Bad semidefinite programs: they all look the same

Gábor Pataki*

October 8, 2018

Abstract

Conic linear programs, among them semidefinite programs, often behave pathologically: the optimal values of the primal and dual programs may differ, and may not be attained. We present a novel analysis of these pathological behaviors. We call a conic linear system $\mathcal{A}x \leq_K b$ *badly behaved* if the value of $\sup \{ \langle c, x \rangle | \mathcal{A}x \leq_K b \}$ is finite but the dual program has no solution with the same value for *some* c. We describe simple and intuitive geometric characterizations of badly behaved conic linear systems. Our main motivation is the striking similarity of badly behaved semidefinite systems in the literature; we characterize such systems by certain *excluded matrices,* which are easy to spot in all published examples.

We show how to transform semidefinite systems into a canonical form, which allows us to easily verify whether they are badly behaved. We prove several other structural results about badly behaved semidefinite systems; for example, we show that they are in $\mathcal{NP} \cap$ co-$\mathcal{NP}$ in the real number model of computing. As a byproduct, we prove that all linear maps that act on symmetric matrices can be brought into a canonical form; this canonical form allows us to easily check whether the image of the semidefinite cone under the given linear map is closed.

Key words: conic linear programming; semidefinite programming; duality; closedness of the linear image of a closed convex cone; pathological semidefinite programs

MSC 2010 subject classification: Primary: 90C46, 49N15; secondary: 52A40

OR/MS subject classification: Primary: convexity; secondary: programming-nonlinear-theory

1 Introduction

Many problems in engineering, combinatorial optimization, machine learning, and related fields can be formulated as the primal-dual pair of conic linear programs

$$
\begin{array}{llll}
 & \sup & \langle c, x \rangle & \\
(P_c) & s.t. & \mathcal{A}x \leq_K b & \\
\end{array}
\qquad\qquad
\begin{array}{llll}
\inf & \langle b, y \rangle & & \\
s.t. & y \geq_{K^*} 0 & & (D_c) \\
 & \mathcal{A}^* y = c, & &
\end{array}
$$

*Department of Statistics and Operations Research, University of North Carolina at Chapel Hill

where $\mathcal{A} : X \rightarrow Y$ is a linear map between finite dimensional Euclidean spaces X and Y, $\mathcal{A}^*$ is its adjoint, $K \subseteq Y$ is a closed, convex cone, K^* is its dual cone, and $s \leq_K t$ means $t - s \in K$. Note that the subscript c refers to the objective of the primal problem.

Problems (P_c) and (D_c) generalize linear programs and share some of the duality theory of linear programming. For instance, a pair of feasible solutions always satisfies the weak duality inequality $\langle c, x\rangle \leq \langle b, y\rangle$. However, in conic linear programming pathological phenomena occur: the optimal values of (P_c) and of (D_c) may differ, and they may not be attained.

In particular, semidefinite programs (SDPs) and second order conic programs (SOCPs) — probably the most useful and pervasive conic linear programs — often behave pathologically: for a variety of examples we refer to the textbooks [6, 37, 11, 3, 42] surveys [44, 41, 27] and research papers [34, 1, 43]. Pathological conic LPs are both theoretically interesting and often difficult, or impossible to solve numerically.

These pathologies arise, since the linear image of a closed convex cone is not always closed. For recent studies about when such sets are closed (or not), see e.g., [4, 2, 29]. Three approaches (which we review in detail below) can help to avoid or remedy the pathologies: one can impose a constraint qualification (CQ), such as Slater's condition; one can regularize (P_c)-(D_c) using a facial reduction algorithm [16, 46, 31]; or write an *extended dual* [34, 24], which uses extra variables and constraints. However, such CQs often do not hold, and neither facial reduction algorithms nor extended duals can help solve all pathological instances.

We started this research observing that pathological SDPs in the literature look curiously similar and one of our main goals is to find the root cause of the similarity. We focus on the system underlying (P_c) and call

$$\mathcal{A}x \leq_K b \tag{P}$$

badly behaved if there exists c such that (D_c) either does not attain its value or its value differs from the value of (P_c). We call (P) *well behaved* if it is not badly behaved.

Main contributions of the paper:

(1) In Theorem 1 of Section 2 we characterize when the system (P) is badly or well behaved. At the heart of Theorem 1 is a simple geometric condition that involves the set of feasible directions at $z \in K$, i.e.,

$$\{\, y \mid z + \epsilon y \in K \text{ for some } \epsilon > 0 \,\},$$

and z is chosen as a certain *slack* in (P).

In Theorem 1 we unify two well-known (and seemingly unrelated) conditions for (P) to be well behaved: the first is Slater's condition, and the second requires K to be polyhedral.

Theorem 1 relies on a result on the closedness of the linear image of a closed convex cone from [29] (which we recap in Lemma 1).

(2) In Section 3 we characterize when a semidefinite system

$$\sum_{i=1}^{m} x_i A_i \preceq B \tag{P_{SD}}$$

is badly behaved via certain *excluded matrices*. We assume (with no loss of generality) that a maximum rank positive semidefinite matrix of the form $B - \sum_i x_i A_i$ is

$$Z = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \text{ for some } 0 \leq r \leq n. \tag{1.1}$$

We prove (in Theorem 2) that (P_{SD}) is badly behaved iff there is a matrix V which is a linear combination of the A_i and B of the form

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}, \tag{1.2}$$

where V_{11} is $r \times r$, V_{22} is positive semidefinite, and $\mathcal{R}(V_{12}^T) \not\subseteq \mathcal{R}(V_{22})$. Here $\mathcal{R}()$ stands for rangespace.

The excluded matrices Z and V are easy to spot in all published badly behaved semidefinite systems (we counted about 20 in the above references). The simplest such system is

$$x_1 \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{1.3}$$

where α is any real number: in (1.3) the right hand side serves as Z and the matrix on the left hand side serves as V.

Theorem 3 similarly characterizes well behaved semidefinite systems.

Theorems 2 and 3 follow from Theorem 1, and Lemma 3, which characterizes the set of feasible directions and related sets in the semidefinite cone.

(3) How do we verify that (P_{SD}) is badly or well behaved? In other words, how do we convince a nonexpert reader that an instance of (P_{SD}) is badly or well behaved? Theorems 4 and 5 in Section 4 show how to transform (P_{SD}) into an equivalent standard system, whose bad or good behavior is self-evident. The transformation is surprisingly simple, as it relies mostly on elementary row operations — the same operations that are used in Gaussian elimination. A natural analogy (and our inspiration) is how one transforms an infeasible linear system of equations $Ax = b$ to derive the obviously infeasible equation $\langle 0, x \rangle = 1$.

Here we also prove that i) badly/well behaved semidefinite systems are in $\mathcal{NP} \cap$co-$\mathcal{NP}$ in the real number model of computing ii) for a well behaved semidefinite system we can restrict optimal dual matrices to be block-diagonal, and iii) roughly speaking, we can partition a well behaved system into a strictly feasible part, and a linear part.

As a byproduct, we prove that all linear maps that act on symmetric matrices can be brought into a canonical form; this canonical form allows us to easily check whether the image of the semidefinite cone under the given linear map is closed.

(4) In Section 5 we sketch analogous results for conic linear programs and SDPs in the dual form and prove that all badly behaved semidefinite systems can be reduced, by a sequence of natural operations, to the system (1.3).

(5) Since most examples in the main body of the paper have at most three variables and 3×3 matrices, in Appendix A we give a larger illustrative example with four variables and 4×4 matrices. We prove other technical results in Appendix B.

We illustrate our results by many examples. The only technical proofs in the main body of the paper are those of Theorem 1 and of Lemma 5, and these can be safely skipped at first reading.

Related work A fundamental question in convex analysis is whether the linear image of a closed convex cone is closed. In this paper we rely on Theorem 1.1 from [29], which we summmarize in Lemma 1. This result gives several necessary conditions, and exact characterizations for the class of *nice* cones. We refer to Bauschke and Borwein [4] for the closedness of the continuous image of a closed convex cone; to Auslender [2] for the closedness of the linear image of an arbitrary closed convex set;

and to Waksman and Epelman [47] for another related result. For perturbation results we refer to Borwein and Moors [13, 14]; the latter paper shows that the set of linear maps under which the image of a closed convex cone is *not* closed is small both in terms of measure and category. For a more general problem, whether the intersection of an infinite sequence of nested sets is nonempty, Bertsekas and Tseng [7] gave a sufficient condition. Their characterization is in terms of a certain *retractiveness* property of the set sequence.

We say that (D_c) is a *strong dual of* (P_c) if they have the same value, and (D_c) attains this value, when it is finite. Thus in general (D_c) is not a strong dual of (P_c). Using this terminology, (P) is well behaved exactly if (D_c) is a strong dual of (P_c) for *all* c. We say that (P) satisfies *Slater's condition*, if there is x such that $b - \mathcal{A}x$ is in the relative interior of K; if this condition holds, then (P) is well behaved.

Ramana in [34] proposed a strong dual for SDPs, which uses polynomially many extra variables and constraints. His result implies that semidefinite feasibility is in $\mathcal{NP} \cap$ co-$\mathcal{NP}$ in the real number model of computing. Klep and Schweighofer in [24] constructed a Ramana-type strong dual for SDPs, which, interestingly, is based on ideas from algebraic geometry, rather than from convex analysis.

The facial reduction algorithm of Borwein and Wolkowicz in [16, 15] converts (P) into a system that satisfies Slater's condition, and is hence well behaved. The algorithm relies on a sequence of reduction steps. For more recent, simplified facial reduction algorithms, see Waki and Muramatsu [46] and Pataki [31]. Ramana, Tunçel, and Wolkowicz in [36] proved the correctness of Ramana's dual from the facial reduction algorithm of [16, 15], showing the connection of these two seemingly unrelated concepts. We refer to Ramana and Freund [35] for a proof that the Lagrange dual of Ramana's dual has the same value as the original problem. Generalizations of Ramana's dual are known for conic LPs over *nice* cones [31]; and for conic LPs over homogeneous cones (Pólik and Terlaky [33]).

For a generalization of the concept of strict complementarity (a concept that plays an important role in our work), we refer to Pena and Roshchina [32]. Schurr et al in [40] characterize *universal duality* — when strong duality holds for all right hand sides and objective functions. Tunçel and Wolkowicz in [43] related the lack of strict complementarity in a homogeneous conic linear system to the existence of an objective function with a positive gap.

We finally remark that the technique of *reformulating* equality constrained SDPs (relying mostly on elementary row operations), to easily verify their infeasibility was used recently by Liu and Pataki [25].

1.1 Preliminaries. When is the linear image of a closed convex cone closed?

We now review some basics in convex analysis, relying mainly on references [38, 23, 12, 5]. In Lemma 1 we also give a short and transparent summary of a result on the closedness of the linear image of a closed convex cone from [29].

