Maximum number of edges in graph classes under degree and matching constraints

Erik Kvam Måland

Master Thesis in Computer Science
May 2015

Department of Informatics
University of Bergen
Norway

Dedicated with love to Marianne and Sverre, my parents.
Thanks for all you have teached me.

Acknowledgements

I would like to thank my supervisor Pinar Heggernes for great feedback and interesting discussions. Your kind attitude has been a driving force through my work on this thesis.

I would also like to say many thanks to all my classmates, both for comments on the text but most importantly for being good friends. You know very well who you are.

Contents

1 Introduction **1**
- 1.1 Terminology 1
- 1.2 Graph classes 3
  - 1.2.1 Chordal graphs 3
  - 1.2.2 Interval graphs 4
  - 1.2.3 Bipartite graphs 4
  - 1.2.4 Split graphs 5
  - 1.2.5 Line graphs 5
- 1.3 Overview of the thesis 6

2 Extremal graph theory **7**
- 2.1 The edge-extremal problem on general graphs 7
- 2.2 Relation to Ramsey numbers 10

3 Mathematical prelimenaries **13**
- 3.1 Bipartite graphs 13
- 3.2 The floor and ceiling functions 14
- 3.3 Optimization Diophantine equations 15

4 Split graphs and disjoint union of split graphs **19**
- 4.1 Split graphs 19
- 4.2 Disjoint union of split graphs 29

5 Unit interval graphs **35**

6 Factor-critical chordal graphs **45**

7 Conclusive remarks **51**
7.1 Summary 51
7.2 Comment on working with the thesis 52
7.3 Open problems 53

Chapter 1

Introduction

Graphs are mathematical objects that formalize the behavior of many different concepts. The problems defined on them arise from practical situations, i.e. internet traffic flow or schedule planning, or provide mathematical insight. A class of problems that might fit both of these characterizations are extremal problems on graphs. These problems seek to determine how large or small a set of elements defined on a graph can be, while satisfying some set of conditions. In this text we will ask how many edges a graph can have under restrictions on its maximum degree and matching number. We will call this the *edge-extremal problem*. The answer to this problem is known for arbitrary graphs. We will find the corresponding answer when the graphs belong to chosen graph classes. A *graph class* is a collection of graphs sharing some common property. Graph classes provide a systematic way to study how extremal values change when we impose some structure on the given graphs. If the maximum number of edges changes upon narrowing the graph class, we might be able to pinpoint exactly which structural features of the class allow the solution. The edge extremal problem is related to the notoriuosly hard problem of Ramsey numbers on graphs, as will be explained further on. This serves as a motivation to solve the problem in its own rights.

1.1 Terminology

A *graph* $G = (V, E)$ is a set of nodes V and edges $E \subseteq V \times V$. We may emphasize that V or E is the set of nodes or edges in G by writing $V(G)$ or $E(G)$, respectively. For an edge $e = (u, v)$, u and v are called its *endpoints*.

We say that e is *incident* with u and v. For a node u, the number of edges incident with u is called its *degree*, denoted by $deg_G(u)$. We may omit the subscript if it is clear of which graph we are speaking. The maximum degree in G is denoted by $\Delta(G)$. If there is an edge that has nodes u and v as endpoints, u and v are *adjacent*. A node adjacent to all nodes of G except itself is *universal* in G. The set of all nodes adjacent to u in G is called the *open neighbourhood* of u, denoted by $N_G(u)$. The *closed neighbourhood* of u is its open neighbourhood and itself, denoted by $N_G[u]$. For a set of nodes $U \subseteq V$, the *neighourhood of* U is defined by $N(U) = (\cup_{u \in U} N(u)) \setminus U$. The subscript G will be dropped if it is clear of which graph we are speaking. A set of nodes $U \subseteq V$ where no two nodes are adjacent, is an *independent set*. Two edges are said to be *adjacent* if they share a common endpoint. Edges that are not adjacent are *independent*. A set of independent edges of G is called a *matching*. The size of a largest matching in G is called its matching number, denoted by $\nu(G)$. A matching is said to *saturate* all nodes that are incident to some edge in the matching. A matching that saturates all nodes of G, is a *perfect matching*. If it sataurates all but one node of G, it is a *near-perfect matching*. If $G \setminus u$ has a perfect matching for all $u \in V(G)$, G is *factor-critical*. The *complement* of G, denoted by $\overline{G}$, is the graph on $V(G)$ where two nodes are adjacent if and only if they are not adjacent in G.

A graph $H = (V_H, E_H)$ with $V_H \subseteq V$ and $E_H \subseteq E$ is a *subgraph* of G. If all edges of G between nodes in V_H are present in H, then H is an *induced subgraph* of G. For a set of nodes $U \subseteq V$ the induced subgraph of G with U as the set of nodes is denoted by $G[U]$. U is said to induce the subgraph $G[U]$ in G. If G does not contain some graph H as an induced subgraph, G is H*-free*. A subgraph H, with $V(H) = V(G)$, is a *spanning subgraph*, or a graph that *spans* G. A graph on n nodes where all possible edges are present, is called a *complete graph*, denoted by K_n. A subset of V that induces a complete graph in G is a *clique*. A bipartite graph with partition (U, W) where every node of U is adjacent to every node of W, is called a *complete bipartite graph*, denoted by $K_{|U|,|W|}$. The complete bipartite graph $K_{1,i-1}$ is called an i*-star*. $K_{1,3}$ is called a *claw*.

The *union* of $G = (V, E)$ and $G' = (V', E')$ is defined by the new graph $G \cup G' = (V \cup V', E \cup E')$. Nodes are labeled in some way, and two nodes are equal if their labels are equal. If we label the nodes such that $V \cap V' = \emptyset$, the union of G and G' is equal to the *disjoint union* of G and G', denoted by $G \uplus G'$.

A sequence of distinct nodes $v_1, v_2, v_3, ..., v_{i-1}, v_i$ with $(v_{j-1}, v_j) \in E$, for

$2 \leq j \leq i$ is a *path* in G. The *length* of a path is the number of edges it contains. If the first and last nodes in a path are the same, it is a *cycle*. A cycle on n nodes is denoted bt C_n. An edge between non-consecutive nodes of a cycle is a *chord*. A cycle with no chord is an *induced cycle*.

A node $u \in G$ is said to *cover* all edges incident to it. A set $U \subseteq V(G)$, covering all edges of G, is called a *vertex cover*. The size of the smallest vertex cover of G is denoted by $\tau(G)$.

1.2 Graph classes

A *graph class* is a set of graphs sharing a common property. When considering a problem on graphs, it is often useful to restrict the instances to a certain graph class. The common structural feature of the class might change the character of the problem. This approach is often applied in algorithmic graph theory. We will not be concerned with algortihms in this text. However, we will be considering the egde extremal problem on different graph classes. These classes will be presented in this section. Further properties will be given when needed.

1.2.1 Chordal graphs

A graph is *chordal* if it does not contain an induced cycle of length 4 or more. This is equivalent to saying that a graph is chordal if every cycle of length 4 or more has a chord. Figure 1.1 illustrates this.

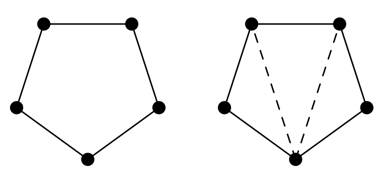

Figure 1.1: The left graph is not chordal since it contains and induced cycle with length 4 or more. Filling in the dashed chords, shown to the right, makes it chordal

1.2.2 Interval graphs

Many graph classes are defined in terms of a model based on some mathematical structure. This may be geometrical figures or even other graphs. In the class of *interval graphs*, each node corresponds to a closed interval on the real line. There is an edge between two nodes if and only if the corresponding intervals overlap. The set of intervals representing the nodes is called the *interval representation* of the graph. This is illustrated in Figure 1.2. If we restrict the length of the invervals to be the same, we have a *unit interval graph*.

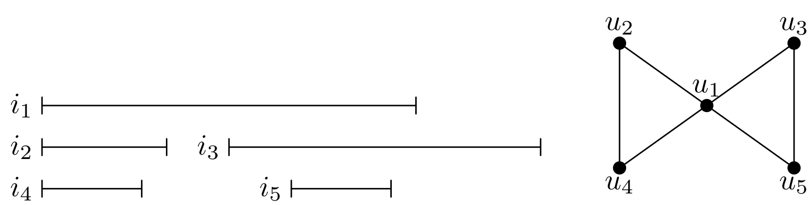

Figure 1.2: An interval graph and its interval representation. Node u_k corresponds to interval i_k

1.2.3 Bipartite graphs

A graph is *bipartite* if its nodes can be partitioned into two sets U and W such that every edge is between a node in U and a node in W. Equivalently, one may say that its nodes can be colored black or white such that each edge is between nodes of different colors. Figure 1.3 shows some examples of bipartite graphs.

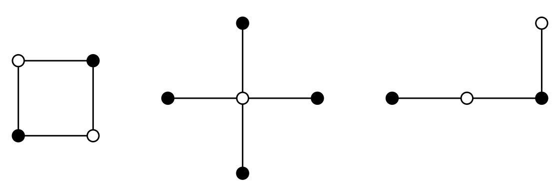

Figure 1.3: Bipartite graphs. The colors of the nodes are indicated

An important subclass of bipartite graphs is the class of *forests*, graphs that do not contain a cycle. A connected forest is a *tree*.

1.2.4 Split graphs

A *split graph* is a graph G whose set of nodes can be partitioned into sets I and C, such that $G[I]$ is an independent set and $G[C]$ is complete. We say that G has *split partition* (C, I). A *full split graph* is a split graph where all possible edges between C and I are present. The class of split graphs is a subclass of chordal graphs. Figure 1.4 shows some examples.

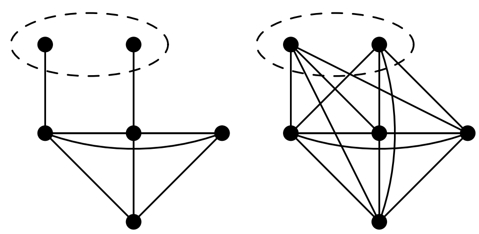

Figure 1.4: Split graphs. The graph to the right is a full split graph. The independent set is circled in each instance

1.2.5 Line graphs

Let $G = (V, E)$ be a graph. Construct a new graph where each node corresponds to an edge in G and two nodes are adjacent if and only if their corresponding edges in G have a common endpoint. This is called the *line graph* of G, denoted by $L(G)$. In general, a graph H is a line graph if $H = L(G)$ for some G. This is illustrated with an example in Figure 1.5.

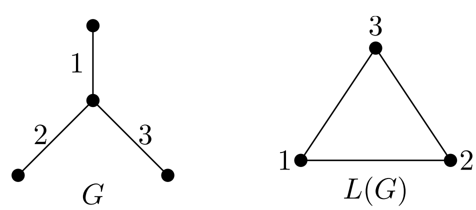

Figure 1.5: A graph G and its line graph $L(G)$. The labels show the corresonding edges and node

1.3 Overview of the thesis

The necassary terminology and overview of graph classes are now presented. In the next chapter, Chapter 2, we will present the main problem of the thesis, the edge-extremal problem. We will do so by giving its solution on general graphs. This result is not worked out by us. The connection with Ramsey numbers will also be explained in this chapter. From Chapter 3, the results presented are ours. In Chapter 3 we present the solution of the edge-extremal problem on bipartite graphs. This is our way of introducing the problem on narrower graph classes than general graphs. Also, some necassary mathematical machinery will be given in this chapter. Chapter 4 gives the solution of the edge-extremal problem on split graphs and disjoint union on split graphs. In Chapter 5 we solve the edge-extremal problem on unit interval graphs. A result obtained as a by-product of the work on the edge-extremal problem on chordal graphs is presented in Chapter 6. This result is a characterization of factor-critical chordal graphs in terms of spanning subgraphs. Finally, summary, comments on the work and open problems are given in Chapter 7.

Chapter 2

Extremal graph theory

Given that a graph does not contain K_r as a subgraph, for some r, what is the maximum number of edges it can have? How large may the matching number of an acyclic graph on n nodes be? These are examples of questions studied in extremal graph theory. Generally, it is the study of how large or small some parameter of a graph can be, satisfying certain constraints. The main problem of this thesis is of this nature. Here we seek to maximize the number of edges, given constraints on maximum degree and matching number. We will consider this problem for graphs from different graph classes. This section gives more details and presents the solution for the problem on general graphs, which is already known. For a more elaborate treatment of extremal graph theory, see [2].

2.1 The edge-extremal problem on general graphs

In this section we have our first look at the main problem studied in this thesis, the edge-extremal problem. Its solution on general graphs is already known from [1], and we give a brief description of it here. For a more detailed presentation, we direct the reader to this paper.

Let $\mathcal{S}$ denote a graph class. $\mathcal{M}_{\mathcal{S}}(i,j)$ denotes all graphs G from $\mathcal{S}$ satisfying $\Delta(G) < i$ and $\nu(G) < j$. Also, no other graph than K_1 in $\mathcal{M}_{\mathcal{S}}(i,j)$ contains isolated nodes. The edge-extremal problem studied in this thesis is the following: Given i,j and $\mathcal{S}$, what is the maximum number of edges a graph in $\mathcal{M}_{\mathcal{S}}(i,j)$ can have? A graph achieving the maximum number of edges is called *edge-extremal* in $\mathcal{M}_{\mathcal{S}}(i,j)$.

Let $\mathcal{GEN}$ denote the class of general graphs. Recall that a graph G is factor critical if $G \setminus u$ has a perfect matching for all $u \in V(G)$. For a proof of the next lemma, see [6].

Lemma 2.1. *[Gallai's lemma] Let G be a connected graph. If $\nu(G) = \nu(G \setminus v)$ for all $v \in V(G)$, then G is factor-critical.*

For the next lemma, recall that an i-star is the complete bipartite graph $K_{1,i-1}$. The proof is recited from [1].

