4OR-Q J Oper Res (2012) 10:321–346
DOI 10.1007/s10288-012-0210-3

INVITED SURVEY

Semidefinite relaxations for partitioning, assignment and ordering problems

F. Rendl

Received: 16 April 2012 / Revised: 17 July 2012 / Published online: 11 August 2012

Abstract Semidefinite optimization is a strong tool in the study of NP-hard combinatorial optimization problems. On the one hand, semidefinite optimization problems are in principle solvable in polynomial time (with fixed precision), on the other hand, their modeling power allows to naturally handle quadratic constraints. Contrary to linear optimization with the efficiency of the Simplex method, the algorithmic treatment of semidefinite problems is much more subtle and also practically quite expensive. This survey-type article is meant as an introduction for a non-expert to this exciting area. The basic concepts are explained on a mostly intuitive level, and pointers to advanced topics are given. We provide a variety of semidefinite optimization models on a selection of graph optimization problems and give a flavour of their practical impact.

Keywords Combinatorial optimization · Semidefinite programming · Graph partition problem · Ordering problem

MSC classification 90C27 · 90C22

1 Introduction

Semidefinite optimization has turned out to be a strong modeling tool in applied mathematics. One of the first survey articles, by Vandenberghe and Boyd (1996), features a variety of applications which receive scientific attention also now, more than 15 years after publication.

F. Rendl (✉)
Institut für Mathematik, Alpen-Adria Universität Klagenfurt,
9020 Klagenfurt, Austria
e-mail: rendl@aau.at

Here we are going to focus on applications of semidefinite optimization in connection with combinatorial optimization, more precisely in approximating NP-hard problems using relaxations based on semidefinite matrices. The main impulses for progress in this area come from the algorithmic developments in the early 1990's to solve semidefinite optimization problems based on interior-point methods, see for instance the work of Nesterov and Nemirovski (1994), and independently from the approximation analysis of randomized hyperplane rounding for Max-Cut by Goemans and Williamson (1995), which impressively demonstrates the theoretical power of semidefinite relaxations in combinatorial optimization.

1.1 Linear and semidefinite relaxations

In order to explain semidefinite relaxations, it is instructive to first briefly review purely linear relaxations. To set the stage, we introduce an abstract combinatorial optimization problem (COP) in the following way. Given some finite set E we consider a family $\mathcal{F}$ of subsets of E to represent the feasible solutions of COP. In the case of the *Minimum spanning tree problem* for example, the set E would be the edge set of the graph G under consideration and $\mathcal{F}$ consists of all subsets of E corresponding to trees in G. We also associate costs c_e for each $e \in E$. If we define the cost of $F \in \mathcal{F}$ to be

$$c(F) := \sum_{e \in F} c_e,$$

then COP amounts to minimizing (or maximizing) $c(F)$ over $F \in \mathcal{F}$,

$$\text{(COP)} \quad z_{COP} := \min\{c(F) : \ F \in \mathcal{F}\}.$$

To obtain linear relaxations of COP, we embed it in $\mathbb{R}^{|E|}$, by assigning to each feasible solution F its characteristic vector $x_F \in \{0, 1\}^{|E|}$ with $(x_F)_e = 1$ if $e \in F$ and $(x_F)_e = 0$ otherwise. The main task now consists in getting at least a partial description of

$$\mathcal{P} := conv\{x_F : F \in \mathcal{F}\},$$

the convex hull of all characteristic vectors x_F of feasible solutions. In general, a representation of $\mathcal{P}$ in the form

$$\mathcal{P} = \{x : Ax \leq b\}$$

of linear inequalities is unlikely to be tractable in the case that COP is itself an NP-hard problem. In practise however, a partial description of $\mathcal{P}$ can be quite useful. Looking at the *Traveling salesman problem* for instance, such a partial description was developed in the 1970's and has turned out to be a remarkably strong basis to solve this problem to optimality. The comprehensive monograph by Applegate et al. (2006) features the computational state of the art to solve traveling salesman problems.

For other problems, like the *Max-Cut problem*, this approach is less successful and turns out to be practical mostly on instances where the underlying graph is very sparse, see for instance Barahona et al. (1989), Pardalos and Rodgers (1990). The Max-Cut problem is also an instructive instance to study a variety of ways to set up integer programming formulations of the problem. Given the weighted $n \times n$ adjacency matrix $A = (a_{ij})$ of a graph, the task is to partition the vertices into two sets so as to maximize the total weight of edges joining the two partition blocks, see Sect. 3 below for details.

A natural way to formulate this problem in a polyhedral setting consists in introducing 0–1 edge variables y_{ij}, indicating whether or not the edge $[i, j]$ is cut. This leads to the linear objective function $\sum_{i<j} a_{ij} y_{ij}$ to be maximized. The catch here is that the polytope given by the cut vectors y of edges which join two vertex partitions is unlikely to have a tractable representation by linear inequalities.

In contrast, a vertex based way to model edge cuts assigns to each vertex of one partition block the value +1 and the vertices of the other block get the value −1. Thus edge cut vectors y are in one-to-one correspondence with vectors $x \in \{-1, 1\}^n$ through $y_{ij} = x_i x_j$ for $i < j$. The simplicity of modeling cuts here is paid for by the fact that now the value of the cut y, given through x, becomes a quadratic function in x, see again Sect. 3 below.

Hence either the convex hull of cut vectors in the edge model needs to be studied to get at least a partial description of its convex hull, or the objective function in the vertex based model has to be linearized in some way to apply polyhedral approaches. We refer to Barahona et al. (1985), Barahona and Mahjoub (1986) for some basic polyhedral investigations based on the edge model, and to Padberg (1989), De Simone (1990) for linearizations of the quadratic cost function underlying the vertex based model. In both cases the computational evidence indicates that the polyhedral approach leaves room for improvement, especially in case that the underlying graph is dense, i.e. if the adjacency matrix A contains only few zero entries.

Let us now turn to the key features of semidefinite relaxations of a COP, constrasting the linear approach just described. Given the characteristic vector $x_F \in \{0, 1\}^{|E|}$, we now consider the polytope

$$\mathcal{M} := conv \left\{ x_F x_F^T : F \in \mathcal{F} \right\}. \tag{1}$$

If the ground set E consists of n elements, the set $\mathcal{F}$ is contained in the space of symmetric $n \times n$ matrices. Since

$$\operatorname{diag}(x_F x_F^T) = x_F \tag{2}$$

holds for any 0–1 vector x_F, the projection of elements in $\mathcal{M}$ to their main diagonal yields $\mathcal{P}$, so moving to $\mathcal{M}$ can indeed be viewed as a generalization of the polyhedral approach. (A summary of the notation is given at the end of this section, $\operatorname{diag}(M)$ denotes the main diagonal of the matrix M). The key additional features moving to $\mathcal{M}$ lie in the following two simple facts.

Semidefiniteness Any matrix $M \in \mathcal{M}$ is positive semidefinite.

Quadratic becomes linear Any quadratic term in x_F becomes linear for $\mathcal{M}$. More precisely, the quadratic form $x_F^T A x_F = \langle x_F x_F^T, A \rangle = \langle X, A \rangle$ for some $X \in \mathcal{M}$.

To provide an example for this approach, we consider the problem QP of *quadratic 0–1 optimization*, defined through a symmetric matrix A (of order n) as

$$z_{QP} := \max\{x^T Ax : x \in \{0, 1\}^n\}. \tag{3}$$

The polytope

$$conv \left\{ \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T : x \in \{0, 1\}^n \right\}$$

is contained in the convex set

$$\left\{ Y : Y = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0, \ \operatorname{diag}(X) = x \right\}$$

and we get the following semidefinite relaxation of QP

$$z_{QP} \le z_{QP}^{sdp} := \max \left\{ \langle A, X \rangle : \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0, \ \operatorname{diag}(X) = x \right\}. \tag{4}$$

It should be noted that instead of working with the vector x we introduced an additional homogenization variable which is set to 1 yielding $\begin{pmatrix} 1 \\ x \end{pmatrix}$. The purpose of this additional variable will become clear from (19) below. The equation $\operatorname{diag}(X) = x$ reflects (2) and the objective function becomes linear in the matrix variable X, which has taken the role of xx^T.

The idea of approaching COP by moving to relaxations based on $\mathcal{M}$ has a long history. Lovász (1979) explores this idea to get tractable bounds for the chromatic number and the clique number in graphs. Later, Lovász and Schrijver (1991) and Sherali and Adams (1990, 1994) provide a systematic study of this approach in connection with optimization in binary variables and Shor (1987) investigates relaxations of quadratic programming. Finally, the analysis of hyperplane rounding for Max-Cut Goemans and Williamson (1995) leading to the performance measure (26) below also got the theoretical computer science community interested in semidefinite relaxations.