If x and y are elements of the same Euclidean space, we sometimes write x^*y for $\langle x, y \rangle$. For a set C we denote its linear span, the orthogonal complement of its linear span, its closure, and interior by $\operatorname{lin} C$, $C^\perp$, $\operatorname{cl} C$, and $\operatorname{int} C$, respectively. For a convex set C we denote its relative interior by $\operatorname{ri} C$. For a convex set C and $x \in C$ we define

$$\begin{aligned} \operatorname{dir}(x, C) &= \{ y \,|\, x + \epsilon y \in C \text{ for some } \epsilon > 0 \}, && (1.4) \\ \operatorname{ldir}(x, C) &= \operatorname{dir}(x, C) \cap -\operatorname{dir}(x, C), && (1.5) \\ \tan(x, C) &= \operatorname{cl} \operatorname{dir}(x, C) \cap -\operatorname{cl} \operatorname{dir}(x, C). && (1.6) \end{aligned}$$

Here $\operatorname{dir}(x, C)$ is the set of feasible directions at x in C, and $\tan(x, C)$ is the tangent space at x in C.

A set C is a *cone* if $\lambda x \in C$ holds for all $x \in C,$ and $\lambda \geq 0$. Let C be a closed convex cone. Its dual cone is

$$C^* = \{ y \mid \langle y, x \rangle \geq 0 \ \forall x \in C \}.$$

For $E,$ a convex subset of $C,$ we say that E is a *face* of $C,$ if $x_1, x_2 \in C,$ and $1/2(x_1 + x_2) \in E$ implies that x_1 and x_2 are in E.

For $x \in C$ and $u \in C^*$, we say that u is *strictly complementary to* x if $u \in \mathrm{ri}(C^* \cap x^\perp)$. If C is the semidefinite cone, or the second order cone, then u is strictly complementary to x iff x is strictly complementary to u; in other cones, however, this may not be the case (see a discussion in [28]).

We say that a closed convex cone C is *nice*, if

$$C^* + E^\perp \text{ is closed for all } E \text{ faces of } C.$$

We know that polyhedral, semidefinite, and p-order cones are nice [16, 15, 29]; the intersection of a nice cone with a linear subspace and the linear preimage of a nice cone are nice [18]; hence homogeneous cones are nice, as they are the intersection of a semidefinite cone with a linear subspace (see [17, 22]). In [30] we characterized nice cones, proved that they must be facially exposed and conjectured that all facially exposed cones are nice. However, Roshchina [39] disproved this conjecture.

We denote the rangespace, nullspace, and adjoint operator of a linear operator $\mathcal{M}$ by $\mathcal{R}(\mathcal{M})$, $\mathcal{N}(\mathcal{M})$ and $\mathcal{M}^*$, respectively. We denote by $\mathcal{S}^n$ the set of n by n symmetric matrices, and by $\mathcal{S}^n_+$ the set of $n \times n$ symmetric positive semidefinite (psd) matrices. For symmetric matrices A and B we write $A \preceq B$ $[A \prec B]$ to denote that $B - A$ is positive semidefinite [positive definite], and we write $A \bullet B$ to denote the trace of AB. We have $(\mathcal{S}^n_+)^* = \mathcal{S}^n_+$ with respect to the $\bullet$ inner product.

We will use the fact that for an $H \subseteq \mathcal{S}^n$ affine subspace

$$\mathrm{ri}(H \cap \mathcal{S}^n_+) = \{ X \in \mathcal{S}^n_+ \mid X \text{ is a maximum rank psd matrix in } H \}.$$

For $A, B \in \mathcal{S}^n$ and an invertible matrix T we will use the identity

$$T^T A T \bullet T^{-1} B T^{-T} = A \bullet B. \tag{1.7}$$

For matrices A_1 and $A_2,$ we let

$$A_1 \oplus A_2 = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

and for sets of matrices X_1 and X_2 we define

$$X_1 \oplus X_2 = \{ A_1 \oplus A_2 \mid A_1 \in X_1, \, A_2 \in X_2 \}.$$

For instance, $\mathcal{S}^r_+ \oplus \{0\}$ (where the order of the 0 matrix will be clear from context) is the set of matrices with the upper left $r \times r$ block positive semidefinite and the rest of the components zero.

We write I_r for the identity matrix of order r.

The following question is fundamental in convex analysis: when is the linear image of a closed convex cone closed? We state and illustrate a short version of Theorem 1.1 from [29], which gives easily checkable conditions which are "almost" necessary and sufficient. We will use Lemma 1 later on to prove Theorem 1.

Lemma 1. *Let $\mathcal{M}$ be a linear map, C a closed convex cone, and $w \in \mathrm{ri}(C \cap \mathcal{R}(\mathcal{M}))$. Conditions (1) and (2) below are equivalent to each other, and* necessary *for $\mathcal{M}^*C^*$ to be closed. If C is nice, then they are* necessary and sufficient.

(1) $\mathcal{R}(\mathcal{M}) \cap \big(\mathrm{cl}\,\mathrm{dir}(w, C) \setminus \mathrm{dir}(w, C)\big) = \emptyset.$

(2) There is $w' \in \mathcal{N}(\mathcal{M}^*) \cap C^*$ *strictly complementary to* w*, and*

$$\mathcal{R}(\mathcal{M}) \cap \big(\tan(w, C) \setminus \mathrm{ldir}(w, C)\big) = \emptyset.$$

$\square$

Our first example which illustrates Lemma 1 is very simple:

Example 1. Let $C = C^* = \mathcal{S}^2_+$ and define the map $\mathcal{M} : \mathbb{R}^2 \to \mathcal{S}^2$ as

$$\mathcal{M}(x_1, x_2) \;=\; \begin{pmatrix} x_1 & x_2 \\ x_2 & 0 \end{pmatrix}.$$

Then $\mathcal{M}^* Y \;=\; (y_{11}, 2y_{12})^T$ where $Y \;\in\; \mathcal{S}^2$, and $\mathcal{M}^* C^*$ is not closed: a direct computation shows $\mathcal{M}^* C^* = (\mathbb{R}_{++} \times \mathbb{R}) \cup (0, 0)$, where $\mathbb{R}_{++}$ stands for the set of strictly positive reals.

Lemma 1 also proves that $\mathcal{M}^* C^*$ is not closed: to see how, let

$$w = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; v = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then $w \in \mathrm{ri}(\mathcal{R}(\mathcal{M}) \cap C)$, since it is a maximum rank psd matrix in $\mathcal{R}(M)$. Also, $v \in \mathcal{R}(\mathcal{M}) \cap (\mathrm{cl}\,\mathrm{dir}(w, C) \setminus \mathrm{dir}(w, C))$, since $v \notin \mathrm{dir}(w, C)$ follows from the definition, and $v \in \mathrm{cl}\,\mathrm{dir}(w, C)$ follows, since putting any $\epsilon > 0$ into the $(2, 2)$ position of v makes it a feasible direction. So condition (1) of Lemma 1 is violated, hence $\mathcal{M}^* C^*$ is not closed.

The next, more involved example illustrates the key point of Lemma 1: the image set $\mathcal{M}^* C^*$ usually has much more complicated geometry than C and C^*. Lemma 1 sheds light on the geometry of $\mathcal{M}^* C^*$ via the geometry of the simpler set C.

Example 2. Let $C = C^* = \mathcal{S}^3_+$ and define the map $\mathcal{M} : \mathbb{R}^3 \to \mathcal{S}^3$ as

$$\mathcal{M}(x_1, x_2, x_3) = \begin{pmatrix} x_1 & 2x_2 & x_3 \\ 2x_2 & x_2 + x_3 & 0 \\ x_3 & 0 & 0 \end{pmatrix}.$$

Thus $\mathcal{M}^* Y = (y_{11}, y_{22} + 4y_{12}, y_{22} + 2y_{13})$, where $Y \in \mathcal{S}^3$.

It is a straightforward computation (which we omit) to show

$$\begin{array}{rcl} \mathrm{cl}(\mathcal{M}^* C^*) & = & \{(\alpha, \beta, \gamma) :\; \alpha \geq 0, 4\alpha + \beta \geq 0\}, \\ \mathrm{cl}(\mathcal{M}^* C^*) \setminus \mathcal{M}^* C^* & = & \{(0, \beta, \gamma) :\; \gamma \neq \beta \geq 0\}. \end{array} \tag{1.8}$$

The set $\mathcal{M}^* C^*$ is shown on Figure 1 in blue, and $\mathrm{cl}\,(\mathcal{M}^* C^*) \setminus \mathcal{M}^* C^*$ in green. (Note that the blue diagonal segment on the green facet actually belongs to $\mathcal{M}^* C^*$.)

Lemma 1 easily proves that $\mathcal{M}^* C^*$ is not closed, even without computing the sets in (1.8); indeed, let

$$w \;:=\; \mathcal{M}(6, 1, 0) \;=\; \begin{pmatrix} 6 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; v := \mathcal{M}(0, 0, 1) \;=\; \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

and observe i) $w \in \mathrm{ri}(\mathcal{R}(\mathcal{M}) \cap C)$, since it is a maximum rank psd matrix in $\mathcal{R}(\mathcal{M})$; ii) $v \notin \mathrm{dir}(w, C)$ follows from the definition; and iii) $v \in \mathrm{cl}\,\mathrm{dir}(w, C)$, since putting any $\epsilon > 0$ into the $(3,3)$ position of v makes it a feasible direction.

Thus condition (1) in Lemma 1 is violated, so $\mathcal{M}^* C^*$ is not closed.

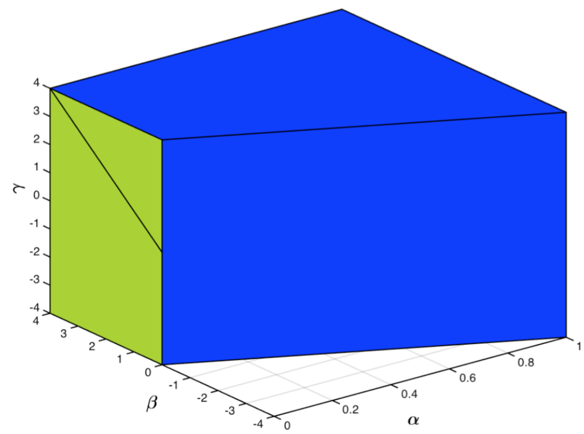

Figure 1: The set $\mathcal{M}^* C^*$ is in blue, and $\mathrm{cl}\,(\mathcal{M}^* C^*) \setminus \mathcal{M}^* C^*$ is in green

We mention in passing that the second part of condition (2) in Lemma 1 is stated in Theorem 1.1 in [29] as $\mathcal{R}(\mathcal{M}) \cap ((E^{\triangle})^{\perp} \setminus \mathrm{lin}\, E) = \emptyset$, where E is the smallest face of C that contains w and $E^{\triangle} = C^* \cap w^{\perp}$. However, this is an equivalent formulation, as implied by the characterization of $\mathrm{lin}\, E$ and $(E^{\triangle})^{\perp}$, see e.g., Lemma 7 in Appendix B.

Throughout the paper we assume that (P) is feasible. Recall that we say that (P) satisfies *Slater's condition* if there exists x such that $b - \mathcal{A}x \in \mathrm{ri}\, K$.

2 When is a conic linear system badly or well behaved?

In this section we present our main characterization of when (P) is badly or well behaved (these concepts are defined in the Introduction). We first need a definition.

Definition 1. *A* slack *in (P) is a vector in*

$$(\mathcal{R}(\mathcal{A}) + b) \cap K,$$

and a maximum slack *is a vector in the relative interior of all slacks.*

We start with a basic lemma:

Lemma 2. *The system (P) is well behaved, if and only if the set*

$$\begin{pmatrix} \mathcal{A}^* & 0 \\ b^* & 1 \end{pmatrix} \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix}$$

is closed.

