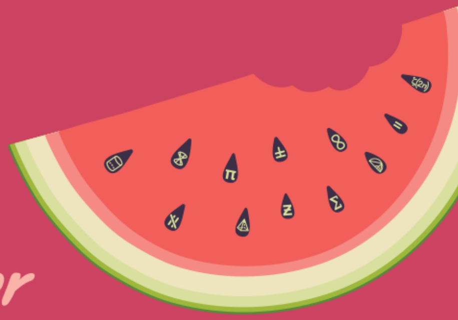
AMSI SUMMER RESEARCH SCHOLARSHIPS 2025–26

Get a taste for Research this Summer



Graph Colouring via Convex Optimisation

Edward Mirco

Supervised by Dr. Hoa Bui and Dr. Sandy Spiers

Curtin University

Jane Street

AMSI Australian Mathematical Sciences Institute

Abstract

This research explores the Graph Colour Problem (GCP) via Semidefinite Programming and projection methods. The GCP is a foundational problem in graph theory of simple description: for a given graph, determine an assignment of 'colours' to its vertices so no adjacent vertices are assigned the same colour, and determine the minimum amount of colours that can be used to do this. Being an NP-complete problem, we formulate several semidefinite approximations, as well as exact approaches using integer semidefinite solvers and projective methods in a non-convex setting, where we use the semidefinite approximation of a solution, then project this solution onto the non-convex space of exact solutions.

Statement of Authorship

- Edward Mirco developed the theory behind this report, the code in Julia, reported and interpreted the results, andwrote this report.
- Dr. Hoa Bui and Dr Sandy Spiers Supervised the work, assisted with the code and the theory behind this report, and proofread this report.

Contents

1 Introduction

1.1 Graph Colouring

Graph colouring (which is used synonymously with proper vertex colouring henceforth) was first introduced in the mid 1800's by Frances Guthrie, and it has since been a source of great theoretical and applied interest in

Jane Street® AMSI

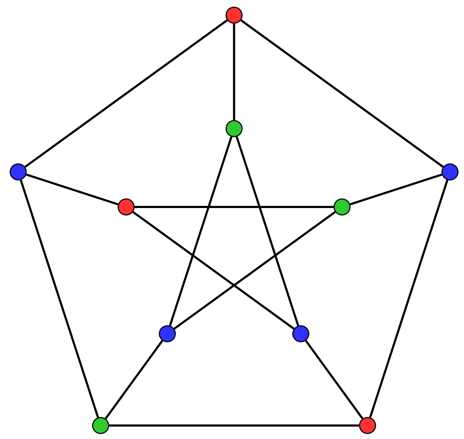


Figure 1: An optimal colouring of the Petersen graph

mathematics, computer science, and an extensive variety of other technical fields (Lewis 2021). A graph consists of a set of vertices, and a set of pairwise edges connecting them, and the graph colouring problem (GCP) is to find an assignment of 'colours' to the vertices such that no two vertices connected by an edge are assigned the same colour. This minimal number of colours for which such a colouring exists for a given graph G is called the chromatic number of the graph, denoted by $\chi(G)$.

The GCP has minimal intrinsic structure, and therefore it is ubiquitous, being an apt formulation for any problem where a minimal partitioning of a collection of objects is to be made while satisfying some pairwise incompatibilities between them. Such tasks applied in practice include scheduling (transport, class timetables, sports fixtures), task allocation (people to tasks, concurrent computing via multithreading) (Ahmed 2012; Zais and Laguna 2016), and resource allocation (frequency allocation of radio towers and WiFi networks, allocation of computational resources in a cloud setup) (Suman 2022). The GCP also arises in many theoretical contexts from quantum contextuality (Cabello, Severini, and Winter 2014) (which is of central importance to quantum computing (Howard et al. 2014)), Sudoku and logic puzzles, and broader mathematics including lie algebras (Bar-Natan 1996). Additionally, the famous four colour theorem; that all planar graphs can be coloured with four or less colours, saw over 160 years as a conjecture and the genesis of powerful computer assisted proof techniques before it was proven in 1976 (Brun 2002). The graph colouring problem is NP-complete, meaning that any NP problem is reducible to a graph colouring problem in polynomial time (Karp 1972), and as such the computation used in all existing approaches becomes restrictively large for graphs beyond a relatively small size; $n \approx 100$. Because of this, in practice if some internal structure/property of the graph to be coloured is known then this can often be exploited to find optimal colourings more efficiently - even in polynomial time (Grtschel, Lovasz, and Schrijver 2012), otherwise heuristic methods are employed to find colourings that are not necessarily optimal; possibly using a number of colour classes $k > \chi$. The most common approach to solving the GCP exactly is to formulate it as a mixed integer linear program (MILP), which declares a set of binary decision variables x_{ik} and y_k, such that $x_{ik} = 1$ iff the vertex labelled i is assigned the colour labelled with

k, and $y_k = 1$ iff the colour labelled k is assigned to any vertex. However, labelling the colours in this way introduces unwanted symmetries, specifically, that colourings that only differ by the labels of each colour are considered truly distinct, which dramatically increases the size of the solutions space, and hence the time to converge on a solution. More recently, due to the advent of efficient methods of solving semidefinite programs, new methods have been devised to solve or approximate the GCP and other similar discrete and combinatoric problems as semidefinite programs, and this has been a fruitful approach that we have followed.

1.2 Semidefinite Programming

A matrix P is said to be positive semidefinite if $x^T P x \geq 0$ for all $x \in \mathbb{R}^n$, or equivalently if all of the eigenvalues of P are real and greater than or equal to 0. Figure 1.2 illustrates that the set of positive semidefinite matrices is a *convex cone*, meaning it is closed under linear combinations with positive coefficients. A number of efficient, polynomial time approaches have been developed to solve semidefinite programming problems: minimisation of a linear function over the positive semidefinite cone subject to linear constraints, thanks to its specific conical structure.

