

Algorithmic and Structural Aspects of Graph Partitioning and Related Problems

Xiaoyan Zhang



UNIVERSITEIT TWENTE.

**Algorithmic and Structural
Aspects of Graph Partitioning
and Related Problems**

Xiaoyan Zhang

The research was funded by and carried out at the group of DMMP, Faculty of Electrical Engineering, Mathematics and Computer Science of the University of Twente, the Netherlands.

The financial support from University of Twente for this research work and publication is gratefully acknowledged.

Typeset with L^AT_EX.

Printed by CPI-Wöhrmann Print Service-Zutphen.

Copyright ©Xiaoyan Zhang, Enschede, 2014.

ISBN 978-90-365-3626-4

DOI 10.3990/1.9789036536264

All rights reserved. No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission from the copyright owner.

**ALGORITHMIC AND STRUCTURAL
ASPECTS OF GRAPH PARTITIONING
AND RELATED PROBLEMS**

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 9 mei 2014 om 14:45 uur

door

Xiaoyan Zhang

geboren op 3 oktober 1978
te Hebei, China

Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr. ir. H. J. Broersma

Prof. dr. M. Uetz

Preface

This work is the result of almost five years of research on graph partitioning and related problems in the field of theoretical computer science and graph theory between 2010 and 2014. After an introductory chapter the reader will find six chapters, each of which is written as a self-contained journal paper. The first three of these chapters deal with the complexity of some vertex partitioning problems. The final three chapters deal with structural properties of some problems related to matching problems which can be regarded as edge partitioning problems. These six chapters are based on the six joint papers that are listed below and have been submitted to journals for publication. Since the work has been written as a collection of more or less independent papers, the reader will find a certain amount of repetition of relevant concepts, definitions and background. The author apologizes for any inconvenience.

Papers underlying this research

- [1] On the complexity of edge-colored subgraphs partitioning problems in network optimization, preprint. (with Z. Zhang and H. J. Broersma)
(Chapter 2)
- [2] On the complexity of injective colorings and its generalizations, *Theoretical Computer Science* **491** (2013), 119-126. (with J. Jin and B. Xu)
(Chapter 3)

- [3] An SDP randomized approximation algorithm for max hypergraph cut with limited unbalance, preprint. (with B. Xu, X. Yu, and Z. Zhang) (Chapter 4)
- [4] Minimum size of n -factor-critical graphs and k -extendable graphs, *Graphs and Combinatorics* **28** (2012), 433-448. (with Z. Zhang, D. Lou, and X. Wen) (Chapter 5)
- [5] Directed Hamilton cycles in digraphs and matching alternating Hamilton cycles in bipartite graphs, *SIAM Journal on Discrete Mathematics* **27** (2013), 274-289. (with Z. Zhang and X. Wen) (Chapter 6)
- [6] Triangle strings: structures for augmentation of vertex disjoint triangle sets, *Information Processing Letters* (2014). In press. (with Z. Zhang) (Chapter 7)

Some other joint papers by the author:

- [1] The minimum all-ones problem for trees, *SIAM Journal on Computing* **33** (2004), 379-392. (with W.Y.C. Chen, X. Li, and C. Wang)
- [2] The edge split reconstruction problem for chemical trees is NP-complete, *MATCH Communications in Mathematical and in Computer Chemistry* **51** (2004), 205-210. (with X. Li)
- [3] Contractible cliques in k -connected graphs, *Graphs and Combinatorics* **22** (2006), 361-370. (with X. Huang, Z. Jin, and X. Yu)
- [4] Contractible subgraphs in k -connected graphs, *Journal of Graph Theory* **55** (2007), 121-136. (with Z. Jin and X. Yu)
- [5] On the minimum monochromatic or multicolored subgraph partition problems, *Theoretical Computer Science* **385** (2007), 1-10. (with X. Li)

- [6] The σ all-ones problem for trees, *Discrete Applied Mathematics* **56** (2008), 1790-1801. (with X. Li and C. Wang)
- [7] Improved bounds on linear coloring of plane graphs, *Science China Mathematics* **53** (2010), 1895-1902. (with D. Wei and B. Xu)
- [8] Degree and connectivity conditions for IM-extendibility and vertex-deletable IM-extendibility, *Ars Combinatoria* **95** (2010), 437-444. (with Z. Zhang, X. Lu, and J. Li)
- [9] Maximal independent sets in bipartite graphs with at least one cycle, *Discrete Mathematics & Theoretical Computer Science* **15** (2013), 243-258. (with S. Li and H. Zhang)

Contents

Preface	i
1 Introduction	1
1.1 Algorithmic aspects of some vertex partitioning problems . . .	4
1.1.1 Monochromatic clique and rainbow cycle partitions . . .	4
1.1.2 Injective coloring problems	7
1.1.3 Max hypergraph cut with limited unbalance	9
1.2 Structural aspects of some edge partitioning and related problems	15
1.2.1 Minimum size of n -factor-critical and k -extendable graphs	15
1.2.2 Matching alternating Hamilton cycles and directed Hamilton cycles	16
1.2.3 Structures for augmentation of vertex-disjoint triangle sets	19
2 Minimum monochromatic clique partition and rainbow cycle partition	23
2.1 Inapproximability of MCLP on monochromatic- K_4^- -free graphs	24
2.2 An approximation algorithm for WMCLP	27
2.3 RCYP is NP-complete for triangle-free graphs	31
2.4 Concluding remarks	34
3 On the complexity of injective coloring	37
3.1 Off-line injective coloring	38

3.1.1	NP-hardness of injective coloring bipartite graphs . . .	38
3.1.2	On the inapproximability of injective coloring bipartite graphs	40
3.1.3	An approximation algorithm for the max-injective coloring problem	41
3.2	On-line injective coloring	43
3.2.1	P_3 -free graphs	45
3.2.2	Triangle-free graphs and bipartite graphs	46
3.2.3	Concluding remarks	49
4	An approximation algorithm for max hypergraph cut with limited unbalance	51
4.1	An SDP relaxation of MHC-LU	52
4.2	Bound on the expected contribution of an edge by Steps 1-4 . .	57
4.3	Bounding $\mathbb{E}[\omega(\tilde{V}_1, V \setminus \tilde{V}_1)]$ after Step 5	69
4.4	Bounding $\mathbb{E}[\tilde{V}_1 (m - \tilde{V}_1)]$	72
4.5	The quality of the SDP approximation algorithm.	75
5	Minimum size of n-factor-critical and k-extendable graphs	85
5.1	Minimum size of n -factor-critical graphs and k -extendable bipartite graphs	86
5.2	Minimum size of 1-extendable non-bipartite graphs and 2-extendable non-bipartite graphs	90
5.3	Concluding remarks	97
6	Directed Hamilton cycles and matching alternating Hamilton cycles	101
6.1	Main results	101
6.2	Proof of Theorem 6.1.2	104
6.3	Concluding remarks	119
7	Triangle strings: structures for augmentation of vertex-disjoint	

triangle sets	121
7.1 Triangle strings	122
7.2 Union graph of two triangle sets and an augmenting theorem .	123
7.3 Triangle sets in triangle strings: an algorithm and a condition for triangle factors	130
Summary	135
Bibliography	138
Acknowledgements	153
Index	157

Chapter 1

Introduction

As the title of the thesis indicates, the common theme of the chapters is best described as graph partitioning problems, although the problems that we will encounter come in a huge diversity, and some of them do not have much in common. Moreover, the first three chapters after this introductory chapter deal with algorithmic aspects, while the final three chapters deal with structural aspects, also reflected in the title.

Graph partitioning problems constitute a large and important class of well-studied problems in the fields of combinatorial optimization and graph theory. In general terms, graph partitioning problems are defined on data represented in the form of a directed or undirected (sometimes weighted) graph (or sometimes hypergraph) $G = (V, E)$, with vertex set V and edge set E , and the question is whether it is possible to partition V or E into smaller subsets with specific properties. Well-known and well-studied exponents of graph partitioning problems – that have been studied almost since the start of graph theory and that appear in every text book on graph theory – are *vertex* and *edge coloring*. In these problems the aim is to partition the vertex set (edge set) of an undirected graph in a small (sometimes minimum) number of subsets such that each subset contains no pair of adjacent vertices (no pair of edges with a common end vertex). These problems and their variations have been applied in numerous application areas, e.g., in the disguise of frequency assignment problems or other assignment problems, and as time-tabling and

other scheduling problems. Other important applications of graph partitioning include scientific computing, partitioning various stages of a VLSI design circuit, task scheduling in multi-processor systems and clustering and detection of cliques in social, pathological and biological networks, to name just a few [13, 118, 129].

Distinguishing between the two sets of a graph, the class of graph partitioning problems can be divided into two subclasses, i.e., one consisting of vertex partitioning problems and the other of edge partitioning problems.

For instance, the problem of finding a certain *clique partition* can be considered as a vertex partitioning problem in which the vertex set has to be divided into k subsets that induce vertex-disjoint complete subgraphs. An *optimum* partition is defined as one in which the number k is as small as possible. Another example of a vertex partitioning problem is the aforementioned graph *vertex coloring* problem, which is a special case of graph *labeling* problems. In such problems one is aiming at an assignment of labels – traditionally called “colors” – to the vertices of a graph subject to certain constraints. The distinct colors used in an assignment then induce a partition of the vertices into disjoint sets. The max cut problem and the max hypergraph cut problem can be considered as some other examples of vertex partitioning problems. The goal in the *max cut* problem is to find a partition of the vertices of a graph into two subsets that maximizes the number of edges with end vertices in both subsets. This problem has a natural analogue in *hypergraphs*, where an edge can consist of more than two (adjacent) vertices and is called a *hyperedge*. In this setting, a hyperedge is considered not to be cut only in case all of its vertices are in one class of the partition, and cut exactly once otherwise, no matter how many vertices are in different classes of the partition.

Many vertex partitioning problems have natural and interesting counterparts for edge partitioning, like the aforementioned vertex and edge coloring problems. Other examples of well-studied edge partitioning problems with many applications are matching problems. A *matching* in a graph is a set of edges in which no pair of edges shares a common end vertex, so each edge of such a matching matches exactly two unique vertices of the graph. A *maximal matching* is a matching that is not a proper subset of any other matching in the graph. A *maximum matching* is a (maximal) matching that contains the largest possible number of edges. A *perfect matching* (also called a *1-factor*)

is a maximum matching that matches all the vertices of the graph. Matching problems have been studied in many different variants depending on additional constraints, weights (costs) on the edges and different optimization criteria.

Graph partitioning problems enjoy many practical applications as well as algorithmic and theoretical challenges. This motivates the topics of this thesis that is composed of two parts. One part is focussed on the algorithmic aspects of some vertex partitioning problems while the other part is focussed on the structural aspects of some problems related to maximum matching and perfect matching problems.

The first part of the thesis consists of Chapters 2 to 4. In this part, we present results on the complexity and inapproximability of some vertex partitioning problems, and we give approximation algorithms and on-line algorithms for some other vertex partitioning problems. We will start by investigating the inapproximability and complexity of the problems of finding the minimum number of monochromatic cliques and rainbow cycles that, respectively, partition $V(G)$, where the graph G avoids some forbidden induced subgraphs. Secondly, we study the complexity, and develop approximation algorithms and on-line algorithms for injective coloring problems, to be defined later. Finally, we consider the design of a semidefinite programming based approximation algorithm for a variant of the max hypergraph cut problem.

The second part of the thesis consists of Chapters 5 to 7. In this part, we turn our attention to structural properties of some problems that are related to matching problems.

Firstly, we determine the minimum size of a k -extendable bipartite graph and that of an n -factor-critical graph, and we study the problem of determining the minimum size of a k -extendable non-bipartite graph. We solve this problem for $k = 1$ and $k = 2$, and we pose a conjecture related to the problem for general k . Secondly, we improve two equivalent structural results due to Woodall and Las Vergnas on the existence of a directed Hamilton cycle in a digraph and the containment of every perfect matching in a Hamilton cycle in a balanced (undirected) bipartite graph, respectively. Finally, we study a generalization of the maximum matching problem called the maximum triangle set problem, in which the aim is to find the maximum number of vertex-disjoint triangles of a given graph. We present a necessary and sufficient condition for

augmenting triangle sets, analogous to the well-known augmenting path result for matchings.

In the remainder of this introductory chapter, we will present, together with the relevant terminology and notations, a survey of the main results of the thesis against a background of related results. We assume that the reader is familiar with the essentials of combinatorial optimization and graph theory. Most of the terminology and notations can be found in [25], [36] and [19].

1.1 Algorithmic aspects of some vertex partitioning problems

In the first part of this section and in Chapter 2, we mainly deal with the algorithmic aspects of the minimum monochromatic clique partitioning and the minimum rainbow cycle partitioning problem.

1.1.1 Monochromatic clique and rainbow cycle partitions

In Chapter 2, we study graph problems related to coloring and partitioning restricted to graphs avoiding certain fixed induced subgraphs. The research on problems regarding coloring and partitioning has a relatively long history, and many important and impressive results have been obtained (See, e.g., Erdős et al. [41], Gyárfás and Simonyi [67], Gyárfás et al. [66], Alon et al. [6], Brualdi and Hollingsworth [30], Alon and Gutin [7], Feder et al. [43], Feder and Motwani [44], Suzuki [130], Akbari and Alipour [3], and Gourvès et al. [60]). Several variations of such problems, and in particular their computational complexity, have been well-studied as well. MacGillivray and Yu [110] studied a general graph partitioning problem including graph coloring, homomorphism to H , conditional coloring, contractibility to H , and partition into cliques as special cases, and investigated its complexity. Yegnanarayanan [139] considered three coloring parameters of a graph G in connection with the computational complexity, partitions, algebra, projective plane geometry and analysis. For more general coloring and partitioning problems, the reader could refer to Garey and Johnson [54], and Kano and Li [86].

Several papers focus on a class of related problems in which the aim is to determine the minimum number k such that the vertex set of an edge-colored graph can be covered by at most k vertex-disjoint monochromatic or rainbow subgraphs, such as paths, trees, cycles and cliques. Erdős, Gyárfás and Pyber [41] introduced the notion of the monochromatic cycle partition number of an r -edge-colored graph G , which is the minimum number k such that the vertices of G can be covered by k vertex-disjoint monochromatic cycles. Here monochromatic means that all edges of the subgraph (cycle) have the same color. They proved that the monochromatic cycle partition number of an r -edge-colored complete graph is less than $cr^2 \log r$ for some constant c . In [66], Gyárfás et al. significantly improved this result for large n . Jin et al. [82] investigated the computational complexity of the problem of partitioning complete multipartite 2-edge-colored graphs into the minimum number of vertex-disjoint monochromatic cycles, paths and trees, respectively. Rainbow partitioning problems and related problems were studied by Alon et al. [6], Brualdi and Hollingsworth [30], Alon et al. [8], Suzuki [130], and Akbari and Alipour [3]. Here rainbow means that all edges of the subgraph have distinct colors.

More recently, researchers studied problems on graphs avoiding some induced subgraphs (See, e.g., Ota and Sueiro [117], Li et al. [99], and Broersma et al. [29] for some examples). In particular, many algorithmic problems concerning coloring and forbidden induced subgraphs have been studied lately. Král et al. [94] gave a complete characterization of all graphs H for which the problem of coloring H -free graphs is polynomial and for which it is NP-complete. Here a graph is called H -free if it does not contain a copy of H as an induced subgraph. They further initiated a study of this problem for pairs of forbidden subgraphs. Motivated by the Strong Perfect Graph Conjecture, Hoàng and Le [78] studied problems related to P_4 -free colorings and showed that P_4 -free k -coloring on comparability graphs is NP-hard. Fiala et al. [51] showed that weakly P_3 -free 2-coloring (they used the term 2-subcolorability in their paper) for triangle-free planar graphs is NP-hard. Broersma et al. [28] considered the problem of coloring a planar graph with the minimum number of colors such that each color class avoids one or more forbidden subgraphs, and studied the computational complexity of this problem in detail. Li and

Zhang [102] showed that both the problems of determining the minimum number of monochromatic cliques and the minimum number of rainbow cycles that partition $V(G)$ for K_4^- -free graphs G are NP-complete. Here K_4^- denotes the graph obtained by deleting one edge from a K_4 .

Clique and cycle partition problems have important applications in engineering problems as well as in operations research problems such as VLSI design automation, resource allocation, and periodic scheduling. In Chapter 2, we focus on clique and cycle partition problems on edge-colored graphs. We abbreviate the problems of partitioning the vertex set of a (not necessarily properly) edge-colored graph into a minimum number of monochromatic cliques and rainbow cycles to MCLP and RCYP, respectively. We investigate the inapproximability and complexity of MCLP and RCYP on graphs avoiding certain induced subgraphs, and derive a tight approximation algorithm for the weighted MCLP on monochromatic- K_4^- -free graphs. Before we present our results, we list some terminology and notations that will be used in Chapter 2.

Let $G = (V, E)$ be a connected undirected simple graph. If G is assigned a mapping $\ell : E \rightarrow \mathbb{N}$, we say that G is an *edge-colored graph*. We call $\ell(e)$ the color of the edge $e \in E$, and we use $\ell(H)$ to denote the number of different colors in the set $\{\ell(e) | e \in E(H)\}$ for a subgraph H of G . A *complete graph* is a graph in which every two distinct vertices are adjacent. We denote by K_m a complete graph on m vertices, and by C_m a cycle on m vertices. A *clique* of G is a nonempty subset of $V(G)$ that induces a complete subgraph of G . A clique CL of G is called a *monochromatic clique* if all the edges of the corresponding subgraph of G have the same color. A cycle CY of G is called a *rainbow cycle* if $\ell(CY) = |E(CY)|$, i.e., if no two edges of CY have the same color. Note that a single vertex can be viewed as a degenerate monochromatic clique or rainbow cycle. We simply call it a *vertex-clique* or *vertex-cycle*.

Let K_4^- denote the graph obtained by deleting one edge from a K_4 . A graph G is said to be *K_4^- -free* if it does not contain K_4^- as an induced subgraph. And a graph G is called *monochromatic- K_4^- -free* if any monochromatic subgraph of G does not contain a K_4^- as an induced subgraph. Note that the properties of being K_4^- -free and monochromatic- K_4^- -free do not imply each other. For example, a K_4 with one edge colored ℓ_1 and the others colored ℓ_2 is K_4^- -free, but not monochromatic- K_4^- -free. However, a monochromatic cycle on 4

vertices with a chord of a different color is monochromatic- K_4^- -free, but not K_4^- -free. A vertex u is *color-adjacent* to a vertex v of a monochromatic clique CL if the edge (u, v) has the same color as the edges of CL . A clique CL of G is called a *maximal monochromatic clique* if there is no vertex $u \in V(G) \setminus V(CL)$ color-adjacent to each vertex of CL .

We show that the minimum monochromatic clique partition problem is APX-hard on K_4^- -free graphs and monochromatic- K_4^- -free graphs, and APX-complete on monochromatic- K_4^- -free graphs in which the size of a maximum monochromatic clique is bounded by a constant. We also show that the minimum rainbow cycle partition problem is NP-complete, even if the input graph G is triangle-free. Moreover, for the weighted version of the minimum monochromatic clique partition problem on monochromatic- K_4^- -free graphs, we derive an approximation algorithm with (tight) approximation guarantee $\ln |V(G)| + 1$.

In the second part of this section and in Chapter 3, we study the algorithmic aspects of the injective coloring problem and some variations.

1.1.2 Injective coloring problems

An *injective k -coloring* of a graph G is a (partition) mapping $c : V(G) \rightarrow \{1, \dots, k\}$ such that $c(u) \neq c(v)$ for any two distinct vertices u and v in $V(G)$ that have a common neighbor. The *injective chromatic number* of a graph G , denoted by $\chi_i(G)$, is the smallest k such that G admits an injective k -coloring. This concept originates from complexity theory on Random Access Machines, and can be applied in the theory of error-correcting codes. In [68], Hahn et al. introduced these concepts, and they proved that $\Delta(G) \leq \chi_i(G) \leq \Delta(G)(\Delta(G) - 1) + 1$ (where $\Delta(G)$ denotes the maximum degree of G). They also characterized the extremal graphs. It is an NP-hard problem to determine the injective chromatic number of a given input graph [68], and it remains NP-hard if the instances are restricted to the class of chordal graphs [77].

Some difficult combinatorial problems that are NP-hard in general admit polynomial time solutions when restricted to instances that avoid certain fixed forbidden subgraphs (see [101] and [103] for examples). In [77], Hell et al. showed that the injective chromatic number of a tree can be computed

efficiently. However, we show in Chapter 3 that it remains NP-hard to determine the injective chromatic number when the instances are restricted to bipartite graphs with some special properties. Furthermore, we show that for every $\epsilon > 0$, it is impossible to efficiently approximate the injective chromatic number of any bipartite graph within a factor of $n^{\frac{1}{3}-\epsilon}$ unless $ZPP = NP$.

We also studied a variation on the injective coloring problem that was motivated by the following weighted analogue of vertex coloring. Let G be a graph with a weight function $\omega : V(G) \rightarrow \mathbf{N}$. The *max-coloring problem* seeks to find a partition of the vertices of G into independent sets that minimizes the sum of the weights of the heaviest vertices in each independent set, involving one vertex of each set in the summation. Pemmaraju, Raman and Varadarajan [119] showed that the max-coloring problem is NP-hard even when restricted to interval graphs, and they devised a simple 2-approximation algorithm for max-coloring on interval graphs. Guan and Zhu [62] showed that the max-coloring problem can be solved in polynomial time on graphs of bounded path-width. Motivated by the allocating buffer application and digital signal processing applications, Govindarajan and Rengarajan [61] experimentally evaluated a first-fit strategy which produces a solution of the max-coloring problem on circular-arc graphs, with weight no more than 102.1% of the optimal weight. In Chapter 3 we also study the max-injective coloring problem, which can be viewed as a combination of max-coloring and injective coloring. Given a graph G with a weight function $\omega : V(G) \rightarrow \mathbf{N}$, the *max-injective coloring problem* is aimed at finding an injective vertex coloring with color classes C_1, C_2, \dots, C_k of G that minimizes $\sum_{i=1}^k \max_{v \in C_i} \omega(v)$. When $\omega(v) = 1$ for all $v \in V(G)$, $\min \sum_{i=1}^k \max_{v \in C_i} \omega(v)$ is simply $\chi_i(G)$. We prove that there is a constant approximation algorithm for solving the max-injective coloring problem on power chordal graphs (graphs of which all powers are chordal) with bounded injective chromatic number, and we devise a constant approximation algorithm for max-injective coloring restricted to some subclass of bipartite graphs.

Motivated by the vast existing literature on the on-line versions of coloring problems, we also study the on-line version of injective coloring. The injective coloring problem gets more complicated in the on-line situation. In this case, vertices of a graph are presented one at a time, and the algorithm has to assign a color irrevocably to a vertex as it comes in. The procedure depends only on the knowledge of the subgraph that has been revealed so

far. To be precise, an *on-line injective coloring* algorithm for a graph G is an algorithm that injectively colors G by receiving its vertices in some order v_1, v_2, \dots, v_n , where the color of v_i is assigned depending only on the subgraph of G induced by $\{v_1, v_2, \dots, v_i\}$ and the colors assigned to these vertices. As usual, turning to the on-line variant makes the problem more complicated. Actually, in Chapter 3 we show that the worst-case performance ratio between on-line and off-line injective coloring of a path on $2n$ vertices is at least $\frac{n}{2}$. Gyárfás and Lehel [65] introduced the concept of on-line coloring while translating a rectangle packing problem in dynamical storage allocation into a coloring problem. Since then, on-line coloring of graphs has been investigated extensively. The optimal competitive ratio of on-line coloring is only slightly sublinear in general [109]. However, a constant competitive ratio is possible on interval graphs [65,92], and a logarithmic ratio can be achieved on bipartite graphs and sparse graphs [79]. Broersma, Capponi and Paulusma [27] proved that there exists an on-line competitive algorithm for the class of P_6 -free bipartite graphs and P_7 -free bipartite graphs, where the number of colors used is bounded by roughly twice and roughly eight times the on-line chromatic number, respectively. In contrast, there exists no competitive on-line algorithm to color P_6 -free bipartite graphs, i.e., for which the number of colors is bounded by any function depending only on the chromatic number [65].

In Chapter 3, we prove that the on-line algorithm known as *First Fit* (FF for short) [65] optimally injectively colors P_3 -free graphs. We also show that the number of colors used by FF* on a bipartite graph G is bounded by $\frac{3}{2}$ times its on-line injective chromatic number, where FF* is an on-line algorithm equivalent to proper coloring the complement of G by FF. Moreover, we present an improved algorithm BFF, and prove that it is optimal for on-line injectively coloring bipartite graphs.

At the end of this section and in Chapter 4, we consider the design of a semidefinite programming (SDP) based approximation algorithm for the problem of max hypergraph cut with limited unbalance (MHC-LU).

1.1.3 Max hypergraph cut with limited unbalance

A *hypergraph* is an ordered pair $H = (V, E)$ in which $V := \{1, 2, \dots, m\}$ is a finite nonempty set and $E := \{S_1, S_2, \dots, S_n\}$ is a collection of distinct

nonempty subsets of V . V and E are the sets of *vertices* and *edges* of H , respectively. A weighted hypergraph is a hypergraph together with a nonnegative real function $\omega : E \rightarrow \mathbb{R}^+$. For convenience, we write $\omega_j := \omega(S_j)$. Given a partition $V = V_1 \cup V_2$, the edge S_j is said to be a *cut edge* with respect to this partition if $S_j \cap V_i \neq \emptyset$ for $i = 1, 2$. The *max hypergraph cut with limited unbalance* problem (MHC-LU) asks for a partition $V = V_1 \cup V_2$ such that $||V_2| - |V_1|| \leq u$ for some given $u \geq 0$ and the total weight of the cut edges is maximized. Note that MHC-LU with $u = 0$ is also known as *max hypergraph bisection*, and MHC-LU with $u = m$ is also studied under the names *max set splitting* and *max hypergraph cut*.

Hypergraph partitioning problems arise naturally in important practical problems, including circuit design and network planning, etc. [76, 87, 128]. For most of the applications, the constraints on unbalancedness make sense. For example, from the point of view of the circuit designer, the suitability of a partition of a circuit is not hugely affected if one relaxes the bisection constraint to limited unbalance (with a small u) in order to get better results in terms of approximating the optimum weight of the cut [53]. Since the partitioning of hypergraphs is critical in several practical application areas, many heuristic algorithms were developed [35, 106]. In Chapter 4, we present a polynomial time SDP randomized approximation algorithm for MHC-LU with guaranteed performance ratio.

When the hypergraph is 2-uniform (a standard graph), MHC-LU is known as the *maximum cut with limited unbalance* problem (MC-LU). Galbiati and Maffioli [53] developed polynomial time randomized approximation algorithms with nontrivial performance guarantees for MC-LU.

The well-known *max cut* problem is equivalent to MHC-LU with $u = m$ and $|S_j| = 2$ for all j . Goemans and Williamson [57], in a major breakthrough, used semidefinite programming relaxation and hyperplane rounding to obtain an approximation algorithm for the Max Cut problem with expected performance guarantee 0.87856. This well-known algorithmic paradigm, with more sophisticated techniques, has been applied to many previously studied problems [49, 50, 52, 53, 70, 71, 138, 143].

When $u = 0$ and $|S_j| = 2$ for all j , MHC-LU is known as the *max bisection* problem. Frieze and Jerrum [52], and Andersson [12] presented a

0.65-approximation algorithm for the Max Bisection problem, and Ye [138] obtained a 0.699-approximation algorithm for this problem. Halperin and Zwick [70] improved the performance ratio for the Max Bisection problem to 0.701, which was further improved by Feige and Langberg [50] to 0.7028 using the RPR² rounding technique. For the case of regular graphs, Feige et al. improved the approximation ratio to 0.795 (0.834 for 3-regular graphs) [46, 47]. Recently, Raghavendra and Tan [127] significantly improved all the above results to 0.85 by using the Lasserre Hierarchy. This algorithm was further improved by Austrin et al. [15] to 0.8776 by also using a relaxation based on the Lasserre Hierarchy.

When $u = 1$, MHC-LU asks for *balanced* bipartitions, i.e. partitions $V = V_1 \cup V_2$ such that $||V_1| - |V_2|| \leq 1$. Bollobás and Scott [23] conjectured that if G is a graph with minimum degree at least 2, then $V(G)$ admits a balanced bipartition V_1, V_2 such that for each $i \in \{1, 2\}$ at most $|E(G)|/3$ edges have both ends in V_i . The minimum degree condition is necessary. Bollobás and Scott [24] established this conjecture for regular graphs. Xu, Yan and Yu [135, 136] proved this conjecture for graphs G with $\Delta(G) \leq \frac{7}{5}\delta(G)$ or with $\delta(G) \geq 5$, where $\Delta(G)$ and $\delta(G)$ are maximum and minimum degrees of G , respectively. Lee, Loh and Sudakov [98] proved a nice asymptotic result stating that every graph with m edges and minimum degree $2k$ or $2k + 1$ admits a balanced bipartition V_1, V_2 such that $\max\{e(V_1), e(V_2)\} \leq (\frac{k+1}{2(2k+1)} + o(1))m$ (when $k = 1$, its main term is $\frac{m}{3}$). The conjecture has been confirmed recently by Xu and Yu [137]. They proved that every graph G with m edges and minimum degree at least 2 admits a balanced bipartition V_1, V_2 with $\max\{e(V_1), e(V_2)\} < m/3$ unless G is a triangle.

For $u = m$, MHC-LU becomes the so-called *max set splitting* problem. Andersson and Engebretsen [11] presented a 0.72-approximation algorithm for this problem, and the approximation ratio was improved to 0.7499 by Zhang, Ye and Han [143]. Gaur and Krishnamurti [55] gave a $k/(k+1)$ -approximation algorithm for the problem, where $k \geq 3$ is the minimum number of vertices in a hyperedge. Arora, Karger and Karpinski [14] designed a PTAS for dense instances of this problem. When restricted to k -uniform hypergraphs the Max Set Splitting problem is known as the *max Ek -set splitting* problem. For any fixed $k \geq 2$, Lovász [108] and Petrank [120] have shown that the max Ek -set splitting problem is NP-hard and APX-complete, respectively.

When $k = 2$, the max E_k -set splitting problem is equivalent to the max cut problem. When $k = 3$, the performance guarantee has been improved by Zwick to 0.90871 [148]. More precisely, Zwick obtained a 0.90871-approximation algorithm for *MAX NAE- $\{3\}$ -SAT* [147, 148], which is the restriction of *MAX NAE SAT* to instances in which all clauses are of size at most 3, where *MAX NAE SAT* is a variant of the well known *MAX SAT*. The objective of *MAX NAE SAT* is to maximize the clauses which contain both true and false literals. Obviously, max E_3 -set splitting is a special case of *MAX NAE- $\{3\}$ -SAT* in which all literals appear unnegated, and thus max E_3 -set splitting can also be approximated with the 0.90871 performance guarantee. When $k \geq 4$, Alimonti [4] and Kann, Lagergren and Panconesi [85] showed that the max E_k -set splitting problem can be approximated within $1 - 2^{1-k}$, which is best possible [64, 75].

Ageev and Sviridenko [1, 2] considered MHC-LU for $u = m - 2k$ with the strict condition that $||V_2| - |V_1|| = u$ (hence $|V_i| = k$ and $|V_{3-i}| = m - k$ for some $i \in \{1, 2\}$), and they gave a 0.5-approximation algorithm based on linear programming. For graphs, Hassin and Rubinfeld [74] presented a different 0.5-approximation with a better running time. Feige and Langberg [49] combined the method in [1] with the semidefinite programming approach to design a $(0.5 + \epsilon)$ -approximation algorithm, where ϵ is some unspecified small positive number. Han et al. [71] and Jäger and Srivastav [81] applied semidefinite programming to obtain better approximation factors than previously known.

In Chapter 4, we apply semidefinite programming to the more general MHC-LU, building on several earlier ideas of Galbiati and Maffioli [53] (for MC-LU), Ye [138] (for Max Bisection), Andersson and Engebretsen [11] and Zhang et al. [143] (for Max Set Splitting). By solving the semidefinite programming relaxation of MHC-LU, we obtain an (almost) optimal vector solution $(v_0^*, v_1^*, \dots, v_m^*)$. In previous work, these vectors are usually rounded by applying an important technique named outward rotations [148], which combines the classical hyperplane rounding method [57] with independent random choice to partition the coordinates into two parts. Motivated by Halperin and Zwick [70] (for maximum graph bisection problems), we also apply the idea of outward rotations to a linear randomized rounding method [70]. This way we obtain better performance ratios for MHC-LU when the minimum number of vertices in a hyperedge is 3. Moreover, we present a generalized formula for

the performance ratio for MHC-LU using some additional parameters. However, it remains open to find a best set of the parameters that optimizes the generalized formula. In practice, the final ratios need to be obtained via a computer search over the parameter space. For fast computations and easy verifications of computational results, we may assign simple values to some of the parameters and simply perform 1-dimensional searches on the remaining parameters. This turns out to be sufficient to improve the following numerical results, where $\tau = u/m$ is known as the *unbalance parameter*:

- For $\tau = 0$, we obtain approximation ratio 0.6271, improving the 0.5-approximation ratio of Ageev and Sviridenko [2] for Max Hypergraph Bisection which was based on linear programming (see Table 1 for $\tau = 0$). The improvement is a consequence of our strengthened SDP relaxation based approximation algorithm.
- For $\tau = 1$, i.e., the version of the Max Set Splitting problem, our strengthened SDP relaxation and the generalized formula for the performance ratio gives approximation ratio 0.7524 which improves the ratio 0.7499 in [143] (see Table 1 for $\tau = 1$).

We show the lower bounds on the approximation ratios for MHC-LU, for the range $0 < \tau < 1$, in Table 1. The corresponding approximation values obtained previously are shown in the last row of Table 1. Note that “none” in the last row of Table 1, indicates no known previous results (as far as we know).

- We obtain approximation ratio 0.7741 for MHC-LU when the minimum number of vertices in a hyperedge is 3, which improves the result 0.75 in Gaur and Krishnamurti [55] (see Table 2 for $\tau = 1$). This improvement is also due to the strengthened SDP relaxation and an improved rounding method by combining the outward rotations of random hyperplane rounding procedure with that of linear randomized rounding procedure. The lower bounds on the approximation ratios for MHC-LU when the minimum number of vertices in a hyperedge is 3 are shown in Table 2, for the range $0 \leq \tau \leq 1$.
- We show that one can further improve the performance ratios like those in [53] for Max Cut with Limited Unbalance when $0.5 < \tau < 1$, using

a new formula for the performance ratio by a tighter analysis than that in [53]. In Table 3, we show the lower bounds on the approximation ratios for Max Cut with Limited Unbalance, for the same values of τ as in [53] when $0.5 < \tau < 1$. The corresponding approximation ratios in [53] are shown in the last row of Table 3.

At the end of Section 6, we present the first worst-case performance ratio 0.6271 of the SDP-algorithm for approximating MHC-LU regardless of the value of τ .

Table 1.1: New results R for MHC-LU compared with previous results R' for some τ

τ	0	0.25	0.5	0.75	0.9	0.999	1
R	0.6271	0.7105	0.7130	0.7194	0.7353	0.7522	0.7524
R'	0.5	none	none	none	none	none	0.7499

Table 1.2: New results R for MHC-LU compared with previous results R' for some τ when the minimum number of vertices in a hyperedge is 3

τ	0	0.25	0.5	0.75	0.9	0.999	1
R	0.7042	0.7459	0.7495	0.7564	0.7656	0.7740	0.7741
R'	none	none	none	none	none	none	0.75

Table 1.3: New results R for MC-LU compared with previous results R' in [53] for some $0.5 < \tau < 1$

τ	0.6	0.7	0.8	0.85	0.9	0.95	0.9999
R	0.7987	0.8052	0.8191	0.8291	0.8417	0.8584	0.8785
R'	0.795	0.793	0.790	0.8126	0.834	0.856	0.878

1.2 Structural aspects of some edge partitioning and related problems

In the second part of the thesis, we turn our attention to structural aspects of selected edge partitioning problems. In the first section of this part and in Chapter 5, we determine the minimum sizes of a k -extendable bipartite graph, an n -factor-critical graph and a k -extendable non-bipartite graph.

1.2.1 Minimum size of n -factor-critical and k -extendable graphs

All graphs considered in this section are finite, connected and simple. Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. The number of vertices in G is denoted by $\nu(G)$ or ν . The number of edges in G is called its *size*. The connectivity, the edge connectivity, the independence number, the minimum degree and the maximum degree of G are denoted by $\kappa(G)$, $\kappa'(G)$, $\alpha(G)$, $\delta(G)$ and $\Delta(G)$, respectively. The neighborhood of a vertex v in G is denoted by $N_G(v)$ or $N(v)$. For other terminologies not defined here, the reader is referred to [25].

A connected graph G is said to be k -*extendable*, if it contains a matching of size k and every matching in G of size k is contained in a perfect matching of G , where k is an integer such that $0 \leq k \leq (\nu(G) - 2)/2$. The concept of k -extendable graphs was introduced by Plummer [122]. A graph G is said to be n -*factor-critical*, or n -*critical*, if $G - S$ has a perfect matching for every $S \subseteq V(G)$ with $|S| = n$, where $0 \leq n \leq \nu(G) - 2$. When $n = 1$ or 2 , we say that G is *factor-critical* or *bicritical*. The concept of n -factor-critical graphs was introduced by Yu [140] and Favaron [42], independently. Extensive researches have been done on these two classes of graphs. The reader may trace the important developments on this field by referring to the surveys [124], [125] and [126] by Plummer, as well as Chapter 6 and Chapter 7 of the book [141] by Yu and Liu. Furthermore, a good description of the application of extendibility and factor-criticality in job assignment can be found in [104].

In [72], Harary defined Harary graphs $H_{m,\nu}$, which are m -connected graphs on ν vertices with $\lceil m\nu/2 \rceil$ edges, for $2 \leq m < \nu$. By considering Harary graphs and related graphs, we obtain the following results.

Theorem 1.2.1. Let $r \geq 2$ and $\nu > 2r$ be two integers. Then $H_{2r,\nu}$ is $(2r - 1)$ -factor-critical if ν is odd and $(2r - 2)$ -factor-critical if ν is even.

Theorem 1.2.2. Let $r \geq 2$ and $\nu > 2r + 1$ be two integers. Then $H_{2r+1,\nu}$ is $2r$ -factor-critical if ν is even and $(2r - 1)$ -factor-critical if ν is odd.

Theorem 1.2.3. Let $s \geq 2$ be an integer. Then $H_{2,2s-1}$ and $H_{3,2s+1}$ are factor-critical. $H_{2,2s}$ is 1-extendable. $H_{3,2s}$ is bicritical if s is even, and 2-extendable if s is odd.

Let $k \geq 1$ be an integer, and denote by $\varepsilon(\nu, k)$ the minimum size of a k -extendable non-bipartite graph G on $\nu \geq 2k + 2$ vertices. We then have the following results.

Theorem 1.2.4. For an even number $\nu \geq 4$, $\varepsilon(\nu, 1) = \nu + 2$.

Theorem 1.2.5. For an even integer $\nu \geq 6$,

$$\varepsilon(\nu, 2) = \begin{cases} 15, & \text{if } \nu = 6, \\ 16, & \text{if } \nu = 8, \\ 19, & \text{if } \nu = 10, \\ 20, & \text{if } \nu = 12, \\ 3\nu/2, & \text{if } \nu \geq 14. \end{cases}$$

In the next section and in Chapter 6, we consider the problems of matching alternating Hamilton cycles in bipartite graphs and directed Hamilton cycles in digraphs.

1.2.2 Matching alternating Hamilton cycles and directed Hamilton cycles

Hamiltonian problems, and their many variations, have been studied extensively for more than half a century. The readers could refer to the surveys of Gould [58,59], Kawarabayashi [89] and Broersma [26] to trace the development in this field. Recently, approximate solutions to many traditional hamiltonian problems and conjectures in digraphs came forth [34, 90, 91, 93], which are surveyed by Kühn and Osthus [95].

Hamiltonicity and related properties are also important in practical applications. For example, in network design the existence of Hamilton cycles

in the underlying topology of an interconnection network provides advantages for routing algorithms to make use of a ring structure, while the existence of a hamiltonian decomposition allows the load to be equally distributed, making the network more robust [22].

There are lots of degree and degree sum conditions known for guaranteeing hamiltonicity. In most cases, the lower bounds in such conditions are best possible. However, we could still consider reducing the bounds and trying to identify all exceptional graphs, that is, the extremal graphs for the conditions. This approach often leads to the discovery of interesting topological structures. In this thesis, we apply this idea to Woodall's condition for the existence of directed Hamilton cycles in digraphs.

In Chapter 6 we consider finite, simple and connected graphs, and finite and simple digraphs. Let D be a digraph with vertex set $V(D)$ and arc set $A(D)$, and let u and v be distinct vertices of D . We denote by $|D|$ the order of $V(D)$, and by $d^+(u)$ and $d^-(u)$ the out-degree and in-degree of u , respectively. The degree of u is the sum of its out-degree and in-degree. The minimum out-degree and in-degree of the vertices in D is denoted by $\delta^+(D)$ and $\delta^-(D)$, respectively. We let $\delta^0(D) = \min\{\delta^+(D), \delta^-(D)\}$. If $(u, v) \in A(D)$ or $(v, u) \in A(D)$, we say that u and v are adjacent. A *transitive tournament* is an orientation of a complete graph for which the vertices can be numbered in such a way that (i, j) is an arc if and only if $i < j$.

Below are three directed analogues of two early theorems due to Dirac [40] and Ore [116] for the existence of Hamilton cycles in undirected graph.

Theorem 1.2.6. (Ghouila-Houri [56]) Let D be a strong digraph. If the degree of every vertex of D is at least $|D|$, then D has a directed Hamilton cycle.

Theorem 1.2.7. (Corollary 5.6.3 [19]) If D is a digraph with $\delta^0(D) \geq |D|/2$, then D has a directed Hamilton cycle.

Theorem 1.2.8. (Woodall [134]) Let D be a digraph. If for every vertex pair u and v , where there is no arc from u to v , we have $d^+(u) + d^-(v) \geq |D|$, then D has a directed Hamilton cycle.

It is not hard to verify that the bounds in above theorems are tight. Nash-Williams [114] raised the problem of describing all the extremal digraphs in

Theorem 1.2.6, that is, all digraphs with minimum degree at least $|D| - 1$, that do not have a directed Hamilton cycle. As a partial solution to this problem, Thomassen proved a structural theorem on the extremal graphs.

Theorem 1.2.9. (Thomassen [131]) Let D be a strong non-hamiltonian digraph, with minimum degree $|D| - 1$. Let C be a longest directed cycle in D . Then any two vertices of $D - C$ are adjacent, every vertex of $D - C$ has degree $|D| - 1$ (in D), and every component of $D - C$ is complete. Furthermore, if D is strongly 2-connected, then C can be chosen such that $D - C$ is a transitive tournament.

Darbinyan characterized the digraphs of even order that are extremal for both Theorem 1.2.6 and Theorem 1.2.7.

Theorem 1.2.10. (Darbinyan [38]) Let D be a digraph of even order such that the degree of every vertex of D is at least $|D| - 1$ and $\delta^0(D) \geq |D|/2 - 1$. Then either D is hamiltonian or D belongs to a non-empty finite family of non-hamiltonian digraphs.

We study the extremal graphs of Theorem 1.2.8 in Chapter 6. In contrast to Theorem 1.2.9 and Theorem 1.2.10, we can completely determine all the extremal graphs.

For other results on degree sum conditions for the existence of Hamilton cycles in digraphs see [16], [17], [18], [38], [39], [111], [113], [145], [146], and a good summary in Chapter 5 of [19].

Another interesting aspect of directed Hamilton cycle problems is their connection with the problem of matching alternating Hamilton cycles in bipartite graphs. Given a bipartite graph G with a perfect matching M , if we orient the edges of G towards the same part, then contracting all edges in M , we get a digraph D . An M -alternating Hamilton cycle of G corresponds to a directed Hamilton cycle of D , and vice versa. Hence, Theorem 1.2.8 is equivalent to the following theorem.

Theorem 1.2.11. (Las Vergnas [97]) Let $G = (B, W)$ be a balanced bipartite graph of order ν . If for any $b \in B$ and $w \in W$, where b and w are nonadjacent, we have $d(w) + d(b) \geq \nu/2 + 2$, then for every perfect matching M of G , there is an M -alternating Hamilton cycle.

Hence, we also determine the extremal graphs for the result of Las Vergnas in Chapter 6 of this thesis.

Theorem 1.2.11 is an instance of the problem of cycles containing matchings, which studies the conditions that enforce certain matchings to be contained in certain cycles. Some related works can be found in [9], [10], [21], [69], [80], [88], [121] and [133]. In particular, Berman proved the following.

Theorem 1.2.12. (Berman [21]) Let G be a graph on $\nu \geq 3$ vertices. If for any pair of independent vertices $x, y \in V(G)$, we have $d(x) + d(y) \geq \nu + 1$, then every matching lies in a cycle.

Similarly to the above-mentioned works, Jackson and Wormald determined all the extremal graphs of a generalized version of Berman's result.

Theorem 1.2.13. (Jackson and Wormald [80]) Let G be a graph on ν vertices and M be a matching of G such that (1) $d(x) + d(y) \geq \nu$ for all pairs of independent vertices x, y that are incident with M . Then M is contained in a cycle of G unless equality holds in (1) and several exceptional cases happen.

Motivated by the above results, in Chapter 6 we reduce both the lower bounds presented in Theorem 1.2.8 and Theorem 1.2.11 by 1, and prove that the conclusions still hold, with only a few exceptional cases that can be clearly characterized.

Finally, in the concluding section and in Chapter 7 we consider the structures for augmentation of vertex-disjoint triangle sets.

1.2.3 Structures for augmentation of vertex-disjoint triangle sets

We consider undirected, simple graphs in this section. Let G be a graph. A set \mathcal{T} of vertex-disjoint triangles in G is called a *vertex-disjoint triangle set* of G , or a *triangle set* for short. The number of triangles in \mathcal{T} , denoted by $|\mathcal{T}|$, is called its *size*. A triangle set of G with the maximum size is called a *maximum triangle set* of G . We say that a vertex u is *covered* by a triangle set \mathcal{T} , if u is a vertex of a triangle in \mathcal{T} . If \mathcal{T} covers all vertices of G , we say that \mathcal{T} is a *perfect triangle set*, or a *triangle factor* of G .

