$\sim\alpha(G)\&\chi(G)\ via SDP$

N. Gvozdenović

Approximating $\alpha(G)$ and $\chi(G)$ of $G$ via Semidefinite Programming

$\sim\alpha(G)\&\chi(G)\ via SDP$

# Approximating *the* Stability Number *and the* Chromatic Number *of a* Graph *via* Semidefinite Programming

Nebojša Gvozdenović

# Approximating the Stability Number and the Chromatic Number of a Graph via Semidefinite Programming

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. D.C. van den Boom
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Agnietenkapel
op donderdag 10 april 2008, te 12.00 uur

door

Nebojša Gvozdenović

geboren te Subotica, Servië.

Promotiecommissie:

Promotor:
prof.dr. A. Schrijver

Co-promotor:
dr. M. Laurent

Overige leden:
prof.dr. H.M. Buhrman
prof.dr. A.M.H. Gerards
prof.dr. F. Rendl
prof.dr. C. Roos
prof.dr. G.J. Woeginger

Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Universiteit van Amsterdam

CWI

NWO
Nederlandse Organisatie voor Wetenschappelijk Onderzoek

The research reported in this thesis has been carried out at the Centrum voor Wiskunde en Informatica (CWI).

It was supported by the Netherlands Organization for Scientific Research (NWO) under project number 639.032.203.



Printed and bound by Ponsen & Looijen b.v., Wageningen.

ISBN: 90–6196–546–2

U sećanje na Caleta

# Preface

As an unexperienced researcher, I have been lucky and privileged to spend the last four years working with some world's experts in the field I am interested in. At the same time, my office has been only couple of meters away from one of the best libraries for mathematics and computer science on the globe. Officially, I have been a PhD student in the PNA-1 group at the Centrum voor Wiskunde en Informatica in Amsterdam.

I have been also lucky to have Monique Laurent as my supervisor. First of all, I am grateful to her for giving me the opportunity to work in a stimulating research environment. I appreciate very much her patience and the time spent on discussing all (ridiculous and reasonable) ideas I have had. Her guidance in the research process has been very important and has resulted in several joint papers and this thesis. Finally, I am indebted for her understanding and help in my extraordinary life circumstances.

I am beholden to professor Alexander Schrijver for offering me the possibility to defend my doctoral thesis at the University of Amsterdam.

Under the favour of professor Franz Rendl I have conducted two productive working visits at the University of Klagenfurt. I want to thank his former students Angelika Wiegele, Janez Povh and Igor Dukanovic for their hospitality, cooperation, and disinterested support when I needed tailor-made versions of their codes.

During four years in Amsterdam many positive characters entered my life. I am thankful to all of them and I apologise to those that are not mentioned below. I have learned a lot from the discussions and seminars I had with my colleagues Hartwig Bosse, Jarek Byrka, Rob Economopoulos, Dion Gijswijt, Erik Jan van Leeuwen, Gábor Maróti, Beniamin Mounits, Fernando Mario de Oliveira Filho, Steven Kelk, Frank Vallentin and many others. My flatmates, Leif Azzopardi, Matthieu Jonckheere, Eike Kiltz, Lasse Leskela, Krzysztof Pietrzak and Falk Unger have made my life, far from my family, easier, and are, indubitably, great companions (sorry Hardy, you are already mentioned above).

I am thankful to Dion for his assistance with 'Samenvatting' and to Rob for proof reading some of the sections of this thesis.

Special thanks go to my family, to Maja, Ilija and Bogdan, for their love, patience and support.

Amsterdam, January 2008.

# Contents

# Chapter 1

# Introduction

In this thesis, we consider two well known graph parameters: the stability number and the chromatic number of a graph. We study semidefinite relaxations of integer and copositive programs defining these parameters. We extensively use techniques, known in the literature as block diagonalization and symmetry reduction, for reducing the sizes of matrices and the number of variables in semidefinite programs.

We give first a brief sketch of the results presented in this work, and go back to details in Section 1.3 after introducing background in Section 1.2.

## 1.1 Overview of results

We compare several hierarchies of semidefinite programming upper bounds for the stability number of a graph. The first order bounds in all these hierarchies coincide with the Lovász theta number. Combining the approaches of Lovász and Schrijver [65], and Lasserre [56, 57], we define a new hierarchy. As a relaxation of the hierarchy of Lasserre, it has an advantage that the semidefinite programs defining its bounds can be block diagonalized. Moreover, it is less costly and at least as strong as the hierarchy of Lovász and Schrijver. Besides, we introduce the hierarchy of de Klerk and Pasechik [50] and show that it is dominated by the hierarchy of Lasserre.

We next define and study the corresponding hierarchies of lower bounds for the (fractional) chromatic number. We introduce a special operator $\Psi$ which maps upper bounds for the stability number to lower bounds for the chromatic number. As an application, we prove that there is no polynomial time computable graph parameter nested between the fractional chromatic and the chromatic number of a graph, unless P=NP.

We compute bounds in the new block diagonal hierarchy for some interesting graph classes. In particular, we are able to compute the bounds, up to order three, for Paley graphs with at most 800 vertices, using the properties of their automorphism groups; and the bounds, of order one and two, for Hamming and Kneser graphs with up to $2^{20}$ vertices, using the explicit block diagonalization of the Terwilliger algebra of the Hamming scheme given by Schrijver in [85]. Finally, we introduce yet another variation of the second order bound in the hierarchy of Lasserre via a semidefinite program which can be also block diagonalized, and we report computational results for some DIMACS benchmark instances.

## 1.2 Background and motivation

### Graphs, stable sets and colourings

A graph consists of vertices and edges. An edge connects two vertices. A stable set is a set of vertices of a graph in which no two of them are connected with an edge. An assignment of colours to the vertices of a graph, such that no two connected vertices share same colour, is called a vertex colouring. Stable sets and vertex colourings are closely related. It is straightforward to see that vertex colouring of a graph is equivalent to partitioning of the set of vertices into stable sets.

Some problems of practical interest can be modelled as stable set or colouring problems, e.g. time tabling, scheduling, frequency assignment, register allocation, pattern matching or coding. In these applications, one is usually interested in finding a maximum-size stable set in a graph or a vertex colouring of a graph which uses the least possible number of colours.

For instance, one of the fundamental problems in coding theory is finding a code, i.e. a subset of possible words which differ from each other significantly, of a maximum size. Here words are all sequences of letters, from a given alphabet $\mathcal{A}$, of some predefined length $n$. Their difference is usually quantified as the number of places in which they differ, called the Hamming distance. Thus, given a positive integer $d$, two words differ significantly if their Hamming distance is at least $d$. Consider the graph whose vertices are the words, i.e. the elements of $\mathcal{A}^n$, two of them being connected with an edge if their Hamming distance is smaller than $d$. Finding an optimal code is now equivalent to finding a maximum size stable set in this graph.

The stability number $\alpha(G)$ of a graph $G$ is the cardinality of a maximum size stable set in the graph. The chromatic number $\chi(G)$ of $G$ is the minimum number of colours that have to be used in a vertex colouring of the graph. Determining $\alpha(G)$ and $\chi(G)$ are hard combinatorial optimization problems.

### Complexity and combinatorial optimization

Optimization problems are problems in which one tries to find a best solution, satisfying certain properties, with respect to a given criterion. It is common to express the criterion of a problem as a function, called an objective function. The goal is then to find its optimal value (usually the maximum or the minimum) on a given domain, known as the feasibility domain or the set of feasible solutions. A combinatorial optimization problem is an optimization problem whose set of feasible solutions is finite. One of the most important issues when dealing with combinatorial optimization problems is their complexity.

A decision problem is a question whose answer depends on some input parameters and can be either 'yes' or 'no'. When we fix input parameters we get an instance of a decision problem. An oracle for a decision problem is a machine (black-box) which is able to solve it in a single operation. The complement of a decision problem is the decision problem resulting from reversing the 'yes' and 'no' answers.

If a 'yes' answer to a decision problem is provided with a certificate, which can be checked in polynomial time (in the size of the input), the problem is said to belong to the class NP ('Non-deterministic Polynomial time'). For example, given a graph $G$ and a positive integer $k$, consider the following two problems:

(S) Does there exist a stable set in $G$ of size at least $k$?

(C) Does there exist a vertex colouring in $G$ which uses at most $k$ colours?

The question (S), known also as the 'stable set problem', is in NP. Namely, a stable set of size at least $k$ is a certificate for a 'yes' answer since we can quickly check that it contains at least $k$ vertices and that no two of them are connected. Accordingly, a vertex colouring which uses at most $k$ colours is a certificate for a 'yes' answer of the problem (C) since we can quickly check that no two connected vertices received same colour. Hence the problem (C), known as the 'colouring problem', is also in NP.

A decision problem is said to belong to the class co-NP if its complement is in NP.

A problem is said to be in class P ('Polynomial time') if it can be solved, i.e. the correct answer can be found, in polynomial time. It is also common to say that such a problem is 'easy'. There exist decision problems in NP for which we still do not know if they are easy or not. In other words, we do not know if P=NP. This is considered to be the most important open question in complexity theory.

A decision problem $A$ can be reduced to a decision problem $B$ in polynomial time if there exists a polynomial time algorithm $f$ which transforms instances of $A$ into instances of $B$, such that for any instance $a$ of $A$ the answer to the instance $f(a)$ of $B$ is 'yes' if and only if the answer to the instance $a$ is 'yes'. As an example we give a transformation from the colouring problem to the stable set problem. Given a graph $G$ with $n$ vertices and a nonnegative integer $k$, make $k$ copies of $G$, for every vertex of $G$ connect all pairs of its copies and call the constructed graph $G_k$. Then, $G$ can be coloured with $k$ colours if and only if there exists a stable set in $G_k$ of size (at least) $n$ (see Section 2.5 for details).

If a problem is in NP, and every other problem from NP can be reduced to it in polynomial time, the problem is said to be NP-complete. A problem $A$ is said to be NP-hard if and only if there is an NP-complete problem $B$ that can be solved in polynomial time with an oracle for $A$.

The problem of determining if for a combinatorial optimization problem there exists a feasible solution, with the objective value of a given quality (usually greater or smaller than some prescribed threshold value), is the decision counterpart of the combinatorial optimization problem. Problem (S) is thus the decision counterpart of the problem of finding $\alpha(G)$, and problem (C) is the decision counterpart of the problem of finding $\chi(G)$. Combinatorial optimization problems whose decision counterparts are NP-complete are NP-hard, i.e., they are at least as hard as any problem in NP. The problems of determining the stability number and the chromatic number of a graph are both NP-hard (cf. [30]).

## Combinatorial optimization and semidefinite relaxations

We cannot expect to find a polynomial time algorithm for an NP-hard combinatorial optimization problem. Still, we can try to solve it approximately by considering some relaxation of it, for which an efficient algorithm exists. In this thesis, we consider two approaches for modelling combinatorial optimization problems and related semidefinite relaxations. The main motivation for using semidefinite relaxations is the fact that semidefinite programs can be solved in polynomial time (up to a certain precision, see Subsection 2.3.2 for more details).

**Semidefinite programs** are programs in which one aims to optimize a linear function over the intersection of an affine subspace and a cone of semidefinite matrices. It is common to write a semidefinite program as

$$\min \langle C, X\rangle \quad \text{subject to} \quad \begin{array}{l} \langle A_j, X\rangle = b_j \ (j = 1, \ldots, m), \\ \text{and } X \in \mathbb{R}^{n\times n} \text{ is positive semidefinite,} \end{array} \tag{1.1}$$

where $\langle \cdot, \cdot\rangle$ denotes the standard inner product on $\mathbb{R}^{n\times n}$.

The set of positive semidefinite matrices in $\mathbb{R}^{n\times n}$, known as the semidefinite cone, is convex. Semidefinite programs thus belong to the class of convex optimization problems, and moreover, they generalize linear, quadratic and second order cone programming problems. If we restrict, for example, the matrix $X$ in (1.1) to be diagonal, we obtain a linear program.

**Integer programming approach**. A classical approach is to model a combinatorial optimization problem as an integer linear program

$$\max\ c^T x \text{ subject to } Ax \leq b,\ x \in \{0, 1\}^n, \tag{1.2}$$

where $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. The hardness of the problem is then hidden in the constraint $x \in \{0, 1\}^n$. (In general, an integer linear program is NP-hard. See e.g. [30].)

The way to define a basic semidefinite relaxation for (1.2) is to introduce the matrix variable $X = \binom{1}{x}\binom{1}{x}^T$. We have that $x \in \{0, 1\}^n$ if and only if the first row of $X$ equals its diagonal. The next step is to express the objective function $c^T x$ and the linear constraints $Ax \leq b$, respectively, as a linear function and linear constraints in terms of the entries of the matrix $X$. (It can be done in several ways, see e.g. [61].) In other words, we can rewrite (1.2) in terms of $X$, where instead of the condition $x \in \{0, 1\}^n$ we require that $X$ is a symmetric, rank one matrix, whose first row equals its diagonal, and whose left upper corner equals one. In this way, the hardness of the problem is moved into the 'rank one' constraint. Finally, the basic semidefinite relaxation is then obtained by dropping the rank constraint, and by requiring positive semidefiniteness for $X$ instead. Note that, apart from the semidefinite constraint, all constraints are then linear.

The first such semidefinite relaxation was proposed by Lovász for the stable set problem in his seminal paper [64]. He introduced a parameter $\vartheta(G)$ of a graph $G$, nowadays called the Lovász theta number, nested between the stability number $\alpha(G)$ and the chromatic number $\chi\left(\overline{G}\right)$ of the complement $\overline{G}$ of $G$. We will study this bound in details in Chapter 3, and we will see in Chapters 4 and 5 how one can strengthen it to obtain stronger semidefinite bounds for $\alpha(G)$ and $\chi\left(\overline{G}\right)$.

Another breakthrough result, obtained by using basic semidefinite relaxations for combinatorial optimization problems, is the paper [34] by Goemans and Williamson. They presented an approximation algorithm for the max cut problem[1]. Given a graph instance, the algorithm finds a solution whose (expected) objective value is not less than 0.878 times the optimum.

**Copositive programming approach**. A more recent approach is based on modelling combinatorial optimization problems as copositive programs. In a copositive program the goal is to minimize a linear function with respect to linear constraints. The variable is a square matrix, restricted to be copositive. In this approach, the hardness of a problem is put into the copositivity constraint. (Testing if a given matrix is not copositive is an NP-complete problem. Cf. [30].)

The way to relax this hard condition is to replace the copositive cone by some tractable subcone of it. For example, replacement by the semidefinite cone would give a semidefinite program. In [75], Parrilo defines a hierarchy of tractable subcones of the copositive cone, where the first subcone is the sum of the semidefinite cone and the cone of symmetric nonnegative matrices. We will see in Section 4.2 the application of Parrilo's idea to the stability number, which is due to de Klerk and Pasechnik [50], and in Subsection 5.2.4 the application to the chromatic number, due to Dukanovic and Rendl [24].

## 1.3 Outline of the thesis and contributions

### Chapter 2: Notation and preliminaries

In Chapter 2 we recall basic linear algebra results, the general framework of conic programming, important facts about semidefinite programming, necessary definitions from graph theory, and some useful polynomial optimization tools.

The key ideas of this thesis introduced in Chapter 2 can be listed as follows:

- In Section 2.4 we focus on block diagonalization and symmetry reduction techniques. We prove Lemma 2.4.5 which enables us to block diagonalize the new hierarchy defined in Subsection 4.1.4. We recall from [85] an explicit block diagonalization of the Terwilliger algebra of the Hamming scheme.
- In Section 2.5 we give a reduction from the colouring problem to the stable set problem. It is the key observation for building the hierarchies of lower bounds for the chromatic number of a graph in Chapter 5.

[1] Given a graph and weights assigned to its edges, the max cut problem is the problem of splitting the vertex set of the graph into two sets such that the size of the cut, i.e. the sum of the weights of the edges connecting vertices from different sets, is maximized.

- In Section 2.6 we explain the sum of squares approach to polynomial optimization problems, that is used in Section 4.2 to construct a series of subcones of a copositive cone. We also recall the dual approach based on moment matrices, whose applications are considered in Subsections 4.1.3, 5.2.1 and 5.2.2.

## Chapter 3: The Lovász theta number

Throughout the whole thesis we work with hierarchies of semidefinite bounds. All these hierarchies have either the Lovász theta number $\vartheta(G)$ of a graph $G$, or some variation of it, as a starting point. Chapter 3 is devoted to this number.

Chapter 3 contains a tour through several semidefinite programming formulations for $\vartheta(G)$, the proof of 'the sandwich theorem', and the definitions of some variations of $\vartheta(G)$ obtained by adding nonnegativity and triangle constraints. We make a few small contributions that are not published, and which are new to the best of our knowledge:

- a simple proof, based on Lemma 3.1.1, of the equivalence between the standard definition of $\vartheta(G)$ and the definition related to the theta body, which is stated in Proposition 4.1.1;

- an observation that nonnegativity constraints do not improve $\vartheta(G)$ if $G$ contains some edge symmetry, given in Proposition 3.3.2;

- and an explanation of the phenomenon, appearing in computational results reported by Dukanovic and Rendl in [25], that triangle constraints do not improve $\vartheta(G)$ if $G$ is a Hamming graph, given in Proposition 3.3.3.

## Chapter 4: Semidefinite programming upper bounds for the stability number

This chapter deals with the hierarchies of semidefinite upper bounds for the stability number $\alpha(G)$ of a graph $G$.

We first recall in Section 4.1 the definitions of the stable set polytope $\mathrm{STAB}(G)$ and its well known relaxations. Among others, we describe the theta body $\mathrm{THETA}(G)$. We then introduce and compare three hierarchies of semidefinite relaxations of $\mathrm{STAB}(G)$, the matrix cut hierarchy of Lovász and Schrijver [65] (Subsection 4.1.2), the moment matrix hierarchy of Lasserre [57] (Subsection 4.1.3), and the new block diagonal hierarchy nested between the previous two (Subsection 4.1.4). They all start from $\mathrm{THETA}(G)$ and converge to $\mathrm{STAB}(G)$ in finitely many steps, for any fixed graph $G$.

The main contributions of Section 4.1 are as follows:

- Theorem 4.1.4 about the convergence in $\alpha(G) - 1$ steps of the hierarchy of Lovász and Schrijver [65] applied to the clique-constrained polytope;

- Subsection 4.1.4, which explains in detail the application of the hierarchy, proposed in the paper [40] by Gvozdenović, Laurent and Vallentin, to the stable set problem.

In Section 4.2 we introduce the hierarchy of de Klerk and Pasechnik [50] based on the sum of squares approach to copositive programs. We discuss some convergence properties of this hierarchy and we compare it with the hierarchy of Lasserre.

The contributions related to the hierarchy of de Klerk and Pasechnik are derived from the papers [38, 39] by Gvozdenović and Laurent (the paper [39] is the journal version of the paper [38]). The most important contributions are as follows:

- Theorem 4.2.13 that partially solves a conjecture of de Klerk and Pasechnik about the convergence of their hierarchy.
- Theorem 4.2.17 and Proposition 4.2.23 in which this hierarchy is compared with the hierarchy of Lasserre and the new block diagonal hierarchy.

### Chapter 5: Semidefinite programming lower bounds for the chromatic number

Although a vast literature exists about hierarchies of relaxations for the stability number $\alpha(G)$ of a graph $G$, to the best of our knowledge no such hierarchy for the chromatic number $\chi(G)$ had been studied before we started our research.

In Chapter 5 we essentially follow the work of Gvozdenović and Laurent [37]. We start with the Lovász theta number $\vartheta\left(\overline{G}\right)$ of the complement of a graph $G$, and try to strengthen it towards the fractional chromatic number $\chi^*(G)$ and the chromatic number $\chi(G)$ of $G$.

In Section 5.1 we apply the reduction from the colouring problem to the stable set problem, given in Section 2.5. We first introduce an operator $\Psi$. It is monotone nonincreasing and maps any graph parameter nested between $\alpha(\cdot)$ and $\overline{\chi}(\cdot)$ to a parameter lying between the clique number $\omega(\cdot)$ and $\chi(\cdot)$. Moreover, if a graph parameter is polynomial time computable, the same holds for its image under $\Psi$. As a direct consequence of the properties of $\Psi$, there is no polynomial time computable graph parameter nested between $\chi^*(\cdot)$ and $\chi(\cdot)$ unless P=NP. We conclude the section with quadratic and copositive programming formulations for $\chi(G)$.

In Section 5.2, we define and study hierarchies of lower bounds for $\chi^*(G)$ and $\chi(G)$, which are closely connected to the hierarchies presented in Chapter 4. In particular,

- we present the hierarchies based on the moment matrix approach of Lasserre [56, 57];
- we define, using the same framework, new hierarchies corresponding to the new block diagonal hierarchy.

Finally, we recall the hierarchy of Dukanovic and Rendl [24] for $\chi^*(G)$. Their hierarchy corresponds to the hierarchy of de Klerk and Pasechnik [50]. We observe that, for vertex transitive graphs, it is dominated by the moment matrix based hierarchy.

It should be mentioned that none of the hierarchies for $\chi^*(G)$ and $\chi(G)$ was known before we started to work on this topic.

## Chapter 6: Computational results

We show in Chapter 6 how to compute the semidefinite bounds studied in this thesis.

Section 6.1 contains results for Paley graphs from Gvozdenović, Laurent and Vallentin [40]. We compute the bounds on the stability number from the new block diagonal hierarchy, up to order three, for Paley graphs with at most 800 vertices. The properties of the automorphism groups of these graphs allow us to significantly reduce the number of variables and the number of blocks in the semidefinite programs that define these bounds.

In the remaining sections of Chapter 6 we follow the paper [36] by Gvozdenović and Laurent.

We consider lower bounds for the (fractional) chromatic numbers of Hamming and Kneser graphs in Sections 6.2 and 6.3, respectively. The bounds, of order one and two, from the new block diagonal hierarchies are computed for graphs with up to $2^{20}$ vertices. As the key ingredient, we use the explicit block diagonalization of the Terwilliger algebra of the Hamming scheme given by Schrijver in [85].

In Section 6.4, we introduce a new lower bound for the chromatic number of a graph. It is a variation of the second order bound in the Lasserre type hierarchy, suitable for nonsymmetric graphs. We report experimental results on some DIMACS benchmark instances. For several instances, our bounds improve the best known lower bounds.

# Chapter 2

# Notation and preliminaries

We summarize in this chapter the mathematical background used in this thesis in order to make it self-contained. We first introduce some notation and recall basic linear algebra results. Using these notions we describe the general framework of conic programming, and recapitulate the most important facts about semidefinite programming. The last but one section focuses on graphs. In particular, we consider stable sets and vertex colourings. Polynomial optimization tools, like sums of squares and moment matrices, are summed up in the last section. More detailed information about conic and semidefinite programming can be found in Ben-Tal and Nemirovski [4], Rockafellar [82], Helmberg [44], or de Klerk [49]. Laurent [58] gives a survey on polynomial optimization techniques based on moment matrices and sums of squares of polynomials. For a classic graph theory text we recommend Diestel [23], while Schrijver [84] contains relevant details about colourings and stable sets.

## 2.1 Sets, vectors and matrices

By $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{R}_+$, we denote, respectively, the sets of integers, nonnegative integers, complex numbers, real numbers and nonnegative real numbers. For $z \in \mathbb{C}$, $\overline{z}$ denotes its complex conjugate.

Given a finite set $V$, its size is denoted by $|V|$, whereas the collection of all its subsets is denoted by $\mathcal{P}(V)$.

For finite sets $V$ and $W$ and a field $R$ ($\mathbb{R}$ or $\mathbb{C}$) we consider the vector spaces $R^V$ and $R^{V\times W}$. Thus, the elements of $R^V$ are vectors indexed by $V$, and the elements of $R^{V\times W}$ are matrices with rows indexed by $V$ and columns indexed by $W$. For $V = \{1, 2, \ldots, n\}$ and $W = \{1, 2, \ldots, m\}$ we write $R^n$ instead of $R^V$, and $R^{n\times m}$ instead of $R^{V\times W}$. In general, we often identify $R^V$ with $R^{|V|}$, and $R^{V\times W}$ with $R^{|V|\times|W|}$. We sometimes consider vectors as $V \times W$ matrices with $|W| = 1$ and matrices as vectors in $R^V$ if $|W| = 1$. We also use the ordinary product of matrices of compatible dimensions.

Matrices and vectors can be indexed by subsets of $\mathcal{P}(V)$. In particular, given an integer $r$, we often use

$$\mathcal{P}_{\le r}(V) := \{I \in \mathcal{P}(V) \mid |I| \le r\} \text{ and } \mathcal{P}_{=r}(V) := \{I \in \mathcal{P}(V) \mid |I| = r\}.$$

Note that $\mathcal{P}_{\leq r}(V)$ contains the empty subset of $V$ which we will denote as $\mathbf{0}$; thus, for instance, $\mathcal{P}_{\leq 1}(V) = \{\mathbf{0}, \{i\}\ (i \in V)\}$. Sometimes we identify $\mathcal{P}_{=1}(V)$ with $V$, i.e., we write $i$ instead of $\{i\}$. Furthermore, we sometimes write $ij$ instead of $\{i, j\}$ and $ijk$ instead of $\{i, j, k\}$, etc. The standard unit vectors in $R^{\mathcal{P}_{\leq 1}(V)}$ we denote by $e_{\mathbf{0}}$, $e_i$, $i \in V$.

Throughout, the letters $\mathbf{I}$, $\mathbf{J}$ and $e$ denote, respectively, the *identity matrix*, the *all-ones matrix* and the *all-ones vector* (of suitable size). If $\mathbf{I}, \mathbf{J} \in R^{n \times n}$ we also write $\mathbf{I}_n$ instead of $\mathbf{I}$, and $\mathbf{J}_n$ instead of $\mathbf{J}$.

For a matrix $M$, we express its $i$th row, its $j$th column, and their intersection by $M_{i\bullet}$, $M_{\bullet j}$ and $M_{ij}$, respectively. The *transpose* of a matrix $M \in \mathbb{R}^{V \times W}$ (of a vector if $|W| = 1$) is the matrix $M^T \in \mathbb{R}^{W \times V}$ with $M^T_{ji} := M_{ij}$ for $i \in V$, $j \in W$. The *conjugate transpose* of a matrix $M \in \mathbb{C}^{V \times W}$ (of a vector if $|W| = 1$) is the matrix $M^* \in \mathbb{C}^{W \times V}$ with $M^*_{ji} := \overline{M}_{ij}$ for $i \in V$, $j \in W$.

For a matrix $M \in \mathbb{R}^{V \times V}$,

- $M$ is *nonsingular* if it has an inverse matrix $M^{-1} \in \mathbb{R}^{V \times V}$ such that $MM^{-1} = M^{-1}M = \mathbf{I}$;
- $M$ is *symmetric* if $M^T = M$;
- $M$ is *diagonal* if $M_{ij} = 0$ for all $i, j \in V$ such that $i \neq j$;
- $M$ is *orthogonal* if $M^T = M^{-1}$, i.e., $MM^T = M^TM = \mathbf{I}$;
- $\mathrm{diag}(M)$ is the vector $(M_{ii})_{i \in V} \in \mathbb{R}^V$;
- the *trace* of $M$ is $\mathrm{Tr}(M) := e^T\mathrm{diag}(M) = \sum_{i \in V} M_{ii}$;
- a nonzero vector $v \in \mathbb{R}^V$ is an *eigenvector* of $M$ with *eigenvalue* $\lambda$ if $Mv = \lambda v$.

For a matrix $M \in \mathbb{C}^{V \times V}$,

- $M$ is *Hermitian* if $M^* = M$;
- $M$ is *unitary* if $M^* = M^{-1}$, i.e., $MM^* = M^*M = \mathbf{I}$.

Given a vector $v \in \mathbb{R}^V$ and a set $S \subseteq V$,

- $\mathrm{Diag}(v)$ denotes the diagonal matrix $M \in \mathbb{R}^{V \times V}$ with $\mathrm{diag}(M) = v$;
- $\chi^S$ denotes the *characteristic vector* of $S$ in $\mathbb{R}^V$ defined by

$$\left(\chi^S\right)_i := \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{if } i \in V \backslash S; \end{cases}$$

- $v(S) := v^T\chi^S = \sum_{i \in S} v_i$.

For matrices $M \in R^{V \times W}$, $N \in R^{V' \times W'}$ their *tensor product* $M \otimes N$ is the matrix indexed by $(V \times V') \times (W \times W')$, with

$$(M \otimes N)_{(i,i'),(j,j')} := M_{i,j}N_{i',j'},$$

for $(i,i') \in V \times V'$, $(j,j') \in W \times W'$.

We also define the *(standard) inner product* of matrices $M, N \in \mathbb{R}^{V\times W}$ (of vectors if $|W| = 1$), denoted $\langle M, N\rangle$ as

$$\langle M, N\rangle := \mathrm{Tr}(M^T N) = \sum_{i\in V, j\in W} M_{ij}N_{ij}.$$

Observe that

$$\langle M, N\rangle = \langle N, M\rangle, \tag{2.1}$$

$$\langle M, N_1N_2\rangle = \langle MN_2^T, N_1\rangle = \langle N_1^T M, N_2\rangle \tag{2.2}$$

if $M, N, N_1, N_2$ are matrices with compatible dimensions.

For $M \in R^{V\times V}$ we have

- $\langle M, \mathbf{I}\rangle = \langle e, \mathrm{diag}(M)\rangle = \mathrm{Tr}(M) = \sum_{i\in V} M_{ii}$;
- $\langle M, \mathbf{J}\rangle = \sum_{i,j\in V} M_{ij}$.

## 2.2 Cones

Given a set $\mathcal{K} \subseteq \mathbb{R}^{V\times W}$, a matrix $M$ belongs to Int $\mathcal{K}$, the *interior* of $\mathcal{K}$, if there exists $\epsilon > 0$ such that $M + N \in \mathcal{K}$ for every $N \in \mathbb{R}^{V\times W}$ satisfying $\langle N, N\rangle \leq \epsilon$. The set $\mathcal{K}$ is *convex* if $\lambda M + (1-\lambda)N \in \mathcal{K}$, for all $M, N \in \mathcal{K}$ and all $\lambda \in (0,1)$. It is *closed* if $\mathrm{Int}(\mathbb{R}^{V\times W}\backslash\mathcal{K}) = \mathbb{R}^{V\times W}\backslash\mathcal{K}$.

A nonempty set $\mathcal{K} \subseteq \mathbb{R}^{V\times W}$ is a *cone* if it is closed under nonnegative scalar multiplication, i.e., $x \in \mathcal{K} \Rightarrow \lambda x \in \mathcal{K}$ for all $\lambda \in \mathbb{R}_+$. A cone $\mathcal{K} \subseteq \mathbb{R}^{V\times W}$ is

- *pointed* if $\mathcal{K} \cap -\mathcal{K} = \{0\}$, where $-\mathcal{K} := \{M \mid -M \in \mathcal{K}\}$;
- *solid* if Int $\mathcal{K} \neq \emptyset$;

The *dual cone* of a nonempty set $\mathcal{K} \subseteq \mathbb{R}^{V\times W}$ is the set

$$\mathcal{K}^* := \{M \mid \langle M, N\rangle \geq 0 \text{ for all } N \in \mathcal{K}\}. \tag{2.3}$$

The name 'dual cone' is motivated by the following result (cf. [4]):

**Theorem 2.2.1.** *Let $\mathcal{K} \subseteq \mathbb{R}^{V\times W}$ be a nonempty set and let $\mathcal{K}^*$ be as in (2.3). Then*

*(a) $\mathcal{K}^*$ is a closed convex cone;*

*(b) if* Int $\mathcal{K} \neq \emptyset$*, then $\mathcal{K}^*$ is pointed;*

*(c) if $\mathcal{K}$ is a closed convex pointed cone then* Int $\mathcal{K}^* \neq \emptyset$*;*

*(d) if $\mathcal{K}$ is a closed convex cone, then so is $\mathcal{K}^*$, and $(\mathcal{K}^*)^* = \mathcal{K}$.*

A cone $\mathcal{K}$ is *self-dual* if $\mathcal{K}^* = \mathcal{K}$. For example, one can easily see that the *nonnegative orthant* $\mathbb{R}^n_+$ is self-dual.

Throughout we mainly work with subcones of the cone of symmetric matrices $\mathcal{S}_n := \{M \in \mathbb{R}^{n\times n} \mid M^T = M\}$. This cone is not pointed but has numerous nice properties. Namely, any matrix $M \in \mathcal{S}_n$ can be decomposed as $M = U\Lambda U^T$ where $U \in \mathbb{R}^{n\times n}$ is an orthogonal matrix and $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix (the spectral decomposition theorem). Since $MU_{\bullet i} = \Lambda_{ii}U_{\bullet i}$, the diagonal entries of $\Lambda$ are the eigenvalues of $M$, and the columns of $U$ form a set of *orthonormal* eigenvectors ($U^T_{\bullet i}U_{\bullet i} = 1$, $U^T_{\bullet i}U_{\bullet j} = 0,\ i \neq j$).

We consider the following subcones of $\mathcal{S}_n$ that are closed, convex, pointed and solid:

- the cone of *(positive) semidefinite matrices*,

$$\mathcal{S}^+_n := \{M \in \mathcal{S}_n \mid v^T M v \geq 0 \text{ for all } v \in \mathbb{R}^n\};$$

- the cone of *copositive matrices* (or simply the copositive cone),

$$\mathcal{C}_n := \{M \in \mathcal{S}_n \mid v^T M v \geq 0 \text{ for all } v \in \mathbb{R}^n_+\};$$

- the cone of *nonnegative matrices*,

$$\mathcal{N}_n := \{M \in \mathcal{S}_n \mid M \geq 0\};$$

- the cone of *doubly nonnegative matrices*,

$$\mathcal{D}_n := \mathcal{S}^+_n \cap \mathcal{N}_n.$$

The cones $\mathcal{S}^+_n$ and $\mathcal{N}_n$ are self-dual. The dual of the copositive cone is the cone of *completely positive matrices*

$$\mathcal{C}^*_n = \{M \in \mathcal{S}_n \mid M = N^T N \text{ for some } N \in \mathbb{R}^{k\times n}_+ \text{ and } k \in \mathbb{N}\}.$$

Finally, one can easily prove that

$$\mathcal{D}^*_n = \mathcal{S}^+_n + \mathcal{N}_n := \{M + N \mid M \in \mathcal{S}_n,\ N \in \mathcal{N}_n\}.$$

We will also use the cones of sums of squares of polynomials and their duals defined in Section 2.6.

For two matrices $M, N \in \mathbb{R}^{m\times n}$, we write $M \geq N$ or $N \leq M$ if $M - N \in \mathbb{R}^{m\times n}_+$. Similarly, for two matrices $M, N \in \mathcal{S}_n$, we write (Löwner partial order) $M \succeq N$ or $N \preceq M$ or $M - N \succeq 0$ if $M - N \in \mathcal{S}^+_n$.

## 2.3 Linear conic programming

A convex programming problem consists of minimizing a convex objective function over a convex feasible set. It is known that every convex program can be restated as a problem of optimizing a linear objective function over an intersection of an affine space and a convex cone (see Remark 2.3.1 below). Problems of this type are called *linear conic programs*. Here we recall some basic definitions and facts about linear conic programs.

REMARK 2.3.1. *(see e.g. Povh [79]) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, and $S \subseteq \mathbb{R}^n$ be a convex set. The convex program $\min_{x\in S} f(x)$ can be rewritten as $\min_{(z,x)\in S'} z$, where $S' := \{(z,x) \mid f(x) \leq z,\ x \in S\} \subseteq \mathbb{R}^{n+1}$. The set $S'$ is convex and can be represented as the intersection of the convex cone $\{\lambda(1,z,x) \mid (z,x) \in S',\ \lambda \geq 0\} \subseteq \mathbb{R}^{n+2}$ and the affine space $\{u \in \mathbb{R}^{n+2} \mid u_1 = 1\}$.*

### 2.3.1 Primal and dual programs

In this thesis the ambient space is $\mathbb{R}^{V\times W}$, where $V$ and $W$ are some finite sets. An *affine space* is given by $\{X \in \mathbb{R}^{V\times W} \mid \mathcal{A}(X) = b\}$, where $\mathcal{A} : \mathbb{R}^{V\times W} \to \mathbb{R}^m$ is a *linear operator* defined by $\mathcal{A}(X)_j := \langle A_j, X\rangle$ for some matrices $A_j \in \mathbb{R}^{V\times W}$, $1 \leq j \leq m$. The *adjoint* operator of $\mathcal{A}$ is the operator $\mathcal{A}^T : \mathbb{R}^m \to \mathbb{R}^{V\times W}$ that satisfies $\langle \mathcal{A}(X), y\rangle = \langle X, \mathcal{A}^T(y)\rangle$, for all $X \in \mathbb{R}^{V\times W}$ and $y \in \mathbb{R}^m$. Since $\langle \mathcal{A}(X), y\rangle = \sum_{j=1}^m y_j\langle A_j, X\rangle = \langle \sum_{j=1}^m y_jA_j, X\rangle$ we get $\mathcal{A}^T(y) = \sum_{j=1}^m y_jA_j$.

Let $\mathcal{K} \subseteq \mathbb{R}^{V\times W}$ be a solid, pointed, closed and convex cone. The *primal linear conic optimization problem* is

$$\text{(P)} \qquad \begin{array}{rll} p^* := & \inf & \langle C, X\rangle \\ & \text{s.t.} & \mathcal{A}(X) = b \\ & & X \in \mathcal{K}, \end{array}$$

where $C \in \mathbb{R}^{V\times W}$, $b \in \mathbb{R}^m$, and $\mathcal{A} : \mathbb{R}^{V\times W} \to \mathbb{R}^m$ is a linear operator.

The Lagrange *dual* of (P) is the conic program

$$\text{(D)} \qquad \begin{array}{rll} d^* := & \sup & \langle b, y\rangle \\ & \text{s.t.} & \mathcal{A}^T(y) + Z = C \\ & & Z \in \mathcal{K}^*,\ y \in \mathbb{R}^m. \end{array}$$

We say that the primal program (P) is *feasible* (resp. *strictly feasible*) if there exists $X \in \mathcal{K}$ (resp. $X \in \text{Int } \mathcal{K}$) such that $\mathcal{A}(X) = b$. Similarly, the dual (D) is *feasible* (resp. *strictly feasible*) if there exists $Z \in \mathcal{K}^*$ (resp. $Z \in \text{Int } \mathcal{K}^*$) and $y \in \mathbb{R}^m$ such that $\mathcal{A}^T(y) + Z = C$. For a primal feasible solution $X$ and a dual feasible solution $(y, Z)$ the *duality gap* $\langle C, X\rangle - \langle b, y\rangle$ satisfies:

$$\langle C, X\rangle - \langle b, y\rangle = \langle C, X\rangle - \langle \mathcal{A}(X), y\rangle\langle C - \mathcal{A}^T(y), X\rangle = \langle Z, X\rangle \geq 0,$$

since $X \in \mathcal{K}$ and $Z \in \mathcal{K}^*$. Hence $\langle C, X\rangle \geq \langle b, y\rangle$, which is known as *weak duality*; moreover equality holds if and only if $\langle Z, X\rangle = 0$ in which case we say that *strong duality* holds.

A primal feasible solution $X^*$ is a *primal optimal* solution if $\langle C, X^*\rangle = p^*$. Similarly, a dual feasible solution $(y^*, Z^*)$ is a *dual optimal* solution if $\langle b, y^*\rangle = d^*$. Note that weak duality implies $p^* \geq d^*$. In general, with respect to this inequality several scenarios might occur (see [79] for the details). For example, both programs (P) and (D) might have optimal solutions such that $p^* > d^*$. On the other hand, it may happen that $p^* = d^*$ but an optimal solution for (P) (or for (D)) does not exist. Sufficient conditions for the equality $p^* = d^*$, i.e., for so called *perfect duality*, are given in the following theorem (see e.g. [4]).

**Theorem 2.3.2** (Conic duality theorem)**.** *Consider the programs (P) and (D).*

*a) If the program (P) is strictly feasible and $p^*$ is finite, then $p^* = d^*$, and there exists a dual optimal solution.*

*b) If the program (D) is strictly feasible and $d^*$ is finite, then $p^* = d^*$, and there exists a primal optimal solution.*

**Corollary 2.3.3.** *If (P) and (D) are both strictly feasible, then $p^* = d^*$ and there exist primal and dual optimal solutions.*

In the programs considered in this thesis primal and dual optimal solutions almost always exist, and moreover $p^* = d^*$. Thereafter, we often replace 'sup' by 'max' and 'inf' by 'min'.

### 2.3.2 Algorithms, complexity and practical efficiency

When the cone $\mathcal{K}$ in the program (P) (or (D)) equals $\mathbb{R}^n_+$, $\mathcal{S}^+_n$, $\mathcal{C}_n$ or $\mathcal{C}^*_n$, we call the corresponding conic program linear, semidefinite, copositive or completely positive, respectively. We briefly recall some known results about the complexity of these types of programs. We avoid technical details which are beyond the scope of this thesis.

**The ellipsoid method.** The history of solving convex conic programs goes back to late fourties, when Dantzig exploited the linear programming duality to design the simplex method. All proposed variations of the simplex method required an exponential number of iterations, and it took more than 30 years until Khachiyan [48] in 1979 presented the ellipsoid method for linear programs which runs in polynomial time.

The ellipsoid method is based on constructing a sequence of ellipsoids, all containing the set of all optimal solutions, whose volumes rapidly converge to zero. In each iteration one has to either show that the center of a considered ellipsoid is feasible, or to construct a hyperplane which separates the center from the feasible region. Provided with an answer, one can cut the ellipsoid into two halves, where one of the halves contains the set of all optimal solutions. This half can be now circumscribed by an ellipsoid whose volume is constant time less than the previous one. With this procedure one can produce a feasible solution at most $\epsilon$ far from an optimal solution, for any given precision $\epsilon$, after a polynomial number of iterations.

Roughly speaking, the ellipsoid method works in polynomial time if checking feasibility and constructing a separating hyperplane can be done in polynomial time (See Grötschel, Lovász and Schrijver [35] for details.). For example, testing if a given matrix is positive semidefinite, and providing a separating hyperplane if it is not, can be done in polynomial time by using Gaussian elimination. Hence, semidefinite programs can be solved (to any given precision) in polynomial time via the ellipsoid method (under certain assumptions about feasibility).

**Interior point methods.** Practical experiences with the ellipsoid method were disappointing. Fortunately in 1984, Karmarkar [47] proposed a polynomial time algorithm which became a basis for, what we nowadays call, interior point (or barrier) methods (see e.g. [91, 49]). In contrast to the ellipsoid method, the barrier methods can be efficiently used in practice. Nesterov and Nemirovski [74] showed that an arbitrary linear problem over a convex cone $\mathcal{K}$ can be solved with these methods in polynomially many iterations by using a self-concordant barrier function for $\mathcal{K}$. Self-concordant barrier functions are smooth on Int $\mathcal{K}$, and go to infinity as the boundary of the cone is approached. Moreover, they are convex and can be minimized efficiently by Newton's method.

The general idea in barrier methods is to leave out the conic constraint, e.g. in the formulation (P), to add the barrier term $f(X)$ to the objective, and try to solve the sequence of problems of the form

$$\inf_{X \in \mathrm{Int}\mathcal{K}} \langle C, X\rangle + \mu f(X) \quad \text{s.t.} \quad \mathcal{A}(X) = b, \tag{2.4}$$

where the parameter $\mu$ is sequentially decreased to zero. Roughly speaking, the strategy is to follow the curve generated by the optimal solutions $X_\mu$ ($\mu \in (0, \infty)$) of the problem (2.4), known also as the *central path*. It converges to the analytic center of the set of all optimal solutions of (P), when $\mu$ goes to zero. The sufficient prerequisite for using these methods is having a computable (up to a certain precision) barrier function. For example, the function $f(X) = -\log(\det(X))$ is such a barrier for the cone $\mathcal{S}_n^+$. Barrier methods can thus be used for solving semidefinite programs.

Interior point methods are the most used algorithms for solving semidefinite programs in practice due to their fast convergence towards an optimal solution. Still, the largest problems that can be solved with these methods involve matrices of a size of the order of a thousand rows, and a thousand constraints. As an alternative some cheaper algorithms were designed, implemented and tested, e.g., bundle methods [45] or first order methods [14, 80]. In practice, these algorithms are sometimes able to tackle and digest much larger problems than interior point algorithms. Their main drawbacks are slow convergence and weaker precision.

**Copositive programs.** While linear and semidefinite programs can be solved efficiently (to any given precision), the optimization over a copositive cone and its dual cone is hard. Namely, some NP hard problems (see e.g. Section 4.2) can be modelled as copositive (completely positive) programs, yielding that optimization over $\mathcal{C}_n$ ($\mathcal{C}_n^*$) is not tractable unless P=NP. More precisely, testing whether a given matrix is not copositive is NP-complete [72], hence unless co-NP=NP, we can not have a polynomial time certificate for copositivity.

Although we do not have an efficient algorithm for solving copositive programs, we will present, in Section 4.2, a strategy which will help us to solve them approximately.

## 2.4 Block diagonal semidefinite programs

Here we focus on semidefinite programs, the main tool for our work. Throughout, we often use the abbreviations PSD for 'positive semidefinite matrix', and

SDP for 'semidefinite program'. We first recall some well known facts about positive semidefinite matrices and then derive several results which are extensively used in this thesis. In particular, we explain the 'block diagonalization' concept and show how to exploit group symmetry. The main idea is to transform an SDP into a simpler problem with a block diagonal structure. It will play a crucial role in our work. Most of the SDP solvers available today support such a structure. In fact, they are much more efficient if fed with an SDP in a block diagonal form.

**Characterizations.** Let $M \in \mathcal{S}_n$. Positive semidefiniteness of $M$ can be characterized in several ways. The following definitions are equivalent:

- $v^T M v \geq 0$ for all $v \in \mathbb{R}^n$;
- all eigenvalues of $M$ are nonnegative;
- $M = LL^T$ for some $L \in \mathbb{R}^{n\times n}$.

As an immediate consequence of the first item above we have

$$M \succeq 0 \Longrightarrow NMN^T \succeq 0 \tag{2.5}$$

for all $N$ with appropriate dimensions, and moreover

$$M \succeq 0 \Longleftrightarrow NMN^T \succeq 0 \tag{2.6}$$

for all nonsingular $N \in \mathbb{R}^{n\times n}$. The next two observations will be used often, and can be easily derived from (2.5):

- if $M$ is PSD then every principal submatrix[1] of $M$ is PSD;
- if $M$ has two identical rows, $M$ is PSD if and only if its principal submatrix, obtained by deleting one of the two identical rows and the corresponding column, is PSD.

If the matrix $N$ in (2.6) is diagonal with $v := \mathrm{diag}(N)$, then $NMN^T$ is the matrix obtained from $M$ by multiplying its $i$th row and its $i$th column by $v_i$ $(i = 1, \ldots, n)$.

A matrix $M$ is positive definite if it is PSD and nonsingular, or equivalently

- $v^T M v > 0$ for all $v \in \mathbb{R}^n \backslash \{0\}$;
- all eigenvalues of $M$ are positive;
- $M = L^T L$ for some nonsingular $L \in \mathbb{R}^{n\times n}$.

It is known that the set of positive definite matrices $\mathcal{S}_n^{++}$ is the interior of the cone of positive semidefinite matrices, i.e., Int $\mathcal{S}_n^{+} = \mathcal{S}_n^{++}$.

**Primal dual pair.** Following the notation from Subsection 2.3.1, given a finite set $V$, the index set of matrices, and $m$, the number of primal linear constraints, a general primal dual pair of SDPs can be written as

$$\text{(PSDP)} \qquad \begin{array}{ll} \inf & \langle C, X\rangle \\ \text{s.t.} & \mathcal{A}(X) = b \\ & X \succeq 0, \end{array}$$

[1] A principal submatrix of a square matrix $M$ is a matrix obtained by deleting the rows and columns of $M$ indexed by a proper subset of the index set of $M$.

$$\text{(DSDP)} \qquad \begin{array}{ll} \sup & \langle b, y\rangle \\ \text{s.t.} & C - \mathcal{A}^T(y) \succeq 0, \end{array}$$

where $C \in \mathbb{R}^{V\times V}$, $b \in \mathbb{R}^m$, and $\mathcal{A} : \mathbb{R}^{V\times V} \to \mathbb{R}^m$ is a linear operator with $\mathcal{A}(X)_j = \langle A_j, X\rangle$ $(j = 1, 2, \dots, m)$.

We can assume, w.l.o.g., that the matrices $C$, $A_j$ $(j = 1, 2, \dots, m)$ are symmetric. Note that the number of variables is $\binom{|V|+1}{2}$ in (PSDP) and $m$ in (DSDP).

### 2.4.1 Block diagonalization

The matrices in this thesis often have a block structure. Namely if $M \in \mathbb{R}^{V\times V}$, and $\{V_p \mid p = 1, ..., k\}$ is a partition[2] of $V$, we can consider $M$ as a *block matrix* $[M(p, q)]_{1\le p,q\le k}$, where $M(p, q) \in \mathbb{R}^{V_p\times V_q}$ is the submatrix of $M$ with row indices in $V_p$ and column indices in $V_q$. We call such a matrix *block diagonal* if $M(p, q) = 0$ for all $p, q$ such that $p \ne q$. Note that $\lambda$ is an eigenvalue of a block diagonal matrix $M$ if and only if it is an eigenvalue of one of its diagonal blocks. As a consequence of this fact we have the following:

**Lemma 2.4.1.** *If $M = [M(p, q)]_{1\le p,q\le k}$ is block diagonal, then*

$$M \succeq 0 \Longleftrightarrow M(p, p) \succeq 0 \text{ for all } p = 1, ..., k. \tag{2.7}$$

While operating with block matrices the Schur complement of a matrix can be very helpful as well.

**Lemma 2.4.2** (Schur complement). *For any $A \in \mathcal{S}_n^{++}$, $C \in \mathcal{S}_m$ and $B \in \mathbb{R}^{n\times m}$ the following holds*

$$M = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0 \Longleftrightarrow C - B^T A^{-1} B \succeq 0. \tag{2.8}$$

*Proof.* By setting $N := \begin{pmatrix} \mathbf{I} & 0 \\ -B^T A^{-1} & \mathbf{I} \end{pmatrix}$ and using (2.6) we have

$$M \succeq 0 \Longleftrightarrow NMN^T = \begin{pmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{pmatrix} \succeq 0. \tag{2.9}$$

The result now follows directly from Lemma 2.4.1. □

**Orthogonal transformation and equivalence.** We can sometimes 'block diagonalize' a program given in (PSDP) form. With this we mean that it can be transformed into an SDP which involves only block diagonal matrices. The basis of all such transformations is the following observation.

**Lemma 2.4.3.** *If $U \in \mathbb{R}^{V\times V}$ is orthogonal then*

$$\langle UMU^T, UNU^T\rangle = \langle M, N\rangle \tag{2.10}$$

*for every $M, N \in \mathbb{R}^{V\times V}$.*

*Proof.* Directly from (2.2). □

[2] $\{V_p \mid p = 1, ..., k\}$ is a partition of $V$ if $V_1, V_2, \dots, V_k$ are pairwise disjoint and $\cup_{p=1}^k V_p = V$.

Assume that $U$ is an orthogonal matrix which 'block diagonalizes' the matrices $C$ and $A_j$ $(j = 1, \dots, m)$ from (PSDP) and (DSDP) simultaneously, i.e., $UCU^T$ and $UA_jU^T$ $(j = 1, \dots, m)$ are all block diagonal with respect to a given partition $\{V_p \mid p = 1, ..., k\}$ of $V$. Then we can define the following primal dual pair of SDPs:

$$\text{(PSDP')} \qquad \begin{array}{ll} \inf & \sum_{p=1}^{k} \langle C_p, X_p \rangle \\ \text{s.t.} & \sum_{p=1}^{k} \mathcal{A}_p(X_p) = b \\ & X_p \in \mathcal{S}^+_{|V_p|}, \ p = 1, \dots, k, \end{array}$$

$$\text{(DSDP')} \qquad \begin{array}{ll} \sup & \langle b, y \rangle \\ \text{s.t.} & C_p - \mathcal{A}_p^T(y) \in \mathcal{S}^+_{|V_p|}, \ p = 1, \dots k, \end{array}$$

where

- $C_p \in \mathbb{R}^{V_p \times V_p}$ is the $p$th diagonal block of $UCU^T$, i.e., $C_p := [UCU^T](p,p)$, and
- $\mathcal{A}_p : \mathbb{R}^{V_p \times V_p} \to \mathbb{R}^m$ is given by $\mathcal{A}_p(X)_j := \langle [UA_jU^T](p,p), X \rangle$, for $j = 1, 2, \dots, m$.

Observe that (PSDP) and (PSDP') are equivalent. Namely, given a (strictly) feasible solution $X$ of (PSDP), the diagonal blocks $X_p := [UXU^T](p,p)$ $(p = 1, \dots, k)$ of $UXU^T$ are (strictly) feasible for (PSDP'). Moreover, from (2.10) we get $\langle C, X \rangle = \sum_{j=1}^{k} \langle C_p, X_p \rangle$. On the other hand, given a block diagonal matrix $X'$ whose diagonal blocks $X_p$ $(p = 1, \dots, k)$ are (strictly) feasible for (PSDP'), the corresponding (strictly) feasible solution for (PSDP) is $U^T X' U$. Analogously, (DSDP) and (DSDP') are equivalent.

In practice, we usually gain in efficiency if we transform a problem, given in the (PSDP) form, into the (PSDP') form (or from (DSDP) into (DSDP')). Indeed, some problems which are too big for nowadays solvers can be solved after being transformed into a block diagonal form. It can be done, e.g. when a program is invariant under a group action, as we will explain in the next subsection. (See also Sections 6.2 and 6.3.) Before that, we illustrate the 'block diagonalization' procedure by proving two useful results.

**Symmetric PSD example.** The following lemma plays an important role in Chapter 5, where we deal with lower bounds for the chromatic number of a graph. It enables us to explore the symmetry in SDPs defining those bounds. See Sections 5.1.2 and 5.2.3 for the details.

