A Projection Technique for Partitioning the Nodes of a Graph

Franz Rendl
Technische Universität Graz
Institut für Mathematik
Kopernikusgasse 24, A-8010 Graz, Austria
and
Henry Wolkowicz
University of Waterloo
Department of Combinatorics and Optimization
Waterloo, Ontario, N2L 3G1, Canada

May 24, 1994

Abstract

Let $G = (N, E)$ be a given 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 a given undirected graph with node set $N = \{1, \dots, n\}$ and edge set E. A common problem in circuit board and micro-chip design, computer program segmentation, floor planning and other layout problems is to partition the node set N into k disjoint subsets $S_1, \dots, S_k$ of specified sizes $m_1 \geq m_2 \geq \dots \geq m_k$, $\sum_{j=1}^{k} m_j = n$, so as to minimize the number 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 [1].

The graph partitoning problem can be formulated as a 0-1 quadratic programming problem (see e.g. [2]): let $X \in \Re^{n \times k}$ with the columns

$$x_j = (x_{1j}\ x_{2j} \dots x_{nj})^t$$

[0]The authors would like to thank the Natural Sciences and Engineering Research Council of Canada and the Austrian Science Foundation (FWF) for their support.

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

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

Let $A_0 = (a_{ij})$ be the *adjacency matrix* for G, i.e. a_{ij} denotes the number of edges connecting nodes i and j, $i \neq j$, $a_{ii} = 0$, $i = 1, \dots, n$. Then

$$\frac{1}{2} \sum_{r=1}^{n} \sum_{s=1}^{n} a_{rs} x_{rj} x_{sj} = \frac{1}{2} x_j^t A_0 x_j \tag{1.1}$$

is the number 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

$$\begin{array}{l} \sum_{i=1}^{n} x_{ij} = m_j, j = 1, \dots, k, \\ \sum_{j=1}^{k} x_{ij} = 1, \quad i = 1, \dots, n. \end{array}$$

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

$$(\mathbf{P}) \qquad \begin{cases} \underline{\max} & \frac{1}{2} \text{ tr } X^t A X \\ \underline{\text{s.t.}} & X u_k = u_n \\ & X^t u_n = m \\ & X \text{ is a } 0,1 \text{ matrix }, \end{cases}$$

where *tr* denotes trace, $u_j \in \Re^j$ is the vector of ones, and $m = (m_1, \dots, m_k)^t$ is the ordered vector of specified sizes. We allow for general matrices A in the objective function rather than restrict $A = A_0$. (The matrix A may contain weights for the edges.)

For a given partition X, we say that $T = (t_{ij}) \in \Re^{n \times n}$ 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. The nonzero eigenvalues of T are $m_1, \dots, m_k$ with normalized eigenvectors $\frac{1}{\sqrt{m_j}} x_j, j = 1, \dots, k$, where the x_j are the columns of X. Thus

$$T = VMV^t, \quad ||T||^2 = ||M||^2 , \tag{1.2}$$

where $M = \text{diag}(m) \in \Re^{k \times k}$, $||M||^2 = trM^tM$ is the Frobenius norm, and the columns of V are $v_j = \frac{1}{\sqrt{m_j}} x_j$. Note that $T = VMV^t = XM^{-1/2}MM^{-1/2}X^t = XX^t$,

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

and $\frac{1}{2} trAT$ is the number of edges not cut by the partition. In addition

$$\begin{aligned} trAT &= tr\ AXX^t \\ &= tr\ X^tAX\ . \end{aligned}$$

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

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

The formulation (P) is very similar to the quadratic assignment problem, QAP. Continuous optimization techniques are employed in [3, 4, 5] to find bounds for the QAP. In particular, a projection technique is used in [4] to eliminate the constraints on the row and column sums of X. An iterative improvement of QAP bounds, based on "reductions", is presented in [5]. In this paper we extend the continuous optimization techniques from [4, 5] to the graph partitioning problem.

The rest of 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 first scale (P) to get an equivalent scaled program (SP) for which the vector m of specified sizes becomes a vector of ones. The orthogonal relaxation of this program yields the eigenvalue bound for (P) proposed in [6].

In Section 3 we extend the projection technique from [4] 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 ($n \leq 20$) 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.

Overview of previous related research

In [7, 6] spectral information of A is used to bound the objective function of (P). Boppana [8] considers graph bisection, i.e. $m_1 = m_2 = \frac{n}{2}$, and improves the eigenvalue bound from [6] for this special case. The papers [9, 10] describe Branch and Bound approaches to solve the partitioning problem in the case $k = 2$ and for general weighted graphs. Both methods seem to work only for extremely thin graphs (average degree not more than 4).