The study on triangle sets and triangle factors has a long history. Important results include sufficient conditions for the existence of triangle factors in graphs, and bounds on the size of the maximum triangle sets in graphs. For example, the following fundamental result is a special case of a theorem in [37].

Theorem 1.2.14. (Corrádi and Hajnal, [37]) If G is a graph with $3k$ vertices and minimum degree at least $2k$, then G contains a triangle factor.

In balanced tripartite graphs the minimum degree bound can be reduced.

Theorem 1.2.15. (Johansson [83]) Let G be a tripartite graph with $3k$ vertices, k in each class, such that each vertex is adjacent to at least $\frac{2}{3}k + \sqrt{k}$ of the vertices in each of the other two classes. Then G has a triangle factor.

Another example is a result on the size of triangle sets in claw-free graphs.

Theorem 1.2.16. (Wang [132]) For any integer $k \geq 2$, if G is a claw-free graph of order at least $6(k-1)$ and with minimum degree at least 3, then G contains a triangle set of size k unless G is of order $6(k-1)$ and G belongs to a known class of graphs.

The problem of determining a maximum triangle set in a given graph, usually called the vertex-disjoint triangles problem, or VDT for short, catches much attention. The VDT problem has many variants such as computing maximum triangle sets in edge-weighted graphs [73], in degree-bounded graphs [31, 32, 84], or in some special classes of graphs [63].

While trying to compute a triangle set in a graph G , we can clearly ignore the edges that are not contained in any triangle without affecting the results. Therefore, henceforth we assume that all edges of the graph G we consider are contained in some triangle. Under this assumption, the VDT problem on tripartite graphs is equivalent to the following 3-dimensional matching problem. Given three finite and disjoint sets W , X and Y , and a subset \mathcal{T} of $W \times X \times Y$, a *3-dimensional matching* \mathcal{M} is a subset of \mathcal{T} such that every element of W , X and Y appears in the triples in \mathcal{M} at most once. The 3-dimensional matching problem (3DM) asks for a maximum 3-dimensional matching of \mathcal{T} , and is a well-known NP-hard problem [54]. In [84], Kann further showed that even if the appearance of every element in \mathcal{T} is bounded by a constant B ,

where $B \geq 3$, the 3DM problem is MAX SNP-complete. In [32], Chlebík and Chlebíková proved that even in case $B = 2$, it is NP-hard to achieve an approximation factor of $\frac{95}{94}$ for the 3DM problem.

Triangle sets can be viewed as a generalization of matchings in graphs. For matching problems, Berge's famous characterization states that a matching M in a graph G is maximum if and only if G has no M -augmenting path [20]. However, for triangle sets in graphs, we do not know of any analogous augmenting structure result. In Chapter 7, we describe a class of structures called *triangle strings*, which corresponds to the union of the graphs of two triangle sets. Based on the concept of triangle strings, we give a sufficient and necessary condition under which a triangle set \mathcal{T} of a graph G can be augmented. We describe an algorithm for deciding whether a given graph G with degree bound 4 is a triangle string; moreover, if G is a triangle string, the algorithm finds a maximum triangle set of G . Finally we give a sufficient and necessary condition under which a triangle string has a triangle factor.

Chapter 2

Minimum monochromatic clique partition and rainbow cycle partition

Given a (not necessarily properly) edge-colored graph $G = (V, E)$, a subgraph H is said to be *monochromatic* if all its edges have the same color, and called *rainbow* if all its edges have distinct colors. We investigate the computational complexity of the problems of determining the minimum number of monochromatic cliques (MCLP) and rainbow cycles (RCYP) that, respectively, partition $V(G)$. We show that the minimum monochromatic clique partition problem is APX-hard on K_4^- -free graphs and on monochromatic- K_4^- -free graphs, and APX-complete on monochromatic- K_4^- -free graphs in which the size of a maximum monochromatic clique is bounded by a constant. We also show that the minimum rainbow cycle partition problem is NP-complete, even if the input graph G is triangle-free. Moreover, for the weighted version of the minimum monochromatic clique partition problem on monochromatic- K_4^- -free graphs, we derive an polynomial approximation algorithm with (tight) approximation guarantee $\ln |V(G)| + 1$.

2.1 Inapproximability of MCLP on monochromatic- K_4^- -free graphs

Li and Zhang [102] have proved that MCLP is NP-complete, even when the input is restricted to K_4^- -free graphs. They also presented a polynomial algorithm to find an approximate solution for MCLP in K_4^- -free graphs with performance ratio $\ln m + 1$, where m is the size of a maximum monochromatic clique in the input graph. Actually, the algorithm works for monochromatic- K_4^- -free graphs instead of K_4^- -free graphs.

Hence, if the input graph for MCLP is monochromatic- K_4^- -free with the size of a maximum monochromatic clique bounded by a constant, we have an approximation algorithm with constant performance ratio.

We further investigate the inapproximability of MCLP. Alimonti and Kann [5] have shown that the Vertex Cover problem restricted to 3-regular connected graphs is APX-complete. This implies that there is some small $\epsilon > 0$ such that the existence of a polynomial time approximation algorithm for finding a minimum cardinality vertex cover in a connected 3-regular graph with performance guarantee $1 + \epsilon$ would imply $P = NP$. We now give an approximation preserving L-reduction from the Vertex Cover problem in 3-regular connected graphs to MCLP and draw the following conclusions.

Theorem 2.1.1. MCLP is

- (1) APX-hard on monochromatic- K_4^- -free graphs, and
- (2) APX-complete on monochromatic- K_4^- -free graphs with the size of a maximum monochromatic clique bounded by a constant.

Proof. Consider an arbitrary instance of the Vertex Cover problem in 3-regular connected graphs. So let $G = (V, E)$ be a 3-regular connected graph with $|V| = 2n$ and $|E| = 3n$ for some $n \in \mathbb{Z}^+$. A corresponding MCLP instance on an edge-colored graph $H = (V_H, E_H)$ is constructed from G in the following way. H is obtained from G by replacing every edge $(u, v) \in E$ by a gadget $g(u, v)$ consisting of the vertices u and v as well as two new vertices $e_{u,v}^1$ and $e_{u,v}^2$, and the edges $(u, e_{u,v}^1)$, $(u, e_{u,v}^2)$, $(v, e_{u,v}^1)$ and $(v, e_{u,v}^2)$. Furthermore, for a vertex u with neighbors v, w and x , the vertices $e_{u,v}^1$, $e_{u,w}^1$ and $e_{u,x}^1$ are made mutually adjacent in H . For every vertex $u \in V$, we define a color $\ell(u)$, and for every edge $(u, v) \in E$, we define two colors $\ell(u, e_{u,v})$ and $\ell(v, e_{u,v})$, where

all the colors we define are different. For an edge $(u, v) \in E$, the corresponding edges in H are colored as follows. The edge $(u, e_{u,v}^1)$ is assigned color $\ell(u)$ and the edge $(v, e_{u,v}^1)$ is assigned color $\ell(v)$. The edge $(u, e_{u,v}^2)$ is assigned color $\ell(u, e_{u,v})$ and the edge $(v, e_{u,v}^2)$ is assigned color $\ell(v, e_{u,v})$. For a vertex u with neighbors v, w and x in G , the edges $(e_{u,v}^1, e_{u,w}^1)$, $(e_{u,w}^1, e_{u,x}^1)$ and $(e_{u,x}^1, e_{u,v}^1)$ in H are all assigned color $\ell(u)$. This completes the construction and edge-coloring of the graph H . It is easy to observe that a largest monochromatic clique in H corresponds to a K_4 , and that H is monochromatic- K_4^- -free, with maximum degree 6. Note that the degree of the vertices u, v , and $e_{v,u}^1$ is exactly 6 for every gadget $g(u, v)$.

Let V_c^* be a minimum vertex cover of G , and let P^* be a minimum monochromatic clique partition of H . Since every vertex in G is incident with exactly three edges, V_c^* has at least $|E|/3 = n$ vertices. There are $|V| + 2|E| = 8n$ vertices in H , so H can be partitioned into $8n$ vertex-cliques. Hence, $|P^*| \leq 8n \leq 8|V_c^*|$.

Suppose P is an arbitrary monochromatic clique partition of H . We claim that P can always be turned into a new monochromatic clique partition P' such that $|P'| \leq |P|$ and for every edge (u, v) of G , $e_{u,v}^1 \in K(u)$ or $e_{u,v}^1 \in K(v)$ holds and there is no vertex-clique u or v in P' . Here $K(v)$ denotes a vertex-clique v or a (nontrivial) monochromatic clique containing v . We now prove this claim below.

First suppose that $K(e_{u,v}^1)$ is a vertex-clique in P , or is a monochromatic clique with color $\ell(u)$ that does not contain u . We can execute one of the following operations on P to merge $K(e_{u,v}^1)$ into $K(u)$ or u into $K(e_{u,v}^1)$, without increasing the cardinality of P . If u forms a vertex-clique or is contained in a monochromatic clique with color $\ell(u)$, then $K(e_{u,v}^1)$ can be combined with $K(u)$ to obtain a larger monochromatic clique with color $\ell(u)$. If u is contained in a monochromatic clique with a color different from $\ell(u)$, then u can be taken away from $K(u)$ and combined with $K(e_{u,v}^1)$ to form a new clique with color $\ell(u)$.

Therefore, we may assume that $e_{u,v}^1 \in K(u)$ or $K(v)$ for all edges (u, v) in G . If there exists a vertex-clique v in a gadget $g(u, v)$ after executing the above operations, then $e_{u,v}^1 \in K(u)$, and $e_{u,v}^2$ forms a vertex-clique. Hence, v can be combined with $e_{u,v}^2$ to form a new monochromatic clique with color $\ell(v, e_{u,v})$, and the cardinality of the partition is decreased.

After applying the above operations we have obtained a new partition P' , with $|P'| \leq |P|$, satisfying the conditions claimed.

Let $g(u, v)$ be a gadget in H . Without loss of generality we may assume that $e_{u,v}^1 \in K(u)$ with color $\ell(u)$ in P' . Then $e_{u,v}^2$ forms either a vertex-clique, or a clique with v of color $\ell(v, e_{u,v}^2)$.

Let be given a set V_c which is composed of all the vertices $u \in V$ such that for some edge $(u, v) \in E$, $e_{u,v}^1 \in K(u)$ in P' . Since for every edge $(u, v) \in E$, $e_{u,v}^1 \in K(u)$ or $e_{u,v}^1 \in K(v)$ in P' , at least one of u and v is in V_c . Hence V_c is a vertex cover of G .

For every edge $(u, v) \in E$, $e_{u,v}^2$ forms either a vertex-clique or a monochromatic clique together with v or u in P' . There are totally $|E|$ such cliques. Each of the other cliques in P' contains exactly one vertex $u \in V$ and at least one vertex $e_{u,v}^1$ for some neighbor v of u in G , and hence corresponds to a vertex $u \in V_c$. Consequently,

$$|V_c| = |P'| - |E| = |P'| - 3n \leq |P| - 3n. \quad (1)$$

On the other hand, we can obtain a monochromatic clique partition \tilde{P} of H from a minimum vertex cover V_c^* of G , as follows. For a gadget $g(u, v)$ in H , if (u, v) is covered by exactly one end vertex in V_c^* , say u , then let $e_{u,v}^1$ be in the same clique with u in \tilde{P} , and hence $K(u)$ is of color $\ell(u)$. If (u, v) is covered by both u and v , then let $e_{u,v}^1$ be in the same clique with either u or v in \tilde{P} arbitrarily. Since every edge is covered by at least one vertex, every vertex of type $e_{u,v}^1$ is contained in either $K(u)$ or $K(v)$. We claim that for every vertex $u \in V_c^*$, $K(u)$ contains at least one vertex of type $e_{u,v}^1$. For if there exists a $u_0 \in V_c^*$ such that $K(u_0)$ contains no vertex $e_{u_0,v}^1$ for every neighbor v of u_0 in G , then $V_c^* \setminus \{u_0\}$ is a vertex cover of G with cardinality less than V_c^* , contradicting the minimality of V_c^* . Since no two vertices in G can be in the same clique in \tilde{P} , there are exactly $|V_c^*|$ cliques in \tilde{P} containing vertices of type $e_{u,v}^1$.

For any vertex v of G that is not contained in V_c^* , let v form a clique in \tilde{P} with a vertex $e_{u,v}^2$ for a neighbor u of v in G . Note that such a vertex $e_{u,v}^2$ is always available for v , since any neighbor of v in G must be in V_c^* .

Finally, we let the remaining vertices of type $e_{u,v}^2$ be vertex-cliques in \tilde{P} . \tilde{P} consists of $|E|$ cliques containing vertices of type $e_{u,v}^2$, and $|V_c^*|$ cliques containing vertices of type $e_{u,v}^1$, therefore

$$|V_c^*| = |\tilde{P}| - |E| = |\tilde{P}| - 3n. \quad (2)$$

Thus \tilde{P} is a minimum monochromatic clique partition; for if there exists a monochromatic clique partition P with $|P| < |\tilde{P}|$, then by the above discussion we can always obtain a vertex cover V_c of G with $|V_c| \leq |P| - 3n < |\tilde{P}| - 3n = |V_c^*|$, contradicting the minimality of V_c^* .

Given an instance of the Vertex Cover problem restricted to 3-regular connected graphs, we can turn it into an instance of MCLP. We can assume that the monochromatic clique partition P we find satisfies the condition that every vertex of type $e_{u,v}^1$ is contained in $K(u)$ or in $K(v)$, and that there is no vertex-clique u or v in P . Then from P we can obtain a solution V_c for the instance of the Vertex Cover problem, in the way we discussed above. We keep using the notations P^* and V_c^* to denote the optimal solutions for both problems. We have $|P| = |V_c| + 3n$ and $|P^*| = |V_c^*| + 3n$. Further, using $|P^*| \leq 8|V_c^*|$ we have $3n \leq 7|V_c^*|$.

Suppose there exists a small positive ϵ such that $|P| \leq (1 + \epsilon)|P^*|$. Substituting V_c and V_c^* into the inequality, we get $|V_c| + 3n \leq (1 + \epsilon)(|V_c^*| + 3n)$, that is, $|V_c| \leq (1 + \epsilon)|V_c^*| + 3n\epsilon \leq (1 + \epsilon)|V_c^*| + 7\epsilon|V_c^*| = (1 + 8\epsilon)|V_c^*|$.

Therefore, the existence of a polynomial time approximation scheme for MCLP restricted to K_4^- -free and monochromatic- K_4^- -free graphs with maximum monochromatic clique K_4 would imply the existence of a polynomial time approximation scheme for the Vertex Cover problem restricted to 3-regular connected graphs. Since the latter problem is APX-complete, we have that MCLP is APX-hard on K_4^- -free graphs, and APX-hard on monochromatic- K_4^- -free graphs. Finally, we have the algorithm from [102] that works out a solution with a constant approximation ratio for MCLP in monochromatic- K_4^- -free graphs in which the size of a maximum monochromatic clique is bounded by a constant. Therefore, statement (2) holds. \square

2.2 An approximation algorithm for WMCLP

We generalize MCLP to its weighted version WMCLP. Let G be an edge-colored graph with colors $\ell(G)$. Each color $c \in \ell(G)$ is associated with a non-negative cost $w(c)$. Every monochromatic clique CL of G with at least two vertices has the same non-negative cost as its color, denoted by $w(CL)$. As any vertex v of G is viewed as a degenerate monochromatic clique, we also assign it a non-negative cost $w(v)$, with $w(v) \leq \min\{w(c) | c \in \ell(G)\}$. WMCLP asks for a monochromatic clique partition such that the sum of the costs of all

cliques in the partition is minimal among all the possible partitions. Obviously, MCLP is the special case of WMCLP in which all the costs are equal to 1.

Li and Zhang [102] presented a polynomial algorithm, denoted by $\text{Alg}(\text{clique})$, which was claimed to calculate all maximal monochromatic cliques in a K_4^- -free graph and return a maximum one. Actually, their claim holds when K_4^- -free is replaced by monochromatic- K_4^- -free.

In this chapter, we use $\text{Alg}(\text{clique})$ to find all the maximal monochromatic cliques by applying an $(\ln |V(G)| + 1)$ -approximation algorithm (Algorithm 1) for solving WMCLP restricted to monochromatic- K_4^- -free graphs. In Algorithm 1, $\text{Alg}(\text{clique})$ is implemented from Step 2 to Step 9.

Let G be a monochromatic- K_4^- -free graph, and let CL_1 and CL_2 be two distinct maximal monochromatic cliques in G . Suppose that there is at least one common edge (u, v) of CL_1 and CL_2 . Since CL_1 and CL_2 are maximal, there must be at least one vertex $w \in V(CL_1) \setminus V(CL_2)$ and one vertex $x \in V(CL_2) \setminus V(CL_1)$ such that w and x are nonadjacent. But then u, v, w and x span a monochromatic- K_4^- in G , a contradiction. Therefore, any two distinct maximal monochromatic cliques in a monochromatic- K_4^- -free graph do not share a common edge.

We note that MCLP can be considered as a variant of the Set Cover problem, in which the (possibly exponentially many) subsets are the vertex sets of all the monochromatic cliques and vertex-cliques of the input graph G , and the objective is to find a minimum collection of pairwise disjoint subsets covering the vertex set of G . Hence, it is natural that our design of a greedy approximation algorithm for WMCLP is inspired by the greedy algorithm for the weighted Set Cover problem in [33].

Theorem 2.2.1. Algorithm 1 runs in polynomial time and achieves the performance ratio $\ln |V(G)| + 1$ for WMCLP on a monochromatic- K_4^- -free graph G .

Proof. In this proof, we do not distinguish between a clique and its vertex set.

First we claim that the set C contains all maximal monochromatic cliques after the execution of the loop from Step 2 to Step 9 in Algorithm 1. Since any two maximal monochromatic cliques do not share an edge in G , every edge belongs to one maximal monochromatic clique. Hence, we can start from the end vertices of any edge, and find out the maximal monochromatic clique

Algorithm 1 An approximation algorithm for WMCLP on monochromatic- K_4^- -free graphs

Input: A monochromatic- K_4^- -free graph G ;

Output: A monochromatic vertex-disjoint clique partition D of G ;

- 1: Let $C := \emptyset, D := \emptyset$;
 - 2: **repeat**
 - 3: Select an edge $(v_i, v_j) \in E(G)$;
 - 4: Let $S := \{v_i, v_j\}$;
 - 5: **while** there is a vertex v_k which is color-adjacent to each vertex of the monochromatic clique $G[S]$ **do**
 - 6: $S = S \cup \{v_k\}$;
 - 7: **end while**
 - 8: $C = C \cup S, E(G) = E(G) \setminus E(S)$, where $E(S)$ denotes the edges of G with both end vertices in S ;
 - 9: **until** no edge in $E(G)$.
 - 10: Let $Q := V(G) \cup C$;
 - 11: **repeat**
 - 12: Pick $q \in Q$ such that the ratio $w(q)/|q|$ is minimum, where $w(q)$ denotes the weight of the monochromatic clique $G[q]$;
 - 13: Let $C' := \emptyset$;
 - 14: **for all** $c \in C$ **do**
 - 15: $c = c \setminus q, C' = C' \cup c$;
 - 16: **end for**
 - 17: $D = D \cup \{q\}, V(G) = V(G) \setminus q, Q = V(G) \cup C', C = C'$;
 - 18: **until** $V(G) = \emptyset$.
 - 19: **return** D .
-

containing the edge through the loop from Step 5 to Step 7. Then, all edges of this clique are removed from $E(G)$. Repeating this process until $E(G)$ becomes empty, all maximal monochromatic cliques of G are found. The running time of the loop from Step 2 to Step 9 is at most $O(|E||V|^2) = O(|V|^4)$.

Assume that the loop from Step 11 to Step 17 is iterated r times. Let the vertex-clique or maximal monochromatic clique picked in Step 12 at the i -th iteration of the loop be q_i , for $1 \leq i \leq r$. Let $G = G_1$ and $G_{i+1} = G_1 \setminus \{q_1 \cup q_2 \cup \dots \cup q_i\} = G_i \setminus q_i$, for $1 \leq i \leq r - 1$. The algorithm outputs $D = \{q_i, 1 \leq i \leq r\}$ as a solution.

It is easy to prove by induction that Q contains all the vertex-cliques and

maximal monochromatic cliques of G_{i+1} at the i -th iteration after the execution of Step 16, for $1 \leq i \leq r-1$. We denote by P_i an optimal monochromatic clique partition of G_i and $w(P_i)$ the cost of P_i , for $1 \leq i \leq r$. Note that P_1 is an optimal solution of the problem. Let j be an integer such that $1 \leq j \leq r-1$. Let the number of cliques in P_j be t , and $P_j = \{p_{j1}, p_{j2}, \dots, p_{jt}\}$. Then $P'_j = \{p_{ji} \setminus q_j : 1 \leq i \leq t\}$ is a monochromatic clique partition of G_{j+1} . Hence, $w(P_{j+1}) \leq w(P'_j) = \sum_{i=1}^t w(p_{ji} \setminus q_j)$. Furthermore, for each $p_{ji} \in P_j$, if $|p_{ji} \setminus q_j| \geq 2$, $w(p_{ji}) = w(p_{ji} \setminus q_j)$; otherwise $w(p_{ji}) \geq w(p_{ji} \setminus q_j)$. Therefore, $w(P_{j+1}) \leq w(P'_j) \leq \sum_{i=1}^t w(p_{ji}) = w(P_j)$.

Note that for any monochromatic clique q and $q' \subseteq q$ with $|q'| \geq 2$, the relation $\frac{w(q)}{|q|} \leq \frac{w(q')}{|q'|}$ holds. Therefore, the clique q_i picked in Step 12 at the i -th iteration has the minimum ratio $\frac{w(q_i)}{|q_i|}$ over all vertex-cliques and monochromatic cliques of G_i . So we have $\frac{w(q_i)}{|q_i|} \leq \frac{w(p_{ij})}{|p_{ij}|}$, for $1 \leq j \leq t$. Therefore,

$$w(P_i) = \sum_{j=1}^t \frac{|w(p_{ij})|}{|p_{ij}|} |p_{ij}| \geq \frac{w(q_i)}{|q_i|} \sum_{j=1}^t |p_{ij}| = \frac{w(q_i)}{|q_i|} |V(G_i)|,$$

or, $w(q_i) \leq \frac{w(P_i)|q_i|}{|V(G_i)|}$. Hence,

$$\begin{aligned} w(D) &= \sum_{i=1}^r w(q_i) \leq \sum_{i=1}^r \frac{w(P_i)|q_i|}{|V(G_i)|} \leq w(P_1) \sum_{i=1}^r \frac{|q_i|}{|V(G_i)|} \\ &\leq w(P_1) \sum_{i=1}^r \sum_{k=0}^{|q_i|-1} \frac{1}{|V(G_i)| - k} = w(P_1) H(|V(G)|) \\ &\leq w(P_1) (\ln |V(G)| + 1), \end{aligned}$$

where $H(|V(G)|)$ is the $|V(G)|$ -th harmonic number. \square

To show that the above approximation ratio is tight, we present an example to demonstrate that the approximation algorithm may find a solution with cost $H(n)$ times the optimum, where n is the number of vertices of the graph.

Let G be an edge-colored complete graph with vertices v_1, v_2, \dots, v_n . And let the cost of every vertex-clique v_i be $\frac{1}{i+\epsilon}$ for $i = 1, 2, \dots, n$, where ϵ is a very small positive number. All edges $e \in E(G)$ have the same color of cost 1. It is not difficult to verify that Algorithm 1 finds a solution consisting of all vertex-cliques in the order v_n, v_{n-1}, \dots, v_1 , with total cost $\sum_{i=1}^n \frac{1}{i+\epsilon}$, and

hence arbitrarily close to $H(n)$, whereas the optimal solution picks G directly, with cost 1.

2.3 RCYP is NP-complete for triangle-free graphs

We first consider several trivial cases of RCYP. Let G be an edge-colored graph on n vertices. If G is colored with a small number of colors, say $\ell(G) = 1$ or $\ell(G) = 2$, then we can only partition G into n vertex-cycles. When G is colored by the largest possible number of colors, that is, $\ell(G) = |E(G)|$, then finding the minimum rainbow cycle partition of G is at least as hard as finding a Hamiltonian cycle of G , which is well-known to be NP-complete. For $\ell(G) \geq 3$, Li and Zhang [102] showed that RCYP is NP-complete, even if the input graph G is K_4^- -free. Their proof is based on a reduction from the Exact Cover By 3-Sets problem. We achieve a further strengthening by showing that RCYP is NP-complete, even if the input graphs are restricted to triangle-free graphs, a proper subclass of K_4^- -free graphs.

Our proof is based on a reduction from the Minimum Set Cover problem, which is well-known to be NP-complete. In an instance of the Minimum Set Cover problem, a universe U of n elements, a collection of subsets of U , $S = \{s_1, \dots, s_m\}$ where $\cup_{i=1}^m s_i = U$, and a positive integer $k \leq \min\{m, n\}$ are given. The question is whether there exists a subcollection C of S with $|C| \leq k$ that covers all the elements of U .

Theorem 2.3.1. RCYP is NP-complete when restricted to triangle-free graphs.

Proof. RCYP on triangle-free graphs is clearly in NP: a nondeterministic algorithm needs only to guess a set of cycles of the input graph, and to check in polynomial time whether the cycles in the set are vertex-disjoint rainbow cycles that cover all vertices of the graph, and whether the number of cycles in the set is not larger than a given positive number.

Suppose now that we are given an instance of the Minimum Set Cover problem, with the universe $U = \{u_i | 1 \leq i \leq n\}$ and the subset collection $S = \{s_j | 1 \leq j \leq m\}$, where $\cup_{i=1}^m s_i = U$. We construct an edge-colored triangle-free graph G as follows.

The vertex set of G is the union of the sets $U_i = \{u_i^1, u_i^2\} \cup \{u_{i_j}^1, u_{i_j}^2 | 1 \leq j \leq m\}$, $1 \leq i \leq n$, and $S_i = \{s_{i_j} | 0 \leq j \leq n+1\}$, $1 \leq i \leq m$.

We define the following colors for the edges of G .

1. For $1 \leq i \leq n$, $1 \leq j \leq 2m + 1$, define the colors c_{ij} .
2. For $0 \leq i \leq n$, define the colors ℓ_i .
3. For $1 \leq i \leq m$, define the colors ℓ_{i1} and $\ell_{i(n+1)}$.
4. For $1 \leq i \leq m$, supposing the set s_i contains t_i elements, define the colors $d_{i1}, d_{i2}, \dots, d_{i(t_i-1)}$.

The edges of G and their colors are given below.

1. For $1 \leq i \leq m$, form the cycles $CY_i(s_{i0}, s_{i(n+1)}) = s_{i0}s_{i1} \dots s_{i(n+1)}$. Assign color ℓ_{i1} to edge (s_{i0}, s_{i1}) , and color $\ell_{i(n+1)}$ to edge $(s_{i0}, s_{i(n+1)})$. Each of the other n edges on the cycle is assigned a color from the set $\{c_{j1} : u_j \notin s_i, 1 \leq j \leq n\} \cup \{\ell_0\} \cup \{d_{ij} : 1 \leq j \leq t_i - 1\}$, so that no two edges have the same color. Note that we have exactly n colors in the color set, because t_i is the number of u_j 's contained in s_i . We denote by $P_i(s_{i1}, s_{i(n+1)})$ the path $s_{i1}s_{i2} \dots s_{i(n+1)}$.
2. For $1 \leq i \leq n$, form the cycles $CY_i(u_i^1, u_i^2) = u_i^1u_{i1}^1u_{i2}^1 \dots u_{im}^1u_{i1}^2u_{i2}^2 \dots u_{im}^2u_i^1$. Assign color ℓ_i to the edge (u_i^1, u_i^2) . Denote the paths $u_i^1u_{i1}^1u_{i2}^1 \dots u_{im}^1u_{i1}^2u_{i2}^2 \dots u_{im}^2u_i^2$ by $P_i(u_i^1, u_i^2)$. Assign colors c_{ij} , $1 \leq j \leq 2m + 1$ to the $2m + 1$ edges of $P_i(u_i^1, u_i^2)$ successively.
3. For $1 \leq j \leq n$, $1 \leq i \leq m$, if s_i contains u_j , join u_j^1 and s_{i1} by an edge and assign color ℓ_{i1} to the edge, and join u_j^2 and $s_{i(n+1)}$ and assign color $\ell_{i(n+1)}$ to the edge.
4. For $1 \leq j < i \leq n$, if there exists a set s_p , $1 \leq p \leq m$, containing both u_i and u_j , then join u_j^1 and u_i^2 , and u_j^2 and u_i^1 , and assign color ℓ_j to the edges.

It is easy to verify that G is triangle-free and that the construction can be accomplished in polynomial time.

Given a positive integer $k \leq \min\{m, n\}$, we prove that there is a covering $C \subseteq S$ of U with no more than k subsets, if and only if G has a rainbow cycle partition P with $k' + m \leq k + m$ rainbow cycles.

The following properties of P are crucial for our proof.

1. For $1 \leq i \leq m$, the vertex s_{i0} either forms a vertex-cycle, or is contained in the cycle $CY_i(s_{i0}, s_{i(n+1)})$ of P , and hence we have m cycles in P , each containing one s_{i0} . To see this, suppose that s_{i0} is contained in a rainbow cycle CY . CY must contain the edges (s_{i0}, s_{i1}) and $(s_{i(n+1)}, s_{i0})$. Since every edge associated with s_{i1} has color ℓ_{i1} except the edge (s_{i1}, s_{i2}) , CY must contain (s_{i1}, s_{i2}) . Similarly CY must contain the edge $(s_{in}, s_{i(n+1)})$ and hence $CY = CY_i(s_{i0}, s_{i(n+1)})$.
2. For $1 \leq i \leq m$, the path $P_i(s_{i1}, s_{i(n+1)})$ is contained in a rainbow cycle; otherwise, the vertices $s_{i0}, s_{i1}, \dots, s_{i(n+1)}$ must form $n+2$ vertex-cycles in P . However, we have m rainbow cycles to cover all s_{i0} by Property 1. Together we need at least $m+n+2-1 = m+n+1 > k' + m$ cycles in P , a contradiction.
3. For any $1 \leq i \neq j \leq m$, the two paths $P_i(s_{i1}, s_{i(n+1)})$ and $P_j(s_{j1}, s_{j(n+1)})$ cannot be contained in the same rainbow cycle, since both of them have an edge of color ℓ_0 .
4. For $1 \leq i \leq n$, the path $P_i(u_i^1, u_i^2)$ is contained in a rainbow cycle; otherwise, the $2m$ vertices u_{ij}^1 and u_{ij}^2 , $j = 1, \dots, m$ would form $2m$ vertex-cycles in P . Then, there would be at least $2m+1 > k' + m$ cycles in P , a contradiction.
5. If there is a rainbow cycle CY in P containing only vertices in $\cup_{j=1}^n U_j$, then $CY = CY_i(u_i^1, u_i^2)$, for some $1 \leq i \leq n$. To see this, let j_0 be the smallest index such that CY contains some vertices in U_{j_0} . By Property 4, CY contains the paths $P_{j_0}(u_{j_0}^1, u_{j_0}^2)$. If CY contains some more vertices in $\cup_{j=1}^n U_j$, then there must be two edges joining $P_{j_0}(u_{j_0}^1, u_{j_0}^2)$ to the other part of CY , which must be associated with $u_{j_0}^1$ and $u_{j_0}^2$. However, both edges have the same color ℓ_{j_0} , contradicting the rainbow property of CY . Therefore $CY = CY_{j_0}(u_{j_0}^1, u_{j_0}^2)$.
6. If s_j does not contain u_i , then the path $P_i(u_i^1, u_i^2)$ cannot be in a rainbow cycle which contains the path $P_j(s_{j1}, s_{j(n+1)})$. The reason is that the first edge of $P_i(u_i^1, u_i^2)$ is colored c_{i1} , while there is also an edge of color c_{i1} on the path $P_j(s_{j1}, s_{j(n+1)})$.

According to the above properties, there are only four possible kinds of rainbow cycles in the partition P . The first kind consists of the cycles $CY_i(s_{i0}, s_{i(n+1)})$, for some $1 \leq i \leq m$. The second kind consists of the cycles which contain a path $P_i(s_{i1}, s_{i(n+1)})$ and several paths $P_j(u_j^1, u_j^2)$, where

$u_j \in s_i$ for some $1 \leq i \leq m$ and $1 \leq j \leq n$. The third kind are the vertex-cycles formed by s_{i0} for some $1 \leq i \leq m$. The fourth kind consists of the cycles $CY_i(u_i^1, u_i^2)$, for some $1 \leq i \leq n$.

Now suppose G has a rainbow cycle partition P with $k' + m \leq k + m$ vertex-disjoint rainbow cycles. By Property 1, there must be m cycles of the first or third kind in P , each covering one s_{i0} , for $1 \leq i \leq m$. Let t and t' denote the number of rainbow cycles in P of the second kind and the fourth kind, respectively. Then $t + t' = k' \leq k$. Every rainbow cycle of the second kind contains a path $P_i(s_{i1}, s_{i(n+1)})$ for some $s_i \in S$; there are t such s_i 's. For every rainbow cycle $CY_i(u_i^1, u_i^2)$ of the fourth kind, we can always find a subset s_j in S containing u_i , since $\cup_{j=1}^m s_j = U$; there are at most t' such s_j 's. The subcollection of S composed of the s_i 's and s_j 's covers U , and has at most $t + t' \leq k' \leq k$ elements.

Conversely, let there be a covering C of U with $k' \leq k$ subsets. Without loss of generality, let $C = \{s_1, \dots, s_k\}$. Let $s'_1 = s_1$, and $s'_i = s_i \setminus \cup_{j=1}^{i-1} s_j$, $i \geq 2$. Then s'_1, \dots, s'_k are k' disjoint subsets (some of which may be empty) whose union covers U . For every $1 \leq i \leq k'$, if $s'_i \neq \emptyset$, take a rainbow cycle of the second kind constituted by the path $P_i(s_{i1}, s_{i(n+1)})$ and the paths $P_j(u_j^1, u_j^2)$ for all $u_j \in s'_i$. Since $\cup_{i=1}^{k'} s'_i$ covers U , all vertices of G in U_i , $1 \leq i \leq n$, are covered by these rainbow cycles. Finally take m rainbow cycles of the first kind $CY_i(s_{i0}, s_{i(n+1)})$ or the third kind s_{i0} , covering all vertices of G that are left. Then a partition of at most $k' + m \leq k + m$ rainbow cycles for G is obtained. \square

2.4 Concluding remarks

In this chapter we obtained results on the inapproximability and complexity of MCLP and RCYP restricted to graphs avoiding some induced subgraphs. We also presented a $(1 + \ln |V(G)|)$ -approximation algorithm for WMCLP restricted to monochromatic- K_4^- -free graphs. A natural suggestion for further research is to consider possible approximation algorithms for RCYP or its weighted version on triangle-free graphs, and alternative algorithms that might improve the approximation ratio $(1 + \ln |V(G)|)$ for WMCLP on monochromatic- K_4^- -free graphs. Another interesting direction is to study the computational complexity of similar problems, e.g., on minimum monochromatic or rainbow path and tree partitions, restricted to graphs avoiding some

induced subgraphs.

Chapter 3

On the complexity of injective coloring

In this chapter, we prove that it is NP-complete to determine the injective chromatic number even when we restrict the problem to some special bipartite graphs. Furthermore, we show that for every $\epsilon > 0$, it is impossible to efficiently approximate the injective chromatic number of bipartite graphs within a factor of $n^{\frac{1}{3}-\epsilon}$ unless $ZPP = NP$. For the max-injective coloring problem, we prove that there is a constant approximation algorithm on power chordal graphs with bounded injective chromatic number. We also devise a constant approximation algorithm for max-injective coloring some bipartite graphs. For the on-line injective coloring problem, we prove that First Fit (FF) injectively colors P_3 -free graphs optimally, where First Fit is an on-line algorithm that simply assigns the smallest eligible color to each vertex. We also prove that the number of colors used by FF* for bipartite graph G is bounded by $\frac{3}{2}$ times the on-line injective chromatic number, where FF* is an on-line algorithm equivalent to FF proper coloring the complement of G . Moreover, we present an improved algorithm BFF, and prove that it is optimal for on-line injectively coloring bipartite graphs.

3.1 Off-line injective coloring

3.1.1 NP-hardness of injective coloring bipartite graphs

Recall the problem of injective k -coloring a graph (see Section 1.1.2). In this section, we show that the problem is NP-hard even when it is restricted to some special class of bipartite graphs.

Problem A

Instance: A balanced bipartite graph $G = (U, V; E)$ where $|U| = |V|$ and a positive integer k .

Question: Is there an injective k -coloring of G ?

Theorem 3.1. *The Problem A is NP-complete.*

Proof. It is easy to see **Problem A** is in NP. We use a reduction from the problem of injective coloring a general graph $G = (V; E)$ which has been proved to be NP-complete [68]. Let k be a positive integer. We will construct a bipartite graph \tilde{G} with the property that $\chi_i(G) \leq k$ if and only if $\chi_i(\tilde{G}) \leq k$. Here $\chi_i(G)$ denotes the injective chromatic number (see Section 1.1.2).

Let $G' = (V'; E')$ be a copy of G , where $V' = \{u' : u \in V\}$ and $E' = \{u'v' : uv \in E\}$. We construct a bipartite graph \tilde{G} with bipartite sets V and V' such that both $uv' \in E(\tilde{G})$ and $u'v \in E(\tilde{G})$ iff $uv \in E(G)$. The graph \tilde{G} can clearly be constructed in polynomial time.

Let $c : V \rightarrow [k]$ be an injective k -coloring of G . We define a coloring $\tilde{c} : V(\tilde{G}) \rightarrow [k]$ as follows. Set $\tilde{c}(u) = c(u)$ for each $u \in V$, and $\tilde{c}(u') = \tilde{c}(u)$ for each $u' \in V'$ (recall u' of V' is the corresponding vertex of u). Let u and v be two vertices in \tilde{G} that have a common neighbor x . By the definition of \tilde{G} , $u, v \in V$ and $x \in V'$, or, $u, v \in V'$ and $x \in V$. In the first case, let w be the vertex in V corresponding to x . Then, uw and $vw \in E$. In the second case, let w be the vertex in V' corresponding to x . Then, uw and $vw \in E'$. Therefore, $\tilde{c}(u) \neq \tilde{c}(v)$, and hence \tilde{c} is an injective coloring of \tilde{G} .

Conversely, let \tilde{c} be an injective k -coloring of \tilde{G} . Since any two vertices of V that have a common neighbor in G must have a common neighbor in \tilde{G} , and \tilde{c} is injective, the restriction of \tilde{c} to V injectively colors G with not more than k colors. \square

Problem B

Instance: A bipartite graph $G = (U, V; E)$ with $|U| \geq |V|$, where G has an induced balanced complete bipartite subgraph with order $2|V|$, and a positive integer k .

Question: Is there an injective k -coloring of G ?

Theorem 3.2. *The **Problem B** is NP-complete.*

Proof. It is easy to see that **Problem B** is in NP. We use a reduction from the problem of proper coloring a graph $G = (V; E)$. We will construct a bipartite graph \tilde{G} with the required property such that $\chi(G) \leq k$ if and only if $\chi_i(\tilde{G}) \leq k + |E|$.

Let $V(\tilde{G}) = V \cup V_E \cup V'_E$, where $V_E = \{x_{uv} : uv \in E\}$, and $V'_E = \{y_{uv} : uv \in E\}$. Let $E(\tilde{G}) = E_A \cup E_B$, where $E_A = \{ux_{uv}, vx_{uv} : uv \in E\}$, and $E_B = \{x_{uv}y_{st} : uv \in E, st \in E\}$.

\tilde{G} is obtained from G by the following two steps. Firstly, we get a bipartite graph $G' = (V, V_E)$ by subdividing each edge of the graph G ; secondly, $\tilde{G} = (V \cup V'_E, V_E)$ is obtained from G' by adding $|E|$ new vertices and $|E|^2$ edges to ensure that $\tilde{G}[V'_E \cup V_E]$ is a complete bipartite graph. This complete bipartite subgraph guarantees that any two vertices of V'_E have common neighbors in V_E , and that each vertex of V'_E and each vertex of V have a common neighbor in V_E (such as u in V and y in V'_E have a common neighbor x_{uv} in V_E). The graph \tilde{G} can clearly be constructed in polynomial time.

Let $c : V \rightarrow [k]$ be a proper k -coloring of G . We define an injective coloring $\tilde{c} : V(\tilde{G}) \rightarrow [k] \cup E$, using $k + |E|$ colors as follows. Set $\tilde{c}(u) = c(u)$ for each $u \in V$, and $\tilde{c}(x_{uv}) = \tilde{c}(y_{uv}) = uv$ for each $x_{uv} \in V_E$ and each $y_{uv} \in V'_E$. It is easy to see that all vertices in V_E have distinct colors (so does V'_E), and that vertices in different bipartite sets have no common neighbor. We only need to consider the vertices in $V \cup V'_E$. Since the vertices of V with common neighbor in \tilde{G} are adjacent in G , and since c is a proper coloring of G , V is injectively colored with respect to \tilde{c} . Since the colors on vertices in V'_E are distinct from that on V , \tilde{c} is injective.

Conversely, let \tilde{c} be an injective l -coloring of \tilde{G} . We claim that the restriction of \tilde{c} on V is a proper coloring of G using no more than $l - |E|$ colors. Since $\tilde{G}[V'_E \cup V_E]$ is a complete bipartite graph, all vertices in V'_E receive distinct colors in \tilde{c} . Note that $N_{\tilde{G}}(u) \cap N_{\tilde{G}}(v) \neq \emptyset$ for each vertex $u \in V$ and each

vertex $v \in V'_E$. With respect to \tilde{c} , the vertices in V receive at most $l - |E|$ colors which all are distinct from those $|E|$ colors assigned to the vertices of V'_E . Since \tilde{c} is injective and each pair of adjacent vertices of G have a common neighbor in \tilde{G} , the restriction of \tilde{c} to V properly colors G . \square

Although the injective chromatic number of a tree can be computed efficiently [77], the problems of injective coloring some special bipartite graphs are NP-hard. Following the idea of [77], we consider the inapproximability of injective coloring bipartite graphs.

3.1.2 On the inapproximability of injective coloring bipartite graphs

A *k-fold coloring* of a graph G is an assignment of sets of size k to vertices of a graph such that adjacent vertices receive disjoint sets. A *k-fold b-coloring* is a *k-fold coloring* out of b available colors. The *k-fold chromatic number* $\chi_k(G)$ is the least b such that G admits a *k-fold b-coloring*. The *fractional chromatic number* $\chi_f(G)$ is defined to be $\chi_f(G) = \inf_k \frac{\chi_k(G)}{k}$. It has been shown that $\chi_f(G) \geq \frac{|V(G)|}{\alpha(G)}$ [96].

The *common neighbor graph* of G , denoted by $G^{[2]}$, is the graph on the vertices of G in which two vertices are adjacent if they have a common neighbor in G . It is not difficult to check that $\chi(G^{[2]}) = \chi_i(G)$.

We have observed that it is NP-hard to compute the injective chromatic number even restricted to some special classes of bipartite graphs. Now we further prove an inapproximability result which shows that it is not tractable to approximate the injective chromatic number of a bipartite graph $G = (U, V; E)$, $n = |U| + |V|$ within a factor of $n^{\frac{1}{3}-\epsilon}$ for all $\epsilon > 0$.

Theorem 3.3. *If $ZPP \neq NP$, then, for every $\epsilon > 0$, it is not possible to efficiently approximate $\chi_i(G)$ within a factor of $n^{\frac{1}{3}-\epsilon}$, for any bipartite graph G .*

Proof. In [48], it was shown that for a given graph G and any fixed $\epsilon > 0$, the problem of deciding whether $\chi(G) \leq n^\epsilon$ or $\alpha(G) < n^\epsilon$ is not possible to be solved efficiently if $ZPP \neq NP$.

Let G be a connected graph. Let \tilde{G}_k be the bipartite graph constructed from k copies of \tilde{G} which is defined in the proof of Theorem 3.2 by identifying, for each $uv \in E(G)$, all copies of x_{uv} , and all copies of y_{uv} . That is, if v_1, v_2, \dots, v_n are the vertices of G , then $V(\tilde{G}_k) = \bigcup_{i=1}^k \{v_1^i, v_2^i, \dots, v_n^i\} \cup \{x_{uv}, y_{uv} : uv \in E(G)\}$, $E(\tilde{G}_k) = \bigcup_{i=1}^k \{u^i x_{uv}, v^i x_{uv} : uv \in E(G)\} \cup \{x_{uv} y_{st} : uv, st \in E(G)\}$.

If $uv \in E(G)$, then $\tilde{G}_k^{[2]}[\bigcup_{i=1}^k u^i \cup \bigcup_{i=1}^k v^i]$ is a clique. Hence it follows that from any coloring of $\tilde{G}_k^{[2]}$ one can construct a fractional k -fold coloring of G by projecting each color class of $\tilde{G}_k^{[2]}$ to G , i.e., mapping each u^i to u . So $\chi(\tilde{G}_k^{[2]}) \geq \chi_k(G) + e(G)$ ($e(G) = |E(G)|$). Since $\chi_f(G) = \inf_k \frac{\chi_k(G)}{k}$, we then obtain the following inequalities.

$$\frac{k \cdot n}{\alpha(G)} + e(G) \leq k \cdot \chi_f(G) + e(G) \leq \chi_k(G) + e(G) \leq \chi(\tilde{G}_k^{[2]}) = \chi_i(\tilde{G}_k) \leq k \cdot \chi(G) + e(G).$$

Therefore if $\chi(G) \leq n^\epsilon$ then $\chi_i(\tilde{G}_k) \leq k \cdot n^\epsilon + e(G)$, and if $\alpha(G) < n^\epsilon$ then $\chi_i(\tilde{G}_k) > k \cdot n^{1-\epsilon} + e(G)$. Now we fix $k = e(G)$, and denote by N the number of vertices in \tilde{G}_k . For $n \geq 2^{1/\epsilon}$, we obtain the following results.

$$\frac{e(G) \cdot n^{1-\epsilon} + e(G)}{e(G) \cdot n^\epsilon + e(G)} \geq \frac{1}{2} n^{1-2\epsilon} \geq n^{1-3\epsilon} \geq (e(G) \cdot n + 2e(G))^{\frac{1}{3}(1-3\epsilon)} = N^{\frac{1}{3}-\epsilon}.$$

Hence if we can efficiently $N^{\frac{1}{3}-\epsilon}$ -approximate the injective coloring of \tilde{G}_k then we can decide whether $\chi(G) \leq n^\epsilon$ or $\alpha(G) < n^\epsilon$. This completes the proof. \square

3.1.3 An approximation algorithm for the max-injective coloring problem

We denote the k -th power of G by G^k which is obtained from G by joining any two vertices of distance at most k in G . A graph is chordal if each of its cycles on four or more vertices has a chord, which is an edge that is not part of the cycle but connects two vertices of the cycle. A *power chordal graph* is a graph of which all powers are chordal. Given a graph $G = (V; E)$ and a weight function $\omega : V \rightarrow \mathbf{N}$, the *max-injective coloring problem* is to find an injective vertex coloring c of G with k colors that minimizes $\sum_{i=1}^k \max_{v \in c^{-1}(i)} \omega(v)$. Here $\sum_{i=1}^k \max_{v \in c^{-1}(i)} \omega(v)$ is referred to as the weight of the injective coloring. When $\omega(v) = 1$ for all $v \in V(G)$, $\min \sum_{i=1}^k \max_{v \in c^{-1}(i)} \omega(v)$ is simply

the injective chromatic number of G . So the problem of max-injective coloring bipartite graphs is also NP-hard. Hell et al. proved that the injective chromatic number of a power chordal graph can be computed in polynomial time [77]. Now, we show that for power chordal graphs with bounded injective chromatic number, an optimal injective coloring will lead to a constant ratio approximation algorithm for max-injective coloring.

