The min-cut and vertex separator problem

Fanz Rendl · Renata Sotirov

Abstract We consider graph three-partitions with the objective of minimizing the number of edges between the first two partition sets while keeping the size of the third block small. We review most of the existing relaxations for this min-cut problem and focus on a new class of semidefinite relaxations, based on matrices of order $2n+1$ which provide a good compromise between quality of the bound and computational effort to actually compute it. Here, n is the order of the graph. Our numerical results indicate that the new bounds are quite strong and can be computed for graphs of medium size ($n \approx 300$) with reasonable effort of a few minutes of computation time. Further, we exploit those bounds to obtain bounds on the size of the vertex separators. A vertex separator is a subset of the vertex set of a graph whose removal splits the graph into two disconnected subsets.

We also present an elegant way of convexifying non-convex quadratic problems by using semidefinite programming. This approach results with bounds that can be computed with any standard convex quadratic programming solver.

Keywords vertex separator · minimum cut · semidefinite programming · convexification

Fanz Rendl
Universitätsstrasse 65-67, 9020 Klagenfurt, Austria
Tel.: +43 463 2700 3114
Fax: +43 463 2700 993114
E-mail: Franz.Rendl@aau.at

Renata Sotirov
Warandelaan 2, 5000 LE, Tilburg, The Netherlands
Tel.: +31 13 466 3178
Fax: +31 13 466 3280
E-mail: r.sotirov@uvt.nl

1 Introduction

The vertex separator problem (VSP) for a graph is to find a subset of vertices (called vertex separator) whose removal disconnects the graph into two components of roughly equal size. The VSP is NP-hard. Some families of graphs are known to have small vertex separators. Lipton and Tarjan [27] provide a polynomial time algorithm which determines a vertex separator in n-vertex planar graphs of size $O(\sqrt{n})$. Their result was extended to some other families of graphs such as graphs of fixed genus [28]. It is also known that trees, 3D-grids and meshes have small separators. However, there are graphs that do not have small separators.

The VSP problem arises in many different fields such as VLSI design [5] and bioinformatics [17]. Finding vertex separators of minimal size is an important issue in communications network [25] and finite element methods [30]. The VSP also plays a role in divide-and-conquer algorithms for minimizing the work involved in solving system of equations, see e.g., [28, 29].

The vertex separator problem is related to the following graph partition problem. Let $A = (a_{ij})$ be the adjacency matrix of a graph G with vertex set $V(G) = \{1, \ldots, n\}$ and edge set $E(G)$. Thus A is a symmetric zero-one matrix of order n with zero diagonal. We are interested in 3-partitions (S_1, S_2, S_3) of $V(G)$ with the property that

$$|S_i| = m_i \geq 1.$$

Given A and $m = (m_1, m_2, m_3)^{\mathrm{T}}$ we consider the following *minimum cut* (MC) problem:

$$(\text{MC}) \quad \text{OPT}_{\text{MC}} := \min\{ \sum_{i \in S_1, j \in S_2} a_{ij} : (S_1, S_2, S_3) \text{ partitions } V, \ |S_i| = m_i, \forall i \}.$$

It asks to find a vertex partition (S_1, S_2, S_3) with specified cardinalities, such that the number of edges joining vertices in S_1 and S_2 is minimized. We remark that this min-cut problem is known to be NP-hard [18]. It is clear that if $\text{OPT}_{\text{MC}} = 0$ for some $m = (m_1, m_2, m_3)^{\mathrm{T}}$ then S_3 separates S_1 and S_2. On the other hand, $\text{OPT}_{\text{MC}} > 0$ shows that no separator S_3 for the cardinalities specified in m exists.

A natural way to model this problem in 0–1 variables consists in representing the partition (S_1, S_2, S_3) by the characteristic vectors x_i corresponding to S_i. Thus $x_i \in \{0, 1\}^n$ and $(x_i)_j = 1$ exactly if $j \in S_i$. Hence partitions with prescribed cardinalities are in one-to-one correspondence with $n \times 3$ zero-one matrices $X = (x_1, x_2, x_3)$ such that $X^{\mathrm{T}}e = m$ and $Xe = e$. (Throughout e denotes the vector of all ones of appropriate size.) The first condition takes care of the cardinalities and the second one insures that each vertex is in exactly one partition block. The number of edges between S_1 and S_2 is now given by $x_1^{\mathrm{T}} A x_2$. Thus (MC) is equivalent to

$$\min\{x_1^{\mathrm{T}} A x_2 : X = (x_1, x_2, x_3) \in \{0, 1\}^{n \times 3}, X^{\mathrm{T}}e = m, \ Xe = e\}. \quad (1)$$

It is the main purpose of this paper to explore computationally efficient ways to get tight approximations of $\mathrm{OPT}_{\mathrm{MC}}$. These will be used to find vertex separators of small size.

In section 2 we provide an overview of various techniques from the literature to get lower bounds for the min-cut problem. Section 3 contains a series of new relaxations based on semidefinite programming (SDP). We also consider convexification techniques suitable for Branch-and-Bound methods based on convex quadratic optimization, see section 4. In section 5.1 we investigate reformulations of our SDP relaxations where strictly feasible points exist. This is crucial for algorithms based on interior-point methods. We also show equivalence of some of the here introduced SDP relaxations with the SDP relaxations from the literature, see section 5.2. In particular, we prove that SDP relaxations with matrices of order $2n+1$ introduced here are equivalent to the SDP relaxations with matrices of order $3n+1$ from the literature. This reduction in the size of the matrix variable enables us to further improve and compute SDP bounds by adding the facet defining inequalities of the boolean quadric polytope. Symmetry ($m_1 = m_2$) is investigated in section 6. We also address the problem of getting feasible 0-1 solutions by standard rounding heuristics, see section 7. Section 8 provides computational results on various classes of graphs taken from the literature, and section 9 final conclusions.

Example 1 The following graph will be used to illustrate the various bounding techniques discussed in this paper. The vertices are selected from a 17×17 grid using the following MATLAB commands to make them reproduceable.

```
rand('seed',27072015), Q=rand(17)<0.33
```

This results in $n = 93$ vertices which correspond to the nonzero entries in Q. These are located at the grid points (i, j) in case $Q_{ij} \neq 0$. Two vertices are joined by an edge whenever their distance is at most $\sqrt{10}$. The resulting graph with $|E| = 470$ is displayed in Figure 1. For $m = (44,\ 43,\ 6)^{\mathrm{T}}$ we find a partition which leaves 7 edges between S_1 and S_2. We will in fact see later on that this partition is optimal for the specific choice of m. Vertices in S_3 are marked by '*', the edges in the cut between S_1 and S_2 are plotted with the thickest lines, those with one endpoint in S_3 are dashed.

Notation.
The space of $k \times k$ symmetric matrices is denoted by $\mathcal{S}_k$, and the space of $k \times k$ symmetric positive semidefinite matrices by $\mathcal{S}_k^+$. The space of symmetric matrices is considered with the trace inner product $\langle A, B \rangle = \mathrm{tr}(AB)$. We will sometimes also use the notation $X \succeq 0$ instead of $X \in \mathcal{S}_k^+$, if the order of the matrix is clear from the context.

We will use matrices having a block structure. We denote a sub-block of a matrix Y such as in equation (6) by Y_{ij} or Y_i. In contrast we indicate the (i, j) entry of a matrix Y by $Y_{i,j}$. For two matrices $X, Y \in \mathbf{R}^{p \times q}$, $X \geq Y$ means $X_{i,j} \geq Y_{i,j}$, for all i, j.
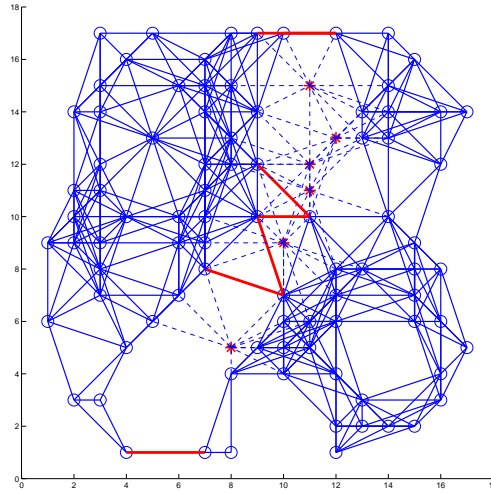
Fig. 1 The example graph: vertices in S_3 are marked by '*', the edges in the cut are plotted with the thickest lines

To denote the ith column of the matrix X we write $X_{:,i}$. J and e denote the all-ones matrix and all-ones vector respectively. The size of these objects will be clear from the context. We set $E_{ij} = e_i e_j^{\mathrm{T}}$ where e_i denotes column i of the identity matrix I.

The 'vec' operator stacks the columns of a matrix, while the 'diag' operator maps an $n \times n$ matrix to the n-vector given by its diagonal. The adjoint operator of 'diag' is denoted by 'Diag'.

The Kronecker product $A \otimes B$ of matrices $A \in \mathbf{R}^{p \times q}$ and $B \in \mathbf{R}^{r \times s}$ is defined as the $pr \times qs$ matrix composed of pq blocks of size $r \times s$, with block ij given by $A_{i,j}B$, $i = 1, \ldots, p$, $j = 1, \ldots, q$, see e.g. [19]. The Hadamard product of two matrices A and B of the same size is denoted by $A \circ B$ and defined as $(A \circ B)_{ij} = A_{i,j} \cdot B_{i,j}$ for all i, j.

2 Overview of relaxations for (MC)

Before we present our new relaxations for (MC) we find it useful to give a short overview of existing relaxation techniques. This allows us to set the stage for our own results and also to describe the rich structure of the problem which gives rise to a variety of relaxations.

2.1 Orthogonal relaxations based on the Hoffman-Wielandt inequality

The problem (MC) can be viewed as optimizing a quadratic objective function over the set P_m of partition matrices where

$$P_m := \{X \in \{0,1\}^{n \times 3} : X^{\mathrm{T}}e = m, Xe = e\}.$$

Historically the first relaxations exploit the fact that the columns of $X \in P_m$ are pairwise orthogonal, more precisely $X^{\mathrm{T}}X = \mathrm{Diag}(m)$. The objective function $x_1^{\mathrm{T}}Ax_2$ can be expressed as $\frac{1}{2}\langle A, XBX^{\mathrm{T}} \rangle$ with

$$B = \begin{pmatrix} 0\ 1\ 0 \\ 1\ 0\ 0 \\ 0\ 0\ 0 \end{pmatrix}. \tag{2}$$

We recall the Hoffman-Wielandt theorem which provides a closed form solution to the following type of problem.

