Annals of Operations Research 58(1995)155–179

A projection technique for partitioning the nodes of a graph

Franz Rendl
Technische Universität Graz, Institut für Mathematik, Kopernikusgasse 24, A-8010 Graz, Austria

Henry Wolkowicz
Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

Let $G = (N, E)$ be an undirected graph. We present several new techniques for partitioning the node set N into k disjoint subsets of specified sizes. These techniques involve eigenvalue bounds and tools from continuous optimization. Comparisons with examples taken from the literature show these techniques to be very successful.

1. Introduction

Let $G = (N, E)$ be an undirected graph with node set $N = \{1,\ldots,n\}$, edge set E and edge weights $\{c(e) : e \in E\}$. (If the graph is unweighted, then we assume $c(e) = 1$ for $e \in E$.) A common problem in circuit board and microchip design, computer program segmentation, floor planning and other layout problems is to partition the node set N into k disjoint subsets $S_1,\ldots,S_k$ of specified sizes $m_1 \geq m_2 \geq \ldots \geq m_k$, $\sum_{j=1}^{k} m_j = n$, so as to minimize the weight of edges connecting nodes in distinct subsets of the partition. We refer to an edge, which connects nodes in distinct subsets of the partition, as being cut by the partition. A recent survey on the graph partitioning problem and further related problems is contained in [21].

There are several possibilities to model graph partition problems. The *polyhedral approach* relies on linear programming. In this approach, variables are introduced on the edge set E, leading to a linear objective function. The hard part consists in finding linear descriptions of those edge sets which correspond to (feasible) partitions, see e.g. [6]. In the present paper, we do not pursue this approach but use a model based on the node set of the graph. In this case, the description of partitions is straightforward, but the objective function will turn out to be quadratic.

The main contribution of this paper will be a theoretically and practically efficient procedure to deal with the quadratic objective function under suitable relaxations of the feasible region.

Introducing variables on the node set, we obtain a quadratic 0–1 program as follows, see e.g. [3]:

Let $X \in \mathbb{R}^{n \times k}$ with the columns

$$x_j = (x_{1j}\, x_{2j} \ldots x_{nj})^{\mathrm{T}}$$

being the *characteristic vector* for the set S_j, $j = 1,\ldots,k$, i.e.

$$x_{ij} = \begin{cases} 1 & \text{if } i \in S_j, \\ 0 & \text{if } i \notin S_j. \end{cases}$$

We denote by $A = (a_{ij})$ the *weighted adjacency matrix* for G, i.e. a_{ij} denotes the weight of the edge connecting nodes i and j, and otherwise $a_{ij} = 0$. Moreover, we assume without loss of generality that G has no loops, thus $a_{ii} = 0$, $i = 1,\ldots,n$. Since G is undirected, the adjacency matrix is symmetric. We observe that

$$\frac{1}{2}\sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs} x_{rj} x_{sj} = \frac{1}{2} x_j^{\mathrm{T}} A x_j \tag{1}$$

is the total weight of edges with both endpoints in S_j. Moreover, the nonnegative integer matrix X defines a partition if and only if its elements satisfy the transportation problem constraints

$$\begin{aligned} \textstyle\sum_{i=1}^{n} x_{ij} &= m_j, \quad & j &= 1,\ldots,k, \\ \textstyle\sum_{j=1}^{k} x_{ij} &= 1, & i &= 1,\ldots,n. \end{aligned}$$

To minimize the weight of edges cut in a partition, we can maximize the weight of edges not cut. Our problem becomes

$$\textbf{(P)} \qquad \begin{aligned} &\text{maximize} && \tfrac{1}{2}\, tr\, X^{\mathrm{T}} A X \\ &\text{subject to} && X u_k = u_n, \\ &&& X^{\mathrm{T}} u_n = m, \\ &&& X \text{ is a } 0,1 \text{ matrix.} \end{aligned}$$

Throughout, tr denotes trace, $u_j \in \mathbb{R}^j$ is the vector of ones, and $m = (m_1 \ldots, m_k)^{\mathrm{T}}$ is the ordered vector of specified sizes. Moreover, $M := \operatorname{diag}(m) \in \mathbb{R}^{k \times k}$. We will use the Frobenius norm for matrices, i.e. $\|A\|^2 = tr\, A^{\mathrm{T}} A$.

For a given partition X, we say that $T = (t_{ij}) \in \mathbb{R}^{n \times n}$ represents the partition X if

$$t_{ij} = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ belong to the same subset,} \\ 0 & \text{otherwise.} \end{cases}$$

Then each partition is identified with a matrix T. Note also that by definition

$$T = XX^{\mathrm{T}} \quad \text{and} \quad M = X^{\mathrm{T}}X.$$

Thus,

$$\|T\|^2 = \|M\|^2.$$

Note that

$$\begin{aligned} \|A - T\|^2 &= \|A\|^2 + \|T\|^2 - 2\, tr\, AT \\ &= \|A\|^2 + \|M\|^2 - 2\, tr\, AT, \end{aligned}$$

and

$$tr\, AT = tr\, AXX^{\mathrm{T}} = tr\, X^{\mathrm{T}}AX.$$

Therefore, an equivalent formulation to (**P**) is the best matrix approximation problem

(**F**) $$\min\{\|A - T\| : T \text{ represents a partition}\}.$$

The formulation (P) is very similar to the quadratic assignment problem, QAP. Continuous optimization techniques are employed in [16, 17, 28] to find bounds for the QAP. In particular, a projection technique is used in [17] to eliminate the linear constraints on the row and column sums of X. An iterative improvement of QAP bounds, based on "reductions", is presented in [28]. In this paper, we extend the continuous optimization techniques from [17, 28] to the graph partitioning problem.

The paper is organized as follows. This section is concluded with an overview of existing results for the graph partitioning problem that are relevant in the present context.

In section 2, we formulate the main mathematical tools to derive our bounds. In section 3, we extend the projection technique from [17] to (P) to get an equivalent program (EP), where the constraints on the row and column sums of X are implicitly satisfied. The program (EP) is the key to several new bounds. These will be presented in section 4. We also discuss several special cases where the bound can be further strengthened.

In section 5, we exploit the concept of diagonal perturbations to improve the bounds. We use an iterative improvement technique to find the best perturbations. Section 6 shows how to find feasible solutions using information from the bounding techniques.

We conclude with some numerical experiments in section 7, both on published data and on randomly generated graphs. We are able to solve smaller problems ($n \le 20$) to optimality in many cases using the new bounds. In general, the best upper bounds proposed in this paper constitute a substantial improvement over the existing bounding rules. A substantial computational study using the bounds developed in this paper is contained in [11].

1.1. OVERVIEW OF PREVIOUS RELATED RESEARCH

In [2,10], spectral information of A is used to bound the objective function of (P). Boppana [4] considers graph bisection, i.e. $m_1 = m_2 = n/2$, and improves the eigenvalue bound from [10] for this special case. The papers [5,29] describe branch and bound approaches to solve the partitioning problem in the case $k = 2$ and for general weighted graphs. Both methods seems to work only for extremely thin graphs (average degree not more than 4).

Recently, interior point methods have also been proposed to obtain bounds for graph bisection, see [19,20]. Judging from these paper, it appears that the eigenvalue approach of the present paper is superior to interior point techniques applied to graph bisection.

A different type of partitioning problem consists of separating the node set into just two sets (of arbitrary cardinalities), so as to maximize the total weight of the edges cut (max-cut problem). In [8,9,22,25,25], eigenvalue related techniques for this problem are analysed. The absence of cardinality constraints seems to make this problem easier to handle than the general partitioning problem studied in this paper.

