Universal Duality in Conic Convex Optimization*

Simon P. Schurr† André L. Tits‡ Dianne P. O'Leary§

Revised: April 18, 2005

Abstract

Given a primal-dual pair of linear programs, it is well known that if their optimal values are viewed as lying on the extended real line, then the duality gap is zero, unless both problems are infeasible, in which case the optimal values are $+\infty$ and $-\infty$. In contrast, for optimization problems over nonpolyhedral convex cones, a nonzero duality gap can exist when either the primal or dual is feasible.

For a pair of dual conic convex programs, we provide simple conditions on the "constraint matrices" and cone under which the duality gap is zero for *every* choice of linear objective function and constraint right-hand side. We refer to this property as "universal duality". Our conditions possess the following properties: (i) they are necessary and sufficient, in the sense that if (and only if) they do not hold, the duality gap is nonzero for some linear objective function and constraint right-hand side; (ii) they are metrically and topologically generic; and (iii) they can be verified by solving a single conic convex program. We relate to universal duality the fact that the feasible sets of a primal convex program and its dual cannot both be bounded, unless they are both empty. Finally we illustrate our theory on a class of semidefinite programs that appear in control theory applications.

Keywords. Conic convex optimization, constraint qualification, duality gap, universal duality, generic property.

1 Introduction and background

It is well known that if a linear program and its Lagrangian dual are both feasible, then strong duality holds for that pair of problems. That is, there is a zero duality gap, and both (finite) optimal values are attained. A key to proving this result is Farkas' Lemma. It is also well known that for nonpolyhedral convex cones, simple generalizations of Farkas' Lemma

*This work was supported by a fellowship at the University of Maryland, in addition to NSF grants DEMO-9813057, DMI0422931, CUR0204084, and DoE grant DEFG0204ER25655.

†Applied Mathematics Program, University of Maryland, College Park, MD 20742; email: spschurr@math.umd.edu .

‡Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742; email: andre@umd.edu . Correspondence should be addressed to this author.

§Computer Science Department and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742; email: oleary@cs.umd.edu .

do not necessarily hold. (However an "asymptotic" Farkas Lemma does hold; see e.g., [12, Theorem 3.2.3].) In fact there exist conic programs[1] that admit a nonzero, and possibly finite, duality gap when either the primal or dual is feasible; see [8, Section 6.1], [12, Section 3.2], or [19, Section 4] for examples. The reason for the failure of simple extensions of Farkas' Lemma to nonpolyhedral convex cones is the potential nonclosedness of the linear image of a closed convex cone. (When the cone is polyhedral, its linear image is always closed.) Other conditions under which closedness is guaranteed to occur can be found, e.g., in [10], and the references therein.

As a consequence of the above-mentioned failure, in optimization over nonpolyhedral convex cones, a regularity condition is often assumed in order to guarantee a zero duality gap. An example of such a condition is the generalized Slater constraint qualification (GSCQ). A sufficient condition for strong duality of a pair of dual conic programs is that both problems satisfy the GSCQ. If the GSCQ holds for only one of the two problems, then a zero duality gap still results, but the optimal values need not both be attained. Further results on duality in linear and nonlinear programming can be found in, e.g., [9, Chapter 8].

In some contexts, one wishes to study a family of optimization problems parameterized by their objective function or constraint right-hand side. For example, in a network optimization problem, it may be the case that the network "structure" remains fixed, but say, the arc costs or arc capacities vary. Under such circumstances, it would be desirable for the "constraint matrices" (corresponding to the network "structure") to be such that the duality gap of the associated optimization problem and its dual to always be zero, regardless of the objective function or constraint right-hand side (which may correspond to arc costs or arc capacities).

In this work, motivated by such considerations, we give necessary and sufficient conditions on the "constraint matrices" and cone that ensure, for every linear objective function and constraint right-hand side, a zero duality gap holds for a conic program and its dual. We refer to this property as "universal duality". We explain how universal duality essentially implies stability of families of optimization problems parameterized by their objective function and constraint right-hand side.

Genericness of certain types of nondegeneracy of conic programs is a useful property, for both theoretical and numerical reasons. It was shown in [2] that primal and dual nondegeneracy, and strict complementarity, holds for Lebesgue almost all semidefinite programs. These results were extended in [11] to the more general case of conic programs in "standard form". As a further contribution of the present paper, we show that universal duality holds generically in a metric as well as a topological sense.

Finally, we show that universal duality—which gives duality information about an infinite family of conic programs—can be verified by solving a single conic program with essentially the same size and structure as that of the original "primal" problem.

The layout of this paper is as follows. Section 2 is devoted to notation and preliminaries. In Section 3, we formally define universal duality and prove simple necessary and sufficient conditions for it to hold. We also use these conditions to show how universal duality relates to the boundedness or lack thereof of the feasible sets of a pair of dual conic programs. We show in Section 4 that universal duality is a metrically generic and topologically generic property, and

[1] Throughout the paper, we shall refer to conic convex programs as simply conic programs.

in Section 5 that universal duality for a pair of dual conic programs can be verified by solving a single conic program. In Section 6 we apply our theory of universal duality to a semidefinite program found in control theory. Finally, in Section 7 we state some conclusions.

2 Notation and preliminaries

Given a set $S \subseteq \mathcal{R}^n$, we will write $\mathrm{ri}(S)$, $\mathrm{int}(S)$, and $\mathrm{cl}(S)$, to denote its relative interior, interior, and closure, respectively. We endow $\mathcal{R}^n$ with the inner product $\langle\cdot,\cdot\rangle$, which induces a vector norm and a corresponding operator norm, both denoted by $\|\cdot\|$. The dual of S is given by $S^* = \{x \in \mathcal{R}^n \mid \langle x, y\rangle \geq 0\ \forall\, y \in S\}$ and the orthogonal complement of S is given by $S^\perp = \{x \in \mathcal{R}^n \mid \langle x, y\rangle = 0\ \forall\, y \in S\}$. The adjoint of a linear operator A is denoted by A^*. We denote the space of linear operators from $\mathcal{R}^n$ to $\mathcal{R}^m$ by $\mathcal{R}^{m\times n}$, and denote by I_n the identity operator from $\mathcal{R}^n$ to $\mathcal{R}^n$, and the identity matrix of order n. (When the domain or range of the identity operator are clear, we may omit the subscript.)

A set $K \subseteq \mathcal{R}^n$ is said to be a *convex cone* if for all $x_1, x_2 \in K$ and $t_1, t_2 \geq 0$, we have $t_1 x_1 + t_2 x_2 \in K$. The dual of any set is a closed convex cone. A cone whose interior is nonempty is said to be *solid*. If K contains no lines, i.e., its lineality space $K \cap -K$ is the origin, then K is said to be *pointed*. A cone that is closed, convex, solid, and pointed, is said to be *full*.[2] A convex cone K induces an quasi-ordering $\succeq_K$, where $x \succeq_K y$ is defined by $x - y \in K$. (If K is also pointed, then $\succeq_K$ is a partial ordering.) If K is also solid, we write $x \succ_K y$ to mean $x - y \in \mathrm{int}(K)$, We will use the standard convention that the infimum (supremum) of an empty set is $+\infty$ $(-\infty)$, and the infimum (supremum) of a subset of the real line unbounded from below (above) is $-\infty$ $(+\infty)$. The nullspace and range of a finite dimensional linear operator A will be denoted by $\mathcal{N}(A)$ and $\mathrm{Range}(A)$ respectively. We will write $A(S) = \{Ax \mid x \in S\}$ to denote the image of a set S under a linear operator A.

We will use the following theorem of the alternative contained in [10, Theorem 4]. It is a generalization of Stiemke's theorem of the alternative for linear equalities and inequalities; see e.g., [17, p. 95].

Lemma 2.1. *Let A be a linear operator and K be a closed convex cone. The following two statements are equivalent.*
(a) *There exists a solution x to the system $Ax = 0,\ x \in \mathrm{ri}(K)$.*
(b) $A^*y \in K^* \ \Rightarrow\ A^*y \in K^\perp$.

3 Universal duality in conic optimization

Any primal-dual pair of convex programs can be expressed in the "standard form"

$$u_P = \inf_x \{\langle f, x\rangle \mid Ax = b,\ x \succeq_K 0\}, \tag{1}$$

$$u_D = \sup_{y,w} \{\langle b, y\rangle \mid A^*y + w = f,\ w \succeq_{K^*} 0\}, \tag{2}$$

[2]Some authors call such a cone *proper* or *regular*.

or in the more general form

$$v_P = \inf_x \{\langle f,x\rangle \mid Ax=b,\ Cx \succeq_K d\}, \tag{3}$$

$$v_D = \sup_{y,w} \{\langle b,y\rangle + \langle d,w\rangle \mid A^*y + C^*w = f,\ w \succeq_{K^*} 0\}, \tag{4}$$

where $A : \mathcal{R}^n \to \mathcal{R}^m$ and $C : \mathcal{R}^n \to \mathcal{R}^p$ are linear operators, and $K \subset \mathcal{R}^p$ is a full cone.[3] The primal formulation (3) can be found in, e.g., [3, Section 4.6.1], and for the case where K is the positive semidefinite cone, in [21, Sections 3.1,4.2]. As is the case for all primal-dual pairs of convex (and even nonconvex) programs, weak duality holds for (1)–(2) and (3)–(4), viz., $u_P \geq u_D$ and $v_P \geq v_D$. The feasible sets of (1) and (2) will be denoted by

$$\begin{aligned} \mathcal{F}_P &= \{x \mid Ax=b,\ x \succeq_K 0\}, \\ \mathcal{F}_D &= \{(y,w) \mid A^*y + w = f,\ w \succeq_{K^*} 0\}. \end{aligned}$$

Unless otherwise stated, the following assumption will be in effect throughout.

