 \twoheadrightarrow \cdots \twoheadrightarrow a_{k-1} \twoheadrightarrow a_0$ be a special cycle in G. Then:*

(a) *k is even,*

(b) *for each vertex $a_i \in V(C)$, $a_i \text{---} a_{i+\frac{k}{2}}$ (subscripts modulo k),*

(c) *$V(C)$ forms a module in G.*

(See Figure 7.3 for an illustration of the lemma).

Proof. Using Corollary 13(a), we have that every vertex of C has exactly one non-neighbor in C. This proves (a).

(b) Let a_j be the non-neighbor of a_i in C. Suppose that $j \neq i + \frac{k}{2}$ (modulo k). Then one of the sets $\{a_{i+1}, a_{i+2}, \ldots, a_{j-1}\}, \{a_{j+1}, a_{j+2}, \ldots, a_{i-1}\}$ (subscripts modulo k) is larger than the other. We shall assume without loss of generality that $|\{a_{i+1}, a_{i+2}, \ldots, a_{j-1}\}| > |\{a_{j+1}, a_{j+2}, \ldots, a_{i-1}\}|$. This means that there exists $a_p, a_q \in \{a_{i+1}, a_{i+2}, \ldots, a_{j-1}\}$ such that $a_p \text{---} a_q$, where a_p occurs before a_q in the ordering $a_{i+1}, a_{i+2}, \ldots, a_{j-1}$. By Corollary 13$(b)$, we know that $a_i \to a_q$. Now consider the cycle $C' = a_q \twoheadrightarrow a_{q+1} \twoheadrightarrow \cdots \twoheadrightarrow a_{i-1} \twoheadrightarrow a_i \to a_q$ (subscripts modulo k). There is no non-neighbor of a_q in C' (as a_p is the only non-neighbor of a_q), which contradicts Lemma 22(a).

(c) Since every vertex of C has a non-neighbor in C, for any $x \in V(G) \setminus V(C)$, x is a neighbor of every vertex in $V(C) = \{a_0, a_1, \ldots, a_{k-1}\}$. This implies that if $x \rightarrow a_i$ for any $i \in \{0, 1, \ldots, k-1\}$, then we also have $x \rightarrow a_{i+1}$ as otherwise, $x \rightarrow a_i \twoheadrightarrow a_{i+1} \rightarrow x$ would be a directed triangle containing a special arc (subscripts modulo k). Therefore applying this observation repeatedly starting from a_0, we get $N^-_G(a_0) \setminus V(C) \subseteq N^-_G(a_1) \setminus V(C) \subseteq N^-_G(a_2) \setminus V(C) \subseteq \cdots \subseteq N^-_G(a_{k-2}) \setminus V(C) \subseteq N^-_G(a_{k-1}) \setminus V(C) \subseteq N^-_G(a_0) \setminus V(C)$. Similarly, if $a_i \rightarrow x$ for any $i \in \{0, 1, \ldots, k-1\}$, then we also have $a_{i-1} \rightarrow x$, as otherwise $x \rightarrow a_{i-1} \twoheadrightarrow a_i \rightarrow x$ would be a directed triangle containing a special arc (subscripts modulo k). Again applying this observation repeatedly starting from a_0, we get $N^+_G(a_0) \setminus V(C) \subseteq N^+_G(a_{k-1}) \setminus V(C) \subseteq N^+_G(a_{k-2}) \setminus V(C) \subseteq \cdots \subseteq N^+_G(a_2) \setminus V(C) \subseteq N^+_G(a_1) \setminus V(C) \subseteq N^+_G(a_0) \setminus V(C)$. This shows that for any two vertices $a_i, a_j \in V(C)$, we have $N^+_G(a_i) \setminus V(C) = N^+_G(a_j) \setminus V(C)$ and $N^-_G(a_i) \setminus V(C) = N^-_G(a_j) \setminus V(C)$, implying that $V(C)$ forms a module in G. $\square$

The relation R and the digraph $\Delta(G)$ Let M be the set $\{(x, y) \colon x \mathrel{---} y\}$. We define a relation R on M as follows. For distinct $(a, b), (c, d) \in M$, we say that $(a, b)R(c, d)$ if and only if there exists the four cycle $a \rightarrow c \twoheadrightarrow b \rightarrow d \twoheadrightarrow a$ in G (refer to Figure 7.4). Note that $(a, b)R(c, d)$ if and only if $(b, a)R(d, c)$. Following Fidler and Yuster [41], we now define an auxiliary digraph $\Delta(G)$ whose vertices are the missing edges of G. This graph has the vertex set $V(\Delta(G)) = \{\{a, b\} \colon a \mathrel{---} b\}$ and edge set $E(\Delta(G)) = \{(\{a, b\}, \{c, d\}) \colon (a, b)R(c, d)\}$. In other words, there is an edge between vertices $\{a, b\}$ and $\{c, d\}$ in $\Delta(G)$ if and only if either $(a, b)R(c, d)$ or $(a, b)R(d, c)$. Note that from the definition of R, we cannot have both $(a, b)R(c, d)$ and $(a, b)R(d, c)$.

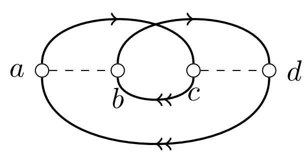


Figure 7.4: Situation that leads to $(a, b)R(c, d)$.

Lemma 24 ([41]). *For any vertex $e \in V(\Delta(G))$, we have $|N^+(e)| \leq 1$ and $|N^-(e)| \leq 1$.*

Proof. Let $e = \{a, b\}$. Suppose that it has two out-neighbors in $\Delta(G)$, say $e_1 = \{c_1, d_1\}$, $e_2 = \{c_2, d_2\}$. Recalling the definition of $\Delta(G)$, we can assume without loss of generality that $(a, b)R(c_1, d_1)$ and $(a, b)R(c_2, d_2)$. That is, we have $a \rightarrow c_1 \twoheadrightarrow b \rightarrow d_1 \twoheadrightarrow a$ and $a \rightarrow c_2 \twoheadrightarrow b \rightarrow d_2 \twoheadrightarrow a$ in G. As d_1 is already a non-neighbor of c_1, we cannot have $c_1 \mathrel{---} d_2$. Now if $c_1 \rightarrow d_2$ then we have the directed triangle $a \to c_1 \to d_2 \twoheadrightarrow a$ containing a special arc, which is a contradiction. Similarly, if $d_2 \to c_1$ then $b \to d_2 \to c_1 \twoheadrightarrow b$ is a directed triangle containing a special arc, which is again a contradiction. Thus, $|N^+(e)| \leq 1$.

Now suppose $e = \{a, b\}$ has two in-neighbors in $\Delta(G)$, say $e_1 = \{c_1, d_1\}$, $e_2 = \{c_2, d_2\}$. Again, we can assume without loss of generality that $(c_1, d_1)R(a, b)$ and $(c_2, d_2)R(a, b)$. Then we have $c_1 \to a \twoheadrightarrow d_1 \to b \twoheadrightarrow c_1$ and $c_2 \to a \twoheadrightarrow d_2 \to b \twoheadrightarrow c_2$ in G. As before, we cannot have $c_1 \dashrightarrow d_2$. If $c_1 \to d_2$ then we have the directed triangle $c_1 \to d_2 \to b \twoheadrightarrow c_1$ containing a special arc and if $d_2 \to c_1$, we have another directed triangle $d_2 \to c_1 \to a \twoheadrightarrow d_2$ containing a special arc. Since we have a contradiction in both cases, we conclude that $|N^-(e)| \leq 1$. $\square$

Therefore, $\Delta(G)$ is a disjoint union of directed paths and directed cycles. Let $\mathcal{P}$ denote the collection of these directed paths and $\mathcal{C}$ denote the collection of these directed cycles.

For a cycle $Q \in \mathcal{C}$, we let $\Gamma(Q) = \bigcup_{\{u,v\} \in V(Q)} \{u, v\}$. That is, if $Q = \{a_1, b_1\}\{a_2, b_2\} \cdots \{a_t, b_t\}$ $\{a_1, b_1\}$, then $\Gamma(Q) = \{a_1, b_1, a_2, b_2, \ldots, a_t, b_t\}$.

Lemma 25. *Let $Q \in \mathcal{C}$. Then there exists a special cycle C in G such that $V(C) = \Gamma(Q)$.*

Proof. Let $Q = \{a_1, b_1\}\{a_2, b_2\} \cdots \{a_k, b_k\}\{a_1, b_1\}$. Note that $a_i \dashrightarrow b_i$, for $1 \leq i \leq k$. We shall assume that k is even as the case when k is odd is similar. Also, we can assume without loss of generality that for every $i \in \{1, 2, \ldots, k-1\}$, $(a_i, b_i)R(a_{i+1}, b_{i+1})$ (since we can always exchange the labels of a_i and b_i, if required, so that this condition is satisified). Then by the definition of R, we have $a_i \to a_{i+1} \twoheadrightarrow b_i \to b_{i+1} \twoheadrightarrow a_i$ for each $i \in \{1, 2, \ldots, k-1\}$. Now if $(a_k, b_k)R(a_1, b_1)$ then we have $a_k \to a_1 \twoheadrightarrow b_k \to b_1 \twoheadrightarrow a_k$ (so $k > 2$, implying that $k \geq 4$). This together with the previous observation implies that $C = a_1 \twoheadrightarrow b_k \twoheadrightarrow a_{k-1} \twoheadrightarrow b_{k-2} \twoheadrightarrow a_{k-3} \twoheadrightarrow \cdots \twoheadrightarrow b_2 \twoheadrightarrow a_1$ (as k is even) is a special cycle in G, which contains only those a_i's where i is odd and those b_i's where i is even. This contradicts Corollary 13(a), as for any odd i, the only non-neighbor b_i of a_i is not contained in C. Therefore, we have $(a_k, b_k)R(b_1, a_1)$. Then, $a_k \to b_1 \twoheadrightarrow b_k \to a_1 \twoheadrightarrow a_k$, which when combined with the previous observations gives us that $C = a_1 \twoheadrightarrow a_k \twoheadrightarrow b_{k-1} \twoheadrightarrow a_{k-2} \twoheadrightarrow b_{k-3} \twoheadrightarrow \cdots \twoheadrightarrow a_2 \twoheadrightarrow b_1 \twoheadrightarrow b_k \twoheadrightarrow a_{k-1} \twoheadrightarrow b_{k-2} \twoheadrightarrow a_{k-3} \twoheadrightarrow \cdots \twoheadrightarrow b_2 \twoheadrightarrow a_1$ is a special cycle in G with $V(C) = \Gamma(Q)$. $\square$

Corollary 14. *Let $Q \in \mathcal{C}$ and $u \in \Gamma(Q)$. Then:*

(a) *there exists $v \in \Gamma(Q)$ such that $u \dashrightarrow v$, and*

(b) *$\Gamma(Q)$ forms a module in G.*

Proof. The proof of (a) is immediate from Lemma 25 and Corollary 13(a). Similarly, (b) is a direct consequence of Lemma 25 and Lemma 23(c). $\square$

Lemma 26. *Let $Q \in \mathcal{C}$. Then for each $u \in \Gamma(Q)$, we have $|N_G^+(u) \cap \Gamma(Q)| = |N_G^{++}(u) \cap \Gamma(Q)|$.*

Proof. As $Q \in \mathcal{C}$, by Lemma 25 there exists a special cycle C in G such that $V(C) = \Gamma(Q)$. Let this cycle be $C = a_0 \twoheadrightarrow a_1 \twoheadrightarrow a_2 \twoheadrightarrow \cdots \twoheadrightarrow a_{2l-1} \twoheadrightarrow a_0$ (note that by Lemma 23(a), C has even length; also note that $l \geq 2$). Consider a vertex $a_i \in V(C)$. By Lemma 23(b), we have $a_i \dashrightarrow a_{i+l}$ and by Corollary 13(b), $N_G^+(a_i) \cap V(C) = \{a_{i+1}, a_{i+2}, \ldots, a_{i+l-1}\}$ (subscripts modulo $2l$). Recalling that $V(C) = \Gamma(Q)$, we now get $|N_G^+(a_i) \cap \Gamma(Q)| = l - 1$. Now, consider any $a_p \in \{a_{i+l}, a_{i+l+1}, \ldots, a_{i+2l-2} = a_{i-2}\}$. Clearly, $a_p \notin N_G^+(a_i)$. Note that for any choice of a_p, the vertex $a_{p+l+1} \in N_G^+(a_i) \cap V(C)$. By Lemma 23$(b)$, we have that $a_p \dashrightarrow a_{p+l}$. Now applying Corollary 13(b) to a_p, we have that $a_{p+l+1} \in N_G^-(a_p) \cap V(C)$. This gives us that $a_p \in N_G^{++}(a_i) \cap \Gamma(Q)$ for each choice of $a_p \in \{a_{i+l}, a_{i+l+1}, \ldots, a_{i+2l-2} = a_{i-2}\}$, implying that $|N_G^{++}(a_i) \cap \Gamma(Q)| \geq l - 1$. Noting that the vertex $a_{i-1} \notin N_G^{++}(a_i)$ (as $a_{i-1} \twoheadrightarrow a_i$), we can now conclude $|N_G^{++}(a_i) \cap \Gamma(Q)| = l - 1 = |N_G^+(a_i) \cap \Gamma(Q)|$. $\square$

Unforced and singly-forced missing edges We now label some missing edges of G as unforced and some others as singly-forced.

Definition 33 (Singly forced missing edge)**.** *A missing edge $e = a \dashrightarrow b$ is said to be* singly-forced *if exactly one of the following conditions hold.*

(1) *There exists $v \in V(G)$ such that $b \twoheadrightarrow v \rightarrow a$ in G.*

(2) *There exists $u \in V(G)$ such that $a \twoheadrightarrow u \rightarrow b$ in G.*

If (1) *holds then we say that e is* forced in the direction b to a, *and if* (2) *holds then we say that e is* forced in the direction a to b. *If neither* (1) *nor* (2) *hold, then e is* unforced. *Note that it is possible for a missing edge to be forced in both directions.* (See Figure 7.5).