Lemma 2.2. *Let G be an edge-extremal graph in $\mathcal{M}_{\mathcal{GEN}}(i,j)$ containing the maximum number of i-stars. Then all connected components of G not an i-star are factor critical.*

Proof. The proof is by contradiction. Assume that a connected component H of G, not an i-star, is not factor critical. This implies that there exists a node $v \in H$, such that $\nu(H) > \nu(H \setminus v)$, by Gallai's lemma. Consider the graph $G' = (G \setminus v) \uplus K_{1,i-1}$. Since $\nu(K_{1,i-1}) = 1$ and $\nu(G') < \nu(G) < j$, it follows that $\nu(G') < j$. Clearly, $\Delta(G') < i$. Also, since $deg(v) < i$, we have

$$\begin{aligned} |E(G')| &= |E(K_{1,i-1})| + |E(G \setminus v)| = (i-1) + |E(G \setminus v)| \\ &\geq (i-1) + (|E(G)| - (i-1)) = |E(G)|. \end{aligned}$$

So G' is also edge-extremal. This contradicts the assumption that G contains the largest possible number of i-stars, and completes the proof. $\square$

It follows that edge-extremal graphs G in $\mathcal{M}_{\mathcal{GEN}}(i,j)$ have two types of connected components; factor-critical and i-stars. Let $\mathcal{C}$ be a factor-critical component of G. Observe that Gallai's lemma implies that $\mathcal{C}$ has a near-perfect matching and $|V(\mathcal{C})| = 2\nu(\mathcal{C}) + 1$. The value of $|E(\mathcal{C})|$ can not be larger than $\lfloor \frac{|V(\mathcal{C})|(i-1)}{2} \rfloor = \lfloor \frac{(2\nu(\mathcal{C})+1)(i-1)}{2} \rfloor$. We can make this bound tighter in the case where $|V(\mathcal{C})| = 2\nu(\mathcal{C}) + 1 < i$. Then $|E(\mathcal{C})|$ is no more than $\frac{|V(\mathcal{C})|(|V(\mathcal{C})|-1)}{2} = \frac{(2\nu(\mathcal{C})+1)2\nu(\mathcal{C})}{2}$. This gives

$$|E(\mathcal{C})| \leq \min\{(2\nu(\mathcal{C})+1)\nu(\mathcal{C}), \lfloor \frac{(2\nu(\mathcal{C})+1)(i-1)}{2} \rfloor\}. \tag{2.1}$$

If G has A i-stars and B factor critical components, this implies

$$|E(G)| \leq (i-1)A + \Sigma_{i=1}^{B}\{(2\nu(\mathcal{C})+1)\nu(\mathcal{C}), \lfloor \frac{(2\nu(\mathcal{C})+1)(i-1)}{2} \rfloor\}. \tag{2.2}$$

The bound on matching number is expressed by

$$A + \Sigma_{i=1}^{B} \nu(\mathcal{C}_i) < j, \tag{2.3}$$

where $\mathcal{C}_i$ is the ith factor-critical component. This leads to a linear program, which gives an uppper bound on edges in edge-extremal graphs. For details, see [1].

The solution obtained with this procedure depends on i being odd or even. For odd i, G is a disjoint union of K_i and i-stars, where the number of K_i is as large as possible. Without going into the details of the calculation, the number of edges in edge-extremal G is given by

$$|E(G)| = (i-1)(j-1) + \frac{i-1}{2}\lfloor\frac{j-1}{\frac{i-1}{2}}\rfloor. \tag{2.4}$$

An edge-extremal instance in $\mathcal{M}_{\mathcal{GEN}}(5,8)$ is shown in Figure 2.1. For this example, Equation 2.4 gives $|E(G)| = (5-1)(8-1) + \frac{5-1}{2}\lfloor\frac{8-1}{\frac{5-1}{2}}\rfloor = 34$.

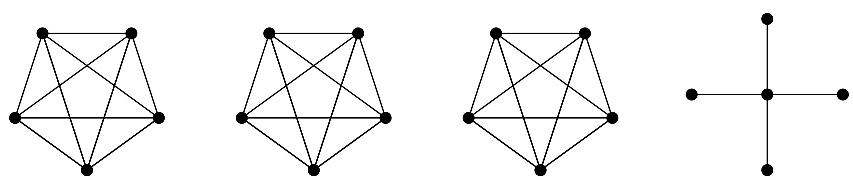

Figure 2.1: An edge-extremal instance in $\mathcal{M}_{\mathcal{GEN}}(5,8)$

In the case where i is even, the factor-critical components are created by the following process: Remove a maximum matching from K_i, introduce a new node v and add an edge from v to any of the $i-1$ nodes in the modified graph. This is illustrated in Figure 2.2.

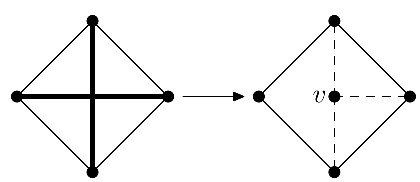

Figure 2.2: Transforming K_4 to a factor-critical component in the edge-extremal instance for $i = 4$. The removed matching is bold and added edges are dashed

We refer to this modified K_i as K_i'. An edge-extremal instance for even i is a disjoint union of K_i' and i-stars, where the number of K_i' is as large as possible. Again, without giving details, the number of edges in this case is given by

$$|E(G)| = (i-1)(j-1) + \frac{i}{2}\left\lfloor\frac{j-1}{\frac{i}{2}}\right\rfloor. \tag{2.5}$$

An edge-extremal graph in $\mathcal{M}_{\mathcal{GEN}}(4,8)$ is shown in Figure 2.3. For this example, Equation 2.5 gives $|E(G)| = (4-1)(8-1) + \frac{4}{2}\left\lfloor\frac{8-1}{\frac{4}{2}}\right\rfloor = 24$.

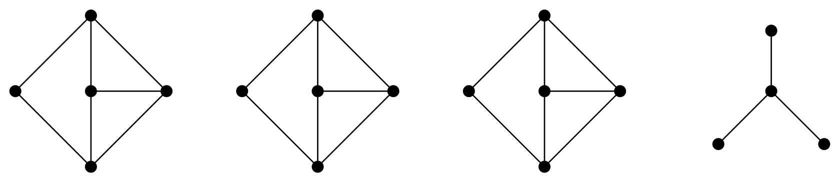

Figure 2.3: An edge-extremal instance in $\mathcal{M}_{\mathcal{GEN}}(4,8)$

2.2 Relation to Ramsey numbers

In *Ramsey theory* we investigate how global assumptions imply local properties. In particular, we ask the following question: for some m, n, how many nodes does a graph G need to have to make sure that it contains K_m or $\overline{K_n}$ as an induced subgraph? This number of nodes is called the *Ramsey number* $R(m,n)$. As an example, consider a graph G on 6 nodes. Look at a node $u \in V(G)$. There has to be at least three nodes $v_1, v_2, v_3 \in V(G)$ adjacent to u, or else G contains $\overline{K_3}$. At least two of v_1, v_2, v_3 have to be adjacent. Assume without loss of generality that $(v_1, v_2) \in E(G)$. But then u, v_1, v_2 induces K_3. This shows that $R(3,3) \leq 6$. This bound is tight, which is seen by considering a cycle on 5 nodes, C_5. This graph does not contain K_3 or $\overline{K_3}$, so 6 nodes is the smallest number to ensure this. Some examples are shown in Figure 2.4.

Ramsey theory has been formulated in many different manners. What is presented in this section, is commonly referred to as Ramsey theory on graphs. The calculation of Ramsey numbers is a notoriously difficult problem, and no general method is known. A bound due to Erdõs and Szekeres [3] is given by:

$$R(m,n) \leq \binom{m+n-2}{m-1}. \tag{2.6}$$

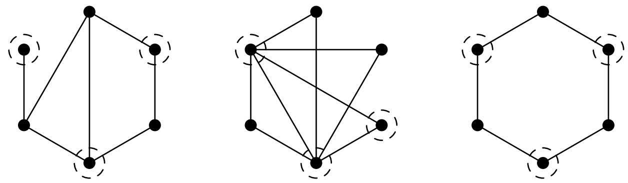

Figure 2.4: Some graphs on 6 nodes. Nodes inducing K_3 or $\overline{K_3}$ are circled

So why are Ramsey numbers presented in this thesis? The problem we are trying to solve can actually be formulated in terms of determining Ramsey numbers on line graphs. To see this, consider the problem on general graphs. We are looking for the maximum number of edges in graphs G, with $\Delta(G) < i$ and $\nu(G) < j$. Let $L(G)$ be a line graph of G. Edges that meet in a node in G are mutually adjacent nodes in $L(G)$ and vice versa. By limiting how many edges that can meet in a single node in G, we are limiting how many nodes that can be mutually adjacent in $L(G)$, which is the same as limiting the size of the largest clique in $L(G)$. Similarly, independent edges in G, a matching, are independent nodes in $L(G)$ and vice versa. So by limiting the matching number of G, we are limiting the size of the largest independent set in G. Maximizing edges in G is the same as maximizing nodes in $L(G)$. The edge-extremal graphs G with $\Delta(G) < i$ and $\nu(G) < j$ correspond to the line graphs with the largest possible number of nodes that do not contain K_i or $\overline{K_j}$ as an induced subgraph. This is exactly $R(i, j) - 1$ for $L(G)$. By considering a particular graph class, we are in essence finding the Ramsey numbers for the line graphs of this graph class. We will not elaborate any further on this through the text.

Chapter 3

Mathematical prelimenaries

This chapter will start with solving the edge extremal problem on bipartite graphs. Hopefully, this will familiarize the reader with the problem on other graph classes than general graphs, before embarking on more demanding classes. Upon solving the problem on unit interval graphs in Chapter 5, we will encounter some mathematical problems that are best treated separately. The rest of the section will be devoted to these. This includes some properties of the floor and ceiling functions and a special type of equation we have chosen to call optimization Diophantine equations. Where no references are given, the results are worked out by us.

3.1 Bipartite graphs

The class of bipartite graphs will be denoted by $\mathcal{BIP}$. To solve the edge-extremal problem on this graph class, we start by introducing a classical theorem in graph theory. For a proof of this, see [6].

Theorem 3.1. *[König's theorem] Let G be a bipartite graph. Then $\nu(G) = \tau(G)$.*

We will also be needing the next observation, given as a lemma.

Lemma 3.2. *Let G be a graph with $\Delta(G) < i$. Then $|E(G)| \leq \tau(G)(i-1)$*

Proof. Let VC be a vertex cover of G. Every edge of G is incident with a node in VC. Also, every node in VC is associated with at most $i-1$ edges, and the lemma follows. □

The next theorem gives a tight bound on the number of edges for graphs in $\mathcal{M}_{\mathcal{BIP}}(i,j)$.

Theorem 3.3. *Let G be a graph in $\mathcal{M}_{\mathcal{BIP}}(i,j)$. Then $|E(G)| \leq (i-1)(j-1)$. This bound is tight.*

Proof. From Theorem 3.1 we know that $\tau(G) = \nu(G) \leq j-1$. Lemma 3.2 then implies $|E(G)| \leq \tau(G)(i-1) \leq (j-1)(i-1)$. The disjoint union of $j-1$ i-stars is bipartite and has $(i-1)(j-1)$ edges. Thus, the bound is tight and the proof is complete. □

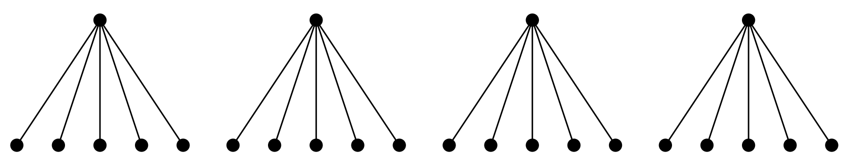

Figure 3.1: An edge-maximal graph in $\mathcal{M}_{\mathcal{BIP}}(6,5)$

An edge-extremal graph in $\mathcal{M}_{\mathcal{BIP}}(i,j)$ is shown in Figure 3.1.

3.2 The floor and ceiling functions

In this section we define and describe properties of the floor and ceiling functions that will be useful later. The first lemma is clear, so we have omitted its proof.

Lemma 3.4. *Let $k \in \mathbb{N}$. If $n = 2k$, then*

$$\left\lfloor \frac{n}{2} \right\rfloor = \left\lceil \frac{n}{2} \right\rceil = k. \tag{3.1}$$

If $n = 2k+1$, then

$$\left\lceil \frac{n}{2} \right\rceil - 1 = \left\lfloor \frac{n}{2} \right\rfloor = k. \tag{3.2}$$

Lemma 3.5. *For $n, x, y \in \mathbb{N}$, and $n = x + y$ we have*

$$\left\lfloor \frac{n}{2} \right\rfloor \geq \left\lfloor \frac{x}{2} \right\rfloor + \left\lfloor \frac{y}{2} \right\rfloor \tag{3.3}$$

Proof. We split the proof into two parts; n odd and n even. First, assume that n is odd, such that $n = 2k + 1$ for some $k \in \mathbb{N}$. Also, assume, without loss of generality, that x is odd and y is even, such that $x = 2l + 1$, $y = 2m$ for $l, m \in \mathbb{N}$. We have $n = x + y = 2k + 1 = 2(l + m) + 1$, so $k = l + m$. According to Lemma 3.4 this gives

$$\left\lfloor \frac{n}{2} \right\rfloor = k = l + m = \left\lfloor \frac{x}{2} \right\rfloor + \left\lfloor \frac{y}{2} \right\rfloor \tag{3.4}$$

For n even, $n = 2k$, we have two choices; x and y both even, or both odd. This gives, respectively,

$$\left\lfloor \frac{n}{2} \right\rfloor = k = \left\lfloor \frac{x}{2} \right\rfloor + \left\lfloor \frac{y}{2} \right\rfloor = \left\lfloor \frac{2l}{2} \right\rfloor + \left\lfloor \frac{2m}{2} \right\rfloor = l + m, \tag{3.5}$$

and

$$\left\lfloor \frac{n}{2} \right\rfloor = k = l + m + 1 > \left\lfloor \frac{x}{2} \right\rfloor + \left\lfloor \frac{y}{2} \right\rfloor = l + m. \tag{3.6}$$

The second equality is because $n = 2k = x + y = 2l + 1 + 2m + 1 = 2(l + m + 1) \Rightarrow k = l + m + 1$. □

3.3 Optimization Diophantine equations

In this section we will look at equations in variables x, y and r on the form

$$j = ax + by + r, \tag{3.7}$$

where $j, a, b \in \mathbb{N}$ and $a - b = 1$. We will call these *optimization Diophantine equations*. To solve this equation we must assign non-negative integer values to x, y and r such that $0 \leq r < b$. Only if $b = 0$, the value of r may be equal to b. We will refer to r as the *remainder*. We are interested in solutions where r is as small or as large as possible. More specifically, we seek the values or r in these cases. If there is a solution such that $r = r_1$, we say that r_1 is an *admissable value* of r. Before we work out this equation, we need a lemma from [4]. For a proof, we direct the reader there.