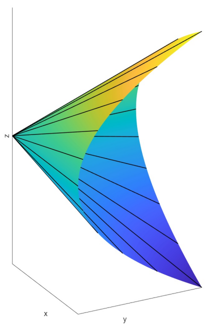


Figure 2: The convex cone of 2×2 symmetric positive semidefinite matrices represented by $\begin{bmatrix} x & z \\ z & y \end{bmatrix}$

Using semidefinite programming to attain approximations of NP-hard combinatorial problems in polynomial time has gained popularity; specifically for graph colouring the Lovász (Lovasz 1979) number $\vartheta(G)$ of a graph G can be found using semidefinite programming and provides a lower bound for the chromatic number that is often tight. In fact, the Lovász "sandwich theorem" proves that $\vartheta(\bar{G})$ (where $\bar{G}$ is the complement graph of G, swapping edges and non-edges between vertices) lies between the clique number $\omega(G)$; the size of the largest subgraph of G in which every vertex is adjacent to every other vertex, and the chromatic number $\chi(G)$: $\omega(G) \leq \vartheta(\bar{G}) \leq \chi(G)$ (Knuth 1993), both of which are NP-complete to compute. In fact, it is NP-hard to decide whether $\chi(G) - \omega(G) > 0$, even when $\chi(G)$ is known Dukanovic and Rendl 2007 and consequently unless P=NP, there is no polynomial time upper bound on $\omega(G)$ provably better than $\vartheta(\bar{G})$. $\vartheta(\bar{G})$ can be formulated to be seen as a relaxation of the constraints of an exact optimisation problem to find $\chi(G)$, essentially just removing the constraint that all nonzero entries of the semidefinite matrix it seeks to find are equal, and it is in this way that our formulations of the GCP are derived, with exact solvers being thought of as recovering an exact solution from the approximation.

2 Formulation

There are multiple variations of the GCP; here we study the minimisation variant, which is to say the objective of our methods will be to take an arbitrary simple graph as an input and return a colouring of that graph using the least number of colours possible.

Definition 1. *For a simple undirected graph $G = (V, E)$, an equivalence relation (that is, a binary relation with the properties of reflexivity, symmetry, and transitivity) on the set of vertices $(V, \sim)$ is called a* proper vertex colouring *or* colouring *of G if $\forall \{i, j\} \in E; i \nsim j$. We call the equivalence classes of colourings* colour classes, *where we denote the colour class which a vertex i belongs to with $[i]$, and if there are exactly k distinct colour classes in a colouring then we call it a k*-colouring *of G.*

Definition 2. *For a simple undirected graph $G = (V, E)$, we call the minimum k such that there exists a k-colouring of G the* chromatic number *of G, denoted by $\chi(G)$.*

2.1 Binary Colouring Matrix

Definition 3. *For a simple undirected graph $G = (V = \{1, ..., n\}, E)$ with n vertices and a proper vertex colouring of it, we call the $n \times n$ matrix with binary entries that encodes the colouring the* binary colouring matrix, *denoted by P. That is;*

$$\forall i, j \in V; \quad P_{ij} = \begin{cases} 1 & \text{if } i \sim j \\ 0 & \text{otherwise.} \end{cases}$$

A binary colouring matrix encoding a k-colouring of G is called a binary k-colouring matrix.

For a more clear intuition, we can observe that a binary colouring matrix is simply a permuted block matrix, with each block representing a colour class. For a particular k-colouring of a graph, its vertices may be relabelled such that they are partitioned according to colour classes. We arbitrarily enumerate the k colour classes, and n_q denotes the number of vertices belonging to the q^{th} colour class, then we can label the vertices belonging to the first colour class with the set of integers $\{1, ..., n_1\}$, the vertices belonging to the second colour class with $\{n_1 + 1, ..., n_1 + n_2\}$ etc. A permutation matrix Q_σ corresponding to such a permutation σ of the set of vertex labels can transform the binary colouring matrix as $P' = Q_\sigma P Q_\sigma^T$, such that both the columns and rows of P' are permuted to have the new vertex ordering. Let J_n be the $n \times n$ matrix whose elements are all 1. P' is then a block matrix, consisting of J_{n_q} blocks along its main diagonal, and therefore its rank is equal to the number of blocks, or equivalently the number of colour classes.

$$P' = Q_\sigma P Q_\sigma = \begin{bmatrix} J_{n_1} & & & \\ & J_{n_2} & & \\ & & \ddots & \\ & & & J_{n_k} \end{bmatrix} : \sum_{q=1}^{k} n_q = n$$

This makes it clear to see that the i^{th} column sum (and the row sum due to symmetry) is equal to $|[i]|$; the cardinality of the colour class to which the vertex labelled i belongs, and this is true for any binary colouring matrix. This transformation is useful later in deriving an objective function to minimise the number of colours used, as well as finding another formulation of the problem.

Lemma 1. *A binary colouring matrix P encoding a k-colouring is symmetric positive semidefinite, and has rank k.*

Proof. Let $G = (V, E)$, where $|V| = n$, and let P be an $n \times n$ binary colouring matrix encoding a k-colouring of G. It is trivial that $P_{ij} = P_{ji}$ as $i \sim j$ if and only if $j \sim i$ by the symmetric property of the equivalence relation. It is sufficient to show there exists a matrix X of rank k such that $P = XX^T$ to prove that P is positive semidefinite and has rank k. We can arbitrarily enumerate the colour classes with the integers from 1 to k, and construct the $n \times k$ matrix denoted X, where X_{iq} is 1 if the vertex labelled i belongs to the colour class labelled q, and is 0 otherwise. We evaluate the entries of XX^T as

$$(XX^T)_{ij} = \sum_{q=1}^{k} X_{iq} X_{jq}$$

For each $i \in V$, $X_{iq} = 1$ for exactly one $q \in \{1, ..., k\}$, so the sum over q is 1 if and only if there is a value of q such that $X_{iq} = X_{jq} = 1$, which is to say i and j both belong to the colour class labelled q, and 0 otherwise. Therefore, $(XX^T)_{ij}$ is 1 if $i \sim j$ and 0 otherwise, hence $P = XX^T$. Finally, X's column vectors are all linearly independent, as each colour class is occupied, so each column has non-zero entries, and since each vertex belongs to exactly one colour class, an entry of 1 for a given row and column implies that the rest of the entries in that row are 0, thus X and P have a rank equal to k. □

The adjacency matrix A of a graph G is defined such that $A_{ij} = 1$ if i and j are adjacent and 0 otherwise, and this allows for a very simple way to ensure that a colouring matrix does not encode a colouring in which $i \sim j$ if i is adjacent to j.