$\square$

To put Lemma 2 into perspective, note that the image set in Lemma 2 is closed if $(\mathcal{A}, b)^* K^*$ is closed (one can argue this directly or by modifying the proof of Lemma 2). In turn, if $(\mathcal{A}, b)^* K^*$ is closed, then the duality gap between (P_c) and (D_c) is zero, even if K lives in an infinite dimensional space — see, e.g., Theorem 7.2 in [3] (where our primal is called the dual). The proof of Lemma 2 is standard, and we give it in Appendix B.

The main result of this section follows (recall the definition of $\operatorname{dir}(z, K)$ and related sets from (1.4)–(1.6)). We write $\mathcal{R}(\mathcal{A}, b)$ for the rangespace of the operator $(x, t) \to \mathcal{A}x + bt$.

Theorem 1. *Let z be a maximum slack in (P). Conditions (1) and (2) below are equivalent to each other, and* necessary *for (P) to be well behaved. If K is nice, then they are* necessary and sufficient.

(1) $\mathcal{R}(\mathcal{A}, b) \cap (\operatorname{cl} \operatorname{dir}(z, K) \setminus \operatorname{dir}(z, K)) = \emptyset$.

(2) There is $u \in \mathcal{N}\big((\mathcal{A}, b)^*\big) \cap K^*$ *strictly complementary to* z, *and*

$$\mathcal{R}(\mathcal{A}, b) \cap \big(\tan(z, K) \setminus \operatorname{ldir}(z, K)\big) = \emptyset.$$

$\square$

To build intuition we show how Theorem 1 unifies two classical, seemingly unrelated, *sufficient* conditions for (P) to be well-behaved.

Corollary 1. *Suppose that K is a nice cone. If K is polyhedral or (P) satisfies Slater's condition, then (P) is well-behaved.*

Proof Let z be a maximum slack in (P). If K is polyhedral, then so is $\operatorname{dir}(z, K)$. If (P) satisfies Slater's condition, then clearly $z \in \operatorname{ri} K$, so $\operatorname{dir}(z, K) = \operatorname{lin} K$. In both cases $\operatorname{dir}(z, K)$ is closed, hence Condition (1) holds, so (P) is well behaved. $\square$

Though Lemma 2 is a bit simpler to state than Theorem 1, the latter will be more useful. On the one hand, Lemma 2 relies on the closedness of the linear image of $K^* \times \mathbb{R}_+$, which may not be easy to check. On the other hand, Theorem 1 relies on the geometry of the cone K itself, and not on the geometry of its linear image. The geometry of typical cones that occur in optimization — e.g. the geometry of the semidefinite cone — is well understood. Thus Theorem 1, among other things, will lead to a proof that badly behaved semidefinite systems are in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$ in the real number model of computing. Lemma 2, by itself, affords no such corollary.

Note that if $b = 0$, then by Lemma 2 the system (P) is well behaved iff $\mathcal{A}^* K^*$ is closed. Thus in this case Lemma 1 and Theorem 1 are equivalent. To prove the general case of Theorem 1 we use a homogenization argument.

Proof of Theorem 1: We consider the homogenized system

$$\begin{array}{rcl} \mathcal{A}x - bx_0 & \leq_K & 0 \\ -x_0 & \leq & 0, \end{array} \qquad (P_h)$$

and first prove the following claim:

Claim There is a z maximum slack in (P) such that $(z, 1)$ is also a maximum slack in (P_h).

To prove the Claim we first note that if z is a maximum slack in (P) and z' is some other slack, then $\lambda z + (1-\lambda)z'$ is also a maximum slack for all $0 < \lambda \leq 1$ (by Theorem 6.1 in [38]). A similar result holds for (P_h).

Now let z_1 be a maximum slack in (P), then $(z_1, 1)$ is a slack in (P_h). Next, let (z_2, x_0) be a maximum slack in (P_h). By the properties of the relative interior, and since $(z_1, 1)$ is a slack in (P_h), we have that $(z_2, x_0) - \epsilon(z_1, 1)$ is a slack in (P_h) for some $\epsilon > 0$. So $x_0 > 0$ must hold, and (after normalizing) we can assume $x_0 = 1$. Hence z_2 is a slack in (P) and

$$z := \frac{1}{2}(z_1 + z_2)$$

will do. This completes the proof of the claim.

To proceed with the proof of the theorem, we note that the set of maximum slacks in (P) is a relatively open set, so by Theorem 18.2 in [38] it is contained in $\operatorname{ri} F$, where F is some face of K. Therefore $\operatorname{dir}(z, K) = K + \operatorname{lin} F$ for any maximum slack z (see e.g. Lemma 2.7 in [28]) so the sets $\operatorname{dir}(z, K)$ and $\tan(z, K)$ depend only on F. Hence we are free to use any maximum slack of (P) in our proof, and we will use the particular maximum slack provided in the preceding Claim.

For convenience we define the linear map

$$\mathcal{A}_h = \begin{pmatrix} \mathcal{A} & b \\ 0 & 1 \end{pmatrix}$$

which corresponds to the homogenized conic linear system (P_h).

We first note that (trivially)

$$\operatorname{dir}\big((z,1), K \times \mathbb{R}_+\big) \quad = \quad \operatorname{dir}(z, K) \times \mathbb{R} \text{ holds.}$$

Equations (1.4)–(1.6) imply that the same statement holds, if we replace the operator $''$ dir$''$ by $''$ cl dir$''$, “tan$''$, or $''$ ldir$''$.

Hence the following equations hold:

$$\operatorname{cl}\operatorname{dir}\big((z,1), K \times \mathbb{R}_+\big) \setminus \operatorname{dir}\big((z,1), K \times \mathbb{R}_+\big) \quad = \quad \big(\operatorname{cl}\operatorname{dir}(z, K) \setminus \operatorname{dir}(z, K)\big) \times \mathbb{R}, \tag{2.9}$$

$$\tan\big((z,1), K \times \mathbb{R}_+\big) \setminus \operatorname{ldir}\big((z,1), K \times \mathbb{R}_+\big) \quad = \quad \big(\tan(z, K) \setminus \operatorname{ldir}(z, K)\big) \times \mathbb{R}. \tag{2.10}$$

Consider now the following variants of conditions (1) and (2):

(1′) $\mathcal{R}(\mathcal{A}_h) \cap \big[\operatorname{cl}\operatorname{dir}\big((z,1), K \times \mathbb{R}_+\big) \setminus \operatorname{dir}\big((z,1), K \times \mathbb{R}_+\big)\big] = \emptyset$.

(2′) There is $(u, u_0) \in \mathcal{N}(A_h^*) \cap (K \times \mathbb{R}_+)^*$ strictly complementary to $(z, 1)$ and

$$\mathcal{R}(\mathcal{A}_h) \cap \big[\tan\big((z,1), K \times \mathbb{R}_+\big) \setminus \operatorname{ldir}\big((z,1), K \times \mathbb{R}_+\big)\big] = \emptyset.$$

Since $(z, 1)$ is a maximum slack in (P_h), we have $(z,1) \in \operatorname{ri}\big(\mathcal{R}(\mathcal{A}_h) \cap (K \times \mathbb{R}_+)\big)$. Hence by Lemma 1 with $C = K \times \mathbb{R}_+$, $\mathcal{M} = \mathcal{A}_h$, $w = (z, 1)$ we find

$$\mathcal{A}_h^*(K \times \mathbb{R}_+)^* \text{ is closed } \Rightarrow (1') \Leftrightarrow (2')$$

and that equivalence holds when $K \times \mathbb{R}_+$ is nice.

We next note that by (2.9) condition (1′) is equivalent to (1). Also, if (u, u_0) is as specified in (2′), then

$$\langle (u, u_0), (z, 1) \rangle \; = \; \langle u, z \rangle + u_0 \; = \; 0,$$

and since both terms above are nonnegative, we must have $u_0 = 0$. Thus using (2.10) we find that statement ($2'$) is equivalent to condition (2) in Theorem 1. Thus we have

$$\mathcal{A}_h^*(K \times \mathbb{R}_+)^* \text{ is closed } \Rightarrow (1) \Leftrightarrow (2)$$

with equivalence holding when $K \times \mathbb{R}_+$ is nice. Finally, K is nice if and only if $K \times \mathbb{R}_+$ is, thus invoking Lemma 2 completes the proof. $\square$

We can easily modify the proof of Theorem 1 to show that conditions 1 and 2 suffice for (P) to be well behaved, even under a weaker condition than K being nice: it is enough for $K^* + F^\perp$ to be closed, where F is the smallest face of K that contains z. This more general version of Theorem 1 implies that Corollary 1 holds even if we do not assume that K is nice – we refer the interested reader to version 3 of the paper on arxiv.org.

3 When is a semidefinite system badly or well behaved?

We now specialize the results of Section 2, and characterize when the semidefinite system (P_{SD}) is badly or well behaved. To this end, we consider the primal-dual pair of SDPs

$$\begin{array}{llll} & \sup & \sum_{i=1}^m c_i x_i \\ (SDP_c) & s.t. & \sum_{i=1}^m x_i A_i \preceq B \end{array} \qquad \begin{array}{lll} \inf & B \bullet Y & \\ s.t. & Y \succeq 0 & (SDD_c) \\ & A_i \bullet Y = c_i \ (i = 1, \dots, m), & \end{array}$$

where $A_1, \dots, A_m, B \in \mathcal{S}^n$, and $c_1, \dots, c_m$ are scalars.

Specializing Definition 1 to the semidefinite system (P_{SD}), we find that i) a slack in (P_{SD}) is a matrix of the form $S = B - \sum_i x_i A_i \succeq 0$, and ii) a maximum slack in (P_{SD}) is a maximum *rank* slack. We also note that the cone of positive semidefinite matrices is nice [16, 15, 29].

We make the following

Assumption 1. *The maximum rank slack in (P_{SD}) is*

$$Z = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \text{for some } 0 \leq r \leq n. \tag{3.11}$$

We can easily satisfy Assumption 1, at least from a theoretical point of view, as follows. If Z is any maximum rank slack in (P_{SD}), Q is a matrix of suitably scaled eigenvectors of Z, and we apply the rotation $Q^T()Q$ to all A_i and B, then the maximum rank slack in the rotated system is in the required form. (We do not make a claim about actually computing Z or Q; we discuss this point more at the end of Section 4).

In the interest of the reader we first state and illustrate the main results, then prove them.

Theorem 2. *The system (P_{SD}) is badly behaved if and only if there is a matrix V which is a linear combination of the A_i and B of the form*

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}, \tag{3.12}$$

where V_{11} is $r \times r$, $V_{22} \succeq 0$, and $\mathcal{R}(V_{12}^T) \not\subseteq \mathcal{R}(V_{22})$. $\square$

The Z and V matrices provide a *certificate* of the bad behavior of (P_{SD}).

Example 3. In the problem

$$\begin{array}{ll} \sup & x_1 \\ s.t. & x_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{array} \tag{3.13}$$

the only feasible solution is $x_1 = 0$. The dual program, in which we denote the components of Y by y_{ij}, is equivalent to

$$\begin{array}{ll} \inf & y_{11} \\ s.t. & \begin{pmatrix} y_{11} & 1/2 \\ 1/2 & y_{22} \end{pmatrix} \succeq 0, \end{array}$$

which has a 0 infimum but does not attain it.