Lemma 3.6. *If m is a positive integer and n is any integer, there exist unique integers q and p such that*

$$n = mq + p, \tag{3.8}$$

where $0 \leq p < m$.

The following two lemmas give the smallest and largest admissable value of r for a given optimization Diophantine equation.

Lemma 3.7. *Given an equation on the form of Equation 3.7, let*

$$j = aq_1 + r_1 = bq_2 + r_2$$

such that $0 \leq r_1 < a$ *and* $0 \leq r_2 < b$. *The smallest admissable value of* r *is* $\max\{0, r_2 - q_2\}$.

Proof. Let us first show that $\max\{0, r_2 - q_2\}$ is an admissable value of r. We know that $a - b = 1$, so $a - b - 1 = 0$. We add this k times to both sides of $j = bq_2 + r_2$ for some non-negative k. This gives

$$j + 0 \cdot k = bq_2 + r_2 + k(a - b - 1) = (q_2 - k)b + ak + (r_2 - k).$$

$q_2 - k$ and $r_2 - k$ have to be non-negative in a solution. So we can not add $a - b - 1$ more than $\min\{q_2, r_2\}$ times. We have two cases; if $q_2 \geq r_2$, we can set $k = r_2$ to get a remainder of 0. If $q_2 < r_2$, we get the smallest remainder by setting $k = q_2$. This gives a remainder of $r_2 - q_2$. Since $r_2 - q_2 \leq r_2 < b$, the remainder is strictly smaller than b in this case, as demanded. Note that when $q_2 < r_2$, the smallest value of r is the remainder when j is divided by a. This shows the existence of a solution where $r = \max\{0, r_2 - q_2\}$.

Let us now show that this is the smallest admissable value of r. In the case where r takes value 0 there is nothing to prove; it can not be less than this by definition. So assume that $q_2 < r_2$ and assume for contradiction that r' is the smallest admissable value of r, where $0 \leq r' < r_2 - q_2$. There must exist non-negative integers x' and y' such that $j = ax' + by' + r'$. Note that y' must be equal to 0, or else we would add $a - b - 1$ to both sides to obtain an even smaller value of r than r'. So we must have $j = ax' + r'$ with $0 \leq r' < r < b$. As noted above, $r_2 - q_2$ is the remainder of j divided by a. But we know that the value of the remainder by division is unique, so r' has to be equal to $r_2 - q_2$. This contradicts the assumption, and the lemma follows. □

Lemma 3.8. *Given an equation on the form of Equation 3.7, let*

$$j = aq_1 + r_1 = bq_2 + r_2 \tag{3.9}$$

such that $0 \leq r_1 < a$ *and* $0 \leq r_2 < b$. *The largest admissable value of* r *in this equation is* $\min\{b - 1, q_1 + r_1\}$.

Proof. The proof is similar to that of the previous lemma. Let us first show that $\min\{b-1, q_1+r_1\}$ is an admissable value of r. We know that $a-b=1$, so $-a+b+1=0$. We add this k times to both sides of $j = aq_1 + r_1$ k for some non-negative integer k. This gives

$$j + 0 \cdot k = aq_1 + r_1 + k(-a+b+1) = (q_1-k)a + bk + (r_1+k). \quad (3.10)$$

To keep $q_1 - k$ non-negative, we must have $k \leq q_1$. Also, since the remainder must be strictly less than b, we must have $r_1 + k \leq b-1$. As in Lemma 3.7, there are two cases; if $r_1 + q_1 \geq b-1$, the largest value of the remainder is achieved by setting $k = (b-1)-r_1$. This will give $r = b-1$. If $r_1+q_1 < b-1$, k is constrained by $k \leq q_1$, and the largest value of r is achieved by setting $k = q_1$. Inserting this into the last term of 3.10 gives $r = r_1 + q_1$. Note that when $r_1+q_1 < b-1$, the largest value of r is the remainder when j is divided by b. This show the existence of a solution where $r = \min\{b-1, q_1+r_1\}$.

Let us now show that this is the largest admissable value of r. In the case where r takes value $b-1$, there is nothing to prove; it can not be larger than this by definition. So assume that $r_1+q_1 < b-1$ and assume for contradiction that r' is the largest admissable value of r, where $r < r' < b$. There must exist non-negative integers x' and y' such that $j = ax' + by' + r'$. Note that a has to be equal to 0, or else we would add $-a+b+1$ to both sides to obtain an even larger value of r than r'. So we must have $j = by' + r'$ with $0 \leq r < r' < b-1$. As noted above, $q_1 + r_1$ is the remainder of j divided by b. But we know that the value of the remainder by division is unique, so r' has to be equal to $q_1 + r_1$. This contradicts the assumption, and the lemma follows. □

Chapter 4

Split graphs and disjoint union of split graphs

In this chapter, we solve the edge-extremal problem on split graphs and disjoint union of split graphs.

4.1 Split graphs

In this section we will solve the edge-extremal problem on split graphs. Since we do not allow isolated nodes, this graph class has a fundamental difference from general graphs; they are connected. The flavour of the edge-extremal problem therefore becomes a little different from the cases we have already seen. For instance, we can not use factor-criticality, because this relies on the possibility of adding stars to the graph.

We will follow a general scheme when solving the edge-extremal problem on split graphs. First, we will introduce lemmas that impose some structural properties on the edge-extremal instances of split graphs. When we have enough information about the structure of these instances, we will approach the problem analytically. This will result in different variants of edge-extremal instances and a closed formula for the number of edges in each case.

Some necassary results on bipartite graphs will be presented now. These are already known. Let G be a bipartite graph with bipartition (A, B) and let X be a subset of A. The *deficiency* of the set X in G, written as $def_G(X)$, is defined by

$$def_G(X) = |X| - |N(X)|. \tag{4.1}$$

The subscript G is omitted when it is clear of which graph we are speaking. This is exemplified in Figure 4.1. The maximum deficiency of all subsets of A is called the A-*deficiency* of G, or just the deficiency of G, denoted by $def(G)$. Note that $def_G(\emptyset) = |\emptyset| - |N(\emptyset)| = 0$, so $def(G) \geq 0$.

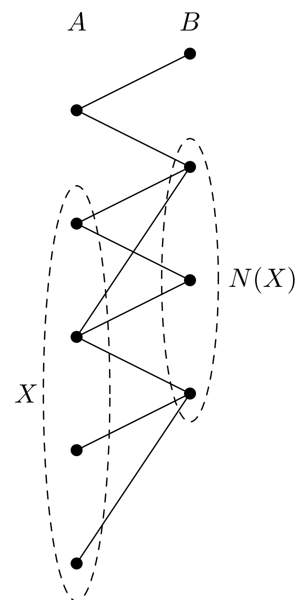

Figure 4.1: The deficiency of X is $def_G(X) = |X| - |N(X)| = 4 - 3 = 1$

The A-deficiency of a bipartite graph is closely related to the size of the maximum matching, as shown by the next lemma. This lemma is given in [6], but the proof is omitted there. We have worked out a proof here based on König's theorem, Theorem 3.1..

Lemma 4.1. *Let G be a bipartite graph with bipartition (A, B). Then $\nu(G) = |A| - def(G)$.*

Proof. We will prove the lemma by showing that the size of the minimum vertex cover of G is equal to $|A| - def(G)$. From König's theorem we then know that $\nu(G) = |A| - def(G)$. First we show that there exists a vertex cover of size $|A| - def(G)$. Let $X \subseteq A$ be the set for which $|X| - |N(X)|$ is maximized, so $def(G) = |X| - |N(X)|$. The set $VC = (A \setminus X) \cup N(X)$ forms a vertex cover of G. The size of VC is given by $|VC| = |A| - |X| + |N(X)| = |A| - (|X| - |N(X)|) = |A| - \mathit{def}(G)$.

Next we show that this is the minimum possible vertex cover can have. We do this by contradiction, so assume that there is a minimum vertex cover VC with $|VC| < |A| - \mathit{def}(G)$. Let $V_A = VC \cap A$. Then $A \setminus V_A$ is the set of nodes in A that are not in VC. We can form a vertex cover by adding these nodes to $N(V_A)$, obtaining a vertex cover of size $|A| - |V_A| + |N(V_A)|$. Since VC is minimum, we have $|VC| \leq |A| - |V_A| + |N(V_A)|$. This implies the following strict inequalities

$$\begin{aligned} |A| - |V_A| + |N(V_A)| &< |A| - \mathit{def}(G) \\ -|V_A| + |N(V_A)| &< -\mathit{def}(G) \\ -\mathit{def}(V_A) &< -\mathit{def}(G) \\ \mathit{def}(V_A) &> \mathit{def}(G) \end{aligned}$$

We have inserted $-|V_A| + |N(V_A)| = -\mathit{def}(V_A)$. This is a contradiction, since $\mathit{def}(G)$ is the maximum deficiency of all subsets of A, and the proof is complete. □

For the rest of this section, let $\mathcal{SPLIT}$ be the class of split graphs and G a graph in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$ with split partition (C, I). The rest of the results in this section are new. We start with two easy properties, stated as lemmas for the sake of reference.

Lemma 4.2. *Assume that we add or remove edges between nodes in C and I in G, while keeping (C, I) a valid partition. This will never make a node in I violate $\Delta(G) < i$.*

Proof. We can safely assume that I and C are non-empty, since in that case there is nothing to prove. Graphs in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$ have no isolated nodes. Therefore there is a node $c \in C$ that is adjacent to some node in I and $deg_G(c) \geq |C|$. A node in $i_1 \in I$ can not have higher degree than $|C|$ as long as (C, I) is kept a valid partition. It follows from $deg_G(c) < i$ and $deg_G(i_1) \leq deg_G(c)$ that $deg_G(i_1) < i$ no matter how the connectivity between I and C is altered. The proof is complete. □

We will make frequent references to the bipartite subgraph of G with bipartition $(N(I), I)$ and all edges with an endpoint in I. Denote this subgraph by $BIP(G)$. This is illustrated in Figure 4.2.

It is central in deciding the maximum matching of G.

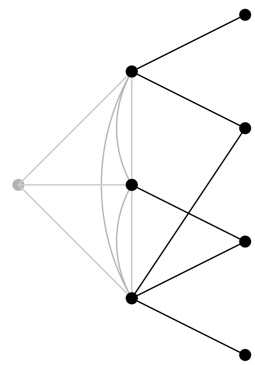

Figure 4.2: The subgraph $BIP(G)$ is colored black

Lemma 4.3. $\nu(G) = \nu(BIP(G)) + \left\lfloor \frac{|C| - \nu(BIP(G))}{2} \right\rfloor$.

Proof. The formula in the statement expresses the size of the maximum matching is equal to that which saturates as many nodes in $BIP(G)$ as possible. To see that this is indeed maximum, look at a maximal matching that does not saturate the maximum possible number of nodes in $BIP(G)$. Rearranging the matching, keeping it maximal, such that more nodes in $BIP(G)$ are saturated and will never decrease the size of the maximal matching. It will only liberate nodes in C. □

Our goal is to transform G to a form suitable for analysis. The transformation must preserve membership in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$ and not decrease the number of edges. As a first step towards this, we introduce the following operation:

Definition 4.4. *The transformation ϕ takes four parameters: split graph G, node $x \in C$ and nodes $y_1, y_2 \in I$. It returns a modified graph $\phi(G, x, y_1, y_2)$ according to the following scheme:*

If $(x, y_1) \in E(G)$ and $(x, y_2) \notin E(G)$:

Remove (x, y_1) and add (x, y_2). In addition, for all nodes $x' \in C$ that are adjacent to y_1 and not y_2, remove (x', y_1) and add (x', y_2).

If not:

Do nothing.

If this isolates any nodes of G, remove them.

This is illustrated in Figure 4.3. For every edge that is removed by ϕ, another edge is added, so it is clear that this operation conserves the number of edges of G.

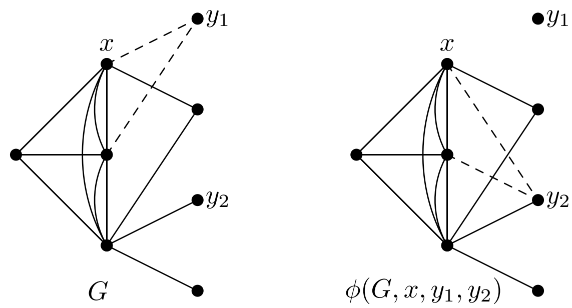

Figure 4.3: An example of ϕ. The dashed edges to the left are removed, and those to the right are added. We also need to remove the isolated node y_1.

The next lemma proves that ϕ conserves membership in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$.

Lemma 4.5. *If* $G \in \mathcal{M}_{\mathcal{SPLIT}}(i, j)$*, then* $\phi(G, x, y_1, y_2) \in \mathcal{M}_{\mathcal{SPLIT}}(i, j)$ *for all* $x \in C$ *and* $y_1, y_2 \in I$*.*

Proof. We have to show that $\Delta(\phi(G, x, y_1, y_2)) < i$ and $\nu(\phi(G, x, y_1, y_2)) < j$. For the sake of notational simplicity we write $G_\phi = \phi(G, x, y_1, y_2)$.

We first show that the bound on maximum degree is conserved. The operation ϕ only alters the connectivity between nodes in C and nodes in I. Lemma 4.2 says that this does not alter $\Delta(G)$, so $\Delta(\phi(G, c, i_1, i_2)) < i$ is satisfied.