Proposition 10. *Let $e = a \dashrightarrow b$. If there exist $u, v \in V(G)$ such that $b \twoheadrightarrow v \rightarrow a$ and $a \twoheadrightarrow u \rightarrow b$, then $(u, v)R(b, a)$. Consequently, if any missing edge is forced in both directions in G, then it has an in-neighbor in $\Delta(G)$.*

Proof. Note that $u \neq v$. Now, if $v \rightarrow u$ or $u \rightarrow v$, then $u \rightarrow b \twoheadrightarrow v \rightarrow u$ or $v \rightarrow a \twoheadrightarrow u \rightarrow v$ would form a directed triangle containing a special arc, which is a contradiction. Therefore, $u \dashrightarrow v$. Then, the fact that $u \rightarrow b \twoheadrightarrow v \rightarrow a \twoheadrightarrow u$ implies that $(u, v)R(b, a)$ and hence $\{u, v\}$ is an in-neighbor of e in $\Delta(G)$. $\square$

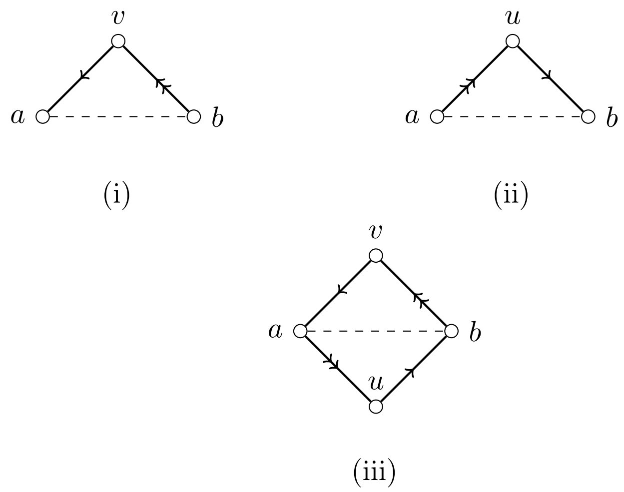


Figure 7.5: A singly forced missing edge is either of the form (i) or (ii) and (iii) shows a missing edge that is forced in both the directions

Proposition 11. *Every singly-forced missing edge is the starting vertex of some path in $\mathcal{P}$.*

Proof. Let $a \text{---} b$ be a singly-forced missing edge. It is enough to prove that $\{a, b\}$ doesn't have any in-neighbor in $\Delta(G)$. Assume to the contrary that $\{a, b\}$ has an in-neighbor $\{c, d\}$ in $\Delta(G)$. Then by definition of $\Delta(G)$ we can assume without loss of generality that $(c, d)R(a, b)$, i.e., there exists a cycle $c \rightarrow a \twoheadrightarrow d \rightarrow b \twoheadrightarrow c$ in G. Note that now we have both $b \twoheadrightarrow c \rightarrow a$ and $a \twoheadrightarrow d \rightarrow b$, implying that both conditions (1) and (2) of Definition 33 hold. This contradicts the fact that $a \text{---} b$ is a singly-forced missing edge. $\square$

Completions and special in-neighbors A tournament T is said to be a *completion* of G if $V(G) = V(T)$ and $E(G) \subseteq E(T)$. It is easy to see that a completion of G can be obtained by "orienting" every missing edge of G, i.e., by adding an oriented edge in place of each missing edge of G. Our strategy will be to show that there exists a way to orient the missing edges of G so that the resulting completion T of G has the property that the feed vertex of any median order of T has a large second neighborhood not just in T, but also in G. A missing edge $a \text{---} b$ of G that has been oriented from a to b in T is denoted by $a \dashrightarrow b$.

Definition 34 (Type-I and Type-II special in-neighbors)**.** *Given a completion T of G and a vertex $v \in V(T)$, we say that an in-neighbor b of v is a* special in-neighbor *if $b \twoheadrightarrow v$ and $b \in N_T^{++}(v)$. Further, we say that a special in-neighbor b of v is of* Type-I *if there exists $a \in V(T)$ such that $v \rightarrow a \dashrightarrow b \twoheadrightarrow v$. Similarly, we say that a special in-neighbor b of v is of* Type-II *if there exists $a \in V(T)$ such that $v \dashrightarrow a \rightarrow b \twoheadrightarrow v$. Note that any special in-neighbor of v is either Type-I or Type-II or both.* (See Figure 7.6).

(i) (ii)

Figure 7.6: (i) b is a Type-I in-neighbor of v and (ii) b is a Type-II in-neighbor of v

Lemma 27. *Let T be a completion of G and let $v \in V(T)$. If there exists a vertex x such that $x \in N_T^{++}(v) \setminus N_G^{++}(v)$ then x is a special in-neighbor of v.*

Proof. Consider $x \in N_T^{++}(v) \setminus N_G^{++}(v)$. As $x \in N_T^{++}(v)$, $x \in N_T^{-}(v)$, implying that we have either $x \to v$ or $x \dashrightarrow v$. Furthermore, there exists $a \in V(T)$ such that $a \in N_T^{+}(v) \cap N_T^{-}(x)$. Since $x \notin N_G^{++}(v)$, we know that either $v \dashrightarrow a$ or $a \dashrightarrow x$. As the missing edges of G form a matching, this implies that $x \to v$. Again using the fact that $x \notin N_G^{++}(v)$, we conclude that $x \twoheadrightarrow v$. This shows that x is a special in-neighbor of v. $\square$

Lemma 28. *Let T be a completion of G and L a median order of T such that the feed vertex d of L does not have a special in-neighbor of Type-I. Then d is a vertex with large second neighborhood in G.*

Proof. We claim that there exists a completion T' of G such that L is a median order of T' and d has no special in-neighbors in T'. If there does not exist a vertex $a \in V(T)$ such that $d \dashrightarrow a$, then clearly $T' = T$ is a completion of G satisfying our requirements. So we shall assume that there exists $a \in V(T)$ with $d \dashrightarrow a$. Now, consider the completion T' of G obtained from T by reorienting the missing edge $d \dashrightarrow a$ as $a \dashrightarrow d$. By Proposition 6, L is a median order of T' as well. Further, it can be easily seen that d does not have any special in-neighbors of Type-I in T' either. As the only missing edge incident on d is oriented towards d in T', d does not have any special in-neighbors of Type-II in T'. This proves our claim.

By Lemma 27 applied on T' and L, we have $N_{T'}^{++}(d) \subseteq N_G^{++}(d)$. By Theorem 33, $|N_G^{+}(d)| = |N_{T'}^{+}(d)| \leq |N_{T'}^{++}(d)|$ (the first equality is because $a \dashrightarrow d$ in T'). Combining this with the previous observation, we have $|N_G^{+}(d)| \leq |N_G^{++}(d)|$. $\square$

Safe completions We now construct a completion T of G by orienting the missing edges of G in a particular fashion. We start by orienting the missing edges that are the starting vertices of paths in $\mathcal{P}$. Among them, we orient the singly-forced missing edges in the direction in which they are forced and the others in an arbitrary direction. Then, repeatedly do the following until every missing edge that is in a path in $\mathcal{P}$ is oriented: if $a \dashrightarrow b$ is unoriented and has an in-neighbor $\{c,d\}$ in $\Delta(G)$ which has been oriented as $c \dashrightarrow d$, then orient $a \dashrightarrow b$ if $(c,d)R(a,b)$ and orient it as $b \dashrightarrow a$ if $(c,d)R(b,a)$. The remaining unoriented missing edges are those that belong to cycles in $\mathcal{C}$. Orient them in arbitrary directions. By Proposition 11, this strategy orients every singly-forced missing edge in the direction in which it is forced.

Definition 35 (Safe completion)**.** *A completion T of G is said to be* safe *if it can be obtained from G by applying the above strategy. Formally, a completion T of G is a safe completion if it satisfies the following two conditions:*

(1) *If $a \text{---} b$ is a singly-forced missing edge that is forced in the direction from a to b, then $a \dashrightarrow b$ in T, and*

(2) *if $\{a,b\}$ does not lie in any cycle in $\mathcal{C}$, $(c,d)R(a,b)$ and $c \dashrightarrow d$ in T, then $a \dashrightarrow b$ in T.*

Recall that $(c,d)R(a,b)$ if and only if $(d,c)R(b,a)$. Therefore, if $\{a,b\}, \{c,d\}$ are two missing edges that do not lie on any cycle in $\mathcal{C}$ and $(c,d)R(a,b)$, then in any safe completion, $c \dashrightarrow d$ if and only if $a \dashrightarrow b$.

As the above strategy of constructing a safe completion of G never fails, we have the following remark.

Remark 6. *Every oriented graph whose missing edges form a matching has a safe completion.*

Lemma 29. *Let T be a safe completion of G. Let $v \in V(T)$ and b be a Type-I special in-neighbor of v. Then there exist $a, u \in V(T)$ such that $v \to a \dashrightarrow b \twoheadrightarrow v$, $a \twoheadrightarrow u \to b$ and $u \text{---} v$. Moreover, b is the only Type-I special in-neighbor of v.*

Proof. As b is a Type-I special in-neighbor of v, there exists $a \in V(T)$ such that $v \to a \dashrightarrow b \twoheadrightarrow v$ in T. Then by Definition 33, $a \text{---} b$ is forced in the direction b to a. But as we have $a \dashrightarrow b$ in T, and every singly-forced missing edge of G was oriented in T in the direction in which it was forced (as T is a safe completion), it must be the case that $a \text{---} b$ is also forced in the direction a to b. That is, there exists $u \in V(T)$ such that $a \twoheadrightarrow u \to b$ (refer to Definition 33). Using Proposition 10, we can now conclude that $(u,v)R(b,a)$, which further implies that $u \text{---} v$. If there exists a Type-I special in-neighbor b' of d such that $b' \neq b$, then the same arguments can be used to infer that there exist $a', u' \in V(T)$ such that $(u',v)R(b',a')$ (which means that $u' \text{---} v$). Since v has at most one non-neighbor, we have that $u' = u$, which gives $(u,v)R(b',a')$. As it can be easily seen that $\{a',b'\} \neq \{a,b\}$, the missing edge $\{u,v\}$ has more than one out-neighbor in $\Delta(G)$, which is a contradiction to Lemma 24. Hence b is the only Type-I special in-neighbor of v. $\square$

Lemma 30. *Let T be a safe completion of G and let L be a median order of T with feed vertex d. If d has a Type-I special in-neighbor b and there exists $w \in V(T)$ such that $d \dashrightarrow w$, then:*

(a) $N_T^{++}(d) \setminus \{b\} \subseteq N_G^{++}(d)$, and

(b) d is a vertex with large second neighborhood in G.

Proof. By Lemma 29, there exist $a, u \in V(T)$ such that $d \rightarrow a \dashrightarrow b \twoheadrightarrow d$, $a \twoheadrightarrow u \rightarrow b$ and $u \text{ - - - } d$. As the only non-neighbor of d is w, we have $u = w$.

(a) Consider a vertex $x \in N_T^{++}(d) \setminus \{b\}$. Suppose for the sake of contradiction that $x \notin N_G^{++}(d)$. Then by Lemma 27, we know that x is a special in-neighbor of d. Since $x \neq b$, we know by Lemma 29 that x cannot be a Type-I special in-neighbor of d. Therefore, x is a Type-II special in-neighbor of d, i.e., $d \dashrightarrow w \rightarrow x \twoheadrightarrow d$ (as w is the only non-neighbor of d). It is easily verified that $a \neq x$. Further, $\{a, x\}$ cannot be a missing edge since $a \text{ - - - } b$ and $x \neq b$. If $x \rightarrow a$ or $a \rightarrow x$, then either $a \twoheadrightarrow u = w \rightarrow x \rightarrow a$ or $d \rightarrow a \rightarrow x \twoheadrightarrow d$ would be a directed triangle containing a special arc, which is a contradiction. This proves (a).

(b) We have $|N_G^+(d)| = |N_T^+(d)| - 1 \leq |N_T^{++}(d)| - 1 = |N_T^{++}(d) \setminus \{b\}| \leq |N_G^{++}(d)|$ (the first equality is because $d \dashrightarrow w$, the second inequality by Theorem 33, the third equality is because $b \in N_T^{++}(d)$, and the fourth inequality by (a)). $\square$

Consider a module I in G such that $|I| \geq 2$ and a vertex $v \in I$. Clearly, any non-neighbor of v outside I has to be a non-neighbor of every vertex in I. As $|I| \geq 2$ and the missing edges of G form a matching, this can only mean that v has no non-neighbors outside I. We thus have the following remark.

Remark 7. *If I is a module in G such that $|I| \geq 2$ and $v \in I$, then v has no non-neighbors outside I.*

Lemma 31. *Let T be a safe completion of G and let I be a module in G with $|I| \geq 2$. Then for any $v \in I$, $N_T^{++}(v) \setminus I \subseteq N_G^{++}(v) \setminus I$.*

Proof. First, suppose that there exists a Type-I special in-neighbor x of v outside I. By Lemma 29, there exists a vertex $u \in V(T)$ such that $u \text{ - - - } v$ and $u \rightarrow x$. By Remark 7, $u \in I$. Now we have $x \rightarrow v$ and $u \rightarrow x$, which contradicts the fact that u and v belong to the module I in G and x is outside that module. Therefore, v has no Type-I special in-neighbors outside I. Next, suppose that there exists a Type-II special in-neighbor x of v outside I. Then, there exists a vertex y such that $v \dashrightarrow y \rightarrow x \twoheadrightarrow v$. By Remark 7, we know that $y \in I$. Then we have $x \rightarrow v$ and $y \rightarrow x$, which contradicts the fact that v and y belong to the module I in G (recall that x is outside I). Therefore, we can conclude that v has no special in-neighbors outside I. This implies, by Lemma 27, that $N_T^{++}(v) \setminus I \subseteq N_G^{++}(v) \setminus I$. $\square$

Corollary 15. *Let T be a safe completion of G and let d be the feed vertex of some median order of T. Let I be a module in G containing d where $|I| \geq 2$. Then for any $v \in I$, $|N_G^+(v) \setminus I| \leq |N_G^{++}(v) \setminus I|$.*

Proof. It is easy to see that as the missing edges of G form a matching, every module in G is also a module in T. Therefore I is a module in T containing d. Then we have from Proposition 9 and Lemma 31 that $|N_G^+(v) \setminus I| \leq |N_T^+(v) \setminus I| \leq |N_T^{++}(v) \setminus I| \leq |N_G^{++}(v) \setminus I|$. $\square$

Prime vertices We define

$$I(u) = \begin{cases} \Gamma(Q) & \text{if } \exists Q \in \mathcal{C} \text{ such that } u \in \Gamma(Q), \\ \{u\} & \text{otherwise.} \end{cases}$$

Note that as any vertex u can be a part of at most one missing edge, there can be at most one cycle $Q \in \mathcal{C}$ such that $u \in \Gamma(Q)$, and therefore $I(u)$ is well defined. We define a vertex u in G to be *prime*, if $I(u) = \{u\}$; in other words, a vertex u is said to be prime if $u \notin \Gamma(Q)$ for any $Q \in \mathcal{C}$.

Note that if u is prime, we have $I(u) = \{u\}$ and therefore, $|N_G^+(u) \cap I(u)| = |N_G^{++}(u) \cap I(u)| = 0$. On the other hand, if $u \in \Gamma(Q)$ for some $Q \in \mathcal{C}$, then $I(u) = \Gamma(Q)$, and by Lemma 26, we get that $|N_G^+(u) \cap I(u)| = |N_G^{++}(u) \cap I(u)|$. We thus have the following.