Yet another type of partitioning problem consists of separating a graph by removing vertices (rather than cutting edges). Eigenvalue related techniques to derive bounds on the minimum size of "vertex separators" are investigated in [18,27]. The latter paper also contains computational experiments on bisection problems arising from real-world applications.

Several articles are devoted to finding "good" partitions using spectral information from A. In [1], the formulation (F) is used and a transportation problem is proposed to find a feasible X. The transportation costs are determined by the (pairwise orthogonal and normed) eigenvectors of A, corresponding to the k largest eigenvalues. The formulation (P) is used in [3]. Therein, A is shifted by a diagonal matrix D so that $A + D$ is positive semidefinite. Then the Cholesky decomposition of $A + D$ is used to improve a given partition.

Finally, a survey on various aspects of the graph partitioning problem and further references are contained in chapter 6 of [21].

EXAMPLE 1

We will illustrate our results, as we progress through the paper, on the following example from [10], see also [1]. The graph is unweighted and has 20 nodes. We partition it into two equal parts, i.e. $m_1 = m_2 = 10$. The edges are represented in table 1. We point out that the cardinality $|E| = 51$. This provides a trivial upper bound on the number of edges not cut by any partition.

2. Preliminaries

We first present some notation and basic results. We let $\mathcal{O}_{k,l}$ (or $\mathcal{O}$ when the meaning is clear) denote the set of $k \times l$ *orthogonal matrices*, i.e. $Q \in \mathrm{O}_{k,l}$ if $Q^{\mathrm{T}}Q = I$.

Table 1

Edge set of example 1.

Node	Connections to
1	7, 12, 13, 14, 15, 16, 17
2	12, 17, 18, 20
3	5, 11, 13, 14, 18, 19, 20
4	6, 9
5	7, 9, 10, 12, 16, 19
6	16, 18, 20
7	8, 9, 11, 16
8	15, 18
9	11, 15, 19
11	14, 17, 18, 20
12	14
13	18, 20
14	16, 18, 20
16	18
17	18
18	20

The vector of ones is $u_l = (1,\ldots,1)^{\mathrm{T}} \in \mathbb{R}^l$. If the meaning is clear, we omit indicating the subscript to indicate the dimension. $r(K) = Ku_l$ is the vector of row sums of a $k \times l$ matrix K, while $s(K) = u_k^{\mathrm{T}} K u_l$ is the sum of all the elements of K. We denote by $m = (m_1,\ldots,m_k)$ the vector of specified sizes of the partition and assume without loss of generality that m is ordered nonincreasingly. We let the (positive) diagonal matrix $M = \mathrm{diag}(m)$, while for a given matrix M, $\mathrm{diag}(M)$ denotes the vector formed from the diagonal of M. We denote the eigenvalues of a symmetric $n \times n$ matrix A by

$$\lambda_1(A) \geq \ldots \geq \lambda_n(A).$$

The set of matrices satisfying the transportation constraints of (P) forms an affine space of matrices and is denoted by $\mathscr{E}$:

$$\mathscr{E} = \{X \in \mathbb{R}^{n\times k} : Xu_k = u_n, X^{\mathrm{T}}u_n = m\}. \tag{2}$$

The set of nonnegative matrices is

$$\mathscr{N} = \{X \in \mathbb{R}^{n\times k} : X \geq 0 \ \text{ elementwise}\}.$$

The feasible set of matrices for (P) is

$$F = \{X \in \mathbb{R}^{n\times k} : x_{ij} \ \text{ is 0 or 1}\} \cap \mathscr{E}. \tag{3}$$

LEMMA 2.1

The feasible set satisfies

$$F = \mathscr{E} \cap \mathscr{N} \cap \{X \in \mathbb{R}^{n \times k} : X^{\mathrm{T}} X = M\}$$
$$= \mathscr{E} \cap \mathscr{N} \cap \{X \in \mathbb{R}^{n \times k} : tr\, X^{\mathrm{T}} X = n\}.$$

Proof

We prove only that the second set on the right is contained in F. (The rest is clear from the definitions.) We see that

$$X \in \mathscr{N}, X u_k = u_n \Rightarrow 0 \leq x_{ij} \leq 1 \Rightarrow x_{ij}^2 \leq x_{ij}.$$

Hence, $n = tr X^{\mathrm{T}} X \leq s(X) = u_n^{\mathrm{T}} X u_k = n$ implies $x_{ij}^2 = x_{ij}$, $\forall i, j$. □

The above lemma suggests several relaxations of the constraints of the graph partitioning problem (P). First, the relaxation to $X \in \mathscr{E} \cap \mathscr{N}$ corresponds to Quadratic Programming. Note that A is in general indefinite, since G has no loops, so the (global) maximum is difficult to find. Relaxing to $X \in \{X \in \mathrm{R}^{n \times k}: X^{\mathrm{T}} X = M\}$ leads to the eigenvalue bound derived in [10], see also theorem 2.2 below.

In this paper, we strengthen the eigenvalue bound by maximizing over

$$X \in \{X \in \mathbb{R}^{n \times k} : X^{\mathrm{T}} X = M\} \cap \mathscr{E}.$$

One of the main tools is the following basic result from matrix analysis, see e.g. [28].

THEOREM 2.1

Let $A = A^{\mathrm{T}}$ be $n \times n$, $B = B^{\mathrm{T}}$ be $k \times k$ and suppose $k \leq n$. Then

$$\max\{tr\, AXBX^{\mathrm{T}} : X^{\mathrm{T}} X = I_k\}$$
$$= \max\left\{ \sum_{i=1}^{k} \lambda_i(B) \lambda_{\phi(i)}(A) : \phi : \{1, \ldots, k\} \mapsto N \text{ injective} \right\}.$$

Suppose in addition that the maximum on the right is attained for injection ϕ. Then the maximum on the left is attained for $X = U_1 U_2^{\mathrm{T}}$, where

$$U_1^{\mathrm{T}} A U_1 = diag(\lambda_{\phi(1)}(A), \ldots, \lambda_{\phi(k)}(A)), \ \ U_1^{\mathrm{T}} U_1 = I_k,$$

and

$$U_2^{\mathrm{T}} B U_2 = diag(\lambda_1(B), \ldots, \lambda_k(B)), \ \ U_2^{\mathrm{T}} U_2 = I_k.$$

Remark

The above formulation is not the one commonly seen in the literature. The classical "Hoffman–Wielandt" inequality is formulated for *normal* matrices with $k = n$ and the above optimization is replaced by an inequality. Ky Fan's theorem allows $k < n$, but $B = I_k$ is assumed. Finally, John von Neumann considers a more general version with A and B arbitrary and a bilinear objective function. A recent summary on all the above-mentioned variations of the Hoffman–Wielandt inequality is contained in [24]. The present formulation seems to be the "right" generalization of the classical Hoffman–Wielandt inequality in the context of graph optimization problems, see also [18].

The proof of this theorem follows easily from e.g. [28]. A full proof is also contained in [18]. In the present application, the matrix B will always be positive semidefinite, so that the best injection ϕ can be given explicitly, using "maximal scalar products".

COROLLARY 2.1

Under the conditions of theorem 2.1, assume that B is positive semidefinite. Then

$$\max\{tr\, AXBX^{\mathrm{T}} : X^{\mathrm{T}}X = I_k\} = \sum_{i=1}^{k} \lambda_i(B)\lambda_i(A).$$

We conclude this section with an eigenvalue based upper bound on $|E_{uncut}|$, the weight of edges not cut by any partition. This bound was proposed by Donath and Hoffman in 1973 and is the starting point of the present paper. The validity of this result follows immediately from the above corollary. This derivation is different from the one contained in [10].

