An exact duality theory for semidefinite programming and its complexity implications

Motakuri V. Ramana

Center for Applied Optimization, 303 Weil Hall, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32608, USA

Received 7 January 1995

Abstract

In this paper, an exact dual is derived for Semidefinite Programming (SDP), for which strong duality properties hold without any regularity assumptions. Its main features are: (i) The new dual is an explicit semidefinite program with polynomially many variables and polynomial size coefficient bitlengths. (ii) If the primal is feasible, then it is bounded if and only if the dual is feasible. (iii) When the primal is feasible and bounded, then its optimum value equals that of the dual, or in other words, there is no duality gap. Further, the dual attains this common optimum value. (iv) It yields a precise theorem of the alternative for semidefinite inequality systems, i.e. a characterization of the *infeasibility* of a semidefinite inequality in terms of the *feasibility* of another polynomial size semidefinite inequality.

The standard duality for linear programming satisfies all of the above features, but no such explicit gap-free dual program of polynomial size was previously known for SDP, without Slater-like conditions being assumed. The dual is then applied to derive certain complexity results for SDP. The decision problem of Semidefinite Feasibility (SDFP), which asks to determine if a given semidefinite inequality system is feasible, is the central problem of interest. he complexity of SDFP is unknown, but we show the following: (i) In the Turing machine model, the membership or nonmembership of SDFP in NP and Co-NP is simultaneous; hence SDFP is not NP-Complete unless NP = Co-NP. (ii) In the real number model of Blum, Shub and Smale, SDFP is in NP∩Co-NP. © 1997 The Mathematical Programming Society, Inc. Published by Elsevier Science B.V.

Keywords: Semidefinite programming; Strong duality; Complexity classes; Theorems of the alternative

1. Introduction

1.1. Problem of interest

We consider the following (Primal) Semidefinite Program (SDP):

$$\begin{aligned} \sup \quad & c^{\mathrm{T}} x \\ \text{s.t.} \quad & \sum_{i=1}^{m} x_i Q_i \preceq Q_0, \end{aligned} \tag{P}$$

where $Q_0, Q_1, \ldots, Q_m$ are given real symmetric matrices, and $\preceq$ denotes the Löewner partial order, i.e., $B \preceq A$ iff $A - B$ is positive semidefinite. (Throughout this paper, (P) will refer to this semidefinite program.)

Formally, one may define,

Definition 1 (*The Semidefinite Programming Problem*). Determine and/or compute the following for (P)

(i) Is the feasible region nonempty?
(ii) If so, is the objective function bounded?
(iii) If so, is the optimum attained?
(iv) If so, compute an (approximate) optimum solution.

Note that the first three parts are decision problems and the last is of numerical computation nature. Of the three decision problems, the first one is most central, and hence we state it separately:

Definition 2 (*Semidefinite Feasibility Problem* (*SDFP*)). Determine whether there exists an $x \in \mathbb{R}^m$ such that $Q(x) \succeq 0$, where $Q(x) = Q_0 - \sum_i x_i Q_i$, for given real symmetric matrices Q_i, $i = 0, \ldots, m$.

At the moment, the complexity of SDFP is open in both the Turing machine as well as the real number models. The existing algorithms for SDP can be characterized as those that find *approximate* optimal solutions to SDP, and for precisely this reason are incapable of solving the decision problems (i)–(iii) in polynomial time. It will follow from the results of this paper that there are polynomial reductions [9] to SDFP from boundedness and attainment problems, as well as other decision problems concerning SDP. Thus, SDFP assumes a central role in a rigorous complexity theoretic treatment of semidefinite programming.

The purpose of this paper is to derive a polynomial size dual program for the problem (P) (Section 1.5). This dual is a semidefinite program, whose coefficients are completely and explicitly determined from the primal data, and it enjoys zero duality gap as well as other strong duality properties. It is emphasized here that we do not make any kind of assumptions concerning the problem instance, and in particular, we do not suppose that any form of constraint qualifications, Slater or otherwise, hold.

As an application of the dual, we derive (Section 3) characterizations for several properties concerning semidefinite systems, including an assumption-free theorem of the alternative. These characterizations are then applied to obtain certain complexity relations for SDFP (Section 4.3), one of which is a proof that semidefinite feasibility is *not* NP-Complete unless NP = Co-NP.

1.2. On semidefinite programming

Semidefinite programming is a generalization of linear programming, obtained by replacing the nonnegativity constraints of LP by the semidefiniteness of some matrix variables and maps, or in other words, the role of the nonnegative orthant is now played by the cone of positive semidefinite matrices.

Historically, semidefinite programming has been studied in more general contexts such as convex and cone programming (see [2] for references).

However, the more recent surge of interest in SDP was perhaps inspired by the work of [11] (see [12, Chapter 9]). In that work, the authors associate with every graph G, a convex set denoted by $\mathrm{TH}(G)$, and show that when G is perfect, this set equals the stable set polytope. Then they demonstrate that one can optimize over $\mathrm{TH}(G)$ in polynomial time, and hence the stable set problem (along with many other related problems) can be solved in polynomial time for perfect graphs.

The algorithms of [11] employ the ellipsoid method, and are not considered to be efficient in practice. In [1] it was shown that many of the known interior point methods for LP readily extend to polynomial time algorithms for solving SDP *approximately* (see Section 4.1). In [18], Nesterov and Nemirovskii developed efficient interior point methods for a wider class of convex programs, by employing self-concordant barrier functions. Other early papers in the area include [15] and [19]. In [21], the relationship between SDP and multiquadratic programming (quadratic programming with quadratic constraints) was studied and certain geometric and algorithmic results were developed for SDP.

A natural generalization of the standard LP duality has been considered in [2] and [18]. This can also be seen as a specialization of the Lagrangian dual of (P) when this problem is viewed as a cone program. However, there are simple instances of SDPs (see Section 4.2) for which this Lagrangian dual exhibits a duality gap. As already mentioned, we remedy this situation in this paper, by deriving an explicit dual for SDP which has no duality gap.

A recent result of Goemans and Williamson [10] showing that one can use a semidefinite relaxation to obtain a .878-approximation algorithm for the Max-Cut problem, gave further impetus to this subject. Their result employs a clever randomized rounding scheme, and it has inspired other results on the application of SDP to combinatorial optimization problems.

An extensive bibliography on semidefinite programming can be found in the survey articles [2,24,29].

1.3. Notation

Most of the notation is taken from [14] and [27]. The main matrix spaces of interest here are:

- $\mathcal{M}_n$: The space of $n \times n$ real matrices.
- $\mathcal{S}_n$: The subspace of symmetric matrices in $\mathcal{M}_n$.

For $A, B \in \mathcal{S}_n$, we write $A \succeq B$ (resp. $\succ$), if $A - B$ is positive semidefinite (resp. positive definite), i.e., all the eigenvalues of $A - B$ are nonnegative (resp. positive). The term"positive semidefinite" will sometimes be abbreviated by PSD and "positive definite" by PD.

The inner product on $\mathcal{M}_n$ (and $\mathcal{S}_n$) is given by

$$A \bullet B = \sum_{i,j} A_{ij} B_{ij}.$$

The trace $A \bullet I$ of a matrix A is denoted by $\mathrm{Tr}(A)$. Given $A \in \mathcal{M}_m, B \in \mathcal{M}_k$, their *direct sum* is the block partitioned $(m+k) \times (m+k)$ matrix

$$A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

The *spectral decomposition* of a symmetric matrix A is $A = VDV^{\mathrm{T}}$, where V is an orthogonal matrix (i.e., $V^{\mathrm{T}}V = I$) whose columns are eigenvectors of A, and D is the diagonal matrix of the eigenvalues of A. The *spectral radius* of a not necessarily symmetric matrix A, denoted by $\rho(A)$, is the largest of the magnitudes of the eigenvalues of A. The null space of A is denoted by $\mathrm{Null}(A)$.

For $A, B \subseteq \mathbb{R}^n$, $A + B$ denotes the *Minkowski Sum* (also called the *set sum*). For a scalar $\alpha \in \mathbb{R}$, and a set A in $\mathbb{R}^n$, $\alpha A = \{\alpha x \mid x \in A\}$. For $A \subseteq \mathbb{R}^n$, $\mathrm{Conv}(A)$ (resp. $\mathrm{Aff}(A)$) denotes the smallest convex set (resp. affine subspace) containing A. The dimension of A is $\dim(\mathrm{Aff}(A))$, and $B(x, r)$ is the ball of radius r around x. We will use the convention that $\subset$ denotes strict containment and $\subseteq$ denotes nonstrict containment (similar usage for $\supset$ and $\supseteq$).

For a convex set $G \subseteq \mathbb{R}^n$, the interior and the relative interior of G are denoted by $\mathrm{Int}(G)$ and $\mathrm{ri}(G)$ respectively. The *recession cone* of G is defined by

$$0^+(G) = \{v \in \mathbb{R}^n \mid \forall x \in G,\ t \geqslant 0,\ x + tv \in G\},$$

and the *lineal hull* of G (contrast with linear hull) is the subspace $0^+(G) \cap 0^+(-G)$. The *geometric polar* (*polar* for short) of a closed convex set G containing the origin is defined as

$$G^\circ = \{y \mid y^{\mathrm{T}}x \leqslant 1\ \forall x \in G\}.$$

Polars play a very important role in the analysis to follow.

We define a *Semidefinite Program* to be an optimization problem of the form:

$$\begin{aligned} \sup \quad & c^{\mathrm{T}}x \\ \text{s.t.} \quad & \sum_{i=1}^{m} x_i Q_i \preceq Q_0, \end{aligned} \tag{P}$$

where Q_i, $i = 0, \ldots, m$ are in $\mathcal{S}_n$ ($n \times n$ symmetric matrices). Note that any semidefinite program as considered in [2] can be cast in the above form. We say that (P) is *homogeneous* if $Q_0 = 0$.

Definition 3. A *spectrahedron* is defined to be a closed convex set of the following type,

$$G = \{x \mid Q(x) \succeq 0\},$$

where $Q(x)$ is an affine symmetric matrix map.

It is clear that spectrahedra are precisely the feasible regions of semidefinite programs.

1.4. On the existing duality theories for SDP

The dual to be proposed (see Section 1.5) in this paper satisfies the following criteria:

(i) The dual is an explicit semidefinite program in polynomially many variables (involves $O(m)$ matrix variables) and constraints (these are $O(m^2n^2)$ in number). It can be written down mechanically (i.e., without performing any computations), and the bitlengths of its coefficients are polynomial in those of the primal.

(ii) The duality gap, which is the difference between the primal and the dual optimal objective function values, is zero whenever the primal is feasible and bounded. Also, in this case, the dual attains its optimum.

