A variant of the Lovász-Theta number based on projection matrices *

Francesco Silvestri †‡
March 15, 2018

Abstract

We introduce a new model for the chromatic number \( \chi(G) \) based on what we call combinatorial projection matrices, which is a special class of doubly stochastic symmetric projection matrices. Relaxing this model yields an SDP whose optimal value is the projection theta number \( \hat{\vartheta}(G) \), which is closely related to the Szegedy number \( \vartheta^+(G) \), a variant of the Lovász theta number. We characterize that in general, \( \hat{\vartheta}(G) \leq \vartheta^+(G) \), with equality if \( G \) is vertex-transitive. While this seems to imply that working with binary matrices is a better paradigm than working with binary eigenvalues in this context, our approach is slightly faster than computing the Szegedy number on vertex-transitive graphs.

Keywords: Graph colouring, Lovász \( \vartheta \)-Function, Semidefinite Programming

1 Introduction

Background

The chromatic number \( \chi(G) \) of a graph \( G \) is the minimum number of colours needed to colour the nodes of \( G \) in such a way that adjacent nodes receive distinct colours. A stable set in \( G \) is a set of nodes \( S \subseteq V \) such that no pair of nodes is adjacent, so a colouring is a partition of \( V \) into stable sets. Both characterizing stable sets and computing \( \chi(G) \) are classical NP-hard problems [13]. Furthermore, it is even NP-hard to approximate \( \chi(G) \) within \( |V(G)|^{1/14-\varepsilon} \), for any \( \varepsilon > 0 \) [1].

In the standard Integer Program Formulation of this problem \([3]\), each stable set \( S_j \) is represented by a characteristic vector \( x_j \) and the number of such vectors being used is minimized. This gives the problem a highly symmetrical structure, as any permutation of the colours \( \{x_j\}_{j \in [k]} \) results in new colourings of equal objective. As a consequence, methods based on convex relaxations like the LP relaxation usually perform badly unless they break these symmetries. Unfortunately, even though symmetry breaking constraints are available \([3][6]\), LP relaxations usually still suffer from poor bounds due to the hardness of the problem.

On the other hand, the colouring problem can also be formulated as a rank constrained matrix problem which implicitly models the colouring \( X \) as a sum of matrices \( X_j \) corresponding to stable sets \([11]\). This removes the symmetry, as any permutation of the colours \( X_j \) will only change the order of summation, which does not change the colouring \( X \). This formulation can then be relaxed to an SDP relaxation and the resulting bound \( \overline{\vartheta}(G) \) is well-know as the Lovász theta number \( \vartheta(G) \) applied on the complement graph \([9]\). In general, and in particular for perfect graphs, \( \overline{\vartheta}(G) \) is usually much better than the LP bound \([13]\).

*The authors gratefully acknowledge support by the German Research Foundation (DFG), grant GRK 1653.
†Institut für Informatik, Heidelberg University, INF205, 69120 Heidelberg, Germany
‡IWR, Heidelberg University, INF205, 69120 Heidelberg, Germany
Summary

In this paper introduce and analyse a new formulation for $\chi(G)$ based on combinatorial projection matrices, which is constructed to admit an SDP relaxation. The resulting SDP relaxation is closely related to the Szegedy-number, a variant of the Lovász theta number, which we compare against. While the latter relaxes matrices with binary entries, our approach relaxes binary eigenvalues instead, and we are able to prove explicit relations between both relaxations.

Related Work

A lot of work has been dedicated to variants of $\overline{\theta}(G)$ with the hope of improving towards $\chi(G)$. Early work mostly consists of adding linear inequalities based on non-negativity and triangle inequalities, summarized in [4]. Other work [5][8] shows that converging SDP hierarchies of growing sizes are available, but these relaxations quickly grow too big for actual computations. Finally, [7] compares the preceding results in a unified way while also using a reduction of colouring onto the stable set problem to yield even better bounds.

Organization of the paper

In Section 2, we shortly recall the relevant results about colouring. Section 3 reviews what we call combinatorial projection matrices and states a few basic properties. In Section 4 we establish a new formulation for the colouring number $\chi(G)$ based on combinatorial projection matrices and introduce the projection $\vartheta$-number $\vartheta(G)$ as SDP relaxation of this formulation. Additionally, we compare $\vartheta(G)$ with the Szegedy-number $\vartheta^+(G)$ analytically and show that in general, $\vartheta^+(G) \leq \vartheta(G)$, while both numbers agree on vertex-transitive graphs. Finally, Section 5 investigates if the method of moments can help to close the gap between the two relaxations.

1.1 Notation

Small letters like $a, x, \lambda$ are scalars or elements of a set and capital letters like $S, T, U$ describe sets. Small bold letters are used for vectors like $\mathbf{a}, \mathbf{x}, \lambda$ and capital bold letters for matrices like $\mathbf{A}, \mathbf{X}, \mathbf{A}$. Instead of writing index sets of small size like $\{i\}$ or $\{i, j\}$, we may write $i$ or $ij$ respectively.

For any $n \in \mathbb{N}$, let $[n] := \{1, \ldots, n\}$. $\mathcal{G}_n$ denotes the set of simple graphs $G$ whose node set is $[n]$ and whose edge set is $E(G)$. The adjacency matrix of $G \in \mathcal{G}_n$ is denoted by $\mathbf{A}_G$.

Given a set $I \subseteq [n]$, its characteristic vector $\mathbf{e}_I \in \{0, 1\}^n$ is coordinate-wise defined as

$$
(e_I)_i = \begin{cases} 
1 & \text{if } i \in I, \\
0 & \text{else.}
\end{cases}
$$

Geometrically, the map $I \mapsto \mathbf{e}_I$ bijectively maps the power set $2^n$ to the vertices of the $n$-dimensional unit cube $[0, 1]^n$. If $n$ is clear from context, we will write $\mathbf{e}$ instead of $\mathbf{e}_n$. The matrix $\mathbf{I}_n$ is the $n$-dimensional identity, $\mathbf{J}_{n_1,n_2} = \mathbf{e}_{[n_1]} \mathbf{e}_{[n_2]}^\top$ is the matrix of all-ones, and we might write $\mathbf{J}_n$ or simply $\mathbf{J}$ instead of $\mathbf{J}_{n,n}$ if clear from context. For a matrix $\mathbf{A}$, $\text{col}(\mathbf{A})$ denotes the set of its columns, and if it is quadratic, $\text{spt}(\mathbf{A})$ denotes its spectrum and $\text{tr}(\mathbf{A})$ its trace.