Theorem 1 (Hoffman-Wielandt Theorem) *If $C,\ D \in \mathcal{S}_n$ with eigenvalues $\lambda_i(C), \lambda_j(D)$, then*

$$\min\{\langle C, XDX^{\mathrm{T}} \rangle :\ X^{\mathrm{T}}X = I\} = \min\{\sum_i \lambda_i(C)\lambda_{\phi(i)}(D) :\ \phi \text{ permutation}\}.$$

The minimum on the right hand side above is attained for the permutation which recursively maps the largest eigenvalue of C to the smallest eigenvalue of D.

Donath and Hoffman [10] use this result to bound (MC) from below,

$$\mathrm{OPT}_{\mathrm{MC}} \geq \frac{1}{2}\min\{\langle A, XBX^{\mathrm{T}} :\ X^{\mathrm{T}}X = \mathrm{Diag}(m)\}.$$

The fact that in this case A and B do not have the same size can easily be overcome, see for instance [34].

To get a further tightening, we introduce the *Laplacian L*, associated to the adjacency matrix A, which is defined as

$$L := \mathrm{Diag}(Ae) - A. \tag{3}$$

By definition, we have $-L = A$ outside the main diagonal, and moreover $\mathrm{diag}(XBX^{\mathrm{T}}) = 0$ for partition matrices X. Therefore the objective function of our problem satisfies $\langle -L, XBX^{\mathrm{T}} \rangle = \langle A, XBX^{\mathrm{T}} \rangle$. The vector e is eigenvector of L, in fact $Le = 0$, which is used in [34] to investigate the following relaxation

$$\mathrm{OPT}_{\mathrm{HW}} := \frac{1}{2}\min\{\langle -L, XBX^{\mathrm{T}} \rangle :\ X^{\mathrm{T}}X = \mathrm{Diag}(m),\ Xe = e,\ X^{\mathrm{T}}e = m\}. \tag{4}$$

This relaxation also has a closed form solution based on the Hoffman-Wielandt theorem. To describe it, we need some more notation. Let λ_2 and λ_n denote the second smallest and the largest eigenvalue of L, with normalized eigenvectors

v_2 and v_n. Further, set $\tilde{m} = (\sqrt{m_1},\ \sqrt{m_2},\ \sqrt{m_3})^{\mathrm{T}}$ and let W be an orthonormal basis of the orthogonal complement to $\tilde{m}$, $W^{\mathrm{T}}W = I_2$, $W^{\mathrm{T}}\tilde{m} = 0$. Finally, define $\tilde{B} = \sqrt{m_1 m_2} W^{\mathrm{T}}BW$ with factorization

$$Q^{\mathrm{T}}\tilde{B}Q = \mathrm{Diag}([\tau_1\ \tau_2]),$$

and eigenvalues $\tau_1 = \frac{1}{n}(-m_1 m_2 - \sqrt{m_1 m_2(n - m_1)(n - m_2)})$ and $\tau_2 = \frac{1}{n}(-m_1 m_2 + \sqrt{m_1 m_2(n - m_1)(n - m_2)})$.

Theorem 2 ([23]) *In the notation above we have* $\mathrm{OPT}_{\mathrm{HW}} = -\frac{1}{2}(\lambda_2 \tau_1 + \lambda_n \tau_2)$ *and the optimum is attained at*

$$X = \frac{1}{n}em^{\mathrm{T}} + (v_2,\ v_n)Q^{\mathrm{T}}W^{\mathrm{T}}\mathrm{Diag}(\tilde{m}). \tag{5}$$

This approach has been investigated for general graph partition with specified sizes $m_1, \ldots, m_k$. We refer to [34] and [14] for further details. More recently, Pong et al. [32] explore and extend this approach for the generalized min-cut problem.

The solution given in closed form through the eigenvalues of the input matrices makes it attractive for large-scale instances, see [32]. The drawback lies in the fact that it is difficult to introduce additional constraints into the model while maintaining the applicability of the Hoffman-Wielandt theorem. This can be overcome by moving to relaxations based on semidefinite optimization.

2.2 Relaxations using Semidefinite Optimization

The relaxations underlying the Hoffman-Wielandt theorem can equivalently be expressed using semidefinite optimization. We briefly describe this connection and then we consider more general models based on semidefinite optimization. The key tool here is the following theorem of Anstreicher and Wolkowicz [1], which can be viewed as an extension of the Hoffman-Wielandt theorem.

Theorem 3 ([1]) *Let* $C, D \in \mathcal{S}_n$. *Then*

$$\min\{\langle C, XDX^{\mathrm{T}}\rangle : X^{\mathrm{T}}X = I\} = \max\{\mathrm{tr}(S) + \mathrm{tr}(T) : D \otimes C - S \otimes I - I \otimes T \succeq 0\}.$$

Based on this theorem, Povh and Rendl [33] show that the optimal value of (4) can equivalently be expressed as the optimal solution of the following semidefinite program with matrix Y of order $3n$.

Theorem 4 ([33]) *We have*

$$\mathrm{OPT}_{\mathrm{HW}} = \min \frac{1}{2}\mathrm{tr}(-L)(Y_{12} + Y_{12}^{\mathrm{T}})$$
$$\text{s.t. } \mathrm{tr}Y_i = m_i,\ \ \mathrm{tr}(JY_i) = m_i^2,\ \ i = 1, 2, 3$$
$$\mathrm{tr}(Y_{ij} + Y_{ij}^{\mathrm{T}}) = 0,\ \ \mathrm{tr}J(Y_{ij} + Y_{ij}^{\mathrm{T}}) = 2m_i m_j,\ \ j > i$$
$$Y = \begin{pmatrix} Y_1 & Y_{12} & Y_{13} \\ Y_{12}^{\mathrm{T}} & Y_2 & Y_{23} \\ Y_{13}^{\mathrm{T}} & Y_{23}^{\mathrm{T}} & Y_3 \end{pmatrix} \succeq 0. \tag{6}$$

A proof of this result is given in [33]. The proof implicitly shows that also the following holds.

Theorem 5 *The following problem also has the optimal value* OPT_{HW}*:*

$$\mathrm{OPT}_{\mathrm{HW}} = \min \ \frac{1}{2}\mathrm{tr}(-L)(Y_{12} + Y_{12}^{\mathrm{T}})$$

$$\text{s.t. } \mathrm{tr}(Y_i) = m_i, \ \mathrm{tr}(JY_i) = m_i^2, \ i = 1, 2$$

$$\mathrm{tr}(Y_{12} + Y_{12}^{\mathrm{T}}) = 0, \ \mathrm{tr}(J(Y_{12} + Y_{12}^{\mathrm{T}})) = 2m_1 m_2$$

$$Y = \begin{pmatrix} Y_1 & Y_{12} \\ Y_{12}^{\mathrm{T}} & Y_2 \end{pmatrix} \succeq 0.$$

We provide an independent proof of this theorem, which simplifies the arguments from [33]. To maintain readability, we postpone the proof to section 10. The significance of these two results lies in the fact that we can compute optimal solutions for the respective semidefinite programs by simple eigenvalue computations.

The SDP relaxation from Theorem 4 can be viewed as moving from $X \in P_m$ to $Y = xx^{\mathrm{T}} \in \mathcal{S}_{3n}$ with $x = \mathrm{vec}(X) \in \mathbf{R}^{3n}$ and replacing the quadratic terms in x by the corresponding entries in Y. The constraint $\mathrm{tr}(Y_1) = m_1$ follows from $\mathrm{tr}(Y_1) = \mathrm{tr}(x_1 x_1^{\mathrm{T}}) = x_1^{\mathrm{T}} x_1 = m_1$. Similarly, $\mathrm{tr}(Y_{12}) = x_1^{\mathrm{T}} x_2 = 0$ and $\mathrm{tr}(JY_1) = (e^{\mathrm{T}} x_1)^2 = m_1^2$. Thus these constraints simply translate orthogonality of X into linear constraints on Y.

In order to derive stronger SDP relaxations than the one from Theorem 4, one can exploit the fact that for $X \in P_m$ it follows that $\mathrm{diag}(Y) = \mathrm{diag}(xx^{\mathrm{T}}) = x$ with $x = \mathrm{vec}(X)$. Now, the constraint $Y - \mathrm{diag}(Y)\mathrm{diag}(Y)^{\mathrm{T}} = 0$ may be weakened to $Y - \mathrm{diag}(Y)\mathrm{diag}(Y)^{\mathrm{T}} \succeq 0$ which is well known to be equivalent to the following convex constraint

$$\begin{pmatrix} Y & \mathrm{diag}(Y) \\ \mathrm{diag}(Y)^{\mathrm{T}} & 1 \end{pmatrix} \succeq 0.$$

The general case of $k-$partition leads to SDP relaxations with matrices of order $(nk + 1)$, see for instance Zhao et al. [42] and Wolkowicz and Zhao [41]. In our notation, the model (4.1) from [41] has the following form:

$$\min \ \tfrac{1}{2}\mathrm{tr}A(Y_{12} + Y_{12}^{\mathrm{T}})$$

$$\text{s.t. } \mathrm{tr}Y_i = m_i, \ \mathrm{tr}(JY_i) = m_i^2, \ \ i = 1, 2, 3$$

$$\mathrm{diag}(Y_{ij}) = 0, \ \mathrm{tr}J(Y_{ij} + Y_{ij}^{\mathrm{T}}) = 2m_i m_j, \ \ j > i \qquad (7)$$

$$Y = \begin{pmatrix} Y_1 & Y_{12} & Y_{13} \\ Y_{12}^{\mathrm{T}} & Y_2 & Y_{23} \\ Y_{13}^{\mathrm{T}} & Y_{23}^{\mathrm{T}} & Y_3 \end{pmatrix}, \ \ y = \mathrm{diag}(Y), \ \ Y - yy^{\mathrm{T}} \succeq 0.$$

Literally speaking, the model (4.1) from [41] does not include the equations involving the m_i above, but uses information from the barycenter of the feasible region to eliminate these constraints by reducing the dimension of the matrix variable Y. We make this more precise in section 5 below.

Further strengthening is suggested by asking $Y \geq 0$ leading to the strongest bound contained in [41].

The min-cut problem can also be seen as a special case of the quadratic assignment problem (QAP), as noted already by Helmberg et al. [23]. This idea is further exploited by van Dam and Sotirov [13] where the authors use the well known SDP relaxation for the QAP [42], as the SDP relaxation for the min-cut problem. The resulting QAP-based SDP relaxation for the min-cut problem is proven to be equivalent to (7), see [13].

2.3 Linear and Quadratic Programming Relaxations

The model (MC) starts with specified sizes $m = (m_1, m_2, m_3)^{\mathrm{T}}$ and tries to separate $V(G)$ into S_1, S_2 and S_3 so that the number of edges between S_1 and S_2 is minimized. This by itself does not yield a vertex separator, but it can be used to experiment with different choices of m to eventually produce a separator.

