Bilinear Positivity Certificates as Convex-in-Polyhedron Set Containments

Brendon G. Anderson
California Polytechnic State University

1 C-in-P Bilinear Programs

Let $C \subseteq \mathbb{R}^n$ be a nonempty convex set and let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a nonempty convex polyhedron with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The primary focus of this paper is on the bilinear program

$$p^\star := \inf_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \left\{ y^\top (b - Ax) : y \geq 0,\ x \in C \right\}. \tag{1}$$

In particular, we seek to provide certificates of positivity for the bilinear function $f \colon (x, y) \mapsto y^\top (b - Ax)$ over the product set $C \times \mathbb{R}^m_+$, which amounts to showing that $p^\star \geq 0$. We now show that certifying the positivity of $p^\star$ exactly corresponds to verifying the convex-in-polyhedron set containment $C \subseteq P$.

Theorem 1. *It holds that $p^\star \geq 0$ if and only if $C \subseteq P$.*

Proof. The condition that $p^\star \geq 0$ is equivalent to the condition that

$$\inf_{(x,y,z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n} \left\{ y^\top b - z^\top x : y \geq 0,\ x \in C,\ A^\top y = z \right\} \geq 0,$$

which in turn is equivalent to the condition that

$$\inf_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \left\{ y^\top b - z^\top x : y \geq 0,\ x \in C,\ A^\top y = z \right\} \geq 0 \text{ for all } z \in \mathbb{R}^n.$$

This holds if and only if

$$\sup_{x \in \mathbb{R}^n} \left\{ z^\top x : x \in C \right\} \leq \inf_{y \in \mathbb{R}^m} \left\{ y^\top b : y \geq 0,\ A^\top y = z \right\} \text{ for all } z \in \mathbb{R}^n. \tag{2}$$

The infimum on the right-hand side of the above inequality is a linear program whose dual problem is

$$\sup_{\lambda \in \mathbb{R}^n} \left\{ z^\top \lambda : A\lambda \leq b \right\} = \sup_{\lambda \in \mathbb{R}^n} \left\{ z^\top \lambda : \lambda \in P \right\}.$$

Since P is nonempty, this dual problem is feasible, and hence strong duality holds. Thus, (2) holds if and only if

$$\sup_{x \in \mathbb{R}^n} \left\{ z^\top x : x \in C \right\} \leq \sup_{\lambda \in \mathbb{R}^n} \left\{ z^\top \lambda : \lambda \in P \right\} \text{ for all } z \in \mathbb{R}^n.$$

Since C and P are both closed and convex, Lemma 9 ensures that this condition holds if and only if $C \subseteq P$. $\square$

In light of Theorem 1, we call the optimization $p^\star$ a *C-in-P bilinear program.* We can also characterize C-in-P bilinear positivity geometrically in $\mathbb{R}^m$, as opposed to in $\mathbb{R}^n$:

Proposition 1. *$p^\star \geq 0$ if and only if $-AC + \{b\} = \{b - Ax \in \mathbb{R}^m : x \in C\}$ is contained in $\mathbb{R}^m_+$.*

Proof. It is clear that $\{b - Ax \in \mathbb{R}^m : x \in C\} \subseteq \mathbb{R}^m_+$ if and only if $b - Ax \geq 0$ for all $x \in C$, which holds if and only if $C \subseteq P$. Thus, applying Theorem 1 concludes the proof. □

1.1 C-in-P Positivity Certificates via Convex Optimization

One major benefit of characterizing bilinear positivity certification as a convex-in-polyhedron set containment is that certificates may be computed by solving a finite number of convex optimization problems. Furthermore, such convex optimization-based certificates are tight, despite $p^\star$ itself being nonconvex:

Corollary 1. *Let $a_i \in \mathbb{R}^n$ denote the ith row of A for all $i \in \{1, \dots, m\}$. It holds that $p^\star \geq 0$ if and only if*

$$\sup_{x \in C} a_i^\top x \leq b_i \text{ for all } i \in \{1, \dots, m\}. \tag{3}$$

Proof. Theorem 1 gives that $p^\star \geq 0$ if and only if $C \subseteq P$, which holds if and only if

$$a_i^\top x \leq b_i \text{ for all } x \in C \text{ and all } i \in \{1, \dots, m\},$$

which is equivalent to (3). □

We can also dualize the above convex optimization problems to compute positivity certificates by solving a single convex feasibility problem.

Corollary 2. *Assume that $C = \{x \in \mathbb{R}^n : f_j(x) \leq 0 \text{ for all } j \in \{1, \dots, p\}\}$, with $p \in \mathbb{N}$ and every f_j being a real-valued convex function defined on $\mathbb{R}^n$. Let $a_i \in \mathbb{R}^n$ denote the ith row of A for all $i \in \{1, \dots, m\}$. For all $i \in \{1, \dots, m\}$, define the Lagrange dual function $g_i \colon \mathbb{R}^p \to \mathbb{R} \cup \{-\infty\}$ by*

$$g_i(\lambda) = \inf_{x \in \mathbb{R}^n} \left(-a_i^\top x + \sum_{j=1}^p \lambda_j f_j(x) \right). \tag{4}$$

If the convex optimization problem

$$\inf_{(\lambda_1, \dots, \lambda_m) \in (\mathbb{R}^p)^m} \{0 : g_i(\lambda_i) + b_i \geq 0 \text{ and } \lambda_i \geq 0 \text{ for all } i \in \{1, \dots, m\}\} \tag{5}$$

is feasible, then $p^\star \geq 0$. Furthermore, if the functions $f_1, \dots, f_p$ satisfy Slater's condition, then $p^\star \geq 0$ if and only if (5) is feasible.

Proof. We first remark that (5) is indeed a convex optimization problem, as every g_i is a concave function, since it is the pointwise infimum of a family of affine functions.

Now, denote the left-hand side of (3) by

$$\mu_i^\star := \sup_{x \in C} a_i^\top x.$$

It holds that

$$\begin{aligned}
\mu_i^\star &= -\inf_{x \in C} \left(-a_i^\top x\right) \\
&= -\inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^p} \left\{ -a_i^\top x + \sum_{j=1}^p \lambda_j f_j(x) : \lambda \geq 0 \right\} \\
&\leq -\sup_{\lambda \in \mathbb{R}^p} \inf_{x \in \mathbb{R}^n} \left\{ -a_i^\top x + \sum_{j=1}^p \lambda_j f_j(x) : \lambda \geq 0 \right\} \\
&= -\sup_{\lambda \in \mathbb{R}^p} \left\{ g_i(\lambda) : \lambda \geq 0 \right\} \\
&= \inf_{\lambda \in \mathbb{R}^p} \left\{ -g_i(\lambda) : \lambda \geq 0 \right\}.
\end{aligned} \tag{6}$$

Thus, if (5) is feasible, then there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}^p$ such that $g_i(\lambda_i) + b_i \geq 0$ and $\lambda_i \geq 0$ for all $i \in \{1, \ldots, m\}$, implying that

$$\mu_i^\star \leq \inf_{\lambda \in \mathbb{R}^p} \{-g_i(\lambda) : \lambda \geq 0\} \leq b_i \text{ for all } i \in \{1, \ldots, m\},$$

and hence $p^\star \geq 0$ by Corollary 1.

If Slater's condition holds for the functions $f_1, \ldots, f_p$, then the inequality (6) becomes an equality. In this case, Corollary 1 gives that $p^\star \geq 0$ if and only if $\mu_i^\star \leq b_i$ for all $i \in \{1, \ldots, m\}$, which holds if and only if there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}^p$ such that $\lambda_i \geq 0$ and $-g_i(\lambda_i) \leq b_i$ for all $i \in \{1, \ldots, m\}$, which is equivalent to (5) being feasible. □

Remark 1. Corollary 2 provides a method for reformulating robust optimization problems with constraints defined by C-in-P bilinear programs into more tractable forms. Specifically, consider a robust optimization problem taking the form

$$\inf_{z \in Z} \left\{ f(z) : y^\top (b(z) - A(z)x) \geq 0 \text{ for all } y \in \mathbb{R}_+^m \text{ and all } x \in C \right\},$$

with Z being some constraint set and $f(z) \in \mathbb{R}$, $A(z) \in \mathbb{R}^{m \times n}$, and $b(z) \in \mathbb{R}^m$ depending on the optimization variable z. Such a problem has an infinite number of nonconvex constraints, and is therefore generally intractable as formulated. Such a problem is equivalent to

$$\inf_{(z, \lambda_1, \ldots, \lambda_m) \in Z \times (\mathbb{R}^p)^m} \{f(z) : g_i(z, \lambda_i) + b_i(z) \geq 0 \text{ and } \lambda_i \geq 0 \text{ for all } i \in \{1, \ldots, m\}\},$$

where g_i is defined as in (4), which now depends on z through $a_i(z)$ (the ith row of $A(z)$). This reformulation is a finite-dimensional convex optimization problem if Z is a convex subset of some Euclidean space $\mathbb{R}^q$, $f \colon z \mapsto f(z)$ is convex, $A \colon z \mapsto A(z)$ is affine, and every $b_i \colon z \mapsto b_i(z)$ is concave. This problem also serves as an equivalent reformulation of the optimization problem $\inf_{z \in Z} \{f(z) : C \subseteq \{x \in \mathbb{R}^n : A(z)x \leq b(z)\}\}$ with a convex-in-polyhedron set containment constraint, per Theorem 1. Set containment constraints appear frequently in applications, such as reachability constraints in control theory (Sadraddini and Tedrake, 2019).

We emphasize the computational advantage of our convex optimization-based positivity certificates (Corollary 1 and Corollary 2) over the common sum of squares decompositions for more general polynomials, which typically amount to solving impractically large semidefinite programs (Lasserre, 2009, Chapter 2). In particular, our convex feasibility problem (5) is capable of avoiding semidefinite cone constraints, which can be prohibitively expensive in large-scale settings (Ahmadi and Majumdar, 2019). As we will see in the specific examples below, our certificates often reduce to highly scalable problems such as linear programs, convex quadratically constrained quadratic programs, and second-order cone programs.

1.1.1 Examples

The following examples show that our convex optimization-based positivity certificates naturally adapt their complexity to the underlying problem instance. For example, when C is a polyhedron, our positivity certificates reduce to linear programs, when C is an intersection of ellipsoids, our certificates reduce to either a convex quadratically constrained quadratic program or a second-order cone program, and when C is a spectrahedron (defined by a linear matrix inequality), our certificates reduce to semidefinite programs. This natural adaptation is in contrast to the powerful yet computationally expensive "one-size-fits-all" approach of sum of squares decompositions and large-scale semidefinite programming.

Positivity over Polyhedra. Consider $C = \{x \in \mathbb{R}^n : Px \leq q\}$ with $P \in \mathbb{R}^{p\times n}$ and $q \in \mathbb{R}^p$. In this case, Theorem 1 shows that the C-in-P bilinear positivity certificate amounts to the classical, well-studied, and efficiently solvable H-in-H polyhedral containment problem, i.e., the problem of verifying whether one polyhedron in H-representation is contained in another in H-representation (Freund and Orlin, 1985; Gritzmann and Klee, 1994). Instantiating Corollary 1 results in the linear programming-based certificate that $p^\star \geq 0$ if and only if

$$\sup_{x\in\mathbb{R}^n} \left\{a_i^\top x : Px \leq q\right\} \leq b_i \text{ for all } i \in \{1, \dots, m\},$$

which, by Corollary 2 is equivalent to the linear program

$$\inf_{(\lambda_1,\dots,\lambda_m)\in(\mathbb{R}^p)^m} \left\{0 : q^\top \lambda_i \leq b_i,\ P^\top \lambda_i = a_i, \text{ and } \lambda_i \geq 0 \text{ for all } i \in \{1, \dots, m\}\right\}$$

being feasible. In other words, we have *exactly* reduced the nonconvex C-in-P bilinear positivity certification problem to a linear program in the case that C is a polyhedron in H-representation.