$\mathbb{R}_+^n$ is the cone of vectors in $\mathbb{R}^n$ with non-negative entries with conic order $\leq$ and $\mathbb{S}^n_+$ is the cone of positive semidefinite $n \times n$ matrices with conic order $\preceq$. The standard simplex is denoted as

$$
\Delta^n := \{ \mathbf{x} \in \mathbb{R}_+^n \mid \langle \mathbf{x}, \mathbf{e} \rangle = 1 \}.
$$

The Kronecker-delta $\delta_\psi$ assumes the truth values $0/1$ of the statement $\psi$, where $\delta_{i,j}$ is short for $\delta_{i,j}$.
2 Preliminaries

This section serves as a short recapitulation of the basics about graph colouring. For an extensive survey on convex relaxations for graph colouring, consider [7].

2.1 Stable Sets

Definition 1:
A stable set of a graph \( G \in \mathcal{G}_n \) is a subset \( S \subseteq \{0,1\}^n \) such that the subgraph induced by \( S \) does not contain any edges. The set of stable sets in \( G \) is denoted by \( \mathcal{S}_G \subseteq 2^{[n]} \) and satisfies the properties of an independence system. The stable set number \( \alpha(G) \) is the biggest size \( |S| \) of any stable set in \( G \) and defines a function \( \alpha : \mathcal{G}_n \rightarrow \mathbb{N} \). The clique number \( \omega(G) \) of \( G \) is the biggest size \( |S| \) of a stable set in the complement graph \( \overline{G} \), so \( \omega(G) := \alpha(\overline{G}) \).

Formally, we can express \( \alpha(G) \) as the solution to an integer problem by working with characteristic vectors \( x \in \{0,1\}^n \) via

\[
\alpha(G) = \max \{ \langle x, e \rangle \mid x, x_j = 0 \quad \forall (i,j) \in E, \ x \in \{0,1\}^n \} \\
= \max \{ \langle x, e \rangle \mid x^\top A_G x = 0, \ x \in \{0,1\}^n \},
\]

where \( A_G \) denotes the adjacency matrix of \( G \).

2.2 Graph Colourings

Definition 2:
A \( k \)-colouring of a graph \( G \in \mathcal{G}_n \) is a \( k \)-partition \( T = (T_1, \ldots, T_k) \) of the node set \( [n] \), such that each \( T_i \) is a stable set in \( G \). The chromatic number \( \chi(G) \) is the smallest value \( k \) such that \( G \) has a \( k \)-colouring. The set of \( k \)-colourings of \( G \) is denoted by \( \mathcal{C}_k(G) \)

In the following, we will describe how to represent \( k \)-colourings as assignment matrices by treating their characteristic vectors as columns of an \( n \times k \) binary matrices. To this end, let \( S^*_G := \mathcal{S}_G \setminus \{\emptyset\} \) to denote

\[
V(S_G) = \{e_I \in \{0,1\}^n \mid I \in \mathcal{S}_G^* \} = \{x \in \{0,1\}^n \mid \text{supp}(x) \in \mathcal{S}_G^* \}
\]

be the vertex-set associated with \( S_G \) and define the set of \( S_G \)-constrained assignment matrices

\[
\mathcal{U}_{n,k}(S_G) = \{U \in \{0,1\}^{n \times k} \mid U e_{[k]} = e_{[n]}, \ \text{col}(U) \subseteq V(S_G) \}. \tag{2}
\]

Additionally, we will also use the shorthand

\[
\mathcal{U}_{n,k} := \mathcal{U}_{n,k}(2^{[n]}) = \{U \in \{0,1\}^{n \times k} \mid U e_{[k]} = e_{[n]}, \ 0 \notin \text{col}(U) \}.
\]

With this notation at hand, we can express the colouring number simply as

\[
\chi(G) = \min \{k \mid \mathcal{U}_{n,k}(S_G) \neq \emptyset \}.
\]

Example 3:
Let \( G \in \mathcal{G}_3 \) be the graph given by the adjacency matrix

\[
A_G := \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
1 & 1 & 0
\end{pmatrix}.
\]

Then \( \{\{1,2\}, \{3\}\} \) is a valid colouring, and ordering the corresponding characteristic vectors lexicographically leads to the assignment matrix

\[
U = \begin{pmatrix}
1 & 0 \\
1 & 0 \\
0 & 1
\end{pmatrix} \in \mathcal{U}_{3,2}(S_G).
\]
The following hardness result shows that an explicit description of neither $U_{n,k}(S_G)$ nor $S_G$ can easily be found:

**Theorem 4 ([13, 1]):**
Computing $\alpha(G)$ or $\chi(G)$ is NP-hard. Furthermore, it is NP-hard to approximate $\chi(G)$ within $n^{1/4-\varepsilon}$ for any $\varepsilon > 0$.

In spite of being NP-hard to compute in general, $\chi(G)$ behaves much nicer restricted to special classes of graphs with regards to the following construction.

**Definition 5 (Lovász theta number [9]):**
The Lovász theta number $\vartheta(G)$ is the optimal value of the following primal/dual SDP problems

\[
\begin{align*}
\max \{ \langle X, I_n \rangle \mid \langle X, I_n \rangle = 1, \langle X, A_G \rangle = 0, X \succeq 0 \}, \quad & (\vartheta-P) \\
\min \left\{ k \left| \text{diag}(Y) = e, \langle Y, A_G \rangle = 0, \begin{pmatrix} k & e^T \\ e & Y \end{pmatrix} \succeq 0 \right. \right\}. \quad & (\vartheta-D)
\end{align*}
\]

**Theorem 6 (Lovász sandwich Theorem [10]):**
For all $G \in \mathcal{G}_n$, we have

$$\omega(G) \leq \vartheta(G) \leq \chi(G).$$

This theorem is a celebrated result for perfect graphs, where $\omega(H) = \chi(H)$ for all induced subgraphs $H \subseteq G$, and $\chi(G)$ can be computed in polynomial time. For more on the theory of perfect graphs, consider the recent survey [15].