**Lemma 2.4.4.** *Let $X \in \mathbb{R}^{nt}$ be a $t \times t$ block matrix, having an $n \times n$ matrix $A$ as its diagonal blocks, and an $n \times n$ matrix $B$ as nondiagonal blocks, i.e.*

$$X = \underbrace{\begin{pmatrix} A & B & \dots & B \\ B & A & \dots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \dots & A \end{pmatrix}}_{t \text{ blocks}}. \tag{2.11}$$

*Then, $X \succeq 0 \Longleftrightarrow A - B \succeq 0$ and $A + (t-1)B \succeq 0$.*

*Proof.* We define a $t \times t$ block matrix $U_t$ having the same block structure as the matrix $X$. For $p, q = 1, \dots, t$, let $U_t(p,q)$ denote the $(p,q)$th block of $U_t$, defined by

$$U_t(p,q) := \begin{cases} \frac{1}{\sqrt{t}}\mathbf{I}_n & \text{if } p = 1 \text{ or } q = 1, \\ \left(\frac{1}{\sqrt{t}+t} - 1\right)\mathbf{I}_n & \text{if } p = q \geq 2, \\ \frac{1}{\sqrt{t}+t}\mathbf{I}_n & \text{otherwise.} \end{cases} \tag{2.12}$$

Notice that $U_t$ is symmetric and orthogonal. Let $Y := (U_t)^T X U_t$. Then, $Y \succeq 0$ if and only if $X \succeq 0$ and a simple calculation gives

$$Y = \begin{pmatrix} A + (t-1)B & 0 & \dots & 0 \\ 0 & A - B & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A - B \end{pmatrix}, \tag{2.13}$$

which shows the lemma. □

**Zeta matrix.** In the proof of the next result we use tools introduced in [59, 65]. In particular, given a finite set $T$, $|T| = t$, we use the *zeta matrix* $Z$ defined in the following way:

- $Z \in \mathbb{R}^{n2^t \times n2^t}$ is the $\mathcal{P}(T) \times \mathcal{P}(T)$ block matrix with $Z(S,S') := \mathbf{I}_n$ if $S \subseteq S'$ and $Z(S,S') := 0$ otherwise, for $S, S' \subseteq T$, and
- its inverse[3] $Z^{-1}$ is the $\mathcal{P}(T) \times \mathcal{P}(T)$ block matrix with $Z^{-1}(S,S') = (-1)^{|S' \setminus S|}\mathbf{I}_n$ if $S \subseteq S'$ and $Z^{-1}(S,S') = 0$ otherwise, for $S, S' \subseteq T$.

The fact that $Z^{-1}$ is the inverse matrix of $Z$ follows from

$$\begin{aligned} [ZZ^{-1}](S,S') &= \sum_{S'' \subseteq T} Z(S,S'')Z^{-1}(S'',S') = \sum_{S \subseteq S'' \subseteq S'} Z(S,S'')Z^{-1}(S'',S') \\ &= \sum_{S \subseteq S'' \subseteq S'} (-1)^{|S'' \setminus S|}\mathbf{I}_n = \sum_{S_1 \subseteq S' \setminus S} (-1)^{|S_1|}\mathbf{I}_n, \end{aligned}$$

which is equal to $\mathbf{I}_n$ if $S = S'$ and to 0 otherwise. Here $[ZZ^{-1}](S,S')$ denotes the $(S,S')$th block of the matrix $ZZ^{-1}$.

We use the zeta matrix and its inverse to block diagonalize moment matrices. In particular, we use them in Section 2.6 to prove the finite convergence of a sequence of approximations for 0/1 polynomial programs, and in Section 4.1.4 to block diagonalize SDPs defining a new hierarchy of bounds for the stable set problem. Both applications are based on the following result.

**Lemma 2.4.5.** *Let $T$, $|T| = t$, be a finite set and let $A_S \in \mathbb{R}^{n \times n}$ ($S \in \mathcal{P}(T)$), be matrices. Let $M \in \mathbb{R}^{n2^t \times n2^t}$ be the $\mathcal{P}(T) \times \mathcal{P}(T)$ block matrix defined by $M(S,S') := A_{S \cup S'}$ (for $S, S' \subseteq T$). Then,*

$$M \succeq 0 \iff \sum_{S \subseteq S' \subseteq T} (-1)^{|S' \setminus S|} A_{S'} \succeq 0 \textit{ for all } S \subseteq T.$$

[3] It is called the Möbius matrix. See [65] and the references therein.

*Proof.* Define the $\mathcal{P}(T) \times \mathcal{P}(T)$ block diagonal matrix $D$ with diagonal blocks

$$D(S,S) := \sum_{S \subseteq S' \subseteq T} (-1)^{|S' \setminus S|} A_{S'} \text{ (for } S \subseteq T).$$

Then $D = Z^{-1} M (Z^{-1})^T$. In order to prove this, consider an $(S,S')$th block of $Z^{-1} M (Z^{-1})^T$. It reads

$$\begin{aligned}
\left[Z^{-1} M (Z^{-1})^T\right](S,S') &= \sum_{S_2 \subseteq T} \left( \sum_{S_1 \subseteq T} Z^{-1}(S,S_1) M(S_1,S_2) \right) Z^{-1}(S',S_2) \\
&= \sum_{S' \subseteq S_2 \subseteq T} (-1)^{|S_2 \setminus S'|} \left( \sum_{S \subseteq S_1 \subseteq T} (-1)^{|S_1 \setminus S|} A_{S_1 \cup S_2} \right). \qquad (2.14)
\end{aligned}$$

Given a set $S_3 \subseteq T$, such that $S_3 \supseteq S \cup S_2$, consider the coefficient in front of $A_{S_3}$ in the sum $\sum_{S \subseteq S_1 \subseteq T} (-1)^{|S_1 \setminus S|} A_{S_1 \cup S_2}$. It reads

$$\sum_{(S_3 \setminus S_2) \cup S \subseteq S_1 \subseteq S_3} (-1)^{|S_1 \setminus S|} = \sum_{S_2' \subseteq (S_2 \setminus S)} (-1)^{|S_2'|}, \qquad (2.15)$$

since $[(S_3 \setminus S_2) \cup S] \cup S_2 = S_3$ and $S_3 \setminus [(S_3 \setminus S_2) \cup S] = S_2 \setminus S$. If $S_2 \setminus S \neq \emptyset$ the value in (2.15) equals zero. Consequently, if $S' \setminus S \neq \emptyset$, all coefficients in (2.14) are equal to zero. This proves that $[Z^{-1} M (Z^{-1})^T](S,S') = 0$ if $S \neq S'$, taking into account the symmetry of (2.14) with respect to $S$ and $S'$.

It is now straightforward to verify that

$$\left[Z^{-1} M (Z^{-1})^T\right](S,S) = \sum_{S \subseteq S' \subseteq T} (-1)^{|S' \setminus S|} A_{S'} = D(S,S) \text{ for all } S \subseteq T.$$

Therefore, $M \succeq 0 \iff D \succeq 0 \iff D(S,S) \succeq 0$ for all $S \subseteq T$, which gives the result. ☐

The following consequence, which is simply the last result when $n = 1$, was presented in [59].

**Corollary 2.4.6.** *Let $T$, $|T| = t$, be a finite set and let $a_S \in \mathbb{R}$ ($S \in \mathcal{P}(T)$) be scalars. Let $M \in \mathbb{R}^{\mathcal{P}(T) \times \mathcal{P}(T)}$ be the matrix defined by $M_{S,S'} := a_{S \cup S'}$ (for $S, S' \subseteq T$). Then,*

$$M \succeq 0 \iff \sum_{S \subseteq S' \subseteq T} (-1)^{|S' \setminus S|} a_{S'} \geq 0 \textit{ for all } S \subseteq T.$$

### 2.4.2 Exploiting group symmetry in semidefinite programs

In Lemma 2.4.4 we considered a matrix of the form (2.11) which is obviously invariant under some permutations of rows and columns. We have seen that testing if such a matrix is PSD can be simplified to checking if two matrices, with considerably smaller dimensions, are PSD. We now put this idea into a more general framework and show how to simplify an SDP invariant under a group action[4].

It has been shown recently that the size of an invariant SDP can be reduced using the regular $*$-representation of the algebra of invariant matrices (see de Klerk, Pasechnik and Schrijver [52]), or by finding an irreducible matrix representation of a group (see Gaterman and Parrilo [31], Vallentin [92], Schrijver [85]). We focus here on the second approach and its application to the Terwilliger algebra of a Hamming scheme presented in [85]. (See also Gijswijt [32] and the references therein.) It will be extensively used in Sections 6.2 and 6.3. Namely, it will enable us to compute bounds for the chromatic numbers of Hamming and Kneser graphs with as many as $2^{20}$ vertices.

**Action of a group.** Let $V$ be a finite set and $\mathcal{G}$ be a subgroup of $\mathrm{Sym}(V)$, the group of permutations of $V$ (also denoted as $\mathrm{Sym}(n)$ if $|V| = n$). The group $\mathcal{G}$ acts on vectors and matrices indexed by $V$ by letting $\sigma(x) := (x_{\sigma(i)})_{i\in V}$ and $\sigma(M) := (M_{\sigma(i),\sigma(j)})_{i,j\in V}$ for $\sigma \in \mathcal{G}$, $x \in \mathbb{C}^V$, $M \in \mathbb{C}^{V\times V}$. In other words, for a given $\sigma \in \mathcal{G}$

$$\sigma(x) = P_\sigma x \quad\text{and}\quad \sigma(M) = P_\sigma^T M P_\sigma,$$

where $P_\sigma \in \mathbb{C}^{V\times V}$ is the *permutation* matrix defined by

$$(P_\sigma)_{ij} := \begin{cases} 1 & \text{if } \sigma(i) = j, \\ 0 & \text{otherwise.} \end{cases}$$

The permutation matrices are orthogonal, i.e. $P_\sigma P_\sigma^T = \mathbf{I}$ ($\sigma \in \mathcal{G}$). Hence, for every two matrices $M$ and $N$ and for all $\sigma \in \mathcal{G}$, we have

$$M \succeq 0 \Longleftrightarrow \sigma(M) \succeq 0 \quad\text{and}\quad \langle M, N\rangle = \langle \sigma(M), \sigma(N)\rangle.$$

The group $\mathcal{G}$ decomposes the set $V \times V$ into orbits, i.e. into the nonempty sets $R_r$ ($r = 1, \ldots, L$) satisfying

$$(i, j) \in R_r \Longleftrightarrow R_r = \{(\sigma(i), \sigma(j)) \mid \sigma \in \mathcal{G}\},$$

for all $i, j \in V$ and $r = 1, \ldots, L$. We also define the matrices $D_r \in \mathbb{C}^{V\times V}$ ($r = 1, \ldots, L$) by

$$(D_r)_{ij} := \begin{cases} 1 & \text{if } (i, j) \in R_r \\ 0 & \text{otherwise.} \end{cases}$$

A matrix $M$ is invariant under action of $\mathcal{G}$ if $\sigma(M) = M$ for all $\sigma \in \mathcal{G}$, or equivalently, if $M = \sum_{r=1}^{L} \lambda_r D_r$ for some $\lambda \in \mathbb{C}^L$. Particularly, the matrix $\frac{1}{|\mathcal{G}|}\sum_{\sigma\in\mathcal{G}} \sigma(M)$, the 'symmetrization' of $M$ obtained by applying the Reynolds operator, is invariant under action of $\mathcal{G}$. Due to the convexity of the cone of semidefinite matrices, we have:

$$M \succeq 0 \Longrightarrow \frac{1}{|\mathcal{G}|}\sum_{\sigma\in\mathcal{G}} \sigma(M) \succeq 0. \tag{2.16}$$

**Reducing the number of variables in an invariant SDP.** Consider now an SDP given in (PSDP) form, i.e.

$$\inf\ \langle C, X\rangle \quad \text{s.t. } \langle A_j, X\rangle = b_j \ (j = 1, \ldots, m),\ X \succeq 0, \tag{2.17}$$

[4]The process of simplifying an invariant SDP is also known as 'symmetry reduction'.

where $C$, $A_j$ $(j = 1, \dots, m)$ are matrices indexed by $V$. The program (2.17) is *invariant under action* of $\mathcal{G}$ if for every feasible solution $X$ and for every $\sigma \in \mathcal{G}$ the matrix $\sigma(X)$ is feasible and $\langle C, X\rangle = \langle C, \sigma(X)\rangle$. Assume now that (2.17) is invariant under action of $\mathcal{G}$, and let $X$ be a feasible solution of (2.17). The matrix $\frac{1}{|\mathcal{G}|}\sum_{\sigma\in\mathcal{G}} \sigma(X)$ is then also feasible for (2.17), due to (2.16). Moreover, $\langle C, X\rangle = \langle C, \frac{1}{|\mathcal{G}|}\sum_{\sigma\in\mathcal{G}} \sigma(X)\rangle$. Therefore, we can restrict $X$ in (2.17) to be invariant under action $\mathcal{G}$, i.e. we can assume that $X = \sum_{r=1}^{L} x_r D_r$ for some $x \in \mathbb{R}^L$. The program (2.17) is thus equivalent to

$$\begin{aligned} \inf\ \textstyle\sum_{r=1}^{L} c_r x_r \quad \text{s.t.} \quad & \textstyle\sum_{r=1}^{L} a_{jr} x_r = b_j\ (j = 1, \dots, m),\ x_r = x_s\ (\text{if } D_r^T = D_s), \\ & \textstyle\sum_{r=1}^{L} x_r D_r \succeq 0, \end{aligned} \tag{2.18}$$

where $c_r := \langle C, D_r\rangle$ and $a_{jr} := \langle A_j, D_r\rangle$ $(r = 1, \dots, L;\ j = 1, \dots, m)$.

We should point out that the transformation of the program (2.17) to the form (2.18) changes, and often considerably reduces, the number of variables from $\binom{|V|}{2}$ to $L$, the number of orbits of $V \times V$ under action of $\mathcal{G}$. We show next how to reduce the sizes of matrices in an invariant SDP.

**Matrix $*$-algebras.** A nonempty set of matrices in $\mathbb{C}^{V\times V}$ is called a *matrix $*$-algebra* if it is closed under addition, scalar multiplication, matrix multiplication and under taking the conjugate transpose. In particular, the set of matrices in $\mathbb{C}^{V\times V}$ invariant under action of $\mathcal{G}$,

$$\mathcal{A} := \{M \in \mathbb{C}^{V\times V} \mid \sigma(M) = M \text{ for all } \sigma \in \mathcal{G}\}, \tag{2.19}$$

is a matrix $*$-algebra. It contains the identity matrix and it is spanned by the matrices $D_r$ $(r = 1, \dots, L)$. Hence, the dimension of the algebra $\mathcal{A}$ is simply the number of orbits of $V \times V$. Every algebra of this type is isomorphic to a certain block diagonal algebra. It is a consequence of the following theorem.

**Theorem 2.4.7.** *[2] If $\mathcal{A} \subseteq \mathbb{C}^{V\times V}$ is a matrix $*$-algebra that contains the identity matrix, then there exists a unitary matrix $U \in \mathbb{C}^{V\times V}$, and positive integers $p_0, p_1, \dots, p_t$ and $q_0, q_1, \dots, q_t$ such that*

$$U^*\mathcal{A}U := \{U^*AU \mid A \in \mathcal{A}\}$$

*consists of all block diagonal matrices*

$$\begin{pmatrix} C_0 & 0 & \dots & 0 \\ 0 & C_1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & C_t \end{pmatrix}$$

*where each $C_k$ $(k = 0, 1, \dots, t)$ is a block diagonal matrix*

$$\begin{pmatrix} B_k & 0 & \dots & 0 \\ 0 & B_k & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & B_k \end{pmatrix}$$

*with $q_k$ identical blocks $B_k \in \mathbb{C}^{p_k\times p_k}$ on the diagonal.*

Obviously $\sum_{k=0}^{t} p_k q_k = |V|$, $L = \sum_{k=0}^{t} p_k^2$, and $\mathcal{A}$ is isomorphic to

$$\bigoplus_{k=0}^{t} \mathbb{C}^{p_k \times p_k} := \left\{ \begin{pmatrix} B_0 & 0 & \dots & 0 \\ 0 & B_1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & B_t \end{pmatrix} \;\middle|\; \begin{array}{c} B_k \in \mathbb{C}^{p_k \times p_k} \\ \\ \text{for } k = 0, 1, \dots, t \end{array} \right\}.$$

In this thesis we are focused on real matrices. In particular, while working with an algebra of type (2.19), we consider only its real counterpart, i.e. its subalgebra

$$\mathcal{B} := \{M \in \mathbb{R}^{V \times V} \mid \sigma(M) = M \text{ for all } \sigma \in \mathcal{G}\}, \tag{2.20}$$

and we use the following observation:

*If the matrix $U$, which block diagonalizes $\mathcal{A}$ as in Theorem 2.4.7, is orthogonal, then it also block diagonalizes $\mathcal{B}$, and moreover $U^T \mathcal{B} U$ is isomorphic to $\bigoplus_{k=0}^{t} \mathbb{R}^{p_k \times p_k}$.*

**Block diagonalization.** Let now $U$ be an orthogonal matrix which block diagonalizes $\mathcal{A}$ (and consequently $\mathcal{B}$) as in Theorem 2.4.7, and let $B_k(M)$ ($k = 0, ..., t$) denote the corresponding $B_k$-th block the a matrix $U^T M U$, where $M \in \mathcal{A}$. Since for every $M \in \mathcal{B}$,

$$M \succeq 0 \Longleftrightarrow B_k(M) \succeq 0 \text{ for all } k = 0, 1, \dots, t, \tag{2.21}$$

the program (2.18) reduces to

$$\begin{gathered} \inf \; \sum_{r=1}^{L} c_r x_r \quad \text{s.t.} \quad \sum_{r=1}^{L} a_{jr} x_r = b_j \; (j = 1, \dots, m), \; x_r = x_s \text{ (if } D_r^T = D_s), \\ B_k \left( \sum_{r=1}^{L} x_r D_r \right) \succeq 0 \; (k = 0, \dots, t). \end{gathered} \tag{2.22}$$

Sometimes it can be more convenient to work on the dual side. Namely, consider the dual program of (2.17), i.e.

$$\sup \; \langle b, y \rangle \quad \text{s.t.} \quad C - \sum_{j=1}^{m} y_j A_j \succeq 0, \tag{2.23}$$

and assume that $X$ in the primal program (2.17) is restricted to be invariant under action of $\mathcal{G}$. Then we may also assume, w.l.o.g., that the matrices $C$ and $A_j$ ($j = 1, 2, \dots, m$) are invariant. If they are not invariant, we get an equivalent SDP if we replace them with $\frac{1}{|\mathcal{G}|} \sum_{\sigma \in \mathcal{G}} \sigma(C)$ and $\frac{1}{|\mathcal{G}|} \sum_{\sigma \in \mathcal{G}} \sigma(A_j)$ ($j = 1, 2, \dots, m$). Since then $C$ and $A_j$ ($j = 1, 2, \dots, m$) belong to $\mathcal{A}$, the program (2.23) is therefore equivalent to

$$\begin{array}{ll} \sup & \langle b, y \rangle \\ \text{s.t.} & B_k \left( C - \sum_{j=1}^{m} y_j A_j \right) \succeq 0 \text{ for all } k = 0, 1, \dots, t. \end{array}$$

It is common to call this procedure, for reducing the sizes of matrices in (primal and/or dual) SDPs, 'block diagonalization'. It permits us to solve some huge, but at the same time highly symmetric, SDPs. For example, in Chapter

6 we deal with SDPs that involve matrices of size $O(2^n)$, whereas at the same time they belong to an algebra, known as the Terwilliger algebra of the binary Hamming scheme, of dimension $O(n^3)$. As a crucial ingredient we use the block diagonalization of the Terwilliger algebra given by Schrijver in [85].

**The block diagonalization of the Terwilliger algebra.** Given a positive integer $n$, set $N := \{1, 2, \dots, n\}$. For $i, j, p = 0, \dots, n$, let $M_{i,j}^{p,n}$ denote the 0/1 matrix indexed by $\mathcal{P}(N)$ whose $(I, J)$-th entry is 1 if $|I| = i$, $|J| = j$, $|I \cap J| = p$, and equal to 0 otherwise. The set

$$\mathcal{A}_n := \left\{ \sum_{i,j,p=0}^{n} x_{i,j}^p M_{i,j}^{p,n} \mid x_{i,j}^p \in \mathbb{R} \right\}$$

is a matrix $*$-algebra, known as the *Terwilliger algebra* of the binary Hamming scheme. For $k = 0, \dots, n$, let $M_k^n$ be the matrix indexed by $\mathcal{P}(N)$ whose $(I, J)$-th entry is 1 if $|I \triangle J| = k$ and 0 otherwise. The set

$$\mathcal{B}_n := \left\{ \sum_{k=0}^{n} x_k M_k^n \mid x_k \in \mathbb{R} \right\}$$

is an algebra, known as the *Bose-Mesner algebra* of the binary Hamming scheme. Obviously, $\mathcal{B}_n \subseteq \mathcal{A}_n$, since $M_k^n = \sum_{i,j,p|i+j-2p=k} M_{i,j}^{p,n}$. As is well known, $\mathcal{B}_n$ is a commutative algebra and thus all matrices in $\mathcal{B}_n$ can be simultaneously diagonalized (cf. Delsarte [21]). The Terwilliger algebra is not commutative, thus it cannot be diagonalized, however it can be block-diagonalized as we saw in Theorem 2.4.7. We recall the main result below.

Given integers $i, j, k, p = 0, \dots, n$, set

$$\beta_{i,j,k}^{p,n} := \sum_{u=0}^{n} (-1)^{p-u} \binom{u}{p} \binom{n-2k}{n-k-u} \binom{n-k-u}{i-u} \binom{n-k-u}{j-u}, \tag{2.24}$$

$$\alpha_{i,j,k}^{p,n} := \beta_{i,j,k}^{p,n} \binom{n-2k}{i-k}^{-\frac{1}{2}} \binom{n-2k}{j-k}^{-\frac{1}{2}}. \tag{2.25}$$

**Theorem 2.4.8.** *[85] For a matrix $M = \sum_{i,j,p} M_{i,j}^{p,n} x_{i,j}^p$ in the Terwilliger algebra,*

$$M \succeq 0 \Longleftrightarrow M_k := \left( \sum_p \alpha_{i,j,k}^{p,n} x_{i,j}^p \right)_{i,j=k}^{n-k} \succeq 0 \ \ \textit{for } k = 0, 1, \dots, \left\lfloor \frac{n}{2} \right\rfloor. \tag{2.26}$$

To show this, Schrijver [85] constructs an explicit matrix $U$ which satisfies the conditions given in Theorem 2.4.7. In fact, his $U$ is orthogonal, it block diagonalizes

$$\left\{ \sum_{i,j,p=0}^{n} x_{i,j}^p M_{i,j}^{p,n} \mid x_{i,j}^p \in \mathbb{C} \right\},$$

and consequently its subalgebra $\mathcal{A}_n$. The matrices $M_k$ from (2.26) play the role of the blocks $B_k$ appearing in the formulation of Theorem 2.4.7. Each of them is repeated $\binom{n}{k} - \binom{n}{k-1}$ times, for $k = 0, \ldots, \lfloor n/2 \rfloor$. Finally, (2.26) follows from (2.21).

The result extends to a block matrix whose blocks all lie in the Terwilliger algebra and which has a border of a special form. We state Lemma 2.4.9 for a $2 \times 2$ block matrix but the analogous result holds obviously for any number of blocks.

**Lemma 2.4.9.** *Let $A, B, C \in \mathcal{A}_n$; say, $A = \sum_{i,j,p} a^p_{i,j} M^{p,n}_{i,j}$, $B = \sum_{i,j,p} b^p_{i,j} M^{p,n}_{i,j}$, $C = \sum_{i,j,p} c^p_{i,j} M^{p,n}_{i,j}$ and define accordingly*

$$A_k := \left( \sum_p \alpha^{p,n}_{i,j,k} a^p_{i,j} \right)_{i,j=k}^{n-k}, \quad B_k := \left( \sum_p \alpha^{p,n}_{i,j,k} b^p_{i,j} \right)_{i,j=k}^{n-k},$$

$$C_k := \left( \sum_p \alpha^{p,n}_{i,j,k} c^p_{i,j} \right)_{i,j=k}^{n-k},$$

*for $k = 0, 1, \ldots, \lceil n/2 \rceil$. Then,*

$$\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0 \iff \begin{pmatrix} A_k & B_k \\ B_k^T & C_k \end{pmatrix} \succeq 0 \text{ for all } k = 0, 1, \ldots, \left\lfloor \frac{n}{2} \right\rfloor.$$

*Proof.* Directly from the above using the orthogonal matrix $\begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix}$. $\square$

**Lemma 2.4.10.** *(see Lemma 1 in [60]) Let $M = \sum_{i,j,p=0}^n x^p_{i,j} M^{p,n}_{i,j} \in \mathcal{A}_n$, $c = \sum_{i=0}^n c_i \chi^{(i)}$, where $\chi^{(i)} \in \{0,1\}^{\mathcal{P}(N)}$ with $\chi^{(i)}_I = 1$ if $|I| = i$ (for $I \in \mathcal{P}(N)$), and $d \in \mathbb{R}$. Let $M_k$ ($k = 0, 1, \ldots, \lfloor n/2 \rfloor$) be as in (2.26). Then,*

$$\begin{pmatrix} d & c^T \\ c & M \end{pmatrix} \succeq 0 \iff \begin{cases} M_k \succeq 0 \quad \text{for } k = 1, \ldots, \left\lfloor \frac{n}{2} \right\rfloor, \\ \tilde{M}_0 := \begin{pmatrix} d & \tilde{c}^T \\ \tilde{c} & M_0 \end{pmatrix} \succeq 0 \end{cases}$$

*after setting $\tilde{c}^T := \left( c_i \sqrt{\binom{n}{i}} \right)_{i=0}^n$.*

## 2.5 Graphs, stable sets and colourings

In this thesis we deal only with finite simple undirected graphs. A *graph* is an ordered pair $G = (V, E)$, where $V$ is a finite set and $E \subseteq \mathcal{P}_{=2}(V)$. The elements of $V$ and $E$ are usually called, respectively, the *vertices* and the *edges* of $G$. Two vertices $i, j \in V$ are *adjacent* if $\{i, j\} \in E$. Throughout we also write $ij \in E$. If $G$ is a graph we denote its vertex set as $V(G)$, and its edge set as $E(G)$.

The adjacency matrix of a graph $G$ is the 0/1 matrix indexed by $V(G)$, denoted by $A_G$ (or by $A$ if the graph is clear from the context) and defined by

$$(A_G)_{ij} := \begin{cases} 1 & \text{if } ij \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

The *degree* of a vertex $i \in V(G)$ is denoted $\deg(i)$ and defined by $\deg(i) := \sum_j (A_G)_{ij}$. We also define

$$\Delta(G) := \max_{i \in V(G)} \deg(i),$$

the *maximum degree* of a graph $G$.

Let $G$ and $H$ be two graphs such that $V(G)$ and $V(H)$ are disjoint. The *direct sum* of $G$ and $H$, denoted by $G + H$, is the graph with $V(G+H) := V(G) \cup V(H)$ and $E(G+H) := E(G) \cup E(H)$. The *strong product* of $G$ and $H$, denoted by $G \cdot H$, is the graph with $V(G \cdot H) := V(G) \times V(H)$, in which $(u, v)$ and $(u', v')$ are adjacent if $(u, v) \neq (u', v')$, $\{u, u'\} \in E(G) \cup V(G)$ and $\{v, v'\} \in E(H) \cup V(H)$. The *Cartesian product* of $G$ and $H$, denoted by $G \Box H$, is the graph with $V(G \Box H) := V(G) \times V(H)$, in which two vertices $(u, v)$ and $(u', v')$ are adjacent if they are adjacent in $G \cdot H$ and either $u = u'$ or $v = v'$.

Two graphs $G$ and $G'$ are *isomorphic* if there exists a bijection $f : V(G) \to V(G')$ such that

$$ij \in E(G) \Longleftrightarrow f(i)f(j) \in E(G').$$

Such a bijection is called an *isomorphism*. A graph $G'$ is a *subgraph* of a graph $G$ if $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. A graph $G'$ is an *induced subgraph* of a graph $G$ if $V(G') \subseteq V(G)$ and $E(G') = E(G) \cap \mathcal{P}_{=2}(V(G'))$.

A graph $G$, with $V(G) = \{v_1, v_2, ..., v_n\}$, is called

- a *cycle* if $n \geq 3$ and $E(G) = \{v_1v_2,\ v_2v_3, \ldots,\ v_{n-1}v_n,\ v_nv_1\}$;
- a *complete graph* if $E(G) = \mathcal{P}_{=2}(V(G))$.

A graph on $n$ vertices is often denoted by $C_n$ if it is a cycle, and by $K_n$ if it is a complete graph.

A set of vertices of a graph

- no two of which are adjacent is called a *stable set*;
- any two of which are adjacent is called a *clique*.

The maximum size of

- a stable set in $G$ is denoted by $\alpha(G)$ and called the *stability number*[5] of $G$;
- a clique in $G$ is denoted by $\omega(G)$ and called the *clique number* of $G$.

A partition of a vertex set of a graph $G$ into

- stable sets is a *(vertex) colouring* of $G$;
- cliques is a *clique cover* of $G$.

The stable sets of a vertex colouring are called the *colours* of the colouring. The minimum number of colours in a vertex colouring of $G$, denoted by $\chi(G)$, is called the *chromatic number* of $G$. The minimum number of cliques in a clique cover of $G$, denoted by $\overline{\chi}(G)$, is called the *clique cover number* of $G$.

[5] also known as the independence number of the graph $G$.

The *graph complement* of a graph $G = (V, E)$ is the graph $\overline{G} = (V, \overline{E})$, where $\overline{E} := \mathcal{P}_{=2}(V) \backslash E$. Note that every stable set in $G$ is a clique in $\overline{G}$, and analogously, a clique in $G$ is a stable set in $\overline{G}$. Similarly if a partition of the vertex set of a graph $G$ is a clique cover of $G$, it is at the same time a colouring of $\overline{G}$. Consequently,

- $\alpha(G) = \omega\left(\overline{G}\right)$;
- $\omega(G) = \alpha\left(\overline{G}\right)$;
- $\chi\left(\overline{G}\right) = \overline{\chi}(G)$.

In accordance with the third item above we use the following convention: given a graph parameter $\beta(\cdot)$, $\overline{\beta}(\cdot)$ is the graph parameter defined by $\overline{\beta}(G) := \beta\left(\overline{G}\right)$ for any graph $G$.

Some of the parameters mentioned above can be compared. If a graph $G$ contains a clique $C$ of size $\omega(G)$, then the vertices of $C$ are coloured by $\omega(G)$ distinct colours in any vertex colouring of $G$. Hence,

$$\omega(G) \leq \chi(G), \tag{2.27}$$

and equivalently

$$\alpha(G) \leq \overline{\chi}(G). \tag{2.28}$$

There exist graphs for which $\omega(G) < \chi(G)$, e.g., the pentagon $C_5$ is the smallest graph of this type (see section 3.2). On the other hand, for some graphs the equality might hold for all their induced subgraphs. Claude Berge [5, 6] was first to study these graphs. He calls a graph $G$ *perfect*[6] if

$$\alpha(G') = \overline{\chi}(G') \tag{2.29}$$

holds for every induced subgraph $G'$ of $G$. He conjectures that a graph $G$ is perfect if and only if any of the following two items hold:

- its complement $\overline{G}$ is perfect,
- $G$ does not contain an induced subgraph isomorphic[7] to $C_k$ or to $\overline{C}_k$ for some odd $k > 3$.

The equivalence with the first item was proved by Lovász in [62], and it is known as *the perfect graph theorem*. The equivalence with the second item turned out to be a much more complicated question. It was finally proved by Chudnovsky, Robertson, Seymour and Thomas in [19], and it is known as the *strong perfect graph theorem*.

As we have mentioned in the previous chapter, the problems of determining $\alpha(G)$ and $\chi(G)$ of a given graph $G$ are both NP-hard problems (see e.g. [30]). Namely, given a graph $G$ and an integer $k$, their decision counterparts (recall problems (S) and (C) from page 3):

a) decide if $\alpha(G) \geq k$, called the *stable set problem*;

[6] He calls it '$\alpha$-perfect', but this term was replaced by 'perfect' after Lovász proved the perfect graph theorem.

[7] An induced subgraph isomorphic to $C_k$ (respectively $\overline{C}_k$), $k > 3$, is usually called a hole (respectively antihole).

b) decide if $\chi(G) \leq k$, called the *colouring problem*;

are NP-complete problems. Besides, it is hard to approximate $\chi(G)$ within $|V(G)|^{1/14-\epsilon}$ for any $\epsilon > 0$ (see [3]), and $\alpha(G)$ can not be approximated within a factor $|V(G)|^{1/2-\epsilon}$ for any $\epsilon \geq 0$ unless[8] P=NP (see [42]).

Several generalizations of these problems were proposed in the literature. For example, given a graph $G = (V, E)$, a weight function $w : V \to \mathbb{R}$ and an integer $k$, the problem of deciding if there exists a stable set $S$ in $G$ such that $w(S) \geq k$ is called the *weighted stable set problem*.

In Chapter 5 we use the following reduction from the colouring problem to the stable set problem. Given a graph $G$ and an integer $t \geq 0$ consider $K_t \Box G$, the Cartesian product of the graphs $G$ and $K_t$. Note that

$$V(K_t \Box G) := V(K_t) \times V(G) = \bigcup_{p=1}^{t} V_p, \quad \text{where } V_p := \{pi \mid i \in V(G)\} \tag{2.30}$$

and $\{pi, qj\}$ is an edge if $i = j$ and $p \neq q$, or if $p = q$ and $ij \in E(G)$. Chvátal [20] observed the following:

$$\chi(G) \leq t \Longleftrightarrow \alpha(K_t \Box G) = |V(G)|. \tag{2.31}$$

To see this note first that $\alpha(K_t \Box G) \leq |V(G)|$ for every $t$. Let $S_p$ ($p = 1, 2, \ldots, t$) be the colours of a vertex colouring of $G$, and set $S_{pp} := \{pi \mid i \in S_p\}$ for $p = 1, 2, \ldots, t$. Now $\cup_{p=1}^{t} S_{pp}$ is a stable set of size $|V(G)|$ in $K_t \Box G$, which proves $\chi(G) \leq t \Longrightarrow \alpha(K_t \Box G) = |V(G)|$. For the other direction let $S$ be a stable set of size $|V(G)|$ in $K_t \Box G$. Set $S_p := \{i \in V(G) \mid pi \in S \cap V_p\}$ for $p = 1, 2, \ldots, t$ and observe that this gives a vertex colouring of $G$. (Reverse reductions, from the stability number to the chromatic number, can be found in Poljak [77], and in Schrijver [84].)

We call a graph 'symmetric' if it has a nontrivial automorphism group. Given a graph $G$, its *automorphism group* $\mathrm{Aut}(G)$ is the set of all bijections $\sigma : V(G) \to V(G)$ satisfying $ij \in E(G) \Longleftrightarrow \sigma(i)\sigma(j) \in E(G)$ for all $i, j \in V(G)$. A graph $G$ is *vertex transitive* if for every pair of vertices $i$ and $j$ there exists $\sigma \in \mathrm{Aut}(G)$ such that $\sigma(i) = j$.

## 2.6 Polynomials and optimization

Let $n$ be positive integer. For a vector $\beta \in \mathbb{N}^n$, we set

- $|\beta| := \sum_{i=1}^{n} \beta_i$,
- $\beta! := \beta_1! \cdots \beta_n!$,
- $S(\beta) := \{i \mid \beta_i \neq 0\}$, and
- $S_{odd}(\beta) := \{i \mid \beta_i \text{ is odd}\}$.

[8] More precisely, $\alpha(G)$ can not be approximated within $|V(G)|^{1-\epsilon}$ unless NP=ZPP

One says that $\beta$ is even when $S_{odd}(\beta) = \emptyset$. We also set

$$I_{\leq}(n,t) := \{\beta \in \mathbb{N}^n \mid |\beta| \leq t\} \text{ and } I_{=}(n,t) := \{\beta \in \mathbb{N}^n \mid |\beta| = t\} \text{ for } t \in \mathbb{N}.$$

The elements of $I_{\leq}(n,1)$ are the zero vector 0 and the standard unit vectors $e_i := \chi^{\{i\}}$ $(i = 1,\ldots,n)$. Given $t \in \mathbb{N}$, the mapping $f_t : I_{=}(n+1,t) \to I_{\leq}(n,t)$, given by $f_t(\beta_1,\ldots,\beta_n,\beta_{n+1}) := (\beta_1,\ldots,\beta_n)$, is bijective hence $|I_{\leq}(n,t)| = |I_{=}(n+1,t)| = \binom{n+t}{t}$.

Let $\mathbb{R}[x_1,x_2,\ldots,x_n] = \mathbb{R}[x]$ denote the ring of polynomials in $n$ variables. For $x \in \mathbb{R}^n$ and $\beta \in \mathbb{N}^n$ we set $x^\beta := \Pi_{i=1}^n x_i^{\beta_i}$. Every polynomial $p(x) \in \mathbb{R}[x]$ can be written as

$$p(x) = \sum_{\beta \in I_{\leq}(n,t)} p_\beta x^\beta \tag{2.32}$$

for some $t \in \mathbb{N}$, where $p = (p_\beta)_{\beta \in I_{\leq}(n,t)} \in \mathbb{R}^{I_{\leq}(n,t)}$ is the vector of coefficients. The *degree*[9] of $p(x)$, $p \neq 0$, is

$$\deg(p) := \max\{d \in \mathbb{N} \mid p_\beta \neq 0 \text{ for some } \beta \in I_{\leq}(n,d)\}.$$

A polynomial of the form

$$p(x) = \sum_{\beta \in I_{=}(n,t)} p_\beta x^\beta \in \mathbb{R}[x]$$

is said to be *homogeneous* of degree $t$. Given a polynomial in $n$ variables of degree $t$ in the form (2.32) its *homogenization* is the polynomial in $n+1$ variables defined by

$$p(\tilde{x}) := \sum_{\beta \in I_{\leq}(n,t)} p_\beta x^\beta x_{n+1}^{t-|\beta|} \tag{2.33}$$

where $\tilde{x} = [x_1,\ldots,x_n,x_{n+1}]$.

Following this, we sometimes identify a polynomial with its coefficient vector, i.e., the set of polynomials of degree at most $t$ with the vector space $\mathbb{R}^{I_{\leq}(n,t)}$, and the set of homogeneous polynomials of degree $t$ with the space $\mathbb{R}^{I_{=}(n,t)}$. Note that the zero polynomial lies in both sets, in $\mathbb{R}^{I_{\leq}(n,t)}$ and in $\mathbb{R}^{I_{=}(n,t)}$.

## Sums of squares

In polynomial optimization, the basic question is to decide if a given polynomial is nonnegative. Instead of trying to answer such a hard question directly, one can, instead, try to determine if a given polynomial is a sum of squares of polynomials. The idea is based on the following two facts:

(i) every polynomial which a sum of squares of polynomials is nonnegative, and

(ii) testing if a polynomial is a sum of squares of polynomials can be done via semidefinite programming.

[9] It is sometimes convenient to set $\deg(0) = -\infty$.

The reverse of (i) is, however, not true. It was proved by Hilbert already in 1888 (See Reznick [81] for a nice survey on this topic.).

An $n$ variable polynomial $p(x) \in \mathbb{R}[x]$ is a *sum of squares of polynomials*, sometimes abbreviated as '$p(x)$ is SOS', if $p(x) = \sum_{i=1}^{k}(u_i(x))^2$ for some $u_i(x) \in \mathbb{R}[x]$ $(i = 1, 2, \dots, k)$. If a polynomial of degree at most $2t$ is SOS then it can be written as a sum of squares of polynomials, each of degree at most $t$. Moreover, if a homogeneous polynomial of degree $2t$ is SOS then it can be written as a sum of squares of homogeneous polynomials, each of degree $t$. In fact, a polynomial is SOS if and only if its homogenization is SOS. We can now observe the following:

(i) $p(x) = \sum_{\beta \in I_{\leq}(n,2t)} p_\beta x^\beta$ is SOS if and only if there exist $X \in \mathbb{R}^{I_{\leq}(n,t) \times I_{\leq}(n,t)}$ such that

$$X \succeq 0 \quad \text{and} \quad \sum_{\substack{\gamma,\delta \in I_{\leq}(n,t) \\ \gamma+\delta=\beta}} X_{\gamma,\delta} = p_\beta \text{ for all } \beta \in I_{\leq}(n,2t); \tag{2.34}$$

(ii) $p(x) = \sum_{\beta \in I_{=}(n,2t)} p_\beta x^\beta$ is SOS of polynomials if and only if there exist $X \in \mathbb{R}^{I_{=}(n,t) \times I_{=}(n,t)}$ such that

$$X \succeq 0 \quad \text{and} \quad \sum_{\substack{\gamma,\delta \in I_{=}(n,t) \\ \gamma+\delta=\beta}} X_{\gamma,\delta} = p_\beta \text{ for all } \beta \in I_{=}(n,2t). \tag{2.35}$$

In other words, deciding if a polynomial $p(x)$ is SOS is equivalent to deciding if the system (2.34) (respectively (2.35) if $p(x)$ is homogeneous) is feasible.

We also define

- $\Sigma(n,2t) := \{p \in \mathbb{R}^{I_{\leq}(n,2t)} \mid p(x) \text{ is SOS}\}$, the cone of polynomials of degree at most $2t$ which are sums of squares, and
- $\Sigma_{=}(n,2t) := \{p \in \mathbb{R}^{I_{=}(n,2t)} \mid p(x) \text{ is SOS}\}$, the cone of homogeneous polynomials of degree $2t$ which are sums of squares.

The cones defined above are closed, convex and solid (see e.g. Reznick [81]). Observe also that the cones $\Sigma_{=}(n+1,2t)$ and $\Sigma(n,2t)$ are basically identical (isomorphic) due to the properties of homogenization.

## Sequences and moment matrices

Given a sequence

- $(y_\beta)_{\beta \in I_{\leq}(n,2t+s)} \in \mathbb{R}^{I_{\leq}(n,2t+s)}$, where $s \in \mathbb{N}$ , its *moment matrix of order* $t$ is the matrix

  $N_t(y) \in \mathbb{R}^{I_{\leq}(n,t) \times I_{\leq}(n,t)}$ defined by $(N_t(y))_{\beta,\gamma} := y_{\beta+\gamma}$, for $\beta, \gamma \in I_{\leq}(n,t)$;

- $(y_\beta)_{\beta \in I_{=}(n,2t)} \in \mathbb{R}^{I_{=}(n,2t)}$, its moment matrix is the matrix

  $N_{=t}(y) \in \mathbb{R}^{I_{=}(n,t) \times I_{=}(n,t)}$ defined by $(N_{=t}(y))_{\beta,\gamma} := y_{\beta+\gamma}$, for $\beta, \gamma \in I_{=}(n,t)$.

Given a polynomial in the form (2.32) and a sequence $y \in \mathbb{R}^{I_{\leq}(n,2t+s)}$, where $s \in \mathbb{N}$, define the new sequence

$$py := N_t(y)p \in \mathbb{R}^{I_{\leq}(n,t)}, \text{ with } \beta\text{th entry } (py)_\beta := \sum_{\gamma \in I_{\leq}(n,t)} p_\gamma y_{\beta+\gamma},\ \beta \in \mathbb{R}^{I_{\leq}(n,t)}.$$

Define, analogously, $py := N_{=t}(y)p \in \mathbb{R}^{I_=(n,t)}$ when $p \in \mathbb{R}^{I_=(n,t)}$ and $y \in \mathbb{R}^{I_=(n,2t)}$ are given.

In this thesis we often deal with vectors indexed by $\mathcal{P}(V)$, where $V$ is some finite set. Given a vector $y \in \mathbb{R}^{\mathcal{P}_{\leq 2t+s}(V)}$, its *combinatorial moment matrix of order* $t$ is the matrix

$$M_t(y) \in \mathbb{R}^{\mathcal{P}_{\leq t}(V) \times \mathcal{P}_{\leq t}(V)} \text{ defined by } (M_t(y))_{IJ} := y_{I \cup J}\ (I, J \in \mathcal{P}_{\leq t}(V)).$$

Given vectors $p \in \mathbb{R}^{\mathcal{P}_{\leq t}(V)}$ and $y \in \mathbb{R}^{\mathcal{P}_{\leq 2t+s}(V)}$, define the new vector

$$py := M_t(y)p \in \mathbb{R}^{\mathcal{P}_{\leq t}(V)}, \text{ with } I\text{th entry } (py)_I := \sum_{J \in \mathcal{P}_{\leq t}(V)} p_J y_{I \cup J},\ I \in \mathbb{R}^{\mathcal{P}_{\leq t}(V)}.$$

We also define the following cones of sequences:

- $\mathcal{N}(n, 2t) := \{y \in \mathbb{R}^{I_{\leq}(n,2t)} \mid N_t(y) \succeq 0\}$;
- $\mathcal{N}_=(n, 2t) := \{y \in \mathbb{R}^{I_=(n,2t)} \mid N_{=t}(y) \succeq 0\}$.

Observe again that the cones $\mathcal{N}_=(n+1, 2t)$ and $\mathcal{N}(n, 2t)$ are basically identical.

## Duality between sums of squares and moment sequences

Recall first the definition of a dual cone from Section 2.2. The dual cone of a cone $\mathcal{K} \subseteq \mathbb{R}^V$ is

$$\mathcal{K}^* := \{p \in \mathbb{R}^V \mid \langle p, q \rangle \geq 0 \text{ for all } q \in \mathcal{K}\}. \tag{2.36}$$

We present now the known duality between the cones of sums of squares of polynomials and the cones of sequences with a positive semidefinite moment matrix. We will use this result in Section 4.2, where we define a hierarchy of cones of matrices nested between the semidefinite cone and the copositive cone. In particular, we will characterize the cones in the dual hierarchy by applying the following theorem.

**Theorem 2.6.1.** *For the cones defined above we have the following duality relations*

*(i)* $(\mathcal{N}(n, 2d))^* = \Sigma(n, 2d)$, $(\Sigma(n, 2d))^* = \mathcal{N}(n, 2d)$;

*(ii)* $(\mathcal{N}_=(n, 2d))^* = \Sigma_=(n, 2d)$, $(\Sigma_=(n, 2d))^* = \mathcal{N}_=(n, 2d)$;

*Proof.* We prove only (ii) which obviously implies (i). Let $p \in \Sigma_=(n, 2d)$ be the square of a polynomial $u \in \mathbb{R}^{I_=(n,d)}$, i.e.,

$$p(x) = \sum_{\beta \in I_=(n,2d)} p_\beta x^\beta = u(x)^2 = \sum_{\beta \in I_=(n,2d)} \sum_{\substack{\gamma,\delta \in \mathbb{R}^{I_=(n,d)} \\ \gamma+\delta=\beta}} u_\gamma u_\delta x^\beta.$$

Given a sequence $y \in \mathcal{N}_{=}(n, 2d)$ we have

$$\langle p, y\rangle = p^T y = \sum_{\beta \in I_{=}(n,2d)} p_\beta y_\beta = \sum_{\beta \in I_{=}(n,2d)} \sum_{\substack{\gamma,\delta \in \mathbb{R}^{I_{=}(n,d)} \\ \gamma+\delta=\beta}} u_\gamma u_\delta y_\beta = u^T N_{=d}(y) u.$$

Hence, $(\Sigma_{=}(n, 2d))^* = \mathcal{N}_{=}(n, 2d)$. Since $\Sigma_{=}(n, 2d)$ is closed convex cone, from Theorem 2.2.1 we get $(\mathcal{N}_{=}(n, 2d))^* = ((\Sigma_{=}(n, 2d))^*)^* = \Sigma_{=}(n, 2d)$. $\square$

For a more general result consult [58] and the references therein.

## Polynomial optimization problem

Given polynomials $p(x)$ and $g_j(x)$ $(j = 1, \ldots, m)$ the polynomial optimization problem is the problem

$$p^* = \inf p(x) \ \text{ s.t. } x \in K, \tag{2.37}$$

where $K := \{x \in \mathbb{R}^n | \ g_j(x) \geq 0 \ (j = 1, 2, \ldots, m)\}$. We set $d_p := \lceil \deg(p)/2 \rceil$, $d_j := \lceil \deg(g_j)/2 \rceil$ $(j = 1, \ldots, m)$ and $d_K := \max(d_1, \ldots, d_m)$. We also assume $\deg(p) \geq 1$.

Among numerous applications let us mention that the polynomial optimization problem captures linear, semidefinite and copositive programs. In particular, a matrix $M \in \mathbb{R}^{n\times n}$ is copositive if and only if $p(x) := x^T M x \geq 0$ for all $x \in K = \mathbb{R}^n_+ = \{y \in \mathbb{R}^n \mid y_i \geq 0 \text{ for all } i = 1, \ldots, n\})$. It can be also used for modelling combinatorial optimization problems since

$$\begin{aligned} \{0,1\}^n &= \{x \in \mathbb{R}^n \mid x_i^2 - x_i \geq 0, \ -x_i^2 + x_i \leq 0 \text{ for all } i = 1, \ldots, n\} \\ &= \{x \in \mathbb{R}^n \mid x_i^2 = x_i \text{ for all } i = 1, \ldots, n\}. \end{aligned}$$

For example, given a graph $G = (V, E)$, the stability number $\alpha(G)$ of $G$ can be formulated as

$$\alpha(G) = \max_{x \in \mathbb{R}^V} e^T x \ \text{ s.t. } \ x_i^2 = x_i \ (i \in V), \ x_i x_j = 0 \ (ij \in E). \tag{2.38}$$

As we have mentioned earlier, testing if a given matrix is not copositive, and deciding if $\alpha(G) \geq k$ for a given $k$, are NP-complete problems. Consequently, the problem (2.37) is a hard problem in general.

**Relaxations.** Several types of relaxations for (2.37) were proposed recently. One way to relax is by replacing 'nonnegativity' by the 'sum of squares' condition. First note that (2.37) can formulated as

$$p^* = \sup_{\rho \in \mathbb{R}} \rho \ \text{ s.t. } p(x) - \rho \geq 0, \text{ for all } x \in K.$$

Following Lasserre [56], one way to relax the condition $p(x) - \rho \geq 0$, for all $x \in K$, is to require $p(x) - \rho = s_0(x) + \sum_{j=1} s_j(x) g_i(x)$, where $s_0, s_j$ $(j = 1, \ldots, m)$ are SOS. Thus we can define

$$p^\Sigma := \sup \rho \text{ s.t. } p(x) - \rho = s_0(x) + \sum_{j=1} s_j(x) g_j(x), \ \rho \in \mathbb{R}, \ s_0, s_j \text{ are SOS.}$$

If we further bound the degrees of polynomials $s_0, s_j$, we get the sequence of semidefinite programming bounds

$$p_t^\Sigma := \sup \rho \text{ s.t. } \begin{array}{l} p(x) - \rho = s_0(x) + \sum_{j=1} s_j(x) g_j(x),\ \rho \in \mathbb{R},\ s_0, s_j \text{ are SOS}, \\ \deg(s_0), \deg(s_j g_j) \leq 2t, \end{array} \tag{2.39}$$

for every $t$ which satisfies $2t \geq \max(\deg(p), \deg(g_1), \ldots, \deg(g_m))$. The bounds satisfy $p_t^\Sigma \leq p_{t+1}^\Sigma \leq p^\Sigma \leq p^*$ and $\lim_{t\to\infty} p_t^\Sigma = p^\Sigma$.

Next we recall the dual approach proposed also in [56], based on moment matrices. Following [56], for $t \geq \max(d_p, d_K)$, we define

$$\begin{array}{rcll} p_t & := & \inf p^T y & \text{s.t.} \quad y \in \mathbb{R}^{I_\leq(n,2t)},\ y_0 = 1,\ N_t(y) \succeq 0, \\ & & & \phantom{\text{s.t.} \quad} N_{t-d_j}(g_j y) \succeq 0 \ (j = 1, \ldots, m). \end{array} \tag{2.40}$$

Obviously, $p_t \leq p_{t+1}$. To see that (2.40) is a relaxation of (2.37) we need to prove

$$p_t \leq p^*, \text{ for all } t \geq \max(d_p, d_K). \tag{2.41}$$

For $x \in \mathbb{R}^n$ define its *zeta vector*

$$\zeta_{t,x} \in \mathbb{R}^{I_\leq(n,t)} \text{ by } (\zeta_{t,x})_\beta := x^\beta \ (\beta \in I_\leq(n,t)).$$

Consider the vector $y := \zeta_{2t,x} \in \mathbb{R}^{I_\leq(n,2t)}$. Then, $y_0 = x^0 = 1$ and $N_t(y) = \zeta_{t,x}\zeta_{t,x}^T \succeq 0$. Moreover, if $x \in K$ we have $N_{t-d_j}(g_j y) = g_j(x) N_{t-d_j}(y) \succeq 0$, hence $y$ is feasible for (2.40) with $p^T y = p(x)$, implying $p_t \leq p^*$.

The programs (2.39) and (2.40) are dual SDPs (see Lasserre [56]), which implies $p_t^\Sigma \leq p_t$.

**0/1 case.** Under certain conditions on $K$, the bounds $p_t$ converge to $p^*$ in finitely many steps. In other words, there exists $t \in \mathbb{N}$ such that $p_t = p^*$. We prove that this holds for $t \geq n + d_K$ when $K \subseteq \{0,1\}^n$. For an exposition on more general results see e.g. [58]. Let

$$K = \{x \in \mathbb{R}^V \mid g_j(x) \geq 0 \ (j = 1, 2, \ldots, m);\ h_i(x) \geq 0,\ -h_i(x) \geq 0,\ (i \in V)\}, \tag{2.42}$$

where $V = \{1, 2, \ldots, n\}$, $g_j(x)$ $(j = 1, 2, \ldots, m)$ are polynomials and $h_i(x) = x_i^2 - x_i$ $(i \in V)$.

Set $n := |V|$. The program (2.40), with $K$ of form (2.42), now reads

$$\begin{array}{rcll} p_t & := & \inf p^T y & \text{s.t.} \quad y \in \mathbb{R}^{I_\leq(n,2t)},\ y_0 = 1,\ N_t(y) \succeq 0, \\ & & & \phantom{\text{s.t.} \quad} N_{t-d_j}(g_j y) \succeq 0 \ (j = 1, \ldots, m), \\ & & & \phantom{\text{s.t.} \quad} N_{t-1}(h_i y) \succeq 0,\ N_{t-1}(-h_i y) \succeq 0 \ (i = 1, \ldots, n). \end{array} \tag{2.43}$$

For $x \in K$ we have $x_i^2 = x_i$ $(i \in V)$, and thus we may assume, w.l.o.g., that each variable occurs in $p(x)$ and $g_j(x)$ $(j = 1, \ldots, m)$ with a degree at most one. In other words we may assume that $p_\beta \neq 0$ (analogously $(g_j)_\beta \neq 0$ $(j = 1, \ldots, m)$) only if $\beta \in \{0,1\}^n$, i.e. if $\beta = \chi^I$ for some $I \subseteq V$. (Recall the definition of a characteristic vector $\chi^I$.) Motivated by this we define $\widehat{p} \in \mathbb{R}^{\mathcal{P}(V)}$ by $\widehat{p}_I := p_{\chi^I}$ $(I \subseteq V)$, and $\widehat{g}_j \in \mathbb{R}^{\mathcal{P}(V)}$ $(j = 1, \ldots, m)$ by $(\widehat{g}_j)_I := (g_j)_{\chi^I}$ $(I \subseteq V)$. The program (2.37), with $K$ given by (2.42), can thus be reformulated as

$$p^* = \min\ \widehat{p}^T \widehat{x} \ \text{ s.t. } \ \widehat{x} \in \{0,1\}^{\mathcal{P}(V)},\ \widehat{g}_j^T \widehat{x} \geq 0 \ (j = 1, \ldots, m). \tag{2.44}$$

We now make similar observations for a vector $y$, which is assumed to be feasible for (2.43).