Assumption 3.1. *The equality constraints $Ax=b$ in (1) and (3), and the "inequality" constraints $Cx \succeq_K d$ in (3) are nonvacuous, i.e., $m,p>0$. (Of course, it is assumed also that $n>0$.)*

In Remarks 3.11 and 5.5, we consider the cases where $m=0$ or $p=0$.

The problem (1) (resp. (3)) is said to be *strongly feasible*[4] if $\{x \mid Ax=b,\ x \succ_K 0\}$ (resp. $\{x \mid Ax=b,\ Cx \succ_K d\}$) is nonempty. Its dual (2) (resp. (4)) is said to be *strongly feasible* if $\{(y,w) \mid A^*y+w=f,\ w \succ_{K^*} 0\}$ (resp. $\{(y,w) \mid A^*y + C^*w = f,\ w \succ_{K^*} 0\}$) is nonempty. Strong feasibility is equivalent to the GSCQ, and the following result holds; see e.g., [13, Theorem 30.4].

Lemma 3.2. *Fix A, C, and K. If (1) is strongly feasible for some b, then $u_P = u_D$ for every f. If (2) is strongly feasible for some f, then $u_P = u_D$ for every b. Likewise, if (3) is strongly feasible for some b and d, then $v_P = v_D$ for every f, and if (4) is strongly feasible for some f, then $v_P = v_D$ for every b and d.*

Lemma 3.2 gives sufficient conditions under which a zero duality gap occurs for a family of conic problems parameterized by the linear objective function of the primal or dual. We will investigate conditions under which a zero duality gap occurs for every linear objective function *and* every right-hand side of (1)–(2) or (3)–(4).

The following notation will be used heavily throughout. Given linear operators A and C and a closed convex cone $\mathcal{G}$ whose dimensions are compatible, define the sets

$$\boxed{S_o(C,\mathcal{G}) = \{x \mid Cx \succ_{\mathcal{G}} 0\}, \qquad S_c(C,\mathcal{G}) = \{x \mid Cx \succeq_{\mathcal{G}} 0\},}$$

and the conditions

Property $\mathrm{P_o}(A,C,\mathcal{G})$:	$\mathcal{N}(A) \cap S_o(C,\mathcal{G})$ is nonempty, and A is onto;
Property $\mathrm{P_c}(A,C,\mathcal{G})$:	$\mathcal{N}(A) \cap S_c(C,\mathcal{G}) = \{0\}$.

[3] It follows that K^* is also a full cone.

[4] Some authors refer to strong feasibility as *strict* feasibility.

(The subscripts o and c remind the reader that $S_{\rm o}(C,\mathcal{G})$ is open and $S_{\rm c}(C,\mathcal{G})$ is closed.) Note that properties $\mathrm{P_o}(A,C,\mathcal{G})$ and $\mathrm{P_c}(A,C,\mathcal{G})$ are mutually exclusive. Figure 1 shows three possible geometrical positions for the sets $\mathcal{N}(A)$ and $S_{\rm c}(C,K)$, for an instance of (3)–(4) in which $S_{\rm o}(C,K)$ is nonempty.

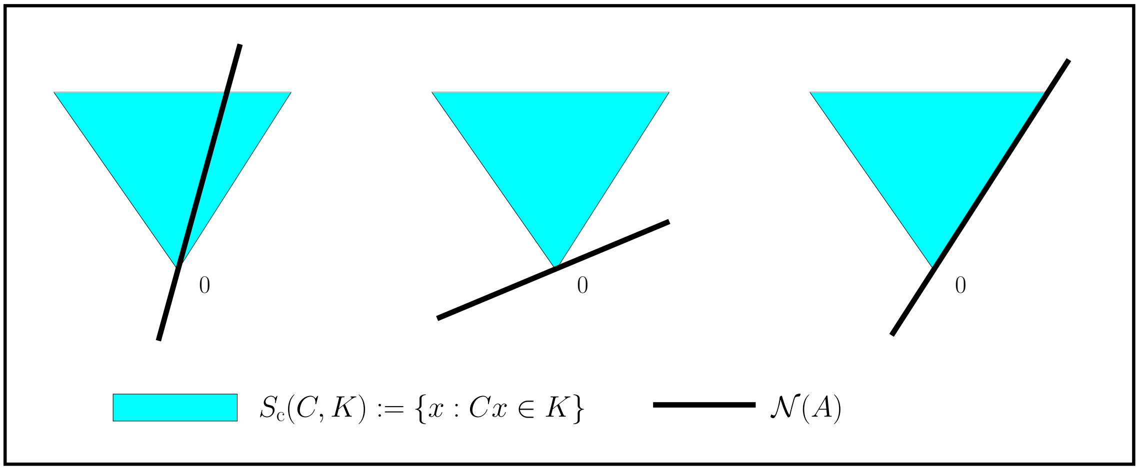

Figure 1: *The first and second plots from the left depict situations where properties* $\mathrm{P_o}(A,C,K)$ *and* $\mathrm{P_c}(A,C,K)$ *hold, respectively. The third plot shows how both properties can fail to hold.*

Before proceeding to define and characterize universal duality, we conclude the introductory portion of this section by proving some useful properties that relate $S_{\rm o}$, $S_{\rm c}$, $\mathrm{P_o}$, and $\mathrm{P_c}$, for the matrices and cone in (3)–(4).

Lemma 3.3. *If* $S_{\rm o}(C,K)$ *is nonempty, then* $S_{\rm o}(C,K) = \mathrm{int}(S_{\rm c}(C,K))$.

Proof. Suppose that C and K are such that $S_{\rm o}(C,K)$ is nonempty, and let $S_{\rm o} := S_{\rm o}(C,K)$ and $S_{\rm c} := S_{\rm c}(C,K)$. Clearly $S_{\rm o} = \mathrm{int}(S_{\rm o}) \subseteq \mathrm{int}(S_{\rm c})$, so it suffices to show that $\mathrm{int}(S_{\rm c}) \subseteq S_{\rm o}$. To prove this, let $x_{\rm c} \in \mathrm{int}(S_{\rm c})$ and $x_{\rm o} \in S_{\rm o}$. Then there exists $\alpha > 0$ such that $x_{\rm c} - \alpha x_{\rm o} \in S_{\rm c}$ and, since K is a cone, $\alpha x_{\rm o} \in S_{\rm o}$. Therefore $C(x_{\rm c} - \alpha x_{\rm o}) \succeq_K 0$ and $\alpha C x_{\rm o} \succ_K 0$. Since $x_{\rm c} = (x_{\rm c} - \alpha x_{\rm o}) + \alpha x_{\rm o}$, it follows that $Cx_{\rm c} \succ_K 0$, i.e., $x_{\rm c} \in S_{\rm o}$. ■

Lemma 3.4. *The following relations between* $\mathrm{P_o}$ *and* $\mathrm{P_c}$ *hold.*
(a) *Property* $\mathrm{P_o}(A,C,K)$ *holds if and only if property* $\mathrm{P_o}\left(\left[\begin{smallmatrix} A & 0 \\ C & -I \end{smallmatrix}\right], I, \mathcal{R}^n \times K\right)$ *holds.*
(b) *Property* $\mathrm{P_c}(A,C,K)$ *holds if and only if property* $\mathrm{P_c}\left(\left[\begin{smallmatrix} A & 0 \\ C & -I \end{smallmatrix}\right], I, \mathcal{R}^n \times K\right)$ *holds.*
(c) *Property* $\mathrm{P_o}(A,C,K)$ *holds if and only if property* $\mathrm{P_c}([A^* \ C^*], I, \mathcal{R}^m \times K^*)$ *holds.*
(d) *Property* $\mathrm{P_c}(A,C,K)$ *holds if and only if property* $\mathrm{P_o}([A^* \ C^*], I, \mathcal{R}^m \times K^*)$ *holds.*

Proof. Statements (a) and (b) can be easily verified from the definitions of $\mathrm{P_o}$ and $\mathrm{P_c}$. We now prove statements (c) and (d). It follows from Lemma 2.1 that, for a linear operator L and a *solid* closed convex cone K, the system $Lx = 0,\ x \succ_K 0$ has a solution x if and only if $L^*y \in K^* \Rightarrow L^*y = 0$. So property $\mathrm{P_o}(L,I,K)$, which amounts to "$Lx = 0,\ x \succ_K 0$ has a solution x, and L is onto", is equivalent to "$L^*y \in K^* \Rightarrow L^*y = 0$, and L onto", which in turn is readily seen to be equivalent to the implication $L^*y \in K^* \Rightarrow y = 0$. So property $\mathrm{P_o}(L,I,K)$

is equivalent to the implication $L^*y \in K^* \Rightarrow y = 0$. Now replacing L and K by $\left[\begin{smallmatrix} A & 0 \\ C & -I \end{smallmatrix}\right]$ and $\mathcal{R}^n \times K$ respectively, and using (a), yields (c) after simplification. Replacing A, C, and K in (c) by $[A^* \ C^*]$, I, and $\mathcal{R}^m \times K^*$ respectively, we obtain statement (d) after simplification. ■

3.1 Universal duality

We first focus on universal duality for the standard form (1)–(2).