Despite the hardness results of Theorem 4, a lot of publications are dedicated to improve $\vartheta(G)$ for imperfect graphs. We exemplarily introduce the following simple improvements taken from [4] and [7], where several other relaxations for $\chi(G)$ may be found.

**Remark 7:**
Adding $X \succeq 0$ to $(\vartheta-P)$ and $Y \succeq 0$ to $(\vartheta-D)$ results in the Schrijver number $\vartheta^-(G)$ [12] and the Szegedy number $\vartheta^+(G)$ [14] respectively. In particular, we get

$$\omega(G) \leq \vartheta^-(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \chi(G).$$

### 3 Combinatorial projection matrices

For the rest of this paper, let $R = (r_1, \ldots, r_n) = \{(r_{ij})_{i,j \in [n]} \in \mathbb{R}^{n \times n}$. We first recall some basic properties about projection matrices arising from orthogonal projections onto subspaces.

**Definition 8:**
Let $X \in \mathbb{R}^{n \times k}$ be a matrix with full column rank to define its corresponding projection matrix as

$$\rho(X) := X(X^T X)^{-1} X^T \in \mathbb{R}^{n \times n}.$$

The set of symmetric projection matrices of size $n$ is defined as

$$\text{SPM}^n := \{ R \in \mathbb{R}^{n \times n} \mid R^2 = R, R = R^T \} \subseteq \mathcal{S}_n.$$

Note that by assumption, $X^T X$ is invertible and $\rho(X) \in \text{SPM}^n$. We also recall some important properties of projection matrices:

**Lemma 9:**
Let $R \in \text{SPM}^n$ and $X \in \mathbb{R}^{n \times k}$ with full column rank. Then the following holds:

(i) Binary eigenvalues: $\text{spec}(R) \subseteq \{0, 1\}$,

(ii) Rank equation: $\text{tr}(R) = \text{rank}(R)$,

(iii) Projection property: $\rho(X) X = X$, $\rho(X) \cdot \mathbb{R}^n = X \cdot \mathbb{R}^k$. 


(iv) Isometry invariance: $\rho(X) = \rho(XQ)$ for all orthogonal $Q \in \mathbb{R}^{n \times n}$.

Proof. (i) follows from the Cayley-Hamilton theorem applied to the matrix polynomial $R^2 = R$. For (ii), note that $\text{tr}(R)$ is equivalent to the sum of its eigenvalues, which count the rank due to (i). (iii) and (iv) follow directly from the definition.

Example 10:
In Example 3, we considered the assignment matrix of the 2-colouring $\{\{1, 2\}, \{3\}\}$ given by

$$U = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Applying $\rho$ to $U$ leads us to the corresponding projection matrix

$$\rho(U) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Denoting the columns as $\rho(U) = (r_1, r_2, r_3)$, we observe for the following that

- $U^T U$ is a diagonal matrix which contains the size of parts of the partition,
- the columns $\{r_1, r_2, r_3\}$ consist of all eigenvectors of $\rho(U)$ with repetition,
- $\{r_1, r_2, r_3\} \subseteq \Delta^3$,
- $\|r_i\|_0 \cdot \|r_i\|_\infty = 1$ holds for each column $i \in [3]$,
- $r_i = r_j$ if and only if $(r_i)_j = (r_j)_i > 0$.

The preceding example motivates the following definition of a special class of doubly stochastic matrices.

**Definition 11:**
The set of combinatorial projection matrices $\text{CPM}_k^n(S_G)$ is given as

$$\text{CPM}_k^n(S_G) = \left\{ R \in \mathbb{R}^{n \times n} \mid \begin{array}{c} \text{col}(R) \subseteq \Delta^n_{S_G} \\ r_{ij} \cdot (r_i - r_j) = 0 \forall i, j \in [n] \\ R = R^T \\ \text{tr}(R) = k \end{array} \right\},$$

where

$$\Delta^n_{S_G} := \{ r \in \Delta^n \mid \text{supp}(r) \in S_G \}.$$  

The quadratic equations

$$r_{ij} \cdot (r_i - r_j) = 0 \quad \forall i, j \in [n]$$

will be called block-inducing in the following.

As expected, this set contains the projection matrices corresponding to assignment matrices, which is shown by the following two lemmas.

**Lemma 12:**
Combinatorial projection matrices are projection matrices

$$\text{CPM}_k^n(S_G) \subseteq \text{SPM}^n$$

and have strictly positive diagonal $\text{diag}(R) > 0$. In particular, each column $r_i$ is uniquely determined by its support via

$$r_{ji} = \delta_{j \in \text{supp}(r_i)} \frac{1}{\|r_i\|_0}$$

and satisfies the equation

$$\|r_i\|_0 \cdot \|r_i\|_\infty = 1.$$
Proof. Let $R \in \text{CPM}_k^n(S_G)$ and choose any $i,j,l \in [n]$. By assumption, the block-inducing equations $r_{il} \cdot (r_i - r_l) = 0$ and $r_{ij} \cdot (r_i - r_j) = 0$ show the identity

$$r_{il}r_{jl} = r_{il}r_{ij} = r_{ij}r_{jl},$$

using the symmetry of $R$. Then $R^2 = R$ follows from

$$(R^2)_{ij} = \sum_{t \in [n]} r_{it}r_{jt} = \sum_{t \in [n]} r_{ij}r_{jt} = r_{ij}\left(\sum_{t \in [n]} r_{jt}\right) = r_{ij} = (R)_{ij} \quad \forall i,j \in [n].$$

Furthermore, the block-inducing equations $r_{ij} \cdot (r_i - r_j) = 0$ show $r_{ii}(r_{ij} - r_{ii}) = 0$ for all $j \in [n]$, which implies $r_{ij} \in \{0, r_{ii}\}$. Since $r_i \in \Delta^n$, we necessarily have $r_{ii} \cdot \|r_i\|_0 = 1$, which completes the proof.

**Theorem 13:**
The combinatorial projection matrices are precisely the projection matrices corresponding to assignment matrices:

$$\rho(\mathcal{U}_{n,k}(S_G)) = \text{CPM}^n_k(S_G) \subseteq \text{SPM}^n.$$

Proof. We first show $\rho(\mathcal{U}_{n,k}(S_G)) \subseteq \text{CPM}^n_k(S_G)$, so let $U = (u_1, \ldots, u_k) \in \mathcal{U}_{n,k}(S_G)$. We already know $\rho(U) \in \text{SPM}^n$ and the corresponding properties from Lemma 9, so we only need to check the first two properties outlined in Definition 11.