(iii) It yields a precise theorem of the alternative for semidefinite inequality systems, i.e., a characterization of the *infeasibility* of a semidefinite inequality in terms of the feasibility of another polynomial size semidefinite inequality.

We will now examine two previously known duality theories for SDP.

Lagrangian dual. The dual considered in [2, 18] satisfies (i) above, but needs to assume certain Slater-like constraint qualifications for (ii) and (iii) to hold.

Minimal cone based dual. On the other hand, the duality approach presented in [3] (when specialized to SDP) while theoretically satisfying (ii), has the following shortcoming: this dual is not an explicit semidefinite program as it requires certain computations to be performed in order to extract the dual. Moreover, these computations are not polynomial time.

Let us take a closer look at these two approaches.

1.4.1. The Lagrange–Slater dual for SDP

Here is the dual as stated [2] and [18] (note that we are using primal and dual in reversed sense here as compared to [2]). This dual is nothing but the Lagrangian dual obtained by treating SDP as a cone program. We will call it the *Lagrange–Slater Dual* (LSD) of (P).

$$\begin{aligned} \inf \quad & U \bullet Q_0 \\ \text{s.t.} \quad & U \bullet Q_i = c_i \quad \forall i = 1, \dots, m, \\ & U \succeq 0. \end{aligned} \tag{LSD}$$

Let us now state a condition, called the *Slater Constraint qualification*, under which the dual LSD entertains no duality gap:

$$\text{there exists an } x \text{ such that } \sum_i x_i Q_i \prec Q_0.$$

While it is not clear whether this condition or other similar regularity conditions are restrictive in practical applications, they severely limit the applicability of LSD for many theoretical concerns, such as theorems of the alternative (see Section 1.4.3).

The following is a very simple SDP, for which the LSD has a nonzero duality gap (slight modification of an example in [28]):

Example 4 (*SDP for which LSD has a duality gap*). Let the primal be

$$\begin{aligned} \sup \quad & x_2 \\ \text{s.t.} \quad & \begin{bmatrix} x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \preceq \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

where $\alpha > 0$. Then the LSD becomes

$$\begin{aligned} \inf \quad & \alpha U_{11} \\ \text{s.t.} \quad & U_{22} = 0, \\ & U_{11} + 2U_{23} = 1, \\ & U \succeq 0. \end{aligned}$$

Any primal feasible solution has $x_2 = 0$, and hence the primal optimum value is 0. Whereas, in the dual, $U_{22} = 0$ forces $U_{23} = 0$, so that $U_{11} = 1$ for every dual feasible solution. Therefore the dual optimum value is α, resulting in a duality gap of α.

An example of failure of a Farkas' Lemma that arises from LSD is given in Section 1.4.3 (Example 5). The remarkable simplicity of these examples reveals some shortcomings of Lagrangian duality as applied to semidefinite programming.

1.4.2. The minimal cone based approach

This approach was proposed by Borwein and Wolkowicz [3]. In simplified terms, they consider a *cone programming* problem

$$\sup\{c^{\mathrm{T}}x \mid b - Ax \in S\},$$

where S is a convex cone, which in this case will be the cone of PSD matrices. For a convex subset T of S, they define the *minimal cone* $S^{\mathrm{f}}(T)$ (slightly different notation being used here) of T to be the intersection of all the faces of S containing T. (A face F of a convex set S is a subset such that every line segment whose relative interior meets F and is contained in S is contained in F as well; see [27].) The following is an equivalent problem to the above cone-program:

$$\sup\{c^{\mathrm{T}}x \mid b - Ax \in S^{\mathrm{f}}(T)\},$$

where $T = \{b - Ax \mid x \in \mathbb{R}^m\} \cap S$. The authors have shown that the "standard Lagrange multiplier theorem" as applied to this modified problem has no duality gap. The authors then develop a theoretical algorithm for computing the minimal cone $S^{\mathrm{f}}(T)$. This approach is mathematically very interesting. And since the first appearance of the results here, certain geometric relations between the minimal cone methodology and our approach here have been demonstrated [26]. However, the main drawback of the approach is that it does not yield a polynomial size dual program. We explain this fact as follows.

Firstly, their algorithm when specialized to the SDP case requires the extraction of *exact* solutions to a sequence of semidefinite inequality systems. This task cannot be accomplished in polynomial time, because semidefinite systems, even when the data is rational, may not possess polynomial size solutions (in bitlength), as we will see in Section 4.2. Finally, as in the case of linear programming, it is evidently highly advantageous to be able to write down a polynomial size gap-free dual for any given SDP and not have to perform computations of some kind or another. For, after all, a duality theory for a mathematical program is an analytical tool and should provide a starting point in an algorithmic treatment of the problem and not vice versa.

1.4.3. Farkas' Lemma and certificates

A related issue is that of Farkas' Lemma and other *Theorems of the Alternative*. In the case of Linear Programming, Farkas' Lemma and its relatives characterize the solvability of one linear inequality system in terms of the unsolvability of another linear inequality system: The system $Ax \leqslant b$ is infeasible iff the system $A^{\mathrm{T}}y = 0$, $b^{\mathrm{T}}y = -1$, $y \geqslant 0$ is feasible. Such pairs of systems are called *Systems of the Alternative*.

Systems of the alternative essentially give "certificates" of infeasibility. Prior to the publication of the Ellipsoid Algorithm (in 1979) which showed that Linear Programming was polynomial time solvable, the LP Farkas' Lemma enabled one to conclude that the feasibility problem for linear inequalities is in NP ∩ Co-NP (see [9, Chapter 7]). The significance of the membership of a problem in NP ∩ Co-NP is derived from the widely held belief that P = NP ∩ Co-NP (see [9]).

Let us now turn to semidefinite inequality systems: Suppose that we want to characterize the feasibility/infeasibility of the system

$$Q(x) \succeq 0. \tag{4}$$

Then a direct analog of the LP Farkas' Lemma [2] gives the following as a system of alternative for (4).

$$Q_i \bullet U = 0 \quad \forall i = 1,\dots,m, \qquad Q_0 \bullet U = -1, \qquad U \succeq 0. \tag{5}$$

However, this extension is inadequate as witnessed in the following example, in which both (4) and (5) are simultaneously infeasible.

Example 5 (*Failure of "Farkas' Lemma"*). Consider the inequality

$$Q(x) = \begin{bmatrix} x & 1 & 0 \\ 1 & y & 0 \\ 0 & 0 & -x \end{bmatrix} \succeq 0,$$

and observe that it is infeasible. The system (5) for this system is given by

$$U_{33} - U_{11} = 0, \qquad -U_{22} = 0, \qquad 2U_{12} = -1, \qquad U \succeq 0,$$

which is also infeasible since a positive semidefinite U with $U_{22} = 0$ will force U_{12} to be zero.

We will derive a polynomial size explicit semidefinite system of the alternative for (4), and employ it to establish the following complexity results:

(i) If SDFP ∈ NP, then SDFP ∈ Co-NP, and vice versa.

(ii) In the Turing Machine model [9], SDFP is not NP-complete unless NP = Co-NP,

(iii) SDFP is in NP∩Co-NP in the real number model of Blum, Shub and Smale [5].

(iv) There are polynomial time reductions from the following problems to SDFP:

  (a) Checking whether a feasible SDP is bounded (i.e., it has a finite optimal value).

  (b) Checking whether a feasible and bounded SDP attains the optimum.

  (c) Checking the optimality of a given feasible solution.

1.5. The proposed dual for SDP

1.5.1. Statement of the dual

Let us define the following.

- $G := \{x \mid \sum_i x_i Q_i \preceq Q_0\}$ is the feasible region of (P).
- $\hat{Q}(x) = \sum_{i=1}^{m} x_i Q_i$.
- $Q(x) = Q_0 - \hat{Q}(x)$ (so that $G = \{x \mid Q(x) \succeq 0\}$).
- $Q^* : \mathcal{M}_n \to \mathbb{R}^m$ is defined by $Q^*(U) = U \bullet Q_i$, $i = 1,\dots,m$. (A comment is in order here. Usually, the superscript $*$ is used to denote the adjoints of linear maps. Here, Q is an *affine* map. We think of Q^* as the adjoint of the linear part of Q. We do not use $\hat{Q}^*$ to avoid cumbersome notation.)
- $Q^\# : \mathcal{M}_n \to \mathbb{R}^{m+1}$, defined by

$$Q^\#(U) = \begin{pmatrix} Q_0 \bullet U \\ Q^*(U) \end{pmatrix}.$$

We now introduce a dual semidefinite program, named the *Extended Lagrange–Slater Dual* (ELSD) for SDP.

$$\begin{aligned}\inf\quad & (U+W_m)\bullet Q_0\\ \text{s.t.}\quad & Q^*(U+W_m)=c,\\ & Q^{\#}(U_i+W_{i-1})=0,\quad i=1,\dots,m,\\ & U_i\succeq W_iW_i^{\mathrm{T}},\quad i=1,\dots,m,\\ & U\succeq 0,\\ & W_0=0.\end{aligned}\qquad\text{(ELSD)}$$

Note that the constraint $U_i \succeq W_iW_i^{\mathrm{T}}$ can alternately be written as (see Proposition 7(4) of Section 2.2).

$$\begin{bmatrix} I & W_i^{\mathrm{T}}\\ W_i & U_i\end{bmatrix}\succeq 0,$$

and consequently (ELSD) is indeed a semidefinite program. A glossary of the domains of the variables is: $U\in\mathcal{S}_n, U_i\in\mathcal{S}_n$ $\forall i=1,\dots,m$ and $W_i\in\mathcal{M}_n$ $\forall i=1,\dots,m$ (and we use an auxiliary matrix variable $W_0=0$ for notational convenience).

Size of ELSD. Note that ELSD has $\mathrm{O}(mn^2)$ number of variables and $\mathrm{O}(m^2n^2)$ number of constraints, and the coefficient data is essentially that of the primal. Hence ELSD is a polynomial size semidefinite program.