It is the purpose of this paper to summarize some recent experience with *matrix liftings* based on $\mathcal{M}$ in the context of several NP-hard combinatorial optimization problems. At this point it should be clear that relaxations using $\mathcal{M}$ rather than $\mathcal{P}$ are not advisable in case that the integer programming formulation of COP is purely linear, as is the case for the (linear) *knapsack problem*. For a recent progress report in this direction, see for instance the monograph edited by Lee and Leyffer (2012). In the following sections we will investigate the use of matrix space embeddings for a variety of problems having some quadratic structure in their integer programming formulations.

1.2 Literature and notation

We close the introduction with a few pointers to survey-type papers on semidefinite optimization, and some notation used throughout.

After the rapid development of interior point methods for semidefinite programs (SDP), stimulated by the work of Nesterov and Nemirovski (1994) in the 1990's, one of the first monographs dedicated to SDP under these new theoretical and algorithmic aspects Wolkowicz et al. (2000) features the state of the art up to the year 2000. The textbook by de Klerk (2002) gives a more concise presentation of the material. More recently, Tunçel (2010) covers progress related to combinatorial optimization and the monograph edited by Anjos and Lasserre (2012) can be viewed as a follow up to Wolkowicz et al. (2000) by focussing on the recent developments up to 2011.

Several survey type papers on the topic are also available and reflect the wide range of applications of SDP in applied sciences. The surveys by Helmberg (2002), Lovász (2003), Laurent and Rendl (2005) all deal with combinatorial applications in the general spirit of the present paper.

Finally, it may also be of interest to mention SDP surveys which cover topics not included in the present paper. Anjos (2005) looks at SDP in connection with satisfiability problems and Vallentin (2008) investigates SDP which are invariant under group action and shows some applications of SDP in connection with geometric packing problems. Finally, in the last years there was strong interest in the connection between SDP, the sum of squares representations of polynomials, and the moment problem from probability theory, see for instance the monograph by Lasserre (2010). The study of sets which can be represented as the feasible region of an SDP joins convex algebraic geometry and optimization. This is currently a hot research area, and the reader is referred to the monograph (Blekherman et al. 2012) in preparation.

Notation: The all-ones vector is denoted by e, and $J = ee^T$ is the all-ones matrix of appropriate size. The inner product in $\mathbb{R}^n$ as well as in the space of $n \times n$ matrices is denoted by $\langle \cdot, \cdot \rangle$. thus $\langle a, b \rangle = a^T b$ for vectors $a, b \in \mathbb{R}^n$ and $\langle A, B \rangle = \mathrm{tr}(A^T B)$ for matrices A, B. If M is a square matrix, then $\mathrm{diag}(M)$ denotes the vector of the main diagonal entries of M, and for a vector $m \in \mathbb{R}^n$, $\mathrm{Diag}(m)$ is the diagonal matrix with m on the main diagonal.

The set of positive semidefinite matrices contains all symmetric matrices X (of a given order n) such that $a^T X a \geq 0 \; \forall a \in \mathbb{R}^n$. We also write $X \succeq 0$ in this case. The statement $X - Y \succeq 0$ is also expressed as $X \succeq Y$. Positive definiteness is denoted by $X \succ 0$ and expresses $a^T X a > 0 \; \forall a \neq 0$. We will consider undirected and loopless graphs $G = (V, E)$ given through their vertex set V and edge set E. If $e \in E$, we also write $e = [i, j]$ in case that e joins the vertices i and j.

2 Semidefinite optimization

We briefly recall some basic facts about linear semidefinite optimization problems (SDP), and refer to Anjos and Lasserre (2012), Wolkowicz et al. (2000) for a more elaborated treatment of the topic. In the simplest form of SDP we are given symmetric matrices $C, A_1, \ldots, A_m$ and $b \in \mathbb{R}^m$. The problem

$$\inf\{\langle C, X\rangle : \langle A_i, X\rangle = b_i \; i = 1, \ldots, m, \;\; X \succeq 0\} \tag{5}$$

is called a (linear) semidefinite program. The feasible region is given by the intersection of an affine space with the cone of semidefinite matrices, hence it is convex but in general not polyhedral. The study of sets which have such a representation is currently an active area of research, see for instance Helton and Nie (2010).

The optimal value for the problem (5) is not necessarily attained, therefore we take the infimum. Examples where the optimum is not attained can be found for instance in Vandenberghe and Boyd (1996). The duality theory of linear programming generalizes naturally to SDP, but becomes substantially more subtle. The (Lagrangian) dual associated to (5) is given as

$$\sup\left\{b^T y : C - \sum_i y_i A_i \succeq 0\right\}. \tag{6}$$

Weak duality (sup $\leq$ inf) holds by construction of the dual. Strong duality (sup $=$ inf), as well as attainment of the respective optima requires some sort of regularity of the feasible regions. We refer to Duffin (1956), and to the handbook Wolkowicz et al. (2000) on semidefinite programming for a thorough discussion of SDP with respect to theoretical properties and algorithms. The existence of feasible points in the interior of the primal and dual feasible region ensures the following characterization of optimality. We write $A(X) = b$ for the equations in (5). The linear map A has an adjoint A^T, defined through the adjoint identity

$$\langle A(X), y\rangle = \langle X, A^T(y)\rangle.$$

The adjoint can be expressed as $A^T(y) = \sum_i y_i A_i$.

Theorem 1 (*Duffin 1956; Wolkowicz et al. 2000*) *Suppose there exists $X_0 \succ 0$ such that $A(X_0) = b$ and there is y_0 such that $C - A^T(y_0) \succ 0$. Then the optima in (5) and (6) are attained. Moreover, X and y are optimal if and only if $A(X) = b$, $X \succeq 0$, $Z := C - A^T(y) \succeq 0$ and the optimal objective values coincide, $\langle X, Z\rangle = 0$.*

In view of this result we are interested to formulate relaxations which contain strictly feasible points in both the primal and the dual problem. Moreover, it will turn out useful to allow also linear inequalities in (5).

The most reliable way to solve (5), (6) in the presence of strictly feasible points is based on interior-point methods. In mathematical terms, this amounts to applying the Newton method to the optimality conditions of the theorem, with $\langle X, Z\rangle = 0$ replaced by (some variant of) the parametrized equation $ZX = \mu I$, where the parameter $\mu > 0$ is driven to zero. The computational effort in each iteration of the Newton method depends on the matrix order n and on the number m of constraints in (5). It includes several $O(n^3)$ operations with matrices of order n as well as solving a linear system of order m. To give the big picture, interior-point methods tend to work nicely as long as n and m are not too large, $n \leq 200$ and $m \leq 2{,}000$ should do. Once $n \geq 1{,}000$ or $m \geq 10{,}000$, the use of interior-point methods is likely to be impractical on standard computing facilities. The website maintained by Mittelmann[1] provides up to date information about the current state of the art of SDP solvers, and it also features benchmark computations using a variety of codes and problem classes.

3 Partitioning and clustering as quadratic binary optimization

In this section we show how semidefinite optimization can be used to model various partition and clustering problems which can be cast in the following general form. We are given a symmetric $n \times n$ matrix $A = (a_{ij})$. For simplicity we assume that all a_{ij} are integers. Let us associate a weighted graph with vertex set $\{1, \ldots, n\}$ to A in the usual way: If $i < j$ and $a_{ij} \neq 0$ we introduce the edge $[i, j]$ with weight a_{ij} in the graph. Thus we interpret A as the weighted adjacency matrix of a graph G with vertex set $V(G) = \{1, \ldots, n\}$.

In graph-theoretical terms partition problems ask to partition $V(G)$ so as to optimize the total weight of edges joining different partition blocks. The precise nature of the problem depends on how feasible partitions are specified. This may involve a limit on the number k of partition blocks as well as possible limits on the cardinalities of the individual partition blocks. We first describe the problems as quadratic optimization problems in binary variables and then consider relaxations involving semidefinite matrix variables, following the basic outline described in the introduction.

A k-partition $(V_1, \ldots, V_k)$ of $V = \{1, \ldots, n\}$ consists of nonempty subsets $V_i \subseteq V$ whose union is V. Thus k-partitions are in one-to-one correspondence with $n \times k$ matrices $S = (s_1, \ldots, s_k)$ with $s_1 + \cdots + s_k = e$ and each $s_i \in \{0, 1\}^n$. We recall that e is the all-ones vector of appropriate size. The column s_i is the incidence vector of partition block V_i. Let us define

$$P_{n,k} := \{S : \ S = (s_1, \ldots, s_k), \ \forall i \ s_i \in \{0, 1\}^n, \ Se = e\}. \tag{7}$$

Matrices $S \in P_{n,k}$ are therefore called *partition matrices*.