For the first, assume $(u_j)_i = 1$ to see

$$\rho(U)e_i = U(U^TU)^{-1}U^T e_i = U(U^TU)^{-1} = \frac{e_j}{(u_j, u_j)} = \frac{u_j}{(e, u_j)} \in \Delta^n_{S_G},$$

where we used that $u_j$ is binary in the second to last equation.

For the second, assume $e_i^\top \rho(U)e_j > 0$, since $\rho(U) \geq 0$ and there is nothing to show otherwise. Assuming $(u_{i'})_j = 1$ and $(u_{i'})_i = 1$, we have

$$0 < e_i^\top \rho(U)e_j = e_i^\top U(U^TU)^{-1} U^T e_j = e_i^\top(U^TU)^{-1} e_{i'} = \delta_{i', j'} \cdot (U^TU)_{i', i'},$$

since $(U^TU)^{-1}$ is diagonal. It follows that $i' = j'$, which in turn means

$$U^T e_i = e_{i'} = e_{j'} = U^T e_j.$$ In particular,

$$\rho(U)(e_i - e_j) = U(U^TU)^{-1} (U^T (e_i - e_j)) = 0,$$

which we wanted to show.

The reverse inclusion $\rho(\mathcal{U}_{n,k}(S_G)) \supseteq \text{CPM}^n_k(S_G)$ follows if we can show that $\rho$ is surjective. To show this, we argue that for $R \in \text{CPM}^n_k(S_G)$, the map

$$\psi : R \rightarrow \psi(R) =: \{\text{supp}(r_i) \mid i \in [n]\} \in C_k(G)$$

is well-defined and $\rho \circ \psi$ yields the identity on $\text{CPM}^n_k(S_G)$.

Since $r_i \in \Delta^n_{S_G}$, we have $\psi(R) \subseteq S_G$. Disjointness of the sets in $\psi(R)$ follows from the implications of the block-inducing equations, and coverage of $[n]$ follows due to $R$ being doubly stochastic and so $\psi(R) \in C_k(G)$ is well-defined.

Finally, the map $\rho \circ \psi$ is the identity on $\text{CPM}^n_k(S_G)$ since $R$ is uniquely reconstructed from $\psi(R)$ by starting with $i \in \text{supp}(r_i)$ and applying Lemma 12. $\square$
3.1 Convexification

In the next section, we will reduce the computation of \( \chi(G) \) to the task of optimizing a linear function over \( \text{CPM}_k^n(S_G) \), which shows that a compact description of \( \text{conv}(\text{CPM}_k^n(S_G)) \) is most likely out of reach due to Theorem 4. For this reason, we are interested in studying relaxations of this convex hull.

The set \( \text{CPM}_k^n(S_G) \) is constructed to make an SDP relaxation immediately available by replacing the non-linear constraints

\[
    r_{ij} \cdot (r_i - r_j) = 0 \quad \forall i, j \in [n]
\]

with the psd. constraint \( R \succeq 0 \), which is implied by the constraint \( R^2 = R \) from being a projection matrix. We thus get the set

\[
    \text{RPM}_k^n(S_G) := \{ R \in \mathbb{R}_+^{n \times n} \mid \text{tr}(R) = k, \ Re = e, \ (R, A_G) = 0, \ R \succeq 0 \}.
\]

4 The projection \( \vartheta \)-number

We now propose a new formulation in terms of projection matrices and relate it to \( \vartheta(G) \).

**Theorem 14:**

We have the following characterization of the chromatic number:

\[
\chi(G) = \min \left\{ k \mid \text{CPM}_k^n(S_G) \neq \emptyset \right\} = \min \left\{ \text{tr}(R) \mid \text{col}(R) \subseteq \Delta^n, \ r_{ij} = 0 \quad \forall(i, j) \in E(G), \ r_{ij} \cdot (r_i - r_j) = 0 \quad \forall i, j \in [n], \ R = R^\top \right\}.
\]

**Proof.** The first equation follows by definition and Theorem 13, so we will show the second. Let \( \chi'(G) \) denote the optimal value of the second optimization problem and let \( T \) be a minimal colouring with \( \chi(G) \) colours. Then the projection matrix \( R(T) \) corresponding to \( T \) is feasible for the second optimization problem, since \( \text{supp}(R(T) e_j) \in S_G \) and \( \text{diag}(R(T)) > 0 \) imply \( e_j^\top R(T) e_j = 0 \) whenever \( (i, j) \in E(G) \), and so \( \chi'(G) \leq \chi(G) \).

To see the other inequality, first note that \( \chi'(G) \geq 0 \) is integral, since whenever \( R \) is feasible, \( R \in \text{SPM}^n \) as shown in the proof of Lemma 12, and so the eigenvalues of \( R \) are binary. It suffices to show that \( R \in \text{CPM}_k^n(R(T)) \) then, and the only thing that is left to prove is \( \text{supp}(r_i) \in S_G \).

Suppose that \( \text{supp}(r_i) \notin S_G \), so there is \( (j, l) \in E(G) \) such that \( (r_i)_j, (r_i)_l > 0 \). It follows from \( r_{ij} \cdot (r_i - r_j) = 0 \) that \( (r_j)_l = (r_i)_l > 0 \), which contradicts \( (r_j)_l = 0 \). \( \square \)

The preceding theorem is important in that it makes the constraint \( \text{supp}(r_i) \in S_G \) tractable, so that we can approximate \( \chi(G) \) by relaxing the set \( \text{CPM}_k^n(S_G) \). Following Section 3.1, we can immediately state the following relaxation.

**Definition 15:**

The projection \( \vartheta \)-number is given as

\[
\vartheta(G) := \min \left\{ k \in \mathbb{R}_+^n \mid \text{RPM}_k^n(S_G) \neq \emptyset \right\} = \min \{ \text{tr}(R) \mid R \succeq 0, \ Re = e, \ (R, A_G) = 0, \ R \succeq 0 \}.
\]

We should note that by non-negativity of \( R \), we have

\[
(R, A_G) = 0 \quad \iff \quad r_{ij} = 0 \quad \forall(i, j) \in E(G),
\]

which we will prefer for notation. For the following, we explicitly recall the Szegedy number we have already seen in Section 2, which is given as

\[
\vartheta^+(G) := \min \left\{ x_0 \mid \begin{pmatrix} x_0 & e^\top \\ e & X \end{pmatrix} \succeq 0, \ \text{diag}(X) = e, \ (X, A_G) = 0, \ X \succeq 0 \right\}.
\]
It makes sense to compare the quality of both relaxations, since they have similar computational complexity and similar constraints. The following theorem shows that unfortunately, the classical approach is always at least as good as using projection matrices.