Several articles are devoted to finding "good" partitions using spectral information from A. In [11] 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 [2]. Therein A_0 is shifted by a diagonal matrix D so that $A = A_0 + D$ is positive semidefinite. Then the Cholesky decomposition of A is used to improve a given partition.

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

Table 1.1: Edge set of Example 1

Finally a survey on various aspects of the graph partitioning problem and further references are contained in Chapter 6 of [1].

Example 1. We will illustrate our results, as we progress through the paper, on the following example from [6]. The graph has 20 nodes and we partition it into two equal parts, i.e. $m_1 = m_2 = 10$. The connections are given in Table 1.1. Note 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. $Q \in \mathcal{O}_{k,l}$ if $Q^tQ = I$. The vector of ones is $u_l = (1, \ldots, 1)^t \in \Re^l$; $r(K) = Ku_l$ is the vector of row sums of a $k \times l$ matrix K, while $s(K) = u_k^t K u_l$ is the sum of all the elements of K. We denote by $m = (m_1, \cdots, m_k)$ 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 $M = \text{diag}\ (m)$, while for a given matrix M, diag (M) denotes the vector formed from the diagonal of M. For a given Hermitian matrix A, $\lambda_j(A)$ denotes the j^{th} largest eigenvalue of A.

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

$$\mathcal{E} = \{X \in \Re^{n\times k} : Xu_k = u_n, X^t u_n = m\}. \tag{2.1}$$

The set of nonnegative matrices is

$$\mathcal{N} = \{X \in \Re^{n\times k} : X \geq 0 \text{ elementwise }\}.$$

The feasible set of matrices for (P) is

$$F = \{X \in \Re^{n\times k} : x_{ij} \text{ is 0 or 1}\} \cap \mathcal{E}. \tag{2.2}$$

Lemma 2.1 *The feasible set satisfies*

$$\begin{aligned} F &= \mathcal{E} \cap \mathcal{N} \cap \{X \in \Re^{n\times k} : X^t X = M\} \\ &= \mathcal{E} \cap \mathcal{N} \cap \{X \in \Re^{n\times k} : trX^t X = n\}. \end{aligned}$$

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

$$\begin{aligned} X \in \mathcal{N}, Xu_k = u_n &\Rightarrow 0 \leq x_{ij} \leq 1 \\ &\Rightarrow x_{ij}^2 \leq x_{ij}. \end{aligned}$$

And so

$$n = trX^t X \leq s(X) = u_n^t X u_k = n$$

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 $X \in \mathcal{E} \cap \mathcal{N}$ corresponds to Quadratic Programming. Note that A is in general indefinite (e.g. $diag(A_o) = 0$), so the (global) maximum is difficult to find. Relaxing to $X \in \{X \in \Re^{n\times k} : X^t X = M\}$ leads to the eigenvalue bound derived in [6], see also Theorem 2.1 below. In this paper we strengthen the bound by maximizing over

$$X \in \{X \in \Re^{n\times k} : X^t X = M\} \cap \mathcal{E}.$$

We now scale X in (P) in order to change M to the identity I. We let

$$X = YM^{1/2} \tag{2.3}$$

and define

$$\bar{m} = M^{1/2} u_k = (\sqrt{m_1}, \ldots, \sqrt{m_k})^t.$$

Note that

$$\bar{m} = M^{-1/2}m$$

and

$$\| \bar{m} \| = \| u_n \| = \sqrt{n}.$$

The constraint $Xu_k = u_n$ is equivalent to $(XM^{-1/2})(M^{1/2}u_k) = u_n$ while $X^t u_n = m$ is equivalent to $(XM^{-1/2})^t u_n = M^{-1/2}m = \bar{m}$. Therefore the problem (P) is equivalent to the scaled problem

$$(\mathbf{SP}) \quad \begin{cases} \underline{\max} & \frac{1}{2} \text{ tr } MY^tAY \\ \underline{\text{s.t.}} & Y\bar{m} = u_n \\ & Y^t u_n = \bar{m} \\ & YM^{1/2} \text{ is } 0,1. \end{cases}$$

Note that Y is $n \times k$ and

$$Y^tY = (XM^{-1/2})^t(XM^{-1/2}) = I.$$

Moreover $Y^tY\bar{m} = \bar{m}$ and $YY^tu_n = u_n$, i.e. 1 is a singular value of Y with right and left singular vectors $\bar{m}$ and u_n.