Let us define for an arbitrary positive integer k,

$$\begin{aligned}\mathcal{C}_k&=\{(U_i,W_i)_{i=1}^k \mid Q^{\#}(U_i+W_{i-1})=0,\ U_i\succeq W_iW_i^{\mathrm{T}},\ \forall i=1,\dots,k,\ W_0=0\},\\ \mathcal{U}_k&=\{U_k\mid (U_i,W_i)_{i=1}^k\in\mathcal{C}_k\},\\ \mathcal{W}_k&=\{W_k\mid (U_i,W_i)_{i=1}^k\in\mathcal{C}_k\},\end{aligned}$$

It will be shown in Lemma 10 that the sequences of sets $\mathcal{U}_k$ and $\mathcal{W}_k$ are increasing with respect to containment, i.e., $\mathcal{U}_1\subseteq\mathcal{U}_2\subseteq\cdots$, etc. In terms of this notation, we can write ELSD in a more compact form. Future references to (ELSD) will be to this form of the dual.

$$\begin{aligned}\inf\quad & (U+W)\bullet Q_0\\ \text{s.t.}\quad & Q^*(U+W)=c,\\ & W\in\mathcal{W}_m,\\ & U\succeq 0.\end{aligned}\qquad\text{(ELSD)}$$

A pair (U,W) is said to be *dual feasible*, if it satisfies the above constraints. For the purposes of our proofs, we need to consider the following weakening of (ELSD), which we call the Weak-ELSD.

$$\begin{aligned} \inf \quad & (U+W) \bullet Q_0 \\ \text{s.t.} \quad & Q^*(U+W) = c, \\ & W \in \mathcal{W}_{m-1}, \\ & U \succeq 0. \end{aligned} \qquad \text{(Weak-ELSD)}$$

The term "weak" is used since, while the optimal value of this dual is the same as primal and (ELSD) (by the duality theorem to follow), it is not known if this dual attains its optimal value. A feasible solution (U, W) to (Weak-ELSD) will be called *weakly dual feasible.*

1.5.2. Strong duality theorem for SDP

In the sections to follow, we will prove the following strong duality theorem, which closely resembles duality theorems for linear programming.

Theorem 6 (Duality Theorem). *The following hold for the primal problem* (P) *and the dual problems* (ELSD) *and* (Weak-ELSD):

(i) (Weak duality) *If x is primal feasible and (U, W) is dual feasible (or weakly dual feasible), then $c^{\mathrm{T}}x \leqslant (U+W) \bullet Q_0$.*

(ii) (Primal boundedness) *If the primal is feasible, then its optimal value is finite if and only if the dual* (ELSD) *(or* (Weak-ELSD)*) is feasible.*

(iii) (Zero gap) *If both the primal and the dual* (ELSD) *(or* (Weak-ELSD)*) are feasible, then the optimal values of all the three programs are equal.*

(iv) (Dual attainment) *Whenever the common optimal value of the primal and* (ELSD) *are finite, the latter attains this value.*

The only difference between a standard LP Duality theorem and the above is borne by the fact that some semidefinite programs (i.e., primal) may not achieve their optimum values (Example 24 in Section 4.2). The proof of the theorem is given in Section 2.5.

2. Derivation of the dual

2.1. A brief sketch of the derivation

The technique used here emerged from the results of [23], where certain geometric properties of spectrahedra (G) were investigated. In particular, in order to understand the polars (G°) of spectrahedra, an object called the algebraic polar(G^*) was introduced (see Section 2.4 here). This is an "algebraic approximation" of the polar, which is a geometrically defined object. The relations established there were: (1) $G^\circ = \mathrm{Cl}(G^*)$ and (2) $G^\circ = G^* + \mathrm{Aff}(G)^\perp$ (Lemmas 13 and 15 here). A closer examination of the proof of (2) revealed that one can extract very useful information concerning $\mathrm{Aff}(G)$ whenever G^* turns out not to be closed (and hence not equal to G°). This key observation is exploited here in Section 2.4.1.

In the next subsection, we list a few simple results concerning positive semidefinite matrices and convex sets, which are used in our proofs. Then we establish weak duality for ELSD in Section 2.3. In Section 2.4, we establish an exact algebraic description of G°. This description readily yields the strong duality properties of the dual ELSD (proof in Section 2.5). Subsequently, a theorem of the alternative for semidefinite systems, and further properties of SDP are characterized in Section 3. Finally, in Section 4, we establish the previously stated complexity results for SDP.

2.2. *Some useful facts about PSD matrices*

We now collect as a proposition some useful and well-known properties of semidefinite matrices.

Proposition 7. *The following hold*:

(i) *If $U \succeq 0$, and $U_{ii} = 0$, then $U_{ij} = 0\ \forall j$.*

(ii) *If $A \in \mathcal{S}_n$ and $S \in \mathcal{M}_n$ and nonsingular, then $A \succeq 0$ iff $S^\mathrm{T}AS \succeq 0$.*

(iii) *If $A \succeq 0$ and $u \in \mathbb{R}^n$, then $u^\mathrm{T}Au = 0$ iff $Au = 0$.*

(iv) *Given a block partitioned matrix*

$$U = \begin{bmatrix} A & B^\mathrm{T} \\ B & C \end{bmatrix},$$

where A and C are square, and A is nonsingular, then the Schur Complement *of A in U is the matrix $S = C - BA^{-1}B^\mathrm{T}$. We have that, if $A \succ 0$, then $U \succeq 0 \Leftrightarrow S \succeq 0$ and $U \succ 0 \Leftrightarrow S \succ 0$.*

(v) *$A \in \mathcal{S}_n$ is PSD iff $A \bullet B \geqslant 0$ for all $B \succeq 0$.*

(vi) *If $A, B \succeq 0$, then $A \bullet B = 0$ if and only if $AB = 0$.*

(vii) *If $U - WW^\mathrm{T} \succeq 0$, then there exists a matrix $H \in \mathcal{M}_n$ such that $W = UH$.*

(viii) *If $U_1, U_2 \succeq 0$, then* $\mathrm{Null}(U_1 + U_2) = \mathrm{Null}(U_1) \cap \mathrm{Null}(U_2)$.

Proof. Most of these results are quite classical, and may be found in texts on matrix analysis [14]. We make critical use of (vi) and (vii). The proof of (vi) may be found in [2], and that of (vii) is given here: Suppose that $U - WW^\mathrm{T} \succeq 0$. Then, for an arbitrary $x \in \mathbb{R}^n$, if $Ux = 0$, then $x^\mathrm{T}Ux = 0$, and $x^\mathrm{T}(U - WW^\mathrm{T})x \geqslant 0$. Hence $0 \geqslant x^\mathrm{T}WW^\mathrm{T}x$, implying that $W^\mathrm{T}x = 0$. Thus, $Ux = 0 \Rightarrow W^\mathrm{T}x = 0$, and hence there must exist a matrix H such that $W^\mathrm{T} = H^\mathrm{T}U$, and result follows from the symmetry of U. □

Another simple folklore result we need is the following.

Proposition 8. *If $K \subseteq \mathbb{R}^m$ is a convex set such that* $\mathrm{Cl}(K) = \mathbb{R}^m$, *then $K = \mathbb{R}^m$.*

Proof. From Theorem 6.3 of [27],

$$\mathrm{ri}(K) = \mathrm{ri}(\mathrm{Cl}(K)) = \mathbb{R}^m,$$

and hence $K = \mathbb{R}^m$. □

2.3. Weak duality

In this section, we will establish weak duality for (ELSD). We need the following lemma.

Lemma 9. *Let $k \leqslant m$. Then for any $x \in G$ and $U \in \mathcal{U}_k$, $W \in \mathcal{W}_k$, we have*

$$Q(x)U = 0, \qquad Q(x)W = 0,$$

and hence $Q(x) \bullet U = 0 = Q(x) \bullet W$. In particular, if $0 \in G$, then $Q_0 \bullet U = 0 = Q_0 \bullet W$.

Proof. Let $x \in G$ and

$$(U_i, W_i)_{i=1}^{k} \in \mathcal{C}_k.$$

To start with, note that as $W_0 = 0$, and so $Q^{\#}(U_1) = 0$, we have

$$Q(x) \bullet U_1 = Q_0 \bullet U_1 - \sum_{i=1}^{m} x_i (Q_i \bullet U_1) = 0,$$

Since $Q(x), U_1 \succeq 0$, by Proposition 7(vi),

$$Q(x)U_1 = 0.$$

Since $U_1 \succeq W_1 W_1^{\mathrm{T}}$, by Proposition 7(vii), we infer that there exists a matrix H such that

$$W_1 = U_1 H,$$

and consequently,

$$Q(x)W_1 = 0 \quad \text{and} \quad Q(x) \bullet W_1 = 0.$$

Since $Q^{\#}(U_2 + W_1) = 0$, the above gives

$$Q(x) \bullet U_2 = 0,$$

which as above gives $Q(x)U_2 = 0$. It is clear that by continuing in this fashion, one can conclude the truth of the lemma. □

That weak duality holds for the primal–dual pairs (P)–(ELSD) and (P)–(Weak-ELSD) is shown below.

Proof of Weak Duality (Theorem 6(i)). Let x be primal feasible, $k \leqslant m$, and (U, W) satisfy

$$Q^{*}(U + W) = c,$$

$$W \in \mathcal{W}_k,$$

$$U \succeq 0.$$

We then have

$$\begin{aligned} c^{\mathrm{T}}x &= x^{\mathrm{T}}Q^*(U+W) \\ &= -Q(x)\bullet(U+W)+Q_0\bullet(U+W) \\ &= -Q(x)\bullet U - Q(x)\bullet W + Q_0\bullet(U+W). \end{aligned}$$

Since $Q(x), U \succeq 0$, $-Q(x)\bullet U \leqslant 0$. From the above lemma, $Q(x)\bullet W = 0$, and we get

$$c^{\mathrm{T}}x \leqslant Q_0\bullet(U+W).$$

Applying the above with $k = m$ yields that weak duality holds for the pair (P)–(ELSD). □

We will establish the following lemma for use in future sections. It sheds some light on the incremental structure of the dual (ELSD).

Lemma 10. *The following hold*:

(i) *The sequences $\mathcal{U}_k$ and $\mathcal{W}_k$ are increasing (set) sequences.*
(ii) *For any $k \leqslant m$, $\mathcal{W}_k \subseteq \mathcal{M}_n$ and $Q^*(\mathcal{W}_k) \subseteq \mathbb{R}^m$ are linear subspaces.*
(iii) $Q^*(\mathcal{W}_1)\subseteq\cdots\subseteq Q^*(\mathcal{W}_m)$.
(iv) *If $0\in G$, then $G\subseteq (Q^*(\mathcal{W}_k))^{\perp}$ for every $k = 1,\dots,m$.*

Proof. The proof of (i) is as follows: suppose that

$$(U_i, W_i)_{i=1}^{k} \in \mathcal{C}_k,$$

then set

$$U_i' = U_{i-1}, \qquad W_i' = W_{i-1} \quad \forall i = 2,\dots,k+1, \quad \text{and} \quad U_1' = 0 = W_1',$$

and observe that

$$(U_i', W_i')_{i=1}^{k+1} \in \mathcal{C}_{k+1},$$

and hence $W_k \in \mathcal{W}_k$ and $U_k \in \mathcal{U}_k$.