Now to the bound on the matching number. From Lemma 4.3 we know that, given G, $\nu(G)$ is only dependent on $\nu(BIP(G))$. Furthermore, from Lemma 4.1 we know that $\nu(BIP(G))$ is determined by $\mathit{def}(BIP(G))$, or $\nu(BIP(G)) = |I| - \mathit{def}(BIP(G))$. So if ϕ conserves $\mathit{def}(BIP(G))$, then $\nu(G_\phi) < j$. We will show that this is the case. $\mathit{def}(BIP(G))$ is changed by ϕ if and only if $|C'| - |N(C')|$ is changed by ϕ, for some $C' \subseteq C$. Since ϕ only changes the connectivity of G, the only magnitude that can change in this expression is $|N(C')|$. It is only necassary to consider those $C' \subseteq C$ for which $N(C')$ contains y_1 or y_2, since every other set is left unchanged by ϕ. There are three possibilities to consider; $N(C')$ contains only y_1, only y_2 or both. If $N(C')$ contains only y_1 in G, then it does not contain y_1 in G_ϕ. However, it does contain y_2 in G_ϕ, so $|N(C')|$ is the same in G and G_ϕ. If

$N(C')$ contains only y_2 in G, then $G = G_\phi$ and $|N(C')|$ is the same in G and G_ϕ. Lastly, if $N(C')$ contains y_1 and y_2 in G, then it only contains y_2 in G_ϕ, so $|N(C')|$ is less in G_ϕ than in G. $|N(C')|$ does not increase in any of these instances, which implies that *def*$(BIP(G))$ does not decrease, which again implies that $\nu(G_\phi) \leq \nu(G)$ and $G_\phi < j$. This completes the proof. □

We will be using ϕ to transform G to a form that is suitable to approach analytically. But first we need to extend the operation.

Definition 4.6. *The transformation χ takes a split graph G with partition (C, I) and two nodes $y_1, y_2 \in I$ and returns G after the following loop:*

> *for all $x \in C$:*
> $\quad G = \phi(G, x, y_1, y_2)$

A little informally, $\chi(G, y_1, y_2)$ is $\phi(G, x, y_1, y_2)$ performed for all nodes $x \in C$.

Assume that α is some ordering of the nodes of I, $(i_1, i_2, ..., i_{|I|})$. Let α_k denote the tuple $(i_k, i_{k+1}, ..., i_{|I|})$. We define a last operation, building on χ:

Definition 4.7. *The transformation ψ_k takes a split graph G with partition (C, I) and one node $y_1 \in I$ and returns G after the following loop:*

> *for all $y_2 \in \alpha_k$:*
> $\quad G = \chi(G, y_1, y_2)$

In programming terminology, ψ_k this is similar to a double for-loop with ϕ in the innermost loop. From Lemma 4.5 it is clear that ψ_k conserves membership in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$; after all, ψ_k is just ϕ performed repeatedly.

An ordering of the nodes in I, $(i_1, i_2, ..., i_{|I|})$, such that $N(i_j) \subseteq N(i_{j-1})$, for $2 \leq j \leq |I|$ will be called a *inclusion ordering* of G. Note that this does not exist in all split graphs. A node in I that is adjacent to all nodes in $N(I)$ will be called *locally universal*. We are now ready to prove the next lemma, which is the most technically demanding in this section.

Lemma 4.8. *For each i and j, there is an edge-extremal graph $G \in \mathcal{M}_{\mathcal{SPLIT}}(i, j)$ such that G has an inclusion ordering.*

Proof. Assume that G has partition (C, I). The proof is by induction on the nodes of I. Consider the following procedure on G: perform $\psi_k(G, x)$ for $k = 1, 2, 3, ...$ in that order. Our claim is that this leaves G with an inclusion ordering. As the base case, we show that $\psi_1(G, i_1)$ leaves i_1 locally universal.

Consider some node $c \in N(I)$. If it already is adjacent to i_1, then we are done. If not, there exists a node i'_1 which is adjacent to c, since c is in $N(I)$. When $\phi(G, c', i'_1, i_1)$ is performed, c and i_1 becomes adjacent. Since ψ_k never removes edges from i_1, this node is left locally universal by ψ_1. This implies that $N(i_k) \subseteq N(i_1)$ for all $1 \leq k \leq |I|$.

For the inductive case, assume that $N(i_1) \supseteq N(i_2) \supseteq ... \supseteq N(i_k)$. Consider some node $c' \in N(I)$ which is adjacent to i_{k+1} after $\psi_{k+1}(G, i_{k+1})$ is performed. We will show by contradiction that c' is always adjacent to i_k. Note that ψ_{k+1} does not remove any edges from i_l, where $1 \leq l \leq k$. There are two possibilities for c'; it was adjacent with i_{k+1} before ψ_{k+1} was performed, or it was made adjacent by ψ_{k+1}. For each case, assume that c' is not adjacent to i_k. Assume first that c' was adjacent with i_{k+1} before the application of ψ_{k+1}. c' can not be adajcent to any node i_l, with $l \geq k + 2$, as the application of $\psi_k(G, i_k)$, in particular $\phi(G, c', i_k, i_l)$, would have made c' adjacent to i_k. But then $\phi(G, c', i_k, i_{k+1})$ must have been performed as a part of ψ_k. This would have made c' and i_k adjacent, and we have reached the desired contradiction in the first case. Secondly, assume that c' was made adjacent to i_{k+1} by the application of ψ_{k+1}. This means that before the application of ψ_{k+1}, c' is adjacent to some i_l, where $l \geq k + 2$. But this would also be the case when $\psi_k(G, i_k)$ is performed. And since ψ_k is performed before ψ_{k+1}, i_k must have been made adjacent by ψ_k. Again, ψ_{k+1} does not remove any edges from i_l, where $l \leq k + 1$, so we have reached the desired contradiction in the second case. Thus, if c' is adjacent to i_{k+1}, it is also adjacent to i_k. This implies that $N(i_{k+1}) \subseteq N(i_k)$, so we have $N(i_1) \supseteq N(i_2) \supseteq ... \supseteq N(i_k) \supseteq N(i_{k+1})$. The proof by induction is now complete. □

Corollary 4.9. *For each i and j, there is an edge-extremal graph G in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$ that has a universal node.*

Proof. Let G have partition (C, I) and let $i_1, i_2, ..., i_{|I|}$ be the nodes of I. Also, let $\mathcal{U} = N(i_1) \cap N(i_2) \cap ... \cap N(i_{|I|})$. From Lemma 4.8 we can assume that G has an inclusion ordering. This implies that $\mathcal{U} \neq \emptyset$. Every node $c \in \mathcal{U}$ is adjacent to all of I. Also, since $c \in C$, it is also adjacent to all of $C \setminus \{c\}$. The nodes of $\mathcal{U}$ are thus universal and the proof is complete. □

Corollary 4.10. *For each i and j, there is an edge-extremal graph G in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$ such that $|V(G)| \leq i$.*

Proof. From Corollary 4.9 we can assume that G has a universal node. But then $|V(G)| \leq i$, or else the bound on the maximum degree, $\Delta(G) < i$, would be violated. □

Lemma 4.11. *Let G be an edge-extremal graph in $\mathcal{M}_{\mathcal{SPLIT}}(i,j)$ with partition (C,I). If one node in I is locally universal, then all nodes in I are.*

Proof. The proof is by contradiction. Assume that G is an edge-extremal graph in $\mathcal{M}_{\mathcal{SPLIT}}(i,j)$ with partition (C,I) where not all nodes in I are universally local. We can assume that G has an inclusion ordering, so $N(i_1) \supseteq N(i_2) \supseteq ... \supseteq N(i_{|I|})$ for nodes $i_1, i_2, ..., i_{|I|} \in I$. i_1 is universally local, so assume for contradiction that this does not imply that all nodes are so. Also, assume that k is the least subindex such that i_k is not universally local. Recall from Lemma 4.3 that given a split graph G, $\nu(G)$ is decided by $\nu(BIP(G))$. In particular, we have from Lemma 4.1 that $\nu(BIP(G)) = |I| - \mathit{def}(BIP(G))$. Let $X \subseteq I$ be a set for which $|X| - |N(X)|$ is maximized, that is, $\mathit{def}(BIP(G)) = |X| - |N(X)|$. Then X can not contain a universally local node, since removing a node different from it from X would decrease $|X| - |N(X)|$. This implies that all locally universally nodes can be made adjacent to all nodes of C, thus becoming a part of C. Since no locally universal node is part of X, this would not affect $\nu(BIP(G))$ and $\nu(G)$. Also, from Corollary 4.10 we might assume that $|V(G)| \leq i$, so that adding edges to G never violate $\Delta(G) < i$. This is valid even if G is not edge-extremal, because the proof of Corollary 4.10 does not rely on on G having this property. But the fact that we can add edges to G while keeping it in $\mathcal{M}_{\mathcal{SPLIT}}(i,j)$ contradicts the assumption that G is edge-extremal. This completes the proof. □

Lemma 4.12. *Let G be an edge-extremal graph in $\mathcal{M}_{\mathcal{SPLIT}}(i,j)$ with partition (C,I). Either I is empty or $|C| < |I|$.*

Proof. The proof is by contradiction. So assume that I is non-empty and $|C| \geq |I|$. From Lemma 4.8 we can assume that G has an inclusion ordering, which implies that it has a locally universal node. Lemma 4.11 then tells us that all nodes of I are locally universal. Since $|C| \geq |I|$ there is a maximum matching that saturates I, and G has a perfect or near-perfect matching. Since $|V(G)| \leq i$, from Lemma 4.10, we can obtain a graph in $\mathcal{M}_{\mathcal{SPLIT}}(i,j)$ with more edges in G by making y universal, for some $y \in I$. This contradicts the assumption that G is edge-extremal and completes the proof. □

Let G be an edge-extremal graph in $\mathcal{M}_{\mathcal{SPLIT}}(i,j)$ with partition (C,I). The lemmas presented in this section tell us that we can assume that all possible edges between I and C are present. So G is uniquely decided by the sizes of C and I. We do not yet know these. We can immediately decide G in one particalar situation; if $i \leq 2j+1$, we can form a complete graph of size i. Since we know that $|V(G)| \leq i$, this has the maximum number of edges possible. If $i > 2j+1$, it is not clear that a complete graph is edge-extremal. The lemmas in this section have provided us with enough information to attack the problem analytically.

Let $\mathcal{N}$ be a maximum matcing in G. We can make a distinction between two types of edges in $\mathcal{N}$: edges with two endpoints in C and with one endpoint in C and one in I. Call these types of edges C-edges and I-edges, respectively. Also, let C_E and I_E be the set of C- and I-edges in $\mathcal{N}$, respectively. Consistent with earlier notation, we denote $|I_E|$ by $\nu(BIP(G))$. We know that I is either empty or $|C| < |I|$. In any case, $\mathcal{N}$ saturates the nodes in $V(BIP(G))$ that belongs to C. Thus, the number of edges with an endpoint in $C \cap V(BIP(G))$ can not exceed $\Delta(G)\nu(BIP(G))$. The rest of the edges are between nodes in C that are not in $BIP(G)$. These nodes induce a clique in G, so there is no more than $2(\nu - \nu(BIP(G))) + 1$ of them. This means that the rest of the edges in G does not exceed $\frac{(2(\nu-\nu(BIP(G)))+1)2(\nu-\nu(BIP(G)))}{2}$. Thus, we have

$$|E(G)| \leq \Delta(G)\nu(BIP(G)) + \frac{(2(\nu - \nu(BIP(G))) + 1)2(\nu - \nu(BIP(G)))}{2}.$$

This bound can be made tighter. Edges between nodes in $V(BIP(G)) \cap C$ are counted twice in the term $\Delta(G)\nu(BIP(G))$. Because $|C| < |I|$, the nodes of $N(I)$ are saturated by the maximum matching and $|V(BIP(G)) \cap C| = \nu(BIP(G))$. The number of edges counted twice is therefore $\frac{\nu(BIP(G))(\nu(BIP(G))-1)}{2}$. Subtracting this term from the bound on $|E(G)|$ gives

$$\begin{aligned}|E(G)| \leq &\Delta(G)\nu(BIP(G)) - \frac{\nu(BIP(G))(\nu(BIP(G)) - 1)}{2} \\ &+ \frac{(2(\nu - \nu(BIP(G))) + 1)2(\nu - \nu(BIP(G)))}{2}.\end{aligned}$$

We are interested in how this upper bound varies as $\nu(BIP(G))$ varies, and differentiate the expression with respect to $\nu(BIP(G))$. This gives

$$\frac{\partial E}{\partial \nu(BIP(G))} = \Delta(G) + 3\nu(BIP(G)) - 4\nu - \frac{1}{2}. \tag{4.2}$$

To identify possible extremal points, top- or bottom-points, we set the expression for $\frac{\partial E}{\partial \nu(BIP(G))}$ to 0 and solve for $\nu(BIP(G))$:

$$\Delta(G) + 3\nu(BIP(G)) - 4\nu - \frac{1}{2} = 0 \Rightarrow \nu(BIP(G)) = \frac{1}{3}(4\nu - \Delta(G) + \frac{1}{2}).$$

We know that $\Delta > 2\nu + 1$, which implies that in this extremal point, we have $\nu(BIP(G)) < \frac{2\nu}{6} - \frac{1}{6}$. Since $\frac{\partial^2 E}{\partial \nu(BIP(G))^2} = 3 > 0$, this extremal point is a bottom-point. This means that the candidates for top-points are the end-points of the domain of $|E(G)|$, which are $\nu(BIP(G)) = 0$ or $\nu(BIP(G)) = \nu$. If $\nu(BIP(G)) = 0$, then $|E(G)| \leq \frac{(2\nu+1)(2\nu)}{2}$. This bound is tight, as it is realized by a complete graph with $2\nu + 1$ nodes. Note that such a graph will never violate $\Delta(G) < i$, since we have $\Delta(G) > 2\nu + 1$ in this case. If $\nu(BIP(G)) = \nu$, we have $|E(G)| \leq \Delta(G)\nu(BIP(G)) - \frac{\nu(BIP(G))(\nu(BIP(G))-1)}{2}$. This bound is also tight, as it is realized by a full split graph with $N(I) = C$, $|C| = j - 1$ and $|V(G)| = i$.

The following theorem summarizes these results.