Theorem 3.4. *Let s be a positive integer, and let G be a power chordal graph with a weight function $\omega : V(G) \rightarrow \mathbf{N}$. If $\chi_i(G) \leq s$, then an optimal injective coloring of G produces a max-injective coloring for G whose weight is at most $s \cdot \text{opt}_M(G)$, where opt_M is the weight of an optimal max-injective coloring.*

Proof. Let opt_M be the weight of an optimal max-injective coloring, and let A_1, A_2, \dots, A_k be the k injective color classes in opt_M . For each subset S of $V(G)$, we use $\omega(S)$ to denote the weight of the heaviest vertex in S , where $\omega(S) = 0$ if $S = \emptyset$. Under this assumption, $\text{opt}_M = \sum_{i=1}^k \omega(A_i)$. Let C^1, C^2, \dots, C^s be the color classes of an optimal injective coloring c of G , where empty sets may appear if $\chi_i(G) < s$. Now, the weight of c is $\sum_{j=1}^s \omega(C^j)$. Let $C_i^j = A_i \cap C^j, 1 \leq i \leq k, 1 \leq j \leq s$. Then $C^j = \bigcup_{i=1}^k C_i^j, A_i = \bigcup_{j=1}^s C_i^j, \omega(C^j) = \max_{1 \leq i \leq k} \omega(C_i^j) \leq \sum_{i=1}^k \omega(C_i^j)$, and $\omega(A_i) = \max_{1 \leq j \leq s} \omega(C_i^j)$. So, $\sum_{j=1}^s \omega(C^j) \leq \sum_{j=1}^s \sum_{i=1}^k \omega(C_i^j) \leq \sum_{i=1}^k s \cdot \max_{1 \leq j \leq s} \omega(C_i^j) = \sum_{i=1}^k s \cdot \omega(A_i) = s \cdot \text{opt}_M$. \square

Motivated by [119], we obtain a constant-approximation algorithm $MBFF$ for max-injective coloring a class of special bipartite graphs, with an on-line algorithm BFF (BFF will be introduced in Section 3.2) for injective coloring bipartite graphs.

$MBFF(G, \omega)$:

1. Sort the vertices of $V(G)$ in non-increasing order of weights.
2. Present the vertices in this order to BFF .
3. Return the injective coloring produced by BFF .

A class \mathcal{G} of graphs is called *induced-hereditary*, if \mathcal{G} is closed under taking induced subgraphs. Then we draw the following conclusion.

Theorem 3.5. *Let b be a constant and let \mathcal{G} be an induced-hereditary subclass of special bipartite graphs such that $\forall G \in \mathcal{G}, \alpha(G) \leq b\Delta(G)$. Then, for any*

$G \in \mathcal{G}$ and for any weight function $\omega : V(G) \rightarrow \mathbf{N}$, *MBFF* produces an injective coloring for G whose weight is at most $b \cdot \text{opt}_M(G)$, where opt_M is the weight of an optimal max-injective coloring.

Proof. Suppose that A_1, A_2, \dots, A_t is the injective coloring of G produced by *MBFF*. Then $t \leq \chi_i^{BFF}(G)$. Set $a_m = \max_{v \in A_m} \omega(v)$. Without loss of generality we assume that $a_1 \geq a_2 \geq \dots \geq a_t$. Then, the weight of A_1, A_2, \dots, A_t is $\sum_{m=1}^t a_m$.

Let C_1, C_2, \dots, C_k be an injective coloring of G which is an optimal max-injective coloring of G . Set $\omega_m = \max_{v \in C_m} \omega(v)$ and without loss of generality assume that $\omega_1 \geq \omega_2 \geq \dots \geq \omega_k$. Thus $k \geq \chi_i(G) \geq \Delta(G)$. Note that $\text{opt}_M(G) = \sum_{m=1}^k \omega_m$.

By Lemmas 3 and 4 given below in Section 3.2.2, $\chi_i^{BFF}(G) \leq \alpha(G) \leq b\Delta(G)$. So $t \leq \chi_i^{BFF}(G) \leq b\Delta(G) \leq b\chi_i(G) \leq bk$. Set $A_{t+1} = \dots = A_{bk} = \emptyset$, and $a_{t+1} = \dots = a_{bk} = 0$.

We will prove that $\omega_m \geq a_j$, for each m , $1 \leq m \leq k$, and each j , $b(m-1) < j \leq bk$.

For $m = 1$ and each j , $1 \leq j \leq t$, we have $\omega_1 \leq a_j$, since ω_1 is the maximum weight of any vertex in G . For any $m \geq 2$, define $V_m = \{v \mid \omega(v) > \omega_m\} \subseteq V$. The injective coloring C_1, C_2, \dots, C_k of G restricted to V_m is a $m-1$ injective coloring of $G[V_m]$. Since $\chi_i^{BFF}(G[V_m]) \leq \alpha(G[V_m]) \leq b\Delta(G[V_m]) \leq b\chi_i(G[V_m]) \leq b(m-1)$, algorithm *MBFF* injectively colors $G[V_m]$ with no more than $b(m-1)$ colors. Therefore, $\omega_m \geq a_j$, for each m , $1 \leq m \leq k$, and each j , $b(m-1) < j \leq t$.

As $\sum_{m=1}^{b \cdot k} a_m = \sum_{m=1}^k \sum_{j=b(m-1)+1}^{bm} a_j \leq \sum_{m=1}^k \sum_{j=b(m-1)+1}^{bm} \omega_m = \sum_{m=1}^k b \cdot \omega_m = b \cdot \text{opt}_M(G)$, we obtain that $\sum_{m=1}^t a_m \leq b \cdot \text{opt}_M(G)$.

□

3.2 On-line injective coloring

Let again $\alpha(G)$ denote the *independence number* of G , that is the size of a maximum independent set in G . Let $AOL(G)$ be the set of all on-line injective coloring algorithms for a graph G , and let $\Pi(G)$ denote all permutations of the vertices of G . For $A \in AOL(G)$ and $\pi \in \Pi(G)$, we use $\chi_i^A(G, \pi)$ to denote the number of colors used by A when the vertices of G are presented in the order of π . Let $\chi_i^A(G) = \max_{\pi \in \Pi(G)} \chi_i^A(G, \pi)$, and refer to $\chi_i^A(G)$ as the *A-injective*

chromatic number of G . Actually, $\chi_i^A(G)$ measures the worst case behavior of A on G .

The *on-line injective chromatic number* of G , denoted by $\chi_i^{OL}(G)$, is the smallest number of colors used by an algorithm in $AOL(G)$. Thus, $\chi_i^{OL}(G) = \min_{A \in AOL(G)} \chi_i^A(G)$. OL can be considered to be an optimal on-line algorithm for injective coloring G . Obviously, $\chi_i^{OL}(G) \geq \chi_i(G)$.

Let \mathcal{G} be a family of graphs. If $A \in AOL(G)$ for every graph $G \in \mathcal{G}$, we say that A is an *on-line injective coloring algorithm* for \mathcal{G} and write $A \in AOL(\mathcal{G})$. Let $f(x)$ be a function of a single variable x . An algorithm $A \in AOL(\mathcal{G})$ is said to be *f-competitive* for \mathcal{G} if $\chi_i^A(G) \leq f(\chi_i(G))$ for every $G \in \mathcal{G}$, i.e., the number of colors used by A on G is bounded from above by a function depending only on the injective chromatic number of G . When $f(x) = cx$ for some constant c , an *f-competitive* algorithm is called *c-competitive*, and c is referred to as the *competitive ratio*.

In fact, there are many graphs which have no competitive algorithms. Take P_{2n} for example. Any algorithm A uses at least n colors when the vertices of P_{2n} are revealed under an order π where a maximum independent set I of G emerges first, without knowledge of the coming vertices. Compared with 2 (the injective chromatic number of P_{2n}), A is not competitive for P_{2n} . These negative results have led to the definition of a weaker form of competitiveness. An algorithm $A \in AOL(\mathcal{G})$ is said to be *on-line c-competitive* if there exists a constant c such that $\chi_i^A(G) \leq c\chi_i^{OL}(G)$ for every $G \in \mathcal{G}$, and c is referred to as the *on-line competitive ratio*.

Let F be a subset of $V(G)$. The *hue* of F with respect to some coloring c of G , denoted by $H(F)$, is the set of all colors assigned to the vertices in F by c . Let $N_G(v)$ denote the neighbor set of v in G , and let $\bar{N}_G(v) = V(G) \setminus N_G(v)$.

While considering on-line injective coloring a graph of order n , every vertex must be colored differently from the previous vertex to avoid possible coming common neighbors. This leads to n colors used. So it is natural to restrict our attention to some special families of graphs. Here we focus on *P_3 -free graphs* and *triangle-free graphs*. P_3 -free graphs are those that do not contain an induced path on three vertices. In other words, in a P_3 -free graph, nonadjacent vertices have no common neighbor. Triangle-free graphs are those that do not contain an induced triangle. So, adjacent vertices have no common neighbor

in triangle-free graphs.

3.2.1 P_3 -free graphs

Let G be a P_3 -free graph and let C be a component of G . If C is not complete, then there must be two vertices at distance two in C , and hence a P_3 appears. Therefore, G is the union of disjoint cliques, and thus $\chi_i(G) = \omega(G)$, where the clique number $\omega(G)$ of a graph G is the number of vertices in a maximum clique in G .

Since in G , adjacent vertices must have common neighbors, and nonadjacent vertices cannot have common neighbors, in each injective coloring of G , adjacent vertices must be colored differently, and nonadjacent vertices may receive the same color. Thus an injective coloring of G is equivalent to a proper coloring of G .

Now, we can apply FF to on-line injectively color G (FF was used by Gyarfas and Lehel [65] in 1988 to assign the smallest possible integer as color to the current vertex v of the graph). Let $\pi \in \Pi(G)$. For each vertex v of G , we use G_v to denote the subgraph of G induced by v and all the vertices before v under π . We formulate FF as follows.

FF:

If $\tilde{H}(v) := H(\bar{N}_{G_v}(v)) - H(N_{G_v}(v)) = \emptyset$, then a new color is used on v .

Else, the smallest color in $\tilde{H}(v)$ is used on v .

Theorem 3.6. *Let G be a P_3 -free graph. Then FF uses the optimal number of colors on G . That is, $\chi_i^{FF}(G) = \omega(G)$.*

Proof. First, $\chi_i^{FF}(G) \geq \chi_i(G) = \omega(G)$ trivially holds. When we color the current vertex v , it has at most $\Delta(G)$ neighbors (the last vertex of the greatest clique in G has just now $\Delta(G)$ neighbors), and these vertices admit at most $\Delta(G)$ colors, which leads to at most $\Delta(G) + 1$ colors used on vertices v . Therefore, $\chi_i^{FF}(G) \leq \Delta(G) + 1$. Since G is a union of cliques, $\omega(G) = \Delta(G) + 1$. So, $\chi_i^{FF}(G) = \omega(G)$. \square

So FF yields an optimal on-line injective coloring on a P_3 -free graph.

3.2.2 Triangle-free graphs and bipartite graphs

Let \mathcal{G} denote a family of triangle-free graphs, and let $G \in \mathcal{G}$. Since any two adjacent vertices cannot have a common neighbor in G , adjacent vertices can be colored the same, and nonadjacent vertices must be colored differently to avoid the possible common neighbor during the on-line process. So, an on-line injective coloring of G is equivalent to an on-line proper coloring of its complement \bar{G} .

If every component of G is either a single vertex or a single edge, then all vertices can be colored with the same color. So, we only consider graphs with a component of size (the number of edges contained in it) at least 2.

Theorem 3.7. *Let G be a triangle-free graph with a component of size at least 2, and let OL be an optimal on-line injective coloring algorithm for G . Then $\chi_i^{OL}(G) \geq \alpha(G)$.*

Proof. Let A be an on-line injective coloring of G , and let $\pi(G)$ be an ordering of the vertices of G in which the vertices of a maximum independent set S of G emerge first. Then S must be colored by $\alpha(G)$ colors according to the rule. So $\chi_i^A(G, \pi) \geq \alpha(G)$. Because $\chi_i^A(G)$ is not less than $\chi_i^A(G, \pi)$, $\chi_i^A(G) \geq \alpha(G)$. \square

Here, we can also use FF to on-line injectively color any $G \in \mathcal{G}$. Let $\pi \in \Pi(G)$ for G , and assume that $v \in V(G)$. Let G_v denote the subgraph of G induced by v and all vertices before v under π . We express FF* as follows.

FF*:

If $\tilde{H}(v) = H(N_{G_v}(v)) - H(\bar{N}_{G_v}(v)) = \emptyset$, then a new color is used on v .

Else, the smallest color in $\tilde{H}(v)$ is used on v .

In [65], Gyárfás and Lehel proved that FF returns a tight bound to the on-line coloring of a graph G while \bar{G} is bipartite. Here \bar{G} denotes the complement of G .

Lemma 1. [65] $\chi^{FF}(\bar{G}) \leq \frac{3}{2}\alpha(G)$ for each bipartite graph G , and this bound is tight.

Therefore, we have $\chi_i^{FF^*}(G) \leq \frac{3}{2}\alpha(G)$ for each bipartite graph G , and thus

Theorem 3.8. $\chi_i^{FF^*}(G) \leq \frac{3}{2}\chi_i^{OL}(G)$ for each bipartite graph G .

Next we further improve the competitive ratio of FF* by presenting a new optimal on-line injective coloring algorithm BFF for G .

Let $\pi \in \Pi(G)$ for a bipartite graph G , and let v be a vertex of G . We still use G_v to denote the subgraph induced by v and all vertices before v under π . Let $C_v = (I_1, I_2)$ be the component of G_v containing v . Note that (I_1, I_2) is the unique bipartition of C_v . Without loss of generality, we assume that $v \in I_1$.

BFF:

If $\tilde{H}(v) = H(I_2) - H(I_1) - H(G_v \setminus C_v) = \emptyset$, then a new color is used on v .

Else, the smallest color in $\tilde{H}(v)$ is used on v .

Remark: In BFF, we know that the colors of vertices in the same bipartite class of the same component are different from each other, and the same happens on the vertices in different components.

Lemma 2. Let G be a bipartite graph with two components A_1 and A_2 . Then, for any $\pi \in \Pi(G)$, $\chi_i^{BFF}(G, \pi) = \chi_i^{BFF}(A_1, \pi_1) + \chi_i^{BFF}(A_2, \pi_2)$, where π_j is the restriction of π on A_j , $j = 1, 2$.

Proof. We proceed by induction on the order n of G . The lemma is true for $n = 2$, because $G = 2K_1$ in this case. Suppose that $n \geq 2$. Let v be the last vertex of G under π , and let $G' = G \setminus \{v\}$. Without loss of generality, $v \in A_1$. By induction, $\chi_i^{BFF}(G', \pi) = \chi_i^{BFF}(A_1 \setminus \{v\}, \pi_1) + \chi_i^{BFF}(A_2, \pi_2)$.

By the Remark, $\tilde{H}_G(v) = H_G(I_2) - H_G(I_1) - H_G(G_v \setminus C_v) = H_G(I_2) - H_G(I_1)$, where $|H_G(I_2)| = |I_2|$, $|H_G(I_1)| = |I_1|$, and $\tilde{H}_{A_1}(v) = H_{A_1}(I_2) - H_{A_1}(I_1)$, where $|H_{A_1}(I_2)| = |I_2|$, $|H_{A_1}(I_1)| = |I_1|$.

Suppose that v uses a new color in G under π (i.e., $\chi_i^{BFF}(G, \pi) = \chi_i^{BFF}(G', \pi) + 1$). It follows that $\tilde{H}_G(v) = \emptyset$. By BFF, $H_G(I_2) \subseteq H_G(I_1)$. Thus $|I_1| \geq |I_2|$. Assume that v used an old color in A_1 under π_1 , thus $H_{A_1}(I_2) \supset H_{A_1}(I_1)$. Then $|I_2| > |I_1|$, a contradiction. So, $\chi_i^{BFF}(A_1, \pi_1) = \chi_i^{BFF}(A_1 \setminus \{v\}, \pi_1) + 1$.

If v uses an old color in G under π (i.e., $\chi_i^{BFF}(G, \pi) = \chi_i^{BFF}(G', \pi)$), then $\tilde{H}_G(v) \neq \emptyset$. By BFF, $H_G(I_1) \subset H_G(I_2)$. So $|I_1| < |I_2|$. Assume that v used a

new color in A_1 under π_1 . Then $H_{A_1}(I_2) \subseteq H_{A_1}(I_1)$. It follows that $|I_2| \leq |I_1|$, a contradiction. Thus $\chi_i^{BFF}(A_1, \pi_1) = \chi_i^{BFF}(A_1 \setminus \{v\}, \pi_1)$.

We conclude that $\chi_i^{BFF}(G, \pi) = \chi_i^{BFF}(A_1, \pi_1) + \chi_i^{BFF}(A_2, \pi_2)$. \square

It is easy to see that Lemma 2 can be generalized to graphs with three or more components.

Lemma 3. *Let G be a bipartite graph with k components A_1, A_2, \dots, A_k , $k \geq 3$. Then, for any $\pi \in \Pi(G)$, $\chi_i^{BFF}(G, \pi) = \sum_{j=1}^k \chi_i^{BFF}(A_j, \pi_j)$, where π_j is the restriction of π on A_j , $j = 1, \dots, k$.*

Lemma 4. *If G is a connected bipartite graph, then $\chi_i^{BFF}(G) \leq \alpha(G)$.*

Proof. We prove by induction on the order n of G that $\chi_i^{BFF}(G, \pi) \leq \alpha(G)$, for any $\pi \in \Pi(G)$. The statement is true for $n = 3$, because $G = P_3$ and $\chi_i^{BFF}(P_3, \pi) = 2 = \alpha(P_3)$.

Suppose, then, that $n \geq 4$. Let v be the last vertex of G under π and set $G' = G(V \setminus \{v\})$. Let $c(G')$ be the number of components of G' .

First we suppose that $c(G') = 1$. Now, $\chi_i^{BFF}(G', \pi) \leq \alpha(G')$ by induction hypothesis. If $\chi_i^{BFF}(G', \pi) < \alpha(G')$ (i.e., $\chi_i^{BFF}(G', \pi) \leq \alpha(G') - 1$), then $\chi_i^{BFF}(G, \pi) \leq \chi_i^{BFF}(G', \pi) + 1 \leq \alpha(G') \leq \alpha(G)$ since v is colored by at most a new color in BFF . If $\chi_i^{BFF}(G', \pi) = \alpha(G')$ and $\alpha(G') < \alpha(G)$, then we also have $\chi_i^{BFF}(G, \pi) \leq \chi_i^{BFF}(G', \pi) + 1 \leq \alpha(G') + 1 \leq \alpha(G)$. So, we assume that $\chi_i^{BFF}(G', \pi) = \alpha(G')$, and $\alpha(G) = \alpha(G')$.

Because $c(G') = c(G) = 1$, set $G' = (I'_1, I'_2)$ and $G = (I_1, I_2)$. Without loss of generality, suppose that $I_1 = I'_1 \cup \{v\}$ and $I_2 = I'_2$. Then v must be adjacent with some vertex in I'_2 , otherwise it contradicts the connectivity. Assume that v was colored with a new color in BFF . Since G has a unique component, by the rule of BFF , $H(I'_1)$ must contain all the old $\alpha(G')$ colors. Let O denote the set of vertices colored with these $\alpha(G')$ colors in I'_1 . It follows that $O \cup \{v\} \subset I_1$ is an independent set of size $\alpha(G') + 1$, larger than $\alpha(G')$, a contradiction. Thus $\chi_i^{BFF}(G, \pi) = \chi_i^{BFF}(G', \pi) = \alpha(G') = \alpha(G)$.

Now, we suppose that $c(G') = m, m \geq 2$. Set $G' = \bigcup_{k=1}^m A_k$, where A_k denote the k th component of G' . Since the order of A_k is less than n , $\chi_i^{BFF}(A_k, \pi) \leq \alpha(A_k)$ by induction.

If there exists a $j \in \{1, 2, \dots, m\}$ such that $\chi_i^{BFF}(A_j, \pi_j) < \alpha(A_j)$ (i.e., $\chi_i^{BFF}(A_j, \pi_j) \leq \alpha(A_j) - 1$), since v is colored by at most a new color in BFF ,

$\chi_i^{BFF}(G, \pi) \leq \chi_i^{BFF}(G', \pi) + 1 = \sum_{l=1}^m \chi_i^{BFF}(A_l, \pi_l) + 1 \leq \sum_{l=1}^{j-1} \alpha(A_l) + \alpha(A_j) - 1 + \sum_{l=j+1}^m \alpha(A_l) + 1 \leq \sum_{l=1}^m \alpha(A_l) = \alpha(G') \leq \alpha(G)$ by Lemma 3.

So, we suppose that $\chi_i^{BFF}(A_j, \pi_j) = \alpha(A_j)$ for all $j \in \{1, 2, \dots, m\}$. By Lemma 3, $\chi_i^{BFF}(G', \pi) = \sum_{l=1}^m \chi_i^{BFF}(A_l, \pi_l) = \sum_{l=1}^m \alpha(A_l) = \alpha(G')$. If $\alpha(G) = \alpha(G') + 1$, $\chi_i^{BFF}(G, \pi) \leq \chi_i^{BFF}(G', \pi) + 1 = \alpha(G') + 1 = \alpha(G)$ as v is colored by at most a new color in *BFF*. Otherwise, we have $\alpha(G) = \alpha(G')$. Set $A_1 = (A_{11}, A_{12}), \dots, A_m = (A_{m1}, A_{m2})$. Set $G = (I_1, I_2)$. Since G is connected, it follows that v must be adjacent to some vertex in A_{jk} , for every $j \in \{1, \dots, m\}$, and $k = 1$ or 2 . Without loss of generality, we suppose that $k = 2$ for all j . Set $I'_1 = \bigcup_{j=1}^m A_{j1}$ and $I'_2 = \bigcup_{j=1}^m A_{j2}$. Let $I_1 = \bigcup_{j=1}^m A_{j1} \cup \{v\}$ and $I_2 = I'_2$. Assume that v was colored with a new color in *BFF*. By the rule of *BFF*, $H(I'_1)$ must contain all the old $\alpha(G')$ colors. Let O denote the set of vertices colored with the $\alpha(G')$ colors. Then $O \cup \{v\} \subset I_1$ is an independent set with size $\alpha(G') + 1$, larger than $\alpha(G')$ (i.e., $\alpha(G)$), a contradiction. Thus $\chi_i^{BFF}(G, \pi) = \chi_i^{BFF}(G', \pi) = \alpha(G') = \alpha(G)$.

We conclude that for any $\pi \in \Pi(G)$, $\chi_i^{BFF}(G, \pi) \leq \alpha(G)$, i.e., $\chi_i^{BFF}(G) \leq \alpha(G)$. \square

Combining Lemma 4 with Lemma 3, we have that $\chi_i^{BFF}(G) \leq \alpha(G)$ for each bipartite graph G , and thus

Theorem 3.9. *If G is a bipartite graph with at least one component of size greater than one, then $\chi_i^{BFF}(G) = \chi_i^{OL}(G)$.*

3.2.3 Concluding remarks

It is interesting to notice that it is easy to determine the chromatic number of bipartite graphs, whereas it is NP-hard to determine the injective chromatic number of bipartite graphs. And for every $\epsilon > 0$, it is even impossible to efficiently approximate the injective chromatic number of bipartite graphs within a factor of $n^{\frac{1}{3}-\epsilon}$ if $ZPP \neq NP$. In on-line situations, to the best of our knowledge, it is not known whether there is an optimal on-line algorithm for proper coloring bipartite graphs. However, we obtain an optimal on-line algorithm for injective coloring bipartite graphs.

The first open question is the problem of deciding the complexity of max-injective coloring trees and the existence of a constant approximation algorithm if it is NP-complete. Another interesting question is to investigate

the existence of an on-line competitive algorithm for on-line injective coloring triangle-free graphs.

Chapter 4

An approximation algorithm for max hypergraph cut with limited unbalance

We consider the design of a semidefinite programming (SDP) based approximation algorithm for the problem *max hypergraph cut with limited unbalance* (MHC-LU): find a partition of the vertices of a weighted hypergraph $H = (V, E)$ into two subsets V_1, V_2 with $||V_2| - |V_1|| \leq u$ for some given u and maximizing the total weight of the edges meeting both V_1 and V_2 . The problem MHC-LU generalizes several other combinatorial optimization problems including *max cut*, *max cut with limited unbalance* (MC-LU), *max set splitting*, *max Ek -set splitting* and *max hypergraph bisection*.

By generalizing several earlier ideas, we present an SDP randomized approximation algorithm for MHC-LU with guaranteed worst-case performance ratios for various unbalance parameters $\tau = u/|V|$. We also give the worst-case performance ratio of the SDP-algorithm for approximating MHC-LU regardless of the value of τ . Our strengthened SDP relaxation and rounding method improve a result of Ageev and Sviridenko [2] on max hypergraph bisection (MHC-LU with $u = 0$), and results of Andersson and Engebretsen [11], Gaur and Krishnamurti [55] and Zhang et al. [143] on max set splitting (MHC-LU with $u = |V|$). Furthermore our new formula for the performance ratio by a tighter analysis compared with that in Galbiati and Maffioli [53] is responsible

for the improvement of a result of Galbiati and Maffioli [53] on MC-LU for some range of τ .

4.1 An SDP relaxation of MHC-LU

We first formulate MHC-LU as an integer program. Let $H = (V, E)$ be a hypergraph with $V = \{1, \dots, m\}$ and $E = \{S_1, \dots, S_n\}$, and let ω be a non-negative real function on E with $\omega_j := \omega(S_j)$.

Let $x_0 \in \{-1, 1\}$ be a reference variable. Let (V_1, V_2) be a cut of H such that $\|V_1\| - \|V_2\| \leq u$. For each vertex i , let $x_i = x_0$ if $i \in V_1$ and otherwise let $x_i = -x_0$. Then

$$\left| \sum_{i=1}^m x_0 x_i \right| = \left| \sum_{i=1}^m x_i \right| = \|V_1\| - \|V_2\| \leq u;$$

so

$$\left(\sum_{i=1}^m x_i \right)^2 \leq u^2 \text{ and } -u \leq \sum_{i=1}^m x_0 x_i \leq u.$$

Also note that

$$\left(\sum_{i \in S_j} x_i \right)^2 = |S_j| + 2 \sum_{i < k \in S_j} x_i x_k,$$

which implies

$$\sum_{i < k \in S_j} x_i x_k \geq -\lfloor |S_j|/2 \rfloor,$$

since $|\sum_{i \in S_j} x_i| \geq 1$ when $|S_j|$ is odd.

Let $z_j = 1$ if S_j is in the cut (V_1, V_2) , and let $z_j = 0$ otherwise. Then the total weight of the cut is

$$\omega(V_1, V_2) = \sum_{j=1}^n \omega_j z_j.$$

Note that if $x_i = x_k$ for all $i, k \in S_j$, i.e., $z_j = 0$, then $\sum_{i < k \in S_j} (1 - x_i x_k)/2 = 0$, and if $x_i \neq x_k$ for some $i, k \in S_j$, i.e., $z_j \neq 0$, then $\sum_{i < k \in S_j} (1 - x_i x_k)/2 \geq t(|S_j| - t)$ for some $1 \leq t \leq |S_j| - 1$ (so $t(|S_j| - t) \geq |S_j| - 1$); hence

$$\frac{1}{|S_j| - 1} \sum_{i < k \in S_j} \frac{1 - x_i x_k}{2} \geq z_j.$$

Thus MHC-LU can be formulated as the following quadratic integer program.

$$\begin{aligned} \omega^{opt} &= \max \sum_{j=1}^n \omega_j z_j \\ s.t. & \frac{1}{|S_j| - 1} \sum_{i < k \in S_j} \frac{1 - x_i x_k}{2} \geq z_j \quad \text{for all } S_j \in E \end{aligned} \quad (4.1)$$

$$\sum_{1 \leq i, k \leq m} x_i x_k \leq u^2 \quad (4.2)$$

$$-u \leq \sum_{1 \leq i \leq m} x_0 x_i \leq u \quad (4.3)$$

$$\sum_{i < k \in S_j} x_i x_k \geq -\lfloor |S_j|/2 \rfloor \quad \text{for all } S_j \in E \quad (4.4)$$

$$x_i \in \{-1, 1\}, \quad \text{for } i = 0, 1, 2, \dots, m$$

$$z_j \in \{0, 1\}, \quad \text{for } j = 1, \dots, n$$

As mentioned in the introduction, Max Set Splitting is the restricted versions of MHC-LU. We note that constraint (4.1) is used in Andersson and Engebretsen [11] for Max Set Splitting. Constraint (4.2) is the limited unbalance constraint used in Galbiati and Maffioli [53] for MC-LU. Constraint (4.4) is used in Zhang et al. [143] to strengthen the SDP relaxation in [11] for Max Set Splitting. Note that constraint (4.3) is not used in [53]. Although constraint (4.3) is implied by constraint (4.2), its corresponding constraint (4.8) in our SDP relaxation below, is not implied by constraint (4.7) (corresponding to constraint (4.2) above); and we also add further triangle inequalities (4.10) and (4.11) which were first used by Feige and Goemans [45] to improve the performance ratio of the approximation algorithms for Max 2SAT and Max DICUT. The triangle constraints (4.10) and (4.11) does not appear in the relaxation presented in [11] and [143] for Max Set Splitting. So with constraints (4.8), (4.10) and (4.11) the following SDP relaxation strengthens the SDP relaxations in [53], [11] and [143] for MC-LU and Max Set Splitting, respectively.

The SDP relaxation of the above quadratic integer program becomes:

$$\omega^{SDP} : = \max \sum_{i=1}^n \omega_j z_j \quad (4.5)$$

$$s.t. \quad \frac{1}{|S_j| - 1} \sum_{i < k \in S_j} \frac{1 - X_{ik}}{2} \geq z_j \quad \text{for all } S_j \in E \quad (4.6)$$

$$\sum_{1 \leq i, k \leq m} X_{ik} \leq u^2 \quad (4.7)$$

$$-u \leq \sum_{1 \leq i \leq m} X_{0i} \leq u \quad (4.8)$$

$$\sum_{i < k \in S_j} X_{ik} \geq -\lfloor |S_j|/2 \rfloor \quad \text{for all } S_j \in E \quad (4.9)$$

$$-X_{0i} - X_{0k} + X_{ik} \geq -1 \quad \text{for } 1 \leq i \leq k \leq m \quad (4.10)$$

$$X_{ik} + X_{il} + X_{kl} \geq -1 \quad \text{for } 0 \leq i \leq k \leq l \leq m \quad (4.11)$$

$$X_{ii} = 1 \quad \text{for } i = 0, 1, \dots, m$$

$$X \succeq 0$$

$$0 \leq z_j \leq 1 \quad \text{for } j = 1, 2, \dots, n$$

where $X \succeq 0$ means that the $(n+1) \times (n+1)$ symmetric matrix X is positive semidefinite.

Throughout the chapter, $\omega(V_1, V \setminus V_1)$ denotes the weight of the cut $(V_1, V \setminus V_1)$, and $\tau = u/m$ denotes the unbalance parameter. We now present an SDP-based approximation algorithm for MHC-LU as follows.

SDP-algorithm for MHC-LU:

Step 1. Solve problem (4.5) to obtain an almost optimal positive semidefinite matrix X^* and a vector z^* . Apply a Cholesky decomposition to X^* , $X^* = v \cdot v^T$, to obtain column vectors $(v_0^*, v_1^*, \dots, v_m^*)$ of v . For a given $0 < \epsilon < 1$, repeat the following steps $K = O((1/\epsilon) \log(1/\epsilon))$ times, and output the best set \bar{V}_1 .

Step 2. Choose two rotation coefficients $0 < \theta, \theta' \leq 1$, and a probability $0 \leq \nu \leq 1$. Construct two sets of unit vectors (v_0, v_1, \dots, v_m) and $(v'_0, v'_1, \dots, v'_m)$ such that $v_i \cdot v_j = \theta v_i^* \cdot v_j^*$, and $v'_i \cdot v'_j = \theta' v_i^* \cdot v_j^*$ for every

$0 \leq i < j < m$. Goto Step 3 with probability ν , and independently goto Step 4 with probability $1 - \nu$.

Step 3. Choose a random vector r and for $0 \leq i \leq m$ let $\hat{x}_i = 1$ if $v_i \cdot r \geq 0$, and $\hat{x}_i = -1$ otherwise. Set $\hat{V}_1 = \{i : \hat{x}_i = \hat{x}_0\}$. Goto Step 5.

Step 4. For each $i \in V$, put i in \hat{V}_1 , with probability $(1 + v'_0 \cdot v'_i)/2$, and independently in $V \setminus \hat{V}_1$ with probability $(1 - v'_0 \cdot v'_i)/2$.

Step 5. Given $0 \leq p \leq 1/2$, each element in \hat{V}_1 or in $V \setminus \hat{V}_1$ has a probability of p of being assigned to the opposite subset, independently. This gives a new partition $(V_1, V \setminus V_1)$ of V . Let

$$\tilde{V}_1 = \begin{cases} V_1, & \text{if } |V_1| \geq m/2, \\ V \setminus V_1, & \text{otherwise.} \end{cases}$$

If $|\tilde{V}_1| \leq (m + u)/2$, let $\bar{V}_1 = \tilde{V}_1$, else goto Step 6.

Step 6. Let $\tilde{V}_1 = \{1, 2, \dots, T\}$. For each $1 \leq t \leq T$, define $B(t) := \{S_j : t \in S_j \cap \tilde{V}_1 \neq \emptyset \text{ and } S_j \cap (V \setminus \tilde{V}_1) \neq \emptyset\}$. Set

$$c(t) := \sum_{S_j \in B(t)} \frac{w_j}{|S_j \cap \tilde{V}_1|}$$

as the contribution of vertex t to the weight of the cut $(\tilde{V}_1, V \setminus \tilde{V}_1)$. Assume without loss of generality that $c(1) \leq \dots \leq c(T)$. Let $\bar{V}_1 = \tilde{V}_1 \setminus \{1, 2, \dots, T - (m + u)/2\}$.

Note that if we set $\theta = 1$ in Step 2, the technique of randomized rounding in Step 3 is the hyperplane rounding due to Goemans and Williamson [57]. The technique of outward rotations used in Step 3 is a way of combining the classical hyperplane rounding method with independent random choice which was used in [49, 115, 138, 148].

If we set $\theta' = 1$ in Step 2, the technique of randomized rounding in Step 4 is the linear randomized rounding due to Halperin and Zwick [70]. In Step 4, we

combine the linear randomized rounding method with independent random choice by using a new rotation parameter θ' . It turns out that we achieve better performance ratios by combining the outward rotations of the random hyperplane rounding procedure with that of the linear randomized rounding procedure for MHC-LU when the minimum number of vertices in a hyperedge is 3. We point out that \widehat{V}_1 in Step 3 and Step 4 is only introduced for the rounding analysis in the next Section.

The first part of Step 5 is a probabilistic postprocessing step, which perturbs the initial partition $(\widehat{V}_1, V \setminus \widehat{V}_1)$ obtained in Step 3 or Step 4 to construct a new partition $(V_1, V \setminus V_1)$ correspondingly.

Step 6 adjusts the size of the sets in a partition to satisfy the limited unbalance constraint. We have the following lemma.

Lemma 4.1.1. $\omega(\overline{V}_1, V \setminus \overline{V}_1) \geq \frac{m+u}{2T} \omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)$ if $T > \frac{m+u}{2}$, and $\omega(\overline{V}_1, V \setminus \overline{V}_1) = \omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)$ otherwise.

Proof. Step 6 reduces the number of vertices in \widetilde{V}_1 to $(m+u)/2$ by moving from \widetilde{V}_1 to $V \setminus \widetilde{V}_1$ the $T - (m+u)/2$ vertices with the smallest contributions to the cut. (Note that the condition $\|V_2\| - \|V_1\| \leq u$ may be written as $\frac{m-u}{2} \leq |V_i| \leq \frac{m+u}{2}$ for $i = 1, 2$.) Observe that $\omega(\widetilde{V}_1, V \setminus \widetilde{V}_1) = \sum_{t=1}^T c(t)$. Moreover, the construction of \overline{V}_1 guarantees that

$$\frac{\omega(\overline{V}_1, V \setminus \overline{V}_1)}{(m+u)/2} \geq \frac{\sum_{t=T-(m+u)/2+1}^T c(t)}{(m+u)/2} \geq \frac{\sum_{t=1}^T c(t)}{T} = \frac{\omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)}{T},$$

where the first inequality follows from definition of \overline{V}_1 , and the second inequality holds because the average over the largest numbers of a sequence is at least the average over the entire sequence. Therefore, $\omega(\overline{V}_1, V \setminus \overline{V}_1) \geq \frac{m+u}{2T} \omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)$ if $T > \frac{m+u}{2}$, and $\omega(\overline{V}_1, V \setminus \overline{V}_1) = \omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)$ otherwise. \square

To analyze the quality of the SDP approximation algorithm, we need to establish a lower bound on $E[\omega(\overline{V}_1, V \setminus \overline{V}_1)]/\omega^{opt}$. By Lemma 4.1.1, we then wish to establish a lower bound on $E[\frac{m+u}{2T} \cdot \omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)]/\omega^{opt}$ which is also a lower bound on $E[\omega(\overline{V}_1, V \setminus \overline{V}_1)]/\omega^{opt}$. However, it is not easy to calculate the expected value of $\frac{m+u}{2T} \cdot \omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)$, since both $\frac{m+u}{2T}$ and $\omega(\widetilde{V}_1, V \setminus \widetilde{V}_1)$ are random variables and they are multiplied together where $T = |\widetilde{V}_1|$. Instead, we study a family of random variables $Z(\gamma)$ whose expected values can be easily

estimated and bounded, an idea first used by Frieze and Jerrum [52] for Max Bisection and extended in [49, 70, 71, 138]. For fixed values $\tau = u/m \in [0, 1]$ and $\gamma > 0$,

$$Z(\gamma) = \frac{\omega(\tilde{V}_1, V \setminus \tilde{V}_1)}{\omega^{opt}} + \gamma \frac{|\tilde{V}_1|(m - |\tilde{V}_1|)}{m^2}.$$

Whenever $Z(\gamma)$ fulfills its expectation, (i.e., $Z(\gamma) \geq \alpha^{(p)} + \gamma\beta - o(1)$ (see the last Section)), we have $\frac{m+u}{2T} \cdot \omega(\tilde{V}_1, V \setminus \tilde{V}_1) \geq \bar{R}(\tau) \cdot \omega^{opt}$ which establishes a performance ratio $\bar{R}(\tau)$ of the SDP approximation algorithm in the last Section. Therefore, the aim of the following sections (except the last one) is to give bounds on $\mathbb{E}[\omega(\tilde{V}_1, V \setminus \tilde{V}_1)]$ and $\mathbb{E}[|\tilde{V}_1|(m - |\tilde{V}_1|)]$. These bounds will be used in the last Section to give bounds on $\mathbb{E}[Z(\gamma)]$ and then to obtain the performance ratio $\bar{R}(\tau)$.

4.2 Bound on the expected contribution of an edge by Steps 1-4

To obtain the bound on $\mathbb{E}[\omega(\tilde{V}_1, V \setminus \tilde{V}_1)]$ in terms of the maximum value ω^{opt} , we need to establish the bound on the expected contribution of an edge by executing Steps 1-4. Note that we have an initial cut $(\hat{V}_1, V \setminus \hat{V}_1)$ of H by executing Steps 1-4. An edge S_j is in this cut if there exists $i, k \in S_j$ such that $i \in \hat{V}_1$ and $k \in \hat{V}_2 := V \setminus \hat{V}_1$. The indicator random variables \hat{X}_{ik} and $\hat{U}(S_j)$ are defined as

$$\hat{X}_{ik} = \begin{cases} -1, & \text{if } i \in \hat{V}_i \text{ and } k \in V \setminus \hat{V}_i, \text{ for } i = 1, 2, \\ 1, & \text{otherwise.} \end{cases}$$

$$\hat{U}(S_j) = \begin{cases} 1, & \text{if } S_j \in (\hat{V}_1, V \setminus \hat{V}_1), \\ 0, & \text{otherwise.} \end{cases}$$

From [57, 148], we have the following equality if the algorithm executes Step 3,

$$E[\hat{X}_{ik}] = (2/\pi) \arcsin(\theta X_{ik}^*),$$

where X^*, z^* is an optimal solution of (4.5).

If the algorithm executes Step 4, it is easily seen that

$$\begin{aligned}
E[\widehat{X}_{ik}] &= (-1) \cdot \left(\frac{1 + \theta' X_{0i}^*}{2} \cdot \frac{1 - \theta' X_{0k}^*}{2} + \frac{1 - \theta' X_{0i}^*}{2} \cdot \frac{1 + \theta' X_{0k}^*}{2} \right) + \\
&\quad (+1) \left(\frac{1 + \theta' X_{0i}^*}{2} \cdot \frac{1 + \theta' X_{0k}^*}{2} + \frac{1 - \theta' X_{0i}^*}{2} \cdot \frac{1 - \theta' X_{0k}^*}{2} \right) \\
&= \theta'^2 X_{0i}^* X_{0k}^*.
\end{aligned}$$

Hence, from Steps 1-4, we have

$$E[\widehat{X}_{ik}] = \nu(2/\pi) \arcsin(\theta X_{ik}^*) + (1 - \nu)\theta'^2 X_{0i}^* X_{0k}^*,$$

The main objective here is to establish a lower bound on the expected contribution of an individual edge to the cut before executing Step 5. More precisely, we show that for any fixed $s \geq 2$ and $s \in \mathbb{N}$ there exists a constant $\alpha_s > 0$ such that for each $S_j \in E$ with $|S_j| = s$ we have $\mathbb{E}[\widehat{U}(S_j)] \geq \alpha_s z_j^*$.

For convenience let $S := S_j$ be an edge of H belonging to $(\widehat{V}_1, V \setminus \widehat{V}_1)$. For the same reason as for constraint (4.4), $\sum_{i < k \in S} \widehat{X}_{ik} = \sum_{i < k \in S} \widehat{x}_i \widehat{x}_k \geq -\lfloor |S|/2 \rfloor$. So

$$\begin{aligned}
\sum_{i < k \in S} \frac{1 - \widehat{X}_{ik}}{2} &\leq (1/2) \left[\binom{|S|}{2} + \lfloor |S|/2 \rfloor \right] \\
&\leq L := \begin{cases} |S|^2/4, & \text{if } |S| \text{ is even,} \\ (|S| + 1)(|S| - 1)/4, & \text{otherwise.} \end{cases}
\end{aligned}$$

By the definition of $\widehat{U}(S)$, it follows that

$$\begin{aligned}
 \mathbb{E}[\widehat{U}(S)] &\geq \frac{1}{L} \mathbb{E} \left[\sum_{i < k \in S} \frac{1 - \widehat{X}_{ik}}{2} \right] \\
 &= \sum_{i < k \in S} \frac{1 - \mathbb{E}[\widehat{X}_{ik}]}{2L} \\
 &= \widehat{f}^* \tag{4.12} \\
 &:= \sum_{i < k \in S} \frac{1 - (\nu(2/\pi) \arcsin(\theta X_{ik}^*) + (1 - \nu)\theta'^2 X_{0i}^* X_{0k}^*)}{2L} \tag{4.13}
 \end{aligned}$$

It suffices to show that when $z_j^* > 0$, there exists a constant $\alpha_{|S|} > 0$ (dependent only on $|S|$) such that

$$\widehat{f}^*/z_j^* \geq \alpha_{|S|}.$$

If $|S| = 2$, then $L = 1$ and

$$\begin{aligned}
 \widehat{f}^*/z_j^* &= \frac{1 - (\nu(2/\pi) \arcsin(\theta X_{ik}^*) + (1 - \nu)\theta'^2 X_{0i}^* X_{0k}^*)}{2z_j^*} \\
 &\geq \frac{\nu(1 - (2/\pi) \arcsin(\theta X_{ik}^*)) + (1 - \nu)(1 - \theta'^2 X_{0i}^* X_{0k}^*)}{2 \frac{1 - X_{ik}^*}{2}} \\
 &\geq \alpha_2 := \min_{-1 \leq X < 1, -1 \leq Y \leq 1} \frac{\nu(1 - (2/\pi) \arcsin(\theta X)) + (1 - \nu)(1 - \theta'^2 Y)}{1 - X},
 \end{aligned}$$

where the first inequality holds in view of $z_j^* \leq \frac{1 - X_{ik}^*}{2}$ by (4.6) and the last inequality holds because we assume $z_j^* > 0$ (so $-1 \leq X_{ik}^* < 1$) and $-1 \leq X_{0i}^* X_{0k}^* \leq 1$. We now define

$$\alpha_2 := \min_{-1 \leq X < 1, -1 \leq Y \leq 1} \frac{\nu(1 - (2/\pi) \arcsin(\theta X)) + (1 - \nu)(1 - \theta'^2 Y)}{1 - X} \tag{4.14}$$

When $\theta = 1$ and $\nu = 1$, α_2 was calculated by Goemans and Williamson [57] to obtain a 0.878 approximation result for Max Cut.