**Lemma 2.6.2.** *Let $y$ be feasible for the program (2.43). For all $\beta \in \mathbb{R}^{I_\leq(n,2t)}$*

$$y_\beta = y_\gamma, \text{ where } \gamma_i := \min(\beta_i, 1) \ (i \in V).$$

*Proof.* For all $\beta \in \mathbb{R}^{I_\leq(n-1,2t)}$ and all $i \in S(\beta)$, $y_{\beta+e_i} - y_\beta$ appears as an entry in $N_{t-1}(h_i y)$. For all $i \in V$ we have $N_{t-1}(h_i y) \succeq 0$ and $N_{t-1}(-h_i y) \succeq 0$. This implies $N_{t-1}(h_i y) = 0$ and thus proves the lemma. □

Consequently, two columns of $N_t(y)$ indexed by $\beta, \gamma \in \mathbb{R}^{I_\leq(n,t)}$ are identical if $S(\beta) = S(\gamma)$. Moreover, it further implies that two columns of $N_{t-d_j}(g_j y)$ are identical if their indices have an identical support.

Motivated by Lemma 2.6.2 and the last observations, we define the vector $\hat{y} \in \mathbb{R}^{\mathcal{P}_{\leq 2t}(V)}$ by $(\hat{y})_I := y_{\chi^I}$ $(I \in \mathcal{P}_{\leq 2t}(V))$. Recall its combinatorial moment matrix $M_t(\hat{y}) \in \mathbb{R}^{\mathcal{P}_{\leq t}(V) \times \mathcal{P}_{\leq t}(V)}$ whose entries are given by

$$(M_t(\hat{y}))_{I,J} = \hat{y}_{I \cup J}.$$

The matrix $M_t(\hat{y})$ is a principal submatrix of $N_t(y)$, and moreover

$$M_t(\hat{y}) \succeq 0 \iff N_t(y) \succeq 0. \tag{2.45}$$

Similarly,

$$M_{t-d_j}(\widehat{g}_j \hat{y}) \succeq 0 \iff N_{t-d_j}(g_j y) \succeq 0 \ (j = 1, \ldots, m).$$

The program (2.43) is therefore equivalent to

$$\begin{array}{rclll} p_t & := & \inf \widehat{p}^T \hat{y} & \text{s.t.} & \hat{y} \in \mathbb{R}^{\mathcal{P}_{\leq 2t}(V)}, \ \hat{y}_{\mathbf{0}} = 1, \ M_t(\hat{y}) \succeq 0, \\ & & & & M_{t-d_j}(\widehat{g}_j \hat{y}) \succeq 0 \ (j = 1, \ldots, m). \end{array} \tag{2.46}$$

Observe that the formulation (2.46) involves smaller matrices and less variables than (2.43).

We focus now on the convergence of the bounds $p_t$. Before proving the main result we need the following lemma.

**Lemma 2.6.3.** *Let $\widehat{y}$ be feasible for the program (2.46), and let $Z \in \{0,1\}^{\mathcal{P}(V)}$ be the zeta matrix. If $t \geq n$ then*

$$\widehat{y} \in \text{conv}\{Z_{\bullet I} \mid I \subseteq V\}$$

*Proof.* Assume that $t \geq n$ and recall the definition of the zeta matrix $Z \in \mathbb{R}^{\mathcal{P}(V)}$: $Z_{I,J} = 1$ if $I \subseteq J$ and $Z_{I,J} = 0$ otherwise, for $I, J \subseteq V$. Corollary 2.4.6 gives

$$M_t(\hat{y}) = M_n(\hat{y}) \succeq 0 \iff Z^{-1}\hat{y} = \sum_{I \subseteq I' \subseteq V} (-1)^{|I' \setminus I|} \hat{y}_{I'} \geq 0 \text{ for all } I \subseteq V.$$

Moreover, $\hat{y} = Z\left(Z^{-1}\hat{y}\right) \in \text{conv}\{Z_{\bullet I} \mid I \subseteq V\}$. □

We can finally state the main result.

**Theorem 2.6.4.** *Let $K$ be as in (2.42). If $t \geq n + d_K$ then $p_t = p^*$.*

*Proof.* Assume that $t \geq n + d_K$. Let $\widehat{y}$ be an optimal solution for the program (2.46). From Lemma 2.6.3 we have that $\widehat{y} = \sum_{s=1}^{r} \lambda_s Z_{\bullet I_s}$ for some distinct $I_s \in \mathcal{P}(V)$ $(s = 1, \ldots, r)$, $\lambda_s > 0$ $(s = 1, \ldots, r)$ and $\widehat{y}_{\mathbf{0}} = \sum_{s=1}^{r} \lambda_s = 1$. It is now enough to show that $Z_{\bullet I_s}$ $(s = 1, \ldots, r)$ are feasible for the program (2.44), since then

$$\widehat{p}^T \widehat{y} = \sum_{s=1}^{r} \lambda_s \widehat{p}^T Z_{I_s} \geq \sum_{s=1}^{r} \lambda_s p^* = p^*. \tag{2.47}$$

It can be verified that

$$\widehat{q}^T M_{t-d_j}(\widehat{g}_j \widehat{y}) \widehat{q} = \sum_{s=1}^{r} \lambda_s \widehat{g}_j^T Z_{\bullet I_s} \left(\widehat{q}^T Z_{\bullet I_s}\right)^2 \tag{2.48}$$

for all $\widehat{q} \in \mathbb{R}^{\mathcal{P}(V)}$ and $j = 1, \ldots, m$.

The vectors $Z_{\bullet I_s}$ $(s = 1, \ldots, r)$ are linearly independent since the columns of $Z$ are linearly independent. Therefore, the system

$$\widehat{q}^T Z_{\bullet I_s} = a_s \ (s = 1, \ldots, r), \tag{2.49}$$

has a solution for every choice of $a_s$ $(s = 1, \ldots, r)$.

Pick an $s_0 \in \{1, \ldots, r\}$ and let $\widehat{q}$ be a solution of (2.49) for the choice $a_{s_0} := \frac{1}{\sqrt{\lambda_{s_0}}}$, $a_s := 0$ $(s \neq s_0)$. Then

$$\widehat{q}^T M_{t-d_j}(\widehat{g}_j \widehat{y}) \widehat{q} = \sum_{s=1}^{r} \lambda_s \widehat{g}_j^T Z_{\bullet I_s} \left(\widehat{q}^T Z_{\bullet I_s}\right)^2 = \widehat{g}_j^T Z_{\bullet I_{s_0}} \geq 0,$$

for every $j = 1, \ldots, m$. In other words, $Z_{\bullet I_{s_0}}$ is feasible for the program (2.44).

Finally, (2.47) and (2.41) yield $p^* = p_t$. □

Consider now the problem (2.38). Using the ideas from above, the relaxation (2.46) for this problem reduces to

$$p_t(G) := \max_{\hat{y} \in \mathcal{P}_{\leq 2t}(V)} \sum_{i=1}^{|V|} \hat{y}_i \text{ s.t. } \hat{y}_0 = 1,\ M_t(\hat{y}) \succeq 0,\ \hat{y}_{ij} = 0\ (ij \in E). \tag{2.50}$$

To see this observe that the conditions $x_i x_j = 0$ $(ij \in E(G))$ in (2.38) transform into

$$(\widehat{y}_{I \cup J \cup \{i,j\}})_{I,J \in \mathcal{P}_{\leq (t-1)}(V)} = 0 \ (ij \in E) \tag{2.51}$$

in the relaxation (2.46), and that the conditions (2.51) are implied by $\widehat{y}_{ij} = 0$ $(ij \in E)$ and $M_t(\widehat{y}) \succeq 0$.

**Proposition 2.6.5.** *Let $G = (V, E)$ be a graph. Then $p_{\alpha(G)}(G) = \alpha(G)$.*

*Proof.* From Theorem 2.6.4 we have that $p_{|V|+1}(G) = \alpha(G)$. Let $t \geq \alpha(G)$ and $\widehat{y}$ be feasible for the program (2.50). Then $\widehat{y}_I = 0$ for every $I \in \mathcal{P}_{\leq 2t}(V)$ for which there exists $ij \in E$ such that $\{i, j\} \subseteq I$, due to (2.51). In particular, for every $I \subseteq \mathcal{P}_{\leq t}(V) \backslash \mathcal{P}_{\leq \alpha(G)}(V)$ we have $\widehat{y}_I = 0$. Consequently, the columns of $M_t(\widehat{y})$ indexed by $\mathcal{P}_{\leq t}(V) \backslash \mathcal{P}_{\leq \alpha(G)}(V)$ are zero columns.

Therefore $p_{\alpha(G)}(G) = p_t(G) = p_{|V|+1}(G) = \alpha(G)$ for all $t \geq \alpha(G)$. □

We will discuss the program (2.50) further in Chapter 4.

# Chapter 3

# The Lovász theta number

The Lovász theta function maps a graph $G = (V, E)$ to $\mathbb{R}_+$. It was introduced by Lovász in [64] for bounding the stability number and the Shannon capacity[1] of a graph. Several equivalent formulations using orthonormal representations, eigenvalues, adjacency matrices, semidefinite programming were studied in [54, 61, 64]. Here we define the Lovász theta number via semidefinite programming and we prove 'the sandwich theorem'. We finalize the chapter with some variations of the Lovász theta number obtained by adding nonnegativity and triangle constraints.

## 3.1 Equivalent formulations

Given a graph $G = (V, E)$, set $n := |V|$. We define $\vartheta(G)$, the *Lovász theta number* of $G$, via the SDP

$$\begin{array}{rll} \vartheta(G) := & \max & \langle \mathbf{J}, X \rangle \\ & \text{s.t.} & \mathrm{Tr}(X) = 1 \\ & & X_{ij} = 0 \ \ (ij \in E(G)) \\ & & X \succeq 0, \end{array} \tag{3.1}$$

where $X$ is a symmetric matrix indexed by $V$ (or $\mathcal{P}_{=1}(V)$).

If $V = \emptyset$ we define $\vartheta(G) := 0$. This is however trivial case.

The dual SDP of (3.1) reads

$$\begin{array}{ll} \min & t \\ \text{s.t.} & t\mathbf{I} + \sum_{ij \in E} y_{ij} E^{ij} - \mathbf{J} = Z \\ & Z \succeq 0, \ t \in \mathbb{R}, \ y_{ij} \in \mathbb{R} \ (ij \in E(G)), \end{array} \tag{3.2}$$

where $E^{ij} \in \mathbb{R}^{V \times V}$, $E^{ij}_{pq} := 1$ if $\{p, q\} = \{i, j\}$ and $E^{ij}_{pq} := 0$ otherwise.

Since $\frac{1}{|V|}\mathbf{I}$ is strictly feasible for (3.1) and $(|V| + 1)\mathbf{I} - \mathbf{J}$ is strictly feasible for (3.2) the minimum in (3.2) is equal to $\vartheta(G)$ (directly from Corollary 2.3.3).

[1] The Shannon capacity of a graph $G$ is defined as $\Theta(G) := \lim_{n\to\infty} (\alpha(G^n))^{\frac{1}{n}}$, where $G^n$ is given by $G^1 := G$ and $G^n := G^{n-1} \cdot G$ for $n \geq 2$.

Let $(t, Z)$ be feasible for (3.2). Set

$$Z' := \begin{pmatrix} t & e^T \\ e & \frac{1}{t}Z + \frac{1}{t}ee^T \end{pmatrix}. \tag{3.3}$$

Consider $Z'$ to be indexed by $\mathcal{P}_{\leq 1}(V)$, with the first row corresponding to $\mathbf{0}$ (the empty set). Lemma 2.4.2 yields $Z' \succeq 0 \iff \frac{1}{t}Z \succeq 0 \iff Z \succeq 0$, and thus (3.2) is equivalent to

$$\begin{array}{llll} \vartheta(G) & \min & Z'_{\mathbf{00}} \\ & \text{s.t.} & Z'_{\mathbf{0}i} + Z'_{i\mathbf{0}} = 2 \ \ (i \in V) \\ & & Z'_{ii} = 1 \ \ (i \in V) \\ & & Z'_{ij} = 0 \ \ (ij \in E(\overline{G})) \\ & & Z' \succeq 0 \end{array} \quad = \quad \begin{array}{lll} \min & Z'_{\mathbf{00}} \\ \text{s.t.} & Z'_{ii} = Z'_{\mathbf{0}i} \ \ (i \in V) \\ & Z'_{ii} = 1 \ \ (i \in V) \\ & Z'_{ij} = 0 \ \ (ij \in E(\overline{G})) \\ & Z' \succeq 0 \end{array} \tag{3.4}$$

where variable $Z'$ is a symmetric matrix indexed by $\mathcal{P}_{\leq 1}(V)$.

Strict feasibility of (3.2) implies strict feasibility of (3.4) since for $Z'$ from (3.3) we have $Z' \succ 0 \iff Z \succ 0$. Hence, we can write the dual of (3.4) as

$$\begin{array}{lll} \vartheta(G) & \max & -2\sum_{i \in V} X'_{\mathbf{0}i} - \sum_{i \in V} X'_{ii} \\ & \text{s.t.} & X'_{\mathbf{00}} = 1 \ \ (i \in V) \\ & & X'_{ij} = 0 \ \ (ij \in E(G)) \\ & & X' \succeq 0. \end{array} \tag{3.5}$$

The following lemma leads us to a formulation which is extensively used in the next chapter.

**Lemma 3.1.1.** *If $X'$ is an optimal solution for the program (3.5) then $X'_{\mathbf{0}i} + X'_{ii} = 0$ for all $i \in V$.*

*Proof.* Let $X'$ be feasible for the program (3.5), and suppose $a := X'_{\mathbf{0}i} + X'_{ii} \neq 0$ for some $i \in V$. Then $b := X'_{ii} > 0$ since $X'_{ii} \geq (X'_{\mathbf{0}i})^2$. Set $c_{X'} := -2\sum_{i \in V} X'_{\mathbf{0}i} - \sum_{i \in V} X'_{ii}$. The matrix $X''$ obtained by multiplying the $i$th row and the $i$th column of $X'$ by $1 - \frac{a}{b}$ is feasible for (3.5) with the objective value $c_{X'} + \frac{a^2}{b}$. □

By Lemma 3.1.1 we can restrict the feasible set in (3.5) to the PSD matrices

$$X' = \begin{pmatrix} 1 & -x^T \\ -x & X \end{pmatrix}, \tag{3.6}$$

such that $\operatorname{diag}(X) = x$, and $X_{ij} = 0$ for all $ij \in E(G)$. Moreover, the objective function then reads $2\sum_{i \in V} x_i - \sum_{i \in V} X_{ii} = \sum_{i \in V} x_i$. Multiplying the first row and the first column of $X'$ by $-1$ gives

$$\begin{pmatrix} 1 & -x^T \\ -x & X \end{pmatrix} \succeq 0 \iff \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0. \tag{3.7}$$

The program (3.5) is thus equivalent to:

$$\begin{array}{lll} \vartheta(G) = & \max & \sum_{i \in V} X'_{ii} \\ & \text{s.t.} & X'_{\mathbf{00}} = 1 \\ & & X'_{ii} = X'_{\mathbf{0}i} \ (i \in V) \\ & & X'_{ij} = 0 \ (ij \in E) \\ & & X' \succeq 0. \end{array} \tag{3.8}$$

This program is also strictly feasible. To see this take $x := \frac{1}{|V|+1}e$ and $X := \mathrm{Diag}(x)$ and set $X' := \begin{pmatrix} 1 & x \\ x & X \end{pmatrix}$.

## 3.2 The sandwich theorem

The sandwich theorem [64] compares three graph parameters: the stable set number, the clique cover number and the Lovász theta number.

**Theorem 3.2.1** (The sandwich theorem)**.** *For any graph $G = (V, E)$, one has*

$$\alpha(G) \leq \vartheta(G) \leq \overline{\chi}(G). \tag{3.9}$$

*Proof.* Let $S \subseteq V$ be a stable set in $G$ such that $|S| = \alpha(G)$. Set $X' := \begin{pmatrix} 1 \\ \chi^S \end{pmatrix} \begin{pmatrix} 1 \\ \chi^S \end{pmatrix}^T$. Here $\chi^S$ is the characteristic vector of $S$ in $\mathbb{R}^V$, and $\begin{pmatrix} 1 \\ \chi^S \end{pmatrix} = \chi^{\mathcal{P}_{\leq 1}(S)}$ is the characteristic vector of $\mathcal{P}_{\leq 1}(S)$ in $\mathbb{R}^{\mathcal{P}_{\leq 1}(V)}$. Since $X'$ is feasible for the program (3.8) and $\sum_i X'_{ii} = \alpha(G)$, we have $\alpha(G) \leq \vartheta(G)$.

Let $C_j \subseteq V$ $(j = 1, 2, \ldots, \overline{\chi}(G))$ be disjoint cliques such that $\bigcup_j C_j = V$. Take their characteristic vectors $\chi^{C_j} \in \mathbb{R}^V$ $(j = 1, 2, \ldots, \overline{\chi}(G))$ and set

$$Z' := \sum_j \begin{pmatrix} 1 \\ \chi^{C_j} \end{pmatrix} \begin{pmatrix} 1 \\ \chi^{C_j} \end{pmatrix}^T. \tag{3.10}$$

The matrix $Z'$ is now feasible for (3.4) and $Z'_{\mathbf{00}} = \overline{\chi}(G)$. This proves $\vartheta(G) \leq \overline{\chi}(G)$. ◻

In fact, we have proved $\vartheta(G) \leq \overline{\chi}(G)$ by using the following formulation for the clique cover number:

$$\begin{aligned} \overline{\chi}(G) = \quad \min \quad & \sum_{C \text{ clique}} \lambda_C \\ \text{s.t.} \quad & \sum_{C \text{ clique}} \lambda_C \chi^C = e \\ & \lambda_C \in \{0, 1\} \text{ for each clique } C. \end{aligned} \tag{3.11}$$

Note that $\lambda$ is indexed by the cliques of $G$, and observe that the number of cliques can be exponential in $|V|$. Thus the above program is an example of an integer linear program with an exponential number of variables. By relaxing the discrete variable domain $\{0, 1\}$ to the interval $[0, 1]$ we obtain a lower bound on $\overline{\chi}(G)$ called the *fractional clique cover number* of $G$:

$$\begin{aligned} \overline{\chi^*}(G) := \quad \min \quad & \sum_{C \text{ clique}} \lambda_C & = \quad \max \quad & e^T x \\ \text{s.t.} \quad & \sum_{C \text{ clique}} \lambda_C \chi^C = e & \text{s.t.} \quad & x(C) \leq 1 \; (C \text{ clique}) \\ & \lambda \geq 0 & & x \in \mathbb{R}^V_+. \end{aligned} \tag{3.12}$$

Although the fractional clique cover number is defined to be the optimum of the linear program above, it is known to be NP-hard (see [35]). The difficulty is caused by the fact that the constraints in the 'primal' program (3.12) are not polynomial time checkable. For more details about $\overline{\chi}^*(G)$ consult e.g. [84].

A comparison between $\vartheta(G)$ and $\overline{\chi}^*(G)$ can be derived by adjusting the proof for $\vartheta(G) \leq \overline{\chi}(G)$.

**Proposition 3.2.2.** *For any graph $G = (V, E)$ one has $\vartheta(G) \leq \overline{\chi}^*(G)$.*

*Proof.* Take an optimal $\lambda$ for the program (3.12), set

$$Z' := \sum_{C \text{ clique}} \lambda_C \begin{pmatrix} 1 \\ \chi^C \end{pmatrix} \begin{pmatrix} 1 \\ \chi^C \end{pmatrix}^T, \tag{3.13}$$

and observe that $Z'$ is feasible for (3.4) with $Z'_{\mathbf{00}} = \overline{\chi}^*(G)$. □

The fractional clique cover number of $\overline{G}$ is called the *fractional chromatic number* of $G$ and it is denoted by $\chi^*(G)$. Namely,

$$\begin{array}{rcll} \chi^*(G) := \quad \overline{\chi}^*(\overline{G}) & = & \min & \displaystyle\sum_{S \text{ stable set}} \lambda_S \\ & & \text{s.t.} & \displaystyle\sum_{S \text{ stable set}} \lambda_S \chi^S = e \\ & & & \lambda \geq 0. \end{array} \tag{3.14}$$

We can now summarize the results described in this section. For any graph $G$ we have

$$\alpha(G) \leq \vartheta(G) \leq \overline{\chi}^*(G) \leq \overline{\chi}(G), \tag{3.15}$$

and equivalently

$$\omega(G) \leq \vartheta\left(\overline{G}\right) \leq \chi^*(G) \leq \chi(G), \tag{3.16}$$

after recalling $\omega(G) = \alpha\left(\overline{G}\right)$ and $\chi(G) = \overline{\chi}\left(\overline{G}\right)$.

**Perfect graphs.** For a perfect graph all the inequalities in (3.15) and (3.16) are equalities. Therefore, the stability number of a perfect graph $G$ can be determined by computing an approximated value for $\vartheta(G)$. It can be done by solving, for instance, the SDP (3.1) with precision $< \frac{1}{2}$, and then rounding the objective value to the nearest integer. Similarly, determining the chromatic number of a perfect graph $G$ can be done by computing $\vartheta\left(\overline{G}\right)$. Besides these numerical values, one can find a stable set of the maximum size and an optimal colouring of a perfect graph in polynomial time using the Lovász theta number (for details see [35]).

On the other hand, all the inequalities in (3.15) and (3.16) can be strict. The smallest graph for which this happens is $C_5$, since $\alpha(C_5) = 2$, $\vartheta(C_5) = \sqrt{5}$, $\overline{\chi}^*(C_5) = \frac{5}{2}$ and $\overline{\chi}(C_5) = 3$, and since all graphs smaller than $C_5$ are perfect. Nevertheless, the values $\vartheta(G) - \alpha(G)$, $\overline{\chi}^*(G) - \vartheta(G)$ and $\overline{\chi}(G) - \overline{\chi}^*(G)$ can be arbitrarily large. To see this take, let $G$ be the union of $k$ disjoint copies of $C_5$. We have

- $\alpha(G) = k\alpha(C_5) = 2k$;
- $\vartheta(G) = k\vartheta(C_5) = k\sqrt{5}$;
- $\overline{\chi}^*(G) = k\overline{\chi}^*(C_5) = \frac{5}{2}k$;

- $\overline{\chi}(G) = k\overline{\chi}(C_5) = 3k$;

since the parameters are additive with respect to the direct sum of graphs. Knuth [54] proves additivity for $\vartheta$, while for $\alpha, \overline{\chi^*}$ and $\overline{\chi}$ it follows directly from the definitions.

## 3.3 Nonnegativity and triangle inequalities

Several variations of the Lovász theta number were proposed in order to get sharper upper bounds for the stability number of a graph and sharper lower bounds for the chromatic number of a graph.

Given a graph $G = (V, E)$, $|V| \neq 0$, set $n := |V|$. McEliece, Rodemich, and Rumsey [68] and Schrijver [83] give the bound:

$$\begin{array}{llll} \vartheta'(G) := & \max & \langle \mathbf{J}, X\rangle & = \max & \langle \mathbf{J}, X\rangle \\ & \text{s.t.} & \mathrm{Tr}(X) = 1 & \quad\ \text{s.t.} & \mathrm{Tr}(X) = 1 \\ & & X_{ij} = 0 \ \ (ij \in E(G)) & & X_{ij} = 0 \ \ (ij \in E(G)) \\ & & X \succeq 0, X \geq 0 & & X \in \mathcal{D}_n, \end{array} \tag{3.17}$$

where $X$ is indexed by $V$.

Schrijver [83] proves the relation of $\vartheta'(G)$ with the linear programming bound for codes of Delsarte [21].

Among equivalent formulations for $\vartheta'(G)$ we often use the following:

$$\begin{array}{llll} \vartheta'(G) = & \max & \sum_i X'_{ii} & = \max & \sum_i X'_{ii} \\ & \text{s.t.} & X'_{\mathbf{00}} = 1 & \quad\ \text{s.t.} & X'_{\mathbf{00}} = 1 \\ & & X'_{ij} = 0 \ \ (ij \in E(G)) & & X'_{ij} = 0 \ \ (ij \in E(G)) \\ & & X' \succeq 0, X' \geq 0 & & X' \in \mathcal{D}_{n+1}, \end{array} \tag{3.18}$$

where $X'$ is indexed by $\mathcal{P}_{\leq 1}(V)$. In order to see the link between (3.17) and (3.18) one should follow the steps given in Section 3.1.

Szegedy [90] presented a sharper bound for the fractional clique cover number:

$$\begin{array}{lll} \vartheta^+(G) := & \max & \langle \mathbf{J}, X\rangle \\ & \text{s.t.} & \mathrm{Tr}(X) = 1 \\ & & X_{ij} \leq 0 \ \ (ij \in E(G)) \\ & & X \succeq 0. \end{array} \tag{3.19}$$

The formulation for $\vartheta^+(G)$ which corresponds to (3.4) is:

$$\begin{array}{llll} \vartheta^+(G) = & \min & Z'_{\mathbf{00}} & = \min & Z'_{\mathbf{00}} \\ & \text{s.t.} & Z'_{ii} = Z'_{\mathbf{0}i} \ \ (i \in V) & \quad\ \text{s.t.} & Z'_{ii} = Z'_{\mathbf{0}i} \ \ (i \in V) \\ & & Z'_{ii} = 1 \ \ (i \in V) & & Z'_{ii} = 1 \ \ (i \in V) \\ & & Z'_{ij} = 0 \ \ (ij \in E(\overline{G})) & & Z'_{ij} = 0 \ \ (ij \in E(\overline{G})) \\ & & Z' \succeq 0, \ Z' \geq 0 & & Z' \in \mathcal{D}_{n+1}, \end{array} \tag{3.20}$$

where $Z'$ is indexed by $\mathcal{P}_{\leq 1}(V)$.

The bounds $\vartheta'(G)$ and $\vartheta^+(G)$ were further strengthened by adding triangle inequalities. Dukanovic and Rendl [25] define:

$$\begin{array}{rll} \vartheta'^{\triangle}(G) := \max & \langle \mathbf{J}, X\rangle & \\ \text{s.t.} & \mathrm{Tr}(X) = 1 & \\ & X_{ij} = 0 & (ij \in E(G)) \\ & X_{ij} \le X_{ii} & (i,j \in V) \\ & X_{ik} + X_{jk} \le X_{ij} + X_{kk} & (i,j,k \in V) \\ & X \succeq 0, X \ge 0. & \end{array} \tag{3.21}$$

By adding triangle inequalities in (3.20) Meurdesoif [70] defines:

$$\begin{array}{rll} \vartheta^{+\triangle}(G) = \min & Z'_{\mathbf{00}} & \\ \text{s.t.} & Z'_{ii} = Z'_{\mathbf{0}i} & (i \in V) \\ & Z'_{ii} = 1 & (i \in V) \\ & Z'_{ij} = 0 & (ij \in E(\overline{G})) \\ & Z'_{ij} + Z'_{jk} - Z'_{ki} \le 1 & (i,j,k \in V) \\ & Z' \succeq 0, \ Z' \ge 0, & \end{array} \tag{3.22}$$

where $Z'$ is again indexed by $\mathcal{P}_{\le 1}(V)$.

The parameters defined above satisfy

$$\alpha(G) \le \vartheta'^{\triangle}(G) \le \vartheta'(G) \le \vartheta(G) \le \vartheta^+(G) \le \vartheta^{+\triangle}(G) \le \overline{\chi^*}(G),$$

The last inequality follows from the fact that the matrix defined in (3.13) is feasible for the program (3.22) defining $\vartheta^{+\triangle}(G)$. For proving the first inequality take a maximum size stable set $S$ in $G$ and set $X := \frac{1}{\alpha(G)}\chi^S\left(\chi^S\right)^T$. The matrix $X$ is feasible for (3.21) and $\langle \mathbf{J}, X\rangle = \alpha(G)$.

The links between the bounds defined for a graph $G$ and for its complement $\overline{G}$ are given in the following theorem.

**Theorem 3.3.1.** *For any graph $G$*

*(a)* $\alpha(G)\chi^*(G) \ge |V(G)|$,

*(b)* $\vartheta(G)\overline{\vartheta}(G) \ge |V(G)|$,

*(c)* $\vartheta'(G)\vartheta^+\left(\overline{G}\right) \ge |V(G)|$,

*(d)* $\vartheta'^{\triangle}(G)\vartheta^{+\triangle}\left(\overline{G}\right) \ge |V(G)|$.

*Moreover, the equality holds in (a),(b),(c) and (d) if $G$ is vertex transitive.*

*Proof.* (a) Take an optimal solution $\lambda$ for the program (3.14) defining $\chi^*(G)$. Now $\alpha(G)\chi^*(G) = \alpha(G)\sum_{S \text{ stable}}\lambda_S \ge \sum_{S \text{ stable}}\lambda_S|S| = \sum_{S \text{ stable}}\lambda_S e^T\chi^S = e^Te = |V(G)|$.

Assume that $G$ is vertex transitive and let $S$ be a stable set in $G$ with $|S| = \alpha(G)$. We have $\sum_{\sigma\in \mathrm{Aut}(G)}\chi^{\sigma(S)} = k_Se$ for some $k_S \in \mathbb{R}_+\backslash\{0\}$. Note next that $\sigma(S)$ is stable for all $\sigma \in \mathrm{Aut}(G)$, which yields $\chi^*(G) \le \frac{|\mathrm{Aut}(G)|}{k_s}$. On the other hand we also obtain $k_S|V(G)| = |\mathrm{Aut}(G)|\alpha(G)$, yielding $\frac{|V(G)|}{\alpha(G)} \ge \chi^*(G)$.

(b) Let $Z'$ be optimal for the program defining $\overline{\vartheta}(G)$ (see (3.4)). Then $\frac{1}{\overline{\vartheta}(G)} Z'$ is feasible for (3.8). Now $Z'_{ii} = 1$ $(i \in V(G))$ yields $\vartheta(G) \geq \frac{1}{\overline{\vartheta}(G)}|V(G)|$.

Assume that $G$ is vertex transitive. As the program (3.8) is invariant under action of $\mathrm{Aut}(G)$, we can restrict $X'$ in (3.8) to be invariant under action of $\mathrm{Aut}(G)$. Let now matrix $X'$ be invariant under action of $\mathrm{Aut}(G)$ and optimal for (3.8). Since $X'_{ii} = \frac{\vartheta(G)}{n}$ as $G$ is vertex transitive, $\frac{n}{\vartheta(G)} X'$ is feasible for (3.4) defining $\overline{\vartheta}(G)$. Therefore $\overline{\vartheta}(G) \leq \frac{n}{\vartheta(G)}$.

The proofs for (c) and (d) are analogous to the one of (b). □

The relation (b) was proven in [64], (c) in [90], while [25] contains the proof for (d). We will see in Chapter 5 how this theorem can be generalized to more graph parameters which are of interest in this thesis.

**Negative results.** Dukanovic and Rendl [25] compute the bounds defined above for several graph classes. In fact, they test if adding nonnegativity and triangle constraints leads to stronger bounds. Analyzing computational results they observe the following:

(i) on random graphs (see Section 6.4) adding nonnegativity or triangle inequalities does not improve the Lovász theta number considerably (only in the order of decimals);

(ii) on some vertex transitive graphs the nonnegativity constraints might lead to substantial improvements over the Lovász theta number (see Section 6.2 for some examples), whereas additional inclusion of the triangle constraints often does not give any improvements.

We prove below two negative results about adding additional constraints. We apply the following result in Section 6.1.

**Proposition 3.3.2.** *If $G$ is vertex transitive and for any pair of edges $ij$, $i'j'$ there exist $\sigma \in \mathrm{Aut}(G)$ such that $\sigma(\{i,j\}) = \{i',j'\}$ then $\vartheta'\left(\overline{G}\right) = \vartheta\left(\overline{G}\right)$.*

*Proof.* Since the program (3.1) defining the parameter $\vartheta\left(\overline{G}\right)$ is invariant under the action of $\mathrm{Aut}(G)$, we can assume that the matrix variable $X$ is invariant under the action of $\mathrm{Aut}(G)$. If $G$ is vertex transitive and $E(G) = \{\{\sigma(i),\sigma(j)\} \mid \sigma \in \mathrm{Aut}(G)\}$ the program (3.1) reads

$$\vartheta\left(\overline{G}\right) = \max \langle \mathbf{J}, X\rangle \text{ s.t. } X = \frac{1}{|V(G)|}\mathbf{I} + xA_G \succeq 0,\ X \in \mathbb{R}^{V(G)\times V(G)},\ x \in \mathbb{R}.$$

Since $X := \frac{1}{|V(G)|}\mathbf{I}$ is feasible, any optimal solution satisfies $x \geq 0$. This proves $\vartheta'\left(\overline{G}\right) = \vartheta\left(\overline{G}\right)$. □

We explain now why adding triangle constraints in a program defining the Lovász theta number of a Hamming graph does not give any improvements.

Given an integer $n \geq 1$ and $\mathcal{D} \subseteq N := \{1,\ldots,n\}$, the *Hamming graph* $H(n,\mathcal{D})$ is the graph $G$ with node set $V(G) := \mathcal{P}(N)$ and with an edge $(I,J)$ if $|I \triangle J| := |I\backslash J| + |J\backslash I| \in \mathcal{D}$ (for $I,J \in \mathcal{P}(N)$).

**Proposition 3.3.3.** *Let $G := H(n,\mathcal{D})$, where $n \geq 2$. Then $\vartheta'(G) = \vartheta'^{\triangle}(G)$.*

*Proof.* Observe first that each permutation $\sigma \in \mathrm{Sym}(n)$ induces an automorphism of $G$, by letting $\sigma(I) := \{\sigma(i) \mid i \in I\}$ for $I \in \mathcal{P}(N)$. For any $K \in \mathcal{P}(N)$, the *switching mapping* $s_K$ defined by $s_K(I) := I \triangle K$ (for $I \in \mathcal{P}(N)$) is also an automorphism of $G$. Let $X$ be optimal for (3.17) and invariant under action of $\mathrm{Aut}(G)$. It suffices to show that $X$ is feasible for (3.21).

Since $X$ is invariant under action of $\mathrm{Aut}(G)$, an entry $X_{IJ}$ depends only on $I \triangle J$, i.e. the matrix $X$ belongs to the Bose-Mesner algebra $\mathcal{B}_n$. Therefore $X_{II} = X_{JJ}$ $(I, J \in \mathcal{P}(N))$. We also have $X_{IJ} \leq X_{II}$ $(I, J \in \mathcal{P}(N))$, since

$$\begin{pmatrix} X_{II} & X_{IJ} \\ X_{IJ} & X_{JJ} \end{pmatrix} = \begin{pmatrix} X_{II} & X_{IJ} \\ X_{IJ} & X_{II} \end{pmatrix} \succeq 0 \quad (I, J \in \mathcal{P}(N)).$$

It remains to prove $X_{IK} + X_{JK} \leq X_{IJ} + X_{KK}$ $(I, J, K \in \mathcal{P}(N))$. Let $I, J, K \in \mathcal{P}(N)$. Set $L := (I \triangle J) \triangle K$ and consider the submatrix of $X$ indexed by $I, J, K, L$. It is PSD, and since $L \triangle I = J \triangle K$, $L \triangle J = I \triangle K$, $L \triangle K = I \triangle J$, we have $X_{IL} = X_{JK}$, $X_{JL} = X_{IK}$, $X_{KL} = X_{IJ}$. Finally,

$$\begin{aligned} \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}^T & \begin{pmatrix} X_{KK} & X_{IJ} & X_{IK} & X_{JK} \\ X_{IJ} & X_{KK} & X_{JK} & X_{IK} \\ X_{IK} & X_{JK} & X_{KK} & X_{IJ} \\ X_{JK} & X_{IK} & X_{IJ} & X_{KK} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \\ &= X_{IJ} + X_{KK} - X_{IK} - X_{JK} \geq 0. \end{aligned}$$

$\square$

Note that we did not use nonnegativity in the last proof. Namely, the triangle constraints are implied by the positive-semidefiniteness and the membership in the Bose-Mesner algebra.

# Chapter 4

# Semidefinite programming upper bounds for the stability number

Various techniques have been developed in order to formulate and solve (hard) combinatorial optimization problems. A classical approach in polyhedral combinatorics is to identify the set of feasible solutions of a problem with a finite set $F \subseteq \{0,1\}^n$, and then to find the (or at least a partial) linear inequality description of the polytope $P := \mathrm{conv}(F)$. Research has mainly focused on developing methods for strengthening an initial linear relaxation. Among the first were the method of Gomory and its various extensions for generating strong cutting planes (see e.g. [73, 94]). Another approach is to try to represent a polytope $P$ as the projection of another polytope lying in a higher dimension. Techniques based on this idea are usually referred to in the literature as 'lift-and-project' methods.

A common feature of the lift-and-project methods is the construction of a hierarchy $K^1 \supseteq K^2 \supseteq \cdots \supseteq P$ of relaxations of $P$ such that $K^n = P$. Various hierarchies of linear relaxations were proposed, e.g., in Balas, Ceria and Cornuéjols [1], Sherali and Adams [87], Lovász and Schrijver [65]. In [65] the authors also propose the semidefinite variant of their hierarchy obtained by adding semidefinite conditions to the relaxations based on matrix cuts (see Section 4.1.2). Recently, another hierarchy of semidefinite relaxations was proposed by Lasserre [57]. He shows how the 'sums of squares and moment method' technique for polynomial optimization problems (see Section 2.6) can be used for finding $P$. The various methods were compared with each other in [59]. In particular, it is shown in [59] that the hierarchy of Lasserre refines the hierarchy of Lovász and Schrijver.

Another more recent approach to combinatorial optimization problems is based on modelling them as copositive programs. Optimizing over a copositive cone is hard in general, hence the goal is to find a good approximation for it. Motivated by the theorem of Pólya, Parrilo [75] defines a hierarchy of cones that approximate a copositive cone from inside (see section 2.6). His method has been used for approximating the stability number in de Klerk and Pasechnik [50] and the chromatic number in Dukanovic and Rendl [24]. Most recently, Burer [12] gave a recipe for transforming an integer program, having a quadratic objective function and linear constraints, into a completely positive program.

In this chapter we consider both, polyhedral and copositive programming, approaches to the stable set problem.

In Section 4.1, we recall the application of the matrix-cut method [65] and the moment method [57] to the stable set problem. By combining these two methods we design a new block diagonal hierarchy which is computationally less costly than the hierarchy of Lasserre, and moreover it gives at least as good bounds as the hierarchy of Lovász and Schrijver.

Section 4.2 is devoted to the upper bounds for the stable set number obtained from the hierarchy of de Klerk and Pasechnik [50]. We review results about the convergence of these bounds to the stability number and we prove that they are dominated by the bounds in the hierarchy of Lasserre.

Although most of the techniques presented below can be generalized and applied to the weighted stable set problem, we restrict our attention to the case when all weights are set to one, i.e., to the stable set problem.

## 4.1 Hierarchies of relaxations for the stable set polytope

We can not expect in general to find a 'polynomial size' linear description of the stable set polytope $\mathrm{STAB}(G)$ of a graph $G$. Otherwise, we would be able to optimize (a linear function) over it in polynomial time by solving a linear program, and thus determine the stability number which is an NP-hard graph parameter. As an alternative, one can consider linear or semidefinite relaxations of $\mathrm{STAB}(G)$. We deal with them in this section.

We first review some classical relaxations of $\mathrm{STAB}(G)$ in Subsection 4.1.1. In Subsection 4.1.2, we apply the operators $N$ and $N_+$ proposed by Lovász and Schrijver in [65] to these relaxations.

We have seen in Section 2.6 how the hierarchy of Lasserre [56, 57], which is based on moment matrices, can be applied to combinatorial optimization problems in general, and to the stable set problem in particular. We recall this hierarchy in Subsection 4.1.3.

We define a new hierarchy as a relaxation of the hierarchy of Lasserre in Subsection 4.1.4. Due to its 'block diagonalization' property, it is at least as good as the hierarchy of Lovász and Schrijver and computationally less expensive.

### 4.1.1 The stable set polytope and its classical relaxations

Given a graph $G = (V, E)$, define

$$F := \{\chi^S \mid S \text{ stable in } G\} = \{x \in \{0,1\}^V \mid x_i + x_j \le 1 \ (ij \in E)\}.$$

The stable set polytope is

$$\mathrm{STAB}(G) := \mathrm{conv}(F). \tag{4.1}$$

Well known linear relaxations are the fractional stable set polytope

$$\mathrm{FRAC}(G) := \{x \in \mathbb{R}_+^V \mid x_i + x_j \le 1 \ (ij \in E),\ x_i \le 1 \ (i \in V)\}, \tag{4.2}$$

and its strengthening

$$\mathrm{QSTAB}(G) := \{x \in \mathbb{R}_+^V \mid x(C) \le 1 \ (C \text{ clique in } G)\}, \tag{4.3}$$

obtained by adding the clique inequalities. Next we define the even stronger relaxation

$$\mathrm{THETA}(G) := \{x \in \mathbb{R}^V \mid x = \mathrm{diag}(X);\ X - xx^T \succeq 0;\ X_{ij} = 0 \ (ij \in E)\}, \tag{4.4}$$

called the *theta body*. We have

$$\mathrm{STAB}(G) \subseteq \mathrm{THETA}(G) \subseteq \mathrm{QSTAB}(G) \subseteq \mathrm{FRAC}(G). \tag{4.5}$$

To prove the first inclusion let $x \in \mathrm{STAB}(G)$. Then $x = \sum_{i=1}^k \lambda_i \chi^{S_i}$, for some stable sets $S_i$, where $\lambda \ge 0$ is such that $e^T\lambda = 1$. Set $X' := \sum_{i=1}^k \lambda_i \begin{pmatrix} 1 \\ \chi^{S_i} \end{pmatrix} \begin{pmatrix} 1 \\ \chi^{S_i} \end{pmatrix}^T$ and observe that $X'$ has the form

$$X' = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}, \tag{4.6}$$

where $\mathrm{diag}(X) = x$, and $X_{ij} = 0$ for all $ij \in E$. Since $X' \succeq 0 \iff X - xx^T \succeq 0$ we get $x \in \mathrm{THETA}(G)$.

For the second inclusion let $x \in \mathrm{THETA}(G)$ be such that $x = \mathrm{diag}(X)$ for some $X \in \mathbb{R}^{V \times V}$, satisfying $X - xx^T \succeq 0$ and $X_{ij} = 0$ $(ij \in E)$, and let $C$ be a clique in $G$. Now $(\chi^C)^T(X - xx^T)\chi^C = x(C) - (x(C))^2 \ge 0$ yielding $x(C) \le 1$. Therefore, $x \in \mathrm{QSTAB}(G)$.

**Proposition 4.1.1.** *For any graph $G$ we have*

- $\alpha(G) = \max e^T x$ s.t. $x \in \mathrm{STAB}(G)$,
- $\vartheta(G) = \max e^T x$ s.t. $x \in \mathrm{THETA}(G)$,
- $\overline{\chi^*}(G) = \max e^T x$ s.t. $x \in \mathrm{QSTAB}(G)$.

*Proof.* The first item is obvious and the second is equivalent to (3.8). For the third write the dual program of (3.12):

$$\overline{\chi^*}(G) = \max e^T x \ \text{ s.t. } x \in \mathbb{R}^V,\ x(C) \le 1 \ (C \text{ clique in } G). \tag{4.7}$$

Now, if $x$ is feasible for (4.7) then $\widetilde{x} \in \mathbb{R}_+^V$, defined by $\widetilde{x}_i := \max\{0, x_i\}$ $(i \in V)$, is also feasible, and $e^T\widetilde{x} \ge e^T x$. Thus, the set of optimal solutions of (4.7) remains unchanged if we add the condition $x \in \mathbb{R}_+^V$. $\square$

For more details about the stable set polytope and its basic relaxations we recommend Gröschel, Lovász, Schrijver [35]. One of the most important results there is that the stability number of any perfect graph can be found in polynomial time.

**Theorem 4.1.2.** *[35] A graph $G$ is perfect if and only if*

$$\mathrm{STAB}(G) = \mathrm{QSTAB}(G).$$

Thus, if $G$ is perfect then $\mathrm{STAB}(G) = \mathrm{THETA}(G)$, which implies $\alpha(G) = \vartheta(G)$. The parameter $\vartheta(G)$ can be computed then to $\epsilon$ (e.g. for $\epsilon = \frac{1}{2}$) precision in polynomial time, hence we get $\alpha(G)$ by rounding an obtained solution to the nearest integer.

### 4.1.2 The Lovász-Schrijver hierarchy

Let $V$ be a finite set. Set $n := |V|$. Let $K \subseteq \mathbb{R}^{\mathcal{P}_{\leq 1}(V)}$ be a convex cone contained in the homogeneous unit cube $Q := \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid 0 \leq x_i \leq x_{\mathbf{0}} \ (i \in V)\}$. The objective is to find the complete linear inequality description of the polytope

$$P := \mathrm{conv}\left(x \in \{0,1\}^V \mid \begin{pmatrix} 1 \\ x \end{pmatrix} \in K\right) \subseteq [0,1]^V \tag{4.8}$$

or, equivalently, of the cone $\tilde{P} := \mathbb{R}_+ \left(\begin{pmatrix} 1 \\ x \end{pmatrix} \in K \text{ with } x \in \{0,1\}^V\right)$. Set

$$M_V := \{Y \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V) \times \mathcal{P}_{\leq 1}(V)} \mid Y_{ij} = Y_{ji} \ (i,j \in V),\ Y_{jj} = Y_{\mathbf{0}j} = Y_{j\mathbf{0}} \ (j \in V)\}, \tag{4.9}$$

$$M_{+,V} := \{Y \in M_V \mid Y \succeq 0\} \tag{4.10}$$

and, following Lovász and Schrijver [65], define

$$M(K) := \{Y \in M_V \mid Ye_k,\ Y(e_{\mathbf{0}} - e_k) \in K \text{ for } k \in V\}, \tag{4.11}$$

$$N(K) := \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid x = Ye_{\mathbf{0}} \text{ for some } Y \in M(K)\}, \tag{4.12}$$

$$M_+(K) := \{Y \in M_{+,V} \mid Ye_k,\ Y(e_{\mathbf{0}} - e_k) \in K \text{ for } k \in V\}, \tag{4.13}$$

$$N_+(K) := \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid x = Ye_{\mathbf{0}} \text{ for some } Y \in M_+(K)\}, \tag{4.14}$$

where $e_{\mathbf{0}}, e_k$ ($k \in V$) denote the standard unit vectors in $\mathbb{R}^{\mathcal{P}_{\leq 1}(V)}$. We have $\tilde{P} \subseteq N_+(K) \subseteq N(K) \subseteq K$. The inclusion $\tilde{P} \subseteq N_+(K)$ follows from the fact that, for $\begin{pmatrix} 1 \\ x \end{pmatrix} \in K$ with $x \in \{0,1\}^V$, the matrix $Y := \begin{pmatrix} 1 \\ x \end{pmatrix}\begin{pmatrix} 1 \\ x \end{pmatrix}^T$ belongs to $M_+(K)$ and the inclusion $N(K) \subseteq K$ follows from the definitions (4.11) and (4.12). Define iteratively $N^t(K) := N(N^{t-1}(K))$ and $N_+^t(K) := N_+(N_+^{t-1}(K))$ for $t \geq 2$, setting $N^1(K) := N(K)$ and $N_+^1(K) := N_+(K)$. Then,

$$\tilde{P} \subseteq N^{t+1}(K) \subseteq N^t(K) \subseteq K, \text{ and } \tilde{P} \subseteq N_+^{t+1}(K) \subseteq N_+^t(K) \subseteq K \text{ for } t \geq 1.$$

Lovász and Shrijver [65] showed that the $N$ and $N_+$ operators have a nice algorithmic property. Namely, for any fixed $t$, if one can optimize over $K$ in polynomial time then the same holds for both $N^t(K)$ and $N_+^t(K)$. This fact is a consequence of Proposition 4.1.3 below. Notwithstanding, Lovász and Schrijver also prove that their hierarchies converge in $n$ steps, i.e. $N^t(K) = N_+^t(K) = \tilde{P}$ for $t \geq n$.

**Semidefinite programming formulation**

The above definitions for $N^t(K)$ and $N_+^t(K)$ are recursive. We now 'unfold the recursion' and give an explicit semidefinite programming formulation for $N_+^t(K)$. The analog result holds for $N^t(K)$.

**Proposition 4.1.3.** *[40] Let $x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)}$ and $t \geq 1$. Then, $x \in N_+^t(K)$ if and only if there exists $Y^{(0)} \in M_{+,V}$ with $Y^{(0)}e_{\mathbf{0}} = x$ and, for all $s \in \{1, \ldots, t-1\}$, $k_1, \ldots, k_s \in V$, $\sigma \in \{\pm 1\}^s$, there exist $Y^{(\sigma_1 k_1, \ldots, \sigma_s k_s)} \in M_{+,V}$ satisfying*

$$Y^{(\sigma_1 k_1,\ldots,\sigma_s k_s)}e_{\mathbf{0}} = Y^{(\sigma_1 k_1,\ldots,\sigma_{s-1}k_{s-1})}(e_{\mathbf{0}}(1-\sigma_s)/2 + e_{k_s}\sigma_s) \tag{4.15}$$

*and, for all $k_1, \ldots, k_t \in V$, $\sigma \in \{\pm 1\}^t$,*

$$Y^{(\sigma_1 k_1,\ldots,\sigma_{t-1}k_{t-1})}(e_{\mathbf{0}}(1-\sigma_t)/2 + e_{k_t}\sigma_t) \in K \tag{4.16}$$

*(setting $Y^{(\sigma_1 k_1,\ldots,\sigma_s k_s)} := Y^{(0)}$ if $s = 0$).*

*Proof.* We use the induction on $t$. For $t = 1$ the proposition is equivalent to (4.14). Suppose now that the claim holds for every convex cone contained in $Q$ and some $t \geq 1$. Given such a cone $L$, set $K := N_+(L)$. Then $x \in N_+^{t+1}(L) = N_+^t(K)$ if and only if there exist matrices $Y^{(\sigma_1 k_1,\ldots,\sigma_s k_s)}$ (as described above) satisfying (4.15) and (4.16). The condition (4.16) and $K = N_+(L)$ imply that for all $k_1, \ldots, k_t \in V$ and $\sigma \in \{\pm 1\}^t$ there exist matrices $Y^{(\sigma_1 k_1,\ldots,\sigma_t k_t)} \in M_{+,V}$ such that

$$Y^{(\sigma_1 k_1,\ldots,\sigma_t k_t)}e_{\mathbf{0}} = Y^{(\sigma_1 k_1,\ldots,\sigma_{t-1}k_{t-1})}(e_{\mathbf{0}}(1-\sigma_t)/2 + e_{k_t}\sigma_t),$$

and, for all $k_1, \ldots, k_{t+1} \in V$ and $\sigma \in \{\pm 1\}^{t+1}$,

$$Y^{(\sigma_1 k_1,\ldots,\sigma_t k_t)}(e_{\mathbf{0}}(1-\sigma_{t+1})/2 + e_{k_{t+1}}\sigma_{t+1}) \in L.$$

$\square$

**Application to the stable set problem**

In [65] the operators $N$ and $N_+$ are applied to the stable set problem. To find the linear description of the stable set polytope STAB($G$) of a graph $G = (V, E)$ or the corresponding cone

$$\mathrm{ST}(G) := \left\{ \lambda \begin{pmatrix} 1 \\ x \end{pmatrix} \;\middle|\; x \in \mathrm{STAB}(G),\ \lambda \in \mathbb{R}_+ \right\},$$

one should apply the $N$, or $N_+$, operator to one of its linear, or semidefinite, relaxations. For example to FRAC($G$), QSTAB($G$) or THETA($G$), or more precisely, to their corresponding cones

$$\mathrm{FR}(G) := \{x \in \mathbb{R}_+^{\mathcal{P}_{\leq 1}(V)} \mid x_i + x_j \leq x_{\mathbf{0}}\ (ij \in E),\ x_i \leq x_{\mathbf{0}}\ (i \in V)\},$$

$$\mathrm{QST}(G) := \{x \in \mathbb{R}_+^{\mathcal{P}_{\leq 1}(V)} \mid x(C) \leq x_{\mathbf{0}}\ (C \text{ clique in } G)\},$$

$$\mathrm{TH}(G) := \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid x = Ye_{\mathbf{0}} \text{ for some } Y \in M_V \text{ with } Y_{ij} = 0\ (ij \in E)\}.$$

From (4.5) we have

$$\mathrm{ST}(G) \subseteq \mathrm{TH}(G) \subseteq \mathrm{QST}(G) \subseteq \mathrm{FR}(G).$$

One can verify (see [65]) that $N_+(\mathrm{FR}(G)) \subseteq \mathrm{TH}(G)$, hence applying the $N_+$ operator to $\mathrm{FR}(G)$ yields a relaxation of $\mathrm{ST}(G)$ which is already better than $\mathrm{TH}(G)$. For instance, when $G$ is an odd circuit, $N_+(\mathrm{FR}(G)) = \mathrm{ST}(G)$ is a strict subset of $\mathrm{TH}(G)$. Lovász and Schrijver [65] also prove

$$N^t(\mathrm{FR}(G)) = \mathrm{ST}(G) \quad \text{for } t \geq n - \alpha(G) - 1.$$

$$N_+^t(\mathrm{FR}(G)) = \mathrm{ST}(G) \quad \text{for } t \geq \alpha(G).$$

We prove a bit stronger result for $\mathrm{QST}(G)$. Namely, applying the $N_+$ operator to the relaxation $\mathrm{QST}(G)$ instead of $\mathrm{FR}(G)$, the stable set polytope is found in $\alpha(G) - 1$ steps.

**Theorem 4.1.4.** *$N_+^t(\mathrm{QST}(G)) = \mathrm{ST}(G)$ for any $t \geq \alpha(G) - 1$.*

In the proof we use the following result.

**Theorem 4.1.5.** *(Goemans and Tunçel [33, Th. 3.6]) Let*

$$K := \left\{ \begin{pmatrix} x_{\mathbf{0}} \\ x \end{pmatrix} \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \;\middle|\; Ax \leq bx_{\mathbf{0}},\ 0 \leq x \leq x_{\mathbf{0}} e \right\},$$

*where $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^m$. Let $a \in \mathbb{R}_+^V$, $c \in \mathbb{R}$ and $I_+ := \{i \in V \mid a_i > 0\}$. Assume that the inequality $a^T x \leq c x_{\mathbf{0}}$ is valid for the set $K \cap \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid x_i = x_{\mathbf{0}}\ \forall i \in I\}$ for all sets $I \subseteq I_+$ satisfying*

*(a) $|I| = t$, or*

*(b) $|I| \leq t - 1$ and $\sum_{i \in I} a_i > c$.*

*Then, the inequality $a^T x \leq c x_{\mathbf{0}}$ is valid for $N_+^t(K)$.*

*Proof of Theorem 4.1.4.* We show that $N_+^t(\mathrm{QST}(G)) \subseteq \mathrm{ST}(G)$ for $t := \alpha(G) - 1$. For this consider an inequality $a^T x \leq c x_{\mathbf{0}}$ which is valid for $\mathrm{ST}(G)$. Set $I_+ := \{i \in V \mid a_i > 0\}$ (note $a \geq 0$), $K := \mathrm{QST}(G)$, and $K_I := K \cap \{x \mid x_i = x_{\mathbf{0}}\ \forall i \in I\}$ for $I \subseteq V$. Consider a set $I \subseteq I_+$ satisfying (a) or (b) in the Theorem 4.1.5.