**Theorem 16:**
Let \( G \in \mathcal{G}_n \), then \( \vartheta^+(G) \geq \hat{\vartheta}(G) \).

**Proof.** Let \( x_0 = \vartheta^+(G) \) and \( X \) be an optimal solution to (4). Since \( X \) is symmetrical and non-negative, it follows from [2] that there exists a diagonal matrix \( D = \text{Diag}(d) \) with \( d > 0 \) such that

\[
\begin{pmatrix}
1 & 0 \\
0 & D
\end{pmatrix}
\begin{pmatrix}
\vartheta^+(G) & e^T \\
e & X
\end{pmatrix}
\begin{pmatrix}
1 & 0 \\
0 & D
\end{pmatrix}
= \begin{pmatrix}
\vartheta^+(G) & d^T \\
d & DXD
\end{pmatrix} \succeq 0,
\]

where \( R := DXD \) is doubly stochastic, so \( Re = e \). In particular, \( R \) is feasible for (3), and we have \( \text{tr}(R) = \|d\|_2^2 \). Using positive semidefiniteness, for any \( \mu \in \mathbb{R} \), we have

\[
0 \leq \begin{pmatrix}
-1
\end{pmatrix}^T \begin{pmatrix}
\vartheta^+(G) & d^T \\
d & R
\end{pmatrix} \begin{pmatrix}
-1
\end{pmatrix} = \vartheta^+(G) - 2\mu\|d\|_2^2 + \mu^2(d^T Rd)
\]

which becomes

\[
\vartheta^+(G) \geq 2\mu\|d\|_2^2 - \mu^2(d^T Rd) = \|d\|_2^2 \left(2\mu - \mu^2\left(\frac{d^T Rd}{\|d\|_2^2}\right)\right)
\]

\[
\geq \|d\|_2^2 \left(2\mu - \mu^2\right) = \text{tr}(R) (2\mu - \mu^2).
\]

Since \( \max \left\{ 2\mu - \mu^2 \mid \mu \in \mathbb{R} \right\} = 1 \), it follows that \( R \) is a feasible solution to (3) that satisfies \( \vartheta^+(G) \geq \text{tr}(R) \) and the theorem follows.

To get a better understanding of the projection model, it would be beneficial to get a quantitative bound on the magnitude of the gap \( \vartheta^+(G) - \hat{\vartheta}(G) \). Fortunately, we can state an explicit asymptotic lowerbound.

**Theorem 17:**
The worst case gap absolute \( \vartheta^+(G) - \hat{\vartheta}(G) \) has asymptotic behaviour

\[
\max \left\{ \vartheta^+(G) - \hat{\vartheta}(G) \mid G \in \mathcal{G}_n \right\} = \Omega(n).
\]

In particular,

\[
\limsup_{n \to \infty} \max \left\{ \frac{\vartheta^+(G) - \hat{\vartheta}(G)}{n} \mid G \in \mathcal{G}_n \right\} \geq \left(\frac{3}{\sqrt{2}} - 2\right) \approx 0.1213.
\]

**Proof.** We will explicitly construct a graph family for which the bound on the gap holds true asymptotically. To this end, consider for any two integers \( n_1, n_2 \in \mathbb{N} \), the graph

\[ G(n_1, n_2) := K_{n_1} \cup K_{n_2}, \]

which we define as the union of two complete graphs with \( n_1 \) and \( n_2 \) nodes respectively. In particular, the corresponding adjacency matrix is given as

\[
A_{G(n_1, n_2)} = \begin{pmatrix}
J_{n_1} - I_{n_1} & 0 \\
0 & J_{n_2} - I_{n_2}
\end{pmatrix},
\]

and since these graphs are perfect, we explicitly have

\[
\chi(G(n_1, n_2)) = \vartheta^+(\overline{G(n_1, n_2)}) = \omega(G(n_1, n_2)) = \max\{n_1, n_2\}.
\]
For $\hat{\vartheta}(G(n_1, n_2))$, we will construct the optimal solution $R$ in closed form. Since this graph has multiple symmetries, we can assume $R$ to be symmetry invariant and parametrize the feasible set with only three parameters $\alpha, \beta, \gamma$ by setting

$$R = \begin{pmatrix} \alpha I_{n_1} \\ \beta J_{n_2, n_1} \\ \gamma I_{n_2} \end{pmatrix}.$$ 

Now that the constraint $(R, A_{G(n_1, n_2)}) = 0$ is satisfied, the remaining affine constraints turn into

$$R \geq 0 \iff \alpha, \beta, \gamma \geq 0,$$

$$Re = e \iff \alpha + n_2\beta = 1, \quad \gamma + n_1\beta = 1.$$ 

Assuming $\alpha, \gamma \neq 0$ to make the condition $R \succeq 0$ non-trivial, we can use the Schur complement to rewrite

$$R \succeq 0 \iff \gamma I_{n_2} - \frac{\beta^2}{\alpha} J_{n_2, n_1} \cdot J_{n_1, n_2} \succeq 0 \iff \gamma \geq \frac{\beta^2}{\alpha} n_1 n_2,$$

where we used the fact that $J_n$ only has one non-zero eigenvalue given by $n$. Lastly, we can explicitly express the objective function as

$$\text{tr}(R) = \alpha n_1 + \gamma n_2 = n_1 + n_2 - 2n_1 n_2\beta,$$

using the affine constraints. Ignoring the constants, the resulting problem of computing the number $\hat{\vartheta}(G(n_1, n_2))$ is equivalent to

$$\max \left\{ \beta \mid \alpha + n_2\beta = 1, \quad \gamma + n_1\beta = 1, \quad \alpha \gamma \geq \beta^2 n_1 n_2, \quad \alpha, \beta, \gamma \geq 0 \right\}.$$ 