Definition 3.5. *Given a linear operator $A : \mathcal{R}^n \to \mathcal{R}^m$ and a full cone $K \subset \mathcal{R}^p$, we say that* universal duality *holds for (A, K) if for all choices of b and f, $u_P = u_D$ holds in (1)–(2). (A common value of $+\infty$ or $-\infty$ is permitted.)*

In characterizing universal duality in terms of properties $\mathrm{P_o}$ and $\mathrm{P_c}$, we will use the following two lemmas. Some parts of these lemmas are well known, but for ease of reference, we give a complete proof for Lemma 3.6 here. Lemma 3.7 is proven in a similar way.

Lemma 3.6. *The following statements are equivalent.*
(a) *Property $\mathrm{P_o}(A, I, K)$ holds.*
(b) *For every f, the set $\mathcal{F}_D$ is bounded (and possibly empty).*
(c) *For every b, (1) is feasible.*
(d) *For every b, (1) is strongly feasible.*

Proof. First, observe from Lemma 3.4(c) that properties $\mathrm{P_o}(A, I, K)$ and $\mathrm{P_c}([A^* \ I], I, \mathcal{R}^m \times K^*)$ are equivalent. Now the latter property can be expressed in the following way: the only $z \in \mathcal{R}^m \times K^*$ satisfying $[A^* \ I]z = 0$ is $z = 0$. But this characterization is none other than that of the nonexistence of a (nonzero) recession direction for $\mathcal{F}_D$. Hence (a)⇔(b). This same characterization of nonexistence can be rewritten as $\{y \mid A^*y \in K^*\} = \{0\}$. Taking the dual of each side of this equality yields the equivalent statement $\mathrm{cl}(A(K)) = \mathcal{R}^m$. Now $\mathrm{cl}(A(K)) = \mathcal{R}^m$ if and only if $A(K) = \mathcal{R}^m$, which is nothing other than feasibility of (1) for every b. Hence (b)⇔(c). Finally, suppose that $A(K) = \mathcal{R}^m$. Taking the relative interior of each side yields the equivalent statement $\mathrm{ri}(A(K)) = \mathcal{R}^m$. Now the relation $\mathrm{ri}(A(K)) = A(\mathrm{ri}(K))$ holds since A is a linear operator A and K is a convex set [13, Theorem 6.6]. Moreover, $\mathrm{ri}(K) = \mathrm{int}(K)$ since K has nonempty interior. So (c) is equivalent to $A(\mathrm{int}(K)) = \mathcal{R}^m$, i.e., (1) is strongly feasible for every b. Hence (c)⇔(d). ■

Lemma 3.7. *The following statements are equivalent.*
(a) *Property $\mathrm{P_c}(A, I, K)$ holds.*
(b) *For every b, the set $\mathcal{F}_P$ is bounded (and possibly empty).*
(c) *For every f, (2) is feasible.*
(d) *For every f, (2) is strongly feasible.*

We are now ready to characterize universal duality for (A, K) in terms of the properties $\mathrm{P_o}$ and $\mathrm{P_c}$.

Theorem 3.8. *Universal duality holds for (A, K) if and only if either property $\mathrm{P_o}(A, I, K)$ or property $\mathrm{P_c}(A, I, K)$ holds.*

Proof. ($\Rightarrow$) If A and K are such that both properties $\mathrm{P_o}(A, I, K)$ and $\mathrm{P_c}(A, I, K)$ fail, then it follows from the implications (c)$\Rightarrow$(a) in Lemmas 3.6 and 3.7 that for some b and f, (1) and (2) are infeasible, i.e., $u_P = +\infty$ and $u_D = -\infty$. Clearly universal duality cannot hold for (A, K).
($\Leftarrow$) If either of the two properties hold, then implications (a)$\Rightarrow$(d) in Lemmas 3.6 and 3.7 imply that strong feasibility holds for every right-hand side of (1) or for every right-hand side of (2). Universal duality for (A, I, K) now follows from Lemma 3.2. ■

We now define a related concept of universal duality for the formulation (3)–(4).

Definition 3.9. *Given linear operators $A : \mathcal{R}^n \to \mathcal{R}^m$ and $C : \mathcal{R}^n \to \mathcal{R}^p$, and a full cone $K \subset \mathcal{R}^p$, we say that* universal duality *holds for the triple (A, C, K) if for all choices of b, d, and f, $v_P = v_D$ holds in (3)–(4). (A common value of $+\infty$ or $-\infty$ is permitted.)*

A characterization of universal duality for (A, C, K) is readily obtained, analogous to Theorem 3.8.

Theorem 3.10. *Universal duality holds for (A, C, K) if and only if either property $\mathrm{P_o}(A, C, K)$ or property $\mathrm{P_c}(A, C, K)$ holds.*

Proof. Let $\bar{A} = \left[\begin{smallmatrix} A & 0 \\ C & -I \end{smallmatrix}\right]$ and $\bar{K} = \mathcal{R}^n \times K$. Observe that Lemma 3.4(a),(b) implies properties $\mathrm{P_o}(\bar{A}, I, \bar{K})$ and $\mathrm{P_o}(A, C, K)$ are equivalent, as are $\mathrm{P_c}(\bar{A}, I, \bar{K})$ and $\mathrm{P_c}(A, C, K)$. So by Theorem 3.8, property $\mathrm{P_o}(A, C, K)$ or property $\mathrm{P_c}(A, C, K)$ holds, if and only if $u_P = u_D$ for every b_1, b_2 and f, where

$$
\begin{aligned}
u_P &= -\langle f, \left[\begin{smallmatrix} 0 \\ b_2 \end{smallmatrix}\right]\rangle + \inf_{x'} \{\langle \left[\begin{smallmatrix} I \\ C \end{smallmatrix}\right]^* f, x'\rangle \mid Ax' = b_1,\ Cx' \succeq_K b_2\}, \\
u_D &= -\langle f, \left[\begin{smallmatrix} 0 \\ b_2 \end{smallmatrix}\right]\rangle + \sup_{y', w'} \{\langle b_1, y'\rangle + \langle b_2, w'\rangle \mid A^* y' + C^* w' = \left[\begin{smallmatrix} I \\ C \end{smallmatrix}\right]^* f,\ w' \succeq_{K^*} 0\}.
\end{aligned}
$$

(Here we have replaced A and K in (1)–(2) by $\bar{A}$ and $\bar{K}$.) The vectors $b_1 \in \mathcal{R}^m$ and $b_2 \in \mathcal{R}^p$ are such that $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, where b is from (1)–(2). Note that $\left[\begin{smallmatrix} I \\ C \end{smallmatrix}\right]^*$ is onto regardless of C, so given any $\bar{f}$, there exists a solution f to the linear system $\bar{f} = \left[\begin{smallmatrix} I \\ C \end{smallmatrix}\right]^* f$. It follows that $u_P = u_D$ for every b_1, b_2 and f if and only universal duality holds for (A, C, K). ■

Remark 3.11. Theorems 3.8 and 3.10 *still apply when $m = 0$ or $p = 0$, under appropriate conventions. We will adopt the convention that if $m = 0$, then A is onto and $\mathcal{N}(A) = \mathcal{R}^n$. Properties $\mathrm{P_o}(A, C, K)$ and $\mathrm{P_c}(A, C, K)$ then become property $\mathrm{P'_o}(C, K) : S_o(C, K)$ is nonempty, and property $\mathrm{P'_c}(C, K) : S_c(C, K) = \{0\}$, respectively. Properties $\mathrm{P'_o}(C, K)$ and $\mathrm{P'_c}(C, K)$ are mutually exclusive. Further, we will adopt the convention that if $p = 0$, then $K = \mathcal{R}^p = \{0\}$, and $S_o(C, K) = S_c(C, K) = \mathcal{R}^n$. Properties $\mathrm{P_o}(A, C, K)$ and $\mathrm{P_c}(A, C, K)$ then become property $\mathrm{P''_o}(A) :$ A is onto, and property $\mathrm{P''_c}(A) :$ A is one-to-one, respectively. If $p = 0$ and A is invertible (so that $m = n$), then clearly properties $\mathrm{P''_o}(A)$ and $\mathrm{P''_c}(A)$ both hold. Otherwise these two properties are mutually exclusive. Under these conventions, Theorems 3.8 and 3.10 hold when $m = 0$ or $p = 0$, with properties $\mathrm{P_o}(A, C, K)$ and $\mathrm{P_c}(A, C, K)$ replaced by their primed versions defined above.*

It is clear from the definitions that universal duality for (A, I, K) implies universal duality for (A, K). The converse also holds. That is, the set of linear operators A for which a zero duality gap is obtained in (1)–(2) for every permissible constraint right-hand side and objective function is unchanged when the primal constraint $x \succeq_K 0$ is replaced by $x \succeq_K d$.