Assume first that $I$ is not stable in $G$. Let $i \neq j \in I$ with $ij \in E$. For $x \in K_I$ we have $x_i = x_j = x_{\mathbf{0}}$ and $x_i + x_j \leq x_{\mathbf{0}}$, implying $x_i = x_j = x_{\mathbf{0}} = 0$. Thus $K_I = \{0\}$ and the inequality $a^T x \leq c x_{\mathbf{0}}$ is trivially valid for $K_I$.

Assume now that $I$ is stable in $G$. Therefore, $a(I) \leq b$ and thus (a) applies, i.e., $|I| = t = \alpha(G) - 1$. Set $S := \{i \in V \setminus I \mid I \cup \{i\}$ is stable in $G\}$ and $T := V \setminus (I \cup S)$. Then, $a_i \leq c - a(I)$ for any $i \in S$ (since $I \cup \{i\}$ is stable in $G$) and $S$ is a clique in $G$ (since, if $i \neq j \in S$ are not adjacent, then $I \cup \{i, j\}$ is a stable set of size $\alpha(G) + 1$). Let $x \in K_I$; we show that $a^T x \leq c x_{\mathbf{0}}$. We have $x(S) \leq x_{\mathbf{0}}$ (as $S$ is a clique) and $x_i = 0$ for all $i \in T$ (as $i \in T$ is adjacent to some $j \in I$ and thus $x_i + x_j \leq x_{\mathbf{0}}$ with $x_j = x_{\mathbf{0}}$, implying $x_i = 0$). Therefore, $a^T x = \sum_{i \in I} a_i x_i + \sum_{i \in S} a_i x_i \leq a(I) x_{\mathbf{0}} + (c - a(I)) x(S) \leq a(I) x_{\mathbf{0}} + (b - a(I)) x_{\mathbf{0}} = c x_{\mathbf{0}}$. ☐

Thus we have:

$$\mathrm{ST}(G) = N_+^{\alpha(G)-1}(\mathrm{TH}(G)) \subseteq N_+^{\alpha(G)-1}(\mathrm{QST}(G)) \subseteq N_+^{\alpha(G)-1}(\mathrm{FR}(G)). \quad (4.17)$$

The right most inclusion is strict, e.g., when $G$ is the line graph of a complete graph with an odd number of nodes (see [89]).

### 4.1.3 The Lasserre hierarchy

Recall first the definition of a combinatorial moment matrix from Section 2.6. Given an integer $t$ and a sequence $y \in \mathbb{R}^{\mathcal{P}_{\leq 2t}(V)}$ the combinatorial moment matrix $M_t(y) \in \mathbb{R}^{\mathcal{P}_{\leq t}(V)\times\mathcal{P}_{\leq t}(V)}$ is defined by

$$M_t(y)_{I,J} := y_{I\cup J} \ (I, J \in \mathcal{P}_{\leq t}(V)).$$

For a given graph $G = (V, E)$ set $n := |V|$ and

$$Q_t(G) := \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid \begin{array}{l} \exists y \in \mathbb{R}^{\mathcal{P}_{\leq 2t}(V)} \text{ satisfying} \\ y_{\mathbf{0}} = x_{\mathbf{0}},\ y_i = x_i \ (i \in V) \\ y_{ij} = 0 \ (ij \in E),\ M_t(y) \succeq 0\}. \end{array} \tag{4.18}$$

Note that $Q_1(G) = \mathrm{TH}(G)$. In Section 2.6 we proved (see (2.50) and the comment below) that $\mathrm{ST}(G) \subseteq Q_{t+1}(G) \subseteq Q_t(G) \subseteq Q_1(G) = \mathrm{TH}(G)$ and $\mathrm{ST}(G) = Q_t(G)$ for $t \geq \alpha(G)$.

#### Comparison with the Lovász-Schrijver hierarchy

The hierarchy $Q_t(G)$ in fact refines the hierarchy obtained by applying the $N_+$ operator to the cone $\mathrm{TH}(G)$. Indeed it is shown in [59] (see the next subsection) that

$$Q_t(G) \subseteq N_+(Q_{t-1}(G)) \quad \text{for } t \geq 2, \tag{4.19}$$

which implies $Q_t(G) \subseteq N_+^{t-1}(Q_1(G)) = N_+^{t-1}(\mathrm{TH}(G)) \subseteq N_+^{t-1}(\mathrm{FR}(G))$. Hence $Q_t(G) = \mathrm{ST}(G)$ for $t \geq \alpha(G)$ (see [59]).

Accordingly, the graph parameter

$$\begin{aligned} \mathrm{las}^{(t)}(G) :=& \ \max \textstyle\sum_{i\in V} x_i \quad \text{s.t.} \quad x \in Q_t(G),\ x_{\mathbf{0}} = 1 \\ =& \ \max \textstyle\sum_{i\in V} y_i \quad \text{s.t.} \quad y \in \mathbb{R}^{\mathcal{P}_{\leq 2t}(V)},\ y_{\mathbf{0}} = 1,\ y_{ij} = 0 \ (ij \in E) \\ & \qquad\qquad\qquad\quad\ M_t(y) \succeq 0 \end{aligned} \tag{4.20}$$

satisfies $\mathrm{las}^{(1)}(G) = \vartheta(G)$, $\alpha(G) \leq \mathrm{las}^{(t+1)}(G) \leq \mathrm{las}^{(t)}(G)$, with $\mathrm{las}^{(t)}(G) = \alpha(G)$ if $t \geq \alpha(G)$. In this way one obtains a hierarchy of upper bounds for the stability number, known as Lasserre's hierarchy (see [57, 59]). Note that the program (4.20) is equivalent to (2.50). The computation of the order $t$ parameter $\mathrm{las}^{(t)}(G)$ is via a semidefinite program which involves a matrix of size $O(n^t)$ and $O(n^{2t})$ variables. It is too expensive for nowadays semidefinite programming solvers. In the next section we introduce a more economical variation of this parameter which remains however at least as good as the parameter obtained by optimizing over the relaxation $N_+^{t-1}(\mathrm{TH}(G))$.

### 4.1.4 A new block diagonal hierarchy

In order to define a more economical hierarchy than the hierarchy of Lasserre, the main idea is to consider, instead of the full matrix $M_t(y)$, some suitable principal submatrices of it. Namely, given an integer $t \geq 1$ and a subset $T \subseteq V$ with $|T| = t - 1$, let $M(T; y)$ denote the principal submatrix of $M_t(y)$ indexed by the set

$$\bigcup_{S\subseteq T} \{S, S \cup \{i\} \ (i \in V)\}. \tag{4.21}$$

This is in fact a multiset as we keep the possible repeated occurrences of indices (e.g., $S = S \cup \{i\}$ if $i \in S$). Hence, in order to define the matrices $M(T; y)$ for all subsets $T \subseteq V$ of cardinality $|T| = t-1$, we need only to know the components of $y$ indexed by $\mathcal{P}_{\leq t+1}(V)$. Define

$$L_t(G) := \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid \begin{array}{l} \exists y \in \mathbb{R}^{\mathcal{P}_{\leq t+1}(V)} \text{ satisfying } y_{\mathbf{0}} = x_{\mathbf{0}}, \\ y_i = x_i \ (i \in V), \ y_{ij} = 0 \ (ij \in E), \\ M(T; y) \succeq 0 \ \forall T \subseteq V \text{ with } |T| = t-1\}, \end{array} \tag{4.22}$$

$$\begin{array}{rlll} \ell^{(t)}(G) := & \max \sum_{i \in V} x_i & \text{s.t.} & x \in L_t(G), \ x_{\mathbf{0}} = 1 \\ = & \max \sum_{i \in V} y_i & \text{s.t.} & y \in \mathbb{R}^{\mathcal{P}_{\leq t+1}(V)}, \ y_{\mathbf{0}} = 1, \ y_{ij} = 0 \ (ij \in E), \\ & & & M(T; y) \succeq 0 \ \forall T \subseteq V \text{ with } |T| = t-1. \end{array} \tag{4.23}$$

Thus, $L_1(G) = Q_1(G) = \text{TH}(G)$ and $\ell^{(1)}(G) = \text{las}^{(1)}(G) = \vartheta(G)$. Obviously, for any $t \geq 1$, $Q_t(G) \subseteq L_t(G)$ and $\text{las}^{(t)}(G) \leq \ell^{(t)}(G)$. For $t = 2$, $\ell^{(2)}(G)$ coincides with the parameter $\ell(G)$ introduced in [60] and further considered in [51].

**Comparison with the Lovász-Schrijver hierarchy**

We show next that the analogue of (4.19) holds for the relaxation $L_t(G)$. We first observe that the edge conditions in the definition of $L_t(G)$ imply in fact that all variables indexed by non-stable sets are identically zero.

**Lemma 4.1.6.** *For $y \in \mathbb{R}^{\mathcal{P}_{\leq t+1}(V)}$, $M(T; y) \succeq 0$ ($T \in \mathcal{P}_{=(t-1)}(V)$),*

$$y_{ij} = 0 \ (ij \in E) \Longrightarrow y_I = 0 \ (I \in \mathcal{P}_{\leq t+1}(V) \textit{ not stable in } G).$$

*Proof.* Let $I \in \mathcal{P}_{\leq t+1}(V)$ containing an edge $ij$. Assume $|I| \geq 3$, let $k \in I \setminus \{i, j\}$, and set $T := I \setminus \{j, k\}$. Then, $M(T; y) \succeq 0$ and both sets $\{i, j\}$ and $T \cup \{k\}$ occur in the index set of $M(T; y)$. As the $(ij, ij)$th entry of $M(T; y)$ is $y_{ij} = 0$, it follows that its $(ij, T \cup \{k\})$th entry is also 0, giving $0 = y_{T \cup \{i,j,k\}} = y_I$. $\square$

Note also that for every $T \subseteq V$ we have

$$M(T; y) \succeq 0 \Longleftrightarrow M(T'; y) \succeq 0 \ (T' \subseteq T).$$

The left implication is trivial, and for the right implication observe that $M(T'; y)$ is a principal submatrix of $M(T; y)$ if $T' \subseteq T$.

**Lemma 4.1.7.** *For $t \geq 2$, $L_t(G) \subseteq N_+(L_{t-1}(G))$.*

*Proof.* Consider $x \in L_t(G)$, i.e., $x = (y_{\mathbf{0}}, y_i \ (i \in V))$ where $y \in \mathbb{R}^{\mathcal{P}_{\leq t+1}(V)}$ satisfies $y_{ij} = 0$ $(ij \in E)$ and $M(T; y) \succeq 0$ for all $T \in \mathcal{P}_{\leq t-1}(V)$. We show $x \in N_+(L_{t-1}(G))$. For this consider the matrix $Y := M_1(y)$. As $x = Ye_{\mathbf{0}}$ it suffices to show that $Y \in M_+(L_{t-1}(G))$. We already have $Y \succeq 0$ and $Y_{jj} = Y_{\mathbf{0}j} = y_j$ $(j \in V)$. Remains to show that $z := (y_k, y_{ik} \ (i \in V))$ and $x - z = (y_{\mathbf{0}} - y_k, y_i - y_{ik} \ (i \in V))$ belong to $L_{t-1}(G)$.

For this define $u \in \mathbb{R}^{\mathcal{P}_{\leq t}(V)}$ by $u_I := y_{I \cup \{k\}}$ for $I \in \mathcal{P}_{\leq t}(V)$. Then, $u_{\mathbf{0}} = z_{\mathbf{0}}$, $(y - u)_{\mathbf{0}} = (x - z)_{\mathbf{0}}$, $u_i = z_i$ and $(y - u)_i = (x - z)_i$ for $i \in V$. Moreover, $u_{ij} = y_{\{i,j,k\}} = 0$ if $ij \in E$. Remains to show $M(U; u) \succeq 0$, $M(U; y - u) \succeq 0$ for

$U \in \mathcal{P}_{\leq t-2}(V)$. Set $T := U \cup \{k\}$. The subsets of $T$ are of the form $S$ or $S \cup \{k\}$ for $S \subseteq U$. Hence the index set (4.21) of $M(T;y)$ can be partitioned into $\mathcal{A} \cup \mathcal{B}$, where $\mathcal{A} = \{S, S \cup \{i\} \mid S \subseteq U,\ i \in V\}$ and $\mathcal{B} = \{S \cup \{k\}, S \cup \{i,k\} \mid S \subseteq U,\ i \in V\}$. With respect to this partition, the matrix $M(T;y)$ has the block form

$$M(T;y) = \begin{array}{c c} & \begin{array}{cc} \mathcal{A} & \mathcal{B} \end{array} \\ \begin{array}{c} \mathcal{A} \\ \mathcal{B} \end{array} & \begin{pmatrix} A & B \\ B & B \end{pmatrix} \end{array}.$$

Now, $M(T;y) \succeq 0$ is equivalent to $B,\ A - B \succeq 0$. The condition $B \succeq 0$ means precisely that $M(U;u) \succeq 0$ and $A - B \succeq 0$ means $M(U; y - u) \succeq 0$. □

**Corollary 4.1.8.** *For $t \geq 1$, $\mathrm{ST}(G) \subseteq L_t(G) \subseteq N_+^{t-1}(\mathrm{TH}(G)) \subseteq N_+^{t-1}(\mathrm{FR}(G))$, with equality $L_t(G) = \mathrm{ST}(G)$ for $t \geq \alpha(G)$.*

Note that this implies (4.19) and $Q_t(G) = \mathrm{ST}(G)$ for $t \geq \alpha(G)$.

*Proof.* Directly from Theorem 4.1.4, Lemma 4.1.7, and $L_1(G) = \mathrm{TH}(G)$. □

Gijswijt [32, Chapter 6] studies SDP upper bounds for the stability numbers of some coding graphs (recall the definition of a Hamming graph from Section 3.3). In particular, he computes the maximum $\sum_{i \in V(G)} x_i$ over strengthenings of $L_2(G)$ and $N_+(\mathrm{TH}(G))$ obtained by adding nonnegativity conditions, and reports some instances where the two maximums differ. Thus, we tend to believe that the inclusion $L_t(G) \subseteq N_+^{t-1}(\mathrm{TH}(G))$ can be strict.

### The new hierarchy is 'block diagonal'

The parameter $\ell^{(t)}(G)$ from (4.23) is expressed via a semidefinite program involving the $\binom{n}{t-1}$ matrices $M(T;y)$ (for $T \subseteq V$ with $|T| = t-1$). We now observe that each matrix $M(T;y)$ has a special block structure with symmetries that can be exploited to block-diagonalize it. Recall from (4.21) that $M(T;y)$ is indexed by the set $\cup_{S \subseteq T} \mathcal{A}_S$, setting $\mathcal{A}_S := \{S, S \cup \{i\} \mid i \in V\}$. With respect to this partition of its index set, the matrix $M(T;y)$ has the block form $(M(S,S'))_{S,S' \subseteq T}$, where $M(S,S')$ denotes the submatrix of $M(T;y)$ with row indices in $\mathcal{A}_S$ and column indices in $\mathcal{A}_{S'}$.

**Lemma 4.1.9.** *For $S, S' \subseteq T$, $M(S,S')$ depends only on $S \cup S'$.*

*Proof.* Directly from the definition of $M(T;y)$ as a moment matrix. □

Hence there exist matrices $A_S$ ($S \subseteq T$) indexed by $\mathcal{P}_{\leq 1}(T)$ with the property that $M(S,S') = A_{S \cup S'}$ for all $S, S' \subseteq T$. We can now use Lemma 2.4.5. Namely, for every $T \subseteq V$ we have

$$M(T;y) \succeq 0 \Longleftrightarrow \sum_{S \subseteq S' \subseteq T} (-1)^{|S' \setminus S|} A_{S'} \succeq 0 \text{ for all } S \subseteq T.$$

For instance, for $T = \{1,2\}$, the matrix $M(T;y)$ which has the block form

$$M(\{1,2\};y) = \begin{array}{c c} & \begin{array}{cccc} \mathcal{A}_{\mathbf{0}} & \mathcal{A}_1 & \mathcal{A}_2 & \mathcal{A}_{12} \end{array} \\ \begin{array}{c} \mathcal{A}_{\mathbf{0}} \\ \mathcal{A}_1 \\ \mathcal{A}_2 \\ \mathcal{A}_{12} \end{array} & \begin{pmatrix} A_{\mathbf{0}} & A_1 & A_2 & A_{12} \\ A_1 & A_1 & A_{12} & A_{12} \\ A_2 & A_{12} & A_2 & A_{12} \\ A_{12} & A_{12} & A_{12} & A_{12} \end{pmatrix} \end{array}, \tag{4.24}$$

is PSD if and only if

$$A_{\mathbf{0}} - A_1 - A_2 + A_{12} \succeq 0,\ A_1 - A_{12} \succeq 0,\ A_2 - A_{12} \succeq 0,\ A_{12} \succeq 0. \tag{4.25}$$

### Complexity comparison

Here we compare the sizes of the semidefinite programs one has to solve in order to compute the bounds obtained by optimizing over the relaxations $Q_t(G) \subseteq L_t(G) \subseteq N_+^{t-1}(\mathrm{TH}(G))$. We use the following convention. Consider a semidefinite program involving the linear matrix inequalities $C^{(j)} + \sum_{i=1}^k y_i B_i^{(j)} \succeq 0$ ($j = 1, \ldots, m$), where each $B_i^{(j)}$ has order $n_j$. The size of the $j$th inequality is defined as $n_j^2$ and the total size of the semidefinite program as $\sum_{j=1}^k n_j^2$. The results are summarized in Table 1 below.

The data from Table 1 for $Q_t(G)$ follow directly from (4.18): the semidefinite program defining $Q_t(G)$ involves $|\mathcal{P}_{\leq 2t}(V)| = O(n^{2t})$ variables, $m := |E(G)|$ linear equations, and one linear matrix inequality (LMI) of order $|\mathcal{P}_{\leq t}(V)| = O(n^t)$.

By (4.22), the semidefinite program defining $L_t(G)$ involves $|\mathcal{P}_{\leq t+1}(V)| = \frac{n^{t+1}}{(t+1)!} + O(n^t)$ variables and $m$ linear equations. By Lemma 2.4.5, it involves $\binom{n}{t-1}2^{t-1}$ LMI's: $\sum_{S \subseteq S' \subseteq T}(-1)^{|S' \setminus S|} A_{S'} \succeq 0$ (for $S \subseteq T \subseteq V$, $|T| = t-1$). The $(S,T)$th LMI has size $(n+1)^2$ since all matrices $A_{S'}$ have order $n+1$. Thus the total size is $\binom{n}{t-1}2^{t-1}(n+1)^2 = n^{t+1}\frac{2^{t-1}}{(t-1)!} + O(n^t)$.

The data for $N_+^{t-1}(\mathrm{TH}(G))$ can be evaluated from the following consequence of Proposition 4.1.3.

**Proposition 4.1.10.** *Let $G = (V, E)$ be a graph, $x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)}$ and $t \geq 1$. Then, $x \in N_+^{t-1}(\mathrm{TH}(G))$ if and only if there exists $Y^{(0)} \in M_{+,V}$, $Y^{(0)}e_{\mathbf{0}} = x$, $Y_{ij}^{(0)} = 0$ ($ij \in E$) and, for all $s \in \{1, \ldots, t-1\}$, $k_1, \ldots, k_s \in V$, $\sigma \in \{\pm 1\}^s$, there exist $Y^{(\sigma_1 k_1, \ldots, \sigma_s k_s)}$ satisfying*

$$Y^{(\sigma_1 k_1, \ldots, \sigma_s k_s)} e_{\mathbf{0}} = Y^{(\sigma_1 k_1, \ldots, \sigma_{s-1} k_{s-1})}(e_{\mathbf{0}}(1 - \sigma_s)/2 + e_{k_s}\sigma_s)$$

*and,*

$$Y^{(\sigma_1 k_1, \ldots, \sigma_s k_s)} \in M_{+,V},\ Y_{ij}^{(\sigma_1 k_1, \ldots, \sigma_s k_s)} = 0\ (ij \in E).$$

Setting $h(n,t) := \sum_{s=0}^{t-1}(2n)^s = 2^{t-1}n^{t-1} + O(n^{t-2})$, we see that the formulation for $N_+^{t-1}(\mathrm{TH}(G))$ involves $\binom{n}{2}h(n,t)$ variables, $h(n,t)$ matrices of order[1] $n+1$, and $O(|E(G)|n^{t-1})$ linear conditions.

We can make the following comments. While the relaxation $L_t(G)$ is at least as good as $N_+^{t-1}(\mathrm{TH}(G))$, its computation is less costly. Indeed, it involves less linear constraints, the number of variables and the size of the SDP have the same order of magnitude, but the constant for the leading term is smaller for $L_t$.

[1] One may observe that, for $1 \leq r \leq s$, the $k_r$th column of the matrix $Y^{(\sigma_1 k_1, \ldots, \sigma_s k_s)}$ is identically zero if $\sigma_r = -1$ and it is equal to the 0th column if $\sigma_r = 1$. Thus $Y^{(\sigma_1 k_1, \ldots, \sigma_s k_s)}$ has order $n - s + 1 = O(n)$ (for fixed $s$).

| Opt. over | size of SDP | # variables | # linear constraints |
|---|---|---|---|
| $Q_t(G)$ | $O(n^{2t})$ | $O(n^{2t})$ | $m$ |
| $L_t(G)$ | $\frac{2^{t-1}}{(t-1)!}n^{t+1} + O(n^t)$ | $\frac{n^{t+1}}{(t+1)!} + O(n^t)$ | $m$ |
| $N_+^{t-1}(\mathrm{TH}(G))$ | $2^{t-1}n^{t+1} + O(n^t)$ | $2^{t-2}n^{t+1} + O(n^t)$ | $O(mn^{t-1})$<br>$m$ if $t = 1$ |

Table 1: Complexity comparison of $Q_t(G)$, $L_t(G)$ and $N_+^{t-1}(G)$.

## 4.2 The de Klerk-Pasechnik hierarchy

Solving a copositive program is not an easy task in general since, as it was mentioned earlier, deciding if a given matrix is not copositive is an NP-complete problem. However, one can try to relax the copositive condition, e.g. by using the recipe of Parrilo [75]. He observed that the copositive cone may be approximated by a series of smaller subcones of it. We briefly recall his idea in Subsection 4.2.1.

De Klerk and Pasechnik [50] used this idea to design a hierarchy of upper bounds $\vartheta^{(t)}(G)$, $t \in \mathbb{N}$, for the stability number $\alpha(G)$ of a graph $G$. They conjecture that $\vartheta^{(t)}(G) = \alpha(G)$ for $t \geq \alpha(G) - 1$. In Subsection 4.2.2, we give an overview of their work and we partially prove their conjecture.

In Subsection 4.2.3 we compare $\vartheta^{(t)}(G)$ with the strengthening of $\mathrm{las}^{(t+1)}(G)$ obtained by adding nonnegativity constraints. Our proof technique enables us to compare $\ell^{(t+1)}(G)$ with $\vartheta^{(t)}(G)$ for $t = 0, 1$.

### 4.2.1 Approximating the copositive cone

To define a tractable subcone observe first that a symmetric matrix $M$ is copositive if and only if

$$p_M(x) := v(x)^T M v(x) = \sum_{i,j=1}^{n} M_{ij} x_i^2 x_j^2 \geq 0, \text{ for all } x \in \mathbb{R}^n, \tag{4.26}$$

where $v(x) \in \mathbb{R}^n_+$ is defined by $v(x)_i := x_i^2$, $i = 1, \ldots, n$. Now, the goal is to decide if the polynomial $p_M(x)$ is nonnegative for all $x \in \mathbb{R}^n$. We cannot hope to answer this question in polynomial time in general, but we can instead ask if $p_M(x)$ is SOS, i.e. if its vector of coefficients $p_M$ belongs to the cone $\Sigma_=(n, 4)$ of sum of squares of polynomials. We have seen in Section 2.6 that this is equivalent to deciding if a certain SDP (see (4.28) below) is feasible. Parrilo [75] observed that the sum of squares condition can be further strengthened. He was motivated by the following result of Pólya [78].

**Theorem 4.2.1.** *[78] Let $q$ be a homogeneous polynomial that is positive on the simplex*

$$\Delta := \{z \in \mathbb{R}^n_+ \mid z^T e = 1\}.$$

*Then, there exists $N \in \mathbb{N}$ such that all the coefficients of the polynomial*

$$\left(\sum_{i=1}^n z_i\right)^N q(z)$$

*are positive.*

Parrilo [75] defines the following hierarchy of subcones of the copositive cone $\mathcal{C}_n$. Given an integer $t \in \mathbb{N}$, the cone $\mathcal{K}_n^{(t)} \subseteq \mathcal{S}_n$ denotes the cone of matrices $M$ for which the polynomial

$$p_M^{(t)}(x) := p_M(x)\left(\sum_{i=1}^n x_i^2\right)^t \tag{4.27}$$

is SOS. Obviously, if $p_M^{(t)}(x)$ is SOS then $p_M^{(t+1)}(x)$ is SOS and $p_M(x) \geq 0$ for all $x \in \mathbb{R}^n$. In other words:

$$\mathcal{K}_n^{(0)} \subseteq \cdots \subseteq \mathcal{K}_n^{(t)} \subseteq \mathcal{K}_n^{(t+1)} \subseteq \cdots \subseteq \mathcal{C}_n,$$

which together with Theorem 4.2.1 implies Int $\mathcal{C}_n \subseteq \bigcup_{t\in\mathbb{N}} \mathcal{K}_n^{(t)} \subseteq \mathcal{C}_n$.

Recall from Section 2.6 that $p_M^{(t)}(x)$ is SOS if and only if there exists $X \in \mathbb{R}^{I_=(n,t+2)\times I_=(n,t+2)}$ such that

$$X \succeq 0 \quad \text{and} \sum_{\substack{\gamma,\delta\in I_=(n,t+2)\\ \gamma+\delta=\beta}} X_{\gamma,\delta} = (p_M)_\beta \text{ for all } \beta \in I_=(n, 2t+4). \tag{4.28}$$

(Recall $I_=(n,r) = \{\beta \in \mathbb{N}^n \mid \sum_{i=1}^n \beta_i = r\}$.) Thus testing if $p_M^{(t)}(x)$ is an SOS can be done for any fixed $t \in \mathbb{N}$ by solving an SDP involving a matrix of size $O(n^{t+2})$ and $O(n^{2t+4})$ variables.

Parrilo [75] showed that $\mathcal{K}_n^{(0)} = \mathcal{S}_n + \mathcal{N}_n$, which implies $\left(\mathcal{K}_n^{(0)}\right)^* = \mathcal{S}_n \cap \mathcal{N}_n = \mathcal{D}_n$. Using ideas from Parrilo [75], Bomze and de Klerk [8] gave the following characterization of $\mathcal{K}_n^{(1)}$.

**Theorem 4.2.2.** *[8] A matrix $M$ belongs to $\mathcal{K}_n^{(1)}$ if and only if there exist symmetric matrices $M(i)$ $(i = 1, \ldots, n)$ such that*

$$\begin{array}{rcll} M - M(i) & \in & \mathcal{K}_n^{(0)} & (i = 1, \ldots, n) \\ M(i)_{jk} + M(j)_{ik} + M(k)_{ij} & \geq & 0 & (i, j, k = 1, \ldots, n). \end{array} \tag{4.29}$$

Although the condition (4.29) can be adapted for $t \geq 2$, we are not able to give a characterization for $\mathcal{K}_n^{(t)}$. However, we can prove the following.

**Lemma 4.2.3.** *[39] Let $M \in \mathcal{S}_n$. If there exist symmetric matrices $M(i)$ $(i = 1, \ldots, n)$ such that*

$$M - M(i) \in \mathcal{K}_n^{(t)} \ (i = 1, \ldots, n) \tag{4.30}$$

*and*

$$M(i)_{jk} + M(j)_{ik} + M(k)_{ij} \geq 0 \ (i, j, k = 1, \ldots, n), \tag{4.31}$$

*then* $M \in \mathcal{K}_n^{(t+1)}$.

*Proof.* Let $M$, $M(i)$ $(i = 1, \ldots, n)$ be symmetric matrices which satisfy (4.30) and (4.31). Decompose $p_M^{(t+1)}(x)$ as

$$\left(\sum_{i=1}^n x_i^2\right)^t \sum_{i=1}^n x_i^2 v(x)^T (M - M(i)) v(x) + \left(\sum_{i=1}^n x_i^2\right)^t \sum_{i=1}^n x_i^2 v(x)^T M(i) v(x). \tag{4.32}$$

The first term in (4.32) is SOS due to (4.30), and the second term is SOS since

$$\sum_{i=1}^n x_i^2 v(x)^T M(i) v(x) = \sum_{i,j,k=1}^n x_i^2 x_j^2 x_k^2 (M(i)_{jk} + M(j)_{ik} + M(k)_{ij})$$

and (4.31) holds. Therefore, the polynomial $p_M^{(t+1)}(x)$ is SOS. $\square$

We next present an explicit description of the dual cone of the cone $\mathcal{K}_n^{(t)}$ $(t \in \mathbb{N})$. For this we need the following observation of Bomze and De Klerk [8, p. 169, Thm. 2.2]:

$$p_M^{(t)}(x) = \sum_{\beta \in I_=(n,t+2)} (p_M^{(t)})_\beta x^{2\beta}, \text{ where } (p_M^{(t)})_\beta := \frac{t!}{\beta!}(\beta^T M \beta - \beta^T \mathrm{diag}(M)). \tag{4.33}$$

**Definition 4.2.4.** *Let* $y = (y_\delta)_{\delta \in I_=(n,2t+4)}$ *be given, and let* $N_{=(t+2)}(y)$ *be its moment matrix indexed by* $I_=(n, t+2)$.

- *(i) For* $\gamma \in I_=(n,t)$, $N^\gamma(y)$ *denotes the principal submatrix of* $N_{=(t+2)}(y)$ *indexed by* $\gamma + 2e_1, \ldots, \gamma + 2e_n$*; that is,* $N^\gamma(y)$ *is the* $n \times n$ *matrix with* $(i,j)$*-th entry* $y_{2\gamma+2e_i+2e_j}$, *for* $i, j = 1, \ldots, n$.
- *(ii) Define the* $n \times n$ *matrix*

$$C(y) := \sum_{\gamma \in I_=(n,t)} \frac{t!}{\gamma!} N^\gamma(y). \tag{4.34}$$

**Definition 4.2.5.** *Define the cone*

$$\mathcal{C}_n^{(t)} := \{Z \in \mathbb{R}^{n \times n} \mid Z = C(y) \text{ for some } y \in \mathbb{R}^{I_=(n,2t+4)} \text{ with } N_{=(t+2)}(y) \succeq 0\}.$$

Notice that the matrix $C(y)$ in (4.34) involves only entries of $y$ indexed by even sequences. Therefore in the definition of the cone $\mathcal{C}_n^{(t)}$ one can assume w.l.o.g. that $y_\delta = 0$ whenever $\delta$ has an odd component.

**Proposition 4.2.6.** *The cones* $\mathcal{K}_n^{(t)}$ *and* $\mathcal{C}_n^{(t)}$ *are dual of each other; that is,* $\mathcal{C}_n^{(t)} = (\mathcal{K}_n^{(t)})^*$ *and* $\mathcal{K}_n^{(t)} = (\mathcal{C}_n^{(t)})^*$.

This duality follows from the duality of the cone $\Sigma_{=}(n, 2t+4)$ of sum of squares of polynomials and the cone $\mathcal{N}_{=}(n, 2dt+4)$ of positive semidefinite moment matrices observed in Section 2.6.

*Proof. (of Proposition 4.2.6).* Let $C(y) \in \mathcal{C}_n^{(t)}$, let $M$ be a symmetric $n \times n$ matrix and let $p_M^{(t)}$ be the associated polynomial via (4.27). Using (4.33), one can verify that

$$\operatorname{Tr}(MC(y)) = y^T p_M^{(t)} \quad \text{for any } y \in \mathbb{R}^{I_{\leq}(n,2t+4)}. \tag{4.35}$$

Indeed,

$$\operatorname{Tr}(MC(y)) = \sum_{i,j=1}^{n} M_{ij} C(y)_{ij} = \sum_{i,j=1}^{n} M_{ij} \sum_{\gamma \in I_{\leq}(n,t)} \frac{t!}{\gamma!} y_{2\gamma+2e_i+2e_j}$$
$$= \sum_{\beta \in I_{\leq}(n,t+2)} \left( \sum_{i|\beta_i \geq 2} \frac{t!}{(\beta-2e_i)!} M_{ii} y_{2\beta} + \sum_{i \neq j|\beta_i,\beta_j \geq 1} \frac{t!}{(\beta-e_i-e_j)!} M_{ij} y_{2\beta} \right)$$
$$= \sum_{\beta \in I_{\leq}(n,t+2)} \left( \sum_{i} \frac{t!\beta_i(\beta_i-1)}{\beta!} M_{ii} y_{2\beta} + \sum_{i \neq j} \frac{t!\beta_i\beta_j}{\beta!} M_{ij} y_{2\beta} \right)$$
$$= \sum_{\beta \in I_{\leq}(n,t+2)} \frac{t!}{\beta!} y_{2\beta} \left( \beta^T M \beta - \beta^T \operatorname{diag}(M) \right) = y^T p_M^{(t)}.$$

Using (4.35) and the equality $(\mathcal{N}_{=}(n, 2t+4))^* = \Sigma_{=}(n, 2t+4)$, one can immediately conclude that $\mathcal{K}_n^{(t)} = (\mathcal{C}_n^{(t)})^*$. The cone $\mathcal{C}_n^{(t)}$ is closed since it consists of linear combinations of positive semidefinite matrices and the positive semidefinite cone is closed. Hence $\mathcal{C}_n^{(t)} = (\mathcal{K}_n^{(t)})^*$. □

### 4.2.2 Sums of squares and the stable set problem

Recall first that minimization of a quadratic function over the standard simplex $\Delta := \{x \in \mathbb{R}^n_+ \mid x^T e = 1\}$ is equivalent to a copositive programming problem.

**Theorem 4.2.7** (Bomze et al. [9])**.** *For any $Q \in \mathcal{S}_n$ we have*

$$\min_{x \in \Delta} x^T Q x = \max \lambda \quad s.t. \ \lambda \in \mathbb{R}, \ Q - \lambda \mathbf{J} \in \mathcal{C}_n.$$

*Proof.* $\min_{x\in\Delta} x^T Q x = \max_{\lambda \in \mathbb{R}} \lambda$ s.t. $x^T Q x - \lambda \geq 0$ for all $x \in \Delta$

$= \max_{\lambda \in \mathbb{R}} \lambda$ s.t. $x^T Q x - \lambda x^T \mathbf{J} x \geq 0$ for all $x \in \Delta$

$= \max_{\lambda \in \mathbb{R}} \lambda$ s.t. $x^T (Q - \lambda \mathbf{J}) x \geq 0$ for all $x \in \Delta$

$= \max \lambda$ s.t. $\lambda \in \mathbb{R}, \ Q - \lambda \mathbf{J} \in \mathcal{C}_n$. □

This can be applied to the following result of Motzkin and Straus.

**Theorem 4.2.8** (Motzkin and Straus [71])**.** *For any graph $G$ with adjacency matrix $A_G$ one has*

$$\frac{1}{\alpha(G)} = \min_{x \in \Delta} x^T (\mathbf{I} + A_G) x.$$

As a consequence of the above two theorems we have:

**Corollary 4.2.9.** *For any graph $G$ with adjacency matrix $A_G$ one has*

$$\alpha(G) = \min \lambda \quad s.t. \ \lambda \in \mathbb{R}, \ \lambda(\mathbf{I} + A_G) - \mathbf{J} \in \mathcal{C}_n. \tag{4.36}$$

If we substitute the copositive cone in (4.36) by some subcone of it we obtain an upper bound for $\alpha(G)$. De Klerk and Pasechnik (4.36) use this observation and define the hierarchy of upper bounds for $\alpha(G)$:

$$\vartheta^{(t)}(G) := \min \lambda \ \text{ s.t. } \lambda \in \mathbb{R}, \ \lambda(\mathbf{I} + A_G) - \mathbf{J} \in \mathcal{K}_n^{(t)}. \tag{4.37}$$

They prove $\vartheta^{(0)}(G) = \vartheta'(G)$, hence $\vartheta^{(0)}(G) \leq \vartheta(G)$. Thus $\alpha(G) = \vartheta^{(0)}(G)$ for a perfect graph $G$.

**Convergence towards $\alpha(G)$**

De Klerk and Pasechnik [50] proved that the bounds (4.37) can be used for approximating $\alpha(G)$.

**Theorem 4.2.10.** *[50] For any graph $G$ we have $\lfloor \vartheta^{(t)}(G) \rfloor = \alpha(G)$ for $t \geq (\alpha(G))^2$.*

We have seen in the previous section some hierarchies of semidefinite relaxation of the stable set polytope of a graph $G$ that converge in finitely many steps. For example we have proved that $Q_t(G) = L_t(G) = N_+^{t-1}(\mathrm{TH}(G)) = \mathrm{ST}(G)$ for $t \geq \alpha(G)$. Motivated by this fact and Theorem 4.2.10, de Klerk and Pasechnik [50] ask if the sequence $\vartheta^{(t)}(G)$ ($t \in \mathbb{N}$) converge in $\alpha(G)$ steps towards $\alpha(G)$.

**Theorem 4.2.11.** *[50] If $G$ has stability number $\alpha(G) \leq 2$ then*

$$\alpha(G) = \vartheta^{(\alpha(G)-1)}(G).$$

**Conjecture 4.2.12.** *[50] For any graph $G$*

$$\alpha(G) = \vartheta^{(\alpha(G)-1)}(G).$$

This conjecture and the hierarchy of bounds (4.37) were further studied in [76, 39]. Although the general case of Conjecture 4.2.12 remains open, we can prove the following strengthening of Theorem 4.2.11.

**Theorem 4.2.13.** *[39] If $G$ has stability number $\alpha(G) \leq 8$ then*

$$\alpha(G) = \vartheta^{(\alpha(G)-1)}(G).$$

The proof of Theorem 4.2.13 is inductive, and it is based on Lemma 4.2.3. It is rather technical and lengthy, so we omit it here. Instead, we prove Theorem 4.2.11 and explain where the main difficulty comes from.

Let $G = (V, E)$ be a graph with $V = \{1, 2, \ldots, n\}$ and $\alpha(G) \geq 2$. Set $\alpha := \alpha(G)$ and $i^\perp := \{i\} \cup \{j \mid ij \in E\}$ ($i \in V$). We can now use the idea from the proof of Lemma 4.2.3. For $M := \alpha(\mathbf{I} + A_G) - \mathbf{J}$, decompose $p_M^{(t)}(x)$, defined in (4.27), as

$$\left(\sum_{i=1}^n x_i^2\right)^{t-1} \sum_{i=1}^n x_i^2 v(x)^T X(i) v(x) + \left(\sum_{i=1}^n x_i^2\right)^{t-1} \sum_{i=1}^n x_i^2 v(x)^T Y(i) v(x), \tag{4.38}$$

where

$$Y(i) := \begin{array}{c c} & \begin{array}{cc} i^\perp & \quad V(G)\backslash i^\perp \end{array} \\ \begin{array}{c} i^\perp \\ V(G)\backslash i^\perp \end{array} & \begin{pmatrix} (\alpha-1)\mathbf{J} & -\mathbf{J} \\ -\mathbf{J} & \alpha(\mathbf{I}+A_{G\backslash i^\perp}) - \mathbf{J} \end{pmatrix} \end{array}$$

$$= \frac{\alpha}{\alpha-1}\begin{pmatrix} 0 & 0 \\ 0 & (\alpha-1)(\mathbf{I}+A_{G\backslash i^\perp}) - \mathbf{J} \end{pmatrix} + \begin{pmatrix} (\alpha-1)\mathbf{J} & -\mathbf{J} \\ -\mathbf{J} & \frac{1}{\alpha-1}\mathbf{J} \end{pmatrix}, \tag{4.39}$$

$X(i) := M - Y(i)$ $(i \in V(G))$ and $G\backslash i^\perp$ denotes the subgraph of $G$ induced by $V\backslash i^\perp$. The following observations are easy:

(i) $\sum_{i=1}^n x_i^2 v(x)^T X(i) v(x)$ is SOS, since matrices $X(i)$ $(i = 1, \ldots, n)$ satisfy (4.31);

(ii) the second matrix in (4.39) is PSD;

(iii) $\alpha(G\backslash i^\perp) \leq \alpha - 1$.

Observe next that only the matrix $\frac{\alpha}{\alpha-1}\begin{pmatrix} 0 & 0 \\ 0 & (\alpha-1)(\mathbf{I}+A_{G\backslash i^\perp}) - \mathbf{J} \end{pmatrix}$ has to be considered, since (i) and (ii) guarantee that the remaining terms are SOS. This matrix equals 0 if $\alpha = 2$, hence we proved Theorem 4.2.11.

A possible strategy when $\alpha \geq 3$ would be induction. Assume thus that the conjecture is true for all graphs whose stability number is strictly smaller than $\alpha$. Then,

$$(\alpha-1)(\mathbf{I}+A_{G\backslash i^\perp}) - \mathbf{J} \in \mathcal{K}^{(\alpha-2)}_{|V(G\backslash i^\perp)|} \ (i \in V(G)).$$

Taking (i)-(iii) into account, it would be enough to prove

$$\begin{pmatrix} 0 & 0 \\ 0 & (\alpha(G)-1)(\mathbf{I}+A_{G\backslash i^\perp}) - \mathbf{J} \end{pmatrix} \in \mathcal{K}^{(\alpha-2)}_n.$$

The main difficulty now comes from the fact that, for $t \geq 1$, the cone $\mathcal{K}^{(t)}$ is not invariant under some simple matrix transformations, like extending a matrix by zero row and column, which obviously preserves positive semidefiniteness and copositivity.

**Example 4.2.14.** *For $G := C_5$, the matrix $2(\mathbf{I}+A_G) - \mathbf{J}$ is in $\mathcal{K}^{(1)}_5$, but adding zero row and column yields a matrix that does not belong to $\mathcal{K}^{(1)}_6$ .*

Thus, the idea of decomposing $p^{(t)}_M$ as (4.38) does not work. One has to do a finer analysis to prove the result of Theorem 4.2.13, which unfortunately works only for the case $\alpha(G) \leq 8$.

### Dual side

Consider first the dual formulation of (4.37):

$$\vartheta^{(t)}(G) = \max \operatorname{Tr}(\mathbf{J}X) \ \text{ s.t. } \operatorname{Tr}((\mathbf{I}+A_G)X) = 1, \ X \in \mathcal{C}^{(t)}_n. \tag{4.40}$$

The set of feasible solutions of (4.40) is bounded due to $\operatorname{Tr}((\mathbf{I}+A_G)X) = 1$ which together with $\mathcal{C}^{(t)}_n \subseteq \mathcal{S}^+_n \cup \mathcal{N}_n$ implies $X_{ij} \in [0,1]$ $(i, j \in V(G))$. The maximum thus exists, since we are optimizing over a compact (closed and bounded) set.

Moreover, the program (4.37) is feasible and (4.40) is strictly feasible (see [50]) hence the equality in (4.40) holds.

Dukanovic and Rendl [24] split the constraint $\mathrm{Tr}((\mathbf{I}+A_G)X)=1$ and consider slightly sharper bounds:

$$\begin{aligned}\widehat{\vartheta}^{(t)}(G) &:= \max \mathrm{Tr}(\mathbf{J}X) \quad \text{s.t. } \mathrm{Tr}(X)=1,\ \mathrm{Tr}(A_GX)=0,\ X\in\mathcal{C}_n^{(t)}\\ &= \min \lambda \quad \text{s.t. } \lambda, y\in\mathbb{R},\ \lambda\mathbf{I}+yA_G-\mathbf{J}\in\mathcal{K}_n^{(t)}.\end{aligned} \tag{4.41}$$

Observe that for $t=0$ the definition coincides with (3.17), yielding $\widehat{\vartheta}^{(0)}(G)=\vartheta^{(0)}(G)=\vartheta'(G)$. They also consider a corresponding hierarchy of lower bounds for the fractional chromatic number of a graph $\chi^*(G)$, which will be recalled in the next chapter in Subsection 5.2.4. They show how to use symmetry, when $G$ is vertex transitive, to compute the bound $\widehat{\vartheta}^{(1)}(G)$. They use the characterization given in Theorem 4.2.2 and prove that instead of $n$ matrices $M(i)$ $(i=1,\dots,n)$ it is enough to consider one matrix when $G$ is vertex transitive. However, as we will see in Proposition 4.2.23 below, $\widehat{\vartheta}^{(1)}(G)$ is dominated by the strengthening of $\ell^{(2)}(G)$ obtained by adding nonnegativity conditions.

To be able to compare the bounds $\vartheta^{(t)}(G)$ and $\widehat{\vartheta}^{(t)}(G)$ with the bounds $\mathrm{las}^{(t+1)}(G)$ from the previous section, we define yet another sequence of relaxations of the stable set polytope and the corresponding bounds. Given a graph $G=(V,E)$ and $t\in\mathbb{N}$, define

$$P^{(t)}(G) := \{x\in\mathbb{R}^n \mid x=\mathrm{diag}(X) \quad \text{for some } X\in\mathcal{C}_n^{(t)} \text{ satisfying } \mathrm{Tr}(A_GX)=0,\ X-xx^T\succeq 0\},$$

and define the parameter:

$$\widetilde{\vartheta}^{(t)}(G) := \max_{x\in P^{(t)}(G)} \sum_{i\in V} x_i. \tag{4.42}$$

Note that the program (4.42) is equivalent to (3.18) and thus $\widetilde{\vartheta}^{(0)}(G)=\widehat{\vartheta}^{(t)}(G)=\vartheta^{(0)}(G)=\vartheta'(G)$.

**Theorem 4.2.15.** $\mathrm{STAB}(G)\subseteq P^{(t)}(G)$ *and* $\alpha(G)\le\widetilde{\vartheta}^{(t)}(G)\le\widehat{\vartheta}^{(t)}(G)\le\vartheta^{(t)}(G)$ *for any integer* $t\ge 0$.

*Proof.* Given a stable set $S$ with incidence vector $x:=\chi^S$, define the vector $y\in\mathbb{R}^{I_=(n,2t+4)}$ with

$$y_\delta := \frac{1}{|S|^t} \text{ if } \delta \text{ is even and } S(\delta)\subseteq S, \text{ and } y_\delta:=0 \text{ otherwise},$$

after recalling $S(\delta)=\{i \mid \delta_i\neq 0\}$, $S_{odd}(\delta)=\{i \mid i \text{ is odd}\}$, and $\delta$ being even if $S_{odd}(\delta)=\emptyset$.

Then, $|S|^tN_{=t+2}(y)$ is a 0/1 block diagonal matrix, whose blocks are indexed by the sets

$$\mathcal{O}_T := \{\alpha\in I_=(n,t+2) \mid S(\alpha)\subseteq S,\ S_{odd}(\alpha)=T\} \text{ for } T\subseteq S,$$

and the set $\mathcal{O}:=\{\alpha\in I_=(n,t+2) \mid S(\alpha)\nsubseteq S\}$.

Each $\mathcal{O}_T \times \mathcal{O}_T$ block is the all-ones matrix, and the $\mathcal{O} \times \mathcal{O}$ block is zero. Hence $N_{=t+2}(y) \succeq 0$. For $\gamma \in I_=(n,t)$,

$$|S|^t N^\gamma(y) = \begin{array}{c} \\ {\scriptstyle S} \\ {\scriptstyle V\setminus S} \end{array} \!\!\begin{array}{c} \begin{array}{cc} {\scriptstyle S} & {\scriptstyle V\setminus S} \end{array} \\ \begin{pmatrix} \mathbf{J} & 0 \\ 0 & 0 \end{pmatrix} \end{array} = xx^T$$

if $S(\gamma) \subseteq S$, and $N^\gamma(y) = 0$ otherwise. Hence $C(y) = \sum_{\gamma \in I_=(n,t)} \frac{t!}{\gamma!} N^\gamma(y) = \sum_{\gamma \in I_=(n,t)} \frac{t!}{\gamma!} \frac{1}{|S|^t} xx^T = xx^T$. Setting $X := C(y) = xx^T$, we have $\mathrm{Tr}(A_G X) = 0$, and $x = \mathrm{diag}(X)$, which shows that $x \in P^{(t)}(G)$. This shows the inclusion: $\mathrm{STAB}(G) \subseteq P^{(t)}(G)$ which in turn implies the inequality: $\alpha(G) \le \widetilde{\vartheta}^{(t)}(G)$.

We now verify the inequality: $\widetilde{\vartheta}^{(t)}(G) \le \widehat{\vartheta}^{(t)}(G)$. For this let $x \in P^{(t)}(G)$. Then, $x = \mathrm{diag}(X)$ with $X \in \mathcal{C}_n^{(t)}$, $\mathrm{Tr}(A_G X) = 0$ and $X - xx^T \succeq 0$. The matrix $X' := \frac{X}{\mathrm{Tr}(X)}$ is feasible for (4.41) with $\mathrm{Tr}(\mathbf{J}X') \ge \mathrm{Tr}(X) = \sum_i x_i$, which shows $\widehat{\vartheta}^{(t)}(G) \ge \widetilde{\vartheta}^{(t)}(G)$. The last inequality is obvious. $\square$

### 4.2.3 Comparison with the hierarchy of Lasserre

We compare here the hierarchy of bounds (4.42) with the Lasserre hierarchy (4.20). For this, we first add some nonnegativity conditions for the variable $y$ in (4.18), and define

$$\begin{aligned} Q_t^+(G) := \{x \in \mathbb{R}^{\mathcal{P}_{\le 1}(V)} \mid \quad & \exists y \in \mathbb{R}^{\mathcal{P}_{\le 2t}(V)} \text{ satisfying } y_{\mathbf{0}} = x_{\mathbf{0}},\ y_i = x_i\ (i \in V), \\ & y_I \ge 0\ (I \in \mathcal{P}_{\le 2t}(V)),\ y_{ij} = 0\ (ij \in E),\ M_t(y) \succeq 0\}. \end{aligned} \tag{4.43}$$

$$\mathrm{las}_+^{(t)}(G) := \max \sum_{i \in V} x_i \quad \text{s.t.} \quad x \in Q_t^+(G),\ x_{\mathbf{0}} = 1. \tag{4.44}$$

We also need the following lemma

**Lemma 4.2.16.** *For any $C(y) \in \mathcal{C}_n^{(t)}$ one has*

$$\mathrm{Tr}(\mathbf{J}C(y)) = \sum_{\beta \in I_=(n,t+2)} \frac{(t+2)!}{\beta!} y_{2\beta}, \ \textit{and} \tag{4.45}$$

$$\mathrm{Tr}(C(y)) = \sum_{\beta \in I_=(n,t+2)} \frac{t!}{\beta!} y_{2\beta} \left( \sum_{i=1}^n \beta_i^2 - \beta_i \right). \tag{4.46}$$

*Proof.* Directly from (4.33) and (4.35). $\square$

Finally, the comparison is given in the following result:

**Theorem 4.2.17.** *For any graph $G = (V,E)$ and $t \ge 1$*

$$\mathrm{las}_+^{(t)}(G) \le \widetilde{\vartheta}^{(t-1)}(G).$$

*Proof.* In view of Theorem 4.2.15, we have to show that $\mathrm{las}_+^{(t)}(G) \leq \widetilde{\vartheta}^{(t-1)}(G)$ for any positive integer $r$. For this, let $x \in \mathbb{R}^{\mathcal{P}_{\leq 2t}(V)}$ be feasible for the program defining $\mathrm{las}_+^{(r)}(G)$. Then, $x_I = 0$ for any $I \in \mathcal{P}_{\leq 2t}(V)$ containing an edge. We may assume that $\sum_{i=1}^n x_i > 0$. For $p = 1, \ldots, t+1$, define

$$\ell_p := \sum_{\beta \in I_=(n,p-1)} \frac{(p-1)!}{\beta!} x_{S(\beta)}.$$

Then, $\ell_1 = 1$, $\ell_p \geq \ell_2 = \sum_{i=1}^n x_i > 0$ for $p \geq 2$. For $p = 1, \ldots, t$, define $y = (y_\delta)_{\delta \in I_=(n,2p+2)}$ as follows: $y_\delta = 0$ if $S_{odd}(\delta) \neq \emptyset$, $y_\delta := \frac{1}{\ell_p} x_{S(\delta)}$ otherwise (then $|S(\delta)| \leq p+1 \leq t+1$).

We first prove some intermediate results.

**Lemma 4.2.18.** $N_{=p+1}(y) \succeq 0$.

*Proof.* For $I \subseteq V$, set $\mathcal{O}_I := \{\beta \in I_=(n,p+1) \mid S_{odd}(\beta) = I\}$ and $N_I := (y_{\beta+\beta'})_{\beta,\beta' \in \mathcal{O}_I}$. Then, $N_{=p+1}(y)$ is a block diagonal matrix with the matrices $N_I$ ($I \subseteq V$) as diagonal blocks. As $\ell_p N_I = (x_{S(\beta) \cup S(\beta')})_{\beta,\beta' \in \mathcal{O}_I}$, $N_I \succeq 0$ since it is obtained from a principal submatrix of $M_t(x)$ by duplicating certain rows/columns (unless $|I| = t+1$ in which case $N_I$ is the $1 \times 1$ matrix with entry $x_I \geq 0$, implying again $N_I \succeq 0$).

□

Therefore, the matrix $Z(p) := C(y) = \sum_{\gamma \in I_=(n,p-1)} \frac{(p-1)!}{\gamma!} N^\gamma(y)$ belongs to the cone $\mathcal{C}_n^{(p-1)}$. Moreover, $Z(p)_{ij} = 0$ if $ij \in E$. Define the matrix

$$\tilde{Z}(p) := \begin{pmatrix} 1 & Z(p)_{11} \ldots Z(p)_{nn} \\ Z(p)_{11} & \\ \vdots & Z(p) \\ Z(p)_{nn} & \end{pmatrix}. \tag{4.47}$$

**Lemma 4.2.19.** $\tilde{Z}(p) \succeq 0$.

*Proof.* The matrix:

$$\begin{aligned} \ell_p \tilde{Z}(p) &= \sum_{\gamma \in I_=(n,p-1)} \frac{(p-1)!}{\gamma!} \begin{pmatrix} x_{S(\gamma)} & y_{2\gamma+4e_1} \cdots y_{2\gamma+4e_n} \\ y_{2\gamma+4e_1} & \\ \vdots & (y_{2\gamma+2e_j+2e_k})_{j,k=1}^n \\ y_{2\gamma+4e_n} & \end{pmatrix} \\ &= \sum_{\gamma \in I_=(n,p-1)} \frac{(p-1)!}{\gamma!} \begin{pmatrix} x_{S(\gamma)} & x_{S(\gamma+e_1)} \cdots x_{S(\gamma+e_n)} \\ x_{S(\gamma+e_1)} & \\ \vdots & (x_{S(\gamma+e_j+e_k)})_{j,k=1}^n \\ x_{S(\gamma+e_n)} & \end{pmatrix} \end{aligned}$$

is positive semidefinite, since the matrices in the above summation are principal submatrices of $M_t(x)$.