THEOREM 2.2 [10]

Let A and m describe a graph partitioning problem. Then

$$\begin{aligned} |E_{uncut}| &\le \max\left\{\tfrac{1}{2}\, tr\, X^{\mathrm{T}}AX : X^{\mathrm{T}}X = M\right\} \\ &= \max\left\{\tfrac{1}{2}\, tr\, MY^{\mathrm{T}}AY : Y^{\mathrm{T}}Y = I_k\right\} = \tfrac{1}{2}\sum_{j=1}^{k} m_j\lambda_j(A). \end{aligned} \tag{4}$$

Proof

Feasible partitions X clearly satisfy $X^{\mathrm{T}}X = M$, so the inequality is obvious. The first equality follows by setting $X = YM^{1/2}$, and the second equality follows from the corollary. Here, we also use the nonincreasing order of m. □

EXAMPLE 1 (continued)

In table 2 and figure 1, we summarize the various bounds for example 1 in detail. Table 2 contains the relevant eigenvalue information and the upper bounds. Since $\lambda_1(A) = 6.0429$ and $\lambda_2(A) = 3.1375$, we get an upper bound of 45.9019 using theorem 2.2. Thus, no partition leaves more than 45 edges uncut. The corresponding maximizers X are represented graphically in figure 1. Ideally, half the components

Table 2

Upper bounds for $|E_{uncut}|$ in example 1. The eigenvalues given are the two largest of A for theorem 2.2 and of $\hat{A}$ for the remaining bounds.

	Bound	Eigenvalues	
Theorem 2.2	45.9019	6.0429	3.1375
Corollary 4.1	42.1269	3.3254	2.1946
Lemma 5.3	38.5516	2.6103	2.6103

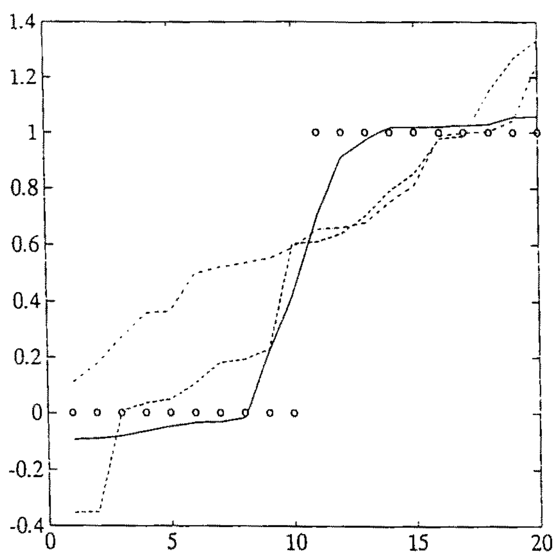

Figure 1. Sorted eigenvector components for the various bounds of table 2. Ideally, half the components should be 0, the other half equal to 1. The dash–dotted line corresponds to theorem 2.2, the dashed line to corollary 4.1, and the solid line to lemma 5.5. Note that the eigenvector corresponding to lemma 5.3 is quite close to a partitioning vector.

of an eigenvector should be 0 and the other half be 1. The sorted eigenvector components for $\lambda_1(A)$ are plotted as a dash–dotted line in figure 1. Since about half of its components are around 0.5, we conclude that the maximizer X does not give any clue on how to obtain a good feasible partition from X.

3. Projection of (P)

We now project the feasible set of the problem (P) onto the linear manifold defined by the contraints $\mathscr{E}$. We do this by eliminating the constraints $\mathscr{E}$ while simultaneously maintaining the trace structure of the objective function and the orthogonality properties of the constraints. This structure allows us to still apply the eigenvalue bounds. This extends the projection technique in [17] for the QAP.

We define

$$\bar{m} := M^{1/2}u_k = \left(\sqrt{m_1}, \dots, \sqrt{m_k}\right)^{\mathrm{T}} = M^{-1/2}m.$$

We let P and Q be orthogonal matrices with

$$P = \left[\frac{u_n}{\sqrt{n}}\ V\right]; \quad Q = \left[\frac{\bar{m}}{\sqrt{n}}\ W\right].$$

Clearly, both V and W are not uniquely determined. The characterizing condition for V and W are

$$V^{\mathrm{T}}V = I_{n-1}, V^{\mathrm{T}}u_n = 0, \text{ and similarly } W^{\mathrm{T}}W = I_{k-1}, W^{\mathrm{T}}\bar{m} = 0.$$

The results derived in the subsequent sections will not depend on the specific choice of the representation for V and W. Note that VV^{T} is the orthogonal projection on $\{u_n\}^{\perp}$, while WW^{T} is the orthogonal projection on $\{\bar{m}\}^{\perp}$. The following lemma contains a parametrization of the set $\mathscr{E}$ that is crucial in deriving improved bounds for graph partitioning.

LEMMA 3.1

Let P, Q and M be as above. Suppose X is $n \times k$, Z is $(n-1) \times (k-1)$, and X and Z are related by

$$X = P\begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix} Q^{\mathrm{T}}M^{1/2}. \tag{5}$$

Then

(a) $X \in \mathscr{E}$.

(b) $X \in \mathscr{N} \Leftrightarrow VZW^{\mathrm{T}} \geq -\frac{1}{n}u_n\bar{m}^{\mathrm{T}}$.

(c) $X^{\mathrm{T}}X = M \Leftrightarrow Z \in \mathscr{O}_{(n-1)\times(k-1)}$.

Conversely, if $X \in \mathscr{E}$, then there exists Z such that the representation (5) holds.

Proof

First note that expanding (5) yields

$$X = \frac{1}{n} u_n u_k^{\mathrm{T}} M + VZW^{\mathrm{T}} M^{1/2}. \tag{6}$$

Now observe that, since $V^{\mathrm{T}} u_n = 0$, we obtain

$$X^{\mathrm{T}} u_n = \frac{1}{n} M u_k u_n^{\mathrm{T}} u_n + M^{1/2} WZ^{\mathrm{T}} V^{\mathrm{T}} u_n = M u_k = m.$$

Similarly,

$$X u_n = \frac{1}{n} u_n u_k^{\mathrm{T}} M u_k + VZW^{\mathrm{T}} M^{1/2} u_k = u_n,$$

because $W^{\mathrm{T}} M^{1/2} u_k = 0$. Thus (a) is proved. By (6), we can write

$$X = \frac{1}{n} u_n \overline{m}^{\mathrm{T}} M^{1/2} + VZW^{\mathrm{T}} M^{1/2}.$$

Thus

$$X \in \mathcal{N} \Leftrightarrow VZW^{\mathrm{T}} \geq -\frac{1}{n} u_n \overline{m}^{\mathrm{T}},$$

because multiplying with the positive diagonal matrix $M^{-1/2}$ does not change the inequality. Finally, note that

$$X^{\mathrm{T}} X = M \Leftrightarrow Q \begin{bmatrix} 1 & 0 \\ 0 & Z^{\mathrm{T}} \end{bmatrix} P^{\mathrm{T}} P \begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix} Q^{\mathrm{T}} = I \Leftrightarrow Z \in \mathcal{O},$$

because P and Q are orthogonal. To conclude, suppose $X \in \mathcal{E}$. Then

$$P^{\mathrm{T}} X M^{-1/2} Q = \begin{bmatrix} 1 & 0 \\ 0 & V^{\mathrm{T}} X M^{-1/2} W \end{bmatrix}.$$