Next, we consider the case when $|S| \geq 3$. For our purpose, we fix parameters $\theta, \theta' \in (0, 1]$ and $\nu \in [0, 1]$. Let $N_{|S|} = |S|(|S| - 1)/2$, $I = [-\lfloor |S|/2 \rfloor, N_{|S|}]$

and $I' = [-|S|/2, N_{|S|}]$. For $-1 \leq X_{0i}, X_{0k}, X_{ik} \leq 1$ which satisfy the triangle inequalities (4.10) and (4.11), we put $\lambda = \sum_{i < k \in S} X_{ik}$ and $\lambda' = \sum_{i < k \in S} X_{0i}X_{0k}$. Let

$$z'(\lambda) = \min\left\{1, \frac{N_{|S|} - \lambda}{2(|S| - 1)}\right\},$$

$$h_1(\lambda) = \sum_{i < k \in S} 1 - (2/\pi) \arcsin(\theta X_{ik}), \text{ where } \lambda \in I,$$

$$h_2(\lambda') = \sum_{i < k \in S} (1 - \theta'^2 X_{0i}X_{0k}), \text{ where } \lambda' \in I'.$$

Moreover, let $\alpha_{|S|} = \min_{\lambda \in I, \lambda' \in I'} \alpha_{|S|}(\lambda, \lambda')$, where $\alpha_{|S|}(\lambda, \lambda') = (\nu h_1(\lambda) + (1 - \nu)h_2(\lambda')) / (2Lz'(\lambda))$. Then

Lemma 4.2.1. $\widehat{f}^*/z_j^* \geq \alpha_{|S|}$.

Proof. Let $\lambda^* = \sum_{i < k \in S} X_{ik}^*$. By the constraints (4.6), (4.9) and the assumption $z_j^* > 0$, we have

$$-|S|/2 \leq \lambda^* \leq N_{|S|} - 2z_j^*(|S| - 1) < N_{|S|}$$

and

$$z_j^* \leq \min\left\{1, \frac{1}{|S| - 1} \sum_{i < k \in S} \frac{1 - X_{ik}^*}{2}\right\} = \min\left\{1, \frac{N_{|S|} - \lambda^*}{2(|S| - 1)}\right\} = z'(\lambda^*) \quad (4.15)$$

Moreover, let $\lambda'^* = \sum_{i < k \in S} X_{0i}^*X_{0k}^*$, then we have

$$-|S|/2 \leq \lambda'^* \leq N_{|S|},$$

since $(\sum_{i \in S} X_{0i}^*)^2 \leq |S| + 2 \sum_{i < k \in S} X_{0i}^*X_{0k}^*$, and $-1 \leq X_{0i}^* \leq 1$ for $i \in S$.

From (4.13) and (4.15), we find

$$\begin{aligned} \widehat{f}^*/z_j^* &\geq \sum_{i < k \in S} \frac{\nu(1 - (2/\pi) \arcsin(\theta X_{ik}^*)) + (1 - \nu)(1 - \theta'^2 X_{0i}^*X_{0k}^*)}{2Lz'(\lambda^*)} \\ &\geq \min_{\lambda^* = \sum_{i < k \in S} X_{ik}^*, \lambda'^* = \sum_{i < k \in S} X_{0i}^*X_{0k}^*} \alpha_{|S|}(\lambda^*, \lambda'^*) \\ &\geq \min_{\lambda \in I, \lambda' \in I'} \min_{\lambda = \sum_{i < k \in S} X_{ik}, \lambda' = \sum_{i < k \in S} X_{0i}X_{0k}} \alpha_{|S|}(\lambda, \lambda') \\ &= \alpha_{|S|}, \end{aligned}$$

where again $-1 \leq X_{ik}, X_{0i}, X_{0k} \leq 1$ satisfy the triangle inequalities. \square

Since $E[\widehat{U}(S)] \geq \widehat{f}^* \geq \alpha_{|S|} z_j^*$ if $z_j^* = 0$, we have by Lemma 4.2.1

Theorem 4.2.2. The expected contribution of the edge S_j to the cut generated by executing Steps 1-4 of our algorithm satisfies

$$E[\widehat{U}(S_j)] \geq \alpha_{|S_j|} z_j^*.$$

To establish the bound on the expected contribution of an edge after executing Step 5, we need to compute $\alpha_{|S|}$ efficiently. However, it is not easy to compute it directly. Therefore we strive to provide a good lower bound for $\alpha_{|S|}$. The key to establish the lower bound on $\alpha_{|S|}$ is to obtain a lower bound on $h_1(\lambda)$ for $\lambda \in I$ first, since it is easy to calculate $h_2(\lambda') = \sum_{i < k \in S} (1 - \theta'^2 X_{0i} X_{0k}) = N_{|S|} - \theta'^2 \lambda'$ for $\lambda' = \sum_{i < k \in S} X_{0i} X_{0k} \in I'$. A lower bound on $h_1(\lambda)$ is obtained by applying the following lemma which has already been proven in [70] using the definitions of concavity and convexity.

Lemma 4.2.3. Given $\theta \in (0, 1]$, define

$$d(x) = 1 - (2/\pi) \arcsin(\theta x).$$

Then $d(x)$ is convex in $[-1, 0]$ and concave in $[0, 1]$. Therefore, $d(x') + d(y') \leq d(x) + d(y)$ for $-1 \leq x \leq x' \leq y' \leq y \leq 0$ with $x + y = x' + y'$, and $d(x) + d(y) \leq d(x') + d(y')$ for $0 \leq x \leq x' \leq y' \leq y \leq 1$ with $x + y = x' + y'$.

With the definition of $d(x)$ in the above lemma, it is easily seen that

$$h_1(\lambda) \geq h'_1(\lambda) := \min \sum_{i < k \in S} d(X_{ik}), \quad (4.16)$$

where the minimum is taken over all X such that $\lambda = \sum_{i < k \in S} X_{ik}$, and $-1 \leq X_{ik} \leq 1$ for $1 \leq i < k \leq |S|$.

Let $X = (X_{ij})$ be a minimum solution to (4.16). Since $d(x)$ is convex in $[-1, 0]$ and concave in $[0, 1]$ by Lemma 4.2.3, we may choose a minimum solution X to (4.16) so that at most one entry of X belongs to $(0, 1)$ and all entries of X in $[-1, 0]$ are equal. This nice structural property of the minimum solution helps us calculate $h'_1(\lambda)$ easily.

To derive a lower bound on $\alpha_{|S|}$, we need the following lemma. Let $\overline{N} =$

$(|S| - 4)(|S| - 1)/2$, N be a nonnegative integer and

$$\begin{aligned}
f_1(x, N) &= N_{|S|} + N(2/\pi) \arcsin(\theta) - (N_{|S|} - N - 1)(2/\pi) \arcsin(\theta) - \\
&\quad (2/\pi) \arcsin(\theta(x - N_{|S|} + 2N + 1)), \\
f_2(x, N) &= N_{|S|} - (N_{|S|} - N)(2/\pi) \arcsin(\theta) - N(2/\pi) \arcsin(\theta x), \\
l_1 &= \min_{N \in [0, N_{|S|} - 1]} \left\{ \min_{\lambda \in I(N), \lambda' \in I'} [\nu f_1(\lambda, N) + (1 - \nu)h_2(\lambda')] \right\} \\
&\quad \text{where } I(N) = [-\lfloor |S|/2 \rfloor, \overline{N}] \cap (N_{|S|} - 2N - 1, N_{|S|} - 2N), \\
l_2 &= \min_{N \in [1, N_{|S|}]} \left\{ \min_{\lambda \in J(N), \lambda' \in I'} [\nu f_2\left(\frac{\lambda - (N_{|S|} - N)}{N}, N\right) + (1 - \nu)h_2(\lambda')] \right\} \\
&\quad \text{where } J(N) = [-\lfloor |S|/2 \rfloor, \overline{N}] \cap [N_{|S|} - 2N, N_{|S|} - N], \\
l_3 &= \min_{N \in [0, N_{|S|} - 1]} \left\{ \min_{\lambda \in K(N), \lambda' \in I'} \frac{\nu f_1(\lambda, N) + (1 - \nu)h_2(\lambda')}{N_{|S|} - \lambda} \right\} \\
&\quad \text{where } K(N) = [\overline{N}, N_{|S|}] \cap (N_{|S|} - 2N - 1, N_{|S|} - 2N), \\
l_4 &= \min_{N \in [1, N_{|S|}]} \left\{ \min_{\lambda \in M(N), \lambda' \in I'} \frac{\nu f_2\left(\frac{\lambda - (N_{|S|} - N)}{N}, N\right) + (1 - \nu)h_2(\lambda')}{N_{|S|} - \lambda} \right\} \\
&\quad \text{where } M(N) = [\overline{N}, N_{|S|}] \cap [N_{|S|} - 2N, N_{|S|} - N].
\end{aligned}$$

Lemma 4.2.4. Given fixed parameters $\theta, \theta' \in (0, 1]$ and $\nu \in [0, 1]$, we have

$$\alpha_{|S|} = \min_{\lambda \in I, \lambda' \in I'} \alpha(\lambda, \lambda') \geq \min\left\{ \frac{1}{2L}l_1, \frac{1}{2L}l_2, \frac{|S| - 1}{L}l_3, \frac{|S| - 1}{L}l_4 \right\}.$$

where again $\lambda = \sum_{i < k \in S} X_{ik}$, $\lambda' = \sum_{i < k \in S} X_{0i}X_{0k}$, and $-1 \leq X_{0i}, X_{0k}, X_{ik} \leq 1$ satisfy the triangle inequalities.

Proof. First, we show that one can choose a minimum solution X to (4.16) (in Section 3) so that at most one entry of X belongs to $(0, 1)$ and all entries of X in $[-1, 0]$ are equal. Let $X = (X_{ij})$ be a minimum solution to (4.16). By Lemma 4.2.3, $d(x)$ is convex in $[-1, 0]$ and concave in $[0, 1]$.

Then we may choose a minimum solution X to (4.16) so that at most one entry of X belongs to $(0, 1)$. For, suppose $X_{ik} < X_{st}$ and $X_{ik}, X_{st} \in (0, 1)$. Let $\epsilon = \min\{1 - X_{st}, X_{ik}\}$. By Lemma 4.2.3, replacing X_{ik}, X_{st} with $X_{ik} - \epsilon, X_{st} + \epsilon$, respectively, does not increase $\sum_{i < k \in S} d(X_{ik})$ (while maintaining the equation constraint). So we obtain a new minimum solution to (4.16) with fewer entries in $(0, 1)$. Continuing this process, we arrive at the desired solution X .

We may further choose X so that all entries of X in $[-1, 0]$ are equal. For, let $X_{ik}, X_{st} \in [-1, 0]$ with $X_{ik} < X_{st}$. Now replacing X_{ik} and X_{st} with $(X_{ik} + X_{st})/2$ we obtain a minimum solution to (4.16) with smaller number of pairs of non-equal entries of X in $[-1, 0]$. Continuing this process, we obtain the desired X .

If there is some entry $X_{pq} \in (0, 1)$, we claim that all other entries of X must be in $\{-1, 1\}$. For, otherwise, let $X_{pq} \in (0, 1)$ and assume that there are some entries of X in $(-1, 0]$ (which must be equal by the choice of X above). Since X is a minimum solution to (4.16), every $X_{ik} \in (-1, 1)$ must satisfy the Karush-Kuhn-Tucker condition: there exists a Lagrange multiplier t for the equality constraint $\lambda = \sum_{i < k \in S}$ in (4.16) such that every $X_{ik} \in (-1, 1)$ satisfies

$$-2\theta/(\pi\sqrt{1 - (\theta X_{ik})}) = t.$$

So

$$X_{ik} = (\pm\sqrt{1 - (2\theta/\pi t)^2})/\theta. \tag{4.17}$$

Now suppose $X_{st} \in (-1, 0]$. Then $X_{st} < X_{pq}$ and both X_{st} and X_{pq} belong to $(-1, 1)$. Therefore, we have $X_{st} = (-\sqrt{1 - (2\theta/\pi t)^2})/\theta$, $X_{pq} = (\sqrt{1 - (2\theta/\pi t)^2})/\theta$, and

$$1 - (2\theta/\pi t)^2 > 0. \tag{4.18}$$

Then

$$\begin{aligned} & d(X_{st}) + d(X_{pq}) \\ &= 1 - (2/\pi) \arcsin(\theta X_{st}) + 1 - (2/\pi) \arcsin(\theta X_{pq}) \\ &= 1 - (2/\pi) \arcsin(-\theta\sqrt{1 - (2\theta/\pi t)^2}) + 1 - (2/\pi) \arcsin(\theta\sqrt{1 - (2\theta/\pi t)^2}) \\ &= 2 - 2 \cdot (2/\pi) \arcsin(0) \\ &= 2 \cdot d(0). \end{aligned}$$

By replacing both X_{st} and X_{pq} with 0 and keeping all other entries of X unchanged, we obtain a new minimal solution with some entries in $(-1, 1)$ not satisfying (4.17) and (4.18), a contradiction. Thus, if there is some entry $X_{pq} \in (0, 1)$, then all other entries of X must be in $\{-1, 1\}$.

The above properties of the minimum solution will help us calculate $h'_1(\lambda)$ (in (4.16)) and then obtain a lower bound on $\alpha_{|S|}$. Recall that $z'(\lambda^*) = \min\{1, \frac{N_{|S|} - \lambda^*}{2(|S| - 1)}\}$ from (4.15) in the proof of Lemma 4.2.1. So $z'(\lambda^*) = 1$ when

$\lambda^* \in I_1 := [-\lfloor |S|/2 \rfloor, \overline{N}]$, where $\overline{N} = (|S| - 4)(|S| - 1)/2$; and $z'(\lambda^*) = \frac{N_{|S|} - \lambda^*}{2(|S| - 1)}$ when $\lambda^* \in I_2 := [\overline{N}, N_{|S|})$, where $N_{|S|} = |S|(|S| - 1)/2$. To obtain a lower bound on $\alpha_{|S|}$, we next distinguish two cases.

Subcase 2.1. $\lambda^* \in I_1$. In this case we have $z'(\lambda^*) = 1$.

From the proof of Lemma 4.2.1, it follows

$$\begin{aligned} & \widehat{f}^*/z_j^* \\ & \geq \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} (X_{0i} X_{0k})} \alpha_{|S|}(\lambda^*, \lambda'^*) \\ & = \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} (X_{0i} X_{0k})} (\nu h_1(\lambda^*) + (1 - \nu) h_2(\lambda'^*)) / (2L z'(\lambda^*)) \\ & = \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} (X_{0i} X_{0k})} (\nu h'_1(\lambda^*) + (1 - \nu) h_2(\lambda'^*)) (2L), \end{aligned}$$

where $-1 \leq X_{ik}, X_{0i}, X_{0k} \leq 1$ satisfy the triangle inequalities.

From (4.16), we have

$$h'_1(\lambda^*) = \min \sum_{i < k \in S} d(X_{ik}), \quad (4.19)$$

where the minimum is taken over all X such that $\lambda^* = \sum_{i < k \in S} X_{ik}$, and $-1 \leq X_{ik} \leq 1$ for $1 \leq i < k \leq |S|$.

As above, we may choose a minimum solution X to (4.19) so that at most one entry of X belongs to $(0, 1)$ and all entries of X in $[-1, 0]$ are equal, and if some entry $X_{pq} \in (0, 1)$ then all other entries of X are in $\{-1, 1\}$. Thus, there exists a nonnegative integer $N_1 \in [0, N_{|S|} - 1]$ such that

$$\lambda^* = N_1(-1) + X_{pq} + (N_{|S|} - N_1 - 1).$$

Note that $N_1 \neq N_{|S|}$ (because of X_{pq}). Since $X_{pq} \in (0, 1)$, $\lambda \in I(N_1) := I_1 \cap (N_1(-1) + 0 + (N_{|S|} - N_1 - 1), N_1(-1) + 1 + (N_{|S|} - N_1 - 1)) = I_1 \cap (N_{|S|} - 2N_1 - 1, N_{|S|} - 2N_1)$. By definitions of $h'_1(\lambda^*)$ and $d(x)$, we have

$$\begin{aligned} h'_1(\lambda^*) &= N_{|S|} - N_1(2/\pi) \arcsin(\theta(-1)) - (N_{|S|} - N_1 - 1)(2/\pi) \arcsin(\theta) \\ &\quad - (2/\pi) \arcsin(\theta X_{pq}) \\ &= N_{|S|} + N_1 2/\pi \arcsin(\theta) - (N_{|S|} - N_1 - 1) 2/\pi \arcsin(\theta) \\ &\quad - 2/\pi \arcsin(\theta(\lambda^* + N_1 - (N_{|S|} - N_1 - 1))) \\ &= f_1(\lambda^*, N_1). \end{aligned}$$

It is easily seen that

$$\begin{aligned}
 & \widehat{f^*}/z_j^* \\
 \geq & \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} X_{0i}X_{0k}} (\nu h'_1(\lambda^*) + (1 - \nu)h_2(\lambda'^*))/(2L), \\
 = & \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} X_{0i}X_{0k}} (\nu f_1(\lambda^*, N_1) + (1 - \nu)h_2(\lambda'^*))/(2L), \\
 \geq & \min_{\lambda \in I(N_1), \lambda' \in I'} \min_{\lambda = \sum_{i < k \in S} X_{ik}, \lambda' = \sum_{i < k \in S} X_{0i}X_{0k}} \{\nu f_1(\lambda, N_1) \\
 & + (1 - \nu)h_2(\lambda')\}/(2L), \\
 \geq & \frac{1}{2L} \min_{N \in [0, N_{|S|} - 1]} \left\{ \min_{\lambda \in I(N), \lambda' \in I'} [\nu f_1(\lambda, N) + (1 - \nu)h_2(\lambda')] \right\} \\
 = & \frac{1}{2L} l_1.
 \end{aligned}$$

where $\lambda = \sum_{i < k \in S} X_{ik}$, $\lambda' = \sum_{i < k \in S} X_{0i}X_{0k}$, and $-1 \leq X_{ik}, X_{0i}, X_{0k} \leq 1$ satisfy the triangle inequalities.

Now assume that there is no entry $X_{pq} \in (0, 1)$. Then by the choice of X as above, except the entries which are equal to 1, all other entries of X are in $[-1, 0]$ and equal. Let $X_{ik} \in [-1, 0]$ be an arbitrary entry of X . Then there exists a nonnegative integer $N'_1 \in [1, N_{|S|}]$ such that

$$\lambda^* = N'_1 X_{ik} + (N_{|S|} - N'_1).$$

Note that $N'_1 \neq 0$, since $\lambda^* < N_{|S|}$. Since $X_{ik} \in [-1, 0]$, we have $\lambda^* \in J(N'_1) := I_1 \cap [N'_1(-1) + (N_{|S|} - N'_1), N_{|S|} - N'_1] = I_1 \cap [N_{|S|} - 2N'_1, N_{|S|} - N'_1]$. By definitions of $h'_1(\lambda^*)$ and $d(x)$, we find

$$\begin{aligned}
 h'_1(\lambda^*) &= N_{|S|} - N'_1(2/\pi) \arcsin(\theta X_{ik}) - (N_{|S|} - N'_1)(2/\pi) \arcsin(\theta) \\
 &= N_{|S|} - N'_1(2/\pi) \arcsin\left(\theta \frac{(\lambda^* - (N_{|S|} - N'_1))}{N'_1}\right) \\
 &\quad - (N_{|S|} - N'_1)(2/\pi) \arcsin(\theta) \\
 &= f_2\left(\frac{\lambda^* - (N_{|S|} - N'_1)}{N'_1}, N'_1\right).
 \end{aligned}$$

Thus,

$$\begin{aligned}
& \widehat{f}^*/z_j^* \\
& \geq \min_{\lambda^*=\sum_{i<k\in S} X_{ik}, \lambda'^*=\sum_{i<k\in S} X_{0i}X_{0k}} (\nu h_1'(\lambda^*) + (1-\nu)h_2(\lambda'^*))/(2L), \\
& = \min_{\lambda^*=\sum_{i<k\in S} X_{ik}, \lambda'^*=\sum_{i<k\in S} X_{0i}X_{0k}} \left(\nu f_2 \left(\frac{\lambda^* - (N_{|S|} - N_1')}{N_1'}, N_1' \right) + (1-\nu)h_2(\lambda'^*) \right) / (2L), \\
& \geq \min_{\lambda \in J(N_1'), \lambda' \in I'} \min_{\lambda=\sum_{i<k\in S} X_{ik}, \lambda'=\sum_{i<k\in S} X_{0i}X_{0k}} \left(\nu f_2 \left(\frac{\lambda - (N_{|S|} - N_1')}{N_1'}, N_1' \right) + (1-\nu)h_2(\lambda') \right) / (2L), \\
& \geq \min_{N \in [1, N_{|S|}]} \left\{ \min_{\lambda \in J(N), \lambda' \in I'} \left(\nu f_2 \left(\frac{\lambda - (N_{|S|} - N)}{N}, N \right) + (1-\nu)h_2(\lambda') \right) \right\} / (2L) \\
& = \frac{1}{2L} l_2.
\end{aligned}$$

where $\lambda = \sum_{i<k\in S} X_{ik}$, $\lambda' = \sum_{i<k\in S} X_{0i}X_{0k}$, and $-1 \leq X_{ik}, X_{0i}, X_{0k} \leq 1$ satisfy the triangle inequalities.

Subcase 2.2. $\lambda^* \in I_2$. In this case, we have $z'(\lambda^*) = \frac{N_{|S|} - \lambda^*}{2(|S| - 1)}$.

From the proof of Lemma 4.2.1, we have

$$\begin{aligned}
& \widehat{f}^*/z_j^* \\
& \geq \min_{\lambda^*=\sum_{i<k\in S} X_{ik}, \lambda'^*=\sum_{i<k\in S} X_{0i}X_{0k}} \alpha_{|S|}(\lambda^*, \lambda'^*) \\
& = \min_{\lambda^*=\sum_{i<k\in S} X_{ik}, \lambda'^*=\sum_{i<k\in S} X_{0i}X_{0k}} (\nu h_1(\lambda^*) + (1-\nu)h_2(\lambda'^*))/(2Lz'(\lambda^*)) \\
& \geq \min_{\lambda^*=\sum_{i<k\in S} X_{ik}, \lambda'^*=\sum_{i<k\in S} X_{0i}X_{0k}} (\nu h_1'(\lambda^*) + (1-\nu)h_2(\lambda'^*)) \frac{|S| - 1}{L(N_{|S|} - \lambda^*)},
\end{aligned}$$

where $-1 \leq X_{ik}, X_{0i}, X_{0k} \leq 1$ satisfy the triangle inequalities.

Similarly, we may choose a minimum solution X to (4.19) so that at most one entry of X belongs to $(0, 1)$ and all entries of X in $[-1, 0]$ are equal, and

if some entry $X_{pq} \in (0, 1)$ then all other entries of X are in $\{-1, 1\}$. Thus in this case there exists a nonnegative integer $N_2 \in [0, N_{|S|} - 1]$ such that

$$\lambda^* = N_2(-1) + X_{pq} + (N_{|S|} - N_2 - 1).$$

Note that $N_2 \neq N_{|S|}$ (because of X_{pq}). Since $X_{pq} \in (0, 1)$, it follows $\lambda^* \in K(N_2) := I_2 \cap (N_2(-1) + N_{|S|} - N_2 - 1, N_2(-1) + 1 + N_{|S|} - N_2 - 1) = I_2 \cap (N_{|S|} - 2N_2 - 1, N_{|S|} - 2N_2)$. By definitions of $h'_1(\lambda^*)$ and $d(x)$, we have

$$\begin{aligned} h'_1(\lambda^*) &= N_{|S|} - N_2(2/\pi) \arcsin(\theta(-1)) - (N_{|S|} - N_2 - 1)(2/\pi) \arcsin(\theta) \\ &\quad - (2/\pi) \arcsin(\theta X_{pq}) \\ &= N_{|S|} + N_2(2/\pi) \arcsin(\theta) - (N_{|S|} - N_2 - 1)(2/\pi) \arcsin(\theta) \\ &\quad - (2/\pi) \arcsin(\theta(\lambda^* + N_2 - (N_{|S|} - N_2 - 1))), \\ &= f_1(\lambda^*, N_2). \end{aligned}$$

It is easily seen that

$$\begin{aligned} &\widehat{f^*}/z_j^* \\ &\geq \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} X_{0i}X_{0k}} (\nu h'_1(\lambda^*) + (1 - \nu)h_2(\lambda'^*)) \frac{|S| - 1}{L(N_{|S|} - \lambda^*)}, \\ &= \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} X_{0i}X_{0k}} (\nu f_1(\lambda^*, N_2) + (1 - \nu)h_2(\lambda'^*)) \frac{|S| - 1}{L(N_{|S|} - \lambda^*)}, \\ &\geq \min_{\lambda \in K(N_2), \lambda' \in I'} \min_{\lambda = \sum_{i < k \in S} X_{ik}, \lambda' = \sum_{i < k \in S} X_{0i}X_{0k}} (\nu f_1(\lambda, N_2) + (1 - \nu)h_2(\lambda')) \frac{|S| - 1}{L(N_{|S|} - \lambda)}, \\ &\geq \frac{|S| - 1}{L} \min_{N \in [0, N_{|S|} - 1]} \left\{ \min_{\lambda \in K(N), \lambda' \in I'} \frac{\nu f_1(\lambda, N) + (1 - \nu)h_2(\lambda')}{N_{|S|} - \lambda} \right\} \\ &= \frac{|S| - 1}{L} l_3. \end{aligned}$$

where $\lambda = \sum_{i < k \in S} X_{ik}$, $\lambda' = \sum_{i < k \in S} X_{0i}X_{0k}$, and $-1 \leq X_{ik}, X_{0i}, X_{0k} \leq 1$ satisfy the triangle inequalities.

Now assume that there is no entry $X_{pq} \in (0, 1)$. Then by the choice of X , except the entries which are equal to 1, all other entries of X are in $[-1, 0]$

and equal. Let $X_{ik} \in [-1, 0]$ be an arbitrary entry of X . Hence there exists a nonnegative integer $N'_2 \in [1, N_{|S|}]$ such that

$$\lambda^* = N'_2 X_{ik} + (N_{|S|} - N'_2).$$

Note that $N'_2 \neq 0$ because $\lambda^* < N_{|S|}$. Since $X_{ik} \in [-1, 0]$, it follows $\lambda^* \in M(N'_2) := I_2 \cap [N'_2(-1) + (N_{|S|} - N'_2), N_{|S|} - N'_2] = I_2 \cap [N_{|S|} - 2N'_2, N_{|S|} - N'_2]$. By definitions of $h'_1(\lambda^*)$ and $d(x)$, we have

$$\begin{aligned} h'_1(\lambda^*) &= N_{|S|} - (N_{|S|} - N'_2)(2/\pi) \arcsin(\theta) - N'_2(2/\pi) \arcsin(\theta X_{ik}) \\ &= N_{|S|} - (N_{|S|} - N'_2)(2/\pi) \arcsin(\theta) \\ &\quad - N'_2(2/\pi) \arcsin\left(\frac{\theta(\lambda^* - (N_{|S|} - N'_2))}{N'_2}\right) \\ &= f_2\left(\frac{\lambda^* - (N_{|S|} - N'_2)}{N'_2}, N'_2\right). \end{aligned}$$

Hence,

$$\begin{aligned} &\widehat{f^*}/z_j^* \\ &\geq \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} X_{0i} X_{0k}} (\nu h'_1(\lambda^*) + (1 - \nu) h_2(\lambda'^*)) \frac{|S| - 1}{L(N_{|S|} - \lambda^*)}, \\ &= \min_{\lambda^* = \sum_{i < k \in S} X_{ik}, \lambda'^* = \sum_{i < k \in S} X_{0i} X_{0k}} \left(\nu f_2\left(\frac{\lambda^* - (N_{|S|} - N'_2)}{N'_2}, N'_2\right) + (1 - \nu) h_2(\lambda'^*) \right) \frac{|S| - 1}{L(N_{|S|} - \lambda^*)}, \\ &\geq \min_{\lambda \in M(N'_2), \lambda' \in I'} \min_{\lambda = \sum_{i < k \in S} X_{ik}, \lambda' = \sum_{i < k \in S} X_{0i} X_{0k}} \left\{ \nu f_2\left(\frac{\lambda - (N_{|S|} - N'_2)}{N'_2}, N'_2\right) \right. \\ &\quad \left. + (1 - \nu) h_2(\lambda') \right\} \frac{|S| - 1}{L(N_{|S|} - \lambda)}, \\ &\geq \frac{|S| - 1}{L} \min_{N \in [1, N_{|S|}]} \left\{ \min_{\lambda \in M(N), \lambda' \in I'} \frac{\nu f_2\left(\frac{\lambda - (N_{|S|} - N)}{N}, N\right) + (1 - \nu) h_2(\lambda')}{N_{|S|} - \lambda} \right\} \\ &= \frac{|S| - 1}{L} l_4. \end{aligned}$$

where $\lambda = \sum_{i < k \in S} X_{ik}$, $\lambda' = \sum_{i < k \in S} X_{0i} X_{0k}$, and $-1 \leq X_{ik}, X_{0i}, X_{0k} \leq 1$ satisfy the triangle inequalities.

Summing up, by the definition of $\alpha_{|S|}$, we have

$$\alpha_{|S|} \geq \min\left\{\frac{1}{2L}l_1, \frac{1}{2L}l_2, \frac{|S|-1}{L}l_3, \frac{|S|-1}{L}l_4\right\}.$$

□

Note that for any fixed $|S|$ and nonnegative integer N , the quantities l_1, l_2, l_3 and l_4 can be evaluated numerically. (For example, we used the optimization toolbox of Matlab.)

It is easily seen that our scheme enumerates at most $O(|S|^2)$ nonnegative integers N with $N \leq N_{|S|}$ for calculating l_1, l_2, l_3 or l_4 . We note that in the case when $|S| = 3$, it is not hard to verify that $\alpha_3 = 0.90871$ by setting $\theta = 0.9789$ and $\nu = 1$ which achieves the same bound as in Zwick [148] for max E3-set splitting as mentioned in the Introduction.

4.3 Bounding $\mathbb{E}[\omega(\tilde{V}_1, V \setminus \tilde{V}_1)]$ after Step 5

The aim of this section is to derive a bound on $\mathbb{E}[\omega(\tilde{V}_1, V \setminus \tilde{V}_1)]$ in terms of ω^{opt} using the results in the previous section. After executing Step 5 in our SDP-algorithm for MHC-LU, we have the second partition $V = \tilde{V}_1 \cup (V \setminus \tilde{V}_1)$.

Let $\tilde{U}(S_j)$ be an indicator random variable, defined as follows

$$\tilde{U}(S_j) = \begin{cases} 1, & \text{if } S_j \in (\tilde{V}_1, V \setminus \tilde{V}_1); \\ 0, & \text{otherwise.} \end{cases}$$

Recall the definition of α_2 (see (4.14)) and from Lemma 4.2.1 the definition of $\alpha_s := \alpha_{|S|}$ (with $|S| = s$) in the previous section. For a suitable nonnegative integer $M \geq 4$ and for any integer s with $2 \leq s < M$, let

$$\begin{aligned} r_s(p) &= \alpha_s(1 - p(1 - p)^{s-1} - p^{s-1}(1 - p)) + (1 - \alpha_s)(1 - p^s - (1 - p)^s) \text{ and} \\ r_M(p) &= 1 - p^M - (1 - p)^M. \end{aligned}$$

Note that p is the probability used in Step 5. The following lemma is due to Andersson and Engebretsen [11], who also used the first part of Step 5 as a

probabilistic postprocessing step to perturb an initial partition to a new one in their algorithm for Max Set Splitting.

Lemma 4.3.1. For any integer $M \geq 4$ and any $|S_j| \geq 2$,

$$\mathbb{E}[\tilde{U}(S_j)] \geq \begin{cases} r_{|S_j|}(p)z_j^*, & \text{if } |S_j| < M ; \\ r_M(p), & \text{otherwise.} \end{cases}$$

As an immediate consequence, we have

Corollary 4.3.2. For any integer $M \geq 4$ and any real number $0 \leq p \leq 1/2$,

$$\sum_{\{j:|S_j|\geq 2\}} \omega_j \mathbb{E}[\tilde{U}(S_j)] \geq \sum_{\{j:2\leq|S_j|<M\}} \omega_j r_{|S_j|}(p)z_j^* + \sum_{\{j:|S_j|\geq M\}} \omega_j r_M(p).$$

Let $\alpha^{(p)} = \min_{2 \leq s \leq M} r_s(p)$. Recall that $\omega^{SDP} = \sum_j \omega_j z_j^* \geq \max \sum_j \omega_j z_j = \omega^{opt}$. So we have the following lower bound on $\mathbb{E}[\omega(\tilde{V}_1, V \setminus \tilde{V}_1)]$.

Corollary 4.3.3. For any integer $M \geq 4$ and any real number $0 \leq p \leq 1/2$,

$$\begin{aligned} \mathbb{E}[\omega(\tilde{V}_1, V \setminus \tilde{V}_1)] &= \sum_{\{j:|S_j|\geq 2\}} \omega_j \mathbb{E}[\tilde{U}(S_j)] \\ &\geq \sum_{\{j:2\leq|S_j|<M\}} \omega_j r_{|S_j|}(p)z_j^* + \sum_{\{j:|S_j|\geq M\}} \omega_j r_M(p)z_j^* \\ &\geq \alpha^{(p)} \omega^{SDP} \\ &\geq \alpha^{(p)} \omega^{opt}. \end{aligned}$$

For the hyperedge S_j with size $|S_j| = 2$, the ratio $r_2(p)$ can be improved due to the following lemma given by Zhang et. al [143]. Let $\varphi_\theta(t) = (1/\pi) \arccos(\theta(1 - 2t))$ and let γ_θ be the minimizer of $\varphi_\theta(t)/t$ in the interval $(0, 1]$. By setting $x = 1 - 2t$ one can show (see [57]) that

$$\frac{\varphi_\theta(\gamma_\theta)}{\gamma_\theta} = \min_{-1 \leq x < 1} \frac{1 - (2/\pi) \arcsin(\theta x)}{1 - x}.$$

If the algorithm executes Step 3 and Step 5 which are also used in [143] (for Max Set Splitting), Zhang et. al [143] give an improved bound on the contribution to $\omega(\tilde{V}_1, V \setminus \tilde{V}_1)$ of the size 2 edges (in terms of their contribution to the objective function) as follows.

Lemma 4.3.4. For $|S_j| = 2$,

$$\omega_j \mathbb{E}[\tilde{U}(S_j)] \geq \min_{\gamma_\theta \leq x \leq 1} \frac{\varphi_\theta(x)(1-2p)^2 + (2p-2p^2)}{x} \omega_j z_j^*.$$

It is easily seen that $r_2(p) = \frac{1-\theta'^2}{2}(1-2p)^2 + (2p-2p^2)$ is the ratio for the hyperedge with size $S_j = 2$ if the algorithm executes Step 4 and Step 5. This follows by the definition of $r_2(p)$, since

$$\min_{-1 \leq X < 1, -1 \leq Y \leq 1} \frac{1 - \theta'^2 Y}{1 - X} = \frac{1 - \theta'^2}{2}$$

is the bound on the expected contribution of an edge with size $S_j = 2$ by executing Steps 1, Step 2 and Step 4 (see Section 3). Since our algorithm executes independently Step 3 with probability ν , and Step 4 with probability $1 - \nu$, we have

Corollary 4.3.5. For $|S_j| = 2$,

$$\begin{aligned} \omega_j \mathbb{E}[\tilde{U}(S_j)] &\geq \nu \min_{\gamma_\theta \leq x \leq 1} \frac{\varphi_\theta(x)(1-2p)^2 + (2p-2p^2)}{x} \omega_j z_j^* + \\ &\quad (1-\nu) \left(\frac{1-\theta'^2}{2}(1-2p)^2 + (2p-2p^2) \right) \omega_j z_j^*. \end{aligned}$$

We then let

$$\begin{aligned} r_2(p) &= \nu \min_{\gamma_\theta \leq x \leq 1} \frac{\varphi_\theta(x)(1-2p)^2 + (2p-2p^2)}{x} \\ &\quad + (1-\nu) \left(\frac{1-\theta'^2}{2}(1-2p)^2 + (2p-2p^2) \right), \end{aligned}$$

in the computation of $\alpha^{(p)}$. Given θ, θ', ν and M , for each $0 \leq p \leq 1/2$, we can compute (using the results in the previous section) the lower bounds on $\alpha^{(p)}$ which is the performance guarantee of the algorithm for MHC-LU when $\tau = 1$, i.e., max set splitting, by Corollary 4.3.3. For fast computations and easy verifications of computational results, we only use the triangle inequalities for computing α_3 and α_4 in the previous section and it is also sufficient for us to fix $M = 15$ for obtaining the desired performance guarantee of our algorithm from numerical results. We then obtain approximation ratio 0.7524

for max set splitting by setting the other parameters $\theta = 0.967$, $p = 0.1$ and $\nu = 1$ which improves the result 0.7499 in Zhang et al. [143]. The improvement is a consequence of our strengthened SDP relaxation and the generalized formula for the performance ratio (see Lemma 4.2.4). If we instead use $\alpha^{(p)} = \min_{3 \leq s \leq M} r_s(p)$ by setting the parameters $\theta = 0.961$, $\theta' = 0.1$, $p = 0.1$ and $\nu = 0.242$, we then obtain an approximation ratio 0.7741 for max set splitting when the minimum number of vertices in a hyperedge is 3. Our algorithm of combining the outward rotations of the random hyperplane rounding procedure with that of a linear randomized rounding procedure improves the result 0.75 in Gaur and Krishnamurti [55].

4.4 Bounding $\mathbb{E}[|\tilde{V}_1|(m - |\tilde{V}_1|)]$

We now proceed to establish a bound on $\mathbb{E}[|\tilde{V}_1|(m - |\tilde{V}_1|)]$, which will be used in the next section to obtain the performance guarantee of our SDP algorithm for MHC-LU when $0 \leq \tau < 1$.

First, let \tilde{X}_{ij} be an indicator random variable after the algorithm executes Step 5, defined as ($\tilde{V}_2 := V \setminus \tilde{V}_1$)

$$\tilde{X}_{ij} = \begin{cases} -1, & \text{if } i \in \tilde{V}_i \text{ and } k \in V \setminus \tilde{V}_i, \text{ for } i = 1, 2, \\ 1, & \text{otherwise.} \end{cases}$$

We then have the following lemma if the algorithm executes Step 3 and Step 5.

Lemma 4.4.1.

$$E[\tilde{X}_{ij}] = \begin{cases} 1, & \text{if } i=j, \\ (2/\pi)(1 - 2p)^2 \arcsin(\theta X_{ij}^*), & \text{otherwise.} \end{cases}$$

Proof. It is straightforward to see that if $i = j$, then $E[\tilde{X}_{ij}] = 1$. Now assume $i \neq j$. By Step 3 of our SDP algorithm in Section 2 and the definition of \hat{X}_{ij} in Section 3, we have

$$\begin{aligned} \Pr[\tilde{X}_{ij} = -1 | \hat{X}_{ij} = -1] &= p^2 + (1 - p)^2 = 1 - 2p(1 - p), \\ \Pr[\tilde{X}_{ij} = -1 | \hat{X}_{ij} = 1] &= 2p(1 - p) = 2p - 2p^2, \end{aligned}$$

and

$$\begin{aligned}
 E[\tilde{X}_{ij}|\hat{X}_{ij}] &= \Pr[\tilde{X}_{ij} = 1|\hat{X}_{ij}] - \Pr[\tilde{X}_{ij} = -1|\hat{X}_{ij}] \\
 &= 1 - \Pr[\tilde{X}_{ij} = -1|\hat{X}_{ij}] - \Pr[\tilde{X}_{ij} = -1|\hat{X}_{ij}] \\
 &= 1 - 2\Pr[\tilde{X}_{ij} = -1|\hat{X}_{ij}].
 \end{aligned}$$

From [57, 148], we have

$$\Pr(\hat{X}_{ij} = -1) = \arccos(\theta X_{ij}^*)/\pi.$$

Thus,

$$\begin{aligned}
 E[\tilde{X}_{ij}] &= \Pr(\hat{X}_{ij} = -1)E[\tilde{X}_{ij}|\hat{X}_{ij} = -1] + \Pr(\hat{X}_{ij} = 1)E[\tilde{X}_{ij}|\hat{X}_{ij} = 1] \\
 &= \Pr(\hat{X}_{ij} = -1)(1 - 2\Pr[\tilde{X}_{ij} = -1|\hat{X}_{ij} = -1]) + \\
 &\quad \Pr(\hat{X}_{ij} = 1)(1 - 2\Pr[\tilde{X}_{ij} = -1|\hat{X}_{ij} = 1]) \\
 &= \Pr(\hat{X}_{ij} = -1)(4p - 4p^2 - 1) + (1 - \Pr(\hat{X}_{ij} = -1))(1 - 4p + 4p^2) \\
 &= (1 - 2p)^2(1 - 2\Pr(\hat{X}_{ij} = -1)) \\
 &= (2/\pi)(1 - 2p)^2(\pi/2 - \arccos(\theta X_{ij}^*)) \\
 &= (2/\pi)(1 - 2p)^2 \arcsin(\theta X_{ij}^*).
 \end{aligned}$$

□

Next, for $1 \leq i \leq m$ let $p(i)$ be the probability of putting i into V_1 if the algorithm executes Step 4 and Step 5. We have the following lemma.

Lemma 4.4.2.

$$p(i) = \frac{1 + (1 - 2p)\theta' X_{0i}^*}{2}.$$

Proof. Recall that \hat{V}_1 is formed in Step 4 by adding, independently, each $1 \leq i \leq m$ with probability $\frac{1 + \theta' X_{0i}^*}{2}$, and V_1 is constructed in Step 5 by independently perturbing each element of \hat{V}_1 with probability p to the opposite subset and each element of the opposite subset to \hat{V}_1 . It is easily seen that

$$\begin{aligned}
 p(i) &= \frac{1 + \theta' X_{0i}^*}{2}(1 - p) + \frac{1 - \theta' X_{0i}^*}{2}p \\
 &= \frac{1 + (1 - 2p)\theta' X_{0i}^*}{2}.
 \end{aligned}$$

□

Recall that m is the number of vertices and $\tau = u/m$ is the unbalance parameter. Let

$$\begin{aligned}\chi &= \min_{-1 \leq x < 1} (2/\pi)(1-2p)^2 \frac{\arcsin(\theta) - \arcsin(\theta x)}{1-x}, \\ \varphi_\chi &= 1 - (2/\pi)(1-2p)^2 \arcsin(\theta),\end{aligned}$$

We can now present the bound on $\mathbb{E}(|\tilde{V}_1|(m - |\tilde{V}_1|))$ as follows.

Lemma 4.4.3. For sufficiently large m ,

$$\mathbb{E}[|\tilde{V}_1|(m - |\tilde{V}_1|)] \geq \frac{m^2}{4} (\nu(\varphi_\chi + (1 - \tau^2)\chi) + (1 - \nu)(1 - \theta'(1 - 2p)\tau)^2 + o(1))$$

Proof. Note that the algorithm executes independently Step 3 with probability ν , and Step 4 with probability $1 - \nu$. In Step 5, we have $\tilde{V}_1 = V_1$ if $|V_1| \geq m/2$, and $\tilde{V}_1 = V \setminus V_1$ otherwise. So $|\tilde{V}_1|(m - |\tilde{V}_1|) = |V_1|(m - |V \setminus V_1|)$ and

$$\begin{aligned}& \mathbb{E}[|\tilde{V}_1|(m - |\tilde{V}_1|)] \\ &= \nu \mathbb{E}[(1/4) \sum_{i,j} (1 - \tilde{X}_{ij})] + (1 - \nu) \sum_{1 \leq i \neq j \leq m} p(i)(1 - p(j)) \\ &= \nu(1/4) \sum_{i \neq j} (1 - (2/\pi)(1-2p)^2 \arcsin(\theta X_{ij}^*)) \\ & \quad + (1 - \nu) \sum_{1 \leq i \neq j \leq m} \left(\frac{1 + (1-2p)\theta' X_{0i}^*}{2} \cdot \frac{1 - (1-2p)\theta' X_{0j}^*}{2} \right) \\ &= \nu(1/4) \sum_{i \neq j} (\varphi_\chi + (2/\pi)(1-2p)^2 (\arcsin(\theta) - \arcsin(\theta X_{ij}^*))) \\ & \quad + (1 - \nu) \sum_{i=1}^m \left(\frac{1 + (1-2p)\theta' X_{0i}^*}{2} \cdot \sum_{j \neq i} \frac{1 - (1-2p)\theta' X_{0j}^*}{2} \right)\end{aligned}$$

$$\begin{aligned}
 &\geq \nu(1/4) \sum_{i \neq j} (\varphi_\chi + (1 - X_{ij}^*)\chi) + (1 - \nu) \sum_{i=1}^m \\
 &\quad \left(\frac{1 + (1 - 2p)\theta' X_{0i}^*}{2} \cdot \sum_{j=1}^m \left(\frac{1 - (1 - 2p)\theta' X_{0j}^*}{2} - \frac{1 - (1 - 2p)\theta' X_{0i}^*}{2} \right) \right) \\
 &\geq \nu(1/4) \left(\sum_{i \neq j} \varphi_\chi + (m^2 - u^2)\chi \right) \\
 &\quad + (1 - \nu) \left(\frac{m}{2} - \frac{(1 - 2p)\theta' u}{2} \right) \left(\sum_{j=1}^m \frac{1 - (1 - 2p)\theta' X_{0j}^*}{2} - 1 \right) \\
 &\geq \nu(1/4) (m(m - 1)\varphi_\chi + m^2(1 - \tau^2)\chi) \\
 &\quad + (1 - \nu)m \left(\frac{1}{2} - \frac{\theta'(1 - 2p)\tau}{2} \right) \left(\frac{m}{2} - \frac{\theta'(1 - 2p)u}{2} - 1 \right) \\
 &= \nu(1/4)m^2 ((1 - 1/m)\varphi_\chi + (1 - \tau^2)\chi) \\
 &\quad + (1 - \nu)m^2 \left(\frac{1}{2} - \frac{\theta'(1 - 2p)\tau}{2} \right) \left(\frac{1}{2} - \frac{\theta'(1 - 2p)\tau}{2} - \frac{1}{m} \right) \\
 &= (1/4)m^2 (\nu(\varphi_\chi + (1 - \tau^2)\chi) + (1 - \nu)(1 - \theta'(1 - 2p)\tau)^2 + o(1)) \\
 &\quad (\text{since } m \text{ is large}),
 \end{aligned}$$

where the second equality follows from Lemma 4.4.1 and Lemma 4.4.2, the first inequality follows from the definition of χ , and the second inequality follows from (4.7) (note that when $i = j$, $X_{ii}^* = 1$, so $\sum_{i \neq j} (1 - X_{ij}^*) \geq m^2 - u^2$), and (4.8) ($\sum_{i=1}^m X_{0i}^* \leq u$).