Positivity over Ellipsoids Consider

$$C = \{x \in \mathbb{R}^n : x^\top P_j x + q_j^\top x + r_j \leq 0 \text{ for all } j \in \{1, \dots, p\}\}$$

with $r_1, \dots, r_p \in \mathbb{R}$, $q_1, \dots, q_p \in \mathbb{R}^n$, and every $P_j \in \mathbb{R}^{n\times n}$ a positive semidefinite matrix. Instantiating Corollary 1 results in the quadratically constrained quadratic programming-based certificate that $p^\star \geq 0$ if and only if

$$\sup_{x\in\mathbb{R}^n} \left\{a_i^\top x : x^\top P_j x + q_j^\top x + r_j \leq 0 \text{ for all } j \in \{1, \dots, p\}\right\} \leq b_i \text{ for all } i \in \{1, \dots, m\}.$$

The associated Lagrange dual function (4) is given by

$$g_i(\lambda) = \sum_{j=1}^{p} \lambda_j r_j - \frac{1}{4}\left(\sum_{j=1}^{p} \lambda_j q_j - a_i\right)^{\top} \left(\sum_{j=1}^{p} \lambda_j P_j\right)^{+} \left(\sum_{j=1}^{p} \lambda_j q_j - a_i\right)$$

if $\sum_{j=1}^{p} \lambda_j P_j \succeq 0$ and $a_i - \sum_{j=1}^{p} \lambda_j q_j \in \text{Range}\left(\sum_{j=1}^{p} \lambda_j P_j\right)$, and $g_i(\lambda) = -\infty$ otherwise. Notice that the linear matrix inequality $\sum_{j=1}^{p} \lambda_j P_j \succeq 0$ is implied by the condition that $\lambda \geq 0$, and hence these semidefinite cone constraints can be dropped from the convex feasibility problem (5). If $a_i \neq 0$ for all $i \in \{1, \dots, m\}$ and $P_j \succ 0$ for all $j \in \{1, \dots, p\}$, then $(\lambda_1, \dots, \lambda_m) \in (\mathbb{R}^p)^m$ is infeasible for (5) whenever $\lambda_i = 0$ for some $i \in \{1, \dots, m\}$, and hence, for every $i \in \{1, \dots, m\}$, it holds that $\sum_{j=1}^{p} (\lambda_i)_j P_j$ is positive definite and invertible for all feasible $(\lambda_1, \dots, \lambda_m)$. Thus, the range constraint $a_i - \sum_{j=1}^{p} \lambda_j q_j \in \text{Range}(\sum_{j=1}^{p} \lambda_j P_j)$ can be replaced by the condition that $\lambda_i \neq 0$, reducing the feasibility problem (5) to

$$\begin{aligned} \inf_{(\lambda_1,\dots,\lambda_m)\in(\mathbb{R}^p)^m} \Bigg\{0 :& \\ &\sum_{j=1}^{p} (\lambda_i)_j r_j - \frac{1}{4}\left(\sum_{j=1}^{p} (\lambda_i)_j q_j - a_i\right)^{\top} \left(\sum_{j=1}^{p} (\lambda_i)_j P_j\right)^{-1} \left(\sum_{j=1}^{p} (\lambda_i)_j q_j - a_i\right) + b_i \geq 0, \\ &\qquad \lambda_i \geq 0,\ \lambda_i \neq 0 \text{ for all } i \in \{1, \dots, m\}\Bigg\}. \end{aligned}$$

This problem is indeed convex, as the left-hand side of every nonlinear constraint is concave (see Example 3.4 in Boyd and Vandenberghe (2004) on matrix fractional functions), and every set $\{\lambda_i \in \mathbb{R}^p : \lambda_i \geq 0,\ \lambda_i \neq 0\} = \mathbb{R}^p_+ \setminus \{0\}$ is convex. In fact, this is a second-order cone program (neglecting the $\lambda_i \neq 0$ constraints), as it can be rewritten into the following form using the techniques for reformulating matrix fractional programs as second-order cone programs from Lobo et al. (1998):

$$\begin{aligned} \inf_{\substack{(\lambda_1,\dots,\lambda_m)\in(\mathbb{R}^p)^m, \\ \{y_{ij} : i\in\{1,\dots,m\},\ j\in\{1,\dots,p\}\}\subseteq\mathbb{R}^n \\ T\in\mathbb{R}^{m\times p}}} &\Bigg\{0 : \frac{1}{4}\sum_{j=1}^{p} T_{ij} \leq \sum_{j=1}^{p} (\lambda_i)_j r_j + b_i, \\ &\sum_{j=1}^{p} P_j^{\frac{1}{2}} y_{ij} = \sum_{j=1}^{p} (\lambda_i)_j q_j - a_i,\ \left\| \begin{bmatrix} 2y_{ij} \\ T_{ij} - (\lambda_i)_j \end{bmatrix} \right\|_2 \leq T_{ij} + (\lambda_i)_j, \\ &\qquad T_{ij} \geq 0,\ \lambda_i \geq 0,\ \lambda_i \neq 0 \text{ for all } i \in \{1, \dots, m\} \text{ and all } j \in \{1, \dots, p\}\Bigg\}. \end{aligned}$$

Per Corollary 2, feasibility of the above second-order cone problem implies that $p^\star \geq 0$. Furthermore, if there exists $x \in C$ such that $x^\top P_j x + q_j^\top x + r_j < 0$ for all $j \in \{1, \dots, p\}$ such that $P_j \neq 0$, then $p^\star \geq 0$ if and only if the second-order cone problem is feasible.

Positivity over Spectrahedra Consider $C = \{x \in \mathbb{R}^n : P_0 + \sum_{i=1}^{n} x_i P_i \preceq 0\}$ with $P_0, \dots, P_n \in \mathbb{R}^{k\times k}$ symmetric (but possibly indefinite). The convex set C is a spectrahedron. The geometry of spectrahedra has been studied in depth (Ramana and Goldman, 1995; Gouveia and Netzer, 2011), as well as their intimate relationships to semidefinite programming, global polynomial optimization, and convex algebraic geometry (Vandenberghe and Boyd, 1996; Lasserre, 2001; Blekherman et al., 2012). Theorem 1 shows that the C-in-P bilinear positivity certificate amounts to verifying that this spectrahedron is contained in a polyhedron. Prior works have studied such spectrahedron-in-polyhedron containment problems, which we discuss in further detail below.

Instantiating Corollary 1 results in the semidefinite programming-based certificate that $p^\star \geq 0$ if and only if

$$\sup_{x\in\mathbb{R}^n} \left\{ a_i^\top x : P_0 + \sum_{i=1}^n x_i P_i \preceq 0 \right\} \leq b_i \text{ for all } i \in \{1,\dots,m\}.$$

Although C does not quite take the form $\{x \in \mathbb{R}^n : g_j(x) \leq 0 \text{ for all } j \in \{1,\dots,p\}\}$ as it is instead defined by a generalized inequality with respect to the positive semidefinite cone, it is easy to see how to generalize Corollary 2 to this case. In particular, the convex feasibility problem (5) becomes

$$\begin{aligned} \inf_{(\Lambda_1,\dots,\Lambda_m)\in(\mathbb{S}^k)^m} \big\{ & 0 : \operatorname{tr}(\Lambda_i P_0) + b_i \geq 0, \;\; \operatorname{tr}(\Lambda_i P_j) = A_{ij}, \\ & \Lambda_i \succeq 0 \text{ for all } i \in \{1,\dots,m\} \text{ and all } j \in \{1,\dots,n\} \big\}, \end{aligned} \tag{7}$$

where $\mathbb{S}^k$ denotes the set of $k \times k$ symmetric matrices with real elements. This is also a semidefinite program. Similar to the case of positivity over the intersection of ellipsoids, feasibility of the above semidefinite program implies that $p^\star \geq 0$, and if $P_0 + \sum_{i=1}^n x_i P_i \prec 0$ for some $x \in \mathbb{R}^n$, then feasibility also becomes necessary for $p^\star \geq 0$.

Kellner (2015) studied general spectrahedron-in-spectrahedron containment problems. Consider the spectrahedra $S_1 = \{x \in \mathbb{R}^n : Q_0 + \sum_{i=1}^n x_i Q_i \succeq 0\}$ and $S_2 = \{x \in \mathbb{R}^n : R_0 + \sum_{i=1}^n x_i R_i \succeq 0\}$, with all $Q_0,\dots,Q_n \in \mathbb{R}^{k\times k}$ and all $R_0,\dots,R_n \in \mathbb{R}^{l\times l}$ symmetric. Then, the main result of Kellner (2015), namely, Theorem 4.3, shows that $S_1 \subseteq S_2$ if there exists a symmetric matrix $M \in \mathbb{R}^{kl\times kl}$ taking the block form

$$M = \begin{bmatrix} M_{11} & \cdots & M_{1k} \\ \vdots & \ddots & \vdots \\ M_{k1} & \cdots & M_{kk} \end{bmatrix}$$

with every $M_{ij} \in \mathbb{R}^{l\times l}$, such that

$$M \succeq 0, \;\; R_0 - \sum_{i,j=1}^k (Q_0)_{ij} M_{ij} \succeq 0, \;\text{ and } R_p = \sum_{i,j=1}^k (Q_p)_{ij} M_{ij} \text{ for all } p \in \{1,\dots,n\}. \tag{8}$$

This is a sufficient condition for containment that can be verified by solving a semidefinite feasibility problem. In particular, this approach may also be used to verify that $C = \{x \in \mathbb{R}^n : P_0 + \sum_{i=1}^n x_i P_i \preceq 0\} \subseteq \{x \in \mathbb{R}^n : Ax \leq b\} = P$ and hence to compute a C-in-P bilinear positivity certificate $p^\star \geq 0$, since all polyhedra are spectrahedra. This can be seen more explicitly by writing the polyhedron as $P = \{x \in \mathbb{R}^n : \operatorname{diag}(b_1 - a_1^\top x, \dots, b_m - a_m^\top x) \succeq 0\}$ with $a_i \in \mathbb{R}^n$ denoting the ith row of A, in which case $l = m$. However, our semidefinite feasibility problem (7) boasts three notable advantages over (8). First, (8) requires resolving a potentially massive $km \times km$ semidefinite cone constraint, whereas (7) consists of m much smaller and parallelizable $k \times k$ semidefinite cone constraints. Second, the condition (8) is only proven to be necessary and sufficient for spectrahedron-in-polyhedron containment when the spectrahedral representations are monic. In our setting, this monic requirement amounts to $P_0 = -I_k$ and $b = 1_m$ (the m-vector of all ones), which automatically requires both C and P to contain the origin. Thus, (8) might be overconservative for C-in-P bilinear certification in general nonmonic cases. On the other hand, our feasibility criterion in (7) is necessary and sufficient for $C \subseteq P$ with general spectrahedra C and general polyhedra P, so long as $P_0 + \sum_{i=1}^n x_i P_i \prec 0$ for some $x \in \mathbb{R}^n$. Third, our spectrahedron-in-polyhedron containment criterion is theoretically stronger (less conservative) than that of Kellner (2015), as we now show:

Proposition 2. *Let $P_0, \dots, P_n \in \mathbb{R}^{k \times k}$ be symmetric, and consider $Q_0 = -P_0, \dots, Q_n = -P_n$ and $R_0 = \operatorname{diag}(b_1, \dots, b_m), R_1 = -\operatorname{diag}(A_{11}, \dots, A_{m1}), \dots, R_n = -\operatorname{diag}(A_{1n}, \dots, A_{mn})$. If there exists $M \in \mathbb{R}^{km \times km}$ symmetric, such that (8) holds, then $(\Lambda_1, \dots, \Lambda_m)$ with $\Lambda_i \in \mathbb{R}^{k \times k}$ defined by $(\Lambda_i)_{pq} = (M_{pq})_{ii}$ for all $i \in \{1, \dots, m\}$ and all $p, q \in \{1, \dots, k\}$ is feasible for (7).*

Proof. Define $\Lambda_i \in \mathbb{R}^{k \times k}$ by $(\Lambda_i)_{pq} = (M_{pq})_{ii}$ for all $i \in \{1, \dots, m\}$ and all $p, q \in \{1, \dots, k\}$. Let $e_i \in \mathbb{R}^m$ denote the ith standard unit vector in $\mathbb{R}^m$ for all $i \in \{1, \dots, m\}$. Then, since M satisfies (8), it holds that

$$\begin{aligned}
y^\top \Lambda_i y &= \sum_{p,q=1}^k y_p (\Lambda_i)_{pq} y_q \\
&= \sum_{p,q=1}^k y_p e_i^\top M_{pq} e_i y_q \\
&= \begin{bmatrix} y_1 e_i^\top & \cdots & y_k e_i^\top \end{bmatrix} M \begin{bmatrix} y_1 e_i \\ \vdots \\ y_k e_i \end{bmatrix} \\
&\geq 0
\end{aligned}$$

for all $i \in \{1, \dots, m\}$ and all $y \in \mathbb{R}^k$, since $M \succeq 0$. Therefore, $\Lambda_i \succeq 0$ for all i. Furthermore, we find that

$$\operatorname{tr}(\Lambda_i P_j) = \sum_{p,q=1}^k (\Lambda_i)_{pq} (P_j)_{pq} = \sum_{p,q=1}^k (-Q_j)_{pq} (M_{pq})_{ii} = -\left(\sum_{p,q=1}^k (Q_j)_{pq} M_{pq} \right)_{ii} = -(R_j)_{ii} = A_{ij}$$

for all $i \in \{1, \dots, m\}$ and all $j \in \{1, \dots, n\}$, and similarly,

$$\operatorname{tr}(\Lambda_i P_0) + b_i = -\left(\sum_{p,q=1}^k (Q_0)_{pq} M_{pq} \right)_{ii} + (R_0)_{ii} \geq 0$$

for all $i \in \{1, \dots, m\}$, since the diagonal elements of a positive semidefinite matrix are nonnegative. Thus, $(\Lambda_1, \dots, \Lambda_m)$ is feasible for (7). $\square$

Proposition 2 shows that the feasible region defined by (8) is more restrictive than the feasible region defined by our semidefinite feasibility program (7). On the other hand, (8) is applicable to more general containment problems than (7) is, as it is capable of verifying spectrahedron-in-spectrahedron containments, which is not possible using (7) unless the circumbody can be reduced to a polyhedron in H-representation.