□

If we substitute the parametrization (6) for X in the objective function of problem (P), we obtain an equivalent formulation of the partitioning problem in the Z-space.

$$\begin{aligned} tr\, X^{\mathrm{T}} A X &= tr\Big(\tfrac{1}{n} M u_k u_n^{\mathrm{T}} + M^{1/2} WZ^{\mathrm{T}} V^{\mathrm{T}}\Big) A \Big(\tfrac{1}{n} u_n u_k^{\mathrm{T}} M + VZW^{\mathrm{T}} M^{1/2}\Big) \\ &= tr\Big[\tfrac{1}{n^2} (u_n^{\mathrm{T}} A u_n)(u_k^{\mathrm{T}} M^2 u_k) + \tfrac{2}{n} Z^{\mathrm{T}} V^{\mathrm{T}} A u_n u_k^{\mathrm{T}} M^{3/2} W \\ &\qquad + (W^{\mathrm{T}} M W) Z^{\mathrm{T}} (V^{\mathrm{T}} A V) Z\Big] \\ &= \tfrac{1}{n^2} s(A) s(M^2) + tr\Big[\hat{M} Z^{\mathrm{T}} \hat{A} Z + \tfrac{2}{n} Z^{\mathrm{T}} V^{\mathrm{T}} r(A) r^{\mathrm{T}}(M^{3/2}) W\Big]. \end{aligned}$$

Translating the feasibility conditions on X to the appropriate conditions on Z, derived in the lemma, we obtain the following program (EP) in the variable $Z \in \mathbb{R}^{(n-1)\times(k-1)}$, which is equivalent to the original graph partitioning problem (P):

$$\textbf{(EP)} \quad \text{maximize} \quad tr\left\{\tfrac{1}{2}\hat{M}Z^{\mathrm{T}}\hat{A}Z + \tfrac{1}{n}Z^{\mathrm{T}}V^{\mathrm{T}}r(A)r^{\mathrm{T}}(M^{3/2})W\right\} + \tfrac{1}{2n^2}s(A)s(M^2)$$

$$\text{subject to} \quad Z^{\mathrm{T}}Z = I,$$

$$VZW^{\mathrm{T}} \geq -\tfrac{1}{n}u_n\overline{m}^{\mathrm{T}}.$$

Throughout this paper, we use the notation

$$\hat{A} = V^{\mathrm{T}}AV, \quad \hat{M} = W^{\mathrm{T}}MW.$$

Summarizing, we have shown that (P) is indeed equivalent to (EP).

THEOREM 3.1

Suppose X and Z are related by (5). Then Z solves (EP) $\Leftrightarrow X$ solves (P). □

We conclude this section with a more technical aspect. The projection technique was based on the decomposition given in lemma 3.1. This technique can be generalized using the singular value decomposition of X. Suppose that $U \in \mathcal{O}_{k,l}$ and $V \in \mathcal{O}_{n,l}$ are orthogonal matrices satisfying

$$XU = V\Sigma, \quad X^{\mathrm{T}}V = U\Sigma,$$

for some matrix $\Sigma \in \mathbb{R}^{l,l}$. Let P and Q be square orthogonal matrices with $P = [V\overline{V}]$, $Q = [U\overline{U}]$. Then

$$XX^{\mathrm{T}}V = XU\Sigma = V\Sigma^2,$$

$$X^{\mathrm{T}}XU = X^{\mathrm{T}}V\Sigma = U\Sigma^2$$

and

$$P^{\mathrm{T}}XQ = \begin{bmatrix} V^{\mathrm{T}}XU & V^{\mathrm{T}}X\overline{U} \\ \overline{V}^{\mathrm{T}}XU & \overline{V}^{\mathrm{T}}X\overline{U} \end{bmatrix}$$

or

$$X = P\begin{bmatrix} \Sigma & 0 \\ 0 & \overline{V}^{\mathrm{T}}X\overline{U} \end{bmatrix}Q^{\mathrm{T}}.$$

We therefore get a decomposition of X which becomes particularly nice if X is orthogonal, for then both Σ and $\overline{V}^{\mathrm{T}}X\overline{V}$ must also be orthogonal.

4. Bounds for (P)

Using the equivalent program (EP) instead of the original problem (P), we obtain new bounds for (P). First note that due to the elimination of the constraint $\mathscr{E}$, we have a linear term in the objective function of (EP), while (P) has a purely quadratic objective function. Maximizing (EP) over orthogonal Z is in general difficult, because the linear term does not allow a direct application of the bound from theorem 2.2. Therefore, we treat the quadratic and linear part separately. The quadratic part is bounded using theorem 2.2, while maximizing a linear function over the constraints (P) is equivalent to a (bipartite) transportation problem, and so can be handled directly.

THEOREM 4.1

Let A and m describe a graph partitioning problem. Assume that the nodes are numbered such that $r(A) = (r_1(A), \ldots, r_n(A))^T$ is in nonincreasing order and define $p_0 := 0$ and the partial sums $p_j := \sum_{i=1}^{j} m_i$ and $R_j(A) := \sum_{i=p_{j-1}+1}^{p_j} r_i(A)$, $j = 1, \ldots, k$. Then

$$|E_{uncut}| \leq \frac{1}{2} \sum_{j=1}^{k-1} \lambda_j(\hat{A})\lambda_j(\hat{M}) + \frac{1}{n} \sum_{j=1}^{k} R_j(A) m_j - \frac{1}{2n^2} s(A)s(M^2). \tag{7}$$

Proof

The quadratic term in (EP) is bounded independently of the linear term by theorem 2.2, contributing the first summand in (7). To bound the linear term of (EP), we observe

$$\begin{aligned} \tfrac{1}{n} tr\, V^T A u_n u_k^T M^{3/2} W Z^T &= \tfrac{1}{n} tr\, A u_n u_k^T M (M X^T - \tfrac{1}{n} M u_k u_n^T) \\ &= \tfrac{1}{n} tr\, \{r(A) r^T(M) X^T\} - \tfrac{1}{n^2} s(A)s(M^2). \end{aligned}$$

It is easy to verify that due to the ordering of $r(A)$ and m, the transportation problem

$$\max\{tr \quad r(A) r^T(M) X^T : X \in F\}$$

has optimal value $\sum_{j=1}^{k} R_j(A) m_j$. (Take the partition where nodes $1, \ldots, m_1$ belong to S_1, nodes $m_1 + 1, \ldots, m_1 + m_2$ belong to S_2, etc.) Summing all the terms completes the proof. □

We point out that in general there will not be a matrix X for which the bound is attained, because we maximize two terms independently. In the following three special cases however, we are able to treat the objective function as a whole.

COROLLARY 4.1

Under the conditions of theorem 4.1, assume that $m_1 = \ldots = m_k$ (i.e. partition into k blocks, each of size n/k). Then

$$|E_{uncut}| \leq \max\left\{\frac{1}{2}\, tr\, X^{\mathrm{T}}AX : X \in \mathscr{E}, X^{\mathrm{T}}X = \frac{n}{k}I\right\} = \frac{n}{2k}\sum_{j=1}^{k-1}\lambda_j(\hat{A}) + \frac{1}{2k}s(A).$$

Moreover, the bound is attained for

$$X = \frac{1}{k}u_n u_k^{\mathrm{T}} + \sqrt{\frac{n}{k}}\, VZW^{\mathrm{T}},$$

where $Z = (z_1, \ldots, z_{k-1}) \in \mathcal{O}$ contains the eigenvectors z_j corresponding to $\lambda_j(\hat{A})$.