The total weight of edges joining distinct partition blocks can be expressed conveniently using the Laplacian L_A, associated to the matrix A, which is defined as follows

$$L_A := \mathrm{Diag}(Ae) - A. \tag{8}$$

In case that the weight matrix A is clear from the context, we simply write L instead of L_A. It should also be noted that the definition of the Laplacian L implies that $Le = 0$, and that $L \succeq 0$ in case that $A \geq 0$ elementwise. Let $S = (s_1, \ldots, s_k) \in P_{n,k}$, then

$$s_i^T L s_i = \sum_{v \in V_i, w \notin V_i} a_{vw} \tag{9}$$

[1] http://plato.asu.edu/bench

gives the weight of edges joining V_i to $V \backslash V_i$. Hence

$$\frac{1}{2}\langle S, LS\rangle \tag{10}$$

gives the total weight of edges joining distinct partition blocks given by the partition matrix S.

3.1 Max-Cut and Max-k-Cut

In the simplest form of the problem, only the number $k \geq 2$ of partition blocks is specified and the objective is to maximize (10). This problem is called *Max-k-Cut*. Thus we get Max-k-Cut as the following optimization problem in 0–1 variables involving linear constraints and a quadratic objective function

$$z_{k\text{-}cut} := \max\left\{\frac{1}{2}\langle S, LS\rangle : \; S \in P_{n,k}\right\}. \tag{11}$$

The case $k = 2$ is usually simply called *Max-Cut*. In view of (9), we get

$$z_{mc} := \max\left\{s^T Ls : s \in \{0, 1\}^n\right\}. \tag{12}$$

The variable transformation $y = 2s - e$ together with the fact that $Le = 0$ gives the more familiar formulation of Max-Cut as a problem in $-1, 1$ variables

$$z_{mc} = \max\left\{\frac{1}{4} y^T Ly : \; y \in \{-1, 1\}^n\right\}. \tag{13}$$

3.2 k-Equicut and cardinality constrained k-partition

In some situations not only the number k of partition blocks, but also the sizes m_i of the partition blocks V_i are specified. In case of k-Equicut, V has to be partitioned into k blocks of equal size, $m_1 = \cdots = m_k = \frac{n}{k}$. In the general case, we are given integers k and $m_1, \ldots, m_k$ such that $\sum m_i = n$ and feasible partitions $(V_1, \ldots, V_k)$ satisfy $|V_i| = m_i$. This leads to the following formulation of the problem

$$z_{k\text{-}gp} = \min\left\{\frac{1}{2}\langle S, LS\rangle : \; S = (s_1, \ldots, s_k) \in P_{n,k}, \; \forall i \; s_i^T e = m_i\right\}. \tag{14}$$

The case $k = 2$ again can be simplified and has natural formulations both in 0,1 variables and in $-1, 1$ variables. With the transformation $2s = y + e$, we get similar to (13) and (12) the following

$$z_{2\text{-}gp} = \min\left\{s^T L s : \ s \in \{0,1\}^n, \ s^T e = m_1\right\}$$
$$= \min\left\{\frac{1}{4} y^T L y : \ y \in \{-1,1\}^n, \ e^T y = 2m_1 - n\right\}.$$

This type of problem has wide applications, for instance in numerical linear algebra. The exchange heuristic introduced by Kernighan and Lin (1970) is still considered a very efficient technique to quickly find good partitions.

3.3 k-densest subgraph

A final variant of the problem has the flavour of clustering. The goal here is to find a subset V_1 of V consisting of k vertices, such that the total weight of edges in the subgraph spanned by V_1 is maximized. This leads to

$$z_{k\text{-}ds} := \max\left\{\frac{1}{2} s^T A s : \ s \in \{0,1\}^n, \ s^T e = k\right\}. \tag{15}$$

Formally, this is 0–1 quadratic programming as in (3), with the additional constraint that exactly k of the variables can be set to one.

All the partition problems considered here share the following general features. A quadratic cost function is optimized over binary variables and possibly subject to some equality constraints. Moreover they are all NP-complete problems. Max-Cut for instance was already among the problems in Karp's original list, see Karp (1972). The k-densest subgraph problem appears under various names in the literature, and it is NP-hard even for very restricted classes of graphs, such as planar graphs (Keil and Brecht 1991) or unweighted bipartite graphs with maximal degree three.

Before looking at their matrix relaxations, we investigate problems which are based on permutations, leading to ordering and assignment problems.

4 Ordering and assignment problems

Assignment and ordering problems have a natural representation through permutations. Let Π_n denote the set of permutations of $N = \{1, \ldots, n\}$. For $\pi \in \Pi_n$, the set $(1, \pi(1)), \ldots, (n, \pi(n))$ defines an assignment of the elements of N in the sense that element k is assigned to position $\pi(k)$ for $k \in N$. Assignment problems ask to find an assignment of minimum cost. In case of the *linear assignment problem (LAP)*, the cost function is built up from individual profits $C = (c_{ij})$ describing the profit c_{ij} of the assignment (i, j), i.e. $\pi(i) = j$. Hence we get the classical (linear) assignment problem (LAP) as

$$z_{LAP} := \max\left\{\sum_i c_{i\pi(i)} : \pi \in \Pi\right\},$$

which is one of the most fundamental tractable problems in combinatorial optimization. An informative survey on the history of LAP was compiled by Schrijver, see chapter 17.5 in Schrijver (2003).

The *quadratic assignment problem (QAP)* has cost elements of the form $D = (d_{ij,kl})$ describing the profit resulting from assigning i to $\pi(i) = k$ and j to $\pi(j) = l$. The QAP therefore has the following combinatorial formulation,

$$z_{QAP} := \max \left\{ \sum_{i,j} d_{ij,\pi(i)\pi(j)} : \pi \in \Pi \right\}.$$

To get an integer programming formulation of assignment problems, it is natural to express permutations π through the associated permutation matrix $X = (x_{ij})$ with $x_{ij} \in \{0, 1\}$ and $x_{ij} = 1$ precisely if $j = \pi(i)$, i.e. if i is assigned to j. This leads to the well-known formulation of LAP as the linear program

$$z_{LAP} = \max \left\{ \sum_{i,j} c_{ij} x_{ij} : \; X = (x_{ij}) \in \Omega \right\}$$

defined over the polytope Ω of doubly stochastic matrices,

$$\Omega := \{X : \; Xe = X^T e = e, \; X \geq 0\}$$

which is by Birkhoff's theorem the convex hull of the set of permutation matrices. Incidentally, this is one of the few cases where the concept of linear relaxations of COP, described in the introduction, works nicely. The quadratic assignment problem has the following integer programming formulation

$$z_{QAP} = \max \left\{ \sum_{i,j,k,l} d_{ij,kl} x_{ik} x_{jl} : X = (x_{ij}) \in \Omega, \; \forall i, j \; x_{ij} \in \{0, 1\} \right\}.$$

While LAP can be solved in $O(n^3)$ time using for instance augmenting path techniques, the QAP is not only NP-complete, but also difficult to approximate. The recent monograph (Burkard et al. 2009) summarizes the current state of the art with respect to both LAP and QAP and further variations on assignment problems. In particular, it provides an overview of the development of algorithms to solve LAP, starting from the classical 'Hungarian Method'.