1.2 Analysis of C-in-P Bilinear Programs

Recall the C-in-P bilinear program of interest:

$$p^\star = \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} \left\{ y^\top (b - Ax) : y \geq 0,\ x \in C \right\}. \tag{1}$$

In this section, we analyze some of the fundamental properties of such optimization problems by leveraging our geometric correspondences with set containments from Theorem 1.

Notice that the feasible set $C \times \mathbb{R}^m_+$ of (1) is not compact and the objective function is not coercive, and therefore we cannot ensure the existence of minimizers using standard arguments based on Weierstrass' extreme value theorem. As the bilinear objective function in (1) is a nonconvex quadratic function, the classical Frank-Wolfe theorem ensures the existence of a minimizer in the special case that C is a polyhedron (not necessarily compact) and $p^\star$ is finite (Frank and Wolfe, 1956). However, follow-up studies have shown that nonconvex quadratic functions do not in general attain minimizers over unbounded convex constraint sets defined by two or more nonlinear constraints, even when the problem is bounded below (Luo and Zhang, 1999). Despite this, we show that the C-in-P bilinear programs of interest always attain solutions when they are bounded below, and we characterize such solutions:

Lemma 1. *If the C-in-P bilinear program* (1) *is bounded below, then* $C \subseteq P$.

Proof. Suppose that (1) is bounded below. For the sake of contradiction, assume that $C \not\subseteq P$. Then, there exists $x \in C$ such that $a_i^\top x > b_i$ for some $i \in \{1, \dots, m\}$. Therefore, taking $y = \alpha e_i$ with $e_i \in \mathbb{R}^m$ the ith standard unit vector in $\mathbb{R}^m$ and $\alpha > 0$ shows that

$$p^\star \leq \alpha(b_i - a_i^\top x) < 0.$$

Thus,

$$p^\star \leq \lim_{\alpha\to\infty} \alpha(b_i - a_i^\top x) = -\infty.$$

This contradicts the fact that (1) is bounded below. Thus, it must be that $C \subseteq P$. $\square$

Proposition 3. *If the C-in-P bilinear program* (1) *is bounded below, then* $p^\star = 0$ *and* $(x, y) \in C \times \mathbb{R}^m_+$ *is a solution to* (1) *if and only if*

$$y_i = 0 \text{ or } a_i^\top x = b_i \text{ for all } i \in \{1, \dots, m\}. \tag{9}$$

In particular, every point $(x, 0) \in \mathbb{R}^n \times \mathbb{R}^m$ *with* $x \in C$ *is a solution to* (1) *when* $p^\star$ *is finite.*

Proof. Suppose that (1) is bounded below. By Lemma 1, it holds that $C \subseteq P$. By Theorem 1, this implies that $p^\star \geq 0$. We see that $p^\star \leq y^\top(b - Ax) = 0 \leq p^\star$ for all $(x, y) \in C \times \mathbb{R}^m_+$ such that $y = 0$. Hence, $p^\star = 0$. Thus, a point $(x, y) \in C \times \mathbb{R}^m_+$ is a solution to (1) if and only if $y^\top(b - Ax) = 0$. Since $C \subseteq P$, it holds that $Ax \leq b$ and hence $b - Ax \in \mathbb{R}^m_+$. Thus, $y_i(b_i - a^\top x)$ is nonnegative for all $i \in \{1, \dots, m\}$, implying that $y^\top(b - Ax) = 0$ if and only if $y_i(b_i - a_i^\top x) = 0$ for all $i \in \{1, \dots, m\}$, which is further equivalent to the condition (9). It is clear that every point $(x, 0) \in \mathbb{R}^n \times \mathbb{R}^m$ with $x \in C$ satisfies (9) and hence solves (1) when $p^\star$ is finite. □

Proposition 3 shows that either $p^\star$ is attained and has optimal value 0 (in the case that $C \subseteq P$), or $p^\star = -\infty$ (in the case that $C \nsubseteq P$); there are no intermediate optimal values:

Corollary 3. *Either $p^\star = 0$ or $p^\star = -\infty$.*

Proof. This follows immediately from Proposition 3. □

According to Proposition 3, feasibility together with the "slackness" condition (9) are necessary and sufficient for optimality in C-in-P bilinear programs. This is in contrast to general nonconvex optimization problems, where similar first-order conditions (e.g., KKT conditions) are only necessary and may require regularity assumptions such as the linear independence constraint qualification. The slackness condition (9) allows us to further refine our characterization of solutions to (1) in the case of strict containment (in the sense of topological interiors):

Corollary 4. *The following are equivalent:*

1. $-AC + \{b\} = \{b - Ax \in \mathbb{R}^m : x \in C\} \subseteq \mathbb{R}^m_{++}$.
2. *$p^\star$ is finite and the set of solutions to* (1) *exactly equals* $C \times \{0\}$.

Denote the ith row of A by $a_i \in \mathbb{R}^n$. If $b_i > 0$ for all $i \in \{1, \dots, m\}$ such that $a_i = 0$, then the above two conditions are also equivalent to:

3. $C \subseteq \operatorname{int}(P)$.

Proof. Certainly, $-AC + \{b\} \subseteq \mathbb{R}^m_{++}$ if and only if $Ax < b$ for all $x \in C$. If this holds, then $C \subseteq P$ and hence $p^\star \geq 0$ by Theorem 1 and the set of solutions to (1) equals

$$\begin{aligned}\{(x,y) \in C \times \mathbb{R}^m_+ : y_i = 0 \text{ or } a_i^\top x = b_i \text{ for all } i \in \{1, \dots, m\}\} \\ = \{(x,y) \in C \times \mathbb{R}^m_+ : y_i = 0 \text{ for all } i \in \{1, \dots, m\}\} \\ = C \times \{0\}\end{aligned}$$

by Proposition 3. Thus, the first condition implies the second.

On the other hand, suppose that $p^\star$ is finite and the set of solutions to (1) equals $C \times \{0\}$. Since $p^\star$ is finite, it must be that $p^\star = 0$ by Corollary 3 and hence Proposition 1 implies that $-AC + \{b\} \subseteq \mathbb{R}^m_+$. If $-AC + \{b\} \nsubseteq \mathbb{R}^m_{++}$, then there would exist $x \in C$ such that $Ax \leq b$ and $a_i^\top x = b_i$ for some $i \in \{1, \dots, m\}$. In this case, $(x, y) = (x, e_i)$, with e_i the ith standard unit vector in $\mathbb{R}^m$, would be optimal for (1) as it is an element of $C \times \mathbb{R}^m_+$ and hence is feasible, and it satisfies (9). However, $(x, e_i) \notin C \times \{0\}$, which is the set of all solutions to (1). Thus, it must be the case that $-AC + \{b\} \subseteq \mathbb{R}^m_{++}$. Therefore, the second condition implies the first.

Now, suppose that $b_i > 0$ for all $i \in \{1, \dots, m\}$ such that $a_i = 0$. Then, by Lemma 6, it holds that $\text{int}(P) = \{x \in \mathbb{R}^n : Ax < b\}$. Hence, $C \subseteq \text{int}(P)$ if and only if $Ax < b$ for all $x \in C$, which holds if and only if $-AC + \{b\} \subseteq \mathbb{R}^m_{++}$. Thus, the first and third conditions are equivalent in this case. □

Next, we show that, despite the nonconvexity of C-in-P bilinear programs, they have no spurious local minima when the containment $C \subseteq P$ holds. That is, for this class of nonconvex optimization problems, every local minimizer is a global minimizer when the optimization is bounded below, and hence algorithms capable of finding local minima are sufficient for global optimization in this case.

Theorem 2. *Let $a_i \in \mathbb{R}^n$ denote the ith row of A. The following both hold:*

1. *Every local minimizer $(x, y) \in C \times \mathbb{R}^m_+$ of the C-in-P bilinear program* (1) *satisfies* (9)*, and hence has objective value $y^\top(b - Ax) = 0$.*

2. *If $C \subseteq P$, then every local minimizer of the C-in-P bilinear program* (1) *is a global minimizer.*

Proof. Let $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ be a local minimizer of (1). Then, it holds that $(x, y) \in C \times \mathbb{R}^m_+$, and there exists $\epsilon > 0$ such that

$$y^\top(b - Ax) \leq \tilde{y}^\top(b - A\tilde{x}) \tag{10}$$

for all $(\tilde{x}, \tilde{y}) \in B((x, y), \epsilon) \cap (C \times \mathbb{R}^m_+)$, where $B((x, y), \epsilon) = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^m : \|(\tilde{x}, \tilde{y}) - (x, y)\|_2 < \epsilon\}$.

Assume for the sake of contradiction that there exists $i \in \{1, \dots, m\}$ such that $y_i > 0$ and $b_i - a_i^\top x \neq 0$. Suppose that $b_i - a_i^\top x > 0$. Then, let $\tilde{x} = x$ and define $\tilde{y}$ by $\tilde{y}_j = y_j$ for all $j \neq i$ and $\tilde{y}_i = y_i - \mu$ for some $\mu \in (0, \min\{\epsilon, y_i\})$. Then, it holds that $\|(\tilde{x}, \tilde{y}) - (x, y)\|_2 = \mu < \epsilon$, so $(\tilde{x}, \tilde{y}) \in B((x, y), \epsilon)$. Furthermore, it holds by construction that $(\tilde{x}, \tilde{y}) \in C \times \mathbb{R}^m_+$. Hence, (10) holds. However, this contradicts the fact that

$$\tilde{y}^\top(b - A\tilde{x}) = y^\top(b - Ax) - \mu(b_i - a_i^\top x) < y^\top(b - Ax)$$

by our construction of $(\tilde{x}, \tilde{y})$. On the other hand, if $b_i - a_i^\top x < 0$, then we can take $\tilde{x} = x$, $\tilde{y}_j = y_j$ for all $j \neq i$, and $\tilde{y}_i = y_i + \mu$ for some $\mu \in (0, \epsilon)$ to arrive at the same contradiction that $(\tilde{x}, \tilde{y}) \in B((x, y), \epsilon) \cap (C \times \mathbb{R}^m_+)$ and $\tilde{y}^\top(b - A\tilde{x}) < y^\top(b - Ax)$. Therefore, it must be the case that $y_i = 0$ or $b_i - a_i^\top x = 0$ for all $i \in \{1, \dots, m\}$. That is, (9) must hold for the local minimizer (x, y). Therefore, if $C \subseteq P$, then (x, y) is a global minimizer of (1) by Theorem 1 and Proposition 3. □

We now characterize the local optima of the C-in-P bilinear program in the case that the containment $C \subseteq P$ may *not* hold.