Proof

By substituting $M = (n/k)I$ in (EP) and using the expansion of the linear term, contained in the proof of theorem 4.1, we get

$$\frac{1}{2}\, tr\, X^{\mathrm{T}}AX = \frac{n}{2k}\, tr\, IZ^{\mathrm{T}}\hat{A}Z + \frac{1}{n}\, tr\, r(A)r^{\mathrm{T}}(M)X^{\mathrm{T}} - \frac{1}{2n^2}s(A)s(M^2).$$

Now note that $X \in \mathscr{E}$ implies

$$r^{\mathrm{T}}(M)X^{\mathrm{T}} = \frac{n}{k}u_k^{\mathrm{T}}X^{\mathrm{T}} = \frac{n}{k}u_n^{\mathrm{T}}.$$

Thus, the linear term is constant:

$$\frac{1}{n}\, tr\, r(A)r^{\mathrm{T}}(M)X^{\mathrm{T}} = \frac{1}{n}\,\frac{n}{k}s(A).$$

Finally, we have

$$s(M^2) = k\frac{n^2}{k^2}.$$

Bounding the quadratic term again by theorem 2.2 and summing the remaining (constant) terms proves the upper bound. The upper bound for the quadratic term is attained for Z containing the (normalized and pairwise orthogonal) eigenvectors corresponding to the first $k-1$ largest eigenvalues of $\hat{A}$. X is recovered using (6). □

EXAMPLE 1 (continued)

$\lambda_1(\hat{A}) = 3.3254$. Using corollary 4.1, we obtain

$$|E_{uncut}| \leq \frac{n}{4}\lambda_1(\hat{A}) + \frac{1}{4}s(A) = 42.1219.$$

Thus, no partition leaves more than 42 edges uncut. Note also (see table 2) that the largest eigenvalue of $\hat{A}$ is simple and so the maximizer Z is unique up to multiplication by -1. It produces a matrix X where already several components are either close to 0 or close to 1, see the dashed line in figure 1.

It is worth mentioning that the bound from corollary 4.1 is equivalent to the bound proposed by Boppana [4] in the case $k = 2$. We leave it as an exercise for the interested reader to establish this equivalence. Boppana's bound, however, does not seem to allow a generalization to $k > 2$.

COROLLARY 4.2

Under the conditions of theorem 4.1, assume that u_n is an eigenvector of A with eigenvalue t. (This occurs, for instance, if the underlying graph is t-regular.) Then

$$|E_{uncut}| \leq \max\left\{\frac{1}{2} tr\, X^{\mathrm{T}} AX : X \in \mathscr{E}, X^{\mathrm{T}} X = M\right\} = \frac{1}{2} \sum_{j=1}^{k-1} \lambda_j(\hat{A})\lambda_j(\hat{M}) + \frac{t}{2n} s(M^2).$$

Moreover, the bound is attained for

$$X = \frac{1}{n} u_n u_k^{\mathrm{T}} M + VZW^{\mathrm{T}} M^{1/2},$$

where $Z = U_1 U_2^{\mathrm{T}}$ and $U_2 \in \mathcal{O}_{n-1,k-1}$ diagonalizes $\hat{M}$ and $U_1 \in \mathcal{O}_{n-1,k-1}$ contains the eigenvectors corresponding to the $k-1$ largest eigenvalues of $\hat{A}$.

Proof

We first show that in this case the linear term in (EP) vanishes. We have, as in the proof of theorem 4.1,

$$\frac{t}{n} tr(u_k^{\mathrm{T}} M)(X^{\mathrm{T}} u_n) - \frac{t}{n} s(M^2) = \frac{t}{n} tr\,(m^{\mathrm{T}} m) - \frac{t}{n} s(M^2) = 0.$$

The upper bound for the quadratic term is attained for $Z = U_1 U_2^{\mathrm{T}}$ where $U_2 \in \mathcal{O}$ diagonalizes $\hat{M}$, and $U_1 \in \mathcal{O}$ contains the eigenvectors corresponding to the $k-1$ largest eigenvalues of $\hat{A}$. X is recovered using (6). □

In the previous two cases, we were able to strengthen theorem 4.1 because in these cases the linear term in the objective function of (EP) was constant for all feasible X. We conclude this section with a nontrivial extension of theorem 4.1 in the case $k = 2$, i.e. partition into two blocks (of possibly different sizes).

COROLLARY 4.3

Under the conditions of theorem 4.1, assume that $k = 2$. Then

$$|E_{uncut}| \le \max\left\{\frac{1}{2} tr\, X^{\mathrm{T}}AX : X \in \mathscr{E}, X^{\mathrm{T}}X = M\right\}$$

$$= \max\{z^{\mathrm{T}}Cz + c^{\mathrm{T}}z + const : z^{\mathrm{T}}z = 1\},$$

where

$$z \in \mathbb{R}^{n-1},\ C = \frac{1}{n} m_1 m_2 \hat{A},\ c = \sqrt{\frac{m_1 m_2}{n}}\ \frac{m_2 - m_1}{n} V^{\mathrm{T}}Au_n,\ const = \frac{1}{2n^2} s(A)s(M^2).$$

Proof

From (EP), it is clear that the matrix $Z = z \in \mathbb{R}^{n-1}$ and $\hat{M}$ is a scalar. Note further that by the definition of W, we can set

$$W = \frac{1}{\sqrt{n}}\left(-\sqrt{m_2}\ \sqrt{m_1}\right)^{\mathrm{T}}.$$

Thus,

$$\hat{M} = W^{\mathrm{T}}MW = \frac{2}{n} m_1 m_2.$$

The quadratic term in (EP) therefore simplifies to $z^{\mathrm{T}}Cz$, and the linear term simplifies to $c^{\mathrm{T}}z$. □

We point out that the (global) maximum of

$$\{z^{\mathrm{T}}Cz + c^{\mathrm{T}}z + const : z^{\mathrm{T}}z = 1\}$$

can be calculated efficiently, see [12–14,23]. The main computational steps involve finding the eigenvalues of the symmetric matrix C and the largest zero of a rational function, see [13,23] for details and computational experiments. The maximizing z can be recovered using the eigenvectors of C.

5. Diagonal perturbations to improve the bounds

It is a trivial observation to note that loops in a graph (i.e. edges joining some $i \in V$ to itself) are not cut by any partition. Therefore, adding "weighted loops" to our graph, i.e. replacing A by $A + diag(d)$ for some $d \in \mathbb{R}^n$ does not affect the graph partitioning problem, see also [3,10]. To be more specific, we will first show that adding a multiple of the identity to A not only leaves the graph partitioning problem unchanged, but also all the bounds described so far.

LEMMA 5.1

Let A and m describe a graph partitioning problem. Let $A(\alpha) := A + \alpha I$ for some $\alpha \in \mathbb{R}$. Then

$$tr\, X^{\mathrm{T}} AX = tr\, X^{\mathrm{T}} A(\alpha)X - \alpha n, \quad \forall X : X^{\mathrm{T}} X = M. \tag{8}$$

Moreover, the upper bounds from theorem 2.2 and from section 4 gives the same result when applied to the left-hand side and to the right-hand side of (8).

Proof

The equality (8) is obvious. Therefore, any bound obtained by maximizing over $X^{\mathrm{T}}X = M$ will be unaltered by the change in A. The only open case is theorem 4.1 because there we maximize two terms independently.