To prove (ii), let $W \in \mathcal{W}_k$ and let

$$(U_i, W_i)_{i=1}^{k} \in \mathcal{C}_k$$

such that $W = W_k$. We will first show that $\mu W \in \mathcal{W}_k$ for every $\mu \in \mathbb{R}$. Define the matrices

$$W_i' = \mu^{2^{k-i}} W_i, \qquad U_i' = \mu^{2^{k-i+1}} U_i$$

for every $i = 1,\dots,k$. Then,

$$Q^{\#}(U_i' + W_{i-1}') = \mu^{2^{k-i+1}} Q^{\#}(U_i + W_{i-1}) = 0,$$

and

$$U_i' - W_i'W_i'^{\mathrm{T}} = \mu^{2^{k-i+1}}(U_i - W_iW_i^{\mathrm{T}}) \succeq 0,$$

and it follows that

$$(U_i', W_i')_{i=1}^k \in \mathcal{C}_k,$$

hence

$$\mu W_k \in \mathcal{W}_k.$$

It can now be concluded that $\mathcal{W}_k$ is a subspace, once we show that $\mathcal{W}_k$ is convex, for which it suffices to show that $\mathcal{C}_k$ is convex, which follows trivially using the fact that one can write the condition $U - WW^{\mathrm{T}} \succeq 0$ as

$$\begin{bmatrix} I & W^{\mathrm{T}} \\ W & U \end{bmatrix} \succeq 0.$$

Clearly (i) implies (iii), and the proof is complete, as (iv) is an easy consequence of Lemma 9. □

Example 4 revisited. In Section 1.4.1, it was seen that the LSD fails for Example 4: there was a duality gap of $\alpha > 0$. Now, let us apply ELSD to this example. Since $m = 2$, we get:

$$\begin{aligned}
\inf\quad & (U + W_2) \bullet Q_0 \\
\text{s.t.}\quad & Q^*(U + W_2) = c, \\
& Q^{\#}(U_2 + W_1) = 0, \\
& Q^{\#}(U_1) = 0, \\
& U_1 \succeq W_1 W_1^{\mathrm{T}}, \\
& U_2 \succeq W_2 W_2^{\mathrm{T}}, \\
& U \succeq 0.
\end{aligned} \qquad \text{(ELSD)}$$

In this case, $Q_0 = \alpha e_1 e_1^{\mathrm{T}}$, $Q_1 = e_2 e_2^{\mathrm{T}}$, $Q_2 = e_1 e_1^{\mathrm{T}} + e_2 e_3^{\mathrm{T}} + e_3 e_2^{\mathrm{T}}$, and $c = (0, 1)^{\mathrm{T}}$. It is not hard to see that the solution $U = U_1 = W_1 = 0$, $U_2 = e_3 e_3^{\mathrm{T}}$ and $W_2 = e_3 e_2^{\mathrm{T}}$ is feasible for the ELSD, and the objective function value of this solution is 0, and by weak duality, it must be optimal to the ELSD. Hence, ELSD gives zero duality gap as well as attains its optimum for the SDP of Example 4, and as we will show in the next two sections, the same holds for any SDP.

2.4. Polars of spectrahedra

In this section, we derive an expression for the polar of an arbitrary spectrahedron defined by

$$G = \{x \mid Q(x) \succeq 0\},$$

and containing the origin, i.e., $Q_0 \succeq 0$.

The polar is defined to be

$$G^\circ = \{y \mid x^{\mathrm{T}}y \leqslant 1,\ \forall x \in G\}.$$

The following is a standard result and the proofs can be found in [27].

Proposition 11. *The following hold for a closed convex set G containing the origin:*
(i) *G° is a closed convex set containing the origin.*
(ii) $G = G^{\circ\circ}$.

The central result of this paper will now be established.

2.4.1. Description of the polar

Theorem 12. *If G contains the origin, then its polar is given by*

$$G^\circ = \{Q^*(U + W) \mid W \in \mathcal{W}_k,\ U \succeq 0,\ U \bullet Q_0 \leqslant 1\}\ \forall k \geqslant m - 1.$$

The proof will be given shortly, but first consider the following object, called the *algebraic polar* of G (with respect to the representation $Q(x) \succeq 0$). It serves as an "initial approximation" for G°:

$$G^* = \{Q^*(U) \mid U \bullet Q_0 \leqslant 1,\ U \succeq 0\}.$$

We reproduce here the proof that G^* indeed approximates G° (missing only a set of measure zero) closely.

Lemma 13. *Assuming that* $0 \in G$, *we have* $G^\circ = \mathrm{Cl}(G^*)$.

Proof. Let us first prove that $G^\circ \supseteq G^*$:

$$\begin{aligned} G^\circ &= \{y \mid \sup\{y^{\mathrm{T}}x \mid Q(x) \succeq 0\} \leqslant 1\} \\ &\supseteq \{Q^*(U) \mid U \bullet Q_0 \leqslant 1,\ U \succeq 0\} \\ &= G^*. \end{aligned}$$

Since G° is always closed, $\mathrm{Cl}(G^*) \subseteq G^\circ$. To show the reverse inclusion, put $H = \mathrm{Cl}(G^*)$ and consider any $w \in H^\circ$. Then, we have

$$w^{\mathrm{T}}(Q^*(U)) \leqslant 1 \quad \text{whenever } U \succeq 0,\ U \bullet Q_0 \leqslant 1. \tag{10}$$

We claim that $Q(w) \succeq 0$. First, if $Q_0 = 0$, then (10) implies

$$Q(w) \bullet U \geqslant -1 \quad \forall U \succeq 0,$$

which happens if and only if $Q(w) \succeq 0$.

Suppose that $Q_0 \neq 0$. Let $V \succ 0$, and choose $\lambda > 0$ such that $\lambda V \bullet Q_0 = 1$, which is possible since $Q_0 \succeq 0$ and nonzero. Then

$$(\lambda V) \bullet Q(w) = (\lambda V) \bullet Q_0 - w^{\mathrm{T}} Q^*(\lambda V) = 1 - w^{\mathrm{T}} Q^*(\lambda V) \geqslant 0 \quad \text{by (10)}.$$

Therefore, $Q(w) \bullet V \geqslant 0 \ \forall V \succ 0$, implying that $w \in G$, and hence $H^{\circ} \subseteq G$. But then,

$$H = H^{\circ\circ} \supseteq G^{\circ},$$

and the proof is complete. □

We will need the following corollary, which generalizes Lemma 13.

Corollary 14. *Let* $0 \in G$, *and* $T = \{x \mid Ax = 0\}$ *be a subspace. Then*

$$(G \cap T)^{\circ} = \mathrm{Cl}(G^* + T^{\perp}).$$

Proof. We give below a spectrahedral representation of $G \cap T$. Let

$$\tilde{Q}(x) := Q(x) \oplus \mathrm{Diag}(Ax) \oplus \mathrm{Diag}(-Ax).$$

Clearly,

$$G \cap T = \{x \mid \tilde{Q}(x) \succeq 0\}.$$

It is not difficult to show that the algebraic polar of $G \cap T$ with respect to the matrix map $\tilde{Q}(x)$ is precisely $G^* + T^{\perp}$. Now we can apply Lemma 13 and conclude the truth of the corollary. □

The following fact was established in [23] and its proof is omitted as we will not use the result directly here.

Lemma 15. *Let* $G = \{x \mid Q(x) \succeq 0\}$ *with* $Q_0 \succeq 0$. *Then*

$$G^{\circ} = \mathrm{Cl}(G^*) = G^* + \mathrm{Aff}(G)^{\perp}.$$

The essence of the lemma is that the difference of the polar G° and its algebraic approximation G^* is offset by the orthogonal complement of the affine hull of G. However, the disadvantage of the above lemma is that we need to first have at our disposal a description of the affine hull of G, say as $\{x \mid Ax = 0\}$ for some A. Since such a representation is not readily available, we attack the problem as follows.

As seen, if G^* turns out to be closed, we already have a description of G°. So suppose that this is not the case. Interestingly enough, from the nonclosedness, we can obtain information on the annihilators of G (as will be witnessed in Claims 16 and 17 below).

Proof of Theorem 12. Let us first consider the algebraic polar G^* of G:

$$G^* = \{Q^*(U) \mid U \bullet Q_0 \leqslant 1,\ U \succeq 0\},$$

and the subspaces (see Lemma 10) of $\mathbb{R}^m$

$$S_0 = \{0\}, \qquad S_k = Q^*(\mathcal{W}_k) \quad \forall k = 1, \ldots, m,$$

which satisfy

$$S_0 \subseteq S_1 \subseteq S_2 \subseteq \cdots \subseteq S_m.$$

(There is some potential for confusion between the matrix space $\mathcal{S}_n$ and the subspaces S_i of $\mathbb{R}^m$ being used here.)

Let us invoke the assumption that $0 \in G$, and apply Lemma 10 (with $x = 0$) to conclude that for any $W \in \mathcal{W}_k$,

$$Q_0 W = 0,$$

and hence by another application of Lemma 10, we get that

$$x^{\mathrm{T}} Q^*(W) = 0 \quad \forall x \in G.$$

This implies easily that

$$G^* + S_k \subseteq G^\circ.$$

Therefore, we have

$$G^* \subseteq G^* + S_1 \subseteq G^* + S_2 \subseteq \cdots \subseteq G^\circ = \mathrm{Cl}(G^*),$$

and we need to show that $G^* + S_{m-1} = G^\circ$. If for some $k \leqslant m$, $G^* + S_k$ is closed, then clearly

$$G^* + S_j = G^\circ \quad \forall j \geqslant k.$$

We will show (Claim 16 below) that whenever $G^* + S_k$ (with $0 \leqslant k \leqslant m - 1$) is not closed, then

$$S_k \subset S_{k+1},$$

and therefore

$$\dim(S_k) < \dim(S_{k+1}).$$

We now assert that the above implies

$$G^\circ = G^* + S_{m-1}.$$

For the sake of a contradiction, suppose that this is not the case.

If $G^* + S_{m-1}$ is not closed, then $G + S_k$ is not closed for every $0 \leqslant k \leqslant m - 1$, which by Claim 16 will result in

$$0 = \dim(S_0) < \dim(S_1) < \dim(S_2) < \cdots < \dim(S_{m-1}) < \dim(S_m),$$

implying that $S_m = \mathbb{R}^m$. But by Lemma 10(iv), this implies that $G = \{0\}$, and hence $G^\circ = \mathbb{R}^m$. But then $\mathrm{Cl}(G^*) = \mathbb{R}^m$, and by Proposition 8, the only convex set whose closure is the whole space is the whole space itself and hence $G^* = \mathbb{R}^m$, a contradiction. □

Let us now prove the claim made in the above proof.

Claim 16. *If $G^* + S_k$ is not closed, then*

$$S_k \subset S_{k+1} \quad \textit{and} \quad \dim(S_k) < \dim(S_{k+1}).$$

Proof. Let A be an $r \times m$ full row rank matrix such that

$$T := S_k^{\perp} = \{x \mid Ax = 0\},$$

and suppose that $G^* + S_k = G^* + T^{\perp}$ is not closed.