Several papers consider the separator problem directly as a linear integer problem of the following form

$$(VS) \quad \begin{aligned} \min\{e^{\mathrm{T}}x_3 : X = (x_1, x_2, x_3) \in \{0,1\}^{n \times 3}, \ Xe = e, \ (x_1)_i + (x_2)_j \leq 1 \\ \forall [i,j] \in E, \ l_i \leq e^{\mathrm{T}}x_i \leq u_i \ i = 1, 2\}. \end{aligned}$$

The constraint $Xe = e$ makes sure that X represents a vertex partition, the inequalities on the edges inforce that there are no edges joining S_1 and S_2 and the last constraints are cardinality conditions on S_1 and S_2. The objective function looks for a separator of smallest size. We refer to Balas and de Souza [6,39] who exploit the above integer formulation within Branch and Bound settings with additional cutting planes to find vertex separators in small graphs. Biha and Meurs [7] introduced new classes of valid inequalities for the vertex separator polyhedron and solved instances from [39] to optimality.

Hager et al., [20,21] investigate continuous bilinear versions and show that

$$\max\{e^{\mathrm{T}}x_1 + e^{\mathrm{T}}x_2 - \langle x_1, (A+I)x_2 \rangle : \ 0 \leq x_i \leq e, \ l_i \leq e^{\mathrm{T}}x_i \leq u_i \ i = 1, 2\}$$

is equivalent to (VS). Even though this problem is intractable, as the objective function is indefinite, it is shown in [21] that this model can be used to produce heuristic solutions of good quality even for very large graphs.

A quadratic programming (QP) relaxation for the min-cut problem is derived in [32]. That convex QP relaxation is based on the QP relaxation for the QAP, see [1–3]. Numerical results in [32] show that QP bounds are weaker, but cheaper to compute than the strongest SDP bounds, see also section 4.

In 2012, Armbruster et al. [4] compared branch-and-cut frameworks for linear and semidefinite relaxations of the minimum graph bisection problem on large and sparse instances. Extensive numerical experiments show that the semidefinite branch-and-cut approach is superior choice to the simplex approach. In the sequel we mostly consider SDP bounds for the min-cut problem.

3 The new SDP relaxations

In this section we derive several SDP relaxations with matrix variables of order $2n+1$ and increasing complexity. We also show that our strongest SDP relaxation provides tight min-cut bounds on a graph with 93 vertices.

Motivated by Theorem 5 and also in view of the objective function $x_1^{\mathrm{T}} A x_2$ of (MC), which makes explicit use only of the first two columns of $X \in P_m$ we propose to investigate SDP relaxations of (MC) with matrices of order $2n$, obtained by moving from $\binom{x_1}{x_2}$ to $\binom{x_1}{x_2}\binom{x_1}{x_2}^{\mathrm{T}}$.

An integer programming formulation of (MC) using only x_1 and x_2 amounts to the following

$$\mathrm{OPT}_{\mathrm{MC}} = \min\{x_1^{\mathrm{T}} A x_2 : \; x_1, x_2 \in \{0,1\}^n, \; x_1^{\mathrm{T}} e = m_1, \; x_2^{\mathrm{T}} e = m_2, \; x_1 + x_2 \le e\}. \tag{8}$$

This formulation has the disadvantage that its linear relaxation ($0 \le x_i \le 1$) is intractable, as the objective function $x_1^{\mathrm{T}} A x_2$ is indefinite. An integer linear version is obtained by linearizing the terms $(x_1)_i(x_2)_j$ in the objective function. We get

$$\mathrm{OPT}_{\mathrm{MC}} = \min\{ \sum_{[ij]\in E(G)} z_{ij} + z_{ji} : u,v \in \{0,1\}^n, z \ge 0, e^{\mathrm{T}}u = m_1, e^{\mathrm{T}}v = m_2,$$

$$u + v \le e, z_{ij} \ge u_i + v_j - 1, z_{ji} \ge u_j + v_i - 1 \; \forall[ij] \in E(G)\}.$$

This is a binary LP with $2n$ binary and $2m$ continuous variables. Unfortunately, its linear relaxation gives a value of 0 (by setting an appropriate number of the nonzero entries in u and v to $\frac{1}{2}$). Even the use of advanced ILP technology, as for instance provided by GUROBI or CPLEX or similar packages, is only moderately successful on this formulation. We will argue below that some SDP models in contrast yield tight approximations to the optimal value of the integer problem.

Moving to the matrix space, we consider

$$Y = \begin{pmatrix} Y_1 & Y_{12} \\ Y_{12}^{\mathrm{T}} & Y_2 \end{pmatrix},$$

where Y_i corresponds to $x_i x_i^{\mathrm{T}}$ and Y_{12} represents $x_1 x_2^{\mathrm{T}}$. The objective function $x_1^{\mathrm{T}} A x_2$ becomes

$$\langle A, Y_{12} \rangle = \langle M, Y \rangle,$$

where we set

$$M := \frac{1}{2} \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}. \tag{9}$$

Looking at Theorem 5 we consider the following simple SDP as our starting relaxation.

$$(SDP_0) \min \langle M, Y \rangle$$

$$\text{s.t. } \operatorname{tr}(Y_i) = m_i, \ \operatorname{tr}(JY_i) = m_i^2, \ i = 1, 2 \tag{10}$$

$$\operatorname{tr}(Y_{12} + Y_{12}^{\mathrm{T}}) = 0, \ \operatorname{tr}(J(Y_{12} + Y_{12}^{\mathrm{T}})) = 2m_1 m_2 \tag{11}$$

$$Y = \begin{pmatrix} Y_1 & Y_{12} \\ Y_{12}^{\mathrm{T}} & Y_2 \end{pmatrix}, \ y = \operatorname{diag}(Y), \tag{12}$$

$$\begin{pmatrix} Y & y \\ y^{\mathrm{T}} & 1 \end{pmatrix} \succeq 0 \tag{13}$$

This SDP captures the constraints from Theorem 5 and has $(2n + 1) + 6$ linear equality constraints. We have replaced $Y \succeq 0$ by the stronger condition $Y - yy^{\mathrm{T}} \succeq 0$, and we also replaced $-L$ by A in the objective function.

There is an immediate improvement by exploiting the fact that $x_1 \circ x_2 = 0$ (elementwise product $(x_1)_i \cdot (x_2)_i = 0$ is zero). Thus we also impose

$$\operatorname{diag}(Y_{12}) = 0, \tag{14}$$

which adds another n equations and makes $\operatorname{tr}(Y_{12} + Y_{12}^{\mathrm{T}}) = 0$ redundant. We call SDP_1 the relaxation obtained from SDP_0 by replacing $\operatorname{tr}(Y_{12} + Y_{12}^{\mathrm{T}}) = 0$ with $\operatorname{diag}(Y_{12}) = 0$. The equations in (14) are captured by the 'gangster operator' in [41]. Moreover, once these constraints are added, it will make no difference whether the adjacency matrix A or $-L$ is used in the objective function.

Up to now we have not yet considered the inequality

$$e - x_1 - x_2 \geq 0, \tag{15}$$

where x_1 (resp. x_2) represents the first n (resp. last n) coordinates of y.

Lemma 6 *Let Y satisfy (12), (13) and (14). Then $\operatorname{diag}(Y_1) + \operatorname{diag}(Y_2) \leq e$.*

Proof The submatrix in (13) indexed by $(i, n + i, 2n + 1)$ and $i \in \{1, \dots, n\}$ is positive semidefinite, i.e.,

$$\begin{pmatrix} y_i & 0 & y_i \\ 0 & y_{i+n} & y_{i+n} \\ y_i & y_{i+n} & 1 \end{pmatrix} \succeq 0.$$

The proof of the lemma follows from the following inequality

$$(-1, -1, 1)^{\mathrm{T}} \begin{pmatrix} y_i & 0 & y_i \\ 0 & y_{i+n} & y_{i+n} \\ y_i & y_{i+n} & 1 \end{pmatrix} (-1, -1, 1) \geq 0$$

□

In order to obtain additional linear constraints for our SDP model, we consider (15) and

$$y \geq 0, \ e - y \geq 0,$$

which we multiply pairwise and apply linearization. A pairwise multiplication of individual inequalities from (15) yields

$$1 - (y_i + y_j + y_{n+i} + y_{n+j}) + Y_{i,j} + Y_{i,n+j} + Y_{j,n+i} + Y_{n+i,n+j} \geq 0 \ \ \forall i < j \in V(G). \tag{16}$$

We also get

$$y_j - Y_{i,j} - Y_{n+i,j} \geq 0, \quad y_{n+j} - Y_{i,n+j} - Y_{n+i,n+j} \geq 0 \ \ \forall i, j \in V(G), \ i \neq j \tag{17}$$

by multiplying individual constraints from (15) and $y \geq 0$. Finally we get in a similar way by multiplying with $e - y \geq 0$

$$1 - y_i - y_{n+i} - y_j + Y_{i,j} + Y_{n+i,j} \geq 0 \tag{18}$$

$$1 - y_i - y_{n+i} - y_{n+j} + Y_{i,n+j} + Y_{n+i,n+j} \geq 0 \ \ \forall i, j \in V(G), \ i \neq j. \tag{19}$$

The inequalities (16)–(19) are based on a technique known as the reformulation-linearization technique (RLT) that was introduced by Sherali and Adams [36]. In order to strengthen our SDP relaxation further, one can add the following facet defining inequalities of the boolean quadric polytope (BQP), see e.g., [31],

$$0 \leq Y_{i,j} \leq Y_{i,i}$$
$$Y_{i,i} + Y_{j,j} \leq 1 + Y_{i,j}$$
$$Y_{i,k} + Y_{j,k} \leq Y_{k,k} + Y_{i,j} \tag{20}$$
$$Y_{i,i} + Y_{j,j} + Y_{k,k} \leq Y_{i,j} + Y_{i,k} + Y_{j,k} + 1.$$

In our numerical experiments we will consider the following relaxations which we order according to their computational effort.

name	constraints	complexity
SDP_0	(10)–(13)	$O(n)$
SDP_1	(10)–(13), (14)	$O(n)$
SDP_2	(10)–(13), (14), $Y_{i,j} \geq 0$ if $M_{i,j} > 0$	$O(n + |E|)$
SDP_3	(10)–(13), (14), $Y_{i,j} \geq 0$ if $M_{i,j} > 0$, (16)–(19)	$O(n^2)$
SDP_4	(10)–(13), (14), (16)–(19), (20)	$O(n^3)$

The first two relaxations can potentially produce negative lower bounds, which would make them useless. The remaining relaxations yield nonnegative bounds due to the nonnegativity condition on Y corresponding to the nonzero entries in M.