□

4.5 The quality of the SDP approximation algorithm.

Recall that $\tau = u/m \in [0, 1]$ is the unbalance parameter. For $\gamma > 0$, let

$$Z(\gamma) = \frac{\omega(\tilde{V}_1, V \setminus \tilde{V}_1)}{\omega^{opt}} + \gamma \frac{|\tilde{V}_1|(m - |\tilde{V}_1|)}{m^2}.$$

Also recall that \bar{V}_1 is given in Step 6 of our SDP algorithm. We now give

a lower bound on $\omega(\bar{V}_1, V \setminus \bar{V}_1)$ in terms of ω^{opt} , whenever $Z(\gamma)$ fulfills its expectation. This will be used to establish the bound on the quality of the SDP approximation algorithm for MHC-LU when $0 \leq \tau < 1$.

First, we show that $\omega(\tilde{V}_1, V \setminus \tilde{V}_1)/\omega^{opt} \leq 2/(1 + \tau)$. If $|\tilde{V}_1| \leq (m + u)/2$ then $\bar{V}_1 = \tilde{V}_1$ (see Step 6 of our algorithm in Section 2), and hence

$$\omega(\tilde{V}_1, V \setminus \tilde{V}_1)/\omega^{opt} \leq \omega(\tilde{V}_1, V \setminus \tilde{V}_1)/\omega(\bar{V}_1, V \setminus \bar{V}_1) = 1 \leq 2/(1 + \tau).$$

So we may assume $|\tilde{V}_1| > (m + u)/2$. Then by Lemma 4.1.1,

$$\omega(\tilde{V}_1, V \setminus \tilde{V}_1)/\omega(\bar{V}_1, V \setminus \bar{V}_1) \leq 2|\tilde{V}_1|/(m + u), \text{ and hence}$$

$$\begin{aligned} \omega(\tilde{V}_1, V \setminus \tilde{V}_1)/\omega^{opt} &\leq \omega(\tilde{V}_1, V \setminus \tilde{V}_1)/\omega(\bar{V}_1, V \setminus \bar{V}_1) \\ &\leq 2|\tilde{V}_1|/(m + u) \\ &\leq 2m/(m + u) \\ &= 2/(1 + \tau). \end{aligned}$$

Since

$$\gamma \frac{|\tilde{V}_1|(m - |\tilde{V}_1|)}{m^2} = \left(-\frac{|\tilde{V}_1|^2}{m^2} + \frac{|\tilde{V}_1|}{m} \right) \gamma \leq \frac{\gamma}{4},$$

we have

$$Z(\gamma) \leq A := \frac{2}{1 + \tau} + \frac{\gamma}{4}. \quad (4.20)$$

Let

$$\beta = (\nu(\varphi_\chi + (1 - \tau^2)\chi) + (1 - \nu)(1 - \theta'(1 - 2p)\tau)^2 + o(1)) / 4,$$

Then by Corollary 4.3.3 and Lemma 4.4.3, we have

$$\begin{aligned} \mathbb{E}(Z(\gamma)) &= \mathbb{E}(\omega(\tilde{V}_1, V \setminus \tilde{V}_1)/\omega^{opt}) + \gamma \mathbb{E}(|\tilde{V}_1|(m - |\tilde{V}_1|))/m^2 \\ &\geq B := \alpha^{(p)} + \gamma\beta. \end{aligned} \quad (4.21)$$

For any given $0 < \epsilon < 1$, let $r = \Pr\{Z(\gamma) < B - \epsilon(A - B)\}$. Then

$$\mathbb{E}[Z(\gamma)] < r(B - \epsilon(A - B)) + (1 - r) \max(Z(\gamma)). \quad (4.22)$$

From (4.20), (4.21), and (4.22), we have

$$r < \frac{A - B}{A - B + \epsilon(A - B)} = \frac{1}{1 + \epsilon}.$$

We run the proposed algorithm $K = O((1/\epsilon) \log(1/\epsilon))$ times, and denote the maximum value of $Z(\gamma)$ in these K runs by Z_K . Then

$$\Pr(Z_K < B - \epsilon(A - B)) \leq r^K < \left(\frac{1}{1 + \epsilon}\right)^K \leq \epsilon.$$

This implies that $Z_K \geq B - O(\epsilon) = \alpha^{(p)} + \gamma\beta - O(\epsilon)$ with probability at least $1 - \epsilon$. Since ϵ can be made arbitrarily small, almost surely

$$\frac{\omega(\tilde{V}_1, V \setminus \tilde{V}_1)}{\omega^{opt}} + \gamma \frac{|\tilde{V}_1|(m - |\tilde{V}_1|)}{m^2} = Z(\gamma) \geq \alpha^{(p)} + \gamma\beta - o(1). \quad (4.23)$$

Let $|\tilde{V}_1| = qm$, where $1/2 \leq q \leq 1$ (see Step 5 in Section 2). Then from (4.23), almost surely,

$$\frac{\omega(\tilde{V}_1, V \setminus \tilde{V}_1)}{\omega^{opt}} \geq \alpha^{(p)} + \gamma\beta - \gamma q(1 - q) - o(1).$$

By Lemma 4.1.1, if $|\tilde{V}_1| \leq \frac{m+u}{2}$, we obtain $\omega(\bar{V}_1, V \setminus \bar{V}_1) = \omega(\tilde{V}_1, V \setminus \tilde{V}_1)$; otherwise, we have (almost surely)

$$\begin{aligned} \frac{\omega(\bar{V}_1, V \setminus \bar{V}_1)}{\omega^{opt}} &\geq \frac{m + u}{2|\tilde{V}_1|} \cdot \frac{\omega(\tilde{V}_1, V \setminus \tilde{V}_1)}{\omega^{opt}} \\ &= \frac{1 + \tau}{2q} \cdot \frac{\omega(\tilde{V}_1, V \setminus \tilde{V}_1)}{\omega^{opt}} \\ &\geq \frac{1 + \tau}{2q} (\alpha^{(p)} + \gamma\beta - \gamma q(1 - q) - o(1)). \end{aligned}$$

Write

$$\mu(q) = \alpha^{(p)} + \gamma\beta - \gamma q(1 - q).$$

Then almost surely,

$$\frac{\omega(\bar{V}_1, V \setminus \bar{V}_1)}{\omega^{opt}} \geq \min \left\{ \mu(q) - o(1), \frac{1 + \tau}{2q} \mu(q) - o(1) \right\}.$$

We now compute the lower bound on $\omega(\overline{V}_1, V \setminus \overline{V}_1)$ in terms of ω^{opt} , whenever $Z(\gamma)$ fulfills its expectation. Let

$$\begin{aligned}\gamma_0 &= \frac{\alpha^{(p)}}{2\beta} \left(\frac{1}{\sqrt{1-4\beta}} - 1 \right), \\ \gamma_1 &= \frac{4\alpha^{(p)}}{(1+2\tau)^2 - 4\beta}, \\ \gamma_2 &= \frac{\alpha^{(p)}(2-2\tau)}{1-2(1-\tau)\beta}.\end{aligned}$$

Note that γ_0, γ_1 and γ_2 are chosen to optimize certain functions. Now we have the following lemma.

Lemma 4.5.1. If the random variable fulfills its expectation, i.e., $Z(\gamma) \geq \alpha^{(p)} + \gamma\beta - o(1)$, then almost surely, $\omega(\overline{V}_1, V \setminus \overline{V}_1)/\omega^{opt} \geq \overline{R}(\tau) - o(1)$, where

$$\overline{R}(\tau) = \begin{cases} \overline{R}_0, & \text{if } \tau = 0, \beta < 1/4 \text{ and } \frac{\alpha^{(p)}}{1-\beta} \leq \gamma = \gamma_0 \leq \frac{4\alpha^{(p)}}{1-4\beta}, \\ \overline{R}_1, & \text{if } 0 < \tau \leq 1/2, \text{ and } \gamma = \gamma_1 \\ \overline{R}_2, & \text{if } 1/2 < \tau < 1, \text{ and } \gamma = \gamma_2, \end{cases}$$

and

$$\begin{aligned}\overline{R}_0 &= \frac{\alpha^{(p)}}{1 + \sqrt{1-4\beta}}, \\ \overline{R}_1 &= \frac{\alpha^{(p)}(1+\tau)\tau}{(1+2\tau)^2/4 - \beta}, \\ \overline{R}_2 &= \frac{\alpha^{(p)}(1+\tau)}{2-4(1-\tau)\beta}.\end{aligned}$$

Proof. When $\tau = 0$, we have

$$\begin{aligned}\omega(\overline{V}_1, V \setminus \overline{V}_1)/\omega^{opt} &\geq \min\{\mu(q) - o(1), \frac{1+\tau}{2q}\mu(q) - o(1)\}, \\ &= \min\{\mu(q) - o(1), \frac{\mu(q)}{2q} - o(1)\}, \\ &\geq \frac{\mu(q)}{2q} - o(1), \\ &= (\alpha^{(p)} + \gamma\beta - \gamma q(1-q))/(2q) - o(1), \\ &\geq \sqrt{\gamma(\alpha^{(p)} + \gamma\beta)} - \gamma/2 - o(1) \\ &\geq \alpha^{(p)}/(1 + \sqrt{1-4\beta}) - o(1) \\ &= \overline{R}_0 - o(1).\end{aligned}$$

The second inequality follows from $q \geq 1/2$. The third inequality follows from the simple calculus that $q = q^* = \sqrt{\alpha^{(p)}/\gamma + \beta}$ yields the minimal value for $\frac{\mu(q)}{2q}$ when $q \in [1/2, 1]$. It is easy to verify that $q^* \in [1/2, 1]$ if $4\beta < 1$ and $\alpha^{(p)}/(1 - \beta) \leq \gamma \leq 4\alpha^{(p)}/(1 - 4\beta)$. Substituting $\gamma = \gamma_0$ into the third inequality yields the maximal value for $\sqrt{\alpha^{(p)} + \gamma\beta} - \gamma/2$ and gives the last inequality. So we obtain the first result in the lemma.

Next, when $\tau > 0$, we select $\gamma = \gamma_1$ and $\gamma = \gamma_2$ when $0 < \tau \leq 1/2$ and $1/2 \leq \tau \leq 1$, respectively, to make the minimal values of $\omega(\overline{V}_1, V \setminus \overline{V}_1)/\omega^{opt}$ equal in both of the cases of $\frac{1+\tau}{2q} \leq 1$ and $\frac{1+\tau}{2q} \geq 1$.

Case 1. Suppose $\frac{1+\tau}{2q} \leq 1$. Then $q \in [\frac{(1+\tau)}{2}, 1]$. and $\omega(\overline{V}_1, V \setminus \overline{V}_1)/\omega^{opt} \geq \frac{1+\tau}{2q}\mu(q) - o(1)$. Simple calculations show that the function $\frac{1+\tau}{2q}\mu(q)$ has $q^* = \sqrt{\alpha^{(p)}/\gamma + \beta}$ as its stationary point and is decreasing in $(0, q^*)$.

If $0 < \tau \leq 1/2$ and $\gamma = \gamma_1$, then $q^* = (1 + 2\tau)/2$; so $\frac{(1+\tau)}{2} < q^* \leq 1$, and the function $\frac{1+\tau}{2q}\mu(q)$ achieves its minimal value at $q^* \in (\frac{(1+\tau)}{2}, 1]$. Thus,

$$\begin{aligned} \frac{\omega(\overline{V}_1, V \setminus \overline{V}_1)}{\omega^{opt}} &\geq \frac{1 + \tau}{2q^*}\mu(q^*) - o(1) \\ &= \left(2\sqrt{(\alpha^{(p)} + \gamma\beta)\gamma} - \gamma\right) \frac{1 + \tau}{2} - o(1) \\ &= \frac{\alpha^{(p)}(1 + \tau)\tau}{(1 + 2\tau)^2/4 - \beta} - o(1) \\ &= \overline{R}_1 - o(1). \end{aligned}$$

If $1/2 < \tau < 1$ and $\gamma = \gamma_2$, then

$$q^* = \sqrt{\frac{1}{2 - 2\tau}}.$$

Since $\tau > 1/2$, $q^* > 1$, it follows that the minimal value of $\frac{1+\tau}{2q}\mu(q)$ is achieved

at 1 by the fact that $\frac{1+\tau}{2q}\mu(q)$ is decreasing in $(0, q^*)$ and $q \in [\frac{1+\tau}{2}, 1]$. Then

$$\begin{aligned} \frac{\omega(\bar{V}_1, V \setminus \bar{V}_1)}{\omega^*} &\geq \frac{1+\tau}{2q}\mu(q) - o(1) \\ &\geq \frac{1+\tau}{2}\mu(1) - o(1) \\ &= \frac{\alpha^{(p)}(1+\tau)}{2-4(1-\tau)\beta} - o(1) \\ &= \bar{R}_2 - o(1). \end{aligned}$$

Case 2. Now assume $\frac{1+\tau}{2q} \geq 1$. Then $1/2 \leq q \leq \frac{1+\tau}{2}$, and $\omega(\bar{V}_1, V \setminus \bar{V}_1)/\omega^{opt} \geq \mu(q) - o(1)$. Simple calculations show that $\mu(q)$ has a minimum value at $q^* = \frac{1}{2}$.

If $0 < \tau \leq 1/2$ and $\gamma = \gamma_1$, we have

$$\begin{aligned} \omega(\bar{V}_1, V \setminus \bar{V}_1)/\omega^{opt} &\geq \mu(q^*) - o(1) \\ &= \alpha^{(p)} + \gamma\beta - \gamma/4 - o(1) \\ &= \frac{\alpha^{(p)}(1+\tau)\tau}{(1+2\tau)^2/4 - \beta} - o(1) \\ &= \bar{R}_1 - o(1). \end{aligned}$$

If $1/2 < \tau < 1$ and $\gamma = \gamma_2$, we have

$$\begin{aligned} \omega(\bar{V}_1, V \setminus \bar{V}_1)/\omega^{opt} &\geq \mu(q^*) - o(1) \\ &= \frac{\alpha^{(p)}(1+\tau)}{2-4(1-\tau)\beta} - o(1) \\ &= \bar{R}_2 - o(1). \end{aligned}$$

This completes the proof. □

For given fixed value τ , we select θ, θ', ν, p as solution of

$$\begin{aligned}
 & \max \bar{R}(\tau) && (4.24) \\
 \text{s.t.} & \quad 0 < \theta \leq 1 \\
 & \quad 0 < \theta' \leq 1 \\
 & \quad 0 \leq \nu \leq 1 \\
 & \quad 0 \leq p \leq 0.5
 \end{aligned}
 \tag{4.25}$$

Then, we have the following conclusion:

Theorem 4.5.2. For given fixed value τ , the worst-case performance ratio of the SDP-algorithm for approximating MHC-LU is at least $\bar{R}(\tau) - o(1)$.

The quality of the approximation depends on the optimization of $\bar{R}(\tau)$ over all parameters θ , θ' , ν and p . However, it is not easy to compute $\max \bar{R}(\tau)$ from (4.24). It would be interesting to have an efficient scheme for finding a best set of these parameters.

In practice, instead of solving (4.24) one could perform 1-dimensional searches on the parameters (with τ fixed) and obtain a lower bound $\bar{R}^*(\tau)$ on the performance guarantee of the algorithm. The use of numerical methods to evaluate the performance guarantee is rather common for SDP-based approximation algorithms, see [45, 49, 50, 52, 53, 57, 70, 71, 138, 143]. For fast computations and easy verifications of computational results, we may assign simple values to some of the parameters and simply perform 1-dimensional searches on the remaining parameters. This turns out to be sufficient to improve some previous known results mentioned in the introduction. Table 4.1 shows the values of parameters used in such computations of $\bar{R}^*(\tau)$ for MHC-LU for the same range of $0 \leq \tau < 1$ as that in Table 1.1. It seems that $\nu = 1$ is the optimal choices for the problem from numerical results. Table 4.2 shows the values of parameters used in such computations of $\bar{R}^*(\tau)$ for MHC-LU for the same range of $0 \leq \tau < 1$ as that in Table 1.2 when the minimum number of vertices in a hyperedge is 3. In such cases, our algorithm of combining the outward rotations of a random hyperplane rounding procedure with that of a linear randomized rounding procedure can always obtain better performance ratios from numerical results.

Note that $\alpha^{(p)} = r_2(p)$ holds when the hypergraph is 2-uniform (a standard graph). If we replace $\alpha^{(p)}$ by $r_2(p)$ in Lemma 4.5.1 and Theorem 4.5.2, we obtain the corresponding results for max cut with limited unbalance. It seems that $\nu = 1$ and $p = 0$ are the optimal choices for the problem from numerical results. Table 4.3 shows the values of parameters used in such computations of $\overline{R}^*(\tau)$ for MC-LU for the same range of τ as that in Table 1.3.

Table 4.1: Values of parameters used in the computation of $\overline{R}^*(\tau)$ for MHC-LU for some τ

τ	0	0.25	0.5	0.75	0.8	0.9	0.999
θ	0.923	0.967	0.967	0.967	0.967	0.967	0.967
p	0.1	0.1	0.1	0.1	0.1	0.1	0.1
γ	7.2165	2.2735	0.9506	0.4106	0.3216	0.1548	0.0015
β	0.2424	0.2315	0.2058	0.1692	0.1604	0.1395	0.11164
$\alpha^{(p)}$	0.7360	0.7524	0.7524	0.7524	0.7524	0.7524	0.7524
$\overline{R}^*(\tau)$	0.6271	0.7105	0.7130	0.7194	0.7245	0.7353	0.7522

Table 4.2: Values of parameters used in the computation of $\overline{R}^*(\tau)$ for MHC-LU for some τ when the minimum number of vertices in a hyperedge is 3

τ	0	0.25	0.5	0.75	0.9	0.999
θ	0.961	0.961	0.961	0.961	0.961	0.961
θ'	0.1	0.1	0.1	0.1	0.1	0.1
p	0.1	0.1	0.1	0.1	0.1	0.1
ν	0.242	0.242	0.242	0.242	0.242	0.242
γ	14.1884	2.3870	0.9993	0.4322	0.1612	0.0015
β	0.2475	0.2382	0.2253	0.2089	0.1974	0.1891
$\alpha^{(p)}$	0.7741	0.7741	0.7741	0.7741	0.7741	0.7741
$\overline{R}^*(\tau)$	0.7042	0.7459	0.7495	0.7564	0.7656	0.7740

We next give a detailed explanation for the reason of numerical improvements shown in Table 1.3. When $0 \leq \tau \leq 0.5$, we have the same performance

Table 4.3: Values of parameters used in the computation of $\overline{R}^*(\tau)$ for MC-LU for some τ

τ	0.6	0.7	0.8	0.85	0.9	0.95	0.9999
θ	0.98	0.99	0.99	0.99	1	1	1
γ	0.7987	0.5684	0.1272	0.2689	0.1772	0.0880	0.0002
β	0.1591	0.1272	0.0964	0.0795	0.0417	0.0214	0.0000
α^p	0.8713	0.8750	0.8750	0.8750	0.8786	0.8786	0.8786
$\overline{R}^*(\tau)$	0.7987	0.8052	0.8191	0.8291	0.8417	0.8584	0.8785

ratios as those in [53], as we use the same idea for estimating the values of performance ratios in these cases. The main idea is to choose a suitable formula of γ such that the minimum value of $\frac{(1+\tau)\mu(q)}{2q}$ is equal to that of $\mu(q)$ which occurs if $q = q^*$ (see the proof of Lemma 4.5.1). However, when $0.5 < \tau < \sqrt{2/3}$, we obtain improved results compared with that in [53], since we have shown in the proof of Lemma 4.5.1 that the minimum value of $\frac{(1+\tau)\mu(q)}{2q}$ is $\frac{(1+\tau)\mu(1)}{2}$. This value is greater than or equal to the value $\frac{(1+\tau)\mu(q^*)}{2q^*}$ used in [53] (due to $q^* > 1$ and the fact that $\frac{1+\tau}{2q}\mu(q)$ is decreasing in $(0, q^*)$).

At last, we explain the reason that our performance ratios are still better than those in [53] despite a new formula ($= 0.87856\frac{1+\tau}{2}$) of the performance ratios for $\tau \geq \sqrt{2/3}$ is used in [53]. In these cases, our formula for the performance ratio is $\overline{R}_2 = \frac{\alpha^{(p)}(1+\tau)}{2-4(1-\tau)\beta} = \frac{r_2(p)}{1-2(1-\tau)\beta} \cdot \frac{1+\tau}{2}$ where $\nu = 1$ and $p = 0$. Then by choosing the other parameter values shown as in Table 6 for $\tau \geq \sqrt{2/3}$ where $1 - 2(1 - \tau)\beta \leq 1$, it is easy to verify that the better numerical results ($\geq 0.87856\frac{1+\tau}{2}$) can be obtained (see Table 1.3).

Due to the use of Lemma 4.4.3, the values reported in the above tables are correct only for sufficiently large m .

Finally, we give the worst-case performance ratio of the SDP-algorithm for approximating MHC-LU regardless of the value of τ .

Theorem 4.5.3. The worst-case performance ratio of the SDP-algorithm for approximating MHC-LU is at least $\overline{R}_0 - o(1)$.

Proof. Since $\tau \geq 0$, we have, with the same arguments as these in the proof

of Lemma 4.5.1,

$$\begin{aligned} \omega(\overline{V}_1, V \setminus \overline{V}_1) / \omega^{opt} &\geq \min\{\mu(q) - o(1), \frac{(1 + \tau)\mu(q)}{2q} - o(1)\} \\ &\geq \min\{\mu(q) - o(1), \frac{\mu(q)}{2q} - o(1)\} \\ &\geq \overline{R}_0 - o(1). \end{aligned}$$

Thus, the worst-case performance ratio of our SDP-algorithm for approximating MHC-LU is at least $\overline{R}_0 - o(1)$, regardless of the value of τ . \square

It is easy to see that the numerical worst-case performance ratio of the SDP-algorithm for approximating MHC-LU, regardless of the value of τ , is 0.6271 which is the same as that for MHC-LU when $\tau = 0$.

The technique of the Lasserre Hierarchy is to use a family of stronger SDP relaxations derived by the so-called Lasserre lift-and-project system whose vector solutions enjoy nice structural properties. It is very interesting to consider whether an application of Lasserre Hierarchy can lead to a significantly improved approximation ratio for MHC-LU.

Chapter 5

Minimum size of n -factor-critical and k -extendable graphs

We determine the minimum size of n -factor-critical graphs and of k -extendable bipartite graphs, by considering Harary graphs and related graphs. Moreover, we determine the minimum size of k -extendable non-bipartite graphs for $k = 1, 2$, and pose a related conjecture for general k .

For obtaining the above results, we need some results as follows.

Lemma 5.0.4. (Plummer [123]) Let G be a connected bipartite graph with bipartition (U, W) , k a positive integer such that $k \leq (\nu(G) - 2)/2$. Then G is k -extendable if and only if $|U| = |W|$ and for all non-empty subset X of U with $|X| \leq |U| - k$, $|N(X)| \geq |X| + k$.

Lemma 5.0.5. (Yu [140], Favaron [42]). A graph G is n -factor-critical if and only if $\nu(G) \equiv n \pmod{2}$ and for any vertex set $S \subseteq V(G)$ with $|S| \geq n$, $o(G - S) \leq |S| - n$, where $o(G - S)$ denote the number of odd components of $G - S$.

Lemma 5.0.6. (Plummer [122]) A k -extendable graph G is $(k+1)$ -connected.

By Lemma 5.0.6, in a k -extendable graph G , $\delta(G) \geq \kappa(G) \geq k + 1$.

Lemma 5.0.7. (Plummer [122]) If G is a k -extendable graph and if u is a vertex of degree $k + 1$ in G , then $N(u)$ is independent.

Lemma 5.0.8. (Favaron [42], Liu and Yu [105]) An n -factor-critical graph G is n -connected, $(n + 1)$ -edge-connected, and $\delta(G) \geq n + 1$.

Lemma 5.0.9. (Lou and Yu [107]) If G is a k -extendable non-bipartite graph on ν vertices with $k \geq \nu/4$, then $\kappa(G) \geq 2k$.

Lemma 5.0.10. (Zhang et al. [144]) Let G be a non-bipartite graph on ν vertices, and k an integer such that $k \geq (\nu + 2)/4$. Then G is k -extendable if and only if it is $2k$ -factor-critical.

Lemma 5.0.11. (Maschlanka and Volkmann [112]) If G is a k -extendable non-bipartite graph, then $\alpha(G) \leq \nu(G)/2 - k$.

A fullerene graph is a planar cubic 3-connected graph with only pentagonal and hexagonal faces. A fullerene graph on n vertices exists for even $n \geq 20$, except $n = 22$.

Lemma 5.0.12. (Zhang and Zhang [142]) Every fullerene graph is 2-extendable.

5.1 Minimum size of n -factor-critical graphs and k -extendable bipartite graphs

In this section, we determine the minimum size of n -factor-critical graphs and k -extendable bipartite graphs.

Let G be a graph on ν vertices. We use numbers $0, 1, \dots, \nu - 1$ to label the ν vertices of a graph G . We call a vertex labelled by an even (odd) number an *even (odd) vertex*. We denote an edge of G joining i to j by (i, j) , and a path or cycle in G by listing all vertices of it in order. Throughout this section, the labels of the vertices are reduced modulo ν .

In [72], Harary defined a class of graphs that were later called Harary graphs. These graphs are m -connected graphs on ν vertices with $\lceil m\nu/2 \rceil$ edges, for $2 \leq m < \nu$. A *Harary graph* $H_{m,\nu}$ with vertex set $\{0, 1, \dots, \nu - 1\}$ is defined as follows.

- (1) $m = 2r$ is even. Two vertices i and j are joined if $i - r \leq j \leq i + r$.
- (2) $m = 2r + 1$ is odd and ν is even. $H_{2r+1,\nu}$ is constructed from $H_{2r,\nu}$ by adding the edges $(i, i + \nu/2)$, for all $0 \leq i \leq \nu/2 - 1$.

(3) Both $m = 2r + 1$ and ν are odd. $H_{2r+1,\nu}$ is constructed from $H_{2r,\nu}$ by adding the edges $(0, (\nu - 1)/2)$, $(0, (\nu + 1)/2)$ and $(i, i + (\nu + 1)/2)$ for all $1 \leq i < (\nu - 1)/2$.

Firstly, we will prove some results on the factor-criticality and extendibility of Harary graphs. We present some useful notations and definitions below. We denote by C the Hamilton cycle $(0, 1, \dots, \nu - 1, 0)$ in $H_{m,\nu}$. For a vertex set $S \subset V(G)$, we define an S -segment to be the maximal segment P of C such that all internal vertices of P belong to S , while the endvertices of P belong to $V(G) \setminus S$. We say that a component of $G - S$, containing an endvertex of P , be *associated* with P . An S -segment $P = (i, i + 1, \dots, j)$ is called an S -link if the vertices i and j belong to different components of $G - S$.

Theorem 5.1.1. Let $r \geq 2$ and $\nu > 2r$ be two integers. Then $H_{2r,\nu}$ is $(2r - 1)$ -factor-critical if ν is odd and $(2r - 2)$ -factor-critical if ν is even.

Proof. Suppose $\nu = 2s + 1$ is odd, and $G = H_{2r,2s+1}$ is not $(2r - 1)$ -factor-critical. By Lemma 5.0.5, there exists a vertex set $S \subset V(G)$ with $|S| \geq 2r - 1$, such that $o(G - S) > |S| - (2r - 1)$. By parity, $o(G - S) \geq |S| - (2r - 1) + 2 = |S| - 2r + 3$. Let c be the number of components of $G - S$. Then $c \geq o(G - S) \geq |S| - 2r + 3 \geq 2$. If $c = 2$ then all equalities must hold and $|S| = 2r - 1$. But this is impossible since G is $2r$ -connected. So, $c \geq 3$.

By the definition of $H_{2r,2s+1}$, every S -link P contains at least r internal vertices. Since $G - S$ has at least two components, a component of $G - S$ must be associated with at least two S -links. Hence there are at least $2c/2 = c$ S -links. Therefore $|S| \geq cr$. So we have

$$c \geq |S| - 2r + 3 \geq cr - 2r + 3.$$

That is, $(c - 2)(1 - r) \geq 1$. However, this is impossible since $c \geq 3$ and $r \geq 2$. Hence, G must be $(2r - 1)$ -factor-critical.

Suppose that $\nu = 2s$ is even, and $G = H_{2r,2s}$ is not $(2r - 2)$ -factor-critical. By Lemma 5.0.5, there exists a vertex set $S \subset V(G)$ with $|S| \geq 2r - 2$, such that $o(G - S) > |S| - (2r - 2)$. By parity, $o(G - S) \geq |S| - 2r + 4$. Using the same notations and analogous analysis in the case that ν is odd, we have $c \geq 3$ and $(c - 2)(1 - r) \geq 2$, which are impossible. \square

Theorem 5.1.2. Let $r \geq 2$ and $\nu > 2r + 1$ be two integers. Then $H_{2r+1,\nu}$ is $2r$ -factor-critical if ν is even and $(2r - 1)$ -factor-critical if ν is odd.

Proof. The proof is similar to that of Theorem 5.1.1. \square

Theorem 5.1.3. Let $s \geq 2$ be an integer. Then $H_{2,2s-1}$ and $H_{3,2s+1}$ are factor-critical. $H_{2,2s}$ is 1-extendable. $H_{3,2s}$ is bicritical if s is even, and 2-extendable if s is odd.

Proof. Since $H_{2,\nu}$ is a cycle of order ν , $H_{2,2s-1}$ is factor-critical and $H_{2,2s}$ is 1-extendable. Since $H_{2,2s+1}$ is a spanning subgraph of $H_{3,2s+1}$, $H_{3,2s+1}$ is factor-critical.

Consider $H_{3,2s}$. If s is even, we can verify by definition that $H_{3,2s}$ is bicritical. If s is odd, $H_{3,2s}$ is a bipartite graph with two parts consisting of the even vertices and odd vertices respectively. Denote the parts of even vertices by U_e , and that of odd vertices by U_o .

Suppose that $H_{3,2s}$ is not 2-extendable. Then by Lemma 5.0.4, there exist a vertex set $U \subset U_e$ with $|U| \leq s-2$, such that $|N(U)| < |U|+2$. By considering the neighborhood of U on C , we can see that this is impossible. \square

By Lemma 5.0.8, if G is n -factor-critical, then $\delta(G) \geq n+1$. So, an n -factor-critical graph G on ν vertices has at least $\nu(n+1)/2$ edges. Note that $\nu(n+1)/2$ is an integer, since n and ν have the same parity. For two odd integers $n = 2r-1 \geq 3$ and $\nu > n$, by Theorem 5.1.1, $H_{n+1,\nu}$ is an n -factor-critical graph on ν vertices with $\nu(n+1)/2$ edges. For two even integers $n = 2r \geq 4$ and $\nu > n$, by Theorem 5.1.2, $H_{n+1,\nu}$ is an n -factor-critical graph on ν vertices with $\nu(n+1)/2$ edges. For $n = 1$ and an odd integer $\nu > 1$, by Theorem 5.1.3, $H_{2,\nu}$ is a factor-critical graph with ν edges. For $n = 2$ and an even integer $\nu > 2$, it is not hard to check that a *wheel* W_ν , which is formed by connecting a single vertex to all vertices of a cycle of length $\nu-1$, is a bicritical graph on ν vertices, with $3\nu/2$ edges. Therefore, the minimum size of an n -factor-critical graph on ν vertices is exactly $\nu(n+1)/2$, for all integers $n \geq 1$ and $\nu > n$, where n and ν has the same parity.

Modifying the construction of Harary graphs slightly, we can get a class of k -extendable bipartite graphs with minimum size.

Let $2 \leq m \leq s$ be integers, $H_{m,2s}^B$ with vertex set $\{0, 1, \dots, 2s-1\}$, is defined as follows.

(1) $m = 2r$, where $r \leq s/2$. Then vertex i is adjacent to j , if i is even, j is odd and $i - 2r + 1 \leq j \leq i + 2r - 1$.

(2) $m = 2r + 1$, where $r \leq s/2$. Then vertex i is adjacent to j , if i is even, j is odd and $i - 2r + 1 \leq j \leq i + 2r + 1$.

It is clear that every $H_{m,2s}^B$ is a balanced bipartite graph whose two parts are consisting of the even vertices and the odd vertices, respectively. We denote the part consisting of the even (odd) vertices by V_e (V_o).

Theorem 5.1.4. For two integers $r \geq 1$ and $s \geq 2r$, $H_{2r,2s}^B$ is $(2r - 1)$ -extendable.

Proof. Let $G = H_{2r,2s}^B$ and assume that G is not $(2r - 1)$ -extendable. By Lemma 5.0.4, there exists $U \subset V_e$, such that $|U| \leq s - (2r - 1)$ and $|N(U)| \leq |U| + (2r - 2)$.

Define a U -consecutive set $U' = \{2i_0, 2i_1, \dots, 2i_{t-1}\}$ as a maximal subset of U so that $2i_l < 2i_{l+1} \leq 2i_l + 4r - 2$, for $0 \leq l \leq t - 2$. Then the neighborhoods of different U -consecutive sets do not intersect, and U can be uniquely divided into U -consecutive sets. For every U -consecutive set $U' = \{2i_0, 2i_1, \dots, 2i_{t-1}\}$, if $\{2i_0 - 2r + 1, 2i_0 - 2r + 3, \dots, 2i_0 - 1\} \cap \{2i_{t-1} + 1, 2i_{t-1} + 3, \dots, 2i_{t-1} + 2r - 1\} \neq \emptyset$, then $N(U') = U_o$, contradicting $|N(U)| \leq |U| + (2r - 2) \leq s - 1$. Hence $\{2i_0 - 2r + 1, 2i_0 - 2r + 3, \dots, 2i_0 - 1\} \cap \{2i_{t-1} + 1, 2i_{t-1} + 3, \dots, 2i_{t-1} + 2r - 1\} = \emptyset$. So $|N(U')| \geq r + r + t - 1 = |U'| + 2r - 1$, and hence $|N(U)| \geq |U| + 2r - 1$, contradicting $|N(U)| \leq |U| + (2r - 2)$. Therefore G is $(2r - 1)$ -extendable. \square

Theorem 5.1.5. For two integers $r \geq 1$ and $s \geq 2r + 1$, $H_{2r+1,2s}^B$ is $(2r)$ -extendable.

Proof. The proof is similar to that of Theorem 5.1.4. \square

For a k -extendable graph G , $\delta(G) \geq k + 1$. Hence, for an integer $k \geq 1$ and an even integer $\nu \geq 2k + 2$, a k -extendable graph G on ν vertices has at least $\nu(k + 1)/2$ edges. By Theorem 5.1.4 and 5.1.5, for all $2 \leq m \leq s$, $H_{m,2s}^B$ are $(m - 1)$ -extendable bipartite graphs having $\nu(k + 1)/2$ edges. Therefore, the minimum size of a k -extendable bipartite graph on ν vertices is exactly $\nu(k + 1)/2$.

5.2 Minimum size of 1-extendable non-bipartite graphs and 2-extendable non-bipartite graphs

In the previous section we constructed k -extendable bipartite graphs with minimum size. Now we consider k -extendable non-bipartite graphs with minimum size. Let G be a non-bipartite graph on ν vertices, where ν even, and k be a positive integer such that $k \geq (\nu + 2)/4$. By Lemma 5.0.10, if G is k -extendable, then it is $2k$ -factor-critical. Therefore, $\delta(G) \geq 2k + 1$, and G has at least $(2k + 1)\nu/2$ edges, which is greater than the lower bound for k -extendable bipartite graphs. Hence, we raise the following problem.

Problem 1. Let $k \geq 1$ be an integer, and G a k -extendable non-bipartite graph on $\nu \geq 2k + 2$ vertices. What is the minimum size of G ?

Denote such a minimum number by $\varepsilon(\nu, k)$. In this section, we solve the problem for $k = 1, 2$.

Theorem 5.2.1. For an even number $\nu \geq 4$, $\varepsilon(\nu, 1) = \nu + 2$.

Proof. We have $\delta(G) \geq 2$ in a 1-extendable graph G . Hence, a 1-extendable graph on ν vertices has at least ν edges. However, a connected graph with ν vertices and ν edges can only be the cycle C_ν , which is bipartite. Therefore $\varepsilon(\nu, 1) \geq \nu + 1$. Take a cycle $C = v_0v_1 \dots v_{\nu-1}v_0$, joint v_0 to v_2 and v_1 to v_3 we get a 1-extendable non-bipartite graph G . Therefore $\varepsilon(\nu, 1) \leq \nu + 2$.

Let G be a 1-extendable non-bipartite graph with ν vertices and $\nu + 1$ edges. By the Handshaking Lemma, G has precisely two vertices of degree 3, while the other vertices are of degree 2. Since G is non-bipartite, there is an odd cycle $Q = v_0v_1 \dots v_{2l}v_0$ in G , and $G - Q$ is not null. By Lemma 5.0.6, $\kappa(G) \geq 2$. So, there is at least two vertices on Q , say v_0 and v_i , where $1 \leq i \leq 2l$, who send edges to $G - Q$, and hence $d(v_0) = d(v_i) = 3$. If i is odd, then v_0v_1 is not contained in any perfect matching of G . If i is even, then $v_{2l}v_0$ is not contained in any perfect matching of G . These contradict that G is 1-extendable. Hence $\varepsilon(\nu, 1) \neq \nu + 1$. So $\varepsilon(\nu, 1) = \nu + 2$. \square

Now we consider 2-extendable non-bipartite graphs. For a 2-extendable graph G , $\delta(G) \geq 3$. Hence, $\varepsilon(\nu, 2) \geq 3\nu/2$. The next theorem shows that the bound can be achieved when ν is large. We will prove the theorem in the rest of this section.

Theorem 5.2.2. For an even integer $\nu \geq 6$,

$$\varepsilon(\nu, 2) = \begin{cases} 15, & \text{if } \nu = 6, \\ 16, & \text{if } \nu = 8, \\ 19, & \text{if } \nu = 10, \\ 20, & \text{if } \nu = 12, \\ 3\nu/2, & \text{if } \nu \geq 14. \end{cases}$$

We will prove several theorems, that will be combined to obtain Theorem 5.2.2.

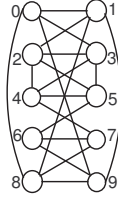


Figure 5.1: G_1 , a 2-extendable non-bipartite graph with 10 vertices and 19 edges

Theorem 5.2.3. $\varepsilon(10, 2) = 19$.

Proof. It can be verified that the graph we show in Figure 5.1 is a 2-extendable non-bipartite graph with 10 vertices and 19 edges. To prove that $\varepsilon(10, 2) = 19$, it suffices to show that there do not exist 2-extendable non-bipartite graphs with 10 vertices and no more than 18 edges. Assume that we can find such a graph G . By Lemma 5.0.11, $\alpha(G) \leq 3$. Firstly we prove the following claims.

Claim 1. There do not exist three independent vertices of degree 3 in G .

We prove Claim 1 by contradiction. Suppose there are three independent vertices u_0, u_1, u_2 of degree 3 in G . By Lemma 5.0.7, the neighborhood of $u_i, 0 \leq i \leq 2$, is independent. Since $\alpha(G) \leq 3$ and $\nu(G) = 10$, we have $|N(u_0) \cup N(u_1) \cup N(u_2)| = 7$.

Suppose that $|N(u_0) \cap (N(u_1) \cup N(u_2))| \leq 1$. Then $(N(u_0) \setminus (N(u_1) \cup N(u_2))) \cup \{u_1, u_2\}$ is an independent set of order at less 4, a contradiction. So $|N(u_0) \cap (N(u_1) \cup N(u_2))| \geq 2$. Similarly, we have $|N(u_1) \cap (N(u_0) \cup N(u_2))| \geq 2$ and $|N(u_2) \cap (N(u_0) \cup N(u_1))| \geq 2$. But then $|N(u_0) \cup N(u_1) \cup N(u_2)| \leq 1 + 1 + 1 + (2 + 2 + 2)/2 = 6$, contradicting $|N(u_0) \cup N(u_1) \cup N(u_2)| = 7$. Therefore, there do not exist three independent vertices of degree 3 in G .

Claim 2. Every vertex in G has at most two neighbors of degree 3.

Suppose there is a vertex u in G , who has at least three neighbors, say v_0 , v_1 and v_2 , of degree 3. By Claim 1, at least two of them, say v_0 and v_1 , are adjacent. Then $N(v_0)$ is not independent, contradicting Lemma 5.0.7. Hence Claim 2 holds.

Suppose there are at least six vertices of degree 3 in G , one of which being u . By Claim 2, u can be adjacent to at most two other vertices of degree 3. Then there are at least three vertices of degree 3 that are not adjacent to u , at least two of which are not adjacent. Such two vertices and u are three independent vertices of degree 3 in G , contradicting Claim 1. Therefore, there are no more than five vertices of degree 3 in G . Moreover, $|E(G)| \geq \lceil (3 \times 5 + 4 \times 5) / 2 \rceil = 18$, and equality must hold.

Suppose there is a vertex v such that $d(v) \geq 6$. By Claim 2, v has at least four neighbors of degree no less than 4. Then G has at least $\lceil (6 + 4 \times 4 + 5 \times 3) / 2 \rceil = 19$ edges, a contradiction. Hence $\Delta(G) \leq 5$.

Now we show that $\Delta(G) < 5$. Suppose that $\Delta(G) = 5$. Since there are no more than five vertices of degree 3, the non-increasing degree sequence of G must be $(5, 4, 4, 4, 4, 3, 3, 3, 3, 3)$. Assume that there exist a vertex u_0 of degree 3, which is adjacent to at most one vertex of degree 3 in G . There are at least other three vertices of degree 3 that are not adjacent to u_0 , at least 2 of which, denoted by u_1 and u_2 , are not adjacent. Then u_0 , u_1 and u_2 are three independent vertices of degree 3 in G , contradicting Claim 1. Hence, every vertex of degree 3 in G has exactly two neighbors of degree 3. Then, the five vertices of degree 3 in G constitute a cycle, denoted by $C_0 = v_0v_1v_2v_3v_4v_0$. Furthermore, for each $0 \leq i \leq 4$, v_i sends an edge to $G - C_0$.

Suppose the five vertices in $G - C_0$ are adjacent to v_0 , v_1 , v_2 , v_3 and v_4 , respectively. Denote the vertex adjacent to v_i by u_i , $0 \leq i \leq 4$. Noticing that v_0 and v_2 are not adjacent, and u_1 , u_3 and u_4 are not adjacent to v_0 or v_2 , by $\alpha(G) \leq 3$, u_1 , u_3 and u_4 must be adjacent to each other. Similar analysis shows that u_1 , u_2 , u_3 , u_4 and u_5 must be mutually adjacent, a contradiction.

Therefore, we can assume that there are two vertices on C_0 , say v_0 and v_2 , share a common neighbor $u_0 \in V(G - C_0)$. Denote the other four vertices in $G - C_0$ by u_1 , u_2 , u_3 and u_4 . Since v_0 and v_2 are not adjacent, and they are not adjacent to u_1 , u_2 , u_3 or u_4 , by $\alpha(G) \leq 3$, u_1 , u_2 , u_3 and u_4 must be mutually adjacent. Each of v_3 and v_4 must send an edge to $\{u_1, u_2, u_3, u_4\}$. Furthermore, v_3 and v_4 cannot have a common neighbor. Without loss of generality, suppose $v_3u_3, v_4u_4 \in E(G)$. Now consider the independent vertices

u_0, v_1 and v_3 . Any other vertex in G must be adjacent to one of them. Therefore, there must be an edge from $\{u_0, v_1\}$ to u_4 . Similarly, there is an edge from $\{u_0, v_1\}$ to u_3 . Then we have $d(u_3), d(u_4) \geq 5$, contradicting the degree sequence of G . Hence, $\Delta(G) < 5$.

So, all vertices in G are of degree 3 or 4. Since $|E(G)| = 18$, a simple calculation shows that there are six vertices of degree 4 and four vertices of degree 3.

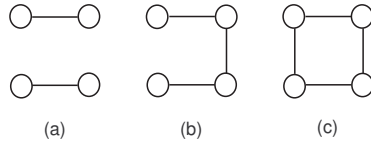


Figure 5.2: Three possible configurations of H

Consider the subgraph H of G induced by all vertices of degree 3. By Lemma 5.0.7, H cannot contain a triangle. By Claim 1, $\alpha(H) \leq 2$. Hence, H can only be (a), (b) or (c) in Figure 5.2. We will discuss them case by case.

Let H be (a) of Figure 5.2. Denote the two edges in H by u_0u_1 and u_2u_3 . Then u_0, u_1 are adjacent to two vertices of degree 4 respectively, and they do not have a common neighbor. Denote the other neighbors of u_0 by v_0, v_1 , those of u_1 by v_2, v_3 . By Lemma 5.0.7, $v_0v_1, v_2v_3 \notin E(G)$. Denote the other two vertices of degree 4 in G by v_4 and v_5 . Since $\{u_0, v_2, v_3\}$ is an independent set of order 3, u_2 and u_3 send edges to $\{v_2, v_3\}$. Moreover, u_2 and u_3 cannot have a common neighbor. Without loss of generality we may assume $u_2v_2, u_3v_3 \in E(G)$. Similarly, we can assume $u_2v_0, u_3v_1 \in E(G)$. Then $\{u_0, u_2, v_3\}, \{u_0, u_3, v_2\}, \{u_1, u_2, v_1\}$ and $\{u_1, u_3, v_0\}$ are independent sets of order 3. Hence v_0, v_1, v_2 and v_3 must be adjacent to v_4 and v_5 . But then $v_4v_5 \notin E(G)$, and $\{u_0, u_2, v_4, v_5\}$ is an independent set of order 4, contradicting $\alpha(G) \leq 3$.