Our main tool is the result in Claim 17 below, according to which there exists a matrix $U \succeq 0$ and a vector $\lambda \in \mathbb{R}^r$ such that

$$Q^*(U) + A^{\mathrm{T}}\lambda = 0, \qquad Q_0 \bullet U = 0$$

and $\{x \mid Ax = 0,\ Q(x)U = 0\}$ is a subspace with a strictly smaller dimension than T. We will use the quantities U, λ as follows. First, note that

$$T^{\perp} = S_k = Q^*(\mathcal{W}_k),$$

and hence we may choose $W \in \mathcal{W}_k$ such that

$$A^{\mathrm{T}}\lambda = Q^*(W).$$

Since $W \in \mathcal{W}_k$, $Q_0 W = 0$ from an application of Lemma 13, and hence

$$Q^*(U + W) = 0, \qquad Q_0 \bullet (U + W) = 0.$$

From the claim, the subspace

$$\hat{T} := \{x \in T \mid Q(x)U = 0\} = \left\{x \in T \;\middle|\; \sum_{i=1}^{m} x_i Q_i U = 0\right\}$$

(the equality follows from $Q_0 U = 0$) is strictly contained in the subspace $T = S_k^{\perp}$. Hence there exists $v \in T$ that is orthogonal to every vector in $\hat{T}$, i.e., whenever

$$\sum_{i=1}^{m} x_i Q_i U = 0,$$

one has $v^{\mathrm{T}}x = 0$. Now, treating the matrix equation in the display as n^2 separate scalar equations, one can conclude that there exists a multiplier (not necessarily symmetric) matrix $V \in \mathcal{M}_n$ such that

$$v_i = V \bullet (Q_i U) = (Q^*(UV))_i \quad \forall i = 1, \ldots, m.$$

The proof of the claim will be completed by showing that the vector $v = Q^*(UV)$ is contained in $S_{k+1} = Q^*(\mathcal{W}_{k+1})$. Let

$$(U_i, W_i)_{i=1}^{k} \in \mathcal{C}_k,$$

such that $W_k = W$. Obviously neither U nor V is zero, for otherwise, $v = 0$ will be in the subspace S_k. Define

$$\mu = 1/(\rho(V)\sqrt{\rho(U)}),$$

and set $U_{k+1} = U$ and $W_{k+1} = \mu UV$. Let us show that $U_{k+1} \succeq W_{k+1}W_{k+1}^{\mathrm{T}}$:

$$\begin{aligned} W_{k+1}W_{k+1}^{\mathrm{T}} &= (1/(\rho(U)\rho(V)^2))UVV^{\mathrm{T}}U \\ &\preceq (1/\rho(U))U^2 \\ &\preceq U = U_{k+1}. \end{aligned}$$

It is clear that

$$(U_i, W_i)_{i=1}^{k+1} \in C_{k+1},$$

and hence

$$\mu v = Q^*(W_{k+1}) \in S_{k+1},$$

and since the latter is a subspace, it contains v. The proof of Claim 16 is complete. □

The proof of the following Claim 17 was inspired and closely parallels that of Lemma 15 given in [23].

Claim 17. *Let $T = \{x \mid Ax = 0\}$, where A is an $r \times m$ full row rank matrix, and let $G = \{x \mid Q(x) \succeq 0\}$ be any spectrahedron containing the origin. If $G^* + T^\perp$ is not closed, then there exist $U \succeq 0$ and $\lambda \in \mathbb{R}^r$ such that*

$$Q^*(U) + A^{\mathrm{T}}\lambda = 0, \qquad Q_0 \bullet U = 0$$

and $\{x \mid Ax = 0,\ Q(x)U = 0\}$ is a subspace whose dimension is strictly smaller than T.

Proof. The proof is by induction on n (the size of the matrices Q_i), with the base case being $n = 1$, which is trivial to verify. Put $Z = G^* + T^\perp$ and suppose that $w \in \mathrm{Cl}(Z) \setminus Z$, and pick a sequence

$$[U(i) \succeq 0,\ \lambda(i) \in \mathbb{R}^r], \quad i = 1, \dots, \infty$$

for which

$$U(i) \bullet Q_0 \leqslant 1 \quad \forall i \quad \text{and} \quad \lim_{i\to\infty} Q^*(U(i)) + A^{\mathrm{T}}\lambda(i) = w \notin Z.$$

If the sequence $[U(i), \lambda(i)]$ has an accumulation point, say, $[\tilde{U}, \tilde{\lambda}]$, then we will have $w = Q^*(\tilde{U}) + A^{\mathrm{T}}\tilde{\lambda}$. Since $\tilde{U} \succeq 0$ and $\tilde{U} \bullet Q_0 \leqslant 1$, this will imply that $w \in Z$, contradicting our hypothesis. Hence, of the two sequences $\{\mathrm{Tr}(U(i)) = U(i) \bullet I\}$, and $\{\|\lambda(i)\|\}$, at least one is unbounded.

Now define the normalized sequence

$$[\hat{U}(i), \hat{\lambda}(i)] = [U(i), \lambda(i)]/(U(i) \bullet I + \|\lambda(i)\|) \quad \forall i,$$

and assume, by passing to a subsequence if necessary, that the sequence converges to $(\hat{U}, \hat{\lambda})$. Note that the following hold:

- $\hat{U} \succeq 0$.
- $Q^*(\hat{U}) + A^{\mathrm{T}}\hat{\lambda} = \lim w/(U(i) \bullet I + \|\lambda(i)\|) = 0$.
- At least one of $\hat{U}$ or $\hat{\lambda}$ is nonzero.
- In fact, $\hat{U} \neq 0$. This can be seen as follows: if $\hat{U} = 0$, then $A^{\mathrm{T}}\hat{\lambda} = 0$, which contradicts the fact that A has full row rank.
- Also, since $0 \leqslant U(i) \bullet Q_0 \leqslant 1$ $\forall i$ (as $Q_0 \succeq 0$ from $0 \in G$), it follows that $\hat{U} \bullet Q_0 = 0$.

Now consider the affine subspace

$$T' = \{x \mid Ax = 0, Q(x)\hat{U} = 0\}.$$

Since $Q_0 \bullet \hat{U} = 0$ (and $Q_0 \succeq 0$), by Proposition 7(vi), $Q_0\hat{U} = 0$, and hence

$$T' = \{x \mid Ax = 0, \ \hat{Q}(x)\hat{U} = 0\},$$

and consequently T' is a linear subspace.

Now, if $\dim(T') < \dim(T)$, we are done. Let us suppose to the contrary. This implies that $T' = T$. Let spectral decomposition of $\hat{U}$ be given by

$$\hat{U} = X \begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix} X^{\mathrm{T}},$$

where D is the diagonal matrix made up of the positive eigenvalues (which are $l > 0$ in number) of $\hat{U}$, and the columns of X are composed of a set of orthonormal eigenvectors of $\hat{U}$. Also, block partition $X^{\mathrm{T}}Q(x)X$ accordingly

$$X^{\mathrm{T}}Q(x)X = \begin{bmatrix} P(x) & R(x)^{\mathrm{T}} \\ R(x) & S(x) \end{bmatrix}, \tag{11}$$

giving

$$Q(x)\hat{U} = X \begin{bmatrix} 0 & R(x)^{\mathrm{T}}D \\ 0 & S(x)D \end{bmatrix} X^{\mathrm{T}}, \quad \text{and}$$

$$T' = \{x \mid Ax = 0, \ R(x) = 0, \ S(x) = 0\}.$$

Since $T' = T$, we have

$$Ax = 0 \quad \Rightarrow \quad R(x) = 0 \text{ and } S(x) = 0,$$

implying that

$$G \cap T = \{x \mid P(x) \succeq 0, \ Ax = 0\} = H \cap T,$$

where H is the spectrahedron defined by

$$H = \{x \mid P(x) \succeq 0\}.$$

Since $Q(x) \succeq 0$ iff $X^{\mathrm{T}}Q(x)X \succeq 0$, and a principal submatrix of a PSD matrix is PSD, it follows from (11) that $G \subseteq H$. From the definition of the algebraic polar given in Section 2.4.1, it is quite evident that the algebraic polar of G with respect to the map $X^{\mathrm{T}}Q(x)X$ is the same as that with respect to $Q(x)$. Since $P(x)$ is a pricipal submatricial map of $X^{\mathrm{T}}Q(x)X$, it follows that the algebraic polar of H with respect to $P(x)$ is contained in G^*, i.e., $H^* \subseteq G^*$, and so

$$H^* + T^{\perp} \subseteq G^* + T^{\perp}.$$

We have,

$$\begin{aligned}\mathrm{Cl}(H^* + T^{\perp}) = (H \cap T)^{\circ} = (G \cap T)^{\circ} = \mathrm{Cl}(G^* + T^{\perp})\\ \supseteq G^* + T^{\perp} \supseteq H^* + T^{\perp},\end{aligned}$$

in which two applications of Corollary 14 are made. If $H^* + T^{\perp}$ is closed, then equality holds throughout, and in particular, $G^* + T^{\perp}$ is closed. Hence we may assume that $H^* + T^{\perp}$ is not closed. But since the dimension of the matrix map $P(x)$ is $n - l < n$, by induction there exists an $(n-l) \times (n-l)$ matrix $V \succeq 0$ and $\tilde{\lambda} \in \mathbb{R}^r$ such that

$$P^*(V) + A^{\mathrm{T}}\tilde{\lambda} = 0, \qquad V \bullet P_0 = 0,$$

and

$$\{x \mid Ax = 0,\ P(x)V = 0\} \subset T.$$

To obtain the U, λ as required by the claim, we just set

$$U = X \begin{bmatrix} V & 0 \\ 0 & D \end{bmatrix} X^{\mathrm{T}},$$

and $\lambda = \hat{\lambda} + \tilde{\lambda}$. The proof of Claim 17 is now complete. □

2.4.2. *Value and gauge functions*

With a complete description of the polar of a spectrahedron at hand, it is quite easy to derive a strong dual for SDP, namely ELSD. The underlying notions behind our proof of strong duality are those of support and gauge functions, which we now define for an arbitrary closed convex set G containing the origin.

The *value function* (also known as support function) of G is defined as

$$v_G(y) = \sup\{y^{\mathrm{T}}x \mid x \in G\},$$

and the *gauge function* is defined by

$$\rho_G(y) = \inf\{t > 0 \mid y \in tG\}.$$

(By convention, infimum is ∞ if for no $t > 0$, $y \in tG$.)

The following lemma, linking polars, support functions and gauges, serves as a prototype for the strong duality proof in the next section.

