Projection: A Unified Approach to Semi-Infinite Linear Programs and Duality in Convex Programming

Amitabh Basu
Kipp Martin
Christopher Thomas Ryan

April 5, 2013

Abstract

Fourier-Motzkin elimination is a projection algorithm for solving finite linear programs. We extend Fourier-Motzkin elimination to semi-infinite linear programs which are linear programs with finitely many variables and infinitely many constraints. Applying projection leads to new characterizations of important properties for primal-dual pairs of semi-infinite programs such as zero duality gap, feasibility, boundedness, and solvability. Extending the Fourier-Motzkin elimination procedure to semi-infinite linear programs yields a new classification of variables that is used to determine the existence of duality gaps. In particular, the existence of what the authors term dirty variables can lead to duality gaps. Our approach has interesting applications in finite-dimensional convex optimization. For example, sufficient conditions for a zero duality gap, such as existence of a Slater point, are reduced to guaranteeing that there are no dirty variables. This leads to completely new proofs of such sufficient conditions for zero duality.

Contents

1 Introduction

Duality is an important theoretical and practical topic in optimization. In order to better understand the structure of an optimization problem (called the primal), and design solution algorithms, it is often useful to consider its dual (or duals). A key determinant of the usefulness of the dual is the *duality gap* which is the difference between the optimal value of a primal and the optimal value of the dual. Establishing that the primal and dual have *zero duality gap* is particularly desirable and is a subject of intense study throughout the field of optimization.

Linear programming is a perfect example. Every linear program has a well-understood dual with the simple property that when the primal is feasible with bounded optimal value, there is zero duality gap. Moreover, optimal solutions to both the primal and dual are guaranteed to exist. For more general problems, additional conditions are needed to establish zero duality gap and the existence of an optimal solution.

Much research has focused on *sufficient* conditions for zero duality gap. Possibly the most well-known sufficient condition for zero duality gap is the *Slater condition* for convex programming. Slater's condition states that when the feasible region of the primal convex program has an interior point (sometimes called a *Slater point*) there is zero duality gap. "Slater-like" conditions are also prevalent in conic programming, where the existence of interior points to the dual conic program guarantees a zero duality gap (see for instance, Gartner and Matousék [8]). Less well-known is the duality theory of *semi-infinite linear programs.* These are linear optimization problems with a finite number of variables and possibly infinitely many constraints. In this paper we use this theory to understand the duality of both convex and conic programs. In semi-infinite linear programming, a variety of sufficient conditions for zero duality gap have been introduced (see for example, Anderson and Nash [1], Charnes, Cooper and Kortanek [2], Duffin and Karlovitz [6], Goberna and López [9], and Karney [10]). We provide an alternate and unifying approach to duality in semi-infinite linear programs.

We extend Fourier-Motzkin elimination (projection) [7, 11] to semi-infinite systems of linear inequalities as a method to study duality. Taking the dictum expressed by Duffin and Karlovitz [6] of "the desirability of omitting topological considerations" to its logical conclusion, the method of projection is purely algebraic. It is simply the aggregation of pairs of linear inequalities using nonnegative multipliers. Applying Fourier-Motzkin elimination to a semi-infinite linear program reveals important properties about the semi-infinite linear program that can only be obtained through this elimination (or projection) process. In particular, Fourier-Motzkin elimination reveals the existence of what the authors term "dirty" variables. Dirty variables are necessary for the existence of a duality gap. The dirty variable characterization also has important implications for finite dimensional problems. For example, sufficient conditions for a zero duality gap in a finite-dimensional convex optimization problem, such as existence of a Slater point, are reduced to guaranteeing that there are no dirty variables in an appropriately defined semi-infinite linear program.

The extension of Fourier-Motzkin elimination to semi-infinite linear programs involves subtleties that do not arise in standard Fourier-Motzkin theory where the number of inequalities is finite. Sections 2 and 3 provide a cogent framework for analyzing semi-infinite linear programs. Applying projection leads to new characterizations of important properties for primal-dual pairs of semi-infinite programs such as zero duality gap, feasibility and boundedness, and solvability. These results have implications for finite-dimensional conic linear programs and convex optimization. See Section 5 and Section 6, respectively. Applications of the results from Section 3 to the generalized Farkas' theorem and additional sufficient conditions for zero duality gap in semi-infinite linear programs are in Section 7 and Section 8, respectively. Concluding remarks are in Section 9.

We begin with a brief notation review and a summary of our results.

Notation

Let Y be a vector space. The *algebraic dual* of Y, denoted Y', is the set of linear functionals with domain Y. Let $\psi \in Y'$. The evaluation of ψ at y is denoted by $\langle y, \psi \rangle$; that is, $\langle y, \psi \rangle = \psi(y)$.

Let P be a convex cone in Y. A convex cone P is *pointed* if and only if $P \cap -P = \{0\}$. A pointed convex cone P in Y defines a vector space ordering $\succeq_P$ of Y, with $y \succeq_P y'$ if $y - y' \in P$. The *dual cone* of P is $P' = \{\psi \in Y : \langle y, \psi \rangle \geq 0 \text{ for all } y \in P\}$. Elements of P' are called *positive linear functionals* in Y. A cone P is *reflexive* if $P'' = P$ under the natural embedding of $Y \hookrightarrow Y''$.

Let A be a linear mapping from vector space X to vector space Y. The *algebraic adjoint* $A' : Y' \to X'$ is defined by $A'(\psi) = \psi \circ A$ and satisfies $\langle x, A'(\psi) \rangle = \langle A(x), \psi \rangle$ where $\psi \in Y'$ and $x \in X$.

Given any set I, $\mathbb{R}^I$ denotes the vector space of real-valued functions u with domain I, i.e., $u : I \to \mathbb{R}$. For $u \in \mathbb{R}^I$ the *support* of u is the set $\text{supp}(u) = \{i \in I : u(i) \neq 0\}$. The subspace $\mathbb{R}^{(I)}$ are those functions in $\mathbb{R}^I$ with finite support. Let $\geq$ denote the standard vector space ordering on $\mathbb{R}^I$. That is, $u \geq v$ if and only if $u(i) \geq v(i)$ for all $i \in I$. The subspace $\mathbb{R}^{(I)}$ inherits this ordering. Let $\mathbb{R}^I_+$ (resp. $\mathbb{R}^{(I)}_+$) denote the pointed cone of $u \in \mathbb{R}^I$ (resp. $u \in \mathbb{R}^{(I)}_+$) with $u \geq 0$. Using the standard embedding of $\mathbb{R}^{(I)}$ into $(\mathbb{R}^I)'$ for $u \in \mathbb{R}^I$ and $v \in \mathbb{R}^{(I)}$, write $\langle u, v \rangle = \sum_{i \in I} u(i)v(i)$. The latter sum is well-defined since v has finite support.

For all $h \in I$, define a function $e^h \in \mathbb{R}^I$ by $e^h(i) = 1$ if $h = i$, and $e^h(i) = 0$ if $h \neq i$ for all $i \in I$. When $I = \{1, 2, \ldots, n\}$, $\mathbb{R}^I$ is $\mathbb{R}^n$ and $e^1, e^2, \ldots e^n$ correspond to the standard unit vectors of $\mathbb{R}^n$.

The optimal value of optimization problem (*) is denoted by $v(*)$.

Our results

The main topic of study is the semi-infinite program

$$\begin{array}{rll} \inf_{x \in \mathbb{R}^n} & c^\top x & \\ \text{s.t.} & \sum_{k=1}^n a^k(i) x_k \geq b(i) & \text{for } i \in I \end{array} \tag{SILP}$$

where I is an arbitrary (potentially infinite) index set, $c \in \mathbb{R}^n$, and $b, a^k \in \mathbb{R}^I$ for $k = 1, \ldots, n$, and its *finite support dual*

$$\begin{array}{rrll} \sup & \sum_{i \in I} b(i)v(i) & & \\ \text{s.t.} & \sum_{i \in I} a^k(i)v(i) & = c_k & \text{for } k = 1, \ldots, n \\ & v & \in \mathbb{R}^{(I)}_+. & \end{array} \tag{FDSILP}$$

Our main results on this primal-dual pair are summarized in Table 1 (see page 23). These include a sufficient condition for primal solvability (Theorem 3.10) and characterizations of both dual solvability (Theorem 3.20) and zero duality gap (Theorem 3.21). Here, zero duality gap means $v(\text{SILP}) = v(\text{FDSILP})$ when (SILP) is feasible.

We identify a special class of semi-infinite linear programs, termed *tidy* semi-infinite linear programs, where zero duality gap is guaranteed to hold (Theorem 3.24). The name *tidy* comes from the fact that the Fourier-Motzkin elimination procedure eliminates (or "cleans up") all primal decision variables. In our terminology, there are no "dirty" decision variables.

Theorem 3.24. If (SILP) is feasible and tidy then

(i) (SILP) is solvable,

(ii) (FDSILP) is feasible and bounded,

(iii) there is a zero duality gap for the primal-dual pair (SILP) and (FDSILP).

In particular, the method of projection is used to prove a result due to Duffin and Karlovitz [6] on duality gaps for semi-infinite linear programs with bounded feasible regions by showing such semi-infinite linear programs are tidy (Theorem 3.25 establishes a slightly more general result).

A number of sufficient conditions for zero duality gap in semi-infinite linear programs due to Karney [10] also follow directly from our results in Section 3. This is shown in Section 8.

Applications in Convex Optimization

The theory of tidy semi-infinite linear programs is leveraged to establish important duality results in conic and convex programming. In conic programming the standard primal is

$$\begin{array}{rl} \inf_{x\in X} & \langle x, \phi\rangle \\ \text{s.t.} & A(x) \succeq_P d \end{array} \tag{ConLP}$$

where X and Y are vector spaces, $A : X \to Y$ is a linear mapping, $d \in Y$, P is a pointed convex cone in Y and ϕ is a linear functional on X. The standard dual (also a conic program) is

$$\begin{array}{rl} \sup_{\psi\in Y'} & \langle d, \psi\rangle \\ \text{s.t.} & A'(\psi) = \phi \\ & \psi \in P'. \end{array} \tag{ConLPD}$$

We study a semi-infinite linear program that is equivalent to (ConLP) and use the method of projection to give a new proof of the following well-known duality result for conic programs.

Theorem 5.12 (Slater's theorem for conic programs)**.** Let X and Y be finite-dimensional vector spaces, and let P be reflexive. Assume the primal conic program (ConLP) is feasible. Suppose there exists $\psi^* \in \text{int}(P')$ with $A'(\psi^*) = c$. Then the primal-dual pair (ConLP)-(ConLPD) has a zero duality gap. Moreover, the primal is solvable.

Our proof uses the interior point ψ^* to construct a set of constraints that show the associated semi-infinite linear program is tidy. Thus, zero duality gap and primal solvability are established in a transparent "algebraic" manner. Alternate proofs (see for instance Gartner and Matousék [8]) are based on a intermediate result called *regular duality*, a core theorem in its own right. According to regular duality, the primal optimal value corresponds to limiting values of sequences of points in the dual space, which are dual feasible in a limiting sense. Regular duality is also established in this paper as a consequence of our method of projection (Theorem 5.10), but this intermediate step is unnecessary in our proof of Theorem 5.12.

In addition, we prove Theorem 5.8 below. A more restricted version of this result is known in the classical conic programming literature (see for example Duffin [5]). Our result is obtained with a completely new proof using projection techniques.

Theorem 5.8 (Zero duality gap via boundedness)**.** Let X be finite-dimensional. If P is reflexive and there exists a scalar γ such the set $\{x : A(x) \succeq_P d \text{ and } \langle x, \phi \rangle \leq \gamma\}$ is nonempty and bounded, then there is no duality gap for the primal-dual pair (ConLP)-(ConLPD).

In particular, conic programs with bounded feasible regions always have zero duality gaps.

Next, consider the following general convex program

$$\begin{array}{rrcll} \max_{x \in \mathbb{R}^n} & f(x) & & & \\ \text{s.t.} & g_i(x) & \geq & 0 & \text{for } i = 1, \ldots, p \\ & x & \in & \Omega & \end{array} \tag{CP}$$

where $f(x)$ and $g_i(x)$ for $i = 1, \ldots, p$ are concave functions, and Ω is a closed, convex set. Define the Lagrangian function $L(\lambda) := \max\{f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) \,:\, x \in \Omega\}$. The Lagrangian dual is

$$\inf_{\lambda \geq 0} L(\lambda). \tag{LD}$$

Slater's condition is a key result in finite-dimensional convex programming.

Theorem 6.4 (Slater's theorem for convex programs)**.** Suppose the convex program (CP) is feasible and bounded. Moreover, suppose there exists $x^* \in \Omega$ such that $g_i(x^*) > 0$ for all $i = 1, \ldots, p$. Then there is zero duality gap between the convex program (CP) and its Lagrangian dual (LD). Moreover, there exists $\lambda^* \geq 0$ such that $v(\text{LD}) = L(\lambda^*)$, i.e., the Lagrangian dual is solvable.

Our proof uses the fact that the Slater point x^* corresponds to a useful constraint in the semi-infinite linear program representing the Lagrangian dual. The structure of this constraint implies the boundedness of the feasible region for a fixed objective value. By Theorem 3.25, this implies zero duality gap and dual solvability. As in the case of conic programs, the result is established in a transparent "algebraic" manner using the method of projection.

Beyond these results in conic and convex programming, the method of projection is used to elegantly prove several foundational results for semi-infinite linear programs. In the process, new structural insights are given. These results include finite approximability of semi-infinite linear programs (see our Theorem 8.3 that generalizes Theorem 2.1 in Karney [10]) and the generalized Farkas' theorem for infinite systems of linear inequalities (see our Theorem 7.1 and Theorem 3.1 in Goberna and López [9]). Goberna and López use the latter result as the main tool for deriving their own set of necessary and sufficient conditions for zero duality in semi-infinite linear programs. Thus, our methodology can, in principle, be used as an alternate starting point to derive their results.

2 Fourier-Motzkin elimination

In this section we extend Fourier-Motzkin elimination to semi-infinite linear systems. For background on Fourier-Motzkin elimination applied to finite linear systems see Fourier [7], Motzkin [11], and Williams [12]. In this section, Fourier-Motzkin elimination elimination is used to characterize the feasibility and boundedness of semi-infinite systems of linear inequalities. In addition, useful properties are shown about the Fourier-Motzkin multipliers which appear while aggregating constraints. These properties prove critical in our approach to duality theory.

Consider the semi-infinite linear system

$$a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n \geq b(i) \quad \text{for } i \in I \tag{2.1}$$

where I is an arbitrary index set. Denote the set of $(x_1, \dots, x_n) \in \mathbb{R}^n$ that satisfy these inequalities by Γ. The projection of Γ into the subspace of $\mathbb{R}^n$ spanned by $\{e^j\}_{j=2}^n$ is

$$P(\Gamma; x_1) := \{(x_2, x_3, \dots, x_n) \in \mathbb{R}^{n-1} \,:\, \exists x_1 \in \mathbb{R} \text{ s.t. } (x_1, x_2, \dots, x_n) \in \Gamma\}. \tag{2.2}$$

Under certain conditions, the projection $P(\Gamma; x_1)$ is characterized by aggregating inequalities in the original system. Define the sets

$$\begin{aligned} \mathcal{H}_+(k) &:= \{i \in I \mid a^k(i) > 0\} \\ \mathcal{H}_-(k) &:= \{i \in I \mid a^k(i) < 0\} \\ \mathcal{H}_0(k) &:= \{i \in I \mid a^k(i) = 0\} \end{aligned} \tag{2.3}$$

based on the coefficients of variable x_k in (2.1).

For now, assume $\mathcal{H}_+(1)$ and $\mathcal{H}_-(1)$ are both nonempty. As in the finite case, eliminate variable x_1 by adding all possible pairs of inequalities with one inequality in $\mathcal{H}_+(1)$ and the other from $\mathcal{H}_-(1)$. Since there are potentially infinitely many constraints this may involve aggregating an infinite number of pairs. The resulting system is

$$\sum_{k=2}^{n} a^k(i)x_k \geq b(i) \qquad \text{for } i \in \mathcal{H}_0(1) \tag{2.4}$$

$$\sum_{k=2}^{n} \left(\frac{a^k(p)}{a^1(p)} - \frac{a^k(q)}{a^1(q)} \right) x_k \geq \frac{b(p)}{a^1(p)} - \frac{b(q)}{a^1(q)} \qquad \text{for } p \in \mathcal{H}_+(1) \text{ and } q \in \mathcal{H}_-(1). \tag{2.5}$$

Denote the set of $(x_2, \dots, x_n) \in \mathbb{R}^{n-1}$ that satisfy the constraints in (2.4)-(2.5) by $FM(\Gamma; x_1)$.

Remark 2.1. One way to view the inequalities (2.5) is the following : pick a pair (p, q) of inequalities with $p \in \mathcal{H}_+(1)$ and $q \in \mathcal{H}_-(1)$. Then form a new constraint by multiplying the first constraint by $\frac{1}{a^1(p)}$, multiplying the second constraint by $-\frac{1}{a^1(q)}$, and adding them together. This "eliminates" x_1 from this pair of constraints. Of course, one can achieve this by choosing any common multiple of $\frac{1}{a^1(p)}$ and $-\frac{1}{a^1(q)}$ as the multipliers prior to adding them together, and achieve a "scaled" inequality describing the same halfspace (with x_1 "eliminated"). $\triangleleft$

A key result is that the inequalities in (2.4)-(2.5) describe the projected set $P(\Gamma; x_1)$.

Theorem 2.2. If $\mathcal{H}_+(1)$ and $\mathcal{H}_-(1)$ are both nonempty, then $P(\Gamma; x_1) = FM(\Gamma; x_1)$.