We first point out that

$$\hat{A}(\alpha) = \hat{A} + \alpha I_{n-1},$$

$$s(A(\alpha)) = s(A) + \alpha n,$$

$$R_j(A(\alpha)) = R_j(A) + \alpha m_j,$$

$$tr\, \hat{M} = tr\, M - \frac{1}{n} tr\, M^2 = n - \frac{1}{n} s(M^2).$$

The last relation follows using $WW^{\mathrm{T}} = I - \frac{1}{n} \overline{m}\,\overline{m}^{\mathrm{T}}$. Let us denote by $EPB(A)$ the eigenvalue bound of theorem 4.1 applied to the matrix A. Bounding the right-hand side in (8), we obtain

$$\begin{aligned}
EPB(A(\alpha)) &- \frac{1}{2}\alpha n \\
&= \frac{1}{2} \sum \lambda_j(\hat{A}(\alpha))\lambda_j(\hat{M}) + \frac{1}{n} \sum R_j(A(\alpha))m_j \\
&\quad - \frac{1}{2n^2} s(A(\alpha))s(M^2) - \frac{1}{2}\alpha n \\
&= \frac{1}{2} \sum \lambda_j(\hat{A})\lambda_j(\hat{M}) + \frac{1}{2}\alpha \sum \lambda_j(\hat{M}) + \frac{1}{n} \sum R_j(A)m_j \\
&\quad + \frac{1}{n}\alpha \sum m_j^2 - \frac{1}{2n^2} s(A)s(M^2) - \frac{1}{2n}\alpha s(M^2) - \frac{1}{2}\alpha n \\
&= EPB(A) + \frac{1}{2}\alpha(tr\, M - \frac{1}{n} tr\, M^2) + \frac{1}{n}\alpha s(M^2) - \frac{1}{2n}\alpha s(M^2) - \frac{1}{2}\alpha n \\
&= EPB(A). \qquad \square
\end{aligned}$$

As mentioned above, a general perturbation of the main diagonal of A does not affect the edges cut by a partition. This has been pointed out and used by several researchers in the past.

LEMMA 5.2 [3,10]

For $d \in \mathbb{R}^n$, let $A(d) = A + diag(d)$. Then

$$tr\, X^{\mathrm{T}} A(d) X = tr\, X^{\mathrm{T}} AX + s(d)$$

for all partitions X.

If d is arbitrary, then $\lambda_j(A(d))$ will in general be different from $\lambda_j(A) + d_j$, so the upper bounds may vary with d. In view of lemma 5.1, it is sufficient to consider perturbations d that sum up to 0. Then the graph partitioning problems with matrices $A(=A(0))$ and $A(d)$ are identival and we may choose any $A(d)$ where $s(d) = 0$ to derive an upper bound.

In the following, we focus on the special case of corollary 4.1, even though the techniques can be extended to the general case (but become more complicated).

Let

$$\hat{A}(d) := \hat{A} + V^{\mathrm{T}} diag(d) V$$

and

$$g(d) := \sum_{j=1}^{k-1} \lambda_j(\hat{A}(d)).$$

Donath and Hoffman [10] point out that

$$\sum_{j=1}^{k} \lambda_j(A + B)$$

is a convex function of B for A fixed, provided both A and B are symmetric. Therefore, $g(d)$ is convex. Moreover, using theorem 4.6 in [7], it is easy to verify that $g(d)$ is differentiable for all d such that

$$\lambda_{k-1}(\hat{A}(d)) \neq \lambda_k(\hat{A}(d)).$$

Note also that under the assumption $s(d) = 0$, it is easily shown that

$$\lim_{\|d\| \to \infty} g(d) = \infty. \tag{9}$$

The above discussion is summarized as follows.

LEMMA 5.3

Suppose $m = (n/k) u_k$. Then

$$|E_{uncut}| \leq \frac{1}{2k} s(A) + \min \left\{ \frac{n}{2k} \sum_{j=1}^{k-1} \lambda_j(\hat{A} + V^{\mathrm{T}} diag(d) V) : d \in \mathbb{R}^n, s(d) = 0 \right\}. \tag{10}$$

□

We point out that the minimum is attained because of (9). We now address the question of differentiability of $g(d)$ in more detail. Suppose first that $\lambda_i(\hat{A}(d))$ is simple with normalized eigenvector x_i, then

$$\frac{\partial}{\partial d_j}\lambda_i(\hat{A}(d)) = \frac{\partial}{\partial d_j}\left(x_i^{\mathrm{T}}V^{\mathrm{T}}\left(A + \sum_{r=1}^{n} d_r e_r e_r^{\mathrm{T}}\right)Vx_i\right)$$
$$= x_i^{\mathrm{T}}V^{\mathrm{T}}e_j e_j^{\mathrm{T}}Vx_i$$
$$= (e_j^{\mathrm{T}}Vx_i)^2, \quad \forall j = 1,\dots,n. \tag{11}$$

Here, e_j denotes the jth canonical unit vector. Otherwise, an element of the eigenspace has to be chosen properly, see theorem 5.1 in [15], to provide the differentials. In the general case, (11) still provides a subgradient.

In summary, the function $g(d)$ to be minimized is convex and we can provide a subgradient for any d. So applying techniques from nonsmooth optimization applied to convex functions, it is possible to find the best possible upper bound in (10). We used the BT method proposed in [30] to carry out the minimization.

EXAMPLE 1 (continued)

The results for the iterative improvement of our example are summarized in table 3. As a stopping criterion, we tested whether an ε-subgradient of norm less than 0.001 was found. This occurred after 45 iterations. We observe that after only a few iterations, we already have a very good upper bound and most of the iterations are spent finding a subgradient of small norm. Moreover, it turns out that at the final

Table 3

Subgradient improved upper bound for example 1. The first column indicates the iteration and the second column the corresponding upper bound. The last column contains the norm of an ε-subgradient for $g(d)$ found at the given iteration.

Iteration	Bound	Norm
1	42.13	1.065
5	38.80	0.323
10	38.61	0.085
15	38.57	0.068
20	38.56	0.028
25	38.56	0.012
30	38.56	0.008
35	38.55	0.004
40	38.55	0.006
45	38.55	0.0007

perturbation d, the largest eigenvalue has multiplicity larger than 1, see also table 2; therefore, $g(d)$ is nondifferentiable for this d. This coincides with the experiences reported in [7]. Finally, we point out that the maximizer X, producing the bound, is already very close to a 0, 1 matrix, see the solid line in figure 1.

We conclude this section with a perturbation of the main diagonal of A that allows an application of corollary 4.2.

THEOREM 5.1

Let A and m describe a graph partitioning problem. Let

$$d_i := \frac{1}{n} s(A) - r_i(A)$$

and

$$A(d) = A + diag(d).$$

Then

$$|E_{uncut}| \leq \frac{1}{2} \sum_{j=1}^{k-1} \lambda_j(\hat{A}(d))\lambda_j(\hat{M}) + \frac{1}{2n^2} s(A)s(M^2).$$

Moreover, the bound is attained for

$$X = \frac{1}{n} u_n u_k^{\mathrm{T}} M + VZW^{\mathrm{T}} M^{1/2},$$

where $Z = U_1 U_2^{\mathrm{T}}$, and $U_2 \in \mathcal{O}_{k-1,k-1}$ diagonalizes $\hat{M}$ and $U_1 \in \mathcal{O}_{n-1,k-1}$ contains the eigenvectors corresponding to the $k-1$ largest eigenvalues of $\hat{A}(d)$.