Lemma 18. *For a closed convex set $0 \in G \subseteq \mathbb{R}^m$, and a vector $c \in \mathbb{R}^m$, the following hold:*

(i) *$c^{\mathrm{T}}x$ is bounded on G iff there exists $t > 0$ such that $tc \in G^\circ$.*

(ii) *The origin maximizes $c^{\mathrm{T}}x$ over G iff $tc \in G^\circ$ for every $t > 0$.*

(iii) *When $c^{\mathrm{T}}x$ is bounded on nonempty G, then $v_G(y) = \rho_{G^\circ}(y)\ \forall y \in \mathbb{R}^m$.*

Proof. Let $y \in \mathbb{R}^m$.

$$\begin{aligned}\rho_{G^\circ}(y) &= \inf\{t > 0 \mid y \in tG^\circ\}\\ &= \inf\{t > 0 \mid x^{\mathrm{T}}y \leqslant t\ \forall x \in G\}\\ &= \sup\{y^{\mathrm{T}}x \mid x \in G\}\\ &= v_G(y).\end{aligned}$$

The lemma trivially follows. □

2.5. Proof of strong duality for ELSD

Part (i) of Theorem 6, namely weak duality, has already been established. We will first prove (ii) and (iii) of Theorem 6, and then apply the (Weak-ELSD) version of (ii) to prove (iv), namely, dual attainment.

First, let us consider the following simplification that allows us to assume that $0 \in G$. In the special case of $Q_0 = 0$, 0 is in G, and in the general case, since we may assume for the purposes of Theorem 1 that the primal is feasible, let $\overline{x} \in G$. Let us perform the simple translation of G that sends $\overline{x}$ to 0: $z = x - \overline{x}$, so that we get

$$G' = G - \overline{x} = \{z \mid Q'(z) \succeq 0\},$$

where

$$Q'(z) := (Q(\overline{x})) - \hat{Q}(z).$$

So the primal SDP (P) transforms to:

$$\begin{aligned}\sup\quad & c^{\mathrm{T}}z + c^{\mathrm{T}}\overline{x}\\ \text{s.t.}\quad & \sum_i z_i Q_i \preceq Q(\overline{x}).\end{aligned} \tag{P'}$$

Note that $\hat{Q}$ has not changed, and it is not difficult to show that the sets $C_k, \mathcal{U}_k$ and $\mathcal{W}_k$ remain unchanged, as the equations $Q^{\#}(X) = 0$ are equivalent to $Q^*(X) = 0, Q(\overline{x}) \bullet X = 0$. And therefore, the (ELSD) corresponding to (P′) is given by

$$\begin{aligned} \inf \quad & (U+W) \bullet Q(\bar{x}) + c^{\mathrm{T}}\bar{x} \\ \text{s.t.} \quad & Q^*(U+W) = c, \\ & W \in \mathcal{W}_m, \\ & U \succeq 0. \end{aligned} \qquad \text{(ELSD}'\text{)}$$

Since $Q^*(U+W)=c$,

$$\begin{aligned} (U+W) \bullet Q(\bar{x}) + c^{\mathrm{T}}\bar{x} &= (U+W)\bullet Q_0 - \bar{x}^{\mathrm{T}} Q^*(U+W) + c^{\mathrm{T}}\bar{x} \\ &= (U+W) \bullet Q_0, \end{aligned}$$

and thus there is an exact correspondence between (ELSD) and (ELSD′) (and a similar statement holds for (Weak-ELSD)), and we may therefore assume without loss of generality that 0 is feasible for (P). Note that this implies

$$\sup\{c^{\mathrm{T}}x \mid x \in G\} \geqslant 0.$$

Proof of parts (ii) *and* (iii): We have the following for any $k \geqslant m-1$ (explanation given below):

$$\begin{aligned} \sup\{c^{\mathrm{T}}x \mid x \in G\} &= \inf\{t > 0 \mid c^{\mathrm{T}}x \leqslant t \ \forall x \in G\} \\ &= \inf\{t > 0 \mid c^{\mathrm{T}}x/t \leqslant 1 \ \forall x \in G\} \\ &= \inf\{t > 0 \mid c/t \in G^{\circ}\} \\ &= \inf\{t > 0 \mid Q^*(U'+W') = c/t, \ W' \in \mathcal{W}_k, \ U' \succeq 0, \ U' \bullet Q_0 \leqslant 1\} \\ &= \inf\{t > 0 \mid Q^*(U+W) = c, \ W \in \mathcal{W}_k, \ U \succeq 0, \ U \bullet Q_0 \leqslant t\} \\ &= \inf\{U \bullet Q_0 \mid Q^*(U+W) = c, \ W \in \mathcal{W}_k, \ U \succeq 0\} \\ &= \inf\{(U+W) \bullet Q_0 \mid Q^*(U+W) = c, \ W \in \mathcal{W}_k, \ U \succeq 0\}. \end{aligned}$$

Explanation: The first equality is from the fact that the primal optimal value is nonnegative. The second is obvious, and the third follows from the definition of polar. The fourth equality is from Theorem 12, and the fifth is obtained by a change of variables $U = tU'$, $W = tW'$, and (since $\mathcal{W}_k$ is a subspace, $W \in \mathcal{W}_k$). The sixth equality is quite obvious and the last one comes from the fact that $W \bullet Q_0 = 0$ whenever $W \in \mathcal{W}_k$ (of course, under the assumption that $0 \in G$).

Thus, we deduce that

(1) Whenever the primal (P) has a finite optimal value, the duals (ELSD) and (Weak-ELSD) are feasible.

(2) In such a case, the optimal values of (P), (ELSD) and (Weak-ELSD) are equal.

Parts (ii) and (iii) of Theorem 6 follow. Note that the duality gap is zero for the weak version of the dual (Weak-ELSD) itself.

Proof of part (iv): Let us now show that the dual (ELSD) attains its optimum value whenever the latter is finite (it does not follow from the proof below that (Weak-

ELSD) attains its optimum). Let $0 \leqslant \alpha < \infty$ be the common optimal value of the three programs (P), (ELSD) and (Weak-ELSD). We distinguish between two cases.

Case 1: $\alpha > 0$. Clearly $c \neq 0$. Consider

$$0 < \alpha = \inf\{t > 0 \mid c/t \in G^{\circ}\}$$

and hence

$$1/\alpha = \sup\{s \geqslant 0 \mid sc \in G^{\circ}\},$$

and since G° is a closed convex set and $c \neq 0$, the above supremum is attained, say for $s = s^*$. Then clearly,

$$\alpha = 1/s^* = \min\{t > 0 \mid c/t \in G^{\circ}\},$$

and it easily follows that (ELSD) attains the optimal value α in this case.

Case 2: $\alpha = 0$. This means that 0 is an optimal solution to the primal. If $c = 0$, then $(U, W) = 0$ will be a dual feasible solution with an optimal value of 0. So assume that $c \neq 0$.

For the remaining case, we need to show that there exists a dual feasible solution (to (ELSD)) with an optimal value of 0, which is equivalent to the feasibility of the following system:

$$(U + W) \bullet Q_0 = 0, \qquad Q^*(U + W) = c, \qquad U \succeq 0, \qquad W \in \mathcal{W}_m. \tag{14}$$

It is easy to see that the above system is the set of feasible solutions of the (Weak-ELSD) of the following *homogeneous* semidefinite program in $m + 1$ variables:

$$\begin{array}{ll} \sup & c^{\mathrm{T}}x \\ \text{s.t.} & \displaystyle\sum_{i=1}^{m} x_i Q_i \preceq x_0 Q_0. \end{array} \tag{15}$$

There is a cause for confusion here as the matrix map has changed slightly after homogenization. We are using W_m (and implicitly Q^* and $Q^{\#}$) as they are defined for the original $Q(x)$.

Since the above primal problem is feasible (with $x = 0, x_0 = 0$), the infeasibility of the (Weak-ELSD) system (14) implies that the SDP (15) is unbounded (this is precisely the motivation behind introducing the weak version of the dual. The improvement obtained in having $W \in \mathcal{W}_{m-1}$ is absorbed by the above increase in the number of variables from m to $m + 1$). From the homogeneity of (15), we conclude that there exists (x, x_0) such that

$$c^{\mathrm{T}}x > 0, \qquad \sum_{i=1}^{m} x_i Q_i \preceq x_0 Q_0.$$

Let us pick an $r \geqslant 0$ such that $r + x_0 > 0$ and set

$$z = x/(r + x_0).$$

We claim that $Q(z) \succeq 0$:

$$\begin{aligned}(r+x_0)Q(z) &= (r+x_0)\Big(Q_0 - \sum_{i=1}^{m} z_i Q_i\Big)\\ &= (r+x_0)Q_0 - \sum_{i=1}^{m} x_i Q_i\\ &= rQ_0 + \Big(x_0 Q_0 - \sum_{i=1}^{m} x_i Q_i\Big)\\ &\succeq 0.\end{aligned}$$

Therefore z is feasible for (P) and quite clearly $c^{\mathrm{T}} z > 0$, which is a contradiction since $\alpha = 0$. Thus, (14) is feasible, and the dual (ELSD) attains its optimum value of 0.

This concludes the proof of our strong duality result. □

3. Related characterizations

3.1. An exact theorem of the alternative

By applying Theorem 6(ii) (the Weak-ELSD version), we can now trivially derive an exact theorem of the alternative for semidefinite systems.

Let $Q(x) = Q_0 - \sum_{i=1}^{m} x_i Q_i$ be a given matrix map, and suppose that we would like characterize the nonemptyness of $G = \{x \mid Q(x) \succeq 0\}$ (or the feasibility of $Q(x) \succeq 0$). Consider the following homogeneous SDP in $m+1$ variables:

$$\begin{aligned}\sup\quad & x_0\\ \text{s.t.}\quad & \sum_{i=1}^{m} x_i Q_i - x_0 Q_0 \preceq 0.\end{aligned}$$

It is quite clear that G is nonempty if and only if the above supremum is ∞, i.e., unbounded. Therefore, we can apply the Duality Theorem (Theorem 6(ii); for Weak-ELSD) to conclude that G is empty iff there exist U, W such that

$$(U+W) \bullet Q_0 = -1, \qquad Q^*(U+W) = 0, \qquad W \in \mathcal{W}_m, \qquad U \succeq 0.$$

We now state this fact as a theorem of the alternative.

Theorem 19 (Farkas' Lemma for SDP). *Let $Q(x) = Q_0 - \sum_{i=1}^{m} x_i Q_i$ be a given affine matrix map. Then exactly one of the following semidefinite systems is consistent:*

(i) $\quad Q(x) \succeq 0$;

$$\text{(ii)}\quad \begin{aligned} &(U+W_m)\bullet Q_0=-1,\\ &Q^*(U+W_m)=0,\\ &Q^\#(U_i+W_{i-1})=0 \quad \forall i=1,\dots,m,\\ &W_0=0,\\ &U\succeq 0,\\ &\begin{bmatrix} I & W_i^{\mathrm{T}}\\ W_i & U_i\end{bmatrix}\succeq 0 \quad \forall i=1,\dots,m.\end{aligned}$$

3.2. Optimality and attainment

Given a semidefinite program (P) and a feasible vector $\bar{x}$, suppose that we wish to test the optimality of this vector. The strong duality theorem, in particular the dual attainment part, gives us an answer.