Remark 8. *For any vertex $u \in V(G)$, $|N_G^+(u) \cap I(u)| = |N_G^{++}(u) \cap I(u)|$.*

Theorem 34. *Let d be the feed vertex of some median order of a safe completion T of G. Then every vertex in $I(d)$ has a large second neighborhood in G.*

Proof. Suppose that d is prime. Then, $I(d) = \{d\}$. If d has no special in-neighbor of Type-I in T, then we are done by Lemma 28. So let us assume that d has a special in-neighbor b of Type-I in T. Then by Lemma 29, there exist $a, u \in V(T)$ such that $d \rightarrow a \dashrightarrow b \twoheadrightarrow d$, $a \twoheadrightarrow u \rightarrow b$, where $u \text{---} d$. This means that $(u, d)R(b, a)$. If $u \dashrightarrow d$, then since T is a safe completion of G, the fact that $a \dashrightarrow b$ implies that $\{u, d\}$ and $\{b, a\}$ lie in some cycle in $\mathcal{C}$, contradicting the assumption that d is prime. Therefore, we have $d \dashrightarrow u$. Then we are done by Lemma 30(b).

Next, consider the case when d is not prime, i.e. $d \in \Gamma(Q)$ for some $Q \in \mathcal{C}$. Note that we then have $I(d) = \Gamma(Q)$ and therefore, $|I(d)| \geq 2$. Consider any vertex $v \in I(d)$. As $I(d) = \Gamma(Q)$ is a module (Corollary 14(b)), we have by Corollary 15 that $|N_G^+(v) \setminus I(d)| \leq |N_G^{++}(v) \setminus I(d)|$. By

Remark 8, $|N_G^+(v) \cap I(d)| = |N_G^{++}(v) \cap I(d)|$. We now have $|N_G^+(v)| = |N_G^+(v) \setminus I(d)| + |N_G^+(v) \cap I(d)| \leq |N_G^{++}(v) \setminus I(d)| + |N_G^{++}(v) \cap I(d)| = |N_G^{++}(v)|$. Hence the theorem. $\square$

As $u \in I(u)$ for every vertex $u \in V(G)$, Remark 6 and Theorem 34 give us the following corollary.

Corollary 16. *Every oriented graph whose missing edges form a matching contains a vertex with a large second neighborhood.*

A useful lemma about special arcs We now state a property of special arcs that are "reverse arcs" in a median order, which will be useful for deriving the results in the next section.

Definition 36 (Reverse special arc)**.** *Given a median order $L = (x_1, x_2, \dots, x_n)$ of any completion T of G, a special arc $x_j \twoheadrightarrow x_i$ is said to be a* reverse special arc *in (T, L) if $i < j$.*

Lemma 32. *Let $L = (x_1, x_2, \dots, x_n)$ be a median order of a completion T of G and $x_j \twoheadrightarrow x_i$ be a reverse special arc in (T, L). Then at least one of the following conditions hold:*

(*a*) *there exists x_k such that $x_i \dashrightarrow x_k \rightarrow x_j$, where $i < k < j$, or*

(*b*) *there exists x_l such that $x_i \rightarrow x_l \dashrightarrow x_j$, where $i < l < j$.*

Moreover, if exactly one of the above conditions holds, then $L' = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_j, x_i, x_{j+1}, \dots, x_n)$ is also a median order of T.

Proof. For the purposes of this proof, for $u \in \{x_i, x_{i+1}, \dots, x_j\}$, we shall abbreviate $N_T^+(u) \cap \{x_i, x_{i+1}, \dots, x_j\}$ and $N_T^-(u) \cap \{x_i, x_{i+1}, \dots, x_j\}$ to just $N_{i,j}^+(u)$ and $N_{i,j}^-(u)$ respectively. By Proposition 4, we have

$$\left|N_{i,j}^+(x_i)\right| \geq \frac{j-i}{2} \quad \text{and} \quad \left|N_{i,j}^-(x_j)\right| \geq \frac{j-i}{2} \tag{7.1}$$

Alternatively,

$$\left|N_{i,j}^-(x_i)\right| \leq \frac{j-i}{2} \quad \text{and} \quad \left|N_{i,j}^+(x_j)\right| \leq \frac{j-i}{2} \tag{7.2}$$

We shall first make an observation about any vertex $x_p \in N_{i,j}^+(x_i) \setminus N_{i,j}^+(x_j)$. Clearly, $x_p \in N_{i,j}^+(x_i) \cap N_{i,j}^-(x_j)$ (recall that $x_j \twoheadrightarrow x_i$). Note that either $x_i \dashrightarrow x_p$ or $x_p \dashrightarrow x_j$, as otherwise $x_i \rightarrow x_p \rightarrow x_j \twoheadrightarrow x_i$ would form a directed triangle containing a special arc, which is a contradiction. Since the missing edges of G form a matching, this implies that either $x_i \dashrightarrow x_p \rightarrow x_j$ or $x_i \rightarrow x_p \dashrightarrow x_j$.

Suppose that neither of the conditions in the lemma hold. Then from the above observation, it is clear that $N^+_{i,j}(x_i) \subseteq N^+_{i,j}(x_j)$. Note that $x_i \notin N^+_{i,j}(x_i)$ but $x_i \in N^+_{i,j}(x_j)$. Therefore we have, $\left|N^+_{i,j}(x_i)\right| < \left|N^+_{i,j}(x_j)\right| \leq \frac{j-i}{2}$ (by (7.2)), which contradicts (7.1). Therefore at least one of the conditions (a) or (b) should hold.

Now suppose that exactly one of the conditions (a) or (b) holds. Note first that from the previous observation and the fact that the missing edges of G form a matching, it follows that if there exist two distinct vertices x_p, x_q in $N^+_{i,j}(x_i) \setminus N^+_{i,j}(x_j)$, then $x_i \dashrightarrow x_p \rightarrow x_j$ and $x_i \rightarrow x_q \dashrightarrow x_j$, implying that both conditions hold. Therefore, there is exactly one vertex in $N^+_{i,j}(x_i) \setminus N^+_{i,j}(x_j)$, i.e., $|N^+_{i,j}(x_i) \setminus N^+_{i,j}(x_j)| = 1$. Since $x_i \in N^+_{i,j}(x_j) \setminus N^+_{i,j}(x_i)$, we have that $|N^+_{i,j}(x_i) \setminus (N^+_{i,j}(x_j) \setminus \{x_i\})| = 1$. This means that $|N^+_{i,j}(x_i)| - (|N^+_{i,j}(x_j)| - 1) \leq 1$, implying that $|N^+_{i,j}(x_i)| \leq |N^+_{i,j}(x_j)|$. Hence, $\frac{j-i}{2} \leq \left|N^+_{i,j}(x_i)\right| \leq \left|N^+_{i,j}(x_j)\right| \leq \frac{j-i}{2}$ (from (7.1) and (7.2)). Therefore, we have $\left|N^+_{i,j}(x_i)\right| = \frac{j-i}{2} = \left|N^-_{i,j}(x_i)\right|$ (which means that $j-i$ is even). Then by Proposition 5(a), $L' = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_j, x_i, x_{j+1}, \ldots, x_n)$ is also a median order of T. □

In Section 7.2.4, we shall use the concepts introduced so far in order to generalize Corollary 16 to show that in any graph whose missing edges can be partitioned into a matching and a star, there exists a vertex with a large second neighborhood. We need the notion of "sedimentation" of median orders, first introduced by Havet and Thomasse [65], to derive our result.

7.2.3 Sedimentation of a good median order

Ghazal modified the notion of sedimentation of median orders to apply to good median orders. We slightly modify this so as to redefine sedimentation without referring to the "good" and "bad" vertices that appear in the work of Havet and Thomasse and Ghazal.

Suppose that $L = (x_1, x_2, \ldots, x_n)$ is a good median order of a tournament T with respect to $\mathcal{I}$, where $\mathcal{I}$ is a partition of $V(T)$ into modules. Let I be the set in $\mathcal{I}$ containing x_n and $t = |I|$. Then $I = \{x_{n-t+1}, x_{n-t+2}, \ldots, x_n\}$. Recall that by Proposition 9, $|N^+(x_n) \setminus I| \leq |N^{++}(x_n) \setminus I|$. Then the *sedimentation of L with respect to $\mathcal{I}$*, denoted by $Sed_{\mathcal{I}}(L)$, is an ordering of $V(T)$ that is defined in the following way. If $|N^+(x_n) \setminus I| < |N^{++}(x_n) \setminus I|$, then $Sed_{\mathcal{I}}(L) = L$. If $|N^+(x_n) \setminus I| = |N^{++}(x_n) \setminus I|$, then $Sed_{\mathcal{I}}(L)$ is defined as follows. Let $b_1, b_2, \ldots, b_k$ be the vertices in $N^-(x_n) \setminus N^{++}(x_n)$ which are outside I and $v_1, v_2, \ldots, v_{n-t-k}$ the vertices in $N^+(x_n) \cup N^{++}(x_n)$ which are outside I, both enumerated in the order in which they appear in L (note that in any tournament, $N^{++}(u) \subseteq N^-(u)$ for any vertex u in it). Then $Sed_{\mathcal{I}}(L)$ is the order $(b_1, b_2, \ldots, b_k, x_{n-t+1}, x_{n-t+2}, \ldots, x_n, v_1, v_2, \ldots, v_{n-t-k})$.

We shall now prove the following proposition and theorem which are adapted from the proof of Havet and Thomasse so as to incorporate our slightly changed definition of sedimentation.

Proposition 12. *Let T be a tournament and $L = (x_1, x_2, \ldots, x_n)$ be a median order of T such that $|N^+(x_n)| = |N^{++}(x_n)|$.*

(a) If $N^-(x_n) = N^{++}(x_n)$, then $(x_n, x_1, x_2, \ldots, x_{n-1})$ is a median order of T, and

(b) if $N^-(x_n) \setminus N^{++}(x_n) \neq \emptyset$ and x_i is the vertex in $N^-(x_n) \setminus N^{++}(x_n)$ that occurs first in L, then $(x_i, x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is a median order of T.

Proof. (a) Since $|N^+(x_n)| = |N^{++}(x_n)|$ and $N^-(x_n) = N^{++}(x_n)$, we have $|N^+(x_n)| = |N^-(x_n)|$. Therefore, by Proposition 5(b) applied on x_1 and x_n, we have that $(x_n, x_1, x_2, \ldots, x_{n-1})$ is a median order of T.

(b) Let $D = \{x_1, x_2, \ldots, x_{i-1}\}$ and $U = \{x_i, x_{i+1}, \ldots, x_n\}$. By Proposition 3($a$), $(x_i, x_{i+1}, \ldots, x_n)$ is a median order of the subtournament $T[U]$ of T. Applying Theorem 33 to this median order of the tournament $T[U]$, we have $|N_T^+(x_n) \cap U| = |N_{T[U]}^+(x_n)| \leq |N_{T[U]}^{++}(x_n)| \leq |N_T^{++}(x_n) \cap U|$. This together with the fact that, $|N_T^+(x_n)| = |N_T^+(x_n) \cap D| + |N_T^+(x_n) \cap U|$, $|N_T^{++}(x_n)| = |N_T^{++}(x_n) \cap D| + |N_T^{++}(x_n) \cap U|$ and $|N_T^+(x_n)| = |N_T^{++}(x_n)|$ (assumption of the lemma) implies that $|N_T^+(x_n) \cap D| \geq |N_T^{++}(x_n) \cap D|$. As $x_i \in N_T^-(x_n) \setminus N_T^{++}(x_n)$, we have that $N_T^+(x_n) \cap D \subseteq N_T^+(x_i) \cap D$ and $N_T^-(x_i) \cap D \subseteq N_T^-(x_n) \cap D$. As x_i is the first vertex in L that belongs to $N_T^-(x_n) \setminus N_T^{++}(x_n)$, we also have that $N_T^-(x_n) \cap D = N_T^{++}(x_n) \cap D$. By Proposition 4($a$) applied to x_1 and x_i, we get $|N_T^+(x_i) \cap D| \leq |N_T^-(x_i) \cap D|$. Combining everything, we have $|N_T^+(x_n) \cap D| \leq |N_T^+(x_i) \cap D| \leq |N_T^-(x_i) \cap D| \leq |N_T^-(x_n) \cap D| = |N_T^{++}(x_n) \cap D|$. Recalling our previous observation that $|N_T^+(x_n) \cap D| \geq |N_T^{++}(x_n) \cap D|$, we then have $|N_T^+(x_i) \cap D| = |N_T^-(x_i) \cap D|$. Now from Proposition 5(b) applied on x_1 and x_i, we get that $(x_i, x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is a median order of T. □

Following is the theorem from [65] that we need.

Theorem 35 ([65])**.** *Let T be a tournament and $L = (x_1, x_2, \ldots, x_n)$ be a median order of it such that $|N^+(x_n)| = |N^{++}(x_n)|$. Let $b_1, b_2, \ldots, b_k$ be the vertices in $N^-(x_n) \setminus N^{++}(x_n)$ and $v_1, v_2, \ldots, v_{n-k-1}$ be the vertices in $N^+(x_n) \cup N^{++}(x_n)$, both enumerated in the order in which they appear in L. Then $(b_1, b_2, \ldots, b_k, x_n, v_1, v_2, \ldots, v_{n-k-1})$ is a median order of T.*

Proof. We prove this by induction on $|N^-(x_n) \setminus N^{++}(x_n)|$. If $|N^-(x_n) \setminus N^{++}(x_n)| = 0$, then we are done by Proposition 12(a). So let us assume that $N^-(x_n) \setminus N^{++}(x_n) \neq \emptyset$ and that $b_1, b_2, \ldots, b_k, v_1, v_2, \ldots, v_{n-k-1}$ are the vertices as defined in the statement of the theorem. By

Proposition 12(b), we know that $\hat{L} = (x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is a median order of T, where $x_i = b_1$. By Proposition 3(a), we know that $L' = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is a median order of $T' = T - \{b_1\}$. It is easy to see that $N^+_{T'}(x_n) = N^+_T(x_n)$, $N^-_{T'}(x_n) = N^-_T(x_n) \setminus \{b_1\}$ and $N^{++}_{T'}(x_n) = N^{++}_T(x_n)$. Therefore, $|N^+_{T'}(x_n)| = |N^{++}_{T'}(x_n)|$ and $N^-_{T'}(x_n) \setminus N^{++}_{T'}(x_n) = \{b_2, b_3, \ldots, b_k\}$. By the induction hypothesis applied on the tournament T' and the median order L', we get that $(b_2, b_3, \ldots, b_k, x_n, v_1, v_2, \ldots, v_{n-k-1})$ is a median order of T'. Now by Proposition 3(b), we can replace the subsequence $(x_1, \ldots, x_n)$ of $\hat{L}$ with any median order of T' to obtain a median order of T. Therefore, $(b_1, b_2, \ldots, b_k, x_n, v_1, v_2, \ldots, v_{n-k-1})$ is a median order of T. $\square$

Given below is the main theorem that we need for sedimentation of median orders. This is a modification of a result of Ghazal [53] to apply to our version of sedimentation (we again want to avoid using the "good vertices" of Havet and Thomasse).