Proof. Since $\mathcal{H}_+(1)$ and $\mathcal{H}_-(1)$ are both nonempty,

$$\begin{aligned} & (x_2, x_3, \dots, x_n) \in P(\Gamma; x_1) \\ \Leftrightarrow\ & \exists x_1 \in \mathbb{R} \text{ such that } a^1(i)x_1 + a^2(i)x_2 + \dots + a^n(i)x_n \geq b(i) \text{ for } i \in I \\ \Leftrightarrow\ & \exists x_1 \in \mathbb{R} \text{ such that } \left\{ \begin{array}{l} \sum_{k=2}^n a^k(i)x_k \geq b(i) \quad \forall i \in \mathcal{H}_0 \text{ and} \\ x_1 \geq \frac{b(p)}{a^1(p)} - \sum_{k=2}^n \frac{a^k(p)}{a^1(p)} x_k,\ \forall p \in \mathcal{H}_+(1) \text{ and} \\ x_1 \leq \frac{b(q)}{a^1(q)} - \sum_{k=2}^n \frac{a^k(q)}{a^1(q)} x_k,\ \forall q \in \mathcal{H}_-(1) \end{array} \right\} \\ \Leftrightarrow\ & \left\{ \begin{array}{l} \sum_{k=2}^n a^k(i)x_k \geq b(i) \quad \forall i \in \mathcal{H}_0 \text{ and} \\ \frac{b(p)}{a^1(p)} - \sum_{k=2}^n \frac{a^k(p)}{a^1(p)} x_k \leq \frac{b(q)}{a^1(q)} - \sum_{k=2}^n \frac{a^k(q)}{a^1(q)} x_k \quad \forall p \in \mathcal{H}_+(1), \forall q \in \mathcal{H}_-(1) \end{array} \right\} \\ \Leftrightarrow\ & (x_2, x_3, \dots, x_n) \in FM(\Gamma; x_1). \end{aligned}$$

Note that the second to last equivalence holds because both $\mathcal{H}_+(1)$ and $\mathcal{H}_-(1)$ are nonempty. $\square$

Equally as important to our theory is how "dual information" is accrued during the process of elimination. The following result captures the essence of this idea.

Corollary 2.3. If $\mathcal{H}_+(1)$ and $\mathcal{H}_-(1)$ are both nonempty, then there exists an index set $\tilde{I}$ and $u^h \in \mathbb{R}_+^{(I)}$ for $h \in \tilde{I}$ such that the projection $P(\Gamma; x_1)$ is

$$P(\Gamma; x_1) = \{(x_2, \ldots, x_n) \mid \tilde{a}^2(h)x_2 + \cdots + \tilde{a}^n(h)x_n \geq \tilde{b}(h) \text{ for } h \in \tilde{I}\}$$

where $\tilde{b}, \tilde{a}^2, \ldots, \tilde{a}^n \in \mathbb{R}^{\tilde{I}}$ are given by

(i) $\tilde{b}(h) = \langle b, u^h \rangle$ for all $h \in \tilde{I}$,

(ii) $\tilde{a}^k(h) = \langle a^k, u^h \rangle$ for all $k = 2, \ldots, n$ and $h \in \tilde{I}$,

(iii) $\langle a^1, u^h \rangle = 0$ for all $h \in \tilde{I}$.

Proof. By Theorem 2.2, $P(\Gamma; x_1) = FM(\Gamma; x_1)$. Show that $FM(\Gamma; x_1)$ has the required representation. Since $\mathcal{H}_+(1)$ and $\mathcal{H}_-(1)$ are both nonempty, take $\tilde{I} = \mathcal{H}_0(1) \cup (\mathcal{H}_+(1) \times \mathcal{H}_-(1))$. For each $h \in \mathcal{H}_0(1)$, take $u^h \in \mathbb{R}_+^{(I)}$ as the function with value 1 at h and 0 otherwise. For each $h = (p, q) \in \mathcal{H}_+(1) \times \mathcal{H}_-(1)$, take $u^h \in \mathbb{R}_+^{(I)}$ as the function $u^h : I \to \mathbb{R}$ defined by

$$u^h(i) = \begin{cases} \frac{1}{a^1(p)}, & \text{when } i = p \\ -\frac{1}{a^1(q)}, & \text{when } i = q \\ 0, & \text{otherwise.} \end{cases}$$

Now define $\tilde{b}, \tilde{a}^2, \ldots, \tilde{a}^n$ using the equations from (i) and (ii) in the statement of the corollary. The proof is then complete by observing that $FM(\Gamma; x_1) = \{(x_2, \ldots, x_n) \mid \tilde{a}^2(h)x_2 + \cdots + \tilde{a}^n(h)x_n \geq \tilde{b}(h) \text{ for } h \in \tilde{I}\}$ with these definitions. □

Below is a formal statement of Fourier-Motzkin elimination, which applies the above procedure sequentially for each variable.

Fourier-Motzkin Elimination Procedure

Input: A semi-infinite linear inequality system

$$a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n \geq b(i) \quad \text{for } i \in I.$$

Output: A semi-infinite linear inequality system

$$\tilde{a}^\ell(h)x_\ell + \tilde{a}^{\ell+1}(h)x_{\ell+1} + \cdots + \tilde{a}^n(h)x_n \geq \tilde{b}(h) \quad \text{for } h \in \tilde{I}. \tag{2.6}$$

The variables $x_\ell, \ldots, x_n$ form a subset of the variables of the input system relabeled according to a permutation $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$. We allow $\ell \in \{1, \ldots, n, n+1\}$, interpreting $\ell = n+1$ to mean that the left-hand side is zero. We also output a set of vectors $\{u^h \in \mathbb{R}_+^{(I)} : h \in \tilde{I}\}$.

Procedure:

1. Initialization: $\mathcal{D} \leftarrow \{1, \ldots, n\}$, $\tilde{I} \leftarrow I$, $\tilde{a}^k \leftarrow a^k$ for all $k \in \mathcal{D}$, $\tilde{b} \leftarrow b$, and $j \leftarrow 1$. For each $h \in \tilde{I} = I$, set $u^h \leftarrow e^h$.

2. Elimination: While $(j \leq n)$ do:

   a. Define the sets $\mathcal{H}_+(j)$, $\mathcal{H}_-(j)$ and $\mathcal{H}_0(j)$ as follows.

$$\mathcal{H}_+(j) := \{h \in \tilde{I} \mid \tilde{a}^j(h) > 0\}$$
$$\mathcal{H}_-(j) := \{h \in \tilde{I} \mid \tilde{a}^j(h) < 0\}$$
$$\mathcal{H}_0(j) := \{h \in \tilde{I} \mid \tilde{a}^j(h) = 0\}$$

   b. If $\mathcal{H}_+(j) \neq \emptyset$ and $\mathcal{H}_-(j) \neq \emptyset$ do:

(i) Set $\tilde{I} \leftarrow \mathcal{H}_0(j) \cup [\mathcal{H}_+(j) \times \mathcal{H}_-(j)]$ and $\mathcal{D} \leftarrow \mathcal{D} \setminus \{j\}$.

(ii) For each $k \in \mathcal{D}$ define $\hat{a}^k : \tilde{I} \to \mathbb{R}$ by

$$\hat{a}^k(h) := \begin{cases} \tilde{a}^k(h) & \text{for } h \in \mathcal{H}_0(j) \\ \frac{\tilde{a}^k(p)}{\tilde{a}^j(p)} - \frac{\tilde{a}^k(q)}{\tilde{a}^j(q)} & \text{for } h = (p,q) \in \mathcal{H}_+(j) \times \mathcal{H}_-(j) \end{cases}$$

(iii) For each $h \in \tilde{I}$, define $\hat{u}^h \in \mathbb{R}^{(I)}_+$ by

$$\hat{u}^h := \begin{cases} u^h & \text{for } h \in \mathcal{H}_0(j) \\ \frac{1}{\tilde{a}^j(p)} u^p - \frac{1}{\tilde{a}^j(q)} u^q & \text{for } h = (p,q) \in \mathcal{H}_+(j) \times \mathcal{H}_-(j) \end{cases}$$

(iv) For each $k \in \mathcal{D}$, set $\tilde{a}^k \leftarrow \hat{a}^k$. For each $h \in \tilde{I}$, set $u^h \leftarrow \hat{u}^h$.

(v) Define $\hat{b} : \tilde{I} \to \mathbb{R}$ by

$$\hat{b}(h) := \begin{cases} \tilde{b}(h) & \text{for } h \in \mathcal{H}_0(j) \\ \frac{\tilde{b}(p)}{\tilde{a}^j(p)} - \frac{\tilde{b}(q)}{\tilde{a}^j(q)} & \text{for } h = (p,q) \in \mathcal{H}_+(j) \times \mathcal{H}_-(j) \end{cases}$$

and set $\tilde{b} \leftarrow \hat{b}$.

end do.

c. If $\mathcal{H}_+(j) \cup \mathcal{H}_-(j) = \emptyset$ then set $\mathcal{D} \leftarrow \mathcal{D} \setminus \{j\}$.

d. $j \leftarrow j + 1$.

end do.

3. Output formatting: Upon termination $\mathcal{D}$ is either empty or, for some $\ell \in \{1, \ldots, n\}$, can be written $\mathcal{D} = \{d_1, \ldots, d_{n-\ell+1}\}$ where $d_i \in \{1, \ldots, n\}$ with $d_i \le d_j$ for $i \le j$. Let $\overline{\mathcal{D}} = \{1, \ldots, n\} \setminus \mathcal{D} = \{\bar{d}_1, \ldots, \bar{d}_{\ell-1}\}$ where $\bar{d}_i \in \{1, \ldots, n\}$ and $\bar{d}_i \le \bar{d}_j$ for $i \le j$. In other words, $\ell - 1$ variables were eliminated and the rest $n - \ell + 1$ variables indexed by the indices in $\mathcal{D}$ are not eliminated.

   a. If $\mathcal{D} = \emptyset$, output the system

$$0 \ge \tilde{b}(h) \quad \text{for } h \in \tilde{I}.$$

   b. Else if $\mathcal{D} \neq \emptyset$, reassign the indices in $\mathcal{D}$ by $d_i \leftarrow \ell - 1 + i$ for $i = 1, \ldots, n - \ell + 1$. If $\overline{\mathcal{D}}$ is nonempty, reassign the indices in $\overline{\mathcal{D}}$ by $\bar{d}_i \leftarrow i$ for $i = 1, \ldots, \ell - 1$. This defines the permutation π described in the output. Now, construct the system

$$\tilde{a}^\ell(h)x_\ell + \tilde{a}^{\ell+1}(h)x_{\ell+1} + \ldots + \tilde{a}^n(h)x_n \ge \tilde{b}(h) \quad \text{for } h \in \tilde{I}.$$

Definition 2.4 (Clean and dirty variables)**.** At the end of the Fourier-Motzkin procedure, the variables $x_1, \ldots, x_{\ell-1}$ are called *clean* variables and the variables $x_\ell, \ldots, x_n$ are called *dirty* variables. Thus, a dirty variable is one that the Fourier-Motzkin procedure could not eliminate and a clean variable is one that the procedure could eliminate.

Definition 2.5 (Canonical form)**.** A semi-infinite linear system (2.1) is said to be in *canonical form* if the permutation π output by the Fourier-Motzkin elimination is the identity permutation.

Lemma 2.6. For every semi-infinite linear system, there exists a permutation of the variables that puts it into canonical form. Moreover, if you apply the Fourier-Motzkin procedure to the original system and to the permuted system, they result in the same system of inequalities in the output.

Proof. The permutation output by the Fourier-Motzkin procedure is one such desired permutation. $\square$

Remark 2.7. In light of Lemma 2.6, we assume without loss, that semi-infinite linear systems are always given in canonical form before applying the Fourier-Motzkin elimination procedure. There may exist multiple permutations of the variables which put a given semi-infinite system into canonical form. Moreover, two different permutations may lead to systems in canonical form with a different number of clean and dirty variables. For our purposes, this will not make a difference and *any* permutation that puts the semi-infinite system into a canonical form will suffice.

Definition 2.8. The finite support element, u^h for every $h \in \tilde{I}$, that is generated by the Fourier-Motzkin elimination procedure is called a *Fourier-Motzkin elimination multiplier*, or simply a *multiplier*.

The key property of the Fourier-Motzkin elimination procedure is that it characterizes geometric projections. For $\ell \leq n$ define

$$P(\Gamma; x_1, \ldots, x_{\ell-1}) := \{(x_\ell, \ldots, x_n) \in \mathbb{R}^{n-\ell+1} \,:\, \exists x_1, \ldots, x_{\ell-1} \text{ s.t. } (x_1, \ldots, x_{\ell-1}, x_\ell, \ldots, x_n) \in \Gamma\}.$$

Theorem 2.9. Apply the Fourier-Motzkin elimination procedure with input inequality system (2.1) to produce output system (2.6). For all $h \in \tilde{I}$, the finite-support multipliers $u^h \in \mathbb{R}^{(I)}_+$ generated by the Fourier-Motzkin procedure satisfy

(i) $\tilde{b}(h) = \langle b, u^h \rangle$,

(ii) $\tilde{a}^k(h) = \langle a^k, u^h \rangle$ for all $k = \ell, \ldots, n$, and

(iii) $\langle a^k, u^h \rangle = 0$ for all $k = 1, \ldots, \ell - 1$.

In addition, if not all variables are eliminated, and in the output system (2.6) $\ell \leq n$, then

$$P(\Gamma; x_1, \ldots, x_{\ell-1}) = \{(x_\ell, \ldots, x_n) \mid (2.6) \text{ holds}\}.$$

Proof. If $\ell = 1$, then only Step 2d. of the Fourier-Motzkin elimination procedure is executed and the original system remains unchanged so $\tilde{I} = I$, $\tilde{a}^k = a^k$, $k = 1, \ldots, n$ and $\tilde{b} = b$. Based on the initialization step, $u^h = e^h$ for $h \in \tilde{I}$ and (i)-(iii) follow. If $\ell \geq 2$, since the system is in canonical form, the result follows from recursively applying Corollary 2.3. $\square$

Corollary 2.10 (Clean projection)**.** Let (2.1) be a semi-infinite linear system and let $1 \leq M < \min\{\ell, n\}$ where ℓ is the index of the first dirty variable in the output system (2.6). After the while loop in Step 2 of the Fourier-Motzkin elimination procedure iterates M times, we have the following intermediate system (recall (2.1) is assumed to be in canonical form)

$$\tilde{a}^{M+1}(h)x_{M+1} + \tilde{a}^{M+2}(h)x_{M+2} + \cdots + \tilde{a}^n(i)x_n \geq \tilde{b}(h) \quad \text{ for } h \in \tilde{I}. \tag{2.7}$$

Then

$$P(\Gamma; x_1, \ldots, x_M) = \{(x_{M+1}, \ldots, x_n) \mid (2.7) \text{ holds}\}.$$

Proof. Follows from a finite number of applications of Corollary 2.3. $\square$

Partition the index set $\tilde{I}$ in (2.6), into two sets $H_1 := \{h \in \tilde{I} : \tilde{a}^k(h) = 0$ for all $k \in \{\ell, \ldots, n\}$ and $H_2 := \tilde{I} \setminus H_1$. Rewrite (2.6) as

$$0 \;\geq\; \tilde{b}(h) \quad \text{for } h \in H_1 \tag{2.8}$$

$$\tilde{a}^\ell(h)x_\ell + \tilde{a}^{\ell+1}(h)x_{\ell+1} + \cdots + \tilde{a}^n(h)x_n \;\geq\; \tilde{b}(h) \quad \text{for } h \in H_2. \tag{2.9}$$

If $H_2 = \emptyset$ (that is, $\ell = n + 1$), then system (2.8)-(2.9) is a *clean* system. Otherwise, if $H_2 \neq \emptyset$, (2.8)-(2.9) is a *dirty* system. In a dirty system, for any $k \in \{\ell, \ldots, n\}$, either $\tilde{a}^k(h) \geq 0$ for all $h \in H_2$, or $\tilde{a}^k(h) \leq 0$ for all $h \in H_2$. Moreover, $\sum_{k=\ell}^n |\tilde{a}^k(h)| > 0$ for $h \in H_2$.

Definition 2.11. Given a dirty system (2.8)-(2.9) and a real number $\delta \geq 0$, let $x(\delta;\ell)$ denote the tuple $(\bar{x}_\ell, \ldots, \bar{x}_n)$ where for each $k \in \{\ell, \ldots, n\}$, $\bar{x}_k = \delta$ if $\tilde{a}^k(h) \geq 0$ for all $h \in H_2$ and $\bar{x}_k = -\delta$ otherwise. Let $x_k(\delta;\ell)$ denote the kth entry of $x(\delta;\ell)$.

Remark 2.12. When I is a finite set, the concept of a dirty variable is unnecessary. In the finite case, there is always a value of δ such that $x(\delta,\ell)$ is a feasible solution to (2.9). It is therefore legitimate to drop the constraints indexed by H_2 from further consideration. Therefore, when implementing the Fourier-Motzkin procedure in the finite case, if variable x_k is dirty, then one would drop all the constraints h for which $\tilde{a}^k(h) > 0$ (or $\tilde{a}^k(h) < 0$). ◁

Theorem 2.13 (Feasibility)**.** Applying Fourier-Motzkin elimination to (2.1) results in system (2.8)-(2.9). If $H_2 \neq \emptyset$ then the system is feasible (i.e. Γ is nonempty) if and only if

(i) $\tilde{b}(h) \leq 0$ for all $h \in H_1$, and

(ii) $\sup_{h \in H_2} \tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}^k(h)| < \infty$.

Moreover, if $H_2 = \emptyset$ then Γ is nonempty if and only if (i) holds.

Proof. If $H_2 \neq \emptyset$, then Γ is nonempty if and only if $P(\Gamma; x_1, \ldots, x_{\ell-1})$ is nonempty. By Theorem 2.9, $P(\Gamma; x_1, \ldots, x_{\ell-1})$ is defined by (2.8)-(2.9). Therefore, it suffices to show (2.8)-(2.9) has a feasible solution if and only if conditions i) and ii) hold.

($\Longrightarrow$) Assume $\bar{x}_\ell, \ldots, \bar{x}_n$ be a feasible solution to (2.8)-(2.9). Let $\delta = \max\{|\bar{x}_\ell|, |\bar{x}_{\ell+1}|, \ldots, |\bar{x}_n|\}$. First, all inequalities in H_1 are satisfied and this gives condition i). Moreover, for all $h \in H_2$, $\tilde{b}(h) \leq \sum_{k=\ell}^n \tilde{a}^k(h)\bar{x}_k \leq \sum_{k=\ell}^n |\tilde{a}^k(h)||\bar{x}_k| \leq \delta(\sum_{k=\ell}^n |\tilde{a}^k(h)|)$. This implies for every $h \in H_2$, $\tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}^k(h)| \leq \delta < \infty$ and this gives condition ii).