Theorem 4.13. *Let G be an edge-extremal graph in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$. We have two different cases:*

1. *If $i - 1 \leq 2(j-1) + 1$, then $G = K_i$ and $|E(G)| = \frac{i(i-1)}{2}$.*
2. *If $i - 1 > 2(j-1) + 1$, then $|E(G)| = \max\{\frac{(2(j-1)+1)(2(j-1))}{2}, (i-1)(j-1) - \frac{(j-1)((j-1)-1)}{2}\}$. We create a subcases, depending on which term is largest:*
   - *(a) When the first term is largest, $G = K_{2(j-1)+1}$.*
   - *(b) When the second term is largest, G is a full split graph with $N(I) = C$, $|C| = j - 1$ and $|V(G)| = i$.*

Let us conclude the section with an example. Consider the case $i = 12$ and $j = 5$. Here, $i-1 > 2(j-1)+1$, so we are in case 2 of Theorem 4.13. We have $\frac{(2(j-1)+1)(2(j-1))}{2} = \frac{(2(5-1)+1)(2(5-1))}{2} = 36$. Also, $(i-1)(j-1) - \frac{(j-1)((j-1)-1)}{2} = (12-1)(5-1) - \frac{(5-1)((5-1)-1)}{2} = 38$, so we are in case 2a. The candidates for edge extremal graphs in this case are shown in Figure 4.4, with the edge-extremal graph to the right.

If we decrease i and let $i = 11$ and $j = 5$, the situation changes. We are in case 2b, since the first term in Theorem 4.13 is largest; $|E(G)| = \max\{36, 34\} =$

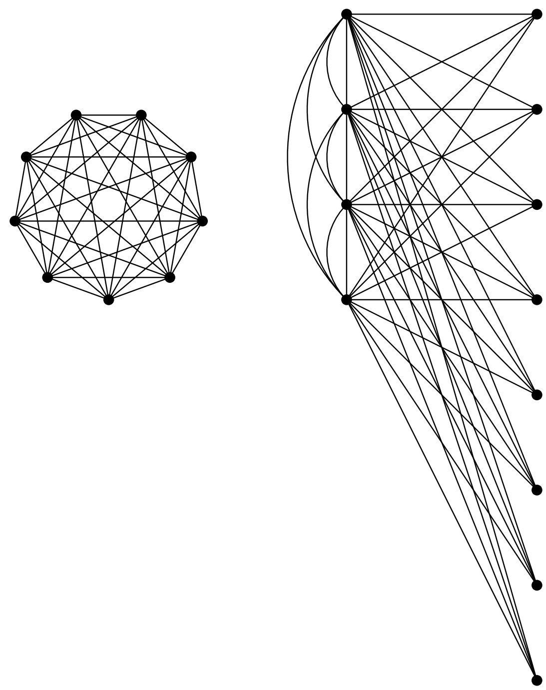

Figure 4.4: Candidates for edge-extremal graph in $\mathcal{M}_{\mathcal{SPLIT}}(12, 5)$

36. The candidates are shown in Figure 4.5, with the edge-extremal graph to the right.

If we decrease i further, such that $i = 10$ and $j = 5$, we reach case 1 of Theorem 4.13. The edge-extremal graph is thus K_9 in this case. In this example, we see that a larger i, relative to $2(j-1)+1$ favorizes the complete split graph being edge-extremal. This is intuitive, since the complete graph would not be exploiting the large matching number; the bound on i will hinder it doing so.

4.2 Disjoint union of split graphs

Recall from Chapter 2 that for general graphs an edge-extremal instance is a disjoint union of factor-critical components and stars. The proof of this relies on adding stars to the graph when some connected component is not factor-critical. This is not allowed to do for graphs in $\mathcal{M}_{\mathcal{SPLIT}}(i, j)$, since they are connected. But what if we remove the requirement of being connected by

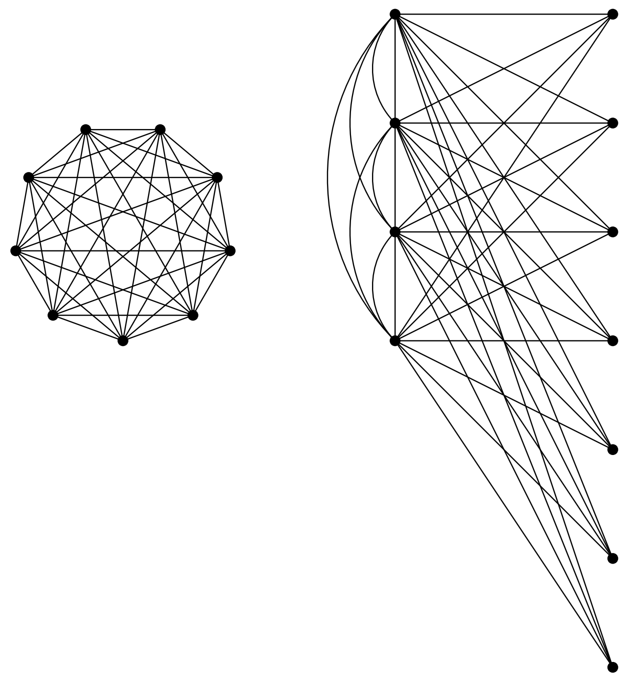

Figure 4.5: Candidates for edge-extremal graph in $\mathcal{M}_{\mathcal{SPLIT}}(11, 5)$

considering a disjoint union of split graphs? What would the edge-extremal instances look like in this case? This question is answered in this section.

Let $\mathcal{DSPLIT}$ be the class of disjoint union of split graphs. For general graphs and odd i, an edge-extremal instances is a disjoint union of complete graphs and stars- This is a disjoint union of split graphs, so the edge-extremal instances for odd i in $\mathcal{M}_{\mathcal{DSPLIT}}(i, j)$ is covered for. In this section we will solve the edge-extremal problem when i is even.

We start with two lemmas that give some structure to the edge-extremal graphs in $\mathcal{M}_{\mathcal{DSPLIT}}(i, j)$.

Lemma 4.14. *There is an edge-extremal graph $G \in \mathcal{M}_{\mathcal{DSPLIT}}(i, j)$ where every factor-critical connected component is complete.*

Proof. Let H be a factor-critical connected component of G with partition (I, C) where I is maximal. We will prove the theorem by giving two operations on H that preserves $|E(G)|$, $\Delta(G) < i$, $\nu(G) < j$ but decreases $|I|$. $|I|$ is equal to 1 if and only if H is complete. So by performing these operations enough times, we can transform G to an instance where every factor-critical component is complete.

If H is complete, there is nothing to prove. If not, let u and v be two nodes of I. Consider first the case $N(u) \cap N(v) = \emptyset$. Since nodes in I are mutually non-adjacent, and the size of C can not exceed $i-1$ without violating $\Delta(G) < i$, we must have $deg(u) + deg(v) < i$. So we can delete u and v and replace them with $K_{1,i-1}$, without decreasing $|E(H)|$. It is clear that $\Delta(G) < i$ is preserved in doing is. Also, H is factor-critical, so removal of u and v makes $\nu(H)$ decrease by 1, which is balanced by adding $K_{1,i-1}$. Thus, the operation does not decrease $|E(G)|$, preserves membership in $\mathcal{M}_{\mathcal{DSPLIT}}(i,j)$ and decreases $|I|$. Note that it does not depend upon H being factor-critical. If it is not, there is a node we can remove to make $\nu(H)$ decrease and then add $K_{1,i-1}$. We can do this until it becomes either factor-critical or a single node, which is complete. This is illustrated in Figure 4.6.

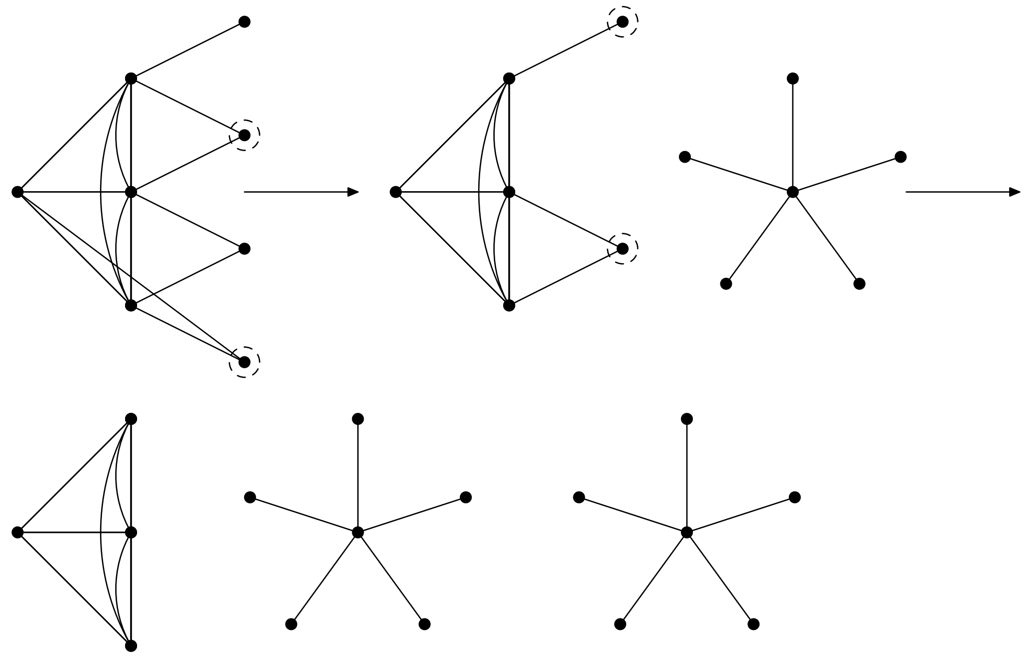

Figure 4.6: We remove circled nodes

Second, consider the case $N(u) \cap N(v) \neq \emptyset$. For every node $w \in C \setminus N(u)$, add the edge (u, w) to H. Should the bound on maximum degree be violated, remove edge (w, z), for some $z \in I \setminus \{u\}$. Such a node z will always exist if the bound is violated. Note that as long as $I \neq \emptyset$, we have $|C| - 1 < i - 1$ or

$|C| < i$. If this was not the case, a node in C adjacent to a node in I would violate $\Delta(G) < i$. So for w to violate $\Delta(G) < i$ it has to be adjacent to some $z \in I \setminus \{u\}$, which implies the existence of an edge that we can remove to keep $deg(w) < i$. The component H is factor-critical, so moving edges does not increase $\nu(H)$. When this operation is performed, the indpendent set and clique become $I \setminus \{u\}$ and $U \cup \{u\}$, respectively. So it decreases $|I|$ but not $|E(H)|$ while conserving membership in $\mathcal{M}_{\mathcal{DSPLIT}}(i,j)$. This process is illustrated in Figure 4.7.

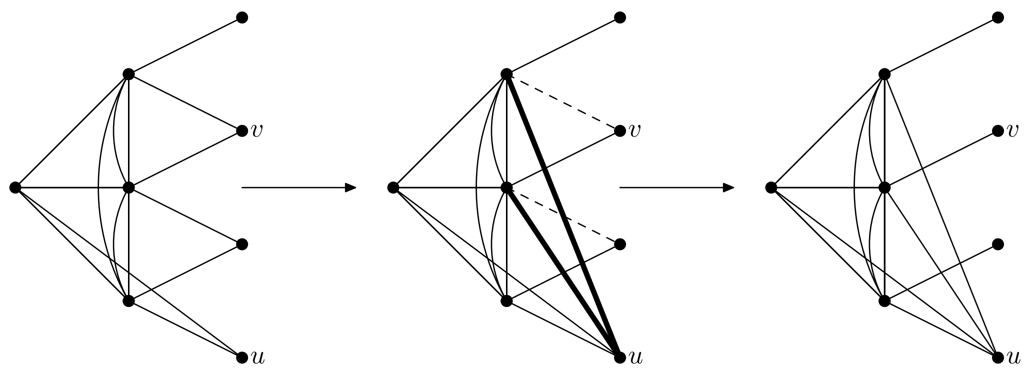

Figure 4.7: The dashed edges are removed and bold edges are added

We can always perform these operations on the components of an edge-extremal graph G to make them complete. Since $|E(G)|$ is not decreased in the process, there will always exists an edge-extremal graph where all factor-critical connected components are complete. This completes the proof. $\square$

Lemma 4.15. *Let $G \in \mathcal{M}_{\mathcal{DSPLIT}}(i,j)$ be a disjoint union of complete graphs. Then $|E(G)| \leq (i-1)(j-1)$.*

Proof. Denote the connected components of G by $H_1, H_2, ..., H_\ell$. Write G_k for $\uplus_{l=1}^{k} H_l$. The proof is by induction on the number of connected components of G, ℓ. For some complete component H_k, we have $|V(H_k)| \leq \Delta(H_k)$ and $|V(H_k)| \leq 2\nu(H_k)+1$. Using this for $\ell = 1$, we get

$$\begin{aligned} |E(G_1)| &\leq \frac{|H_1|(|H_1|-1)}{2} \leq \frac{\Delta(H_1)(2\nu(H_1)+1-1)}{2} \\ &\leq \Delta(H_1)\nu(H_1). \end{aligned}$$

For the inductive case, assume that the lemma is true for $\ell = \ell'$. For $\ell = \ell'+1$, without loss of generality, assume that $\Delta(G_{\ell'}) \geq \Delta(H_\ell)$. Note that this implies $\Delta(G_{\ell'}) = \Delta(G_\ell)$. We also have that $\nu(G_{\ell'}) + \nu(H_\ell) = \nu(G_\ell)$. Using this gives us the following chain of inequalities:

$$\begin{aligned}|E(G_\ell)| &\leq |E(G_{\ell'})| + |E(H_\ell)| \leq \Delta(G_{\ell'})\nu(G_{\ell'}) + \Delta(H_\ell)\nu(H_\ell) \\ &\leq \Delta(G_{\ell'})(\nu(G_{\ell'}) + \nu(H_\ell)) \leq \Delta(G_\ell)(\nu(G_{\ell'}) + \nu(H_\ell)) \\ &\leq \Delta(G_\ell)\nu(G_\ell)\end{aligned}$$

We now know that $|E(G_\ell)| \leq \Delta(G_\ell)\nu(G_\ell)$ for any ℓ. This implies that $G \leq \Delta(G)\nu(G) \leq (i-1)(j-1)$, and the proof is complete. □

We are now ready to bound the number of edges of a graph in $\mathcal{M}_{\mathcal{DSPLIT}}(i, j)$.