Theorem 36 ([53])**.** *Let T be a tournament. If $\mathcal{I}$ is a partition of $V(T)$ into modules and L is a good median order of T with respect to $\mathcal{I}$, then $Sed_{\mathcal{I}}(L)$ is also a good median order of T with respect to $\mathcal{I}$.*

Proof. Let $L = (x_1, x_2, \ldots, x_n)$ and let $I \in \mathcal{I}$ be the module containing x_n. If $|N^+_T(x_n) \setminus I| < |N^{++}_T(x_n) \setminus I|$, then $Sed_{\mathcal{I}}(L) = L$ and there is nothing to prove. Therefore, by Proposition 9, we can assume that $|N^+_T(x_n) \setminus I| = |N^{++}_T(x_n) \setminus I|$. Let $t = |I|$. Then $I = \{x_{n-t+1}, x_{n-t+2}, \ldots, x_n\}$. Let $b_1, b_2, \ldots, b_k$ be the vertices outside I that are in-neighbors of x_n but not its second out-neighbors (i.e., $\{b_1, b_2, \ldots, b_k\} = (N^-_T(x_n) \setminus N^{++}_T(x_n)) \setminus I)$, where $0 \le k \le n - t$, and $v_1, v_2, \ldots, v_{n-t-k}$ the vertices in $(N^+_T(x_n) \cup N^{++}_T(x_n)) \setminus I$, both enumerated in the order in which they appear in L.

For ease of notation, we denote x_{n-t+i} by u_i, for each $i \in \{1, 2, \ldots, t\}$. Then $u_1 = x_{n-t+1}$ and $u_t = x_n$. By Proposition 3(a), $L' = (x_1, x_2, \ldots, x_{n-t+1} = u_1)$ is a median order of $T' = T - \{u_2, u_3, \ldots, u_t\}$. As u_1 and x_n belong to the module I of T, $N^+_{T'}(u_1) = N^+_T(u_1) \setminus I = N^+_T(x_n) \setminus I$ and $N^-_{T'}(u_1) = N^-_T(u_1) \setminus I = N^-_T(x_n) \setminus I$. By Proposition 7, we further have $N^{++}_{T'}(u_1) = N^{++}_T(u_1) \setminus I = N^{++}_T(x_n) \setminus I$. Since $|N^+_T(x_n) \setminus I| = |N^{++}_T(x_n) \setminus I|$, it then follows that $|N^+_{T'}(u_1)| = |N^{++}_{T'}(u_1)|$ and that $N^-_{T'}(u_1) \setminus N^{++}_{T'}(u_1) = \{b_1, b_2, \ldots, b_k\}$.

By Theorem 35 applied on T' and L', we get that $(b_1, b_2, \ldots, b_k, u_1, v_1, v_2, \ldots, v_{n-t-k})$ is a median order of T'. From Proposition 3(b), we know that we can replace the subsequence $(x_1, x_2, \ldots, x_{n-t+1} = u_1)$ of L with this new median order of T' to get another median order $(b_1, b_2, \ldots, b_k, u_1, v_1, v_2, \ldots, v_{n-t-k}, u_2, u_3, \ldots, u_t)$ of T. By repeatedly applying Proposition 8(b)

on the median order $(b_1, b_2, \ldots, b_k, u_1, u_2, \ldots, u_i, v_1, v_2, \ldots, v_{n-t-k}, u_{i+1}, u_{i+2}, \ldots, u_t)$ of T and the vertices u_i and u_{i+1}, for each value of i from 1 to $t-1$, we can conclude that $Sed_{\mathcal{I}}(L) = (b_1, b_2, \ldots, b_k, u_1, u_2, \ldots, u_t, v_1, v_2, \ldots, v_{n-t-k})$ is a median order of T.

It only remains to be proven that $Sed_{\mathcal{I}}(L)$ is a good median order of T with respect to $\mathcal{I}$. It can be easily seen that for any $J \in \mathcal{I}$, if there exists $u \in J$ such that $u \in N^-_T(x_n) \setminus N^{++}_T(x_n) = \{b_1, b_2, \ldots, b_k\}$, then $J \subseteq N^-_T(x_n) \setminus N^{++}_T(x_n)$. As $\{u_1, u_2, \ldots, u_t\} = I \in \mathcal{I}$, this implies that every other module in $\mathcal{I}$ is a subset of either $\{b_1, b_2, \ldots, b_k\}$ or $\{v_1, v_2, \ldots, v_{n-t-k}\}$. Since the vertices in each set in $\mathcal{I}$ occur in $Sed_{\mathcal{I}}(L)$ in the same order as they occur in L, and L is a good median order of T with respect to $\mathcal{I}$, we can conclude that the vertices in each module in $\mathcal{I}$ occur consecutively in $Sed_{\mathcal{I}}(L)$ too. □

Stable and periodic median orders Following Ghazal and Havet and Thomasse, we inductively define $Sed^0_{\mathcal{I}}(L) = L$ and for integer $q \geq 1$, $Sed^q_{\mathcal{I}}(L) = Sed_{\mathcal{I}}(Sed^{q-1}_{\mathcal{I}}(L))$. We say that a good median order L of T with respect to some $\mathcal{I}$ is *stable* if there exists integer $q \geq 0$ such that $Sed^{q+1}_{\mathcal{I}}(L) = Sed^q_{\mathcal{I}}(L)$. Otherwise, we say that L is *periodic*.

7.2.4 Tournaments missing a matching and a star

In this section, we shall show that if the missing edges of an oriented graph can be partitioned into a matching and a star, then it contains a vertex with a large second neighborhood. As noted in the beginning, any sink in an oriented graph is a vertex with a large second neighborhood. Therefore, we only need to show the result for graphs that contain no sink. In fact, we show the following stronger result.

Theorem 37. *Let H be an oriented graph that does not contain a sink and $z \in V(H)$ such that $G = H - \{z\}$ is a tournament missing a matching. Then there exists a vertex in $V(G)$ that has a large second neighborhood in both G and H.*

When H is a tournament missing a matching and a star, and H does not contain a sink, we can apply the above theorem taking z to be the center of the star, to obtain the result that there is a vertex other than z having a large second neighborhood in H.

For the remainder of this section, we assume that H is an oriented graph without a sink containing a vertex $z \in V(H)$ such that $G = H - \{z\}$ is a tournament missing a matching.

We say that the "value" of a tournament is the number of forward arcs in any median order of it. Let T be a safe completion of G with largest value. In other words, T is a safe completion of G with smallest feedback arc set (a set of arcs whose removal makes T acyclic). Henceforth, we shall use T to denote such a completion of G. We immediately have the following observation about any median order of T.

Lemma 33. *If $L = (x_1, x_2, \ldots, x_n)$ is a median order of T, then there cannot exist a missing edge $x_j \dashrightarrow x_i$, where $i < j$, such that $\{x_i, x_j\}$ is an isolated vertex of $\Delta(G)$ and $x_i \,\text{---}\, x_j$ is unforced.*

Proof. Let T' be the tournament obtained from T by reversing the arc $x_j \dashrightarrow x_i$. By Proposition 6, L is a median order of T' as well. Also, as $x_i \,\text{---}\, x_j$ is unforced and $\{x_i, x_j\}$ is an isolated vertex in $\Delta(G)$, the tournament T' is also a safe completion of G. This contradicts our choice of T as T' has higher value than T. □

Lemma 34. *Let L be a median order of T having feed vertex d. If d does not have a large second neighborhood in H, then:*

(*a*) $z \in N_H^+(d)$, *and*

(*b*) $\nexists u \in N_H^+(z)$ *such that* $u \in N_T^-(d) \setminus N_T^{++}(d)$.

Proof. (*a*) If $z \notin N_H^+(d)$, then by Theorem 34 and the fact that $N_G^{++}(d) \subseteq N_H^{++}(d)$, we have $|N_H^+(d)| = |N_G^+(d)| \leq |N_G^{++}(d)| \leq |N_H^{++}(d)|$. This contradicts the assumption that d does not have a large second neighborhood in H.

(*b*) Suppose for the sake of contradiction that such a u exists. From (*a*), $z \in N_H^+(d)$. As $z \in N_H^+(d) \cap N_H^-(u)$ and $u \in N_T^-(d)$, we get $u \in N_H^{++}(d)$. Note that as $u \in N_T^-(d) \setminus N_T^{++}(d)$, we have $u \notin N_G^{++}(d)$. Combining all these together we get,

$$\begin{aligned}
|N_H^+(d)| &= |N_G^+(d)| + 1 \qquad (\text{as } z \in N_H^+(d)) \\
&\leq |N_G^{++}(d)| + 1 \qquad (\text{by Theorem 34}) \\
&= |N_G^{++}(d) \cup \{u\}| \qquad (\text{as } u \notin N_G^{++}(d)) \\
&\leq |N_H^{++}(d)| \qquad (\text{as } N_G^{++}(d) \subseteq N_H^{++}(d) \text{ and } u \in N_H^{++}(d))
\end{aligned}$$

and therefore d has a large second neighborhood in H, which is a contradiction. □

Define $\mathcal{I}(G) = \{I(u) \colon u \in V(G)\} = \{\Gamma(Q) \colon Q \in \mathcal{C}\} \cup \{\{u\} \colon u \text{ is prime}\}$. By Corollary 14(*b*), $\mathcal{I}(G)$ is a partition of $V(G)$ into modules of G. As noted before, it can be easily seen that since the missing edges of G form a matching, every module in G is also a module in T. This implies that $\mathcal{I}(G)$ is a partition of $V(T)$ into modules of T as well. Therefore, by Lemma 21, there exists a good median order of T with respect to $\mathcal{I}(G)$.

Lemma 35. *If there exists a good median order L of T with respect to $\mathcal{I}(G)$ which is periodic, then there exists $x \in V(G)$ such that x has a large second neighborhood in both G and H.*

Proof. For the purposes of this proof, given an ordering of vertices $\hat{L} = (x_1, x_2, \ldots, x_n)$ and a vertex $v \in \{x_1, x_2, \ldots, x_n\}$, we define the "index of v in $\hat{L}$" to be the integer i such that $x_i = v$.

Let us denote the feed vertex of $Sed^i_{\mathcal{I}(G)}(L)$ by d_i. In particular, $d_0 = d$. By Theorem 36, we know that for any integer $i \geq 0$, $Sed^i_{\mathcal{I}(G)}(L)$ is a good median order of T with respect to $\mathcal{I}(G)$. Note that we then have by Theorem 34 that for every integer $i \geq 0$, d_i has a large second neighborhood in G.

As H does not have any sink, there exists $u \in V(G)$ such that $u \in N^+_H(z)$. If there exists an integer $i \geq 0$ such that $d_i = u$, then as $z \notin N^+_H(d_i)$, by Lemma 34(a), $d_i = u$ has a large second neighborhood in H too, and we are done. This means that there exists an integer $q \geq 0$ such that the index of u in $Sed^{q+1}_{\mathcal{I}(G)}(L)$ is less than its index in $Sed^q_{\mathcal{I}(G)}(L)$ (recall that L is periodic). Then u must be in $N^-_T(d_q) \setminus N^{++}_T(d_q)$, which implies by Lemma 34(b) that d_q has a large second neighborhood in H. $\square$

By the above lemma, henceforth we can focus our attention on the case when every good median order of T with respect to $\mathcal{I}(G)$ is stable. That is, for any good median order L of T with respect to $\mathcal{I}(G)$, there exists a median order $Sed^q_{\mathcal{I}(G)}(L)$ (where $q \geq 0$) whose feed vertex d satisfies $|N^{++}_T(d) \setminus I(d)| > |N^+_T(d) \setminus I(d)|$. Therefore, to complete the proof of Theorem 37, we only need to show that if d is the feed vertex of a median order of T satisfying the above property, then d has a large second neighborhood in H. The remainder of the section is devoted to proving this fact, which we state as Lemma 42.

Lemma 36. *Let L be a median order of T having feed vertex d such that $|N^{++}_T(d) \setminus I(d)| > |N^+_T(d) \setminus I(d)|$. If either d has no special in-neighbors or d has a special in-neighbor of Type-I, then d has a large second neighborhood in H.*

Proof. If $z \notin N^+_H(d)$, then we are done by Lemma 34(a). So we can assume that $z \in N^+_H(d)$.

Suppose that d has no special in-neighbors. Then, by Lemma 27, we have $N^{++}_T(d) \subseteq N^{++}_G(d)$. Consequently, $N^{++}_T(d) \setminus I(d) \subseteq N^{++}_G(d) \setminus I(d)$.

Now suppose that d has a special in-neighbor of Type-I and d is not prime. Then there exists $Q \in \mathcal{C}$ such that $I(d) = \Gamma(Q)$. By Lemma 31, we have $N^{++}_T(d) \setminus I(d) \subseteq N^{++}_G(d) \setminus I(d)$.

Therefore, if d has no special in-neighbors or if d has a special in-neighbor of Type-I but d is not prime, we have $N^{++}_T(d) \setminus I(d) \subseteq N^{++}_G(d) \setminus I(d)$. In that case, we get,

$$\begin{aligned}
|N^+_H(d)| &= |N^+_G(d)| + 1 \quad (\text{as } z \in N^+_H(d)) \\
&= |N^+_G(d) \setminus I(d)| + |N^+_G(d) \cap I(d)| + 1 \\
&\leq |N^+_T(d) \setminus I(d)| + |N^{++}_G(d) \cap I(d)| + 1 \quad (\text{since } N^+_G(d) \subseteq N^+_T(d) \text{ and by Remark 8}) \\
&\leq |N^{++}_T(d) \setminus I(d)| + |N^{++}_G(d) \cap I(d)| \quad (\text{as } |N^{++}_T(d) \setminus I(d)| > |N^+_T(d) \setminus I(d)|) \\
&\leq |N^{++}_G(d) \setminus I(d)| + |N^{++}_G(d) \cap I(d)| \quad (\text{as } N^{++}_T(d) \setminus I(d) \subseteq N^{++}_G(d) \setminus I(d)) \\
&= |N^{++}_G(d)| \\
&\leq |N^{++}_H(d)|
\end{aligned}$$

and hence d has a large second neighborhood in H. Now, to prove the lemma, it only remains to consider the case when d has a special in-neighbor b of Type-I and d is prime. Then by Lemma 29, there exist $a, u \in V(T)$ such that $d \to a \dashrightarrow b \twoheadrightarrow d$ and $a \twoheadrightarrow u \to b$, where $u \text{---} d$. This means that $(u, d)R(b, a)$. If $u \dashrightarrow d$, then since T is a safe completion of G, the fact that $a \dashrightarrow b$ implies that $\{u, d\}$ and $\{b, a\}$ lie in some cycle Q in $\mathcal{C}$. But then $d \in \Gamma(Q)$, which contradicts the fact that d is prime. Therefore, we have $d \dashrightarrow u$. Then by Lemma 30(a), we have $N^{++}_T(d) \setminus \{b\} \subseteq N^{++}_G(d)$. Combining all these together, we have

$$\begin{aligned}
|N^+_H(d)| &= |N^+_G(d)| + 1 \quad (\text{as } z \in N^+_H(d)) \\
&= |N^+_T(d)| - 1 + 1 \quad (\text{as } d \dashrightarrow u \text{ in } T) \\
&\leq |N^{++}_T(d) \setminus \{b\}| \quad (\text{as } I(d) = \{d\}, \text{ we have } |N^{++}_T(d)| > |N^+_T(d)|) \\
&\leq |N^{++}_G(d)| \quad (\text{as } N^{++}_T(d) \setminus \{b\} \subseteq N^{++}_G(d)) \\
&\leq |N^{++}_H(d)|
\end{aligned}$$

Hence the lemma. □

The relation F Define a relation F on $V(G)$ as follows. For $x, y \in V(G)$ such that x is prime, we say that xFy if and only if there exists $x' \in V(G)$ such that $x \twoheadrightarrow y \to x'$, $x \text{---} x'$ in G and the missing edge $x \text{---} x'$ is singly-forced. Note that if xFy, then the missing edge $x \text{---} x'$ is forced in the direction x to x', and the condition that $x \text{---} x'$ is singly-forced ensures that it is not forced in the direction x' to x.