($\Longleftarrow$) Assume i) and ii) hold. By ii) there exits a $\delta \geq \max\{0, \sup_{h \in H_2} \tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}^k(h)|\}$. Show that $x(\delta;\ell)$ is a feasible solution to (2.8)-(2.9). It suffices to show (2.9), since (2.8) is implied by condition i). For any $h \in H_2$, $\sum_{k=\ell}^n \tilde{a}^k(h)x_k(\delta;\ell) = \delta(\sum_{k=\ell}^n |\tilde{a}^k(h)|) \geq \tilde{b}(h)$, where the last inequality follows from the fact that $\delta \geq \sup_{h \in H_2} \tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}^k(h)|$. Thus, $x(\delta;\ell)$ is a feasible solution.

Now consider the case $H_2 = \emptyset$. If the inequalities in the original system hold (that is, $\Gamma \neq \emptyset$) then the inequalities $0 \geq \tilde{b}(h)$ for $h \in H_1$ must also hold, since these inequalities are consequences of the original system. Thus, (i) holds. Conversely, suppose $\tilde{b}(h) \leq 0$ for all $h \in H_1$. Now, just before x_n is eliminated in the Fourier-Motzkin elimination procedure (x_n must be eliminated since $H_2 = \emptyset$) the system stored in the algorithm (after a scaling as stated in Remark 2.1) is

$$0 \geq \hat{b}(h) \text{ for } h \in \mathcal{H}_0(n) \tag{2.10}$$

$$x_n \geq \hat{b}(h') \text{ for } h' \in \mathcal{H}_+(n) \tag{2.11}$$

$$-x_n \geq \hat{b}(h'') \text{ for } h'' \in \mathcal{H}_-(n). \tag{2.12}$$

When x_n is eliminated, system (2.8)-(2.9) is derived with $\tilde{b}(h) = \hat{b}(h') + \hat{b}(h'')$ where $h = (h', h'')$ for $h' \in \mathcal{H}_+(n)$ and $h'' \in \mathcal{H}_-(n)$. By hypothesis, $\tilde{b}(h) \leq 0$ for all $h \in H_1$ and this implies $\hat{b}(h') \leq -\hat{b}(h'')$. Then there exists an x_n such that $\hat{b}(h') \leq x_n \leq -\hat{b}(h'')$ for all $h' \in \mathcal{H}_+(n)$ and $h'' \in \mathcal{H}_-(n)$ and this x_n that satisfies (2.11) and (2.12). Note that (2.10) holds by hypothesis since $\mathcal{H}_0(n) \subseteq H_1$. Thus, (2.10)-(2.12) is a feasible system. By Corollary 2.10 this system is the projection $P(\Gamma; x_1, \ldots, x_{n-1})$. Thus, $P(\Gamma; x_1, \ldots, x_{n-1})$ is nonempty and therefore Γ is nonempty. □

Remark 2.14. In the proof of Theorem 2.8 it was shown that when Γ is nonempty and

$$\delta \geq \max\{0, \sup_{h \in H_2} \tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}^k(h)|\},$$

the tuple $x(\delta;\ell)$ as defined in Definition 2.11 is feasible to (2.8)-(2.9) and thus can be extended to a feasible vector in Γ. This fact is used in later development. ◁

We next characterize the boundedness of the feasible set Γ.

Theorem 2.15 (Boundedness)**.** If Γ is a bounded set, then applying Fourier-Motzkin elimination to the system (2.1) that defines Γ, gives the system (2.8)-(2.9) with $H_2 = \emptyset$.

Proof. Prove the contrapositive and assume H_2 is nonempty. This implies the existence of dirty variables. Since $\Gamma \neq \emptyset$, $\sup_{h \in H_2} \tilde{b}(h)/\sum_{k=\ell}^{n} |\tilde{a}^k(h)| < \infty$ by Theorem 2.13. For any

$$\delta \geq \max\{0, \sup_{h \in H_2} \tilde{b}(h)/\sum_{k=\ell}^{n} |\tilde{a}^k(h)|\},$$

$x(\delta;\ell)$ is feasible for the system (2.8)-(2.9) by Remark 2.14. The components of $x(\delta;\ell)$ become arbitrarily large in absolute value as $\delta \to \infty$. Since (2.8)-(2.9) describes the projection $P(\Gamma; x_\ell, \ldots, x_n)$ there are feasible solutions for Γ which take arbitrarily large values in the components $x_\ell, \ldots, x_n$. This contradicts the fact that Γ is bounded. □

Example 2.16. The opposite implication in Theorem 2.15 does not hold in general. For example, consider the linear system $-x_1 - x_2 \geq 0, x_1 + x_2 \geq 0$. The feasible region is the unbounded line $x_1 + x_2 = 0$; but H_2 is empty when applying the Fourier-Motzkin elimination procedure because the output is the degenerate system $0 \geq 0$. ◁

Theorem 2.17 below provides a very useful property about Fourier-Motzkin elimination multipliers that plays a pivotal role in establishing duality results in Section 3.3.

Theorem 2.17. Applying Fourier-Motzkin elimination to (2.1) gives (2.6). Let $\bar{u} \in \mathbb{R}^{(I)}_+$ such that $\langle a^k, \overline{u}\rangle = 0$ for $k = 1, \ldots, M$ with $\ell - 1 \leq M \leq n$. Then, there exists a nonempty finite index set $\bar{I} \subseteq \tilde{I}$ such that for all $h \in \bar{I}$ the Fourier-Motzkin multipliers u^h satisfy $\langle a^k, u^h\rangle = 0$ for $k = 1, \ldots, M$. Moreover, there exist scalars $\lambda_h \geq 0$ for $h \in \bar{I}$ so that $\overline{u} = \sum_{h \in \bar{I}} \lambda_h u^h$.

Proof. Proceed by induction on n. First prove the inductive step on n and then the $n = 1$ step. Assume the result is true for an $n-1$ variable system and show that this implies the result is true for an n variable system. Apply Fourier-Motzkin elimination to the $n-1$ variable system

$$a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_{n-1} \geq b(i) \quad \text{for } i \in I, \tag{2.13}$$

obtained by dropping the last column in system (2.1). The result is

$$\hat{a}^{\ell_{n-1}}(h)x_{\ell_{n-1}} + \hat{a}^{\ell_{n-1}+1}(h)x_{\ell_{n-1}+1} + \cdots + \hat{a}^{n-1}(h)x_{n-1} \geq \hat{b}(h) \quad \text{for } h \in \hat{I} \tag{2.14}$$

where ℓ_{n-1} denotes the first index of the dirty variables in the Fourier-Motzkin elimination output. There are two cases to consider.

<u>*Case 1: $M < n$.*</u> Variable $\ell - 1$ is the last clean variable in (2.1). The assumption that $M < n$, together with the theorem hypothesis that $\ell - 1 \leq M$, implies $\ell - 1 < n$ so the last clean variable in (2.1) is strictly less than variable n. Then the last clean variable in (2.13) is the same as the last clean variable in (2.1). This implies Fourier-Motzkin elimination applied to both systems yields identical multiplier vectors. Invoke the induction hypothesis for the $n-1$ variable system (2.13). Denote by M_{n-1} the value of M and ℓ_{n-1} the value of ℓ when the induction hypothesis is applied to (2.13). Since $\ell - 1 \leq M < n$ and the last clean variable for (2.1) is the same as the last clean variable for (2.13), it is valid to set $M_{n-1} = M$ and $\ell_{n-1} - 1 = \ell - 1$. Because Fourier-Motzkin elimination applied to both systems yields identical multiplier vectors, the induction hypothesis implies that the Fourier-Motzkin multipliers also satisfy the requirements of the theorem for the n variable system.

<u>*Case 2: $M = n$.*</u> In this case $\langle a^k, \overline{u}\rangle = 0$ for $k = 1, \ldots, n$. Therefore it is valid to apply the induction hypothesis to the $n-1$ variable system (2.13) with $M_{n-1} = n-1$ and $\ell_{n-1} = \min\{\ell, n\}$.

Then there exists a finite index set $\{1, \ldots, t\} = \overline{I} \subseteq \hat{I}$ and multipliers w^j such that $\langle a^k, w^j \rangle = 0$ for all $k = 1, \ldots, n-1$ and $j = 1, \ldots, t$ and scalars $\hat{\alpha}_j \geq 0$ such that

$$\bar{u} = \textstyle\sum_{j=1}^{t} \hat{\alpha}_j w^j. \tag{2.15}$$

The multipliers w^j, $j = 1, \ldots, t$, are used to show that column n is clean in (2.1) and that $\overline{u}$ is a nonnegative combination of multipliers that result from eliminating this last column n.

By Theorem 2.9, the scalars $\langle a^n, w^j \rangle$ are among the coefficients on x_n before that variable is processed when Fourier-Motzkin elimination is applied to (2.1). Either

(i) $\langle a^n, w^j \rangle = 0$ for $j = 1, \ldots, t$

or

(ii) there exists $j^+, j^- \in \{1, \ldots, t\}$ such that $\langle a^n, w^{j^+} \rangle > 0$ and $\langle a^n, w^{j^-} \rangle < 0$.

Conditions (i) and (ii) are exhaustive since $0 = \langle a^n, \bar{u} \rangle = \sum_{j=1}^{t} \hat{\alpha}_j \langle a^n, \ w^j \rangle$ for $\hat{\alpha}_j \geq 0$ and so if $\langle a^n, w^j \rangle \geq 0$ for $j = 1, \ldots, t$ (similiarly $\langle a^n, w^j \rangle \leq 0$ for $j = 1, \ldots, t$) then $\langle a^n, w^j \rangle = 0$ for $j = 1, \ldots, t$.

If (i) holds, and $\langle a^n, w^j \rangle = 0$ for $j = 1, \ldots, t$, then $\langle a^k, w^j \rangle = 0$ for $j = 1, \ldots, t$, $k = 1, \ldots, n$; thus w^j for $j = 1, \ldots, t$ are Fourier-Motzkin multipliers when Fourier-Motzkin is applied to (2.1), and $\bar{u} = \sum_{j=1}^{t} \hat{\alpha}_j w^j$ and Case 2 is proved.

If (ii) holds then x_n is a clean variable with respect to the system produced during the Fourier-Motzkin procedure before variable x_n is processed: it has both a positive coefficient $\langle a^n, w^{j^+} \rangle > 0$ and a negative coefficient $\langle a^n, w^{j^-} \rangle < 0$.

Define three sets J^+, J^- and J^0 where $j \in J^+$ if $\langle a^n, w^j \rangle > 0$, $j \in J^-$ if $\langle a^n, w^j \rangle < 0$ and $j \in J^0$ if $\langle a^n, w^j \rangle = 0$. In case (ii) both J^+ and J^- are nonempty. As discussed in case (i), for $j \in J^0$, w^j is already a Fourier-Motzkin multiplier which satisfies $\langle a^k, w^j \rangle = 0$ for $k = 1, \ldots, M$ and so they meet the specifications of the theorem. Now consider the w^j for $j \in J^+$ and $j \in J^-$. Each pair of $(j^+, j^-) \in J^+ \times J^-$ yields a final Fourier-Motzkin multiplier which is a conic combination of w^{j^+} and w^{j^-}. In order to simplify the analysis, normalize the w^j so that $\langle a^n, w^j \rangle = 1$ for $j \in J^+$ and $\langle a^n, w^j \rangle = -1$ for $j \in J^-$. Let α_j be the multipliers after the corresponding scaling of $\hat{\alpha}_j$ for $j \in J^+ \cup J^-$. With this scaling, from Step 2.b.(iii) of the Fourier-Motzkin procedure, the $u^{j^+j^-} = w^{j^+} + w^{j^-}$ for all $(j^+, j^-) \in J^+ \times J^-$ are among the Fourier-Motzkin elimination multipliers for the full system. It suffices to show that there exists multipliers $\theta_{j^+j^-}$ such that

$$\bar{u} = \sum_{j \in J^0} \hat{\alpha}_j w_j + \sum_{j^+ \in J^+} \sum_{j^- \in J^-} \theta_{j^+j^-} u^{j^+j^-} \tag{2.16}$$

and

$$\langle a^k, u^{j^+j^-} \rangle = \langle a^k, w^{j^+} + w^{j^-} \rangle = 0 \text{ for } k = 1, \ldots, M. \tag{2.17}$$

Condition (2.17) follows since $\langle a^k, w^j \rangle = 0$ for $k = 1, \ldots, M-1$ and $\langle a^n, w^{j^+} \rangle = -\langle a^n, w^{j^-} \rangle = 1$ for all $j^+ \in J^+$ and $j^- \in J^-$.

To establish (2.16) consider a transportation linear program with supply nodes indexed by J^+ and demand nodes indexed by J^-. Each supply node $j \in J^+$ has supply α_j. Each demand node $j \in J^-$ has demand $-\alpha_j$. Since

$$0 = \langle a^n, \bar{u} \rangle = \langle a^n, \sum_{j \in J^0} \alpha_j w^j + \sum_{j \in J^+ \cup J^-} \alpha_j w^j \rangle = \sum_{j \in J^+ \cup J^-} \alpha_j \langle a^n, w^j \rangle = \sum_{j \in J^+} \alpha_j - \sum_{j \in J^-} \alpha_j$$

total supply is equal to total demand. Therefore the transportation problem has a feasible solution θ_{j^+,j^-} which is the flow from supply node j^+ to demand node j^-. This feasible flow satisfies

$$\begin{aligned}\sum_{j^-\in J^-}\theta_{j^+,j^-} &= \alpha_{j^+}, \quad \text{for } j^+\in J^+\\ \sum_{j^+\in J^+}\theta_{j^+,j^-} &= \alpha_{j^-}, \quad \text{for } j^-\in J^-\end{aligned}$$

and so

$$\begin{aligned}\sum_{j^+\in J^+}\sum_{j^-\in J^-}\theta_{j^+,j^-}u^{j^+,j^-} &= \sum_{j^+\in J^+}\sum_{j^-\in J^-}\theta_{j^+,j^-}(w^{j^+}+w^{j^-})\\ &= \sum_{j^+\in J^+}\sum_{j^-\in J^-}\theta_{j^+,j^-}w^{j^+}+\sum_{j^+\in J^+}\sum_{j^-\in J^-}\theta_{j^+,j^-}w^{j^-}\\ &= \sum_{j^+\in J^+}\alpha_{j^+}w^{j^+}+\sum_{j^-\in J^-}\alpha_{j^-}w^{j^-}.\end{aligned}$$

Combining this with (2.15) yields (2.16).

Next, consider the base case $n = 1$. By hypothesis, this forces $M = 1$, i.e., $\langle a^1, \overline{u}\rangle = 0$. If the coefficient of x_1 is zero for all the constraints indexed by $\operatorname{supp}(\overline{u})$, then the Fourier-Motzkin procedure initialization step gives multiplers $w^j = e^j$, $j \in \operatorname{supp}(\overline{u})$. Then $\overline{u} = \sum_{j\in\operatorname{supp}(\overline{u})} \overline{u}(j)w^j$. Otherwise, if variable x_1 has nonzero coefficients in the system indexed by $\operatorname{supp}(\overline{u})$, it follows that variable x_1 has both positive and coefficients in this system, since $\overline{u}$ is nonegative and $\langle a^1, \overline{u}\rangle$. Define the usual multiplier vector for each pair of positive and negative coefficients. Again, assume without the loss, the rows are scaled such that the positive coefficients are 1 and the negative coefficients -1. Create a transportation problem as above where each node has supply of $\overline{u}_j$ if j corresponds to a row with $+1$, or demand $-\overline{u}_j$ corresponds to a row with a -1. Solving this transportation problem, and using the same logic as before, gives the coefficients $\theta_{j^+,j^{-1}}$ to be used on the multiplier vectors u^{j^+,j^-} in order to generate $\overline{u}$. This completes the proof. □

3 Solvability and duality theory using projection

3.1 The projected system

The semi-infinite linear program

$$\begin{aligned}\inf_{x\in\mathbb{R}^n} \quad & c^\top x\\ \text{s.t.} \quad & a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n \geq b(i) \quad \text{for } i \in I\end{aligned} \tag{SILP}$$

is the primal problem. Reformulate (SILP) as

$$\inf \quad z \tag{3.1}$$

$$\text{s.t.} \quad -c_1x_1 - c_2x_2 - \cdots - c_nx_n + z \geq 0 \tag{3.2}$$

$$a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n \geq b(i) \quad \text{for } i \in I. \tag{3.3}$$

Let $\Lambda \subseteq \mathbb{R}^{n+1}$ denote the set of $(x_1, \ldots, x_n, z)$ that satisfy (3.2)-(3.3). Consider z as the $(n+1)$st variable and constraint (3.2) as the 0th constraint in the system.

Remark 3.1. This formulation allows for $v(\text{SILP}) = +\infty$ or $v(\text{SILP}) = -\infty$. The former arises when the feasible region is empty. The latter signifies that the primal is unbounded. ◁

Applying Fourier-Motzkin elimination procedure to the input system (3.2)-(3.3) gives the output system (2.6), rewritten as

$$\begin{array}{rcll} 0 & \geq & \tilde{b}(h), & h \in I_1 \\ \tilde{a}^{\ell}(h)x_{\ell} + \tilde{a}^{\ell+1}(h)x_{\ell+1} + \cdots + \tilde{a}^{n}(h)x_n & \geq & \tilde{b}(h), & h \in I_2 \\ z & \geq & \tilde{b}(h), & h \in I_3 \\ \tilde{a}^{\ell}(h)x_{\ell} + \tilde{a}^{\ell+1}(h)x_{\ell+1} + \cdots + \tilde{a}_{n}(h)x_n + z & \geq & \tilde{b}(h), & h \in I_4 \end{array} \tag{3.4}$$

where I_1, I_2, I_3 and I_4 are disjoint with $\tilde{I} = I_1 \cup \cdots \cup I_4$. Note that z can never be eliminated, so system (3.4) is always dirty and $I_3 \cup I_4 \neq \emptyset$. This formatting also assumes that every time a constraint involving z was aggregated, a multiplier of 1 is used. This can always be achieved by Remark 2.1. It is possible that all other variables can be eliminated when $I_2 = I_4 = \emptyset$ (that is, $\ell = n+1$). By construction, $|\sum_{k=\ell}^{n} \tilde{a}^k(h)| > 0$ for all $h \in I_2 \cup I_4$.