Lemma 2. *A binary colouring matrix P encoding a colouring of a graph $G = (V, E)$ with associated adjacency matrix A satisfies the Frobenius inner product equality $\langle P, A \rangle = 0$*

Proof. As both A and P are real matrices, the inner product $\langle P, A \rangle$ is equal to

$$\langle P, A \rangle = \sum_{i,j} P_{ij} A_{ij}$$

Recall the adjacency matrix of a graph $G = (V, E)$ is defined entry-wise such that A_{ij} is equal to 1 iff $\{i, j\} \in E$, and 0 otherwise. By the definition of proper vertex colourings, $\forall \{i, j\} \in E; i \nsim j$, thus $A_{ij} = 1$ implies $P_{ij} = 0$.

Thus the sum of all the terms where $A_{ij} = 1$ is 0, and since the rest of the terms have $A_{ij} = 0$, the total sum must be 0. □

Given that binary colouring matrices are both positive semidefinite and have binary entries, we consider an alternative characterisation of such matrices defined by linear inequalities, given by Proposition 2 of Letchford and Sørensen 2012.

Lemma 3 (Letchford and Sørensen 2012)**.** *A symmetric binary matrix $M \in \{0,1\}^{n\times n}$, with $n \geq 3$, is positive semidefinite only if it satisfies the following inequalities:*

$$M_{ij} \leq M_{ii} \quad (1 \leq i < j \leq n)$$
$$M_{ik} + M_{jk} \leq M_{kk} + M_{ij}$$

Proof. A symmetric binary matrix is positive semidefinite if and only if it is the sum of one or more symmetric rank one binary matrices (Proposition 1 Letchford and Sørensen 2012), so it is sufficient to prove that these inequalities hold for symmetric rank one binary matrices. Symmetric rank one binary matrices are those matrices that can be written as vv^T for some $v \in \{0,1\}^n$). $(vv^T)_{ij} = v_i v_j \leq v_i v_i$ so the first inequality is satisfied by these rank 1 matrices. We can rewrite the second inequality as $v_i v_k + v_j v_k \leq v_k v_k + v_i v_j$: if $v_k = 1$ then we get $v_i + v_j \leq 1 + v_i v_j$ which is satisfied by all $v_i, v_j \in \{0,1\}$, and if $v_k = 0$ then $0 \leq v_i v_j$, which is clearly true. □

We now provide a full characterisation of binary colouring matrices as a positive semidefinite matrix with binary entries satisfying certain linear constraints.

Theorem 1. *A matrix P is a binary colouring matrix of a graph $G = (V, E)$ with adjacency matrix A if and only if all of the following constraints are met:*

(i) $\forall i, j \in V; P_{ij} \in \{0,1\}$

(ii) $\langle P, A\rangle = 0$

(iii) $tr(P) = |V|$

(iv) $P \in S_n^+$

Proof. From the definition of binary colouring matrices P is a binary colouring matrix only if (i), and only if $P_{ii} = 1$ for all i, which (iii) implies in conjunction with (i), as the trace of P reaches its maximum value of n if and only if all of the diagonal entries take their maximum values of 1. We have already shown that P is a binary colouring matrix only if (ii) and (iv) are satisfied in Lemmas 2 and 1 respectively, so all that is left is to prove is the reverse direction.

Conditions (i) and (iv) imply that P is a logical matrix representing a binary relation $\sim$ on the set of integers from 1 to n (the set of vertices V of the graph G). As (i) and (iii) together imply $P_{ii} = 1$ for all i, then $i \sim i$ for all i, so $\sim$ is reflexive, and since (iv) implies $P_{ij} = P_{ji}$ for all i, j then $i \sim j$ iff $j \sim i$, so $\sim$ is symmetric. To prove that the binary relation $\sim$ is transitive and hence an equivalence relation we must show that $i \sim k$

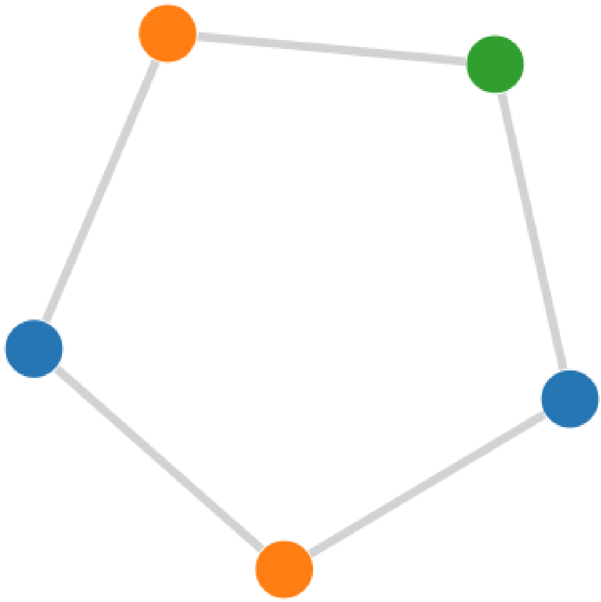


Figure 3: Optimally coloured Pentagon graph/5 cycle graph

and $j \sim k$ implies $i \sim j$. Referring to Lemma 3 and replacing M with P, (iv) implies that $P_{ik} + P_{jk} \leq 1 + P_{ij}$ (given that $P_{ii} = 1$ for all i, which we already proven is implied by the constraints). Thus, $P_{ik} = 1$ and $P_{jk} = 1$ implies $P_{ij} = 1$, or equivalently $i \sim k$ and $j \sim k$ implies $i \sim j$, therefore $\sim$ is an equivalence relation. Condition (ii) can be restated with the following sum:

$$\langle P, A \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} A_{ij} = 0$$

from which it is clear that $A_{ij} = 1$ implies $P_{ij} = 0$, so by the definition of the adjacency matrix $\{i, j\} \in E$ implies $i \nsim j$, and thus the equivalence relation is a colouring of G. □

While this representation of a a graph colouring is similar to that used in the formulation of the Lovász number ϑ, we propose a novel convex function that whose minimised over the set defined by the constraints of Theorem 1 is the chromatic number χ of the graph.