Lemma 37. *Let x be the feed vertex of a median order L of T. Suppose that x is prime and there* *exists $y \in V(G)$ such that xFy. Then, there exists $y' \in V(G)$ such that $y \dashline y'$ and $y \dashrightarrow y' \to x$ in T. Moreover, y is prime.*

Proof. By the definition of xFy, we have that there exists $x' \in V(G)$ such that $x \twoheadrightarrow y \to x'$, where $x \dashline x'$ is singly-forced and is forced in the direction from x to x'. As T is a safe completion, we then have $x \dashrightarrow x'$. Since x is the feed vertex of L, $x \twoheadrightarrow y$ is a reverse special arc in (T, L), where the only missing edge $x \dashline x'$ that is incident on x is oriented as $x \dashrightarrow x'$. This implies that condition (b) of Lemma 32 does not hold. Therefore, condition (a) of the lemma must be true, i.e. there should exist $y' \in V(T)$ such that $y \dashrightarrow y' \to x$ in T. Now, if y is not prime, then we have that, $y \in \Gamma(Q)$ for some $Q \in \mathcal{C}$. Therefore by Corollary 14(a), $y' \in \Gamma(Q)$. But we have a vertex $x \in V(G) \setminus \Gamma(Q)$ (as x is prime) such that $x \to y$ and $y' \to x$. As $y, y' \in \Gamma(Q)$, this contradicts the fact that $\Gamma(Q)$ forms a module in G (by Corollary 14(b)). Therefore we can conclude that y is prime. □

Lemma 38. *The relation F is not cyclic, i.e. there does not exist vertices $x_1, x_2, \ldots, x_k \in V(G)$ such that $x_1 F x_2 F \cdots F x_k F x_1$.*

Proof. Suppose not. Then there exist vertices $x_1, x_2, \ldots, x_k \in V(D)$ such that $x_1 F x_2 F \cdots F x_k F x_1$. Therefore by the definition of F, there is a special cycle $C = x_1 \twoheadrightarrow x_2 \twoheadrightarrow x_3 \twoheadrightarrow \cdots \twoheadrightarrow x_k \twoheadrightarrow x_1$. By Corollary 13$(a)$, we know that there exists $i \in \{3, 4, \ldots, k-1\}$ such that $x_1 \dashline x_i$. Then by the definition of F, the fact that $x_1 F x_2$ implies that the missing edge $x_1 \dashline x_i$ is forced in the direction x_1 to x_i and not forced in the direction x_i to x_1. But since the only non-neighbor of x_i is x_1, the fact that $x_i F x_{i+1}$ similarly implies that the missing edge $x_1 \dashline x_i$ is forced in the direction x_i to x_1, which is a contradiction. □

Lemma 39. *Let L be a median order of T having feed vertex d and let y be a vertex such that dFy. Then there exists a median order of T having feed vertex y.*

Proof. As dFy, by the definition of F, there exists d' such that $d \twoheadrightarrow y \to d'$ where $d \dashline d'$ is singly-forced in the direction from d to d' and hence, $d \dashrightarrow d'$ in T (recall that T is a safe completion). As d is the feed vertex of L, $d \twoheadrightarrow y$ is a reverse special arc in (T, L). Since the only missing edge $d \dashline d'$ that is incident on d is oriented as $d \dashrightarrow d'$, condition (b) of Lemma 32 does not hold. Therefore, condition (a) of Lemma 32 should be true. Now as exactly one of the conditions of the lemma is satisfied, if $L = (x_1, x_2, \ldots, x_n = d)$ and $y = x_i$, for some $i \in \{1, 2, \ldots, n-1\}$, we have that $L' = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n, x_i = y)$ is also a median order of T. □

Lemma 40. *Let L be a median order of T whose feed vertex d is prime. Suppose that there exists $d' \in V(T)$ such that $d \dashrightarrow d'$ in T. Then the missing edge $\{d, d'\}$ does not have an in-neighbor in $\Delta(G)$.*

Proof. Suppose not. Let $\{a, a'\}$ be an in-neighbor of $\{d, d'\}$ in $\Delta(G)$. Then, without loss of generality, by the definition of $\Delta(G)$, we can assume that $(a, a')R(d, d')$, and therefore there exists the four cycle $a \rightarrow d \twoheadrightarrow a' \rightarrow d' \twoheadrightarrow a$. As d is prime, d does not belong to $\Gamma(Q)$ for any $Q \in \mathcal{C}$. This means that $\{d, d'\}$ does not lie in any cycle in $\mathcal{C}$. Then as T is a safe completion, $d \dashrightarrow d'$ and $(a, a')R(d, d')$, we have that $a \dashrightarrow a'$ in T. As d is the feed vertex of L, $d \twoheadrightarrow a'$ is a reverse special arc in (T, L). Note that the only missing edge $d{-}{-}{-}d'$ incident on d is oriented as $d \dashrightarrow d'$, and the only missing edge $a{-}{-}{-}a'$ incident on a' is oriented as $a \dashrightarrow a'$. This implies that neither of the conditions (a) or (b) of Lemma 32 hold, which is a contradiction. □

Lemma 41. *Let d be the feed vertex of a median order L of T and $d' \in V(T)$ be such that $d \dashrightarrow d'$ in T. If d has no Type-I special in-neighbor then d is not prime.*

Proof. Suppose for the sake of contradiction that d has no Type-I special in-neighbor and d is prime. Then, by Lemma 40, the missing edge $\{d, d'\}$ does not have an in-neighbor in $\Delta(G)$. Now, suppose that $\{d, d'\}$ has an out-neighbor, say $\{a, a'\}$ in $\Delta(G)$. Without loss of generality, by the definition of $\Delta(G)$, we can assume that $(d, d')R(a, a')$. That is, there exists the four cycle $d \rightarrow a \twoheadrightarrow d' \rightarrow a' \twoheadrightarrow d$. As d is prime, the missing edge $\{d, d'\}$ does not lie on any cycle in $\mathcal{C}$. Since T is a safe completion, $d \dashrightarrow d'$ and $(d, d')R(a, a')$, we have $a \dashrightarrow a'$ in T. Then, $d \rightarrow a \dashrightarrow a' \twoheadrightarrow d$, implying that a' is a Type-I special in-neighbor of d, which is a contradiction. Therefore, we can conclude that the missing edge $\{d, d'\}$ is an isolated vertex in $\Delta(G)$. Then, by Lemma 33, we have that $d{-}{-}{-}d'$ is not an unforced missing edge. Now, if $d{-}{-}{-}d'$ is forced in both directions, by Proposition 10, we have that $\{d, d'\}$ has an in-neighbor in $\Delta(G)$, which is a contradiction. Therefore, we can conclude that $d{-}{-}{-}d'$ is singly-forced in G.

As $d \dashrightarrow d'$ in T and T is a safe completion, it should be the case that $d{-}{-}{-}d'$ is singly-forced in the direction d to d' in G. Then by Definition 33, there exists $v \in V(G)$ such that, $d \twoheadrightarrow v \rightarrow d'$. This together with the assumption that d is prime implies that dFv. Let $y_1, y_2, \ldots, y_k$ be a sequence of vertices of maximum length such that $dFy_1Fy_2F \cdots Fy_k$, where $y_1 = v$. Note that $k \geq 1$, and since F is acyclic as shown in Lemma 38, such a sequence exists and each vertex in $d, y_1, y_2, \ldots, y_k$ is distinct. Now let $L_0 = L$ and let $L_1, L_2, \ldots, L_k$ be the median orders of T such that for each $i \in \{1, \ldots, k\}$, L_i is obtained by applying Lemma 39 on L_{i-1}. As y_{k-1} is the feed vertex of L_{k-1} and $y_{k-1}Fy_k$, by Lemma 37, we have that there exists $y'_k \in V(T)$ such that

$y_k \dashrightarrow y'_k \rightarrow y_{k-1}$ in T and that y_k is prime.

Since y_k is the feed vertex of L_k, and $y_k \dashrightarrow y'_k$, by Lemma 40, we have that $\{y_k, y'_k\}$ has no in-neighbor in $\Delta(G)$. Now, suppose that the $\{y_k, y'_k\}$ has an out-neighbor $\{b, b'\}$ in $\Delta(G)$. Then, without loss of generality, we can assume that $(y_k, y'_k)R(b, b')$, i.e. there exists the four cycle $y_k \rightarrow b \twoheadrightarrow y'_k \rightarrow b' \twoheadrightarrow y_k$. As y_k is prime, the missing edge $\{y_k, y'_k\}$ does not lie on any cycle in $\mathcal{C}$. Therefore, since T is a safe completion, $y_k \dashrightarrow y'_k$ and $(y_k, y'_k)R(b, b')$, we have that $b \dashrightarrow b'$. Since $y_{k-1}Fy_k$, there exists a vertex y'_{k-1} such that $y_{k-1}\text{---}y'_{k-1}$ is singly-forced in the direction from y_{k-1} to y'_{k-1}. As T is a safe completion, this means that $y_{k-1} \dashrightarrow y'_{k-1}$. As the missing edge incident on b' is oriented towards b' in T, this implies that $b' \neq y_{k-1}$. Clearly, $b \neq y_{k-1}$ (as $y_{k-1} \twoheadrightarrow y_k$, but $y_k \rightarrow b$). Recalling that $b\text{---}b'$, we now have that either $b \rightarrow y_{k-1}$ or $y_{k-1} \rightarrow b$. Now if $b \rightarrow y_{k-1}$, then $b \rightarrow y_{k-1} \twoheadrightarrow y_k \rightarrow b$ would form a directed triangle containing a special arc and if $y_{k-1} \rightarrow b$, then $y_{k-1} \rightarrow b \twoheadrightarrow y'_k \rightarrow y_{k-1}$ would form a directed triangle containing a special arc. As we have a contradiction in both cases, $\{y_k, y'_k\}$ has no out-neighbor in $\Delta(G)$. Therefore, $\{y_k, y'_k\}$ is an isolated vertex in $\Delta(G)$.

By Lemma 33 applied on the median order L_k of T, we have that $y_k\text{---}y'_k$ is not an unforced missing edge. As $\{y_k, y'_k\}$ has no in-neighbor in $\Delta(G)$, by Proposition 10, $y_k\text{---}y'_k$ is not forced in both directions. Therefore, we can conclude that, $y_k\text{---}y'_k$ is singly forced. As $y_k \dashrightarrow y'_k$ and T is a safe completion, we know that $y_k\text{---}y'_k$ is forced in the direction y_k to y'_k, i.e. there exists a vertex u such that $y_k \twoheadrightarrow u \rightarrow y'_k$. As y_k is prime, this further implies that y_kFu. We now have $dFy_1Fy_2F\cdots Fy_{k-1}Fy_kFu$. By Lemma 38, $u \neq d$ and $u \neq y_i$ for any $i \in \{1, 2, \ldots, k\}$. This contradicts the choice of $y_1, y_2, \ldots, y_k$. $\square$

Lemma 42. *Let L be a median order of T having feed vertex d. If $|N^{++}_T(d) \setminus I(d)| > |N^+_T(d) \setminus I(d)|$, then d has a large second neighborhood in H.*

Proof. If $z \notin N^+_H(d)$, then we are done by Lemma 34(a). So we can assume that $z \in N^+_H(d)$.

If d has no special in-neighbors or has a special in-neighbor of Type-I, then we are done by Lemma 36. Therefore, we shall assume that d has no Type-I special in-neighbors but has at least one Type-II special in-neighbor. Let x be any Type-II special in-neighbor of d. Then there exists $d' \in V(T)$ such that $d \dashrightarrow d' \rightarrow x \twoheadrightarrow d$ in T. As d has no Type-I special in-neighbors, by Lemma 41, we get that d is not prime, i.e. $I(d) = \Gamma(Q)$ for some $Q \in \mathcal{C}$. Therefore, by Corollary 14(a), $d' \in \Gamma(Q)$. Now suppose that $x \notin \Gamma(Q)$. Then since $d' \rightarrow x$, $x \rightarrow d$ and $d, d' \in \Gamma(Q)$, we have a contradiction to the fact that $\Gamma(Q)$ is a module in G (by Corollary 14(b)). Therefore, every special in-neighbor of d is contained in $\Gamma(Q) = I(d)$; in other words, there are no special in-neighbors of d in $N_T^{++}(d) \setminus I(d)$. Then by Lemma 27, we have that $N_T^{++}(d) \setminus I(d) \subseteq N_G^{++}(d) \setminus I(d)$. By Remark 8, we have $|N_G^+(d) \cap I(d)| = |N_G^{++}(d) \cap I(d)|$. Combining our observations, we get

$$\begin{aligned}
|N_H^+(d)| &= |N_G^+(d)| + 1 \text{ (as } z \in N_H^+(d)) \\
&= |N_G^+(d) \setminus I(d)| + |N_G^+(d) \cap I(d)| + 1 \\
&\leq |N_T^+(d) \setminus I(d)| + |N_G^+(d) \cap I(d)| + 1 \quad \text{(since } N_G^+(d) \subseteq N_T^+(d)) \\
&\leq |N_T^{++}(d) \setminus I(d)| + |N_G^{++}(d) \cap I(d)| \quad \text{(as } |N_T^{++}(d) \setminus I(d)| > |N_T^+(d) \setminus I(d)|) \\
&\leq |N_G^{++}(d) \setminus I(d)| + |N_G^{++}(d) \cap I(d)| \quad \text{(as } N_T^{++}(d) \setminus I(d) \subseteq N_G^{++}(d) \setminus I(d)) \\
&= |N_G^{++}(d)| \\
&\leq |N_H^{++}(d)|
\end{aligned}$$

Therefore, d has a large second neighborhood in H. □

We are now ready to give a formal proof of Theorem 37.