Let H be (b) of Figure 5.2, a path $u_0u_1u_2u_3$. Denote the other neighbors of u_0 by v_0 and v_1 . Then $\{u_1, v_0, v_1\}$ is an independent set. By $\alpha(G) \leq 3$, u_3 must be adjacent to v_0 or v_1 . Without loss of generality, let $u_3v_0 \in E(G)$. Suppose that $u_3v_1 \in E(G)$. Let w be the neighbor of v_0 which is different from u_0 and u_3 , obviously $w \notin \{u_1, u_2\}$. Then $\{u_1u_2, v_0w\}$ is not contained in any perfect matching of G , contradicting 2-extendibility of G . Hence, $u_3v_1 \notin E(G)$. So u_3 has another neighbor v_2 . Denote the other three vertices of degree 4 in G by v_3, v_4 and v_5 . Since u_0 and u_3 are not adjacent and they send no edge

to $\{v_3, v_4, v_5\}$, v_3, v_4 and v_5 must form a triangle or we get an independent set of order 4 in G . Since $\{u_0, u_2, v_2\}$ is an independent set of order 3, each of v_3, v_4 and v_5 sends an edge to u_2 or v_2 . Similarly, each of v_3, v_4 and v_5 sends an edges to u_1 or v_1 . Then v_3, v_4 and v_5 cannot send any edge to v_0 . So, $N(v_0) = \{u_0, u_3\}$, a contradiction.

Let H be (c) of Figure 5.2, a cycle $u_0u_1u_2u_3u_0$. Suppose that u_0 and u_2 have a common neighbor other than u_1 and u_3 . Then the other five vertices of degree 4 are not adjacent to u_0 or u_2 , so they must be mutually adjacent or we get an independent set of order 4 in G . But then G is not connected, a contradiction. Therefore, $N(u_0) \cap N(u_2) = \{u_1, u_3\}$. Similarly, u_1 and u_3 do not have any common neighbor other than u_0 and u_2 . Hence, each u_i has a neighbor $v_i \notin \{u_0, u_1, u_2, u_3\}$, and $v_i \neq v_j$ for $i \neq j$, $0 \leq i, j \leq 3$. Denote by v_4 and v_5 the other two vertices left. If there are any two vertices in $\{v_0, v_2, v_4, v_5\}$ that are not adjacent, then they form an independent set of order 4 with u_1 and u_3 , a contradiction. Hence v_0, v_2, v_4 and v_5 are mutually adjacent. Similarly v_1, v_3, v_4 and v_5 are mutually adjacent. But then $d(v_4), d(v_5) \geq 5$, a contradiction.

Thus we have led to contradictions in all cases and deny the existence of a 2-extendable non-bipartite graph with 10 vertices and no more than 18 edges, and conclude that $\varepsilon(10, 2) = 19$. \square

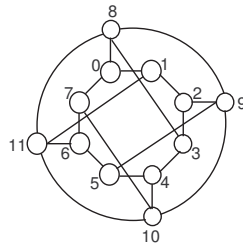


Figure 5.3: G_2 , the unique 2-extendable non-bipartite graphs with 12 vertices and 20 edges

Theorem 5.2.4. $\varepsilon(12, 2) = 20$. Furthermore, there is only one 2-extendable non-bipartite graphs with 12 vertices and 20 edges up to isomorphism, as showed in Figure 5.3.

Proof. It can be checked that the graph showed in Figure 5.3 is 2-extendable. Let G be a 2-extendable non-bipartite graph with 12 vertices and no more

than 20 edges, we prove that G must be isomorphic to the graph showed in Figure 5.3.

By Lemma 5.0.11, $\alpha(G) \leq 4$. Let the number of vertices of degree 3 in G be x , then $(3x + 4(12 - x))/2 \leq |E(G)| \leq 20$, hence $x \geq 8$. We discuss two cases.

Case 1. There are four independent vertices u_0, u_1, u_2 and u_3 , of degree 3 in G .

By $\alpha(G) \leq 4$ and $\nu(G) = 12$, we have $|N(u_0) \cup N(u_1) \cup N(u_2) \cup N(u_3)| = 8$.

By Lemma 5.0.7, the neighbors of u_0 are independent. If there are two neighbors v_0 and v_1 of u_0 that are not adjacent to u_1, u_2 or u_3 , $\{v_0, v_1, u_1, u_2, u_3\}$ is an independent set, contradicting $\alpha(G) \leq 4$. So $|N(u_0) \cap (\cup_{1 \leq j \leq 3} N(u_j))| \geq 2$.

Similarly, for any $0 \leq i \leq 3$, $|N(u_i) \cap (\cup_{0 \leq j \leq 3, j \neq i} N(u_j))| \geq 2$. Hence, $|N(u_0) \cup N(u_1) \cup N(u_2) \cup N(u_3)| \leq 1 \times 4 + 2 \times 4/2 = 8$. Since equality holds, for every u_i , $|N(u_i) \cap (\cup_{0 \leq j \leq 3, j \neq i} N(u_j))| = 2$. Furthermore, any vertices in $V(G) \setminus \{u_0, u_1, u_2, u_3\}$ can be adjacent to at most two vertices in $\{u_0, u_1, u_2, u_3\}$.

Suppose $|N(u_0) \cap N(u_1)| = 2$. Then $N(u_0) \cup \{u_2, u_3\}$ is an independent set of order 5, a contradiction. Hence $|N(u_0) \cap N(u_1)| \leq 1$, and similarly $|N(u_i) \cap N(u_j)| \leq 1$ for any $0 \leq i \neq j \leq 3$. So, for every u_i , $0 \leq i \leq 3$, u_i has common neighbors with u_j and u_k , where $0 \leq j \neq k \leq 3$, and $j, k \neq i$. Without loss of generality, suppose that u_0 has common neighbors with u_1 and u_3 . Then u_2 also has common neighbors with u_1 and u_3 . Hence, u_0, u_1, u_2, u_3 and the common neighbors form a cycle on 8 vertices. Denote the cycle by $C_1 = u_0v_0u_1v_1u_2v_2u_3v_3u_0$ and the other vertices adjacent to u_i by w_i , $0 \leq i \leq 3$.

If $w_0w_1 \notin E(G)$, $\{w_0, w_1, v_0, u_2, u_3\}$ is an independent set of order 5, a contradiction. So $w_0w_1 \in E(G)$. Similarly, $w_1w_2, w_2w_3, w_3w_0 \in E(G)$.

If $v_0v_2 \in E(G)$, then $\{v_0v_2, w_1w_2\}$ is not contained in any perfect matching of G , a contradiction. Therefore $v_0v_2 \notin E(G)$. Furthermore, v_0 can not be adjacent to v_1 or v_3 . Similarly, all v_i , $0 \leq i \leq 3$, cannot be adjacent to each other. Hence every v_i , $0 \leq i \leq 3$, sends edges to some w_j , $0 \leq j \leq 3$, and the number of such edges is at least 4. Then, $E(G) \geq 8 + 4 + 4 + 4 = 20$. By our assumption, equality holds, and each v_i , $0 \leq i \leq 3$, sends exactly one edge to w_j , for one $0 \leq j \leq 3$.

The vertex v_0 can only be adjacent to w_2 or w_3 . Without lose of generality,

suppose $v_0w_3 \in E(G)$. The vertex v_1 can only be adjacent to w_0 or w_3 . If $v_1w_3 \in E(G)$, then $\{u_2v_2, u_0v_3\}$ is not contained in any perfect matching of G , a contradiction. Therefore, we must have $v_1w_0 \in E(G)$. Similarly $v_2w_1, v_3w_2 \in E(G)$. So, G is isomorphic to the graph showed in Figure 5.3.

Case 2. There do not exist four independent vertices of degree 3 in G .

We claim that there exists a vertex of degree 3, whose neighbors are all of degree 3. Suppose to the contrary that there is no such a vertex. Obviously, there exists a vertex u of degree 3, who has two neighbors, denoted by v and w , of degree 3 in G . Then, each of v and w has at most one more neighbor of degree 3. So, there are at least three vertices of degree 3 in G who are not adjacent to v or w . By the condition of Case 2, these three vertices must form a triangle, a contradiction to Lemma 5.0.7. So there is a vertex, say u_0 , of degree 3 in G , whose neighbors are all of degree 3.

By Lemma 5.0.7, $N(u_0)$ is independent. By the condition of Case 2, any other vertex of degree 3 must be adjacent to some vertices in $N(u_0)$. Hence, there is a neighbor u_1 of u_0 , who is adjacent to other two vertices of degree 3. Denote the other neighbors of u_0 by v_0 and v_1 , and the other neighbors of u_1 by v_2 and v_3 . There are at least two more vertices, say w_0 and w_1 , of degree 3 in G .

Since there are no four independent vertices of degree 3, there is at least one edge among v_0, v_1, v_2 and v_3 . But $v_0v_1, v_2v_3 \notin E(G)$. Without loss of generality we assume that $v_1v_2 \in E(G)$. Since $\{u_0, v_2, v_3\}$ is an independent set, by the condition of Case 2, both w_0 and w_1 send some edges to $\{v_2, v_3\}$. Similarly, both w_0 and w_1 send some edges to $\{v_0, v_1\}$.

if $w_0w_1 \in E(G)$, the subgraph of G induced by $\{u_0, u_1, v_0, v_1, v_2, v_3, w_0, w_1\}$ sends at most two edges to the other part of G , so $\kappa(G) \leq \kappa'(G) \leq 2$, contradicting Lemma 5.0.6. Hence $w_0w_1 \notin E(G)$.

If v_0 is not adjacent to w_0 or w_1 , then u_1, v_0, w_0, w_1 are four independent vertices of degree 3, contradicting the condition of Case 2. So v_0 , and similarly v_1, v_2 and v_3 , must be adjacent to w_0 or w_1 . Since both w_0 and w_1 send some edges to $\{v_0, v_1\}$ and $\{v_2, v_3\}$, without loss of generality, we can assume that $w_0v_0, w_1v_1 \in E(G)$. Then w_1 must be adjacent to v_3 , and v_2 must be adjacent to w_0 . But then $\{u_0v_0, w_1v_3\}$ is not contained in any perfect matching of G , contradicting 2-extendibility of G .

Therefore, there can be only one 2-extendable non-bipartite graph with 12 vertices and no more than 20 edges, upto isomorphism, as shown in Figure 5.3. \square

Theorem 5.2.5. When $\nu = 14, 16$ or 18 , $\varepsilon(\nu, 2) = 3\nu/2$.

Proof. We already have $\varepsilon(\nu, 2) \geq 3\nu/2$. To prove the equality, we show 2-extendable graphs on $\nu = 14, 16$ and 18 vertices and $3\nu/2$ edges in Figure 5.4, 5.5 and 5.6, respectively. \square

Theorem 5.2.6. When $\nu \geq 20$, $\varepsilon(\nu, 2) = 3\nu/2$.

Proof. We have $\varepsilon(\nu, 2) \geq 3\nu/2$. To prove the equality we must find 3-regular 2-extendable non-bipartite graphs on ν vertices for all even $\nu \geq 20$. By definition and Lemma 5.0.12, Fullerene graphs are 3-regular 2-extendable non-bipartite graphs, and Fullerene graphs with ν vertices exists for all even $\nu \geq 20$, except $\nu = 22$. So, we only need to construct a 3-regular 2-extendable non-bipartite graph on 22 vertices. One such graph is shown in Figure 5.7. \square

Now we can prove Theorem 5.2.2.

Proof. Let G be a 2-extendable graph on ν vertices with minimum size. By Theorem 5.0.9, when $\nu \leq 8$, $\delta(G) \geq \kappa(G) \geq 4$. For $\nu = 6$, it is not hard to check that G must be K_6 , thus $\varepsilon(6, 2) = 15$. For $\nu = 8$, we have $\varepsilon(8, 2) \geq 16$, and it is obvious that the graph shown in Figure 5.8 is a 2-extendable non-bipartite graph with 8 vertices and 16 edges, hence $\varepsilon(8, 2) = 16$. For other values of ν the results follow from Theorem 5.2.3, 5.2.4, 5.2.5 and 5.2.6. \square

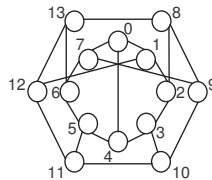


Figure 5.4: G_3 , a 2-extendable non-bipartite graph with 14 vertices and 21 edges

5.3 Concluding remarks

We finish the chapter with some ideas on Problem 1 for general k .

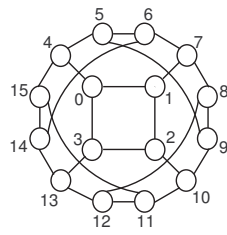


Figure 5.5: G_4 , a 2-extendable non-bipartite graph with 16 vertices and 24 edges

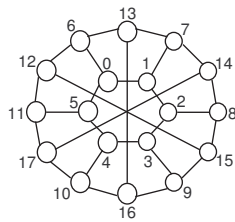


Figure 5.6: G_5 , a 2-extendable non-bipartite graph with 18 vertices and 27 edges

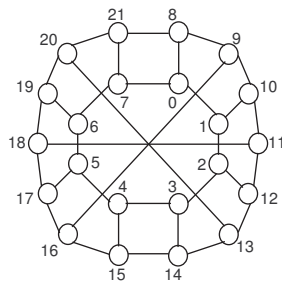


Figure 5.7: G_6 , a 2-extendable non-bipartite graph with 22 vertices and 33 edges

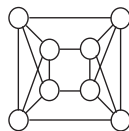


Figure 5.8: a 2-extendable non-bipartite graph with 8 vertices and 16 edges

By Lemma 5.0.10, the set of k -extendable non-bipartite graphs and the set of $2k$ -factor-critical graphs coincide when $\nu(G) \leq 4k - 2$. We have found $2k$ -factor-critical graphs with minimum size among Harary graphs in Section 5.1. So we have

$$\varepsilon(\nu, k) = (2k + 1)\nu/2 \quad \text{if } \nu \leq 4k - 2.$$

By Lemma 5.0.9, when $\nu(G) = 4k$, the connectivity of a k -extendable non-bipartite graph G is $2k$. Therefore $\delta(G) \geq \kappa(G) \geq 2k$. This bound is obtained by the following graph. Let H_1 and H_2 be two copies of K_{2k} where $V(H_1) = \{u_0, u_1, \dots, u_{2k-1}\}$ and $V(H_2) = \{v_0, v_1, \dots, v_{2k-1}\}$. And construct G by joining every u_i to v_i , $0 \leq i \leq 2k - 1$. It is not hard to check that G is a k -extendable graph with $4k$ vertices and regular degree $2k$. Therefore

$$\varepsilon(4k, k) = 4k^2.$$

When $\nu \geq 4k + 2$, an example in [107] shows that the connectivity of a k -extendable graph G with ν vertices can be $k + 1$, that is, the bound given by Lemma 5.0.6. The case that $k = 2$ gives us some hints that there may exist k -extendable regular graphs with degree $k + 1$, and hence $\varepsilon(\nu, k) = (k + 1)\nu/2$, when ν is large. Let ν_0 be the minimum even integer such that $\varepsilon(\nu_0, k) = (k + 1)\nu_0/2$. Then $\nu_0 \geq 4k + 2$. Assume that for a given k , the function $\varepsilon(\nu, k)$ is increasing for even integer ν . We have $(\nu_0 - 4k)/2 \leq \varepsilon(\nu_0, k) - \varepsilon(4k, k) = (k + 1)\nu_0/2 - 4k^2$, that is, $\nu_0 \geq 8k - 4$. Therefore, we pose the following conjecture.

Conjecture 1. For a given integer $k > 0$, if G is a k -extendable non-bipartite graph with edge number $\nu(G)(k + 1)/2$, then $\nu(G) \geq 8k - 4$.

Chapter 6

Directed Hamilton cycles and matching alternating Hamilton cycles

In 1972, Woodall established the following Ore type condition for directed Hamilton cycles in digraphs: Let D be a digraph. If for every vertex pair u and v , where there is no arc from u to v , we have $d^+(u) + d^-(v) \geq |D|$, then D has a directed Hamilton cycle. By a correspondence between bipartite graphs and digraphs, the above result is equivalent to the following result of Las Vergnas: Let $G = (B, W)$ be a balanced bipartite graph. If for any $b \in B$ and $w \in W$, where b and w are nonadjacent, we have $d(w) + d(b) \geq |G|/2 + 1$, then every perfect matching of G is contained in a Hamilton cycle.

The lower bounds in both results are tight. In this chapter, we reduce both bounds by 1, and prove that the conclusions still hold, with only a few exceptional cases that can be clearly characterized.

We will state our main results and their proofs in the following sections.

6.1 Main results

Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. We denote by ν or $|G|$ the order of $V(G)$. For $u \in V(G)$, we denote by $d(u)$ the degree of u ,

and $N(u)$ or $N_G(u)$ the set of neighbors of u in G . For a subgraph H of G and a vertex $u \in V(G - H)$, we also denote by $N_H(u)$ the set of neighbors of u in H . For any two disjoint vertex sets X, Y of G we denote by $e(X, Y)$ the number of edges of G from X to Y . For $u, v \in V(G)$, we denote by $d(u, v)$ the distance between u and v , that is, the length of the shortest path connecting u and v . By $uv+$ ($uv-$) we denote that the vertices u and v are adjacent (nonadjacent). If a vertex u sends (no) edges to X , where X is a subgraph or a vertex subset of G , we write $u \rightarrow X$ ($u \nrightarrow X$). By nK_2 , we denote a graph consisting of n independent edges. Let (u, v) denote an arc from u to v . If $(w, u) \in A(D)$ and $(w, v) \in A(D)$, then we say that the pair $\{u, v\}$ is dominated, if $(u, w) \in A(D)$ and $(v, w) \in A(D)$, then we say that the pair $\{u, v\}$ is dominating. The complete digraph on $n \geq 1$ vertices, denoted by \overleftrightarrow{K}_n , is obtained from the complete graph K_n by replacing every edge xy with two arcs (x, y) and (y, x) . Without causing ambiguity, we use I_n to denote a graph or a digraph consisting of n independent vertices. Let $C = u_0u_1 \dots u_{m-1}u_0$ be a cycle in a graph G . Throughout Chapter 6, the subscript of u_i is reduced modulo m . We always orient C such that u_{i+1} is the successor of u_i .

If $U \subseteq V(C)$, the set of predecessors and successors of U on C is denoted by U^- and U^+ respectively, or u^- and u^+ when U contains only one vertex u . For $0 \leq i, j \leq m - 1$, the path $u_iu_{i+1} \dots u_j$ is denoted by $u_iC^+u_j$, while the path $u_iu_{i-1} \dots u_j$ is denoted by $u_iC^-u_j$. For a path $P = v_0v_1 \dots v_{p-1}$ and $0 \leq i, j \leq p - 1$, the segment of P from v_i to v_j is denoted by v_iPv_j .

A *matching* M of G is a subset of $E(G)$ in which no two elements share a common end vertex. If every $v \in V(G)$ is covered by an edge in M then M is said to be a *perfect matching* of G . For a matching M , an *M -alternating path* (*M -alternating cycle*) is a path (cycle) of which the edges appear alternately in M and $E(G) \setminus M$. We call an edge in M or an M -alternating path starting and ending with edges in M a *closed M -alternating path*, while an edge in $E(G) \setminus M$ or an M -alternating path starting and ending with edges in $E(G) \setminus M$ an *open M -alternating path*.

Let $m, n \geq 1$ be integers. Let \mathcal{D}_1 be the set of all digraphs obtained by identifying one vertex of $\overleftrightarrow{K}_{n+1}$ with one vertex of $\overleftrightarrow{K}_{m+1}$. Let \mathcal{D}_2 be an arbitrary digraph on n vertices, and take a copy of I_{n+1} . Let \mathcal{D}_2 be the set of all digraphs obtained by adding arcs of two directions between every vertex

of I_{n+1} and every vertex of D_2 . Let D_3 be as shown in Figure 6.1, and take a copy of \overleftrightarrow{K}_n . Let \mathcal{D}_3 be the set of all graphs constructed by adding arcs of two directions between v_i , $i = 0, 1$, and every vertex of \overleftrightarrow{K}_n , and possibly, adding any of the arcs (v_0, v_1) and (v_1, v_0) , or both. Finally, let D_4 be the digraph showed in Figure 6.2. Our main result is as below.

Theorem 6.1.1. Let D be a digraph. For every vertex pair u and v , where there is no arc from u to v , we have $d^+(u) + d^-(v) \geq |D| - 1$, then D has a directed Hamilton cycle, unless $D \in \mathcal{D}_1, \mathcal{D}_2$ or \mathcal{D}_3 , or $D = D_4$.

Let \mathcal{G}_1 be the class of graphs G constructed by identifying an edge of one $K_{m+1, m+1}$ and one $K_{n+1, n+1}$, and \mathcal{M}_1 be the set of all perfect matchings of G containing the identified edge. Let \mathcal{G}_2 be the class of graphs G , constructed by taking a copy of $(n+1)K_2$ with bipartition (B, W) , and an arbitrary bipartite graph G_2 with bipartition (B_1, W_1) , where $|B_1| = |W_1| = n$, which has at least one perfect matching, then connecting every vertex in B to every vertex in W_1 , and every vertex in W to every vertex in B_1 . Furthermore, let \mathcal{M}_2 be the set of all perfect matchings of G , containing all the edges in $(n+1)K_2$ (shown thick in Figure 6.3). Let G_3 be as shown in Figure 6.4, and \mathcal{G}_3 the set of the graphs G constructed by taking one copy of $K_{n, n}$ with bipartition (B, W) , and connecting every vertex in B to w_0 and w_1 , every vertex in W to b_0 and b_1 , and possibly, adding any of the edges w_0b_1 , w_1b_0 , or both. Let \mathcal{M}_3 be the set of perfect matchings of G , containing the thick edges in G_3 . Finally, we let graph G_4 be the graph in Figure 6.5, and M_4 the perfect matching of it, consisting of the thick edges. We have the following version of our main theorem.

Theorem 6.1.2. Let $G = (W, B)$ be a bipartite graph with a perfect matching M , such that for every vertex pair $w \in W$ and $b \in B$ with $wb-$, we have $d(w) + d(b) \geq \nu/2 + 1$. Then G has an M -alternating Hamilton cycle, unless one of the following holds.

- (1) $G \in \mathcal{G}_1$, and $M \in \mathcal{M}_1$.
- (2) $G \in \mathcal{G}_2$, and $M \in \mathcal{M}_2$.
- (3) $G \in \mathcal{G}_3$, and $M \in \mathcal{M}_3$.
- (4) $G = G_4$ and $M = M_4$.

Since the two results are equivalent, we only prove Theorem 6.1.2 in the next section. Before that, let us say a few words on the non-existence of

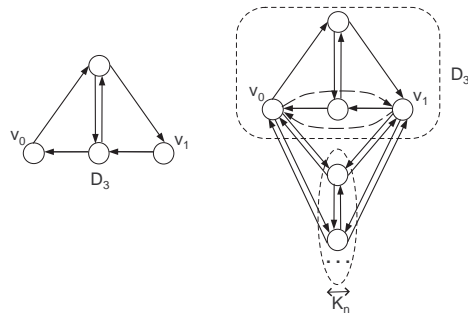


Figure 6.1: Exceptional graph family: \mathcal{D}_3

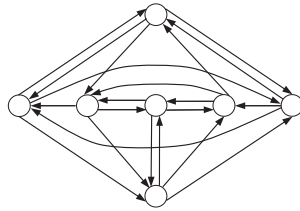


Figure 6.2: Exceptional graph D_4

M -alternating Hamilton cycles in the four exceptional cases. In Case (1), an M -alternating cycle of G must contain the identified edge, whose endvertices form a vertex cut of G , so G does not have an M -alternating Hamilton cycle. In Case (2), if there is an M -alternating Hamilton cycle C of G , then the edges on C that belong to M must be in $(n + 1)K_2$ and G_2 alternately, but there is one more such edge in $(n + 1)K_2$, a contradiction. In Case (3), we can not have an M -alternating Hamilton cycle containing both e_0 and e_1 . Finally in Case (4), the non-existence of any M -alternating Hamilton cycle can be verified directly.

6.2 Proof of Theorem 6.1.2

Let $G = (W, B)$ be a bipartite graph satisfying the condition of the theorem, and let M be a perfect matching of G . Suppose that G does not have an M -alternating Hamilton cycle. We prove the theorem by characterizing G .

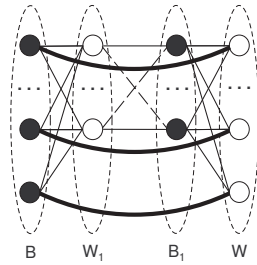


Figure 6.3: Exceptional graph family: \mathcal{G}_2

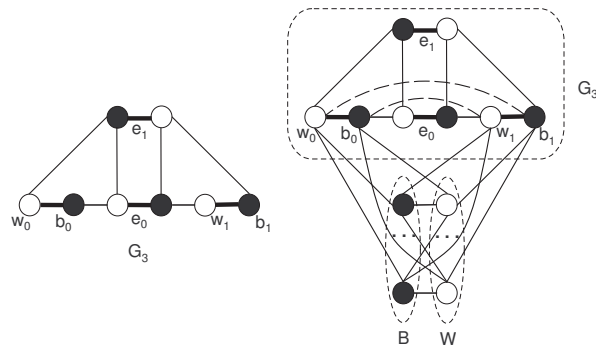


Figure 6.4: Exceptional graph family: \mathcal{G}_3

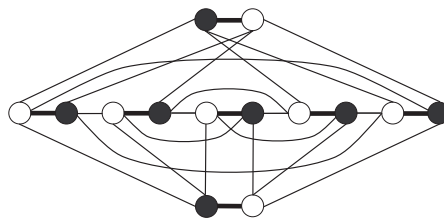


Figure 6.5: Exceptional graph G_4

The following two lemmas will be used in our proof.

Lemma 6.2.1. Let $G = (W, B)$ be a bipartite graph with a perfect matching M . Let $C = u_0u_1 \dots u_{2m-1}u_0$ be a longest M -alternating cycle in G , where $u_{2i} \in W$, $u_{2i+1} \in B$, and $u_{2i}u_{2i+1} \in M$, $0 \leq i \leq m-1$. Let $b \in B$, $w \in W$ be the ending vertices of a closed M -alternating path P in $G - C$. Then, for every $0 \leq i \leq m-1$, either $u_{2i}b-$ or $u_{2i-1}w-$. Furthermore, if $b \rightarrow C$ and $w \rightarrow C$, then $|N_C(b)| + |N_C(w)| \leq m - |P|/2 + 1$.

Proof. If there exists $0 \leq k \leq m-1$, such that $u_{2k}b+$ and $u_{2k-1}w+$, then $u_{2k}C^+u_{2k-1}wPbu_{2k}$ is an M -alternating cycle longer than C , a contradiction. Thus, for $0 \leq i \leq m-1$, either $u_{2i}b-$ or $u_{2i-1}w-$.

If $b \rightarrow C$ and $w \rightarrow C$, let $u_{2r} \in N_C(b)$ and $u_{2s-1} \in N_C(w)$ be such that $P' = u_{2s}C^+u_{2r-1}$ is the shortest. Then, there is no neighbor of w and b on P' . Since C is the longest, we have $|P'| \geq |P|$. So $|N_C(w)| + |N_C(b)| \leq 2 + (|C| - |P'| - 2)/2 = m - |P'|/2 + 1 \leq m - |P|/2 + 1$. \square

Lemma 6.2.2. Let G be a bipartite graph with a perfect matching M . Let $C = u_0u_1 \dots u_{2m-1}u_0$ be a longest M -alternating cycle in G , where $u_{2i}u_{2i+1} \in M$, $0 \leq i \leq m-1$. Let C_1 be an M -alternating cycle in $G - C$. For any vertex set $\{u_{2i-1}, u_{2i}\}$, $0 \leq i \leq m-1$, either $u_{2i-1} \rightarrow C_1$ or $u_{2i} \rightarrow C_1$.

Proof. Suppose there exists $0 \leq k \leq m-1$ such that $u_{2k-1} \rightarrow C_1$ and $u_{2k} \rightarrow C_1$. Let $b \in N_{C_1}(u_{2k})$ and $w \in N_{C_1}(u_{2k-1})$. We can always find a closed M -alternating path, P , as a segment of C_1 , connecting b and w . Then $u_{2k}C^+u_{2k-1}wPbu_{2k}$ is an M -alternating cycle longer than C , contradicting our condition. \square

In our proof, some important intermediate results are shown as claims.

Claim 1. There is an M -alternating cycle in G whose length is at least $\nu/2+1$.

Proof. Let $P = u_0u_1 \dots u_{2p-1}$ be a longest closed M -alternating path in G . Then all neighbors of u_0 and u_{2p-1} in G should be on P .

If $u_0u_{2p-1}+$, then we obtain a cycle $C = u_0u_1 \dots u_{2p-1}u_0$. Since P is the longest, $e(V(C), V(G - C)) = 0$. However, G is connected, so C must be an M -alternating Hamilton cycle and the claim holds.

If $u_0u_{2p-1}-$, by our condition, $d(u_0) + d(u_{2p-1}) \geq \nu/2 + 1$. Without loss of generality, assume that $d(u_0) \geq d(u_{2p-1})$ and let u_{2i-1} be the neighbor of u_0 with the maximum i , $1 \leq i \leq p$. Then, $i \geq (\nu/2 + 1)/2$ and $u_0Pu_{2i-1}u_0$ is an M -alternating cycle with length at least $2i \geq \nu/2 + 1$. This proves our claim. \square

Now let $C = u_0u_1 \dots u_{2m-1}u_0$ be a longest M -alternating cycle in G , where $u_{2i} \in W$, $u_{2i-1} \in B$ and $u_{2i}u_{2i+1} \in M$. Let $G_1 = G - C$. Denote the neighborhood and degree of $v \in V(G_1)$ in G_1 by $N_1(v)$ and $d_1(v)$. By Claim 1, $|G_1| \leq \nu/2 - 1$.

Let $P_1 = v_0v_1 \dots v_{2p_1-1}$ be a longest closed M -alternating path in G_1 , where $v_{2i} \in W$ and $v_{2i+1} \in B$, $0 \leq i \leq p_1 - 1$. Then $N_1(v_0), N_1(v_{2p_1-1}) \subseteq V(P_1)$, and $d_1(v_0), d_1(v_{2p_1-1}) \leq p_1$. Firstly, we prove that $v_0 \rightarrow C$ and $v_{2p_1-1} \rightarrow C$.

If $v_0 \nrightarrow C$ and $v_{2p_1-1} \nrightarrow C$, then $d(v_0) + d(v_{2p_1-1}) \leq 2p_1 \leq |G_1| \leq \nu/2 - 1$. By the condition of our theorem, $v_0v_{2p_1-1}+$, and we get a cycle $C_1 = v_0v_1 \dots v_{2p_1-1}v_0$ in G_1 . By Lemma 6.2.2, for any two vertices u_{2i-1} and u_{2i} on C , at least one of them, say $u_{2i} \nrightarrow C_1$. Then $d(u_{2i}) \leq \nu/2 - p_1$. But then $d(u_{2i}) + d(v_{2p_1-1}) \leq \nu/2$, contradicting the condition of the theorem.

If only one of v_0 and v_{2p_1-1} , say $v_0 \rightarrow C$. Let a neighbor of v_0 on C be u_{2j-1} , by Lemma 6.2.1, u_{2j} sends no edge to P_1 , so $d(u_{2j}) \leq \nu/2 - p_1$, and $d(u_{2j}) + d(v_{2p_1-1}) \leq \nu/2$, again contradicting the condition of the theorem.

Therefore $v_0 \rightarrow C$ and $v_{2p_1-1} \rightarrow C$.

By Lemma 6.2.1, $|N_C(v_0)| + |N_C(v_{2p_1-1})| \leq m - p_1 + 1$. Therefore,

$$\begin{aligned}
 d(v_0) + d(v_{2p_1-1}) &\leq 2p_1 + (m - p_1 + 1) \\
 &= m + p_1 + 1 \\
 &\leq m + |G_1|/2 + 1 \\
 &= \nu/2 + 1.
 \end{aligned} \tag{6.1}$$

If $v_0v_{2p_1-1}-$, then by our condition, $d(v_0) + d(v_{2p_1-1}) \geq \nu/2 + 1$ and hence equalities in (6.1) hold. But then we must have $v_0v_{2p_1-1}+$, a contradiction. So $v_0v_{2p_1-1}+$, and we get a cycle $C_1 = v_0v_1 \dots v_{2p_1-1}v_0$.

If $G_1 - C_1$ is nonempty, then there exists an edge $wb \in M \cap E(G_1 - C_1)$, where $w \in W$ and $b \in B$. By the choice of P_1 , $e(V(C_1), V(G_1 - C_1)) = 0$. By our condition, $d(w) + d(b) + d(v_0) + d(v_{2p_1-1}) \geq 2(\nu/2 + 1) = \nu + 2$. However, by Lemma 6.2.1, $|N_C(w)| + |N_C(b)| \leq m$, and hence $d(w) + d(b) \leq |G_1| - 2p_1 + m$, while $d(v_0) + d(v_{2p_1-1}) \leq m + p_1 + 1$ by (6.1), therefore $d(w) + d(b) + d(v_0) + d(v_{2p_1-1}) \leq |G_1| + 2m - p_1 + 1 = \nu - p_1 + 1 < \nu + 1$, a contradiction. Hence, $G_1 - C_1$ must be empty. Then $|G_1| = 2p_1$ and C_1 is an M -alternating Hamilton cycle of G_1 .

We claim that every vertex of G_1 sends some edges to C . Let v be any vertex in G_1 . Since G_1 has an M -alternating Hamilton cycle C_1 , we can choose a closed M -alternating Hamilton path P_1 of G_1 starting from v . By the above discussion, v sends some edges to C .

For a longest M -alternating cycle C in G , we call the graph $G_1 = G - C$ a critical graph (with respect to C) and a closed M -alternating Hamilton path of G_1 , $P_1 = v_0v_1 \dots v_{2p_1-1}$, where $v_{2i} \in W$ and $v_{2i+1} \in B$, a critical path, or a critical edge if $|P_1| = 2$. For a critical path P_1 , we can always find $u_{2s-1} \in N_C(v_0)$ and $u_{2r} \in N_C(v_{2p_1-1})$, such that $P_2 = u_{2s}C^+u_{2r-1}$ is the shortest. We let $R = u_{2r}C^+u_{2s-1}$.

By Lemma 6.2.2, $u_{2s} \rightarrow G_1$ and $u_{2r-1} \rightarrow G_1$. Further, for any edge $u_{2i-1}u_{2i}$ on R , we must have $e(\{u_{2i-1}, u_{2i}\}, \{u_{2s}, u_{2r-1}\}) \leq 1$, or we get an M -alternating Hamilton cycle

$$u_{2r}C^+u_{2i-1}u_{2s}C^+u_{2r-1}u_{2i}C^+u_{2s-1}v_0P_1v_{2p_1-1}u_{2r}.$$

Hence,

$$d(u_{2s}) + d(u_{2r-1}) \leq |P_2| + 2 + (|R| - 2)/2 = |P_2| + |R|/2 + 1. \quad (6.2)$$

Moreover,

$$d(v_0) + d(v_{2p_1-1}) \leq 2p_1 + 2 + (|R| - 2)/2 = 2p_1 + |R|/2 + 1. \quad (6.3)$$

So,

$$d(u_{2s}) + d(u_{2r-1}) + d(v_0) + d(v_{2p_1-1}) \leq 2p_1 + |P_2| + |R| + 2 = \nu + 2. \quad (6.4)$$

However $v_0u_{2r-1}-$ and $v_{2p_1-1}u_{2s}-$, by our condition,

$$d(u_{2s}) + d(u_{2r-1}) + d(v_0) + d(v_{2p_1-1}) \geq 2(\nu/2 + 1) = \nu + 2. \quad (6.5)$$

So all equalities in (6.2), (6.3), (6.4) and (6.5) must hold. To get equality in (6.3), each of v_0 and v_{2p_1-1} must be adjacent to all vertices in the other class of the bipartition of G_1 , and for any edge $u_{2i-1}u_{2i}$ on R , $e(\{u_{2i-1}, u_{2i}\}, \{v_0, v_{2p_1-1}\}) = 1$. Therefore, for a critical path $P_1 = v_0v_1 \dots v_{2p_1-1}$, we find two closed M -alternating paths R and P_2 as segments of C , such that $V(C) = V(R) \cup V(P_2)$, where the ending vertices of R are adjacent to v_0 and v_{2p_1-1} , respectively, and for any edge $u_{2i-1}u_{2i} \notin M$ on R , $e(\{u_{2i-1}, u_{2i}\}, \{v_0, v_{2p_1-1}\}) = 1$, while $e(V(P_2), \{v_0, v_{2p_1-1}\}) = 0$. We call P_2 the opposite path, and R the central path for P_1 .

Furthermore, to get equality in (6.2), each of u_{2s} and u_{2r-1} must be adjacent to all other vertices on P_2 in the other class of the bipartition.

Claim 2. A critical graph G_1 is complete bipartite.

Proof. Since C_1 is an M -alternating Hamilton cycle of G_1 , for any vertex $v \in V(G_1)$, P_1 can be chosen so that it is starting from v . By the equality of (6.3), v sends edges to every vertex in the other class of the bipartition of G_1 . \square

Let $G_2 = G[V(P_2)]$. We call G_2 the opposite graph. We choose C , G_1 and P_1 so that the opposite path P_2 is the shortest.

Claim 3. $e(V(G_1), V(G_2)) = 0$, and each of u_{2s-1} and u_{2r} is adjacent to every vertex of G_1 in the other class of the bipartition.

Proof. If $|G_1| = 2$ the conclusion holds. We assume that $|G_1| \geq 4$.

For any closed M -alternating Hamilton path P'_1 of G_1 with ending vertices $w \in W$ and $b \in B$, we can find an opposite path P'_2 and a central path R' for P'_1 . Since P_2 is chosen as the shortest, $|P'_2| \geq |P_2|$ and $|R'| \leq |R|$. Similar to (6.3) we have

$$d(w) + d(b) \leq 2p_1 + |R'|/2 + 1 \leq 2p_1 + |R|/2 + 1. \quad (6.6)$$

Together with (6.2), we have

$$d(u_{2s}) + d(u_{2r-1}) + d(w) + d(b) \leq \nu + 2. \quad (6.7)$$

Since $wu_{2r-1}-$ and $bu_{2s}-$, by the condition given,

$$d(u_{2s}) + d(u_{2r-1}) + d(w) + d(b) \geq 2(\nu/2 + 1) = \nu + 2. \quad (6.8)$$

Hence all equalities in (6.6), (6.7) and (6.8) must hold. Therefore $|R| = |R'|$, $|P'_2| = |P_2|$, $d(w) = d(v_0) = \nu/2 + 1 - d(u_{2r-1})$ and $d(b) = d(v_{2p_1-1}) = \nu/2 + 1 - d(u_{2s})$. In other words, all opposite paths have the same length, all central paths have the same length, all vertices in $V(G_1) \cap W$ have the same degree $\nu/2 + 1 - d(u_{2r-1})$, and all vertices in $V(G_1) \cap B$ have the same degree $\nu/2 + 1 - d(u_{2s})$.

Let $b \neq v_{2p_1-1}$ be a vertex in $V(G_1) \cap B$, and assume that b has a neighbor $u_{2r'}$ on P_2 . Since G_1 is complete bipartite we can always find a closed M -alternating path P''_1 connecting v_0 and b in G_1 . Let $P''_2 = u_{2s}C^+u_{2r'-1}$ and $R'' = u_{2r'}C^+u_{2s-1}$. For any vertex pair $\{u_{2i-1}, u_{2i}\}$ on the path R'' , we have $e(\{u_{2i-1}, u_{2i}\}, \{u_{2s}, u_{2r'-1}\}) \leq 1$, or we get an M -alternating cycle

$$u_{2r'}C^+u_{2i-1}u_{2s}C^+u_{2r'-1}u_{2i}C^+u_{2s-1}v_0P''_1bu_{2r'},$$

which is longer than C . Therefore,

$$d(u_{2s}) + d(u_{2r'-1}) \leq |P''_2| + 2 + (|R''| - 2)/2 = |P''_2| + |R''|/2 + 1 < |P_2| + |R|/2 + 1.$$

By $d(v_0) + d(b) = d(v_0) + d(v_{2p_1-1}) = 2p_1 + |R|/2 + 1$, we have $d(u_{2s}) + d(u_{2r'-1}) + d(v_0) + d(b) < (|P_2| + |R|/2 + 1) + (2p_1 + |R|/2 + 1) = \nu + 2$. However, since $u_{2s}b-$ and $u_{2r'-1}v_0-$, by our condition, $d(u_{2s}) + d(u_{2r'-1}) + d(v_0) + d(b) \geq \nu + 2$, a contradiction. Hence b , and similarly any $w \in V(G_1) \cap W$, does not have any neighbor on P_2 . That is, $e(V(G_1), V(G_2)) = 0$.

For any closed M -alternating Hamilton path P'_1 of G_1 with ending vertices $w \in W$ and $b \in B$, let P'_2 be one opposite path of it. Since w and b send no edges to P_2 , P_2 must be part of P'_2 . However, all opposite paths have the same length, so $|P'_2| = |P_2|$, and therefore $P'_2 = P_2$. Then, $wu_{2r-1}+$ and $bu_{2s}+$. Since any vertex in G_1 can be an ending vertex of a closed M -alternating Hamilton path of G_1 , this completes the proof of the second part of the claim. \square

Claim 4. G_2 is complete bipartite, and each of u_{2s-1} and u_{2r} is adjacent to every vertex of G_2 in the other class of the bipartition.

Proof. By the above discussion, we have a cycle $C_2 = u_{2s}C^+u_{2r-1}u_{2s}$. Since $e(V(G_1), V(G_2)) = 0$, for every edge $u_{2j-1}u_{2j}$ on P_2 , where $s+1 \leq j \leq r-1$,

we can replace u_{2r-1} with u_{2j-1} and u_{2s} with u_{2j} in (6.2), (6.4) and (6.5), and all equalities must hold. So, each of u_{2j-1} and u_{2j} must be adjacent to all other vertices in the other class of the bipartition on P_2 , $u_{2j-1}u_{2r}+$ and $u_{2j}u_{2s-1}+$, therefore the claim holds. \square

For convenience we change some notations henceforth. We let $|G_2| = 2p_2$ and the vertices of G_2 be $v'_0, v'_1, \dots, v'_{2p_2-1}$, where $v'_{2j}v'_{2j+1} \in M$, for $0 \leq j \leq p_2 - 1$, and let $R = u_0u_1 \dots u_{2r-1}$.

Now we discuss the situations case by case, with respect to the length of R and the distribution of edges between R and G_i , $i = 1, 2$.

Case 1. $|R| = 2$.

Then $R = u_0u_1$. By Claim 3 and Claim 4, for any $0 \leq i \leq p_1 - 1$ and $0 \leq j \leq p_2 - 1$, $u_0v_{2i+1}+$, $u_0v'_{2j+1}+$, $u_1v_{2i}+$ and $u_1v'_{2j}+$. Therefore $G \in \mathcal{G}_1$ and $M \in \mathcal{M}_1$.

Case 2. $|R| \geq 4$.

Claim 5. For $j = 1, 2$, and every edge $u_{2i-1}u_{2i}$, $1 \leq i \leq r - 1$, exactly one of $u_{2i-1} \rightarrow G_j$ and $u_{2i} \rightarrow G_j$ holds. Furthermore, if $u_{2i-1} \rightarrow G_j$ or $u_{2i} \rightarrow G_j$, it is adjacent to all vertices of G_j in the other class of the bipartition.

Proof. Firstly, we prove that for $j = 1, 2$ and every edge $u_{2i-1}u_{2i}$, $1 \leq i \leq r - 1$, $u_{2i-1} \rightarrow G_j$ or $u_{2i} \rightarrow G_j$. By Lemma 6.2.2, the conclusion holds for G_1 . Now we prove it for G_2 . Suppose to the contrary that there exists $1 \leq l \leq r - 1$ such that $u_{2l-1} \rightarrow G_2$ and $u_{2l} \rightarrow G_2$, and let $v'_{2s} \in N_{G_2}(u_{2l-1})$ and $v'_{2t+1} \in N_{G_2}(u_{2l})$. If $|G_2| = 2$ or $t \neq s$, we can find a closed M -alternating Hamilton path Q of G_2 connecting v'_{2s} and v'_{2t-1} , and hence we have an M -alternating Hamilton cycle

$$u_0Ru_{2l-1}v'_{2s}Qv'_{2t-1}u_{2l}Ru_{2r-1}v_0P_1v_{2p_1-1}u_0$$

of G , contradicting our assumption. If $|G_2| \geq 4$ and $t = s$, let P'_2 be a closed M -alternating Hamilton path of $G_2 - \{v'_{2s}, v'_{2s+1}\}$. Then P'_2 is an opposite path for P_1 , with the central path $u_0Ru_{2l-1}v'_{2s}v'_{2s+1}u_{2l}Ru_{2r-1}$, which is shorter than P_2 , contradicting our choice of P_2 . Hence $u_{2i-1} \rightarrow G_2$ or $u_{2i} \rightarrow G_2$, for $1 \leq i \leq r - 1$.

Arbitrarily choose $0 \leq l \leq p_1 - 1$ and $0 \leq k \leq p_2 - 1$. We have $d(v_{2l}) + d(v_{2l+1}) \leq 2p_1 + 2 + (|R| - 2)/2 = 2p_1 + r + 1$ and similarly $d(v'_{2k}) + d(v'_{2k+1}) \leq$

$2p_2 + r + 1$. So

$$d(v_{2l}) + d(v_{2l+1}) + d(v'_{2k}) + d(v'_{2k+1}) \leq 2p_1 + 2p_2 + 2r + 2 = \nu + 2. \quad (6.9)$$

However $v_{2l}v'_{2k+1}-$ and $v_{2l+1}v'_{2k}-$, by the condition of the theorem,

$$d(v_{2l}) + d(v'_{2k+1}) + d(v_{2l+1}) + d(v'_{2k}) \geq 2(\nu/2 + 1) = \nu + 2, \quad (6.10)$$

and all equalities must hold. To obtain equalities, for $j = 1, 2$, and every edge $u_{2i-1}u_{2i}$, $1 \leq i \leq r - 1$, exactly one of $u_{2i-1} \rightarrow G_j$ and $u_{2i} \rightarrow G_j$ must hold. Furthermore, since l and k are arbitrarily chosen, we prove that if $u_{2i-1} \rightarrow G_j$ or $u_{2i} \rightarrow G_j$, it is adjacent to vertices of G_j in the other class of the bipartition. \square

Let $1 \leq i \leq r - 1$. We define E_1 (E'_1) to be the set of edges $u_{2i-1}u_{2i}$, where $u_{2i-1}v_{2j}+$, for every $0 \leq j \leq p_1 - 1$ ($u_{2i-1}v'_{2k}+$, for every $0 \leq k \leq p_2 - 1$), and E_2 (E'_2) to be the set of edges $u_{2i-1}u_{2i}$, where $u_{2i}v_{2j+1}+$, for every $0 \leq j \leq p_1 - 1$ ($u_{2i}v'_{2k+1}+$, for every $0 \leq k \leq p_2 - 1$).