Theorem 3.12. *Universal duality holds for (A, I, K) if and only if universal duality holds for (A, K).*

Proof. As was just pointed out, the forward implication is a direct consequence of the definitions. We prove that if universal duality fails to hold for (A, I, K), then it also fails to hold for (A, K). Suppose that universal duality does not hold for (A, I, K). Then for some b, d, and f, (3)–(4) with $C = I$ exhibits a nonzero duality gap. Now consider (3)–(4) with $C = I$, and let $\hat{x} = x - d$ and $\hat{b} = b - Ad$. Then the primal constraints become $A\hat{x} = \hat{b}$ and $\hat{x} \succeq_K 0$, and the primal objective function becomes $\langle f, \hat{x}\rangle + \langle f, d\rangle$. Noting that any dual feasible solution (y, w) must satisfy $w = f - A^*y$, we can write the dual objective function as $\langle \hat{b}, y\rangle + \langle f, d\rangle$. So (3) and (4) take the form of (1) and (2) respectively, except for the addition of a common constant term $\langle f, d\rangle$ in each objective function. Hence universal duality does not hold for (A, K). ■

In some applications in which one wishes to study the behavior of the duality gap under perturbations in the right-hand side and objective function coefficient data, it is likely that the perturbed data will be restricted. The following result shows that as long as the set of perturbed data contains the origin in its interior, then a zero duality gap ensues for that set of data (if and) only if universal duality holds. We formally state and prove this result for the standard form.

Theorem 3.13. *Let $B \subseteq \mathcal{R}^m$ and $F \subseteq \mathcal{R}^n$ be neighborhoods of the origin. Universal duality holds for (A, K) if and only if $u_P = u_D$ for every $b \in B$ and $f \in F$.*

Proof. Clearly the forward implication holds. To prove the reverse implication, let b and f in (1)–(2) be arbitrary. There exists $\alpha > 0$ such that $\alpha b \in B$ and $\alpha f \in F$, so

$$
\begin{aligned}
u_P &= \inf_x \{\langle f, x\rangle \mid Ax = b,\ x \succeq_K 0\} \\
&= \frac{1}{\alpha^2} \inf_x \{\langle \alpha f, x\rangle \mid Ax = \alpha b,\ x \succeq_K 0\} \\
&= \frac{1}{\alpha^2} \sup_{y,w} \{\langle \alpha b, y\rangle \mid A^*y + w = \alpha f,\ w \succeq_{K^*} 0\} \\
&= \sup_{y,w} \{\langle b, y\rangle \mid A^*y + w = f,\ w \succeq_{K^*} 0\} \\
&= u_D,
\end{aligned}
$$

where we have used the mappings $x \to \alpha x$ and $(y, w) \to \alpha(y, w)$. ■

3.2 Universal duality and the boundedness of primal and dual feasible sets

It is shown in [4, Theorem 1] that if a convex program has a nonempty bounded feasible set, then its dual must have an unbounded feasible set. This turns out to be a direct corollary of the following result that connects the boundedness or lack thereof of the feasible sets $\mathcal{F}_P$ and

$\mathcal{F}_D$ to properties $\mathrm{P_o}$ and $\mathrm{P_c}$. The results in this section are phrased in terms of the standard form, but are easily extended to (3)–(4).

Theorem 3.14. (a) *If property $\mathrm{P_o}(A, I, K)$ holds, then for every b and f, $\mathcal{F}_D$ is bounded (and possibly empty), and $\mathcal{F}_P$ is unbounded.*
(b) *If property $\mathrm{P_c}(A, I, K)$ holds, then for every b and f, $\mathcal{F}_P$ is bounded (and possibly empty), and $\mathcal{F}_D$ is unbounded.*
(c) *If properties $\mathrm{P_o}(A, I, K)$ and $\mathrm{P_c}(A, I, K)$ both fail, then for every b and f, both $\mathcal{F}_P$ and $\mathcal{F}_D$ are unbounded or empty.*

Proof. To prove (a), suppose that property $\mathrm{P_o}(A, I, K)$ holds. It follows from the implications (a)⇒(b) and (a)⇒(c) in Lemma 3.6 that for every b and f, $\mathcal{F}_D$ is bounded (and possibly empty) and $\mathcal{F}_P$ is nonempty. Now it has already been noted that properties $\mathrm{P_o}(A, I, K)$ and $\mathrm{P_c}(A, I, K)$ are mutually exclusive, so it must be the case that $\mathrm{P_c}(A, I, K)$ fails. It then follows from the implication (b)⇒(a) in Lemma 3.7, that for some b, $\mathcal{F}_P$ is unbounded, i.e., contains a recession direction. Since $\mathcal{F}_P$ is nonempty for every b, we conclude it is unbounded for every b. This concludes the proof of statement (a). Statement (b) is proved similarly. To prove (c), observe that if properties $\mathrm{P_o}(A, I, K)$ and $\mathrm{P_c}(A, I, K)$ both fail, then Lemmas 3.6 and 3.7 imply that for some b and f, both $\mathcal{F}_P$ and $\mathcal{F}_D$ are unbounded. Hence both $\mathcal{F}_P$ and $\mathcal{F}_D$ contain recession directions, and so whenever these sets are nonempty they are unbounded. ■

We conclude this section by giving alternative necessary and sufficient conditions for universal duality of (A, K), which involve boundedness or lack thereof of $\mathcal{F}_P$ and $\mathcal{F}_D$.

Theorem 3.15. (a) *If for some b and f, either $\mathcal{F}_P$ or $\mathcal{F}_D$ is nonempty and bounded, then universal duality holds for (A, K).*
(b) *If universal duality holds for (A, K), then one of $\mathcal{F}_P$ and $\mathcal{F}_D$ is unbounded for every b and f, and the other is bounded (and possibly empty) for every b and f.*

Proof. (a) If for some b and f, either $\mathcal{F}_P$ or $\mathcal{F}_D$ is nonempty and bounded, then the contrapositive of Theorem 3.14(c) shows that either property $\mathrm{P_c}(A, I, K)$ or $\mathrm{P_o}(A, I, K)$ holds. So by Theorem 3.8, universal duality holds for (A, K).
(b) Appealing again to Theorem 3.8, observe that if universal duality holds for (A, K), then exactly one of properties $\mathrm{P_o}(A, I, K)$ and $\mathrm{P_c}(A, I, K)$ holds. Hence either statement (a) or (b) in Theorem 3.14 applies. ■

It has been shown e.g., in [13, Theorem 30.4], that for a pair of dual convex programs, if the set of primal or dual *optimal* solutions is nonempty and bounded, then a zero duality gap results. It may be of interest to compare this result with Theorem 3.15(a).

4 Generic properties of universal duality

On a Euclidean space X, we can speak of a *metrically generic* property that holds at "almost all" points in X, or a *topologically generic* property that holds on a *residual set* in X. Here, "almost all" is in the sense of Lebesgue measure, and a residual set in X is one that contains a countable intersection of open dense subsets in X.[5] Focusing on (3)–(4), we will take X to be

[5]Neither type of genericness is implied by the other. The terminology "topologically generic" and "metrically generic" can be found in, e.g., [18].

$\mathcal{R}^{m\times n} \times \mathcal{R}^{p\times n}$, since this is the domain of the pair of linear operators (A, C). We will show that for a fixed full cone K, universal duality for (A, C, K) is both a metrically generic property and a topologically generic property on X. Universal duality for (A, K) enjoys similar properties.

4.1 Metric genericness of universal duality

In showing that universal duality for (A, C, K) is metrically generic, we will use several lemmas. The first two results are well known; the first follows from Fubini's theorem; see e.g., [7, p. 147, Theorem A]. The third one is proved in the appendix.

Lemma 4.1. *A Lebesgue measurable set $W \subseteq \mathcal{R}^m \times \mathcal{R}^n$ has zero Lebesgue measure if and only if the set $\{x \in \mathcal{R}^m \mid (x, y) \in W\}$ has zero Lebesgue measure for Lebesgue almost every $y \in \mathcal{R}^n$.*

Lemma 4.2. *The set of matrices in $\mathcal{R}^{m\times n}$ containing a square singular submatrix has zero Lebesgue measure. In particular, the set of rank deficient matrices in $\mathcal{R}^{m\times n}$ has zero Lebesgue measure.*

Lemma 4.3. *Let $S \subseteq \mathcal{R}^n$ be a solid closed convex cone, and let p be a positive integer. Then the sets*

$$\begin{aligned}\mathcal{M}_1 &= \{M \in \mathcal{R}^{p\times n} \mid \mathcal{N}(M) \cap \operatorname{int}(S) \text{ is empty, and } \mathcal{N}(M) \cap S \neq \{0\}\},\\ \mathcal{M}_2 &= \{M \in \mathcal{R}^{n\times p} \mid \operatorname{Range}(M) \cap \operatorname{int}(S) \text{ is empty, and } \operatorname{Range}(M) \cap S \neq \{0\}\}\end{aligned}$$

have zero Lebesgue measure.

Theorem 4.4. *Universal duality for (A, C, K) is metrically generic. Specifically, given a full cone K, the set of pairs (A, C) such that universal duality fails to hold for (A, C, K) has zero Lebesgue measure in $\mathcal{R}^{m\times n} \times \mathcal{R}^{p\times n}$.*

Proof. Let T be the set of pairs (A, C) such that universal duality fails to hold for (A, C, K). We consider the two cases $m \geq n$ and $m < n$.

First, suppose that $m \geq n$. Then by Theorem 3.10 we have

$$\begin{aligned}T &\subseteq \{(A, C) \mid \text{property } \mathrm{P_c}(A, C, K) \text{ fails}\}\\ &= \{(A, C) \mid \mathcal{N}(A) \cap S_{\mathrm{c}}(C, K) \neq \{0\}\}\\ &\subseteq \{A \in \mathcal{R}^{m\times n} \mid \mathcal{N}(A) \neq \{0\}\} \times \mathcal{R}^{p\times n}\\ &= \{A \in \mathcal{R}^{m\times n} \mid \operatorname{rank}(A) < n\} \times \mathcal{R}^{p\times n}.\end{aligned}$$

It follows from Lemma 4.2 that $\{A \in \mathcal{R}^{m\times n} \mid \operatorname{rank}(A) < n\}$ has zero Lebesgue measure, and then from Lemma 4.1 that T has zero Lebesgue measure.