Proof

First note that $s(d) = 0$. Moreover, by the definition of d, u_n is an eigenvector of $A(d)$ with corresponding eigenvalue $(1/n)s(A)$. The result now follows using corollary 4.2. □

6. Finding a closest feasible solution

Our bounding techniques find approximate solution matrices X which in general are not feasible for (P). We now present several procedures for finding feasible solutions Y using the information from X. One approach consists of looking for a feasible Y that is as close as possible to X. Alternatively, we propose to use X to linearize the objective function in (P) to derive good feasible solutions.

6.1. CLOSEST IN FROBENIUS NORM

Suppose that the matrix X obtained from our relaxation procedure satisfies $X^{\mathrm{T}}X = M$, but is not a 0, 1 matrix. We want to find a feasible matrix Y for (P) which best approximates X in Frobenius norm. Note that feasibility implies $Y^{\mathrm{T}}Y = M$ as well. Therefore,

$$\begin{aligned}\|X - Y\|^2 &= \|X\|^2 + \|Y\|^2 - 2tr\, X^{\mathrm{T}}Y \\ &= 2tr\, M - 2tr\, X^{\mathrm{T}}Y.\end{aligned}$$

We can now find the best feasible approximate to X in Frobenius norm by solving the following transportation problem in the variable $Y \in \mathbb{R}^{n\times k}$:

$$\min\{-tr\, X^{\mathrm{T}}Y : Y \in F\}.$$

Since the sum of the elements of Y is n, note that the objective function is equivalent to $tr(\frac{1}{2}u_n u_k^{\mathrm{T}} - X)^{\mathrm{T}}Y$. This latter function has an l_1 norm quality.

We point out that this idea is also (implicitly) used by Barnes [1] to derive feasible solutions. Barnes uses the appropriately normalized eigenvectors corresponding to the largest eigenvalues of A for X. It is clear that the above model works for any X, as long as $X^{\mathrm{T}}X = M$. This approximation model has, however, the disadvantage that the structure of the problem, i.e. A, is not used and one just tries to find a feasible Y closest to X. Therefore, it only makes sense if X is already "very close" to an optimal partition.

EXAMPLE 1 (continued)

If we solve the above transportation problem with the X corresponding to the bound from lemma 5.3, then we obtain the feasible solution

$$S_1 = \{1, 2, 3, 11, 12, 13, 14, 17, 18, 20\}, \quad S_2 = N \backslash S_1$$

of value 38. Comparing with the bounds in table 2, we conclude that this partition is already optimal, because the upper bound from lemma 5.3 also becomes 38 after rounding down. We point out that this solution can also be obtained by simply rounding the maximizer X to the nearest integers, see figure 1.

6.2. LINEAR APPROXIMATION

Expanding the objective function at X, we obtain

$$tr\, Y^{\mathrm{T}}AY = tr\, X^{\mathrm{T}}AX + 2tr\, X^{\mathrm{T}}A(Y - X) + tr(Y - X)^{\mathrm{T}}A(Y - X).$$

If we use the linear approximation, we get the transportation problem in Y:

$$\max\{tr\, X^{\mathrm{T}}AY : Y \in F\}.$$

If A is positive definite, then the weighted Frobenius norm

$$\| A^{1/2}(X - Y)\|^2 = \| A^{1/2}X\|^2 + \| A^{1/2}Y\|^2 - 2tr\, X^{\mathrm{T}}AY,$$

i.e. the above problem is equivalent to the weighted Frobenius norm approximation problem if we ignore the quadratic term in Y.

Barnes et al. [3] use a feasible X and try to improve it. They propose a diagonal perturbation that changes A to a positive semidefinite matrix and then set up a transportation problem to find a better partition Y. A careful analysis of their objective function shows that it corresponds precisely to the linearized model above. Their model makes essential use of semidefiniteness and feasibility of X. Our model shows that both these assumptions are not necessary. It raises the interesting question, however, to maximize the above model over all diagonal shifts. This will not be pursued further in this paper.

7. Computational results

In this section, we first present computational experiences for the various new bounds on small graphs that have been previously used in the literature. We use the following graphs described in table 4 and note that G2 is our running example 1.

Table 4

Graphs from the literature.

Name	$\|V\|$	$\|E\|$	Source
G1	20	55	[10], table 2 p. 425
G2	20	51	[10], table 3 p. 425
G3	20	46	[7], table 1 p. 52
G4	21	48	[3], figure 2 p. 305

In table 5, we summarize the results in the case of partitions into sets of equal size. Comparing the last two columns, we see that the feasible solutions obtained are in fact optimal for all graphs except G3. The solution of G3 is at most "one edge off" from optimality.

Table 5

Partitioning into k sets of equal size.

Graph	k	Thm 2.2	Cor. 4.2	Lemma 5.3	Feas. X
G1	2	47.13	45.65	42.85	42
G2	2	45.90	42.13	38.55	38
G3	2	37.71	35.54	34.22	33
G4	3	47.98	47.19	45.47	45

Next, we investigate our bounds for the weighted graph from [5, p. 67]. This graph has 40 nodes and $m_1 = m_2 = 20$. Two sets C_1 and C_2 of edge costs are given in [5]. The underlying graph is 3-regular. According to the authors, the costs C_1 and C_2 are drawn uniformly from $\{1,\ldots,10\}$. We examine the following three variants V1, V2 and V3 for this problem:

V1: all edge costs are 1,

V2: use C_1 for the edge costs,

V3: use C_2 for the edge costs.

In table 6, the results for the various bounds are summarized. Note that in the case of V1, the bound from lemma 5.3 hardly improves the classical bound of theorem 2.2. One reason for this may lie in the regular structure of the graphs, which contains many optimal partitions. We also point out that the bound from theorem 5.1 in quite competitive with the subgradient improved bound from lemma 5.3 for the variants V2 and V3. The optimal solution values are from [5].

Table 6

Three variants of a 40-node graph.

Variant	Σa_{ij}	Thm 2.2	Cor. 4.1	Lemma 5.3	Thm 5.1	Opt.
V1	60	58.52	58.52	57.35	58.52	54
V2	316	345.77	326.44	307.10	309.88	297
V3	341	397.18	367.84	330.42	334.34	322

To further examine the performance of these bounds, we generated a series of pseudo-random graphs of larger sizes. We generated five graphs, each of average degree 5 for $n \in \{30, 40, 50\}$. The results are summarized in table 7. (Further computational results on much larger graphs are described in a companion paper [11].)

The columns labeled "Thm 2.2" and "Thm 2.2 + impr." in table 7 represent the Donath Hoffman bound without and with diagonal perturbations to improve the bound. Comparing again the last two columns with the column of theorem 2.2 plus improvement, it turns out that new bounding rules constitute a significant improvement over the previously known techniques. The gap for the problem with 30 nodes is never larger than 2 edges, for problems with 50 nodes it never exceeds 5 edges. Moreover, the gap as compared to previously known bounds is typically reduced by about 30 to 50% on all problems considered.