Using the equations, one can show that the unique solution is

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \frac{1}{n_1 + n_2} \begin{pmatrix} n_1 \\ 1 \\ n_2 \end{pmatrix}$$

and

$$\hat{\vartheta}(G(n_1, n_2)) = \frac{n_1^2 + n_2^2}{n_1 + n_2}.$$ 

In particular, we now have the gap

$$\Delta(n_1, n_2) := \vartheta^{*}(G(n_1, n_2)) - \hat{\vartheta}(G(n_1, n_2)) = \max(n_1, n_2) - \frac{n_1^2 + n_2^2}{n_1 + n_2}. \quad (5)$$

W.l.o.g., let $n_1 = m, n_2 = \mu m \in \mathbb{N}$ for some $\mu \in [0, 1]$. Then (5) reads

$$\Delta(m, \mu m) = m - \frac{(1 + \mu^2)m^2}{(1 + \mu)m} = \mu \left( \frac{1 - \mu}{1 + \mu} \right) m.$$ 

Finally, optimizing the choice $\mu \in [0, 1]$ yields the biggest theoretical gap for $\mu = \sqrt{2} - 1$ and

$$\max_{\mu \in [0, 1]} \Delta(m, \mu m) = (3 - 2\sqrt{2})m.$$ 

For growing $m$, we can approximate this gap arbitrarily well by choosing $n_1 = m$ and $n_2 = \lceil (\sqrt{2} - 1)m \rceil$ to get a graph of size $\lceil \sqrt{2}m \rceil$ with asymptotic relative gap

$$\lim_{m \to \infty} \sup \frac{\Delta(m, \lceil (\sqrt{2} - 1)m \rceil)}{\sqrt{2}m} = \frac{3 - 2\sqrt{2}}{\sqrt{2}} \approx 0.1213.$$ 

$\square$
As a corollary of the preceding theorem, the following result sheds some light on the discrepancy between the two relaxations by relating $\hat{\vartheta}(G)$ to Theorem 6.

**Corollary 18:**
The inequation $\omega(G) \leq \hat{\vartheta}(G)$ does not hold in general. In particular, $\hat{\vartheta}(G)$ is not necessarily exact for perfect graphs $G$.

Despite this disadvantage over the classical formulation, we can show some useful properties for the class of vertex-transitive graphs. We first cite the property in question.

**Theorem 19** ([14]):
For all $G \in G_n$, the inequality
$$\vartheta^+(G) \cdot \vartheta^-(G) \geq n$$
holds, with equality if $G$ is vertex-transitive.

As shown in the following, this inequality can be sharpened by using $\hat{\vartheta}(G)$.

**Lemma 20:**
For all $G \in G_n$, the inequality $\hat{\vartheta}(G) \cdot \vartheta^-(G) \geq n$ holds.

**Proof.** Let $R$ be an optimal solution to (3). Recalling the definition
$$\vartheta^-(G) = \max \{ \langle X, J \rangle \mid \text{tr}(X) = 1, \langle X, A_G \rangle = 0, X \succeq 0, X \succeq 0 \}$$
as $(\vartheta-P)$ with non-negativity constraints, we see that $X := \frac{1}{\text{tr}(R)} R$ is a feasible solution and as desired,
$$\vartheta^-(G) \geq \langle X, J \rangle = \frac{\langle R, J \rangle}{\text{tr}(R)} = \frac{n}{\hat{\vartheta}(G)}.$$

**Theorem 21:**
For vertex-transitive $G \in G_n$, we have
$$\vartheta^+(G) = \hat{\vartheta}(G).$$
In particular, $\omega(G) = \hat{\vartheta}(G) = \chi(G)$ for vertex-transitive perfect $G \in G_n$.

**Proof.** Using Theorem 16, Lemma 20 and Theorem 19, we see that
$$\vartheta^+(G) \geq \hat{\vartheta}(G) \geq \frac{n}{\vartheta^-(G)} = \hat{\vartheta}(G)$$
whenever $G \in G_n$ is vertex-transitive.

Vertex-transitive graphs are known examples for which the relaxations $\vartheta(G)$ and $\vartheta^+(G)$ perform badly, and it is surprising to see that an analogue of Theorem 6 for $\hat{\vartheta}(G)$ can be recovered in this case. In particular, any advantage of $\vartheta(G)$ over $\hat{\vartheta}(G)$ seems to be related to exploiting the lacking symmetries of a given graph.

## 5 Extension with the Method of Moments

The relaxation $\vartheta(G)$ is inherently tied to the method of moments, a technique from polynomial optimization. This method constructs a hierarchy of convex relaxations of growing size for polynomial optimization problems, which converge towards the global optimum. Applying it to (1) yields a formulation that can be transformed into $(\vartheta-P)$ as a first step in this hierarchy.

Following this approach, we can consider $\text{CPM}^\omega_n(S_G)$ as an algebraic variety and apply the method of moments to it as well. For our purposes, it will be enough to consider the first stage of the relaxation hierarchy; we will skip the intricacies of the general method of moments and
construct the relaxation of the first stage explicitly, which is motivated by the following line of thought:

Recalling that \( \mathbf{R} = (r_1, \ldots, r_n) \), we would like to work with the rank-1 matrices

\[
\mathbf{M}(\mathbf{R}) = \begin{pmatrix}
1 & \mathbf{r}_1^\top \\
\mathbf{r}_1 & \mathbf{r}_1 & \mathbf{r}_1^\top \\
\vdots & \vdots & \vdots \\
\mathbf{r}_n & \mathbf{r}_n & \mathbf{r}_n & \mathbf{r}_n^\top \\
\end{pmatrix},
\]

which contains all the products of entries of \( \mathbf{R} \), since all constraints in \( \text{CPM}_k^n(S_G) \) turn into linear constraint over \( \mathbf{M} \). Since this can not be done efficiently in practice, the method of moments proceeds by using all these linear constraints and applies it to the tractable matrix variable

\[
\mathbf{M}_i(\mathbf{R}) = \begin{pmatrix}
1 & \mathbf{r}_i^\top \\
\mathbf{r}_i & \mathbf{R}_{ii} & \mathbf{r}_i^\top \\
\vdots & \vdots & \vdots \\
\mathbf{r}_n & \mathbf{R}_{n1} & \mathbf{r}_n & \mathbf{r}_n^\top \\
\end{pmatrix} \succeq 0.
\]