Suppose now $m < n$. Consider the following conditions on A and C: (i) $\mathcal{N}(A) \cap S_{\mathrm{o}}(C, K)$ is empty, and (ii) $\mathcal{N}(A) \cap S_{\mathrm{c}}(C, K) \neq \{0\}$. Consider also the sets

$$\begin{aligned}T_1 &= \{(A, C) \mid \text{conditions (i) and (ii) hold, and } S_{\mathrm{o}}(C, K) \text{ is empty}\},\\ T_2 &= \{(A, C) \mid \text{conditions (i) and (ii) hold, and } S_{\mathrm{o}}(C, K) \text{ is nonempty}\}.\end{aligned}$$

Noting the relationship between property $\mathrm{P_o}(A, C, K)$ and condition (i), and between property $\mathrm{P_c}(A, C, K)$ and condition (ii), we see that Theorem 3.10 implies

$$T \subseteq \{(A, C) \mid A \text{ is not onto}\} \,\cup\, T_1 \,\cup\, T_2. \tag{5}$$

The first set on the right-hand side of (5) has zero Lebesgue measure by Lemmas 4.2 and 4.1. We now proceed to show that T_1 and T_2 also have zero Lebesgue measure. In view of Lemma 4.2 we can restrict our attention to matrices C (and A) having full rank. If $\operatorname{rank}(C) = p$, then $\operatorname{Range}(C) = \mathcal{R}^p$, so that $S_o(C, K)$ is nonempty. Therefore we can assume that any C such that $(A, C) \in T_1$ satisfies $\operatorname{rank}(C) = n < p$. Now

$$\begin{aligned} T_1 &= \{(A, C) \mid S_o(C, K) \text{ is empty, and (ii) holds}\} \\ &\subseteq \{(A, C) \mid S_o(C, K) \text{ is empty, and } S_c(C, K) \neq \{0\}\,\} \\ &= \{(A, C) \mid \operatorname{Range}(C) \cap \operatorname{int}(K) \text{ is empty, and } \operatorname{Range}(C) \cap K \neq \{0\}\}, \end{aligned}$$

where the last equality holds due to C having full column rank. (This condition implies that $S_c(C, K) \neq \{0\}$ if and only if $\operatorname{Range}(C) \cap K \neq \{0\}$.) It follows from Lemmas 4.3 and 4.1 that T_1 has zero Lebesgue measure. Now in view of Lemma 3.3, any C satisfying $(A, C) \in T_2$ will also satisfy $S_o(C, K) = \operatorname{int}(S_c(C, K))$, and hence $S_c(C, K)$ will be solid. So any $A \in \mathcal{R}^{m \times n}$ such that $(A, C) \in T_2$ lies in the set

$$\{A \mid \mathcal{N}(A) \cap \operatorname{int}(S_c(C, K)) \text{ is empty, and } \mathcal{N}(A) \cap S_c(C, K) \neq \{0\}\}.$$

By Lemma 4.3 this set has zero Lebesgue measure, so it follows from Lemma 4.1 that T_2 also has zero Lebesgue measure. ■

For the standard form (1)–(2), a metric genericness result of the following form can be obtained.

Theorem 4.5. *Given a full cone K, the set*

$$\{A \mid \text{universal duality fails for } (A, K)\}$$

has zero Lebesgue measure in $\mathcal{R}^{m \times n}$.

Proof. Similar to that of Theorem 4.4. (We set $C = I$ and also use Theorem 3.12.) ■

Note that Theorem 4.5 neither implies nor is implied by Theorem 4.4.

4.2 Topological genericness of universal duality

Theorem 4.6. *Universal duality for (A, C, K) is topologically generic. In fact, given a full cone K, the set of pairs (A, C) for which universal duality holds for (A, C, K) is open and dense in $\mathcal{R}^{m \times n} \times \mathcal{R}^{p \times n}$.*

Proof. The complement of a set having zero Lebesgue measure is dense. (If not, then that set would contain an open hypercube, which must have positive measure.) So Theorem 4.4 implies that the set of pairs (A, C) such that universal duality holds for (A, C, K) is dense in $\mathcal{R}^{m \times n} \times \mathcal{R}^{p \times n}$. We now show that:
(a) The set of pairs (A, C) such that property $\mathrm{P_o}(A, C, K)$ holds is open in $\mathcal{R}^{m \times n} \times \mathcal{R}^{p \times n}$;

and
(b) The set of pairs (A, C) such that property $\mathrm{P_c}(A, C, K)$ holds is open in $\mathcal{R}^{m\times n} \times \mathcal{R}^{p\times n}$.

To prove (a), suppose that (A, C) is such that property $\mathrm{P_o}(A, C, K)$ holds. When $m > n$, A cannot be onto, so it must be the case that $m \le n$. Further, if $m = n$, then $\mathcal{N}(A) \cap S_o(C, K)$ is empty whenever A is onto, so property $\mathrm{P_o}(A, C, K)$ fails to hold. Hence $m < n$. Now let $\{(A^i, C^i)\}_i$ be an infinite sequence such that $(A^i, C^i) \to (A, C)$. Since the set of full rank matrices is open, then A^i is onto for i large enough. So it is enough to show that $\mathcal{N}(A^i) \cap S_o(C^i, K)$ is nonempty for i large enough. Let $x \in \mathcal{N}(A) \cap S_o(C, K)$ and let x^i be the orthogonal projection of x onto $\mathcal{N}(A^i)$. Then $\lim_{i\to\infty} x^i = x$. Now writing $C^i x^i - Cx = C^i(x^i - x) + (C^i - C)x$, we have

$$\|C^i x^i - Cx\| \le \|C^i\|\|x^i - x\| + \|C^i - C\|\|x\|. \tag{6}$$

As $i \to \infty$, the right-hand side of (6), and hence the left-hand side, tends to zero. It follows from $Cx \succ_K 0$ that $C^i x^i \succ_K 0$ for i large enough. That is, $x^i \in \mathcal{N}(A^i) \cap \mathrm{S_o}(C^i, K)$ for such i. This proves statement (a).

To prove (b), let S be the set of pairs (A, C) such that property $\mathrm{P_c}(A, C, K)$ holds. Proceeding by contradiction, we suppose that S is not open in $\mathcal{R}^{m\times n} \times \mathcal{R}^{p\times n}$. Then for some $(A, C) \in S$, there exists a sequence $\{(A^i, C^i)\}_i$ with $(A^i, C^i) \notin S$ for all i, but $(A^i, C^i) \to (A, C)$. So for each i, there exists a nonzero $x^i \in \mathcal{N}(A^i) \cap S_c(C^i, K)$. Now set $y^i = x^i/\|x^i\|$ so that $\|y^i\| = 1$, $A^i y^i = 0$ and $C^i y^i \succeq_K 0$ for all i. Since $\{y^i\}$ is a bounded sequence, it contains a convergent subsequence. Passing to such a subsequence if necessary, we conclude that there exists a limit point $y \ne 0$. Since K is closed, $y \in \mathcal{N}(A) \cap S_c(C, K)$, so that $(A, C) \notin S$—a contradiction. ■

For the standard form (1)–(2), a topological genericness result of the following form can be obtained.

Theorem 4.7. *Given a full cone K, the set*

$$\{A \mid \text{universal duality holds for } (A, K)\}$$

is open and dense in $\mathcal{R}^{m\times n}$.

Proof. Similar to that of Theorem 4.6. (We set $C = I$ and also use Theorem 3.12.) ■

Note that Theorem 4.7 neither implies nor is implied by Theorem 4.6.

5 Verifying universal duality

We show that universal duality for (A, C, K) or (A, K) can be checked by solving a single conic program with essentially the same size and "structure" as that in (3). We first state two well known results.

Lemma 5.1. *Let the set S be such that S^* (defined with respect to the inner product $\langle\cdot,\cdot\rangle$) has nonempty interior. Then for any $y \in S$ and $z \in \mathrm{int}(S^*)$, $\langle y, z\rangle \le 0$ implies that $y = 0$.*

Lemma 5.2. *If $K \subset \mathcal{R}^p$ is a full cone, then $\mathrm{int}(K) \cap \mathrm{int}(K^*)$ is nonempty.*

We now show how properties $\mathrm{P_o}(A, C, K)$ and $\mathrm{P_c}(A, C, K)$, and hence universal duality for (A, C, K), can be verified by solving a single conic program.