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. In the subsequent sections we will provide various improved versions of this bound. The bound follows from the scaled program (SP) by relaxing the constraints to $Y^tY = I$. This is a different derivation than the one in [6].

Theorem 2.1 *[6] Let A and m describe a graph partitioning problem. Then*

$$|E_{uncut}| \leq \max\{\frac{1}{2} trMY^tAY : Y^tY = I\} = \frac{1}{2}\sum_{j=1}^{k} \lambda_j(M)\lambda_j(A) = \frac{1}{2}\sum_{j=1}^{k} m_j\lambda_j(A). \qquad (2.3)$$

The proof, using Lagrange multipliers, is similar to that given in [5] for general square matrices A and M. X can be recovered using (2.3). Note that the transportation constraints corresponding to $X \in \mathcal{E}$ as well as nonnegativity constraints are dropped and only $X^tX = M$ is maintained.

Example 1 (continued) In Tables 2.1 and 2.1, we summarize the various bounds for Example 1 in detail. Table 2.1 contains the relevant eigenvalue information and the upper bounds. The corresponding maximizers X are summarized in Table 2.1. Since $\lambda_1(A) = 6.0429$ and $\lambda_2(A) = 3.1375$, we get an upper bound of 45.9019 using Theorem 2.1. Thus no partition leaves more than 45 edges uncut. We point out that the maximizer X does not give any clue on how to obtain a good feasible partition from X.

	Bound	eigenvalues	
Theorem 2.1	45.9019	6.0429	3.1375
Corollary 4.1	42.1219	3.3254	2.1946
Lemma 5.3	38.5516	2.6103	2.6103

Table 2.1: Upper bounds for $|E_{uncut}|$ in Example 1.1. The eigenvalues given are the two largest of A for Theorem 2.1 and of $\hat{A}$ for the remaining bounds.

X_{i1}	X_{i2}	X_{i1}	X_{i2}	X_{i1}	X_{i2}
0.7530	0.4120	0.6406	0.3594	0.4282	0.5718
0.5510	-0.6441	0.0478	0.9522	-0.0952	1.0952
0.9867	-0.2220	0.1052	0.8948	-0.0355	1.0355
0.1746	0.3196	0.7922	0.2078	1.0256	-0.0256
0.6558	1.3613	1.2658	-0.2658	1.0287	-0.0287
0.5349	-0.3579	0.2295	0.7705	0.7049	0.2951
0.6799	1.2116	1.1458	-0.1458	1.0567	-0.0567
0.3638	0.3119	0.7042	0.2958	1.0558	-0.0558
0.5199	1.3605	1.3322	-0.3322	1.0194	-0.0194
0.1085	0.4339	0.8536	0.1464	1.0205	-0.0205
1.0031	-0.0850	0.1915	0.8085	-0.0641	1.0641
0.4967	0.2494	0.6112	0.3888	0.2216	0.7784
0.6593	-0.5396	0.0341	0.9659	-0.0824	1.0824
1.0421	-0.3468	0.0058	0.9942	-0.0477	1.0477
0.2708	0.6644	0.9772	0.0228	1.0196	-0.0196
0.8131	0.4408	0.6033	0.3967	0.9723	0.0277
0.5882	-0.3871	0.1800	0.8200	-0.0918	1.0918
1.2476	-0.8974	-0.3540	1.3540	-0.0163	1.0163
0.3578	0.7968	0.9851	0.0149	0.9113	0.0887
0.9970	-0.9857	-0.3511	1.3511	-0.0316	1.0316

Table 2.1: Three Maximizers X producing the bounds of the previous Table 2.1.

3 Projection of (P)

We now project the feasible set of the problem (P) onto the linear manifold defined by the constraints $\mathcal{E}$. We do this by eliminating the constraints $\mathcal{E}$ 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 [4] for the QAP.