We can interpret $\pi \in \Pi_n$ also as defining the *ordering*

$$\pi(1) \prec \pi(2) \prec \cdots \prec \pi(n).$$

Ordering problems ask for orderings of maximal profit. In the simplest case there is a profit c_{ij} in case that i comes before j in the ordering. Thus the *linear ordering problem (LOP)* is given as

$$z_{LOP} := \max \left\{ \sum_{i<j} c_{\pi(i)\pi(j)} : \pi \in \Pi_n \right\}.$$

The *quadratic ordering problem* has profits $d_{ij,kl}$ for $i \neq j, k \neq l$ which are gained in case that $i \prec j$ and $k \prec l$ in the ordering. Thus it amounts to

$$z_{QOP} := \max \left\{ \sum_{i<j,k<l} d_{\pi(i)\pi(j),\pi(k)\pi(l)} : \pi \in \Pi_n \right\}.$$

To get an integer formulation of LOP, permutation matrices are no longer useful, because we not only need to know the position of item i in the ordering, but also its relative position to other items j in the ordering. The natural integer programming formulation therefore uses indicator variables $x_{ij} \in \{0, 1\}$, indicating whether or not i is before j (for $i \neq j$). It is not hard to verify that such a set of variables x_{ij} describes an ordering if there is no configuration of the form that i comes before j, j comes before k and k comes before i. In other words, x_{ij}, x_{jk} and x_{ki} can not all at the same time have the same value. Moreover $x_{ij} + x_{ji} = 1$ because either i is before j or the other way around. This leads to the 0–1 integer programming formulation of LOP

$$z_{lop} := \max \left\{ \sum_{i \neq j} c_{ij} x_{ij} : x_{ij} \in \{0, 1\}, \ \forall i < j \ x_{ji} = 1 - x_{ij}, \right.$$
$$\left. \forall i < j < k \ 0 \leq x_{ij} + x_{jk} - x_{ik} \leq 1 \right\}.$$

The linear relaxation is obtained by dropping the integrality condition on the variables. The monograph by Marti and Reinelt (2011) provides a recent status report about the linear ordering problem.

In the context of matrix liftings, it will turn out useful to switch to $-1, 1$ variables through $y_{ij} = 2x_{ij} - 1$. We use $y_{ij} + y_{ji} = 0$ to eliminate the variables y_{ij} with $i > j$. The set F_{OP} of feasible solutions y for the ordering problem is now described as follows

$$F_{OP} = \{ y \in \mathbb{R}^{\binom{n}{2}} : y_{ij} \in \{-1, 1\}, \ \forall 1 \leq i < j < k \leq n \ |y_{ij} + y_{jk} - y_{ik}| = 1 \}.$$

The equation expresses the fact that out of the three terms y_{ij}, y_{jk}, y_{ki} exactly two are equal to one another. Thus QOP has the following integer programming formulation

$$z_{QOP} = \max \left\{ \sum_{i<j,k<l} d_{ij,kl} y_{ij} y_{kl} : y \in F_{OP} \right\}.$$

Again this problem has a quadratic cost function. The feasible region described by F_{OP} will yield a natural matrix lifting relaxation, as shown in the following section. The quadratic ordering problem has applications in several areas. Since it captures the effect of distinguishing whether i comes before j and k before l or not, it can be used to minimize the number of crossings when drawing a graph with vertices lying in prespecified layers. This problem is referred to as *multi-level-crossing minimization*. Jünger and Mutzel (1997) consider the two-layer version based on polyhedral combinatorics in combination with Branch and Bound and Buchheim et al. (2009) present recent computational experiments based on SDP relaxations. The multi-layer version was considered in Jünger et al. (1997) using linear programming techniques. SDP approaches for the multi-layer problem were investigated only recently, see Chimani et al. (2011) and the recent dissertation (Hungerländer 2012). Other application areas include the *linear arrangement problem* and the *single-row facility location problem*.

5 Matrix relaxations

5.1 Partitioning and clustering

All the problems (11), (13), (14) and (15) are well-known to be NP-complete. It is therefore natural to consider tractable relaxations. Polyhedral relaxations for these problems were studied in the 1980's, exploiting the computational efficiency of the simplex method to solve linear programs. We refer to Barahona et al. (1989) for polyhedral studies as well as computational experience.

The inherent 'quadratic flavour' of all these combinatorial optimization problems may be an intuitive explanation for the (mostly) moderate computational success of purely polyhedral approaches.

Let us therefore focus on models based on matrix relaxations, which provide a natural way to linearize quadratic functions. All these problems have a quadratic objective but otherwise the feasible region, consisting of binary variables, is linearly constrained. It is therefore a natural idea to use the matrix lifting idea from (1) to linearize the objective function as follows. Since

$$\langle S, LS\rangle = \langle L, SS^T\rangle,$$

we rewrite Max-k-Cut as

$$\max\left\{\frac{1}{2}\langle L, Y\rangle : Y \in \{SS^T : S \in P_{n,k}\}\right\}$$
$$= \max\left\{\frac{1}{2}\langle L, Y\rangle : Y \in \text{conv}\{SS^T : S \in P_{n,k}\}\right\}.$$

The first maximization is of combinatorial type, just like (11), while the second one is a linear program. The catch is of course that the polytope

$$P_{k\text{-}cut} = \text{conv}\{SS^T : S \in P_{n,k}\}$$

is given through its vertices rather than through a system of linear inequalities (facets). In view of the intractability of Max-k-Cut it should also be expected that such a description through facets is intractable as well. Following the general strategy described in the introduction, we need to find a partial description of $P_{k\text{-}cut}$ as the feasible region of a (linear) semidefinite program.

For the case of Max-k-Cut we have the following basic SDP description.

Proposition 2 *Let $S \in P_{n,k}$ and set $Y = SS^T$. Then*

$$Y \geq 0,\ diag(Y) = e, \quad kY - J \succeq 0.$$

The last condition shows in particular that Y is not only semidefinite, but satisfies the stronger condition $Y \succeq \frac{1}{k} J \succeq 0$. It follows from

$$\sum_{i=1}^{k} \begin{pmatrix} 1 \\ s_i \end{pmatrix} \begin{pmatrix} 1 \\ s_i \end{pmatrix}^T = \begin{pmatrix} k & e^T \\ e & Y \end{pmatrix} \succeq 0.$$

The variable transformation

$$X = \frac{1}{k-1}(kY - J),$$

together with $LJ = 0$, leads to the following basic semidefinite relaxation for Max-k-Cut, as investigated by Frieze and Jerrum (1997) and also by de Klerk et al. (2004)

$$z^{sdp}_{k\text{-}cut} := \max \left\{ \frac{k-1}{2k} \langle L, X \rangle : \ \text{diag}(X) = e,\ X \succeq 0,\ \forall i < j\ x_{ij} \geq -\frac{1}{k-1} \right\}. \tag{16}$$

In case $k = 2$, the lower bounds $x_{ij} \geq -1$ are implied by $X \succeq 0$ and $\text{diag}(X) = e$. Thus we get the well-known basic SDP relaxation for Max-Cut

$$z^{sdp}_{mc} := \max \left\{ \frac{1}{4} \langle L, X \rangle : \ \text{diag}(X) = e,\ X \succeq 0 \right\}. \tag{17}$$

This relaxation was introduced in dual form by Delorme and Poljak (1993) and received the attention of the scientific community due the seminal work of Goemans and Williamson (1995), who provided an a-priori approximation estimate of this relaxation, which has become the starting point for various other approximation results in combinatorial optimization.

In case of cardinality constrained k-partition, we need to model the additional constraints $s_i^T e = m_i$. Let us first consider the case of k-Equicut, where $m_1 = \ldots = m_k = \frac{n}{k}$, hence $S^T e = \frac{n}{k} e$. This implies that

$$Ye = SS^T e = \frac{n}{k} Se = \frac{n}{k} e.$$

Thus, in view of the previous proposition, we get the following basic SDP relaxation for k-Equicut

$$z_{k\text{-}gp}^{sdp} := \min \left\{ \frac{1}{2}\langle L, Y\rangle : \ \operatorname{diag}(Y) = e, \ Ye = \frac{n}{k}e, \ Y \succeq 0, \ Y \geq 0 \right\}. \tag{18}$$

It can easily be verified that matrices Y feasible for (18) also satisfy

$$\frac{1}{k}J \preceq Y \preceq \frac{n}{k}I,$$

see Rendl (2009) for further details. This relaxation was also derived from a 'dual' point of view by considering partition matrices $S \in P_{n,k}$ and exploiting the fact that the columns of S are pairwise orthogonal, and

$$S^T S = \operatorname{Diag}(m_1, \ldots, m_k).$$

We do not exploit this connection any further but refer to Karisch and Rendl (1998) and also to Rendl (2009).