Theorem 5.3. *Let $e \in \mathrm{int}(K) \cap \mathrm{int}(K^*)$. Universal duality for (A, C, K) can be verified by solving the conic program*

$$\bar{r} = \sup_{x,\, r} \{r \mid Ax = 0,\ Cx \succeq_K re,\ \langle Cx, e\rangle = 1\}. \tag{7}$$

Specifically,[6]
(a) *Property $\mathrm{P_o}(A, C, K)$ holds if and only if $\bar{r} > 0$ and A is onto;*
(b) *Property $\mathrm{P_c}(A, C, K)$ holds if and only if $\bar{r} < 0$ and $\mathcal{N}(A) \cap \mathcal{N}(C) = \{0\}$.*

Proof. We first show that (a) holds. It suffices to show that $\mathcal{N}(A) \cap S_\mathrm{o}(C, K)$ is nonempty if and only if $\bar{r} > 0$.
($\Rightarrow$) The nonemptiness of $\mathcal{N}(A) \cap S_\mathrm{o}(C, K)$ implies that there exists $\tilde{x}$ such that $A\tilde{x} = 0$ and $C\tilde{x} \succ_K 0$. Hence $C\tilde{x} - \tilde{r}e \succ_K 0$ for some $\tilde{r} > 0$ sufficiently small. Since $C\tilde{x} - \tilde{r}e \in K$ and $e \in K^*$, we have $k := \langle C\tilde{x}, e\rangle = \langle C\tilde{x} - \tilde{r}e, e\rangle + \tilde{r}\langle e, e\rangle > 0$. Hence $(\tilde{x}/k, \tilde{r}/k)$ is feasible for (7), so that $\bar{r} \geq \tilde{r}/k > 0$.
($\Leftarrow$) Suppose $\bar{r} > 0$. Then there exists $\tilde{r} > 0$ and $\tilde{x}$ such that $A\tilde{x} = 0$ and $C\tilde{x} \succeq_K \tilde{r}e \succ_K 0$, i.e., $\tilde{x} \in \mathcal{N}(A) \cap S_\mathrm{o}(C, K)$.

We now prove statement (b).
($\Rightarrow$) Suppose that $\mathrm{P_c}(A, C, K)$ holds. Then $0 \subseteq \mathcal{N}(A) \cap \mathcal{N}(C) \subseteq \mathcal{N}(A) \cap S_\mathrm{c}(C, K) = \{0\}$, so $\mathcal{N}(A) \cap \mathcal{N}(C) = \{0\}$. It remains to prove that $\bar{r} < 0$.

If (x, r) with $r \geq 0$ satisfies the constraints $Ax = 0$ and $Cx \succeq_K re$ ($\succeq_K 0$) in (7), then $x \in \mathcal{N}(A) \cap S_\mathrm{c}(C, K)$. Since $\mathrm{P_c}(A, C, K)$ holds, we must have $x = 0$, but this violates the constraint $\langle Cx, e\rangle = 1$. Hence every pair (x, r) with $r \geq 0$ is infeasible for (7). It follows that $\bar{r} \leq 0$. We now rule out the case $\bar{r} = 0$.

If (7) is infeasible, there is nothing to prove, so suppose that (7) is feasible for $(\hat{x}, \hat{r})$ with $\hat{r} < 0$. Consider the set T of feasible points (x, r) satisfying $\hat{r} \leq r \leq 0$. Suppose there exists a recession direction $(d_x, d_r) \in \mathcal{R}^n \times \mathcal{R}$ for T. Since r is bounded in T, $d_r = 0$, and d_x satisfies $Ad_x = 0$, $Cd_x \succeq_K 0$, and $\langle Cd_x, e\rangle = 0$. Since $e \succ_{K^*} 0$, then by Lemma 5.1, the last two conditions imply that $Cd_x = 0$. So $d_x \in \mathcal{N}(A) \cap \mathcal{N}(C)$, which was shown to be the origin. Hence the nonempty set T is bounded. It follows from the closedness of K that the feasible set of (7), and hence T, is closed. So $\bar{r}$, which equals the supremum of a linear function over the compact set T, is achieved. Since we showed that (x, r) is infeasible for every $r \geq 0$, it follows that $\bar{r} < 0$.
($\Leftarrow$) If $\bar{r} < 0$, then $r = 0$ is infeasible for (7), so there does not exist an x such that $Ax = 0$, $Cx \succeq_K 0$, and $\langle Cx, e\rangle > 0$. That is, any x satisfying $Ax = 0$ and $Cx \succeq_K 0$ must also satisfy $\langle Cx, e\rangle \leq 0$, which implies $Cx = 0$ by Lemma 5.1, since $e \succ_{K^*} 0$. In other words, $\mathcal{N}(A) \cap S_\mathrm{c}(C, K) = \mathcal{N}(A) \cap \mathcal{N}(C)$. Since $\mathcal{N}(A) \cap \mathcal{N}(C) = \{0\}$, property $\mathrm{P_c}(A, C, K)$ holds. ■

Remark 5.4. *Theorem 5.3 can be used to check universal duality for (A, K) by setting $C = I$ and using Theorem 3.12.*

[6] The set of instances for which $\bar{r} < 0$ includes those for which (7) is infeasible, i.e., $\bar{r} = -\infty$. In contrast, it is not possible for A, C, and K to be such that $\bar{r} = +\infty$. In fact the constraints in (7) imply that $\bar{r} \leq 1/\|e\|^2$.

Remark 5.5. *If $p = 0$, then Remark 3.11—with properties $\mathrm{P_o}(A, C, K)$ and $\mathrm{P_c}(A, C, K)$ replaced by $\mathrm{P''_o}(A)$ and $\mathrm{P''_c}(A)$—tells us that* for a pair of dual problems containing linear equality constraints only, universal duality holds for (A, C, K) *if and only if A is onto or one-to-one. Of course there is no need to solve a conic program to verify whether A satisfies these conditions. If $m = 0$, then Theorem 5.3—with properties $\mathrm{P_o}(A, C, K)$ and $\mathrm{P_c}(A, C, K)$ replaced by $\mathrm{P'_o}(C, K)$ and $\mathrm{P'_c}(C, K)$—holds under the convention specified in Remark 3.11.*

6 An application of universal duality

In this section, we consider certain semidefinite programs (SDPs) derived from the Kalman-Yakubovich-Popov (KYP) lemma, which are of interest in control theory and signal processing. Specifically we study one type of KYP-SDP from [21, Section 2.2]. First let us define the necessary notation. Denote the space of symmetric matrices of order n by $\mathcal{S}^n$ and the cone of positive semidefinite matrices by $\mathcal{S}^n_+$. The standard inner product defined on $\mathcal{S}^n$ is given by $\langle M, N\rangle = \mathrm{trace}(MN)$ for $M, N \in \mathcal{S}^n$. It can be shown that the positive semidefinite cone is a full cone that is also *self-dual*, i.e., $(\mathcal{S}^n_+)^* = \mathcal{S}^n_+$. The interior of $\mathcal{S}^n_+$ is the set of positive definite matrices. In this section, $\succeq$ refers to the ordering induced by the positive semidefinite cone. That is, given matrices $M, N \in \mathcal{S}^n$, $M \succeq N$ means that $M - N$ is a positive semidefinite matrix.

Consider the continuous-time dynamical system

$$\dot{x}(t) = Ax(t) + Bu(t), \quad t \geq 0, \qquad x(0) = x_0,$$

where $A \in \mathcal{R}^{n\times n}$, $B \in \mathcal{R}^{n\times m}$, $x(t) \in \mathcal{R}^n$, and $u(t) \in \mathcal{R}^m$. Given a matrix $M \in \mathcal{S}^{n+m}$, we seek the optimal state vector x and control vector u such that the cost functional

$$J(u) = \int_0^\infty \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T M \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \, dt$$

is minimized over (say) the space of piecewise continuous controls u, subject to the above differential equation and the constraint that $x(t) \to 0$ as $t \to \infty$. Define the linear operator $\mathcal{L} : \mathcal{S}^n \to \mathcal{S}^{n+m}$ by

$$\mathcal{L}(P) = \begin{bmatrix} A^TP + PA & PB \\ B^TP & 0 \end{bmatrix}.$$

This optimal control problem is closely linked ([21, Section 2.2]) to the SDP

$$v = \sup_{P\in\mathcal{S}^n} \{\langle P, x_0x_0^T\rangle \mid \mathcal{L}(P) \succeq -M\}, \tag{8}$$

in that, whenever the optimal value of J exists, it is given by v. A *sufficient* condition for this to hold is that the pair of matrices (A, B) is *controllable*, i.e., the matrix

$$[B \; AB \; A^2B \; \cdots A^{n-1}B] \tag{9}$$

has full (row) rank. (This condition is not *necessary* however, as the following trivial example shows: $n = m = 1$, $A = -1$, $B = 0$, $x_0 = 1$, $M = I_2$, with $v = 1/2$.) The following characterization of controllability of (A, B) in terms of L can be found in [20, Lemma 1].