The certificates of the bad behavior of the system in (3.13) are

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Example 4. The problem

$$\begin{array}{ll} \sup & x_2 \\ s.t. & x_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \tag{3.14}$$

again has an attained 0 supremum. The reader can easily check that the value of the dual program is 1 (and it is attained), so there is a finite, positive duality gap.

In (3.14) the right hand side is the maximum slack, and we can choose the coefficient matrix of x_2 as the V matrix of Theorem 2.

We next characterize well behaved semidefinite systems:

Theorem 3. *The system (P_{SD}) is well behaved if and only if conditions (1) and (2) below hold:*

(1) There is a matrix U of the form

$$U = \begin{pmatrix} 0 & 0 \\ 0 & U_{22} \end{pmatrix}, \tag{3.15}$$

with $U_{22} \in \mathcal{S}^{n-r}$, $U_{22} \succ 0$ and

$$A_1 \bullet U = \ldots = A_m \bullet U = B \bullet U = 0. \tag{3.16}$$

(2) For all V matrices, which are a linear combination of the A_i and B and are of the form

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & 0 \end{pmatrix},$$

with $V_{11} \in \mathcal{S}^r$, we must have $V_{12} = 0$.

□

Example 5. The system

$$x_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.17}$$

is well behaved; we can easily prove this either directly or via Theorem 3. To do the latter, note that the right hand side of (3.17) is the maximum rank slack, condition (1) of Theorem 3 holds with $U = 0 \oplus I_2$, and condition (2) holds vacuously (the $(1,2),(1,3)$ block of both constraint matrices is zero).

Example 6. This example illustrates both badly and well behaved semidefinite systems, depending on the value of the parameter α :

$$x_1 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -3 \\ 1 & -3 & 8 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 & -3 \\ 1 & 0 & 1 \\ -3 & 1 & -6 \end{pmatrix} + x_3 \begin{pmatrix} 1 & 1 & \alpha-3 \\ 1 & 1 & -2 \\ \alpha-3 & -2 & 2 \end{pmatrix} \preceq \begin{pmatrix} 2 & 2 & \alpha-5 \\ 2 & 2 & -4 \\ \alpha-5 & -4 & 4 \end{pmatrix} \tag{3.18}$$

Let us write A_i for the constraint matrices on the left, and B for the right hand side matrix in (3.18). We first observe that $Z = I_1 \oplus 0$ is the maximum rank slack; indeed i) $Z = B - A_1 - A_2 - A_3$, so it is a slack, and ii) the matrix

$$U = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 10 & 3 \\ 0 & 3 & 1 \end{pmatrix} \tag{3.19}$$

satisfies $B \bullet U = A_i \bullet U = 0$ for all i. Hence U is orthogonal to any slack matrix, so the rank of any slack matrix is at most 1.

If $\alpha \neq 1,$ then (3.18) is badly behaved; as proof, observe that

$$V := A_3 - A_2 - A_1 = \begin{pmatrix} 1 & 0 & \alpha-1 \\ 0 & 0 & 0 \\ \alpha-1 & 0 & 0 \end{pmatrix}$$

is a certificate matrix as required by Theorem 2.

If $\alpha = 1,$ then (3.18) is well behaved, and we can verify this using Theorem 3 as follows. The U matrix in (3.19) satisfies condition (1) of Theorem 3. As to condition (2), if the lower right 2×2 block of $V := \sum_{i=1}^{3} \lambda_i A_i + \mu B$ is zero, then $(\lambda_1, \lambda_2, \lambda_3, \mu)$ must be a linear combination of

$$(0,0,2,-1) \text{ and } (5,5,-1,-2),$$

so for all such $(\lambda_1, \lambda_2, \lambda_3, \mu)$ the upper left 1×2 block of V is also zero.

We return to Examples 3–6 in Section 4. As we will see there, the bad or good behavior of semidefinite systems can be verified using only an elementary linear algebraic argument, *without ever referring to Theorems 2 or 3.* We will use Examples 3–6 as illustrations.

The reader may find it interesting to spot the Z and V excluded matrices in other pathological SDPs in the literature, e.g., in the instances in [6, 11, 3, 44, 41, 34, 43, 27].

Theorems 2 and 3 simply follow from Theorem 1 and from Lemma 3 below, which describes the set of feasible directions and related sets in the semidefinite cone:

Lemma 3. *Let Z be as in Assumption 1, and recall the definition of the set of feasible directions, and related sets from (1.4)-(1.6). Then*

$$\begin{aligned}
\operatorname{ldir}(Z, \mathcal{S}^n_+) &= \mathcal{S}^r \oplus \{0\}, && (3.20)\\
\operatorname{cl}\operatorname{dir}(Z, \mathcal{S}^n_+) &= \left\{ \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & Y_{22} \end{pmatrix} \,\Big|\, Y_{22} \in \mathcal{S}^{n-r}_+ \right\}, && (3.21)\\
\tan(Z, \mathcal{S}^n_+) &= \left\{ \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & 0 \end{pmatrix} \,\Big|\, Y_{11} \in \mathcal{S}^r \right\}, && (3.22)\\
\operatorname{dir}(Z, \mathcal{S}^n_+) &= \left\{ \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & Y_{22} \end{pmatrix} \,\Big|\, Y_{22} \in \mathcal{S}^{n-r}_+, \mathcal{R}(Y_{12}^T) \subseteq \mathcal{R}(Y_{22}) \right\}. && (3.23)
\end{aligned}$$

□

The proof of Lemma 3 is given in Appendix B.

Proof of Theorem 2 By condition (1) of Theorem 1 we see that (P_{SD}) is badly behaved, iff there is a matrix $V \in \operatorname{lin}\{A_1, \dots, A_m, B\}$ such that

$$V \in \operatorname{cl}\operatorname{dir}(Z, \mathcal{S}^n_+) \setminus \operatorname{dir}(Z, \mathcal{S}^n_+).$$

Thus our result follows from parts (3.21) and (3.23) in Lemma 3. □

Proof of Theorem 3 We apply Theorem 1 to the system (P_{SD}). We first observe that a matrix $U \succeq 0$ is strictly complementary to Z if and only if

$$U = \begin{pmatrix} 0 & 0 \\ 0 & U_{22} \end{pmatrix}, \text{ with } U_{22} \in \mathcal{S}^{n-r},\, U_{22} \succ 0.$$

Next we note that the first part of condition (2) in Theorem 1 holds iff there is such a U that satisfies (3.16). By (3.20) and (3.22) in Lemma 3 the second part of condition (2) in Theorem 1 holds iff all $V \in \operatorname{lin}\{A_1, \dots, A_m, B\}$ which are of the form

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & 0 \end{pmatrix}$$

satisfy $V_{12} = 0$. This completes the proof. □

To summarize, Theorems 2 and 3 are a "combinatorial version" of Theorem 1.

We note that for semidefinite systems that are strictly feasible, a matrix similar to the V matrix in Theorem 2 can make sure that the optimal primal-dual solution pair fails strict complementarity; see [48].

Although we focus on feasible systems, we obtain natural corollaries about *weakly infeasible* SDPs, a class of pathological infeasible SDPs. To describe the connection, note that the alternative system

$$Y \succeq 0,\ A_i \bullet Y = 0\,(i = 1, \dots, m),\ B \bullet Y = -1 \tag{3.24}$$

gives a natural proof of infeasibility of (P_{SD}): if (3.24) is feasible, then (P_{SD}) is trivially infeasible. However, (P_{SD}) and (3.24) may both be infeasible, in which case we call the semidefinite system (P_{SD}) *weakly infeasible.*

As background on weakly infeasible SDPs, we mention that Waki [45] recently described a method for generating weakly infeasible SDPs based on Lasserre's relaxation for polynomial optimization problems; Klep and Schweighofer [24] analyzed weakly infeasible SDPs using real algebraic geometry techniques; and Lourenco et al [26] proved that any weakly infeasible SDP with order n matrices has a weakly infeasible subsystem with dimension at most $n-1$.

To apply our machinery to weakly infeasible SDPs, we homogenize (P_{SD}) to obtain the system

$$\sum_{i=1}^{m} x_i A_i - x_0 B \preceq 0. \tag{3.25}$$

Assume that the system (3.25) satisfies Assumption 1. First, suppose that (P_{SD}) is weakly infeasible. Then (3.25) is badly behaved, since

$$\sup\{\, x_0 \mid (x, x_0) \,\text{is feasible in}\, (3.25)\} = 0, \tag{3.26}$$

but there is no solution feasible in the dual of (3.26) (such a dual solution would be feasible in (3.24)). Hence by Theorem 2 the excluded matrices Z and V appear in (3.25). In turn, if (3.25) satisfies the conditions of Theorem 3 and hence it is well behaved, then (P_{SD}) cannot be weakly infeasible.

4 Reformulations. Badly behaved semidefinite systems are in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$

4.1 Reformulations

To motivate the discussion of this section, we recall a basic result from the theory of linear equations:

> "The system $Ax = b$ is infeasible if and only if its row echelon form contains the equation $\langle 0, x \rangle = \alpha$, where $\alpha \neq 0$."

Since the "if" direction is trivial, we will — informally — say that the row echelon form is an *easy-to-verify certificate, or witness*, of infeasibility.

In this section we describe analogous results for a very different problem : we show how to transform (P_{SD}) into an equivalent system whose bad or good behavior is trivial to verify. As a corollary we prove that badly (and well) behaved semidefinite systems are in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$ in the real number model of computing. (In this model we can store arbitrary real numbers in unit space and perform arithmetic operations in unit time; see e.g. [9]. We do not claim that badly behaved semidefinite systems are in $\mathcal{P}$, i.e., we do not provide a polynomial time algorithm to decide whether (P_{SD}) is badly behaved. We discuss this point in more detail at the end of this section.)

We first define the type of transformation that we use on (P_{SD}).

Definition 2. We obtain an *elementary reformulation,* or simply a *reformulation*, of (SDP_c) by a sequence of the following operations:

(1) Apply a rotation $T^T()T$ to all A_i and B, where $T = I_r \oplus M$ and M is invertible.

(2) Replace B by $B + \sum_{j=1}^{m} \mu_j A_j$, where $\mu \in \mathbb{R}^m$.

(3) Exchange (A_i, c_i) and (A_j, c_j), where $i \neq j$.

(4) Replace (A_i, c_i) by $(\sum_{j=1}^m \lambda_j A_j, \sum_{j=1}^m \lambda_j c_j),$ where $\lambda \in \mathbb{R}^m$, $\lambda_i \neq 0$.

We obtain an elementary reformulation of the system (P_{SD}) by applying the preceding operations with some c.

Clearly, in all reformulations of (P_{SD}) the maximum rank slack is the same.

Where do these operations come from? Operations (3) and (4) are equivalent to elementary row operations (inherited from Gaussian elimination) done on (SDD_c) :

- Operation (3) exchanges the dual equations $A_i \bullet Y = c_i$ and $A_j \bullet Y = c_j$; and
- Operation (4) replaces the dual equation $A_i \bullet Y = c_i$ by $\sum_{j=1}^m (\lambda_j A_j) \bullet Y = \sum_{j=1}^m \lambda_j c_j$.