Proof of Theorem 37

Proof. Let T be the completion of G chosen as explained before: i.e. T is a safe completion of G which has maximum value, where the value of a tournament is the number of forward arcs in any median order of it. By Lemma 21, there exists a good median order L of T with respect to $\mathcal{I}(G)$. If L is periodic, then we are done by Lemma 35. Therefore, we can assume that L is stable. Then, by the definition of a stable median order, there exists an integer $q \geq 0$ such that $Sed^{q+1}_{\mathcal{I}(G)}(L) = Sed^{q}_{\mathcal{I}(G)}(L)$. By Theorem 36, $L' = Sed^{q}_{\mathcal{I}(G)}(L)$ is a median order of T. Let d be the feed vertex of L'. By Theorem 34, d has a large second neighborhood in G. As $Sed_{\mathcal{I}(G)}(L') = L'$, we have $|N_T^{++}(d) \setminus I(d)| > |N_T^+(d) \setminus I(d)|$. We can then conclude by Lemma 42 that d has a large second neighborhood in H as well. □

Corollary 17. *Every oriented graph whose missing edges can be partitioned into a matching and a star contains a vertex with a large second neighborhood.*

Corollary 18. *Every oriented graph whose missing edges form a matching and does not contain a sink contains at least two vertices with large second neighborhoods.*

Proof. Let H be an oriented graph whose missing edges form a matching and does not contain a sink. By Corollary 16, we know that there exists a vertex z in H with a large second neighborhood. As $H - \{z\}$ is an oriented graph whose missing edges form a matching, by Theorem 37, we can infer that there exists a vertex $z' \in V(H) \setminus \{z\}$ that has a large second neighborhood in H. $\square$

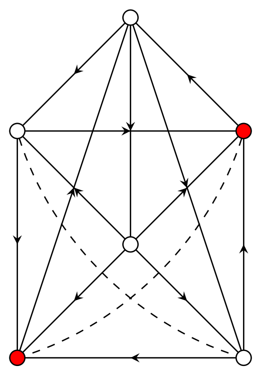


Figure 7.7: A tournament missing a matching with no sink and exactly two vertices with large second neighborhoods (those vertices are shown in dark red).

The graph shown in Figure 7.7 is a tournament missing a matching without a sink that contains exactly two vertices with large second neighborhoods. Therefore, Corollary 18 is tight.

Chapter 8

SSNC for Graphs with Constraints on Out-degree

8.1 Introduction

In the study of the Seymour Second Neighborhood Conjecture so far, we can see that other researchers have tried to attack special cases of the Second Neighborhood Conjecture without using the median order approach. For example, Lladó [82] proved the conjecture in regular oriented graphs with high connectivity. Kaneko and Locke [74] verified the conjecture for oriented graphs with minimum out-degree at most 6. We state their result below as we use it later.

Theorem 38 ([74]). *Every oriented graph with minimum out-degree less than 7 has a vertex with a large second neighborhood.*

Let G be an oriented graph with a minimum degree vertex, say v and having the property that, the out-neighborhood of v, $N^+(v)$ is an independent set. Let $w \in N^+(v)$. Then as $N^+(v)$ is independent, we have that $N^+(v) \cap N^+(w) = \emptyset$. This implies that $N^+(w) \subseteq N^{++}(v)$. Also by the choice of v, we have that $|N^+(v)| \leq |N^+(w)| \leq |N^{++}(v)|$. Therefore we can conclude that v is a vertex with large second neighborhood. This proves that the conjecture is true for bipartite graphs (in fact, it is true if the underlying undirected graph is triangle-free). It appears difficult to prove the conjecture even for oriented graphs whose underlying undirected graph is 3-colorable. In this chapter, we show that the conjecture is true for any oriented graph G such that $V(G)$ is the disjoint union of two sets A and B where $G[A]$ is 2-degenerate and $G[B]$ is an independent set.The proof relies on some counting arguments.

8.2 Graphs with constrained out-degree

An undirected graph H is said to be *2-degenerate* if every subgraph of H has a vertex of degree at most two. We say that an oriented graph is 2-degenerate if its underlying undirected graph is 2-degenerate. Note that every subgraph of a 2-degenerate graph is also 2-degenerate.

Proposition 13. *Let $H = (V, E)$ be an oriented graph on n vertices which is 2-degenerate. Then,*

(a) *If $n \geq 2$ then $|E(H)| \leq 2n - 3$.*

(b) *H has at least one vertex with out-degree at most 1.*

Proof. (a) We prove this by induction on $|V(H)| = n$. It is trivially true in the base case where $n = 2$. Assume that the statement is true for all 2-degenerate graphs with less than n vertices. As H is 2-degenerate, it has a vertex of degree at most 2, say x. Now the subgraph $H - \{x\}$ of H is itself 2-degenerate and has only $n - 1$ vertices. Therefore, by the induction hypothesis, $|E(H - \{x\})| \leq 2(n - 1) - 3$. As x has at most 2 edges incident to it, we have $|E(H)| \leq |E(H - \{x\})| + 2 \leq 2n - 3$.

(b) Note that $|E(H)| = \sum_{u \in V(H)} |N^+(u)|$. Therefore, if $|N^+(u)| \geq 2$ for every vertex $u \in V(H)$, then we would get $|E(H)| \geq 2n$, contradicting (a). $\square$

For the remainder of this section, we denote by $G = (V, E)$ an oriented graph whose vertex set has a partition (A, B) such that B is an independent set of G and $G[A]$ is 2-degenerate. Refer to Figure 8.1 for an example of such an oriented graph.

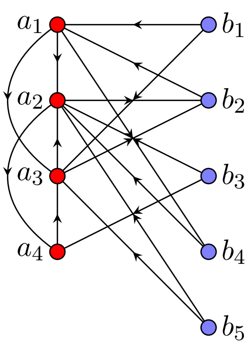


Figure 8.1: An example

Let d be the minimum out-degree of G.

Lemma 43. *If there is a vertex in B with out-degree d in D, then G has a vertex with large second neighborhood.*

Proof. Suppose not. Let $v \in B$ be a vertex such that $|N^+(v)| = d$. We can assume that $d \geq 2$, as otherwise either D contains a sink or v can be easily verified to be a vertex with large second neighborhood. As $v \in B$ and B is an independent set, we have $N^+(v) \subseteq A$. Let $N^{++}(v) = X \cup Y$, where $X \subseteq A$, $Y \subseteq B$. Also, let $|X| = x$ and $|Y| = y$. As v does not have a large second neighborhood and $|N^+(v)| = d$, we have $x + y \leq d - 1$. Consider the subgraph $H = G[N^+(v) \cup X \cup Y]$. As $N^+(v) \cup X \subseteq A$ and $G[A]$ is 2-degenerate, by Proposition 13(a), the maximum number of edges in $G[N^+(v) \cup X]$ is $2(d + x) - 3$. Together with the at most dy edges between $N^+(v)$ and Y, we get that the number of edges in H that have at least one end-point in $N^+(v)$ is at most $2(d + x) + dy - 3$. i.e., $|\{(p, q) \in E(H) \colon \{p, q\} \cap N^+(v) \neq \emptyset\}| \leq 2(d + x) + dy - 3$. Also since each vertex $u \in N^+(v)$ has out-degree at least d, we have that $|\{(p, q) \in E(H) \colon p \in N^+(v)\}| \geq d^2$. Therefore we can conclude that,

$$2d + 2x + dy - 3 \geq d^2 \tag{8.1}$$

Suppose that $y \leq d - 2$. Then we have,

$$\begin{aligned} 2d + 2x + dy - 3 &= 2d + 2(x + y) + (d - 2)y - 3 && \text{(adding and subtracting } 2y) \\ &\leq 2d + 2(d - 1) + (d - 2)^2 - 3 && (\text{since } x + y \leq d - 1 \text{ and } y \leq d - 2) \\ &= d^2 - 1 < d^2 \end{aligned}$$

which is a contradiction to (8.1). Therefore, $y \geq d - 1$. Since $x + y \leq d - 1$, this implies that $x = 0$ and $y = d - 1$. As $N^+(v) \subseteq A$, we know that $G[N^+(v)]$ is 2-degenerate. Then by Proposition 13(b), there exists a vertex $w \in N^+(v)$ whose out-degree in $G[N^+(v)]$ is at most 1. In fact, the out-degree of w in $G[N^+(v)]$ is exactly 1, as otherwise $N^+(w) \subseteq Y$, implying that $y \geq |N^+(w)| \geq d$, which contradicts the fact that $y = d - 1$. Let w' be the unique out-neighbor of w in $G[N^+(v)]$. Note that since $w \in N^+(v)$, we have $N^+(w) \subseteq N^+(v) \cup N^{++}(v)$, or in other words, $N^+(w) \subseteq N^+(v) \cup X \cup Y$. Then the fact that $N^+(w) \cap N^+(v) = \{w'\}$ and $x = 0$ implies that $N^+(w) \subseteq \{w'\} \cup Y$. Since $y = d - 1$, this further implies that $|N^+(w)| = d$; in particular, $N^+(w) = Y \cup \{w'\}$. Again, as with w, it can be seen that $N^+(w') \subseteq N^+(v) \cup X \cup Y$. As $x = 0$ and w' has at most $d - 1$ out-neighbors in $N^+(v)$, it is clear that w' should have at least one out-neighbor in Y, say z. Then $z \in N^+(w) \cap N^+(w')$. As Y is an independent set, we have $N^+(z) \subseteq A \setminus \{w, w'\}$, implying that $N^+(z)$ is disjoint from $N^+(w) = Y \cup \{w'\}$. This means that $N^+(z) \subseteq N^{++}(w)$, which gives $|N^{++}(w)| \geq |N^+(z)| \geq d = |N^+(w)|$. Hence w has a large second neighborhood in G, which is a contradiction. □

Lemma 44. *If the out-degree of every vertex in B is at least $d+1$, then G has a vertex with large second neighborhood.*

Proof. Suppose not. Note that from Theorem 38, we have $d > 6$. Clearly, there is a vertex $v \in A$ such that $|N^+(v)| = d$. Let $N^+(v) = X \cup Y$, where $X \subseteq A$ and $Y \subseteq B$, and $N^{++}(v) = X' \cup Y'$, where $X' \subseteq A$ and $Y' \subseteq B$. Also, let $|X| = x, |Y| = y, |X'| = x'$ and $|Y'| = y'$. Note that $x + y = d$, and since v does not have a large second neighborhood, $x' + y' \leq d - 1$. Since each vertex of Y has at least $d+1$ out-neighbors, all of which lie in $X \cup X'$, we further have $x + x' \geq d + 1$.

Claim 1. $x \geq 3$.

Assume to the contrary that $x \leq 2$. Then since $x' \leq d-1$ and $x + x' \geq d+1$, it should be the case that $x = 2$, $x' = d-1$, $y' = 0$ and $x + x' = d+1$. This implies that $N^+(u) = X \cup X'$ for all $u \in Y$. Then, neither vertex in X can have an out-neighbor in Y. Now if $w \in X$ is a vertex that has no out-neighbor in X (clearly, such a vertex exists as $x = 2$), the fact that $y' = 0$ implies that $N^+(w) \subseteq X'$. But $x' = d-1$, implying that $|N^+(w)| < d$, which is a contradiction. This proves the claim.

Now, consider the subgraph $H = G[X \cup Y \cup X' \cup Y']$. As $X \cup X' \subseteq A$, $x + x' \geq 2$ and $G[A]$ is 2-degenerate, by Proposition 13(a), the maximum number of edges in $G[X \cup X']$ is $2(x + x') - 3$. Together with the at most xy edges between X and Y, the at most xy' edges between X and Y' and the at most yx' edges between Y and X', we get that the number of edges in H with at least one end-point in $N^+(v) = X \cup Y$ is at most $2(x + x') - 3 + xy + xy' + x'y$, i.e., $|\{(p,q) \in E(H)\colon \{p,q\} \cap N^+(v) \neq \emptyset\}| \leq 2(x + x') - 3 + xy + xy' + x'y$. There are at least d edges going out from each vertex of X and at least $d+1$ edges going out from each vertex of Y. Therefore, $|\{(p,q) \in E(H)\colon p \in X\}| \geq dx$ and $|\{(p,q) \in E(H)\colon p \in Y\}| \geq (d+1)y$. Altogether, we have $|\{(p,q) \in E(H)\colon p \in N^+(v)\}| \geq dx + (d+1)y = d^2 + y$ (as $x + y = d$). Hence we can conclude that,

$$2(x + x') - 3 + xy + xy' + x'y \geq d^2 + y \tag{8.2}$$

Claim 2. At most one of x and x' can be greater than or equal to $\frac{d}{2} + 1$.

Suppose for the sake of contradiction that $x = \frac{d}{2} + r$ and $x' = \frac{d}{2} + s$, where $r, s \geq 1$. As $x + y = d$ and $x' + y' \leq d - 1$ we have $y = \frac{d}{2} - r$ and $y' \leq \frac{d}{2} - s - 1$. By substituting these in the LHS of the equation (8.2) we have,

$$\begin{aligned} 2(x + x') - 3 + xy + xy' + x'y &\leq 2(d + r + s) - 3 + \left(\frac{d}{2} + r\right)\left(\frac{d}{2} - r\right) \\ &\quad + \left(\frac{d}{2} + r\right)\left(\frac{d}{2} - s - 1\right) + \left(\frac{d}{2} + s\right)\left(\frac{d}{2} - r\right) \\ &\leq \frac{3d^2}{4} + 2d + r - 3 - r^2 - \frac{d}{2} \quad \text{(as } r \geq 1 \text{ we have } rs \geq s\text{)} \end{aligned}$$

Combining the last inequality with (8.2), we get

$$\begin{aligned} \frac{3d^2}{4} + 2d + r - 3 - r^2 - \frac{d}{2} &\geq d^2 + y \\ 6d + 4r &\geq d^2 + 4y + 12 + 4r^2 \\ d^2 + 4r &> d^2 + 4y + 12 + 4r^2 \quad \text{(as } d > 6\text{)} \\ 4r &> 4y + 12 + 4r^2 \end{aligned}$$

This is a contradiction as $r \geq 1$ and $y \geq 0$. This proves the claim.

Now, consider the LHS of (8.2).

$$\begin{aligned} 2(x + x') - 3 + xy + xy' + x'y &= 2x + 2x' - 3 + xy + y'(x + y) + x'(x + y) - xx' - yy' \\ &\qquad \text{(adding and subtracting } xx' + yy'\text{)} \\ &= 2x + 2x' - 3 + xy + d(x' + y') - xx' - yy' \\ &\qquad \text{(as } x + y = d\text{)} \\ &\leq 2x + 2x' - 3 + xy + d(d - 1) - xx' - yy' \qquad (8.3) \\ &\qquad \text{(as } x' + y' \leq d - 1\text{)} \end{aligned}$$

Now, suppose that $x' \geq y + 2$. Then (8.4) implies,

$$\begin{aligned} 2(x + x') - 3 + xy + xy' + x'y &\leq 2x + 2x' - 3 + xy + d(d - 1) - x(y + 2) - yy' \\ &= d^2 + 2x' - 3 - d - yy' \end{aligned}$$

Combining this inequality with (8.2), we have

$$\begin{aligned} d^2 + 2x' - 3 - d - yy' &\geq d^2 + y \\ 2x' - 3 - d - yy' &\geq y \qquad (8.4) \end{aligned}$$

Therefore we get,

$$\begin{aligned} 2x' - 3 - d - yy' + 2x &\geq y + 2x \\ 2(x + x') - 3 - d - yy' &\geq 2d - y \quad (\text{as } x + y = d) \end{aligned}$$

As $\max\{x, x'\} = d$ and by Claim 2, $\min\{x, x'\} \leq \frac{d}{2} + 1$, we have $x + x' \leq \frac{3d}{2} + 1$. Combining this with the above inequality, we have

$$\begin{aligned} 3d - 1 - d - yy' &\geq 2d - y \\ y &\geq yy' + 1 \end{aligned}$$

This implies that $y' = 0$. Then (8.4) becomes

$$\begin{aligned} 2x' - 3 - d &\geq y \\ 2x' &\geq x + 2y + 3 \quad (\text{as } d = x + y) \\ x' &\geq y + 3 \quad (\text{as } x \geq 3 \text{ by Claim 1}) \end{aligned}$$

Substituting this together with $y' = 0$ in the RHS of (8.4) we get,

$$\begin{aligned} 2(x + x') - 3 + xy + xy' + x'y &\leq 2x + 2x' - 3 + xy + d(d-1) - x(y+3) \\ &= d^2 + 2x' - 3 - d - x \end{aligned}$$

Combining this with (8.2), we have

$$\begin{aligned} d^2 + 2x' - 3 - d - x &\geq d^2 + y \\ x' &\geq d + \frac{3}{2} \quad (\text{as } x + y = d) \end{aligned}$$

which contradicts the fact that $x' + y' \leq d - 1$. Therefore, we can assume that $x' \leq y + 1$.