Theorem 20 (Optimality condition). *A given feasible vector $\bar{x}$ is optimal to SDP* (P) *if and only there exist U, W such that*

$$(U+W)\bullet Q_0=c^{\mathrm{T}}\bar{x},\qquad Q^*(U+W)=0,\qquad W\in\mathcal{W}_m,\qquad U\succeq 0.$$

Thus the *optimality detection problem* for SDP reduces to semidefinite feasibility. The proof is an obvious application of Theorem 6. On a related note, let us now characterize the condition that a (feasible and bounded) semidefinite program (P) attains its optimum value (as seen in Example 24 of Section 4.2, not all SDPs attain their optima).

Theorem 21 (Primal attainment). *Suppose that* (P) *as well* (ELSD) *are feasible. Then the primal SDP attains the optimal value iff the following system is feasible*:

$$Q(x)\succeq 0,\quad c^{\mathrm{T}}x=(U+W)\bullet Q_0,\quad Q^*(U+W)=0,\quad W\in\mathcal{W}_m,\quad U\succeq 0.$$

4. On the complexity of SDP

4.1. An overview of known complexity results

By applying ellipsoid and interior point methods, one can deduce the following complexity results for SDP. First let

$$Q(x)=Q_0-\sum_{i=1}^{m}x_iQ_i,$$

where Q_i, $i=0,\dots,m$ are given rational symmetric matrices, c is a rational vector, and let

$$G = \{x \mid Q(x) \succeq 0\}$$

be the feasible region of (P). The maximum of the bitlengths of the entries of the Q_i and the components of c will be denoted by L, and define for $\varepsilon > 0$,

$$S(G,\varepsilon) = G + B(0,\varepsilon) \quad \text{and} \quad S(G,-\varepsilon) = \{x \mid B(x,\varepsilon) \subseteq G\}.$$

- If a positive integer R is known a priori such that either $G = \emptyset$ or $G \cap B(0,R) \neq \emptyset$, then there is an algorithm that solves the "weak optimization" problem, i.e., for any rational $\varepsilon > 0$, the algorithm either finds a point $y \in S(G,\varepsilon)$ that satisfies $c^{\mathrm{T}}x \leqslant c^{\mathrm{T}}y + \varepsilon\ \forall x \in S(G,-\varepsilon)$, or asserts that $S(G,-\varepsilon)$ is empty [12]. The complexity of the algorithm is polynomial in n, m, L, and $\log(1/\varepsilon)$.
- There is an algorithm which, given any rational $\varepsilon > 0$ and an x_0 such that $Q(x_0) \succ 0$, computes a rational vector $\overline{x}$ such that $Q(\overline{x}) \succ 0$, and $c^{\mathrm{T}}\overline{x}$ is within an additive factor ε of the optimum value of SDP. The arithmetic complexity of the algorithm is polynomial in $n, m, L, \log(1/\varepsilon), \log(R)$ and the bitlength of x_0, where R is an integer such that the feasible region of the SDP lies inside the ball of radius R around the origin [2,18]. However, it should be mentioned that a polynomial bound has not been established for the bitlengths of the intermediate numbers occurring in these algorithms.
- For any fixed m, there is a polynomial time algorithm (in n, L), that checks whether there exists an x such that $Q(x) \succ 0$, and if so, computes such a vector [21]. For the nonstrict case of $Q(x) \succeq 0$, the feasibility can be verified in polynomial time for the fixed dimensional problem as shown in [17].

In a recent work [8], Freund discusses interior-point algorithms for SDPs in which no regularity (Slater-like) conditions are assumed.

4.2. "Ill-conditioned" SDPs

Progression from "weak optimization" to "strong optimization" is made difficult as there exist SDPs with certain undesirable features as described in the following.

(i) *SDPs with no rational optimal solutions.* For a symmetric matrix A, consider the SDP

$$\max\{t \mid A - tI \succeq 0\},$$

and note that its optimal value equals the least eigenvalue of A, and this is usually irrational even for rational matrices A, for instance, if

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix},$$

then the least eigenvalue is $(3 - \sqrt{5})/2$.

(ii) *Feasible regions with small volume.*

Example 22 (*Khachiyan's example*). Consider the map

$$Q(x) = Q_1(x) \oplus Q_2(x) \oplus \cdots \oplus Q_n(x) \oplus [1 - x_n],$$

where

$$Q_i(x) = \begin{bmatrix} x_0 & 2x_{i-1} \\ 2x_{i-1} & x_i \end{bmatrix},$$

and $\oplus$ denotes the direct sum operation for matrices (Section 2.1).

It is then easily shown that for any x with $Q(x) \succeq 0$, we have $x_i \leqslant 1$, $4^{2^i-1} x_0 \leqslant x_i \; \forall i = 1, \ldots, n$, and $x_i \geqslant 0, i = 0, \ldots, n$. Hence, the volume of G is doubly exponentially small in n. This example was found by L.G. Khachiyan.

(iii) *SDPs with doubly exponential optimal solutions.* Rational optimal solutions exist, but they require exponentially many bits: In the above example $\max\{x_0 \mid Q(x) \succeq 0\} = 4^{1-2^n}$.

(iv) *Feasible region too far from the origin.* In the above example, if we replace $x_n \leqslant 1$ by $x_0 \geqslant 1$, then the feasible region is at a doubly exponential distance from the origin. Another example of this type is the following.

Example 23. Consider the map

$$Q(x) = [x_1 - 2] \oplus Q_1(x) \oplus \cdots \oplus Q_{n-1}(x),$$

where

$$Q_i(x) = \begin{bmatrix} 1 & x_i \\ x_i & x_{i+1} \end{bmatrix}, \quad i = 1, \ldots, n-1.$$

Then

$$G = \{Q(x) \succeq 0\} = \{x \mid x_1 \geqslant 2, \; x_2 \geqslant x_1^2, \; \ldots, \; x_n \geqslant x_{n-1}^2\},$$

and hence for any $x \in G$,

$$x_i \geqslant 2^{2^{i-1}} \quad \forall i,$$

which is doubly exponential and therefore any rational solution in G has exponential bitlength and is at an exponential distance from the origin. Further, there does not exist any rational point satisfying $Q(x) \succeq 0$ whose bitlength is single exponential in n. Hence, any polynomial algorithm that attempts to check whether $\{x \mid Q(x) \succeq 0\} \neq \emptyset$ cannot actually output such an x.

(v) *SDP with unattained optimum.*

Example 24. Consider the map

$$Q(x) = \begin{bmatrix} x_1 & 1 \\ 1 & x_2 \end{bmatrix} \succeq 0,$$

and note that $\inf\{x_1 \mid Q(x) \succeq 0\} = 0$ (this is the infimum of x_1 over the top right lobe of the hyperbola $x_1 x_2 \geqslant 1$), but it is not attained.

The success of ellipsoid and interior point methods in establishing the polynomial complexity for the strong versions of linear and convex quadratic programming relies on the fact that the above idiosyncrasies do not arise for these problems. For this reason, to show that the SDFP is polynomial time solvable, if at all this is the case, it seems to require a completely new approach. This point is elaborated as follows.

As seen, the optimal as well as feasible solutions of SDPs could potentially be exclusively irrational, double exponential in bitlength, or sometimes the optimum may not even be attained, so that one cannot hope to write down explicitly an optimal solution or, for that matter, an arbitrary feasible solution of general SDPs in polynomial time. But, there might be alternate *implicit* polynomial size representations for these solutions. For instance, the optimal solution of the SDP in (iii) above is the *unique* solution to the system

$$x_n = 1, \qquad x_0 x_i = 4x_{i-1}^2, \quad i = 1, \ldots, n,$$

and it takes polynomially many bits to write down this system. Similarly, the optimal value of $\max\{t \mid A - tI \succeq 0\}$ is the smallest root of the characteristic polynomial of A, and clearly it is the *unique* solution to the polynomial system

$$S_k(A - tI) = 0 \quad \forall k = 1, \ldots, n,$$

where $S_k(\bullet)$ denotes the kth elementary symmetric function of a matrix [14]. Now, if A is an integral matrix whose entries have bitlength at most L, then the above polynomial system has size polynomial in n, L (as each of the $S_k(A - tI)$ is the sum of single exponentially many subdeterminants of $A - tI$, which are polynomials of polynomial bitlength).

4.3. Complexity consequences of ELSD

In this section, we will address certain complexity issues concerning Semidefinite Programming. The two computational models of our interest are

The Turing Machine model [9] (abbreviated as *TM Model* here). We will mainly borrow notation in Chapters 2 and 7 of Garey and Johnson [9].

The real number model of Blum, Shub and Smale [5] (abbreviated as *BSS Model*).

The primary decision problem of interest is:

Semidefinite Feasibility Problem (SDFP).

Instance: Positive integers m and n, and $n \times n$ symmetric matrices $Q_0, \ldots, Q_m$ with integer entries.

Question: Are there real numbers such that $x_1, \ldots, x_m$ such that $Q(x) = Q_0 - \sum_{i=1}^{m} x_i Q_i$ is positive semidefinite?

Sizes of instances. In the TM model, the size of an instance of SDFP is $m + n$ plus the total number of bits required to write down the matrices. In the BSS model, the size is $m + n$ (bitlengths are not allowed as a part of the size here).

At present, it is not known whether SDFP is in NP (and of course, it is open whether SDFP is P) for the TM Model. In the BSS model, from the fact that checking for positive semidefiniteness can be accomplished using Gaussian elimination, it follows that SDFP is in NP, and it is shown here that it is also in Co-NP. Let us define a language [9] (or a decision problem) to be *semidefinite reducible*, if there exists a polynomial reduction from that language to SDFP.

Our main result of the section is the following.

Theorem 25. *The following hold concerning* SDFP:

(i) *If* SDFP $\in$ NP, *then* SDFP $\in$ Co $-$ NP *and vice-versa.*

(ii) *In the TM model, unless* NP = Co $-$ NP, SDFP *is neither* NP-*complete nor* Co-NP-*complete.*

(iii) *In the BSS model,* SDFP *is in* NP $\cap$ Co $-$ NP.

Proof. The result is essentially an application of the theorem of alternative, namely Theorem 19.