Example 2 We continue with the example from the introduction and provide the various lower bounds introduced so far, see Table 1. We also include the value of the best known feasible solution, which we found using a rounding heuristic described in section 7 below. In most of these cases OPT_{eig}, SDP_0 and SDP_1 bounds are negative but we know that $\mathrm{OPT}_{\mathrm{MC}} \geq 0$. In contrast, the strongest bound SDP_4 proves optimality of all the solutions found by our heuristic. Here we do not solve SDP_3 and SDP_4 exactly. The SDP_3 (resp. SDP_4) bounds are obtained by adding the most violated inequalities of type (16)–(19) (resp. (16)–(19) and (20)) to SDP_2. The cutting plane scheme adds at most $2n$ violated valid constraints in each iteration and performs at most 15 iterations. It takes about 6 minutes to compute SDP_4 bound for fixed m. We compute SDP_2 (resp. SDP_3) bound in about 5 seconds (resp. 2 minutes) for fixed m.

For comparison purposes, we computed also linear programming (LP) bounds. The LP bound RLT_3 incorporates all constraints from SDP_3 except the SDP constraint, including of course standard linearization constraints. The RLT_3 bound for $m = (45, 44, 4)^{\mathrm{T}}$ is zero. Similarly, we derive the linear programming bound RLT_4 that includes all constraints from SDP_4 except the SDP constraint. We solve RLT_4 approximately by cutting plane scheme that first adds all violated (16)–(19) constraints and then at most $4n$ violated constraints of type (20), in each iteration of the algorithm. After 100 such iterations the bound RLT_4 was still zero.

We find it remarkable that even the rather expensive model SDP_3 is not able to approximate the optimal solution value within 'reasonable' limits. On the other hand, the 'full' model SDP_4 is strong enough to actually solve these problems. We will see in the computational section, that only the full model is strong enough to actually get good approximations also on instances from the literature.

Table 1 Lower bounds for the example graph.

m^{T}	(45,44,4)	(44, 43,6)	(42,42,9)	(42,41,10)
OPT_{eig}	-7.39	-15.66	-27.32	-31.03
SDP_0	-14.61	-23.66	-34.58	-37.71
SDP_1	3.16	-4.17	-13.93	-16.92
SDP_2	5.69	1.53	0.00	0.00
SDP_3	7.37	3.43	0.05	0.00
SDP_4	11.82	6.41	1.63	0.26
feas. sol.	12	7	2	1

4 Bounds based on convex quadratic optimization

Convex quadratic programming bounds are based on the fact that minimizing a convex quadratic function over a polyhedron is tractable and moreover there are several standard solvers available for this type of problem.

In our case the objective function $f(x)$ is not convex. It can however be reformulated as a convex quadratic $L(x)$ such that $f(x) = L(x)$ for all feasible 0-1 solutions by exploiting the fact that $x \circ x = x$ for 0-1 valued x. This convexification is based on Lagrangian duality and has a long history in nonlinear optimization, see for instance Hammer and Rubin [22] and Shor [37,38]. Lemarechal and Oustry [26], Faye and Roupin [15] and Billionet, Elloumi and Plateau [8] consider convexification of quadratic problems and the connection to semidefinite optimization. We briefly summarize the theoretical background behind this approach.

4.1 Convexification of 0-1 problems

We first recall the following well-known facts from convex analysis. Let $f(x) := x^{\mathrm{T}}Qx + 2q^{\mathrm{T}}x + q_0$ for $q_0 \in \mathbf{R}, q \in \mathbf{R}^n$ and $Q \in S_n$. Then

$$\inf_x f(x) > -\infty \iff Q \succeq 0 \text{ and } \exists \xi \in \mathbf{R}^n \text{ such that } q = Q\xi,$$

due to the first order $(Qx + q = 0)$ and second order $(Q \succeq 0)$ necessary optimality conditions. The following proposition summarizes what we need later for convexification.

Proposition 7 *Let $Q \succeq 0$ and $q = Q\xi$ for some $\xi \in \mathbf{R}^n$ and $q_0 \in \mathbf{R}$, and $f(x) = x^{\mathrm{T}}Qx + 2q^{\mathrm{T}}x + q_0$. Set $f^* := q_0 - \xi^{\mathrm{T}}Q\xi$. Then*

1. $\inf\{f(x) : x \in \mathbf{R}^n\} = f^*,$
2. $\inf\{\langle Q, X \rangle + 2q^{\mathrm{T}}x + q_0 : X - xx^{\mathrm{T}} \succeq 0\} = f^*,$
3. $\sup\{q_0 + \sigma : \begin{pmatrix} Q & q \\ q^{\mathrm{T}} & -\sigma \end{pmatrix} \succeq 0\} = f^*.$

Proof For completeness we include the following short arguments. The first statement follows from $\nabla f(x) = 0$. To see the last statement we use the factorization

$$\begin{pmatrix} Q & Q\xi \\ (Q\xi)^{\mathrm{T}} & -\sigma \end{pmatrix} = \begin{pmatrix} Q^{\frac{1}{2}} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & Q^{\frac{1}{2}}\xi \\ (Q^{\frac{1}{2}}\xi)^{\mathrm{T}} & -\sigma \end{pmatrix} \begin{pmatrix} Q^{\frac{1}{2}} & 0 \\ 0 & 1 \end{pmatrix}.$$

Using a Schur-complement argument, this implies

$$\begin{pmatrix} I & Q^{\frac{1}{2}}\xi \\ (Q^{\frac{1}{2}}\xi)^{\mathrm{T}} & -\sigma \end{pmatrix} \succeq 0 \iff -\sigma - \xi^{\mathrm{T}}Q\xi \geq 0,$$

which shows that the supremum is attained at $-\xi^{\mathrm{T}}Q\xi$ with optimal value $q_0 - \xi^{\mathrm{T}}Q\xi$. Finally, the second problem is the dual of the third, with strictly feasible solution $X = I$, $x = 0$. $\square$

We are going to describe the convexification procedure for a general problem of the form

$$\min\{f(x): \ x \in \{0,1\}^n, \ Dx = d, \ Cx \geq c\} \qquad (21)$$

for suitable data D, d, C, c. In case of (MC) we have $x = \binom{x_1}{x_2} \in \mathbf{R}^{2n}$ with $n+2$ constraints $e^{\mathrm{T}}x_1 = m_1, \ e^{\mathrm{T}}x_2 = m_2, \ x_1 + x_2 \leq e$.

The key idea is to consider a relaxation of the problem where integrality of x is expressed by the quadratic equation $x \circ x = x$. Let us consider the following simple relaxation

$$\inf\{f(x): \ x \circ x = x, \ Dx = d\} \qquad (22)$$

to explain the details. Its Lagrangian is given by

$$L(x; u, \alpha) = f(x) + \langle u, x \circ x - x \rangle + \langle \alpha, d - Dx \rangle.$$

The associated *Lagrangian dual* reads

$$\sup_{u,\alpha} \ \inf_x \ L(x; u, \alpha).$$

Ignoring values (u, α) where the infimum is $-\infty$, this is by Proposition 7 equivalent to

$$\sup_{u,\alpha} \ \sup_\sigma \{q_0 + \langle \alpha, d \rangle + \sigma : \ \begin{pmatrix} Q + \mathrm{Diag}(u) & q - \frac{1}{2}D^{\mathrm{T}}\alpha - \frac{1}{2}u \\ (.)^{\mathrm{T}} & -\sigma \end{pmatrix} \succeq 0\}. \qquad (23)$$

This is a semidefinite program with strictly feasible points (by selecting u and $-\sigma$ large enough and $\alpha = 0$). Hence its optimal value is equal to the value of the dual problem, which reads

$$\inf\{\langle Q, X \rangle + 2q^{\mathrm{T}}x + q_0 : \ \begin{pmatrix} X & x \\ x^{\mathrm{T}} & 1 \end{pmatrix} \succeq 0, \ \mathrm{diag}(X) = x, \ Dx = d\}. \qquad (24)$$

Let $(u^*, \alpha^*, \sigma^*)$ be an optimal solution of (23). Then we have

$$q_0 + \langle \alpha^*, d \rangle + \sigma^* = \sup_\sigma \{q_0 + \langle \alpha^*, d \rangle + \sigma : \ \begin{pmatrix} Q + \mathrm{Diag}(u^*) & q - \frac{1}{2}D^{\mathrm{T}}\alpha^* - \frac{1}{2}u^* \\ (.)^{\mathrm{T}} & -\sigma \end{pmatrix} \succeq 0\}.$$

Using Proposition 7 again we get the following equality

$$\inf_x L(x; u^*, \alpha^*) = q_0 + \langle \alpha^*, d \rangle + \sigma^*. \qquad (25)$$

The proposition also shows that $L(x; u^*, \alpha^*)$ is convex (in x) and moreover $L(x; u^*, \alpha^*) = f(x)$ for all integer feasible solutions x. The convex quadratic programming relaxation of problem (21), obtained from (22), consists in minimizing $L(x; u^*, \alpha^*)$ over the polyhedron

$$0 \leq x \leq e, \ Dx = d, \ Cx \geq c.$$

We close the general description of convexification with the following observation which will be used later on.

Proposition 8 *Let (X^*, x^*) be optimal for (24) and $(u^*, \alpha^*, \sigma^*)$ be optimal for (23). Then*

$$\inf_x L(x; u^*, \alpha^*) = L(x^*, u^*, \alpha^*) = q_0 + \langle \alpha^*, d \rangle + \sigma^*.$$

Thus the main diagonal x^* from the semidefinite program (24) is also a minimizer for the unconstrained minimization of $L(x; .)$.

Proof From $X^* - x^* x^{*\mathrm{T}} \succeq 0$ and $Q + \mathrm{Diag}(u^*) \succeq 0$ we get

$$x^{*\mathrm{T}}(Q + \mathrm{Diag}(u^*))x^* \leq \langle Q + \mathrm{Diag}(u^*), X^* \rangle. \tag{26}$$

We also have, using (25) and strong duality

$$\inf_x L(x; u^*, \alpha^*) = q_0 + \langle \alpha^*, d \rangle + \sigma^* = \langle Q, X^* \rangle + 2q^{\mathrm{T}}x^* + q_0.$$

Feasibility of (X^*, x^*) for (24) shows that the last term is equal to

$$\langle Q + \mathrm{Diag}(u^*), X^* \rangle + 2q^{\mathrm{T}}x^* + q_0 - u^{*\mathrm{T}}x^* + \alpha^{*\mathrm{T}}(d - Dx^*).$$

Finally, using (26), this term is lower bounded by $L(x^*; u^*, \alpha^*)$.   □

The relaxation (22) is rather simple. In [8] it is suggested to include all equations obtained by multiplying the constraints $e^{\mathrm{T}}x_1 = m_1, e^{\mathrm{T}}x_2 = m_2$ with $x_{(i)}$ and $1 - x_{(i)}$, where $x_{(i)}$ denotes the i−th component of x. The inclusion of quadratic constraints is particularly useful, as their multipliers provide additional degrees of freedom for the Hessian of the Lagrangian function.