2.1.1 Convex rank approximation

Lemma 4. *The rank of a binary colouring matrix P is equal to the following expression:*

$$rank(P) = \sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{n} P_{ij}} \tag{1}$$

Proof. For a given graph $G = (V, E)$ and a colouring of its vertices represented by the binary colouring matrix P, there exists a permutation matrix Q_σ, where σ denotes a permutation of the vertex labels, such that $P' = Q_\sigma P Q_\sigma^T$ is a block matrix consisting of J_{n_q} matrices along its main diagonal (where n_q is the cardinality of the q^{th} colour class for a chosen enumeration of the colour classes). Note, that since this transformation is only a permutation of the vertex labels, the column and row sums of the relabelled vertices are equal to the column and row sums of the original labelling. The rank is equal to the number of blocks in P', so to count each block as 1 we sum over the multiplicative inverses of the row sums. For each vertex belonging to the colour class labelled q, the inverse of its row sum is $\frac{1}{n_q}$, and since there are n_q of these identical rows in P' (all contained in the one J_{n_q} block) we count this value n_q times for all n_q vertices in the colour class. Each colour class then contributes a total value of 1 to the sum, and thus the total sum is equal to the number of colour classes, which is equal to the rank. □

Definition 4. *Let D_n be the set $\{M \in [0,1]^{n\times n} | tr(M) = n\}$. We define φ to be the function whose domain and codomain are D_n and $\mathbb{R}$ respectively such that for all $M \in D_n$:*

$$\varphi(M) = \sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{n} M_{ij}} \tag{2}$$

Lemma 5. *φ is convex over its domain.*

Proof. If f and g are both convex functions, then both $f+g$ and $f \circ g$ are also convex. Thus it is sufficient that each term of the sum over i is convex. The reciprocal function is convex over the strictly positive real numbers, and since the domain of φ enforces that all elements are non-negative and that at least one of the elements of each column is equal to 1, then the reciprocal function is indeed convex. Each term of the sum over i is the composition of a strictly positive summation, which is convex, and the reciprocal function in the domain in which it is convex, thus φ is convex. □

We can now present the following optimisation problem, whose minimum value with respect to the constraints is the chromatic number χ of the graph with adjacency matrix A:

$$\begin{aligned} \min_{P \in \mathbb{R}^{n\times n}} \quad & \varphi(P) = \sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{n} P} \\ \text{s.t.} \quad & \forall i,j; P_{ij} \in \{0,1\} \\ & \forall i,j; P_{ij} = P_{ji} \\ & \langle A, P \rangle = 0 \\ & \text{tr}(P) = n \\ & P \succeq 0 \end{aligned} \tag{3}$$

The permuted binary colouring matrix $P' = Q_\sigma^T P Q_\sigma$ as constructed earlier has its column vectors being eigenvectors of the matrix itself. The block matrices along the diagonal make this easier to see, and since the i^{th} column/row sum is equal to $|[i]|$ for the vertices as originally labelled, then it is clear that $P(Pe_i) = |[i]|Pe_i$, as the block matrix to which the column vector is associated with counts once for each vertex in the colour class for each entry.

2.2 Normalised Colouring Matrix

Definition 5. *For a binary colouring matrix P of a graph G, let D be the diagonal matrix whose ith diagonal element is the ith column sum of P (or equivalently, the ith diagonal element of D is equal to $|[i]|$; the cardinality of the colour class to which the vertex labelled i belongs). Then, the matrix PD^{-1} is called a* normalised colouring matrix *of G that encodes the same colouring as P, denoted by $\hat{P}$.*

Remark 1. *Given that we have constructed a normalised colouring matrix $\hat{P}$ by multiplying a binary colouring matrix P by a necessarily full rank matrix D^{-1}, the rank of $\hat{P}$ is equal to the rank of P. We may then call a normalised colouring matrix of rank k a* normalised k-colouring matrix.

Lemma 6. *Let $\hat{P}$ be a symmetric normalised colouring matrix. Then $\hat{P}$ is idempotent: $\hat{P}^2 = \hat{P}$*

Proof. $\hat{P} = PD^{-1}$, where $D_{jj} = \sum_{k=1}^{n} P_{kj}$. $\hat{P}$, P and D are symmetric, so $\hat{P} = PD^{-1} = \hat{P}^T = (PD^{-1})^T = D^{-1}P$. By considering the permutation of the indices such that P is a block diagonal matrix composed of blocks of J matrices, it is clear that the columns (and rows due to symmetry) of P are its own eigenvectors as : $P(Pe_j) = \lambda_j Pe_j$ where λ_j is the j^{th} column sum, which is equal to D_{jj}, and thus $P^2 = PD$. So we can now evaluate $\hat{P}^2$ as the following:

$$\hat{P}^2 = (PD^{-1})(PD^{-1}) = D^{-1}P^2D^{-1} = D^{-1}PDD^{-1} = D^{-1}P = PD^{-1} = \hat{P}$$

□

Theorem 2. *A matrix $\hat{P}$ is a normalised colouring matrix of the graph with adjacency matrix A if and only if all of the following constraints are met:*