We let P and Q be orthogonal matrices with

$$P = [\frac{u_n}{\sqrt{n}}\ V_n]\ ;\ Q = [\frac{\bar{m}}{\sqrt{n}}\ V_k].$$

For example, we could apply the Gram-Schmidt process to the columns of the full rank matrix

$$\begin{bmatrix} \frac{\bar{m}}{\sqrt{n}} & \begin{matrix} 0 \\ I \end{matrix} \end{bmatrix}$$

to obtain V_k. Note that $V_n V_n^t$ is the orthogonal projection on $\{u_n\}^\perp$ while $V_k V_k^t$ is the orthogonal projection on $\{\bar{m}\}^\perp$. Let

$$P^t AP = \begin{bmatrix} \alpha & a^t \\ a & \hat{A} \end{bmatrix},\ Q^t MQ = \begin{bmatrix} \beta & b^t \\ b & \hat{M} \end{bmatrix},$$

where

$$\begin{array}{ll} \hat{A} = V_n^t A V_n, & \hat{M} = V_k^t M V_k, \\ a = V_n^t r(A)/\sqrt{n}, & b = V_k^t r(M^{3/2})/\sqrt{n}, \\ \alpha = s(A)/n, & \beta = s(M^2)/n = \sum m_j^2/n. \end{array}$$

We define the following program in the variable $Z \in \Re^{(n-1)\times(k-1)}$:

$$(\mathbf{EP}) \quad \left\{ \begin{array}{ll} \underline{\max} & \frac{1}{2}\ \mathrm{tr}\ (\hat{M} Z^t \hat{A} Z + 2 Z^t a b^t + \alpha\beta) \\ \underline{\text{s.t}} & Z^t Z = I, \\ & V_n Z V_k^t \geq -\frac{1}{n} u_n \bar{m}^t. \end{array} \right.$$

We will see that (EP) is equivalent to (P). This follows from the following characterization of the feasible set F of (P).

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*

$$X = P \begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix} Q^t M^{1/2}. \tag{3.1}$$

Then

(a) $X \in \mathcal{E}$

(b) $X \in \mathcal{N} \iff V_n Z V_k^t \geq -\frac{1}{n} u_n \bar{m}^t$

(c) $X^tX = M \iff Z \in \mathcal{O}_{(n-1)\times(k-1)}$.

Proof. First note that expanding (3.1) yields

$$X = \frac{1}{n} u_n u_k^t M + V_n Z V_k^t M^{1/2}. \tag{3.2}$$

Now observe that, since $V_n^t u_n = 0$, we get

$$X^t u_n = \frac{1}{n} M u_k u_n^t u_n + M^{1/2} V_k Z^t V_n^t u_n = M u_k = m.$$

Similarly,

$$X u_k = \frac{1}{n} u_n u_k^t M u_k + V_n Z V_k^t M^{1/2} u_k = u_n,$$

because $V_k^t M^{1/2} u_k = 0$. Thus a) is proved. By (3.2) we can write

$$X = \frac{1}{n} u_n \bar{m}^t M^{1/2} + V_n Z V_k^t M^{1/2}.$$

Thus

$$X \in \mathcal{N} \iff V_n Z V_k^t \geq -\frac{1}{n} u_n \bar{m}^t,$$

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

$$X^tX = M \iff Q \begin{bmatrix} 1 & 0 \\ 0 & Z^t \end{bmatrix} P^t P \begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix} Q^t = I \iff Z \in \mathcal{O},$$

because P and Q are orthogonal. □

This lemma shows in particular that, given an arbitrary $(n-1)\times(k-1)$ matrix Z, there exists a (unique) matrix $X \in \mathcal{E}$, satisfying (3.1). Conversely, any $X \in \mathcal{E}$ uniquely determines Z by

$$Z = V_n^t X M^{-1/2} V_k$$

using (3.2). We prove now that (P) is indeed equivalent to (EP).

Theorem 3.1 *Suppose X and Z satisfy (3.1). Then Z solves $(EP) \iff X$ solves (P).*

Proof. Lemmas 2.1 and 3.1 show that X is feasible for (P) if and only if Z is feasible for (EP). The result follows upon noting that the objective function values of (P) and (EP) are equal for matrices X and Z related by (3.2), i.e.

$$\begin{aligned} tr\ X^tAX &= tr(\tfrac{1}{n} M u_k u_n^t + M^{1/2} V_k Z^t V_n^t) A (\tfrac{1}{n} u_n u_k^t M + V_n Z V_k^t M^{1/2}) \\ &= tr\left[\tfrac{1}{n^2}(u_n^t A u_n)(u_k^t M^2 u_k) + \tfrac{2}{n} Z^t V_n^t A u_n u_k^t M^{3/2} V_k + (V_k^t M V_k) Z^t (V_n^t A V_n) Z\right] . \\ &= \tfrac{1}{n^2} s(A) s(M^2) + tr\left[\hat{M} Z^t \hat{A} Z + \tfrac{2}{n} Z^t V_n^t r(A) r^t(M^{3/2}) V_k\right] \end{aligned}$$