In fact, $x' = y + 1$, as otherwise, $x + x' < x + y + 1 = d + 1$, which is a contradiction to our earlier observation that $x + x' \geq d + 1$. Now, substituting $x' = y + 1$ in the RHS of (8.4), we get

$$\begin{aligned} 2(x + x') - 3 + xy + xy' + x'y &\leq 2x + 2(y + 1) - 3 + xy + d^2 - d - x(y + 1) - yy' \\ &= d^2 + d - 1 - x - yy' \quad (\text{as } x + y = d) \end{aligned}$$

Now, combining this with (8.2), we have

$$\begin{aligned} d^2 + d - 1 - x - yy' &\geq d^2 + y \\ yy' + 1 &\leq 0 \quad (\text{as } x + y = d) \end{aligned}$$

which is a contradiction. This proves the lemma. □

Theorem 39. *Let $G = (V, E)$ be an oriented graph whose vertex set $V(G)$ has a partition (A, B), such that B is an independent set and $G[A]$ is 2-degenerate. Then G has a vertex with a large second neighborhood.*

Proof. The proof is immediate from Lemma 43 and Lemma 44. □

8.3 Some related conjectures

The question of whether there exists two vertices with large second neighborhoods in any oriented graph without a sink seems to be open.

Conjecture 2. *Any oriented graph without a sink contains at least two vertices with large second neighborhoods.*

Clearly, Conjecture 2 implies Conjecture 1 (the Seymour Second Neighborhood Conjecture). We propose the following conjecture, which though apparently weaker at first sight, can be shown to be equivalent to Conjecture 2.

Conjecture 3. *If an oriented graph contains exactly one vertex with a large second neighborhood, then that vertex is a sink.*

Proposition 14. *Conjecture 3 implies Conjecture 1.*

Proof. Suppose that Conjecture 3 is true but Conjecture 1 is not. Let G be a minimal counterexample to Conjecture 1: i.e., G is an oriented graph with minimum number of vertices and edges in which no vertex has a large second neighborhood. In particular, G cannot have a sink. Let $(u, v) \in E(G)$. Consider the graph G' obtained by removing the edge (u, v) from G. As G is a minimal counterexample to Conjecture 1, we know that G' contains at least one vertex with a large second neighborhood. We claim that G' contains at least two vertices with large second neighborhoods. Suppose not. Then by Conjecture 3, G' has a sink in it. As there is no sink in G and every vertex other than u has the same out-neighborhood in both G and G', this means that u must be a sink in G'. Then, $N^+_G(u) = \{v\}$ and $N^+_G(v) \subseteq N^{++}_G(u)$. Since v is not a sink in G, we also have that $|N^+_G(v)| \geq 1$, which gives us $|N^{++}_G(u)| \geq 1$. Therefore, u has a large second neighborhood in G, which is a contradiction. This proves that G' contains at least two vertices with large second neighborhoods. Then there exists a vertex $w \neq u$ in G' such that $|N^+_{G'}(w)| \leq |N^{++}_{G'}(w)|$. As $w \neq u$, by the definition of G', we have that $N^+_G(w) = N^+_{G'}(w)$ and $N^{++}_{G'}(w) \subseteq N^{++}_G(w)$. Combining this with the previous observation, we get $|N^+_G(w)| \leq |N^{++}_G(w)|$, implying that w has a large second neighborhood in G, which is a contradiction. $\square$

By the above proposition, it can be easily seen that Conjecture 3 implies Conjecture 2 and therefore they are equivalent. We do not know if these conjectures are equivalent to Conjecture 1 or if they hold for the class of graphs studied in this chapter.

Chapter 9

Conclusion

9.1 A brief summary of the work

In this work, we explored a few graph theoretical problems using the approach of some special vertex orderings. The major problems that we encountered in this thesis and the vertex ordering techniques that we adapted for solving those problems are briefly summarized as follows:

(*a*) The threshold cover problem – our results majorly relied on the lexicographic method introduced by Hell and Huang [68]. We also make use of a property of Lex-BFS orderings.

(*b*) On the kernel and related problems in interval digraphs – this work includes a vertex ordering characterization (using some forbidden patterns) for reflexive interval digraphs, that can be used to solve some of the problems that we deal in this part. Some other problems that we studied here for interval bigraphs and some subclasses of interval digraphs also make use of the vertex orderings based on the end-points of the intervals representing vertices.

(*c*) Extending the Seymour Second Neighborhood Conjecture for some special graph classes – here we make use of the notions such as median order of tournaments and the sedimentation of median order.

9.2 Some open problems

Here we will list some of the interesting open problems related to the work done in this thesis.

1. As we have seen in Chapter 2, Chvátal and Hammer first asked the question of whether there is any graph G such that $\mathrm{th}(G) > \chi(G^*)$, where $\mathrm{th}(G)$ denotes the size of a minimum threshold cover of G. Cozzens and Leibowitz [27] showed that for every $k \geq 4$, there exists

a graph G such that $\chi(G^*) = k$ but $\mathrm{th}(G) > k$. Raschle and Simon [102] settled the conjecture for the case $k = 2$, by proving that for any graph G such that $\chi(G^*) = 2$, we have $\mathrm{th}(G) = \chi(G^*)$. But *the status of the above question for the case* $k = 3$ *is still open.*

2. In Chapter 3, we defined the 2-chain partition problem – that asks for any input bipartite graph G, whether there exist two chain graphs H_1 and H_2 such that $E(G) = E(H_1) \cup E(H_2)$ and $E(H_1) \cap E(H_2) = \emptyset$. We have seen that the 2-chain partition problem is polynomial-time solvable as it can be reduced to the recognition problem of interval bigraphs which can be solved in polynomial time [95]. It would be interesting to *explore the complexity of the* k*-chain partition problem for* $k > 2$.

3. Using a close relationship between interval digraphs and interval bigraphs, Müller [95] gave a polynomial time algorithm that can be used to recognize both the classes interval digraphs and interval bigraphs (refer to Chapter 3 for details). As this algorithm takes $O(nm^6(n+m)\log n)$ time, it is a long standing open problem in the literature to *find a simpler and efficient algorithm for recognizing the classes of interval bigraphs and interval digraphs. The question of a forbidden structure characterization for these classes are also open.*

4. The recognition algorithm for interval digraphs given by Müller [95] can be easily adapted to recognize reflexive interval digraphs (as in addition, it is only required to check whether each vertex has a loop). We have witnessed in Chapter 4 how the reflexivity of interval digraphs has a significant role in allowing efficient algorithms for all the computational problems that we studied for interval digraphs. Therefore it is interesting to see *whether there exist a simpler and more efficient algorithm for the recognition problem for the class of reflexive interval digraphs.*

5. Many of the subclasses of interval digraphs, like adjusted interval digraphs [114], chronological interval digraphs [31], interval catch digraphs [99], and interval point digraphs [100] have simpler and much more efficient recognition algorithms. But the complexity status of the *recognition problem for the class of interval nest digraphs is still open.*

6. In Chapter 4, we in fact solved the problems Independent Set, Kernel, Min-Kernel and Max-Kernel for a super class of reflexive interval digraphs called DUF-digraphs. But we could solve the Absorbing-Set problem only for reflexive interval digraphs. Therefore *the complexity status of the* Absorbing-Set *problem for the class of DUF-digraphs remains*

open.

7. In Chapter 6, we solved the Weak Independent-Set (resp. Feedback Vertex-Set) problem for some subclasses of interval digraphs. *But the complexities of the* Weak Independent-Set *(resp.* Feedback Vertex-Set*) problem for the class of reflexive interval digraphs and more generally for the class of interval digraphs remain open.*

8. In Chapter 7 and Chapter 8 we studied the Seymour Second Neighborhood Conjecture (SSNC) for some special graph classes. Therefore, an obvious problem that we can think here is *extending SSNC for other classes of oriented graphs.* As it is easy to verify SSNC to be true for the oriented graphs whose underlying graph is bipartite, it would be interesting to think about *SSNC for the class of oriented graphs whose underlying graph is 3-colorable.*

9. As a strengthening of the result of Havet and Thomasse [65] that "any tournament without a sink has at least two vertices with large second neighborhood", in Chapter 7 we could prove that "any tournament whose missing edges form a matching and has no sink has at least two vertices with large second neighborhood". *The question of whether there exists two vertices with large second neighborhoods in any oriented graph without a sink seems to be open.*

   In regard to this, we propose the following two conjectures.

   - **Conjecture 1:** *Any oriented graph without a sink contains at least two vertices with large second neighborhoods.*
   - **Conjecture 2:** *If an oriented graph contains exactly one vertex with a large second neighborhood, then that vertex is a sink.*

   Note that clearly, Conjecture 1 implies SSNC and we have proved in Chapter 8 that Conjecture 2 also implies SSNC. It is also shown that Conjecture 1 and Conjecture 2 are equivalent.

Bibliography

[1] Moncef Abbas and Youcef Saoula. Polynomial algorithms for kernels in comparability, permutation and P_4-free graphs. *4OR*, 3(3):217–225, 2005.

[2] Jochen Alber, Hans L. Bodlaender, Henning Fernau, Ton Kloks, and Rolf Niedermeier. Fixed parameter algorithms for dominating set and related problems on planar graphs. *Algorithmica*, 33(4):461–493, 2002.

[3] Noga Alon, Graham Brightwell, Hal A. Kierstead, Alexandr V. Kostochka, and Peter Winkler. Dominating sets in k-majority tournaments. *Journal of Combinatorial Theory, Series B*, 96(3):374–387, 2006.

[4] Srinivasa Rao Arikati and C. Pandu Rangan. Linear algorithm for optimal path cover problem on interval graphs. *Information Processing Letters*, 35(3):149–153, 1990.

[5] Jean-Pierre Barthelemy, Alain Guenoche, and Olivier Hudry. Median linear orders: heuristics and a branch and bound algorithm. *European Journal of Operational Research*, 42(3):313–325, 1989.

[6] Seymour Benzer. On the topography of the genetic fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, 47(3):403–415, 1961.

[7] Claude Berge. *Graphs and Hypergraphs*. North-Holland Publishing, 1973.

[8] Claude Berge and Pierre Duchet. Recent problems and results about kernels in directed graphs. *Discrete Mathematics*, 86(1-3):27–31, 1990.

[9] Mostafa Blidia, Pierre Duchet, and Frédéric Maffray. On kernels in perfect graphs. *Combinatorica*, 13(2):231–233, 1993.

[10] Kenneth P. Bogart and Ann N. Trenk. Bounded bitolerance digraphs. *Discrete Mathematics*, 215(1-3):13–20, 2000.

[11] Kellogg S. Booth and George S. Lueker. Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms. *Journal of Computer and System Sciences*, 13(3):335–379, 1976.

[12] Endre Boros and Vladimir Gurvich. Perfect graphs are kernel solvable. *Discrete Mathematics*, 159(1-3):35–55, 1996.

[13] Endre Boros and Vladimir Gurvich. Perfect graphs, kernels, and cores of cooperative games. *Discrete Mathematics*, 306(19-20):2336–2354, 2006.

[14] Andreas Brandstädt, Feodor F. Dragan, and Falk Nicolai. LexBFS-orderings and powers of chordal graphs. *Discrete Mathematics*, 171(1-3):27–42, 1997.

[15] Andreas Brandstädt, Van Bang Le, and Jeremy P. Spinrad. *Graph classes: a survey*. SIAM, 1999.

[16] Louis Caccetta and Roland Häggkvist. *On minimal digraphs with given girth*. Department of Combinatorics and Optimization, University of Waterloo, 1978.

[17] Irène Charon, Alain Guénoche, Olivier Hudry, and Frédéric Woirgard. New results on the computation of median orders. *Discrete Mathematics*, 165:139–153, 1997.

[18] Jianer Chen, Yang Liu, Songjian Lu, Barry O'Sullivan, and Igor Razgon. A fixed-parameter algorithm for the directed feedback vertex set problem. In *Proceedings of the fortieth ACM Symp. on Theory of Computing*, pages 177–186, 2008.

[19] Miroslav Chlebík and Janka Chlebíková. The complexity of combinatorial optimization problems on d-dimensional boxes. *SIAM Journal on Discrete Mathematics*, 21(1):158–169, 2007.

[20] Maria Chudnovsky, Ringi Kim, Chun-Hung Liu, Paul Seymour, and Stéphan Thomassé. Domination in tournaments. *Journal of Combinatorial Theory, Series B*, 130:98–113, 2018.

[21] Václav Chvátal. On the computational complexity of finding a kernel. *Report CRM-300, Centre de Recherches Mathématiques, Université de Montréal*, 592, 1973.

[22] Václav Chvátal and Peter L. Hammer. Aggregations of inequalities in integer programming. *Annals of Discrete Mathematics*, 1:145–162, 1977.

[23] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to algorithms*. MIT Press, 2009.

[24] Derek G. Corneil. Lexicographic breadth first search – a survey. In *Proceedings of the thirtieth International Conference on Graph-Theoretic Concepts in Computer Science*, WG '04, pages 1–19, 2004.

[25] Derek G. Corneil, Stephan Olariu, and Lorna Stewart. The LBFS structure and recognition of interval graphs. *SIAM Journal on Discrete Mathematics*, 23(4):1905–1953, 2010.

[26] Margaret B. Cozzens and Mark D. Halsey. The relationship between the threshold dimension of split graphs and various dimensional parameters. *Discrete Applied Mathematics*, 30(2):125–135, 1991.

[27] Margaret B Cozzens and Rochelle Leibowitz. Threshold dimension of graphs. *SIAM Journal on Algebraic Discrete Methods*, 5(4):579–595, 1984.

[28] Pierluigi Crescenzi. A short guide to approximation preserving reductions. In *Proceedings of Twelfth Annual IEEE Conference on Computational Complexity*, pages 262–273. IEEE, 1997.

[29] Peter Damaschke. Forbidden ordered subgraphs. In Rainer Bodendiek and Rudolf Henn, editors, *Topics in Combinatorics and Graph Theory: Essays in Honour of Gerhard Ringel*, pages 219–229. Physica-Verlag HD, Heidelberg, 1990.