Lemma 4. *The system (P_{SD}) is well behaved if and only if its elementary reformulations are.*

Proof Operations (1)-(4) of Definition 2 keep the value of (SDP_c) finite, if it is finite; and infinite, if it is infinite. Suppose now that Y is feasible in (SDD_c) with value, say, α, and we apply operations (1) and (2) with rotation matrix T and vector μ. Then identity (1.7) implies that $T^{-T} Y T$ is feasible in the dual of the reformulated problem with value $\alpha + \sum_{j=1}^m \mu_j c_j$. Operations (3) and (4) preserve the feasibility and objective value of a solution of (SDD_c). Thus if (P_{SD}) is well behaved, so are its reformulations, and this completes the proof of the "Only if" direction. Since (P_{SD}) is a reformulation of its reformulations, the "If" direction follows as well. $\square$

4.2 Reformulating (P_{SD}) to verify maximality of the maximum rank slack

Recall that Z is the maximum rank slack in (P_{SD}) described in Assumption 1. We reformulate (P_{SD}) in two steps. In the first step, given in Lemma 5, we reformulate (P_{SD}) so the resulting system has easy-to-verify witnesses that Z is a maximum rank slack. (The Y_j matrices in Lemma 5 will be the witnesses.)

In Lemma 5 we rely on a facial reduction algorithm (see [16, 15, 46, 31]). It is important that in Lemma 5 we only use rotations, i.e., type (1) operations of Definition 2.

Lemma 5. *The system (P_{SD}) has a reformulation*

$$\sum_{i=1}^m x_i A_i' \preceq B' \qquad (P_{SD}')$$

and there exist symmetric matrices of the form

$$Y_j = \begin{array}{c} \begin{array}{ccc} & \overbrace{}^{r_j} & \overbrace{}^{r_{j-1}+\cdots+r_1} \end{array} \\ \begin{pmatrix} 0 & 0 & \times \\ 0 & I & \times \\ \times & \times & \times \end{pmatrix} \end{array} (j = 1, \ldots, \ell) \qquad (4.27)$$

where $\ell \geq 0, r_1 > 0, \ldots, r_\ell > 0,\ r_1 + \cdots + r_\ell = n - r,$ and

$$Y_j \bullet B' = Y_j \bullet A_i' = 0 \qquad (4.28)$$

holds for all i and j. Here the $\times$ symbols denote blocks with arbitrary elements in the Y_j matrices.

$\square$

If $Z = I$, i.e., (P_{SD}) satisfies Slater's condition, then we just take $B' = B$, $A'_i = A_i$ for all i and $\ell = 0$ in Lemma 5.

To build intuition, we first establish why the Y_j matrices indeed prove that the rank of any slack matrix is at most r. Let S be a slack in (P_{SD}), and $Y_1, \dots, Y_\ell$ as in the statement of Lemma 5. Then $S = B' - \sum_i x_i A'_i$ for some $x \in \mathbb{R}^m$. So $Y_1 \bullet S = 0$ and $S \succeq 0$, hence the last r_1 rows and columns of S are zero; $Y_2 \bullet S = 0$ and $S \succeq 0$ imply that the next r_2 rows and columns of S are zero, and so on. Inductively we find that the last $r_1 + \cdots + r_\ell = n - r$ rows and columns of S are zero, hence S must have rank at most r.

Thus we can prove that Z is a maximum rank slack in (P'_{SD}) (hence also in (P_{SD})) using

(1) a vector $x \in \mathbb{R}^m$ such that $Z = B' - \sum_{i=1}^m x_i A'_i$, and

(2) the Y_j matrices of Lemma 5.

We next illustrate Lemma 5.

Example 7. (Examples 3, 4, 5 and 6 continued) In all these examples it is easy to show why the maximum rank slack is indeed a slack. Also, in Example 3

$$Y_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

is orthogonal to all constraint matrices (using the $\bullet$ inner product), so it proves that the rank of any slack matrix is at most one.

In Example 4 the matrix $Y_1 = 0 \oplus I_1$ proves that the rank of any slack is at most one, and in Example 5 the matrix $Y_1 = 0 \oplus I_2$ proves that the rank of any slack is at most two. (So the first three examples do not even need to be reformulated to have a convenient proof that Z is a maximum rank slack.)

In Example 6 we let

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix},$$

and apply the rotation $T^T()T$ to all matrices to obtain the system

$$x_1 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 & 1 & \alpha \\ 1 & 1 & 1 \\ \alpha & 1 & -1 \end{pmatrix} \preceq \begin{pmatrix} 2 & 2 & \alpha+1 \\ 2 & 2 & 2 \\ \alpha+1 & 2 & -2 \end{pmatrix}. \tag{4.29}$$

Now $Y_1 = 0 \oplus I_2$ is orthogonal to all constraint matrices in (4.29), and this proves that the rank of any slack is at most one, so $Z = I_1 \oplus 0$ is a maximum rank slack.

In Appendix A we give a larger example, in which we need two Y_j matrices to prove that any slack matrix has rank at most 2.

Proof of Lemma 5

To find the reformulation assume that $k \geq 0$, we have a reformulation of the form (P'_{SD}) and matrices $Y_1, \ldots, Y_k$ such that (4.28) holds for all i and for $j = 1, \ldots, k$. At the start $k = 0$ and $B' = B$, $A'_i = A_i$ for all i. For brevity, let $s_k = r_1 + \cdots + r_k$. We claim that

$$s_k \leq n - r \text{ holds.}$$

This indeed follows since if $S \succeq 0$ is a slack in (P'_{SD}) then (using the same argument that we used before) the last s_k rows and columns of S must be zero.

If $s_k = n - r$, we set $\ell = k$, and stop; otherwise, we define the cone $K = \mathcal{S}^n_+ \cap Y_1^\perp \cdots \cap Y_k^\perp$. Clearly, K and its dual cone K^* are of the form

$$K = \left\{ \begin{pmatrix} Y_{11} & 0 \\ 0 & 0 \end{pmatrix} \,|\, Y_{11} \in \mathcal{S}^{n-s_k}_+ \right\}, \; K^* = \left\{ \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & Y_{22} \end{pmatrix} \,|\, Y_{11} \in \mathcal{S}^{n-s_k}_+ \right\}.$$

Next, define the affine subspace

$$H = \operatorname{lin}\{ A'_1, \ldots, A'_m \} + B'.$$

Since Z is also a maximum rank slack in (P'_{SD}), and $r < n - s_k$, we have $H \cap K \neq \emptyset$, $H \cap \operatorname{ri} K = \emptyset$, hence $H^\perp \cap (K^* \setminus K^\perp) \neq \emptyset$ by a classic theorem of the alternative (see e.g. Lemma 1 in [31]).

Let

$$Y_{k+1} \in H^\perp \cap (K^* \setminus K^\perp).$$

Since $Y_{k+1} \bullet Z = 0$, we have

$$Y_{k+1} = \begin{array}{c} \begin{array}{ccc} \overbrace{}^{r} & & \overbrace{}^{s_k} \end{array} \\ \begin{pmatrix} 0 & 0 & \times \\ 0 & Y' & \times \\ \times & \times & \times \end{pmatrix} \end{array}$$

for some $Y' \succeq 0$. (Again, the $\times$ symbols stand for submatrices with arbitrary elements). Let r_{k+1} be the number of positive eigenvalues of Y'; since $Y_{k+1} \notin K^\perp$, we have $r_{k+1} > 0$.

Let Q be an invertible matrix such that $Q^T Y' Q = 0 \oplus I_{r_{k+1}}$, and $T = I_r \oplus Q \oplus I_{s_k}$. We apply the rotation $T^T()T$ to $Y_1, \ldots, Y_{k+1}$, and the rotation $T^{-1}()T^{-T}$ to all A'_i and to B'.

By (1.7) the equation (4.28) holds for all i and for $j = 1, \ldots, k+1$. By the form of T now $Y_1, \ldots, Y_{k+1}$ are in the required shape (see equation (4.27)). We then set $k = k + 1$ and continue.

Clearly, our algorithm terminates in finitely many steps, so the proof is complete. □

4.3 Reformulating (P_{SD}) to verify that it is badly behaved

In Theorem 4 we give the *final* reformulation of (P_{SD}) to prove its bad behavior. We point out that in Theorem 4 the proof of the "if" direction is elementary, thus the reformulated system ($P_{\mathrm{SD,bad}}$) is an easy-to-verify certificate that (P_{SD}) is badly behaved.

Theorem 4. *The system (P_{SD}) is badly behaved if and only if it has a reformulation*

$$\sum_{i=1}^{k} x_i \begin{pmatrix} F_i & 0 \\ 0 & 0 \end{pmatrix} + \sum_{i=k+1}^{m} x_i \begin{pmatrix} F_i & G_i \\ G_i^T & H_i \end{pmatrix} \preceq \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = Z, \qquad (P_{\mathrm{SD,bad}})$$

where

(1) matrix Z is the maximum rank slack, and its maximality can be verified by matrices $Y_1, \ldots, Y_\ell$, as given by Lemma 5.

(2) The matrices

$$\begin{pmatrix} G_i \\ H_i \end{pmatrix} \; (i = k+1, \ldots, m)$$

are linearly independent.

(3) $H_m \succeq 0$.

Proof (If) By Lemma 4 it is enough to prove that $(P_{\mathrm{SD,bad}})$ is badly behaved. Let x be feasible in $(P_{\mathrm{SD,bad}})$ with a corresponding slack S. Note that the last $n - r$ rows and columns of S must be zero, otherwise $\frac{1}{2}(S + Z)$ would be a slack with larger rank than Z. Hence, by condition (2) we must have $x_{k+1} = \ldots = x_m = 0$. Next, let us consider the SDP

$$\sup \{ -x_m \,|\, x \text{ is feasible in } (P_{\mathrm{SD,bad}}) \}, \tag{4.30}$$

which, by the above argument, has optimal value 0. We prove that its dual cannot have a feasible solution with value 0, so suppose that

$$Y = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & Y_{22} \end{pmatrix} \succeq 0$$

is such a solution. By $Y \bullet Z = 0$ we get $Y_{11} = 0$, hence by psdness of Y we deduce $Y_{12} = 0$. Thus

$$\begin{pmatrix} F_m & G_m \\ G_m^T & H_m \end{pmatrix} \bullet Y = H_m \bullet Y_{22} \geq 0,$$

which contradicts the assumption that Y is feasible in the dual of (4.30).

Proof (Only if) We start with the system (P'_{SD}) given by Lemma 5 and further reformulate it. For brevity we denote the constraint matrices on the left hand side by A'_i throughout the process.

We first replace B' by Z in (P'_{SD}). Since the resulting system is still badly behaved, by Theorem 2 there is a matrix of the form

$$V = \lambda_0 Z + \sum_{i=1}^{m} \lambda_i A'_i = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix},$$

with $V_{11} \in \mathcal{S}^r, V_{22} \succeq 0$, and $\mathcal{R}(V_{12}^T) \not\subseteq \mathcal{R}(V_{22})$. By the form of Z we can assume $\lambda_0 = 0$ (otherwise we can replace V by $V - \lambda_0 Z$).