Theorem 3. *Let $a_i \in \mathbb{R}^n$ denote the ith row of A, and let e_i denote the ith standard unit vector in $\mathbb{R}^m$. Every local minimizer $(x, y) \in C \times \mathbb{R}^m_+$ of* (1) *satisfies all of the following three conditions:*

1. *(x, y) satisfies* (9) *for all $i \in \{1, \dots, m\}$.*

2. $x \in P$*.*

3. For all $i \in \{1, \dots, m\}$*, at least one of the following holds:*

(a) $a_i^\top x < b_i$*.*

(b) $A^{-1}(\{e_i\}) \cap \bigcup_{\mu>0} \frac{1}{\mu}(\{-x\} + C) = \emptyset$*, where* $A^{-1}(\{e_i\}) = \{v \in \mathbb{R}^n : Av = e_i\}$*.*

When $A \neq 0$*, a point* $(x, y) \in C \times \mathbb{R}^m_+$ *is a local minimizer of* (1)*, if, in addition to the above three conditions, it satisfies the following:*

3'. For all $i \in \{1, \dots, m\}$*, at least one of the following holds:*

(a) $a_i^\top x < b_i$*.*

(b) $(\operatorname{span}\{a_i\}^\perp)^c \cap \bigcup_{\mu>0} \frac{1}{\mu}(\{-x\} + C) = \emptyset$*, where* $\operatorname{span}\{a_i\}^\perp = \{v \in \mathbb{R}^n : a_i^\top v = 0\}$*.*

Proof. We start by proving the proposed necessary conditions.

Necessary conditions for local minimizers. Suppose that $(x, y) \in C \times \mathbb{R}^m_+$ is a local minimizer of (1), so that there exists $\epsilon > 0$ such that (10) holds for all $(\tilde{x}, \tilde{y}) \in B((x, y), \epsilon) \cap (C \times \mathbb{R}^m_+)$, where $B((x, y), \epsilon) = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^m : \|(\tilde{x}, \tilde{y}) - (x, y)\|_2 < \epsilon\}$. Then, by Theorem 2, it holds that (x, y) satisfies (9) for all $i \in \{1, \dots, m\}$.

For the sake of contradiction, suppose that $x \notin P$. Then, there exists $i \in \{1, \dots, m\}$ such that $b_i - a_i^\top x < 0$. Let $\tilde{x} = x$ and define $\tilde{y}$ by $\tilde{y}_j = y_j$ for all $j \neq i$ and $\tilde{y}_i = y_i + \mu$ for some $\mu \in (0, \epsilon)$. Then, it holds that $\|(\tilde{x}, \tilde{y}) - (x, y)\|_2 = \mu < \epsilon$, so $(\tilde{x}, \tilde{y}) \in B((x, y), \epsilon)$. Furthermore, it holds by construction that $(\tilde{x}, \tilde{y}) \in C \times \mathbb{R}^m_+$. Hence, (10) must hold for $(\tilde{x}, \tilde{y})$. However, this contradicts the fact that

$$\tilde{y}^\top (b - A\tilde{x}) = y^\top (b - Ax) + \mu(b_i - a_i^\top x) < y^\top (b - Ax)$$

by our construction of $(\tilde{x}, \tilde{y})$. Therefore, it must be the case that $x \in P$.

Finally, suppose for the sake of contradiction that there exists $i \in \{1, \dots, m\}$ such that $a_i^\top x \geq b_i$ and $A^{-1}(\{e_i\}) \cap \bigcup_{\mu>0} \frac{1}{\mu}(\{-x\} + C) \neq \emptyset$. Then, there exists $v \in \mathbb{R}^n \setminus \{0\}$ and $\mu > 0$ such that $Av = e_i$ and $x + \mu v \in C$. Let

$$\tilde{x} = x + \mu' v$$

with $\mu' \in \left(0, \min\{\mu, \frac{\epsilon}{\sqrt{2}\|v\|_2}\}\right)$. Furthermore, define $\tilde{y}$ by $\tilde{y}_j = y_j$ for all $j \neq i$ and $\tilde{y}_i = y_i + \mu''$ with $\mu'' \in \left(0, \frac{\epsilon}{\sqrt{2}}\right)$. Then, it holds that

$$\begin{aligned}
\|(\tilde{x}, \tilde{y}) - (x, y)\|_2^2 &= \|\tilde{x} - x\|_2^2 + \|\tilde{y} - y\|_2^2 \\
&= (\mu')^2 \|v\|_2^2 + (\mu'')^2 \\
&< \epsilon^2,
\end{aligned}$$

and hence $(\tilde{x}, \tilde{y}) \in B((x,y),\epsilon))$. Furthermore, since $0 < \mu' < \mu$, we have that $\mu' = \mu(1-\theta)$ for some $\theta \in (0,1)$, and therefore

$$\begin{aligned}\tilde{x} &= x + \mu' v \\ &= x + \mu(1-\theta)v \\ &= \theta x + (1-\theta)(x + \mu v) \\ &\in C\end{aligned}$$

by convexity of C. Therefore, it holds that $(\tilde{x}, \tilde{y}) \in B((x,y),\epsilon) \cap (C \times \mathbb{R}^m_+)$. Hence, (10) must hold for $(\tilde{x}, \tilde{y})$. However, this contradicts the fact that

$$\begin{aligned}\tilde{y}^\top(b - A\tilde{x}) &= \tilde{y}^\top(b - Ax - \mu' Av) \\ &= y^\top(b - Ax - \mu' e_i) + \mu''(b_i - a_i^\top x - \mu') \\ &= y^\top(b - Ax) - \mu' y_i + \mu''(b_i - a_i^\top x) - \mu'\mu'' \\ &< y^\top(b - Ax),\end{aligned}$$

where the inequality follows from the fact that $y_i \geq 0$ and $a_i^\top x = b_i$, as we know that x satisfies both $a_i^\top x \geq b_i$ by assumption and $x \in P$ by our analysis above. Therefore, it must be the case that $a_i^\top x < b_i$ or $A^{-1}(\{e_i\}) \cap \bigcup_{\mu>0} \frac{1}{\mu}(\{-x\} + C) = \emptyset$, for all $i \in \{1,\dots,m\}$.

We now move on to proving the proposed sufficient conditions.

Sufficient conditions for local minimizers. Assume that $A \neq 0$. Suppose that $(x,y) \in C \times \mathbb{R}^m_+$ is such that (x,y) satisfies (9) for all $i \in \{1,\dots,m\}$, $x \in P$, and, for all $i \in \{1,\dots,m\}$, it holds that $a_i^\top x < b_i$ or $(\operatorname{span}\{a_i\}^\perp)^c \cap \bigcup_{\mu>0} \frac{1}{\mu}(\{-x\} + C) = \emptyset$. Then, by (9), we conclude that

$$y^\top(b - Ax) = 0.$$

Since $A \neq 0$, we have that

$$\max\{\|a_1\|_2, \dots, \|a_m\|_2\} > 0.$$

Define $I(x) = \left\{i \in \{1,\dots,m\} : a_i^\top x \neq b_i\right\}$, which may be empty. Clearly, it holds that $\inf\{b_i - a_i^\top x : i \in I(x)\} > 0$. Define

$$\epsilon = \frac{\inf\{b_i - a_i^\top x : i \in I(x)\}}{\max\{\|a_1\|_2, \dots, \|a_m\|_2\}} \in (0, \infty].$$

Let $(\tilde{x}, \tilde{y}) \in B((x,y),\epsilon) \cap (C \times \mathbb{R}^m_+)$, where $B((x,y),\epsilon) = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^m : \|(\tilde{x}, \tilde{y}) - (x,y)\|_2 < \epsilon\}$. Consider an arbitrary index $i \in \{1,\dots,m\}$. If $i \in I(x)$, then $\epsilon < \infty$ and

$$\begin{aligned}b_i - a_i^\top \tilde{x} &= b_i - a_i^\top x + a_i^\top(x - \tilde{x}) \\ &\geq b_i - a_i^\top x - \|a_i\|_2 \|x - \tilde{x}\|_2 \\ &\geq b_i - a_i^\top x - \max\{\|a_1\|_2, \dots, \|a_m\|_2\}\|x - \tilde{x}\|_2 \\ &> b_i - a_i^\top x - \max\{\|a_1\|_2, \dots, \|a_m\|_2\}\epsilon \\ &= b_i - a_i^\top x - \min\{b_j - a_j^\top x : j \in I(x)\} \\ &\geq 0.\end{aligned}$$

On the other hand, if $i \notin I(x)$, then $a_i^\top x = b_i$. By our assumption, this implies that $(\operatorname{span}\{a_i\}^\perp)^c \cap \bigcup_{\mu>0} \frac{1}{\mu}(\{-x\} + C) = \emptyset$. Therefore, for all $v \in \mathbb{R}^n$, it must be the case that $a_i^\top v = 0$ or $x+\mu v \notin C$ for all $\mu > 0$. Let $\mu > 0$ be arbitrary. Since $\tilde{x} = x+\mu\left(\frac{1}{\mu}(\tilde{x}-x)\right) \in C$, we find that $\frac{1}{\mu}a_i^\top(\tilde{x}-x) = 0$, and hence that $a_i^\top \tilde{x} = a_i^\top x = b_i$. Therefore, we conclude that $b_i - a_i^\top \tilde{x} = 0$ in this case. Thus, it holds that $b_i - a_i^\top \tilde{x} \geq 0$ for all $i \in \{1, \dots, m\}$, implying that

$$\tilde{y}^\top(b - A\tilde{x}) \geq 0 = y^\top(b - Ax).$$

Since this holds for arbitrary $(\tilde{x}, \tilde{y}) \in B((x,y),\epsilon) \cap (C \times \mathbb{R}_+^m)$, we conclude that (x,y) is a local minimizer of (1). □

Corollary 5. *Let $a_i \in \mathbb{R}^n$ denote the ith row of A, and assume that $C \not\subseteq P$. The following both hold:*

1. *Every point in $\{(x,y) \in C \times \{0\} : Ax < b\}$ is a local minimizer of (1).*
2. *No point in $\bigcup_{i=1}^m \{(x,y) \in C \times \mathbb{R}_+^m : b_i - a_i^\top x < 0\}$ is a local minimizer of (1).*

Proof. It is easy to see that a point $(x,y) \in C \times \{0\}$ such that $Ax < b$ satisfies the sufficient conditions of Theorem 3, and hence is a local minimizer of (1). On the other hand, a point $(x,y) \in C \times \mathbb{R}_+^m$ such that $b_i - a_i^\top x < 0$ for some $i \in \{1, \dots, m\}$ is not a local minimizer of (1), as it violates the second necessary condition of Theorem 3 (namely, that $x \in P$). □

Corollary 5 shows that, unlike the case where $C \subseteq P$, the C-in-P bilinear program (1) may have local minima that are not global minima in the case where $C \not\subseteq P$, and in such a case, there are no global minima since $p^\star = -\infty$.

Example 1. As a concrete example, consider $C = \{x \in \mathbb{R}^2 : \|x\|_2 \leq \sqrt{2}\}$ and

$$A = \begin{bmatrix} I_2 \\ -I_2 \end{bmatrix}, \qquad b = 1_4.$$

The polyhedron $P = \{x \in \mathbb{R}^2 : Ax \leq b\}$ equals the box $\{x \in \mathbb{R}^2 : \|x\|_\infty \leq 1\}$, which is contained in the scaled ball C. Hence, $C \not\subseteq P$, and therefore $p^\star = -\infty$. However, Corollary 5 shows that every pair $(x,0) \in \mathbb{R}^2 \times \mathbb{R}^4$ with x in the interior of the box P is a local minimizer of (1), and that no pair $(x,y) \in \mathbb{R}^2 \times \mathbb{R}^4$ with $x \in C \setminus P$ is a local minimizer of (1). It is easy to see that every point on the perimeter of the box P is arbitrarily close to some feasible point in C that lies outside of P. At such points in $C \cap P^c$, the C-in-P bilinear program's objective value can be made arbitrarily small by increasing some element of y. Thus, the points on the perimeter of the box are not local minima either.