The main insight from the analysis so far can be summarized as follows. Given a relaxation of the original problem, such as (22) above, form the associated semidefinite relaxation obtained from relaxing $X - xx^{\mathrm{T}} = 0$ to $X - xx^{\mathrm{T}} \succeq 0$. Then the optimal solution of the dual problem yields the desired convexification.

4.2 Convexifying (8)

Let us now return to (MC) given in (8). We have $x = \binom{x_1}{x_2} \in \mathbf{R}^{2n}$ and $f(x) = x^{\mathrm{T}} M x$ where M is given in (9). The following models are natural candidates for convexification. We list them in increasing order of computational effort to determine the convexification.

- Starting with $\inf\{f(x) : x \circ x = x\}$ we get the Lagrangian relaxation

$$\min\{\langle M, Y \rangle : Y - yy^{\mathrm{T}} \succeq 0, y = \mathrm{diag}(Y)\} = \max_{\sigma, u}\left\{\sigma : \begin{pmatrix} M + \mathrm{diag}(u) & -\frac{1}{2}u \\ -\frac{1}{2}u^{T} & -\sigma \end{pmatrix} \succeq 0\right\}.$$

Let σ^*, u^* be an optimal solution to the last problem above. The Lagrangian $L(x; u^*) = x^{\mathrm{T}} M x + \langle u^*, x \circ x - x \rangle = x^{\mathrm{T}}(M + \mathrm{Diag}(u^*))x - u^{*\mathrm{T}}x$ is convex

Table 2 Convexification for the example graph

m^{T}	(44, 43,6)	(42,42,9)	(42,41,10)
(28)	-117.50	-117.50	-117.50
(27)	-34.63	-48.16	-52.28
SDP_0	-23.66	-34.58	-37.71
convex QP	-15.65	-26.76	-30.13
SDP_1	-4.17	-13.93	-16.92
feas. sol.	7	2	1

and $f(x) = L(x; u^*)$ for all $x \in \{0,1\}^{2n}$. The convex QP bound based on this convexification is therefore given by

$$\min\{L(x; u^*) : x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ e^{\mathrm{T}}x_1 = m_1, \ e^{\mathrm{T}}x_2 = m_2, \ x_1 + x_2 \le e, \ x \ge 0\}. \quad (27)$$

The unconstrained minimization of L would result in

$$\inf\{L(x; u^*) : x \in \mathbf{R}^{2n}\}. \quad (28)$$

• In the previous model, no information of the m_i is used in the convexification. We next include also the equality constraints

$$e^{\mathrm{T}}x_i = m_i, \ (e^{\mathrm{T}}x_i)^2 = m_i^2, \ (i = 1,2), \ (e^{\mathrm{T}}x_1)(e^{\mathrm{T}}x_2) = m_1 m_2, \ x_1^{\mathrm{T}}x_2 = 0.$$

The Lagrangian relaxation corresponds to SDP_0. The dual solution of SDP_0 yields again the desired convexification.

• Finally, we replace $x_1^{\mathrm{T}}x_2$ by $x_1 \circ x_2 = 0$ above and get SDP_1 as Lagrangian relaxation. In this case we know from Lemma 6 and Proposition 8 that the convex QP bound is equal to the value of SDP_1 which in turn is equal to the unconstrained minimum of the Lagrangian L.

Example 3 We apply the convexification as explained above to the example graph from the introduction, see Table 2. In the first two cases we provide the unconstrained minimum along with the resulting convex quadratic programming bound. In case of SDP_1 we know from the previous proposition that the unconstrained minimum agrees with the optimal value of SDP_1. These bounds are not very useful, as we know a trivial lower bound of zero in all cases. On the other hand, the convex QP bound is computationally cheap compared to solving SDP, and may be useful in a Branch-and-Bound process.

Convex quadratic programming bounds may play crucial role in solving non-convex quadratic 0-1 problems to optimality, see for instance the work of Anstreicher et al. [2] on the quadratic assignment problem. Here we presented a general framework for obtaining convexifications, partly iterating the approach from [8]. Compared to Table 1, we note that the bounds based on convexification and using convex QP are not competitive to the strongest SDP bounds. On the other hand, these bounds are much cheaper to compute, so

their use in a Branch and Bound code may still be valuable. Implementing such bounds within a Branch and Bound framework is out of the scope of the present paper, so we will leave this for future research.

5 The Slater feasible versions of the SDP relaxations

In this section we present the Slater feasible versions of here introduced SDP relaxations. In particular, in section 5.1 we derive the Slater feasible version of the relaxation SDP_1, and in section 5.2 present the Slater feasible version of the SDP relaxation (7) from [41]. In section 5.2 we prove that the SDP relaxations SDP_1 and (7) are equivalent, and that SDP_3 with additional nonnegativity constraints on all matrix elements is equivalent to the strongest SDP relaxation from [41]. We actually show here that our strongest SDP relaxation with matrix variable of order $2n+1$, i.e., SDP_4 dominates the currently strongest SDP relaxation with matrix variable of order $3n+1$.

5.1 The projected new relaxations

In this section we take a closer look at the feasible region of our basic new relaxation SDP_0. The following lemma will be useful.

Lemma 9 *Suppose $X - xx^\mathrm{T} \succeq 0$, $x = \mathrm{diag}(X)$, and there exists $a \neq 0$ such that $a^\mathrm{T}(X - xx^\mathrm{T})a = 0$. Set $t := a^\mathrm{T}x$. Then $(a^\mathrm{T}, -t)^\mathrm{T}$ is eigenvector of $\begin{pmatrix} X & x \\ x^\mathrm{T} & 1 \end{pmatrix}$ to the eigenvalue 0.*

Proof If $X - xx^\mathrm{T} \succeq 0$ and $a^\mathrm{T}(X - xx^\mathrm{T})a = 0$, then $(X - xx^\mathrm{T})a = 0$, hence $Xa = tx$. From this the claim follows. $\square$

We introduce some notation to describe the feasible region of SDP_0. Let Z be a symmetric $(2n+1) \times (2n+1)$ matrix with the block form

$$Z = \begin{pmatrix} Y_1 & Y_{12} & y_1 \\ Y_{12}^\mathrm{T} & Y_2 & y_2 \\ y_1^\mathrm{T} & y_2^\mathrm{T} & 1 \end{pmatrix} \tag{29}$$

as in the definition of SDP_0. We define

$$F_1 := \{Z : Z \text{ satisfies } (29), (10), (12), (13) \text{ and } \mathrm{tr}(J(Y_{12} + Y_{12}^\mathrm{T})) = 2m_1m_2\}. \tag{30}$$

The set F_1 differs from the feasible region of SDP_0 only in the constraint $\mathrm{tr}(Y_{12} + Y_{12}^\mathrm{T}) = 0$ which is not included in F_1.

Lemma 10 *Let $Z \in F_1$. Then $T = \begin{pmatrix} e & 0 \\ 0 & e \\ -m_1 & -m_2 \end{pmatrix}$ is contained in the nullspace of Z.*

Proof The vector $a^{\mathrm{T}} := (e^{\mathrm{T}}, 0, \ldots, 0)$ satisfies $a^{\mathrm{T}}y = m_1$ and $a^{\mathrm{T}}(Y - yy^{\mathrm{T}})a = 0$. Therefore, using the previous lemma, the first column of T is eigenvector of Z to the eigenvalue 0. A similar argument applies to $a^{\mathrm{T}} := (0, \ldots, 0, e^{\mathrm{T}})$. $\square$

Let V_n denotes a basis of $e^{\perp}$, for instance

$$V_n := \begin{pmatrix} I \\ -e^{\mathrm{T}} \end{pmatrix} \in \mathbf{R}^{n \times (n-1)}. \tag{31}$$

Then the matrix

$$W := \begin{pmatrix} V_n & 0 & \frac{m_1}{n}e \\ 0 & V_n & \frac{m_2}{n}e \\ 0 & 0 & 1 \end{pmatrix} \tag{32}$$

forms a basis of the orthogonal complement to T, $W^{\mathrm{T}}T = 0$. Using the previous lemma, we conclude that $Z \in F_1$ implies that $Z = WUW^{\mathrm{T}}$ for some $U \in \mathcal{S}_{2n-1}^{+}$. Let us also introduce the set

$$F_2 := \{WUW^{\mathrm{T}} : U \in \mathcal{S}_{2n-1}^{+}, U_{2n-1,2n-1} = 1, \operatorname{diag}(WUW^{\mathrm{T}}) = WUW^{\mathrm{T}}e_{2n+1}\}.$$

Here e_{2n+1} is the last column of the identity matrix of order $2n + 1$. In the following theorem we prove that sets F_1 and F_2 are equal. Similar results are also shown in the connection with the quadratic assignment problem, see e.g., [42].

Theorem 11 $F_1 = F_2$.

Proof. We first show that $F_1 \subseteq F_2$ and take $Z \in F_1$. The previous lemma implies that Z is of the form $Z = WUW^{\mathrm{T}}$ and $U \succeq 0$. $Z_{2n+1,2n+1} = 1$ implies that $U_{2n-1,2n-1} = 1$ due to the way W is defined in (32). The main diagonal of Z is equal to its last column, which translates into $\operatorname{diag}(Z) = Ze_{2n+1}$, so $Z \in F_2$.

Conversely, consider $Z = WUW^{\mathrm{T}} \in F_2$ and let it be partitioned as in (29). We have $W^{\mathrm{T}}T = 0$, so $ZT = 0$. Multiplying out columnwise and using the block form of Z we get

$$Y_i e = m_i y_i, \ y_i^{\mathrm{T}} e = m_i \ \text{and} \ Y_{12}^{\mathrm{T}} e = m_1 y_2, \ Y_{12} e = m_2 y_1.$$

From this we conclude that $\operatorname{tr}(JY_i) = e^{\mathrm{T}}Y_i e = m_i e^{\mathrm{T}}y_i = m_i^2$ and $\operatorname{tr}(J(Y_{12} + Y_{12}^{\mathrm{T}})) = e^{\mathrm{T}}(Y_{12} + Y_{12}^{\mathrm{T}})e = 2m_1 m_2$. Finally, $\operatorname{diag}(Z) = Ze_{2n+1}$ yields $\operatorname{diag}(Y_i) = y_i$ and we have $\operatorname{tr}Y_i = e^{\mathrm{T}}\operatorname{diag}(Y_i) = e^{\mathrm{T}}y_i = m_i$, hence $Z \in F_1$. $\square$

We conclude by arguing that F_2 contains matrices where $U \succ 0$. To see this we note that the barycenter of the feasible set is

$$\hat{Z} = \begin{pmatrix} \frac{m_1}{n}I + \frac{m_1(m_1-1)}{n(n-1)}(E - I) & \frac{m_1 m_2}{n(n-1)}(E - I) & \frac{m_1}{n}e \\ \frac{m_1 m_2}{n(n-1)}(E - I) & \frac{m_2}{n}I + \frac{m_2(m_2-1)}{n(n-1)}(E - I) & \frac{m_2}{n}e \\ \frac{m_1}{n}e^{\mathrm{T}} & \frac{m_2}{n}e^{\mathrm{T}} & 1 \end{pmatrix}.$$

Since $\hat{Z} \in F_2$ it has the form $\hat{Z} = W\hat{U}W^{\mathrm{T}}$. It can be derived from the results in Wolkowicz and Zhao [41, Theorem 3.1.] that it has a two-dimensional nullspace, so $\hat{U} \succ 0$.