By Theorem 2.9, system (3.4) describes the projection $P(\Lambda; x_1, \ldots, x_{\ell-1})$ (recall the assumption that the system of inequalities (3.2)-(3.3) is in canonical form). Therefore, to solve (SILP) it suffices to consider the optimization problem

$$\begin{array}{rl} \inf_{z, x_\ell, \ldots, x_n} & z \\ \text{s.t.} & (3.4). \end{array} \tag{3.5}$$

A further step (Lemma 3.7) is to examine the geometric projection of Λ onto the z-variable space in terms of the data from the output system (3.4). It is easier to characterize the boundedness and solvability of (SILP) in this one-dimensional space.

3.2 Primal results

3.2.1 Primal feasibility

Feasibility of (SILP) is determined by looking at the constraints indexed by I_1, I_2, I_3 and I_4.

Theorem 3.2 (Primal Feasibility)**.** (SILP) is feasible if and only if

(i) $\tilde{b}(h) \leq 0$ for all $h \in I_1$,

(ii) $\displaystyle\sup_{h \in I_2} \frac{\tilde{b}(h)}{\sum_{k=\ell}^{n} |\tilde{a}^k(h)|} < \infty$,

(iii) $\displaystyle\sup_{h \in I_3} \tilde{b}(h) < \infty$,

(iv) $\displaystyle\sup_{h \in I_4} \frac{\tilde{b}(h)}{\sum_{k=\ell}^{n} |\tilde{a}^k(h)| + 1} < \infty$.

Proof. The result follows directly from applying Theorem 2.13 to the dirty system (3.4) with $H_1 = I_1$ and $H_2 = I_2 \cup I_3 \cup I_4$. □

Remark 3.3. Some readers may find it counter-intuitive that primal feasibility involves consideration of constraints involving z (those indexed by I_3 and I_4), a variable that does not appear in the initial description (3.3) of the feasible region. However, this is indeed the case since during Fourier-Motzkin elimination the constraints involving only $x_1, \ldots, x_n$ can be mixed with constraints involving z in such a way that careful consideration of all four types of constraints (those indexed by I_1 through I_4) is necessary. Example 3.4 below demonstrates this.

Alternatively, one could apply Fourier-Motzkin elimination to (2.1) and obtain (2.8)-(2.9). By Theorem 2.13, the conditions for feasibility are

(i) $\tilde{b}(h) \leq 0$ for all $h \in H_1$,

(ii) $\sup_{h \in H_2} \tilde{b}(h) / \sum_{k=\ell}^{n} |\tilde{a}^k(h)| < \infty$.

The key point is that the inequalities indexed by H_1 and H_2 in (2.8)-(2.9) are not identical to the inequalities indexed by I_1 and I_2 in (3.4).

Example 3.4. Consider the following instance of (SILP)

$$
\begin{array}{rrrrl}
\inf & x_2 & & & \\
& x_2 & \geq & i & \text{for } i = 1, 2, \ldots \quad (3.6)\\
x_1 & & \geq & 0 & \quad (3.7)\\
-x_1 & + x_3 & \geq & 0 & \quad (3.8)\\
x_1 & - x_3 & \geq & -1. & \quad (3.9)
\end{array}
$$

The two sets of interest are: $\Gamma = \left\{x \in \mathbb{R}^3 : (3.6) - (3.9)\right\}$ and $\Lambda = \{(x, z) \in \Gamma : -x_2 + z \geq 0\}$. Applying Fourier-Motzkin elimination to the inequalities describing Γ gives

$$
\begin{array}{rrrl}
0 & \geq & -1 & \\
x_2 & \geq & i & \text{for } i = 1, 2, \ldots \\
x_3 & \geq & 0. &
\end{array}
$$

and x_2 is a dirty variable. Applying Fourier-Motzkin elimination to the inequalities describing Λ gives

$$
\begin{array}{rrrl}
0 & \geq & -1 & \\
x_3 & \geq & 0 & \\
z & \geq & i & \text{for } i = 1, 2, \ldots
\end{array}
$$

and x_2 is now a clean variable. Note that $|H_2| = \infty$ but $|I_2| = 1$. $\triangleleft$

Corollary 3.5 below states some consequences of primal feasibility for (SILP) which are useful later. The proof is analogous to the proof of Theorem 2.13. First introduce the function

$$
\omega(\delta) := \sup_{h \in I_4} \left\{\tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)|\right\} \tag{3.10}
$$

that is used throughout the paper. Note ω can take values in the extended reals. If $I_4 = \emptyset$ then $\omega(\delta) = -\infty$, and $\omega(\delta)$ can diverge to $+\infty$. Observe ω is a nonincreasing function of δ since $\sum_{k=\ell}^{n} |\tilde{a}^k(h)| \geq 0$.

Corollary 3.5. If (SILP) is feasible then

(i) $\delta_2 := \sup_{h \in I_2} \dfrac{\tilde{b}(h)}{\sum_{k=\ell}^{n} |\tilde{a}^k(h)|} < \infty$,

(ii) $\delta_3 := \sup_{h \in I_3} \tilde{b}(h) < \infty$,

(iii) $\lim_{\delta \to \infty} \omega(\delta) < \infty$,

(iv) $(x(\bar{\delta}; \ell), \bar{z}) \in P(\Lambda; x_1, \ldots, x_{\ell-1})$ for all $\bar{\delta}, \bar{z} \in \mathbb{R}$ such that $\bar{\delta} \geq \max\{0, \delta_2\}$ and $\bar{z} \geq \max\{\delta_3, \omega(\bar{\delta})\}$. Moreover, by conditions i), ii) and iii) above, at least one such pair $(\bar{\delta}, \bar{z})$ of real number exists.

Proof. Conditions i)-ii) follow immediately from Theorem 3.2. Condition iii) follows from the claim below and condition iv) of Theorem 3.2.

Claim 3.6. $\sup_{h \in I_4} \dfrac{\tilde{b}(h)}{\sum_{k=\ell}^{n} |\tilde{a}^k(h)| + 1} < \infty \iff \lim_{\delta \to \infty} \omega(\delta) < \infty.$

Proof of Claim 3.6. ($\Longrightarrow$) Let $\bar{\delta} = \sup_{h\in I_4} \tilde{b}(h)/(\sum_{k=\ell}^n |\tilde{a}^k(h)| + 1) < \infty$. This implies $\bar{\delta} \geq \tilde{b}(h)/(\sum_{k=\ell}^n |\tilde{a}^k(h)| + 1)$ for every $h \in I_4$. Rearranging, $\bar{\delta}(\sum_{k=\ell}^n |\tilde{a}^k(h)| + 1) \geq \tilde{b}(h)$, which implies $\bar{\delta} \geq \tilde{b}(h) - \bar{\delta}(\sum_{k=\ell}^n |\tilde{a}^k(h)|)$ for all $h \in I_4$. Thus, $\bar{\delta} \geq \sup\{\tilde{b}(h) - \bar{\delta}(\sum_{k=\ell}^n |\tilde{a}^k(h)|) : h \in I_4\} = \omega(\bar{\delta})$. Thus, $\infty > \bar{\delta} \geq \omega(\bar{\delta})$ and since $\omega(\delta)$ is a nonincreasing function, this yields $\lim_{\delta\to\infty} \omega(\delta) < \infty$.

($\Longleftarrow$) Since $\lim_{\delta\to\infty} \omega(\delta) < \infty$ and $\omega(\delta)$ is a nonincreasing function, there exists $\bar{\delta} < \infty$ such that $\bar{\delta} \geq \omega(\bar{\delta})$. The reasoning is as follows: $\lim_{\delta\to\infty} \omega(\delta) < \infty$ implies there exists a $\hat{\delta}$ such that $\omega(\hat{\delta}) = c < \infty$. Take $\bar{\delta} \geq \max\{\hat{\delta}, c\}$. Since $\omega(\delta)$ is nonincreasing in δ, $\omega(\bar{\delta}) \leq \omega(\hat{\delta}) = c \leq \bar{\delta}$. Now, because $\bar{\delta} \geq \omega(\bar{\delta})$ it follows $\bar{\delta} \geq \sup\{\tilde{b}(h) - \bar{\delta}(\sum_{k=\ell}^n |\tilde{a}^k(h)|) : h \in I_4\}$. Hence, $\bar{\delta} \geq \tilde{b}(h) - \bar{\delta}(\sum_{k=\ell}^n |\tilde{a}^k(h)|)$ for all $h \in I_4$. Rearranging, $\bar{\delta} \geq \tilde{b}(h)/(\sum_{k=\ell}^n |\tilde{a}^k(h)| + 1)$ for every $h \in I_4$ and so $\infty > \bar{\delta} \geq \sup_{h\in I_4} \tilde{b}(h)/(\sum_{k=\ell}^n |\tilde{a}^k(h)| + 1)$. □

Prove condition iv) in the statement of the corollary by verifying that the constraints indexed by I_1, I_2, I_3 and I_4 are satisfied by $(x(\bar{\delta}; \ell), \bar{z}) \in P(\Lambda; x_1, \ldots, x_{\ell-1})$ when $\bar{\delta} \geq \max\{0, \delta_2\}$ and $\bar{z} \geq \max\{\delta_3, \omega(\bar{\delta})\}$. Since δ_2, δ_3 and $\lim_{\delta\to\infty} \omega(\delta)$ are all finite, and $\omega(\delta)$ is a nonincreasing function, there exists at least one such pair $(\bar{\delta}, \bar{z})$ of real numbers.

Since (SILP) is feasible, the constraints in I_1 are satisfied by condition i) in Theorem 3.2. By definition, $\delta_2 \geq \tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}_k(h)|$ for all $h \in I_2$, which implies $\bar{\delta}\sum_{k=\ell}^n |\tilde{a}_k(h)| \geq \tilde{b}(h)$ for all $h \in I_2$. Since $\sum_{k=\ell}^n \tilde{a}_k(h)x_k(\bar{\delta}; \ell) = \bar{\delta}\sum_{k=\ell}^n |\tilde{a}_k(h)|$ by construction of $x(\bar{\delta}; \ell)$, $(x(\bar{\delta}; \ell), \bar{z})$ satisfies the constraints indexed by I_2 in (3.4).

Since $\bar{z} \geq \delta_3$, all the constraints indexed by I_3 are satisfied. Finally, since $\bar{z} \geq \omega(\bar{\delta})$, $\bar{z} \geq \sup_{h\in I_4}\{\tilde{b}(h) - \bar{\delta}\sum_{k=\ell}^n |a^k(h)|\}$ and so for all $h \in I_4$, $\bar{z} + \sum_{k=\ell}^n a^k(h)x_k(\bar{\delta}; \ell) = \bar{z} + \bar{\delta}\sum_{k=\ell}^n |a^k(h)| \geq \tilde{b}(h)$. Conclude $(x(\bar{\delta}; \ell), \bar{z}))$ satisfies the constraints indexed by I_4, and therefore feasible to (3.4). Thus, $(x(\bar{\delta}; \ell), \bar{z})) \in P(\Lambda; x_1, \ldots, x_\ell)$ by Theorem 2.9. □

3.2.2 Primal boundedness

To establish boundedness and solvability, we start by giving a characterization of the closure of the projection of the feasible region described by (3.4) onto the z-variable space.

Lemma 3.7. Assume (SILP) is feasible and applying Fourier-Motzkin elimination to (3.2)-(3.3) gives (3.4). Let $P(\Lambda; x_1, \ldots, x_n)$ denote the projection of Λ into the z-variable space. Then, the closure of $P(\Lambda; x_1, \ldots, x_n)$ is given by the system of inequalities

$$z \geq \sup_{h\in I_3} \tilde{b}(h) \tag{3.11}$$

$$z \geq \lim_{\delta\to\infty} \omega(\delta). \tag{3.12}$$

Proof. Since (SILP) is feasible, conditions ii) and iii) in Corollary 3.5 imply that $\sup_{h\in I_3} \tilde{b}(h) < \infty$ and $\lim_{\delta\to\infty} \omega(\delta) < \infty$. Let δ_2 and δ_3 be as defined in i)-ii) of Corollary 3.5.

First, suppose $\bar{z}$ satisfies (3.11)-(3.12) and show $\bar{z} \in \text{cl}(P(\Lambda; x_1, \ldots, x_n))$. Consider the following two exhaustive cases.

Case 1: $\bar{z} > \lim_{\delta\to\infty} \omega(\delta)$. Since $\omega(\delta)$ is nonincreasing in δ, there exists $\hat{\delta} \in \mathbb{R}$ such that $\bar{z} > \omega(\hat{\delta})$. Choose $\bar{\delta} \geq \max\{0, \hat{\delta}, \delta_2\}$. By (3.11), $\bar{z} \geq \sup_{h\in I_3} \tilde{b}(h) = \delta_3$. Also, $\bar{z} > \omega(\hat{\delta}) \geq \omega(\bar{\delta})$ since $\omega(\delta)$ is nonincreasing. Thus, $(x(\bar{\delta}; \ell), \bar{z})$ satisfies the hypotheses of condition iv) of Corollary 3.5. Therefore $(x(\bar{\delta}; \ell), \bar{z}) \in P(\Lambda; x_1, \ldots, x_{\ell-1})$ and this implies $\bar{z} \in P(\Lambda; x_1, \ldots, x_n)$.

Case 2: $\bar{z} = \lim_{\delta\to\infty} \omega(\delta)$. Since $\omega(\delta)$ nonincreasing in δ, there exists a sequence of real numbers $(\bar{\delta}_m)_{m\in\mathbb{N}}$ such that for every $m \in \mathbb{N}$, $\bar{\delta}_m \geq \max\{0, \delta_2\}$ and $z_m := \omega(\bar{\delta}_m) \to \bar{z}$. Since $\omega(\delta)$ is nonincreasing and $\bar{z}$ satisfies (3.11), $z_m = \omega(\bar{\delta}_m) \geq \lim_{\delta\to\infty} \omega(\delta) = \bar{z} \geq \sup_{h\in I_3} \tilde{b}(h)$. Hence $z_m \geq \max\{\delta_3, \omega(\bar{\delta}_m)\}$ and by Corollary 3.5(iv), $(x(\bar{\delta}_m; \ell), z_m) \in P(\Lambda; x_1, \ldots, x_{\ell-1})$. Therefore $z_m \in P(\Lambda; x_1, \ldots, x_n)$ and $z_m \to \bar{z}$. This implies $\bar{z} \in \text{cl}(P(\Lambda; x_1, \ldots, x_n))$.

Conversely, let $\bar{z} \in \text{cl}(P(\Lambda; x_1, \ldots, x_n))$ and show $\bar{z}$ satisfies (3.11) and (3.12). Since $z \in \text{cl}(P(\Lambda; x_1, \ldots, x_n))$ there exists a sequence $z_m \in P(\Lambda; x_1, \ldots, x_n)$ where $z_m \to \bar{z}$. Since $z_m \in P(\Lambda; x_1, \ldots, x_n)$ there exists an $x^m = (x_\ell^m, \ldots, x_n^m)$ such that (x^m, z_m) satisfies the constraints of system (3.4). This implies $z_m \geq \sup_{h \in I_3} \tilde{b}(h)$. Since $z_m \to \bar{z}$, conclude $\bar{z} \geq \sup_{h \in I_3} \tilde{b}(h)$.

Also, since (x^m, z_m) satisfies (3.4), $z_m \geq \sup_{h \in I_4} \{\tilde{b}(h) - \sum_{k=\ell}^n a^k(h) x_k^m\}$. Letting $\bar{\delta}_m = \max_{k=\ell,\ldots,n} |x_k^m|$ gives $z_m \geq \sup_{h \in I_4} \{\tilde{b}(h) - \sum_{k=\ell}^n a^k(h) x_k^m\} \geq \sup_{h \in I_4} \{\tilde{b}(h) - \bar{\delta}_m \sum_{k=\ell}^n |a^k(h)|\} = \omega(\bar{\delta}_m)$. Thus, $z_m \geq \omega(\bar{\delta}_m) \geq \lim_{\delta \to \infty} \omega(\delta)$ for all m, where the last inequality holds since $\omega(\delta)$ is nonincreasing. Since $z_m \to \bar{z}$, conclude $\bar{z} \geq \lim_{\delta \to \infty} \omega(\delta)$. Hence $\bar{z}$ is a feasible solution to system (3.11)-(3.12). $\square$

By Lemma 3.7, if (SILP) is feasible, then its optimal value is found by solving the optimization problem

$$\begin{array}{ll} \inf_z & z \\ \text{s.t.} & (3.11) - (3.12). \end{array} \tag{3.13}$$

This follows because the optimal value of a continuous objective function over a convex feasible region is the same the optimal value of that objective when optimized over the closure of the region. The next two results follow directly from this observation.

Lemma 3.8. If (SILP) is feasible then $v(\text{SILP}) = \max\left\{\sup_{h \in I_3} \tilde{b}(h), \lim_{\delta \to \infty} \omega(\delta)\right\}$.

Theorem 3.9 (Primal boundedness)**.** A feasible (SILP) is bounded if and only if $I_3 \neq \emptyset$ or $\lim_{\delta \to \infty} \omega(\delta) > -\infty$.

Proof. By contrapositive in both directions. By Lemma 3.8, $v(\text{SILP}) = -\infty$ if and only if $\max\{\sup_{h \in I_3} \tilde{b}(h), \lim_{\delta \to \infty} \omega(\delta)\} = -\infty$ if and only if $\sup_{h \in I_3} \tilde{b}(h) = -\infty$ and $\lim_{\delta \to \infty} \omega(\delta) = -\infty$. Note that $\sup_{h \in I_3} \tilde{b}(h) = -\infty$ if and only if $I_3 = \emptyset$. $\square$

3.2.3 Primal solvability

An instance of (SILP) is solvable if the infimum value of its objective is attained. Note that an optimal solution $v(\text{SILP})$ may exist to (3.13) even though an optimal solution to (SILP) does not exist (see for instance Example 3.11 below). This is due to the fact that (3.13) is an optimization problem over the *closure* of the projection $P(\Lambda; x_1, \ldots, x_n)$, and hence an optimal solution to (3.5) may exist in the closure but not the projection itself. Thus, the solution may not "lift" to an optimal solution of (SILP). A sufficient condition for when this "lifting" can occur is given in Theorem 3.10.