□

**Lemma 4.2.20.** $\mathrm{Tr}(\mathbf{J}Z(p)) = \frac{\ell_{p+2}}{\ell_p}$ *and* $\mathrm{Tr}(Z(p)) = \frac{\ell_{p+1}}{\ell_p}$.

*Proof.* As $Z(p) = C(y) \in \mathcal{C}_n^{(p-1)}$, one can use (4.45) and (4.46). Namely,

$$Tr(\mathbf{J}Z(p)) = \sum_{\beta \in I_=(n,p+1)} \frac{(p+1)!}{\beta!} y_{2\beta} = \frac{1}{\ell_p} \sum_{\beta \in I_=(n,p+1)} \frac{(p+1)!}{\beta!} x_{S(\beta)} = \frac{\ell_{p+2}}{\ell_p}.$$

Moreover,

$$\begin{aligned}\mathrm{Tr}(Z(p)) &= \sum_{\beta \in I_=(n,p+1)} \frac{(p-1)!}{\beta!} y_{2\beta} \sum_{i=1}^{n} (\beta_i^2 - \beta_i) \\ &= \tfrac{1}{\ell_p} \sum_{i=1}^{n} \sum_{\beta \in I_=(n,p+1)} \frac{(p-1)!}{\beta!} \beta_i(\beta_i - 1) x_{S(\beta)}.\end{aligned}$$

We can restrict the inner summation to $\beta$ with $\beta_i \geq 2$. Then, $\delta := \beta - e_i$ has the same support as $\beta$ and

$$\mathrm{Tr}(Z(p)) = \frac{1}{\ell_p} \sum_{i=1}^{n} \sum_{\delta \in I_=(n,p)} \frac{(p-1)!}{\delta!} \delta_i x_{S(\delta)} = \frac{1}{\ell_p} \sum_{\delta \in I_=(n,p)} \frac{(p-1)!}{\delta!} |\delta| x_{S(\delta)} = \frac{\ell_{p+1}}{\ell_p}.$$

□

**Lemma 4.2.21.** $\frac{\ell_{p+2}}{\ell_{p+1}} \geq \frac{\ell_{p+1}}{\ell_p}$.

*Proof.* By Lemma 4.2.19, $\tilde{Z}(p) \succeq 0$, implying $Z(p) - \mathrm{diag}(Z(p))\mathrm{diag}(Z(p))^T \succeq 0$. Therefore, $e^T(Z(p) - \mathrm{diag}(Z(p))\mathrm{diag}(Z(p))^T)e \geq 0$, yielding $Tr(JZ(p)) \geq (Tr(Z(p)))^2$. The result now follows using Lemma 4.2.20. □

We can now conclude the proof of Theorem 4.2.17. From Lemmas 4.2.20 and 4.2.21, we deduce that $\sum_{i=1}^n Z(t)_{ii} = \frac{\ell_{t+1}}{\ell_t} \geq \frac{\ell_2}{\ell_1} = \sum_{i=1}^n x_i$. The vector $z := \mathrm{diag}(Z(t))$ is feasible for the program (4.42) defining the parameter $\widetilde{\vartheta}^{(t-1)}(G)$. Hence, $\widetilde{\vartheta}^{(t-1)}(G) \geq \sum_{i=1}^n z_i = \mathrm{Tr}(Z(t)) \geq \sum_{i=1}^n x_i$. This shows that $\widetilde{\vartheta}^{(t-1)}(G) \geq \mathrm{las}^{(r)}(G)$ and finishes the proof of Theorem 4.2.17. □

**Corollary 4.2.22.** *For any graph $G = (V, E)$ and $t \geq 1$*

$$\mathrm{las}_+^{(t)}(G) \leq \widetilde{\vartheta}^{(t-1)}(G) \leq \widehat{\vartheta}^{(t-1)}(G) \leq \vartheta^{(t-1)}(G).$$

*Proof.* Directly from Theorems 4.2.15 and 4.2.17. □

Observe that in the proof above we use only components of vector $x$ indexed by $\mathcal{P}_{\leq t+1}(V)$. Hence, one may wonder if we can compare $\ell^{(t)}$ with $\widetilde{\vartheta}^{(t-1)}(G)$. For this define

$$\begin{aligned} L_t^+(G) := \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid \quad & \exists y \in \mathbb{R}_+^{\mathcal{P}_{\leq t+1}(V)} \text{ satisfying } y_{\mathbf{0}} = x_{\mathbf{0}}, \\ & y_i = x_i \ (i \in V),\ y_{ij} = 0 \ (ij \in E), \\ & M(T; y) \succeq 0 \ \forall T \subseteq V \text{ with } |T| = t-1\}, \end{aligned} \tag{4.48}$$

and

$$\ell_+^{(t)}(G) := \max \sum_{i \in V} x_i \quad \text{s.t} \quad x \in L_t^+(G),\ x_{\mathbf{0}} = 1. \tag{4.49}$$

**Proposition 4.2.23.** *For any graph $G = (V, E)$ and $t \in \{1, 2\}$*

$$\mathrm{las}_+^{(t)}(G) \le \ell_+^{(t)}(G) \le \widetilde{\vartheta}^{(t-1)}(G) \le \widehat{\vartheta}^{(t-1)}(G) \le \vartheta^{(t-1)}(G).$$

The proof is along the same lines as for the proof of Theorem 4.2.17. One should only take into account that the matrices $\ell_p N_I(y)$ in the proof of Lemma 4.2.18 are now principal submatrices (after duplicating some rows and columns) of some $M(T; x)$, with $|T| = t - 1$. Notwithstanding, this proof does not work for $t \ge 3$ since some of the matrices $N_I(y)$ have a higher order than the matrices $M(T; x)$. In particular, for $t = 3$, the matrix $\ell_3 N_{\mathbf{0}}(y)$ is not a submatrix of any of the matrices $M(T; x)$, with $|T| = 2$.

### 4.2.4 Weighted case

Let $G = (V, E)$ be a graph. Busygin [15] shows the following extension to the weighted case of the Motzkin-Straus theorem.

**Theorem 4.2.24.** *[15] Given $w_i > 0$ ($i \in V$), set $w_{\min} := \min_{i \in V} w_i$. Then,*

$$\frac{w_{\min}}{\alpha_w(G)} = \min_{x \in \Delta} x^T \left( w_{\min}(\mathrm{Diag}(w))^{-1} + A_G \right) x.$$

In other words, the matrix $\alpha_w(G) \left( (\mathrm{Diag}(w))^{-1} + \frac{1}{w_{\min}} A_G \right) - \mathbf{J}$ is copositive or, equivalently, the matrix $\alpha_w(G) \left( \mathrm{Diag}(w) + A_{G,w} \right) - ww^T$ is copositive, where $A_{G,w}$ is the matrix whose $ij$-th entry is $\frac{w_i w_j}{w_{\min}}$ if $ij \in E$ and 0 otherwise. Set

$$w_{\max} := \max_{i \in V} w_i, \;\; W_G := \frac{(w_{\max})^2}{w_{\min}}. \tag{4.50}$$

The matrix $\alpha_w(G) \left( \mathrm{Diag}(w) + W_G A_G \right) - ww^T$ is also copositive, since the entries of $A_{G,w}$ are at most $W_G$. This leads us to define the following weighted analogue of the parameter $\vartheta^{(r)}(\cdot)$:

$$\vartheta_w^{(r)}(G) := \min \; t \;\; \text{subject to } t(\mathrm{Diag}(w) + W_G A_G) - ww^T \in \mathcal{K}_n^{(r)}. \tag{4.51}$$

This definition reduces to the original definition (4.37) when all weights are equal to 1.

Recall next the definition of $P^{(t)}(G)$ from previous subsection. As we have seen, the sets $P^{(t)}(G)$ provide a hierarchy of semidefinite relaxations for $\mathrm{STAB}(G)$. Theorem 4.1.2 implies $\mathrm{STAB}(G) = P^{(0)}(G)$ when $G$ is a perfect graph. With respect to this, a natural question to ask is whether the analogue of Conjecture 4.2.12 may hold, asserting that $\mathrm{STAB}(G) = P^{(t)}(G)$ for $t \ge \alpha(G) - 1$. We are able to give a positive answer only in the case $t = 1$. For this, given positive node weights $w \in \mathbb{R}_+^V$, we have to compare the *weighted stability number* $\alpha_w(G) := \max\limits_{x \in \mathrm{STAB}(G)} w^T x$ and the weighted parameter:

$$\widetilde{\vartheta}_w^{(t)}(G) := \max_{x \in P^{(t)}(G)} w^T x. \tag{4.52}$$

**Lemma 4.2.25.** *The parameters (4.52) and (4.51) satisfy: $\tilde{\vartheta}_w^{(r)}(G) \le \vartheta_w^{(r)}(G)$.*

*Proof.* Assume $M := t(\mathrm{Diag}(w) + W_G A_G) - ww^T \in \mathcal{K}_n^{(r)}$ and let $x = \mathrm{diag}(X)$ where $X \in \mathcal{C}_n^{(r)}$, $\mathrm{Tr}(A_G X) = 0$, $X - xx^T \succeq 0$. Then, $0 \leq \mathrm{Tr}(MX) = tw^T x - w^T X w$, yielding $tw^T x \geq w^T X w \geq (w^T x)^2$ and thus $t \geq w^T x$. This gives the desired inequality. □

**Lemma 4.2.26.** *For $r = 0$, $\tilde{\vartheta}_w^{(0)}(G) = \vartheta_w^{(0)}(G)$. Therefore, $\vartheta_w^{(0)}(G) = \alpha_w(G)$ when $G$ is a perfect graph.*

*Proof.* It remains to show the inequality: $\vartheta_w^{(0)}(G) \leq \tilde{\vartheta}_w^{(0)}(G)$. For this, we first observe that

$$\vartheta_w^{(0)}(G) \leq \phi(G) := \min\ t \text{ subject to } t\mathrm{Diag}(w) + yA_G - ww^T \in \mathcal{K}_n^{(0)}. \quad (4.53)$$

Our argument is similar to the one used by de Klerk and Pasechnik [50] in the unweighted case. Assume $M := t\mathrm{Diag}(w) + yA_G - ww^T \in \mathcal{K}_n^{(0)}$. Then, $M = P + N$, where $P \succeq 0$, $N \geq 0$, $\mathrm{diag}(N) = 0$. Hence, $t(\mathrm{Diag}(w) + W_G A_G) - ww^T = M + (tW_G - y)A_G = P + N + (tW_G - y)A_G$. It suffices now to verify that $N' := N + (tW_G - y)A_G \geq 0$. For this pick an edge, say $12 \in E$. As $P \succeq 0$, we have $P_{11} + P_{22} \geq 2P_{12}$, yielding $t(w_1 + w_2) - 2(y - N_{12}) \geq (w_1 - w_2)^2$. Finally, $2N'_{12} = 2N_{12} + 2t\frac{w_{\max}^2}{w_{\min}} - 2y \geq t(w_1 + w_2) - 2(y - N_{12}) \geq 0$ proves (4.53).

Next, using conic duality, we obtain that

$$\phi(G) = \max\ w^T X w \ \text{ subject to } \mathrm{Tr}(\mathrm{Diag}(w)X) = 1,\ \mathrm{Tr}(A_G X) = 0,\ X \in \mathcal{C}_n^{(0)}.$$

Set $u := (\sqrt{w_i})_{i=1}^n$. Rescaling $X$ by $Y = \mathrm{Diag}(u)X\mathrm{Diag}(u)$, we find that

$$\phi(G) = \max\ u^T Y u \ \text{ subject to } \mathrm{Tr}(Y) = 1,\ \mathrm{Tr}(A_G Y) = 0,\ Y \in \mathcal{C}_n^{(0)}.$$

(As $\mathcal{C}_n^{(0)}$ consists of the nonnegative positive semidefinite matrices, it is closed under the above rescaling.) We can now conclude that $\phi(G) \leq \tilde{\vartheta}_w^{(0)}(G)$. This is the same proof as for Theorem 67.11 in [84] (which gives the result with the cone $\mathcal{C}_n^{(0)}$ being replaced by the cone of positive semidefinite matrices). □

**Theorem 4.2.27.** *[39] For a graph $G$ with positive node weights $w \in \mathbb{R}^V$,*

$$\vartheta_w^{(1)}(G) \leq \max_{i \in V}\ (w_i + \vartheta_w^{(0)}(G \backslash i^\perp)). \quad (4.54)$$

We omit here the proof of this result. It is similar to the proof of Theorem 4.2.11, i.e. it is based on Theorem 4.2.2.

**Corollary 4.2.28.** $\mathrm{STAB}(G) = P^{(1)}(G)$ *if $G \backslash i^\perp$ is perfect for all $i \in V$; this holds in particular if $\alpha(G) = 2$.*

## 4.3 Conclusions

We presented several hierarchies of semidefinite upper bounds for the stability number $\alpha(G)$ of a graph $G$ in this chapter.

In Section 4.1 we were dealing with three hierarchies of relaxations of the stable set polytope $\mathrm{STAB}(G)$. We proposed the new block diagonal hierarchy which is, as a relaxation of the hierarchy of Lasserre [57], at least as good as the hierarchy of Lovász and Schrijver [65], and computationally less costly. All hierarchies converge to STAB($G$) in $\alpha(G)$ steps.

We will see in the next chapter how these hierarchies can be transformed into hierarchies of lower bounds for the fractional chromatic number $\chi^*(G)$ and the chromatic number $\chi(G)$. In Chapter 6 we will show how to compute bounds, up to order three, in the new block diagonal hierarchy for Paley graphs. For that we exploit the properties of the automorphism groups of these graphs.

In Section 4.2 we considered the copositive programming formulation for $\alpha(G)$ of de Klerk and Pasechnik [50] and, by replacing the copositive cone by $\mathcal{K}_n^{(t)}$ ($t \in \mathbb{N}$), we obtained a hierarchy of upper bounds for $\alpha(G)$. This hierarchy is conjectured to converge to $\alpha(G)$ in $\alpha(G)$ steps. We could prove it for $\alpha(G) \leq 8$, but the case $\alpha(G) > 8$ remains open. Using the dual formulations for bounds in this hierarchy we defined a slightly sharper hierarchy via $P^{(t)}(G)$ ($t \in \mathbb{N}$), the sequence of relaxations of STAB($G$). We have shown that these hierarchies of bounds are dominated (after adding nonnegativity conditions) by the hierarchy of Lasserre, and moreover that the first two bounds are dominated by the new block diagonal, hierarchy bounds. We also introduced the generalization of the hierarchy of de Klerk and Pasechnik, which consists of upper bounds for the weighted stability number $\alpha_w(G)$.

In the next chapter we will see how $\chi^*(G)$ can be formulated as a copositive program. By once again applying the recipe of Parrilo [75], we will construct a hierarchy of lower bounds for $\chi^*(G)$, which was proposed recently by Dukanovic and Rendl [24].

## Acknowledgements

We thank E. d Klerk for communicating Example 4.2.14 to us.

# Chapter 5

# Semidefinite programming lower bounds for the chromatic number

The chromatic number $\chi(G)$ of a graph $G$ can be formulated via a 0/1 linear program (see, e.g., [69]) involving $O(n^2)$ variables. One can thus apply the classical lift and project procedures, proposed for example in [57, 65, 87], to derive hierarchies of semidefinite approximations finding $\chi(G)$ in $O(n^2)$ steps. However, in this chapter we do not take a 0/1 formulation as a starting point. Instead, we start with the Lovász theta number $\overline{\vartheta}(G)$ of the complement $\overline{G}$ defined in Chapter 3. As a lower bound for $\chi(G)$, at least as strong as $\omega(G)$, $\overline{\vartheta}(G)$ was already used, e.g., for approximately colouring the graph (see [53, 26, 46]).

We have seen in Section 3.3 how $\overline{\vartheta}(G)$ can be strengthened by adding nonnegativity and triangle constraints. In this chapter we go beyond. We obtain approximations for $\chi(G)$ by applying the techniques presented in Chapter 4 and by generalizing the reduction (recall (2.31))

$$\chi(G) \leq t \Longleftrightarrow \alpha(K_t \Box G) = |V(G)|, \tag{5.1}$$

from the colouring problem to the stable set problem from Section 2.5. We present hierarchies of semidefinite bounds, which start with $\overline{\vartheta}(G)$ or some variations of it, and converge either to the fractional chromatic number $\chi^*(G)$ or to the chromatic number $\chi(G)$.

In Section 5.1, we generalize the reduction (5.1) to all graph parameters lying between $\omega(\cdot)$ and $\chi(\cdot)$. We introduce an operator $\Psi$, which is monotone nonincreasing and maps a parameter sandwiched between $\frac{|V(\cdot)|}{\chi(\cdot)}$ and $\overline{\chi}(\cdot)$ to an integer parameter that lies between $\omega(\cdot)$ and $\chi(\cdot)$. Moreover, if a graph parameter is polynomial time computable, the same holds for its image. We show that $\Psi$ maps

- $\vartheta(\cdot)$ to $\left\lceil \overline{\vartheta}(\cdot) \right\rceil$ by exploring the symmetry in $K_t \Box G$, i.e. in the semidefinite programs defining the image of $\vartheta(\cdot)$;

- the whole interval $[\frac{|V(\cdot)|}{\chi(\cdot)}, \overline{\chi}(\cdot)]$ to $\omega(\cdot)$, and hence there is no polynomial time computable graph parameter sandwiched between $\frac{|V(\cdot)|}{\chi(\cdot)}$ (or $\chi^*(\cdot)$) and $\overline{\chi}(\cdot)$.

In addition, by applying the idea to the Motzkin-Straus formulation for $\alpha(G)$ we give (quadratically constrained) quadratic and copositive programming formulations for $\chi(G)$.

In Section 5.2 we introduce hierarchies of semidefinite bounds which converge to either $\chi^*(G)$, or $\chi(G)$. We construct hierarchies, converging to $\chi(G)$ in $|V(G)|$ steps, by applying the operator $\Psi$ to Lasserre's hierarchy and to the new block diagonal hierarchy presented in Chapter 4. We show how to explore symmetry in semidefinite programs defining $\Psi_{\ell^{(2)}}(G)$.

To approximate $\chi^*(G)$ we choose a convenient formulation for $\overline{\vartheta}(G)$ (equivalent to (3.4)), and use the moment matrix approach. In this way we build hierarchies for $\chi^*(G)$, which converge to $\chi^*(G)$ in $\alpha(G)$ steps. At the end of the section, we also briefly revisit the approach of Dukanovic and Rendl [24] and indicate links to our hierarchies. They give completely positive programming formulation for $\chi^*(G)$, and apply the ideas from Parrilo [75] and de Klerk and Pasechnik [50].

## 5.1 The operator $\Psi$ and its applications

### 5.1.1 Basic properties of $\Psi$

Using relation (5.1), we see that the chromatic number of a graph $G$ can be defined as the optimum solution of the following program

$$\chi(G) = \min_{t \in \mathbb{N}} t \ \text{ s.t. } \alpha(K_t \Box G) = |V(G)|, \tag{5.2}$$

where $K_t \Box G$ is the Cartesian product of the graphs $G$ and $K_t$. Recall that

$$V(K_t \Box G) := V(K_t) \times V(G) = \bigcup_{p=1}^{t} V_p, \quad \text{where } V_p := \{pi \mid i \in V(G)\} \tag{5.3}$$

and $\{pi, qj\}$ is an edge if $i = j$ and $p \neq q$, or if $p = q$ and $ij \in E(G)$. This fact motivates the following definition.

**Definition 5.1.1.** *Given a graph parameter $\beta(\cdot)$ satisfying*

$$\frac{|V(\cdot)|}{\chi(\cdot)} \leq \beta(\cdot) \leq \overline{\chi}(\cdot), \tag{5.4}$$

*define the graph parameter $\Psi_\beta(\cdot)$ by*

$$\Psi_\beta(G) := \min_{t \in \mathbb{N}} t \ \text{ s.t. } \beta(K_t \Box G) = |V(G)|. \tag{5.5}$$

**Lemma 5.1.2.** *(a) The graph parameter $\Psi_\beta(G)$ is well defined if $\beta(\cdot)$ satisfies (5.4).*

*(b) The operator $\Psi$ is monotone nonincreasing; that is, $\Psi_{\beta_2}(\cdot) \leq \Psi_{\beta_1}(\cdot)$ if $\beta_1(\cdot)$, $\beta_2(.)$ satisfy (5.4) and $\beta_1(\cdot) \leq \beta_2(\cdot)$.*

*(c)* $\Psi_\alpha(G) = \chi(G)$.

*(d)* $\Psi_\beta(G) = \omega(G)$ *for* $\beta(\cdot) := \frac{|V(\cdot)|}{\omega(\cdot)}$.

*(e)* $\Psi_{\overline{\chi}}(G) = \omega(G)$.

*(f)* $\Psi_\beta(G) = \chi(G)$ *for* $\beta(\cdot) := \frac{|V(\cdot)|}{\chi(\cdot)}$.

*(g) If* $\beta(\cdot)$ *satisfies (5.4), then*

$$\omega(\cdot) \leq \Psi_\beta(\cdot) \leq \chi(\cdot). \tag{5.6}$$

*Proof.* (a) Assume $\beta(\cdot)$ satisfies (5.4) and let $1 \leq t \leq n := |V(G)|$. As $\chi(K_t \square G) \leq n$ we have $\beta(K_t \square G) \geq \frac{|V(K_t \square G)|}{\chi(K_t \square G)} \geq \frac{nt}{n} = t$. On the other hand, $\beta(K_t \square G) \leq \overline{\chi}(K_t \square G) \leq n$. Therefore, $\beta(K_n \square G) = n$, thus showing that $\Psi_\beta(G)$ is well defined.

(b) If $\beta_1(\cdot) \leq \beta_2(\cdot)$ satisfy (5.4), then $\beta_1(K_t \square G) = n$ implies $\beta_2(K_t \square G) = n$, which gives $\Psi_{\beta_2}(G) \leq \Psi_{\beta_1}(G)$.

(c) The identity $\Psi_\alpha(G) = \chi(G)$ follows directly from (5.2).

(d) For $\beta(\cdot) := \frac{|V(\cdot)|}{\omega(\cdot)}$, the identity $\Psi_\beta(G) = \omega(G)$ follows from the fact that $\omega(K_t \square G) = \max(t, \omega(G))$.

(e) We verify that $\Psi_{\bar{\chi}}(G) = \omega(G)$. As $\overline{\chi}(\cdot) \geq \frac{|V(\cdot)|}{\omega(\cdot)}$, we deduce using (b) and (d) that $\Psi_{\overline{\chi}}(G) \leq \Psi_{|V|/\omega}(G) = \omega(G)$. To show the reverse inequality, consider a clique $C$ in $G$ of size $\omega(G)$ and let $C_t$ be the subset of $V(K_t \square G)$ consisting of all the copies of the nodes in $C$. Thus $C_t$ is covered by $t$ cliques of $K_t \square G$. As the remaining nodes of $K_t \square G$ can be covered by $n - |C|$ cliques, we have $\overline{\chi}(K_t \square G) \leq t + n - |C|$. Therefore $\overline{\chi}(K_t \square G) = n$ implies $t \geq |C| = \omega(G)$, which shows $\Psi_{\bar{\chi}}(G) \geq \omega(G)$.

(f) Consider now the parameter $\beta(\cdot) := \frac{|V(\cdot)|}{\chi(\cdot)}$. As $\beta(\cdot) \leq \alpha(\cdot)$, we deduce using (b) that $\Psi_\beta(G) \geq \Psi_\alpha(G) = \chi(G)$, and equality holds since one can easily verify that $\beta(K_t \square G) = n$ for $t := \chi(G)$.

(g) Relation (5.6) now follows directly using again (b). □

**Corollary 5.1.3.** *If* $\beta(\cdot)$ *is a graph parameter satisfying* $\frac{|V(\cdot)|}{\omega(\cdot)} \leq \beta(\cdot) \leq \bar{\chi}(\cdot)$, *then* $\Psi_\beta = \omega$. *In particular,* $\Psi_{\overline{\chi^*}} = \omega$. *Moreover, if* $\frac{|V(\cdot)|}{\chi(\cdot)} \leq \beta(\cdot) \leq \alpha(\cdot)$ *then* $\Psi_\beta = \chi$.

*Proof.* Directly from Lemma 5.1.2 (b),(c),(d),(e),(f) and Theorem 3.3.1 (a). □

Therefore, the operator $\Psi$ takes a graph parameter $\beta(G)$ (nested e.g. between $\alpha(G)$ and $\overline{\chi}(G)$) and produces the integer lower bound $\Psi_\beta(G)$ (nested between $\omega(G)$ and $\chi(G)$) for the chromatic number $\chi(G)$; figure 5.1 illustrates how the operator $\Psi$ acts on various parameters. As $\alpha(G)\chi^*(G) \geq |V(G)|$,

$$\beta(G) \geq \alpha(G) \Longrightarrow \chi(G) \geq \chi^*(G) \geq \frac{|V(G)|}{\beta(G)}.$$

The next lemma shows that, under the mild assumption (5.7), $\Psi_\beta(G)$ is at least as good as the obvious lower bound $|V(G)|/\beta(G)$ for $\chi(G)$. However, $\Psi_\beta(G)$ may be equal to $\chi(G)$ while $\frac{|V(G)|}{\beta(G)}$ always remains below the fractional chromatic number $\chi^*(G)$.

**Lemma 5.1.4.** *Assume that a graph parameter $\beta(\cdot)$ satisfies: $\alpha(\cdot) \leq \beta(\cdot) \leq \overline{\chi}(\cdot)$ and*

$$\beta(K_t \square G) \leq t\beta(G) \quad \textit{for all } t \in \mathbb{N}. \tag{5.7}$$

*Then, $\Psi_\beta(G) \geq \frac{|V(G)|}{\beta(G)}$.*

*Proof.* If $\beta(K_t \square G) = |V(G)|$, then $|V(G)| \leq t\beta(G)$, i.e., $t \geq \frac{|V(G)|}{\beta(G)}$. $\square$

The condition 5.7 holds e.g. when $\beta(\cdot)$ is additive with respect to the direct sum of graphs and monotone nonincreasing with respect to adding edges to a graph, i.e. if $\beta(\cdot)$ satisfies

(i) $\beta(G + G') = \beta(G) + \beta(G')$, for any $G = (V, E)$ and $G' = (V', E')$ with $V \cap V' = \emptyset$; and

(ii) $E \subseteq E' \Longrightarrow \beta(G) \geq \beta(G')$, for any $G = (V, E)$ and $G' = (V, E')$.

Apparently, if (i) and (ii) hold for $\beta(\cdot)$ then $t\beta(G) = \sum_{i=1}^t \beta(G_i) \geq \beta(K_t \square G)$, where $G_i = (V_i, E_i)$ ($i = 1, \ldots, t$) are the copies of $G$ in $K_t \square G$, i.e. the subgraphs of $K_t \square G$ induced by $V_i$ ($i = 1, \ldots, t$) (recall the definition of $K_t \square G$ from (2.30)).

One can easily verify that conditions (i) and (ii), and consequently (5.7), hold for $\alpha(\cdot)$, $\overline{\chi}(\cdot)$, $\overline{\chi}^*(\cdot)$, $\vartheta(\cdot)$, $\vartheta'(\cdot)$, and the parameters $\text{las}^{(r)}(\cdot)$ and $\ell^{(r)}(\cdot)$ defined in (4.20) and (4.23) respectively.

Remark 5.1.5. *If $\beta(\cdot) \in [\alpha(\cdot), \overline{\chi}(\cdot)]$, then $\Psi_\beta(G) - \frac{|V(G)|}{\beta(G)} \leq \chi(G) - \frac{|V(G)|}{\overline{\chi}(G)}$, with equality e.g. when $G$ is a perfect graph (since then $\alpha(G) = \overline{\chi}(G) = \beta(G)$ and $\omega(G) = \chi(G) = \Psi_\beta(G)$). Hence the gap $\Psi_\beta(G) - \frac{|V(G)|}{\beta(G)}$ can be made arbitrarily large. For instance, this gap is equal to $n - \frac{2n}{n+1} = n\frac{n-1}{n+1}$ when $G$ is the disjoint union of a clique of size $n$ and $n$ isolated points.*

$$\begin{array}{ccccccccccccc}
\frac{|V|}{\chi} & \leq & \frac{|V|}{\chi^*} & \leq & \alpha & \leq & \vartheta' & \leq & \vartheta & \leq & \overline{\chi}^* & \leq & \overline{\chi} \\
 & \searrow & \downarrow & \swarrow & & & \downarrow & & \downarrow & & \downarrow & \swarrow & \\
 & & \chi & & \geq & & \left\lceil \overline{\vartheta^+} \right\rceil & \geq & \left\lceil \overline{\vartheta} \right\rceil & \geq & \omega & & \\
 & & & & & & & & & \nearrow & \uparrow & \nwarrow & \\
 & & & & & & & & \frac{|V|}{\omega} & \leq & \overline{\chi}^* & \leq & \overline{\chi}
\end{array}$$

Figure 5.1: Converting graph parameters by the operator $\Psi$

### 5.1.2 Action of the operator $\Psi$ on the theta number

We investigate here how the operator $\Psi$ applies to the theta number $\vartheta(\cdot)$ and its strengthening $\vartheta'(\cdot)$. In particular we prove that the operator $\Psi$ maps the theta number $\vartheta(\cdot)$ to $\left\lceil \overline{\vartheta}(\cdot) \right\rceil$, and its strengthening $\vartheta'(\cdot)$ to $\left\lceil \overline{\vartheta^+}(\cdot) \right\rceil$. De Klerk et al. [53] consider a graph parameter closely related to $\Psi_\vartheta$ for which they can also show that it coincides with $\left\lceil \overline{\vartheta}(\cdot) \right\rceil$.

**Theorem 5.1.6.** *For any graph $G$ the following holds:*

*(i)* $\Psi_\vartheta(G) = \left\lceil \overline{\vartheta}(G) \right\rceil$,

*(ii)* $\Psi_{\vartheta'}(G) = \left\lceil \overline{\vartheta^+}(G) \right\rceil$.

We need the following lemma for the proof of Theorem 5.1.6.

**Lemma 5.1.7.** *For a positive semidefinite $n \times n$ matrix $X$, $n\mathrm{Tr}(X) \geq \langle \mathbf{J}, X \rangle$, with equality if and only if $X = c\mathbf{J}$ for some nonnegative scalar $c$.*

*Proof.* As $X \succeq 0$, its entries satisfy $X_{ii} + X_{jj} \geq 2X_{ij}$ for all $i, j \in \{1, ..., n\}$. Thus, $n\sum_{i=1}^n X_{ii} \geq \sum_{i,j=1}^n X_{ij}$. Equality holds if and only if $X_{ii} + X_{jj} = 2X_{ij}$ for all $i, j$, which gives $X_{ii} = X_{jj} = X_{ij}$ for all $i, j$. □

*Proof of Theorem 5.1.6.* If $G$ has no edges (i) and (ii) trivially hold. Assume thus that $G$ has at least one edge. Then, $\vartheta(G) < n$ and consequently $\Psi_\vartheta(G) \geq 2$.

(i) Let $(t, X)$ be a feasible solution for the program defining $\Psi_\vartheta(G)$; that is,

$$X \succeq 0,\ X_{uv} = 0\ (uv \in E(K_t \square G)),\ \mathrm{Tr}(X) = 1,\ \langle \mathbf{J}, X \rangle = n. \tag{5.8}$$

Here the matrix $X$ is indexed by $V(K_t \square G) = \cup_{p=1}^t V_p$ (recall (2.30)) and $t \in \mathbb{N}$, $t \geq 2$. As the program (5.8) is invariant under action of the group $\mathrm{Sym}(t)$, one may assume that $X$ is invariant under action of $\mathrm{Sym}(t)$. Then $X$ has the block form (2.11). Using Lemma 2.4.4, (5.8) can be rewritten as

$$\begin{gathered} A - B \succeq 0,\ A + (t-1)B \succeq 0,\ A_{ij} = 0\ (ij \in E(G)),\ \mathrm{diag}(B) = 0, \\ \mathrm{Tr}(A) = \tfrac{1}{t},\ \langle \mathbf{J}, A + (t-1)B \rangle = \tfrac{n}{t}. \end{gathered} \tag{5.9}$$

Lemma 5.1.7 implies $A + (t-1)B = \frac{1}{nt}\mathbf{J}$. Setting $U := nt(t-1)(A - B)$, we find

$$U = nt^2 A - \mathbf{J}. \tag{5.10}$$

One can verify that $(t, U)$ is feasible for the program (3.2) defining the parameter $\overline{\vartheta}(G)$. As $t \in \mathbb{N}$ this implies $\Psi_\vartheta(G) \geq \left\lceil \overline{\vartheta}(G) \right\rceil$. Conversely, let $(t, U)$ be feasible for (3.2) with $t$ integer. Define the matrices $A, B$ via the equations

$$A - B = \frac{1}{nt(t-1)}U \ \text{ and } A + (t-1)B = \frac{1}{nt}\mathbf{J} \tag{5.11}$$

and let $X$ be the corresponding block matrix as in (2.11). One can verify that (5.9) holds and thus (5.8) holds too. That is, $(t, X)$ is feasible for (5.8). Thus we have shown:

$$\Psi_\vartheta(G) = \min_{t \in \mathbb{N}} \ t \ \text{ s.t. } \mathrm{diag}(U) = (t-1)e,\ U_{ij} = -1\ (ij \in E(G)),\ U \succeq 0. \tag{5.12}$$

We now show $\Psi_\vartheta(G) \le \left\lceil \overline{\vartheta}(G) \right\rceil$. For this, set $t := \overline{\vartheta}(G)$ and take an optimal solution $U$ to the program (3.2). Then, setting $Y := \frac{1}{\lceil t \rceil - 1} U + \frac{\lceil t \rceil - t}{\lceil t \rceil - 1} \mathbf{I}$, the pair $(\lceil t \rceil, Y)$ is feasible for (5.12) with objective value $\lceil t \rceil$, which implies $\lceil t \rceil \ge \Psi_\vartheta(G)$. Thus equality $\left\lceil \overline{\vartheta}(G) \right\rceil = \Psi_\vartheta(G)$ holds.

The proof of (ii) is analogous to that of (i). Simply note that adding the condition $X \ge 0$ to (5.8) amounts to adding the condition $A, B \ge 0$ to (5.9) and thus, in view of (5.10), to adding the condition $U_{ij} \ge -1$ $(i, j \in V)$ to (5.12). $\square$

### 5.1.3 Computing $\Psi_\beta$

We consider here issues related to the computation of $\Psi_\beta(G)$. We assume throughout that $\beta(\cdot)$ satisfies (5.4). There is an obvious way to find $\Psi_\beta(G)$; namely, by computing $\beta(K_t \Box G)$ for each $t = 1, \ldots, n$. We now observe that, when $\beta(\cdot)$ is monotone nondecreasing (with respect to taking induced subgraphs), one can use binary search and it suffices to compute $\beta(K_t \Box G)$ for $O(\log n)$ instances of $t$.

**Lemma 5.1.8.** *Assume*

$$\beta(K_t \Box G) \le \beta(K_{t+1} \Box G) \text{ for all } t \in \mathbb{N}. \tag{5.13}$$

*Then* $\beta(K_t \Box G) = n \iff \Psi_\beta(G) \le t$.

*Proof.* The 'only if' part follows from the definition of $\Psi_\beta(G)$. For the 'if' assume $t_0 := \Psi_\beta(G) \le t$. Then $\beta(K_{t_0} \Box G) = n \le \beta(K_t \Box G)$ implies $\beta(K_t \Box G) = n$, since $\beta(K_t \Box G) \le \overline{\chi}(G) \le n$. $\square$

Under assumption (5.13) one can use binary search for computing $\Psi_\beta(G)$. Namely, given $t_0 \in [1, n]$, compute $\beta(K_{t_0} \Box G)$. There are two cases:
- Either $\beta(K_{t_0} \Box G) < n$. Then $\Psi_\beta(G) \ge t_0 + 1$ (by the above lemma) and we can now restrict the search to $t \in [t_0 + 1, n]$.
- Or $\beta(K_{t_0} \Box G) = n$. Then $\Psi_\beta(G) \le t_0$ and we can restrict the search to $t \in [1, t_0]$.

Therefore, one can find $\Psi_\beta(G)$ by computing $\beta(K_t \Box G)$ for $O(\log n)$ queries of $t$.

Observe that one may restrict the range of search for $t$. Suppose we know a lower bound $t_1$ and an upper bound $t_2$ on $\chi(G)$; that is, $t_1 \le \chi(G) \le t_2$. Then we may assume $t \le t_2$ in the definition of $\Psi_\beta(G)$ and if we add the condition $t \ge t_1$ then one still obtains a lower bound for $\chi(G)$. Therefore, we may restrict the binary search to $t \in [t_1, t_2]$. For instance, one can choose $t_1 = 3$ if $G$ is not bipartite, or $t_1(G) = \omega(G)$, and $t_2 = \Delta(G) + 1$ (or even $\Delta(G)$ by Brook's theorem (see [84]) if $G$ is not a clique or an odd circuit).

We next present an easy but quite surprising consequence of Lemma 5.1.2 concerning the complexity of graph parameters nested between the fractional chromatic and chromatic numbers or, more generally, in the interval $\left[\frac{|V(\cdot)|}{\omega(\cdot)}, \overline{\chi}(\cdot)\right]$. The key observation is that the operator $\Psi$ maps the whole interval to a single graph parameter (namely, the clique number $\omega(\cdot)$), which is hard to compute.

**Theorem 5.1.9.** *If $\beta(\cdot)$ is a graph parameter satisfying $\frac{|V(\cdot)|}{\omega(\cdot)} \leq \beta(\cdot) \leq \overline{\chi}(\cdot)$, then there is no algorithm permitting to compute $\beta(G)$ in time polynomial in $|V(G)|$ unless P=NP. As $\frac{|V(\cdot)|}{\omega(\cdot)} \leq \overline{\chi^*}(\cdot) \leq \overline{\chi}(\cdot)$, the same conclusion holds if $\overline{\chi^*}(\cdot) \leq \beta(\cdot) \leq \overline{\chi}(\cdot)$.*

*Proof.* Applying Lemma 5.1.2, we find that $\Psi_\beta(\cdot) = \omega(\cdot)$. Suppose one can compute $\beta(G)$ in time $f(n)$ where $f$ is a polynomial in $n = |V(G)|$. Then one can compute $\Psi_\beta(G) = \omega(G)$ in time $\sum_{l=1}^n f(ln)$, thus polynomial in $n$. As computing the clique number is an NP-hard problem [30], this implies P=NP. $\square$

**Corollary 5.1.10.** *If $\beta(\cdot)$ is a graph parameter satisfying $\frac{|V(\cdot)|}{\chi(\cdot)} \leq \beta(\cdot) \leq \alpha(\cdot)$, then there is no algorithm permitting to compute $\beta(G)$ in time polynomial in $|V(G)|$ unless P=NP.*

Let us mention a few graph parameters that are known to lie within the 'hard' interval $[\chi^*, \chi]$. Hence none of them can be computed in polynomial time unless P=NP.

Such result was proved already for the *circular chromatic number* (or *star chromatic number*) $\chi_c(G)$ in [10]. It was introduced by Vince [93] and further studied in [10], [95]. Given $r \in \mathbb{R}$, $r \geq 2$, a function $f : V(G) \to [0, r)$ is said to be a $r$-colouring if $1 \leq |f(u) - f(v)| \leq r - 1$ for all edges $uv \in E(G)$. Then $\chi_c(G)$ is defined as the infimum of all $r$ for which $G$ has a $r$-colouring. The following holds: $\chi(G) - 1 < \chi_c(G) \leq \chi(G)$ and $\chi^*(G) \leq \chi_c(G) \leq \chi(G)$ (see e.g. [95]).

Another graph parameter which lies in $[\chi^*, \chi]$ is the *local chromatic number* $\chi_{\text{loc}}(G)$, introduced in [27] as the minimum over all proper colourings of $G$ of the largest number of colours used to colour the neighborhood $N_G(v) = \{w \in V(G) \mid vw \in E(G)\}$ of any vertex $v \in V(G)$. Obviously, $\chi_{\text{loc}}(G) \leq \chi(G)$ (the gap between the two parameters can in fact be arbitrarily large [27]) and Körner et al. [55] show that $\chi^*(G) \leq \chi_{\text{loc}}(G)$.

The *independence ratio* of a graph $G$ is $i(G) := \frac{\alpha(G)}{|V(G)|}$ and its *Hall ratio* is $\rho(G) := \max_{H \subseteq G} \frac{|V(H)|}{\alpha(H)}$, where the maximum is taken over all subgraphs of $G$. Set $G^{\square 1} = G$, and for an integer $k \geq 2$ set $G^{\square k} := G \square G^{\square (k-1)}$. Note that $G^{\square k}$ is the graph obtained by taking the Cartesian product of $k$ copies of $G$. The *ultimate independence ratio* $I(G)$ and the *ultimate Hall ratio* $h_\square(G)$ are defined respectively as $I(G) := \lim_{k \to \infty} i(G^{\square k})$ and $h_\square(G) := \lim_{k \to \infty} \rho(G^{\square k})$. These graph parameters are studied e.g. in [41], [43], [88]. In particular, the following relations with fractional and circular chromatic numbers are shown there:

$$\chi^*(G) \leq \frac{1}{I(G)} = h_\square(G) \leq \chi_c(G) \leq \chi(G)$$

(see [95] for the inequality $1 \leq I(G)\chi_c(G)$).

### 5.1.4 Semidefinite programming formulation for the new bounds

Next we show that $\Psi_\beta(G)$ can be formulated via a single semidefinite program when $\beta(\cdot)$ is given by a semidefinite program satisfying certain assumptions. Namely, our construction applies to the case when the semidefinite program defining $\beta(\cdot)$ involves at least one equality constraint of the form $\langle A, X\rangle = 1$ with $A \succeq 0$. Then one may assume without loss of generality that all other (in)equality constraints in the program are homogeneous, i.e., of the form $\langle B, X\rangle \geq 0$. (Write any equation $\langle B, X\rangle = 0$ as two opposite inequalities $\langle -B, X\rangle \geq 0$ and $\langle B, X\rangle \geq 0$.) So let us assume that, for an arbitrary graph $H$, we can express $\beta(H)$ as

$$\beta(H) = \max\ \langle C(H), X(H)\rangle \ \text{ s.t. } \begin{array}{l} \langle A(H), X(H)\rangle = 1 \\ \mathcal{B}(H)(X(H)) \geq 0 \\ X(H) \succeq 0, \end{array} \tag{5.14}$$

where $C(H)$ and $A(H)$ are constant symmetric $n \times n$ matrices, $\mathcal{B}(H) : S_n \to \mathbb{R}^{d(H)}$ is a linear operator, and $X(H)$ is the matrix variable. Note that $d(\cdot)$ depends on $H$, e.g. $d(H) = 2|E(H)|$ in the formulation of $\vartheta(H)$. Moreover we assume that

$$A(H) \succeq 0, \tag{5.15}$$

$$\langle A(H), X(H)\rangle = 0 \Longrightarrow \langle C(H), X(H)\rangle = 0. \tag{5.16}$$

Note that Assumptions (5.13), (5.14), (5.15),(5.16) hold, e.g., for $\vartheta(\cdot)$, or for the hierarchy of Lasserre considered in Section 4.1.3, and the new block diagonal hierarchy presented in Section 4.1.4. To see this observe first that the bounds (4.20) and (4.23) satisfy (5.13). Next, in the SDPs (4.20) and (4.23) we can identify $\langle A(H), X(H)\rangle$ with $y_0$, $\langle C(H), X(H)\rangle$ with $\sum_{i\in V(H)} y_i$ and the conditions $\mathcal{B}(H)(X(H)) \geq 0$ with $y_{ij} = 0$ ($ij \in E$). Then the condition (5.15) holds trivially while $y_0 = 0$ implies $y_i = 0$ ($i \in V(H)$) and thus (5.16).

Recall that our operator $\Psi$ maps $\beta(\cdot)$ in the following way:

$$\Psi_\beta(G) \ := \ \begin{array}{ll} \min & t \\ \text{s.t.} & \beta(G_t) = n \end{array} \ = \ \begin{array}{ll} \min & t \\ \text{s.t.} & \langle C(G_t), X(G_t)\rangle = n \\ & \langle A(G_t), X(G_t)\rangle = 1 \\ & \mathcal{B}(G_t)(X(G_t)) \geq 0 \\ & X(G_t) \succeq 0. \end{array} \tag{5.17}$$

Here we use the more concise notation $G_t := K_t \square G$. Let us define

$$\Phi_\beta(G) := \min\ \sum_{t=1}^n t\langle A(G_t), X(G_t)\rangle \ \text{ s.t. } \begin{array}{l} \sum_{t=1}^n \langle C(G_t), X(G_t)\rangle = n \\ \sum_{t=1}^n \langle A(G_t), X(G_t)\rangle = 1 \\ \mathcal{B}(G_t)(X(G_t)) \geq 0 \ (t = 1, ..., n) \\ X(G_t) \succeq 0 \ (t = 1, ..., n). \end{array} \tag{5.18}$$

**Theorem 5.1.11.** *Under assumptions (5.14),(5.15) and (5.16),* $\Phi_\beta(G) = \Psi_\beta(G)$.

*Proof.* Take a feasible solution $(t, X(G_t))$ for the program (5.17) and for $k \neq t$ set $X(G_k) := 0$. In this way one obtains a feasible solution for (5.18) with the same objective value as (5.17), which shows $\Phi_\beta(G) \leq \Psi_\beta(G)$. Conversely, let $X(G_t)$ ($t = 1, ..., n$) be a feasible solution for (5.18) and set $a_t := \langle A(G_t), X(G_t)\rangle$. Thus $a_t \geq 0$ since $A(G_t) \succeq 0$ (by assumption (5.15)) and $\sum_t a_t = 1$. Consider $t$ for which $a_t > 0$. As $\left\langle A(G_t), \frac{X(G_t)}{a_t}\right\rangle = 1$, $\frac{X(G_t)}{a_t}$ is feasible for (5.14) (with $H = G_t$) which implies $\left\langle C(G_t), \frac{X(G_t)}{a_t}\right\rangle \leq \beta(G_t) \leq n$;

moreover, equality $\left\langle C(G_t), \frac{X(G_t)}{a_t}\right\rangle = n$ implies $\beta(G_t) = n$ and thus $\Psi_\beta(G) \leq t$. Now we have

$$n = \sum_t \langle C(G_t), X(G_t)\rangle = \sum_{t|a_t>0} a_t \left\langle C(G_t), \frac{X(G_t)}{a_t}\right\rangle \leq \left(\sum_{t|a_t>0} a_t\right) n = n.$$

(Here we used assumption (5.16) for the second equality.) Therefore, equality holds throughout which implies $\Psi_\beta(G) \leq t$ whenever $a_t > 0$. Hence, $\sum_t ta_t = \sum_{t|a_t>0} ta_t \geq \Psi_\beta(G)(\sum_{t|a_t>0} a_t) = \Psi_\beta(G)$ which gives $\Phi_\beta(G) \geq \Psi_\beta(G)$. □

Hence, under the assumptions (5.14),(5.15) and (5.16), the parameter $\Psi_\beta(G)$ can be formulated via the semidefinite program (5.18) which involves a block-diagonal matrix with diagonal blocks $X(G_1), \ldots, X(G_n)$, each $X(G_t)$ being the matrix variable involved in the program (5.14) for the graph $H = G_t$. For instance, if (5.14) involves a matrix variable of order $f(V(H))$, then (5.18) involves a block-diagonal matrix with block sizes $f(n), f(2n), \ldots, f(n^2)$. As explained in Section 5.1.3 one can reduce the size of the program (5.18) by restricting the range of $t$ in program (5.18) to $t \in [t_1, t_2]$ where $t_1 \leq \chi(G) \leq t_2$.

### 5.1.5 Quadratic programming formulation for $\chi(G)$

The technique used in Section 5.1.4 can also be applied to derive (quadratically constrained) quadratic and copositive programming formulations for the chromatic number.

Recall from Section 4.2 that

$$\frac{1}{\alpha(G)} = \min\ x^T(\mathbf{I} + A_G)x \ \text{ s.t. } x \in \mathbb{R}_+^{V(G)},\ e^Tx = 1 \tag{5.19}$$

or, equivalently,

$$\alpha(G) = \min\ t \ \text{ s.t. } t(\mathbf{I} + A_G) - \mathbf{J} \in \mathcal{C}_n. \tag{5.20}$$

Using (5.19), we can rewrite the program (5.2) as

$$\chi(G) = \min\ t \text{ s.t. } x_t^T(\mathbf{I} + A_{G_t})x_t = \tfrac{1}{n},\ e_t^Tx_t = 1,\ x_t \in \mathbb{R}_+^{V(G_t)}. \tag{5.21}$$

Here and below $e_t$ denotes the all-ones vector in $\mathbb{R}^{V(G_t)}$. Using the idea from Section 5.1.4 let us define

$$\begin{aligned} \Phi_1(G) \quad := \quad \min \quad & \sum_{t=1}^n t(e_t^Tx_t)^2 \\ \text{s.t.} \quad & \sum_{t=1}^n (e_t^Tx_t)^2 = 1 \\ & \sum_{t=1}^n x_t^T(\mathbf{I} + A_{G_t})x_t = \frac{1}{n} \\ & x_t \in \mathbb{R}_+^{V(G_t)}\ (t = 1, \ldots, n). \end{aligned} \tag{5.22}$$

**Proposition 5.1.12.** $\Phi_1(G) = \chi(G)$.

*Proof.* Taking a feasible solution $(t, x_t)$ for the program (5.21) and setting $x_k = 0$ for $k \neq t$, we obtain a feasible solution for (5.22) with objective value $t$. Thus, $\Phi_1(G) \leq \chi(G)$. Conversely, let $x_t$ $(t = 1, ..., n)$ be feasible for (5.22). Then

$$\frac{1}{n} = \sum_t x_t^T(\mathbf{I}+A_{G_t})x_t = \sum_{t|x_t\neq 0} \frac{x_t^T}{e_t^T x_t}(\mathbf{I}+A_{G_t})\frac{x_t}{e_t^T x_t}(e_t^T x_t)^2 \geq \frac{1}{n}\sum_{t|x_t\neq 0}(e_t^T x_t)^2 = \frac{1}{n}.$$

We have used $\frac{x_t^T}{e_t^T x_t}(\mathbf{I}+A_{G_t})\frac{x_t}{e_t^T x_t} \geq \frac{1}{\alpha(G_t)} \geq \frac{1}{n}$. Hence equality holds throughout, which implies $\alpha(G_t) = n$ if $x_t \neq 0$ and thus $\chi(G) \leq t$ if $x_t \neq 0$. Therefore,

$$\sum_t t(e_t^T x_t)^2 = \sum_{t|x_t\neq 0} t(e_t^T x_t)^2 \geq \chi(G)\sum_{t|x_t\neq 0}(e_t^T x_t)^2 = \chi(G).$$

This shows $\Phi_1(G) \geq \chi(G)$. $\square$

Up to rescaling, we obtain the following formulation for $\chi(G)$ involving only quadratic constraints:

$$\begin{aligned} \chi(G) \quad = \quad \min \quad & \tfrac{1}{n^2}\sum_{t=1}^n t(e_t^T x_t)^2 \\ \text{s.t.} \quad & \sum_{t=1}^n (e_t^T x_t)^2 = n^2 \\ & \sum_{t=1}^n x_t^T(\mathbf{I}+A_{G_t})x_t = n \\ & x_t \in \mathbb{R}_+^{V(G_t)} \ (t = 1, \ldots, n). \end{aligned} \tag{5.23}$$

It is not difficult to verify that the above program remains a formulation of $\chi(G)$ if we replace the condition $x_t \geq 0$ (for all $t$) by the condition $x_t$ is 0/1 valued (for all $t$). Therefore this gives a 0/1 (quadratically constrained) quadratic programming formulation for the chromatic number involving $O(n^3)$ variables.

### 5.1.6 Copositive programming formulation for $\chi(G)$

Starting from (5.23), we can now derive a copositive programming formulation for $\chi(G)$. Namely, consider the program

$$\begin{aligned} \Phi_2(G) \quad := \quad \min \quad & \tfrac{1}{n^2}\textstyle\sum_{t=1}^n t\langle \mathbf{J}, X_t\rangle \\ \text{s.t.} \quad & \sum_{t=1}^n \langle \mathbf{J}, X_t\rangle = n^2 \\ & \sum_{t=1}^n \langle \mathbf{I}+A_{G_t}, X_t\rangle = n \\ & X_t \in \mathcal{C}_{nt}^* \ (t = 1, \ldots, n). \end{aligned} \tag{5.24}$$

**Proposition 5.1.13.** $\Phi_2(G) = \chi(G)$.

*Proof.* The formulation (5.23) for $\chi(G)$ implies directly $\Phi_2(G) \leq \chi(G)$. Conversely, let $X_t$ $(1 \leq t \leq n)$ be a feasible solution for (5.24). Consider $t$ for which $X_t \neq 0$. Say, $X_t = \sum_{i_t} x_{i_t}x_{i_t}^T$ where $x_{i_t} \geq 0$, $x_{i_t} \neq 0$ for all $i_t$.

Thus $\lambda_{i_t} := \sqrt{\langle \mathbf{J}, x_{i_t} x_{i_t}^T \rangle} = e_t^T x_{i_t} > 0$. Set $y_{i_t} := \frac{x_{i_t}}{\lambda_{i_t}}$. By assumption, we have $\sum_t \langle n(\mathbf{I} + A_{G_t}) - \mathbf{J}, X_t \rangle = 0$. By (5.20), each matrix $n(\mathbf{I} + A_{G_t}) - \mathbf{J}$ is copositive, since $n \geq \alpha(G_t)$. This implies $\langle n(\mathbf{I} + A_{G_t}) - \mathbf{J}, X_t \rangle = 0$ and thus $\langle n(\mathbf{I} + A_{G_t}) - \mathbf{J}, x_{i_t} x_{i_t}^T \rangle = 0$ for all $i_t$. From this follows that $\langle \mathbf{I} + A_{G_t}, y_{i_t} y_{i_t}^T \rangle = \frac{1}{n}$ for all $i_t$. As $e_t^T y_{i_t} = 1$, $y_{i_t}$ is feasible for the program (5.21), implying $\chi(G) \leq t$ whenever $X_t \neq 0$. Now, $(1/n^2) \sum_t t \langle \mathbf{J}, X_t \rangle \geq (1/n^2) \chi(G) \sum_t \langle \mathbf{J}, X_t \rangle = \chi(G)$, giving $\Phi_2(G) \geq \chi(G)$. □

Rewriting the condition $\sum_t \langle \mathbf{I} + A_{G_t}, X_t \rangle n$ as $\sum_t \langle n(\mathbf{I} + A_{G_t}) - \mathbf{J}, X_t \rangle = 0$, the dual conic program of (5.24) reads:

$$\max_{y,z} \; y \text{ s.t. } \frac{1}{n^2}(t-y)\mathbf{J} + z(n(\mathbf{I} + A_{G_t}) - \mathbf{J}) \in \mathcal{C}_{nt}, \; 1 \leq t \leq n. \tag{5.25}$$

There is no duality gap since the program (5.25) is strictly feasible. Thus (5.25) is yet another formulation of $\chi(G)$. This opens the road to another type of hierarchy of relaxations for $\chi(G)$, obtained by approximating the copositive cone by tractable subcones as suggested by Parrilo [75]. This type of approach based on copositive programming has been studied e.g. in [9] for standard quadratic optimization problems, in [50, 39, 76] for the stable set problem and recently in [24] for the colouring problem. We will come back to it in Section 5.2.4.