□

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

$$XU = V\Sigma$$

$$X^tV = U\Sigma \ ,$$

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

$$XX^tV = XU\Sigma = V\Sigma^2$$
$$X^tXU = X^tV\Sigma = U\Sigma^2$$

and

$$P^tXQ = \begin{bmatrix} V^tXU & V^tX\bar{U} \\ \bar{V}^tXU & \bar{V}^tX\bar{U} \end{bmatrix}$$

or

$$X = P \begin{bmatrix} \Sigma & 0 \\ 0 & \bar{V}^tX\bar{U} \end{bmatrix} Q^t \ .$$

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

4 Bounds for (P)

Using the equivalent program (EP) instead of of the original problem (P), we get new bounds for (P). First note that due to the elimination of the constraints $\mathcal{E}$ 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 2.1. Therefore we treat the quadratic and linear part separately. The quadratic part is bounded using Theorem 2.1 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 $r(A) = (r_1(A), \ldots, r_n(A))^t$, the vector of row-sums of A, 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*

$$|E_{uncut}| \leq \frac{1}{2}\sum_{j=1}^{k-1} \lambda_j(\hat{A})\lambda_j(\hat{M}) + \frac{1}{n}\sum_{j=1}^{k} R_j(A)m_j - \frac{1}{2n^2}s(A)s(M^2). \tag{4.1}$$

Proof. The quadratic term in (EP) is bounded independently of the linear term by Theorem 2.1, contributing the first summand in (4.1). To bound the linear term of (EP) we observe

$$\begin{aligned} trab^tZ^t &= \tfrac{1}{n}trV_n^tAu_nu_k^tM^{3/2}V_kZ^t \\ &= \tfrac{1}{n}trAu_nu_k^tM^{3/2}(M^{-1/2}X^t - \tfrac{1}{n}M^{1/2}u_ku_n^t) \\ &= \tfrac{1}{n}tr\ \{r(A)r^t(M)X^t\} - \tfrac{1}{n^2}s(A)s(M^2). \end{aligned}$$

The first equality follows from the definition of a and b, the second from (3.2). It is easy to verify that due to the ordering of $r(A)$ and m, the transportation problem

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

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 $\frac{n}{k}$). Then*

$$|E_{uncut}| \leq \max\{\frac{1}{2}trX^tAX : X \in \mathcal{E}, X^tX = \frac{n}{k}I\} = \frac{n}{2k}\sum_{j=1}^{k-1}\lambda_j(\hat{A}) + \frac{1}{2k}s(A).$$

Moreover the bound is attained for

$$X = \frac{1}{k}u_nu_k^t + \sqrt{\frac{n}{k}}V_nZV_k^t$$

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 $M = \frac{n}{k}I$ in (EP) and using the expansion of the linear term, contained in the proof of Theorem 4.1, we get

$$\frac{1}{2}trX^tAX = \frac{n}{2k}trIZ^t\hat{A}Z + \frac{1}{n}tr\ r(A)r^t(M)X^t - \frac{1}{2n^2}s(A)s(M^2).$$

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

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

Thus the linear term is constant:

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

Finally we have

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

Bounding the quadratic term again by Theorem 2.1 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 (3.2). □

It is worth mentioning that the bound from Corollary 4.1 is equivalent to the bound proposed by Boppana [8] 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 $A = A_o$ and the underlying graph is $t-$regular.) Then*

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

Moreover the bound is attained for

$$X = \frac{1}{n} u_n u_k^t M + V_n W U^t V_k^t M^{1/2}$$

where $U \in \mathcal{O}_{k-1,k-1}$ diagonalizes $\hat{M}$ and $W \in \mathcal{O}_{n-1,k-1}$ 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,

$$\begin{aligned} \mathrm{tr}(ab^t Z^t) &= \tfrac{t}{n} tr(u_k^t M)(X^t u_n) - \tfrac{t}{n} s(M^2) \\ &= \tfrac{t}{n} \mathrm{tr}(m^t m) - \tfrac{t}{n} s(M^2) \\ &= 0. \end{aligned}$$

The upper bound for the quadratic term is attained for $Z = WU^t$ where $U \in \mathcal{O}$ diagonalizes $\hat{M}$, and $W \in \mathcal{O}$ contains the eigenvectors corresponding to the $k-1$ largest eigenvalues of $\hat{A}$. X is recovered using (3.2). □