Theorem 4.16. *Let G be a graph in $\mathcal{M}_{\mathcal{DSPLIT}}(i, j)$. Then we have $|E(G)| \leq (i-1)(j-1)$. This bound is tight.*

Proof. We know that G is a disjoint union of factor-critical components and i-stars. From Lemma 4.14, we know that there exists an edge-extremal graph where every factor-critical component is complete. Assume that G is edge-extremal and on this form. Let $\mathcal{FC}$ and $\mathcal{S}$ be the graph consisting of the factor-critical components and i-stars of G, respectively. We know from Lemma 4.15 that $|E(\mathcal{FC})| \leq \Delta(\mathcal{FC})\nu(\mathcal{FC})$. It is clear that $|E(\mathcal{S})| \leq \Delta(\mathcal{S})\nu(\mathcal{S})$. Assume, without loss of generality, that $\Delta(\mathcal{S}) \geq \Delta(\mathcal{FC})$. Using that $\nu(\mathcal{FC}) + \nu(\mathcal{S}) \leq \nu(G) \leq j-1$, $\Delta(\mathcal{FC}) \leq i-1$ and $\Delta(\mathcal{S}) \leq i-1$ we get the following chain of inequalities:

$$\begin{aligned}|E(G)| &\leq |E(\mathcal{FC})| + |E(\mathcal{S})| \leq \Delta(\mathcal{FC})\nu(\mathcal{FC}) + \Delta(\mathcal{S})\nu(\mathcal{S}) \\ &\leq \Delta(\mathcal{S})(\nu(\mathcal{FC}) + \nu(\mathcal{S})) \leq \Delta(G)(\nu(\mathcal{FC}) + \nu(\mathcal{S})) \\ &\leq \Delta(G)\nu(G) \leq (i-1)(j-1).\end{aligned}$$

To see that the bound is tight, consider the graph where $\mathcal{FC} = \emptyset$. It consists of $j-1$ i-stars, and has $(i-1)(j-1)$ edges. This completes the proof. □

Chapter 5

Unit interval graphs

In this section we solve the edge-extremal problem on the class of unit interval graphs. All results presented here are worked out by us. When proving properties of unit interval graphs, we will often make use of the underlying interval representation. The intervals have a natural ordering, and it is often useful to reference a node which is represented by a certain interval in this ordering. To make this clear, we use the *unit interval ordering*, defined next.

Definition 5.1. *Let G be a unit interval graph, and I its interval representation. Let $(i_1, i_2, ..., i_n)$ be an ordering of I, sorted in non-decreasing order on the left endpoint of the intervals. Let $(v_1, v_2, ..., v_n)$ be an ordering of the nodes in G, such that v_j corresponds to i_j. This is called a* unit interval ordering *of G.*

Figure 5.1 illustrates this definition.

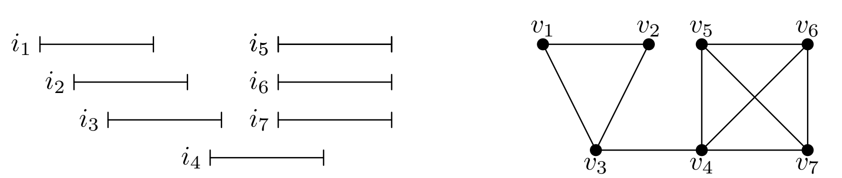

Figure 5.1: A unit interval graph with its corresponding unit interval ordering

Let G be a unit interval graph and $(v_1, v_2, ..., v_n)$ the unit interval ordering of its nodes. A well known observation that we will make use of, is the following: Let v_k and v_l be nodes in a unit interval graph, with $k < l$. If there is an edge between these nodes, i_k and i_l intersect. All i_m with

$k < m < l$, will lie between i_k and i_l in the unit interval ordering. Thus, all v_m are adjacent to v_l. This is illustrated in figure 5.2.

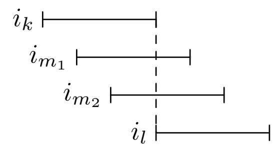

Figure 5.2: i_{m_1} and i_{m_2} intersect i_l

We start the solving of the edge-extremal problem on unit interval graph with a useful lemma:

Lemma 5.2. *Every connected unit interval graph has a Hamiltonian path.*

Proof. Let G be a unit interval graph, with unit interval ordering $(v_1, v_2, ..., v_n)$. We will show that $v_1v_2...v_n$ is a Hamiltonian path in G.

Every node v_i, $1 \leq i \leq n-1$ has an edge to v_{i+1}. If not, v_i could not have any edges to v_j, $j > i$, but this would contradict the fact that G is connected. So we can always traverse the nodes in the order of the unit interval ordering without visiting any node twice. □

Figure 5.3 shows a Hamiltonian path in a unit interval graph.

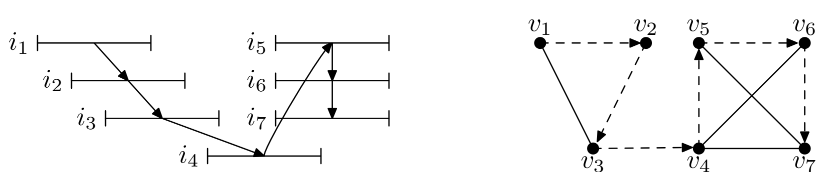

Figure 5.3: Constructing a Hamiltonian path from unit interval ordering

Corollary 5.3. *Let* $G = (V, E)$ *be a connected unit interval graph. Then* $\nu(G) = \lfloor \frac{|V|}{2} \rfloor$.

Proof. Let G have unit interval ordering $(v_1, v_2, ..., v_n)$. By Lemma 5.2, $v_1, v_2, ..., v_n$ is a Hamiltonian path in G. We can construct a matching by using every other edge of this path, i.e. $(v_1, v_2), (v_3, v_4), ...$ If $|V|$ is even, the matching is perfect and has size $\lfloor \frac{|V|}{2} \rfloor$. If $|V|$ is odd, we have a near-perfect matching, with size $\frac{|V|-1}{2} = \lfloor \frac{|V|}{2} \rfloor$. The corollary follows. □

Through the rest of the section, let $\mathcal{UNIT}$ denote the class of unit interval graphs.

Corollary 5.4. *Every edge-extremal graph in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$ with at most i nodes is complete.*

Proof. The proof is by contradiction. Let G be edge-extremal in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$ with $|V(G)| \leq i$ and assume that G is not complete. So there are two nodes $u, v \in V(G)$, with $(u,v) \notin E(G)$. Add (u,v) to G to obtain a new graph G'. We know by Corollary 5.3 that G has a perfect or near-perfect matching. Adding (u,v) to G can not increase $\nu(G)$, so $\nu(G') = \nu(G) < j$. Also, since G' has no more than i nodes, G' can not violate the bound $\Delta(G') < i$. This implies that G' is in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$ but $|E(G')| > |E(G)|$. We assumed that G was edge-extremal, so this is a contradiction, implying that G has to be complete. This completes the proof. $\square$

For a unit interval graph $G = (V,E)$, let $V_k(G)$ denote the first k nodes in the unit interval ordering of G.

Lemma 5.5. *There is an edge-extremal graph $G \in \mathcal{M}_{\mathcal{UNIT}}(i,j)$ where no connected component has more than i nodes.*

Proof. Assume for contradiction that G has a connected component H with more than i nodes. We will show that G can be transformed in a way that conserves membership in $G \in \mathcal{M}_{\mathcal{UNIT}}(i,j)$ without decreasing the number of edges. The transformation consists of two steps:

1. Disconnect $V_i(H)$ from H to obtain the connected components H_1' and H_2' on $V_i(H)$ and $V(H) \setminus V_i(H)$, respectively.

2. Add all possible edges to H_1', making it isomorphic to K_i.

Let G' be the result of the transformation; $G' \simeq (G \setminus V_i) \uplus K_i$. The transformation is exemplified in Figure 5.4. Clearly, G' does not violate $\Delta(G') < i$. Let us show that $\nu(G') < j$ We have $\nu(G) = \left\lfloor \frac{|V|}{2} \right\rfloor$ and $\nu(G') = \left\lfloor \frac{i}{2} \right\rfloor + \left\lfloor \frac{n-i}{2} \right\rfloor$. Since $n = i + (n-i)$, we have $\nu(G') \leq \nu(G) < j$ by Lemma 3.5. The result of the operation is the same as if we had removed V_i from G and disjointly added K_i. The class of unit interval graphs are closed under deletion of nodes, and K_i is unit interval. It follows that G' is unit interval and $G' \in \mathcal{M}_{\mathcal{UNIT}}(i,j)$.

It remains to show that $|E(G')| \geq |E(G)|$. Refer to edges between nodes in $V_i(G)$ as *internal*, and edges between $V_i(G)$ and $V(G) \setminus V_i(G)$ as *external*.

In step 1 of the transformation, all external edges are removed. We will show that the edges added in step 2 is at least as large as those removed in step 1. In G, let v_k be a node in $V_i(G)$ with an external edge, and say that v_k has d edges to nodes in $V_i(G)$. So there are $(i-1)-d$ nodes in $V_i(G)$ that does not have an edge to v_k. We call these nodes available, and their number equals the number of edges added in step 2 of the transformation. We must therefore show that the number of available nodes is greater or equal to the number of external edges adjacent to v_k. There are $deg(v_k)-d$ external edges adjacent to v_k. The desired inequality is thus $(i-1)-d \geq deg(v_k)-d$. This is easily obtained by adding $-d$ to both sides of $deg(v_k) \leq i-1$. This is valid for all $v_k \in V_i$, so $|E(G')| \geq |E(G)|$. This completes the proof.

□

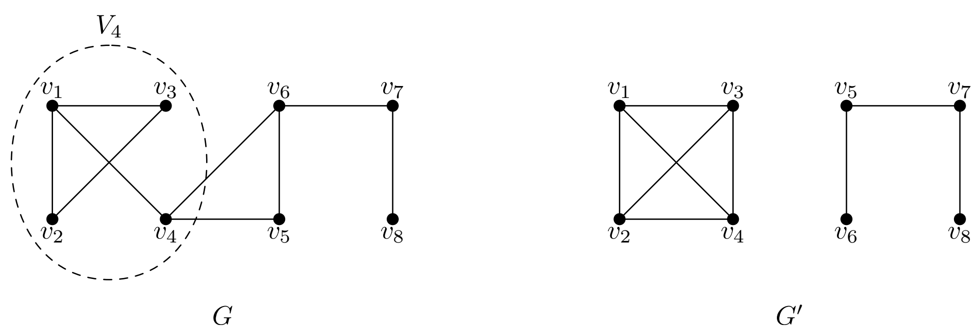

Figure 5.4: Illustrating the transformation for $G \in \mathcal{M}_{\mathcal{UNIT}}(4,5)$

Lemma 5.6. *No edge-extremal graph in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$ have more than one connected component with strictly less than $i-1$ nodes.*

Proof. Let G be an edge-extremal graph in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$. Assume for contradiction that G has two or more connected components with strictly less than $i-1$ nodes. Let us call two of these connected components H_1 and H_2. All nodes in H_1 or H_2 have strictly less than $i-2$ neighbours, so $\Delta(H_1) < i-2$ and $\Delta(H_2) < i-2$. By Corollary 5.4 we know that H_1 and H_2 must be complete. Assume, without loss of generality, that $\Delta(H_1) \geq \Delta(H_2)$ and perform the following operation: Add two universal nodes to H_1 and remove two nodes from H_2. This yields two new complete connected components H_1' and H_2' of a new graph G'. The operation is exemplified in Figure 5.5. We will show that G' is in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$ and that $|E(G')| > |E(G)|$, contradicting the edge-extremality of G. The operation does not make $\Delta(H_1')$ increase with more than 2, and $\Delta(H_2')$ does not increase. So $\Delta(H_1') < (i-2)+2 = i$, $\Delta(H_2') < i$ and thus $\Delta(G') < i$. For the combined maximum matching sizes we have

$$\begin{aligned}
\nu(H_1') + \nu(H_2') &= \lfloor \frac{|V(H_1')|}{2} \rfloor + \lfloor \frac{|V(H_2')|}{2} \rfloor = \lfloor \frac{|V(H_1)|+2}{2} \rfloor + \lfloor \frac{|V(H_2)|-2}{2} \rfloor \\
&= \lfloor \frac{|V(H_1)|}{2} + 1 \rfloor + \lfloor \frac{|V(H_2)|}{2} - 1 \rfloor \\
&= \lfloor \frac{|V(H_1)|}{2} \rfloor + 1 + \lfloor \frac{|V(H_2)|}{2} \rfloor - 1 = \lfloor \frac{|V(H_1)|}{2} \rfloor + \lfloor \frac{|V(H_2)|}{2} \rfloor \\
&= \nu(H_1) + \nu(H_2).
\end{aligned}$$

So $\nu(G') = \nu(G) < j$. Also, since G' is clearly a unit interval graph, it is in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$.

Now we show that $|E(G')| > |E(G)|$. In adding the nodes to H_1 we gain $|V(H_1)| + (|V(H_1)|+1) = 2|V(H_1)|+1$ edges and we lose $|V(H_2)| + (|V(H_2)|-1) = 2|V(H_2)|-1$ edges when removing the nodes from H_2. Since $|V(H_1)| \geq |V(H_2)|$, we have that $2|V(H_1)|+1 > 2|V(H_2)|-1$.