By Claim 5, for every $1 \leq i \leq r - 1$, $u_{2i-1}u_{2i} \in E_1 \cap E'_1$, $E_1 \cap E'_2$, $E_2 \cap E'_1$ or $E_2 \cap E'_2$. Accordingly, we say that $u_{2i-1}u_{2i}$ is an edge of type I, II, III or IV for G_1 , G_2 and R . Let the number of edges $u_{2i-1}u_{2i}$ belonging to $E_1 \cap E'_1$, $E_1 \cap E'_2$, $E_2 \cap E'_1$ and $E_2 \cap E'_2$ be t_{11} , t_{12} , t_{21} and t_{22} , respectively. We have $d(v_0) = t_{11} + t_{12} + p_1 + 1$, $d(v_1) = t_{22} + t_{21} + p_1 + 1$, $d(v'_0) = t_{11} + t_{21} + p_2 + 1$ and $d(v'_1) = t_{22} + t_{12} + p_2 + 1$.

Since equalities hold in (6.9) and (6.10), we have $d(v_{2l}) + d(v'_{2k+1}) = d(v_{2l+1}) + d(v'_{2k}) = \nu/2 + 1$ for any $0 \leq l \leq p_1 - 1$ and $0 \leq k \leq p_2 - 1$. Hence

$$\begin{aligned} t_{11} + t_{22} + 2t_{12} + p_1 + p_2 + 2 &= d(v_0) + d(v'_1) \\ &= \nu/2 + 1 \\ &= d(v_1) + d(v'_0) \\ &= t_{11} + t_{22} + 2t_{21} + p_1 + p_2 + 2. \end{aligned} \quad (6.11)$$

So $t_{12} = t_{21}$.

We let $t_1 = t_{11}$, $t_2 = t_{22}$ and $t_0 = t_{12} = t_{21}$. Then $t_1 + t_2 + 2t_0 = r - 1$.

We summarise some structural results in the form of observations.

Observation 1. *If there exists $1 \leq j < i \leq r - 1$, such that $u_{2i-1}u_{2i} \in E_1 (E'_1)$ and $u_{2j-1}u_{2j} \in E'_2 (E_2)$, then $u_{2j-1}u_{2i}-$.*

Proof. If $u_{2j-1}u_{2i}+$, we obtain an M -alternating Hamilton cycle

$$u_0Ru_{2j-1}u_{2i}Ru_{2r-1}v'_0P_2v'_{2p_2-1}u_{2j}Ru_{2i-1}v_0P_1v_{2p_1-1}u_0 \\ (u_0Ru_{2j-1}u_{2i}Ru_{2r-1}v_0P_1v_{2p_1-1}u_{2j}Ru_{2i-1}v'_0P_2v'_{2p_2-1}u_0),$$

contradicting our assumption. \square

Observation 2. *If there exists $1 \leq i \leq r - 2$, such that $u_{2i-1}u_{2i} \in E_1$ and $u_{2i+1}u_{2i+2} \in E_2$, then $u_{2i}u_{2i+1}$ is a critical edge, $|G_1| = |G_2| = 2$, and exactly one of $u_{2i}v'_1+$ and $u_{2i+1}u_0+$ ($u_{2i+1}v'_0+$ and $u_{2i}u_{2r-1}+$) holds.*

If there exists $1 \leq i \leq r - 2$, such that $u_{2i-1}u_{2i} \in E'_1$ and $u_{2i+1}u_{2i+2} \in E'_2$, then $u_{2i}u_{2i+1}$ is a critical edge, $|G_1| = 2$, and exactly one of $u_{2i}v_1+$ and $u_{2i+1}u_0+$ ($u_{2i+1}v_0+$ and $u_{2i}u_{2r-1}+$) holds.

Proof. Suppose there exists $1 \leq i \leq r - 2$, such that $u_{2i-1}u_{2i} \in E_1$ and $u_{2i+1}u_{2i+2} \in E_2$. Then $u_{2i}u_{2i+1}$ is a critical edge with respect to the M -alternating cycle

$$u_0Ru_{2i-1}v_0P_1v_{2p_1-1}u_{2i+2}Ru_{2r-1}v'_0P_2v'_{2p_2-1}u_0,$$

where P_1 is an opposite path. Since G_1 is critical, $|G_1| = 2$. Since $|P_1| = 2$, and P_2 is a shortest opposite path, $|G_2| = 2$. Since $u_0v'_1$ ($u_{2r-1}v'_0$) are on a central path for the critical edge $u_{2i}u_{2i+1}$ and the opposite path v_0v_1 , exactly one of $u_{2i+1}u_0+$ and $u_{2i}v'_1+$ ($u_{2i+1}v'_0+$ and $u_{2i}u_{2r-1}+$) holds.

Now suppose there exists $1 \leq i \leq r - 2$, such that $u_{2i-1}u_{2i} \in E'_1$ and $u_{2i+1}u_{2i+2} \in E'_2$. Then $u_{2i}u_{2i+1}$ is a critical edge with respect to the M -alternating cycle

$$u_0Ru_{2i-1}v'_0P_2v'_{2p_2-1}u_{2i+2}Ru_{2r-1}v_0P_1v_{2p_1-1}u_0,$$

where P_2 is an opposite path. Since G_1 is critical, $|G_1| = 2$. Since u_0v_1 ($u_{2r-1}v_0$) are on a central path for the critical edge $u_{2i}u_{2i+1}$ and the opposite path P_2 , exactly one of $u_{2i+1}u_0+$ and $u_{2i}v_1+$ ($u_{2i+1}v_0+$ and $u_{2i}u_{2r-1}+$) holds. \square

Observation 3. *If there exists $1 \leq i < k < j \leq r - 1$, such that $u_{2i-1}u_{2i} \in E_1 (E'_1)$, $u_{2j-1}u_{2j} \in E_2 (E'_2)$, $u_{2k-1}u_{2k} \in E'_2 (E_2)$ and $u_{2k-1}u_0+$, then $u_{2i}u_{2j-1}-$.*

Proof. If $u_{2i}u_{2j-1}+$, we obtain an M -alternating Hamilton cycle

$$u_0Ru_{2i-1}v_0P_1v_{2p_1-1}u_{2j}Ru_{2r-1}v'_0P_2v'_{2p_2-1}u_{2k}Ru_{2j-1}u_{2i}Ru_{2k-1}u_0,$$

contradicting our assumption.

By symmetry, the statement holds under the other situation. \square

Claim 6. $|G_1| = 2$.

Proof. Suppose $|G_1| \geq 4$. By Observation 2, there does not exist $1 \leq i \leq r-1$, such that $u_{2i-1}u_{2i} \in E_1 (E'_1)$ and $u_{2i+1}u_{2i+2} \in E_2 (E'_2)$. Therefore, there can not exist $i < j$, such that $u_{2i-1}u_{2i} \in E_1 (E'_1)$ and $u_{2j-1}u_{2j} \in E_2 (E'_2)$. In other words, there exists an integer $0 \leq k_1 \leq r-1$ ($0 \leq k_2 \leq r-1$), such that for all $i \leq k_1$ ($j \leq k_2$), $u_{2i-1}u_{2i} \in E_2$ ($u_{2j-1}u_{2j} \in E'_2$) and for all $i > k_1$ ($j > k_2$), $u_{2i-1}u_{2i} \in E_1$ ($u_{2j-1}u_{2j} \in E'_1$). It is easily seen that $t_0 = 0$ and $k_1 = k_2$. We let $k = k_1 = k_2$.

Suppose that $t_1, t_2 \neq 0$, or equally, $1 \leq k \leq r-2$. Consider the vertices u_{2k-1} and u_{2k+2} . By Observation 1, for all $j \geq k+1$, $u_{2k-1}u_{2j-}$, and for all $j \leq k$, $u_{2k+2}u_{2j-1}$. Particularly, $u_{2k-1}u_{2k+2}$. But then we have $d(u_{2k-1}) \leq k+1$, $d(u_{2k+2}) \leq r-k$ and $d(u_{2k-1})+d(u_{2k+2}) \leq r+1 < \nu/2+1$, contradicting our condition.

Suppose one of t_1 and t_2 , say $t_1 = 0$. Then for $1 \leq i \leq r-1$, $d(u_{2i-1}) \leq r$. Moreover $d(v_0) = p_1 + 1$, so $d(u_{2i-1}) + d(v_0) \leq r + p_1 + 1 < \nu/2 + 1$ but v_0u_{2i-1} , a contradiction.

So we must have $|G_1| = 2$. \square

Claim 7. Either $t_0 = 0$, or $t_1 = t_2 = 0$.

Proof. Suppose that $t_0 > 0$, and one of t_1 and t_2 is greater than 0. Without lost of generality, we may assume that $t_1 \geq t_2$, and so $t_1 > 0$.

Let $u_{2i-1}u_{2i} \in E_1 \cap E'_1$, $1 \leq i \leq r-1$, be such that i is the maximum. Then by our condition, $d(u_{2i}) + d(v_1) \geq \nu/2 + 1$. Hence, $d(u_{2i}) \geq \nu/2 + 1 - d(v_1) = \nu/2 + 1 - (t_2 + t_0 + 2) = t_1 + t_0 + \nu/2 - r$. By Observation 1, u_{2i} can not be adjacent to any u_{2j-1} , where $u_{2j-1}u_{2j} \in E_2 \cup E'_2$ and $j < i$. Hence u_{2i} sends at least $t_1 + t_0 + \nu/2 - r - (t_1 + 1) = t_0 + \nu/2 - r - 1$ edges to $\{u_{2r-1}\} \cup \{u_{2j-1} : u_{2j-1}u_{2j} \in E_2 \cup E'_2, j > i+1\}$. Since $t_0 > 0$ and $\nu/2 - r \geq 2$, $u_{2i} \rightarrow \{u_{2j-1} : u_{2j-1}u_{2j} \in E_2 \cup E'_2, j > i+1\}$, so there exists at least one $u_{2j-1}u_{2j}$ such that $j > i+1$ and $u_{2j-1}u_{2j} \in E_2 \cup E'_2$.

By our choice of $u_{2i-1}u_{2i}$, $u_{2i+1}u_{2i+2} \in E_2 \cup E'_2$. If $u_{2i+1}u_{2i+2} \in E_2$, then by Observation 2, $u_{2i}u_{2i+1}$ is a critical edge, and exactly one of $u_{2i}v'_1+$ and $u_{2i+1}u_0+$ holds. By $u_{2i-1}u_{2i} \in E'_1$ we have $u_{2i}v'_1-$, therefore $u_{2i+1}u_0+$. If $u_{2i+1}u_{2i+2} \in E'_2$, then again by Observation 2, $u_{2i}u_{2i+1}$ is a critical edge, and exactly one of $u_{2i}v_1+$ and $u_{2i+1}u_0+$ holds. By $u_{2i-1}u_{2i} \in E_1$ we have $u_{2i}v_1-$, hence $u_{2i+1}u_0+$.

Now we discuss different situations of $u_{2i+1}u_{2i+2}$.

If $u_{2i+1}u_{2i+2} \in E_2 \cap E'_2$, let $j > i + 1$ be such that $u_{2i}u_{2j-1}+$, $u_{2j-1}u_{2j} \in E_2 \cup E'_2$. By Observation 3, $u_{2i}u_{2j-1}-$, a contradiction.

If $u_{2i+1}u_{2i+2} \in E_1 \cap E'_2$ or $E_2 \cap E'_1$, without lost of generality, we may assume that $u_{2i+1}u_{2i+2} \in E_1 \cap E'_2$. Since $u_{2i}u_{2i+1}$ is a critical edge and $u_{2i+1}v_0+$, by Observation 2, we have $u_{2i}u_{2r-1}-$. For $j > i + 1$, where $u_{2j-1}u_{2j} \in E_2$, by Observation 3, $u_{2i}u_{2j-1}-$. Therefore u_{2i} sends at least $t_0 + \nu/2 - r - 1 \geq t_0 + 1$ edges to $\{u_{2j-1} : u_{2j-1}u_{2j} \in E_1 \cap E'_2, j > i + 1\}$. However, the number of such u_{2j-1} is at most t_0 , a contradiction. \square

Case 2.1. $t_0 = 0$.

Without lost of generality, we may assume that $t_1 > 0$, and let $u_{2i-1}u_{2i} \in E_1 \cap E'_1$.

If there exists $u_{2j-1}u_{2j}$, $j < i$, such that $u_{2j-1}u_{2j} \in E_2 \cap E'_2$, then $u_{2j-1}u_{2i}-$ by Observation 1.

If there exists $u_{2j-1}u_{2j}$, $j > i + 1$, such that $u_{2j-1}u_{2j} \in E_2 \cap E'_2$, then there exists $i \leq k \leq j - 1$, such that $u_{2k-1}u_{2k} \in E_1 \cap E'_1$ and $u_{2k+1}u_{2k+2} \in E_2 \cap E'_2$. By Observation 2, $u_{2k}u_{2k+1}$ is a critical edge, and since $u_{2k+1}v_0-$ and $u_{2k}v_1-$, we have $u_{2k}u_{2r-1}+$ and $u_{2k+1}u_0+$. By Observation 3, $u_{2i}u_{2j-1}-$.

Hence, for all $u_{2j-1}u_{2j} \in E_2 \cap E'_2$, $j \neq i + 1$, $u_{2i}u_{2j-1}-$. So, $d(u_{2i}) \leq t_1 + 2$. But then

$$\begin{aligned}
\nu/2 + 1 &\leq d(u_{2i}) + d(v_1) \\
&\leq t_1 + 2 + t_2 + 2 \\
&= (\nu - 2p_2 - 2 - 2)/2 + 4 \\
&= \nu/2 - p_2 + 2.
\end{aligned} \tag{6.12}$$

Since $p_2 \geq 1$, all equalities must hold, hence $p_2 = 1$ and $2r - 1 = \nu - 5$. Furthermore, to get $d(u_{2i}) = t_1 + 2$, we must have the following.

- (a) $u_{2i+1}u_{2i+2} \in E_2 \cap E'_2$, hence $u_{2i-1}u_{2i} \neq u_{\nu-7}u_{\nu-6}$.
- (b) $u_{2i}u_{2j-1}+$, for all $u_{2j-1}u_{2j} \in E_1 \cap E'_1$.
- (c) $u_{2i}u_{\nu-5}+$.

By (a), $t_2 \geq 0$, and similarly, for any $u_{2i-1}u_{2i} \in E_2 \cap E'_2$, we can prove the following.

- (d) $u_{2i-3}u_{2i-2} \in E_1 \cap E'_1$, hence $u_{2i-1}u_{2i} \neq u_1u_2$.
- (e) $u_{2i-1}u_{2j}+$, for all $u_{2j-1}u_{2j} \in E_2 \cap E'_2$.
- (f) $u_{2i-1}u_0+$.

So, the edges $u_{2i-1}u_{2i}$, $1 \leq i \leq \nu/2 - 3$, belong to $E_1 \cap E'_1$ and $E_2 \cap E'_2$ alternatively. Moreover, $u_1u_2 \in E_1 \cap E'_1$ and $u_{\nu-7}u_{\nu-6} \in E_2 \cap E'_2$. Hence we must have $\nu = 4n + 2$, for some integer $n \geq 2$, $u_{4j+1}u_{4j+2} \in E_1 \cap E'_1$ and $u_{4j+3}u_{4j+4} \in E_2 \cap E'_2$ for $0 \leq j \leq n - 2$. The vertex sets $\{u_{4j+1}, u_{4j+2} : 0 \leq j \leq n - 2\} \cup \{v_0, v'_0, u_{4n-3}\}$, and $\{u_{4j+3}, u_{4j+4} : 0 \leq j \leq n - 2\} \cup \{v_1, v'_1, u_0\}$, induce complete bipartite subgraphs, respectively.

Let $B_1 = \{u_{4j+1} : 0 \leq j \leq n - 1\}$, $W = \{u_{4j+2} : 0 \leq j \leq n - 2\} \cup \{v_0, v'_0\}$, $B = \{u_{4j+3} : 0 \leq j \leq n - 2\} \cup \{v_1, v'_1\}$ and $W_1 = \{u_{4j} : 0 \leq j \leq n - 1\}$. By the above discussion, there can be no more edge between B and W . But we can add edges between B_1 and W_1 freely, to obtain all graphs $G \in \mathcal{G}_2$, with $M \in \mathcal{M}_2$.

Case 2.2. $t_1 = t_2 = 0$. Since $t_1 + t_2 + 2t_0 = r - 1$, we have $r = 2t_0 + 1$ and r must be odd.

If there exists $1 \leq i \leq r - 2$, such that $u_{2i-1}u_{2i} \in E_1 \cap E'_2$ and $u_{2i+1}u_{2i+2} \in E_2 \cap E'_1$ ($u_{2i-1}u_{2i} \in E_2 \cap E'_1$ and $u_{2i+1}u_{2i+2} \in E_1 \cap E'_2$), we say that an A-change (B-change) occurs at u_{2i-1} . If there exist i and j , such that $2 \leq i + 1 < j \leq r - 2$, and there is an A-change (B-change) occurring at u_{2i-1} and a B-change (A-change) occurring at u_{2j-1} , we say that a change couple occurs at (u_{2i-1}, u_{2j-1}) .

Case 2.2.1. $|G_2| \geq 4$.

There can not be any A-change, or by Observation 2, $|G_1| = |G_2| = 2$. To avoid any A-change, for $1 \leq i \leq (r - 1)/2$, $u_{2i-1}u_{2i} \in E_2 \cap E'_1$ and for $(r + 1)/2 \leq i \leq r - 1$, $u_{2i-1}u_{2i} \in E_1 \cap E'_2$.

Suppose that $r = 3$. It is not hard to see that u_0u_3- and u_2u_5- , while each of u_0u_5 and u_1u_4 can exist or not. Hence we obtain all the graphs in class \mathcal{G}_3 , except those with $n = 1$.

If $r \geq 5$, then $u_{r-1}u_r$ is a critical edge, with central path $u_{r+1}Ru_{2r-1}v_0v_1u_0Ru_{r-2}$ and opposite graph G_2 (Figure 6.6). Consider the edge v_1u_0 and u_1u_2 on the central path. We have $v_1u_{r-1}+$, $u_0 \rightarrow G_2$, $u_1 \rightarrow G_2$, and by Claim 7, u_2u_r+ . But then an A-change occurs at v_1 , a contradiction.

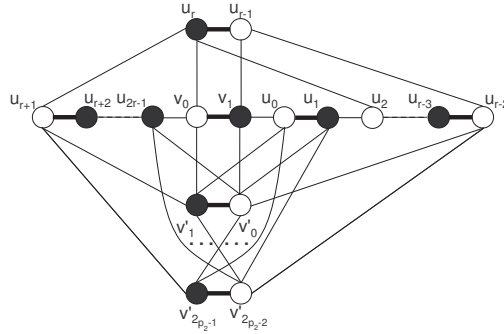


Figure 6.6: Contradiction in Case 2.2.1

Case 2.2.2. $|G_2| = 2$.

Then $\nu = 4n + 6$, for some $n \geq 1$. For $n = 1$, it is not hard to verify that $G \in \mathcal{G}_3$, $M \in \mathcal{M}_3$, and we obtain all graphs in \mathcal{G}_3 together with Case 2.2.1. For $n = 2$, it can be checked that $G = G_4$, $M = M_4$. Henceforth we assume that $n \geq 3$, and then $r = 2n + 1 \geq 7$.

We call G_1 and G_2 a critical edge pair with central path R . Since we have discussed all other cases, we may assume that for every critical edge pair and the central path, every edge of the central path that is not in M is of type II or III.

Let there be a change couple occurring at (u_{2i-1}, u_{2j-1}) . Without loss of generality, suppose that an A-change occurs at u_{2i-1} and a B-change occurs at u_{2j-1} , then $u_{2i}u_{2i+1}$ and $u_{2j}u_{2j+1}$ are critical edges. Since $u_{2i}u_{2i+1}$ and v_1v_0 is a critical edge pair, with the central path $u_{2i+2}Ru_{2r-1}v'_0v'_1u_0Ru_{2i-1}$, by our assumption, $u_{2j-1}u_{2j}$ and $u_{2j+1}u_{2j+2}$ are of type II or III. By $u_{2j}v_1+$ and $u_{2j+1}v_0+$, we have $u_{2j-1}u_{2i}+$ and $u_{2j+2}u_{2i+1}+$. Similarly, we have $u_{2i-1}u_{2j}+$

and $u_{2i+2}u_{2j+1}+$. However, then we get an M -alternating Hamilton cycle

$$u_0Ru_{2i-1}u_{2j}u_{2j+1}u_{2i+2}Ru_{2j-1}u_{2i}u_{2i+1}v'_0v'_1u_{2j+2}Ru_{2r-1}v_0v_1u_0,$$

a contradiction. Therefore, there exists no any change couple.

By symmetry, we may assume that $u_1u_2 \in E_1 \cap E'_2$, and let $r_0 > 0$, $r_1 > r_0$ and $r_2 \geq r_1$ be such that $u_1u_2, \dots, u_{2r_0-1}u_{2r_0} \in E_1 \cap E'_2$, $u_{2r_0+1}u_{2r_0+2}, \dots, u_{2r_1-1}u_{2r_1} \in E_2 \cap E'_1$, $u_{2r_1+1}u_{2r_1+2}, \dots, u_{2r_2-1}u_{2r_2} \in E_1 \cap E'_2$ and if $u_{2r_2+1}u_{2r_2+2}$ exists, $u_{2r_2+1}u_{2r_2+2} \in E_2 \cap E'_1$.

If $r_1 - r_0 \geq 2$ and $r_2 - r_1 \geq 1$, then a change couple occurs at (u_{2r_0-1}, u_{2r_1-1}) , a contradiction. Hence, $r_1 - r_0 = 1$ or $r_2 = r_1$.

If $r_1 - r_0 = 1$, then $r_2 > r_1$, and the edge $u_{2r_2+1}u_{2r_2+2}$ exists. If $r_2 - r_1 \geq 2$, a change couple occurs at (u_{2r_1-1}, u_{2r_2-1}) , a contradiction. Therefore $r_2 = r_1 + 1$. Moreover, if any B-change occurs at u_{2j-1} where $j \geq r_2 + 1$, we obtain a change couple (u_{2r_0-1}, u_{2j-1}) , again leading to a contradiction. Hence, we must have $u_{2r_2+1}u_{2r_2+2}, \dots, u_{2r-3}u_{2r-2} \in E_2 \cap E'_1$, and then $r_0 = (r - 3)/2$, $r_1 = (r - 1)/2$ and $r_2 = (r + 1)/2$.

Then $u_{r+1}u_{r+2}$ and v_1v_0 is a critical edge pair, with the central path $u_{r+3}Ru_{2r-1}v'_0v'_1u_0Ru_r$. Again we may assume that the edges of the central path not in M are of type II or III. Consider the edges $u_{r-4}u_{r-3}$ and $u_{r-2}u_{r-1}$. Since $u_{r-4}v_0+$ and $u_{r-1}v_1+$, we must have $u_{r-3}u_{r+2}+$ and $u_{r-2}u_{r+1}+$. Since $r \geq 7$, $2r - 3 > r + 3$. Consider the edge $u_{2r-3}u_{2r-2}$. Since $v_1u_{2r-2}+$, we must have $u_{2r-3}u_{r+1}+$. But then we find a change couple occurring at (u_{2r-3}, u_{r-4}) , a contradiction (Figure 6.7).

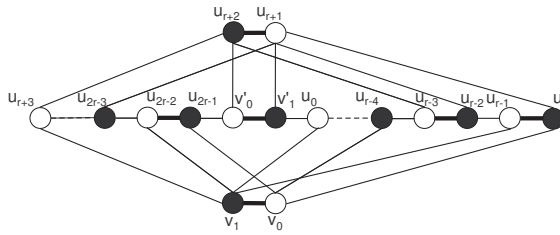


Figure 6.7: Critical pair $u_{r+1}u_{r+2}$ and v_1v_0

If $r_2 = r_1$, then $u_1u_2, \dots, u_{r-2}u_{r-1} \in E_1 \cap E'_2$ and $u_ru_{r+1}, \dots, u_{2r-3}u_{2r-2} \in$

$E_2 \cap E'_1$. Then, $u_{r-1}u_r$ and v_0v_1 is a critical pair, with the central path $u_{r+1}Ru_{2r-1}v'_0v'_1u_0Ru_{r-2}$. For the edges $u_{2i-1}u_{2i}$ with $(r+3)/2 \leq i \leq r-1$, $v_1u_{2i}+$, so we must have $u_{2i-1}u_{r-1}+$. For the edges $u_{2i-1}u_{2i}$ with $1 \leq i \leq (r-3)/2$, $v_0u_{2i-1}+$, so we must have $u_{2i}u_r+$. For the edge $u_{2r-1}v'_0$ and v'_1u_0 , we have $u_{2r-1}v_0+$, v'_0u_r+ , $v'_1u_{r-1}+$ and u_0v_1+ . Thus we reach a similar configuration as in the case that $r_1 - r_0 = 1$.

6.3 Concluding remarks

Most of the degree sum conditions for hamiltonian problems involve independent vertex sets. In our work, we try to strengthen the condition of our main theorem, by replacing “for every vertex pair u and v , where there is no arc from u to v ” with “for every vertex pair u and v ”. Naturally, if the former condition guarantees hamiltonicity without exception, then such a strengthening brings nothing. But in the case where there are exceptions, we do find some differences. Let \mathcal{D}'_1 be a subset of \mathcal{D}_1 , in which $n = m$. Let \mathcal{D}'_3 be a subset of \mathcal{D}_3 , where $n = 1$. We have the following result.

Theorem 6.3.1. Let D be a digraph. If for every vertex pair u and v , we have $d^+(u) + d^-(v) \geq |D| - 1$, then D has a directed Hamilton cycle, unless $D \in \mathcal{D}'_1$, \mathcal{D}_2 or \mathcal{D}'_3 , or $D = D_4$.

As a corollary, we can improve the Ore condition as well. Given an (undirected) graph G , if we replace every edge $uv \in E(G)$ with two arcs uv and vu , we have a digraph D . Applying Theorem 6.1.1 on D , we obtain the following result.

Let $n, m \geq 1$, and \mathcal{G}_5 be the set of graphs obtained by identifying one vertex of a complete graph K_{m+1} with one vertex of a complete graph K_{n+1} . Let \mathcal{G}_6 be the set of all graphs obtained by joining every vertex of a graph I_{n+1} to every vertex of an arbitrary graph on n vertices.

Corollary 6.3.2. Let G be a graph. If for every distinct nonadjacent vertex pair u and v , we have $d(u) + d(v) \geq |G| - 1$, then G has a Hamilton cycle, unless $G \in \mathcal{G}_5$, or $G \in \mathcal{G}_6$.

A slightly stronger result can be found in [100]. There is only one exceptional class, for it considers only 2-connected graphs.

Theorem 6.3.3. (Li, Li and Feng) Let G be a 2-connected graph with $|G| \geq 3$. If $d(u) + d(v) \geq |G| - 1$ for every pair of vertices u and v with $d(u, v) = 2$, then G has a Hamilton cycle, unless $|G|$ is odd and $G \in \mathcal{G}_6$.

Stimulated by the above results, we conjecture that the lower bound on the degree sum in the following result can be reduced by 1, with only a few exceptional cases which can be characterized clearly.

Theorem 6.3.4. (Bang-Jensen, Gutin and Li [17]) Let D be a strong digraph such that for every pair of dominating non-adjacent and every pair of dominated non-adjacent vertices $\{u, v\}$, we have $\min\{d^+(u) + d^-(v), d^-(u) + d^+(v)\} \geq |D|$. Then D has a directed Hamilton cycle.

Chapter 7

Triangle strings: structures for augmentation of vertex-disjoint triangle sets

Vertex-disjoint triangle sets (triangle sets for short) have been studied extensively. Many theoretical and computational results have been obtained. While the maximum triangle set problem can be viewed as generalization of the maximum matching problem, there seems no parallel result to Berge's augmenting path characterization on maximum matching. In this chapter, we describe a class of structures called triangle string, which turns out to be equivalent to the class of union of two triangle sets in a graph. Based on the concept of triangle sets, a sufficient and necessary condition for a triangle set to be augmented is given. Furthermore, we provide an algorithm to determine whether a graph G with maximum degree 4 is a triangle string, and if G is a triangle string, to compute a maximum triangle set of G . Finally, we give a sufficient and necessary condition for a triangle string to have a triangle factor.

We use $\langle \mathcal{T} \rangle$ to denote the graph consisting of all the triangles in a triangle set \mathcal{T} , and say that it is the *graph* of \mathcal{T} . We often consider $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$, the union of the graphs of two triangle sets \mathcal{T}_1 and \mathcal{T}_2 , which is called the *union graph* of \mathcal{T}_1 and \mathcal{T}_2 in this thesis.

Let u be a vertex of degree d in a graph G . Then we say that u is a *d-vertex* in G . Let $T = uvw$ be a triangle in G , where the degree of u , v and w are

d_u , d_v and d_w in G . Then we say that T is a (d_u, d_v, d_w) -triangle in G . We also say that the *degree sequence* of T in G is (d_u, d_v, d_w) .

7.1 Triangle strings

In this section we describe the triangle string structure. The following operations are to be used in the construction of triangle strings.

Let G be a graph. The operations (1), (2) and (3) below are called *triangle-additions*.

(1) Let w be a 2-vertex in G , add two new vertices u and v to G , and add edges to form a triangle on u , v and w .

(2) Let v and w be two nonadjacent 2-vertices at odd distance in G , add a new vertex u to G , and add edges to form a triangle on u , v and w .

(3) Let u , v and w be three pairwise nonadjacent 2-vertices at pairwise odd distance in G , add edges to form a triangle on u , v and w .

Denote by K_4^- the graph obtained from K_4 by removing an edge. An *hourglass* is a graph isomorphic to $K_5 - E(C_4)$, i.e. the graph obtained by removing the edges of a 4-cycle from a K_5 . The following operations (4) and (5) are called K_4^- -insertions.

(4) Let w be a 2-vertex in G and u, v the neighbors of w . Replace w with a K_4^- denoted by K , and let u, v be adjacent to one 2-vertex w_1 in K .

(5) Let w be the 4-vertex in an induced hourglass in G . Let the neighbors of w be u_1, u_2, v_1 and v_2 , where $u_1v_1 \in E(G)$ and $u_2v_2 \in E(G)$. Replace w with a K_4^- denoted by K , and let u_1 and v_1 be adjacent to a 2-vertex w_1 in K , u_2 and v_2 be adjacent to another 2-vertex w_2 in K .

We call a K_4^- in a graph G connected to the other parts of G through the 2-vertices in it a *swing* K_4^- . Hence, operation (4) and (5) insert swing K_4^- 's into G .

A *vertex-jointed triangle string* S is either a triangle, or obtained from another vertex-jointed triangle string S' by performing one triangle addition.

A *triangle string* S is either a K_4^- , or a vertex-jointed triangle string, or obtained from another triangle string S' by performing a K_4^- -insertion.

It is not hard to see the following properties of a triangle string S .

- (a) Every vertex in S should be of degree 2, 3 or 4.
- (b) S is vertex-jointed if and only if it contains no vertex of degree 3.
- (c) Every 2-vertex in S is contained in one triangle.
- (d) Every 3-vertex in S is a 3-vertex in a swing K_4^- , and hence it is contained in two triangles.
- (e) Every 4-vertex in S is the 4-vertex in an induced hourglass, and hence it is contained in two triangles. Furthermore, operation (5) is valid for every 4-vertex in S .
- (f) A triangle in S can only have one of the following six degree sequences in S : $(2, 2, 2)$, $(2, 2, 4)$, $(2, 4, 4)$, $(4, 4, 4)$, $(2, 3, 3)$ and $(3, 3, 4)$.

7.2 Union graph of two triangle sets and an augmenting theorem

Let \mathcal{T}_1 and \mathcal{T}_2 be two triangle sets in a graph G , and $T_1 = uvw$ be a triangle in \mathcal{T}_1 . Considering how the vertices of T_1 be covered by the triangles in \mathcal{T}_2 , we have the following cases on the structure near T_1 and the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$.

(1) All vertices of T_1 are not covered by \mathcal{T}_2 . Then T_1 forms a component of $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$, and the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ is $(2, 2, 2)$.

(2) There is one vertex u of T_1 covered by \mathcal{T}_2 , as shown in Figure 7.1. Then, u is a vertex of degree 4 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$, and the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ is $(2, 2, 4)$.

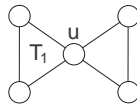


Figure 7.1: Case (2): One vertex of T_1 covered by \mathcal{T}_2

(3) There are two vertices u and v of T_1 covered by \mathcal{T}_2 . There are two possible situations.

(3.1) u and v are covered by different triangles in \mathcal{T}_2 , as shown in Figure 7.2. Then the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ is $(2, 4, 4)$.

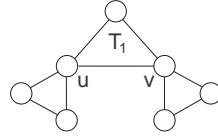


Figure 7.2: Case (3.1): Two vertices of T_1 covered by different triangles in \mathcal{T}_2

(3.2) u and v are covered by the same triangle in \mathcal{T}_2 , as shown in Figure 7.3. Then the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ is $(2, 3, 3)$.

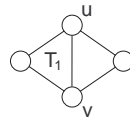


Figure 7.3: Case (3.2): Two vertices of T_1 covered by the same triangle in \mathcal{T}_2

(4) All vertices u, v and w of T_1 are covered by \mathcal{T}_2 . There are three possible situations.

(4.1) u, v and w are covered by the same triangle $T_1 \in \mathcal{T}_2$. Then the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ is $(2, 2, 2)$, and T_1 forms a component of $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$.

(4.2) Two vertices of T_1 , say u and v , are covered by one triangle in \mathcal{T}_2 , while w is covered by another triangle in \mathcal{T}_2 , as shown in Figure 7.4. Then the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ is $(4, 3, 3)$.

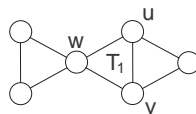


Figure 7.4: Case (4.2): Two vertices of T_1 covered by one triangle, and the other one covered by another triangle in \mathcal{T}_2

(4.3) u, v and w are covered by different triangles in \mathcal{T}_2 , as shown in Figure 7.5. Then the degree sequence of T_1 in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ is $(4, 4, 4)$.

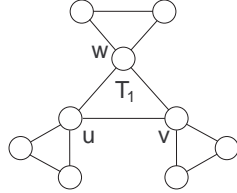


Figure 7.5: Case (4.3): All three vertices of T_1 covered by different triangles in \mathcal{T}_2

By the above discussion, we can see that (a), (c), (d), (e) and (f) also hold for the union graph $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$.

We define two operations that will be used in our proofs.

Let T be a triangle of degree sequence $(2, 2, 4)$ or $(2, 4, 4)$ or $(4, 4, 4)$ in a graph G . We removed the edges of T and the 2-vertices in T . Such an operation is called a *triangle-removal* on T . We also say that T is removed from G .

Suppose the G is a graph not isomorphic to K_4^- . Let K be a swing K_4^- in G with 2-vertices u_1 and u_2 , and 3-vertices v_1 and v_2 . Remove K from G , add a new vertex k to G , and join the neighbors of u_1 and u_2 in G other than v_1 and v_2 to k . Such an operation is called a K_4^- -*absorption* on K .

A K_4^- -absorption is obviously the reverse operation of a K_4^- -insertion. But a triangle-removal on T is not necessarily the reverse operation of a triangle-addition, because the 4-vertices in T , which become 2-vertices after removing T , may not be at pairwise odd distance in the graph obtained. However, we will prove that a triangle-removal in a union graph of two triangle sets is indeed the reverse operation of a triangle-addition.

The following lemma is an easy observation and we omit the proof.

Lemma 5. *Let \mathcal{T}_1 and \mathcal{T}_2 be two triangle sets in a graph G , then any triangle T in $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ must be an element of \mathcal{T}_1 or \mathcal{T}_2 .*

Theorem 7.1. *Let \mathcal{T}_1 and \mathcal{T}_2 be two triangle sets in a graph G . Let H be a component of $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$. Then, H is a triangle string.*

Proof. We apply mathematical induction on the number of edges in H . It is not hard to verify the conclusion for H having at most 6 edges. Assume that H has at least 7 edges, and the conclusion holds for any components of a union graph of two triangle sets with fewer edges than H .

If there is at least one 3-vertex in H , by (d), there is at least one swing K_4^- , denoted by K , in H . Let the 2-vertices in K be u_1 and u_2 , and the 3-vertices in K be v_1 and v_2 . Without loss of generality, assume that the triangles $u_1v_1v_2u_1 \in \mathcal{T}_1$ and $u_2v_1v_2u_2 \in \mathcal{T}_2$. We perform a K_4^- -absorption by replacing K with a new vertex k and obtain the graph H' . Note that u_1 is either of degree 4 or 2 in H . If $d_H(u_1) = 2$, we delete the triangle $u_1v_1v_2u_1$ from \mathcal{T}_1 . If u_1 is of degree 4 then it is contained in another triangle $u_1w_1w_2u_1 \in \mathcal{T}_2$, where $\{w_1, w_2\} \cap \{v_1, v_2\} = \emptyset$. We delete the triangle $u_1v_1v_2u_1$ from \mathcal{T}_1 and replace the triangle $u_1w_1w_2u_1$ with kw_1w_2k in \mathcal{T}_2 . Similarly we handle the triangles containing u_2 . Let \mathcal{T}'_1 and \mathcal{T}'_2 be the two triangle sets we obtain. Then H' is a component of $\langle \mathcal{T}'_1 \rangle \cup \langle \mathcal{T}'_2 \rangle$. By our induction hypothesis, H' is a triangle string. Then H is a triangle string, since it can be obtained from H' by one K_4^- -insertion.

Therefore, we can assume that there is no vertex of degree 3 in H .

Let H' be a connected graph obtained from H by performing one triangle-removal which removes the triangle T . By Lemma 5, T is either in \mathcal{T}_1 or in \mathcal{T}_2 . Without loss of generality assume that $T \in \mathcal{T}_1$. Then $H' = \langle \mathcal{T}_1 \setminus T \rangle \cup \langle \mathcal{T}_2 \rangle$, and by our induction hypothesis, H' is a triangle string.

If H contains a triangle of degree sequence $(2, 2, 2)$, then H itself must be this triangle. But this is impossible since we have assumed that H contains at least 7 edges.

If H contains a triangle T of degree sequence $(2, 2, 4)$, then we can perform a triangle-removal on T . The graph H' obtained is connected and hence a triangle string. Therefore, H is a triangle string since it is obtained from H' by a triangle-addition adding T to it.

So, we can further assume that H contains $(2, 4, 4)$ -triangles and $(4, 4, 4)$ -triangles only.

Suppose H contains at least one $(2, 4, 4)$ -triangle. We prove that there exists one $(2, 4, 4)$ -triangle T , such that removing T from H results in a connected graph H' . Assume that such a triangle does not exist, then removing

any $(2, 4, 4)$ -triangle from H results in a graph with two components. We choose a $(2, 4, 4)$ -triangle T such that removing T from H we obtain a component H_0 of the minimum order. We claim that H_0 has only one 2-vertex, which is a vertex of T with degree 4 in H . Otherwise, we have at least another $(2, 4, 4)$ -triangle T_0 in H_0 , which also has degree sequence $(2, 4, 4)$ in H . Removing T_0 from H we get a component of order smaller than H_0 , contradicting the choice of H_0 . Hence H_0 has one 2-vertex, and the other vertices are of degree 4. By Lemma 5 T is in \mathcal{T}_1 or \mathcal{T}_2 , so H_0 can be represented as the union graph of two triangle sets, denoted by \mathcal{T}'_1 and \mathcal{T}'_2 . Since there is no 3-vertex in H_0 , no triangle shares a common edge. Let the number of triangles in \mathcal{T}'_1 and \mathcal{T}'_2 be t_1 and t_2 , and assume without loss of generality that the unique 2-vertex is covered by \mathcal{T}'_1 . Since every 4-vertex is covered by one triangle from \mathcal{T}'_1 and one from \mathcal{T}'_2 , counting the edges of the triangles in \mathcal{T}'_1 and \mathcal{T}'_2 , we have $3t_1 = \frac{1}{2}(2(|H_0| - 1) + 2) = |H_0|$ and $3t_2 = \frac{1}{2}(2(|H_0| - 1)) = |H_0| - 1$, which is impossible since t_1, t_2 and $|H_0|$ are integers. Therefore, there must exist a $(2, 4, 4)$ -triangle T such that removing it from H results in a connected graph H' . Then H' is a triangle string.

Let the vertices of T with degree 4 in H be u_1 and u_2 . Then u_1 and u_2 are 2-vertices in H' . We prove that u_1 and u_2 are at odd distance in H' . Let the shortest path between u_1 and u_2 in H' be $P = x_0x_1 \dots x_l$, where $x_0 = u_1$ and $x_l = u_2$. Since P is the shortest, there could not be any edge x_ix_j in H' for $0 \leq i, j \leq l$ and $j > i + 1$. If x_1 has degree 2 in H' , then x_0 and x_2 must be connected in H' , a contradiction. Therefore x_1 , and similarly x_2, \dots, x_{l-1} , must have degree 4 in H' . Let y_0 be the vertex other than x_0 and x_1 in the triangle covering x_0 . Then y_0 cannot be joined to x_i , $2 \leq i \leq l$, or P cannot be the shortest. Since x_1 is of degree 4, there must be another triangle $x_1x_2y_1$ covering it. Similarly, y_1 cannot be joined to any x_i , $0 \leq i \leq l$ and $i \neq 1, 2$. By analogous arguments we conclude that there are triangles $x_0x_1y_0x_0, x_1x_2y_1x_1, \dots, x_{l-1}x_ly_{l-1}x_{l-1}$ in H' , where all y_i , $0 \leq i \leq l - 1$, are different. All the triangles must also be in H , and by Lemma 5, be in \mathcal{T}_1 or \mathcal{T}_2 . Any two triangles sharing a common vertex must be in different triangle sets, hence the triangles $x_{2m}x_{2m+1}y_{2m}x_{2m}$, $0 \leq m \leq \lfloor (l - 1)/2 \rfloor$ must be in the same triangle set. Since $x_0x_1y_0x_0$ and $x_{l-1}x_ly_{l-1}x_{l-1}$ are both adjacent to T in H , they must be in the same triangle set. Therefore $l - 1$ must be even, and l , the length of P , is odd. See Figure 7.6.

Because the distance between u_1 and u_2 in H' is odd, H is obtained from the triangle string H' by a triangle-addition adding T to it. By definition,

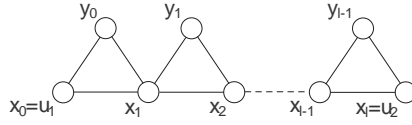


Figure 7.6: The triangles on the edges of P

H is a triangle string. Finally, we assume that H contains only $(4, 4, 4)$ -triangles. Removing a $(4, 4, 4)$ -triangle T from H , the resulting graph H' must be connected, or we will get a component of H' with one vertex of degree 2, and all other vertices of degree 4, which is impossible by the above discussion. Then H' is a triangle string. We can use arguments similar to the above to prove that the distances between any two vertices of T in H' is odd. Hence H is obtained from H' by a triangle-addition, and it is a triangle string. \square

Theorem 7.2. *Every triangle string S can be represented as the union graph of two triangle sets. Furthermore, if S contains at least two triangles the representation are unique.*

Proof. It is easy to verify the conclusion for triangle strings with at most three triangles. Let S be a triangle string with at least four triangles, and the conclusion holds for all triangle strings with fewer edges than S . S must be obtained from another triangle string S' by a triangle-addition or a K_4^- -insertion, and S' contains at least two triangles.

Suppose that S is a vertex-jointed triangle string. Then S is obtained from S' through any of the triangle-additions (1), (2) or (3), and S' is vertex-jointed. By the induction hypothesis, S' is uniquely represented as the union graph of two triangle sets \mathcal{T}'_1 and \mathcal{T}'_2 . It suffices to prove that any two 2-vertices at odd distance in S' are covered by the same triangle set, say \mathcal{T}'_1 . Then we can put the added triangle into \mathcal{T}'_2 but not into \mathcal{T}'_1 to form two triangle sets whose union graph is exactly S . And the uniqueness of these two triangle sets are guaranteed by the fact that we can only add the triangle to one of \mathcal{T}'_1 or \mathcal{T}'_2 .

Let u_1 and u_2 be two 2-vertices at odd distance in S' . We prove that u_1 and u_2 are covered by the same triangle set. Let $P = x_0x_1 \dots x_l$ be a shortest path between u_1 and u_2 , where $x_0 = u_1$, $x_l = u_2$ and l is odd. Since S' is vertex-jointed, it contains only 2-vertices and 4-vertices. By similar argument as those in the proof of Theorem 7.1, we get a serial of triangles $x_0x_1y_0x_0$, $x_1x_2y_1x_1, \dots, x_{l-1}x_ly_{l-1}x_{l-1}$, where all y_i , $0 \leq i \leq l - 1$, are different. By

Lemma 5, the triangles $x_0x_1y_0x_0, x_1x_2y_1x_1, \dots, x_{l-1}x_ly_{l-1}x_{l-1}$ must be in either \mathcal{T}'_1 or \mathcal{T}'_2 , and since they are adjacent one by one they must be in \mathcal{T}'_1 and \mathcal{T}'_2 alternatively. Since l is odd, $x_0x_1y_1x_0$ and $x_{l-1}x_ly_{l-1}x_{l-1}$ must be in the same triangle set, which implies that $u_1 = x_0$ and $u_2 = x_l$ are covered by the same triangle set.