[30] Peter Damaschke, Jitender S. Deogun, Dieter Kratsch, and George Steiner. Finding Hamiltonian paths in cocomparability graphs using the bump number algorithm. *Order*, 8(4):383–391, 1991.

[31] Sandip Das, Mathew Francis, Pavol Hell, and Jing Huang. Recognition and characterization of chronological interval digraphs. *The Electronic Journal of Combinatorics*, 20(3):P5, 2013.

[32] Sandip Das, Malay K. Sen, A. B. Roy, and Douglas B. West. Interval digraphs: An analogue of interval graphs. *Journal of Graph Theory*, 13(2):189–202, 1989.

[33] Olivier Durand de Gevigney, Frédéric Meunier, Christian Popa, Julien Reygner, and Ayrin Romero. Solving coloring, minimum clique cover and kernel problems on arc intersection graphs of directed paths on a tree. *4OR*, 9(2):175–188, 2011.

[34] W. Fernandez De La Vega. Kernels in random graphs. *Discrete Mathematics*, 82(2):213–217, 1990.

[35] Nathaniel Dean and Brenda J. Latka. Squaring the tournament—an open problem. *Congressus Numerantium*, pages 73–80, 1995.

[36] Gabriel Andrew Dirac. On rigid circuit graphs. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 25, pages 71–76. Springer, 1961.

[37] Michael Dom, Daniel Lokshtanov, and Saket Saurabh. Incompressibility through colors and IDs. In *International Colloquium on Automata, Languages, and Programming*, pages 378–389. Springer, 2009.

[38] Pierre Duchet. A sufficient condition for a digraph to be kernel-perfect. *Journal of Graph Theory*, 11(1):81–85, 1987.

[39] Jack Edmonds. Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17:449–467, 1965.

[40] Tomás Feder, Pavol Hell, Jing Huang, and Arash Rafiey. Interval graphs, adjusted interval digraphs, and reflexive list homomorphisms. *Discrete Applied Mathematics*, 160(6):697–707, 2012.

[41] Dror Fidler and Raphael Yuster. Remarks on the second neighborhood problem. *Journal of Graph Theory*, 55(3):208–220, 2007.

[42] Peter C. Fishburn. *Interval Orders and Interval Graphs: A Study of Partially Ordered Sets*. A Wiley Interscience Publication. A Wiley Interscience Publication, 1985.

[43] David C. Fisher. Squaring a tournament: a proof of Dean's conjecture. *Journal of Graph Theory*, 23(1):43–48, 1996.

[44] Aviezri S. Fraenkel. Planar kernel and grundy with $d \leq 3$, $d_{out} \leq 2$, $d_{in} \leq 2$ are NP-complete. *Discrete Applied Mathematics*, 3(4):257–262, 1981.

[45] Delbert Fulkerson and Oliver Gross. Incidence matrices and interval graphs. *Pacific Journal of Mathematics*, 15(3):835–855, 1965.

[46] David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[47] Hortensia Galeana-Sánchez and Víctor Neumann-Lara. On kernels and semikernels of digraphs. *Discrete Mathematics*, 48(1):67–76, 1984.

[48] Georges Gardarin and Stefano Spaccapietra. Integrity of data bases: A general lockout algorithm with deadlock avoidance. In *IFIP Working Conference on Modelling in Data Base Management Systems*, pages 395–412, 1976.

[49] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness.* Mathematical Sciences Series. W. H. Freeman, 1979.

[50] Fănică Gavril. Algorithms for minimum coloring, maximum clique, minimum covering by cliques, and maximum independent set of a chordal graph. *SIAM Journal on Computing*, 1(2):180–187, 1972.

[51] Fănică Gavril. The intersection graphs of subtrees in trees are exactly the chordal graphs. *Journal of Combinatorial Theory, Series B*, 16(1):47–56, 1974.

[52] Salman Ghazal. Seymour's second neighborhood conjecture for tournaments missing a generalized star. *Journal of Graph Theory*, 71(1):89–94, 2012.

[53] Salman Ghazal. A remark on the second neighborhood problem. *Electronic Journal of Graph Theory and Applications*, 3(2):182–190, 2015.

[54] Paul C. Gilmore and Alan J. Hoffman. A characterization of comparability graphs and of interval graphs. *Canadian Journal of Mathematics*, 16:539–548, 1964.

[55] Martin Charles Golumbic. Interval graphs and related topics. *Discrete Mathematics*, 55(2):113–121, 1985.

[56] Martin Charles Golumbic. *Algorithmic graph theory and perfect graphs.* Elsevier, 2004.

[57] Martin Charles Golumbic, Tirza Hirst, and Moshe Lewenstein. Uniquely restricted matchings. *Algorithmica*, 31(2):139–154, 2001.

[58] Martin Charles Golumbic, Clyde L. Monma, and William T. Trotter Jr. Tolerance graphs. *Discrete Applied Mathematics*, 9(2):157–170, 1984.

[59] Martin Charles Golumbic and Ann N. Trenk. *Tolerance graphs*, volume 89 of *Cambridge studies in advanced mathematics.* Cambridge University Press, 2004.

[60] Louis Guttman. A basis for scaling qualitative data. *American Sociological Review*, 9(2):139–150, 1944.

[61] Michel Habib, Ross M. McConnell, Christophe Paul, and Laurent Viennot. LexBFS and partition refinement, with applications to transitive orientation, interval graph recognition and consecutive ones testing. *Theoretical Computer Science*, 234(1-2):59–84, 2000.

[62] György Hajós. Über eine art von graphen. *Internationale Mathematische Nachrichten*, 11:65, 1957.

[63] Peter L. Hammer, Toshihide Ibaraki, and Uri N. Peled. Threshold numbers and threshold completions. In *Annals of Discrete Mathematics*, volume 9, pages 103–106. Elsevier, 1980.

[64] Frank Harary, Jerald A. Kabell, and Frederick R. McMorris. Bipartite intersection graphs. *Commentationes Mathematicae Universitatis Carolinae*, 23(4):739–745, 1982.

[65] Frédéric Havet and Stéphan Thomassé. Median orders of tournaments: A tool for the second neighborhood problem and Sumner's conjecture. *Journal of Graph Theory*, 35(4):244–256, 2000.

[66] Teresa W. Haynes, Stephen T. Hedetniemi, and Peter J. Slater. *Domination in Graphs: Volume 2: Advanced Topics*. Chapman & Hall/CRC Pure and Applied Mathematics. Taylor & Francis, 1998.

[67] Ryan B. Hayward, Jeremy P. Spinrad, and R. Sritharan. Weakly chordal graph algorithms via handles. In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '00, pages 42–49, 2000.

[68] Pavol Hell and Jing Huang. Lexicographic orientation and representation algorithms for comparability graphs, proper circular arc graphs, and proper interval graphs. *Journal of Graph Theory*, 20(3):361–374, 1995.

[69] Pavol Hell and Jing Huang. Interval bigraphs and circular arc graphs. *Journal of Graph Theory*, 46(4):313–327, 2004.

[70] Toshihide Ibaraki and Uri N. Peled. Sufficient conditions for graphs to have threshold number 2. In *Annals of Discrete Mathematics*, volume 59, pages 241–268. North-Holland, 1981.

[71] Kyriaki Ioannidou, George B. Mertzios, and Stavros D. Nikolopoulos. The longest path problem has a polynomial solution on interval graphs. *Algorithmica*, 61(2):320–341, 2011.

[72] Lars Jaffke, O-joung Kwon, and Jan Arne Telle. Classes of intersection digraphs with good algorithmic properties. *arXiv preprint arXiv:2105.01413*, 2021.

[73] Arthur B. Kahn. Topological sorting of large networks. *Communications of the ACM*, 5(11):558–562, 1962.

[74] Yoshihiro Kaneko and Stephen C. Locke. The minimum degree approach for Paul Seymour's distance 2 conjecture. *Congressus Numerantium*, 148:201–206, 2001.

[75] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer, 1972.

[76] Ekkehard Köhler and Lalla Mouatadid. A linear time algorithm to compute a maximum weighted independent set on cocomparability graphs. *Information Processing Letters*, 116(6):391–395, 2016.

[77] Jan Kratochvíl, Andrzej Proskurowski, and Jan Arne Telle. Complexity of graph covering problems. *Nordic Journal of Computing*, 5:173–195, 1998.

[78] Dieter Kratsch, Ross M. McConnell, Kurt Mehlhorn, and Jeremy P. Spinrad. Certifying algorithms for recognizing interval graphs and permutation graphs. *SIAM Journal on Computing*, 36(2):326–353, 2006.

[79] Dieter Kratsch and Lorna Stewart. Domination on cocomparability graphs. *SIAM Journal on Discrete Mathematics*, 6(3):400–417, 1993.

[80] C. Lekkerkerker and J. Ch. Boland. Representation of a finite graph by a set of intervals on the real line. *Fundamenta Mathematicae*, 51(1):45–64, 1962.

[81] Ching-Hao Liu, Sheung-Hung Poon, and Jin-Yong Lin. Independent dominating set problem revisited. *Theoretical Computer Science*, 562:1–22, 2015.

[82] Anna Lladó. On the second neighborhood conjecture of Seymour for regular digraphs with almost optimal connectivity. *European Journal of Combinatorics*, 34(8):1406–1410, 2013.

[83] Tze-Heng Ma. On the threshold dimension 2 graphs. Technical Report, Institute of Information Science, Academia Sinica, Nankang, Taipei, Republic of China, 1993.

[84] Tze-Heng Ma and Jeremy P. Spinrad. On the 2-chain subgraph cover and related problems. *Journal of Algorithms*, 17(2):251–268, 1994.

[85] Hiroshi Maehara. A digraph represented by a family of boxes or spheres. *Journal of Graph Theory*, 8(3):431–439, 1984.

[86] Frédéric Maffray. On kernels in i-triangulated graphs. *Discrete Mathematics*, 61(2-3):247–251, 1986.

[87] Frédéric Maffray. Kernels in perfect line-graphs. *Journal of Combinatorial Theory, Series B*, 55(1):1–8, 1992.

[88] Nadimpalli V. R. Mahadev and Uri N. Peled. *Threshold graphs and related topics*, volume 56. Annals of Discrete Mathematics, 1995.

[89] Ross M. McConnell and Jeremy P. Spinrad. Modular decomposition and transitive orientation. *Discrete Mathematics*, 201(1-3):189–241, 1999.

[90] Terry A. McKee and Fred R. McMorris. *Topics in intersection graph theory*. SIAM, 1999.

[91] Nimrod Megiddo and Uzi Vishkin. On finding a minimum dominating set in a tournament. *Theoretical Computer Science*, 61(2-3):307–316, 1988.

[92] George B. Mertzios and Derek G. Corneil. A simple polynomial algorithm for the longest path problem on cocomparability graphs. *SIAM Journal on Discrete Mathematics*, 26(3):940–963, 2012.

[93] Sounaka Mishra. On the maximum uniquely restricted matching for bipartite graphs. *Electronic Notes in Discrete Mathematics*, 37:345–350, 2011.

[94] Oskar Morgenstern and John Von Neumann. *Theory of games and economic behavior*. Princeton University Press, 1953.

[95] Haiko Müller. Recognizing interval digraphs and interval bigraphs in polynomial time. *Discrete Applied Mathematics*, 78(1-3):189–205, 1997.

[96] Stephan Olariu. An optimal greedy heuristic to color interval graphs. *Information Processing Letters*, 37(1):21–25, 1991.

[97] Christos H. Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43(3):425–440, 1991.

[98] Adèle Pass-Lanneau, Ayumi Igarashi, and Frédéric Meunier. Perfect graphs with polynomially computable kernels. *Discrete Applied Mathematics*, 272:69–74, 2020.

[99] Erich Prisner. A characterization of interval catch digraphs. *Discrete Mathematics*, 73(3):285–289, 1989.

[100] Erich Prisner. Algorithms for interval catch digraphs. *Discrete Applied Mathematics*, 51(1-2):147–157, 1994.

[101] Ganesan Ramalingam and C. Pandu Rangan. A unified approach to domination problems on interval graphs. *Information Processing Letters*, 27(5):271–274, 1988.

[102] Thomas Raschle and Klaus Simon. Recognition of graphs with threshold dimension two. In *Proceedings of the Twenty-seventh Annual ACM Symposium on Theory of Computing*, STOC '95, pages 650–661, 1995.

[103] K. Brooks Reid, Alice A. McRae, Sandra Mitchell Hedetniemi, and Stephen T. Hedetniemi. Domination and irredundance in tournaments. *Australasian Journal of Combinatorics*, 29:157–172, 2004.

[104] Moses Richardson. On weakly ordered systems. *Bulletin of the American Mathematical Society*, 52(2):113–116, 1946.

[105] Moses Richardson. Extension theorems for solutions of irreflexive relations. *Proceedings of the National Academy of Sciences of the United States of America*, 39(7):649–655, 1953.

[106] Moses Richardson. Solutions of irreflexive relations. *Annals of Mathematics*, 58(3):573–590, 1953.

[107] J. Riguet. Sur les ensembles reguliers de rélations binaires, les relations de Ferrers. *Comptes rendus de l'Académie des Sciences*, 231:936–937, 1950.

[108] Fred S. Roberts. Indifference graphs. In *Proof techniques in graph theory: Proceedings of the Second Ann Arbor Graph Theory Conference*, pages 139–146. Academic Press, New York, 1969.

[109] Donald J. Rose, Robert E. Tarjan, and George S. Lueker. Algorithmic aspects of vertex elimination on graphs. *SIAM Journal on Computing*, 5(2):266–283, 1976.

[110] Barun K. Sanyal and Malay K. Sen. New characterizations of digraphs represented by intervals. *Journal of Graph Theory*, 22(4):297–303, 1996.

[111] Jeremy P. Spinrad. *Efficient Graph Representations*, volume 19 of *Fields Institute Monographs*. American Mathematical Soc., 2003.

[112] Andrea Sterbini and Thomas Raschle. An $O(n^3)$ time algorithm for recognizing threshold dimension 2 graphs. *Information Processing Letters*, 67(5):255–259, 1998.

[113] Edward Szpilrajn-Marczewski. Sur deux propriétés des classes d'ensembles. *Fundamenta Mathematicae*, 33(1):303–307, 1945. An English translation is available at http://webdocs.cs.ualberta.ca/~stewart/Pubs/MarczewskiTranslation.pdf.

[114] Asahi Takaoka. A recognition algorithm for adjusted interval digraphs. *Discrete Applied Mathematics*, 294:253–256, 2021.

[115] Robert Tarjan. Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2):146–160, 1972.

[116] William T. Trotter Jr. Interval graphs, interval orders, and their generalizations. *Applications of Discrete Mathematics, SIAM, Philadelphia, PA*, pages 45–58, 1988.

[117] Karsten Weihe. Covering trains by stations or the power of data reduction. In *Proceedings of Algorithms and Experiments, ALEX*, pages 1–8, 1998.

[118] Mihalis Yannakakis. The complexity of the partial order dimension problem. *SIAM Journal on Algebraic Discrete Methods*, 3(3):351–358, 1982.