Theorem 3.10 (Primal solvability)**.** If (SILP) is feasible and $\sup_{h \in I_3} \tilde{b}(h) > \lim_{\delta \to \infty} \omega(\delta)$, then (SILP) has an optimal solution with value $v(\text{SILP}) = \sup_{h \in I_3} \tilde{b}(h)$.

Proof. Let $z^* = v(\text{SILP})$. Since (SILP) is feasible, by part (ii) of Corollary 3.5 it follows that $\infty > \sup_{h \in I_3} \tilde{b}(h) > \lim_{\delta \to \infty} \omega(\delta)$. Moreover, by Lemma 3.8, $z^* = \sup_{h \in I_3} \tilde{b}(h)$. Let δ_2 be as defined in Corollary 3.5. Since $\omega(\delta)$ is a nonincreasing function, there exists $\delta^* \geq \max\{0, \delta_2\}$ such that $\omega(\delta^*) < \sup_{h \in I_3} \tilde{b}(h) = z^*$. Then, $(x(\delta^*; \ell), z^*)$ satisfies the hypotheses of condition iv) in Corollary 3.5 and so $(x(\delta^*; \ell), z^*) \in P(\Lambda; x_\ell, \ldots, x_n)$, showing that there exists a feasible point $(x_1, \ldots, x_n, z)$ in Λ where $z = z^*$. Thus there is a feasible point for (SILP) with value $z^* = v(\text{SILP})$. $\square$

In light of the previous result, an immediate question is whether primal solvability holds when $\lim_{\delta \to \infty} \omega(\delta) = \sup_{h \in I_3} \tilde{b}(h)$. The following two examples demonstrate that such problems can be either solvable or not solvable.

Example 3.11. Consider the following instance of (SILP)

$$\begin{array}{lll} \inf x_1 & & \\ \quad x_1 + \frac{1}{t^2}x_2 & \geq & \frac{1}{t^2} + \frac{1}{t} \quad \text{for } t \geq 1 \\ \quad x_1 & \geq & 0. \end{array} \tag{3.14}$$

Applying Fourier-Motzkin elimination to

$$\begin{array}{rcl} -x_1 \qquad + z & \geq & 0 \\ x_1 + \frac{1}{t^2}x_2 \qquad & \geq & \frac{1}{t^2} + \frac{1}{t} \quad \text{for } t \geq 1 \\ x_1 \qquad & \geq & 0 \end{array} \tag{3.15}$$

yields (by eliminating x_1)

$$\begin{array}{rcl} \frac{1}{t^2}x_2 + z & \geq & \frac{1}{t^2} + \frac{1}{t} \quad \text{for } t \geq 1 \\ z & \geq & 0. \end{array} \tag{3.16}$$

The only I_3 constraint is $z \geq 0$ so $\sup_{h \in I_3} \tilde{b}(h) = 0$. Show that for $\delta \geq 3/2$,

$$\omega(\delta) = \sup_{t \geq 1} \left\{ \frac{1}{t^2} + \frac{1}{t} - \frac{\delta}{t^2} \right\} = \sup_{t \geq 1} \left\{ \frac{(1-\delta)}{t^2} + \frac{1}{t} \right\} = \frac{1}{4(\delta - 1)}.$$

When $\delta \geq 1$ and $t \neq 0$, the function $\frac{(1-\delta)}{t^2} + \frac{1}{t}$ is concave and quadratic in $\frac{1}{t}$. The supremum is achieved by $t^* = -2(1-\delta)$. When $\delta \geq 3/2$, $t^* \geq 1$ and substituting the optimal value of t^* into $\frac{(1-\delta)}{t^2} + \frac{1}{t}$ gives $\frac{1}{4(\delta-1)}$. Clearly, $\lim_{\delta \to \infty} \omega(\delta) = 0 = \sup_{h \in I_3} \tilde{b}(h)$ and so by Lemma 3.8 the optimal value is 0.

However, for $z = 0$ the system (3.16) has no possible feasible assignment for x_2. Indeed, for any proposed $\bar{x}_2$ take $t \geq \bar{x}_2$. This implies $\frac{1}{t^2}\bar{x}_2 + 0 \leq \frac{1}{t} < \frac{1}{t^2} + \frac{1}{t}$, which means $(\bar{x}_2, 0)$ is infeasible to (3.16) and the primal is not solvable. ◁

Example 3.12. Consider the following instance of (SILP)

$$\begin{array}{rcl} \inf\ x_1 & & \\ x_1 \qquad & \geq & 0 \\ - x_2 & \geq & -1 \\ x_1 - \frac{1}{i}x_2 & \geq & 0 \quad \text{for } i = 3, 4, \ldots \end{array}$$

Applying Fourier-Motzkin elimination to

$$\begin{array}{rcl} -x_1 \qquad + z & \geq & 0 \\ x_1 \qquad & \geq & 0 \\ - x_2 \qquad & \geq & -1 \\ x_1 - \frac{1}{i}x_2 \qquad & \geq & 0 \quad \text{for } i = 3, 4, \ldots \end{array}$$

yields (after projecting out x_1)

$$\begin{array}{rcl} -x_2 \qquad & \geq & -1 \\ z & \geq & 0 \\ -\frac{1}{i}x_2 + z & \geq & 0 \quad \text{for } i = 3, 4, \ldots \end{array}$$

Observe $I_3 = \{1\}$ and $\sup_{h \in I_3} \tilde{b}(h) = 0$. Note

$$\begin{aligned} \omega(\delta) &= \sup \left\{ \tilde{b}(h) - \delta \sum_{h \in I_4} |\tilde{a}^k(h)| \ : \ h \in I_4 \right\} \\ &= \sup \left\{ 0 - \delta/h \ : \ h = 3, 4, \ldots \right\} = 0. \end{aligned}$$

Thus, $\lim_{\delta \to \infty} \omega(\delta) = 0 = \sup_{h \in I_3} \tilde{b}(h)$. By Lemma 3.8, this implies $v(\text{SILP}) = 0$ and this value is obtained for the feasible solution $x_1 = x_2 = 0$ and the primal is solvable. ◁

3.3 Dual results

The next step is to develop a duality theory for (SILP) using Fourier-Motzkin elimination. The standard dual problem in the semi-infinite linear programming literature (see for instance [2]) is the finite support (Haar) dual introduced in Section 1 and reproduced here for convenience.

$$\begin{array}{lll} \sup & \sum_{i\in I} b(i)v(i) & \\ \text{s.t.} & \sum_{i\in I} a^k(i)v(i) & = c_k \quad \text{for } k=1,\ldots,n \\ & v & \in \mathbb{R}^{(I)}_+ \end{array} \tag{FDSILP}$$

In this section, we characterize when (FDSILP) is feasible, bounded, and solvable. Later in Section 3.3.4 we characterize when there is zero duality gap between (SILP) and (FDSILP); that is, $v(\text{SILP}) = v(\text{FDSILP})$.

As in the case of the primal, allow $v(\text{FDSILP})$ to take on values in the extended reals. (FDSILP) is unbounded when $v(\text{FDSILP}) = \infty$. When $v(\text{FDSILP}) = -\infty$, the problem is infeasible.

In the remainder of this section, assume Fourier-Motzkin elimination has been applied to (3.2)-(3.3) yielding (3.4). Our attention turns to the multipliers generated in Step 2.b.(iii) of the Fourier-Motzkin elimination procedure. These multipliers generate solutions to (FDSILP).

First a small, but important, distinction. The multipliers u^h generating (3.4) are real-valued functions defined on the set $\{0\} \cup I$ where the inequality (3.2) has index 0. However, solutions to (FDSILP) are real-valued functions defined only on I. Thus, it is useful to work with the restriction $v^h : I \to \mathbb{R}$ of u^h to I. That is, $v^h(i) = u^h(i)$ for $i \in I$. Conversely, given a function $v : I \to \mathbb{R}$ and a real number v_0, let $u = (v_0, v)$ denote the *extension* of v onto the index set $\{0\} \cup I$ where $u(0) = v_0$ and $u(i) = v(i)$ for all $i \in I$. Lemma 3.13 gives basic properties of v^h that are used later.

Lemma 3.13. If Fourier-Motzkin elimination is applied to (3.2)-(3.3) yielding (3.4), then

(i) for every $h \in I_1 \cup I_2 \cup I_3 \cup I_4$, $\tilde{b}(h) = \langle b, v^h \rangle$.

(ii) for $h \in I_1$, $u^h(0) = 0$ and v^h is a recession direction for the feasible region of (FDSILP).

(iii) for $h \in I_2$, $u^h(0) = 0$ and v^h satisfies $\sum_{i\in I} a^k(i)v^h(i) = 0$ for $k = 1, \ldots, \ell - 1$, and $\sum_{i\in I} a^k(i)v^h(i) = \tilde{a}^k(h)$ for $k = \ell, \ldots, n$.

(iv) for $h \in I_3$, $u^h(0) = 1$ and v^h is a feasible solution to (FDSILP), and

(v) for $h \in I_4$, $u^h(0) = 1$ and v^h satisfies $\sum_{i\in I} a^k(i)v^h(i) - c_k = 0$ for $k = 1, \ldots, \ell - 1$, and $\sum_{i\in I} a^k(i)v^h(i) - c_k = \tilde{a}^k(h)$ for $k = \ell, \ldots, n$.

Proof. We establish parts (i), (ii) and (iv) only. The other parts are seen analogously.

(i) By Theorem 2.9, for all h: $\tilde{b}(h) = \langle (0, b), u^h \rangle = \langle b, v^h \rangle$.

(ii) The constraints indexed by I_1 do not involve z and so the multipliers u^h for $h \in I_1$ must have $u^h(0) = 0$. By Theorem 2.9(ii), for $h \in I_1$, $0 = \langle (-c_k, a^k), u^h \rangle = \langle a^k, v^h \rangle$ for all $k = 1, \ldots, n$. This implies v^h satisfies $\sum_{i\in I} a^k(i)v^h(i) = 0$. In addition, $u^h \geq 0$ implies $v^h \geq 0$ and v^h is a recession direction for the feasible region of (FDSILP).

(iv) The constraints indexed by I_3 must involve z and so the multipliers u^h for $h \in I_3$ must have $u^h(0) > 0$. Assume $u^h(0) = 1$, which is without loss by Remark 2.1. By Theorem 2.9(ii), for $h \in I_3$, $0 = \langle (-c_k, a^k), u^h \rangle = -c_k + \langle a^k, v^h \rangle$ for all $k = 1, \ldots, n$. This implies v^h satisfies the equality constraints of (FDSILP). In addition, $u^h \geq 0$ implies $v^h \geq 0$ and v^h is a feasible solution to (FDSILP). ☐

3.3.1 Dual feasibility

The next two subsections relate dual feasibility and boundedness to properties of the projected system (3.4). Theorem 2.17 and Lemma 3.13 play pivotal roles in the proofs.

Theorem 3.14 (Dual Feasibility)**.** (FDSILP) is feasible if and only if $I_3 \neq \emptyset$.

Proof. ($\Longrightarrow$) If (FDSILP) is feasible, there is a $\overline{v} \geq 0$ with finite support such that $\sum_{i \in I} a_k(i)\overline{v}_i = c_k, k = 1, \ldots, n$ and this implies $\langle (-c_k, a^k), (1, \overline{v}) \rangle = 0, k = 1, \ldots, n$. Then, by applying Theorem 2.17 to (3.2)-(3.3) with $M = n$, there exists an index set $\bar{I} \subseteq (I_1 \cup I_3)$ and multipliers $u^h : \{0\} \cup I \to \mathbb{R}$ for $h \in \bar{I}$ such that

$$\begin{aligned} (1, \overline{v}) &= \sum_{h \in \bar{I}} \lambda_h u^h \\ &= \sum_{h \in \bar{I} \cap I_1} \lambda_h u^h + \sum_{h \in \bar{I} \cap I_3} \lambda_h u^h \\ &= \sum_{h \in \bar{I} \cap I_1} \lambda_h (0, v^h) + \sum_{h \in \bar{I} \cap I_3} \lambda_h (1, v^h) \end{aligned}$$

where $\lambda_h \geq 0$ for all $h \in \bar{I}$ and v^h is the restriction of u^h onto I. The third equality follows from Lemma 3.13(ii) and (iv). Now, the 1 in the first component of $(1, \overline{v})$ implies that $\bar{I} \cap I_3$ cannot be empty, and hence I_3 cannot be empty.

($\Longleftarrow$) Take any u^h with $h \in I_3$. By Lemma 3.13(iv), v^h is a feasible solution to (FDSILP). ☐

3.3.2 Dual boundedness

To characterize dual boundedess, first establish weak duality.

Lemma 3.15 (Weak Duality)**.** Suppose $\tilde{b}(h) \leq 0$ for all $h \in I_1$. If $\overline{v}$ is a feasible dual solution to problem (FDSILP) then

(i) there exists an $\bar{h} \in I_3$ such that $\tilde{b}(\bar{h}) \geq \langle b, \overline{v} \rangle$,

(ii) $\langle b, \overline{v} \rangle$ is a lower bound on the optimal solution value of (SILP).

Proof. Applying Theorem 2.17 as in the proof of Theorem 3.14 implies there exists an index set $\bar{I} \subseteq I_1 \cup I_3$ such that

$$(1, \overline{v}) = \sum_{h \in \bar{I} \cap I_1} \lambda_h (0, v^h) + \sum_{h \in \bar{I} \cap I_3} \lambda_h (1, v^h).$$

Reasoning about the components of $(1, \bar{v})$ separately gives,

$$\bar{v} = \sum_{h \in \bar{I} \cap I_1} \lambda_h v^h + \sum_{h \in \bar{I} \cap I_3} \lambda_h v^h \tag{3.17}$$

and $1 = \sum_{h \in \bar{I} \cap I_3} \lambda_h$. Lemma 3.13(i) and the hypothesis $\tilde{b}(h) \leq 0$ for all $h \in I_1$ imply $\langle b, v^h \rangle \leq 0$ for all $h \in I_1$. Thus, (3.17) gives $\langle b, \overline{v} \rangle \leq \sum_{h \in \bar{I} \cap I_3} \lambda_h \langle b, v^h \rangle \leq \langle b, v^{\bar{h}} \rangle = \tilde{b}(\bar{h})$ for some $\bar{h} \in \bar{I} \cap I_3$, where the second inequality follows because the λ_h for $h \in I_3$ are nonnegative and sum to 1. This implies i). Now ii) follows immediately from Lemma 3.8. ☐

Theorem 3.16 (Dual boundedness)**.** Suppose (FDSILP) is feasible. Then (FDSILP) is bounded if and only if

(i) $\tilde{b}(h) \leq 0$ for all $h \in I_1$ and

(ii) $\sup_{h \in I_3} \tilde{b}(h) < \infty$.

Proof. ($\Longleftarrow$) By contrapositive. Suppose (FDSILP) is unbounded and show that if condition i) holds, then ii) does not hold. Assume $\tilde{b}(h) \leq 0$ for all $h \in I_1$. Since (FDSILP) is unbounded, for every $M \in \mathbb{N}$ there exists a feasible $\bar{v}_M$ with $\langle b, \bar{v}_M \rangle \geq M$. By Lemma 3.15, there exists some $h_M \in I_3$ such that $\tilde{b}(h_M) \geq \langle b, \bar{v}_M \rangle \geq M$. Thus, $\sup_{h \in I_3} \tilde{b}(h) \geq \tilde{b}(h_M) \geq M$ for all $M \in \mathbb{N}$ and this implies $\sup_{h \in I_3} \tilde{b}(h) = \infty$. Therefore, (ii) does not hold.

($\Longrightarrow$) By contrapositive. Assume condition i) does not hold. Thus, there exists $h^* \in I_1$ such that $\tilde{b}(h^*) > 0$ and by Lemma 3.13(ii), $\langle a^k, v^{h^*} \rangle = 0$ for all $k = 1, \ldots, n$. Now, consider any $\bar{v}$ feasible to (FDSILP), which exists since (FDSILP) is feasible. Then, $\bar{v} + \lambda v^{h^*}$ is also feasible for all $\lambda \geq 0$. Now, the objective value for these feasible solutions equal $\langle b, \bar{v} + \lambda v^{h^*} \rangle = \langle b, \bar{v} \rangle + \lambda \langle b, v^{h^*} \rangle$. Since $\langle b, v^{h^*} \rangle = \tilde{b}(h^*) > 0$, letting $\lambda \to \infty$, yields unbounded values for the objective value of (FDSILP).

Next assume condition ii) does not hold. This implies there is a sequence of $\{h_m\}_{m \in \mathbb{N}}$ in I_3 such that, by Lemma 3.13(i), $\langle b, v^{h_m} \rangle = \tilde{b}(h_m) \to \infty$. By Lemma 3.13(iii), each v^{h_m} is a feasible solution to (FDSILP) and thus (FDSILP) is unbounded. □

Remark 3.17. Observe that there are two distinct ways for a feasible (FDSILP) to be unbounded. The first is when there there is a recession direction to the feasible region that drives the objective value to $+\infty$. From Lemma 3.13(ii) every $h \in I_1$ yields a recession direction v^h. In addition, if $\tilde{b}(h) > 0$ then $\langle b, v^h \rangle > 0$ and so moving within the feasible region along recession direction v^h drives the objective to $+\infty$. This argument was given in full detail in the proof of Theorem 3.16.

Contrary to our intuition from finite dimensions, the second way (FDSILP) may have an unbounded objective value can occur when the feasible region itself is bounded. This happens when there are no recession directions and $\sup_{h \in I_3} \tilde{b}(h) = \infty$. This occurs when (FDSILP) has a sequence of feasible solutions whose values converge to $+\infty$. Consider the following example

$$\begin{array}{rcll} \inf x_1 & & & \\ \text{s.t. } x_1 & \geq & i & \text{for } i \in \mathbb{N} \end{array}$$

with finite support dual

$$\begin{array}{rcll} \sup \sum_{i \in \mathbb{N}} i v(i) & & & \\ \text{s.t. } \sum_{i \in \mathbb{N}} v(i) & = & 1 & \\ v(i) & \geq & 0 & \text{for } i \in \mathbb{N} \end{array}$$

The feasible region of the finite support dual is bounded (note that $0 \leq v(i) \leq 1$ for all i) and there is no recession direction. However, the problem is still unbounded. Consider the sequence of feasible extreme point solutions e^m. Clearly, $\sup_{m \to \infty} \sum_{i \in \mathbb{N}} i e^m(i) = m \to \infty$. Thus (FDSILP) is unbounded.