Note that the block of V comprising the last $n - r$ columns must be nonzero. We pick an i such that $\lambda_i \neq 0$, replace A'_i by V, then switch A'_i and A'_m. Next we choose a maximal subset of the A'_i matrices so their blocks comprising the last $n - r$ columns are linearly independent. We let A'_m to be one of these matrices (this can be done, since A'_m is now the V certificate matrix), and permute the A'_i so this special subset becomes $A'_{k+1}, \ldots, A'_m$ for some $k \geq 0$.

We finally add suitable multiples of $A'_{k+1}, \ldots, A'_m$ to $A'_1, \ldots, A'_k$ to zero out the last $n - r$ columns and rows of the latter, and arrive at the required reformulation. $\square$

Example 8. (Examples 3, 4 and 6 continued) The first two of these examples are already in the standard form ($P_{\mathrm{SD,bad}}$). Suppose now $\alpha \neq 1$ in Example 6, i.e., the system (3.18) is badly behaved. Recall that by a rotation we brought (3.18) to the simpler form (4.29). Then in (4.29) we set

$$\begin{aligned} B &:= B - A_1 - A_2 - A_3, \\ A_3 &:= A_3 - A_1 - A_2, \end{aligned}$$

and obtain the system

$$x_1 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 & 0 & \alpha - 1 \\ 0 & 0 & 0 \\ \alpha - 1 & 0 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.31}$$

which is in the standard form ($P_{\mathrm{SD,bad}}$) (with $k = 0$). The objective function $\sup -x_3$ yields a zero optimal value over (4.31) but there is no dual solution with the same value: we can argue this as in the proof of the "if" direction in Theorem 4.

Note that the certificate matrix V of Theorem 2 appears in the system ($P_{\mathrm{SD,bad}}$) as the last matrix on the left hand side.

4.4 Reformulating (P_{SD}) to verify that it is well behaved

We now turn to well behaved semidefinite systems, and in Theorem 5 we show how to reformulate them to easily verify their good behavior. In Theorem 5 we also show block-diagonality of dual optimal solutions. Note that the proof of the "if" direction in Theorem 5 is easy, so the system ($P_{\mathrm{SD,good}}$) is an easy-to-verify certificate of good behavior.

Theorem 5. *The system (P_{SD}) is well behaved if and only if it has a reformulation*

$$\sum_{i=1}^{k} x_i \begin{pmatrix} F_i & 0 \\ 0 & 0 \end{pmatrix} + \sum_{i=k+1}^{m} x_i \begin{pmatrix} F_i & G_i \\ G_i^T & H_i \end{pmatrix} \preceq \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = Z, \tag{$P_{\mathrm{SD,good}}$}$$

where

(1) the matrix Z is the maximum rank slack.

(2) The matrices $H_i \, (i = k+1, \dots, m)$ are linearly independent.

(3) $H_{k+1} \bullet I = \cdots = H_m \bullet I = 0$.

Also, if (P_{SD}) is well behaved, and the value of (SDP_c) is finite, then there is an optimal dual matrix in $\mathcal{S}_+^r \oplus \mathcal{S}_+^{n-r}$.

Proof (If and block-diagonality) Let c be such that

$$v := \sup \{ \sum_{i=1}^{m} c_i x_i \,|\, x \text{ is feasible in } (P_{\mathrm{SD,good}}) \} \tag{4.32}$$

is finite. By the proof of Lemma 4 it suffices to prove that the dual of (4.32) has a block-diagonal solution with value v. An argument like in the proof of Theorem 4 proves that $x_{k+1} = \cdots = x_m = 0$ holds for any x feasible in (4.32), so

$$v = \sup \{ \sum_{i=1}^{k} c_i x_i \mid \sum_{i=1}^{k} x_i F_i \preceq I_r \}. \tag{4.33}$$

Since (4.33) satisfies Slater's condition, there is Y_{11} feasible in its dual with $Y_{11} \bullet I_r = v$.

As the H_i are linearly independent, we can choose $Y_{22} \in \mathcal{S}^{n-r}$ (which is possibly not psd) such that

$$Y := \begin{pmatrix} Y_{11} & 0 \\ 0 & Y_{22} \end{pmatrix}$$

satisfies the equality constraints of the dual of (4.32). We then add a positive multiple of the identity to Y_{22} to make Y psd. Taking condition (3) into account we can see that after this operation Y is feasible in the dual of (4.32) and clearly $Y \bullet Z = v$ holds. The proof is now complete.

Proof (Only if) We again start with the system (P'_{SD}) that Lemma 5 provides; now (P'_{SD}) is well behaved. (We also note that the U matrix of Theorem 3 became the $Y_1 = 0 \oplus I_{n-r}$ matrix of Lemma 5, after we rotated it.) We first replace B' by Z. Next we choose a maximal subset of the A'_i whose lower principal $(n-r) \times (n-r)$ blocks are linearly independent. We permute the A'_i if needed, to make this subset $A'_{k+1}, \ldots, A'_m$ for some $k \geq 0$.

To complete the process we add multiples of $A'_{k+1}, \ldots, A'_m$ to $A'_1, \ldots, A'_k$ to zero out the lower principal $(n-r) \times (n-r)$ block of the latter. By Theorem 3 the upper right $r \times (n-r)$ block of $A'_1, \ldots, A'_k$ and the symmetric counterpart also become zero. This concludes the proof. □

Example 9. (Examples 5 and 6 continued) In Example 5 the system (3.17) is already in the form of $(P_{\mathrm{SD,good}})$.

Suppose now $\alpha = 1$ in Example 6, i.e., (3.18) is well behaved. Recall that we transformed this system into the system (4.31) (in Example 9; note that this can be done independently of the value of α). We then switch the first and third matrices in (4.31) to get

$$x_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.34}$$

in the standard form $(P_{\mathrm{SD,good}})$ (with $k = 1$).

We next discuss some implications of Theorem 5. First, as the proof of the "if" direction shows, we can compute an optimal solution of (4.32) from an optimal solution of the reduced problem (4.33); to do so, we only need to solve a linear system of equations (to find Y_{22}) and do a linesearch (to make Y_{22} psd).

Second, loosely speaking, the system $(P_{\mathrm{SD,good}})$ can be partitioned into a strictly feasible part, and a linear part, which corresponds to variables $x_{k+1}, \ldots, x_m$.

Third, how do we generate a well behaved semidefinite system? Theorem 5 can help us to do this: we can choose matrices Z, H_i, G_i, F_i to obtain a system in the form $(P_{\mathrm{SD,good}})$, then arbitrarily reformulate it, while keeping it well behaved. In fact, according to Theorem 5, we can obtain *any* well behaved semidefinite system in this manner.

In related work, Bomze et al in [10] describe methods to generate pathological conic LP instances from other pathological conic LPs. Their results differ from ours, since they need to start with a pathological conic LP.

We also note that using Lemma 1 the authors in Theorem 3.2 in [19] characterized the situation when the projection of $\mathcal{S}_+^n$ onto some entries is closed; we can view Theorem 5 as a generalization of this result.

4.5 Badly behaved semidefinite systems are in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$. Certificates to verify (non)closedness of the linear image of the semidefinite cone

We now state our main complexity result:

Theorem 6. *Badly (and well) behaved semidefinite systems are in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$ in the real number model of computing.*

Proof We give the following certificates to check the status of (P_{SD}): (1) a reformulation of (P_{SD}) into the form $(P_{\text{SD,bad}})$ or $(P_{\text{SD,good}})$; (2) the Y_j matrices of Lemma 5 to verify that Z is indeed a maximum rank slack; (3) a matrix $T = I_r \oplus M,$ and $\mu \in \mathbb{R}^m,$ which were used to transform (P_{SD}) into $(P_{\text{SD,bad}})$ or $(P_{\text{SD,good}})$.

The verifier first checks that $(P_{\text{SD,bad}})$ or $(P_{\text{SD,good}})$ is indeed a reformulation of (P_{SD}); then verifies the properties of $(P_{\text{SD,bad}})$ or $(P_{\text{SD,good}})$ as given in Theorems 4 or 5; then the proof of the "If" part in Theorems 4 or 5 shows that these systems are well- or badly behaved. $\square$

Assume that we are working with the real number model of computing. We don't claim to have a polynomial time algorithm to decide whether (P_{SD}) is badly behaved; in particular, we don't have a polynomial time algorithm to compute the Z and V excluded matrices of Theorem 2, or one to compute the reformulated systems $(P_{\text{SD,bad}})$ or $(P_{\text{SD,good}})$.

In analogy, if (P_{SD}) is feasible, we can verify this in polynomial time (by plugging in a feasible x). If (P_{SD}) is infeasible, we can also verify this in polynomial time, using one of the infeasibility certificates in [34, 24, 46, 25]. However, we don't know how to decide in polynomial time whether (P_{SD}) is feasible.

Thus feasibility of a semidefinite system is similar to the bad behavior of a feasible system: both properties are in $\mathcal{NP} \cap \text{co-}\mathcal{NP},$ but neither is known to be in $\mathcal{P}$.

To conclude this section, we briefly discuss easy-to-verify certificates for the (non)closedness of the linear image of $\mathcal{S}_+^n$. All linear maps that map from $\mathcal{S}^n$ to $\mathbb{R}^m$ are of the form $\mathcal{A}^* : \mathcal{S}^n \to \mathbb{R}^m,$ where

$$\mathcal{A}(x) = \sum_{i=1}^m x_i A_i,\ \mathcal{A}^*(Y) = (A_1 \bullet Y, \ldots, A_m \bullet Y)^T$$

and $A_i \in \mathcal{S}^n$ for all i. We know that $\mathcal{A}^*(\mathcal{S}_+^n)$ is closed if and only if the homogeneous system

$$\sum_{i=1}^m x_i A_i \preceq 0 \tag{4.35}$$

is well behaved (this is immediate from Lemma 2). Thus reformulating this homogeneous system into the standard forms of $(P_{\text{SD,bad}})$ or $(P_{\text{SD,good}})$ gives easy-to-verify certificates of the closedness or nonclosedness of $\mathcal{A}^*(\mathcal{S}_+^n)$.

To illustrate this point we revisit Examples 1 and 2. The semidefinite system

$$- \mathcal{M}(x) \preceq 0, \tag{4.36}$$

where $\mathcal{M}$ is the linear map defined there, is badly behaved (since the image of the semidefinite cone under $\mathcal{M}^*$ is not closed). We can apply the machinery of this paper to study the system (4.36); e.g., we can find the Z and V excluded matrices of Theorem 2, and reformulate (4.36) into the standard form ($P_{\mathrm{SD,bad}}$). We leave the details to the reader.

5 Concluding remarks

Theorem 2 gives the Z and V excluded matrices to characterize bad behavior of (P_{SD}). We can carry this idea further, and prove the following result:

Corollary 2. *Suppose that in addition to the operations of Definition 2 we allow a sequence of the following operations:*

(1) Delete row i and column i from all matrices, where $i \in \{1, \dots, n\}$.

(2) Delete a constraint matrix.

Then we can bring any badly behaved semidefinite system to the form of

$$x_1 \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{5.37}$$

where α is some real number.