## 5.2 Hierarchies of semidefinite bounds for $\chi^*(G)$ and $\chi(G)$

We have seen in the previous section how to construct semidefinite programming lower bounds for the chromatic number of a graph from semidefinite programming upper bounds on the stability number. Several hierarchies of such upper bounds for the stability number have been presented in the previous chapter. As we have seen, Lasserre's hierarchy proposed in [57] gives the tightest bounds. For this reason we focus in this section on this hierarchy and its cheaper variant from Section 4.1.4. We show how they can be used and transformed to produce hierarchies of lower bounds for the (fractional) chromatic number. We also discuss the link with another hierarchy recently proposed by Dukanovic and Rendl [24] based on copositive programming.

### 5.2.1 Lasserre type hierarchies towards $\chi^*(G)$

Given a graph $G = (V, E)$ and an integer $r \geq 1$, define the parameters

$$\begin{aligned} \psi^{(r)}(G) := \; & \min t \quad \text{s.t.} \quad x \in Q_r(G), \; x_{\mathbf{0}} = t, \; x_i = 1 \; (i \in V), \\ = \; & \min t \quad \text{s.t.} \quad y \in \mathbb{R}^{\mathcal{P}_{\leq 2r}(V)}, \; y_{\mathbf{0}} = t, \; y_i = 1 \; (i \in V), \\ & \qquad\qquad\quad y_{ij} = 0 \; (ij \in E), \; M_t(y) \succeq 0 \end{aligned} \tag{5.26}$$

and

$$\begin{aligned}\widetilde{\psi}^{(r)}(G) :=\ & \min t \quad \text{s.t.} \quad x \in L_r(G),\ x_{\mathbf{0}} = t,\ x_i = 1\ (i \in V),\\ =\ & \min t \quad \text{s.t.} \quad y \in \mathbb{R}^{\mathcal{P}_{\leq 2r}(V)},\ y_i = 1\ (i \in V),\ y_{ij} = 0\ (ij \in E),\\ & \qquad\qquad\quad y_{\mathbf{0}} = t, M(T; y) \succeq 0\ \forall T \subseteq V \text{ with } |T| = t-1.\end{aligned} \tag{5.27}$$

Note that one can avoid the variable $t$, simply by replacing $t$ by $x_{\mathbf{0}}$ in the objective function. We choose these formulations in order to have a unified presentation of the various bounds.; compare e.g. with (5.2), (5.5), (5.28), (5.29), (5.41). The minimum is attained in programs (5.26) and (5.27), and moreover, for fixed $r$, one can compute $\psi^{(r)}(G)$ and $\widetilde{\psi}^{(r)}(G)$ to an arbitrary precision in polynomial time. Obviously, $\widetilde{\psi}^{(r)}(G) \geq \psi^{(r)}(G)$ since $Q_r(G) \subseteq L_r(G)$, with equality for $r = 1$ as $Q_1(G) = L_1(G)$.

**Theorem 5.2.1.** *The parameters $\psi^{(r)}(G)$ satisfy:*

*(a)* $\psi^{(r)}(G) \leq \psi^{(r+1)}(G)$,

*(b)* $\psi^{(1)}(G) = \overline{\vartheta}(G)$,

*(c)* $\overline{\vartheta^{+\triangle}}(G) \leq \psi^{(2)}(G)$,

*(d)* $\psi^{(r)}(G) \leq \chi^*(G)$, *with equality if* $r \geq \alpha(G)$,

*(e)* $\psi^{(r)}(G)\mathrm{las}^{(r)}(G) \geq |V(G)|$, *with equality if $G$ is vertex-transitive.*

*Proof.* (a) is obvious. For (b), let $x$ be optimal for (5.26) with $r = 1$ and let $y$ be such that $y \in \mathbb{R}^{\mathcal{P}_{\leq 2}(V)}, y_{\mathbf{0}} = x_{\mathbf{0}},\ y_i = x_i\ (i \in V),\ y_{ij} = 0\ (ij \in E),\ M_1(y) \succeq 0$ (recall the definition of $Q_r$ from (4.43)). Then $M_1(y) = \begin{pmatrix} t & e^T \\ e & M \end{pmatrix} \succeq 0$ or, equivalently, $M - \frac{1}{t}ee^T \succeq 0$. After setting $U := t(M - \frac{1}{t}ee^T) = tM - \mathbf{J}$, we can rewrite the program for $\psi^{(1)}(G)$ in the following way

$$\psi^{(1)}(G) = \min\ t \ \text{ s.t. } \ \begin{array}{l} U_{ii} = t-1 \\ U_{ij} = -1\ (ij \in E) \\ U \succeq 0. \end{array}$$

Thus, in view of (3.2), $\psi^{(1)}(G) = \overline{\vartheta}(G)$.
(c) Assume $(t, x)$ is feasible for the program defining $\psi^{(2)}(G)$ and let $y$ be such that $y \in \mathbb{R}^{\mathcal{P}_4(V)}, y_{\mathbf{0}} = x_{\mathbf{0}},\ y_i = x_i\ (i \in V),\ y_{ij} = 0\ (ij \in E),\ M_2(y) \succeq 0$. Consider the principal submatrix $Y$ of $M_2(y)$ indexed by $\{k, ij, ik, jk\}$ where $i, j, k$ are distinct elements of $V$ and the vector $w := (1, 1, -1, -1)^T$. Then, $w^T Y w \geq 0$ gives $y_{ik} + y_{jk} - y_{ij} \leq 1$. Setting $U := \frac{t}{t-1}\left((y_{ij})_{i,j=1}^n - \frac{1}{t}\mathbf{J}\right)$, one can now verify that $(t, U)$ is feasible for the program defining $\overline{\vartheta^{+\triangle}}(G)$, which shows the result.
(d) Let $\lambda$ be an optimum solution for the minimization program defining $\chi^*(G)$ (recall (3.12)). That is, $e^T\lambda = \chi^*(G)$, $\sum_{S \text{ stable}} \lambda_S \chi^S = e$ and $\lambda \geq 0$. For $r \in \mathbb{N}$, define $y := \sum_{S \text{ stable}} \lambda_S \chi^{S,r}$ and $x \in \mathcal{P}_{\leq 1}(V)$ with $x_I := y_I$ for $I \in \mathcal{P}_{\leq 1}(V)$. Observe that $x$ is feasible for (5.26) with objective value $\chi^*(G)$. This shows $\psi^{(r)}(G) \leq \chi^*(G)$. Assume now $r \geq \alpha(G)$ and consider an optimum solution $x$ for (5.26). Setting $\tilde{x} := \frac{1}{\psi^{(r)}(G)} x$, we see that $\tilde{x}$ is feasible for the program

(4.20), i.e., $\tilde{x} \in Q_r(G) = \mathrm{ST}(G)$. Hence $\tilde{x} = \sum_{S \text{ stable}} \lambda_S \chi^S$ for some $\lambda_S \geq 0$ with $\sum_S \lambda_S = 1$. Since

$$\psi^{(r)}(G)\tilde{y} = \psi^{(r)}(G) \sum_S \lambda_S \begin{pmatrix} 1 \\ \chi^S \end{pmatrix} = \begin{pmatrix} \psi^{(r)}(G) \\ e \end{pmatrix},$$

we have $\chi^*(G) \leq \psi^{(r)}(G)$.

(e) Take again an optimum solution $x$ for (5.26) and let $n = |V(G)|$. Since $\frac{1}{\psi^{(r)}(G)}x$ is feasible for (4.20) with objective value $\frac{n}{\psi^{(r)}(G)}$, we get $\mathrm{las}^{(r)}(G) \geq \frac{n}{\psi^{(r)}(G)}$. Assume that $G$ is vertex-transitive. Then there exists an optimum solution $x$ for (4.20) which is invariant under action of the automorphism group of $G$. In particular, $x_i = x_j$ for all $i, j \in V$ and thus $x_i = \frac{\mathrm{las}^{(r)}(G)}{n}$ for all $i \in V$. Then $\frac{n}{\mathrm{las}^{(r)}(G)}x$ is feasible for (5.26), yielding $\psi^{(r)}(G) \leq \frac{n}{\mathrm{las}^{(r)}(G)}$. $\square$

**Theorem 5.2.2.** *The parameters $\widetilde{\psi}^{(r)}(G)$ satisfy:*

*(a)* $\widetilde{\psi}^{(r)}(G) \leq \widetilde{\psi}^{(r+1)}(G)$*,*

*(b)* $\widetilde{\psi}^{(1)}(G) = \overline{\vartheta}(G)$*,*

*(c)* $\overline{\vartheta^{+}}(G) \leq \widetilde{\psi}^{(2)}(G)$*,*

*(c')* $\overline{\vartheta^{+\triangle}}(G) \leq \widetilde{\psi}^{(3)}(G)$*,*

*(d)* $\widetilde{\psi}^{(r)}(G) \leq \chi^*(G)$*, with equality if* $r \geq \alpha(G)$*,*

*(e)* $\widetilde{\psi}^{(r)}(G)\ell^{(r)}(G) \geq |V(G)|$*, with equality if $G$ is vertex-transitive.*

The proof of Theorem 5.2.2 is along the same lines with the proof of Theorem 5.2.1. The only difference is that we are not able to compare $\overline{\vartheta^{+\triangle}}(G)$ with the second bound in the hierarchy (5.27), but with the third one instead. However, the second bound $\widetilde{\psi}^{(2)}(G)$, introduced as $\psi(G)$ in [37] and computed for some Hamming and Kneser graph instances in [36] (see Chapter 6), remains at least as strong as $\overline{\vartheta^{+}}(G)$.

Theorem 5.2.1 and Theorem 5.2.2 show that the reciprocity relations from Theorem 3.3.1 for the pairs $(\vartheta, \overline{\vartheta}) = (\mathrm{las}^{(1)}, \psi^{(1)}) = (\ell^{(1)}, \widetilde{\psi}^{(1)})$ and $(\alpha, \chi^*) = (\mathrm{las}^{(r)}, \psi^{(r)}) = (\ell^{(r)}, \widetilde{\psi}^{(r)})$ (for $r$ large, $r \geq \alpha(G)$) extend to any order $r$ pair $(\mathrm{las}^{(r)}, \psi^{(r)})$ and $(\ell^{(r)}, \widetilde{\psi}^{(r)})$ in the hierarchies.

### 5.2.2 The hierarchies $\Psi_{\mathrm{las}^{(r)}}(G)$ and $\Psi_{\ell^{(r)}}(G)$ towards the chromatic number

By applying the operator $\Psi$ to the hierarchies $\mathrm{las}^{(r)}(\cdot)$ and $\ell^{(r)}(\cdot)$ introduced in Chapter 4, we obtain the following hierarchies of lower bounds for $\chi(G)$:

$$\begin{array}{rlll} \Psi_{\mathrm{las}^{(r)}}(G) & = \min \ t & \text{s.t.} & \mathrm{las}^{(r)}(G_t) = n \\ & = \min \ t & \text{s.t.} & y_{\mathbf{0}} = 1, \ \sum_{u \in V(G_t)} y_u = n, \\ & & & y_{uv} = 0 \ (uv \in E(G_t)), \\ & & & M_r(y) \succeq 0 \end{array} \tag{5.28}$$

where the variable $y$ is indexed by $\mathcal{P}_{\leq 2r}(V(G_t))$, and

$$\begin{array}{lll} \Psi_{\ell^{(r)}}(G) & = \min\ t\ \text{ s.t.} & \ell^{(r)}(G_t) = n \\ & = \min\ t\ \text{ s.t.} & y_{\mathbf{0}} = 1,\ \sum_{u \in V(G_t)} y_u = n, \\ & & y_{uv} = 0\ (uv \in E(G_t)), \\ & & M(T; y) \succeq 0\ (T \subseteq V(G_t),\ |T| = r-1) \end{array} \tag{5.29}$$

where the variable $y$ is indexed by $\mathcal{P}_{\leq r+1}(V(G_t))$. Recall that the hierarchy $\text{las}^{(r)}(\cdot)$ refines the hierarchy $\ell^{(r)}(\cdot)$ and that the operator $\Psi$ is monotone non-increasing. Thus $\Psi_{\text{las}^{(r)}}(G) \geq \Psi_{\ell^{(r)}}(G)$. As $\alpha(G_t) \leq n$, from Corollary 4.1.8 we get $\ell^{(n)}(G_t) = \text{las}^{(n)}(G_t) = \alpha(G_t)$ for all $t \in \mathbb{N}$. Therefore, (2.31) implies:

**Proposition 5.2.3.** $\Psi_{\text{las}^{(n)}}(G) = \Psi_{\ell^{(n)}}(G) = \chi(G)$. □

In fact, the hierarchy $\Psi_{\text{las}^{(r)}}$ refines the hierarchy $\psi^{(r)}$, and $\Psi_{\ell^{(r)}}$ refines the hierarchy $\widetilde{\psi}^{(r)}$.

**Proposition 5.2.4.** *For any integer $r \geq 1$,*

*(a)* $\psi^{(r)}(G) \leq \Psi_{\text{las}^{(r)}}(G)$, *and*

*(b)* $\widetilde{\psi}^{(r)}(G) \leq \Psi_{\ell^{(r)}}(G)$.

The proof uses the following lemma.

**Lemma 5.2.5.** *If $(t, y)$ is feasible for the program (5.28), where $y$ is invariant under action of the symmetric group* $\text{Sym}(t)$*, then $y_u = \frac{1}{t}$ for all $u \in V(G_t)$.*

*Proof.* Let $(t, y)$ be feasible for (5.28), such that $y$ is invariant under action of the symmetric group $\text{Sym}(t)$. Observe that $y \in \mathbb{R}^{\mathcal{P}_{\leq 2r}(V(G_t))}$ satisfies $y_{\mathbf{0}} = 1$, $y_{uv} = 0$ $(uv \in E(G_t))$, $\sum_{u \in V(G_t)} y_u = n$, and $M_r(y) \succeq 0$. Let $X$ denote the principal submatrix of $M_r(y)$ indexed by $\mathcal{P}_{\leq 1}(V(G_t))$. As $y$ is invariant, with respect to the partition of $\mathcal{P}_{\leq 1}(V(G_t)) \sim \{\mathbf{0}\} \cup V(G_t)$ into $\{\mathbf{0}\} \cup V_1 \cup \ldots \cup V_t$ (recall (2.30)), the matrix $X$ has the block form

$$\begin{pmatrix} 1 & a^T & a^T & \ldots & a^T \\ a & A & B & \ldots & B \\ a & B & A & \ldots & B \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & \underbrace{B \quad B \quad \ldots \quad A}_{t \text{ blocks}} \end{pmatrix} \tag{5.30}$$

where $a = \text{diag}(A)$, $\text{diag}(B) = 0$, $A_{ij} = 0$ for $ij \in E(G)$, and $e^T a = \frac{n}{t}$. By taking the Schur complement with respect to the left upper corner and using Lemma 2.4.4, we have $A + (t-1)B - taa^T \succeq 0$. This implies $\langle \mathbf{J}, A + (t-1)B \rangle \geq t(e^T a)^2 = \frac{n^2}{t}$. On the other hand, by Lemma 5.1.7, $\langle \mathbf{J}, A + (t-1)B \rangle \leq n\text{Tr}(A + (t-1)B) = n\text{Tr}(A) = \frac{n^2}{t}$. Hence equality holds, implying $A + (t-1)B = \frac{1}{t}\mathbf{J}$ and thus $a = \frac{1}{t}e$. This shows $y_u = \frac{1}{t}$ for all $u \in V(G_t)$. □

We now prove Proposition 5.2.4 (a) and the proof of (b) is analogous. Note that we may assume w.l.o.g. that $y$ in programs (5.28) (or (5.29)) is invariant under action of the symmetric group $\text{Sym}(t)$.

*Proof of Proposition 5.2.4 (a).* Let $(t, y)$ be feasible for (5.28) and $y$ be invariant under action of the symmetric group $\mathrm{Sym}(t)$. Define the vector $\tilde{y} \in \mathbb{R}^{\mathcal{P}_{\leq 2r}(V)}$ with $I$th entry $\tilde{y}_I := ty_{\{pi|i\in I\}}$ for $I \in \mathcal{P}_{\leq 2r}(V)$ (where $p$ is any fixed integer in $\{1, \ldots, t\}$) and $y$ is feasible (5.28) and invariant under action of $\mathrm{Sym}(t)$. Then, $M_r(\tilde{y}) \succeq 0$, since it coincides with the principal submatrix of $M_r(ty)$ indexed by $\{\mathbf{0}\} \cup \{\{pi \mid i \in I\} \mid I \in \mathcal{P}_{\leq r}(V) \backslash \{\mathbf{0}\}\}$. Define $x \in \mathcal{P}_{\leq 1}(V)$ with $x_I := \tilde{y}_I$ for $I \in \mathcal{P}_{\leq 1}(V)$. Since, $x_{\mathbf{0}} = t$ and $x_i = 1$ for $i \in V$, we have that $(t, x)$ is feasible for the program (5.26), which implies $\psi^{(r)}(G) \leq \Psi_{\mathrm{las}^{(r)}}(G)$. $\square$

Summarizing we have shown the following relations among the graph parameters:

$$\frac{|V(G)|}{\mathrm{las}^{(r)}(G)} \leq \psi^{(r)}(G) \leq \Psi_{\mathrm{las}^{(r)}}(G) \leq \chi(G), \tag{5.31}$$

$$\frac{|V(G)|}{\ell^{(r)}(G)} \leq \widetilde{\psi}^{(r)}(G) \leq \Psi_{\ell}^{(r)}(G) \leq \chi(G). \tag{5.32}$$

Let us point out again that, while $\psi^{(r)}(G)$ and $\widetilde{\psi}^{(r)}(G)$ remain below the fractional chromatic number $\chi^*(G)$, $\Psi_{\mathrm{las}}^{(r)}(G)$ and $\Psi_{\ell^{(r)}}(G)$ may reach the chromatic number $\chi(G)$.

### 5.2.3 Exploiting symmetry to compute the bound $\Psi_{\ell^{(2)}}(G)$

We group here some observations about the complexity of computing the graph parameter $\Psi_{\ell^{(2)}}(\cdot)$. We show how one can exploit symmetry, present in the structure of the matrix involved in the program defining the parameter or in the graph instance, in order to reduce the size of the program. In this way we will be able to compute the graph parameters for certain large graphs (with as much as $2^{20}$ nodes for certain Hamming graphs). We give more details and report about computational results for Hamming and Kneser graphs in Chapter 6.

In order to determine $\Psi_{\ell^{(2)}}(G)$, we need to compute the parameter

$$\begin{aligned} \ell^{(2)}(G_t) = \quad & \max \textstyle\sum_{i\in V(G_t)} x_i \quad \text{s.t.} \quad x \in L_2(G_t),\ x_{\mathbf{0}} = 1 \\ = \quad & \max \textstyle\sum_{i\in V(G_t)} y_i \quad \text{s.t.} \quad y \in \mathbb{R}^{\mathcal{P}_{\leq 3}(V(G_t))},\ y_{\mathbf{0}} = 1, \\ & \qquad\qquad\qquad\qquad\quad y_{ij} = 0\ (ij \in E(G_t)), \\ & \qquad\qquad\qquad\qquad\quad M(\{u\}; y) \succeq 0\ (u \in V(G_t)). \end{aligned} \tag{5.33}$$

(recall (4.23)) for several queries of $t \in \mathbb{N}$. To avoid trivial technicalities we assume $t \geq 2$ throughout the section. As was observed in the previous section, the program (5.33) is invariant under action of $\mathrm{Sym}(t)$, hence we may assume that $y$ is invariant under action of $\mathrm{Sym}(t)$. Moreover, it suffices to require the condition $M(\{u\}; y) \succeq 0$ for all $u \in V_1$ instead of for all $u \in V(G_t)$. (Recall from (5.3) that $V_1 = \{1i \mid i \in V\}$ denotes the 'first layer' of the nodeset $V(G_t) = \{pi \mid p = 1, \ldots, t,\ i \in V\}$ of $G_t$.) Furthermore, when $G$ is vertex-transitive, it suffices to require $M(\{u\}; y) \succeq 0$ for *one* choice of $u \in V_1$ instead of for *all* $u \in V_1$.

Recall from Section 4.1.4 that $M(\{u\}; y)$ is the matrix indexed by $\mathcal{A}_{\mathbf{0}} \cup \mathcal{A}_{\{u\}} = \mathcal{P}_{\leq 1}(V) \cup \{\{u\}, \{u, v\} \mid v \in V(G_t)\}$ with the following structure:

$$M(\{u\}; y) = \begin{pmatrix} A_{\mathbf{0}} & A_{\{u\}} \\ A_{\{u\}} & A_{\{u\}} \end{pmatrix}.$$

We have also seen that $M(\{u\};y)$ can be block diagonalized, i.e.

$$M(\{u\};y) \succeq 0 \Longleftrightarrow A_{\mathbf{0}} - A_{\{u\}} \succeq 0, A_{\{u\}} \succeq 0. \tag{5.34}$$

Let $M_2(u;y)$ be the principal submatrix of $M(\{u\};y)$ indexed by $\mathcal{A}_{\mathbf{0}} \cup \mathcal{A}_{\{u\}} = \mathcal{P}_{\leq 1}(V) \cup \{\{u,v\} \mid v \in V(G_t)\} = \{\mathbf{0}\} \cup \{\{v\} \mid v \in V(G_t)\} \cup \{\{u,v\} \mid v \in V(G_t)\}$. In other words, $M_2(u;y)$ is obtained from $M(\{u\};y)$ by deleting the second column indexed by $\{u\}$. We now show, using the invariance of $y$ under action of $\mathrm{Sym}(t)$, that the matrix $M_2(u;y)$ has a special block structure, whose symmetry can be used to block-diagonalize it. To begin with, with respect to the partition $\{\mathbf{0}\} \cup \{\{v\} \mid v \in V(G_t)\} \cup \{\{u,v\} \mid v \in V(G_t)\}$ of its index set, the matrix $M_2(u;y)$ has the block form

$$M_2(u;y) = \begin{pmatrix} y_{\mathbf{0}} & c^T & d^T \\ c & C & D \\ d & D & D \end{pmatrix}.$$

with $a, c, d, C, D$ being defined in terms of $y$. In view of (5.34), we have:

$$M_2(u;y) \succeq 0 \Longleftrightarrow \begin{pmatrix} y_{\mathbf{0}} - y_u & c^T - d^T \\ c - d & C - D \end{pmatrix} \succeq 0 \text{ and } D \succeq 0. \tag{5.35}$$

Next we observe that the invariance of $y$ under $\mathrm{Sym}(t)$ implies a special block structure for the matrices $C$ and $D$.

**Lemma 5.2.6.** *Consider the partition $V(G_t) = V_1 \cup \ldots \cup V_t$ of the nodeset of graph $G_t$, where $V_p := \{pi \mid i \in V\}$ for $p = 1, \ldots, t$. With respect to this partition, the matrices $C$ and $D$ have the block form:*

$$C = \begin{pmatrix} A^1 & A^2 & \cdots & A^2 \\ A^2 & A^1 & \cdots & A^2 \\ \vdots & \vdots & \ddots & \vdots \\ A^2 & \cdots & \cdots & A^1 \end{pmatrix}, \; D = \begin{pmatrix} B^1 & B^2 & B^2 & \cdots & B^2 \\ (B^2)^T & B^3 & B^4 & \cdots & B^4 \\ (B^2)^T & B^4 & B^3 & \cdots & B^4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (B^2)^T & B^4 & \cdots & \cdots & B^3 \end{pmatrix}, \tag{5.36}$$

*where[1] $A^1, \ldots, B^4 \in \mathbb{R}^{n \times n}$. Moreover, setting $a_1 := \mathrm{diag}(A^1)$, $b_1 := \mathrm{diag}(B^1)$, $b_3 := \mathrm{diag}(B^3)$, we have $c = [a_1^T \ldots a_1^T]^T$, and $d = [b_1^T\, b_3^T\, b_3^T \ldots b_3^T]^T$.*

*Proof.* Consider $i, j \in V$ and $p, q, p', q' \in \{1, \ldots, t\}$ with $p = q$ if and only if $p' = q'$. Then $C_{pi,qj} = y_{\{pi,qj\}} = y_{\{p'i,q'j\}} = C_{p'i,q'j}$; indeed, as there exists $\sigma \in \mathrm{Sym}(t)$ mapping $\{p, q\}$ to $\{p', q'\}$, the equality $y_{\{pi,qj\}} = y_{\{p'i,q'j\}}$ follows from the fact that $y$ is invariant under action of $\mathrm{Sym}(t)$. This shows that $C$ has the form indicated in (5.36); the argument is analogous for matrix $D$. ☐

To fix ideas, set $u = 1h \in V_1$ (where $h \in V$ is a given node of $G$). Then the entries of $A^1, \ldots, B^4$ are given by

$$\begin{gathered} A^1_{ij} = y_{\{1i,1j\}},\; A^2_{ij} = y_{\{1i,2j\}},\; B^1_{ij} = y_{\{1i,1h,1j\}}, \\ B^2_{ij} = y_{\{1i,1h,2j\}},\; B^3_{ij} = y_{\{2i,1h,2j\}},\; B^4_{ij} = y_{\{2i,1h,3j\}} \end{gathered} \tag{5.37}$$

[1] Here $A^i$ or $B^i$ should not be interpreted as powers of $A$ or $B$. Namely, $i$ is just an upper index.

for $i,j \in V$. (Recall that $y_{\{1i,1j\}} = y_{\{pi,pj\}}$, $y_{\{1i,2j\}} = y_{\{pi,qj\}}$, $y_{\{1i,2j,3h\}} = y_{\{pi,qj,rh\}}$ for any distinct $p,q,r \in \{1,\ldots,t\}$ since $y$ is invariant under action of $\mathrm{Sym}(t)$.) Moreover, the edge constraints $y_{uv} = 0$ (for $uv \in E(G_t)$) in (4.23) can be reformulated as

$$\begin{array}{l} A^1_{ij} = 0 \ \text{ if } \ ij \in E(G), \\ B^1_{ij} = 0 \ \text{ if } \ \{i,j,h\} \text{ contains an edge of } G, \\ B^2_{ij} = 0 \ \text{ if } \ hi \in E(G) \text{ or } j \in \{i,h\}, \\ B^3_{ij} = 0 \ \text{ if } \ ij \in E(G) \text{ or if } h \in \{i,j\}, \\ B^4_{ij} = 0 \ \text{ if } \ h \in \{i,j\}, \\ \mathrm{diag}(A^2) = \mathrm{diag}(B^2) = \mathrm{diag}(B^4) = 0, \end{array} \tag{5.38}$$

for distinct $i,j \in V$.

The next lemma indicates how one can further block-diagonalize the two matrices appearing at the right hand side of the equivalence in (5.35).

**Lemma 5.2.7.** *We have*

$$D \succeq 0 \Longleftrightarrow \begin{pmatrix} B^1 & (t-1)B^2 \\ (t-1)(B^2)^T & (t-1)B^3 + (t-1)(t-2)B^4 \end{pmatrix}, \ B^3 - B^4 \succeq 0.$$

*Moreover,*

$$\begin{pmatrix} y_{\mathbf{0}} - y_u & c^T - d^T \\ c - d & C - D \end{pmatrix} \succeq 0 \Longleftrightarrow \ A^1 - B^3 - A^2 + B^4 \succeq 0 \ \text{ and}$$

$$\begin{pmatrix} y_{\mathbf{0}} - y_u & a_1^T - b_1^T & (t-1)(a_1^T - b_3^T) \\ & A^1 - B^1 & (t-1)(A^2 - B^2) \\ & & (t-1)(A^1 - B^3) + (t-1)(t-2)(A^2 - B^4) \end{pmatrix} \succeq 0.$$

*(We wrote only the upper triangular part in the above (symmetric) matrix.)*

*Proof.* Consider the orthogonal matrices

$$M := \begin{pmatrix} \mathbf{I}_n & 0 \\ 0 & U_{t-1} \end{pmatrix}, \ N := \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix},$$

where $U_{t-1}$ is defined as in the proof of Lemma 2.11, i.e. $U_{t-1}$ is a $(t-1)\times(t-1)$ block-matrix where, for $p,q = 1,\ldots,t-1$, its $(p,q)$th block $U_{t-1}(p,q)$ is the $n\times n$ matrix defined as

$$U_{t-1}(p,q) := \begin{cases} \frac{1}{\sqrt{t-1}}\mathbf{I}_n & \text{if } p = 1 \text{ or } q = 1, \\ \left(\frac{1}{\sqrt{t-1}+t-1} - 1\right)\mathbf{I}_n & \text{if } p = q \geq 2, \\ \frac{1}{\sqrt{t-1}+t-1}\mathbf{I}_n & \text{otherwise.} \end{cases} \tag{5.39}$$

Recall that $U_{t-1}$ is symmetric and orthogonal. A simple calculation shows that

$$MDM = \begin{pmatrix} B^1 & \sqrt{t-1}B^2 & 0 & \ldots & 0 \\ \sqrt{t-1}(B_2)^T & B^3 + (t-2)B^4 & 0 & \ldots & 0 \\ 0 & 0 & B^3 - B^4 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & B^3 - B^4 \end{pmatrix}.$$

The first assertion of the lemma now follows after multiplying the second row/column block by $\sqrt{t-1}$. Next we have

$$N \begin{pmatrix} y_{\mathbf{0}} - y_u & c^T - d^T \\ c - d & C - D \end{pmatrix} N = \begin{pmatrix} y_{\mathbf{0}} - y_u & (c-d)^T M \\ M(c-d) & M(C-D)M \end{pmatrix}.$$

As the matrix $C - D$ has the same type of block shape as $D$, we deduce from the above that $M(C-D)M$ is block-diagonal. More precisely, the first diagonal block has the form

$$\begin{pmatrix} A^1 - B^1 & \sqrt{t-1}(A^2 - B^2) \\ \sqrt{t-1}(A^2 - B^2)^T & (A^1 - B^3) + (t-2)(A^2 - B^4) \end{pmatrix}$$

and the remaining $t-2$ diagonal blocks are all equal to $A^1 - B^3 - A^2 + B^4$. One can moreover verify that $(c-d)^T M = \big(a_1^T - b_1^T, \sqrt{t-1}(a_1^T - b_3^T), 0 \ldots 0\big)$. From this follows the second assertion of the lemma. □

Summarizing, we have obtained the following more compact SDP for the parameter $\ell^{(2)}(G_t)$

$$\begin{gathered} \ell^{(2)}(G_t) = \max\ te^T a_1 \text{ s.t. } A^1, A^2, B^1, B^2, B^3, B^4 \in \mathbb{R}^{n\times n} \text{ satisfy (5.38)}, \\ a_1 = \operatorname{diag}(A^1),\ b_1 = \operatorname{diag}(B^1),\ b_3 = \operatorname{diag}(B^3), \\ \begin{pmatrix} 1 - (a_1)_h & a_1^T - b_1^T & (t-1)(a_1^T - b_3^T) \\ & A^1 - B^1 & (t-1)(A^2 - B^2) \\ & & (t-1)(A^1 - B^3) + (t-1)(t-2)(A^2 - B^4) \end{pmatrix} \succeq 0, \\ \begin{pmatrix} B^1 & (t-1)B^2 \\ & (t-1)B^3 + (t-1)(t-2)B^4 \end{pmatrix} \succeq 0, \\ A^1 - A^2 - B^3 + B^4 \succeq 0, \\ B^3 - B^4 \succeq 0. \end{gathered} \tag{5.40}$$

This formulation applies when $G$ is vertex-transitive; here $h$ is any fixed node of $G$. Recall that $\Psi_{\ell^{(2)}}(G)$ can be obtained by computing $\ell^{(2)}(G_t)$ for $O(\log n)$ queries of the parameter $t$ and, for $G$ vertex-transitive, the computation of each $\ell^{(2)}(G_t)$ is via an SDP with four LMI's involving matrices of size $2n+1$, $2n$, $n$, $n$, respectively. The above reductions obviously apply to the bound $\Psi_{\ell_+^{(2)}}$ (obtained by adding nonnegativity). We use the formulation (5.40) in Sections 6.2 and 6.3 to compute bounds for Hamming and Kneser graphs.

### 5.2.4 Link with copositive programming based hierarchies

We have seen two possible constructions for hierarchies of bounds towards $\alpha(G)$ and $\chi^*(G)$, based on the method of Lasserre and its 'block diagonal' variation. As mentioned in Chapter 4 there are several other possible constructions for approximating the stable set problem. However, to the best of our knowledge, such constructions were much less investigated for the colouring problem. Recently Dukanovic and Rendl [24] investigated a hierarchy of lower bounds for

$\chi^*(G)$, which is closely related to the hierarchy of de Klerk and Pasechnik [50] for $\alpha(G)$ presented in Section 4.2. Both are based on copositive programming and some of its tractable relaxations in terms of sums of squares of polynomials, proposed by Parrilo [75]. Dukanovic and Rendl [24] propose an analogous hierarchy toward the fractional chromatic number. To start with they show the following copositive programming formulation for $\chi^*(G)$:

**Theorem 5.2.8.** *For any graph G*

$$\chi^*(G) = \min\ t \text{ s.t. } \begin{array}{l} X_{ii} = t\ (i \in V),\ X_{ij} = 0\ (ij \in E(G)) \\ X \in \mathcal{C}_n^*,\ X - \mathbf{J} \succeq 0. \end{array} \tag{5.41}$$

Using the idea of Parrilo [75] they replace the cone $\mathcal{C}_n$ by its subcone $K_n^{(r)}$ in (5.41), i.e. for $r \in \mathbb{Z}_+$ they replace $\mathcal{C}_n^*$ by $\mathcal{C}_n^{(r)}$ and define

$$\widehat{\psi}^{(r)}(G) := \min\ t \text{ s.t. } \begin{array}{l} X_{ii} = t\ (i \in V),\ X_{ij} = 0\ (ij \in E(G)) \\ X \in \mathcal{C}_n^{(r)},\ X - \mathbf{J} \succeq 0. \end{array} \tag{5.42}$$

Thus, $\widehat{\psi}^{(r)}(G) \leq \widehat{\psi}^{(r+1)}(G) \leq \chi^*(G)$. Moreover, it is proved in [24] that $\widehat{\psi}^{(0)}(G) = \overline{\vartheta^+}(G)$ and that the pair $(\widehat{\vartheta}^{(r)}, \widehat{\psi}^{(r)}(G))$ satisfies the reciprocity relation:

$$\widehat{\vartheta}^{(r)}(G)\widehat{\psi}^{(r)}(G) \geq |V(G)|,\ \text{with equality if } G \text{ is vertex-transitive,} \tag{5.43}$$

thus extending the result given in Theorem 3.3.1(c).

Now one may wonder what is the link between the two hierarchies $\text{las}^{(r)}$ and $\vartheta^{(r)}$ for $\alpha$, and between the two hierarchies $\psi^{(r)}$ and $\widehat{\psi}^{(r)}(G)$ for $\chi^*$. Recall definitions for $Q_r^+(G)$ and $L_r^+(G)$ from (4.43) and (4.48) respectively and define:

$$\psi_+^{(r)}(G) := \min t \ \text{ s.t. } x \in Q_r^+(G),\ x_{\mathbf{0}} = t,\ x_i = 1\ (i \in V), \tag{5.44}$$

and

$$\widetilde{\psi}_+^{(r)}(G) := \min t \ \text{ s.t. } x \in L_r^+(G),\ x_{\mathbf{0}} = t,\ x_i = 1\ (i \in V). \tag{5.45}$$

The analogue of Theorem 5.2.1 (e) holds for the pairs $\left(\text{las}_+^{(r)}, \psi_+^{(r)}\right)$ and $\left(\ell_+^{(r)}, \widetilde{\psi}_+^{(r)}\right)$ as well and we have

$$\text{las}_+^{(1)}(G) = \ell_+^{(1)}(G) = \vartheta'(G) = \widehat{\vartheta}^{(0)}(G),$$

$$\psi_+^{(1)}(G) = \widetilde{\psi}_+^{(1)}(G) = \overline{\vartheta^+}(G) = \widehat{\psi}^{(0)}(G).$$

The reciprocity relations, Corollary 4.2.22 and Proposition 4.2.23 imply:

**Proposition 5.2.9.** *For any vertex-transitive graph G*

$$\widehat{\psi}^{(1)}(G) \leq \widetilde{\psi}_+^{(2)}(G),\ \textit{and}$$

$$\widehat{\psi}^{(r-1)}(G) \leq \psi_+^{(r)}(G)\ \textit{when}\ r \geq 1.$$

*Proof.* Assume that $G$ is vertex transitive. Then for any $r \in \mathbb{N}\backslash\{0\}$ we have $\widehat{\psi}^{(r-1)}(G) = \frac{|V(G)|}{\widehat{\vartheta}^{(r-1)}(G)}$, $\widetilde{\psi}^{(r)}(G) = \frac{|V(G)|}{\ell^{(r)}(G)}$, and $\psi^{(r)}(G) = \frac{|V(G)|}{\mathrm{las}^{(r)}(G)}$. From Corollary 4.2.22 we have $\widehat{\psi}^{(r-1)}(G) = \frac{|V(G)|}{\widehat{\vartheta}^{(r-1)}(G)} \leq \frac{|V(G)|}{\mathrm{las}^{(r)}(G)} = \psi_+^{(r)}(G)$. Finally, Proposition 4.2.23 implies $\widehat{\psi}^{(1)}(G) = \frac{|V(G)|}{\widehat{\vartheta}^{(1)}(G)} \leq \frac{|V(G)|}{\ell^{(2)}(G)} = \widetilde{\psi}^{(2)}(G)$. □

It is an open question to determine whether the above inequalities remain valid when $G$ is not vertex-transitive. In [24] the bounds $\widehat{\psi}_+^{(1)}(G)$, $\widetilde{\psi}^{(2)}(G)$ and $\widetilde{\psi}_+^{(2)}(G)$ were compared for some instances of Hamming graphs (which are indeed vertex-transitive). In most of the instances presented there the first inequality in Proposition 5.2.9 is strict.

## 5.3 Conclusions

In this chapter we defined and studied hierarchies of lower bounds for the fractional chromatic number $\chi^*(G)$ and the chromatic number $\chi(G)$ of a graph $G$.

In Section 5.1 we used reduction (2.31) from the colouring problem to the stable set problem. This motivated the definition of the operator $\Psi$ which maps (polynomial time) upper bounds for $\alpha(G)$ to (polynomial time) lower bounds for $\chi(G)$. In particular, $\Psi_\alpha(G) = \chi(G)$, $\Psi_{\overline{\chi}}(G) = \Psi_{\overline{\chi^*}}(G) = \omega(G)$ and $\Psi_\vartheta(G) = \left\lceil \overline{\vartheta}(G) \right\rceil$. An interesting implication of these properties of $\Psi$, is that a polynomial time computable graph parameter nested between $\chi^*(G)$ and $\chi(G)$ cannot exist, unless P=NP. We also gave copositive and quadratic programming formulations for $\chi(G)$, which can be seen as analogs of Motzkin-Straus' result for the stability number.

In Section 5.2 we presented hierarchies of semidefinite lower bounds, corresponding to the 'stability number' hierarchies from Chapter 4, for $\chi^*(G)$ and $\chi(G)$. Two hierarchies $\psi^{(t)}(G)$ and $\widetilde{\psi}^{(t)}(G)$ ($t \in \mathbb{N}\backslash\{0\}$), based on the 'moment matrix' approach of Lasserre and its new block diagonal version, that converge to $\chi^*(G)$ in $\alpha(G)$ steps, and their counterparts $\Psi_{\mathrm{las}^{(t)}}(G)$ and $\Psi_{\ell^{(t)}}(G)$ that converge to $\chi(G)$ in $V(G)$ steps, were proposed and studied. We showed how to use symmetry to reduce the sizes of the semidefinite programs defining $\Psi_{\ell^{(2)}}(G)$. The hierarchy $\widehat{\psi}^{(t)}(G)$ ($t \in \mathbb{N}$), closely linked to the hierarchy of de Klerk and Pasechnik, based on the copositive formulation for $\chi(G)$ was also revisited.

We will show how to compute the second order bounds $\widetilde{\psi}^{(2)}(G)$ and $\Psi_{\ell^{(2)}}(G)$ when $G$ is a Hamming or Kneser graph in Chapter 6. In Section 6.27 we will propose another relaxation $\psi_K$, $K$ being a clique in $G$, of $\psi^{(2)}(G)$, which leads to strong bounds for $\chi(G)$ when $G$ is a nonsymmetric graph.

## Acknowledgements

We thank Alexander Schrijver for pointing out the reduction (5.1) to us, and for suggesting to apply the operator $\Psi$ to the interval (5.4), whereas we applied it to the subinterval $\left[\min\left(\alpha(\cdot), \frac{|V(\cdot)|}{\omega(\cdot)}\right), \overline{\chi}(\cdot)\right]$ in [37].

# Chapter 6

# Computational results

In this chapter we investigate how to compute the bounds presented in this thesis for certain graph classes. Our focus is the new block diagonal hierarchy, and in particular, the bounds $\ell^{(r)}(\cdot)$ and $\widetilde{\psi}^{(r)}(\cdot)$ where $r \leq 3$, and the bound $\Psi_{\ell^{(2)}}(\cdot)$. To compute these bounds we use symmetry reduction and block diagonalization techniques extensively.

In Section 6.1 we present results about Paley graphs. As well known instances of pseudo-random graphs, Paley graphs are very similar to typical graphs in $\mathcal{G}(q, 1/2)$, the class of random graphs on $q$ nodes with edge probability 1/2. Paley graphs are also used e.g. by Shearer [86] for bounding the Ramsey number. For a detailed study of Paley graphs see Bollobas [7, Chap. 13.2]. We show how to compute upper bounds for the stability numbers of Paley graphs using the compact SDP formulation for the bounds $\ell^{(r)}(\cdot)$ defined in Subsection 4.1.4. We reduce the number of blocks in SDPs considerably by exploring an edge transitivity property.

Sections 6.2 and 6.3 contain results on Hamming and Kneser graphs. colouring Hamming graphs is of interest e.g. to the Borsuk problem (see [96]), while the chromatic number of Kneser graphs was computed in the celebrated paper of Lovász [63] using topological methods (see e.g. [67] for a study of topological lower bounds for the chromatic number).

In Section 6.2 we compute the bounds $\widetilde{\psi}^{(2)}(\cdot)$ and $\Psi_{\ell^{(2)}}(\cdot)$ for Hamming graphs. As a crucial ingredient we use the block diagonalization of the Terwilliger algebra given by Schrijver [85]. In [85] the author proposes an upper bound for the stability number of Hamming graphs. Laurent [60] shows that this bound is just a relaxation of the bound $\ell^{(2)}(\cdot)$, and uses the same block diagonalization technique to compute $\ell^{(2)}(\cdot)$. The same recipe works for computing the parameter $\widetilde{\psi}^{(2)}(\cdot)$ of Hamming graphs, due to its reciprocity relation to $\ell^{(2)}(\cdot)$. For some instances, the parameter $\widetilde{\psi}^{(2)}(G)$ improves the theta number $\bar{\vartheta}(G)$ substantially. It can be further improved by adding nonnegativity. Still, $\Psi_{\ell^{(2)}}(G)$ hardly improves upon $\widetilde{\psi}^{(2)}(G)$ for Hamming graphs.

We use the technique from [85] extended to constant-weight codes to compute the bound $\Psi_{\ell^{(2)}}(\cdot)$ for Kneser graphs in Section 6.3. The fractional chromatic and the chromatic number of a Kneser graph are known (see [64] and [63]). Although the bound $\widetilde{\psi}^{(2)}(\cdot)$ coincides with the fractional chromatic number, $\Psi_\ell^{(2)}(\cdot)$ is at least as good and sometimes reaches $\chi(G)$.

We introduce a further variation $\psi_K(G)$, $K$ being a clique in $G$, of our bound $\psi^{(2)}(\cdot)$ in Section 6.4. It can be especially useful for graphs without apparent symmetries. Using a simple block diagonalization argument, $\psi_K(G)$ can be formulated via a semidefinite program involving $|K|+1$ matrices of size $|V(G)|+1$. We report experimental results on some DIMACS benchmark instances. The bound $\psi_K$ is quite strong for some dense random graphs although, as a variation of $\widetilde{\psi}^{(2)}(\cdot)$, it remains below the fractional chromatic number.

## 6.1 Experimental results for Paley graphs

Here we present some computational results for Paley graphs. We compute the bounds $\ell^{(t)}(G)$ $(t \leq 3)$ from (4.23) for the stability number of Paley graphs $G$ with at most 809 nodes.

Consider a finite field $\mathbb{F}_q$ where $q$ is a prime power satisfying $q = 1 \pmod 4$. The *Paley graph* $P_q$ is the graph whose vertices are the elements of $\mathbb{F}_q$, two elements $u \neq v \in \mathbb{F}_q$ being adjacent if $v-u$ is a square in $\mathbb{F}_q$. This defines an undirected graph since, as $q = 1 \pmod 4$, $-1$ is a square in $\mathbb{F}_q$ and thus $v-u$ is a square if and only if $u-v$ is a square in $\mathbb{F}_q$.

The Paley graph $P_q$ is isomorphic to its complementary graph, strongly regular, and vertex transitive. The automorphism group of $P_q$ consists of the affine mappings $\phi_{ab} : \mathbb{F}_q \to \mathbb{F}_q$, defined by $\phi_{ab}(u) = au + b$ $(u \in \mathbb{F}_q)$, where $a, b \in \mathbb{F}_q$, $a \neq 0$ square in $\mathbb{F}_q$. It is easy to prove that

$$E(P_q) = \{\{\sigma(h), \sigma(k)\} \mid \sigma \in \mathrm{Aut}(P_q)\} \text{ if } hk \in E(P_q); \text{ and} \tag{6.1}$$

$$E(\overline{P}_q) = \{\{\sigma(h), \sigma(k)\} \mid \sigma \in \mathrm{Aut}(P_q)\} \text{ if } hk \in E(\overline{P}_q). \tag{6.2}$$

Recall from Theorem 3.3.1 that $\vartheta(G)\vartheta(\overline{G}) = |V(G)|$ when $G$ is a vertex-transitive graph. Since the Paley graph $P_q$ is isomorphic to its complementary graph, we have $\vartheta(P_q) = \vartheta\left(\overline{P}_q\right) = \sqrt{q}$. This gives the well known analytical upper bound $\sqrt{q}$ for $\alpha(P_q)$ (cf. e.g. [7, Thm. 13.14]). When $q$ is a square, equality $\alpha(P_q) = \sqrt{q}$ (also equal to the chromatic number of $P_q$) is proved in [11]. A small improvement is proposed by Maistrelli and Penman [66] who show that $\alpha(P_q) \leq \sqrt{q-4}$ when $q$ is not a square and $q \neq 5$.

J.B. Shearer has computed $\alpha(P_q)$ for all prime $q \leq 7000$. His results can be found at http://www.research.ibm.com/people/s/shearer/indpal.html9. In order to illustrate the quality of the relaxations $L_t^+(P_q)$, we have computed the bounds $\ell_+^{(t)}(P_q)$ for $t = 2, 3$. For the first bound we have $\ell_+^{(1)}(P_q) = \vartheta'(P_q) = \vartheta(P_q) = \sqrt{q}$, which follows from (6.2) and Proposition 3.3.2.

We now give some details about computing $\ell_+^{(2)}(P_q)$ and $\ell_+^{(3)}(P_q)$. The basic tool we use is the invariance of the semidefinite program involved for optimizing $\sum_{u\in\mathbb{F}_q} x_u$ over $L_t^+(P_q)$.

Here we consider the Paley graph $G = P_q$ and the group $\mathcal{G} = \mathrm{Aut}(P_q)$. The semidefinite program $\max_{x\in L_t^+(P_q)} \sum_{u\in\mathbb{F}_q} x_u$ is invariant under the action of $\mathcal{G}$ and thus we can assume w.l.o.g. that the variable $y \in \mathbb{R}^{\mathcal{P}_{\leq t+1}(\mathbb{F}_q)}$ is invariant under the action of $\mathcal{G}$ (recall the definition of $L_t(G)$ from (4.22)). As $P_q$ is vertex transitive,

$$M(\{h\}; y) \succeq 0 \iff M(\{h'\}; y) \succeq 0 \text{ for any } h, h' \in \mathbb{F}_q.$$

Moreover, for two pairs $\{h, k\}$, $\{h', k'\}$ which are simultaneously edges or non-edges in $P_q$ the properties (6.1) and (6.2) yield

$$M(\{h, k\}; y) \succeq 0 \iff M(\{h', k'\}; y) \succeq 0.$$

Therefore, in the definition of $L_2(P_q)$, it suffices to require $M(\{h\}; y) \succeq 0$ for *one* choice of $h \in \mathbb{F}_q$ and, in the definition of $L_3(P_q)$, it suffices to require the conditions $M(\{h_1, k_1\}; y) \succeq 0$ and $M(\{h_2, k_2\}; y) \succeq 0$ for *one* choice of a non-edge $\{h_1, k_1\}$ and *one* choice of an edge $\{h_2, k_2\}$ of $P_q$.

| $q$ | $\ell_+^{(1)}(P_q) = \vartheta(P_q) = \sqrt{q}$ | $\ell_+^{(2)}(P_q)$ | $\ell_+^{(3)}(P_q)$ | $\alpha(P_q)$ |
|---|---|---|---|---|
| 61 | 7.810 | 5.465 | 5.035 | 5 |
| 73 | 8.544 | 5.973 | 5.132 | 5 |
| 89 | 9.434 | 6.304 | 5.391 | 5 |
| 97 | 9.849 | 7.398 | 6.596 | 6 |
| 101 | 10.050 | 6.611 | 5.496 | 5 |
| 109 | 10.440 | 7.366 | 6.578 | 6 |
| 113 | 10.630 | 7.599 | 7.009 | 7 |
| 137 | 11.705 | 8.200 | 7.047 | 7 |
| 149 | 12.207 | 8.231 | 7.136 | 7 |
| 157 | 12.530 | 8.707 | 7.485 | 7 |
| 173 | 13.153 | 9.426 | 8.050 | 8 |
| 181 | 13.454 | 9.112 | 7.606 | 7 |
| 193 | 13.892 | 9.210 | 7.651 | 7 |
| 197 | 14.036 | 9.226 | 8.063 | 8 |
| 229 | 15.133 | 10.290 | 9.076 | 9 |
| 233 | 15.264 | 10.182 | 8.245 | 7 |
| 241 | 15.524 | 9.891 | 8.272 | 7 |
| 257 | 16.031 | 10.247 | 8.131 | 7 |
| 269 | 16.401 | 10.624 | 8.778 | 8 |
| 277 | 16.643 | 10.340 | 8.670 | 8 |
| 281 | 16.763 | 10.605 | 8.397 | 7 |
| 293 | 17.117 | 10.937 | 9.183 | 8 |
| 313 | 17.692 | 11.551 | 9.451 | 8 |
| 317 | 17.804 | 12.337 | 10.363 | 9 |
| 337 | 18.358 | 11.658 | 9.464 | 9 |
| 401 | 20.025 | 12.753 | 10.023 | 9 |
| 509 | 22.561 | 14.307 | 11.185 | 9 |
| 601 | 24.515 | 16.077 | 12.478 | 11 |
| 701 | 26.476 | 16.857 | 12.824 | 10 |
| 809 | 28.443 | 17.371 | 13.494 | 11 |

Table 2: Optimizing over $L_t^+(P_q)$ for Paley graphs

We can simplify the programs even further by applying Lemma 2.4.5. Since

$$M(\{h\}; y) \succeq 0 \iff A_{\mathbf{0}}(y) - A_{\{h\}}(y) \succeq 0 \text{ and } A_{\{h\}}(y) \succeq 0,$$

optimization over $L_2(P_q)$ can be done via a semidefinite program involving two matrices, each of size $q + 1$. Each of the conditions $M(\{h_1, k_1\}; y) \succeq$

0, $M(\{h_2, k_2\}; y) \succeq 0$ can be reformulated as four conditions using (4.25). As $\{h_2, k_2\}$ is an edge in $P_q$, we deduce from Lemma 4.1.6 that the matrix $A_{\{h_2,k_2\}}(y)$ is identically zero. By invariance of the variable $y$,

$$A_{\{h_1\}}(y) \succeq 0 \Longrightarrow A_{\{h\}}(y) \succeq 0 \text{ for all } h \in \mathbb{F}_q, \text{ and}$$

$$A_{\{h_1\}}(y) - A_{\{h_1,k_1\}}(y) \succeq 0 \Longrightarrow A_{\{k_1\}}(y) - A_{\{h_1,k_1\}}(y) \succeq 0.$$

Summarizing, optimization over $L_3(P_q)$ can be done via a semidefinite program involving the conditions:

$$\begin{gathered} A_{\mathbf{0}}(y) - A_{\{h_1\}}(y) - A_{\{k_1\}}(y) + A_{\{h_1,k_1\}}(y) \succeq 0, \\ A_{\{h_1\}}(y) - A_{\{h_1,k_1\}}(y) \succeq 0,\ A_{\{h_1,k_1\}}(y) \succeq 0, \\ A_{\mathbf{0}}(y) - A_{\{h_2\}}(y) - A_{\{k_2\}}(y) \succeq 0, \end{gathered}$$

thus involving four matrices, each of size $q+1$. Moreover, for optimization over $L_2(P_q)$, we have variables corresponding to nodes, non-edges and stable sets of size 3, while for optimization over $L_3(P_q)$, we additionally have variables corresponding to all stable sets of size 4. For example, the program for optimizing over $L_3(P_{809})$ involves 34 (resp., 842) variables corresponding to the stable sets of size 3 (resp., of size 4).

Computational results for Paley graphs $P_q$ for various $q \leq 809$ are given in Table 2. We can observe that for all instances the bounds $\ell^{(2)}(P_q)$ and $\ell^{(3)}(P_q)$ are much smaller than $\vartheta(P_q) = \sqrt{q}$. Moreover, for all Paley graphs $P_q$ with $q < 230$, the bound $\lfloor \ell^{(3)}(P_q) \rfloor$ coincides with $\alpha(G)$.

## 6.2 Bounds for Hamming graphs

Given an integer $n \geq 1$ and $\mathcal{D} \subseteq N := \{1, \ldots, n\}$, the *Hamming graph* $H(n, \mathcal{D})$ is the graph $G$ with node set $V(G) := \mathcal{P}(N)$ and with an edge $(I, J)$ if $|I \triangle J| := |I \backslash J| + |J \backslash I| \in \mathcal{D}$ (for $I, J \in \mathcal{P}(N)$). Hence $|V(G)| = 2^n$.