It seems more difficult to find good bounds if the block sizes m_i are not equal. Since we have proposed several new bounding techniques also for this situation, we conclude this section with a numerical study of partitioning into sets of different

Table 7

Partitioning of pseudo-random graphs into two blocks of equal size.

n	$\|E\|$	Thm 2.2	Cor. 4.1	Thm 2.2 + impr.	Lemma 5.3	Feas. X
30	75	66.65	62.51	60.39	58.10	56
30	75	72.51	65.58	62.31	59.77	58
30	84	77.61	72.15	69.10	65.25	63
30	73	67.43	62.47	60.45	58.06	56
30	69	64.54	60.26	58.70	56.73	54
40	110	101.23	93.84	90.57	87.17	82
40	102	97.33	89.70	85.90	82.98	79
40	102	102.93	95.81	90.26	87.52	86
40	91	89.96	83.11	79.34	76.56	73
40	101	95.52	87.32	84.21	80.86	77
50	139	128.17	118.17	114.71	110.43	105
50	117	117.41	106.28	100.25	95.45	90
50	123	117.51	108.57	104.12	100.86	96
50	128	120.53	109.73	108.06	103.38	98
50	138	126.97	120.56	115.22	112.47	108

Table 8

Partitioning G2 into two blocks of sizes m_1 and $20 - m_1$.

m_1	Thm 2.2	Thm 4.1	Cor. 4.3	Thm 5.1
19	58.98	53.00	55.71	50.14
17	56.07	52.98	53.20	48.82
15	53.17	51.09	49.41	47.80
13	50.26	47.64	45.87	47.11
11	47.35	44.01	43.10	46.77

sizes. We take the graph G2, or running example, and partition it into two blocks of different sizes. The numerical results are summarized in table 8. Since G2 has only 51 edges, we conclude that in the case of "very unequal" block sizes, all bounds except theorem 5.1 fail. We have to note, however, that the bound in theorem 5.1 does not improve after a diagonal perturbation, but all other bounds can be further improved in general. As m_1 decreases, the bound from corollary 4.3 turns out to be the favorite.

8. Summary and conclusions

We have presented several new bounds for the graph partitioning problem. The main mathematical tool was the variation of the Hoffman–Wielandt inequality used in corollary 2.1. As a second crucial ingredient, we used a parametrization of matrices X describing partitions that allowed us to constrain our relaxation to the subspace of matrices, having prescribed row and column sums (lemma 3.1). As a last important idea, we used the possibility of adding weighted loops to the graph to further strengthen the relaxations.

We presented two general new bounds (theorems 4.1 and 5.1) and studied the following special cases in more detail:

- partitioning into blocks of equal size (corollary 4.1, lemma 5.3),
- partitioning of regular graphs (corollary 4.2),
- partitioning into only two blocks (corollary 4.3).

Our approach allows some extensions. If a partitioning of the node set is sought where the number k of subsets is prescribed, but the sizes m_i of the sets can vary in some interval,

$$h_i \le m_i \le l_i,$$

then our bounds can still be used. One would have to consider the bounds as functions of the m_i's. If k is reasonably small, it is still feasible to optimize the resulting bounds as functions of the m_i.

Finally, we point out that substantial computational experiments together with implementation details using the bounds of the present paper are contained in a companion paper [11], and show that in most of the cases it is possible to obtain lower and upper bounds on equipartitions differing only by a few percentage points.

Acknowledgements

The authors thank the Natural Sciences and Engineering Research Council of Canada and the Christian Doppler Laboratory "Diskrete Optimierung" for their support. They also thank Helga Schramm for the permission to use her code for the BT method.

References

[1] E.R. Barnes, An algorithm for partitioning the nodes of a graph, SIAM J. Alg. Discr. Math. 3(1982)541–550.

[2] E.R. Barnes and A.J. Hoffman, Partitioning, spectra, and linear programming, in: *Progress in Combinatorial Optimization*, ed. W.E. Pulleyblank (Academic Press, 1984) pp. 13–25.

[3] E.R. Barnes, A. Vannelli and J.Q. Walker, A new heuristic for partitioning the nodes of a graph, SIAM J. Discr. Math. 1(1988)299–305.

[4] R.B. Boppana, Eigenvalues and graph bisection: An average case analysis, in: *Proc. 28th Annual Symp. on Computer Science* (IEEE, 1987) pp. 280–285.

[5] N. Christofides and P. Brooker, The optimal partitioning of graphs, SIAM J. Appl. Math. 30(1976) 55–69.
[6] M. Conforti, M.R. Rao and S. Sassano, The equipartition polytope, Math. Progr. 49(1990) 49–90.
[7] J. Cullum, W.E. Donath and P. Wolfe, The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices, Math. Progr. Study 3(1975)35–55.
[8] Ch. Delorme and S. Poljak, Combinatorial properties and the complexity of a max-cut approximation, Euro. J. Combinatorics 14(1993)313–333.
[9] Ch. Delorme and S. Poljak, Laplacian eigenvalues and the maximum cut problem, Math. Progr. 62(1993)557–574.
[10] W.E. Donath and A.J. Hoffman, Lower bounds for the partitioning of graphs, IBM J. Res. Develop. 17(1973)420–425.
[11] J. Falkner, F. Rendl and H. Wolkowicz, A computational study of graph partitioning, Math. Progr. 66(1994)211–240.
[12] W. Gander, Least squares with a quadratic constraint, Numer. Math. 36(1981)291–307.
[13] W. Gander, G.H. Golub and U. von Matt, A constrained eigenvalue problem, Lin. Alg. Appl. 114/115(1989)815–839.
[14] D.M. Gay, Computing optimal locally constrained steps, SIAM J. Sci. Stat. Comput. 2(1981) 186–197.
[15] B. Gollan, Eigenvalue perturbations and nonlinear parametric optimization, Math. Progr. Studies 30(1987)67–81.
[16] S.W. Hadley, Continuous optimization approaches for the quadratic assignment problem, Ph.D. Thesis, University of Waterloo, Canada (1989).
[17] S.W. Hadley, F. Rendl and H. Wolkowicz, A new lower bound via projection for the quadratic assignment problem, Math. Oper. Res. 17(1992)727–739.
[18] Ch. Helmberg, B. Mohar, S. Poljak and F. Rendl, A spectra approach to bandwidth and separator problems in graphs, Technical Report, Graz (1992).
[19] A. Kamath and N. Karmakar, A continuous approach to computer upper bounds in quadratic maximization problems with integer constraints, *Proc. Conf. on Recent Advances in Global Optimization*, ed. Floudas and Pardalos (1991) pp. 125–140.
[20] A. Kamath and N. Karmakar, A continuous method for computing bounds in integer quadratic optimization problems, J. Global Optim. 2(1992)229–241.
[21] T. Lengauer, *Combinatorial Algorithms for Integrated Circuit Layout* (Wiley, Chichester, 1990).
[22] B. Mohar and S. Poljak, Eigenvalues and the max-cut problem, Czech. Math. J. 40(115)(1990) 343–352.
[23] J.J. Moré and D.C. Sorensen, Computing a trust region step, SIAM J. Sci. Statist. Comput. 4(1983)553–572.
[24] M.L. Overton and R.S. Womersley, On the sum of the largest eigenvalues of a symmetric matrix, SIAM J. Matrix Anal. Appl. 13(1992)41–45.
[25] S. Poljak and F. Rendl, Solving the max-cut problem using eigenvalues, Report 91735, Bonn (1991).
[26] S. Poljak and F. Rendl, Nonpolyhedral relaxations of graoh bisection problems, Report CDLDO-13, University of Technology, Graz (1992).
[27] A. Pothen, H.D. Simon and K.-P. Liou, Partitioning sparse matrices with eigenvectors of graphs, SIAM J. Matrix Anal. Appl. 11(1990)430–452.
[28] F. Rendl and H. Wolkowicz, Applications of parametric programming and eigenvalue maximization to the quadratic assignment problem, Math. Progr. 53(1992)63–78.
[29] C. Roucairol and P. Hansen, Cut cost minimization in graph partitioning, in: *Numerical and Applied Mathematics*, ed. C. Brezinski (J.C. Baltzer, 1989) pp. 585–587.
[30] H. Schramm and J. Zowe, A version of the Bundle idea for minimizing a nonsmooth function: Conceptual idea, convergence analysis, numerical results, SIAM J. Optim. 2(1992)121–152.