It is easy to see that the objective function's Hessian has a strictly negative eigenvalue everywhere, but this result alone does not guarantee the strictly negative curvature along feasible directions :

Proposition 4. *Denote the objective function of (1) by $f \colon (x,y) \mapsto y^\top(b - Ax)$. If $A \neq 0$, then the Hessian $\nabla^2 f(x,y)$ has a strictly negative eigenvalue at every point $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m$.*

Proof. Suppose that $A \neq 0$. Let $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. The Hessian is given by

$$\nabla^2 f(x, y) = \begin{bmatrix} 0 & -A^\top \\ -A & 0 \end{bmatrix}.$$

Since $A \neq 0$, it has a nonzero singular value $\sigma > 0$ such that $\sigma = \sqrt{\lambda}$, where $\lambda > 0$ is an eigenvalue of $A^\top A$. It holds that $-\sigma < 0$ is an eigenvalue of $\nabla^2 f(x, y)$, since

$$\begin{aligned} \det(-\sigma I_{n+m} - \nabla^2 f(x, y)) &= \det \begin{bmatrix} -\sigma I_n & A^\top \\ A & -\sigma I_m \end{bmatrix} \\ &= \det(-\sigma I_m) \det(-\sigma I_n - A^\top(-\sigma I_m)^{-1} A) \\ &= (-\sigma)^m \det\left(\left(-\frac{1}{\sigma}\right)(\sigma^2 I_n - A^\top A)\right) \\ &= \frac{(-\sigma)^m}{(-\sigma)^n} \det(\lambda I_n - A^\top A) \\ &= 0. \end{aligned}$$

□

1.3 On a Related Bilinear Program

Instead of the C-in-P bilinear program (1), one may also consider the related bilinear program in which x is constrained to be nonnegative while y is constrained to be in a convex set (i.e., the objective remains the same, but the constraints are flipped):

$$\tilde{p} \coloneqq \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} \left\{ y^\top(b - Ax) : x \geq 0,\ y \in C \right\}.$$

Here, C is a nonempty convex subset of $\mathbb{R}^m$. Now, the polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ lives in an entirely different space than C, so it makes no sense to consider the relation of this problem to the set containments $C \subseteq P$ or $P \subseteq C$. As the following example shows, the problem $\tilde{p}$ can have optimal values that are not in $\{0, -\infty\}$, and therefore by Corollary 3, $\tilde{p}$ is generally not a C-in-P bilinear program.

Example 2. Consider $\tilde{p}$ with convex set $C = \{y \in \mathbb{R}^m : y = \alpha e_1,\ \alpha \in [1, 2]\}$ and problem parameters given by

$$A = -e_1 \in \mathbb{R}^m, \qquad b = e_1 \in \mathbb{R}^m,$$

where e_1 is the 1st standard unit vector in $\mathbb{R}^m$. Thus, $n = 1$. The set C can be visualized as the line segment $[1, 2]$ contained within the first coordinate axis. We see that

$$\begin{aligned} \tilde{p} &= \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}^m} \left\{ y^\top((x+1)e_1) : x \geq 0,\ y \in C \right\} \\ &= \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}^m} \left\{ (x+1)\alpha : x \geq 0,\ \alpha \in [1, 2] \right\}. \end{aligned}$$

Clearly, we have that $(x + 1)\alpha \geq \alpha \geq 1$ for all feasible (x, y) with $y = \alpha e_1$, and this lower bound is attained by $(x, y) = (0, e_1)$, which is feasible. Hence,

$$\tilde{p} = 1 \notin \{0, -\infty\}.$$

Thus, $\tilde{p}$ is not a C-in-P bilinear program, even though in this case, we have the following convex-in-polyhedral containment:

$$C \subseteq \{z \in \mathbb{R}^m : z = b - Ax,\ x \geq 0\} = \{(1+x)e_1 : x \geq 0\} = \{\alpha e_1 : \alpha \in [1, \infty)\}.$$

The problem $\tilde{p}$ still admits nice geometric interpretations:

Proposition 5. *It holds that $\tilde{p} \geq 0$ if and only if $C \subseteq (-A\mathbb{R}^n_+ + \{b\})^*$, where $(-A\mathbb{R}^n_+ + \{b\})^*$ is the dual cone of $-A\mathbb{R}^n_+ + \{b\} = \{b - Ax \in \mathbb{R}^m : x \in \mathbb{R}^n_+\}$.*

Proof. It holds that $\tilde{p} \geq 0$ if and only if

$$\inf_{(x,y,z)\in\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^m} \left\{y^\top z : x \geq 0,\ y \in C,\ z = b - Ax\right\} \geq 0,$$

which is equivalent to

$$\inf_{(y,z)\in\mathbb{R}^m\times\mathbb{R}^m} \left\{y^\top z : y \in C,\ z \in (-A\mathbb{R}^n_+ + \{b\})\right\} \geq 0.$$

This is in turn equivalent to the condition that, for all $y \in C$, it holds that

$$y^\top z \geq 0 \text{ for all } z \in (-A\mathbb{R}^n_+ + \{b\}).$$

Since $y^\top z \geq 0$ for all $z \in (-A\mathbb{R}^n_+ + \{b\})$ if and only if $y \in (-A\mathbb{R}^n_+ + \{b\})^*$, we find that $\tilde{p} \geq 0$ if and only if

$$C \subseteq (-A\mathbb{R}^n_+ + \{b\})^*.$$

□

In fact, we can exactly characterize the dual cone $(-A\mathbb{R}^n_+ + \{b\})$ as a polyhedron, which shows that the nonnegativity of $\tilde{p}$ amounts to a C-in-P bilinear positivity certification problem, even if $\tilde{p}$ itself is not a C-in-P bilinear program:

Proposition 6. *It holds that*

$$(-A\mathbb{R}^n_+ + \{b\})^* = \tilde{P} \coloneqq \{y \in \mathbb{R}^m : A^\top y \leq 0 \text{ and } b^\top y \geq 0\},$$

and therefore $\tilde{p} \geq 0$ if and only if the optimal value of the following C-in-P bilinear program is zero:

$$\inf_{(y,u)\in\mathbb{R}^m\times\mathbb{R}^{n+1}} \left\{u^\top \left(\tilde{b} - \tilde{A}y\right) : u \geq 0,\ y \in C\right\},$$

where

$$\tilde{A} \coloneqq \begin{bmatrix} A^\top \\ -b^\top \end{bmatrix} \in \mathbb{R}^{(n+1)\times m}, \qquad \tilde{b} \coloneqq 0 \in \mathbb{R}^{n+1}.$$

Proof. We have that

$$\begin{aligned}(-A\mathbb{R}^n_+ + \{b\})^* &= \left\{y \in \mathbb{R}^m : y^\top z \geq 0 \text{ for all } z \in (-A\mathbb{R}^n_+ + \{b\})\right\} \\ &= \left\{y \in \mathbb{R}^m : \inf\{y^\top z : z \in (-A\mathbb{R}^n_+ + \{b\})\} \geq 0\right\} \\ &= \left\{y \in \mathbb{R}^m : \inf_{(x,z)\in\mathbb{R}^n\times\mathbb{R}^m} \left\{y^\top z : z = b - Ax,\ x \geq 0\right\} \geq 0\right\}.\end{aligned}$$

The optimization $\inf_{(x,z)\in\mathbb{R}^n\times\mathbb{R}^m}\{y^\top z : z = b - Ax,\ x \geq 0\}$ is a feasible linear program, and therefore by strong duality, we find that

$$\begin{aligned}\inf_{(x,z)\in\mathbb{R}^n\times\mathbb{R}^m}\{y^\top z : z = b - Ax,\ x \geq 0\} &= \sup_{(\lambda,\mu)\in\mathbb{R}^m\times\mathbb{R}^n}\left\{-\lambda^\top b : y + \lambda = 0,\ A^\top\lambda = \mu,\ \mu \geq 0\right\} \\ &= \begin{cases} b^\top y & \text{if } A^\top y \leq 0, \\ -\infty & \text{otherwise.}\end{cases}\end{aligned}$$

Hence,

$$(-A\mathbb{R}^n_+ + \{b\})^* = \left\{y \in \mathbb{R}^m : b^\top y \geq 0 \text{ and } A^\top y \leq 0\right\} = \tilde{P}.$$

Notice that we may rewrite this polyhedron as

$$\tilde{P} = \left\{y \in \mathbb{R}^m : \tilde{A}y \leq \tilde{b}\right\}$$

with

$$\tilde{A} = \begin{bmatrix} A^\top \\ -b^\top \end{bmatrix} \in \mathbb{R}^{(n+1)\times m}, \qquad \tilde{b} = 0 \in \mathbb{R}^{n+1}.$$

Hence, by Proposition 5, it holds that $\tilde{p} \geq 0$ if and only if $C \subseteq \tilde{P} = \{y \in \mathbb{R}^m : \tilde{A}y \leq \tilde{b}\}$, which, by Theorem 1, holds if and only if

$$\inf_{(y,u)\in\mathbb{R}^m\times\mathbb{R}^{n+1}}\left\{u^\top\left(\tilde{b} - \tilde{A}y\right) : u \geq 0,\ y \in C\right\} \geq 0.$$

This optimization problem is clearly a C-in-P bilinear program of the form (1), and hence has nonnegative optimal value if and only if its optimal value equals zero per Corollary 3. $\square$

2 Lower-Bounding Separable Bilinear Programs

In this section, we consider extending our framework to lower-bound general separable bilinear programs, constrained to products of polyhedra with convex sets:

$$\overline{p} := \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m}\left\{y^\top Ax + b_x^\top x + b_y^\top y + c : x \in C,\ y \in P\right\}, \tag{11}$$

where $C \subseteq \mathbb{R}^n$ is convex, $P \subseteq \mathbb{R}^m$ is a polyhedron, $A \in \mathbb{R}^{m\times n}$, $b_x \in \mathbb{R}^n$, $b_y \in \mathbb{R}^m$, and $c \in \mathbb{R}$. Notice that (11) generalizes the C-in-P bilinear program (1), as the objective now has an additional linear term in x, and the polyhedral constraints on y are now more general than the prior nonnegativity conditions. Separable bilinear programs of the form (11) have been studied quite extensively, although it is most common to assume that C is polyhedral; see, e.g., Gallo and Ülkücü (1977); Czochralska (1982); Sherali and Alameddine (1992); Nahapetyan (2009). A particularly interesting form of (11) is the case in which the constraints reduce to a box: $l \leq (x,y) \leq u$ for some $l, u \in \mathbb{R}^n \times \mathbb{R}^m$. Such problems are instances of box-constrained nonconvex quadratic programs, which are NP-hard problems that have been studied extensively and have classical relaxations that are known to yield effective lower bounds (Burer and Letchford, 2009). However, obtaining nontrivial lower bounds on $\overline{p}$ for the general bilinear program (11) can be quite challenging, as even Lagrangian duality-based approaches tend to fail:

Proposition 7. *Suppose that* $C = \{x \in \mathbb{R}^n : f_i(x) \leq 0,\ i \in \{1, \ldots, p\}\}$ *and* $P = \{y \in \mathbb{R}^m : Qy \leq q\}$ *for some convex functions* $f_i \colon \mathbb{R}^n \to \mathbb{R}$*, some* $Q \in \mathbb{R}^{k \times m}$*, and some* $q \in \mathbb{R}^k$*. If* $A \neq 0$ *and* $f_i(tv) = o(t^2)$ *as* $t \to \infty$ *for all* $i \in \{1, \ldots, p\}$ *and all* $v \in \mathbb{R}^n$*, then the Lagrange dual problem associated to* (11) *is infeasible, and hence the Lagrangian duality gap is infinite.*

Proof. The Lagrangian of (11) is given by

$$L(x, y, \lambda, \mu) = y^\top Ax + b_x^\top x + b_y^\top y + c + \sum_{i=1}^{p} \lambda_i f_i(x) + \mu^\top (Qy - q).$$

Since $A \neq 0$, it has a nonzero singular value $\sigma > 0$ with associated left singular vector $u \in \mathbb{R}^m$ and associated right singular vector $v \in \mathbb{R}^n$. Then, it holds for all $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^k$ that

$$\begin{aligned} L(tv, -tu, \lambda, \mu) &= -t^2 u^\top Av + tb_x^\top v - tb_y^\top u + c + \sum_{i=1}^{p} \lambda_i f_i(tv) - \mu^\top (tQu + q) \\ &= -t^2\sigma + tb_x^\top v - tb_y^\top u + c + \sum_{i=1}^{p} \lambda_i f_i(tv) - \mu^\top (tQu + q) \\ &= -t^2\sigma + o(t^2) \text{ as } t \to \infty. \end{aligned}$$

Therefore, $\lim_{t\to\infty} L(tv, -tu, \lambda, \mu) = -\infty$ for all $(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^k$, so the Lagrangian is unbounded below in its first two arguments. Thus, the Lagrange dual problem associated to (11) reduces to

$$\overline{d} \coloneqq \sup_{(\lambda,\mu)\in\mathbb{R}_+^p\times\mathbb{R}_+^k} \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} L(x, y, \lambda, \mu) = -\infty.$$

This infinite duality gap is equivalent to the Lagrange dual problem being infeasible. ☐

Corollary 6. *Consider* $C = \{x \in \mathbb{R}^n : f_i(x) \leq 0$ *for all* $i \in \{1, \ldots, p\}\}$ *and* $P = \{y \in \mathbb{R}^m : Qy \leq q\}$ *as in Proposition 7. If* $A \neq 0$ *and every* f_i *is Lipschitz continuous, then the Lagrange dual problem associated to* (11) *is infeasible, and hence the Lagrangian duality gap is infinite.*

Proof. Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. If f_i is Lipschitz continuous, then there exists $L \in \mathbb{R}$ (possibly dependent on n) such that

$$\begin{aligned} |f_i(tv)| &\leq |f_i(tv) - f_i(0)| + |f_i(0)| \\ &\leq L|t|\|v\| + |f_i(0)| \\ &= o(t^2) \text{ as } t \to \infty \end{aligned}$$

for all $v \in \mathbb{R}^n$. Thus, the result follows from Proposition 7. ☐

Corollary 7. *If* C *and* P *are both polyhedra, then the bilinear program* (11) *has an infinite Lagrangian duality gap whenever* $A \neq 0$.