(i) $\forall i, j \in V; \hat{P}_{ij} \geq 0$

(ii) $\langle A, \hat{P} \rangle = 0$

(iii) $\hat{P}^T = \hat{P}$

(iv) $\forall j \in V; \sum_{i \in V} \hat{P}_{ij} = 1$

(v) $\hat{P}^2 = \hat{P}$

Proof. We prove the 'only if' part by constructing a normalised colouring matrix from a binary colouring matrix as per Definition 5, and showing it satisfies these constraints. Let P be a binary colouring matrix encoding a colouring of the graph with adjacency matrix A, let D be the diagonal matrix with $D_{jj} = \sum_{k=1}^{n} P_{kj}$, and let $\hat{P}$ be the normalised colouring matrix constructed from P as $\hat{P} = PD^{-1}$. $P_{ij} \in \{0, 1\}$ for all i, j, and clearly scaling P with a diagonal matrix with positive diagonal entries cannot make any of the entries negative, so (i) is satisfied. Indeed, $\hat{P}_{ij} > 0$ if and only if $P_{ij} > 0$, and since $\langle A, P \rangle = 0$, then also $\langle A, \hat{P} \rangle = 0$, so constraint (ii). Permuting the indices of P to be in diagonal block matrix form, we scale each column vector of P by the inverse of the sum of its elements, which is equal for each column of the block, so the block is scaled uniformly, and thus the symmetry of P implies $\hat{P}$ satisfies constraint (iii). The columns of P are scaled by the inverse of the column sum, and so clearly the column sum of $\hat{P}$ is 1, and thus P satisfies constraint (iv), and Lemma **??** shows that $\hat{P}$ satisfies constraint (v).

We now prove the 'if' part by constructing a binary colouring matrix from an arbitrary matrix satisfying these constraints. Let M be a matrix that satisfies (i), (iii), (iv) and (v). From the constraint (v):

$$(M^2)_{ii} = \sum_{k=1}^{n} M_{ik}M_{ki} = \sum_{k=1}^{n}(M_{ik})^2 = M_{ii}$$

Therefore $(M_{ij})^2 \leq M_{ii}$, so $M_{ij} > 0$ implies $M_{ii} > 0$ and $M_{jj} > 0$. Now, we define a binary relation $(\{1, ..., n\}, \sim)$ where $i \sim j$ if there exists a path from i to j ($i = i_0, i_1, i_2, ..., i_k = j$) such that for all $l < k$ $M_{i_i i_{l+1}} > 0$. This binary relation is symmetric, as if there exists from i to j then there exists a path from j to i by reversing its order, and the relation is also reflexive: there is the path from i to i consisting of only itself, as $M_{ii} > 0$ for all i (due to non-negativity, column sum unity, and diagonal dominance: $M_{ij} \leq M_{ii}$). The relation is also transitive: if $i \sim j$ and $j \sim k$, then we can concatenate the path from j to k to the path from i o j to create a path from i to k, so $i \sim k$. Thus $\sim$ defines an equivalence relation, and we so we label the equivalence classes $C_1, C_2, ..., C_r$, and we can permute the indices of M such that the equivalence classes appear consecutively. Thus, we can write M as a block diagonal matrix:

$$M = \begin{bmatrix} M^{C_1} & 0 & \cdots \\ 0 & M^{C_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

where each M^{C_i} block satisfies (i), (iii), (iv) and (v), and is irreducible (the graph of the matrix is strongly connected). If $\mathbf{1}$ is the vector who entries are all 1, then from condition (iv) we find that $\mathbf{1}$ is an eigenvector of M and of M^{C_i} with an eigenvalue of 1:

$$(M\mathbf{1})_j = \sum_{k=1}^{n} M_{jk}\mathbf{1}_k = \sum_{k=1}^{n} M_{jk} = 1 \to M\mathbf{1} = \mathbf{1}$$

It is known that the eigenvalues of any idempotent matrix (one such that $M^2 = M$) can only be equal to 0 or 1. The Perron–Frobenius theorem states that a nonnegative irreducible matrix has a unique eigenvector whose entries are all positive such that its eigenvalue has a multiplicity of 1, and has an absolute value strictly greater than the rest of the eigenvalues. Given that the eigenvalues of M^{C_i} are less than or equal to 1 due to it being idempotent, $\nVdash$ must be that unique eigenvector, and given the multiplicity of its eigenvalue being 1, all of the other eigenvalues must be equal to 0, therefore $\text{rank}(M^{C_i}) = 1$ and $M^{C_i} = u_{C_i}u_{C_i}^T$ for some $u_{C_i} \in \mathbb{R}^{|C_i|}$. Using $M\mathbf{1} = \mathbf{1}$, $(u_{C_i}u_{C_i}^T)\mathbf{1} = u_{C_i}(u_{C_i}^T\mathbf{1}) = \mathbf{1}$ thus u must be proportional to $\mathbf{1}$. Normalising u_{C_i}, we get $u_{C_i} = \frac{1}{|C_i|}\mathbf{1}$, and so we can rewrite M as:

$$M = \begin{bmatrix} \frac{1}{n_1}\mathbf{1}\mathbf{1}^T & 0 & \cdots \\ 0 & \frac{1}{n_2}\mathbf{1}\mathbf{1}^T & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

□

where n_i is the size of the i^{th} diagonal block. Now we can construct a binary colouring matrix P by multiplying M by a diagonal matrix that scales each diagonal entry to be equal to 1, and hence scales each non-zero element of M to 1, and thus $P_{ij} \in \{0,1\}$ for all i, j; satisfying constraint (i) of Theorem 1. Since every diagonal element of P is equal to 1, then clearly $tr(P) = n$ satisfying constraint (iii) of Theorem 2, and since M has binary eigenvalues, scaling its columns by positive factors preserves the non-negativity of the eigenvalues, and thus P is positive semidefinite, satisfying constraint (iv) of Theorem 2. Finally, note that $M_{ij} > 0$ if and only if $P_{ij} > 0$, and so since $\langle A, M\rangle = 0$ by constraint (ii) which means $A_{ij} = 1$ implies $M_{ij} = 0$, then indeed $A_{ij} = 1$ implies $P_{ij} = 0$, and thus $\langle A, P\rangle = 0$, satisfying constraint (ii) of Theorem 2.