Proof Suppose that (P_{SD}) is badly behaved and let us recall the form of the maximum rank slack in Assumption 1. We first add multiples of the A_i to B to make sure that the right hand side is the maximum rank slack. Next we let V to be a certificate matrix as given by Theorem 2; we can assume that V is the linear combination of the A_i only; we reformulate, so V becomes a constraint matrix.

As we show in Lemma 3, we can apply a rotation $T^T()T$ to V (where $T = I_r \oplus M$ for some invertible M) to bring V to the form

$$V = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{12}^T & I_s & 0 \\ V_{13}^T & 0 & 0 \end{pmatrix}, \tag{5.38}$$

where V_{11} is $r \times r$, $s \geq 0$ and $V_{13} \neq 0$. We apply the rotation $T^T()T$ to all constraint matrices, and after this operation V is of the form specified in (5.38). Suppose now that $v_{ij} \neq 0$, where $1 \leq i \leq r$ and $r+s+1 \leq j \leq n$. We rescale V to make sure that $v_{ij} = 1$ holds, then delete all rows and columns from the constraint matrices whose index is not i nor j, to obtain system (5.37). $\square$

Excluded minor results in graph theory, such as Kuratowski's theorem, show that a graph lacks a certain fundamental property, if and only if it can be reduced to a minimal such graph by a sequence of elementary operations. Corollary 2 resembles such results, since system (5.37) is trivially badly behaved.

We can define the well- or badly behaved nature of conic linear systems in a different form, and characterize such systems. For instance, we call the dual system

$$\mathcal{A}^* y = c,\, y \in K^*, \tag{5.39}$$

well behaved, if for all b dual objective functions the values of (D_c) and of (P_c) agree, and the latter value is attained, when it is finite. System (5.39) can be recast in the primal form

$$\mathcal{B}x \leq_{K^*} y_0, \tag{5.40}$$

where $\mathcal{B}$ and y_0 satisfy $\mathcal{R}(\mathcal{B}) = \mathcal{N}(\mathcal{A}^*)$ and $\mathcal{A}^* y_0 = c$. It is straightforward to show that (5.39) is well behaved, if and only if (5.40) is, and to translate the conditions of Theorem 1 to characterize when (5.39) is well- or badly behaved. We leave the details to the reader.

In the special case of semidefinite systems we can obtain the following result:

Theorem 7. *Suppose that in the system*

$$Y \succeq 0,\, A_i \bullet Y \;=\; c_i \,(i = 1, \dots, m) \tag{5.41}$$

the maximum rank feasible matrix is

$$\bar{Y} \;=\; \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \text{for some}\, r \geq 0.$$

Then (5.41) is badly behaved if and only if there is a matrix V and a real number λ such that

$$A_i \bullet V \;=\; \lambda c_i \;(i = 1, \dots, m),$$

and

$$V \;=\; \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix},$$

where V_{11} is r by r, $V_{22} \succeq 0$, and $\mathcal{R}(V_{12}^T) \not\subseteq \mathcal{R}(V_{22})$. □

We can apply similar arguments to conic linear systems in a subspace form

$$K \cap (L + x_0),$$

to characterize their well- or badly behaved status.

We can also characterize badly behaved second order conic systems similarly as we did it for (P_{SD}) in Theorem 2. This result is in version 2 of the online version of the paper on arxiv.org.

We finally mention a subject for possible future work. The interplay of algebraic geometry and optimization is an active research area: see for instance the recent monograph by Blekherman et al [8], and the paper of Klep and Schweighofer [24]. It would be interesting to see how our certificates of bad and good behavior can be interpreted in the language of algebraic geometry.

A A larger badly behaved semidefinite system

In this appendix we give a larger badly behaved semidefinite system to illustrate the standard form reformulation $(P_{\mathrm{SD,bad}})$. What is nice about this example is that the bad behavior of the original (not reformulated) system is very difficult to verify by an *ad hoc* argument, whereas the bad behavior of the reformulated system is self-evident.

Example 10. Consider the badly behaved semidefinite system

$$x_1 \begin{pmatrix} 4 & 3 & -5 & -3 \\ 3 & -2 & 0 & -2 \\ -5 & 0 & -12 & -8 \\ -3 & -2 & -8 & -4 \end{pmatrix} + x_2 \begin{pmatrix} 14 & 10 & -15 & -9 \\ 10 & -6 & 0 & -6 \\ -15 & 0 & -36 & -24 \\ -9 & -6 & -24 & -12 \end{pmatrix} + x_3 \begin{pmatrix} 8 & 6 & -5 & -3 \\ 6 & -4 & 0 & -2 \\ -5 & 0 & -12 & -8 \\ -3 & -2 & -8 & -4 \end{pmatrix} + x_4 \begin{pmatrix} 20 & 15 & -25 & -13 \\ 15 & -10 & -1 & -9 \\ -25 & -1 & -58 & -38 \\ -13 & -9 & -38 & -18 \end{pmatrix}$$
$$\preceq \begin{pmatrix} 45 & 32 & -55 & -31 \\ 32 & -19 & -1 & -21 \\ -55 & -1 & -130 & -86 \\ -31 & -21 & -86 & -42 \end{pmatrix}. \tag{A.42}$$

We show how to bring (A.42) into the form of ($P_{\mathrm{SD,bad}}$), so let us denote the constraint matrices on the left by $A_i \, (i = 1, \ldots, 4),$ and the right hand side matrix by B. Let

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 1/2 \\ 0 & 0 & 3/2 & -1/2 \end{pmatrix},$$

apply the rotation $T^T()T$ to all A_i and $B,$ then perform the following operations:

$$\begin{array}{rcl} B & := & B - A_1 - 2A_2 + A_3 - A_4, \\ A_4 & := & -5A_1 + A_4, \\ A_3 & := & -2A_1 + A_3, \\ A_2 & := & -3A_1 + A_2, \\ A_1 & := & A_1 - 2A_2 + A_3. \end{array}$$

We obtain the system

$$x_1 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & -1 \\ 2 & 3 & 0 & 2 \\ 1 & -1 & 2 & 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 & 0 & 3 & -1 \\ 0 & 0 & 2 & -1 \\ 3 & 2 & 2 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}$$
$$\preceq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{A.43}$$

In (A.43) the matrices

$$Y_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ and } Y_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \tag{A.44}$$

are orthogonal to all the constraint matrices, thus they prove that the rank of any slack is at most two. So in (A.43) the right hand side is the maximum rank slack.

It is easy to see that (A.43) is badly behaved: following the proof of the "If" implication in Theorem 4, one can see that the objective function $\sup -x_4$ yields a value of 0 over (A.43), but there is no dual solution with the same value.

B Proof of Lemmas 2 and 3

In this section we prove Lemmas 2 and 3.

First we need some definitions and notation. For optimization problems we use the symbol val() to denote their optimal value. For program (D_c) we say that $\{y_i\} \subseteq K^*$ is an *asymptotically feasible (AF)* solution, if $\mathcal{A}^* y_i \to c$, and the *asymptotic value of* (D_c) is

$$\text{aval}(D_c) \quad = \quad \inf\{\lim b^* y_i \,|\, \{y_i\} \text{ is asymptotically feasible in } (D_c)\},$$

where the infimum is taken over those AF solutions for which $\lim b^* y_i$ exists.

We prove Lemma 2 by adapting an argument from [20]. We also rely on the following lemma due to Duffin:

Lemma 6. *(Duffin [21]) Problem* (P_c) *is feasible with* $\text{val}(P_c) < +\infty$, *iff* (D_c) *is asymptotically feasible with* $\text{aval}(D_c) > -\infty$, *and if these equivalent statements hold, then*

$$\text{val}(P_c) = \text{aval}(D_c).$$

$\square$

Proof of Lemma 2 We will use the notation

$$\mathcal{A}_h = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$$

(which is also used in the proof of Theorem 1).

Proof (If) Suppose that $\mathcal{A}_h^*(K \times \mathbb{R}_+)^*$ is closed and let c be an objective vector, such that $c_0 := \text{val}(P_c)$ is finite. Then $\text{aval}(D_c) = c_0$ holds by Lemma 6, so there is $\{y_i\} \subseteq K^*$ s.t. $\mathcal{A}^* y_i \to c$, and $b^* y_i \to c_0$, i.e.,

$$(c, c_0) \in \text{cl}(\mathcal{A}, b)^* K^* \;\subseteq\; \text{cl}\, \mathcal{A}_h^*(K^* \times \mathbb{R}_+) \;=\; \mathcal{A}_h^*(K^* \times \mathbb{R}_+).$$

Hence there is $y \in K^*$, $s \geq 0$ such that $\mathcal{A}^* y = c$, and $b^* y + s = c_0$; by weak duality $b^* y = c_0$ must hold. So y is a feasible solution of (D_c) with value c_0, and this completes the proof.

Proof (Only if) To obtain a contradiction, suppose that $\mathcal{A}_h^*(K \times \mathbb{R}_+)^*$ is not closed; then we will show that (P) is badly behaved. Let us choose c and c_0 such that

$$(c, c_0) \quad \in \quad \text{cl}\, \mathcal{A}_h^*(K^* \times \mathbb{R}_+) \setminus \mathcal{A}_h^*(K^* \times \mathbb{R}_+).$$

By $(c, c_0) \in \text{cl}\, \mathcal{A}_h^*(K^* \times \mathbb{R}_+)$ there is $\{(y_i, s_i)\} \subseteq K^* \times \mathbb{R}_+$ s.t. $\mathcal{A}^* y_i \to c$, and $b^* y_i + s_i \to c_0$. Hence

$$\text{val}(P_c) \;=\; \text{aval}(D_c) \leq c_0,$$

where the equality comes from Lemma 6.

However, $(c, c_0) \notin \mathcal{A}_h^*(K^* \times \mathbb{R}_+)$ shows that no feasible solution of (D_c) can have value $\leq c_0$. Hence either $\text{val}(D_c) > c_0$ (this includes the case $\text{val}(D_c) = +\infty$, i.e., when (D_c) is infeasible), or $\text{val}(D_c)$ is not attained. $\square$

To prove Lemma 3 we need another lemma, which is mostly based on results surveyed in [28].

Lemma 7. *Let C be a closed convex cone, $x \in C$, and E the smallest face of C that contains x. Then*

$$\begin{aligned} \operatorname{dir}(x, C) &= C + \operatorname{lin} E, && \text{(B.45)} \\ \operatorname{ldir}(x, C) &= \operatorname{lin} E, && \text{(B.46)} \\ \operatorname{cl} \operatorname{dir}(x, C) &= (C^* \cap x^\perp)^*, && \text{(B.47)} \\ \tan(x, C) &= (C^* \cap x^\perp)^\perp. && \text{(B.48)} \end{aligned}$$

Proof Statements (B.45) and (B.47) are in Lemma 3.2.1 in [28] (in Lemma 2.7 in the online version). We also proved statement (B.48) there, assuming that C is nice. In fact, it follows from (B.47) and (1.6) in general.

In (B.46) the containment $\supseteq$ is trivial. To see $\subseteq$ let $y \in \operatorname{ldir}(x, C)$, then $x \pm \epsilon y \in C$ for some $\epsilon > 0$. Hence $x \pm \epsilon y \in E$, so $\epsilon y \in \operatorname{lin} E$, and this completes the proof. $\square$

Proof of Lemma 3 Let F be the smallest face of $\mathcal{S}_+^n$ that contains Z. Then clearly $F = \mathcal{S}_+^r \oplus \{0\}$, and $\mathcal{S}_+^n \cap Z^\perp = \{0\} \oplus \mathcal{S}_+^{n-r}$. Hence statements (3.20)-(3.22) follow by taking $C = \mathcal{S}_+^n$, $x = Z$, $E = F$ in Lemma 7.