So $G' \in \mathcal{M}_{\mathcal{UNIT}}(i,j)$ and $|E(G')| > |E(G)|$, which is the desired contradiction. Our assumption of G having more than two connected components with strictly less than $i-1$ nodes must be false. This completes the proof. $\square$

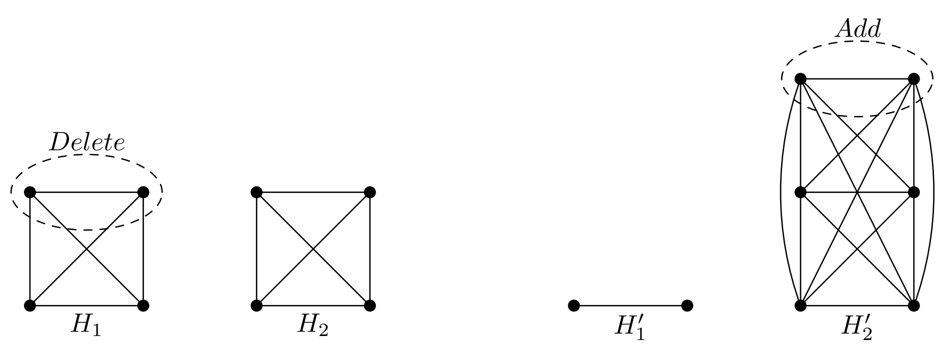

Figure 5.5: The operation described in Lemma 5.6. Edges lost is 5, edges gained is 9

Up to now, have not made the distinction between odd or even i in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$. But we are going to treat the two cases differently from now on. To determine the edge-extremal instances when i is odd, we need one more lemma:

Lemma 5.7. *Let i be odd and G an edge-extremal graph in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$. Then all connected components of G have an odd number of nodes.*

Proof. The proof is once more by contradiction. We know from previous lemmas that all connected components of G are complete components on no more than i nodes. Assume for contradiction that some connected component H of G has an even number of nodes, k. Since i is odd, we know that $k+1 \leq i$, so we can add a universal node to H without violating $\Delta(H) < i$. Call the result of this operation H'. We have $\nu(H') = \left\lfloor \frac{k+1}{2} \right\rfloor = \left\lfloor \frac{k}{2} \right\rfloor = \nu(H)$. But H' has more edges than H so G is not edge-extremal. This is a contradiction and the lemma follows. □

Now we have enough information to determine the maximum number of edges for odd i:

Theorem 5.8. *Let i be odd and G an edge-extremal graph in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$. Then G has $i(j-1) + (2r+1-i)r$ edges, where $r = (j-1) \bmod \left(\frac{i-1}{2}\right)$.*

Proof. From Lemma 5.5 we know that no component has more than i nodes, and Lemma 5.4 tells us that they are all complete. Lemma 5.7 says that all connected components of G have an odd number of nodes. Since $i-1$ is even, Lemma 5.6 implies that there is no more than one connected component with fewer than i nodes. This leaves us with just one possibility; H is a disjoint union of K_i's and possibly one smaller complete graph. Every connected component has matching size $\frac{i-1}{2}$. The smaller component exists when $j-1$ does not divide $\frac{i-1}{2}$, and we have to make an extra component to "exploit" all the matching edges. The matching size of this component is then $(j-1) \bmod \frac{i-1}{2}$. It can not be larger, as this would violate $\nu(G) < j$. Also, it can not be smaller, since then we could add nodes to it without violating $\nu(G) < j$ and $\Delta(G) < i$, obtaining a new graph with more edges.

We want to find a closed expression for the number of nodes in such a graph. There exist q and r such that

$$j - 1 = \left(\frac{i-1}{2}\right) q + r, \qquad 0 \leq r < \frac{i-1}{2} \tag{5.1}$$

From the above discussion, we see that q must be the number of K_i's and $r = (j-1) \bmod \left(\frac{i-1}{2}\right)$ is the matching number of the smaller component. This smaller component will then have $2r+1$ nodes, and so $\frac{(2r+1)2r}{2} = (2r+1)r$

edges. The K_i's have $\frac{i(i-1)}{2}$ edges. Therefore the number of edges of G is given by

$$|E(G)| = \left(\frac{i(i-1)}{2}\right) q + (2r+1)r = i\left(\frac{i-1}{2}q\right) + (2r+1)r. \tag{5.2}$$

Inserting $\frac{i-1}{2}q = (j-1) - r$, implied by Equation 5.1, we get

$$|E(G)| = i((j-1)-r) + (2r+1)r = i(j-1) + (2r+1-i)r \tag{5.3}$$

Note that from Equation 5.1 we have $r < \frac{i-1}{2} \Rightarrow 2r+1-i < 0$. So the second term is never larger than 0, and only 0 when $j-1$ divides $\frac{i-1}{2}$. This is the desired closed expression and the proof is complete. □

Let us consider an example. Look at an edge-extremal graph from the family $\mathcal{M}_{\mathcal{UNIT}}(5,7)$. Here we have $r = (j-1) \bmod \left(\frac{i-1}{2}\right) = 6 \bmod 2 = 0$, so $|E(G)| = i(j-1)+(2r+1-i)r = 5 \cdot 6 + 0 = 30$. G is a disjoint union of 3 K_i. If we increase j and let G be edge-extremal in $\mathcal{M}_{\mathcal{UNIT}}(5,8)$, $r = 7 \bmod 2 = 1$. In this case we have $|E(G)| = 5 \cdot (8-1) + (2 \cdot 1 + 1 - 5)1 = 33$. G in this case is shown in Figure 5.6.

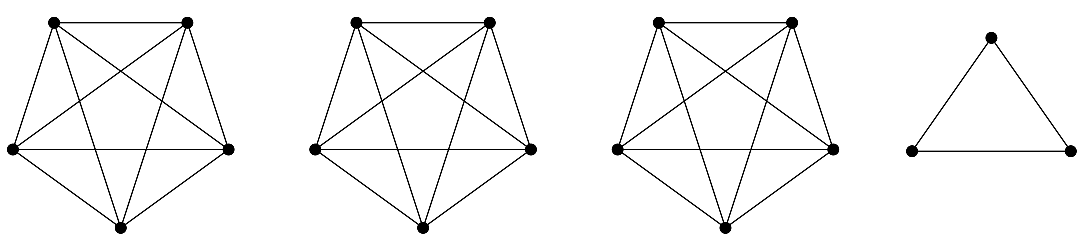

Figure 5.6: Edge-extremal instance in $\mathcal{M}_{\mathcal{UNIT}}(5,8)$

The uniqueness of the edge-extremal graph follows quite easily from the previous theorem.

Corollary 5.9. *The extremal unit interval graph for odd i is unique.*

Proof. From the proof of Theorem 5.1, we know that the extremal graph is uniquely decided by q and r in the expression $j = \left(\frac{i-1}{2}\right)q + r$, with $0 \le r < \frac{i-1}{2}$. These are unique by Lemma 3.6, and so is the solution. □

We now turn our attention to the case where i is even. In this case, we know that an edge-extremal graph consists of complete connected components on i and $i-1$ nodes and possibly one smaller complete component. For odd i, there were only two different components, as opposed to three in this case. This makes the edge-extremal problem for even i considerably harder than for odd i, as reflected in the solution. As we will see, we want to compose the collection of these components such that the smaller one is either as small or as large as possible. The following theorem makes this more precise and solves the case for even i:

Theorem 5.10. *Let i be even and G an edge-extremal graph in $\mathcal{M}_{\mathcal{UNIT}}(i,j)$. Also, let $a = \frac{i}{2}$ and $b = \frac{i-2}{2}$, $j-1 = q_1 a + r_1 = q_2 b + r_2$. Then G has $(i-1)(j-1) - \mathcal{R}$ edges, where $\mathcal{R} = \min\{(i-2(p+1))p, (i-2(s+1))s\}$, where $p = \max\{0, r_2 - q_2\}$ and $s = \min\{q_1 + r_2, b-1\}$.*

Proof. From previous lemmas, we know that G consists of complete connected components on i and $i-1$ nodes, possibly toghether with one smaller complete connected component. Let A and B be the number of connected components with $\frac{i}{2}$ and $\frac{i-2}{2}$ nodes, respectively. Also, let the r be the matching number of the smaller component. This component then has $\frac{(2r+1)2r}{2} = (2r+1)r$ edges. The number of edges of G can then be expressed by

$$|E(G)| = A\frac{i(i-1)}{2} + B\frac{(i-1)(i-2)}{2} + (2r+1)r. \tag{5.4}$$

Also, we can safely assume that $\nu(G) = j-1$ which implies $j-1 = A\frac{i}{2} + B\frac{i-2}{2} + r$. We write this euation as $(j-1) - r = A\frac{i}{2} + B\frac{i-2}{2}$ and insert into Equation 5.4:

$$\begin{aligned} E &= (i-1)\left(A\left(\frac{i}{2}\right) + B\left(\frac{i-2}{2}\right)\right) + (2r+1)r \\ &= (i-1)(j-1-r) + (2r+1)r = (i-1)(j-1) - (i-2(r+1))r. \end{aligned}$$

By varying A and B, we can change only the second term, $(i-2(r+1))r$. Since $2r+1 < i-1$ and thus $2r+2 < i$, this term will always be negative, so we want to chose A and B so to minimize this terms absolute value. To see how it behaves when r varies, we differentiate it with respect to r:

$$\frac{\partial}{\partial r}(i - (2(r+1))r) = i - 4r + 2. \tag{5.5}$$

Setting this equal to 0, to locate possible extremal points, we get $r = \frac{i-2}{4}$. Repeated differentiation reveals that this is in fact a top point; $\frac{\partial}{\partial r}(i-4r+2) = -4$. So the bottom points lie on the edge of the interval $[0, \frac{i-2}{2}]$. In other words, the term $(i-(2(r+1))r$ has its lowest value either when r is as small or as large as possible. What then are the smallest and largest values r can take? Note that the equation $j = A\frac{i}{2} + B\frac{i-2}{2} + r$ is the type of equation described in Section 3.3. This case was actually the sole motivation for writing that section. According to Lemma 3.7 and 3.8, the smallest and largest values r can take are $\max\{0, r_2 - q_2\}$ and $\min\{q_1 + r_2, b-1\}$, respectively. We choose the one that minimizes $(i-(2(r+1))r$. The theorem follows. $\square$

Consider an edge-extremal instance G from the graph family $\mathcal{M}_{\mathcal{UNIT}}(10,7)$. In this case we have

- $a = \frac{10}{2} = 5$, $b = \frac{10-2}{2} = 4$
- $j-1 = 6 = 1 \cdot 5 + 1 = 1 \cdot 4 + 2$, $q_1 = 1$, $r_1 = 1$, $q_2 = 1$, $r_2 = 2$
- $p = \max\{0, 2-1\} = 1$, $s = \min 3, 3 = 3$
- $\mathcal{R} = \min\{12, 6\} = 6$

This gives $|E(G)| = (i-1)(j-1) - 6 = 9 \cdot 7 - 6 = 48$. This is realized by a disjoint union of K_{10} and K_3.

Chapter 6

Factor-critical chordal graphs

Split graphs, disjoint union of split graphs and unit interval graphs are all subclasses of chordal graphs. The edge-extremal instances for odd i in the general case, are chordal. It would have been interesting to see how the edge-extremal chordal graphs for even i compared to these classes. Although we are not able to present such a solution, we conjecture that the edge-extremal bound for chordal graphs is the same as in disjoint union of split graphs.

Our effort on trying to find a solution on chordal graphs resulted in some interesting observations. In particular, we are able to give a characterization of factor-critical chordal graphs in terms of spanning subgraphs. Since we think this might provide insight in its own right, we present the result in this section.

We start with presenting an already known result from [6]. Let $G_1 = (V_1, E_1)$ and $G_2 = (V_1, E_1)$ be two graphs. The operation of adding the graphs together produces a new graph defined by $G_1 + G_2 = (V(G_1) \cup V(G_2), V(G_1) \cup V(G_2))$. The following lemma is presented as a theorem in [6]. We omit the proof here, it is given in [6].

Lemma 6.1. *Every factor-critical graph G can be represented as $P_0 + P_1 + \ldots + P_r$ where $P_0 = K_1$ and for each i, P_{i+1} is either (1) an odd path having only its two endpoints in common with $P_0 + P_1 + \ldots + P_i$ or (2) P_{i+1} is an odd cycle with precisely one node in common with $P_0 + P_1 + \ldots + P_i$.*

The two cases are illustrated in Figure 6.1.

The rest of the section consists of new results. We start with a recursive definition of a graph class we have chosen to call *triangle trees.*

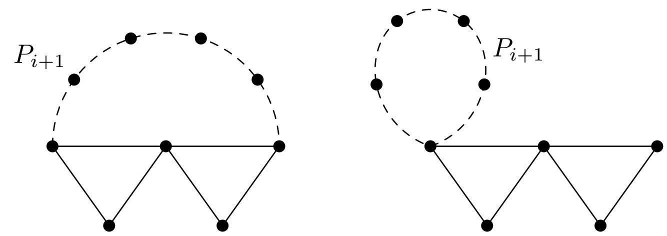

Figure 6.1: Case 1 to the left, case 2 to the right

Definition 6.2. *The empty graph and K_1 are triangle trees. Assume that T is a triangle tree. We can make a new triangle tree by adding two nodes, u, v, to T, together with edges (u, v), (u, w) and (v, w) for some $w \in V(T)$.*

An example of a sequence of triangle trees are shown in Figure 6.2. The following theorem presents the main result of the chapter.

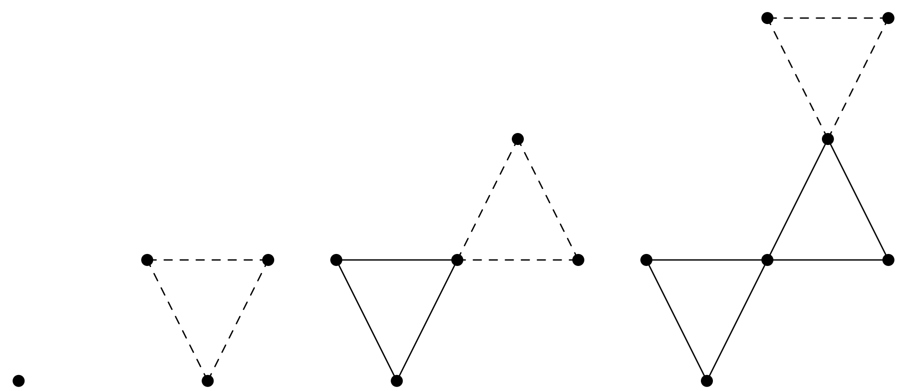

Figure 6.2: A sequence of triangle trees. Added edges in each step are dashed

Theorem 6.3. *Every chordal and factor-critical graph is spanned by a triangle tree.*

Proof. Let G be a chordal and factor-critical graph. Also, let $P_0 + P_1 + \ldots + P_r$ be a decomposition of G as described in Lemma 6.1 and i some non-negative integer with $i \leq r$. We then know from Lemma 6.1 that P_{i+1} is either (1) an odd path with two nodes in common with $P_0 + P_1 + \ldots + P_i$ or (2) an odd cycle with one node in common with $P_0 + P_1 + \ldots + P_i$. We will prove the theorem by showing that in each case P_{i+1} can be spanned by repeatedly adding triangles to the graph as in the recursive step in Definition 6.2. That is, we build each cycle by adding pairs of nodes u, v and edges (u, v), (u, w) and (v, w) for some w from the part of the graph already spanned by triangle trees. This process is illustrated in Figure 6.3. Since $P_0 = K_1$ is a triangle tree, we will then know that G can be spanned by triangle trees by spanning each individual P_{i+1}.