Proof. This follows immediately from Corollary 6, as a polyhedral set C is defined by a finite number of affine (and hence Lipschitz continuous) inequality constraint functions. ☐

2.1 C-in-P Relaxation of Separable Bilinear Programs

Instead of relying on the above (potentially meaningless) duality-based lower bounds, we will now show that the general bilinear program (11) can be relaxed to a C-in-P bilinear program by solving a simple linear system of equations. In doing so, we can appeal to our convex optimization-based C-in-P bilinear positivity certificates (Corollary 1 and Corollary 2) to efficiently compute positivity certificates for the general bilinear program (11).

Theorem 4. *Suppose that $P = \{y \in \mathbb{R}^m : Qy \leq q\}$ for some $Q \in \mathbb{R}^{k\times m}$ and some $q \in \mathbb{R}^k$. If there exist $\overline{A} \in \mathbb{R}^{k\times n}$ and $\overline{b} \in \mathbb{R}^k$ such that*

$$\begin{bmatrix} I_n \otimes Q^\top & 0_{mn\times k} \\ -I_n \otimes q^\top & 0_{n\times k} \\ 0_{m\times nk} & -Q^\top \\ 0_{1\times nk} & q^\top \end{bmatrix} \begin{bmatrix} \overline{a}_1 \\ \vdots \\ \overline{a}_n \\ \overline{b} \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \\ b_x \\ b_y \\ c \end{bmatrix}, \tag{12}$$

where $a_i \in \mathbb{R}^m$ and $\overline{a}_i \in \mathbb{R}^k$ respectively denote the ith columns of A and $\overline{A}$, then the following C-in-P bilinear programming lower bound holds:

$$\inf_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R}^k} \left\{\lambda^\top(\overline{b} - \overline{A}x) : \lambda \geq 0,\ x \in C\right\} \leq \overline{p}. \tag{13}$$

Proof. Suppose that there exist $\overline{A} \in \mathbb{R}^{k\times n}$ and $\overline{b} \in \mathbb{R}^k$ that satisfy (12). It is easy to see that the linear system of equations (12) can be rewritten into the following form:

$$\begin{aligned} A &= Q^\top \overline{A}, \\ b_x &= -\overline{A}^\top q, \\ b_y &= -Q^\top \overline{b}, \\ c &= q^\top \overline{b}. \end{aligned}$$

Thus, we find that

$$\begin{aligned} &\inf_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R}^k} \left\{\lambda^\top(\overline{b} - \overline{A}x) : \lambda \geq 0,\ x \in C\right\} \\ &\leq \inf_{(x,y,\lambda)\in\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^k} \left\{\lambda^\top(\overline{b} - \overline{A}x) : \lambda \geq 0,\ x \in C,\ \lambda = q - Qy\right\} \\ &= \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} \left\{(q - Qy)^\top(\overline{b} - \overline{A}x) : q - Qy \geq 0,\ x \in C\right\} \\ &= \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} \left\{y^\top Q^\top \overline{A}x - q^\top \overline{A}x - \overline{b}^\top Qy + q^\top \overline{b} : x \in C,\ y \in P\right\} \\ &= \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} \left\{y^\top Ax + b_x^\top x + b_y^\top y + c : x \in C,\ y \in P\right\} \\ &= \overline{p}, \end{aligned}$$

which completes the proof. □

Corollary 8. *Consider the polyhedron P as in Theorem 4. If there exist $\overline{A} \in \mathbb{R}^{k\times n}$ and $\overline{b} \in \mathbb{R}^k$ that satisfy* (12)*, then $\overline{p} \geq 0$ if $C \subseteq \overline{P} := \{x \in \mathbb{R}^n : \overline{A}x \leq \overline{b}\}$.*

Proof. The result follows immediately from Theorem 4 together with Theorem 1. $\square$

Notice that (12) may have an infinite number of solutions $(\overline{A}, \overline{b})$, and in such cases Corollary 8 shows that it is advantageous to choose a solution that makes $\overline{P} = \{x \in \mathbb{R}^n : \overline{A}x \leq \overline{b}\}$ "large enough" to cover C. In fact, when $C = \{x \in \mathbb{R}^n : f_j(x) \leq 0 \text{ for all } j \in \{1, \dots, p\}\}$ with $f_1, \dots, f_p$ real-valued convex functions on $\mathbb{R}^n$ satisfying Slater's condition, we may always find such a solution leading to $C \subseteq \overline{P}$, if one exists, by appealing to our convex optimization-based C-in-P positivity certificate of Corollary 2 and showing the feasibility of

$$\inf_{\substack{(\lambda_1,\dots,\lambda_k)\in(\mathbb{R}^p)^k \\ (\overline{A},\overline{b})\in\mathbb{R}^{k\times n}\times\mathbb{R}^k}} \left\{0 : g_i(\lambda_i) + \overline{b}_i \geq 0 \text{ and } \lambda_i \geq 0 \text{ for all } i \in \{1, \dots, k\},\ (\overline{A}, \overline{b}) \text{ satisfies (12)}\right\}, \tag{14}$$

where the Lagrange dual functions $g_i \colon \mathbb{R}^p \to \mathbb{R} \cup \{-\infty\}$ are given by

$$g_i(\lambda) = \inf_{x\in\mathbb{R}^n} \left(-\overline{a}_i^\top x + \sum_{j=1}^p \lambda_j f_j(x)\right) \tag{15}$$

with $\overline{a}_i \in \mathbb{R}^n$ denoting the ith row of $\overline{A}$.[1] This approach can be viewed as combining the search for the bound-inducing parameters $(\overline{A}, \overline{b})$ together with the actual bounding procedure into a single convex optimization problem.

The following example shows that the C-in-P bilinear programming lower bound (13) is capable of granting positivity certificates for the bilinear program (11), even when the Lagrangian dual problem associated to (11) is infeasible.

Example 3. Suppose that $n = 2$, and consider the bilinear program (11) with $A \neq 0$, and with

$$\begin{aligned}
b_x &= -A^\top l, \\
C &= [-1, 1]^2 = \left\{x \in \mathbb{R}^2 : \begin{bmatrix} I_2 \\ -I_2 \end{bmatrix} x \leq 1_4\right\}, \\
P &= \{y \in \mathbb{R}^m : y \geq l\},
\end{aligned}$$

for some vector $l \in \mathbb{R}^m$. We may write $P = \{y \in \mathbb{R}^m : Qy \leq q\}$ with $Q = -I_m$ and $q = -l$. Since C and P are both polyhedra, Corollary 7 gives that (11) has an infinite Lagrangian duality gap. On the other hand, we have that

$$\begin{aligned}
\overline{p} &= \inf_{(x,y)\in\mathbb{R}^2\times\mathbb{R}^m} \left\{y^\top Ax + b_x^\top x + b_y^\top y + c : x \in C,\ y \in P\right\} \\
&= c - \overline{c} + \inf_{(x,y)\in\mathbb{R}^2\times\mathbb{R}^m} \left\{y^\top Ax + b_x^\top x + b_y^\top y + \overline{c} : x \in C,\ Qy \leq q\right\}.
\end{aligned}$$

[1] Beware of the overloaded notation here: $\overline{a}_i$ denotes the ith column of $\overline{A}$ in (12), while $\overline{a}_i$ denotes the ith row of $\overline{A}$ in (15).

where $\bar{c} \coloneqq -l^\top b_y$. Since $\overline{A} \coloneqq -A$ and $\bar{b} \coloneqq b_y$ satisfy the linear system of equations

$$\begin{aligned} A &= Q^\top \overline{A} = -\overline{A}, \\ b_x &= -\overline{A}^\top q = -A^\top l, \\ b_y &= -Q^\top \bar{b} = \bar{b}, \\ \bar{c} &= q^\top \bar{b} = -l^\top b_y, \end{aligned}$$

we conclude from Theorem 4 that

$$\overline{p} \geq c + l^\top b_y + \inf_{(x,\lambda)\in\mathbb{R}^2\times\mathbb{R}^m} \left\{\lambda^\top (b_y + Ax) : \lambda \geq 0,\ x \in C\right\}.$$

Therefore, Theorem 1 gives that $\overline{p} \geq c + l^\top b_y$ if $C \subseteq \overline{P} \coloneqq \{x \in \mathbb{R}^n : -Ax \leq b_y\}$. For example, if $m = 2$, $A = I_2$, and $b_y \in \mathbb{R}^2$ is any vector satisfying $b_y \geq 1_2$, then $C = [-1, 1]^2 \subseteq \overline{P} = \{x \in \mathbb{R}^n : x \geq -b_y\}$, so $\overline{p} \geq c + l^\top b_y$. In this case, $\overline{p} \geq 0$ whenever $c + l^\top b_y \geq 0$, giving a positivity certificate for this instantiation of (11), even though the Lagrangian duality gap is infinite. In fact, it is straightforward to see that this C-in-P programming-based lower bound is tight (meaning $\overline{p} = c + l^\top b_y$) if $\overline{p}$ is solved by some (x, y) with $y = l$.

2.1.1 Feasibility of the Linear System

Theorem 4 shows that, if a solution exists for the linear system of equations (12), then we can provably lower-bound (11) using a C-in-P bilinear program. Therefore, it is natural to ask when such solutions exist. Below, we show that solutions exist whenever the augmented coefficient matrix of P is full column rank. We also show that this occurs whenever P is a polytope (i.e., whenever the polyhedron P is bounded).

Proposition 8. *Suppose that $P = \{y \in \mathbb{R}^m : Qy \leq q\}$ for some $Q \in \mathbb{R}^{k\times m}$ and some $q \in \mathbb{R}^k$. There exists a solution $(\overline{A}, \bar{b}) \in \mathbb{R}^{k\times m} \times \mathbb{R}^k$ to (12) whenever*

$$\operatorname{rank} \begin{bmatrix} Q & q \end{bmatrix} = m + 1.$$

Furthermore, a unique solution exists if and only if $k = m + 1$ and $\begin{bmatrix} Q & q \end{bmatrix}$ is invertible.

Proof. For notational simplicity, denote the coefficient matrix in (12) by

$$M \coloneqq \begin{bmatrix} I_n \otimes Q^\top & 0_{mn\times k} \\ -I_n \otimes q^\top & 0_{n\times k} \\ 0_{m\times nk} & -Q^\top \\ 0_{1\times nk} & q^\top \end{bmatrix} \in \mathbb{R}^{(mn+m+n+1)\times(nk+k)}.$$

Suppose that $\operatorname{rank} \begin{bmatrix} Q & q \end{bmatrix} = m + 1$. It is clear that the rank of the lower-right $(m + 1) \times k$ block is

$$\operatorname{rank} \begin{bmatrix} -Q^\top \\ q^\top \end{bmatrix} = m + 1,$$

and hence the rank of the upper-left $(nm+n) \times nk$ block is

$$\begin{aligned}
\operatorname{rank} \begin{bmatrix} I_n \otimes Q^\top \\ -I_n \otimes q^\top \end{bmatrix} &= \operatorname{rank} \begin{bmatrix} Q^\top & 0 & \cdots & 0 \\ 0 & Q^\top & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q^\top \\ -q^\top & 0 & \cdots & 0 \\ 0 & -q^\top & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -q^\top \end{bmatrix} \\
&= \operatorname{rank} \begin{bmatrix} Q^\top & 0 & \cdots & 0 \\ -q^\top & 0 & \cdots & 0 \\ 0 & Q^\top & \cdots & 0 \\ 0 & -q^\top & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q^\top \\ 0 & 0 & \cdots & -q^\top \end{bmatrix} \\
&= n(m+1).
\end{aligned}$$

Therefore, the overall rank is

$$\operatorname{rank} M = m + 1 + n(m+1) = mn + m + n + 1,$$

so M is full row rank. Therefore, (12) admits a solution $(\overline{A}, \overline{b}) \in \mathbb{R}^{k \times m} \times \mathbb{R}^k$.