Now suppose S is not a vertex-jointed triangle string. Then S is obtained from S' by one K_4^- -insertion, which replaces a vertex k in S' with a swing $K_4^- K$, with 2-vertices u_1 and u_2 , and 3-vertices v_1 and v_2 . By the induction hypothesis, S' can be uniquely represented as $S' = \langle \mathcal{T}'_1 \rangle \cup \langle \mathcal{T}'_2 \rangle$, where \mathcal{T}'_1 and \mathcal{T}'_2 are two triangle sets. If k is of degree 2, let the triangle containing k in S' be kw_1w_2k . Without loss of generality assume that $kw_1w_2k \in \mathcal{T}'_1$, and w_1 and w_2 are adjacent to u_1 in S . Let $\mathcal{T}_1 = (\mathcal{T}'_1 \setminus \{kw_1w_2k\}) \cup \{u_1w_1w_2u_1, u_2v_1v_2u_2\}$ and $\mathcal{T}_2 = \mathcal{T}'_2 \cup \{u_1v_1v_2u_1\}$, we have $S = \langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$, and the representation is unique since the way we put the triangles into the triangle sets is unique. If the degree of k is 4, then it must be a 4-vertex in an induced hourglass in S' . By arguments similar as in the case that k is of degree 2, we can construct \mathcal{T}_1 and \mathcal{T}_2 , such that $S = \langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ and the representation is unique. Thus the theorem is proved. \square

Let \mathcal{T} be a triangle set of a graph G . We call a triangle string S in G a \mathcal{T} -alternating triangle string if the following conditions hold.

- (1) One of every two adjacent triangles in S belongs to \mathcal{T} , and
- (2) if a 2-vertex in S is covered by a triangle $T \in \mathcal{T}$ then T is a subgraph of S .

Let S be a \mathcal{T} -alternating triangle string, and the set of the triangles in S that belong to \mathcal{T} be \mathcal{T}_1 , the set of the other triangles in S be \mathcal{T}_2 . Because one of every two adjacent triangles belongs to \mathcal{T} , the triangles in \mathcal{T}_2 are independent. So \mathcal{T}_2 is a triangle set. Since every edge in S is contained in a triangle, $S = \langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$.

Now we can state our augmentation theorem for triangle sets. Let $V(\mathcal{T})$ denote the set of the vertices of the triangles in \mathcal{T} . Let $V_2(S)$ denote the set of the 2-vertices of S .

Theorem 7.3. *Let \mathcal{T} be a triangle set of a graph G . There exists a triangle set \mathcal{T}' of G with $|\mathcal{T}'| > |\mathcal{T}|$, if and only if there exists a \mathcal{T} -alternating triangle string S , such that $|V(\mathcal{T}) \cap V_2(S)| < |V_2(S)|/2$.*

Proof. Suppose a triangle string S as described exists. If S is a triangle then it is easily seen that the conclusion holds. Assume that S contains more than two triangles. By Theorem 7.2, S can be uniquely represented as the union graph of two triangle sets, say $S = \langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$, where $\mathcal{T}_1 \subseteq \mathcal{T}$. The 3-vertices and 4-vertices of S are covered by both \mathcal{T}_1 and \mathcal{T}_2 , and every 2-vertex in S is covered by exactly one of \mathcal{T}_1 and \mathcal{T}_2 . Note that by condition (2) in the definition of \mathcal{T} -alternating triangle string, $V(\mathcal{T}_1) \cap V_2(S) = V(\mathcal{T}) \cap V_2(S)$. Therefore, $|V(\mathcal{T}_2) \cap V_2(S)| = |V_2(S)| - |V(\mathcal{T}_1) \cap V_2(S)| > |V(\mathcal{T}_1) \cap V_2(S)|$. Let $\mathcal{T}' = (\mathcal{T} \setminus \mathcal{T}_1) \cup \mathcal{T}_2$. \mathcal{T}' is a triangle set of G which covers more vertices than \mathcal{T} , hence $|\mathcal{T}'| > |\mathcal{T}|$.

Now suppose that we find a triangle set \mathcal{T}' such that $|\mathcal{T}'| > |\mathcal{T}|$. By Theorem 7.1, every component of $\langle \mathcal{T} \rangle \cup \langle \mathcal{T}' \rangle$ is a triangle string. Since $|\mathcal{T}'| > |\mathcal{T}|$, there must be a component S with more triangles from \mathcal{T}' than from \mathcal{T} , which is exactly the triangle string we are looking for. \square

7.3 Triangle sets in triangle strings: an algorithm and a condition for triangle factors

When trying to compute a triangle set in a graph G , we can often ignore the edges that are not contained in any triangle, without affecting the results. Therefore, we assume that all edges of the graph G we consider are contained in some triangles henceforth.

A triangle string has maximum degree no more than 4. It has already been known that VDT problem is APX-hard even in graphs with maximum degree 4. We prove in this section that the VDT problem is linear time solvable in triangle strings. Precisely saying, we provide an algorithm to determine whether a given graph G with maximum degree 4 is a triangle string, and if G is a triangle string the algorithm computes its maximum triangle set. A sufficient and necessary condition for a triangle string to have a triangle factor is also described.

We define the *triangle graph* $T(G)$ of G , whose vertex set consists of all triangles in G , and two vertices in $T(G)$ are adjacent if and only if the triangles they represented are adjacent in G .

Theorem 7.4. *G is a triangle string if and only if $T(G)$ is a bipartite graph.*

Proof. Suppose G is a triangle string, by Theorem 7.2, G can be represented as $\langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$, where \mathcal{T}_1 and \mathcal{T}_2 are triangle sets. By Lemma 5, every triangle in G is an element of either \mathcal{T}_1 or \mathcal{T}_2 . Therefore, $T(G)$ is a bipartite graph whose two parts consist of vertices representing the triangles in \mathcal{T}_1 and \mathcal{T}_2 respectively

For the reverse, if $T(G)$ is bipartite, let \mathcal{T}_1 be the triangle set consisting of all triangles represented by one part of $T(G)$, and \mathcal{T}_2 be the triangle set consisting of all triangles represented by the other part of $T(G)$. By the assumption that every edge of G is contained in a triangle, we have that $G = \langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$, and by Theorem 7.1, G is a triangle string. \square

By Theorem 7.4, we can construct the triangle graphs $T(G)$ of a graph G and test whether it is bipartite to determine whether G is a triangle string. Furthermore, the part of $T(G)$ with more vertices represents the maximum triangle set of G . To construct $T(G)$, we run a breadth first search on G , and when we first visit a vertex of a triangle in G , we add one vertex in $T(G)$ representing the triangle. When we visit a vertex v , we add edges between every two triangles containing v . The ideas are implemented in Algorithm 2.

Let n be the number of vertices in G . Step 1 takes $\Theta(n)$ time. Since a vertex in G has degree no more than 4, we need to check the existence of at most 6 edges in Step 5. A vertex in a triangle string is contained in at most two triangles. Therefore, if we find a vertex contained in more than two triangles, we exit from the algorithm with a negative answer in Step 6. In the loop from Step 8 to Step 12, we need to check at most two triangles. In Step 10, we need to find a vertex in $T(G)$ representing a triangle T in G . This could be accelerated by maintaining some pointers for every vertex v in G , pointing to the vertices in $T(G)$ representing the (at most two) triangles in G containing v . Then it takes constant time to find the vertex in $T(G)$. Hence, every operation from Step 4 to Step 15 takes $O(1)$ time. Further, the loop from Step 3 to Step 16 is repeated $O(n)$ times. Therefore, the total running time of the loop from Step 3 to Step 16 is $O(n)$.

In Step 17, we need to test the bipartiteness of $T(G)$. One classical method on this problem running in $O(m)$ time, where m is the number of edges in $T(G)$, is as follows.

Run a breadth first search on $T(G)$, and use two colors red and blue to

color all vertices of $T(G)$. The first visited vertex is colored red, and when visiting a vertex v , color all its uncolored neighbors different with v . After the breadth first search, all vertices of $T(G)$ are colored. Check all edges of $T(G)$. If there exists at least one edge whose two endvertices colored the same, then $T(G)$ is not bipartite. If the endvertices of every edge receive different colors, $T(G)$ is bipartite.

Since a vertex in G is contained in at most two triangles, the number of triangles in G is less than n . Furthermore, a triangle in G is adjacent to at most three triangles. Therefore, m is linear to n . Hence, Step 17 takes $O(n)$ time and Algorithm 2 is a linear time algorithm.

Finally we give a sufficient and necessary condition that a triangle string has a triangle factor.

Theorem 7.5. *A triangle string S has a triangle factor if and only if it satisfies the following.*

- (1) *The number of vertices of S is divided by 3.*
- (2) *The distance between every two 2-vertices is odd.*

Proof. Firstly we prove the theorem for vertex-jointed triangle strings.

Suppose that S is a vertex-jointed triangle string satisfying the given conditions. By Theorem 7.2, $S = \langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ for two triangle sets \mathcal{T}_1 and \mathcal{T}_2 . All 4-vertices in S is covered by both \mathcal{T}_1 and \mathcal{T}_2 . By the proof of Theorem 7.2, two 2-vertices in a vertex-jointed triangle string at odd distance are covered by the same triangle set. Hence by (2), one of \mathcal{T}_1 and \mathcal{T}_2 covered all vertices of S , and it is a triangle factor of S .

Now let S be a vertex-jointed triangle string who has a triangle factor. Obviously (1) holds. By Theorem 7.2, $S = \langle \mathcal{T}_1 \rangle \cup \langle \mathcal{T}_2 \rangle$ for two triangle sets \mathcal{T}_1 and \mathcal{T}_2 . Without loss of generality we may assume that \mathcal{T}_1 is the triangle factor of S . Using ideas similar to those in the proof of Theorem 7.1, which consider a shortest path between two 2-vertices covered by \mathcal{T}_1 , we can prove that they are at odd distance.

Now suppose that S is not a vertex-jointed triangle string. Then S contained at least one swing K_4^- . We use mathematical induction on the number of edges in S to prove the theorem. Performing a K_4^- -absorption on S , replacing a K_4^- denoted by K with the vertex k , we obtain another triangle string

S' .

Suppose that S satisfies the conditions. Since $|S'| = |S| - 3$, and a K_4^- -absorption in a triangle string does not affect the parity of the distance between any two 2-vertices, S' satisfies the conditions as well. Therefore, by the induction hypothesis, S' has a triangle factor \mathcal{T}' . It is easy to see that a triangle factor \mathcal{T} of S can then be obtained by adding one triangle of K to \mathcal{T}' , and replacing the vertex k in a triangle in \mathcal{T}' with one vertex in K .

Suppose that S has a triangle factor \mathcal{T} . Then a triangle factor \mathcal{T}' of S' can be obtained by deleting one triangle in K from \mathcal{T} , and replacing one vertex, both in K and in a triangle in \mathcal{T} , with k . By the induction hypothesis, S' satisfies the conditions. Since $|S'| = |S| - 3$, and a K_4^- -absorption in a triangle string does not affect the parity of the distance between any two 2-vertices, S satisfies the conditions. \square

Algorithm 2 Determining a triangle string and looking for its maximum triangle set

Input: A graph G with maximum degree 4;

Output: The judgement that whether G is a triangle string (YES or NO), and the maximum triangle set of G if it is a triangle string;

- 1: Mark all vertices of G as unvisited; Initial $T(G)$ as an empty graph; and Q as an empty queue;
 - 2: Take a vertex of G and push it into the rear of Q ;
 - 3: **while** Q is not empty **do**
 - 4: Pop an element v from the front of Q ;
 - 5: Find in G all triangles containing v ;
 - 6: **if** v is contained in more than two triangles in G **then return** NO;
 - 7: **end if**
 - 8: **for all** triangle T containing v , **do**
 - 9: **if** all two vertices in T other than v are not visited **then** add a new vertex representing T in $T(G)$;
 - 10: **else** find in $T(G)$ the vertex representing T ;
 - 11: **end if**
 - 12: **end for**
 - 13: Add an edge in $T(G)$ between the vertices representing the triangles containing v ;
 - 14: Push the unvisited neighbors of v into the rear of Q ;
 - 15: Mark v as visited;
 - 16: **end while**
 - 17: Test whether $T(G)$ is bipartite;
 - 18: **if** $T(G)$ is not bipartite **then return** NO;
 - 19: **else return** YES, and the triangle set represented by the part of $T(G)$ with more vertices;
 - 20: **end if**
-

Summary

This research contains results of the complexity, algorithmic and structural properties of some graph partitioning and related problems. After an introductory chapter, in the first part of the work, Chapter 2 through 4, we concentrate on the complexity, inapproximability, approximation algorithms and on-line algorithms of some graph vertex partitioning problems. In the second part of the work, Chapters 5 through 7, we focus on the structural properties of some graph problems related to matching problems which can be regarded as edge partitioning problems. We refer to the listed chapters for the details of the results.

Chapter 1 contains a short general introduction to the topics of the thesis and gives an overview of the main results, together with some motivation and connections to and relationships with older results. Specific terminology and notation can be found just before each of the topics.

In Chapter 2, we first investigate the computational complexity of the problems of determining the minimum number of monochromatic cliques or rainbow cycles that, respectively, partition the vertex set $V(G)$ of a graph G . We show that the minimum monochromatic clique partition problem is APX-hard on K_4^- -free graphs and monochromatic- K_4^- -free graphs, and APX-complete on monochromatic- K_4^- -free graphs in which the size of a maximum monochromatic clique is bounded by a constant. We also show that the minimum rainbow cycle partition problem is NP-complete, even if the input graph G is triangle-free. Moreover, for the weighted version of the minimum monochromatic clique partition problem on monochromatic- K_4^- -free graphs,

we derive an approximation algorithm with (tight) approximation guarantee $\ln |V(G)| + 1$.

In Chapter 3, we first prove that it is NP-complete to determine the injective chromatic number even restricted to some special bipartite graphs. Furthermore, we show that for every $\epsilon > 0$, it is impossible to efficiently approximate the injective chromatic number of any bipartite graph within a factor of $n^{\frac{1}{3}-\epsilon}$ unless $ZPP = NP$. Moreover, for the max-injective coloring problem, we prove that there is a constant approximation algorithm on power chordal graphs with bounded injective chromatic number. We also devise a constant approximation algorithm for max-injective coloring some bipartite graphs. For the on-line injective coloring problem, we prove that First Fit injectively colors P_3 -free graphs optimally. We also prove that the number of colors used by the derived algorithm FF^* for bipartite graph G is bounded by $\frac{3}{2}$ times the on-line injective chromatic number. Moreover, we present an improved algorithm BFF, and prove that it is optimal for on-line injectively coloring bipartite graphs.

In Chapter 4, by generalizing several earlier ideas, we present an SDP based randomized approximation algorithm for max hypergraph cut with limited unbalance (MHC-LU) with guaranteed worst-case performance ratios for various unbalance parameters $\tau = u/|V|$. We also give the worst-case performance ratio of the SDP-algorithm for approximating MHC-LU regardless of the value of τ .

In Chapter 5, we first determine the minimum size of n -factor-critical graphs and then by considering Harary graphs and related graphs, we consider the minimum size of k -extendable bipartite graphs. Moreover, we determine the minimum size of k -extendable non-bipartite graphs for $k = 1, 2$, and pose a related conjecture for general k .

In Chapter 6, we consider the problems of directed Hamilton cycles in digraphs and matching alternating Hamilton cycles in bipartite graphs. We reduce both lower bounds presented in Woodall [134] and Las Vergnas [97] by 1, and prove that their conclusions still hold, with only a few exceptional cases that can be clearly characterized.

In Chapter 7, we consider the maximum triangle set problem which can be viewed as a generalization of the maximum matching problem. We describe a class of structures called triangle string, which turns out to be equivalent to the class of union of two triangle sets in a graph. Based on the concept of triangle sets, a sufficient and necessary condition that a triangle set can be augmented is given. Furthermore, we provide an algorithm to determine whether a graph G with maximum degree 4 is a triangle string, and if G is a triangle string, compute a maximum triangle set of it. Finally, we give a sufficient and necessary condition for a triangle string to have a triangle factor.

Bibliography

- [1] A. A. Ageev and M. I. Sviridenko, Approximation algorithms for maximum coverage and max cut with given sizes of parts, in: Proceedings of Integer Programming and Combinatorial Optimization, 1999, 17-30.
- [2] A. A. Ageev and M. I. Sviridenko, An approximation algorithm for Hypergraph Max k-Cut with given sizes of parts, in: Proceedings of the 8th Annual European Symposium on Algorithms(ESA'00), M. Paterson, ed., Lecture Notes in Comput. Sci. 1879, Springer-Verlag, New York, 2000, 32-41.
- [3] S. Akbari and A. Alipour, Multicolored trees in complete graphs, J. Graph Theory **54** (2007), 221-232.
- [4] P. Alimonti, "Non-oblivious local search for graph and hypergraph coloring problems", in: Proceedings of 21st International Workshop on Graph-Theoretic Concepts in Computer Science, Lecture Notes in Computer Sciences, Vol. 1017, Springer, Berlin, 1995, 167-180.
- [5] P. Alimonti and V. Kann, Some APX-completeness results for cubic graphs, Theor. Comput. Sci. **237** (2000), 123-134.
- [6] N. Alon, R. A. Brualdi, and B.L. Shader, Multicolored forests in bipartite decompositions of graphs, J. Combin. Theory, Ser. B **53** (1991), 143-148.
- [7] N. Alon and G. Gutin, Properly colored hamiltonian cycles in edge-colored complete graphs, Random Structures Algorithms **11** (1997), 179-186.

- [8] N. Alon, T. Jiang, Z. Miller, and D. Pritikin, Properly colored subgraphs and rainbow subgraphs in edge colorings with local constraints, *Random Structures Algorithms* **23** (2003), 409-433.
- [9] D. Amar, E. Flandrin, and G. Gancarzewicz, A degree condition implying that every matching is contained in a hamiltonian cycle, *Discrete Math.* **309** (2009), 3703-3713.
- [10] D. Amar, E. Flandrin, G. Gancarzewicz, and A. P. Wojda, Bipartite graphs with every matching in a cycle, *Discrete Math.* **307** (2007), 1525-1537.
- [11] G. Andersson and L. Engebretsen, Better approximation algorithms for Set splitting and Not-All-Equal SAT, *Inform. Process. Letters* **65** (1998), 305-311.
- [12] G. Andersson, An approximation algorithm for max p-section, in: *Proceedings of 16th International Symposium on Theoretical Aspects of Computer Science*, volume 1563 of *Lecture Notes in Computer Science*, 237-247. Springer-Verlag, Berlin, 1999.
- [13] K. Andreev and H. Räcke, Balanced Graph Partitioning, in: *Proceedings of the sixteenth annual ACM symposium on Parallelism in algorithms and architectures*, 120-124. Barcelona, Spain, 2004.
- [14] S. Arora, D. Karger, and M. Karpinski, Polynomial Time Approximation Schemes for Dense Instances of NP-hard problems, *J. of Comput. and Syst. Sci.* **58** (1999), 193-210.
- [15] P. Austrin, S. Benabbas, and K. Georgiou, Better Balance by Being Biased: A 0.8776-Approximation for Max Bisection, Full version available as arXiv eprint 1205.0458.
- [16] J. Bang-Jensen, G. Gutin, and J. Huang, A sufficient condition for a semicomplete multipartite digraph to be Hamiltonian, *Discrete Math.* **161** (1996), 1-12.
- [17] J. Bang-Jensen, G. Gutin, and H. Li, Sufficient conditions for a digraph to be Hamiltonian, *J. Graph Theory* **22** (1996), 181-187.
- [18] J. Bang-Jensen, Y. Guo, and A. Yeo, A new sufficient condition for a digraph to be Hamiltonian, *Discrete Appl. Math.* **95** (1999), 61-72.

- [19] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, Algorithms, and Applications*, Springer-Verlag, London, 2001.
- [20] C. Berge, Two theorems in graph theory. *Proc. Nat. Acad. Sci. U.S.A.* **43** (1957), 842-844.
- [21] K. A. Berman, Proof of a conjecture of Häggkvist on cycles and independent edges, *Discrete Math.* **46** (1993), 9-13.
- [22] J. C. Bermond, E. Darrot, O. Delmas, and S. Perennes, Hamilton circuits in the directed Butterfly network, *Discrete Appl. Math.* **84** (1998), 21-42.
- [23] B. Bollobás and A. D. Scott, Problems and results on judicious partitions, *Random Struct. Alg.* **21** (2002), 414-430.
- [24] B. Bollobás and A. D. Scott, Judicious partitions of bounded-degree graphs, *J. Graph Theory* **46** (2004), 131-143.
- [25] J. A. Bondy and U. S. R. Murty, *Graph theory with applications*, Macmillan Press, London, 1976.
- [26] H. J. Broersma, On some intriguing problems in hamiltonian graph theory - a survey, *Discrete Math.* **251** (2002), 47-69.
- [27] H. J. Broersma, A. Capponi, and D. Paulusma, A new algorithm for on-line coloring bipartite graphs, *SIAM J. on Discrete Mathematics* **22** (1) (2008), 72-91.
- [28] H. J. Broersma, F. V. Fomin, J. Kratochvíl, and G. J. Woeginger, Planar graph coloring avoiding monochromatic subgraphs: trees and paths make it difficult, *Algorithmica* **44** (2006), 343-361.
- [29] H. J. Broersma, P. A. Golovach, D. Paulusma, and J. Song, Determining the chromatic number of triangle-free $2P_3$ -free graphs in polynomial time, *Theor. Comput. Sci.* **423** (2012), 1-10.
- [30] R. A. Brualdi and S. Hollingsworth, Multicolored trees in complete graphs, *J. Combin. Theory Ser. B* **68** (1996), 310-313.
- [31] M. Chlebík and J. Chlebíková, Approximation hardness for small occurrence instances of NP-hard problems, *Proc. Fifth Conf. Algorithms and Complexity (CIAC)*, Lecture Notes in Computer Science, Vol. 2653, Springer, Berlin, 2003, 152-164.

- [32] M. Chlebík and J. Chlebíková, Complexity of approximating bounded variants of optimization problems, *Theoret. Comput. Sci.* **354** (2006), 320-338.
- [33] V. Chvátal, A greedy heuristic for the set-covering problem, *Mathematics of Operations Research*, **4** (1979), 233-235.
- [34] D. Christofides, P. Keevash, D. Kühn and D. Osthus, A semiexact degree condition for Hamilton cycles in digraphs, *SIAM J. Discrete Math.* **24** (2010), 709-756.
- [35] J. Cong, H. P. Li, S. K. Lim, T. Shibuya and D. Xu, Large Scale Circuit Partitioning with Loose Stable Net Removal and Signal Flow Based Clustering, *ICCAD* (1997), 441-446.
- [36] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization*, John Wiley and Sons, New York, 1998.
- [37] K. Corrádi and A. Hajnal, On the maximal number of independent circuits in a graph, *Acta Math. Acad. Sci. Hungary* **14** (1963), 423-439.
- [38] S. K. Darbinyan, A sufficient condition for the Hamiltonian property of digraphs with large semidegrees (in Russian), *Akad. Nauk. Armyan. SSR Dokl.* **82** (1986), 6-8.
- [39] S. K. Darbinyan, On hamiltonian bypasses in digraphs satisfying Meyniel-like conditions (in Russian), *Math Problems in Computer Science* **20** (1998), 7-19.
- [40] G. A. Dirac, Some theorems on abstract graphs, *Proc. Lond. Math. Soc.* **2** (1952), 69-81.
- [41] P. Erdős, A. Gyárfás, and L. Pyber, Vertex coverings by monochromatic cycles and trees, *J. Combin. Theory, Ser. B* **51** (1991), 90-95.
- [42] O. Favaron, On k -factor-critical graphs, *Discuss. Math. Graph Theory* **16** (1996), 41-51.
- [43] T. Feder, P. Hell, S. Klein, and R. Motwani, Complexity of graph partition problems, in: *Proceedings of Thirty-First Annual ACM Symposium on Theory of Computing*, 1999, 464-472.

- [44] T. Feder and R. Motwani, Clique partitions, graph compression and speeding-up algorithms, *J. Computer and System Sciences* **51** (1995), 261-272.
- [45] U. Feige and M. X. Goemans, Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT, *Proc. of 3rd Israel Symposium on the Theory of Computing and Systems*, 1995, 182-189.
- [46] U. Feige, M. Karpinski, and M. Langberg, A note on approximating Max-Bisection on regular graphs, *Inform. Process. Lett.* **79** (2001), 181-188.
- [47] U. Feige, M. Karpinski, and M. Langberg, Improved approximation of max-cut on graphs of bounded degree, *J. Algorithms* **43** (2002), 201-219.
- [48] U. Feige and J. Killian, Zero Knowledge and the Chromatic Number, *J. of Computer and System Sciences* **57** (1998), 187-199.
- [49] U. Feige and M. Langberg, Approximation algorithms for maximization problems arising in graph partitioning, *J. of Algorithms* **41** (2001), 174-211.
- [50] U. Feige and M. Langberg, The RPR² rounding technique for semidefinite programs, *J. of Algorithms* **60** (2006), 1-23.
- [51] J. Fiala, K. Jansen, V. B. Le, and E. Seidel, Graph subcoloring: complexity and algorithms, *SIAM J. Discrete Math.* **16** (2003), 635-650.
- [52] V. Frieze and M. Jerrum, Improved approximation algorithms for max k-cut and max bisection, *Algorithmica* **18** (1997), 67-81.
- [53] G. Galbiati and F. Maffioli, Approximation algorithms for maximum cut with limited unbalance, *Theor. Comput. Sci.* **385** (2007), 78-87.
- [54] M. R. Garey and D. S. Johnson, *Computers and Intractability: A guide to the theory of NP-completeness*, W. H. Freeman and co., New York, 1979.
- [55] D. R. Gaur and R. Krishnamurti, Simple approximation algorithms for MAXNAEESP and Hypergraph 2-colorability, *J. Comb. Optim.* **5** (2001), 167-173.

- [56] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit hamiltonien, *C. R. Acad. Sci. Paris Sér. A-B* **251** (1960), 495-497.
- [57] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *J. ACM* **42** (1995), 1115-1145.
- [58] R. J. Gould, Updating the hamiltonian problem: a survey, *J. Graph Theory* **15** (1991), 121-157.
- [59] R. J. Gould, Advances on the hamiltonian problem: A survey, *Graphs Combin.* **19** (2003), 7-52.
- [60] L. Gourvès, A. Lyra, C. A. Martinhon, and J. Monnot, Complexity of trails, paths and circuits in arc-colored digraphs, *Discrete Appl. Math.* **161** (2013), 819-828.
- [61] R. Govindarajan and S. Rengarajan, Buffer allocation in regular data flow networks: An approach based on coloring circular-arc graphs, in: *Proceedings of the 2nd International Conference on High Performance Computing*, 1996.
- [62] D. Guan and X. Zhu, A coloring problem for weighted graphs, *Inform. Process. Lett.* **61** (1997), 77-81.
- [63] V. Guruswami, C. Rangan, M. S. Chang, G. J. Chang, and C. K. Wong, The vertex-disjoint triangles problem, In J. Hromkovič and O. Sýkora (Eds.): *Proceedings of WG'98*, LNCS 1517, 1998, 26-37.
- [64] V. Guruswami, Inapproximability results for set splitting and satisfiability problems with no mixed clauses, *Algorithmica* **38** (2003), 451-469.
- [65] A. Gyárfás and J. Lehel, On-Line and First Fit Coloring of Graphs, *J. of Graph Theory* **12** (1988), 217-227.
- [66] A. Gyárfás, M. Ruszinkó, G.N. Sárközy, and E. Szemerédi, An improved bound for the monochromatic cycle partition number, *J. Combin. Theory Ser. B* **96** (2006), 855-873.
- [67] A. Gyárfás and G. Simonyi, Edge colorings of complete graphs without tricolored triangles, *J. Graph Theory* **46** (2004), 211-216.

- [68] G. Hahn, J. Kratochvíl, J. KŠiráň, and D. Sotteau, On the injective chromatic number of graphs, *Discrete Mathematics* **256** (2002), 179-192.
- [69] R. Häggkvist, On F-hamiltonian graphs, *Graph Theory and Related Topics*, J.A. Bondy and U.S.R. Murty (Editors), Academic Press, New York, 1979, 219-231.
- [70] E. Halperin and U. Zwick, A unified framework for obtaining improved approximation algorithms for maximum graph bisection problems, *Random Struct. Alg.* **20** (2002), 382-402.
- [71] Q. Han, Y. Ye, and J. Zhang, An improved rounding method and semidefinite programming relaxation for graph partition, *Mathematical Programming* **92** (2002), 509-535.
- [72] F. Harary, The maximum connectivity of a graph. *Proc. Nat. Acad. Sci. U.S.A.* **48** (1962), 1142-1164.
- [73] R. Hassin and S. Rubinstein, An approximation algorithm for maximum triangle packing, *Discrete Appl. Math.* **154** (2006), 971-979.
- [74] R. Hassin and S. Rubinstein, Approximation algorithms for maximum linear arrangement, in: *Proceedings of the 7th Scandinavian Workshop on Algorithm Theory (SWAT'00)*, M. M. Halldórsson, ed., *Lecture Notes In Comput. Sci.* 1851, Springer-Verlag, New York, 2000, 231-236.
- [75] J. Håstad, Some optimal inapproximability results, *J. ACM* **48** (2001), 798-859.
- [76] S. Hauck and G. Borriello, An Evaluation of Bipartitioning Techniques, *IEEE Trans. on CAD*, **16** (1997), 849-866.
- [77] P. Hell, A. Raspaud, and J. Stacho, On injective colourings of chordal graphs, *Lecture Notes in Comput. Sci.* **4957** (2008), 520-530.
- [78] C. T. Hoàng and V. B. Le, P_4 -free colorings and P_4 -bipartite graphs, *Discrete Math. Theor. Comput. Sci.* **4** (2001), 109-122.
- [79] S. Irani, Coloring inductive graphs on-line, *Algorithmica* **11** (1994), 53-72.
- [80] B. Jackson and N.C. Wormald, Cycles containing matchings and pairwise compatible Euler tours, *J. Graph Theory* **14** (1990), 127-138.

- [81] G. Jäger and A. Srivastav, Improved approximation algorithms for maximum graph partitioning problems, *J. Comb. Optim.* **10** (2005), 133-167.
- [82] Z. Jin, M. Kano, X. Li, and B. Wei, Partitioning 2-edge-colored complete multipartite graphs into monochromatic cycles, paths and trees, *J. Comb. Optim.* **11** (2006), 445-454.
- [83] R. Johansson, Triangle-factors in a balanced blown-up triangle. *Discrete Math.* **211** (2000), 249-254.
- [84] V. Kann, Maximum bounded 3-dimensional matching is MAX SNP-complete, *Inform. Process. Lett.* **37** (1991), 27-35.
- [85] V. Kann, J. Lagergren, and A. Panconesi, Approximability of maximum splitting of k-sets and some other APX-complete problems, *Inform. Process. Lett.* **58** (1996), 105-110.
- [86] M. Kano and X. Li, Monochromatic and heterochromatic subgraphs in edge-colored graphs - a survey, *Graphs Combin.* **24** (2008), 237-263.
- [87] G. Karypis, R. Aggarwal, V. Kumar, and S. Shekhar, Multilevel Hypergraph Partitioning: Applications in VLSI Design, *DAC*, (1997), 526-529.
- [88] K. Kawarabayashi, One or two disjoint cycles cover independent edges: Lovász-Woodall Conjecture, *J. Combin. Theory Ser. B* **84** (2002), 1-44.
- [89] K. Kawarabayashi, A survey on Hamiltonian cycles, *Interdisciplinary Information Sciences* **7** (2001), 25-39.
- [90] P. Keevash, D. Kühn, and D. Osthus, An exact minimum degree condition for Hamilton cycles in oriented graphs, *J. Lond. Math. Soc.* **79** (2009), 144-166.
- [91] L. Kelly, D. Kühn, and D. Osthus, A Dirac type result on Hamilton cycles in oriented graphs, *Combin. Probab. Comput.* **17** (2008), 689-709.
- [92] H. A. Kierstead and W. T. Trotter, An extremal problem in recursive combinatorics, *Congress. Numer* **33** (1981), 143-153.
- [93] D. Kühn, D. Osthus, and A. Treglown, Hamiltonian degree sequences in digraphs, *J. Combin. Theory Ser. B* **100** (2010), 367-380.

- [94] D. Král, J. Kratochvíl, Z. Tuza, and G. J. Woeginger, Complexity of coloring graphs without forbidden induced subgraphs, *Lecture Notes in Computer Science* **2204** (2001), 254-262.
- [95] D. Kühn and D. Osthus, A survey on Hamilton cycles in directed graphs, *Eur. J. Comb.* **33** (2012), 750-766.
- [96] M. Larsen, J. Propp, and D. Ullman, The fractional chromatic number of mycielski's graphs, *J. of Graph Theory* **19** (1995), 411-416.
- [97] M. Las Vergnas, Problèmes de couplages et problèmes hamiltoniens en théorie des graphes, Ph.D. Thesis, Université Paris XI, 1972.
- [98] C. Lee, P. Loh, and B. Sudakov, Bisections of graphs, *J. Combinatorial Theory Ser. B* **103** (2013), 599-629.
- [99] B. Li, H. Broersma, and S. Zhang, Pairs of forbidden induced subgraphs for homogeneously traceable graphs, *Discrete Math.* **312** (2012), 2800-2818.
- [100] S. Li, R. Li, and J. Feng, An efficient condition for a graph to be Hamiltonian, *Discrete Appl. Math.* **155** (2007), 1842-1845.
- [101] X. Li and W. A. Zang, A combinatorial algorithm for minimum weighted colorings of claw-free perfect graphs, *Journal of Combinatorial Optimization* **9** (2005), 331-347.
- [102] X. Li and X. Y. Zhang, On the minimum monochromatic or multicolored subgraph partition problems, *Theor. Comput. Sci.* **385** (2007), 1-10.
- [103] X. Li and W. Zhou, The 2nd Order Conditional 3-Coloring of Claw-free Graphs, *Theoretical Computer Science*, **396** (2008), 151-157.
- [104] Y. Li and Z. Nie, A note on n -critical bipartite graphs and its application, in: Du, D., Hu, X., Pardalos, P.M. (eds.) *COCOA 2009*, LNCS 5573, pp. 279-286. Springer (2009).
- [105] G. Liu and Q. Yu, On n -edge-deletable and n -critical graphs, *Bull. Inst. Combin. Appl.* **24** (1998), 65-72.
- [106] L. T. Liu, M. T. Kuo, S. C. Huang, and C. K. Cheng, A Gradient Method on the Initial Partition of Fiduccia-Mattheyses Algorithm, *ICCAD* (1995), 229-234.

- [107] D. Lou and Q. Yu, Connectivity of k -extendable graphs with large k , *Discrete Appl. Math.* **136** (2004), 55-61.
- [108] L. Lovász, Coverings and colorings of hypergraphs, in: *Proceedings of the 4th Southeastern Conference on Combinatorics, Graph Theory, and Computing*, Utilitas Mathematica Publishing, Winnipeg, 1973, 3-12.
- [109] L. Lovász, M. Saks, and W. T. Trotter, An on-line graph coloring algorithm with sublinear performance ratio, *Discrete Mathematics* **75** (1989), 319-325.
- [110] G. MacGillivray and M. Yu, Generalized partitions of graphs, *Discrete Appl. Math.* **91** (1999), 143-153.
- [111] Y. Manoussakis, Directed hamiltonian graphs, *J. Graph Theory* **16** (1992), 51-59.
- [112] P. Maschlanka and L. Volkmann, Independence number in n -extendable graphs, *Discrete Math.* **154** (1996), 167-178.
- [113] H. Meyniel, Une condition suffisante d'existence d'un circuit hamiltonien dans un graphe orienté, *J. Combin. Theory Ser. B* **14** (1973), 137-147.
- [114] C. ST. J. A. Nash-Williams, Hamilton circuits in graphs and digraphs, *The many facets of graph theory*, Lecture Notes in Math. 110, Springer, Berlin, 1969, 237-243.
- [115] Y. E. Nesterov, Semidefinite relaxation and nonconvex quadratic optimization, *Optimization Methods and Software*, **9** (1998), 141-160.
- [116] O. Ore, A Note on Hamilton circuits, *Amer. Math. Monthly* **67** (1960), 55.
- [117] K. Ota and G. Sueiro, Forbidden induced subgraphs for toughness, *J. Graph Theory*, **73** (2013), 191C202.
- [118] S. B. Patkar and H. Narayanan, An Efficient Practical Heuristic For Good Ratio-Cut Partitioning, *VLSI Design, 2003. International Conference on*, 64-69.
- [119] S. V. Pemmaraju, R. Raman, and K. Varadarajan, Max-coloring and Online Coloring with Bandwidths on Interval Graphs, *ACM Transactions on Algorithms* **7** (2011), 1-23.

- [120] E. Petrank, The hardness of approximation: gap location, *Comput. Complex.* **4** (1994), 133-157.
- [121] A. R. Philpotts, Matchings, factors and cycles in graphs, Ph.D. Thesis, University of Nottingham, 2008.
- [122] M. D. Plummer, On n -extendable graphs, *Discrete Math.* **31** (1980), 201-210.
- [123] M. D. Plummer, Matching extension in bipartite graphs, *Congress. Numer.* **54** (1986), 245-258.
- [124] M. D. Plummer, Extending matchings in graphs: a survey, *Discrete Math.* **127** (1994), 277-292.
- [125] M. D. Plummer, Extending matchings in graphs: an update, *Congress. Numer.* **116** (1996), 3-32.
- [126] M. D. Plummer, Recent progress in matching extension. in: Grötschel, M., Katona, G.O.H. (eds.) *Building Bridges: Between Mathematics and Computer Science*. Bolyai Society Mathematical Studies, Volume 19, pp. 427-454 (2008).
- [127] P. Raghavendra and N. Tan, Approximating CSPs with global cardinality constraints using SDP hierarchies, *InSODA'12, 2012*, 373-387.
- [128] L. Sanchis, Multiple-way Network Partitioning with Different Cost Functions, *IEEE Trans. on Comp.*, **42** (1993), 1500-1504.
- [129] P. Sanders and C. Schulz, Think Locally, Act Globally: Highly Balanced Graph Partitioning, *SEA 2013*, 164-175
- [130] K. Suzuki, A necessary and sufficient condition for the existence of a heterochromatic spanning tree in a graph, *Graphs Combin.* **22** (2006), 261-269.
- [131] C. Thomassen, Long cycles in digraphs, *Proc. Lond. Math. Soc.* **42** (1981), 231-251.
- [132] H. Wang, Vertex-disjoint triangles in claw-free graphs with minimum degree at least three. *Combinatorica* **18** (1998), 441-447.
- [133] A. P. Wojda, Hamiltonian cycles through matchings, *Demonstratio Math.* **XXI** (1983), 547-553.

- [134] D. R. Woodall, Sufficient conditions for circuits in graphs, *Proc. Lond. Math. Soc.* **24** (1972), 739-755.
- [135] B. Xu, J. Yan, and X. Yu, Balanced judicious partitions of graphs, *J. Graph Theory* **63** (2010), 210-225
- [136] B. Xu, J. Yan, and X. Yu, A note on balanced bipartition, *Disc. Math.* **310** (2010), 2613-2617.
- [137] B. Xu and X. Yu, On judicious bisections of graphs, personal communication (2013).
- [138] Y. Ye, A .699-approximation algorithm for max-bisection, *Math. Program. Ser. A* **90** (2001), 101-111.
- [139] V. Yegnanarayanan, Graph colorings and partitions, *Theor. Comput. Sci.* **263** (2001), 59-74.
- [140] Q. Yu, Characterizations of various matching extensions in graphs, *Australas. J. of Combin.* **7** (1993), 55-64.
- [141] Q. Yu and G. Liu, *Graph Factors and Matching Extensions*, Higher Education Press, Beijing (2009).
- [142] H. Zhang and F. Zhang, New lower bound on the number of perfect matchings in fullerene graphs, *J. Math. Chem.* **30** (2001), 343-347.
- [143] J. Zhang, Y. Ye and Q. Han, Improved Approximations for Max Set Splitting and Max NAE SAT, *Disc. Appl. Math.* **142** (2004), 133-149.
- [144] Z. B. Zhang, T. Wang, and D. Lou, Equivalence between extendibility and factor-criticality, *Ars Combin.* **85** (2007), 279-285.
- [145] L. C. Zhao and J. H. Meng, A sufficient condition for hamiltonian cycles in digraphs, *Ars Combin.* **32** (1991), 335-338.
- [146] G. Zhou and K. Zhang, A sufficient condition for a semicomplete multipartite digraph to be hamiltonian, *Australas. J. Combin.* **19** (1999), 231-234.
- [147] U. Zwick, Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint, Ninth SODA, 1998, 201-210.

- [148] U. Zwick, Outward rotations: a tool for rounding solutions of semidefinite programming relaxations, with applications to MAX CUT and other problems, in: Proceedings of the 31st Annual ACM Symposium on Theory of Computing, 1999, 679-687.

Acknowledgements

All work that I have done could not have been finished without the stimulation, cooperation and support of many persons. It was this support that kept me going on this road, and made this trip into one of the most important trips of my life. I am grateful to all who supported me on this journey one way or the other, including some persons which I may forget to mention here, although not intentionally.

First of all, I would like to thank Professor Dr. Xueliang Li who gave me his endless help and established the connection with the DMMP group of the University of Twente for me. And I greatly appreciate my supervisors Professor Dr. Hajo Broersma, Professor Dr. Marc Uetz and Dr. Georg Still for offering me this opportunity to be a PhD. student at the University of Twente. I really enjoyed studying in the DMMP group.

I would like to thank Hajo Broersma whose stimulating enthusiasm and energy were of great help for me to finish the thesis. We have known each other for more than a decade since he visited Nankai University for the first time in 2003. I had the chance to study the theory of computational complexity from him since then. The expert suggestions, technical advice and the motivational encouragement I received from him significantly improved my enthusiasm to academic research. This thesis could not have been written without his stimulation, instruction and cooperation. His profound knowledge and scrupulous scientific spirit have been very helpful in the past and will also be useful in my further research. I am very grateful for his help and supervision.

I would like to thank Marc Uetz, the chair of the DMMP group, who is

so warmhearted and has offered much help for my research. He is easy to approach and is always ready to help others. His expertise suggestions and profound knowledge of combinatorial optimization were very helpful to me. He also guided me into the fascinating field of algorithmic game theory which I am very interested in. My life in Twente has been delightful because he created a free working atmosphere and comfortable living conditions by his innovation and enthusiasm.

I would like to express my deep gratitude to Georg Still whose encouragement and endless help have been accompanying me since the first day I moved into the Netherlands. He provided me with continuous encouragement and advice, while letting me free to explore and research on my own. He has patiently guided me in writing papers in English. I am indebted to him so much not only for his help during my study, but also for his assistance in handling many problems in my daily life. His fatherly love and friendly smile as well as the selfless help will never go out of my memories.

Special thank to Bodo Manthey for his valuable discussions with me on some interesting problems. When my research was in trouble, he unselfishly spent lots of time discussing with me and gave me many valuable comments. I also appreciate Walter Kern for his helpful suggestion on some new and interesting problems in combinatorial optimization.

I want to thank other colleagues in our group, and in particular, I am very grateful to our secretary Marjo Mulder who arranged many wonderful events in our group and offered me a lot of help for my personal life in the Netherlands.

I am also grateful to Professor J. van den Heuvel for his helpful discussion and suggestion on the problems in the future research. And lots of thanks also go to my coauthors, Professor Baogang Xu, Professor Xingxing Yu and Dr. Zanbo Zhang for advising me well to go ahead despite difficulties. I learned a lot from their professional attitude towards work and the spirit to pursue perfection. Special thank to Zanbo for his valuable discussion and suggestion. We cooperated closely and had much fun in discussing academic problems. Many important results in my research were obtained in cooperation with

him.

And I would express my deep thanks to all the friends who I met in Macandra as well as in campus. My life in the Netherlands wouldn't have been so beautiful without them, my great friends, who brought me lots of laughter and happiness. Specially, those happy parties and pleasant journey must have remained in my memory. It was really an enjoyable and memorable time of staying with all of them.

Finally, I want to express my deep thanks to my parents and parents-in-law. They provided me with generous support and encouragement during these years, even though they did not always understand what I was doing. Last but not least, I gratefully acknowledge my wife Jing Tang for her love, support, patience and the best gift I have ever received from her, my son Tangtang.

Xiaoyan Zhang

April 2014, Enschede

Index

- 3-dimensional matching problem, 20
- A-injective chromatic number, 43
- balanced, 11
- bicritical, 15
- max bisection, 10
- clique partition, 2
- closed M -alternating path, 102
- color-adjacent, 7
- common neighbor graph, 40
- competitive ratio, 44
- complete graph, 6
- covered, 19
- edge-colored graph, 6
- factor-critical, 15
- fractional chromatic number, 40
- Harary graph, 86
- hyperedge, 2
- hypergraphs, 2, 9
- hypergraph bisection, 10
- hue, 44
- independence number, 43
- induced-hereditary, 42
- injective chromatic number, 7
- injective k -coloring, 7
- K_4^- -free, 6
- k -extendable, 15
- k -fold b -coloring, 40
- k -fold chromatic number, 40
- k -fold coloring, 40
- M -alternating path, 102
- M -alternating cycle, 102
- max-coloring problem, 8
- max cut, 10
- max Ek -set splitting, 11
- max hypergraph cut, 10
- max hypergraph cut with limited unbalance, 10
- max-injective coloring problem, 8, 41
- MAX NAE- $\{3\}$ -SAT, 12
- MAX NAE SAT, 12
- max set splitting, 11
- maximal matching, 2
- maximal monochromatic clique, 7
- maximum cut with limited unbalance, 10
- maximum matching, 2
- matching, 2, 102
- monochromatic, 23

- monochromatic clique, 6
- monochromatic- K_4^- -free, 6

- n -critical, 15
- n -factor-critical, 15

- on-line competitive ratio, 44
- on-line injective chromatic number, 44
- on-line injective coloring, 9
- on-line injective coloring algorithm, 44
- open M -alternating path, 102

- P_3 -free graphs, 44
- perfect matching, 2,102
- perfect triangle set, 19
- power chordal graph, 41

- rainbow, 23
- rainbow cycle, 6

- transitive tournament, 17
- triangle factor, 19
- triangle-free graphs, 44
- triangle set, 19
- triangle strings, 21,122

- unbalance parameter, 13

- vertex-clique, 6
- vertex coloring, 2
- vertex-cycle, 6
- vertex-disjoint triangle set, 19
- vertex-disjoint triangle string, 122