Fourier-Motzkin elimination can identify which of the conditions of Theorem 3.16 are violated and result in an unbounded problem. Applying Fourier-Motzkin elimination to the system

$$\begin{array}{rcll} -x_1 + z & \geq & 0 & \\ x_1 & \geq & i & \text{for } i = 1, 2, \ldots \end{array}$$

yields (after eliminating x_1)

$$z \geq i \quad \text{for } i = 1, 2, \ldots.$$

Thus, $I_1 = \emptyset$ so there are no recession directions, but $I_3 = \{1, 2, \ldots\}$ and $\sup_{h \in I_3} \tilde{b}(h) = \infty$. ◁

3.3.3 Dual solvability

To characterize dual solvability, begin with a characterization of the optimal dual value.

Theorem 3.18. If $\tilde{b}(h) \leq 0$ for all $h \in I_1$ then $v(\text{FDSILP}) = \sup_{h \in I_3} \tilde{b}(h)$.

Proof. By Lemma 3.15(ii), for every dual feasible solution $\bar{v}$ there exists an $h \in I_3$ with $\tilde{b}(h) \geq \langle b, \bar{v}\rangle$. Hence, $\sup_{h \in I_3} \tilde{b}(h) \geq \langle b, \bar{v}\rangle$ for all feasible $\bar{v}$. This implies $\sup_{h \in I_3} \tilde{b}(h) \geq v(\text{FDSILP})$. Conversely, by Lemma 3.13(iii), every $h \in I_3$ yields a v^h with v^h feasible to (FDSILP) and $\tilde{b}(h) = \langle b, v^h\rangle$. Hence $\tilde{b}(h) = \langle b, v^h\rangle \leq v(\text{FDSILP})$ for all $h \in I_3$. Thus, $\sup_{h \in I_3} \tilde{b}(h) \leq v(\text{FDSILP})$ and the result follows. □

Corollary 3.19. If either (SILP) is feasible or (FDSILP) is feasible and bounded, then $v(\text{FDSILP}) = \sup_{h \in I_3} \tilde{b}(h)$.

Proof. If (SILP) is feasible, then by Theorem 3.2(ii) $\tilde{b}(h) \leq 0$ for all $h \in I_1$. The result follows from Theorem 3.18. If (FDSILP) is feasible and bounded then by Theorem 3.16(i) $\tilde{b}(h) \leq 0$ for all $h \in I_1$. Once again, the result follows from Theorem 3.18. □

Theorem 3.20 (Dual solvability)**.** (FDSILP) has an optimal solution if and only if

(i) $\tilde{b}(h) \leq 0$ for all $h \in I_1$, and

(ii) $\sup_{h \in I_3} \tilde{b}(h)$ is realized for at least one $h \in I_3$.

Proof. ($\Longrightarrow$) Let v^* be an optimal solution to (FDSILP) with optimal value $v(\text{FDSILP}) = \langle b, v^*\rangle$. This implies (FDSILP) is both feasible and bounded. By Theorem 3.16(i), $\tilde{b}(h) \leq 0$ for all $h \in I_1$, establishing condition (i). Apply Lemma 3.15(i) and conclude there exists a v^{h^*} for some $h^* \in I_3$ with $\langle b, v^{h^*}\rangle \geq \langle b, v^*\rangle = v(\text{FDSILP})$. By Lemma 3.13(iv), v^{h^*} is feasible to (FDSILP) and $\langle b, v^{h^*}\rangle \leq v(\text{FDSILP})$. Hence $\tilde{b}(h^*) = \langle b, v^{h^*}\rangle = v(\text{FDSILP}) = \sup_{h \in I_3} \tilde{b}(h)$, where the first equality holds from Lemma 3.13(i), the second equality holds from the arguments in the previous two sentences, and the third equality holds from Corollary 3.19. Thus, $\tilde{b}(h^*) = \sup_{h \in I_3} \tilde{b}(h)$, establishing condition (ii).
($\Longleftarrow$) By hypothesis there is an $h^* \in I_3$ such that $\sup_{h \in I_3} \tilde{b}(h) = \tilde{b}(h^*) < \infty$. The fact that I_3 is nonempty implies (FDSILP) is feasible by Theorem 3.14. Thus, by Theorem 3.16 (FDSILP) is bounded. Since (FDSILP) is feasible and bounded, by Corollary 3.19 $\sup_{h \in I_3} \tilde{b}(h) = v(\text{FDSILP})$. Moreover, Lemma 3.13(i) and (iv) imply that $\tilde{b}(h^*) = \langle b, v^{h^*}\rangle$ and v^{h^*} is a feasible solution to (FDSILP). Putting this together, $v(\text{FDSILP}) = \sup_{h \in I_3} \tilde{b}(h) = \tilde{b}(h^*) = \langle b, v^{h^*}\rangle$ and v^{h^*} is an optimal solution to (FDSILP). □

3.3.4 Zero duality gap and strong duality

The primal-dual pair (SILP) and (FDSILP) has a *zero duality gap* if (SILP) is feasible and $v(\text{SILP}) = v(\text{FDSILP})$.

Theorem 3.21 (Zero Duality Gap)**.** There is a zero duality gap for the primal-dual pair (SILP) and (FDSILP) if and only if

(i) (SILP) is feasible,

(ii) $\sup_{h \in I_3} \tilde{b}(h) \geq \lim_{\delta \to \infty} \omega(\delta)$.

Proof. ($\Longrightarrow$) Assume zero duality gap. Condition (i) holds by definition of zero duality gap. Since (SILP) is feasible, by Corollary 3.19,

$$\sup_{h \in I_3} \tilde{b}(h) = v(\text{FDSILP}) = v(\text{SILP}) = \max\{\sup_{h \in I_3} \tilde{b}(h), \lim_{\delta \to \infty} \omega(\delta)\} \geq \lim_{\delta \to \infty} \omega(\delta),$$

where the third equality holds by Lemma 3.8. Thus condition ii) holds.

($\Longleftarrow$) Now assume conditions (i) and (ii) hold. By i) (SILP) is feasible. By Lemma 3.8, $v(\text{SILP}) = \max\{\sup_{h\in I_3} \tilde{b}(h), \lim_{\delta\to\infty} \omega(\delta)\} = \sup_{h\in I_3} \tilde{b}(h)$, where the second equality follows from condition ii). Also, Corollary 3.19 implies $v(\text{FDSILP}) = \sup_{h\in I_3} \tilde{b}(h)$. Thus, $v(\text{SILP}) = v(\text{FDSILP})$ and there is a zero duality gap. □

Combining solvability and duality, *strong duality* holds if there is a zero duality gap and there is an optimal solution to (SILP) and (FDSILP). Putting several previous results together gives Theorem 3.22.

Theorem 3.22 (Strong Duality)**.** Strong duality holds for the primal-dual pair (SILP) and (FDSILP) if

(i) (SILP) is feasible,

(ii) $\sup_{h\in I_3} \tilde{b}(h) > \lim_{\delta\to\infty} \omega(\delta)$,

(iii) $\sup_{h\in I_3} \tilde{b}(h)$ is realized for at least one $h \in I_3$.

Conversely, if strong duality holds for the primal-dual pair (SILP) and (FDSILP) then (i) and (iii) hold as well as

(ii') $\sup_{h\in I_3} \tilde{b}(h) \geq \lim_{\delta\to\infty} \omega(\delta)$.

Proof. Suppose conditions (i) to (iii) hold. Conditions (i) and (ii) imply primal solvability via Theorem 3.10. Since (SILP) is feasible, by Theorem 3.2(i), $\tilde{b}(h) \leq 0$ for all $h \in I_1$. Combined with condition (iii) dual solvability follows from Theorem 3.20. Conditions (i) and (ii) imply the sufficient conditions for zero duality gap given in Theorem 3.21 and the duality gap is zero.

Conversely, suppose strong duality holds. Then there is a zero duality gap and so Theorem 3.21, (i) and (ii') hold. Theorem 3.20(ii) implies condition (iii). □

In the example below strong duality holds but condition ii) in Theorem 3.22 is not satisfied.

Example 3.23 (Example 3.12 revisited)**.** In this example the primal is solvable with objective value $v(\text{SILP}) = 0$. Recall also that $\sup_{h\in I_3} \tilde{b}(h) = 0$ is attained since I_3 is a singleton. This implies it is dual solvable and there is zero duality gap. This problem satisfies strong duality. However, $\sup_{h\in I_3} \tilde{b}(h) = \lim_{\delta\to\infty} \omega(\delta)$. Therefore condition (ii) in Theorem 3.22 is not satisfied, but condition (ii') is satisfied. ◁

3.4 Summary of primal and dual results

Table 1 summarizes the main results of this section. For brevity in displaying conditions, define $S := \sup_{h\in I_3} \tilde{b}(h)$ and $L := \lim_{\delta\to\infty} \omega(\delta)$.

Result	Sets involved	Characterization
Primal feasibility (Thm 3.2)	I_1, I_2, I_3, I_4	Conditions i)-iv) of Theorem 3.2
Primal boundedness (Thm 3.9)	I_3, I_4	Primal feas. and ($I_3 \neq \emptyset$ OR $L > -\infty$)
Primal solvability* (Thm 3.10)	I_3, I_4	Primal feasible and $S > L$
Dual feasibility (Thm 3.14)	I_3	$I_3 \neq \emptyset$
Dual boundedness (Thm 3.16)	I_1, I_3	Dual feas., $\tilde{b}(h) \leq 0$ for all $h \in I_1$, $S < \infty$
Dual solvability (Thm 3.20)	I_1, I_3	$\tilde{b}(h) \leq 0$ for all $h \in I_1$, sup defining S realized
Zero duality gap (Thm 3.21)	I_3, I_4	$S \geq L$ and Primal feasible

Table 1: *Summary of results from Section 3*. All results are characterizations except primal solvability, where a sufficient conditions is given.

The next two subsections illustrate insights that are gained by applying the results in Table 1 to two special cases of (SILP).

3.5 Tidy semi-infinite linear programs

An instance of (SILP) is *tidy* if, after applying Fourier-Motzkin elimination to (3.2)-(3.3), z is the only dirty variable remaining. Tidy semi-infinite linear programs play a fundamental role in applications of our theory in later sections. The key properties of tidy systems are summarized in the following theorem.

Theorem 3.24 (Tidy semi-infinite linear programs)**.** If (SILP) is feasible and tidy then

(i) (SILP) is solvable,

(ii) (FDSILP) is feasible and bounded,

(iii) there is a zero duality gap for the primal-dual pair (SILP) and (FDSILP).

Proof. Since (SILP) is tidy, $I_2 = I_4 = \emptyset$. Since z cannot be eliminated, $I_4 = \emptyset$ implies $I_3 \neq \emptyset$. In addition, $I_4 = \emptyset$ means $\omega(\delta) = -\infty$ for all δ and $\lim_{\delta\to\infty} \omega(\delta) = -\infty$. Moreover, since $I_3 \neq \emptyset$ it follows that $\sup_{h\in I_3} \tilde{b}(h) > -\infty$. Then, $\sup_{h\in I_3} \tilde{b}(h) > \lim_{\delta\to\infty} \omega(\delta)$ and Theorem 3.10 implies the primal is solvable. This establishes (i).

Since $I_3 \neq \emptyset$, (FDSILP) is feasible by Theorem 3.14. Since the primal is feasible, Theorem 3.2(i) and (ii) imply that the dual is bounded via Theorem 3.16. This establishes (ii).

Since the primal is feasible and $\sup_{h\in I_3} \tilde{b}(h) > \lim_{\delta\to\infty} \omega(\delta)$, Theorem 3.21 implies that there is a zero duality gap. This establishes (iii). □

The following result provides a sufficient condition for the tidiness of a semi-infinite linear program.

Theorem 3.25 (Bounded System)**.** If there exists a $\gamma \in \mathbb{R}$ such that the system

$$\begin{array}{rcll} -c_1x_1 - c_2x_2 - \cdots - c_nx_n & \geq & -\gamma & \\ a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n & \geq & b(i) & \text{for } i \in I \end{array} \tag{3.18}$$

is feasible and bounded then (SILP) is feasible and tidy. In particular, if the set of solutions $(x_1, \ldots, x_n)$ that satisfy (3.18) is feasible and bounded for some $\gamma \in \mathbb{R}$, then (SILP) is solvable and there is zero duality gap.

Proof. Let Γ_γ denote the set of those $x \in \mathbb{R}^n$ that satisfy (3.18). Observe that the columns in systems (3.18) and (3.2)-(3.3) are identical for variables $x_1, \ldots, x_n$. This means if x_k is eliminated when Fourier-Motzkin elimination is applied to one system, it will be eliminated in exactly the same order in the other. In particular, at each step of the elimination process, the sets $\mathcal{H}_0(k), \mathcal{H}_+(k)$ and $\mathcal{H}_-(k)$ are identical for the two systems. By hypothesis, Γ_γ is non-empty and bounded so Theorem 2.15 guarantees that applying Fourier-Motzkin elimination to (3.18) results in a clean system. Thus, variables $x_1, \ldots, x_n$ are eliminated during the procedure and so those variables are eliminated when applying Fourier-Motzkin elimination to (2.8)-(2.9). Thus, (SILP) is tidy. Since Γ_γ is non-empty, (SILP) is feasible and tidy and the hypotheses of Theorem 3.24 are met. Then by Theorem 3.24, (SILP) is solvable and there is a zero duality gap for the primal-dual pair (SILP) and (FDSILP). □

3.6 Finite linear programs

Another special case is a semi-infinite linear program with finitely many constraints, i.e. a finite linear program, or just a linear program. Finite linear programs are a special case of (SILP) and our analysis applies directly.

For finite linear programs, I_1, I_2, I_3 and I_4 are always finite sets. This simplifies the characterizations in Table 1 since the supremums are taken over finite sets. Take, for example, primal feasibility (Theorem 3.2). Conditions ii)-iv) always hold from the finiteness of I_2, I_3 and I_4 respectively. Thus to determine primal feasibility it suffices to check if $\tilde{b}(h) \leq 0$ for all $h \in I_1$. This result is well known (see for instance, Motzkin [11]).

As another example, strong duality holds for a finite linear program when the primal is feasible and bounded. Our framework recovers this result.

Theorem 3.26 (Finite Case)**.** If I is a finite index set and (SILP) is feasible and bounded, then strong duality holds for the primal-dual pair (SILP) and (FDSILP).

Proof. Show that conditions (i)-(iii) of Theorem 3.22 hold. By hypothesis (SILP) is feasible and bounded so i) holds. When I is a finite set, I_4 has finite cardinality so $\lim_{\delta \to \infty} \omega(\delta) = -\infty$. Combining this with the hypothesis that the primal is bounded implies $I_3 \neq \emptyset$ by Theorem 3.9. Thus condition (ii) in Theorem 3.22 holds. Finally, (iii) holds since I_3 is finite whenever I is finite. ☐

Beyond this, the analysis of this section reveals important differences between a finite linear programs and a semi-infinite linear program. Consider the following well-known facts about finite linear programs:

(i) if the primal is infeasible then the dual must be either infeasible or unbounded, and

(ii) if the primal has a finite optimal objective value, then the dual must be feasible and bounded with the same objective value (that is, strong duality always holds).

The following two examples demonstrate that (i) and (ii) need not hold for general semi-infinite linear programs.

Example 3.27. Consider the following problem

$$\begin{array}{lrcll} \inf x_1 & & & & \\ & \frac{1}{i}x_2 & \geq & 1 & \text{for } i = 1, 2, \ldots \\ x_1 & & \geq & 0. & \end{array}$$

This problem is infeasible since for any x_2, there exists sufficiently large i such that $\frac{1}{i}x_2 < 1$.

Add the constraint $-x_1 + z \geq 0$, apply Fourier-Motzkin elimination and project out the clean variable x_1 to get the following system.

$$\begin{array}{rcll} \frac{1}{i}x_2 & \geq & 1 & \text{for } i = 1, 2, \ldots \\ z & \geq & 0, & \end{array}$$

In this system $I_1 = I_4 = \emptyset$ and I_3 is a singleton. This implies that $\sup_{h \in I_3} \tilde{b}(h)$ is achieved and the dual is solvable. ◁

Example 3.28 (Example 3.11 revisited)**.** In Example 3.11, the primal problem has a finite optimal value of 0. This optimal value remains greater than or equal to zero even without the non-negativity constraint on x_1 in (3.14). This is because $\omega(\delta)$ still equals $\frac{1}{4(\delta-1)}$ and $\lim_{\delta \to \infty} \omega(\delta) = 0$. Then by Lemma 3.8, the optimal primal value is greater than or equal to zero. However, the finite support dual of this semi-infinite linear program is infeasible. The objective coefficient of x_2 in the primal is 0 and the coefficient of x_2 is strictly positive in the constraints. This implies that the only possible dual element satisfying the dual constraint corresponding to x_2 is $u = 0$; however, the objective coefficient of x_1 is 1 and this dual vector does not satisfy the dual constraint corresponding to x_1. Alternatively, the infeasibility of the dual follows from Theorem 3.14 because in this case $I_3 = \emptyset$. ◁

4 Feasible sequences and regular duality of semi-infinite linear programs

When I_3 is empty in (3.4), there is no a feasible solution to (FDSILP) as shown in Theorem 3.14. Nevertheless, if the primal problem has optimal solution value z^*, we show there is a sequence $\{h_m\} \in I_4$ for $m \in \mathbb{N}$ with the desirable property that for all $k = 1, \ldots, n$, $\tilde{a}^k(h_m)$ converges to zero and $\tilde{b}(h_m)$ converges to z^* as $m \to \infty$. In Theorem 4.3 it is shown that there is a sequence of finite support elements with nice limiting properties, and whose objective values converges to the *primal optimal value*. The terminology for this phenomenon, standard in conic programming, is introduced next. The concepts date back to Duffin [4].