A unique solution to (12) exists if and only if M is invertible. If $k = m+1$ and $\begin{bmatrix} Q & q \end{bmatrix}$ is invertible, then $\operatorname{rank} \begin{bmatrix} Q & q \end{bmatrix} = m+1$, and hence $\operatorname{rank}(M) = m + 1 + n(m+1) = k + nk$ per the analysis above. In this case, we see that M is a full rank square matrix, and is hence invertible, implying the existence of a unique solution to (12). On the other hand, suppose that (12) has a unique solution, so that M is invertible. Then, it holds that $(m+1)(n+1) = mn + m + n + 1 = nk + k = k(n+1)$, implying that $k = m+1$. If $\begin{bmatrix} Q & q \end{bmatrix}$ is not invertible, then $\operatorname{rank} \begin{bmatrix} Q & q \end{bmatrix} < k$, which implies that $\operatorname{rank} M < nk + k$, contradicting the invertibility of M. Hence, it must be that $\begin{bmatrix} Q & q \end{bmatrix}$ is invertible in this case. ☐

Lemma 2. *Suppose that $P = \{y \in \mathbb{R}^m : Qy \le q\}$ for some $Q \in \mathbb{R}^{k \times m}$ and some $q \in \mathbb{R}^k$. If P is bounded and $q \in \operatorname{Range}(Q)$, then $\dim P = 0$.*

Proof. Denote the ith row of Q by $Q_i \in \mathbb{R}^m$. Suppose that P is bounded and that $q \in \operatorname{Range}(Q)$. The boundedness of P implies that $\operatorname{rank}(Q) = m$, per Lemma 7. Assume for the sake of contradiction that $\dim P > 0$. Then, there exist $i \in \{1, \dots, k\}$ and $y_0 \in P$ such that $Q_i^\top y_0 < q_i$ by definition of $\dim P$ together with the fact that $\operatorname{rank}(Q) = m$. Since $q \in \operatorname{Range}(Q)$, it holds that $Qy = q$ for some $y \in \mathbb{R}^m$. Therefore, we find that $Q_i^\top (y_0 - y) = Q_i^\top y_0 - q_i < q_i - q_i = 0$. Furthermore, it holds that $Q(y_0 - y) = Qy_0 - q \le q - q = 0$, as $y_0 \in P$. Hence, we find that $Q(y + \alpha(y_0 - y)) = q + \alpha Q(y_0 - y) \le q$ for all $\alpha \ge 0$, implying that $y + \alpha(y_0 - y) \in P$ for all $\alpha \ge 0$. Since P is bounded, it must be the case that $y = y_0$. However, this contradicts the fact that $Q_i^\top y_0 < q_i \ne Q_i^\top y$. Hence, it must be the case that $\dim P = 0$. ☐

Proposition 9. *Suppose that* $P = \{y \in \mathbb{R}^m : Qy \leq q\}$ *for some* $Q \in \mathbb{R}^{k\times m}$ *and some* $q \in \mathbb{R}^k$. *If* P *is bounded and* $\dim P > 0$, *then* $\operatorname{rank}\begin{bmatrix} Q & q \end{bmatrix} = m+1$, *and hence* (12) *admits a solution.*

Proof. Suppose that P is bounded and that $\dim P > 0$. Then, Lemma 2 gives that $q \notin \operatorname{Range}(Q)$. By Lemma 7, it must be that $\operatorname{rank}(Q) = m$. Therefore, it must be the case that $\operatorname{rank}\begin{bmatrix} Q & q \end{bmatrix} > \operatorname{rank} Q = m$, and hence $\operatorname{rank}\begin{bmatrix} Q & q \end{bmatrix} = m+1$. The system (12) therefore admits a solution, per Proposition 8. $\square$

2.1.2 Exactness of the Relaxation

Here, we characterize when the C-in-P relaxation of a separable bilinear program is exact, i.e., when the lower bound (13) is tight.

Theorem 5. *Suppose that* $P = \{y \in \mathbb{R}^m : Qy \leq q\}$ *for some* $Q \in \mathbb{R}^{k\times m}$ *and some* $q \in \mathbb{R}^k$. *If* $\overline{p}$ *is attained and there exist* $\overline{A} \in \mathbb{R}^{k\times n}$ *and* $\overline{b} \in \mathbb{R}^k$ *that satisfy* (12) *and* $C \subseteq \overline{P} = \{x \in \mathbb{R}^n : \overline{A}x \leq \overline{b}\}$, *then the inequality* (13) *holds with equality if and only if there exists* $(x,y) \in C \times P$ *such that*

$$\overline{a}_i^\top x = \overline{b}_i \text{ or } Q_i^\top y = q_i \text{ for all } i \in \{1,\dots,k\}, \tag{16}$$

where $\overline{a}_i \in \mathbb{R}^n$ *and* $Q_i \in \mathbb{R}^m$ *respectively denote the* i*th rows of* $\overline{A}$ *and* Q.

Proof. Assume that $\overline{p}$ is attained and that there exist $\overline{A} \in \mathbb{R}^{k\times n}$ and $\overline{b} \in \mathbb{R}^k$ that satisfy (12) and $C \subseteq \overline{P} = \{x \in \mathbb{R}^n : \overline{A}x \leq \overline{b}\}$. Then, Theorem 1, Corollary 3, and Theorem 4 give that

$$0 = \inf_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R}^k} \left\{\lambda^\top(\overline{b} - \overline{A}x) : \lambda \geq 0,\ x \in C\right\} \leq \overline{p}.$$

If there exists $(x,y) \in C \times P$ such that (16) holds, then (12) yields that

$$\overline{p} = \inf_{(\tilde{x},\tilde{y})\mathbb{R}^n\times\mathbb{R}^m} \left\{(q - Q\tilde{y})^\top(\overline{b} - \overline{A}\tilde{x}) : \tilde{x} \in C,\ \tilde{y} \in P\right\} \leq (q - Qy)^\top(b - Ax) = 0,$$

and hence $\overline{p} = 0$, so the inequality (13) holds with equality.

On the other hand, suppose that the inequality (13) holds with equality, so that $\overline{p} = 0$. Since $C \subseteq \overline{P} = \{x \in \mathbb{R}^n : \overline{A}x \leq \overline{b}\}$, it holds that $\overline{b} - \overline{A}x \geq 0$ for all $x \in C$. Therefore, $(q_i - Q_i^\top y)(\overline{b}_i - \overline{a}_i^\top x) \geq 0$ for all $(x,y) \in C \times P$ and all $i \in \{1,\dots,k\}$. Let $(x,y) \in C \times P$ be a solution to $\overline{p}$, which exists since $\overline{p}$ is assumed to be attained. Then, we find that

$$(q_i - Q_i^\top y)(\overline{b}_i - \overline{a}_i^\top x) \geq 0 \text{ for all } i \text{ and } \sum_{i=1}^n (q_i - Q_i^\top y)(\overline{b}_i - \overline{a}_i^\top x) = \overline{p} = 0,$$

so it must be the case that $(q_i - Q_i^\top y)(\overline{b}_i - \overline{a}_i^\top x) = 0$ for all i, and hence that (16) holds. $\square$

As we mentioned in our discussion following Corollary 8, the condition that $(\overline{A}, \overline{b})$ satisfies (12) as well as $C \subseteq \overline{P} = \{x \in \mathbb{R}^n : \overline{A}x \leq \overline{b}\}$ is equivalent to showing that the convex optimization problem (14) is feasible, in the case that C is defined by real-valued convex functions satisfying Slater's condition.

As a consequence of Theorem 5, we find the rather negative result that C-in-P relaxations of separable bilinear programs are never exact when P is bounded, $\dim P > 0$, and every $x \in C$ gets mapped to a vector $\overline{b} - \overline{A}x$ in the positive orthant:

Corollary 9. *Suppose that $P = \{y \in \mathbb{R}^m : Qy < q\}$ for some $Q \in \mathbb{R}^{k\times m}$ and some $q \in \mathbb{R}^k$. Assume that $\overline{p}$ is attained and that there exist $\overline{A} \in \mathbb{R}^{k\times n}$ and $\overline{b} \in \mathbb{R}^k$ that satisfy* (12) *and $C \subseteq \overline{P} = \{x \in \mathbb{R}^n : \overline{A}x \leq \overline{b}\}$. If P is bounded, $\dim P > 0$, and $-\overline{A}C + \{\overline{b}\} \subseteq \mathbb{R}^k_{++}$, then the inequality* (13) *holds strictly.*

Proof. Suppose that P is bounded, $\dim P > 0$, and $-\overline{A}C + \{\overline{b}\} \subseteq \mathbb{R}^k_{++}$. Let $(x, y) \in C \times P$ be a solution to $\overline{p}$, which exists since $\overline{p}$ is attained. Lemma 2 ensures that $q \notin \mathrm{Range}(Q)$. Therefore, it must be that $Qy \neq q$, so that $Q_i^\top y < q_i$ for some $i \in \{1, \dots, k\}$, where $Q_i \in \mathbb{R}^m$ denotes the ith row of Q. Since $x \in C$ and $-\overline{A}C + \{\overline{b}\} \subseteq \mathbb{R}^k_{++}$, it holds that $\overline{A}x < \overline{b}$. Therefore, we find that

$$\overline{p} = (q - Qy)^\top(\overline{b} - \overline{A}x) \geq (q_i - Q_i^\top y)(\overline{b}_i - \overline{a}_i^\top x) > 0.$$

Since $C \subseteq \overline{P}$, Theorem 1 and Corollary 3 give that

$$\inf_{(\tilde{x},\lambda)\in\mathbb{R}^n\times\mathbb{R}^k} \left\{\lambda^\top(\overline{b} - \overline{A}\tilde{x}) : \lambda \geq 0,\ \tilde{x} \in C\right\} = 0,$$

and therefore the inequality (13) holds strictly. □

3 Lemmas

Lemma 3. *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a nonempty polyhedron with $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. If $A = 0$, then $P = \mathbb{R}^n$. The same conclusion holds in the case that $P = \{x \in \mathbb{R}^n : Ax < b\}$ is nonempty with $A = 0$.*

Proof. Since P is nonempty, it holds that $0 = 0x = Ax \leq b$ for some $x \in \mathbb{R}^n$. Hence, it holds that $Ax = 0x = 0 \leq b$ for all $x \in \mathbb{R}^n$. Thus, $P = \mathbb{R}^n$. The argument obviously extends to the case of the open polyhedron $P = \{x \in \mathbb{R}^n : Ax < 0\}$. □

Lemma 4. *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron with $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then, it holds that*

$$\{x \in \mathbb{R}^n : Ax < b\} \subseteq \mathrm{int}(P).$$

Proof. If $\{x \in \mathbb{R}^n : Ax < b\} = \emptyset$, then the claim is vacuously true. Otherwise, let $x \in \mathbb{R}^n$ be such that $Ax < b$. If $A = 0$, then Lemma 3 gives that $\{x \in \mathbb{R}^n : Ax < b\} = \mathbb{R}^n = \mathrm{int}(\mathbb{R}^n) = \mathrm{int}(P)$. On the other hand, suppose that $A \neq 0$. Then, it holds that $\max\{\|a_1\|_2, \dots, \|a_m\|_2\} > 0$, where $a_i \in \mathbb{R}^n$ denotes the ith row of A. Furthermore, as $Ax < b$, we see that

$$\epsilon_x := \frac{\min\{b_1 - a_1^\top x, \dots, b_m - a_m^\top x\}}{\max\{\|a_1\|_2, \dots, \|a_m\|_2\}} \in (0, \infty).$$

Define the open ball $U_x \coloneqq \{y \in \mathbb{R}^n : \|y - x\|_2 < \epsilon_x\}$. If $y \in U_x$, then

$$\begin{aligned}
a_i^\top y &= a_i^\top x + a_i^\top (y - x) \\
&\leq a_i^\top x + \|a_i\|_2 \|y - x\|_2 \\
&\leq a_i^\top x + \max\{\|a_1\|_2, \dots, \|a_m\|_2\} \|y - x\|_2 \\
&< a_i^\top x + \max\{\|a_1\|_2, \dots, \|a_m\|_2\} \epsilon_x \\
&= a_i^\top x + \min\{b_1 - a_1^\top x, \dots, b_m - a_m^\top x\} \\
&\leq b_i
\end{aligned}$$

for all $i \in \{1, \dots, m\}$. Hence, it holds that $U_x \subseteq P$. Therefore, we conclude that $\{x \in \mathbb{R}^n : Ax < b\} \subseteq \operatorname{int}(P)$. □