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*

$$|E_{uncut}| \leq \max\{\frac{1}{2} trX^tAX : X \in \mathcal{E}, X^tX = M\} = \max\{z^tCz + c^tz + const : z^tz = 1\},$$

where

$$z \in \Re^{n-1}, \; C = \frac{1}{n} m_1 m_2 \hat{A},$$

$$c = \sqrt{\frac{m_1 m_2}{n}} \frac{m_2 - m_1}{n} V_n^t A u_n,$$

$$const = \frac{1}{2n^2} s(A) s(M^2).$$

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

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

Thus

$$\hat{M} = V_k^t M V_k = \frac{2}{n} m_1 m_2.$$

The quadratic term in (EP) therefore simplifies to z^tCz, and the linear term simplifies to c^tz. □

We point out that the (global) maximum of

$$\{z^tCz + c^tz + const : z^tz = 1\}$$

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

Example 1 (continued) $\lambda_1(\hat{A}) = 3.3254$. Using Corollary 4.1 we get

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

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 Table 3.

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 $d \in \Re^n$ does not affect the graph partitioning problem, see also [2, 6]. 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 $\alpha \in \Re$. Then*

$$trX^tAX = trX^tA(\alpha)X - \alpha n, \quad \forall X : X^tX = M. \tag{5.1}$$

Moreover, the upper bounds from Theorem 2.1 and from Section 4 give the same result when applied to the left hand side and to the right hand side of (5.1).

Proof. The equality (5.1) is obvious. Therefore any bound obtained by maximizing over $X^tX = M$ 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

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

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

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

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

The last relation follows using $V_kV_k^t = I - \frac{1}{n}\bar{m}\bar{m}^t$. Let us denote by EPB(A) the eigenvalue bound of Theorem 4.1 applied to the matrix A. Bounding the right hand side in (5.1) we get

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

□

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 *[2] [6] For $d \in \Re^n$ let $A(d) = A + diag(d)$. Then*

$$trX^tA(d)X = trX^tAX + s(d)$$

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 identical 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

$$\hat{A}(d) := \hat{A} + V_n^t diag(d) V_n,$$

and

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

Donath and Hoffman [6] 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 [15], 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

$$\lim_{\|d\| \to \infty} g(d) = \infty. \tag{5.1}$$

The above discussion is summarized as follows.

Lemma 5.3 *Suppose $m = \frac{n}{k} u_k$. Then*

$$|E_{uncut}| \leq \frac{1}{2k} s(A) + \min\{\frac{n}{2k} \sum_{j=1}^{k-1} \lambda_j(\hat{A} + V_n^t diag(d) V_n) : \ d \in \Re^n, s(d) = 0\} \tag{5.2}$$

□

We point out that the minimum is attained because of (5.1). 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

$$\begin{aligned}\frac{\partial}{\partial d_j}\lambda_i(\hat{A}(d)) &= \frac{\partial}{\partial d_j}(x_i^t V_n^t (A + \sum_{r=1}^{n} d_r e_r e_r^t) V_n x_i) \\ &= x_i^t V_n^t e_j e_j^t V_n x_i \\ &= (e_j^t V_n x_i)^2, \quad \forall j = 1, \ldots, n. \end{aligned} \tag{5.3}$$

Here e_j denotes the j-th canonical unit vector. Otherwise an element of the eigenspace has to be chosen properly, see Theorem 5.1 in [16], to provide the differentials. In the general case, (5.3) 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 (5.2). We used the BT-Method proposed in [17] to carry out the minimization.

Example 1 (continued) The results for the iterative improvement of our example are summarized in Table 5.1. As a stopping criterion, we tested whether a subgradient of norm less than 0.001 was found. This occurred after 67 iterations. We observe that after only a few iterations we have already 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 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 [15]. Finally we point out that the maximizer X, producing the bound, is already very close to a $0,1$ matrix, see Table 3.

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

$$X = \frac{1}{n}u_n u_k^t M + V_n W U^t V_k^t M^{1/2}$$

where $U \in \mathcal{O}_{k-1,k-1}$ diagonalizes $\hat{M}$ and $W \in \mathcal{O}_{n-1,k-1}$ contains the eigenvectors corresponding to the $k-1$ largest eigenvalues of $\hat{A}(d)$.

iter.	bound	norm
1	42.13	1.065
5	41.11	0.598
10	39.38	0.338
15	39.23	0.194
20	38.74	0.124
25	38.63	0.202
30	38.56	0.072
35	38.56	0.019
40	38.55	0.010
45	38.55	0.007
50	38.55	0.005
55	38.55	0.003
60	38.55	0.0014
65	38.55	0.0013
67	38.55	0.0005

Table 5.1: 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 a subgradient for $g(d)$ found at the given iteration.