Remark 2. *A normalised colouring matrix of a graph G, being an idempotent matrix with all nonzero eigenvalues equal to 1, has its rank equal to its trace (the sum of its eigenvalues is equal to the number of nonzero eigenvalues counting multiplicities, which is the rank), so to minimise the amount of colours used to colour G, we may find the matrix in the set of matrices satisfying Theorem 2 to be a normalised colouring matrix with minimal trace rather than minimising the rank directly.*

With this in mind, we can state the following semidefinite program in its optimisation form, that when the function is minimised with respect to the constraints achieves the chromatic number of the graph with adjacency matrix A:

$$\begin{aligned} \min_{\hat{P}} \quad & \operatorname{tr}(\hat{P}) \\ \text{s.t.} \quad & \forall i, j; \hat{P}_{ij} \geq 0 \\ & \langle A, \hat{P}\rangle = 0 \\ & \hat{P}^T = \hat{P} \\ & \forall j; \sum_i \hat{P}_{ij} = 1 \\ & \hat{P}^2 = \hat{P} \end{aligned} \tag{4}$$

and we can consider the relaxed variant where we replace the idempotent constraint (or equivalently that all of the eigenvalues are in the set $\{0,1\}$) with $\hat{P} \succeq 0$, as the column sum constraint ensures that the absolute value of the eigenvalues are less than or equal to 1, and so positive semidefiniteness ensures that the all the eigenvalues belong to the open interval $[0,1]$. In fact, it has been proven that the set of matrices with eigenvalues in the interval $[0,1]$ is the convex hull of the space of idempotent matrices, and furthermore, the set of idempotent matrices is exactly the set of extreme points of this convex hull (Overton and Womersley 1992). Thus, the matrix which attains the minimum value of a convex function over the set of matrices with eigenvalues in the interval $[0,1]$ that does not contain any local minima in the interior of the set will be a projection matrix. Thus, the matrix with minimal trace in the set of matrices with eigenvalues in $[0,1]$ will be a projection matrix. The additional constraints of normalised colouring matrices interfere with this, and thus the matrix which attains the minimal trace with the linear relaxation of the eigenvalue condition may reside on the facets of the linear constraints, and not be a projection matrix, however in general we observe that the minimal trace matrix will tend toward having binary eigenvalues (what that means specifically will be explained later). This gives justification towards this particular relaxation of the constraints. Upon solving this relaxed variant, we might consider how we might project this approximation onto the non-convex space of exact solutions.

2.2.1 Projections

A first guess for how we might project a matrix onto the set of matrices with eigenvalues in $\{0,1\}$ would be to round the eigenvalues while keeping the respective eigenvectors the same: send eigenvalues less that 0.5 to 0 and those above 0.5 to 1, making some arbitrary choice for eigenvalues of exactly 0.5, and this turns out to be a valid projection (or rather pseudoprojection due to non-convexity and thus non-uniqueness of the projections), meaning the distance from the initial matrix to this matrix on the set of idempotent matrices is less than the distance to any other idempotent matrix.

Lemma 7. *Let M be real symmetric matrix with eigendecomposition $M = Q\Lambda Q^T$ for some orthogonal matrix Q whose column vectors are Ms eigenvectors, and some diagonal matrix Λ whose diagonal entries are the respective eigenvalues, which all belong to the set $[0,1]$. Then, $\pi(M) = QBQ^T$, where B is the diagonal matrix whose i^{th} diagonal entry is 0 if $\Lambda_{ii} \leq 0.5$ and equal to 1 otherwise is a valid pseudoprojection: $\|M - \pi(M)\| \leq \|M - P\|$ for all idempotent P.*

Proof. We use the Hoffmman-Wielandt inequality: for $A, B \in \mathbb{C}^{n\times n}$ that are normal and have eigenvalues $\lambda_1(A), ..., \lambda_n(A)$ and $\lambda_1(B), ..., \lambda_n(B)$ respectively, then denoting the permutation group of $\{1, ..., n\}$ with $\mathcal{S}_n$ the following inequality holds

$$\min_{\sigma \in S_n} \sum_{i=1}^{n} |\lambda_i(A) - \lambda_{\sigma(i)}(B)|^2 \leq \|A - B\|^2$$

which is to say the sum of the square of the difference of the optimally matched eigenvalues of A and B is less than $\|A - B\|^2$. We let $A = M$, and we consider $B \in \{P \in \mathbb{R}^{n\times n} | P^2 = P\}$, so that $\lambda_i(B) \in \{0,1\}$ for all i. As each term of the sum is clearly minimised when $\lambda_i(B)$ is $\lambda_i(A)$ rounded to thee nearest integer, then we have the following chain of inequalities:

$$\sum_{i=1}^{n} |\lambda_i(M) - \lambda_i(\pi(M))|^2 \leq \min_{\sigma \in S_n} \sum_{i=1}^{n} |\lambda_i(M) - \lambda_{\sigma(i)}(B)|^2 \leq \|M - B\|^2 \quad \forall B \in \{P \in \mathbb{R}^{n\times n} | P^2 = P\}$$

We now evaluate $\|M - \pi(M)\|$ via the eigendecompositions, using the fact that the matrix norm is invariant upon multiplication by a unitary (or orthogonal) matrix

$$\|M - \pi(M)\|^2 = \left\|Q\Lambda Q^T - QBQ^T\right\|^2 = \left\|Q(\Lambda - B)Q^T\right\|^2 = \|\Lambda - B\|^2 = \sum_{i=1}^{n} (\Lambda_{ii} - B_{ii})^2 = \sum_{i=1}^{n} (\lambda_i(M) - \lambda_i(\pi(M)))^2$$

$$\to \|M - \pi(M)\|^2 = \sum_{i=1}^{n} |\lambda_i(M) - \lambda_i(\pi(M))|^2 \leq \min_{\sigma \in S_n} \sum_{i=1}^{n} |\lambda_i(M) - \lambda_{\sigma(i)}(B)|^2 \leq \|M - B\|^2$$

$$\to \|M - \pi(M)\| \leq \|M - B\| \quad \forall B \in \{P \in \mathbb{R}^{n\times n} | P^2 = P\}$$

□

With this done, we can use an iterative projective method like the Douglas-Rachford algoorithm to find a solution that lies in the intersection of the set of matrices which satisfy conditions (i) - (iv) and the set of idempotent matrices. While Douglas-Rachford has been used to surprising success in a number of nonconvex, discrete cases to solve combinatorial problems including the GCP (Lindstrom and Sims 2020), it has no guarantee of convergence in the nonconvex case, and in the case it does converge there is no guarantee of the optimality of the solution, so it may only be used as a heuristic.