The k-densest subgraph problem gives rise to the following semidefinite relaxations. We first recall the relaxation (4) of QP, described in the introduction. The matrix constraint can alternatively be expressed in the following form using the Schur-complement lemma

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \Leftrightarrow X - xx^T \succeq 0. \tag{19}$$

The condition $X - xx^T \succeq 0$ can be interpreted as a 'convexification' of $X = xx^T$ which is what we would like to enforce, but which unfortunately is a nonconvex quadratic equation, hence of no direct use. Since $x \in \{0, 1\}^n$, we also require $\operatorname{diag}(X) = x$, see (4). The simplest SDP relaxation for $z_{k\text{-}ds}$ would therefore consist of (4) with $e^T x = k$ added. A stronger relaxation can be obtained if we replace the equation $e^T x = k$ by its squared form

$$\langle J, X\rangle = k^2.$$

There are various standard techniques to translate the term $e^T x = k$ into something linear in X and x. The Lovász-Schrijver idea (Lovász and Schrijver 1991) suggests to multiply the equation individually by each x_i. This leads to the following set of equations

$$Xe = kx,$$

which have been investigated in Roupin (2004), Malick and Roupin (2011). It is known, and also pointed out in Roupin (2004), that there are some dependencies among these relaxations. Trivially, the $n + 1$ equations $Xe = kx, \ e^T x = k$ imply that

$\langle J, X\rangle = e^T Xe = k(e^T x) = k^2$. In the following lemma we summarize some further relations among these constraints.

Lemma 3 *Let X and x satisfy $X - xx^T \succeq 0$, $diag(X) = x$. Then the following statements hold.*

(a) $\langle J, X\rangle = k^2$ *implies* $e^T x \le k$,
(b) $e^T x = k$ *implies* $\langle J, X\rangle \ge k^2$,
(c) $e^T x = k$, $\langle J, X\rangle = k^2$ *implies* $Xe = kx$.
(d) $e^T x = k$, $\langle J, X\rangle = k^2$ *implies that* $\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$ *is singular.*

Proof Since $X - xx^T \succeq 0$ we have $e^T Xe - (e^T x)^2 \ge 0$. The definition of J yields $e^T Xe = \langle J, X\rangle$. Thus $\langle J, X\rangle = k^2$ implies $e^T x \le k$. Similarly $e^T x = k$ yields (b). To see (c) we use the following fact from matrix theory. If $Y \succeq 0$ and $a^T Ya = 0$, then $Ya = 0$. To see this we use the Cholesky decomposition $Y = CC^T$ to conclude $a^T Ya = a^T CC^T a = \|C^T a\|^2 = 0$. Thus $C^T a = 0$ and therefore $CC^T a = Ya = 0$. The assumptions in (c) imply $e^T(X - xx^T)e = 0$. We also have $X - xx^T \succeq 0$ so we get $(X - xx^T)e = 0$, which translates into $Xe - kx = 0$. Finally in case (d) we see that

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}\begin{pmatrix} -k \\ e \end{pmatrix} = \begin{pmatrix} -k + e^T x \\ -kx + Xe \end{pmatrix} = 0,$$

so that the matrix is indeed singular. □

Let us define the following basic SDP relaxation for $z_{k\text{-}ds}$

$$z^{sdp}_{k\text{-}ds} := \max\left\{\frac{1}{2}\langle A, X\rangle : \; X - xx^T \succeq 0, \; \text{diag}(X) = x, \; e^T x = k, \; \langle J, X\rangle = k^2\right\}. \tag{20}$$

In view of the lemma, it has the same optimal value as

$$\max\left\{\frac{1}{2}\langle A, X\rangle : \; X - xx^T \succeq 0, \; \text{diag}(X) = x, \; e^T x = k, \; Xe = kx\right\}.$$

This has also been pointed out by Roupin (2004). Note that this relaxation has $2n + 1$ equations, while the first one has only $n+2$ equations. Both however have the drawback that the feasible region does not contain strictly feasible points, see (d) from the lemma. Therefore feasible solutions necessarily are singular. This is bad news for standard interior-point methods to solve the relaxation. There are several ways to fix this problem but we will see below that there is a work-around which avoids this problem altogether and uses an unconstrained formulation of the problem. Before that we consider some generic ways to tighten the SDP relaxations.

5.2 Tightening the relaxations

We have now seen SDP relaxations of various partition problems. A natural question to investigate would be to look for tightenings of these relaxations. One of the most elementary classes of additional constraints is given by the *metric polytope*.

To motivate these constraints, let i, j, k be a triple of distinct numbers and consider the vector $f \in \{0, 1\}^n$ given by $f_i = f_j = f_k = 1$ and $f_l = 0$ otherwise. In this case the scalar product of f with *any* vector $y \in \{-1, 1\}^n$ can only have values $\langle f, y\rangle \in \{-3, -1, 1, 3\}$. Thus $|\langle f, y\rangle| \geq 1$ and therefore

$$\langle ff^T, yy^T\rangle \geq 1$$

holds for all $y \in \{-1, 1\}^n$. This argument is also valid if some of the entries f_i, f_j, f_k are changed to -1. In the semidefinite embedding we may therefore tighten the constraints $Y \succeq 0,\ \operatorname{diag}(Y) = e$ by asking that

$$\forall f \in \{-1, 0, 1\}^n \ \langle ff^T, Y\rangle \geq 1 \tag{21}$$

with exactly 3 nonzeros in f. Let us define the metric polytope

$$\text{MET} := \{Y : Y = Y^T,\ \forall i, j, k\ y_{ij} + y_{ik} + y_{jk} \geq -1,\ y_{ij} - y_{ik} - y_{jk} \geq -1\}.$$

The inequalities (21) therefore translate into

$$Y \in \text{MET}.$$

It should be noted that $\langle Y, ff^T\rangle = f^T Y f \geq 0$ holds for any semidefinite matrix Y. The tightening $\langle Y, ff^T\rangle \geq 1$ therefore may seem marginal, but in practice the inclusion of these constraints often leads to significant improvements of the relaxations, see Fischer et al. (2006), Rendl et al. (2010). Turning the argument around, one could state that $Y \succeq 0$ yielding $\langle ff^T, Y\rangle \geq 0$ can be viewed as enforcing $\langle ff^T, Y\rangle \geq 1$ at least partially.

The equivalence between the formulation of Max-Cut in 0, 1 and -1, 1 variables suggests that there should also be a system of inequalities equivalent to MET, but for the matrix lifting

$$X - xx^T \succeq 0, \operatorname{diag}(X) = x$$

in the 0,1 setting. This is indeed the case and it is given by the *boolean quadric polytope* BQP which is defined through the following set of inequalities

$$0 \leq x_{ij} \leq x_{ii},\ x_{ii} + x_{jj} - x_{ij} \leq 1,\ x_{ik} + x_{jk} \leq x_{kk} + x_{ij},$$
$$x_{ii} + x_{jj} + x_{kk} - x_{ij} - x_{ik} - x_{jk} \leq 1.$$

It was introduced by Padberg (1989) and also by De Simone (1990). Helmberg (2000) provides a comprehensive overview of these equivalences together with explicit transformations between the two models.

The concept of vectors f with support 3, i.e. with exactly 3 nonzeros has an immediate generalization to arbitrary odd support, leading to $2k+1$-clique inequalities

$$\langle ff^T, X\rangle \geq 1.$$

These are defined for all $f \in \{-1, 0, 1\}^n$ with $\|f\|^2 = 2k+1$. For k fixed these can be separated in polynomial time by enumeration. Some computational experience with these classes of cutting planes in combination with SDP relaxations for the Max-Cut problem are reported in Helmberg and Rendl (1998).

The inequalities in MET also have a combinatorial interpretation. Let us consider $y_{ij} + y_{ik} + y_{jk} \geq -1$. This condition is violated by variables taking values in $\{-1, 1\}$ if and only if all three values are -1. We recall that $y_{ij} = -1$ implies that i and j are in distinct sets. The inequality therefore translates into asking that i, j and k are not allowed to lie in three distinct sets. This condition can therefore be included in all bisection problems. In a similar way we see that $y_{ij} - y_{ik} - y_{jk} \geq -1$ is violated precisely if $y_{ij} = -1$ and $y_{ik} = y_{jk} = 1$. This inequality therefore rules out that i and j are separated while k is in the same set with both i and j. It therefore applies to all partition problems.

A tightened version of (17) is therefore given by

$$z_{mc}^{sdp\text{-}met} := \max\left\{\frac{1}{4}\langle L, X\rangle : \ \mathrm{diag}(X) = e, \ \ X \in \mathrm{MET}, \ \ X \succeq 0\right\}. \tag{22}$$

This SDP is computationally challenging, because MET contains roughly $\frac{2}{3}n^3$ inequalities. In Fischer et al. (2006), an iterative approach is described to compute this relaxation also for larger graphs ($n \approx 1,000$). The software package BiqMac, see Rendl et al. (2010), solves Max-Cut and unconstrained binary quadratic programming problems based on the relaxation (22) in combination with Branch and Bound.