Proof. First note that $s(d) = 0$. Moreover, by the definition of d, u_n is an eigenvector of $A(d)$ with corresponding eigenvalue $\frac{1}{n}s(A)$. 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 in 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 $X^tX = M$, 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 $Y^tY = M$ as well. Therefore

$$\begin{aligned} ||X - Y||^2 &= ||X||^2 + ||Y||^2 - 2trX^tY \\ &= 2trM - 2trX^tY \ . \end{aligned}$$

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

$$min\{-trX^tY : Y \in F\}$$

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

We point out that this idea is also (implicitly) used by Barnes [11] 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 $X^tX = M$. This approximation model has the disadvantage however, 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 makes only 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*

$$Y = \begin{bmatrix} 1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,1\,0\,0\,1\,1\,0\,1 \\ 0\,0\,0\,1\,1\,1\,1\,1\,1\,1\,0\,0\,0\,0\,1\,1\,0\,0\,1\,0 \end{bmatrix}^t$$

of value 38. Comparing with the bounds in Table 2 we conclude that Y is already optimal, because the upper bound from Lemma 5.3 also becomes 38 after rounding down. We point out that this solution Y can also be obtained by simply rounding the maximizer X to the nearest integers, see Table 3.

6.2 Linear Approximation

Expanding the objective function at X we get

$$trY^tAY = trX^tAX + 2trX^tA(Y-X) + tr(Y-X)^tA(Y-X) .$$

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

$$max\{trX^tAY : Y \in F\}$$

If A is positive definite, then the weighted Frobenius norm

$$||A^{1/2}(X-Y)||^2 = ||A^{1/2}X||^2 + ||A^{1/2}Y||^2 - 2trX^tAY \ ,$$

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

Barnes, Vanelli and Walker [2] use a feasible X and try to improve it. They propose a diagonal perturbation that changes A to a positiv 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.

7 Computational Results

In this section we present computational experiences for the various new bounds. First we apply our bounds to problems that have been published and studied previously. In particular we use the following graphs described in Table 7.1. Note that G2 is our running Example 1.

In Table 7.1 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.

Next we investigate our bounds for the weighted graph from [9], page 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 [9]. 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 3 variants V1, V2 and V3 for this problem:

V1: all edge costs are 1

Name	$\|V\|$	$\|E\|$	Source
G1	20	55	[6], Table 2 p.425
G2	20	51	[6], Table 3 p.425
G3	20	46	[15], Table 1 p.52
G4	21	48	[2], Figure 2 p.305

Table 7.1: Graphs from the literature

Graph	k	Thm 2.1	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

Table 7.1: Partitioning into k sets of equal size

V2: use C_1 for the edge costs

V3: use C_2 for the edge costs

In Table 7.1 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 2.1. One reason for this may lie in the regular structure of the graph, which contains many optimal partitions. We also point out that the bound from Theorem 5.1 is quite competitive with the subgradient improved bound from Lemma 5.3 for the variants V2 and V3. The optimal solution values are from [9].

To further examine the performance of these bounds we generated a series of pseudorandom graphs of larger sizes. We generated 5 graphs, each of average degree 5 for $n \in \{30, 40, 50\}$. The results are summarized in Table 7.1.

Comparing again the last two columns, it turns out that the new bounding rules constitute

Variant	$\sum a_{ij}/2$	Thm 2.1	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

Table 7.1: Three Variants of a 40 node graph

n	$\|E\|$	Thm 2.1	Cor. 4.1	Lemma 5.3	feas. X
30	75	66.65	62.51	58.10	56
30	75	72.51	65.58	59.77	58
30	84	77.61	72.15	65.25	63
30	73	67.43	62.47	58.06	56
30	69	64.54	60.26	56.73	54
40	110	101.23	93.84	87.17	82
40	102	97.33	89.70	82.98	79
40	102	102.93	95.81	87.52	86
40	91	89.96	83.11	76.56	73
40	101	95.52	87.32	80.86	77
50	139	128.17	118.17	110.43	105
50	117	117.41	106.28	95.45	90
50	123	117.51	108.57	100.86	96
50	128	120.53	109.73	103.38	98
50	138	126.97	120.56	112.47	108

Table 7.1: Partitioning of pseudorandom graphs into two blocks of equal size

a significant improvement over the previously known techniques. The gap for the problems with 30 nodes is never larger than two edges, for problems with 50 nodes it never exceeds 5 edges.