Lemma 6.1. *The pair (A, B) is controllable if and only if the following implication holds:*

$$\mathcal{L}(P) \succeq 0 \implies P = 0. \tag{10}$$

The dual of (8) is the SDP

$$\inf_{Z \in \mathcal{S}^{n+m}} \{\langle M, Z\rangle \mid \mathcal{L}^*(Z) = -x_0 x_0^T,\ Z \succeq 0\}. \tag{11}$$

Note that (11) is in "primal standard form", i.e., in the form of (1), with $\mathcal{L}^*$ and $\mathcal{S}_+^{n+m}$ playing the role of A and K in (1). It turns out that controllability of (A, B) is precisely what is needed for universal duality to hold for the pair $(\mathcal{L}^*, \mathcal{S}_+^{n+m})$:

Theorem 6.2. *Universal duality holds for $(\mathcal{L}^*, \mathcal{S}_+^{n+m})$ if and only if (A, B) is controllable.*

Proof. We show that controllability of (A, B) is equivalent to the absence of a duality gap between the optimal values of (8) and (11), with $x_0 x_0^T$ replaced by Q, for all M and $Q \in \mathcal{S}^n$. Equivalently, in (8), further replace "sup" with "$-$inf" and Q with $-Q$, and view this transformed SDP as being in primal form (3) with vacuous explicit equality constraints. To prove the claim, it then suffices to show that (A, B) is controllable if and only if either property $\mathrm{P}'_{\mathrm{o}}(\mathcal{L}, \mathcal{S}_+^{n+m})$ holds or property $\mathrm{P}'_{\mathrm{c}}(\mathcal{L}, \mathcal{S}_+^{n+m})$ holds—see Remark 3.11. Now property $\mathrm{P}'_{\mathrm{o}}(\mathcal{L}, \mathcal{S}_+^{n+m})$ is equivalent to the existence of a P such that $\mathcal{L}(P)$ is positive definite. But such a P cannot exist since the $(2, 2)$ block of $\mathcal{L}(P)$ is zero. Finally, it can be easily verified that property $\mathrm{P}'_{\mathrm{c}}(\mathcal{L}, \mathcal{S}_+^{n+m})$ is none other than the implication (10). So the claim follows from Lemma 6.1. ■

Corollary 6.3. *If (A, B) is controllable, then the pair (8)–(11) admits a zero duality gap for every M and every x_0.*

Does universal duality for $(\mathcal{L}^*, \mathcal{S}_+^{n+m})$ hold generically on the space of pairs (A, B)? Since the operator $\mathcal{L}$, due to its specific form, is restricted to lie in a subspace of the space of linear operators mapping $\mathcal{S}^n$ to $\mathcal{S}^{n+m}$, the genericness results of Section 4 have no bearing. Still, because it is equivalent to controllability of (A, B), universal duality for $(\mathcal{L}^*, \mathcal{S}_+^{n+m})$ is indeed metrically and topologically generic in the space of all matrix pairs (A, B). This follows from the characterization of controllability in (9).

7 Conclusions

Given a pair of dual convex problems in conic form, we introduced the concept of universal duality, which is said to hold if a zero duality gap occurs for every linear objective function and constraint right-hand side. We obtained simple necessary and sufficient conditions on the "constraint matrices" and cone that guarantee universal duality. We also gave a relationship between universal duality for conic optimization, and boundedness or lack thereof of the primal and dual feasible sets. A corollary of this relationship is the well known result that the feasible sets of a pair of dual conic programs cannot both be bounded (unless they are both empty). Further, we showed that universal duality holds almost everywhere, and holds on an open, dense set of "constraint matrices", and that universal duality, which gives duality information about an infinite family of conic programs, can be verified by solving a single conic optimization problem. Finally, we showed how our universal duality framework could be applied to a problem found in control theory.

Universal duality and its genericness has consequences for the stability (well-posedness) of conic programs. If A, C, and K are such that universal duality holds for (A, C, K), then for any change in the objective function and right-hand side data, a zero duality gap will result, and the optimal Lagrange multipliers will thus be meaningful for any data whenever the optimal value is finite. Since universal duality hold generically as A, C, and K vary, meaningfulness of the Lagrange multipliers is also generic, and thus conic programs are generically well-posed.

Acknowledgments. We are also grateful to Pierre-Antoine Absil, Carlos Berenstein, Benjamin J. Howard, Daniel Hug, John W. Mitchell, Gábor Pataki, Rolf Schneider, and Henry Wolkowicz, for helpful discussions related to this work. We also thank two anonymous referees who made many valuable suggestions. In particular, they pointed out the result in Lemma 2.1, alerted us to [11], and suggested improvements in the organization of Section 3.

A Appendix: Proof of Lemma 4.3

Our aim is to show that the sets $\mathcal{M}_1$ and $\mathcal{M}_2$ have zero Lebesgue measure. These sets are closely related to the set of q-dimensional linear subspaces L of $\mathcal{R}^n$ for which $L \cap \text{int}(S)$ is empty and $L \cap S \neq \{0\}$, where $q = n - p$ and p, respectively. We exploit this correspondence by invoking a deep theorem on convex bodies in [16, p. 93]. (The result there was first stated in [22].) This result is adapted as Lemma A.1 below, which concerns the Hausdorff measure of a particular subset of $G(n, q)$—the metric space of q-dimensional linear subspaces of $\mathcal{R}^n$.[7] Before stating this result, we discuss Hausdorff measure on metric spaces, and the distance function (i.e., metric) we will associate with the metric space $G(n, q)$.

Given a metric space (X, ρ), where ρ is the distance function, and given $t \geq 0$, the t-dimensional *Hausdorff (outer) measure* of $T \subseteq X$ is defined by

$$\mathcal{H}^t_\rho(T) = \lim_{\delta \to 0} \inf \left\{ \sum_i d_\rho(U_i)^t \;\middle|\; \{U_i\} \text{ is a } \delta\text{-cover of } T \right\}. \tag{A-1}$$

(The limit in (A-1) always exists, though its value may be infinite.) Here d_ρ is the diameter function

$$d_\rho(T) = \sup \{\rho(x, y) \mid x, y \in T\},$$

and a δ-*cover* of T is a countable collection of sets $\{U_i\}$ satisfying $T \subseteq \bigcup_i U_i$ and $0 < d_\rho(U_i) \leq \delta$ for each i.

Suppose now that t is a positive integer. It can be shown that on a t-dimensional Euclidean space endowed with the usual Euclidean distance function, the associated t-dimensional Hausdorff measure of a set $T \subseteq \mathcal{R}^t$ is a constant multiple of the Lebesgue outer measure of T in $\mathcal{R}^t$; see e.g., [14, Theorem 30]. Since a set having zero Lebesgue outer measure is Lebesgue

[7] The set $G(n, q)$ together with a specified "differentiable structure", is known as the *Grassmann manifold*. We will not explicitly use any topological properties of this manifold however.

measurable [15, p. 57, Lemma 6], it follows that a set $T \subset \mathcal{R}^t$ has zero t-dimensional Hausdorff measure if and only if T is Lebesgue measurable and its Lebesgue measure $\mathcal{L}(T)$ equals zero.

For any positive integers n and q with $n > q$, the Hausdorff measure on $G(n,q)$ referred to throughout this appendix will be that associated with the "arc-length" distance function ρ, which is the distance function induced by the unique (to scale) "rotation-invariant Riemannian metric" on $G(n,q)$. It is pointed out in [1, Section 3] that this distance function can be expressed as the two-norm of the vector of "principal" or "canonical" angles between linear subspaces. See also [5, p. 337].

In the sequel, a q-dimensional affine subspace $L \subset \mathcal{R}^n$ with $1 \le q \le n-1$ is said to *support* a nonempty closed convex set S, if L is contained in a supporting hyperplane of S, and $L \cap S$ is nonempty.

Lemma A.1. *Let $S \subseteq \mathcal{R}^n$ be a closed convex cone, q be an integer satisfying $1 \le q \le n-1$, and $\ell = q(n-q)$. The set of linear subspaces lying in $G(n,q)$ that support S, and that contain a ray of S, has zero ℓ-dimensional Hausdorff measure.*[8]

Proof. From the theorem in [16, p. 93], the result holds when S is a *convex body*, i.e., S is nonempty, compact, and convex. Now let S' be the intersection of the closed convex cone S with some convex body containing the origin in its interior. Clearly, S' is a convex body, and any linear subspace that supports S will also support S'. Since the result holds when S is replaced by S', it also holds for S itself. ∎

We now state a useful result that specializes [14, Theorem 29].

Lemma A.2. *Let (X,μ) and (Y,ν) be metric spaces, and $T \subseteq X$. Let $f : T \to Y$ be a Lipschitz mapping, viz., there exists a constant $k > 0$ independent of x_1 and x_2 such that*

$$\nu(f(x_1), f(x_2)) \le k\mu(x_1, x_2) \quad \forall\, x_1, x_2 \in T.$$

Then for any $r \ge 0$,

$$\mathcal{H}^r_\nu(f(T)) \le k\mathcal{H}^r_\mu(T).$$

In particular, if $Y = \mathcal{R}^r$ and T is such that $\mathcal{H}^r_\mu(T) = 0$, then $\mathcal{L}(f(T)) = 0$.

Our final preliminary result shows that if $T \subset G(n, n-q)$ has zero $q(n-q)$-dimensional Hausdorff measure, then the set of matrices whose nullspace or range is T has zero Lebesgue measure.

Lemma A.3. *Let n and q with $n > q$ be positive integers, and $\ell = q(n-q)$. Let $T \subset G(n, n-q)$ be such that $\mathcal{H}^\ell_\rho(T) = 0$. Then the set $\{A \in \mathcal{R}^{q\times n} \mid \mathcal{N}(A) \in T\}$ has zero Lebesgue measure. Dually, if $T \subset G(n,q)$ is such that $\mathcal{H}^\ell_\rho(T) = 0$, then $\{A \in \mathcal{R}^{n\times q} \mid \mathrm{Range}(A) \in T\}$ has zero Lebesgue measure.*

Proof. Let

$$U = \{\mathcal{N}([I_q \;\; B]) \text{ for some } B \in \mathcal{R}^{q\times(n-q)}\} \subset G(n, n-q),$$

[8] A stronger result is stated in [16]. In particular, the set of linear subspaces in the lemma has σ-finite $(\ell-1)$-dimensional Hausdorff measure. A set having σ-finite measure can be written as a countable union of sets having finite measure. Here ℓ is both the "topological dimension" and the "Hausdorff dimension" of the entire metric space $G(n,q)$.

and let $\tilde{U}$ denote the complement of U. The set $\{A \in \mathcal{R}^{q\times n} \mid \mathcal{N}(A) \in \tilde{U}\}$ is the set of matrices in $\mathcal{R}^{q\times n}$ whose leading square full-dimensional submatrix is singular. By Lemma 4.2, this set has zero Lebesgue measure, and therefore so does $\{A \in \mathcal{R}^{q\times n} \mid \mathcal{N}(A) \in T \cap \tilde{U}\}$. To complete the proof of the first claim of the lemma, it therefore suffices to show that $\{A \in \mathcal{R}^{q\times n} \mid \mathcal{N}(A) \in T \cap U\}$ has zero Lebesgue measure.