The inequalities in BQP, reflecting the 0,1 model allow similar interpretations. The condition $x_{ij} + x_{kk} \geq x_{ik} + x_{jk}$ in case of $X \in P_{k\text{-}cut}$ describes the following situation. First, $\mathrm{diag}(X) = e$ yields $x_{kk} = 1$. The entries x_{ij} express whether i and j are in the same partition block ($x_{ij} = 1$) or not ($x_{ij} = 0$). Thus the inequality

$$x_{ij} + 1 \geq x_{ik} + x_{jk} \tag{23}$$

is violated only if $x_{ik} = x_{jk} = 1$ and $x_{ij} = 0$. Thus as before, whenever i, k and j, k are in the same set, then also i, j are together. These inequalities are therefore valid for $P_{k\text{-}cut}$. Ghaddar et al. (2011) use these inequalities applied to the Max-k-Cut relaxation (16) and provide computational results to solve the problem to optimality.

The most general class of inequalities of the type (21) is given by the set of *hypermetric inequalities*, which are defined as follows. Let b be an integer vector such that $\min\{|b^T y| : y \in \{-1, 1\}^n\} = 1$. Then it holds that $\langle bb^T, Y\rangle \geq 1$ for all

$Y \in conv\{yy^T : y \in \{-1,1\}^n\}$. If we denote by $\mathcal{B}$ the set of all vectors b satisfying this condition, then

$$\forall b \in \mathcal{B} \ \langle bb^T, Y\rangle \geq 1$$

is valid for the Max-Cut problem. It is currently not known how to separate violated inequalities of this type. The set $\mathcal{B}$ is after all the intersection of an infinite number of halfspaces. It is a nontrivial observation that this set is actually polyhedral, see Deza et al. (1993). Clearly, the inequalities from MET are special cases of hypermetric inequalities, just like the $2k+1$-clique inequalities.

5.3 Bisection and densest subgraph as Max-Cut

We are now going to summarize some rather obvious transformations which show that both cardinality constrained bisection and also k-densest subgraph can be reformulated as Max-Cut problems (in dense graphs), and hence could be solved or approximated using software for Max-Cut. We first recall that the two bivalent formulations (13) and (12) are equivalent, and also the semidefinite relaxations derived from the two models, see for instance Helmberg (2000).

The bisection problem with cardinality constraint $e^T s_1 = m_1$ from (14) amounts to

$$z_{2\text{-}gp} := \min\left\{\frac{1}{4}y^T Ly : y \in \{-1,1\}^n,\ e^T y = 2m_1 - n := k\right\}.$$

We add a penalty term $\gamma(e^T y - k)^2$ with penalty parameter $\gamma > 0$ and sufficiently large to the objective function. This will insure that the cardinality constraint will hold at the optimum. The resulting unconstrained problem looks like Max-Cut, except that the objective function contains a linear term now. This is fixed by the usual homogenization procedure, i.e. introduce $y_0 \in \{-1,1\}$ and replace the penalty term by $\gamma(e^T y - ky_0)^2$, leading to

$$z_{2\text{-}gp} = \min\left\{\frac{1}{4}y^T Ly + \gamma(e^T y - ky_0)^2 :\ y \in \{-1,1\}^n, y_0 \in \{-1,1\}\right\}.$$

Since the objective function is symmetric (y and $-y$ have the same value), we may without loss of generality assume that $y_0 = 1$.

In case of the k-densest subgraph problem, we proceed in a similar way and consider

$$\max\{x^T Ax - \gamma(e^T x - k)^2 :\ x \in \{0,1\}^n\}.$$

The penalty term now has the form $(e^T x - k)^2 = \left\langle \begin{pmatrix} -k \\ e \end{pmatrix}, \begin{pmatrix} 1 \\ x \end{pmatrix} \right\rangle^2$. The resulting basic semidefinite relaxation now reads

$$\max\{\langle A, X\rangle - \gamma(\langle J, X\rangle - 2ke^T x + k^2) :\ X - xx^T \succeq 0,\ \operatorname{diag}(X) = x\}. \tag{24}$$

Table 1 Approximating k-dense subgraph using Max-Cut with penalty parameter γ

γ	Value of (24)
0	5,933.000
1	894.641
10	389.419
100	367.763
1,000	365.731
10,000	365.530
∞	365.508

The graph has $n = 50$ and corresponds to instance $A50$ described below. The total edge weight is 5,933

The optimization will make the nonnegative term involving γ,

$$\left\langle \begin{pmatrix} k^2 & -ke^T \\ -ke & J \end{pmatrix}, \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \right\rangle$$

go to zero as γ is increased. This together with the matrix variable being semidefinite will force

$$\begin{pmatrix} -k + e^T x \\ -kx + Xe \end{pmatrix} \to 0,$$

see also the argument in the proof of the preceeding lemma. This shows that with increasing value of γ, the optimal value of the basic semidefinite relaxation (24) of k-densest subgraph will approach $z^{sdp}_{k\text{-}ds}$ from above. Table 1 provides an illustration of this convergence behaviour on a small instance of size $n = 50$.

It is likely that the special structure of the cost function for these problems may result in difficult Max-Cut instances. We do not suggest to transform the constrained problems to Max-Cut, but any approximation methods specially tailored for these problems should do at least as good as simple Max-Cut based relaxations of the transformed problem.

5.4 Semidefinite relaxations for ordering

Matrix liftings for the quadratic assignment problem are investigated in Zhao et al. (1998, 2010), Rendl and Sotirov (2007). The QAPLIB[2] website provides a collection of QAP instances together with their best known solution and also a comparison of bounds.

In the following we therefore concentrate on matrix liftings of the quadratic ordering problem, as this is a combinatorial optimization problem which is rather hard to approximate by purely linear techniques. It has also found much less scientific attention. There is no need here to distinguish between LOP and QOP because the

[2] http://www.seas.upenn.edu/qaplib/

feasible region for both problems is the same, and after the matrix lifting, the objective functions of both problems will be linear.

Thus we need to investigate outer approximations of

$$\mathcal{M}_{OP} := \text{conv}\left\{ \begin{pmatrix} 1 \\ y \end{pmatrix} \begin{pmatrix} 1 \\ y \end{pmatrix}^T : y \in F_{OP} \right\}.$$

The matrix order now becomes $\binom{n}{2} + 1$. We recall that

$$F_{OP} = \{y \in \mathbb{R}^{\binom{n}{2}} : y_{ij} \in \{-1, 1\}, \ \forall 1 \le i < j < k \le n \ |y_{ij} + y_{jk} - y_{ik}| = 1\}.$$

The rows and columns of matrices $Y \in M_{OP}$ are therefore indexed by 1, corresponding to the first row, and ij with $i < j$ for the remaining rows. Squaring the equation $|y_{ij} + y_{jk} - y_{ik}| = 1$ leads to the following equation, which obviously holds for any $Y \in M_{OP}$,

$$y_{ij,ij}^2 + y_{ik,ik}^2 + y_{jk,jk}^2 + 2(y_{ij,jk} - y_{ij,ik} - y_{ik,jk}) = 1.$$

Since $\text{diag}(Y) = e$, this simplifies to the linear equation valid for elements $Y \in \mathcal{M}_{\mathcal{OP}}$

$$\forall i < j < k \ y_{ij,jk} - y_{ij,ik} - y_{ik,jk} = -1. \tag{25}$$

Thus QOP and LOP have the following SDP relaxations, with appropriately defined cost matrix D and matrix variable Y of order $\binom{n}{2} + 1$

$$z_{QOP}^{sdp} = \max\{\langle D, Y\rangle : \ Y \succeq 0, \ \text{diag}(Y) = e, \ Y \text{ satisfies (25)}\}.$$

Tightenings as previously described are also possible, for instance $Y \in$ MET is valid for $\mathcal{M}_{OP}$. The recent dissertation Hungerländer (2012) and also Hungerländer and Rendl (2012) contain further details. It should be clear that the resulting SDP are a computational challenge, as both the matrix order ($O(n^2)$) and the number of equality constraints in (25) ($O(n^3)$) and the number of potentially violated inequalities in MET ($O(n^6)$) are nontrivial. To be explicit, ordering problems with $n = 100$ items result in SDP with matrices of order $\binom{n}{2} + 1 = 4,951$, and are beyond the limits for practical computations.

The SDP relaxation for QOP has been used in special cases before by Anjos et al. (2005) for the single row facility location problem, and by Buchheim et al. (2009) for two-level crossing minimization.