First note that, the system of alternative given for $Q(x) \succeq 0$, written here in short form:

$$(U + W) \bullet Q_0 = -1, \qquad Q^*(U + W) = 0, \qquad W \in \mathcal{W}_m, \qquad U \succeq 0,$$

can itself be cast in the form "$Q(x) \succeq 0$": Specifically, we can replace the linear constraints $Ax = b$

$$\operatorname{Diag}(Ax - b) \oplus \operatorname{Diag}(b - Ax) \succeq 0,$$

and one can combine two semidefinite inequalities $P(x) \succeq 0, Q(x) \succeq 0$ into one as:

$$P(x) \oplus Q(x) \succeq 0.$$

These two "tricks" are sufficient to show that the system of alternative (ii) of Farkas' Lemma (Theorem 19) can be written as a semidefinite inequality whose size is *polynomial* in the size of the input instance of the SDFP. So, given any instance of SDFP, we write the above alternative system and deduce from the Farkas' Lemma that the former has a "No" answer iff the latter has a "Yes" answer. This directly implies (i).

To prove (ii), suppose that SDFP is NP-complete (in the TM model). Then, in particular, it is in NP, and then (i) implies that SDFP is in NP$\cap$Co-NP. By Theorem 7.2 of [9], the existence of an NP-complete problem in NP $\cap$ Co-NP implies that NP = Co-NP.

The proof of (iii) will follow if we show that SDFP is in NP for the BSS model. If we can check in time polynomial (in m, n), given a "guess vector" x, whether $Q(x) \succeq 0$, then it follows that SDFP is in NP. We can perform this task in $\mathrm{O}(\max\{n^3, mn^2\})$ arithmetic computations as follows: first compute the combination matrix $Q(x)$ in $\mathrm{O}(mn^2)$ steps and then apply partial Cholesky decomposition (see [12] or [21]) to check if this matrix is positive semidefinite in $\mathrm{O}(n^3)$ steps. □

The reason one cannot extend the above proof that SDFP $\in$ NP$\cap$Co-NP for the BSS model to the TM model is because feasible semidefinite inequalities need not always have rational solutions of polynomial size (see examples in Section 4.2). However, part 2 of the theorem is quite interesting as it results in: either SDFP $\in$ NP $\cap$ Co-NP or SDFP $\notin$ NP $\cup$ Co-NP, which may be interpreted as, "either SDFP is too easy or it is too hard".

By an application of the results of last section, we can conclude that there exist polynomial time reductions (in both TM and BSS models) from the following problems to SDFP:

(i) *The Primal Boundedness Problem.* Checking whether a feasible SDP is bounded; By the strong duality theorem, primal boundedness is equivalent to the feasibility of the ELSD.

(ii) *The Primal Attainment Problem.* Verifying whether a feasible and bounded SDP attains its optimum value; this follows from Theorem 21.

(iii) *The SDP Optimality Problem.* Given an SDP and a feasible solution, determine whether this solution is optimal; The reduction is evident from Theorem 20.

5. Concluding remarks

The ELSD dual has $\mathrm{O}(mn^2)$ number of variables. While this number is polynomial, it is rather large from a computational point of view. Hence it is natural to seek refinements of the dual that have fewer variables. It should, however, be mentioned that despite the large number of variables, the dual ELSD has a good deal of structure, which might lend itself to efficient matrix algebraic manipulations, and reduce the effective algebraic complexity of algorithms involving ELSD.

It has been recently shown [22] that the Lagrangian dual (LSD) of ELSD shares the common optimal value of P and ELSD. The relationship between this problem (i.e., the LSD of ELSD dual), denoted by CP, and the original primal P is the following: from P, one can obtain the polynomial size semidefinite program CP, which has the same optimal value as P and whose Lagrangian duality gap is zero. Thus, this process can be thought of as a "correction" of the original primal, and therefore we call the resulting

SDP the *corrected primal* (CP) of P. It appears very likely at this point that certain infeasible interior point algorithms can be built around the pair of programs CP and ELSD, which are Lagrangian duals of each other with Lagrangian duality gap zero.

It should also be noted that the ELSD duality, unlike the Lagrangian duality, is not symmetric, i.e., if one takes the second ELSD dual of (P), we do not recover the latter problem (for instance, we know that the second ESLD dual will attain its optimum value even if (P) does not). It is quite intriguing as to how this second dual relates to the primal.

Perhaps the most outstanding open problem in the theory of Semidefinite Programming is whether SDFP (see Section 4.3) is polynomial time solvable. While Theorem 25(ii) shows that SDFP is not NP-complete (unless NP = Co-NP), the examples of Section 4.2 are reasonably strong reasons to suspect that SDFP may not be in NP to start with; one must verify feasibility without actually computing a feasible solution. But there may exist a representation theory for solutions of semidefinite systems, which offer polynomial time representable and verifiable certificates (see the remarks at the end of Section 4.2).

It is quite common in the SDP literature to liken SDP to Linear Programming, and a good part of the research in the theory of semidefinite programming has been focused on aspects for which the theory known for Linear Programming extended to SDP without too many complications. The duality issues addressed in the current work may be viewed as situations in which direct extensions from LP do not work. In this context, we would like to mention that no satisfactory generalization of the LP Simplex method is yet known for SDP.

Finally, there is some reason to hope that Graph Isomorphism can be reduced to SDFP. This conjecture comes from certain results in [25]. Let A, B be the adjacency matrices of two graphs. Then these graphs are isomorphic iff there exists a permutation matrix P such that $A = PBP^{\mathrm{T}}$, i.e., the following system is feasible

$$A = PBP^{\mathrm{T}}, \qquad Pe = e, \qquad P^{\mathrm{T}}e = e, \qquad P_{ij} \in \{0,1\}, \tag{GI1}$$

where e denotes the vector of all ones. In [25], it was shown that one may relax the last integrality condition, i.e., let P be doubly stochastic instead, without losing exactness, i.e., GI1 is feasible iff the following is feasible

$$A = PBP^{\mathrm{T}}, \qquad Pe = e, \qquad P^{\mathrm{T}}e = e, \qquad 0 \leqslant P \leqslant J, \tag{GI2}$$

where J is the matrix of all ones. It is probably not far-fetched to conjecture that some semidefinite relaxation of GI1 or GI2 is exact. Another seemingly good candidate for reduction to SDFP is the recognition problem for perfect graphs.

Acknowledgements

I have benefited from my discussions with Lóránt Porkoláb, Rob Freund, Gabor Pataki, Tamas Terlaky and Laci Lovász. Special thanks to Rob for Faxing me an instance of SDP for which the Lagrangian dual program has a duality gap, thus inspiring this current work. I would also like to acknowledge many helpful comments and suggestions from two anonymous referees and the *Mathematical Programming* special issue editors Mike Overton and Henry Wolkowicz. Finally, many thanks to Don Hearn and Jim Renegar for their encouragement.

References

[1] F. Alizadeh, Combinatorial optimization with interior point methods and semi-definite matrices, Ph.D. Thesis, Computer Science Department, University of Minnesota, Minneapolis, MN, 1991.

[2] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization, *SIAM J. Optimization* 5 (1) 1995.

[3] J. Borwein and H. Wolkowicz, Regularizing the abstract convex program, *J. Math. Anal. Appl.* 83 (1981).

[4] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Studies in Applied Mathematics, Vol. 15 (SIAM, Philadelphia, PA, 1994).

[5] L. Blum, M. Shub and S. Smale, On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines, *Bull. Amer. Math. Soc. (N.S.)* 21 (1) (1989) 1–46.

[6] A. Brøndsted, *An Introduction to Convex Polytopes* (Springer, New York, 1983).

[7] V. Chvátal, *Linear Programming* (Freeman, New York, 1983).

[8] R. Freund, Complexity of an algorithm for finding an approximate solution of a semi-definite program, with no regularity condition, Working Paper, OR 302-94, 1994.

[9] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness* (Freeman, New York, 1979).

[10] M.X. Goemans and D.P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *J. ACM* 42 (1995) 1115–1145.

[11] M. Grötschel, L. Lovász and A. Schrijver, Polynomial algorithms for perfect graphs, in: C. Berge and V. Chvátal, eds., *Annals of Discrete Mathematics*, Vol. 21 (North-Holland, Amsterdam, 1984).

[12] M. Grötschel, L. Lovász and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization* (Springer, Berlin, 1988).

[13] R.B. Holmes, *Geometric Functional Analysis and its Applications* (Springer, New York, 1975).

[14] R. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, MA, 1985).

[15] F. Jarre, An interior point method for minimizing the maximum eigenvalue of a linear combination of matrices, Report SOL 91-8, Dept. of OR, Stanford University, Stanford, CA, 1991.

[16] D.S. Johnson, C.H. Papadimitriou and M. Yannakakis, How easy is local search?, *J. Computer and System Science* 37 (1988) 79–100.

[17] L. Porkolab and L. Khachiyan, On the complexity of semidefinite programs, RUTCOR Research Report, RRR 40-95, Rutgers University, New Brunswick, NJ-08903; to appear in *J. Global Optimization* (1997).

[18] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Methods for Convex Programming: Theory and Applications* (SIAM, Philadelphia, PA, 1994).

[19] M.L. Overton, Large-scale optimization of eigenvalues, *SIAM J. Optimization* 2 (1992) 88–120.

[20] G. Pataki, Semidefinite and cone programming: Geometry and simplex-type methods, Ph.D. Thesis, GSIA, Carnegie Mellon University, Pittsburgh, under preparation.

[21] M.V. Ramana, An algorithmic analysis of multiquadratic and semidefinite programming problems, Ph.D. Thesis, The Johns Hopkins University, Baltimore, MD, 1993.

[22] M.V. Ramana and R.M. Freund, A gap-free corrected primal for semidefinite programming, in preparation, 1996.

[23] M.V. Ramana and A.J. Goldman, Some geometric results in semidefinite programming, *J. Global Optimization* 7 (1995) 33–50.

[24] M.V. Ramana and P.M. Pardalos, Semidefinite Programming, T. Terlaky, ed., *Interior Point Methods in Mathematical Programming* (Kluwer Academic Publishers, 1996) 369–398.

[25] M.V. Ramana, E.R. Scheinerman and D. Ullman, Fractional isomorphism of graphs, *Discrete Mathematics* 132 (1994) 247–265.

[26] M.V. Ramana, L. Tunçel and H. Wolkowicz, Strong duality in semidefinite programming, *SIAM J. Optimization*, to appear (1997).

[27] T.R. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).

[28] L. Vandenberghe and S. Boyd, Positive-definite programming, in: J.R. Birge and K.G. Murty, eds., *Mathematical Programming: State of the Art 1994* (University of Michigan, 1994).

[29] L. Vandenberghe and S. Boyd, Semidefinite programming, *SIAM Review* 38 (1996) 49–95.

[30] H. Wolkowicz, Some applications of optimization in matrix theory, *Linear Algebra and its Applications* 40 (1981) 101–118.