It seems more difficult to find good bounds if the blocksizes 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 sizes. We take the graph G2, our running example, and partition it into 2 blocks of different sizes. The numerical results are summarized in the following Table 7.1. Since G2 has only 51 edges we conclude that in the case of "very unequal" blocksizes 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 strategies to derive bounds for the graph partitioning problem. Starting from the orthogonal relaxation (Theorem 2.1) the main improvement was achieved by the additional restriction to the set $\mathcal{E}$ of matrices with prescribed row and column sums. We presented two general new bounds (Theorems 4.1 and 5.1), and studied the following special cases in more detail:

m_1	Thm. 2.1	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

Table 7.1: Partitioning G2 into two blocks of sizes m_1 and $20 - m_1$

- 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)

In each of these special cases we were able to further strengthen the general bounds. It is a challenging research task to use these new bounds in a Branch and Bound program to solve the graph partitioning problem to optimality.

Acknowledgement: We thank Helga Schramm for the permission to use her code for the BT method.

References

[1] T. LENGAUER. *Combinatorial algorithms for integrated circuit layout.* John Wiley and Sons, Chicester, 1990.

[2] E.R. BARNES, A. VANNELLI, and J.Q.WALKER. A new heuristic for partitioning the nodes of a graph. *SIAM J. Discrete Mathematics*, 1:299–305, 1988.

[3] S.W. HADLEY. *Continuous optimization approaches for the quadratic assignment problem.* PhD thesis, University of Waterloo, Waterloo, Canada, 1989.

[4] S.W. HADLEY, F. RENDL, and H. WOLKOWICZ. A new lower bound via projection for the quadratic assignment problem. *Mathematics of Operations Research*, 17:727–739, 1992.

[5] F. RENDL and HENRY WOLKOWICZ. Applications of parametric programming and eigenvalue maximization to the quadratic assignment problem. *Mathematical Programming*, 53:63–78, 1992.

[6] W.E. DONATH and A.J. HOFFMAN. Lower bounds for the partitioning of graphs. *IBM J. of Research and Developement*, 17:420–425, 1973.

[7] E.R. BARNES and A.J. HOFFMAN. Paritioning, spectra, and linear programming. In W.E. Pulleyblank, editor, *Progress in Combinatorial Optimization*, pages 13–25. Academic Press, 1984.

[8] R.B. BOPPANA. Eigenvalues and graph bisection: An average case analysis. In *Proceedings of the 28th Annual Symposium on Computer Science*, pages 280–285. IEEE, 1987.

[9] N. CHRISTOFIDES and P. BROOKER. The optimal partitioning of graphs. *SIAM J. Applied Mathematics*, 30:55–69, 1976.

[10] C. ROUCAIROL and P. HANSEN. Cut cost minimization in graph partitioning. In C. Brezinski, editor, *Numerical and Applied Mathematics*, pages 585–587. J.C. Baltzer AG, 1989.

[11] E.R. BARNES. An algorithm for partitioning the nodes of a graph. *SIAM J. Algebraic and Discrete Mathematics*, 3:541–550, 1982.

[12] W. GANDER. Least squares with a quadratic constraint. *Numer. Math.*, 36:291–307, 1981. This is the abbreviated version of the technical report published in 1978.

[13] W. GANDER, G.H. GOLUB, and U. von MATT. A constrained eigenvalue problem. *Linear Algebra and its Applications*, 114/115:815–839, 1989.

[14] J.J. MORÉ and D.C. SORENSEN. Computing a trust region step. *SIAM J. Sci. Statist. Comput.*, 4:553–572, 1983.

[15] J. CULLUM, W.E. DONATH, and P.WOLFE. The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices. *Mathematical Programming Study*, 3:35–55, 1975.

[16] B. GOLLAN. Eigenvalue perturbations and nonlinear parametric optimization. *Mathematical Programming Studies*, 30:67–81, 1987.

[17] H. SCHRAMM and J. ZOWE. A combination of the bundle approach and the trust region concept. In J. Guddat et al., editor, *Advances in Mathematical Optimization.* Akademie Verlag Berlin, 1988.