A sequence $v^m \in \mathbb{R}^{(I)}$, $m \in \mathbb{N}$ of finite support elements is a *feasible sequence* for (FDSILP) if $v^m \geq 0$ for all $m \in \mathbb{N}$, and for every $k = 1, \ldots, n$, $\lim_{m\to\infty}(\sum_{i\in I} a^k(i)v^m(i)) = c_k$. For a feasible sequence $(v^m)_{m\in\mathbb{N}}$, its *value* is defined by $\text{value}((v^m)_{m\in\mathbb{N}}) := \limsup_{m\to\infty} \sum_{i\in I} b(i)v^m(i)$. For a given (FDSILP), its *limit value* (a.k.a. *subvalue*) is

$$\sup\{\text{value}((v^m)_{m\in\mathbb{N}}) \mid (v^m)_{m\in N} \text{ is a feasible sequence for (FDSILP)}\}.$$

Since any feasible solution $v \in \mathbb{R}^{(I)}$ to (FDSILP) naturally corresponds to a feasible sequence (where every element in the sequence is v), the limit value of (FDSILP) is greater than or equal to its optimal value. We prove a remarkable theorem (Theorem 4.3 below) relating the limit value of (FDSILP) and the optimal value of the primal (SILP).

Lemma 4.1 (Weak Duality-II)**.** Let $\bar{x}$ be a feasible solution to the primal (SILP) and let $(v^m)_{m\in\mathbb{N}}$ be a feasible sequence for (FDSILP). Then $c^\top \bar{x} \geq \text{value}((v^m)_{m\in\mathbb{N}})$.

Proof. Since $\bar{x}$ is a feasible solution to the primal (SILP), $a^1(i)\bar{x}_1 + \ldots + a^n(i)\bar{x}_n \geq b(i)$ for every $i \in I$. For each v^m, since $v^m(i) \geq 0$ for all $i \in I$, $v^m(i)a^1(i)\bar{x}_1 + \ldots + v^m(i)a^n(i)\bar{x}_n \geq b(i)v^m(i)$ for every $i \in I$. Therefore, summing over all the indices $i \in I$, gives $(\sum_{i\in I} v^m(i)a^1(i))\bar{x}_1 + \cdots + (\sum_{i\in I} v^m(i)a^n(i))\bar{x}_n \geq \sum_{i\in I} b(i)v^m(i)$ for all $m \in \mathbb{N}$. Thus,

$$\begin{array}{rcl} c_1\bar{x}_1 + \ldots + c_n\bar{x}_n & = & \lim_{m\to\infty}[(\sum_{i\in I} v^m(i)a^1(i))\bar{x}_1 + \cdots + (\sum_{i\in I} v^m(i)a^n(i))\bar{x}_n] \\ & = & \limsup_{m\to\infty}[(\sum_{i\in I} v^m(i)a^1(i))\bar{x}_1 + \cdots + (\sum_{i\in I} v^m(i)a^n(i))\bar{x}_n] \\ & \geq & \limsup_{m\to\infty}[\sum_{i\in I} b(i)v^m(i)] \\ & = & \text{value}((v^m)_{m\in\mathbb{N}}), \end{array}$$

where the first equality follows from the definition of feasible sequence. □

The following lemma is required for the main result of the section (Theorem 4.3). Applying Fourier-Motzkin elimination on (SILP) gives (3.4). Recall the function $\omega(\delta) = \sup\{\tilde{b}(h) - \delta \sum_{k=\ell}^n |\tilde{a}^k(h)| \,:\, h \in I_4\}$ defined in (3.10).

Lemma 4.2. Suppose $\lim_{\delta\to\infty} \omega(\delta) = d$ such that $-\infty < d < \infty$. Then there exists a sequence of indices h_m in I_4 such that such that $\lim_{m\to\infty} \tilde{b}(h_m) = d$ and $\lim_{m\to\infty} \tilde{a}^k(h_m) = 0$ for all $k = \ell, \ldots, n$. Moreover, $\tilde{b}(h_m) \geq d$ for all $m \in \mathbb{N}$.

Proof. Since $\omega(\delta)$ is a nonincreasing function of δ, $\omega(\delta) \geq d$ for all δ. Therefore, $d \leq \sup\{\tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4\}$ for every δ. Let $\bar{I} \subseteq I_4$ be such that for all $h \in \bar{I}$, $\tilde{b}(h) < d$. Show that it is sufficient to consider indices in $I_4 \setminus \bar{I}$. Given any $\delta \geq 0$, $\tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| < d$ for all $h \in \bar{I}$. Since $d \leq \sup\{\tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4\}$, given $\delta \geq 0$,

$$\sup\{\tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4\} = \sup\{\tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4 \setminus \bar{I}\}.$$

Thus, $\omega(\delta) = \sup\{\tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4 \setminus \bar{I}\}$ for all $\delta \geq 0$.

First show that there exists a sequence of indices $h_m \in I_4 \setminus \bar{I}$ such that $\tilde{a}^k(h_m) \to 0$ for all $k = \ell, \ldots, n$. Begin by showing that $\inf\{\sum_{k=\ell}^{n} |\tilde{a}^k(h)| : h \in I_4 \setminus \bar{I}\} = 0$. This implies that there is a sequence $h_m \in I_4 \setminus \bar{I}$ such that $\lim_{m\to\infty} \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| = 0$ which in turn implies that $\lim_{m\to\infty} \tilde{a}^k(h_m) = 0$ for all $k = \ell, \ldots, n$. Suppose to the contrary that $\inf\{\sum_{k=\ell}^{n} |\tilde{a}^k(h)| : h \in I_4 \setminus \bar{I}\} = \beta > 0$. Since $\omega(\delta)$ is nonincreasing and $\lim_{\delta\to\infty} \omega(\delta) = d < \infty$, there exists $\bar{\delta} \geq 0$ such that $\omega(\bar{\delta}) < \infty$. Observe that $d = \lim_{\delta\to\infty} \omega(\delta) = \lim_{\delta\to\infty} \omega(\bar{\delta} + \delta)$. Then, for every $\delta \geq 0$,

$$\begin{aligned}
\omega(\bar{\delta} + \delta) &= \sup\{\tilde{b}(h) - (\bar{\delta} + \delta) \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4 \setminus \bar{I}\} \\
&= \sup\{\tilde{b}(h) - \bar{\delta} \sum_{k=\ell}^{n} |\tilde{a}^k(h)| - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4 \setminus \bar{I}\} \\
&\leq \sup\{\tilde{b}(h) - \bar{\delta} \sum_{k=\ell}^{n} |\tilde{a}^k(h)| - \delta\beta \,:\, h \in I_4 \setminus \bar{I}\} \\
&= \sup\{\tilde{b}(h) - \bar{\delta} \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4 \setminus \bar{I}\} - \delta\beta \\
&= \omega(\bar{\delta}) - \delta\beta.
\end{aligned}$$

Therefore, $d = \lim_{\delta\to\infty} \omega(\bar{\delta} + \delta) \leq \lim_{\delta\to\infty}(\omega(\bar{\delta}) - \delta\beta) = -\infty$, since $\beta > 0$ and $\omega(\bar{\delta}) < \infty$. This contradicts $-\infty < d$. Thus $\beta = 0$ and there is a sequence $h_m \in I_4 \setminus \bar{I}$ such that $\tilde{a}^k(h_m) \to 0$ for all $k = \ell, \ldots, n$.

Now show there is a subsequence of $\tilde{b}(h_m)$ that converges to d. Since $\lim_{\delta\to\infty} \omega(\delta) = d$, there is a sequence $(\delta_p)_{p\in\mathbb{N}}$ such that $\delta_p \geq 0$ and $\omega(\delta_p) < d + \frac{1}{p}$ for all $p \in \mathbb{N}$. It was shown above that the sequence $h_m \in I_4 \setminus \bar{I}$ is such that $\lim_{m\to\infty} \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| = 0$. This implies that for every $p \in \mathbb{N}$ there is an $m_p \in \mathbb{N}$ such that for all $m \geq m_p$, $\delta_p \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| < \frac{1}{p}$. Thus, one can extract a subsequence $(h_{m_p})_{p\in\mathbb{N}}$ of $(h_m)_{m\in\mathbb{N}}$ such that $\delta_p \sum_{k=\ell}^{n} |\tilde{a}^k(h_{m_p})| < \frac{1}{p}$ for all $p \in \mathbb{N}$. Then

$$\begin{aligned}
d + \frac{1}{p} > \omega(\delta_p) &= \sup\{\tilde{b}(h) - \delta_p \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \,:\, h \in I_4 \setminus \bar{I}\} \\
&\geq \tilde{b}(h_{m_p}) - \delta_p \sum_{k=\ell}^{n} |\tilde{a}^k(h_{m_p})|.
\end{aligned}$$

The second inequality, along with $\delta_p \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| < \frac{1}{p}$, and the fact that $h_{m_p} \in I_4 \setminus \bar{I}$ implies $\tilde{b}(h_{m_p}) \geq d$, gives $d + \frac{2}{p} \geq \tilde{b}(h_{m_p}) \geq d$ and $\tilde{b}(h_{m_p}), p \in \mathbb{N}$ is the desired subsequence. □

Theorem 4.3 (Regular duality of semi-infinite linear programs)**.** If (SILP) has an optimal primal value z^*, where $-\infty < z^* < \infty$, then the limit value $\hat{d}$ of (FDSILP) is finite and $z^* = \hat{d}$.

Proof. By Lemma 3.8, $z^* = \max\{\sup\{\tilde{b}(h) : h \in I_3\}, \lim_{\delta\to\infty} \omega(\delta)\}$. If $z^* = \sup\{\tilde{b}(h) : h \in I_3\}$, then by Theorem 3.21, there is a zero duality gap, i.e., $z^* = d^*$ where d^* is the optimal value of (FDSILP). From Lemma 4.1, $\hat{d} \le z^*$, so $z^* = d^*$ implies $\hat{d} \le d^*$. By definition of limit value, $\hat{d} \ge d^*$. Therefore, $d^* = \hat{d} = z^*$.

In the other case when $z^* = \lim_{\delta\to\infty} \omega(\delta)$, by Lemma 4.2 there is a sequence $h_m \in I_4$ such that $\lim_{m\to\infty} \tilde{b}(h_m) = z^*$ and $\lim_{m\to\infty} \tilde{a}^k(h_m) = 0$ for all $k = \ell, \ldots, n$. By Lemma 3.13 there exist $v^{h_m} \in \mathbb{R}^{(I)}_+$ for each $m \in \mathbb{N}$ such that $-c_k + \sum_{i\in I} v^{h_m}(i)a^k(i) = 0$ for $k = 1, \ldots, \ell - 1$, $-c_k + \sum_{i\in I} v^{h_m}(i)a^k(i) = \tilde{a}^k(h_m)$ for $k = \ell, \ldots, n$, and $\sum_{i\in I} b(i)v^{h_m}(i) = \tilde{b}(h_m)$. Since $\lim_{m\to\infty} \tilde{a}^k(h_m) = 0$ for all $k = \ell, \ldots, n$, and $\lim_{m\to\infty} \tilde{b}(h_m) = z^*$, v^{h_m}, $m \in \mathbb{N}$ is a feasible sequence with value z^*. Thus, $\hat{d} \ge z^*$. Again, from Lemma 4.1, $\hat{d} \le z^*$, so $z^* = \hat{d}$. □

5 Application: Conic linear programs

Recall the definition of (ConLP) in Section 1 and its standard dual (ConLPD). They are reproduced here for convenience. The conic primal is

$$\begin{array}{rl} \inf_{x\in X} & \langle x, \phi\rangle \\ \text{s.t.} & A(x) \succeq_P d \end{array} \tag{ConLP}$$

where X and Y are vector spaces, $A : X \to Y$ is a linear mapping, $d \in Y$, P is a pointed convex cone in Y and ϕ is a linear functional on X. The conic dual is

$$\begin{array}{rl} \sup_{\psi\in Y'} & \langle d, \psi\rangle \\ \text{s.t.} & A'(\psi) = \phi \\ & \psi \in P'. \end{array} \tag{ConLPD}$$

Let $F = \{x \in X \mid A(x) \succeq_P d\}$ denote the feasible region of (ConLP). In our development, it is convenient to assume that the algebraic adjoint A' of linear map A is surjective (in order to apply the Open Mapping Theorem). Lemmas 5.1–5.3 and Remark 5.4 show this assumption can be made without loss of generality. For any linear map T defined on X, let $\ker(T) = \{x \in X : T(x) = 0\}$ denote the kernel of T.

Lemma 5.1. Given a linear mapping $A : X \to Y$, $\ker(A) = \{0\}$ if and only if A' is surjective.

Proof. ($\Longrightarrow$) If $\ker(A) = \{0\}$, then A is one-to-one and there is a linear map $A^{-1} : Im(A) \to X$. Let ϕ be an arbitrary linear functional in X'. Show there exists a linear functional $\psi \in Y'$ such that $\phi = A'(\psi)$. Define the linear functional $\phi \circ A^{-1}$ on $Im(A)$ and let ψ be any extension of this linear functional from $Im(A)$ to Y. Thus $\psi \in Y'$. Show $\phi = A'(\psi)$. For any $x \in X$, $\langle x, A'(\psi)\rangle = \langle A(x), \psi\rangle = (\phi \circ A^{-1})(A(x)) = \phi(x) = \langle x, \phi\rangle$. The second equality follows since $A(x) \in Im(A)$.

($\Longleftarrow$) Consider $x \in X$ such that $A(x) = 0$. Show that for every $\phi \in X'$, $\langle x, \phi\rangle = 0$. This would imply that $x = 0$. Since A' is surjective, for every $\phi \in X'$ there exists $\psi \in Y'$ such that $A'(\psi) = \phi$. Thus, $\langle x, \phi\rangle = \langle x, A'(\psi)\rangle = \langle A(x), \psi\rangle = \langle 0, \psi\rangle = 0$. □

Lemma 5.2. If (ConLP) is feasible and bounded, then $\ker(A) \subseteq \ker(\phi)$.

Proof. Prove the contrapositive and assume that there is an $r \in \ker(A) \setminus \ker(\phi)$. Without loss of generality assume $\langle r, \phi\rangle < 0$ (otherwise make the argument with $-r$). Let $\bar{x}$ be a feasible solution to (ConLP), i.e., $A(\bar{x}) \succeq_P d$. Since $r \in \ker(A)$, $A(\bar{x} + \lambda r) \succeq_P d$ for all $\lambda \ge 0$. But since $\langle r, \phi\rangle < 0$, $\langle \bar{x} + \lambda r, \phi\rangle \to -\infty$ as $\lambda \to \infty$, contradicting the boundedness of (ConLP). □

Lemma 5.3. Let X be a finite-dimensional space, so that orthogonal complements of subspaces are well-defined. Let $\bar{\phi} = \phi|_{\ker(A)^\perp}$ be the linear functional on $\ker(A)^\perp$ defined by the restriction of ϕ to $\ker(A)^\perp$. Similarly, let $\bar{A} = A|_{\ker(A)^\perp}$ denote the restriction of the linear map A. Consider the optimization problem

$$\begin{array}{rl} \inf_{x \in \ker(A)^\perp} & \langle x, \bar{\phi} \rangle \\ \text{s.t.} & \bar{A}(x) \succeq_P d. \end{array} \tag{5.1}$$

If (ConLP) is feasible and bounded, the optimal value of (ConLP) equals the optimal value of (5.1). Moreover, if O is the set of optimal primal solutions for (ConLP), and $\overline{O}$ is the set of optimal primal solutions for (5.1), then $O = \overline{O} + \ker(A)$.

Proof. Since (ConLP) is feasible and bounded, $\ker(A) \subseteq \ker(\phi)$ by Lemma 5.2. For any x feasible to (ConLP), let $r \in \ker(A)$ and $\bar{x} \in \ker(A)^\perp$ such that $x = \bar{x} + r$. Since $\ker(A) \subseteq \ker(\phi)$, $\langle r, \phi \rangle = 0$. Thus, $\langle x, \phi \rangle = \langle \bar{x} + r, \phi \rangle = \langle \bar{x}, \phi \rangle = \langle \bar{x}, \bar{\phi} \rangle$, the last equality follows since $\bar{x} \in \ker(A)^\perp$. Similarly, $\bar{A}(\bar{x}) = A(\bar{x}) = A(\bar{x} + r) = A(x) \succeq_P d$. Thus, $\bar{x}$ is a feasible solution to (5.1) with the same objective value as $\langle x, \phi \rangle$. □

Remark 5.4. By Lemma 5.3, when (ConLP) is feasible and bounded, it suffices to consider a restricted optimization problem like (5.1). Note that $\ker(\bar{A}) = \{0\}$. Thus, without loss of generality, it is valid to assume that for an instance of a feasible and bounded (ConLP) in a finite-dimensional space X, the linear map A has zero kernel, i.e., it is one-to-one. This implies that A' is surjective by Lemma 5.1. ◁

Construct the following primal-dual pair of semi-infinite linear programs in the case where X is finite-dimensional and the cone P is reflexive. Recall that a cone P is *reflexive* if $P'' = P$ under the natural embedding of $Y \hookrightarrow Y''$. The significance of this property will become apparent in Theorem 5.5. The primal semi-infinite linear program is

$$\begin{array}{rll} \inf_{x \in \mathbb{R}^n} & c^\top x & \\ \text{s.t.} & a^1(\psi)x_1 + a^2(\psi)x_2 + \cdots + a^n(\psi)x_n \geq b(\psi) & \text{for all } \psi \in P' \end{array} \tag{ConSILP}$$

where X is isomorphic to $\mathbb{R}^n$ with respect to a basis $e^1, \ldots, e^n$ and $c \in \mathbb{R}^n$ represents the linear functional $\phi \in X'$ (also using the isomorphism of X' and $\mathbb{R}^n$). In (ConSILP), the elements $a^j \in \mathbb{R}^{P'}$ $j = 1, \ldots, n$ and $b \in \mathbb{R}^{P'}$, are defined by $a^j(\psi) := \langle A(e^j), \psi \rangle$ and $b(\psi) := \langle d, \psi \rangle$. The finite support dual of (ConSILP) is

$$\begin{array}{rrcll} & \sup \sum_{\psi \in P'} b(\psi)v(\psi) & & & \\ \text{s.t.} & \sum_{\psi \in P'} a^k(\psi)v(\psi) & = & c_k & \text{for } k = 1, \ldots, n \\ & v & \in & \mathbb{R}_+^{(P')}. & \end{array} \tag{ConFDSILP}$$

The close connection of this primal-dual pair to the conic pair (ConLP)–(ConLPD) is shown in Theorem 5.5 and Theorem 5.7.