This puts us in a position to rewrite our relaxations as SDP having Slater points. In case of SDP_1, we only need to add the condition $\mathrm{diag}(Y_{12}) = 0$ to the constraints defining F_2. It can be expressed in terms of Z as $e_i^{\mathrm{T}} Z e_{n+i} = 0$ $i = 1, \ldots n$. Here e_i and e_{n+i} are the appropriate columns of the identity matrix of order $2n + 1$. We extend the matrix in the objective function by a row and column of zeros,

$$\hat{M} = \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{S}_{2n+1}$$

and get the following Slater feasible version of SDP_1 in matrices $U \in \mathcal{S}_{2n-1}$

$$
\begin{aligned}
(SDP_{1\mathrm{project}}) \; \min \; & \langle W^{\mathrm{T}} \hat{M} W, U \rangle \\
\text{s.t.} \; & e_i^{\mathrm{T}} (WUW^{\mathrm{T}}) e_{n+i} = 0, \quad i = 1, \ldots, n \\
& \mathrm{diag}(WUW^{\mathrm{T}}) = (WUW^{\mathrm{T}}) e_{2n+1} \\
& U_{2n-1,2n-1} = 1, \quad U \succeq 0.
\end{aligned}
$$

5.2 The projected Wolkowicz-Zhao relaxation and equivalent relaxations

The Slater feasible version of the SDP relaxation (7) is derived in [41] and further exploited in [32]. The matrix variable Z in (7) is of order $3n + 1$ and has the following structure

$$\bar{Z} = \begin{pmatrix} Y_1 & Y_{12} & Y_{13} & y_1 \\ Y_{21} & Y_2 & Y_{23} & y_2 \\ Y_{31} & Y_{32} & Y_3 & y_3 \\ y_1^{\mathrm{T}} & y_2^{\mathrm{T}} & y_3^{\mathrm{T}} & 1 \end{pmatrix}. \tag{33}$$

As before we can identify a nullspace common to all feasible matrices. In this case it is given by the columns of $\bar{T}$, see [41], where

$$\bar{T} = \begin{pmatrix} e & 0 & 0 & I \\ 0 & e & 0 & I \\ 0 & 0 & e & I \\ -m_1 & -m_2 & -m_3 & -e^{\mathrm{T}} \end{pmatrix}.$$

Note that this is a $(3n+1) \times (n+3)$ matrix. It has rank $n+2$, as the sum of the first three columns is equal to the sum of the last n columns, see also [41]. A basis of the orthogonal complement to $\bar{T}$ is given by

$$\bar{W} := \begin{pmatrix} V_n & 0 & \frac{m_1}{n} e \\ 0 & V_n & \frac{m_2}{n} e \\ -V_n & -V_n & \frac{m_3}{n} e \\ 0 & 0 & 1 \end{pmatrix}. \tag{34}$$

As before, we argue that feasible $\bar{Z}$ are of the form $\bar{Z} = \bar{W}U\bar{W}^{\mathrm{T}}$ with additional suitable constraints on $U \in \mathcal{S}_{2n-1}$. It is instructive to look at the last n columns of $\bar{Z}\bar{T}$ which, due to the block structure of $\bar{Z}$ translate into the following equations:

$$Y_1 + Y_{12} + Y_{13} = y_1 e^{\mathrm{T}}, \; Y_{21} + Y_2 + Y_{23} = y_2 e^{\mathrm{T}}, \; Y_{31} + Y_{32} + Y_3 = y_3 e^{\mathrm{T}}, \; y_1 + y_2 + y_3 = e. \tag{35}$$

Given Y_1, Y_2, y_1, y_2 and Y_{12} these equations uniquely determine y_3, Y_{13}, Y_{23}, and Y_3 and produce the n-dimensional part of the nullspace of $\bar{Z}$ given by the last n columns of $\bar{T}$. We can therefore drop this linear dependent part of $\bar{Z}$ without loosing any information. Mathematically, this is achieved as follows. Let us introduce the $(2n+1) \times (3n+1)$ matrix

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

It satisfies

$$W = P\bar{W}$$

and gives us a handle to relate the relaxations in matrices of order $3n+1$ to our models.

We recall the Slater feasible version of (7) from [41]:

$$
\begin{aligned}
\min \; & \mathrm{tr}(\bar{W}^{\mathrm{T}} \bar{M} \bar{W}) R \\
\text{s.t. } & e_i^{\mathrm{T}} (\bar{W} R \bar{W}^{\mathrm{T}}) e_{n+i} = 0, \; e_i^{\mathrm{T}} (\bar{W} R \bar{W}^{\mathrm{T}}) e_{2n+i} = 0, \\
& e_{n+i}^{\mathrm{T}} (\bar{W} R \bar{W}^{\mathrm{T}}) e_{2n+i} = 0, \qquad i = 1, \dots, n \\
& R_{2n-1, 2n-1} = 1, \; R \succeq 0, \; R \in \mathcal{S}_{2n-1},
\end{aligned}
\tag{36}
$$

where

$$\bar{M} := \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{S}_{3n+1}.$$

In [13] it is proven that the SDP relaxation (36) is equivalent to the QAP-based SDP relaxation with matrices of order $n^2 \times n^2$. Below we prove that (36) is equivalent to the here introduced SDP relaxation.

Theorem 12 *The SDP relaxation* (36) *is equivalent to* $SDP_{1\text{project}}$.

Proof. The feasible sets of the two problems are related by pre- and post-multiplication of the input matrices by P, for instance $P\bar{Z}P^T$ yields the $(2n+1) \times (2n+1)$ matrix variable of our relaxations. This operation basically removes the block of rows and columns corresponding to x_3. Both models contain the constraint $\mathrm{diag}(Y_{12}) = 0$. From (35) we conclude that

$$\mathrm{diag}(Y_1) + \mathrm{diag}(Y_{12}) + \mathrm{diag}(Y_{13}) = y_1.$$

Thus $\mathrm{diag}(Y_{13}) = 0$ in (36) is equivalent to $\mathrm{diag}(Y_1) = y_1$ in $SDP_{1\text{project}}$. Similarly, $\mathrm{diag}(Y_{23}) = 0$ is equivalent to $\mathrm{diag}(Y_2) = y_2$. The objective function is

nonzero only on the part of $\bar{Z}$ corresponding to Y_{12}. Thus the two problems are equivalent. $\square$

The SDP relaxation (36) with additional nonnegativity constraints $\bar{W}R\bar{W}^{\mathrm{T}} \geq 0$ is investigated by Pong et al. [32] on small graphs. The results show that the resulting bound outperforms other bounding approaches described in [32] for graphs with up to 41 vertices.

It is instructive to look at the nonnegativity condition $\bar{Z}_{ij} \geq 0$, where $\bar{Z}$ is has the block form (33), in connection with (35). From $Y_1 + Y_{21} + Y_{31} = ey_1^{\mathrm{T}}$ we get

$$Y_{2n+i,j} = y_j - Y_{i,j} - Y_{n+i,j} \geq 0,$$

which is (17). In a similar way we get from $\bar{Z} \geq 0$ all constraints (16) – (19). This is summarized as follows.

Theorem 13 *The SDP relaxation* (36) *with additional nonnegativity constraints is equivalent to* $SDP_{1\mathrm{project}}$ *with additional constraints* $WUW^{\mathrm{T}} \geq 0$ *and* (16)–(19).

Note that $SDP_{1\mathrm{project}}$ with additional constraints $WUW^{\mathrm{T}} \geq 0$, (16)–(19) is actually SDP_3 with additional nonnegativity constraints.

6 Symmetry reduction

It is possible to reduce the number of variables in SDP_2 when subsets S_1 and S_2 have the same cardinality. Therefore, let us suppose in this section that $m_1 = m_2$. Now, we apply the general theory of symmetry reduction to SDP_2 (see e.g., [11,12]) and obtain the following SDP relaxation:

$$
\begin{aligned}
\min \ & \langle A, Y_2 \rangle \\
\text{s.t. } & \mathrm{tr}(Y_1) = m_1, \ \mathrm{diag}(Y_2) = 0 \\
& \mathrm{tr}(JY_1) = m_1^2, \ \mathrm{tr}(JY_2) = m_1^2 \\
& Y = \begin{pmatrix} Y_1 & Y_2 \\ Y_2 & Y_1 \end{pmatrix}, \ \begin{pmatrix} Y & e \otimes \mathrm{diag}(Y_1) \\ (e \otimes \mathrm{diag}(Y_1))^{\mathrm{T}} & 1 \end{pmatrix} \succeq 0 \\
& Y \geq 0 \ \text{ on support.}
\end{aligned}
\tag{37}
$$

In order to obtain the SDP relaxation (37) one should exploit the fact that for $m_1 = m_2$ the matrix variable is of the following form

$$Y = I \otimes Y_1 + (J - I) \otimes Y_2,$$

for details see [12], Section 5.1. In particular, the above equation follows from (20), page 264 in [12], and the fact that the basis elements A_t $(t = 1, 2)$ in our case are I and $J - I$. Now, (37) follows by direct verification. In a case that a graph under consideration is highly symmetric, the size of the above SDP can be further reduced by block diagonalizing the data matrices, see e.g. [11,12].

In order to break symmetry we may assume without loss of generality that a vertex of the graph is not in the first partition set. This can be achieved by adding a constraint, which assigns zero value to the variable that corresponds to that vertex in the first set. In general, we can perform n such fixings and obtain n different valid bounds. Similar approach is exploited in e.g., [12,13]. If the graph under consideration has a nontrivial automorphism group, then there might be less than n different lower bounds. It is not difficult to show that each of the bounds obtained in the above described way dominates SDP_2. For the numerical results on the bounds after breaking symmetry see section 8.