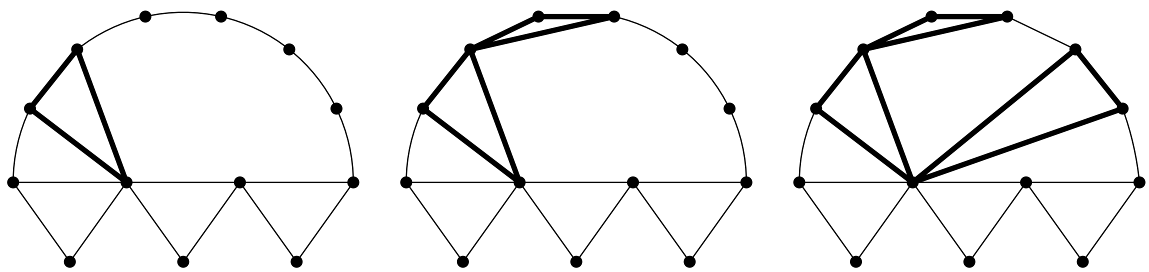

Figure 6.3: Spanning a cycle with triangle trees

Consider first case (2), the case where P_{i+1} is an odd path with exactly one point in common with $P_0 + P_1 + \ldots + P_i$. We will show that this case reduces to case (1). Let $P_{i+1} = u_1, u_2, u_3, ..., u_\ell$, where u_1 is in $P_0 + P_1 + ... + P_i$. Since G is chordal, it does not have a induced cycle of length 4 or more. Consider the path u_ℓ, u_1, u_2, u_3. This is a potential induced cycle of length 4 or more, if not at least one of edges (u_2, u_ℓ) or (u_1, u_3) are present. Assume first that (u_2, u_ℓ) is present. We can then span these nodes with a triangle by adding u_2 and u_ℓ). $u_2, u_3, .., u_\ell$ is then an odd cycle with two nodes in common with $P_0 + P_1 + ... + P_i$, which is case (1). If (u_1, u_3) is present, we add u_2 and u_3 and we again have case (1). Both these cases are illustrated in Figure 6.4.

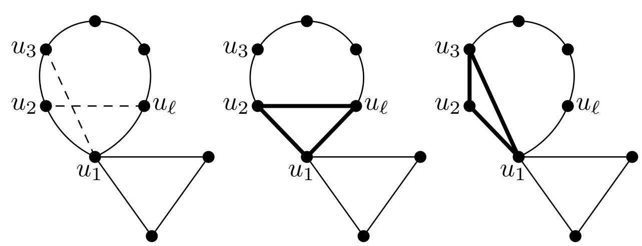

Figure 6.4: At least one of the dashed edges in the leftmost graph must be present. The added triangles are shown with bold edges

Now we show that the odd path in case (1) can be spanned by a triangle tree. We do so by induction on the odd length of the path, $\ell - 1$. Again, let $P_{i+1} = u_1, u_2, u_3, ..., u_\ell$ and u_1 and u_ℓ the two nodes P_{i+1} has in common with $P_0 + P_1 + ... + P_i$. The base case is a path of length 1. We can then add an edge between u_1 and u_2, and the path is spanned by the empty graph, which is a triangle tree.

For the inductive case, we show that if a path of length $k-2$, with $k \geq 3$ can be spanned by triangles, then so can a path of length k. We will do so by showing that there has to be a node $w \in V(P_0 + P_1 + ... + P_i)$ that is adjacent to u_2 and u_3 or u_2 and $u_{\ell-1}$. In the latter case, the edge $(u_2, u_{\ell-1})$ is also present. Either way, there is a triangle we can add to $P_0 + P_1 + ... + P_i$. We are then left with a cycle of length $k-2$, which we know can be spanned by triangles. So assume that $u_\ell, u_{\ell+1}, u_{\ell+2}, ..., u_{\ell+\ell'}, u_1$ is the shortest path from u_ℓ to u_1 in $P_0 + P_1 + ... + P_i$. Look at the path $u_{\ell+\ell'}, u_1, u_2, u_3$. Since we do not have induced cycles of length 4 or more, at least one of the edges (u_1, u_3) or $(u_2, u_{\ell+\ell'})$ must be present. If (u_1, u_3) is present, then we have can add u_2 and u_3 to $P_0 + P_1 + ... + P_i$ to obtain a shorter cycle. If not, $(u_2, u_{\ell+\ell'})$ has to be present. In this case, consider the path $u_{\ell+\ell'-1}, u_{\ell+\ell'}, u_2, u_3$. In the same manner as before, at least one of the edges $(u_2, u_{\ell+\ell'-1})$ or $(u_3, u_{\ell+\ell'})$ is present. If $(u_3, u_{\ell+\ell'})$ is present, we are done. If not, then $(u_2, u_{\ell+\ell'-1})$ has to be present. In general, consider the path $u_{\ell+\ell'-x}, u_{\ell+\ell'-(x-1)}, u_2, u_3$ for increasing x. If edge $(u_3, u_{\ell+\ell'-(x-1)})$ is present, there is a triangle. If not, $(u_2, u_{\ell+\ell'-x})$ has to be present, and so forth, until $\ell + \ell' - x$ equals $\ell - 1$. Then at least one of $(u_2, u_{\ell-1})$ or (u_3, u_ℓ) has to be present. If we have not encountered a triangle for smaller x, then (u_2, u_ℓ) has to be present. If $(u_2, u_{\ell-1})$ is present, the nodes u_2, u_ℓ and $u_{\ell-1}$ constitutes our triangle. Or if (u_3, u_ℓ) is present, the nodes u_2, u_3 and u_ℓ constitutes our triangle. In either case, we obtain a cycle of length $k-2$. This process is illustrated in Figure 6.5. The proof is complete. □

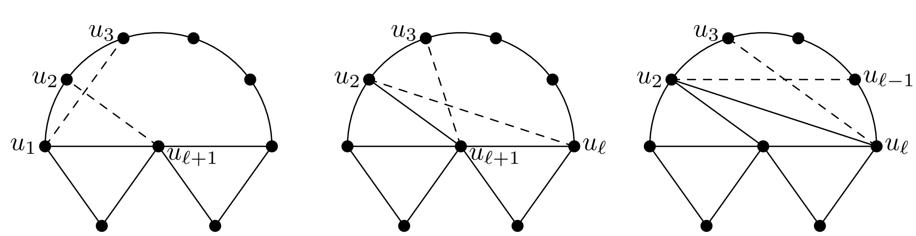

Figure 6.5: Illustrating the proof of Theorem 6.3. At least one of the dashed edges must be present. In the rightmost case, we have a triangle either way

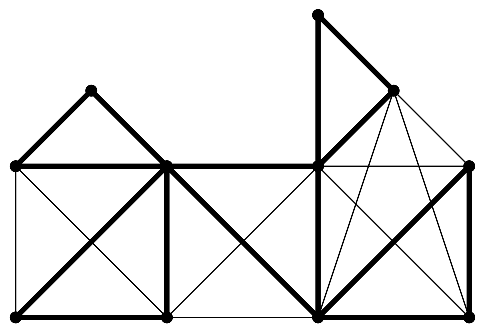

Figure 6.6: A factor-critical chordal graph with a possible spanning triangle tree with bolded edges.

Figure 6.6 shows a chordal factor-critical graph with a possible spanning triangle tree.

If a graph is spanned by a triangle tree, then it is clearly factor-critical but not necassarily chordal. It is not hard to add an edge to a triangle tree which creates an induced cycle of length 4 or more.

The edge-extremal instances for chordal graphs consists of factor-critical components and i-stars. We might imagine that Theorem 6.3 may help us in identifying these factor-critical components or at least bound their number of edges. One possible plan is to build every component by adding edges to triangle trees. The problem we encountered when trying this, was that it is hard to formalize the adding of edges such that we can bound their number.

Chapter 7

Conclusive remarks

This section will summarize the results of the thesis, give some general remarks about the work done and provide some open questions for future work.

7.1 Summary

The thesis have looked at edge-extremal graphs with bounded degree and matching number on specific graph classes. In particular, we let $\mathcal{S}$ be a graph class and $\mathcal{M}_{\mathcal{S}}(i, j)$ all graphs G from $\mathcal{S}$ which satisfy $\Delta(G) < i$ and $\nu(G) < j$. We have asked: what is the maximum number of edges a graph in $\mathcal{M}_{\mathcal{S}}(i, j)$ can achieve, for given $\mathcal{S}$, i and j. Such graphs are called edge-extremal in $\mathcal{M}_{\mathcal{S}}(i, j)$. This is equivalent to asking for the Ramsey number of line graphs for graphs in $\mathcal{S}$. We have answered this question for bipartite graphs, split graphs, disjoint union of split graphs and unit interval graphs. The solution on general graphs has been presented, but this is not our work. In addition we have presented a characterization of factor-critical chordal graphs.

The results for edge-extremal graphs are summarized in Table 7.1.

Table 7.1: Summary of results

Graph class	Condition	Tight bound on edges
Bipartite	-	$E \leq (i-1)(j-1)$
Split	$i-1 \leq 2(j-1)+1$	$E \leq \frac{i(i-1)}{2}$
	$i-1 > 2(j-1)+1$	$E \leq \max\{\frac{(2(j-1)+1)(2(j-1))}{2}, (i-1)(j-1) - \frac{(j-1)((j-1)-1)}{2}\}$
Disjoint union of split	Even i	$E \leq (i-1)(j-1)$
Unit interval	Odd i	$E \leq i(j-1) + (2r+1-i)r$, $r = (j-1) \bmod \left(\frac{i-1}{2}\right)$
	Even i	$E \leq (i-1)(j-1) - \mathcal{R}$, $\mathcal{R} = \min\{(i-2(p+1))p, (i-2(s+1))s\}$, $a = \frac{i}{2}$ and $b = \frac{i-2}{2}$, $j-1 = q_1a + r_1 = q_2b + r_2$ $p = \max\{0, r_2 - q_2\}$ and $s = \min\{q_1 + r_2, b-1\}$

We have only included results worked out by us in the table. The product $(i-1)(j-1)$ is present in many of the tight bounds on edges, possibly together with another term which contributes negatively. This other term seems like a measure of how many edges a graph from that class can have, under restrictions on maximum degree and matching number. This is somewhat loose, but we can still point at one particular tendency; the negatively contributing term is larger for the graph classes that do not allow i-stars and absent for those who do. i-stars seems to play a central role in the bound on edges. We have therefore posed some open questions later in the chapter, concerning how allowing and disallowing stars affects the bound on the number of edges.

7.2 Comment on working with the thesis

The solution of the edge-extremal problem on bipartite graph was used as a familirization to the problem. This solution was produced immediately. When this warm-up was done, our main goal was to solve the edge-extremal problem on chordal graphs. We wanted to begin with subclasses of chordal grahps, starting with interval graphs and gradually move up in the graph hierarchy. Interval graphs did not yield any solution, so we moved futher down in the hierarchy to unit interval graphs. After spending some time with this class, we came up with the solution presented here. However, we did not feel that this provided any more insight to the problem on interval graphs, so we jumped straight up to chordal graphs again. Attacking this problem from any angle we could come up with did not solve it. Along the way, however, we did some interesting observations, one of which is presented in Chapter 6, concerning factor-critical chordal graphs. We have chosen not to present any other insight obtained on the work on chordal graphs, as we did not feel it was solid enough. This led us to start with subclasses of chordal graphs, and we thought that the problem was easy on split grahps. It did turn out to be way more complex than expected. A natural extension of this work was disjoint union of split graphs, relaxing the connectedness of split graphs. Again, the problem on this graph class did not help with the chordal graphs. At that point, we were very pleased with the solutions obtained, and chose to give up chordal graphs. We do however have strong reasons to believe that the solution on chordal graphs is the same as the one on disjoint union of split graphs.

For the most part, a top-down approach has been used; we started with conjecturing properties that the edge-extremal instances might have, trying to prove this. It the result were positive, we inestigated if these properties could be used to anything reasonable. The attempts at proving these conjectured properties were sometimes fruitful, yielding usful lemmas or insight in other ways. Most of the times, however, they resulted in nothing. While the thesis might present the way to the solutions as streamlined, it has been all but.

The process of puzzling with all the subproblems and patching them together to a solution has been very satisfying. We feel that the work presented provides interesting insight and hope that the reader shares this feeling.

7.3 Open problems

The work on the thesis has left som unanswered questions:

- What is the solution to the edge-extremal problem on chordal graphs?

- An interval graph is a unit interval graph if and only if it is claw-free [5]. What happen to the edge-extremal unit interval graphs when we allow them to have claws? In other words, what do the edge-extremal interval graphs look like?

- How are the edge-extremal instances for general graphs affected if we do not allow claws?

Bibliography

[1] Balachandran, N., and Khare, N. Graphs with restricted valency and matching number. *Discrete Mathematics 309*, 12 (2009), 4176 – 4180.

[2] Bollobás, B. *Extremal Graph Theory*. Dover Books on Mathematics. Dover Publications, 2004.

[3] Erdõs, P., and G., S. A combinatorial problem in geometry. *Composito Math 2*, 3 (1935), 463–470.

[4] Fraleigh, J., and Katz, V. *A first course in abstract algebra*. Addison-Wesley world student series. Addison-Wesley, 2003.

[5] Gardi, F. The roberts characterization of proper and unit interval graphs. *Discrete Mathematics 307*, 22 (2007), 2906 – 2908.

[6] Lovász, L., and Plummer, M. *Matching theory*. Elsevier Sciences Publishers B.V., Amsterdam, The Netherlands, 1993.