We indicate here how to compute the parameters $\widetilde{\psi}^{(2)}(G)$ and $\Psi_{\ell^{(2)}}(G)$ when $G$ is a Hamming graph. As the programs defining these parameters for a Hamming graph $H(n, \mathcal{D})$ involve matrices of size $O(2^n)$, they cannot be solved directly for interesting values of $n$. However one can use the fact that $H(n, \mathcal{D})$ has a large automorphism group. Each permutation $\sigma \in \mathrm{Sym}(n)$ induces an automorphism of $G$, by letting $\sigma(I) := \{\sigma(i) \mid i \in I\}$ for $I \in \mathcal{P}(N)$ and, for any $K \in \mathcal{P}(N)$, the *switching mapping* $s_K$ defined by $s_K(I) := I \triangle K$ (for $I \in \mathcal{P}(N)$) is also an automorphism of $G$. Then $\mathrm{Aut}(G) = \{\sigma s_K \mid \sigma \in \mathrm{Sym}(n),\ K \in \mathcal{P}(N)\}$ and $|\mathrm{Aut}(G)| = n!2^n$. It turns out that the block matrices appearing in those programs belong to the Terwilliger algebra of the Hamming graph. Using the explicit block diagonalization of the Terwilliger algebra given in [85] we are able to block-diagonalize the matrices and thus compute $\widetilde{\psi}^{(2)}(G)$ and $\Psi_{\ell^{(2)}}(G)$ for $G = H(n, \mathcal{D})$ for $n$ up to 20.

### 6.2.1 Compact formulation for $\widetilde{\psi}^{(2)}(G)$ for Hamming graphs

As the graph $G = H(n, \mathcal{D})$ is vertex-transitive, we have $\widetilde{\psi}^{(2)}(G) = \frac{2^n}{\ell^{(2)}(G)}$ by Theorem 5.2.2 (e). It is shown in [60] how to compute the parameter $\ell^{(2)}(G)$

(when $\mathcal{D}$ is an interval $[1,d]$ but the reasoning is the same for any $\mathcal{D}$). The basic idea is that the matrix $M(\{h\};x)$, appearing in the program defining $\ell^{(2)}(G)$, is a block matrix whose blocks lie in the Terwilliger algebra and thus it can be block-diagonalized. We recall the details, directly for the parameter $\widetilde{\psi}^{(2)}(G)$ as they will be useful for our treatment of the parameter $\Psi_{\ell^{(2)}}(G)$ in the next section.

Recall first the definition from (5.27):

$$\begin{array}{llll} \widetilde{\psi}^{(2)}(G): & \min t & \text{s.t.} & x \in \mathbb{R}^{\mathcal{P}_{\leq 3}(V(G))},\ x_i = 1\ (i \in V(G)), \\ & & & x_{ij} = 0\ (ij \in E(G)), \\ & & & x_{\mathbf{0}} = t,\ M(\{u\};x) \succeq 0\ (u \in V(G)) \\ = & \min t & \text{s.t.} & x \in \mathbb{R}^{\mathcal{P}_{\leq 3}(V(G))},\ x_i = 1\ (i \in V(G)), \\ & & & x_{ij} = 0\ (ij \in E(G)), \\ & & & x_{\mathbf{0}} = t, M_2(u;x) \succeq 0\ (u \in V(G)). \end{array} \tag{6.3}$$

where $M_2(u;x)$ is the principal submatrix of $M(\{u\};x)$ indexed by $\mathcal{P}_{\leq 1}(V) \cup \{\{u,v\} \mid v \in V(G)\} = \{\mathbf{0}\} \cup \{\{v\} \mid v \in V(G)\} \cup \{\{u,v\} \mid v \in V(G)\}$. (Recall from Subsection 5.1.3 that $M_2(u;x)$ is simply obtained from $M(\{u\};x)$ by deleting the second column indexed by $\{u\}$.) By adding the nonnegativity condition $x \geq 0$ in (6.3) one gets the formulation for $\widetilde{\psi}^{(2)}_{+}(G)$ (compare with (5.45)).

Let $x$ be feasible for the program (6.3). As $G$ is vertex-transitive it suffices to require the condition $M_2(h;x) \succeq 0$ in (6.3) for *one* choice of $h \in V(G)$. Moreover, we may assume that the variable $x$ is invariant under action of the automorphism group of $G$. To fix ideas, let us choose the node $h := \emptyset$ of $G$ (the empty subset of $N$). The matrix $M_2(\emptyset;x)$ has the block form

$$M_2(\emptyset;x) = \begin{pmatrix} t & e^T & b^T \\ e & A & B \\ b & B & B \end{pmatrix} \tag{6.4}$$

where $A, B, e, b$ are indexed by $V(G) = \mathcal{P}(N)$ and $\text{diag}(A) = e$ and $\text{diag}(B) = b$. It can be block diagonalized and (5.35) reads:

$$M_2(\emptyset;x) \succeq 0 \Longleftrightarrow \begin{pmatrix} t-1 & e^T - b^T \\ e-b & A-B \end{pmatrix} \succeq 0 \ \text{ and } \ B \succeq 0. \tag{6.5}$$

As $x$ is invariant under action of $\text{Aut}(G)$, it follows that $A_{I,J} = x_{\{I,J\}} = x_{\{I',J'\}} = A_{I',J'}$ if $|I \triangle J| = |I' \triangle J'|$. In other words, the matrix $A$ lies in the Bose-Mesner algebra; say,

$$A = \sum_{k=0}^{n} x_k M_k^n = \sum_{i,j,p=0}^{n} x_{i+j-2p} M_{i,j}^{p,n} \tag{6.6}$$

for some reals $x_k$. Moreover, $B_{I,J} = x_{\{\emptyset,I,J\}} = x_{\{\emptyset,I',J'\}} = B_{I',J'}$ if $|I'| = |I|$, $|J'| = |J|$ and $|I' \cap J'| = |I \cap J|$. In other words, the matrix $B$ lies in the Terwilliger algebra; say,

$$B = \sum_{i,j,p=0}^{n} x_{i,j}^p M_{i,j}^{p,n} \tag{6.7}$$

for some reals $x_{i,j}^p$. The following relations link the parameters $x_i$, $x_{i,j}^p$.

**Lemma 6.2.1.** *For $i, j, p = 0, \dots, n$,*

$$\begin{aligned} x_i &= x^0_{0,i}, \\ x^p_{i,j} &= x^p_{j,i} = x^{j-p}_{i+j-2p,j} = x^{i-p}_{i+j-2p,i} \end{aligned} \tag{6.8}$$

*and the edge equations read*

$$x^p_{i,j} = 0 \quad \text{if} \quad \{i, j, i+j-2p\} \cap \mathcal{D} \neq \emptyset. \tag{6.9}$$

*Proof.* If $|I| = i$, then $x_i = A_{\emptyset,I} = x_{\{\emptyset,I\}} B_{\emptyset,I} = x^0_{0,i}$. Let $|I| = i$, $|J| = j$ and $|I \cap J| = p$. Then, $x^p_{i,j} = B_{I,J} = B_{J,I} = x^p_{j,i}$. Moreover, $x^p_{i,j} = B_{I,J} = x_{\{\emptyset,I,J\}} x_{\{I,\emptyset,I \triangle J\}} = B_{I,I\triangle J} x^{i-p}_{i+j-2p,i}$. This shows (6.8). The edge conditions read $B_{I,J} = x_{\{I,\emptyset,J\}} = 0$ if $\{|I|, |J|, |I \triangle J|\} \cap \mathcal{D} \neq \emptyset$, giving (6.9). $\square$

We can now use the results about Terwilliger algebra from Subsection 2.4.2 (Theorem 2.4.8 and Lemma 2.4.10) for block-diagonalizing the matrices occurring in (6.5). For $k = 0, \dots, \lfloor n/2 \rfloor$, define the matrices

$$A_k := \left( \sum_p \alpha^{p,n}_{i,j,k} x^0_{0,i+j-2p} \right)_{i,j=k}^{n-k}, \quad B_k := \left( \sum_p \alpha^{p,n}_{i,j,k} x^p_{i,j} \right)_{i,j=k}^{n-k} \tag{6.10}$$

corresponding respectively to the matrices $A$, $B$ in (6.6) and (6.7). Define the vector

$$\tilde{c} : \left( \sqrt{\binom{n}{i}} \left(1 - x^0_{0,i}\right) \right)_{i=0}^{n} \in \mathbb{R}^{n+1}. \tag{6.11}$$

Then the parameter $\widetilde{\psi}^{(2)}(H(n, \mathcal{D}))$ can be reformulated in the following way:

$$\begin{aligned} \widetilde{\psi}^{(2)}(H(n,\mathcal{D})) = \min\ t \ \text{ s.t.} \quad & x^0_{0,0} = 1,\ x^p_{i,j} \text{ satisfy (6.8), (6.9), and} \\ & A_k - B_k \succeq 0 \ (k = 1, \dots, \lfloor n/2 \rfloor), \\ & B_k \succeq 0 \ (k = 0, 1, \dots, \lfloor n/2 \rfloor), \\ & \begin{pmatrix} t-1 & \tilde{c}^T \\ \tilde{c} & A_0 - B_0 \end{pmatrix} \succeq 0, \end{aligned} \tag{6.12}$$

where $A_k, B_k, \tilde{c}$ are as in (6.10) and (6.11). To compute $\widetilde{\psi}^{(2)}_{+}(H(n, \mathcal{D}))$, simply add the nonnegativity condition $x^p_{i,j} \geq 0$ to (6.12).

### 6.2.2 Compact formulation for $\Psi_{\ell^{(2)}}(G)$ for Hamming graphs

We now give a more compact formulation for the parameter $\Psi_{\ell^{(2)}}(G)$ when $G = H(n, \mathcal{D})$. As explained in Subsection 5.1.3, one has to evaluate $\ell^{(2)}(G_t)$ for various choices of $t \in \mathbb{N}$. In Subsection 5.2.3 we explored symmetry in $G_t$. In particular, we formulated $\ell^{(2)}(G_t)$ via the block-diagonal SDP (5.40). As for the parameter $\widetilde{\psi}^{(2)}_{+}(H(n, \mathcal{D}))$, we now observe that $A^1, \dots, B^4$ and thus all blocks in the matrices in (5.40) lie in the Terwilliger algebra.

We fix $h := \emptyset$, the empty subset of $N$. The entries of $A^1, \ldots, B^4$ are given by (5.37):

$$\begin{array}{c} A^1_{IJ} = y_{\{1I,1J\}},\ A^2_{IJ} = y_{\{1I,2J\}},\ B^1_{IJ} = y_{\{1I,1\emptyset,1J\}}, \\ B^2_{IJ} = y_{\{1I,1\emptyset,2J\}},\ B^3_{IJ} = y_{\{2I,1\emptyset,2J\}},\ B^4_{IJ} = y_{\{2I,1\emptyset,3J\}} \end{array} \tag{6.13}$$

for $I, J \in \mathcal{P}(N)$. (Recall that $y_{\{1I,1J\}} = y_{\{pI,pJ\}}$, $y_{\{1I,2J\}} = y_{\{pI,qJ\}}$, $y_{\{1I,2J,3\emptyset\}} = y_{\{pI,qJ,r\emptyset\}}$ for any distinct $p, q, r \in \{1, \ldots, t\}$ since $y$ is invariant under action of $\mathrm{Sym}(t)$.) The edge constraints $y_{uv} = 0$ (for $uv \in E(G_t)$) are given by (5.38)

$$\begin{array}{l} A^1_{IJ} = 0 \ \text{ if } \ |I \triangle J| \in \mathcal{D}, \\ B^1_{IJ} = 0 \ \text{ if } \ \{I, J, \emptyset\} \text{ contains an edge of } H(n, \mathcal{D}), \\ B^2_{IJ} = 0 \ \text{ if } \ |I| \in \mathcal{D} \text{ or } J \in \{I, \emptyset\}, \\ B^3_{IJ} = 0 \ \text{ if } \ |I \triangle J| \in \mathcal{D} \text{ or if } \emptyset \in \{I, J\}, \\ B^4_{IJ} = 0 \ \text{ if } \ \emptyset \in \{I, J\}, \\ \mathrm{diag}(A^2) = \mathrm{diag}(B^2) = \mathrm{diag}(B^4) = 0, \end{array} \tag{6.14}$$

for distinct $I, J \in \mathcal{P}(N)$.

**Lemma 6.2.2.** *The matrices $A^s$ $(s = 1, 2)$ belong to the Bose-Mesner algebra $\mathcal{B}_n$ and the matrices $B^s$ $(s = 1, 2, 3, 4)$ belong to the Terwilliger algebra $\mathcal{A}_n$. Say, $A^s = \sum_{i=0}^n x(s)_i M_i^n$ $(s = 1, 2)$ and $B^s = \sum_{i,j,p=0}^n y(s)^p_{i,j} M^{p,n}_{i,j}$ $(s = 1, 2, 3, 4)$. Then,*

$$\begin{array}{c} x(s)_i = y(s)^0_{0,i} \ \text{ for } s = 1, 2,\ i = 1, \ldots, n, \\ y(s)^p_{i,j} = y(s)^p_{j,i} = y(s)^{j-p}_{i+j-2p,j} = y(s)^{i-p}_{i+j-2p,i} \ (\text{for } s = 1, 4), \\ y(2)^p_{i,j} = y(2)^{i-p}_{i,i+j-2p},\ y(3)^p_{i,j} = y(3)^p_{j,i}, \\ y(3)^p_{i,j} = y(2)^{i-p}_{i+j-2p,i} \ \text{ for } i, j, p = 0, \ldots, n. \end{array} \tag{6.15}$$

*Moreover, the edge conditions can be reformulated as*

$$\begin{array}{ll} y(1)^p_{i,j} = 0 & \text{if } \{i, j, i+j-2p\} \cap \mathcal{D} \neq \emptyset, \\ y(2)^i_{i,i} = y(4)^i_{i,i} = 0 & \text{for } i = 0, \ldots, n, \\ y(2)^p_{i,j} = 0 & \text{if } i \in \mathcal{D} \text{ or } j = 0, \\ y(3)^p_{i,j} = 0 & \text{if } i+j-2p \in \mathcal{D} \text{ or } i = 0 \text{ or } j = 0, \\ y(4)^p_{i,j} = 0 & \text{if } i = 0 \text{ or } j = 0, \end{array} \tag{6.16}$$

*for distinct $i, j \in \{0, 1, \ldots, n\}$.*

*Proof.* We use the fact that $A^1, \ldots, B^4$ satisfy (6.13) and (6.14) where the variable $y$ is assumed to be invariant under action of $\mathrm{Sym}(t) \times \mathrm{Aut}(G) \subseteq \mathrm{Aut}(G_t)$. We have $A^1, A^2 \in \mathcal{B}_n$, since the entries $A^1_{I,J} = y_{\{1I,1J\}}$ and $A^2_{I,J} = y_{\{1I,2J\}}$ depend only on $|I \triangle J|$. (Indeed, if $|I' \triangle J'| = |I \triangle J|$ then there exists $\sigma \in \mathrm{Aut}(G)$ mapping $\{I, J\}$ to $\{I', J'\}$ and thus, by the invariance of $y$ under action of $\sigma$, $y_{\{1I,1J\}} = y_{\{1I',1J'\}}$ and $y_{\{1I,2J\}} = y_{\{1I',2J'\}}$.) Similarly, for $s = 1, \ldots, 4$, $B^s \in \mathcal{A}_n$ since the entry $B^s_{I,J}$ depends only on $|I|, |J|, |I \cap J|$. The proof for the identities $x(s)_i = y(s)^0_{0,i}$ $(s = 1, 2)$ and $y(1)^p_{i,j} = \ldots = y(1)^{i-p}_{i+j-2p,i}$ is identical to the proof of (6.8). Let $I, J \in \mathcal{P}(N)$ with $|I| = i$, $|J| = j$,

$|I \cap J| = p$. Then, $y(4)^p_{i,j} = B^4_{I,J} = y_{\{1\emptyset,2I,3J\}} = y_{\{1\emptyset,3I,2J\}}$ (use the invariance of $y$ under the permutation $(2,3) \in \mathrm{Sym}(t)$), thus equal to $B^4_{J,I} = y(4)^p_{j,i}$. Moreover, $y(4)^p_{i,j} = y_{\{1\emptyset,2I,3J\}} = y_{\{1I,2\emptyset,3I\triangle J\}} = y_{\{2I,1\emptyset,3I\triangle J\}}$ (first apply the switching mapping by $I$ and then permute the indices $1,2$), thus equal to $B^4_{I,I\triangle J} = y(4)^{i-p}_{i,i+j-2p}$. Next we have: $y(2)^p_{i,j} = B^2_{I,J} = y_{\{1I,1\emptyset,2J\}} y_{\{1\emptyset,1I,2I\triangle J\}}$ (apply the switching mapping by $I$), thus equal to $B^2_{I,I\triangle J} = y(2)^{i-p}_{i,i+j-2p}$. Finally, $y(3)^p_{i,j} = B^3_{I,J} y_{\{2I,1\emptyset,2J\}} = B^3_{J,I} = y(3)^p_{j,i}$, and $y(3)^p_{i,j} = y_{\{2I,1\emptyset,2J\}} = y_{\{2\emptyset,1I,2I\triangle J\}} = y_{\{1\emptyset,2I,1I\triangle J\}}$ (first switch by $I$ and then permute $1,2$), thus equal to $B^2_{I\triangle J,I} = y(2)^{i-p}_{i+j-2p,i}$. The identities (6.16) follow directly from (5.38). □

As the blocks of the matrices in the program (5.40) lie in the Terwilliger algebra, the matrices in (5.40) can be block-diagonalized, as explained in Subsection 2.4.2. For this, define the matrices

$$A^s_k := \left(\sum_p \alpha^{p,n}_{i,j,k} y(s)^0_{i+j-2p,0}\right)_{i,j=k}^{n-k}, \quad B^s_k : \left(\sum_p \alpha^{p,n}_{i,j,k} y(s)^p_{i,j}\right)_{i,j=k}^{n-k} \tag{6.17}$$

corresponding, respectively, to the matrices $A^s$ ($s=1,2$) and $B^s$ ($s=1,2,3,4$) and define the vectors

$$\tilde{a} : \left(\sqrt{\binom{n}{i}}\left(y(1)^0_{0,0} - y(1)^i_{i,i}\right)\right)_{i=0}^{n}, \ \tilde{b} := \left(\sqrt{\binom{n}{i}}\left(y(1)^i_{i,i} - y(3)^i_{i,i}\right)\right)_{i=0}^{n} \in \mathbb{R}^{n+1}. \tag{6.18}$$

Using Lemmas 2.4.9 and 2.4.10, we obtain the following reformulation for the parameter $\ell^{(2)}(G_t)$ from (5.40)

$$\begin{gathered}
\ell^{(2)}(G_t) = \ \max\ 2^n t y(1)^0_{0,0} \ \text{ s.t. } y(s)^p_{i,j}\ (s=1,\dots,4) \text{ satisfy } (6.15),(6.16) \text{ and}\\
\begin{pmatrix} 1 - y(1)^0_{0,0} & \tilde{a}^T & (t-1)\tilde{b}^T \\ & A^1_0 - B^1_0 & (t-1)(A^2_0 - B^2_0) \\ & & (t-1)(A^1_0 - B^3_0) + (t-1)(t-2)(A^2_0 - B^4_0) \end{pmatrix} \succeq 0,\\
\begin{pmatrix} A^1_k - B^1_k & (t-1)(A^2_k - B^2_k) \\ & (t-1)(A^1_k - B^3_k) + (t-1)(t-2)(A^2_k - B^4_k) \end{pmatrix} \succeq 0 \ (k = 1, \dots, \lfloor n/2 \rfloor),\\
\begin{pmatrix} B^1_k & (t-1)B^2_k \\ & (t-1)B^3_k + (t-1)(t-2)B^4_k \end{pmatrix} \succeq 0 \ (k = 0, \dots, \lfloor n/2 \rfloor),\\
A^1_k - A^2_k - B^3_k + B^4_k \succeq 0 \ (k = 0, \dots, \lfloor n/2 \rfloor),\\
B^3_k - B^4_k \succeq 0 \ (k = 0, \dots, \lfloor n/2 \rfloor),
\end{gathered} \tag{6.19}$$

where $A^s_k, B^s_k, \tilde{a}, \tilde{b}$ are as in (6.17), (6.18). To compute $\ell^{(2)}_+(G_t)$ simply add the nonnegativity condition $y(s)^p_{i,j} \geq 0$ on all variables.

### 6.2.3 Numerical results for Hamming graphs

We have tested the various bounds on some instances of Hamming graphs. In what follows we use the following convention: For an integer $1 \leq d \leq n$, $H(n,d)$

(resp., $H^-(n,d)$, $H^+(n,d)$) denotes the graph $H(n,\mathcal{D})$ with $\mathcal{D}=\{d\}$ (resp., $\mathcal{D}=\{1,\dots,d\}$, $\{d,\dots,n\}$). The papers [26, 25, 24] give numerical results for the parameters $\overline{\vartheta}(G)$, $\overline{\vartheta^+}(G)$ for such instances.

In Table 3, the symbol '*' indicates the strict inequality $\Psi_{\ell^{(2)}}(G) > \lceil\widetilde{\psi}^{(2)}(G)\rceil$, which happens for $H(10,8)$ and $H^+(10,8)$, and we indicate in bold the values satisfying LB $=\chi(G)$ for the obtained lower bound LB. (Indeed in these instances, LB $=2^{n-1}$, while $\mathcal{P}(V)$ can be covered by the $2^{n-1}$ distinct pairs $\{I, V\setminus I\}$ ($I\subseteq V$) which are stable sets as $n\notin\mathcal{D}$.)

The results in Table 3 indicate that the parameters $\widetilde{\psi}^{(2)}(G)$ and $\widetilde{\psi}^{(2)}_+(G)$ give on some instances a major improvement on Szegedy's bound $\overline{\vartheta^+}(G)$. On the other hand, in most cases, the parameter $\Psi_{\ell^{(2)}}(G)$ gives no improvement since $\Psi_{\ell^{(2)}}(G)=\lceil\widetilde{\psi}^{(2)}(G)\rceil$. It could be that this feature is specific to Hamming graphs. As we will see in the next section, the bound $\Psi_{\ell^{(2)}}(G)$ does improve the bound $\lceil\widetilde{\psi}^{(2)}(G)\rceil$ for Kneser graphs.

| graph | $\overline{\vartheta}(G)$ | $\overline{\vartheta^+}(G)$ | $\widetilde{\psi}^{(2)}(G)$ | $\Psi_{\ell^{(2)}}(G)$ | $\widetilde{\psi}^{(2)}_+(G)$ | $\Psi_{\ell^{(2)}_+}(G)$ |
|---|---|---|---|---|---|---|
| $H^-(7,4)$ | 36 | 42.6667 | **64** | **64** | **64** | **64** |
| $H^-(8,5)$ | 72 | 85.3333 | **128** | **128** | **128** | **128** |
| $H(10,6)$ | 6 | 8.7273 | 10.4366 | 11 | 10.8936 | 11 |
| $H^-(10,6)$ | 207.36 | 320 | **512** | **512** | **512** | **512** |
| $H(10,8)$ | 2.6667 | 3.2 | 3.9232 | $5^*$ | 3.9232 | $5^*$ |
| $H^+(10,8)$ | 3.2 | 3.2 | 3.9232 | $5^*$ | 3.9232 | $5^*$ |
| $H(11,4)$ | 16 | 21.5652 | 25.7351 | 26 | 25.7351 | 26 |
| $H(11,6)$ | 12 | 12 | 12 | 12 | 15.2836 | 16 |
| $H^-(11,7)$ | 414.72 | 640 | **1024** | **1024** | **1024** | **1024** |
| $H^-(11,8)$ | 711.1111 | 819.2 | **1024** | **1024** | **1024** | **1024** |
| $H(11,8)$ | 3.2 | 4.9383 | 5.7805 | 6 | 5.7805 | 6 |
| $H(13,8)$ | 5.3333 | 9.4118 | 12.1429 | 13 | 13.6533 | 14 |
| $H(15,6)$ | 27.7647 | 30.7368 | 46.4371 | 47 | 50.3036 | 51 |
| $H(16,8)$ | 16 | 16 | 16 | 16 | 28.4444 | 29 |
| $H(17,6)$ | 35 | 48.2222 | 86.3086 | 87 | 88.3204 | 89 |
| $H(17,8)$ | 18 | 18 | 32 | 32 | 46.5122 | 47 |
| $H(17,10)$ | 6.6666 | 12.6315 | 15.8750 | 16 | 25.8405 | 26 |
| $H(18,10)$ | 10 | 16 | 18.3076 | 19 | 38.8844 | - |
| $H(20,6)$ | 59.3735 | 59.3735 | 140.9586 | 141 | 140.9586 | - |
| $H(20,8)$ | 41.7143 | 60.9524 | 107.1489 | - | 136.4115 | - |

Table 3: Bounds for the chromatic number of Hamming graphs

## 6.3 Bounds for Kneser graphs

We have seen that the parameter $\widetilde{\psi}^{(2)}(G)$ is bounded by $\chi^*(G)$ and that, for vertex-transitive graphs, it coincides with the bound $|V(G)|/\ell^{(2)}(G)$. On the other hand $\Psi_{\ell^{(2)}}(G)$ can sometimes be strictly greater than $\lceil\widetilde{\psi}^{(2)}(G)\rceil$, e.g., for the Hamming graph $H(10,8)$ (recall Table 3). We present here some numerical results showing that $\Psi_{\ell^{(2)}}(G)$ can in fact be strictly greater than $\lceil\chi^*(G)\rceil$ for Kneser graphs.

Given integers $n \geq 2r$, the Kneser graph $K(n,r)$ is the graph whose vertices are the subsets of size $r$ of a set $N$ with $|N| = n$, two vertices being adjacent if and only if they are disjoint. As shown in [64], $\alpha(K(n,r)) = \binom{n-1}{r-1}$, and thus $\chi^*(K(n,r)) = \frac{n}{r}$ in view of Theorem 3.3.1(a) as $K(n,r)$ is vertex-transitive. Lovász proved that $\chi(K(n,r)) = n - 2r + 2$ in his celebrated paper [63]. Thus the fractional chromatic number and the chromatic number of $K(n,r)$ can differ significantly, and the fractional chromatic number is close to the clique number $\omega(K(n,r)) = \lfloor \frac{n}{r} \rfloor$. Moreover, Lovász [64] proved that for $G = K(n,r)$, $\alpha(G) = \vartheta(G)$. Hence, $\ell^{(2)}(G) = \alpha(G)$, implying $\widetilde{\psi}^{(2)}(G) = \frac{|V(G)|}{\ell^{(2)}(G)} = \chi^*(G) = n/r$. Therefore, $\Psi_{\ell^{(2)}}(G) \geq \lceil n/r \rceil$. We show in this section how to compute $\Psi_{\ell^{(2)}}(G)$.

The Kneser graph $K(n,r)$ coincides with the subgraph of the Hamming graph $H(n, \{2r\})$ induced by the subset $\mathcal{P}_{=r}(N) := \{I \in \mathcal{P}(N) \mid |I| = r\}$. It will be convenient to view the Kneser graph also in the following alternative way. Fix a set $T \subseteq N$ with $|T| = r$ and define

$$\mathcal{P}(N,T) := \{(I', I'') \in \mathcal{P}(T) \times \mathcal{P}(N \setminus T) \mid |I'| = |I''|\}.$$

The mapping

$$\begin{array}{ccc} \mathcal{P}_{=r}(N) & \longrightarrow & \mathcal{P}(N,T) \\ I & \mapsto & (T \setminus I, I \setminus T) \end{array} \tag{6.20}$$

is a bijection and $|I \triangle J| = |(T \setminus I) \triangle (T \setminus J)| + |(I \setminus T) \triangle (J \setminus T)|$ holds for $I, J \in \mathcal{P}_{=r}(N)$. Hence $K(n,r)$ can also be viewed as the graph with nodeset $\mathcal{P}(N,T)$, with two nodes $(I', I''), (J', J'') \in \mathcal{P}(N,T)$ being adjacent if $|I' \triangle J'| + |I'' \triangle J''| = 2r$.

As we will see below, the matrices $A^1, A^2, B^1, \ldots, B^4$ involved in the program (5.40) for the computation of $\Psi_{\ell^{(2)}}(K(n,r))$ lie in $\mathcal{B}_{r,r'}$ ($r' = n - r$), a subalgebra of a tensor product of two Terwilliger algebras, which has also been studied and block-diagonalized by Schrijver [85] (in connection with constant weight codes). We follow the same steps as in Section 6.2 for the computation of $\ell^{(2)}(G_t)$ for Hamming graphs, which we now carry out for Kneser graphs.

### 6.3.1 The subalgebra $\mathcal{B}_{r,r'}$

As above, $|N| = n$ and we fix a subset $T \subseteq N$ with $|T| = r$ and set $r' := n - r$. For $i, j, p = 0, 1, ..., r$ (resp., $i', j', q = 0, 1, ..., r'$), let $M_{i,j}^{p,r}$ (resp., $M_{i',j'}^{q,r'}$) be the matrices indexed by $\mathcal{P}(T)$ (resp., $\mathcal{P}(N \setminus T)$) defining the Terwilliger algebra $\mathcal{A}_r$ (resp., $\mathcal{A}_{r'}$) as in Section 2.4.2. Let now $\mathcal{A}_{r,r'}$ be the algebra generated by the tensor products of matrices in $\mathcal{A}_r$ and $\mathcal{A}_{r'}$; that is,

$$\mathcal{A}_{r,r'} := \left\{ \sum_{i,j,p,i',j',q} x_{i,j,i',j'}^{p,q} M_{i,j}^{p,r} \otimes M_{i',j'}^{q,r'} \mid x_{i,j,i',j'}^{p,q} \in \mathbb{R} \right\}.$$

Matrices in $\mathcal{A}_{r,r'}$ are indexed by the set $\mathcal{P}(T) \times \mathcal{P}(N \setminus T)$. Consider the subalgebra

$$\mathcal{B}_{r,r'} := \left\{ \sum_{i,j,p,q} y_{i,j}^{p,q} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'} \mid y_{i,j}^{p,q} \in \mathbb{R} \right\}.$$

So $\mathcal{B}_{r,r'}$ consists of all matrices from $\mathcal{A}_{r,r'}$ satisfying $x_{i,j,i',j'}^{p,q} = 0$ if $i \neq i'$ or $j \neq j'$. Hence, for $M \in \mathcal{B}_{r,r'}$ and $(I, I'), (J, J') \in \mathcal{P}(T) \times \mathcal{P}(N \setminus T)$, $M_{(I,I'),(J,J')} = 0$

if $|I| \neq |I'|$ or if $|J| \neq |J'|$. Therefore any row/column of $M$ indexed by $(I, I') \notin \mathcal{P}(N,T)$ is identically zero and we may thus restrict matrices in $\mathcal{B}_{r,r'}$ to being indexed by the subset $\mathcal{P}(N,T)$ of $\mathcal{P}(T) \times \mathcal{P}(N \setminus T)$.

For $k \leq r$, let $M_k^{n,r}$ be the matrix indexed by $\mathcal{P}(N,T)$, whose $((I,I'),(J,J'))$th entry is equal to 1 if $|I \triangle J| + |I' \triangle J'| = 2k$, and to 0 otherwise. Thus $M_k^{n,r}$ corresponds to the principal submatrix of $M_{2k}^n$ (in the Bose-Mesner algebra $\mathcal{B}_n$) indexed by the subset $\mathcal{P}_{=r}(N)$ and $M_k^{n,r} \in \mathcal{B}_{r,r'}$ as

$$M_k^{n,r} = \sum_{\substack{i,j,p,q \\ i+j-p-q=k}} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'}.$$

Hence the set

$$\mathcal{B}_n^r := \left\{ \sum_{k=0}^{r} x_k M_k^{n,r} \mid x_k \in \mathbb{R} \right\}$$

is a subalgebra of $\mathcal{B}_{r,r'}$.

Schrijver [85] proved the following analogue of Theorem 2.4.8, giving the explicit block-diagonalization for matrices in $\mathcal{B}_{r,r'}$. For $k = 0, \ldots, \lfloor \frac{r}{2} \rfloor$, $l = 0, \ldots, \left\lfloor \frac{r'}{2} \right\rfloor$ set

$$W_{kl} := \{k, k+1, ..., r-k\} \cap \{l, l+1, ..., r'-l\}.$$

**Theorem 6.3.1.** *[85] For a matrix $M = \sum_{i,j,p,q} y_{i,j}^{p,q} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'}$ in $\mathcal{B}_{r,r'}$,*

$$M \succeq 0 \Longleftrightarrow M_{k,l} := \left( \textstyle\sum_{p,q} \alpha_{i,j,k}^{p,r} \alpha_{i,j,l}^{q,r'} y_{i,j}^{p,q} \right)_{i,j \in W_{kl}} \succeq 0 \quad \textit{for each}$$
$$k = 0, 1, \ldots, \lfloor \tfrac{r}{2} \rfloor \quad \textit{and } l = 0, 1, \ldots, \left\lfloor \tfrac{r'}{2} \right\rfloor. \tag{6.21}$$

We have the following analogues of Lemmas 2.4.9 and 2.4.10.

**Lemma 6.3.2.** *Let $A = \sum_{i,j,p,q} a_{i,j}^{p,q} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'}$, $B = \sum_{i,j,p,q} b_{i,j}^{p,q} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'}$, $C = \sum_{i,j,p,q} c_{i,j}^{p,q} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'}$ be matrices in $\mathcal{B}_{r,r'}$ and define accordingly*

$$A_{kl} = \left( \sum_{p,q} \alpha_{i,j,k}^{p,r} \alpha_{i,j,l}^{q,r'} a_{i,j}^{p,q} \right)_{i,j \in W_{kl}}, \quad B_{kl} = \left( \sum_{p,q} \alpha_{i,j,k}^{p,r} \alpha_{i,j,l}^{q,r'} b_{i,j}^{p,q} \right)_{i,j \in W_{kl}},$$

$$C_{kl} = \left( \sum_{p,q} \alpha_{i,j,k}^{p,r} \alpha_{i,j,l}^{q,r'} c_{i,j}^{p,q} \right)_{i,j \in W_{kl}}.$$

*Then,*

$$\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0 \Longleftrightarrow \begin{pmatrix} A_{kl} & B_{kl} \\ B_{kl}^T & C_{kl} \end{pmatrix} \succeq 0 \; \forall k = 0, 1, \ldots, \left\lfloor \frac{r}{2} \right\rfloor \; \textit{and } l = 0, 1, \ldots, \left\lfloor \frac{r'}{2} \right\rfloor.$$

**Lemma 6.3.3.** *Let $M = \sum_{i,j,p,q=0}^{n} x_{i,j}^{p,q} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'} \in \mathcal{B}_{r,r'}$, $c = \sum_{i=0}^{n} c_i \chi^i$, where $\chi^i \in \{0,1\}^{\mathcal{P}(N,T)}$ with $\chi^i_{(I,I')} = 1$ if $|I| = i$, (for $(I,I') \in \mathcal{P}(N,T)$), and $d \in \mathbb{R}$. Then,*

$$\begin{pmatrix} d & c^T \\ c & M \end{pmatrix} \succeq 0 \Longleftrightarrow \begin{cases} M_{kl} \succeq 0 \quad \textit{for } k = 0, \ldots, \lfloor \frac{r}{2} \rfloor, l = 0, \ \ldots, \left\lfloor \frac{r'}{2} \right\rfloor, \ k+l > 0; \\ \tilde{M}_{00} := \begin{pmatrix} d & \tilde{c}^T \\ \tilde{c} & M_{00} \end{pmatrix} \succeq 0 \end{cases}$$

*after setting* $\tilde{c}^T := \left(c_i\sqrt{\binom{r}{i}\binom{r'}{i}}\right)_{i=0}^{r}$.

### 6.3.2 Compact formulation for $\Psi_{\ell^{(2)}}(G)$ for Kneser graphs

In order to compute $\Psi_{\ell^{(2)}}(G)$ for the Kneser graph $G = K(n,r)$, one has to evaluate $\ell^{(2)}(G_t)$ for various choices of $t$. As $G$ is vertex-transitive, $\ell^{(2)}(G_t)$ can be computed using the program (5.40). We now fix $h := T \in \mathcal{P}_{=r}(N)$ corresponding to $(\emptyset,\emptyset) \in \mathcal{P}(N,T)$ as chosen node of $G$. We now show that the matrices $A^1,\dots,B^4$ appearing in program (5.40) lie in the algebra $\mathcal{B}_{r,r'}$ and thus they can be block-diagonalized using Theorem 6.3.1. The following lemma is the analogue of Lemma 6.2.2.

**Lemma 6.3.4.** *The matrices* $A^s$ $(s = 1,2)$ *belong to* $\mathcal{B}_n^r$ *and the matrices* $B^s$ $(s = 1,2,3,4)$ *belong to* $\mathcal{B}_{r,r'}$. *Say,* $A^s = \sum_{i=0}^r x(s)_i M_i^{n,r}$ $(s = 1,2)$ *and* $B^s = \sum_{i,j,p,q=0}^r y(s)_{i,j}^{p,q} M_{i,j}^{t,r} \otimes M_{i,j}^{q,r'}$ $(s = 1,2,3,4)$*. We have*

$$\begin{gathered}
x(s)_i = y(s)_{0,i}^{0,0} \quad \text{for } s = 1,2,\ i = 1,\dots,r,\\
y(s)_{i,j}^{p,q} = y(s)_{j,i}^{p,q} = y(s)_{i,i+j-p-q}^{i-q,i-p} y_{j,i+j-p-q}^{j-q,j-p} \quad \text{for } s = 1,4,\\
y(2)_{i,j}^{p,q} = y(2)_{i,i+j-p-q}^{i-q,i-p},\ y(3)_{i,j}^{p,q} = y(3)_{j,i}^{p,q},\\
y(3)_{i,j}^{p,q} = y(2)_{i+j-p-q,i}^{i-q,i-p} \quad \text{for } i,j,p,q = 0,\dots,r.
\end{gathered} \tag{6.22}$$

*Moreover, the edge conditions can be reformulated as*

$$\begin{array}{ll}
y(1)_{i,j}^{p,q} = 0 & \text{if } i = r \text{ or } j = r \text{ or } i+j-p-q = r,\\
y(2)_{i,j}^{p,q} = 0 & \text{if } i = r \text{ or } j = 0 \text{ or } i+j-p-q = 0,\\
y(3)_{i,j}^{p,q} = 0 & \text{if } i = 0 \text{ or } j = 0 \text{ or } i+j-p-q = r,\\
y(4)_{i,j}^{p,q} = 0 & \text{if } i = 0 \text{ or } j = 0 \text{ or } i+j-p-q = 0.
\end{array} \tag{6.23}$$

*Proof.* As in the proof of Lemma 6.2.2, the matrices $A^1,\dots,B^4$ satisfy (5.37) and (5.38), where the variable $y$ is invariant under action of $\mathrm{Sym}(t) \times \mathrm{Aut}(G)$. A main difference with the case of the Hamming graph is that, for the Kneser graph $G = K(n,r)$, $\mathrm{Aut}(G) \sim \mathrm{Sym}(n)$, i.e., the only automorphisms of $G$ arise from the permutations of $N$. Recall that $\sigma \in \mathrm{Sym}(n)$ acts on $\mathcal{P}_{=r}(N)$ in the obvious way; namely, $\sigma(I) = \{\sigma(i) \mid i \in I\}$ for $I \in \mathcal{P}_{=r}(N)$.

Let us first show that $A^1 \in \mathcal{B}_n^r$; that is, $A^1_{I,J}$ depends only on $|I \bigtriangleup J|$ (for $I,J \in \mathcal{P}_{=r}(N)$). For this, let $I,J,I',J' \in \mathcal{P}_{=r}(N)$ with $|I \bigtriangleup J| = |I' \bigtriangleup J'|$. Then, $|I \cap J| = |I' \cap J'|$ and thus there exists $\sigma \in \mathrm{Sym}(n)$ such that $\sigma(I) = I'$ and $\sigma(J) = J'$. Hence, $A^1_{I,J} = y_{\{1I,1J\}} = y_{\{1\sigma(I),1\sigma(J)\}} = A^1_{I',J'}$ since $y$ is invariant under action of $\sigma$. The proof for $A^2 \in \mathcal{B}_n^r$, $B^s \in \mathcal{B}_{r,r'}$ is along the same lines.

Let us now prove the identity $y(1)_{i,j}^{p,q} = y(1)_{i,i+j-p-q}^{i-q,i-p}$; the proofs for the remaining identities are along the same lines and thus omitted. Say, $y(1)_{i,j}^{p,q} = B^1_{I,J}$, where $I,J \in \mathcal{P}_{=r}(N)$ with $|T \setminus I| = i$, $|T \setminus J| = j$, $|(T \setminus I) \cap (T \setminus J)| = p$ and $|(I \setminus T) \cap (J \setminus T)| = q$. See Figure 6.1 for the Venn diagram of the sets $I,J,T$. Consider sets $I',J' \in \mathcal{P}_{=r}(N)$ which together with the set $T$ have the Venn diagram shown in Figure 6.1. Then, $B^1_{I',J'} = y(1)_{i,i+j-p-q}^{i-q,i-p}$ and there exists $\sigma \in \mathrm{Sym}(n)$ such that $\sigma(T) = I'$, $\sigma(I) = T$, $\sigma(J) = J'$. Therefore, $y(1)_{i,j}^{p,q} = B^1_{I,J} = y_{\{1I,1J,1T\}} y_{\{1\sigma(I),1\sigma(J),1\sigma(T)\}} y_{\{1T,1J',1I'\}} = B^1_{I',J'} = y(1)_{i,i+j-p-q}^{i-q,i-p}$. $\square$

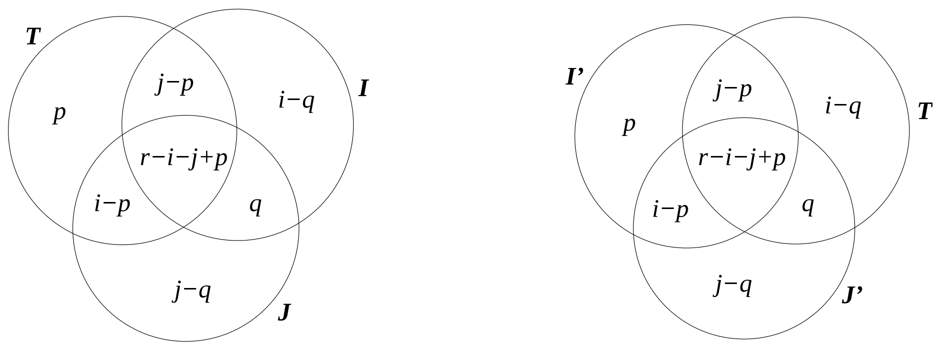


Figure 6.1: Venn diagrams

For $k = 0, ..., \lfloor r/2 \rfloor$, $l = 0, ..., \lfloor r'/2 \rfloor$, define the matrices

$$A^s_{kl} = \left( \sum_{p,q} \alpha^{p,r}_{i,j,k} \alpha^{q,r'}_{i,j,l} y(s)^{0,0}_{0,i+j-p-q} \right)_{i,j \in W_{kl}}, B^s_{kl} = \left( \sum_{p,q} \alpha^{p,r}_{i,j,k} \alpha^{q,r'}_{i,j,l} y(s)^{p,q}_{i,j} \right)_{i,j \in W_{kl}} \tag{6.24}$$

corresponding, respectively, to the matrices $A^s$ $(s = 1, 2)$ and $B^s$ $(s = 1, 2, 3, 4)$ and define the vectors

$$\tilde{a} : \left( \sqrt{\binom{r}{i}\binom{r'}{i}} \left( y(1)^{0,0}_{0,0} - y(1)^{i,i}_{i,i} \right) \right)^r_{i=0}, \tilde{b} := \left( \sqrt{\binom{r}{i}\binom{r'}{i}} \left( y(1)^{i,i}_{i,i} - y(3)^{i,i}_{i,i} \right) \right)^r_{i=0}. \tag{6.25}$$

Using Lemmas 6.3.2 and 6.3.3, we obtain the following reformulation for the parameter $\ell^{(2)}(G_t)$ from (5.40)

$$\ell^{(2)}(G_t) = \max \binom{n}{r} t y(1)^{0,0}_{0,0} \text{ s.t. } y(s)^{p,q}_{i,j},\ s = 1, \dots, 4 \text{ satisfy } (6.22), (6.23) \text{ and}$$

$$\begin{pmatrix} 1 - y(1)^{0,0}_{0,0} & \tilde{a}^T & (t-1)\tilde{b}^T \\ & A^1_{00} - B^1_{00} & (t-1)(A^2_{00} - B^2_{00}) \\ & & (t-1)(A^1_{00} - B^3_{00}) + (t-1)(t-2)(A^2_{00} - B^4_{00}) \end{pmatrix} \succeq 0;$$

$$\begin{pmatrix} A^1_{kl} - B^1_{kl} & (t-1)(A^2_{kl} - B^2_{kl}) \\ & (t-1)(A^1_{kl} - B^3_{kl}) + (t-1)(t-2)(A^2_{kl} - B^4_{kl}) \end{pmatrix} \succeq 0$$

$$\text{for } k = 0, ..., \lfloor r/2 \rfloor, l = 0, ..., \lfloor r'/2 \rfloor,\ k + l > 0;$$

$$\begin{pmatrix} B^1_{kl} & (t-1)B^2_{kl} \\ & (t-1)B^3_{kl} + (t-1)(t-2)B^4_{kl} \end{pmatrix} \succeq 0 \text{ for } k = 0, ..., \lfloor r/2 \rfloor, l = 0, ..., \lfloor r'/2 \rfloor;$$

$$A^1_{kl} - A^2_{kl} - B^3_{kl} + B^4_{kl} \succeq 0 \text{ for } k = 0, ..., \lfloor r/2 \rfloor, l = 0, ..., \lfloor r'/2 \rfloor;$$

$$B^3_{kl} - B^4_{kl} \succeq 0 \text{ for } k = 0, ..., \lfloor r/2 \rfloor, l = 0, ..., \lfloor r'/2 \rfloor, \tag{6.26}$$

where $A^s_{kl}, B^s_{kl}, \tilde{a}, \tilde{b}$ are as in (6.24), (6.25). To compute $\ell^{(2)}_+(G_t)$ simply add the nonnegativity condition $y(s)^{p,q}_{i,j} \geq 0$ on all variables.

### 6.3.3 Numerical results for Kneser graphs

We show in Table 4 below our numerical results for the bounds $\Psi_{\ell^{(2)}}(G)$ and $\Psi_{\ell^{(2)}_+}(G)$ for several instances of Kneser graphs. We indicate in bold the values achieving the chromatic number.

| Graph | $\lceil\chi^*(G)\rceil = \lceil n/r\rceil$ | $\Psi_{\ell^{(2)}}(G)$ | $\Psi_{\ell^{(2)}_+}(G)$ | $\chi(G) = n - 2r + 2$ |
|---|---|---|---|---|
| $K(6,2)$ | 3 | **4** | **4** | **4** |
| $K(7,2)$ | 4 | 4 | **5** | **5** |
| $K(8,3)$ | 3 | **4** | **4** | **4** |
| $K(9,3)$ | 3 | 4 | 4 | 5 |
| $K(10,4)$ | 3 | 3 | **4** | **4** |
| $K(11,3)$ | 4 | 5 | 5 | 7 |
| $K(11,4)$ | 3 | 4 | 4 | 5 |
| $K(12,3)$ | 4 | 5 | 6 | 8 |
| $K(12,4)$ | 3 | 4 | 4 | 6 |
| $K(12,5)$ | 3 | 3 | **4** | **4** |
| $K(13,5)$ | 3 | 4 | 4 | 5 |
| $K(14,5)$ | 3 | 4 | 4 | 6 |
| $K(15,3)$ | 5 | 6 | 6 | 11 |
| $K(16,4)$ | 4 | 5 | 6 | 10 |
| $K(24,6)$ | 4 | 4 | 6 | 14 |
| $K(25,5)$ | 5 | 6 | 7 | 17 |
| $K(34,7)$ | 5 | 6 | 7 | 22 |
| $K(36,6)$ | 6 | 7 | 9 | 26 |

Table 4: Bounds for the chromatic number of Kneser graphs

## 6.4 Computing the new bound $\psi_K$ for DIMACS benchmark graphs

So far we presented some ideas for approximating the chromatic numbers of vertex-transitive graphs. In particular, we considered the bounds $\widetilde{\psi}^{(2)}(\cdot)$ and $\Psi_{\ell^{(2)}}(\cdot)$ and showed how the problem size can be reduced by exploiting symmetry. For the formulation of $\widetilde{\psi}^{(2)}(G)$, it was observed before that, when $G$ is vertex-transitive, it suffices to require in (5.27) positive semidefiniteness of $M(\{h\}, x)$ for *only one* $h \in V(G)$ instead of *for all* $h \in V(G)$. In case of a non-symmetric graph $G$ one would need to require $M(\{h\}, x) \succeq 0$ for *all* $h \in V(G)$; therefore, with $n := |V(G)|$, in order to compute $\widetilde{\psi}^{(2)}(G)$, (resp., $\ell^{(2)}(G_t)$, and thus $\Psi_{\ell^{(2)}}(G)$), one would have to solve a semidefinite program with $2n$ (resp., $4n$) matrices of order $\leq n+1$ (resp., $\leq 2n+1$). For graphs that are of interest, e.g. with $n \geq 100$, this cannot be done with the currently available software for semidefinite programming.

For a non-symmetric graph $G = (V, E)$ we propose another variant of the bound $\widetilde{\psi}^{(2)}(G)$. Namely, given a clique $K$ in $G$ and $x \in \mathbb{R}^{\mathcal{P}_{\leq 4}(V)}$, let $M_2(K; x)$

denote the principal submatrix of $M_2(x)$ indexed by the multiset $\mathcal{P}_{\leq 1}(V) \cup (\cup_{h\in K}\{\{h\},\{h,i\} \mid i \in V\})$. Now define the parameter

$$\psi_K(G) := \min\ t \ \text{ s.t. } \quad \begin{array}{l} x_{\mathbf{0}} = t,\ x_i = 1\ (i \in V),\ M_2(K;x) \succeq 0, \\ x_I = 0 \text{ for all } I \text{ containing an edge.} \end{array} \tag{6.27}$$

Then $\overline{\vartheta}(G) \leq \psi_K(G) \leq \chi^*(G)$. (The inequalities follow from $\overline{\vartheta}(G) = \psi^{(1)}(G) \leq \psi_K(G) \leq \psi^{(2)}(G) \leq \chi^*(G)$ using definition (5.27) and Theorem 5.2.2(b)(d).) Set $k := |K|$ and assume w.l.o.g. that $K = \{1,2,...,k\}$. With respect to the partition of its index set as $\{\mathbf{0},\{i\} \mid i \in V\} \cup \cup_{h=1}^k \{\{h\},\{h,i\} \mid i \in V\}$, the matrix $M_2(K;x)$ has the block form

$$M_2(K;x) = \begin{pmatrix} A_0 & A_1 & A_2 & \dots & A_k \\ A_1 & A_1 & 0 & \dots & 0 \\ A_2 & 0 & A_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ A_k & 0 & \dots & 0 & A_k \end{pmatrix}$$

where $A_0, \dots, A_k$ are indexed by $\mathcal{P}_{\leq 1}(V)$, $(A_0)_{IJ} = x_{I\cup J}$, $(A_h)_{IJ} = x_{\{h\}\cup I\cup J}$ for $h \in K$, $I, J \in V$. The matrix in this form can be block-diagonalized. Define the $(k+1)\times(k+1)$ block matrix $N$ whose blocks $N(i,j)$ $(i,j = 0,\dots,k)$ are $(|V|+1)\times(|V|+1)$ matrices defined by

$$N(i,j) := \begin{cases} \mathbf{I} & \text{if } i = j; \\ -\mathbf{I} & \text{if } j = 0,\ i > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Here $\mathbf{I}$ stands for the identity matrix of order $|V|+1$. Then

$$N^T M_2(K;x) N = \begin{pmatrix} A_0 - \sum_{h=1}^k A_h & 0 & 0 & \dots & 0 \\ 0 & A_1 & 0 & \dots & 0 \\ 0 & 0 & A_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & A_k \end{pmatrix},$$

which yields

$$M_2(K;x) \succeq 0 \iff A_0 - \sum_{h=1}^k A_h \succeq 0,\ A_1, \dots, A_k \succeq 0.$$

Hence $\psi_K(G)$ can be computed via a semidefinite program involving $k+1$ matrices of size $n+1$. (In fact, in the condition $A_h \succeq 0$ $(h = 1,\dots,k)$ we can delete one column indexed by $\{h\}$, and thus get $k$ matrices of size $n$.)

We have conducted experiments for some DIMACS benchmark graphs (studied e.g. in [17, 18, 22, 26, 29, 69]). In Table 5 we present our lower bounds for the chromatic number of the graphs DSJC$a.b$. Recall that DSJC$a.b$ are random graphs with $a$ vertices, two vertices being adjacent with probability $10^{-1}b$. The graph DSJR500.1 is a geometric graph with 500 nodes randomly distributed in the unit square, with an edge between two nodes if their distance is less than 0.1. The graph DSJR500.1c is the complement of DSJR500.1. The graphs can be downloaded from http://mat.gsia.cmu.edu/COLOR03/

In Table 5, the column 'LB' contains the previously best known lower bounds taken from [22, 69], and the values into parentheses come from [16]; the bound 82 for DSJR500.1c is the size of a clique obtained using the heuristic of [13]. The column 'UB' contains the best known upper bounds taken from [17, 29, 28], i.e. the number of colours in the best colourings found so far. The column '$K$' contains the size of the clique used for computing the parameter $\psi_K(G)$ (the clique is found using the heuristic from [13]). We also indicate the value of the theta number $\bar{\vartheta}(G)$ (also computed in [26, 25] for some instances), which already improves the best lower bound in several instances. We indicate in bold best new lower bounds for the chromatic number. On several instances they give a significant improvement on the best known lower bound. Moreover, in two instances, we are able to close the gap as our lower bound matches the upper bound. Namely we find the exact value of the chromatic number for the graphs DSJC125.9 ($\chi(G) = 43$) and DSJR500.1c ($\chi(G) = 85$), which were not known before to the best of our knowledge. These results demonstrate that the bound $\psi_K(G)$ is quite strong.

We should also point out that the semidefinite program for the parameter $\psi_K$, for instance, for the graph DSJR500.1$c$, contains one $501 \times 501$ block and 77 blocks of size at most $500 \times 500$. One cannot hope to solve such a big problem using solvers based on interior point methods. The values in columns ' $\overline{\vartheta}(G)$ ' and ' $\psi_K(G)$ ' were computed using the boundary point method of Povh, Rendl and Wiegele [80]. This method allows to compute the Lovász theta number and its variations (e.g. $\psi_K(G)$) to reasonably high accuracy even for graphs with several hundred nodes.