6 Theoretical approximation results

We have seen a variety of semidefinite relaxations and will now take a short look at their quality. Randomized rounding, introduced by Goemans and Williamson (1995) has lead to the celebrated approximation ratio for Max-Cut, stating that

$$\frac{z_{mc}}{z_{mc}^{sdp}} \geq 0.87856 \tag{26}$$

holds for any graph with nonnegative edge weights, i.e. $A \geq 0$ elementwise. In the case that the Laplacian $L \succeq 0$, even though not necessarily $A \geq 0$, Nesterov (1997) generalizes the error analysis and shows

$$\frac{z_{mc}}{z_{mc}^{sdp}} \geq \frac{2}{\pi} > 0.636.$$

On the other hand, Hastad (1997) shows that unless P = NP, it is NP-hard to find a polynomial time approximation to Max-Cut with performance guarantee 0.94117 or better. In fact, in case that the *Unique Games Conjecture* holds, it would follow that the performance ratio (26) is best possible, see for instance the recent survey paper by Khot (2010).

Frieze and Jerrum (1997) extend the analysis to Max-k-Cut. It is an easy exercise to verify that a random partition of the vertices of a graph with nonnegative weights into k sets has expected cut value $(1 - \frac{1}{k}) \sum_{i<j} a_{ij}$, hence for increasing values of k the approximation ratio $\frac{k-1}{k}$ tends to one, and the problem is not interesting. For small values of k, the following approximation ratios are shown in Frieze and Jerrum (1997) for $A \geq 0$,

$$\frac{z_{3\text{-}cut}}{z_{3\text{-}cut}^{sdp}} \geq 0.8327, \quad \frac{z_{4\text{-}cut}}{z_{4\text{-}cut}^{sdp}} \geq 0.8503, \quad \frac{z_{5\text{-}cut}}{z_{5\text{-}cut}^{sdp}} \geq 0.8742.$$

Later, Goemans and Williamson (2004) and de Klerk et al. (2004) provide the following improvement for $k = 3$,

$$\frac{z_{3\text{-}cut}}{z_{3\text{-}cut}^{sdp}} \geq \frac{7}{12} + \frac{3}{4\pi^2} \arccos^2(-1/4) > 0.836.$$

Finally, Halperin and Zwick (2002) and later Jäger and Srivastav (2005) provide approximations for the k-dense subgraph problem with nonnegative weights and show for instance that $\frac{n}{2}$-dense Subgraph can be approximated with an approximation error of 0.6221 (Halperin and Zwick 2002) and 0.6223 (Jäger and Srivastav 2005).

These approximation results are rather difficult to obtain and involve analytical tools as well as techniques from probability theory. Since they hold for all instances of the problem class under consideration, the actual error of a specific instance may actually be much smaller.

7 Numerical illustrations

To give a flavour of the actual behaviour of the various relaxations, we consider results for the following three input matrices $A50$, $A100$, $A200$. The Matlab commands below generate the matrices $A50$ and $A100$. The first one is a random matrix of order $n = 50$ with integer weights drawn from $\{0, 1, \dots, 10\}$. The second one has order $n = 100$. Half the entries are randomly set to zero, the remaining ones are chosen from $\{-10, -9, \dots, 9, 10\}$. The last instance $A200$ is an unweighted graph from the DIMACS[3] collection (brock200-3).

```
n=50; seed=2012; rand('seed',seed); A=round(10*rand(n)); A=triu(A,1); A50=A+A';
n=100; seed=2012; rand('seed',seed); G=(rand(n)<=.5);
A=round(20*rand(n)-10); A=triu(A,1); A=A.*G; A100=A+A';
```

In Table 2 we provide results of semidefinite relaxations for Max-Cut. The table includes the basic relaxation from (17) and also the relaxation (22), which includes the inequalities from the metric polytope. The last column in the table (labeled bks) provides the value of the *best known solution*, i.e. the value of the largest cut found. It is not known whether these cuts are optimal, but the gap between the relaxations and the cut values is rather small. The ratios $\frac{\text{bound}}{\text{bks}}$ are also included. We see that on these instances the inclusion $X \in \text{MET}$ reduces the gap to bks significantly. The interested reader is referred to Rendl et al. (2010) to get the full picture for the computational experience using (22) to solve Max-Cut to optimality.

In Table 3 we show the bounds on the same graphs for the k-dense subgraph problem with $k = \frac{n}{5}$. The column labeling is similar for all tables. We first give the bound (20) and then the slightly weaker penalty term version (24) with $\gamma = 1{,}000$. In the column labeled MET, we provide the bounds obtained by adding to (24) all the inequalities from MET. The last column features the value of the best known feasible solution. It is interesting to note that this simple reformulation of the problem yields rather strong bounds on all instances. It proves optimality in the smallest case, and leaves a small interval in the other cases. The difference between (20) and (24) is quite small for the instances without negative weights (A50, A200). It is much bigger for A100, which has both positive and negative weights.

Turning to k-GP, we consider partitioning the three instances into 5 sets of equal size. We consider the relaxation (18) which is already nontrivial to compute, as it contains $O(n^2)$ sign constraints $y_{ij} \ge 0$. Thus we also look at the simpler relaxation without these constraints

$$z_{k\text{-}gp}^{sdp\text{-}basic} := \min \left\{ \frac{1}{2}\langle L, Y\rangle : \ \operatorname{diag}(Y) = e, \ Ye = \frac{n}{k}e, \ Y \succeq 0 \right\}. \tag{27}$$

The Table 4 clearly shows that this relaxation is rather weak, and the inclusion of the sign constraints $y_{ij} \ge 0$ in (18) yields a much stronger bound. We recall that this is a minimization problem, so the bounds are smaller than the best known value.

Finally, we also determine approximate solutions to Max-k-Cut, with $k = 5$. As before, the relaxation (16) contains $O(n^2)$ sign constraints $x_{ij} \ge \frac{-1}{k-1}$, so we also consider the model without these constraints

[3] http://dimacs.rutgers.edu/Challenges/

Table 2 Semidefinite relaxations for Max-Cut

Instance	(17)	Ratio	(22)	Ratio	bks
A50	3,466.537	0.986	3,419.398	0.999	3,418
A100	1,654.819	0.860	1,466.311	0.970	1,423
A200	4,560.616	0.977	4,514.753	0.987	4,457

The value of the best known solution (bks) is given in the column labeled bks

Table 3 Semidefinite relaxations for k-dense subgraph

Instance	(20)	(24)	MET	bks
A50	365.508	365.831	333.65	333
A100	607.691	611.070	490.50	489
A200	683.561	683.672	647.97	627

The value k is set to $\frac{n}{5}$. The best known solution (bks) is given in the column labeled bks. The column labeled MET contains the bound after reformulating the problem as in (24) (with $\gamma = 1{,}000$) and with all the triangle inequalities included

Table 4 Semidefinite relaxations for 5-GP

Instance	(27)	(18)	bks
A50	4,231.0	4,401.3	4,475
A100	−2,884.3	−2,322.4	−1,892
A200	8,703.3	8,861.9	9,089

The best known solution (bks) is given in the column labeled bks

Table 5 Semidefinite relaxations for max 5-cut

Instance	(28)	(16)	bks
A50	5,546.4	5,377.6	5,302
A100	2,674.7	2,164.6	1,745
A200	10,667.2	10,532.8	10,320

The best known solution (bks) is given in the column labeled bks

$$z_{k\text{-}cut}^{sdp\text{-}basic} := \max\left\{\frac{k-1}{2k}\langle L, X\rangle : \operatorname{diag}(X) = e,\ X \succeq 0\right\}. \tag{28}$$

The results are not much different from k-GP. The inclusion of the sign constraints yields a significant improvement, especially for the instance with both positive and negative edge weigths (A100) (Table 5).

These results should at least give the qualitative picture for the use of semidefinite relaxations. Even though some of the worst-case error bounds (see the previous section) are rather weak, we see that in practise these bounds provide good approximations to the actual optimal solution. Moreover, the primal optimizers can be used for rounding to generate feasible solutions. It is currently still considered a challenge to solve these relaxations on larger instances, say $n \approx 500$, or to include additional cutting planes, like the triangle inequalities in the relaxation.

Acknowledgments I thank Miguel Anjos and an anonymous referee for giving numerous suggestions to improve the presentation.