7 Feasible solutions

Up to now our focus was on finding lower bounds on $\mathrm{OPT}_{\mathrm{MC}}$. A byproduct of all our relaxations is the vector $y = \binom{y_1}{y_2} \in \mathbf{R}^{2n}$ such that $y_1,\ y_2 \geq 0,\ e^{\mathrm{T}}y_1 = m_1, e^{\mathrm{T}}y_2 = m_2$ and possibly $y_1 + y_2 \leq e$. We now try to generate 0-1 solutions x_1 and x_2 with $x_1 + x_2 \leq e$, $e^{\mathrm{T}}x_1 = m_1$, $e^{\mathrm{T}}x_2 = m_2$ such that $x_1^{\mathrm{T}}Ax_2$ is small.

The hyperplane rounding idea can be applied in our setting. Feige and Langberg [16] propose random projections followed by randomized rounding (RPR^2) to obtain 0-1 vectors x_1 and x_2. In our case, we need to modify this approach to insure that x_1 and x_2 represent partition blocks of requested cardinalities.

It is also suggested to consider y_1 and y_2 and find the closest feasible 0-1 solution x_1 and x_2, which amounts to solving a simple transportation problem, see for instance [34].

It is also common practice to improve a given feasible solution by local exchange operations. In our situation we have the following obvious options. Fixing the set S_3 given by $x_3 := e - x_1 - x_2$, we apply the Kernighan-Lin local improvement [24] to S_1 and S_2 in order to (possibly) reduce the number of edges between S_1 and S_2. After that we fix S_1 and try swapping single vertices between S_2 and S_3 to reduce our objective function.

It turns out that carrying out these local improvement steps by cyclically fixing S_i until no more improvement is found leads to satisfactory feasible solutions. In fact, all the feasible solutions reported in the computational section were found by this simple heuristic.

8 Computational results

In this section we compare several SDP bounds on graphs from the literature. All bounds were computed on an Intel Xeon, E5-1620, 3.70 GHz with 32 GB memory. All relaxations were solved with SDPT3 [40].

We select the partition vector m such that $|m_1 - m_2| \leq 1$. For a given graph with n vertices, the optimal value of the min-cut is monotonically decreasing when m_3 is increasing. We select m_3 small enough so that $\mathrm{OPT}_{\mathrm{MC}} > 0$ and we

can also provide nontrivial (i.e. positive) lower bounds on $\mathrm{OPT_{MC}}$. Thus for given m we provide lower bounds (based on our relaxations) and also upper bounds (using the rounding heuristic from the previous section) on $\mathrm{OPT_{MC}}$. In case of a positive lower bound we also get a lower bound (of $m_3 + 1$) on the size of a strongly balanced ($|m_1 - m_2| \leq 1$) vertex separator. Finally, we also use our rounding heuristic and vary m_3 to actually find vertex separators, yielding also upper bounds for their cardinality. The results are summarized in Table 3.

Matrices can-xx and bcspwr03 are from the library Matrix Market [9], grid3dt matrices are 3D cubical meshes, gridt matrices are 2D triangular meshes, and Smallmesh is a 2D finite-element mesh. Lower bounds for the min-cut problem presented in the table are obtained by approximately solving SDP_4, i.e., by iteratively adding the most violated inequalities of type (16)–(19) and (20) to SDP_2. In particular, we perform at most 25 iterations and each iteration includes at most $2n$ the most violated valid constraints. It takes 59 minutes to compute grid3dt(5) and 170 minutes to compute grid3dt(6).

One can download our test instances from the following link:

https://sites.google.com/site/sotirovr/the-vertex-separator

Table 3 SDP_4 bounds for the min-cut and corresponding bounds for separators

name	n	$|E|$	m_1	m_2	m_3	lb min-cut	ub min-cut	lb separator	ub separator
Example 1	93	470	42	41	10	0.07	1	11	11
bcspwr03	118	179	58	57	3	0.56	1	4	5
Smallmesh	136	354	65	66	5	0.13	1	6	6
can-144	144	576	70	70	4	0.90	6	5	6
can-161	161	608	73	72	16	0.31	2	17	18
can-229	229	774	107	107	15	0.40	6	16	19
gridt(15)	120	315	56	56	8	0.29	4	9	11
gridt(17)	153	408	72	72	9	0.17	4	10	13
grid3dt(5)	125	604	54	53	18	0.54	2	19	19
grid3dt(6)	216	1115	95	95	26	0.28	4	27	30
grid3dt(7)	343	1854	159	158	26	0.60	22	27	37

All the instances in Table 3 have a lower bound of 0 for SDP_2 and SDP_3 while $SDP_4 > 0$. This is a clear indication of the superiority of SDP_4.

Table 4 provides further comparison of the our bounds. In particular, we list SDP_1, SDP_2, SDP_3, SDP_4, SDP_{fix}, and upper bounds for several graphs. SDP_{fix} bound is obtained after breaking symmetry as described in section 6. We choose m such that $SDP_2 > 0$ and $m_1 = m_2$. Thus, we evaluate all bounds obtained by fixing a single vertex and report the best among them. All bounds in Table 4 are rounded up to the closest integer. The results further verify the

quality of SDP_4, and also show that breaking symmetry improves SDP_2 but not significantly.

Table 4 Different bounds for the min-cut problem

name	n	m_1	m_2	m_3	SDP_1	SDP_2	SDP_3	SDP_4	SDP_{fix}	ub
Example 1	93	43	43	7	-7	1	2	5	1	5
can-161	161	78	78	5	3	3	7	21	4	33
gridt(15)	120	59	59	2	2	2	4	9	3	16
grid3dt(6)	216	102	102	12	-5	2	8	29	3	35

9 Conclusion

In this paper we derive several SDP relaxations for the min-cut problem and compare them with relaxations from the literature. Our SDP relaxations have matrix variables of order $2n$, while other SDP relaxations have matrix variables of order $3n$.

We prove that the eigenvalue bound from [23] equals the optimal value of the SDP relaxation from Theorem 5, with matrix variable of order $2n$. In [33] it is proven that the same eigenvalue bound is equal to the optimal solution of an SDP relaxation with matrix variable of order $3n$. Further, we prove that the SDP relaxation SDP_1 is equivalent to the SDP relaxation (36) from [41], see Theorem 12. We also prove that the SDP relaxation obtained after adding all remaining nonnegativity constraints to SDP_3 is equivalent to the strongest SDP relaxation from [41], see Theorem 13. Thus, we have shown that for the min-cut problem one should consider SDP relaxations with matrix variables of order $2n+1$ instead of traditionally considered SDP relaxations with matrices of order $3n+1$. Consequently, our strongest SDP relaxation SDP_4 also has a matrix variable of order $2n+1$ and $O(n^3)$ constraints. SDP_4 relaxation can be solved approximately by the cutting plane schema for graphs of medium size. The numerical results verify the superiority of SDP_4. We further exploit the resulting strong SDP bounds for the min-cut to obtain strong bounds on the size of the vertex separators.

Finally, our general framework for convexifying non-convex quadratic problems (see section 4) results with convex quadratic programming bounds that are cheap to compute. Since convex quadratic programming bounds played in the past crucial role in solving several non-convex problems to optimality, we plan to exploit their potential in our future research.

10 Proof of Theorem 5

We prove this theorem by providing feasible solutions to the primal and the dual SDP which have the same objective function value.

For the primal solution we take the optimizer $X = (x_1 \ x_2 \ x_3)$ from (5) with objective value $\mathrm{OPT_{HW}}$ and define $Y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{\mathrm{T}}$. Feasibility of X with respect to (4) shows that Y is feasible for the SDP. (For instance $\mathrm{tr} Y_1 = x_1^{\mathrm{T}} x_1 = m_1$ and $\mathrm{tr} J Y_1 = (e^{\mathrm{T}} x_1)^2 = m_1^2$).

Next we construct a dual solution with objective value $\mathrm{OPT_{HW}}$. Let

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ E_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \ E_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

With this notation, the primal constraints become

$$\langle E_1 \otimes I, Y \rangle = m_1, \ \langle E_2 \otimes I, Y \rangle = m_2, \ \langle E_{12} \otimes I, Y \rangle = 0.$$

$$\langle E_1 \otimes J, Y \rangle = m_1^2, \ \langle E_2 \otimes J, Y \rangle = m_2^2, \ \langle E_{12} \otimes J, Y \rangle = 2m_1 m_2,$$

and the objective function takes the form $\langle E_{12} \otimes (-\frac{1}{2} L), Y \rangle$. Thus the dual has the following form

$$\max m_1 \alpha_1 + m_2 \alpha_2 + m_1^2 \beta_1 + m_2^2 \beta_2 + 2 m_1 m_2 \beta_{12} \text{ such that}$$

$$E_{12} \otimes (-\frac{1}{2} L) - (\alpha_1 E_1 + \alpha_2 E_2 + \alpha_{12} E_{12}) \otimes I - (\beta_1 E_1 + \beta_2 E_2 + \beta_{12} E_{12}) \otimes J \succeq 0.$$

We recall that $Le = 0$, hence we can select an eigenvalue decomposition of L as $L = P \mathrm{Diag}(\lambda) P^{\mathrm{T}}$ where $P = (\frac{1}{\sqrt{n}} e \ V)$ with $V^{\mathrm{T}} V = I_{n-1}$, $V^{\mathrm{T}} e = 0$ and $\lambda = (0, \lambda_2, \ldots, \lambda_n)^{\mathrm{T}}$ contains the eigenvalues λ_i of L in nondecreasing order. The matrix P also diagonalizes J, $J = P \mathrm{Diag}((n, 0, \ldots, 0)) P^{\mathrm{T}}$. We use this to rewrite $E_{12} \otimes L$ as

$$E_{12} \otimes L = (I_2 E_{12} I_2) \otimes (P \Lambda P^{\mathrm{T}}) = (I_2 \otimes P)(E_{12} \otimes \Lambda)(I_2 \otimes P)^{\mathrm{T}}.$$

The other terms are rewritten similarly by pulling out $I_2 \otimes P$ from left and right. Let

$$s_1 := (n\beta_1 + \alpha_1, \alpha_1, \ldots, \alpha_1)^{\mathrm{T}}, \ s_2 := (n\beta_2 + \alpha_2, \alpha_2, \ldots, \alpha_2)^{\mathrm{T}},$$

$$s_{12} := (\alpha_{12} + n\beta_{12}, \alpha_{12} + \frac{1}{2}\lambda_2, \ldots, \alpha_{12} + \frac{1}{2}\lambda_n)^{\mathrm{T}}$$

be n−vectors and consider the block-diagonal $2n \times 2n$ matrix

$$S = \begin{pmatrix} -\mathrm{Diag}(s_1) & -\mathrm{Diag}(s_{12}) \\ -\mathrm{Diag}(s_{12}) & -\mathrm{Diag}(s_2) \end{pmatrix}.$$