We proceed by first defining the map $\phi : U \to \mathcal{R}^{q\times(n-q)}$ by $\mathcal{N}([I_q \;\; B]) \mapsto B$.[9] Let

$$U_i = \{L \in U \mid \|\phi(L)\| \le i\}$$

for each positive integer i. (Here $\|\cdot\|$ is an operator norm on $\mathcal{R}^{q\times(n-q)}$.) It can be verified that the restriction of ϕ to each U_i is Lipschitz continuous with respect to the arc-length distance function ρ on U_i and the metric induced by the operator norm $\|\cdot\|$ on $\mathcal{R}^{q\times(n-q)}$. Since $\mathcal{H}^\ell_\rho(T \cap U_i) \le \mathcal{H}^\ell_\rho(T) = 0$ for each i, and the range of ϕ has dimension ℓ, it follows from Lemma A.2 that $\mathcal{L}(\phi(T \cap U_i)) = 0$ for each i.

Now let GL_q denote the set of square nonsingular matrices of order q with real entries.[10] Define the map $g : \mathrm{GL}_q \times \mathcal{R}^{q\times(n-q)} \to \mathcal{R}^{q\times n}$ by $(M, B) \mapsto M[I_q \;\; B]$, and let $V = \mathrm{GL}_q \times \phi(T \cap U)$. It can be verified that

$$g(V) = \{A \in \mathcal{R}^{q\times n} \mid \mathcal{N}(A) \in T \cap U\},$$

so we need to show that $\mathcal{L}(g(V)) = 0$.

Now define $\mathrm{GL}_{q,i} = \{M \in \mathrm{GL}_q \mid \|M\| \le i\}$ and $V_i = \mathrm{GL}_{q,i} \times \phi(T \cap U_i)$ for positive integers i. It is clear that the restriction of g to each V_i is Lipschitz continuous. Since $\mathcal{L}(\phi(T \cap U_i)) = 0$ for each i, it follows from Lemma 4.1 that $\mathcal{L}(V_i) = 0$ for each i. Now the domain and range of g are of the same dimension qn, so it follows from Lemma A.2 that $\mathcal{L}(g(V_i)) = 0$ for each i. Finally, since $V = \bigcup_{i=1}^\infty V_i$ is a countable union, we have

$$\mathcal{L}(g(V)) = \mathcal{L}\big(g\big(\cup_i \;\; V_i\big)\big) = \mathcal{L}\big(\cup_i \;\; g(V_i)\big) \le \sum_i \mathcal{L}(g(V_i)) = 0.$$

The dual statement is proved similarly, using

$$U = \left\{\mathrm{Range}\left(\begin{bmatrix} I_q \\ B \end{bmatrix}\right) \quad \text{for some } B \in \mathcal{R}^{(n-q)\times q}\right\} \subset G(n,q),$$

and the maps $\phi : U \to \mathcal{R}^{(n-q)\times q}$ defined by $\mathrm{Range}\left(\begin{bmatrix} I_q \\ B \end{bmatrix}\right) \mapsto B$, and $g : \mathrm{GL}_q \times \mathcal{R}^{(n-q)\times q} \to \mathcal{R}^{n\times q}$ defined by $(M, B) \mapsto \begin{bmatrix} I_q \\ B \end{bmatrix} M$. ■

With these results in hand, we now complete the proof of Lemma 4.3.

Proof. If $p \ge n$, then Lemma 4.1 implies that the sets $\{M \in \mathcal{R}^{p\times n} \mid \mathcal{N}(M) \ne \{0\}\}$ and $\{M \in \mathcal{R}^{n\times p} \mid \mathrm{Range}(M) \ne \mathcal{R}^n\}$ have zero Lebesgue measure. Hence the sets $\mathcal{M}_1$ and $\mathcal{M}_2$ also have zero Lebesgue measure. Now suppose $1 \le p \le n-1$, and define

$$\hat{G}(n,q,S) = \{L \in G(n,q) \mid L \cap \mathrm{int}(S) \text{ is empty, and } L \cap S \ne \{0\}\}$$

[9] To see that ϕ is a single-valued mapping, suppose that $B_1, B_2 \in R^{q\times(n-q)}$ are such that $\phi^{-1}(B_1) = \phi^{-1}(B_2)$, i.e., $\mathcal{N}([I_q \;\; B_1]) = \mathcal{N}([I_q \;\; B_2])$. Then there exists a nonsingular matrix $M \in \mathcal{R}^{q\times q}$ such that $[I_q \;\; B_1] = M[I_q \;\; B_2]$. It follows that $M = I_q$ and $B_1 = B_2$. The map ϕ is one of the canonical "chart mappings" that give the Grassmann manifold its "differentiable structure".

[10] Typically, GL_q is used to denote the general linear group of order q over $\mathcal{R}$, equipped with matrix multiplication. In a slight abuse of notation, we use GL_q to denote the *set* of matrices in this group.

for $q = p, n-p$. Suppose $L \in \hat{G}(n,q,S)$. Since S is solid, it follows from [6, p. 17, Exercise 1] that L supports S. Moreover, $L \cap S$ is the intersection of two convex cones, and is therefore itself a convex cone. Since $L \cap S \neq \{0\}$, this cone must have dimension at least one. That is, L contains a ray of S. It follows from Lemma A.1 that $\mathcal{H}^{\ell}_{\rho}(\hat{G}(n,p,S)) = 0$. Hence from Lemma A.3, the sets $\{M \in \mathcal{R}^{p\times n} \mid \mathcal{N}(M) \in \hat{G}(n,n-p,S)\}$ and $\{M \in \mathcal{R}^{n\times p} \mid \text{Range}(M) \in \hat{G}(n,p,S)\}$ have zero Lebesgue measure. Apart from the requirement that these sets contain only full rank matrices, these sets are $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively. In view of Lemma 4.2, we conclude that $\mathcal{M}_1$ and $\mathcal{M}_2$ also have zero Lebesgue measure. ■

References

[1] Pierre-Antoine Absil, Robert Mahony, and Rodolphe Sepulchre. Riemannian geometry of Grassmann manifolds with a view on algorithmic computation. *Acta Appl. Math.*, 80(2):199–220, 2004.

[2] Farid Alizadeh, Jean-Pierre A. Haeberly, and Michael L. Overton. Complementarity and nondegeneracy in semidefinite programming. *Math. Programming*, 77(2, Ser. B):111–128, 1997.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[4] Richard J. Duffin. Clark's theorem on linear programs holds for convex programs. *Proc. Nat. Acad. Sci. U.S.A.*, 75(4):1624–1626, 1978.

[5] Alan Edelman, Tomás A. Arias, and Steven T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353, 1998.

[6] Günter Ewald. *Combinatorial convexity and algebraic geometry*, volume 168 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996.

[7] Paul R. Halmos. *Measure Theory*. D. Van Nostrand Company, Inc., 1950.

[8] Zhi-Quan Luo, Jos F. Sturm, and Shuzhong Zhang. Conic convex programming and self-dual embedding. *Optim. Methods Softw.*, 14(3):169–218, 2000.

[9] Olvi L. Mangasarian. *Nonlinear programming*, volume 10 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994. Reprint of the 1969 original.

[10] Gabor Pataki. On the closedness of the linear image of a closed convex cone. Research Report TR-02-3, Department of Operations Research, University of North Carolina, Chapel Hill, 2003.

[11] Gábor Pataki and Levent Tunçel. On the generic properties of convex optimization problems in conic form. *Math. Programming.*, 89(3, Ser. A):449–457, 2001.

[12] James Renegar. *A mathematical view of interior-point methods in convex optimization.* MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2001.

[13] R. Tyrrell Rockafellar. *Convex analysis.* Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original.

[14] C. Ambrose Rogers. *Hausdorff measures.* Cambridge University Press, London, 1970.

[15] Halsey L. Royden. *Real analysis, second edition.* The Macmillan Co., New York, 1968.

[16] Rolf Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 44 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1993.

[17] Alexander Schrijver. *Theory of linear and integer programming.* Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons Ltd., Chichester, 1986.

[18] Stanislav Smirnov. Functional Analysis lecture notes. http://www.math.kth.se/∼stas/fa00/part1.ps, 2000.

[19] Mike Todd. Semidefinite optimization. *Acta Numerica*, 10:515–560, 2001.

[20] Lieven Vandenberghe and V. Ragu Balakrishnan. Semidefinite programming duality and linear system theory: connections and implications for computation. In *Proc. IEEE Conference on Decision and Control*, pages 989–994, Phoneix, AZ, December 1999.

[21] Lieven Vandenberghe, V. Ragu Balakrishnan, Ragnar Walliny, Anders Hanssony, and Tae Roh. Interior-point algorithms for semidefinite programming problems derived from the KYP lemma. In A. Garulli and D. Henrion, editors, *Positive Polynomials in Control*, Lectures Notes in Control and Information Science. Springer, 2004.

[22] Viktor A. Zalgaller. k-dimensional directions singular for a convex body F in R^n. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.*, 27:67–72, 1972. English translation: J. Soviet Math. 3 (1975), 437–441.