References

Anjos MF (2005) Semidefinite optimization approaches to satisfiability and maximum-satisfiability problems. J Satisf Boolean Model Comput 1:1–47

Anjos MF, Lasserre JB (eds) (2012) Handbook on semidefinite, conic and polynomial optimization. International series in Operations Research and Management Science. Springer, Berlin

Anjos MF, Kennings A, Vanelli A (2005) A semidefinite optimization approach for the single-row layout problem with unequal dimensions. Discret Optim 2:113–122

Applegate DL, Bixby RE, Chvátal V, Cook WJ (2006) The traveling salesman problem. A computational study. Princeton University Press, Princeton

Barahona F, Grötschel M, Mahjoub A (1985) Facets of the bipartite subgraph polytope. Math Oper Res 10: 340–358

Barahona F, Jünger M, Reinelt G (1989) Experiments in quadratic 0–1 programming. Math Program 44: 127–137

Barahona F, Mahjoub A (1986) On the cut polytope. Math Program 36:157–173

Blekherman G, Parillo PA, Thomas R (2012) Semidefinite optimization and convex algebraic geometry. (forthcoming)

Buchheim C, Wiegele A, Zheng L (2009) Exact algorithms for the quadratic linear ordering problem. INFORMS JComput 22:168–177

Burkard RE, Dell'Amico M, Martello S (2009) Assignment problems. SIAM, Philadelphia, PA

Chimani M, Hungerländer P, Jünger M, Mutzel P (2011) An SDP approach to multi-level crossing minimization. In: Proceedings of algorithm engineering & experiments [ALENEX'2011]

de Klerk E (2002) Aspects of semidefinite programming: interior point algorithms and selected applications. Kluwer Academic Publishers, Dordrecht

de Klerk E, Pasechnik DV, Warners JP (2004) On approximate graph colouring and Max-k-Cut algorithms based on the ϑ-function. J Comb Optim 8(3):267–294

De Simone C (1990) The cut polytope and the Boolean quadric polytope. Discrete Math 79(1):71–75

Delorme C, Poljak S (1993) Laplacian eigenvalues and the maximum cut problem. Math Program 62: 557–574

Deza M, Grishukhin VP, Laurent M (1993) The hypermetric cone is polyhedral. Combinatorica 13:397–411

Duffin RJ (1956) Infinite programs. Ann Math Stud 38:157–170

Fischer I, Gruber G, Rendl F, Sotirov R (2006) Computational experience with a bundle method for semidefinite cutting plane relaxations of max-cut and equipartition. Math Program 105:451–469

Frieze A, Jerrum M (1997) Improved approximation algorithms for Max k-Cut and Max Bisection. Algorithmica 18(1):67–81

Ghaddar B, Anjos MF, Liers F (2011) A branch-and-cut algorithm based on semidefinite programming for the minimum k-partition problem. Ann Oper Res 188:155–174

Goemans MX, Williamson DP (1995) Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J ACM 42:1115–1145

Goemans MX, Williamson DP (2004) Approximation algorithms for max-3-cut and other problems via complex semidefinite programming. J Comput Syst Sci 68(2):442–470

Halperin E, Zwick U (2002) A unified framework for obtaining improved approximation algorithms for maximum graph bisection problems. Random Struct Algorithms 20:382–402

Hastad J (1997) Some optimal inapproximability results. In: Proceedings 29th symposium on the theory of computing, pp 1–10

Helmberg C (2000) Fixing variables in semidefinite relaxations. SIAM J Matrix Anal Appl 21(3): 952–969

Helmberg C (2002) Semidefinite programming. Eur J Oper Res 137:461–482

Helmberg C, Rendl F (1998) Solving quadratic (0,1)-problems by semidefinite programming and cutting planes. Math Program 82:291–315

Helton JW, Nie J (2010) Semidefinite representation of convex sets. Math Program (A) 122:21–64

Hungerländer P (2012) Semidefinite approaches to ordering problems. PhD thesis, Alpen-Adria University Klagenfurt, Austria

Hungerländer P, Rendl F (2012) Semidefinite relaxations of ordering problems. Math Program (B), (forthcoming)

Jäger G, Srivastav A (2005) Improved approximation algorithms for maximum graph partition problems. J Comb Optim 10:133–167

Jünger M, Lee EK, Mutzel P, Odenthal T (1997) A polyhedral approach to the multi-layer crossing minimization problem. In: GD '97: Proceedings of the 5th international symposium on graph drawing, pp 13–24. Springer, Berlin
Jünger M, Mutzel P (1997) 2-layer straightline crossing minimization: performance of exact and heuristic algorithms. J Graph Algorithms Appl 1:1–25
Karisch SE, Rendl F (1998) Semidefinite programming and graph equipartition. Fields Inst Commun 18: 77–95
Karp RM (1972) Reducibility among combinatorial problems. In: Miller RE, Thather JW (eds) Complexity of computer computation. Plenum Press, New York, pp 85–103
Keil JM, Brecht TB (1991) The complexity of clustering in planar graphs. J Comb Math Comb Comput 9:155–159
Kernighan BW, Lin S (1970) An efficient heuristic procedure for partitioning graphs. Bell Syst Tech J 49:291–307
Khot S (2010) On the unique games conjecture. In: Proceedings of the 25th IEEE conference on computational complexity, pp 99–121
Lasserre JB (2010) Moments, positive polynomials and their applications. Imperial College Press, London
Laurent M, Rendl F (2005) Semidefinite programming and integer programming. In: Aardal K, Nemhauser GL, Weismantel R (eds) Discrete optimization. Elsevier, Amsterdam, pp 393–514
Lee J, Leyffer S (eds) (2012) Mixed integer nonlinear programming. IMA volumes in applied mathematics and its applications. Springer, Berlin
Lovász L (1979) On the Shannon capacity of a graph. IEEE Trans Inf Theory 25:1–7
Lovász L (2003) Semidefinite programs and combinatorial optimization. In: Reed BA, Sales CL (eds) Recent advances in algorithms and combinatorics. CMS books in mathematics. Springer, Berlin, pp 137–194
Lovász L, Schrijver A (1991) Cones of matrices and set-functions and 0–1 optimization. SIAM J Optim 1:166–190
Malick J, Roupin F (2011) Solving k-cluster problems to optimality with semidefinite programming. Technical report, CNRS, Grenoble, France
Marti R, Reinelt G (2011) The linear ordering problem: exact and heuristic methods in combinatorial optimization. Applied Mathematical Sciences. Springer, Berlin
Nesterov Y (1997) Quality of semidefinite relaxation for nonconvex quadratic optimization. Technical report, CORE
Nesterov Y, Nemirovski AS (1994) Interior point polynomial algorithms in convex programming. SIAM Publications, SIAM, Philadelphia
Padberg M (1989) The quadric Boolean polytope: some characteristics, facets and relatives. Math Program 45:139–172
Pardalos PM, Rodgers GP (1990) Computational aspects of a branch and bound algorithm for quadratic zero-one programming. Computing 45:131–144
Rendl F (2009) Semidefinite relaxations for integer programming. In: Jünger M, Liebling ThM, Naddef D, Nemhauser GL, Pulleyblank WR, Reinelt G, Rinaldi G, Wolsey LA (eds) 50 years of integer programming 1958–2008. Springer, Berlin, pp 687–726
Rendl F, Rinaldi G, Wiegele A (2010) Solving max-cut to optimality by intersecting semidefinite and polyhedral relaxations. Math Program 212:307–335
Rendl F, Sotirov R (2007) Bounds for the quadratic assignment problem using the bundle method. Math Program (B) 109:505–524
Roupin F (2004) From linear to semidefinite programming: an algorithm to obtain semidefinite relaxations for bivalent quadratic problems. J Comb Optim 8:469–493
Schrijver A (2003) Combinatorial optimization. Polyhedra and efficiency A, volume 24 of algorithms and combinatorics. Springer, Berlin
Sherali HD, Adams WP (1990) A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. SIAM J Discrete Math 3(3):411–430
Sherali HD, Adams WP (1994) A hierarchy of relaxations and convex hull characterizations for mixed-integer zero-one programming problems. Discrete Appl Math 52(1):83–106
Shor NZ (1987) An approach to obtaining global extremums in polynomial mathematical programming problems. Kibernetika 5:102–106
Tunçel L (2010) Polyhedral and semidefinite programming methods in combinatorial optimization. Fields Institute Monographs. American Mathematical Society, Providence, RI

Vallentin F (2008) Lecture notes: semidefinite programs and harmonic analysis. Technical report. Tilburg University, The Netherlands

Vandenberghe L, Boyd S (1996) Semidefinite programming. SIAM Rev 38:49–95

Wolkowicz H, Saigal R, Vandenberghe L (eds) (2000) Handbook of semidefinite programming. Kluwer, Dordrecht

Zhao Q, Karisch SE, Rendl F, Wolkowicz H (1998) Semidefinite programming relaxations for the quadratic assignment problem. J Comb Optim 2:71–109

Zhao X, Sun D, Toh K (2010) A Newton CG augmented Lagrangian method for semidefinite programming. SIAM J Optim 20:1737–1765