instead of \( \mathbf{M}(\mathbf{R}) \), where the matrix blocks \( \mathbf{R}_{ij} \) are new variables that linearize the products \( \mathbf{r}_i \mathbf{r}_j^\top \). The linear constraints extracted from \( \text{CPM}_k^n(S_G) \) can be lineewise converted to the system

\[
\begin{align*}
\mathbf{R}_{ij} \mathbf{e} &= \mathbf{r}_i \\
\mathbf{R}_{ii} \mathbf{e}_i &= \mathbf{r}_{jj} \mathbf{e}_i = \mathbf{r}_{jj} \mathbf{e}_j \\
\langle \mathbf{r}_{ii}, \mathbf{A}_{ij} \rangle &= 0 \\
\sum_{i \in [n]} \text{tr}(\mathbf{R}_{ii}) &= k
\end{align*}
\]

for all choices of \( i, j \in [n] \). By invoking the block inducing equations, one can show that additionally, we can recover the original matrix \( \mathbf{R} \) from the \( \mathbf{R}_{ii} \) through the equation

\[
\sum_{i \in [n]} \mathbf{R}_{ii} = \mathbf{R}.
\]

Since the size of \( \mathbf{M}(\mathbf{R}) \) prohibitive, we relax it further by discarding all matrices \( \mathbf{R}_{ij} \) with \( i \neq j \). This relaxes \( \mathbf{M}_i(\mathbf{R}) \succeq 0 \) to the systems

\[
\mathbf{M}_i(\mathbf{r}_i) = \begin{pmatrix}
1 & \mathbf{r}_i^\top \\
\mathbf{r}_i & \mathbf{r}_i & \mathbf{r}_i^\top \\
\vdots & \vdots & \vdots \\
\mathbf{r}_n & \mathbf{r}_n & \mathbf{r}_n & \mathbf{r}_n^\top \\
\end{pmatrix} \succeq 0 \quad \forall i \in [n],
\]

and since \((-1 \quad \mathbf{e}_i^\top)^\top\) belongs to the kernel of these matrices, we can discard the first row and column from the formulation. This leads us to the strengthening

\[
\text{RPM}_k^n(S_G)' = \{ \mathbf{R} \in F \mid \text{tr}(\mathbf{R}) = k, \mathbf{R} \mathbf{e} = \mathbf{e}, \langle \mathbf{R}, \mathbf{A}_{ij} \rangle = 0 \},
\]

where

\[
F := \left\{ \mathbf{R} \in \mathbb{R}^{n \times n} \mid \mathbf{R} = \sum_{i \in [n]} \mathbf{R}_{ii}, \langle \mathbf{R}_{ii}, \mathbf{R}_{jj} \rangle_{il} = \langle \mathbf{R}_{ij}, \mathbf{R}_{ij} \rangle_{il}, \mathbf{R}_{ii} \in F_i, \forall i, j, l \in [n] \right\}
\]

is build up from the individual sets

\[
F_i := \{ \mathbf{X} \in \mathbb{R}^{n \times n} \mid \langle \mathbf{X}, \mathbf{J}_n \rangle = 1, \mathbf{X} e_i = \text{diag}(\mathbf{X}), \mathbf{X} \succeq 0 \}
\]

that are distinguished by the constraint \( \mathbf{X} e_i = \text{diag}(\mathbf{X}) \). In particular, the symmetry condition in \( F \) makes sure that we can consider each matrix \( \mathbf{R}_{ii} \) as a slice of a symmetrical third-order tensor, as shown in Figure 1.

The advantage of \( \text{RPM}_k^n(S_G)' \) over \( \text{RPM}_k^n(S_G) \) is the inclusion \( F \subseteq \mathbb{R}^{n \times n}_+ \cap S_n^3 \).
Theorem 19: The question if in general, approximating matrices with binary eigenvalues is harder than approximating matrices with binary entries. 

While this process can be used to construct symmetrical tensors of higher orders from appropriate combinatorial moment matrices of rank 1, this approach does not generalize in our setting. The reason for this is the assumption \( \text{rank}(\mathbf{R}) = k > 1 \), which breaks the symmetry; the equation \((\mathbf{R}_{ij})_{lm} = (\mathbf{R}_{kl})_{jm}\) will generally not hold, and so the entries associated with various permutations of \(\{i,j,l,m\}\) can not be assumed to be equal.

Figure 1: Different ways to align the variables of the various \(\mathbf{R}_{ij}\) into the same symmetrical third order tensor, each cube representing one variable.

We note that even by refining \(\hat{\vartheta}(G)\) by using the computationally expensive \(\text{RPM}_k^n(S_G)\)' instead of \(\text{RPM}_k^n(S_G)\), we cannot guarantee a lower-bound of \(\vartheta(G)\). To see this, let

\[
\hat{\vartheta}'(G) := \min \left\{ k \in \mathbb{R}_+ \mid \text{RPM}_k^n(S_G) \neq \emptyset \right\}
\]

be the corresponding strengthening of \(\hat{\vartheta}(G)\).

Empirical evidence shows that \(\hat{\vartheta}'(G) = \chi(G)\) holds for the graph class \(G(n_1, n_2)\) used as counter example in Theorem 17. However, this counter example can be generalized to find another class of perfect graphs for which \(\vartheta(G) > \hat{\vartheta}'(G)\). Let \(n_1, n_2, n_3 \in \mathbb{N}\) such that

\[
G(n_1, n_2, n_3) = K_{n_1} \cup K_{n_2} \cup K_{n_3}
\]

is the union of three complete graphs. This class of graphs is perfect, and as such we get

\[
\vartheta(G(n_1, n_2, n_3)) = \omega(G(n_1, n_2, n_3)) = \max\{n_1, n_2, n_3\}.
\]

Extending this construction, we can define another class of perfect graphs by taking the union of a complete graph on \(n_1\) nodes together with \(m\) isolated nodes. More formally,

\[
G(n_1, e_m) = K_{n_1} \cup \bigcup_{i \in [m]} K_1,
\]

where

\[
\vartheta(G(n_1, e_m)) = \omega(G(n_1, e_m)) = n_1.
\]

Fixing the number of total nodes to 9, the tables in Figure 1 and 2 show the behavior of the various relaxations. While the quality definitely increases while going from \(\hat{\vartheta}(G)\) to \(\hat{\vartheta}'(G)\), the difference between \(\hat{\vartheta}'(G)\) and \(\vartheta(G)\) grows roughly linear with the number of connected components.