Next, fix $Y \in \operatorname{cl} \operatorname{dir}(Z, \mathcal{S}_+^n)$, and partition it as in the right hand side set in (3.21). Then (3.23) is equivalent to

$$Y \in \operatorname{dir}(Z, \mathcal{S}_+^n) \Leftrightarrow \mathcal{R}(Y_{12}^T) \subseteq \mathcal{R}(Y_{22}). \tag{B.49}$$

Let P be an orthogonal matrix, such that $P^T Y_{22} P = I_s \oplus 0$, where s is the number of positive eigenvalues of Y_{22} and $T = I_r \oplus P$.

Define

$$V := T^T Y T = \begin{pmatrix} Y_{11} & Y_{12}P \\ P^T Y_{12}^T & P^T Y_{22} P \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12}P \\ P^T Y_{12}^T & I_s \oplus 0 \end{pmatrix}.$$

Next we claim

$$\begin{aligned} Y \in \operatorname{dir}(Z, \mathcal{S}_+^n) &\Leftrightarrow V \in \operatorname{dir}(Z, \mathcal{S}_+^n), && \text{(B.50)} \\ \mathcal{R}(Y_{12}^T) \subseteq \mathcal{R}(Y_{22}) &\Leftrightarrow \mathcal{R}(P^T Y_{12}^T) \subseteq \mathcal{R}(P^T Y_{22} P). && \text{(B.51)} \end{aligned}$$

Indeed, (B.50) follows from $T^T Z T = Z$, and the definition of feasible directions. As to (B.51), the left hand side statement holds, iff there is a matrix D with

$$Y_{12}^T = Y_{22} D, \tag{B.52}$$

and the right hand side statement holds, iff there is a matrix D' such that

$$P^T Y_{12}^T = P^T Y_{22} P D'. \tag{B.53}$$

If D satisfies (B.52), then $D' := P^{-1} D$ satisfies (B.53). Conversely, if (B.53) holds for D', then $D := PD'$ verifies (B.52).

Next, partition $Y_{12}P$ as (V_{12}, V_{13}), so that V_{12} has s columns; then (B.51) is equivalent to $V_{13} = 0$. So we only need to prove

$$V \in \operatorname{dir}(Z, \mathcal{S}_+^n) \Leftrightarrow V_{13} = 0. \tag{B.54}$$

Consider the matrix $Z + \epsilon V$ for some $\epsilon > 0$. If $V_{13} \neq 0$, then $Z + \epsilon V$ is not positive semidefinite for any $\epsilon > 0$, and this proves the direction $\Rightarrow$. As to $\Leftarrow$, if $V_{13} = 0$, then by the Schur-complement condition for positive semidefiniteness we have that $Z + \epsilon V \succeq 0$ iff

$$(I_r + \epsilon V_{11}) - (\epsilon V_{12})(\epsilon I_s)^{-1}(\epsilon V_{12}^T) \succeq 0,$$

and the latter is clearly true for some small $\epsilon > 0$. $\square$

Acknowledgement My sincere thanks are due to the anonymous referees, the Associate Editor, and Shu Lu for their careful reading of the manuscript, and thoughtful comments. I also thank Minghui Liu for helpful comments, and for his help in proving Theorem 5. My thanks are also due to Asen Dontchev for his support while writing this paper.

References

[1] Erling D. Andersen, Cees Roos, and Tamás Terlaky. Notes on duality in second order and p-order cone optimization. *Optimization*, 51(4):627–643, 2002.

[2] Alfred Auslender. Closedness criteria for the image of a closed set by a linear operator. *Numer. Funct. Anal. Optim.*, 17:503–515, 1996.

[3] Alexander Barvinok. *A Course in Convexity.* Graduate Studies in Mathematics. AMS, 2002.

[4] Heinz Bauschke and Jonathan M. Borwein. Conical open mapping theorems and regularity. In *Proceedings of the Centre for Mathematics and its Applications 36*, pages 1–10. Australian National University, 1999.

[5] Heinz Bauschke and Patrick Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* bb. Springer, 2011.

[6] Aharon Ben-Tal and Arkadii Nemirovskii. *Lectures on modern convex optimization.* MPS/SIAM Series on Optimization. SIAM, Philadelphia, PA, 2001.

[7] Dimitri Bertsekas and Paul Tseng. Set intersection theorems and existence of optimal solutions. *Math. Program.*, 110:287–314, 2007.

[8] Grigoriy Blekherman, Pablo Parrilo, and Rekha Thomas, editors. *Semidefinite Optimization and Convex Algebraic Geometry.* MOS/SIAM Series in Optimization. SIAM, 2012.

[9] Lenore Blum, Felipe Cucker, Michael Shub, and Stephen Smale. *Complexity and Real Computation.* Springer, 1998.

[10] Immanuel Bomze, W. Schachinger, and G. Uchida. Think co(mpletely)positive ! matrix properties, examples and a clustered bibliography on copositive optimization. *J. Global Opt.*, (52), 2012.

[11] Frédéric J. Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems.* Springer Series in Operations Research. Springer-Verlag, 2000.

[12] Jonathan M. Borwein and Adrian S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples.* CMS Books in Mathematics. Springer, 2000.

[13] Jonathan M. Borwein and Warren B. Moors. Stability of closedness of convex cones under linear mappings. *J. Convex Anal.*, 16(3–4), 2009.

[14] Jonathan M. Borwein and Warren B. Moors. Stability of closedness of convex cones under linear mappings II. *J. Nonlin. Anal. Theory and Appl.*, 1(1), 2010.

[15] Jonathan M. Borwein and Henry Wolkowicz. Facial reduction for a cone-convex programming problem. *J. Aust. Math. Soc.*, 30:369–380, 1981.

[16] Jonathan M. Borwein and Henry Wolkowicz. Regularizing the abstract convex program. *J. Math. Anal. App.*, 83:495–530, 1981.

[17] Check-Beng Chua. Relating homogeneous cones and positive definite cones via T-algebras. *SIAM J. Optim.*, 14:500–506, 2003.

[18] Check-Beng Chua and Levent Tunçel. Invariance and efficiency of convex representations. *Math. Program. B*, 111:113–140, 2008.

[19] Dimitry Drusviyatsky, Gábor Pataki, and Henry Wolkowicz. Coordinate shadows of semi-definite and euclidean distance matrices. *SIAM J. Opt.*, 25(2):1160–1178, 2015.

[20] Richard Duffin, Robert Jeroslow, and Les A. Karlovitz. Duality in semi-infinite linear programming. In *Semi-infinite programming and applications (Austin, Tex., 1981)*, volume 215 of *Lecture Notes in Econom. and Math. Systems*, pages 50–62. Springer, Berlin, 1983.

[21] Richard J. Duffin. Infinite programs. In A.W. Tucker, editor, *Linear inequalities and Related Systems*, pages 157–170. Princeton University Press, 1956.

[22] Leonid Faybusovich. On Nesterov's approach to semi-definite programming. *Acta Appl. Math.*, 74:195–215, 2002.

[23] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag, 1993.

[24] Igor Klep and Markus Schweighofer. An exact duality theory for semidefinite programming based on sums of squares. *Math. Oper. Res.*, 38(3):569–590, 2013.

[25] Minghui Liu and Gábor Pataki. Exact duality in semidefinite programming based on elementary reformulations. *SIAM J. Opt.*, 25(3):1441–1454, 2015.

[26] Bruno Lourenco, Masakazu Muramatsu, and Takashi Tsuchiya. A structural geometrical analysis of weakly infeasible SDPs. *Journal of the Operations Research Society of Japan*, 59(3):241–257, 2015.

[27] Zhi-Quan Luo, Jos Sturm, and Shuzhong Zhang. Duality results for conic convex programming. Technical Report Report 9719/A, Erasmus University Rotterdam, Econometric Institute, The Netherlands, 1997.

[28] Gábor Pataki. The geometry of semidefinite programming. In Romesh Saigal, Lieven Vandenberghe, and Henry Wolkowicz, editors, *Handbook of semidefinite programming*. Kluwer Academic Publishers, 2000.

[29] Gábor Pataki. On the closedness of the linear image of a closed convex cone. *Math. Oper. Res.*, 32(2):395–412, 2007.

[30] Gábor Pataki. On the connection of facially exposed and nice cones. *J. Math. Anal. App.*, 400:211–221, 2013.

[31] Gábor Pataki. Strong duality in conic linear programming: facial reduction and extended duals. In David Bailey, Heinz H. Bauschke, Frank Garvan, Michel Théra, Jon D. Vanderwerff, and Henry Wolkowicz, editors, *Proceedings of Jonfest: a conference in honour of the 60th birthday of Jon Borwein*. Springer, also available from http://arxiv.org/abs/1301.7717, 2013.

[32] Javier Pena and Vera Roshchina. A complementarity partition theorem for multifold conic systems. *Math. Program.*, 142:579–589, 2013.

[33] Imre Pólik and Tamás Terlaky. Exact duality for optimization over symmetric cones. Technical report, Lehigh University, Betlehem, PA, USA, 2009.

[34] Motakuri V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Math. Program. Ser. B*, 77:129–162, 1997.

[35] Motakuri V. Ramana and Robert Freund. On the ELSD duality theory for SDP. Technical report, MIT, 1996.

[36] Motakuri V. Ramana, Levent Tunçel, and Henry Wolkowicz. Strong duality for semidefinite programming. *SIAM J. Opt.*, 7(3):641–662, 1997.

[37] James Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization.* MPS-SIAM Series on Optimization. SIAM, Philadelphia, USA, 2001.

[38] Tyrrel R. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, NJ, USA, 1970.

[39] Vera Roshchina. Facially exposed cones are not nice in general. *SIAM J. Opt.*, 24:257–268, 2014.

[40] Simon P. Schurr, André L. Tits, and Dianne P. O'Leary. Universal duality in conic convex optimization. *Math. Program. Ser. A*, 109:69–88, 2007.

[41] Michael J. Todd. Semidefinite optimization. *Acta Numer.*, 10:515–560, 2001.

[42] Levent Tunçel. *Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization.* Fields Institute Monographs, 2011.

[43] Levent Tunçel and Henry Wolkowicz. Strong duality and minimal representations for cone optimization. *Comput. Optim. Appl.*, 53:619–648, 2012.

[44] Lieven Vandenberghe and Steven Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[45] Hayato Waki. How to generate weakly infeasible semidefinite programs via Lasserre's relaxations for polynomial optimization. *Optim. Lett.*, 6(8):1883–1896, 2012.

[46] Hayato Waki and Masakazu Muramatsu. Facial reduction algorithms for conic optimization problems. *J. Optim. Theory Appl.*, 158(1):188–215, 2013.

[47] Z. Waksman and M. Epelman. On point classification in convex sets. *Math. Scand.*, 38:83–96, 1976.

[48] Hua Wei and Henry Wolkowicz. Generating and measuring instances of hard semidefinite programs. *Math. Program.*, 125(1):31–45, 2010.