Lemma 5. *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Denote the ith row of A by $a_i \in \mathbb{R}^n$. If*

$$b_i > 0 \text{ for all } i \in \{1, \dots, m\} \text{ such that } a_i = 0,$$

then it holds that

$$\operatorname{int}(P) \subseteq \{x \in \mathbb{R}^n : Ax < b\}.$$

Proof. Suppose that $b_i > 0$ for all $i \in \{1, \dots, m\}$ such that $a_i = 0$. If $\operatorname{int}(P) = \emptyset$, the claim vacuously holds true. Otherwise, let $x \in \operatorname{int}(P)$. Then, there exists $\epsilon_x > 0$ such that $Ay \leq b$ for all $y \in U_x \coloneqq \{z \in \mathbb{R}^n : \|z - x\|_2 < \epsilon_x\}$. Hence, it holds that $a_i^\top x + a_i^\top (y - x) = a_i^\top y \leq b_i$ for all $i \in \{1, \dots, m\}$ and all $y \in U_x$, implying that

$$a_i^\top x + \sup_{y \in \mathbb{R}^n} \{a_i^\top (y - x) : \|y - x\|_2 < \epsilon_x\} = a_i^\top x + \epsilon_x \|a_i\|_2 \leq b_i$$

for all $i \in \{1, \dots, m\}$. If $a_i \neq 0$, then $\epsilon_x \|a_i\|_2 > 0$, so $a_i^\top x < b_i$. On the other hand, if $a_i = 0$, then it must be that $b_i > 0$, and therefore again we find that $a_i^\top x = 0 < b_i$. Thus, $Ax < b$, implying that $\operatorname{int}(P) \subseteq \{x \in \mathbb{R}^n : Ax < b\}$. □

Lemma 6. *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Denote the ith row of A by $a_i \in \mathbb{R}^n$. If*

$$b_i > 0 \text{ for all } i \in \{1, \dots, m\} \text{ such that } a_i = 0,$$

then it holds that

$$\operatorname{int}(P) = \{x \in \mathbb{R}^n : Ax < b\}.$$

Proof. This follows immediately from combining Lemma 4 together with Lemma 5. □

Notice that we need the above regularity condition that $b_i > 0$ for all $\in \{1, \dots, m\}$ such that $a_i = 0$. To see this, consider the simple case where $P = \{x \in \mathbb{R} : a_1 x \leq b_1,\ a_2 x \leq b_2\}$, with $a_1 = 1$, $a_2 = 0$, and $b_1 = b_2 = 0$. Clearly, $P = (-\infty, 0]$. Thus, $\operatorname{int}(P) = (-\infty, 0)$. However, $\{x \in \mathbb{R} : a_1 x < b_1,\ a_2 x < b_2\} = \{x \in \mathbb{R} : x < 0,\ 0x < 0\} = \emptyset \neq \operatorname{int}(P)$.

Lemma 7. *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a nonempty polyhedron with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. If P is bounded, then* $\operatorname{rank}(A) = n$.

Proof. If P is bounded, then it is pointed, implying that every minimal face is 0-dimensional (consisting of a vertex), and hence A must have rank n (Schrijver, 1998, Section 8.5). □

3.1 Supporting Lemmas

The first lemma below is well-established in the field of robust optimization, and the second was stated in Sadraddini and Tedrake (2019) without proof (we prove it here).

Lemma 8. *Let* $X = \{x \in \mathbb{R}^n : f_i(x) \leq 0 \forall i\}$ *with every* f_i *convex, and let* $Y = \{y \in \mathbb{R}^n : g_j(y) \leq 0 \forall j\}$ *with every* g_j *concave. It holds that*

$$X \subseteq Y \iff p_j^\star := \sup_{x \in X} g_j(x) \leq 0 \forall j,$$

with the optimizations $p_j^\star$ *on the right all being convex problems. Let* $d_j^\star$ *denote the optimal value of the dual of* $p_j^\star$*. If Slater's holds for* X*, then the above is furthermore equivalent to* $d_j^\star \leq 0$ *for all* j*, which can be turned into a feasibility condition with finitely many convex constraints on the dual variables.*

Clearly, X in the above lemma is convex, but Y is nonconvex.

Lemma 9. *Let* $X, Y \subseteq \mathbb{R}^n$ *be closed and convex. Then, it holds that* $X \subseteq Y$ *if and only if*

$$\sup_{x \in X} v^\top x \leq \sup_{y \in Y} v^\top y \ \forall v \in \mathbb{R}^n.$$

Proof. Suppose that $X \subseteq Y$. Let $v \in \mathbb{R}^n$. If $x \in X$, then $x \in Y$, so $v^\top x \leq \sup_{y \in Y} v^\top y$. Thus,

$$\sup_{x \in X} v^\top x \leq \sup_{y \in Y} v^\top y.$$

On the other hand, suppose that the above inequality holds for all $v \in \mathbb{R}^n$. Assume for the sake of contradiction that $X \nsubseteq Y$. Then, there exists $x \in X$ such that $x \notin Y$. Hence, $\{x\}$ is a compact, closed, convex set that is disjoint from the closed and convex set Y. Thus, by the (strict) hyperplane separation theorem, there exist $v \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$ such that

$$v^\top x > c_1 > c_2 > v^\top y \ \forall y \in Y.$$

Hence,

$$v^\top x > \sup_{y \in Y} v^\top y,$$

which contradicts the fact that $\sup_{x \in X} v^\top x \leq \sup_{y \in Y} v^\top y$ for all $v \in \mathbb{R}^n$. Thus, it must be the case that $X \subseteq Y$. $\square$

Notice that in the above lemma, we could have added the condition that $\|v\|_2 \leq \alpha$ for any $\alpha > 0$ by a simple rescaling of the vector defining the separating hyperplane.

In the case that $Y = \{y \in \mathbb{R}^n : g_j(y) \leq 0 \forall j\}$ is defined by affine inequalities $c_j^\top y - d_j \leq 0$, both of the above lemmas apply, and hence we obtain three equivalent conditions.

Another useful fact to recall is the following:

Lemma 10. *If* $P = \{x \in \mathbb{R}^n : a_i^\top x \leq b_i \forall i \in I\}$ *is bounded, then* $|I| \geq n + 1$.

A Nonconvex QPs over Polytopes have Infinite Duality Gaps

Consider minimizing a nonconvex quadratic function over a nonempty, compact, convex polytope $X = \{x \in \mathbb{R}^n : Ax \leq b\}$ with $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$:

$$p^\star := \inf_{x\in\mathbb{R}^n} \left\{x^\top Px + q^\top x + r : Ax \leq b\right\},$$

where $P \in \mathbb{R}^{n\times n}$ is symmetric but not positive semidefinite. Consider the dual problem

$$d^\star := \sup_{\lambda\geq 0} \inf_{x\in\mathbb{R}^n} L(x, \lambda),$$

where the Lagrangian is given by

$$L(x, \lambda) = x^\top Px + q^\top x + r + \lambda^\top(Ax - b).$$

Proposition 10. *It holds that $d^\star = -\infty$, and hence the duality gap between $p^\star$ and $d^\star$ is infinite.*

Proof. Since P is not positive semidefinite, it has a negative eigenvalue α with associated eigenvector $v \in \mathbb{R}^n$. Let $\lambda \in \mathbb{R}^m$ be arbitrary. We see that

$$\lim_{t\to\infty} L(tx, \lambda) = \lim_{t\to\infty}(t^2\alpha\|v\|_2^2 + tq^\top v + r + \lambda^\top(tAv - b)) = -\infty,$$

since $\alpha\|v\|_2^2 < 0$. Thus, $L(\cdot, \lambda)$ is unbounded below; $\inf_{x\in\mathbb{R}^n} L(x, \lambda) = -\infty$. Since λ was chosen arbitrarily, it must be the case that

$$d^\star = \sup_{\lambda\geq 0} \inf_{x\in\mathbb{R}^n} L(x, \lambda) = -\infty.$$

As the feasible set of $p^\star$ is compact and its objective function is continuous, $p^\star$ is attained and hence finite. Thus, the duality gap is infinite. $\square$

B Nonconvex QPs have Unbounded SDP Relaxations

Consider also, the exact reformulation of $p^\star$ as

$$p^\star = p' := \inf_{(x,X)\in\mathbb{R}^n\times\mathbb{R}^{n\times n}} \{\mathrm{tr}(PX) + q^\top x + r : Ax \leq b,\ X = xx^\top\}.$$

The problem p' can be relaxed to the following semidefinite program:

$$\hat{p} := \inf_{(x,X)\in\mathbb{R}^n\times\mathbb{R}^{n\times n}} \{\mathrm{tr}(PX) + q^\top x + r : Ax \leq b,\ X \succeq xx^\top\}.$$

Clearly, $p^\star \geq \hat{p}$.

Proposition 11. *It holds that $\hat{p} = -\infty$; the SDP relaxation is unbounded below.*

Proof. Let $\overline{x}$ be an element of the nonempty polytope $\{x \in \mathbb{R}^n : Ax \leq b\}$. Since P is not positive semidefinite, it has a negative eigenvalue α with associated eigenvector $v \in \mathbb{R}^n$. Let $t \in \mathbb{R}$ and consider $X = (tv)(tv)^\top + \overline{x}\overline{x}^\top = t^2 vv^\top + \overline{x}\overline{x}^\top$. Also, consider $x = \overline{x}$. Then, for all $y \in \mathbb{R}^n$, it holds that

$$y^\top (X - xx^\top) y = t^2 (y^\top v)^2 \geq 0,$$

implying that $X \succeq xx^\top$. Thus, $(x, X) = (\overline{x}, t^2 vv^\top + \overline{x}\overline{x}^\top)$ is feasible for $\hat{p}$ for all $t \in \mathbb{R}^n$. The objective function at such a feasible point equals

$$t^2 \operatorname{tr}(Pvv^\top) + \operatorname{tr}(P\overline{x}\overline{x}^\top) + q^\top \overline{x} + r = t^2 \alpha \|v\|_2^2 + \overline{x}^\top P \overline{x} + q^\top \overline{x} + r \to -\infty$$

as $t \to \infty$, since $\alpha \|v\|_2^2 < 0$. Therefore, we conclude that $\hat{p} = -\infty$. □

References

Sadra Sadraddini and Russ Tedrake. Linear encodings for polytope containment problems. In *58th Conference on Decision and Control (CDC)*. IEEE, 2019.

Jean-Bernard Lasserre. *Moments, Positive Polynomials and Their Applications*. World Scientific, 2009.

Amir Ali Ahmadi and Anirudha Majumdar. DSOS and SDSOS optimization: More tractable alternatives to sum of squares and semidefinite optimization. *Journal on Applied Algebra and Geometry*, 2019.

Robert M. Freund and James B. Orlin. On the complexity of four polyhedral set containment problems. *Mathematical Programming*, 1985.

Peter Gritzmann and Victor Klee. On the complexity of some basic problems in computational convexity: I. Containment problems. *Discrete Mathematics*, 1994.

Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Miguel Sousa Lobo, Lieven Vandenberghe, Stephen Boyd, and Hervé Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 1998.

Motakuri Ramana and Alan J. Goldman. Some geometric results in semidefinite programming. *Journal of Global Optimization*, 1995.

Joao Gouveia and Tim Netzer. Positive polynomials and projections of spectrahedra. *Journal on Optimization*, 2011.

Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Review*, 1996.

Jean-Bernard Lasserre. Global optimization with polynomials and the problem of moments. *Journal on optimization*, 2001.

Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas. *Semidefinite Optimization and Convex Algebraic Geometry*. SIAM, 2012.

Kai Kellner. Containment problems for projections of polyhedra and spectrahedra. *arXiv preprint arXiv:1509.02735*, 2015.

Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 1956.

Zhi-Quan Luo and Shuzhong Zhang. On extensions of the Frank-Wolfe theorems. *Computational Optimization and Applications*, 1999.

Giorgio Gallo and Aydin Ülkücü. Bilinear programming: An exact algorithm. *Mathematical Programming*, 1977.

I. Czochralska. Bilinear programming. *Applicationes Mathematicae*, 1982.

Hanif D. Sherali and Amine Alameddine. A new reformulation-linearization technique for bilinear programming problems. *Journal of Global Optimization*, 1992.

Artyom G. Nahapetyan. Bilinear programming, 2009.

Samuel Burer and Adam N. Letchford. On nonconvex quadratic programming with box constraints. *Journal on Optimization*, 2009.

Alexander Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.