Dual feasibility can be expressed as $(I_2 \otimes P) S (I_2 \otimes P)^{\mathrm{T}} \succeq 0$ which holds if and only if $S \succeq 0$. The special form of S shows, that $S \succeq 0$ breaks down to n semidefinitess conditions in matrices of order two:

$$\begin{pmatrix} -\alpha_1 - n\beta_1 & -\alpha_{12} - n\beta_{12} \\ -\alpha_{12} - n\beta_{12} & -\alpha_2 - n\beta_2 \end{pmatrix} \succeq 0, \tag{38}$$

$$\begin{pmatrix} -\alpha_1 & -\alpha_{12} - \frac{1}{2}\lambda_2 \\ -\alpha_{12} - \frac{1}{2}\lambda_2 & -\alpha_2 \end{pmatrix} \succeq 0, \quad \dots, \quad \begin{pmatrix} -\alpha_1 & -\alpha_{12} - \frac{1}{2}\lambda_n \\ -\alpha_{12} - \frac{1}{2}\lambda_n & -\alpha_n \end{pmatrix} \succeq 0.$$

$$(39)$$

We now select the dual variables in the following way. We set $\alpha_{12} := -\frac{1}{4}(\lambda_n + \lambda_2)$ which insures that

$$-\alpha_{12} - \frac{1}{2}\lambda_2 = \frac{1}{4}(\lambda_n - \lambda_2) \geq \dots \geq -\alpha_{12} - \frac{1}{2}\lambda_n = -\frac{1}{4}(\lambda_n - \lambda_2).$$

Thus all $n-1$ constraints in (39) are satisfied if we select $t < 0$ and set

$$\alpha_1 := \frac{1}{16t}(\lambda_n - \lambda_2) < 0, \ \ \alpha_2 := t(\lambda_n - \lambda_2) < 0.$$

Finally, setting

$$\beta_1 = -\frac{1}{n}\alpha_1, \ \ \beta_2 = -\frac{1}{n}\alpha_2, \ \ \beta_{12} = -\frac{1}{n}\alpha_{12}$$

insures that the matrix in (38) is 0 and hence (38) also holds. We now have a dual feasible solution, and we conclude the proof by selecting $t < 0$ in such a way that the objective function has value OPT_{HW}.

Let $D := m_1 m_2 (n - m_1)(n - m_2)$. We recall that

$$\text{OPT}_{\text{HW}} = (m_1 m_2 + \sqrt{D})\frac{\lambda_2}{2n} + (m_1 m_2 - \sqrt{D})\frac{\lambda_n}{2n} = -\frac{1}{2n}\sqrt{D}(\lambda_n - \lambda_2) + \frac{m_1 m_2}{2n}(\lambda_n + \lambda_2).$$

The dual solution defined above has value

$$m_1 \frac{\lambda_n - \lambda_2}{16t} + m_2 t(\lambda_n - \lambda_2) - \frac{\lambda_n - \lambda_2}{16tn}m_1^2 - \frac{\lambda_n - \lambda_2}{n}tm_2^2 + \frac{\lambda_n + \lambda_2}{4n}2m_1 m_2.$$

Comparing the coefficients of $\lambda_n - \lambda_2$ and $\lambda_n + \lambda_2$ we note that the values agree if

$$-\frac{\sqrt{D}}{2n} = \frac{t}{n}m_2(n - m_2) + \frac{1}{16tn}m_1(n - m_1).$$

This equation holds for

$$t = -\frac{1}{4}\sqrt{\frac{m_1(n - m_1)}{m_2(n - m_2)}}.$$

□

References

1. Anstreicher, K.M., Wolkowicz, H.: On Lagrangian relaxation of quadratic matrix constraints. SIAM Journal on Matrix Analysis and Applications, 22(1), 41–55 (2000)
2. Anstreicher, K.M., Brixius, N., Goux, J.-P., Linderoth, J.: Solving large quadratic assignment problems on computational grids. Math. Program. Ser. B, 91, 563–588 (2002)
3. Anstreicher, K.M., Brixius, N.W.: A new bound for the quadratic assignment problem based on convex quadratic programming. Math. Program. Ser. A, 89(3), 341-357 (2001)
4. Armbruster, M., Helmberg, C., Fügenschuh, M., Martin, A.: LP and SDP branch-and-cut algorithms for the minimum graph bisection problem: a computational comparison. Mathematical Programming Computation, 4(3), 275–306 (2012)
5. Bhatt, S.N., Leighton, F.T.: A framework for solving VLSI graph layout problems. J. Computer System Sci., 28, 300-343 (1984)
6. Balas, E., de Souza, C.C.: The vertex separator problem: a polyhedral investigation. Math. Program. Ser. B, 103, 583–608 (2005)
7. Biha, M.D., Meurs, M.J.: An exact algorithm for solving the vertex separator problem. J. Global Optim., 49, 452–434, (2011)
8. Billionnet, A., Elloumi, S., Plateau, M.: Improving the performance of standard solvers for quadratic $0-1$ programs by a tight convex reformulation: The QCR method. Discrete Appl. Math., 157(6), 1185–1197 (2009)
9. Boisvert, R., Pozo, R., Remington, K., Barrett, R., Dongarra. J.: Matrix Market: A web resource for test matrix collections. In R. Boisvert, ed. The Quality of Numerical Software: Assessment and Enhancement, Chapman and Hall, London, 125–137 (1997)
10. Donath, W.E., Hoffman, A.J.: Lower bounds for the partitioning of graphs. IBM J. of Research and Developement, 17, 420–425 (1973)
11. De Klerk, E., Sotirov, R.: Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. Math. Program. Ser. A, 122(2), 225–246 (2010)
12. De Klerk, E., Pasechnik, D.V., Sotirov, R., Dobre, C.: On semidefinite programming relaxations of maximum k-section. Math. Program. Ser. B, 136(2), 253–278 (2012)
13. Dam van, E.R., Sotirov, R.: On bounding the bandwidth of graphs with symmetry. INFORMS J. on Comp. 27, 75–88 (2014)
14. Falkner, J., Rendl, F., Wolkowicz, H.: A computational study of graph partitioning. Math. Program. Ser. B, 66, 211–239 (1994)
15. Faye, A., Roupin, F.: Partial Lagrangian relaxation for general quadratic programming. 4OR, 5, 75–88 (2007)
16. Feige, U., Langberg, M.: The RPR^2 rounding technique for semidefinite programs. Journal of Algorithms, 60, 1–23 (2006)
17. Fu, B., Oprisan, S.A., Xu, L.: Multi-directional width-bounded geometric separator and protein folding. ISAAC, 995–1006 (2005)
18. Garey, M.R., Johnson, D.S.: Computers and intractability: A guide to the theory of NP-completeness. Freeman, San Francisco (1979)
19. Graham, A.: Kronecker products and matrix calculus with applications, Ellis Horwood Ltd., Chichester (1981)
20. Hager, W.W., Hungerford, J.T.: A continuous quadratic programming formulation of the vertex separator problem. European Journal of Operational Research, 240, 328–337 (2015)
21. Hager, W.W., Hungerford, J.T., Safro, I.: A multilevel bilinear programming algorithm for the vertex separator problem. ArXiv:1410.4885v1, Ocotober (2014)
22. Hammer, P.L., Rubin, A.A.: Some remarks on quadratic programming with 0-1 variables. RAIRO, 3, 67–79 (1970)
23. Helmberg, C., Mohar, B., Poljak, S., Rendl, F.: A spectral approach to bandwidth and separator problems in graphs. Linear and Multilinear Algebra, 39, 73–90 (1995)
24. Kernighan, B. W., Lin, S.: An efficient heuristic procedure for partitioning graphs. Bell System Technical Journal, 49(2), 291-307 (1970)
25. Leighton, F.T.: Complexity issues in VLSI: Optimal layout for the shuffle-exchange graph and other networks. MIT Press, Cambridge, MA (1983)

26. Lemarechal, C., Oustry, F.: Semidefinite relaxations and Lagrangian duality with application to combinatorial optimization. Research Report 3710, INRIA, Rhones-Alpes, (1999)
27. Lipton, R.J., Tarjan, R.E.: A separator theorem for planar graphs. SIAM Journal on Applied Math., 36, 177–189 (1979)
28. Lipton, R.J., Tarjan, R.E.: Applications of a planar separator theorem. SIAM J. Comput., 9, 615-627 (1980)
29. Lipton, R.J., Rose, D.J., Tarjan, R.E.: Generalized nested dissection. SIAM Journal on Numerical Analysis, 16(2), 346–358 (1979)
30. Miller, G.L., Teng, S.-H., Thurstons, W., Vavasis, S.A.: Geometric separators for finite-element meshes. SIAM J. Sci. Comput, 19(2), 364-386 (1998)
31. Padberg, M.W.: The boolean quadric polytope: Some characteristics, facets and relatives. Math. Program., 45, 139–172 (1989)
32. Pong, T.K., Sun, H., Wang, N., Wolkowicz, H.: Eigenvalue, quadratic programming, and semidefinite programming relaxations for a cut minimization problem. Comput. Optim. Appl., 63(2), 333-364 (2016)
33. Povh, J., Rendl, F.: A copositive programming approach to graph partitioning. SIAM J. Optim., 18, 223–241 (2007)
34. Rendl, F., Wolkowicz, H.: A projection technique for partitioning the nodes of a graph. Ann. Oper. Res., 58(3), 155–179 (1995)
35. Sotirov, R.: An efficient semidefinite programming relaxation for the graph partition problem. INFORMS J. Comput., 26(1), 16–30 (2014)
36. Sherali, H.D., Adams, W.P.: 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)
37. Shor, N.Z.: Quadratic Optimization Problems. Soviet Journal of Computer and Systems Sciences, 25, 1-11 (1987)
38. Shor, N.Z.: Dual estimates in multiextremal problems. J. Global Optim., 2, 411–418 (1992)
39. de Souza, C.C., Balas, E.: The vertex separator problem: algorithms and computations. Math. Program., Ser. B, 103, 609–631 (2005)
40. Tütüncü, R.H., Toh, K.C., Todd, M.J.: Solving semidefinite-quadratic-linear programs using SDPT3. Math. Program., Ser. B, 95, 189–217 (2003)
41. Wolkowicz, W., Zhao, Q.: Semidefinite programming relaxations for the graph partitioning problem. Discrete Appl. Math., 96/97, 461–479 (1999)
42. Zhao, Q., Karisch, S.E., Rendl, F., Wolkowicz, H.: Semidefinite programming relaxations for the quadratic assignment problem. J. Comb. Optim., 2, 71–109 (1998)