We suspect that the bad performance of these relaxations is based on the fact that the relaxations are not monotone in terms of subgraphs. In particular, if \(H\) is a subgraph of \(G\), then the implication

\[
H \leq G \quad \Rightarrow \quad \vartheta(H) \leq \vartheta(G),
\]

is not true for either \(\hat{\vartheta}(G)\) or \(\hat{\vartheta}'(G)\). Since it is not easy to add this property to the functions in question, they are left at an inherent disadvantage compared to the original \(\vartheta\)-function. This begs the question if in general, approximating matrices with binary eigenvalues is harder than approximating matrices with binary entries.
Table 1: Relaxations for $G(n_1, n_2, n_3)$.

| $n_1$ | $n_2$ | $n_3$ | $\hat{\vartheta}$ | $\hat{\vartheta}'$ | $\vartheta = \chi$ |
|------|------|------|----------------|----------------|----------------|
| 3    | 3    | 3    | 3              | 3              | 3              |
| 4    | 3    | 2    | 3.222          | 3.968          | 4              |
| 4    | 4    | 1    | 3.666          | 4              | 4              |
| 5    | 2    | 2    | 3.666          | 4.972          | 5              |
| 5    | 3    | 1    | 3.888          | 4.983          | 5              |
| 6    | 2    | 1    | 4.555          | 5.983          | 6              |
| 7    | 1    | 1    | 5.666          | 6.985          | 7              |

Table 2: Relaxations for $G(n_1, e_m)$.

| $n_1$ | $m$ | $\hat{\vartheta}$ | $\hat{\vartheta}'$ | $\vartheta = \chi$ |
|------|-----|----------------|----------------|----------------|
| 2    | 7   | 1.222          | 1.772          | 2              |
| 3    | 6   | 1.666          | 2.792          | 3              |
| 4    | 5   | 2.333          | 3.851          | 4              |
| 5    | 4   | 3.222          | 4.905          | 5              |
| 6    | 3   | 4.333          | 5.951          | 6              |
| 7    | 2   | 5.666          | 6.986          | 7              |
| 8    | 1   | 7.222          | 8              | 8              |

6 Conclusion

We introduced a new model for the chromatic number $\chi(G)$ based on what we call combinatorial projection matrices and proved its correctness. Relaxing this model yielded an SDP whose optimal value was defined as the projection theta number $\hat{\vartheta}(G)$, which was closely related to the Szegedy number $\vartheta^+(G)$, a variant of the Lovász theta number. We characterized that in general, $\hat{\vartheta}(G) \leq \vartheta^+(G)$, with equality if $G$ is vertex-transitive. Additionally, we gave examples for both equality and strict inequality.

Furthermore, we investigated whether the application of the method of moments could help in bridging this gap and showed with counter examples that in general, this is not the case. These results seem to imply that working with binary matrices is a better paradigm than working with binary eigenvalues in this context. This is to be expected, since matrices with binary eigenvalues can be seen as generalization of binary variables, but it is unexpected in the sense that SDP relaxations are usually adept at approximating eigenvalue-problems.

For future work, it will be interesting to closer inspect the relation between matrices with binary entries and matrices with binary eigenvalues. It would be interesting to know whether one can use one modeling paradigm and convert it to the other, and in particular, if SDP relaxations based on binary matrix-entries are always superior to SDP relaxations based on binary eigenvalues.

References

[1] Mihir Bellare and Madhu Sudan. Improved non-approximability results. In Proceedings of the Twenty-sixth Annual ACM Symposium on Theory of Computing, STOC ’94, pages 184–193. ACM, 1994.

[2] Richard A Brualdi, Seymour V Parter, and Hans Schneider. The diagonal equivalence of a nonnegative matrix to a stochastic matrix. Journal of Mathematical Analysis and Applications, 16(1):31 – 50, 1966.

[3] Isabel Méndez Díaz and Paula Zabala. A polyhedral approach for graph coloring. Electronic Notes in Discrete Mathematics, 7:178 – 181, 2001. Brazilian Symposium on Graphs, Algorithms and Combinatorics.

[4] Igor Dukanovic and Franz Rendl. Semidefinite programming relaxations for graph coloring and maximal clique problems. Math. Program., 109(2):345–365, 2007.

[5] Igor Dukanovic and Franz Rendl. Copositive programming motivated bounds on the stability and the chromatic numbers. Mathematical Programming, 121(2):249–268, 2010.

[6] Yuri Faenza and Volker Kaibel. Extended formulations for packing and partitioning orbitopes. Math. Oper. Res., 34(3):686–697, 2009.

[7] Nebojsa Gvozdenovic and Monique Laurent. The operator psi for the chromatic number of a graph. SIAM Journal on Optimization, 19(2):572–591, 2008.
[8] Jean Bernard Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization*, volume 52. Cambridge University Press, 2015.

[9] László Lovász. On the shannon capacity of a graph. *IEEE Trans. Inf. Theor.*, 25(1):1–7, 1979.

[10] László Lovász. *An algorithmic theory of numbers, graphs and convexity*, volume 50. SIAM, 1987.

[11] F. Rendl. Semidefinite relaxations for integer programming. In Michael Jünger, Thomas M. Liebling, Denis Naddef, George L. Nemhauser, William R. Pulleyblank, Gerhard Reinelt, Giovanni Rinaldi, and Laurence A. Wolsey, editors, *50 Years of Integer Programming 1958-2008 - From the Early Years to the State-of-the-Art*, chapter 18. Springer, 2010.

[12] Alexander Schrijver. A comparison of the Delsarte and Lovász bounds. *Information Theory, IEEE Transactions on*, 25(4):425–429, 1979.

[13] Alexander Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency*. Springer, 2003.

[14] Mario Szegedy. A note on the theta number of Lovász and the generalized delsarte bound. In *35th Annual Symposium on Foundations of Computer Science, Santa Fe, New Mexico, USA, 20-22 November 1994*, pages 36–39, 1994.

[15] Nicolas Trotignon. Perfect graphs: a survey. *arXiv preprint arXiv:1301.5149*, 2015.