3 Numerical Analysis

We used *Gurobi* as a mixed integer linear program (MILP) solver to calculate the exact chromatic number of the graphs we tested with the classical MILP GCP formulation, and as a benchmark for the performance of our semidefinite approximations and heuristic solutions. We used the semidefinite proogram solver from *Mosek* to solve the two semidefinite programs stated previously for the binary and normalised formulations with the respective convex relaxations, and we used *Pajarito* to solve the exact binary colouring semidefinite program with integral constraints. We used the *Lovasztheta.jl* package in *Julia* to solve find $\vartheta(\bar{G})$, and all of the other solvers were implemented in the *Julia* programming language. We applied the solvers to graphs with a varying number of vertices, with constant edge density: the number of edges was $\frac{n(n-1)}{3}$ rounded down to the nearest integer, so the edge density was 2/3, which is a critical density for the complexity of calculating the chromatic number of a random graph. We found that the semidefinite approximations were solved exceptionally quickly, especially in relation to the MILP solver, which grew to runtimes of over an hour for some graphs of only 50 vertices. The relaxed semidefinite approximations could be solved in under 10 seconds consistently for graphs of less than 120 vertices, and was only marginally longer for large graphs. As shown in Figure 4, the Lovász

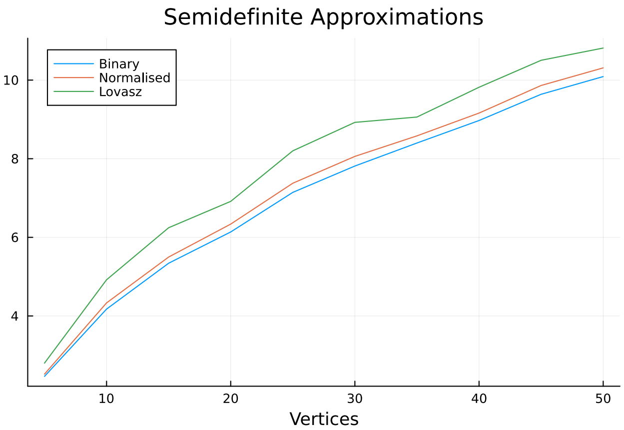


Figure 4: The value of the chromatic number found by the semidefinite approximation schemes. Averaged over 5 random graphs of the specified number of vertices.

number (of the graphs complement) was a tighter lower bound to the actual chromatic number of the graph than either of the formulations we presented, however it did take substantially longer to calculate $\vartheta(\bar{G})$: for instance, for random graphs of 50 vertices (with 2/3 edge density), it took an average of 11.3 seconds to calculate $\vartheta(\bar{G})$, however the maximum time taken by the binary and normalised formulations were 7.4 and 7.9 seconds respectively. The integer semidefinite solver succeeded in finding optimal colourings, but took prohibitively long for more than even just 15 vertices. For the Douglas-Rachford projection heuristic approach, it failed to

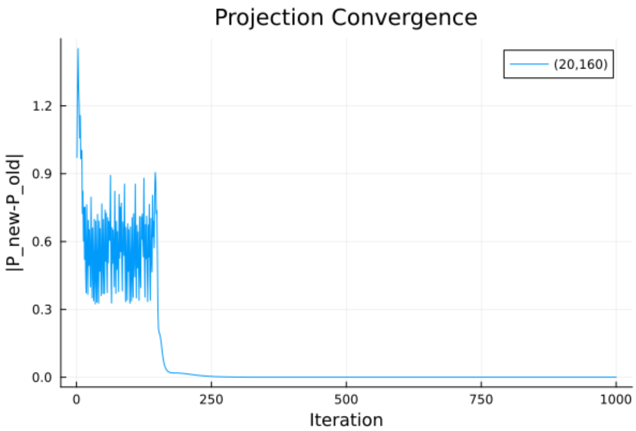


Figure 5: The difference of the norm between iterations in the Douglas-Rachford projection scheme for a random graph with 20 vertices and 160 edges

converge in most cases. Notably, it failed to converge for the Petersen graph and the pentagon graph (the 5 cycle graph). However, in the cases it did converge we identified that the graphs either had a high degree symmetry making the colourings essentially trivial (e.g. cycle graphs with a even number of vertices, complete graphs, tree graphs), or were sufficiently close to these highly symmetric graphs. A notable example is shown in Figure 4, that the convergence to a solution for a graph with 20 vertices and 160 edges was not as immediate as the more symmetric graphs, but was sufficiently close to the graph in which every vertex is adjacent to every other vertex (K_{20}, which has 190 edges) that the pseudoprojection was able to converge on the solution that the relaxed solver was close to, which indeed was optimal using 10 colours, which the MILP solver confirmed was the chromatic number. For every graph we tested in which the iterative projections converged, it converged to a normalised colouring matrix encoding an optimal colouring, and always in less than 500 iterations.