One may wonder why we did not add nonnegativity constraints in the formulation for $\psi_K$. The reason is that for random graphs adding nonnegativity constraints gives only a negligible improvement. This fact was already observed for the Lovász theta number in [26].

| Graph | LB | $\overline{\vartheta}(G)$ | $\lceil\overline{\vartheta}(G)\rceil$ | $K$ | $\psi_K(G)$ | $\lceil\psi_K(G)\rceil$ | UB |
|---|---|---|---|---|---|---|---|
| DSJC125.1 | 5 | 4.1062 | 5 | 4 | 4.337 | **5** | 5 |
| DSJC125.5 | 14 (17) | 11.7844 | 12 | 10 | 13.942 | **14** | 17 |
| DSJC125.9 | 42 | 37.768 | 38 | 34 | 42.53 | **43** | 43 |
| DSJC250.1 | 6 (8) | 4.906 | 5 | 4 | 5.208 | **6** | 8 |
| DSJC250.5 | 14 | 16.234 | 17 | 12 | 19.208 | **20** | 28 |
| DSJC250.9 | 48 | 55.152 | 56 | 43 | 66.15 | **67** | 72 |
| DSJC500.1 | 6 | 6.217 | **7** | 5 | 6.542 | **7** | 12 |
| DSJC500.5 | 13 (16) | 20.542 | 21 | 13 | 27.791 | **28** | 48 |
| DSJC500.9 | 59 | 84.04 | 85 | 56 | 100.43 | **101** | 126 |
| DSJC1000.1 | 6 | 8.307 | **9** | 5 | - | - | 20 |
| DSJC1000.5 | 15 (17) | 31.89 | **32** | 14 | - | - | 83 |
| DSJC1000.9 | 66 | 122.67 | **123** | 65 | - | - | 224 |
| DSJR500.1c | 82 (83) | 83.74 | 84 | 77 | 84.12 | **85** | 85 |

Table 5: Bounds for the chromatic number of DIMACS instances

## 6.5 Conclusions and Remarks

The goal of this chapter was to show that the semidefinite bounds studied in this thesis can be stronger bounds for the stability number and the chromatic number than the Lovász theta number. We confirmed it experimentally, i.e. by the results presented in Tables 2-5.

We mention now several details about our computations.

Experiments were conducted on a single machine with an AMD Athlon 64 3500 processor and 1024 MB of RAM memory.

The computational results reported in Tables 2-4 were carried out using the open source codes for semidefinite programming CSDP 5.0 and DSDP 5.8 available, respectively, at
http://infohost.nmt.edu/ borchers/csdp.html and
http://www-unix.mcs.anl.gov/ benson/dsdp/.

For finding large cliques in the instances in Table 5 we used the heuristic Max-AO (based on [13]) available at
http://dollar.biz.uiowa.edu/ burer/software/Max-AO/index.html.
The boundary point method code is available at
http://www.math.uni-klu.ac.at/or/Software/.

Here is a rough indication of the times needed to compute the bounds in Tables 2-5.

To compute bounds for Paley graph $P_q$ we needed less than a minute when $q \leq 100$ and around 45 minutes to compute $\ell_{+}^{(3)}(P_{809})$.

Each bound in Tables 3-4 could be computed in less than a minute, as it involves a relatively small SDP; for instance, computing $\Psi_\ell(H(20,6))$ is via an SDP with 1502 variables and 47 blocks with sizes ranging from 1 to 43.

It was harder to compute the bounds $\psi_K$ in Table 3. In fact, we had to rerun the boundary point code several times for each instance in order to tailor the parameters of the code and speed up the convergence to an optimal solution. The computation times for the parameter $\psi_K(G)$ vary from a few minutes (e.g. less than 3 minutes for DCJC125.5, about 25 minutes for DCJC125.1) till four days for the most demanding instance DSJR500.1c.

## Acknowledgements

We thank Marco Chiarandini and Michael Trick for telling us about coloring results for DIMACS benchmark graphs. We also thank Janez Povh, Franz Rendl and Angelika Wiegele for adapting their boundary point algorithm code in such a way that it now exploits the block-diagonal structure in semidefinite programs.

# Chapter 7

# Perspectives

For this thesis, conclusions and recapitulations of achievements were given at the ends of the most important chapters. In this chapter we outline possible directions for future investigations.

In this thesis, we have investigated the hierarchies of semidefinite relaxations for the stability number $\alpha(G)$, the fractional chromatic number $\chi^*(G)$ and the chromatic number $\chi(G)$ of a graph $G$. All these hierarchies have either the Lovász theta number $\vartheta(G)$ of a graph $G$, or some variation of it, as the first relaxation. Since the Lovász theta number was introduced in 1979, It has been applied in many ways. For instance, $\overline{\vartheta}(G)$ is used for approximately colouring the graph (see [53, 26, 46]) and the link between $\vartheta(G)$ and the Delsarte bound for codes was observed in [83].

The second order bounds in the new block diagonal hierarchy led to stronger bounds for the stability numbers, e.g. of 'code' graphs in [85, 32, 60], of orthogonality graphs in [51], and of Paley graphs in this thesis. However, these results are just numerical. It will be interesting to see if some analytical bounds can be derived from some of the hierarchies studied in this thesis. For example, $\vartheta(P_q) = \sqrt{q}$ for Paley graph $P_q$ and it is interesting to study if an analogous analytical formula exists, e.g. for the second order bound $\psi^{(2)}(P_q)$ in the new block diagonal hierarchy, since the computational results indicate that it can be much stronger than $\vartheta(P_q)$.

The new block diagonal hierarchy $L_t(G)$ of relaxations of $\mathrm{STAB}(G)$ is just an application, to the stable set problem, of the more general hierarchy $L_t(K)$, $K$ being a convex cone, presented in [40]. It has been shown in [40] that $L_t(K)$ outperforms the hierarchy of Lovász and Schrijver $N^t(K)$. The hierarchy $N^t(K)$ is well studied, and it would be interesting to see if $L_t(K)$ has, apart from being computationally less costly, more properties which are stronger than the properties of $N^t(K)$ observed, e.g., in [65, 59, 33].

One can also think of developing codes for solving the stable set and the colouring problem. We have conducted some preliminary experiments using the bound $\psi_K(G)$. We developed a heuristic for colouring graphs, which can be upgraded to an exact 'branch and bound' algorithm. Preliminary results gave graph colourings with quality comparable to 'state of the art' heuristics.

# Bibliography


[1] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming*, 58:295–324, 1993.

[2] G. P. Barker, L. O. Eifler, and T. P. Kezlan. A non-commutative spectral theorem. *Linear Algebra and Applications*, 20(2):95–100, 1978.

[3] M. Bellare and M. Sudan. $\chi(G)$ is not approximable within $|G|^{1/14-\epsilon}$. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing*, pages 184–193, 1994.

[4] A. Ben-Tal and A. S. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Society for Industrial and Applied Mathematics, Philadelphia, 2001.

[5] C. Berge. Färbung von Graphen, deren sämtliche bzw. deren ungeraden Kreise starr sind (zusammenfassung). *Wissenschaftliche Zeitschrift der Martin-Luther-Universität Halle-Wittenberg Mathematisch-Naturwissenschaftliche Reihe*, 10(114), 1961.

[6] C. Berge. Some classes of perfect graphs. *Six Papers on Graph Theory*, pages 1–21, 1963.

[7] B. Bollobás. *Random Graphs*. Cambridge University Press, Cambridge, 2001.

[8] I. M. Bomze and E. de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24(2):163–185, 2002.

[9] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. J. Quist, and T. Terlaky. On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization*, 18(4):301–320, 2000.

[10] J. A. Bondy and P. Hell. A note on the star chromatic number. *Journal of Graph Theory*, 14(4):479–482, 1990.

[11] I. Broere, D. Döman, and J. N. Ridley. The clique numbers and chromatic numbers of certain Paley graphs. *Quaestiones Mathematicae*, 11(1):91–93, 1988.

[12] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. Preprint, Optimization Online, 2007.

[13] S. Burer, R. Monteiro, and Y. Zhang. Maximum stable set formulations and heuristics based on continuous optimization. *Mathematical Programming*, 94:137–166, 2002.

[14] S. Burer and D. Vandenbussche. Solving lift-and-project relaxations of binary integer programs. *SIAM Journal on Optimization*, 16(3):726–750, 2006.

[15] S. Busygin. A new trust region technique for the maximum weight clique problem. *Discrete Applied Mathematics*, 154(15):1080–1096, 2006.

[16] M. Caramia and P. Dell'Olmo. Bounding vertex coloring by truncated multistage branch and bound. *Networks*, 44(4):231–242, 2004.

[17] M. Caramia and P. Dell'Olmo. Coloring graphs by iterated local search traversing feasible and infeasible solutions. *Discrete Applied Mathematics*, 156(2):201–217, 2008.

[18] M. Chiarandini. *Stochastic Local Search Methods for Highly Constrained Combinatorial Optimisation Problems*. PhD thesis, Darmstadt University of Technology, 2005.

[19] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas. The strong perfect graph theorem. *Annals of Mathematics*, 164(1):51–229, 2006.

[20] V. Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4:305–337, 1973.

[21] P. Delsarte. *An Algebraic Approach to the Association Schemes of Coding Theory*. [Philips Research Reports Supplements (1973) No. 10] Philips Research Laboratories, Eindhoven, 1973.

[22] C. Desrosiers, P. Galinier, and A. Hertz. Efficient algorithms for finding critical subgraphs. *Discrete Applied Mathematics*, 156(2):244–266, 2008.

[23] R. Diestel. *Graph Theory (Graduate Texts in Mathematics)*. Springer, Berlin, 2005.

[24] I. Dukanovic and F. Rendl. Copositive programming motivated bounds on the clique and the chromatic number. Preprint, Optimization Online, 2006.

[25] I. Dukanovic and F. Rendl. Semidefinite programming relaxations for graph coloring and maximal clique problems. *Mathematical Programming*, 109(2):345–365, 2007.

[26] I. Dukanovic and F. Rendl. A semidefinite programming-based heuristic for graph coloring. *Discrete Applied Mathematics*, 156(2):180–189, 2008.

[27] P. Erdös, Z. Füredi, A. Hajnal, P. Komjáth, and V. Rödl. Coloring graphs with locally few colors. *Discrete Mathematics*, 59:21–34, 1986.

[28] P. Galinier, A. Hertz, and N. Zufferey. An adaptive memory algorithm for the k-colouring problem. *Discrete Applied Mathematics*, 156(2):267–279, 2008.

[29] P. Galinier and Hao J. K. Hybrid evolutionary algorithms for graph coloring. *Journal of Combinatorial Optimization*, 3:379–397, 2007.

[30] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, 1979.

[31] K. Gatermann and P. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192:95–128, 2004.

[32] D. Gijswijt. *Matrix Algebras and Semidefinite Techniques for Codes*. PhD thesis, University of Amsterdam, 2005.

[33] M. X. Goemans and L. Tunçel. When does the positive semidefiniteness constraint help in lifting procedures? *Mathematics of Operations Research*, 26(4):796–815, 2001.

[34] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[35] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Algorithms and Combinatorics. Springer-Verlag, Berlin, New York, 1988.

[36] N. Gvozdenović and M. Laurent. Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization. To appear in *SIAM Journal on Optimization*.

[37] N. Gvozdenović and M. Laurent. The operator $\Psi$ for the chromatic number of a graph. To appear in *SIAM Journal on Optimization*.

[38] N. Gvozdenović and M. Laurent. Semidefinite bounds for the stability number of a graph via sums of squares of polynomials, In M. Jünger and V. Kaibel (eds.). *Proceedings of 11th International IPCO Conference, Lecture Notes in Computer Science*, 3509:136–151, 2005.

[39] N. Gvozdenović and M. Laurent. Semidefinite bounds for the stability number of a graph via sums of squares of polynomials. *Mathematical Programming*, 110(1):145–173, 2007.

[40] N. Gvozdenović, M. Laurent, and F. Vallentin. Block-diagonal semidefinite programming hierarchies for 0/1 programming. Preprint, arXiv: 0712.3079, 2007.

[41] G. Hahn, P. Hell, and S. Poljak. On the ultimate independence ration of a graph. *European Journal of Combinatorics*, 16:253–261, 1995.

[42] J. Håstad. Clique is hard to approximate within $\mathrm{n}^{\text{1-epsilon}}$. *Acta Mathematica*, 182(1):105–142, 1999.

[43] P. Hell, X. Yu, and H. Zhou. Independence ratios of graph powers. *Discrete Mathematics*, 127(1-3):213–220, 1994.

[44] C. Helmberg. Semidefinite Programming for Combinatorial Optimization. Habilitationsschrift, TU Berlin, ZIB-Report ZR-00-34, Konrad-Zuse-Zentrum Berlin, 2000.

[45] C. Helmberg and F. Oustry. Bundle methods and eigenvalue functions, In H. Wolkowicz, R. Saigal, L. Vandenberghe (eds.). *Handbook of Semidefinite Programming*, pages 307–337, 2000.

[46] D. R. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM*, 45(2):246–265, 1998.

[47] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.

[48] L. G. Khachiyan. A polynomial algorithm in linear programming, (in russian). *Doklady Akedamii Nauk SSSR*, 244:1093–1096, 1979. (English translation: Soviet Mathematics Doklady, 20, 191–194, 1979).

[49] E. de Klerk. *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications.* Applied Optimization Series. Kluwer Academic, Boston, MA, 2002.

[50] E. de Klerk and D. V. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, 12(4):875–892, 2002.

[51] E. de Klerk and D. V. Pasechnik. A note on the stability number of an orthogonality graph. *European Journal of Combinatorics*, 28(7):1971–1979, 2007.

[52] E. de Klerk, D. V. Pasechnik, and A. Schrijver. Reduction of symmetric semidefinite programs using the regular $*$-representation. *Mathematical Programming*, 109(2-3):613–624, 2007.

[53] E. de Klerk, D. V. Pasechnik, and J. P. Warners. On approximate graph colouring and max-k-cut algorithms based on the theta-function. *Journal of Combinatorial Optimization*, 8(3):267–294, 2004.

[54] D. E. Knuth. The sandwich theorem. *The Electronic Journal of Combinatorics*, 1:1–48, 1994.

[55] J. Körner, C. Pilotto, and G. Simonyi. Local chromatic number and Sperner capacity. *Journal of Combinatorial Theory, Series B*, 95(1):101–117, 2005.

[56] J. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[57] J. B. Lasserre. An explicit exact sdp relaxation for nonlinear 0-1 programs, In K. Aardal and B. Gerards (eds.). *Proceedings of the 8th International IPCO Conference, Lecture Notes in Computer Science*, 2081:293–303, 2001.

[58] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. To appear in IMA volume *Emerging Applications of Algebraic Geometry*, M. Putinar and S. Sullivant (eds.).

[59] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research*, 28(3):470–496, 2003.

[60] M. Laurent. Strengthened semidefinite programming bounds for codes. *Mathematical Programming*, 109(2-3):239–261, 2007.

[61] M. Laurent and F. Rendl. Semidefinite programming and integer programming, In K. Aardal, G. Nemhauser, R. Weismantel (eds.). *Handbook on Discrete Optimization*, pages 393–514, 2005.

[62] L. Lovász. A characterization of perfect graphs. *Journal of Combinatorial Theory, Series B*, pages 95–98, 1972.

[63] L. Lovász. Kneser's conjecture, chromatic number, and homotopy. *Journal of Combinatorial Theory, Series A*, 25(3):319–324, 1978.

[64] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:1–7, 1979.

[65] L. Lovász and A. Schrijver. Cones of matrices and set-functions and $0-1$ optimization. *SIAM Journal on Optimization*, 1:166–190, 1991.

[66] E. Maistrelli and D. B. Penman. Some colouring problems for Paley graphs. *Discrete Mathematics*, 306(1):99–106, 2006.

[67] J. Matoušek and G. M. Ziegler. Topological lower bounds for the chromatic number: A hierarchy. *Jahresbericht der DMV*, 106:71–90, 2004.

[68] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey. The Lovász' bound and some generalizations. *Journal of Combinatorics, Information & System Sciences*, 3:134–152, 1978.

[69] I. Méndez-Díaz and P. Zabala. A branch-and-cut algorithm for graph coloring. *Discrete Applied Mathematics*, 154(5):826–847, 2006.

[70] P. Meurdesoif. Strengthening the Lovász theta($G$) bound for graph coloring. *Mathematical Programming*, 102(3):577–588, 2005.

[71] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Túran. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[72] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.

[73] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization.* Wiley-Interscience, New York, 1988.

[74] Y. E. Nesterov and A. S. Nemirovski. *Interior Point Polynomial Algorithms in Convex Programming.* SIAM Publications. SIAM, Philadelphia, 1994.

[75] P. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization.* PhD thesis, California Institute of Technology, 2000.

[76] J. Peña, J. Vera, and L. F. Zuluaga. Computing the stability number of a graph via linear and semidefinite programming. *SIAM Journal on Optimization*, 18(1):87–105, 2007.

[77] S. Poljak. A note on stable sets and colorings of graphs. *Commentationes Mathematicae Universitatis Carolinae*, 15(2):307–309, 1974.

[78] G. Pólya. *Collected Papers*. MIT Press, Cambridge, Mass., London, Vol. 2:309-313, 1974.

[79] J. Povh. *Application of Semidefinite and Copositive Programming in Combinatorial Optimization*. PhD thesis, Universtity of Ljubljana, 2006.

[80] J. Povh, F. Rendl, and A. Wiegele. A boundary point method to solve semidefinite programs. *Computing*, 78(3):277–286, 2006.

[81] B. Reznick. Sums of even powers of real linear forms. *Memoirs of the American Mathematical Society*, 96(463), 1992.

[82] R. T. Rockafellar. *Covex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[83] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Transactions on Information Theory*, 25:425–429, 1979.

[84] A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency*. Springer-Verlag, Berlin, 2003.

[85] A. Schrijver. New code upper bounds from the Terwilliger algebra and semidefinite programming. *IEEE Transactions on Information Theory*, 51(8):2859–2866, 2005.

[86] J. B. Shearer. Lower bounds for small diagonal Ramsey numbers. *Journal of Combinatorial Theory, Series A*, 42(2):302–304, 1986.

[87] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3):411–430, 1990.

[88] G. Simonyi. Asymptotic values of the Hall-ratio for graph powers. *Discrete Mathematics*, 306(19-20):2593–2601, 2006.

[89] T. Stephen and L. Tunçel. On a representation of the matching polytope via semidefinite liftings. *Mathematics of Operations Research*, 24(1):1–7, 1999.

[90] M. Szegedy. A note on the theta number of Lovász and the generalized Delsarte bound. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 36–39, 1994.

[91] T. Terlaky (Ed.). *Interior Point Methods of Mathematical Programming*. Applied Optimization No.5. Kluwer Academic Publishers, 1996.

[92] F. Vallentin. Symmetry in semidefinite programs. Preprint, arXiv: 0706.4233, 2007.

[93] A. Vince. Star chromatic number. *Journal of Graph Theory*, 12:551–559, 1988.

[94] L. A. Wolsey. *Integer Programming*. John Wiley and Sons, New York, 1998.

[95] X. Zhu. Circular chromatic number: a survey. *Discrete Mathematics*, 229(1-3):371–410, 2001.

[96] G. M. Ziegler. Coloring Hamming graphs, optimal binary codes, and the 0/1-Borsuk problem in low dimensions, In. *Computational Discrete Mathematics, Advanced Lectures, Lecture Notes in Computer Science*, 2122:159–172, 2001.

# Index

# List of notation

## Sets

| | | |
|---|---|---|
| $\mathbb{Z},\ \mathbb{N},\ \mathbb{R},\ \mathbb{C},\ \mathbb{R}_+$ | : | sets of integer, nonnegative integer, real, complex and nonnegative real numbers (respectively). |
| $\|V\|$ | : | size of a finite set $V$; |
| $\mathcal{P}(V)$ | : | collection of all subsets of a finite set $V$; |
| $\mathcal{P}_{=t}(V)\ \ (\mathcal{P}_{\leq t}(V))$ | : | collection of all subsets of size (at most) $t$ of a finite set $V$; |
| $V\backslash S$ | = | $\{i \in V \mid i \notin S\}$; |
| $V \triangle S$ | = | $(V\backslash S) \cup (S\backslash V)$; |
| $\mathcal{P}(V,S)$ | = | $\{(I,J) \in \mathcal{P}(S) \times \mathcal{P}(V\backslash S) \mid \|I\| = \|J\|\}$; |
| $\mathrm{Sym}(V)\,(\mathrm{Sym}(n))$ | : | group of permutations of $V$ (when $\|V\| = n$); |
| $\mathcal{K}^*$ | = | $\{M \in \mathbb{R}^{m\times n} \mid \langle M, N\rangle \geq 0 \text{ for all } N \in \mathcal{K}\}$ (dual cone of a set $\mathcal{K} \subseteq \mathbb{R}^{m\times n}$); |
| $\mathrm{Int}\ \mathcal{K}$ | : | interior of a set $\mathcal{K}$; |
| $\mathrm{conv}(F)$ | = | $\left\{\sum_{x\in F} \lambda_x x \mid \lambda_x \geq 0\ (x \in F),\ \sum_{x\in F} \lambda_x = 1\right\}$ (where $F$ is finite). |

## Vectors and Matrices

| | | |
|---|---|---|
| $M^{-1},\ M^T,\ M^*$ | : | inverse, transpose, conjugate transpose of $M$ (respectively); |
| $M_{ij},\ M_{i\bullet},\ M_{\bullet j}$ | : | $ij$th entry, $i$th row, $j$th column of $M$ (respectively); |
| $M(p,q)$ | : | $pq$th block of a block matrix $M$; |
| $M \geq N$ | : | all entries of $M - N$ are nonnegative; |
| $M \succeq N$ | : | $M - N$ is symmetric positive semidefinite. |

### Sets of vectors and matrices

| | | |
|---|---|---|
| $\mathbb{R}^n\ (\mathbb{C}^n)$ | : | $n$-dimensional real (complex) vector space; |
| $\mathbb{R}^n_+$ | : | positive orthant of $\mathbb{R}^n$; |
| $\mathbb{R}^V\ (\mathbb{R}^{V\times W})$ | : | space of real vectors (matrices) with rows indexed by $V$(and columns indexed by $W$); |

Cones of matrices:

$$
\begin{array}{rcl}
\mathcal{S}_n &=& \{M \in \mathbb{R}^{n\times n} \mid M^T = M\} \text{ (symmetric matrices);}\\
\mathcal{S}_n^+ &=& \{M \in \mathcal{S}_n \mid M \succeq 0\} \text{ (positive semidefinite matrices);}\\
\mathcal{S}_n^{++} &=& \{M \in \mathcal{S}_n \mid M \succ 0\} \text{ (positive definite matrices);}\\
\mathcal{C}_n &=& \{M \in \mathcal{S}_n \mid v^T M v \geq 0 \text{ for all } v \in \mathbb{R}^n_+\}\\
&& \text{(copositive matrices);}\\
\mathcal{C}_n^* &=& \{M \in \mathcal{S}_n \,|\, M = N^T N \text{ for some } N \in \mathbb{R}_+^{k\times n} \text{ and } k \in \mathbb{N}\}\\
&& \text{(completely positive matrices);}\\
\mathcal{N}_n &=& \{M \in \mathcal{S}_n \mid M \geq 0\} \text{ (nonnegative matrices);}\\
\mathcal{D}_n &=& \mathcal{S}_n^+ \cap \mathcal{N}_n \text{ (doubly nonnegative matrices);}\\
\mathcal{D}_n^* &=& \mathcal{S}_n^+ + \mathcal{N}_n = \{M + N \mid M \in \mathcal{S}_n^+,\ N \in \mathcal{N}_n\};\\
M_V &=& \{Y \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)\times\mathcal{P}_{\leq 1}(V)} \mid Y_{ij} = Y_{ji}\ (i,j \in V),\\
&& \qquad\qquad\qquad\qquad Y_{jj} = Y_{\mathbf{0}j} = Y_{j\mathbf{0}}\ (j \in V)\},\\
M_{+,V} &=& \{Y \in M_V \mid Y \succeq 0\}.
\end{array}
$$

For a convex cone $K \subseteq \mathbb{R}^{\mathcal{P}_{\leq 1}(V)}$ :

$$
\begin{array}{rcl}
M(K) &=& \{Y \in M_V \,|\, Ye_k,\ Y(e_{\mathbf{0}} - e_k) \in K \text{ for } k \in V\};\\
N(K) &=& \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \,|\, x = Ye_{\mathbf{0}} \text{ for some } Y \in M(K)\};\\
M_+(K) &=& \{Y \in M_{+,V} \,|\, Ye_k,\ Y(e_{\mathbf{0}} - e_k) \in K \text{ for } k \in V\};\\
N_+(K) &=& \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \,|\, x = Ye_{\mathbf{0}} \text{ for some } Y \in M_+(K)\}.
\end{array}
$$

Matrix algebras:

$$
\begin{array}{rcl}
\mathcal{A}_n &=& \left\{\sum_{i,j,p=0}^n x_{i,j}^p M_{i,j}^{p,n} \mid x_{i,j}^p \in \mathbb{R}\right\}\\
&& \text{(Terwilliger algebra of the Hamming scheme);}\\
\mathcal{B}_n &=& \{\sum_{k=0}^n x_k^p M_k^n \mid x_k \in \mathbb{R}\}\\
&& \text{(Bose-Mesner algebra of the Hamming scheme);}\\
\mathcal{A}_{r,r'} &=& \left\{\sum_{i,j,p,i',j',q} x_{i,j,i',j''}^{p,q} M_{i,j}^{p,r} \otimes M_{i',j'}^{q,r'} \mid x_{i,j,i',j'}^{p,q} \in \mathbb{R}\right\};\\
\mathcal{B}_{r,r'} &=& \left\{\sum_{i,j,p,q} x_{i,j}^{p,q} M_{i,j}^{p,r} \otimes M_{i,j}^{q,r'} \mid x_{i,j}^{p,q} \in \mathbb{R}\right\};\\
\mathcal{B}_n^r &=& \{\sum_{k=0}^r x_k M_k^{n,r} \mid x_k \in \mathbb{R}\}.
\end{array}
$$

## Special vectors and matrices

$$
\begin{array}{rcl}
0 &:& \text{zero vector (or matrix) of size depending on the context;}\\
\mathbf{I}\ (\mathbf{I}_n\ ) &:& \text{identity matrix } (n \times n);\\
\mathbf{J}\ (\mathbf{J}_n\ ) &:& \text{all ones matrix } (n \times n);\\
e &:& \text{all ones vector;}\\
e_{\mathbf{0}} &=& e_{\emptyset},\ \text{standard unit vector indexed by } \mathcal{P}_{\leq 1}(V)\\
&& \text{with 1 at the entry with index } \emptyset;\\
e_i &=& e_{\{i\}},\ \text{standard unit vector indexed by } \mathcal{P}_{\leq 1}(V)\ (i \in V);\\
Z &:& \text{Zeta matrix, } \mathcal{P}(T) \times \mathcal{P}(T) \text{ block matrix, } Z(S,S') = \mathbf{I}_n\\
&& \text{if } S \subseteq S' \text{ and } Z(S,S') = 0 \text{ otherwise;}\\
Z^{-1} &:& \text{Möbius matrix, inverse of } Z,\ Z^{-1}(S,S') = (-1)^{|S'\setminus S|}\mathbf{I}_n\\
&& \text{if } S \subseteq S' \text{ and } Z^{-1}(S,S') = 0 \text{ otherwise;}\\
P_\sigma &:& \text{permutation matrix for a permutation } \sigma;
\end{array}
$$

| | | |
|---|---|---|
| $M_k^n$ | $\in$ | $\mathbb{R}^{\mathcal{P}(N)\times\mathcal{P}(N)}$, $N=\{1,2,\ldots,n\}$, $(M_k^n)_{IJ}=1$ if $\lvert I\triangle J\rvert=k$, and equal to 0 otherwise; |
| $M_{i,j}^{p,n}$ | $\in$ | $\mathbb{R}^{\mathcal{P}(N)\times\mathcal{P}(N)}$, $N=\{1,2,\ldots,n\}$, $(M_{i,j}^{p,n})_{IJ}=1$ if $\lvert I\rvert=i$, $\lvert J\rvert=j$, $\lvert I\cap J\rvert=p$, and equal to 0 otherwise; |
| $M_k^{n,r}$ | $\in$ | $\mathbb{R}^{\mathcal{P}(N,T)\times\mathcal{P}(N,T)}$, $N=\{1,2,\ldots,n\}$, $T\subseteq N$, $\lvert T\rvert=r$, $(M_k^{n,r})_{(I,I')(J,J')}=1$ if $\lvert I\triangle J\rvert+\lvert I'\triangle J'\rvert=2k$, and equal to 0 otherwise; |

## Functions of vectors and matrices

| | | |
|---|---|---|
| $\mathrm{diag}(M)$ | : | vector with entries $M_{ii}$; |
| $\mathrm{Diag}(v)$ | : | diagonal matrix with $M_{ii}=v_i$; |
| $\mathrm{Tr}(M)$ | = | $\sum_i M_{ii}$ (trace of $M$); |
| $\langle M,N\rangle$ | = | $\mathrm{Tr}(M^TN)$ (standard inner product of $M$ and $N$); |
| $M\otimes N$ | : | tensor product of $M$ and $N$, $(M\otimes N)_{(ij),(i'j')}=M_{ij}N_{i'j'}$; |
| $x(S)$ | = | $\sum_{i\in S}x_i$; |
| $\chi^S$ | : | characteristic vector of $S$; |
| $M_t(y)$ | $\in$ | $\mathbb{R}^{\mathcal{P}_{\le t}(V)\times\mathcal{P}_{\le t}(V)}$, combinatorial moment matrix of a vector $y\in\mathbb{R}^{\mathcal{P}_{\le 2t}(V)}$, $M_t(y)_{IJ}=y_{I\cup J}$; |
| $M(T;y)$ | : | $\mathcal{P}(T)\times\mathcal{P}(T)$ block matrix, submatrix of $M_t(y)$ indexed by (multi)set $\bigcup_{S\subseteq T}\{S,S\cup\{i\}\ (i\in V)\}$. |

## Notation for polynomials

| | | |
|---|---|---|
| $\mathbb{R}[x]$ | = | $\mathbb{R}[x_1,x_2,\ldots,x_n]$, ring of polynomials in $n$ variables; |
| $\mathbb{N}^n$ | : | $n$-tuples of nonnegative integers; |
| $I_=(n,t)$ $(I_\le(n,t))$ | : | $n$-tuples of integers whose sum is (at most) $t$. |

For $\beta\in\mathbb{N}^n$, $x\in\mathbb{R}^n$ :

| | | |
|---|---|---|
| $\lvert\beta\rvert$ | = | $\sum_{i=1}^n\beta_i$; |
| $\beta!$ | = | $\Pi_{i=1}^n\beta_i!$; |
| $S(\beta)$ | = | $\{i\mid\beta_i>0\}$; |
| $S_{odd}$ | = | $\{i\mid\beta_i$ is odd$\}$; |
| $x^\beta$ | = | $\Pi_{i=1}^n x_i^{\beta_i}$; |

For $x\in\mathbb{R}^n$, $p\in I_\le(n,t)$ :

| | | |
|---|---|---|
| $\zeta_{t,x}$ | $\in$ | $\mathbb{R}^{I_\le(n,t)}$, zeta vector of $x$, $(\zeta_{t,x})_\beta=x^\beta$; |
| $p(x)$ | = | $p^T\zeta_{t,x}=\sum_{\beta\in I_\le(n,t)}p_\beta x^\beta$. |

For $M\in\mathbb{R}^{n\times n}$ :

| | | |
|---|---|---|
| $p_M(x)$ | = | $\sum_{i,j=1}^n x_i^2x_j^2M_{ij}$; |
| $p_M^{(t)}(x)$ | = | $\left(\sum_{i=1}^n x_i^2\right)^t p_M(x)$; |

Moment matrices:

| | | |
|---|---|---|
| $N_t(y)$ | $\in$ | $\mathbb{R}^{I_\le(n,t)\times I_\le(n,t)}$, moment matrix of $y\in\mathbb{R}^{I_\le(n,2t)}$, $N_t(y)_{\beta,\gamma}=y_{\beta+\gamma}$; |
| $N_{=t}(y)$ | $\in$ | $\mathbb{R}^{I_=(n,t)\times I_=(n,t)}$, moment matrix of $y\in\mathbb{R}^{I_=(n,2t)}$, $N_{=t}(y)_{\beta,\gamma}=y_{\beta+\gamma}$; |

Cones:

$$
\begin{array}{rcl}
\mathcal{N}(n,2t) & = & \{y \in \mathbb{R}^{I_{\le}(n,2t)} \mid N_t(y) \succeq 0\};\\
\mathcal{N}_{=}(n,2t) & = & \{y \in \mathbb{R}^{I_{=}(n,2t)} \mid N_{=t}(y) \succeq 0\};\\
\Sigma(n,2t) & = & (\mathcal{N}(n,2t))^* = \{p \in \mathbb{R}^{I_{\le}(n,2t)} \mid p(x) \text{ is a sum of squares }\},\\
 & & \text{the cone of coefficient vectors of polynomials of degree}\\
 & & \text{at most } 2t \text{ which are sums of squares;}\\
\Sigma_{=}(n,2t) & = & (\mathcal{N}_{=}(n,2t))^* = \{p \in \mathbb{R}^{I_{=}(n,2t)} \mid p(x) \text{ is SOS }\},\\
 & & \text{the cone of coefficient vectors of homogeneous polynomials}\\
 & & \text{of degree } 2t \text{ which are sums of squares;}\\
\mathcal{K}_n^{(t)} & = & \{M \in \mathbb{R}^{n\times n} \mid p_M^{(t)}(x) \text{ is a sum of squares}\}\\
 & & = \{M \in \mathbb{R}^{n\times n} \mid p_M^{(t)} \in \Sigma_{=}(n, 2t+4)\};\\
\mathcal{C}_n^{(t)} & = & \left(\mathcal{K}_n^{(t)}\right)^*.
\end{array}
$$

# Graph theory

$$
\begin{array}{rcl}
G & = & (V,E) \text{ (graph with vertex set } V \text{ and edge set E);}\\
V(G) & : & \text{vertex set of a graph } G;\\
E(G) & : & \text{edge set of a graph } G;\\
ij & = & (i,j) \text{ (edge);}\\
\overline{G} & : & \text{complement of } G;\\
\overline{\beta}(G) & = & \beta\left(\overline{G}\right) \text{ (for a parameter } \beta \text{ mapping graphs to } \mathbb{R}_+);\\
\beta(\cdot) & = & \beta \text{ (for a parameter } \beta \text{ mapping graphs to } \mathbb{R}_+);\\
A_G & : & \text{adjacency matrix of } G;\\
G+H & = & (V(G)\cup V(H), E(G)\cup E(H)) \text{ (direct sum of } G \text{ and } H);\\
G\cdot H & : & \text{strong product of } G \text{ and } H;\\
G\square H & : & \text{Cartesian product of } G \text{ and } H;\\
G_t & = & G\square K_t
\end{array}
$$

## Stable set polytope and relaxations

$$
\begin{array}{rcl}
\mathrm{STAB}(G) & = & \mathrm{conv}(\{\chi^S \mid S \text{ stable in } G\});\\
\mathrm{FRAC}(G) & = & \{x \in \mathbb{R}_+^V \mid x_i + x_j \le 1 \ (ij \in E),\ x_i \le 1 \ (i \in V)\};\\
\mathrm{QSTAB}(G) & = & \{x \in \mathbb{R}_+^V \mid x(C) \le 1 \ (C \text{ clique in } G)\};\\
\mathrm{THETA}(G) & = & \{x \in \mathbb{R}^V \mid x = \mathrm{diag}(X),\ X \succeq xx^T,\ x_{ij} = 0\,(ij \in E)\};\\
\mathrm{ST}(G) & = & \left\{\lambda \begin{pmatrix} 1\\ x\end{pmatrix} \;\middle|\; x \in \mathrm{STAB}(G),\ \lambda \in \mathbb{R}_+\right\};\\
\mathrm{FR}(G) & = & \{x \in \mathbb{R}_+^{\mathcal{P}_{\le 1}(V)} \mid x_i + x_j \le x_{\mathbf{0}} \ (ij \in E),\ x_i \le x_{\mathbf{0}} \ (i \in V)\};\\
\mathrm{QST}(G) & = & \{x \in \mathbb{R}_+^{\mathcal{P}_{\le 1}(V)} \mid x(C) \le x_{\mathbf{0}} \ (C \text{ clique in } G)\};\\
\mathrm{TH}(G) & = & \{x \in \mathbb{R}^{\mathcal{P}_{\le 1}(V)} \mid x = Ye_{\mathbf{0}}, \text{ for some } Y \in M_{+,V}\\
 & & \qquad \text{with } Y_{ij} = 0 \ (ij \in E)\};\\
Q_t(G) & = & \{x \in \mathbb{R}^{\mathcal{P}_{\le 1}(V)} \mid \exists y \in \mathbb{R}^{\mathcal{P}_{\le 2t}(V)} \text{ satisfying } y_i = x_i \ (i \in V),\\
 & & \qquad y_{\mathbf{0}} = x_{\mathbf{0}},\ y_{ij} = 0 \ (ij \in E),\ M_t(y) \succeq 0\};\\
Q_t^+(G) & : & \text{strengthening of } Q_t(G) \text{ by } y \ge 0;
\end{array}
$$

$$
\begin{array}{rcl}
L_t(G) & = & \{x \in \mathbb{R}^{\mathcal{P}_{\leq 1}(V)} \mid \exists y \in \mathbb{R}^{\mathcal{P}_{\leq t+1}(V)} \text{ satisfying } y_i = x_i \ (i \in V), \\
& & \qquad y_{\mathbf{0}} = x_{\mathbf{0}},\ y_{ij} = 0 \ (ij \in E), \\
& & \qquad M(T;y) \succeq 0 \ (T \subseteq V,\ |T| = t-1)\}; \\
L_t^+(G) & : & \text{strengthening of } L_t(G) \text{ by } y \geq 0; \\
P^{(t)}(G) & = & \{x \in \mathbb{R}^n \mid x = \operatorname{diag}(X) \text{ for some } X \text{ satisfying} \\
& & \qquad \operatorname{Tr}(A_G X) = 0,\ X \succeq xx^T,\ X \in \mathcal{C}_n^{(t)}\}.
\end{array}
$$

## Graph parameters

$$
\begin{array}{rcl}
\Delta(G) & : & \text{maximum degree of } G; \\
\alpha(G) & : & \text{stability (independence) number of } G; \\
\omega(G) & : & \text{clique number of } G; \\
\chi(G) & : & \text{chromatic number of } G; \\
\chi^*(G) & : & \text{fractional chromatic number of } G; \\
\vartheta(G) & : & \text{Lovász theta number of } G; \\
\vartheta'(G) & : & \text{Schrijver's number of } G; \\
\vartheta^+(G) & : & \text{Szegedy's number of } G;
\end{array}
$$

Hierarchies (sequences) of upper bounds for $\alpha(G)$ :

$$
\begin{array}{rcl}
\operatorname{las}^{(t)}(G) & = & \max \ \sum_{i \in V(G)} x_i \text{ s.t. } x \in Q_t(G); \\
\operatorname{las}_+^{(t)}(G) & = & \max \ \sum_{i \in V(G)} x_i \text{ s.t. } x \in Q_t^+(G); \\
\ell^{(t)}(G) & = & \max \ \sum_{i \in V(G)} x_i \text{ s.t. } x \in L_t(G); \\
\ell_+^{(t)}(G) & = & \max \ \sum_{i \in V(G)} x_i \text{ s.t. } x \in L_t^+(G); \\
\vartheta^{(t)}(G) & = & \min \ \lambda \text{ s.t. } \lambda(\mathbf{I} + A_G) - \mathbf{J} \in \mathcal{K}_n^{(t)} \\
& & = \max \operatorname{Tr}(\mathbf{J}X) \text{ s.t. } \operatorname{Tr}((\mathbf{I} + A_G)X) = 1,\ X \in \mathcal{C}_n^{(t)}; \\
\widehat{\vartheta}^{(t)}(G) & = & \min \ \lambda \text{ s.t. } \lambda\mathbf{I} + yA_G - \mathbf{J} \in \mathcal{K}_n^{(t)} \\
& & = \max \operatorname{Tr}(\mathbf{J}X) \text{ s.t. } \operatorname{Tr}(X) = 1,\ \operatorname{Tr}(A_G X) = 0,\ X \in \mathcal{C}_n^{(t)}; \\
\widetilde{\vartheta}^{(t)}(G) & = & \max \ \sum_{i \in V(G)} x_i \text{ s.t. } x \in P^{(t)}(G);
\end{array}
$$

Hierarchies (sequences) of lower bounds for $\chi^*(G)$ :

$$
\begin{array}{rcl}
\psi^{(r)}(G) & = & \min \ t \text{ s.t. } x \in Q_r(G),\ x_{\mathbf{0}} = t,\ x_i = 1 \ (i \in V(G)); \\
\psi_+^{(r)}(G) & = & \min \ t \text{ s.t. } x \in Q_r^+(G),\ x_{\mathbf{0}} = t,\ x_i = 1 \ (i \in V(G)); \\
\widetilde{\psi}^{(r)}(G) & = & \min \ t \text{ s.t. } x \in Q_r(G),\ x_{\mathbf{0}} = t,\ x_i = 1 \ (i \in V(G)); \\
\widetilde{\psi}_+^{(r)}(G) & = & \min \ t \text{ s.t. } x \in Q_r^+(G),\ x_{\mathbf{0}} = t,\ x_i = 1 \ (i \in V(G)); \\
\widehat{\psi}^{(r)}(G) & = & \min \ t \text{ s.t. } \operatorname{diag}(X) = te,\ \operatorname{Tr}(A_G X) = 0,\ X - \mathbf{J} \succeq 0, \\
& & \qquad X \in \mathcal{C}_n^{(r)};
\end{array}
$$

Action of the operator $\Psi$ :

$$
\begin{array}{rclcrcl}
\frac{|V(\cdot)|}{\chi(\cdot)} \leq & \beta(\cdot) & \leq \alpha(\cdot) & \Longrightarrow & \Psi_\beta(G) & = & \chi(G); \\
\frac{|V(\cdot)|}{\omega(\cdot)} \leq & \beta(\cdot) & \leq \chi(\cdot) & \Longrightarrow & \Psi_\beta(G) & = & \omega(G); \\
\Psi_\vartheta(G) & = & \left\lceil \overline{\vartheta}(G) \right\rceil; & & \Psi_{\vartheta'}(G) & = & \left\lceil \overline{\vartheta^+}(G) \right\rceil.
\end{array}
$$

## Special graphs

| | | |
|---|---|---|
| $C_n$ | : | cycle, $V(C_n) = \{1, \ldots, n\}$ $(n \geq 3)$,<br>$E(C_n) = \{12, 23, \ldots, (n-1)n, n1\}$; |
| $K_n$ | : | complete graph, $V(K_n) = \{1, \ldots, n\}, E(K_n) = \mathcal{P}_{=2}(V(K_n))$; |
| $P_q$ | : | Paley graph, $q$ prime power, $V(P_q)$ is the set of elements of the finite field $\mathbb{F}_q, uv \in E(P_q)$ if $u - v$ is a square in $\mathbb{F}_q$; |
| $H(n, \mathcal{D})$ | : | Hamming graph, $\mathcal{D} \subseteq N = \{1, \ldots, n\},\ V(H(n, \mathcal{D})) = \mathcal{P}(N)$,<br>$IJ \in E(H(n, \mathcal{D}))$ if $\lvert I \triangle J \rvert \in \mathcal{D}$; |
| $K(n, r)$ | : | Kneser graph, $n \geq 2r, V(K(n,r)) = \mathcal{P}_{=r}(N),\ N = \{1, \ldots, n\}$,<br>$IJ \in V(K(n, r))$ if $I \cap J = \emptyset$. |

# Samenvatting

## Achtergrond en motivatie

Een graaf $G = (V, E)$ bestaat uit een verzameling punten $V$ en een verzameling lijnen $E$. Elke lijn verbindt twee punten met elkaar. Een verzameling punten van de graaf heet onafhankelijk als geen twee van deze punten met elkaar zijn verbonden door een lijn van de graaf. Een puntkleuring van de graaf is een toewijzing van kleuren aan de punten van de graaf, zó dat de twee eindpunten van een lijn altijd verschillend gekleurd zijn. Puntkleuringen en onafhankelijke verzamelingen zijn aan elkaar gerelateerd. Een puntkleuring van een graaf is namelijk precies een opdeling van zijn punten in een aantal onafhankelijke verzamelingen, één voor iedere kleur.

Veel problemen uit de praktijk, zoals het maken van roosters, productieplanning, het toewijzen van (radio)frequenties, patroonherkenning en het construeren van foutcorrigerende codes, kunnen gemodelleerd worden met behulp van puntkleuringen of onafhankelijke verzamelingen in een geschikte graaf. In deze toepassingen is men meestal geïnteresseerd in het vinden van een onafhankelijke verzameling van maximale grootte of een puntkleuring met een minimaal aantal verschillende kleuren.

De maximale grootte van een onafhankelijke verzameling punten in een graaf $G = (V, E)$ wordt aangegeven met $\alpha(G)$. Het chromatisch getal $\chi(G)$ van $G$ is het kleinste aantal kleuren in een puntkleuring van de graaf. Het bepalen van de getallen $\alpha(G)$ en $\chi(G)$ zijn NP-moeilijke problemen uit de combinatorische optimalisering.

Kort gezegd zijn NP-moeilijke problemen optimalisatie-problemen waarvoor geen efficiënte algoritmen bekend zijn en waarvan veel mensen verwachten dat het vinden van zo'n algoritme zelfs onmogelijk is. Met efficiënt wordt bedoeld dat het algoritme het probleem oplost binnen een aantal stappen dat begrensd is door een polynoom in de lengte van de input voor het probleem ("in polynomiale tijd"). Door het ontbreken van efficiënte algoritmen is het in het algemeen nodig om dergelijke problemen middels alternatieve methoden, zoals heuristieken, probabilistische methoden of benaderings-technieken te proberen op te lossen.

## Een overzicht

### Het Lovász theta getal

Een klassieke aanpak is om een combinatorisch optimalisatie-probleem te modelleren als een lineair programma waarbij bovendien geheeltalligheid van de oplossing wordt geëist. De maximale grootte van een onafhankelijke verzameling punten $\alpha(G)$ in een graaf $G = (V, E)$, kan bijvoorbeeld als volgt worden beschreven. We nemen aan dat de punten van $G$ genummerd zijn van 1 tot en met $n$.

$$\alpha(G) = \max \sum_{i \in V} x_i \text{ waarbij } x_i + x_j \leq 1 \ (ij \in E), \ x \in \{0,1\}^n. \tag{7.1}$$

Hier is $x_i$ een variabele behorende bij punt $i$ en staat $ij$ voor de lijn tussen punten $i$ en $j$. Op deze manier is de moeilijkheid van het probleem verstopt in de voorwaarde $x \in \{0,1\}^n$.

Om tot een semidefiniete benadering te komen voor (7.1), wordt met de vector $x$ de matrix $X = \binom{1}{x}\binom{1}{x}^T$ geassocieerd. De rijen en kolommen van $X$ zijn geïndexeerd door de punten $1, \ldots, n$ van de graaf en een extra index 0. Er geldt dat $x \in \{0,1\}^n$, dan en slechts dan als van de matrix $X$ rij 0 gelijk is aan de diagonaal gelijk zijn. Verder is $X$ positief semidefiniet en er geldt $X_{ij} = 0$ als $x_i + x_j \leq 1$ voor punten $i$ en $j$. De semidefiniete benadering $\vartheta(G)$ voor $\alpha(G)$, het zogeheten Lovász theta getal, wordt gegeven door het semidefiniete programma

$$\vartheta(G) := \max \sum_{i \in V} X_{ii} \text{ waarbij } X_{00} = 1, \ X_{ij} = 0 \ (ij \in E), \ X \succeq 0. \tag{7.2}$$

Hier betekent $X \succeq 0$ dat $X$ positief semidefiniet moet zijn. De belangrijkste motivatie om semidefiniete programma's te beschouwen, is dat zij, binnen elke gewenste precisie, in polynomiale tijd kunnen worden opgelost en betere benaderingen geven dan de eenvoudigere lineaire benadering die verkregen wordt door de conditie $x \in \{0,1\}^n$ af te zwakken tot $0 \leq x_i \leq 1$ $(i = 1, \ldots, n)$.

Zoals we reeds hebben gezien, kunnen puntkleuringen van een graaf $G = (V, E)$ gezien worden als een opsplitsing van de verzameling punten in een aantal onafhankelijke verzamelingen. Hiermee kan het chromatisch getal worden geformuleerd als

$$\chi(G) = \min \sum_S \lambda_S \text{ waarbij } \sum_S \lambda_S \chi^S = e, \ \lambda_S \in \{0,1\},$$

waarbij de sommaties lopen over alle onafhankelijke verzamelingen $S \subseteq V$, $e \in \mathbb{R}^V$ de vector van enkel enen is ($e_i = 1$ $(i \in V)$) en $\chi^S$ de karakteristieke vector is van de deelverzameling $S$, dat wil zeggen $\chi_i^S := 1$ als $i \in S$ en $\chi_i^S := 0$ als $i \notin S$. De lineaire benadering $\chi^*(G)$, het gebroken chromatisch getal, wordt verkregen door de 0/1 conditie te versoepelen tot $0 \leq \lambda_S \leq 1$. Beschouw bij een gegeven oplossing $\lambda$ (geïndiceerd door alle onafhankelijke verzamelingen $S$) de matrix

$$Y := \sum_{S \subseteq V \text{ onafh.}} \lambda_S \begin{pmatrix} 1 \\ \chi^S \end{pmatrix} \begin{pmatrix} 1 \\ \chi^S \end{pmatrix}^T.$$

De matrix $Y$ is positief semidefiniet en voldoet aan $Y_{\emptyset\emptyset} = \sum_{S \text{ onafh.}} \lambda_S$, $Y_{ii} = Y_{i0} = 1$ $(i \in V)$ en $Y_{ij} = 0$ $(ij \in E)$. Dit leidt tot een semidefiniete benadering voor $\chi^*(G)$ (en voor $\chi(G)$) door deze eigenschappen als voorwaarde op te leggen aan een matrix variabele $Y$. Het blijkt dat deze benadering gelijk is aan $\vartheta(\overline{G})$, het Lovász theta getal van de complementaire graaf $\overline{G} = (V, \overline{E})$, waar $\overline{E}$ bestaat uit alle paren punten die geen lijn vormen in $G$. Uitgeschreven:

$$\vartheta(\overline{G}) = \min Y_{00} \text{ waarbij } Y_{0i} = Y_{ii} = 1 \ (i \in V),\ Y_{ij} = 0 \ (ij \in E),\ Y \succeq 0.$$

Hiermee is het duidelijk dat dat de volgende bewering geldt, die bekend staat als 'the sandwich theorem':

$$\alpha(G) \leq \vartheta(G) \leq \chi^*(\overline{G}) \leq \chi(\overline{G}).$$

**Resultaten in dit proefschrift**

Er zijn voor combinatorische optimalisatie-problemen verscheidene semidefiniete benaderingen voorgedragen die sterker zijn dan de meest eenvoudige semidefiniete benadering. Daarbij worden extra voorwaarden toegevoegd, zoals in Lovász en Schrijver [65] door het toevoegen van matrix-snedes ('matrix cuts') of zoals in Lasserre [57] middels het gebruik van 'moment matrices' waarbij het probleem opgetild wordt naar een hoger dimensionale ruimte, aldaar wordt opgelost en vervolgens weer wordt terug geprojecteerd naar de oorspronkelijke ruimte. Voor het getal $\alpha(G)$ bijvoorbeeld, geven deze methodes een hiërarchie van grenzen $\nu^{(t)}$ $(t = 1, 2, 3, \ldots)$ die voldoen aan:

$$\alpha(G) \leq \cdots \leq \nu^{(t+1)}(G) \leq \nu^{(t)}(G) \leq \cdots \leq \nu^{(1)}(G) \leq \vartheta(G).$$

In dit proefschrift brengen we verschillende hiërarchieën van semidefiniete programma's in herinnering, samen met de corresponderende bovengrenzen voor $\alpha(G)$ en maken een vergelijking tussen deze verschillende methoden. De grens van eerste orde is in elk van deze hiërarchieën gelijk aan het Lovász theta getal (of een variatie daarop). De grenzen van orde $\alpha(G)$ komen in de meeste hiërarchieën overeen met $\alpha(G)$ zelf. Dat wil zeggen $\nu^{(\alpha(G))} = \alpha(G)$. Door de aanpak van Lovász en Schrijver [65] te combineren met die van Lasserre [56, 57], zijn we in staat om een nieuwe hiërarchie te definieren.

Deze is een relaxatie van de hiërarchie van Lasserre, en heeft daardoor het voordeel dat de optredende semidefiniete programma's in blok-diagonaal vorm kunnen worden gebracht, hetgeen zeer belangrijk is bij daadwerkelijke berekeningen. Bovendien vergt de hiërarchie minder rekenkracht en is tenminste zo sterk als de hiërarchie van Lovász en Schrijver.

Vervolgens definieren en bestuderen we hiërarchieën van bovengrenzen voor het (gebroken) chromatisch getal. Deze hiërarchiën zien er als volgt uit:

$$\vartheta(\overline{G}) \leq \psi^{(1)}(G) \leq \cdots \leq \psi^{(t)}(G) \leq \cdots \leq \psi^{(t+1)}(G) \leq \cdots \leq \chi^*(G);$$

$$\lceil \vartheta(\overline{G}) \rceil \leq \Psi_{\nu^{(1)}}(G) \leq \cdots \leq \Psi_{\nu^{(t)}}(G) \leq \cdots \leq \Psi_{\nu^{(t+1)}}(G) \leq \cdots \leq \chi(G).$$

Hier is $\Psi$ een speciale operator die bovengrenzen voor $\alpha(G)$ afbeeldt op ondergrenzen voor het chromatisch getal. Als toepassing hiervan, bewijzen we dat er geen graafparameter kan bestaan die is ingeklemd tussen het gebroken chromatisch getal $\chi^*(G)$ en het chromatisch getal $\chi(G)$ en die bovendien in polynomiale tijd kan worden berekend, tenzij alle problemen in de complexiteitsklasse NP in polynomiale tijd kunnen worden opgelost, dat wil zeggen, als P=NP.

We berekenen deze grenzen in de nieuwe blok-diagonale hiërarchie voor een aantal interessante klassen van grafen. In het bijzonder zijn we in staat om deze grenzen, tot en met orde 3, te berekenen voor Payleygrafen met tot 800 punten door de eigenschappen van hun automorfismegroepen te benutten. Voor Hamminggrafen en Knesergrafen met tot $2^{20}$ punten, bepalen we de grenzen van orde 1 en 2 door gebruik te maken van de expliciete blokdiagonalisatie van de Terwilliger-algebra van het Hammingschema, zoals gegeven door Schrijver in [85]. Tenslotte introduceren we $\psi_C$ (met $C$ een onafhankelijke verzameling punten in $\overline{G}$), een andere variatie op de tweede orde grens in de hiërarchie van Lasserre. De grens $\psi_C(G)$ wordt gedefinieerd via een semidefiniet programma dat ook in blok-diagonaal vorm kan worden gebracht. We beschrijven computationele resultaten voor een aantal instanties uit de DIMACS benchmark, een lijst geselecteerde probleeminstanties voor het testen van algoritmen voor het puntkleuren van grafen. Deze resultaten wijzen erop dat $\psi_C(G)$ een vrij sterke grens is voor het chromatisch getal $\chi(G)$.