4 Discussion and Conclusions

In this report we presented two novel formulations of the graph colouring problem in the form of semidefinite programs, one of them being an equivalent semidefinite representation of colouring with an alternative convex function to be minimised to the Lovász number, and the other being a novel representation of colourings of a graph. We believe the normalised colouring matrix provides an elegant characterisation of graph colourings, being that we have replaced integral constraints that grow as n^2 with the idempotency constraint, which may be thought of as n integral constraints on the eigenvalues, as well as it having the geometric interpretation as the orthogonal projection matrix that projects the basis vectors representing the vertices from n dimensional space to k dimensional space such that vertices of the some colour class are projected to the same point. Although preliminary results show that the lower bounds established by these approximations are inferior to the Lovász number (which was expected, due to the seemingly fundamental role of the Lovász number as a ”point of closest approach” for polynomial time approximations of the clique number and chromatic number), we believe there is unexplored potential in the normalised colouring matrix formulation. For example, given that the rank of a normalised colouring matrix is not only equal to the trace but also equal to the square of the Frobenius norm, these two quantities being equal if the matrix is a normalised colouring matrix, a splitting algorithm such as ADMM may be used to both minimise the trace and difference between the trace and the square of the norm, thus bringing the solution found by the convex solver closer to the set of exact solutions.

References

Ahmed, Shamim. 2012. Applications of graph coloring in modern computer science. *International Journal of Computer and Information Technology* 3 (2): 1–7.

Bar-Natan, Dror. 1996. *Lie algebras and the four color theorem.* arXiv: q-alg/9606016 [q-alg]. https://arxiv.org/abs/q-alg/9606016.

Brun, Yuriy. 2002. The four-color theorem. *Undergraduate Journal of Mathematics,* 21–28.

Cabello, Adán, Simone Severini, and Andreas Winter. 2014. Graph-theoretic approach to quantum correlations. *Physical Review Letters* 112, no. 4 (January). ISSN: 1079-7114. https://doi.org/10.1103/physrevlett.112.040401. http://dx.doi.org/10.1103/PhysRevLett.112.040401.

Dukanovic, Igor, and Franz Rendl. 2007. Semidefinite programming relaxations for graph coloring and maximal clique problems. *Mathematical Programming* 109 (2): 345–365. ISSN: 1436-4646. https://doi.org/10.1007/s10107-006-0026-z. https://doi.org/10.1007/s10107-006-0026-z.

Grtschel, M., L. Lovasz, and A. Schrijver. 2012. *Geometric algorithms and combinatorial optimization.* Springer Berlin Heidelberg. Book. ISBN: 9783642978814 , lccn= 87036923 , series= Algorithms and Combinatorics. https://books.google.com.au/books?id=hWvmCAAAQBAJ.

Howard, Mark, Joel Wallman, Victor Veitch, and Joseph Emerson. 2014. Contextuality supplies the 'magic' for quantum computation. *Nature* 510 (7505): 351–355. ISSN: 1476-4687. https://doi.org/10.1038/nature13460. https://doi.org/10.1038/nature13460.

Karp, Richard M. 1972. Reducibility among combinatorial problems. In *Complexity of computer computations: proceedings of a symposium on the complexity of computer computations, held march 20–22, 1972, at the ibm thomas j. watson research center, yorktown heights, new york, and sponsored by the office of naval research, mathematics program, ibm world trade corporation, and the ibm research mathematical sciences department,* edited by Raymond E. Miller, James W. Thatcher, and Jean D. Bohlinger, 85–103. Boston, MA: Springer US. ISBN: 978-1-4684-2001-2. https://doi.org/10.1007/978-1-4684-2001-2_9. https://doi.org/10.1007/978-1-4684-2001-2_9.

Knuth, Donald E. 1993. *The sandwich theorem.* arXiv: math/9312214 [math.CO]. https://arxiv.org/abs/math/9312214.

Letchford, Adam N., and Michael M. Sørensen. 2012. Binary positive semidefinite matrices and associated integer polytopes. *Mathematical Programming* 131 (1): 253–271. ISSN: 1436-4646. https://doi.org/10.1007/s10107-010-0352-z. https://doi.org/10.1007/s10107-010-0352-z.

Lewis, Rhyd. 2021. *Guide to graph colouring.* Springer. Book. ISBN: 3030810534.

Lindstrom, Scott B., and Brailey Sims. 2020. *Survey: sixty years of douglas–rachford.* arXiv: 1809.07181 [math.OC]. https://arxiv.org/abs/1809.07181.

Lovasz, L. 1979. On the shannon capacity of a graph. *IEEE Transactions on Information Theory* 25 (1): 1–7. https://doi.org/10.1109/TIT.1979.1055985.

Overton, Michael L., and Robert S. Womersley. 1992. On the sum of the largest eigenvalues of a symmetric matrix. *SIAM Journal on Matrix Analysis and Applications* 13 (1): 41–45. https://doi.org/10.1137/0613006. https://doi.org/10.1137/0613006, eprint = https://doi.org/10.1137/0613006, abstract = Thesumofthelargestkeigenvaluesofasymmetricmatrixhasawell-knownextremalpropertythatwasgivenbyFan in1949[Proc.Nat.Acad.Sci.,35(1949),pp.652655].Asimpleproofofthisproperty,whichseemstohavebeenove rlookedinthevastliteratureonthesubjectanditsmanygeneralizations,isdiscussed.Thekeystepistheobservation, whichisneithernewnorwellknown,thattheconvexhullofthesetofprojectionmatricesofrankkisthesetofsymmetr icmatriceswitheigenvaluesbetween0and1andsummingtok.Theconnectionwiththewell-knownBirkhofftheore mondoublystochasticmatricesisalsodiscussed.Thisapproachprovidesaveryconvenientcharacterizationforthes ubdifferentialoftheeigenvaluesum,describedinaseparatepaper..

Suman, De. 2022. An efficient technique of resource scheduling in cloud using graph coloring algorithm. International Conference on Intelligent Engineering Approach(ICIEA-2022), *Global Transitions Proceedings* 3 (1): 169–176. ISSN: 2666-285X. https://doi.org/https://doi.org/10.1016/j.gltp.2022.03.005. https://www.sciencedirect.com/science/article/pii/S2666285X22000085.

Zais, Mark, and Manuel Laguna. 2016. A graph coloring approach to the deployment scheduling and unit assignment problem. *Journal of Scheduling* 19 (1): 73–90. ISSN: 1099-1425. https://doi.org/10.1007/s10951-015-0434-0. https://doi.org/10.1007/s10951-015-0434-0.