Theorem 5.5 (Primal correspondence)**.** Assume P is reflexive and X is finite-dimensional. Let $e^1, \ldots, e^n$ be the basis of X used to define (ConSILP) and (ConFDSILP). Then, $v(\text{ConLP}) = v(\text{ConSILP})$. Moreover, the set of feasible solutions to (ConLP) is isomorphic to the set of feasible solutions to (ConSILP) under this basis.

Proof. Since X is isomorphic to $\mathbb{R}^n$ with respect to the basis $e^1, \ldots, e^n$ and $c \in \mathbb{R}^n$ represents the linear functional $\phi \in X'$ the objective functions of both problems are identical (under this isomorphism). The result follows if the feasible regions of both problems are isomorphic under this same mapping.

Let F denote the feasible region of (ConLP) and $\hat{F}$ denote the feasible region of (ConSILP). Show F is isomorphic to $\hat{F}$ under the basis $e^1, \ldots, e^n$. First show that if $x = x_1 e^1 + \ldots + x_n e^n \in F$

then $(x_1, \ldots, x_n) \in \hat{F}$. If $x \in F$, then $A(x) \succeq_P d$. Therefore, $A(x) - d \in P$ and so for all $\psi \in P'$, $\langle (A(x) - d), \psi \rangle \geq 0$. Writing $A(x) = \sum_{j=1}^{n} x_j A(e^j)$ and using the linearity of ψ, it follows that $(x_1, \ldots, x_n) \in \hat{F}$.

Next show that if $(x_1, \ldots, x_n) \in \hat{F}$, then $x = x_1 e^1 + \ldots + x^n e^n \in F$. Show the contrapositive, i.e. if $x \notin F$ then $(x_1, \ldots, x_n) \notin \hat{F}$. If $x \notin F$, then $A(x) - d \notin P$ and since P is reflexive, $A(x) - d \notin P''$ (under the natural embedding of $Y \hookrightarrow Y''$). Therefore, there exists $\psi \in P'$ such that $\langle (A(x) - d), \psi \rangle < 0$. Again, using the linearity of ψ it follows that $(x_1, \ldots, x_n) \notin \hat{F}$. $\square$

Remark 5.6. The condition that P is reflexive naturally holds in many important special cases of conic programming. Once such case is when X and Y are finite dimensional spaces and P is a closed, pointed cone in Y. Then P is easily seen to be reflexive. This case includes linear programming, semi-definite programming (SDPs) and copositive programming. The above reformulation as a semi-infinite linear program works for any such instance. $\triangleleft$

Theorem 5.7 (Dual Correspondence)**.** Assume P is reflexive and X is finite-dimensional. Let $e^1, \ldots, e^n$ be the basis of X used to define (ConSILP) and (ConFDSILP). Then, $v(\text{ConLPD}) = v(\text{ConFDSILP})$. Moreover, there exists maps $T : P' \to \mathbb{R}_+^{(P')}$ and $\hat{T} : \mathbb{R}_+^{(P')} \to P'$ such that if $\psi \in P'$ is a feasible solution to (ConLPD) then $T(\psi)$ is a feasible solution to (ConFDSILP). Conversely, if $v \in \mathbb{R}^{(P')}$ is a feasible solution to (ConFDSILP) then $\hat{T}(v)$ is a feasible solution to (ConLPD).

Proof. It suffices to construct maps T and $\hat{T}$ which satisfy the following properties.

(i) $\langle e^k, A'(\psi^*) \rangle = \sum_{\psi \in P'} a^k(\psi) T(\psi^*)(\psi)$, for every $\psi^* \in P'$ and all $k = 1, \ldots, n$.

(ii) $\langle d, \psi^* \rangle = \sum_{\psi \in P'} b(\psi) T(\psi^*)(\psi)$, for every $\psi^* \in P'$.

(iii) $\sum_{\psi \in P'} a^k(\psi) v(\psi) = \langle e^k, A'(\hat{T}(v)) \rangle$, for every $v \in \mathbb{R}_+^{(P')}$ and all $k = 1, \ldots, n$.

(iv) $\sum_{\psi \in P'} b(\psi) v(\psi) = \langle d, \hat{T}(v) \rangle$, for every $v \in \mathbb{R}_+^{(P')}$.

The map T is defined as follows. For any $\psi^* \in P'$, $T(\psi^*)$ is the finite support element $v^* \in \mathbb{R}^{(P')}$ where the only non-zero component of v^* is 1 and corresponds to ψ^*. For any $k \in \{1, \ldots, n\}$, $\sum_{\psi \in P'} a^k(\psi) v^*(\psi) = a^k(\psi^*) = \langle A(e^k), \psi^* \rangle = \langle e^k, A'(\psi^*) \rangle$ and (i) is satisfied. Also, $\sum_{\psi \in P'} b(\psi) v^*(\psi) = b(\psi^*) = \langle d, \psi^* \rangle$ and (ii) is satisfied.

The map $\hat{T}$ is defined as follows. For any $v^* \in \mathbb{R}^{(P')}$, $\hat{T}(v^*) = \sum_{\psi \in P'} v^*(\psi)\psi$ where the sum is well-defined because v^* has finite support. Since v^* has nonnegative entries, $\hat{T}(v^*) \in P'$. Now, $\sum_{\psi \in P'} a^k(\psi) v^*(\psi) = \sum_{\psi \in P'} \langle A(e^k), \psi \rangle v^*(\psi) = \langle A(e^k), \sum_{\psi \in P'} v^*(\psi)\psi \rangle = \langle A(e^k), \hat{T}(v^*) \rangle = \langle e^k, A'(\hat{T}(v^*)) \rangle$ and (iii) is satisfied. Also, $\sum_{\psi \in P'} b(\psi) v^*(\psi) = \sum_{\psi \in P'} \langle d, \psi \rangle v^*(\psi) = \langle d, \sum_{\psi \in P'} v^*(\psi)\psi \rangle = \langle d, \hat{T}(v^*) \rangle$ and (iv) is satisfied. $\square$

5.1 Zero duality gap via boundedness

We prove Theorem 5.8 in this section. A more restricted version of this result is known in the classical conic programming literature (see for example Duffin [5]). Our result is obtained with a completely new proof using projection techniques.

Theorem 5.8 (Zero duality gap via boundedness)**.** Let X be finite-dimensional. If P is reflexive and there exists a scalar γ such the set $\{x : A(x) \succeq_P d \text{ and } \langle x, \phi \rangle \leq \gamma\}$ is nonempty and bounded, then there is no duality gap for the primal-dual pair (ConLP)-(ConLPD).

Proof. By Theorem 5.5, the primal optimal value of (ConLP) is equal to the optimal value of the (ConSILP). By Theorem 3.25 there is a zero duality gap between the primal dual pair (ConSILP)-(ConFDSILP). Finally, from Theorem 5.7, the optimal value of (ConFDSILP) is equal to the optimal value of (ConLPD). $\square$

Corollary 5.9. Semi-definite programs (SDPs) and copositive programs with nonempty, bounded feasible regions have zero duality gap.

5.2 Regular duality for conic programs

We now prove a central result of conic programming, known as *regular duality*, using the machinery of FM elimination. First, some notions from conic programming (see Chapter 4 of Gartner and Matousék [8] for more details). A sequence $(\psi^m)_{m\in\mathbb{N}}$, is called a feasible sequence for the dual program (ConLPD) if $\psi^m \in P'$ for all $m \in \mathbb{N}$ and

$$\lim_{m\to\infty} A'(\psi^m) = \phi.$$

The *value of a feasible sequence* $(\psi^m)_{m\in\mathbb{N}}$ is $\langle d, (\psi^m)_{m\in\mathbb{N}}\rangle = \limsup_{m\to\infty}\langle d, \psi^m\rangle$. The *limit value (a.k.a. subvalue)* of the dual program (ConLPD) is

$$\sup\{\langle d, (\psi^m)_{m\in\mathbb{N}}\rangle \mid (\psi^m)_{m\in\mathbb{N}} \text{ is a feasible sequence for (ConLPD)}\}.$$

A simple proof of regular duality for conic programs is easily obtained using projection (see Theorem 4.7.3 in Gartner and Matousek [8] for the more standard proof technique).

Theorem 5.10 (Regular duality for conic programs)**.** Assume X is finite-dimensional and P is reflexive. If the primal conic program (ConLP) is feasible and has a finite optimal value z^*, then the dual program (ConLPD) has a finite limit value $\hat{d}$ and $z^* = \hat{d}$.

Proof. By Theorem 5.5, the optimal value of (ConSILP) is equal to z^* and z^* is finite since the optimal value of (ConLP) is finite. By Theorem 4.3, the limit value of (ConFDSILP) equals the optimal value of (ConSILP). By Theorem 5.7, every feasible sequence $(\psi^m)_{m\in\mathbb{N}}$ for (ConLPD) maps to a feasible sequence $(T(\psi^m))_{m\in\mathbb{N}}$ for (ConFDSILP). Similarly, every feasible sequence $(u^m)_{m\in\mathbb{N}}$ for (ConFDSILP) maps to a feasible sequence $(\hat{T}(u^m))_{m\in\mathbb{N}}$ for (ConLPD). Thus, the limit value $\hat{d}$ of (ConLPD) is equal z^*, the limit value of (ConFDSILP). □

5.3 Zero duality gap via Slater's condition

Assume throughout this section X and Y are finite-dimensional spaces, and P is reflexive (note that any closed cone is reflexive because Y is finite-dimensional). As before, identify X with $\mathbb{R}^n$. Let $B(x,\epsilon) \subseteq \mathbb{R}^n$ denote the open ball of radius ϵ with center $x \in \mathbb{R}^n$. Identify the objective linear functional $\phi \in X'$ with the vector $c \in \mathbb{R}^n$.

Lemma 5.11. Assume $A' : Y' \to X'$ is surjective and there exists $\psi^* \in \text{int}(P')$ with $A'(\psi^*) = c$. Then there exists $\epsilon > 0$ and $\bar{\psi} \in P'$, such that for all $\bar{c} \in B(c,\epsilon)$, $\bar{c}^\top x \geq \langle d, \bar{\psi}\rangle$ is a constraint in (ConSILP).

Proof. For each $\psi \in P'$, the constraint in (ConSILP) corresponding to ψ is $\sum_{j=1}^n x_j\langle A(e^j), \psi\rangle \geq \langle d, \psi\rangle$. The left hand side of the inequality is the same as $\sum_{j=1}^n x_j\langle e^j, A'(\psi)\rangle = \langle x, A'(\psi)\rangle$. Since A' is a linear map between finite-dimensional spaces, it is continuous and by assumption, surjective. By the Open Mapping theorem, A' maps open sets to open sets. Since $\psi^* \in \text{int}(P')$ there exists an open ball $B^* \subseteq P'$ containing ψ^*. Thus, $A'(B^*)$ is an open set containing c. Therefore, there exists an $\epsilon > 0$ such that $B(c,\epsilon) \subseteq A'(B^*)$. Thus, for every $\bar{c} \in B(c,\epsilon)$, there exists $\bar{\psi} \in B^*$ such that $A'(\bar{\psi}) = \bar{c}$. Since all $\psi \in B^* \subseteq P'$ give constraints $\langle x, A'(\psi)\rangle \geq \langle d, \psi\rangle$ in (ConSILP), for every $\bar{c} \in B(c,\epsilon)$ there is the constraint $\bar{c}^\top x = \langle x, A'(\bar{\psi})\rangle \geq \langle d, \bar{\psi}\rangle$ in (ConSILP). □

Theorem 5.12 (Slater's theorem for conic programs)**.** If the primal conic program (ConLP) is feasible and there exists $\psi^* \in \text{int}(P')$ with $A'(\psi^*) = c$, then there is a zero duality gap for the primal dual pair (ConLP)-(ConLPD). Moreover, the primal is solvable.

Proof. By hypothesis, there exists $\psi^* \in \text{int}(P')$ with $A'(\psi^*) = c$ so the dual conic program (ConLPD) is feasible. Since (ConLP) is also feasible by hypothesis, feasibility of (ConLPD) implies (ConLP) is both feasible and bounded. Then by Remark 5.4, it is valid to assume A' is surjective.

Claim 5.13. The variables $x_1, \ldots, x_n$ remain clean when Fourier-Motzkin elimination is applied to (ConSILP).

Proof. Since $A'(\psi^*) = c$, there is a constraint $c^\top x \geq \langle d, \psi^* \rangle$ in the system (ConSILP). The constraint $-c^\top x + z \geq 0$ is also present when Fourier-Motzkin elimination is performed on a semi-infinite linear program. By Lemma 5.11, there exists $\epsilon > 0$ such that every $\bar{c} \in B(c, \epsilon)$ gives a constraint $\bar{c}^\top x \geq b'$ in (ConSILP). Thus, for any $\delta < \epsilon$, both $(c + \delta e^j)^\top x \geq b_+^j$ and $(c - \delta e^j)^\top x \geq b_-^j$ are constraints for every $j = 1, \ldots, n$, (where b_+^j and b_-^j are some real numbers).

Case 1: $c_j = 0$ for all $j = 1, \ldots, k$. In this case the constraints are $\frac{\epsilon}{2} x_j \geq b_+^j$ and $-\frac{\epsilon}{2} x_j \geq b_-^j$ in the system. During Fourier-Motzkin, for each $j = 1, \ldots, n$, the constraints $\frac{\epsilon}{2} x_j \geq b_+^j$ and $-\frac{\epsilon}{2} x_j \geq b_-^j$ remain in the system until variable x_j is reached. This makes all variables $x_1, \ldots, x_n$ clean throughout the Fourier-Motzkin procedure.
Case 2: $c_j \neq 0$ for some $j = 1, \ldots, k$. Relabel the variables and assume that $c_1 \neq 0$. Thus, the coefficient of x_1 in $-c^\top x + z \geq 0$ has opposite sign to the coefficient of x_1 in the constraints $(c + \delta e^j)^\top x \geq b_+^j$ and $(c - \delta e^j)^\top x \geq b_-^j$ for $j = 2, \ldots, n$. Thefefore x_1 is clean, and when x_1 is eliminated, the constraint $-c^\top x + z \geq 0$ is aggregated with the constraints $(c + \delta e^j)^\top x \geq b_+^j$ and $(c - \delta e^j)^\top x \geq b_-^j$, for each $j = 2, \ldots, n$. This leaves the constraints $\delta x_j + z \geq b_+^j$ and $-\delta x_j + z \geq b_-^j$ in the system for $j = 2, \ldots, n$, after x_1 is eliminated. As in Case 1, these constraints remain in the system variable until x_j is reached. This makes all variables $x_1, \ldots, x_n$ clean throughout the Fourier-Motzkin procedure. ☐

Since variables $x_1, \ldots, x_n$ are clean throughout the Fourier-Motzkin procedure, and (ConSILP) is feasible (since (ConLP) is feasible), the problem is feasible and tidy and by Theorem 3.24, there is a zero duality gap between the pair (ConSILP)-(ConFDSILP), and (ConSILP) is solvable. By Theorems 5.5 and 5.7, this implies that there is zero duality gap for the pair (ConLP)-(ConLPD), and the primal (ConLP) is solvable. ☐

Remark 5.14. Since the dual conic program (ConLPD) is also a conic program, one can consider (ConLPD) as a primal conic program. In this case the dual is (ConLP). By Theorem 5.12, there is a zero duality gap between this primal-dual pair if there is a point x^* such that $A(x^*) - d \in \text{int}(P)$. Moreover, the dual is solvable. ◁

6 Application: Convex programs

Recall the convex program (CP) and its Lagrangian dual (LD) introduced in Section 1. They are reproduced below for convenience. The primal is

$$\begin{array}{rrcll} \max_{x \in \mathbb{R}^n} & f(x) & & & \\ \text{s.t.} & g_i(x) & \geq & 0 & \text{for } i = 1, \ldots, p \\ & x \in \Omega & & & \end{array} \tag{CP}$$

where $f(x)$ and $g_i(x)$, $i = 1, \ldots, p$ are concave functions, and Ω is a closed, convex set. The dual is

$$\inf_{\lambda \geq 0} L(\lambda) \tag{LD}$$

where L is the Lagrangian function

$$L(\lambda) := \max\{f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) \,:\, x \in \Omega\}.$$

Construct a semi-infinite linear program which is shown in Theorem 6.1 to have the same optimal value as the Lagrangian dual (LD). This semi-infinite linear program is

$$\begin{array}{lllll} \inf & \sigma & & & \\ \text{s.t.} & \sigma - \sum_{i=1}^{p} \lambda_i g_i(x) & \geq & f(x) & \text{for } x \in \Omega \\ & \lambda & \geq & 0. & \end{array} \tag{CP-SILP}$$

Construct the finite support dual for (CP-SILP). There are two sets of constraints in (CP-SILP). There are typically an uncountable number of constraints indexed by $x \in \Omega$ and a finite number of nonnegativity, $\lambda \geq 0$, constraints indexed by $\{1, \ldots, p\}$. Thus, the finite support dual elements belong to $\mathbb{R}^{(\Omega \cup \{1,\ldots,p\})}$. The finite support dual defined over $(u, v) \in \mathbb{R}^{(\Omega)} \times \mathbb{R}^p$ is

$$\text{(CP-FDSILP)} \qquad \sup \sum_{x \in \Omega} u(x) f(x) \tag{6.1}$$

$$\text{s.t.} \qquad \sum_{x \in \Omega} u(x) = 1 \tag{6.2}$$

$$-\sum_{x \in \Omega} u(x) g_i(x) + v_i = 0 \quad \text{for } i = 1, \ldots, p \tag{6.3}$$

$$(u, v) \in \mathbb{R}_+^{(\Omega)} \times \mathbb{R}_+^p. \tag{6.4}$$

Recall $v(\text{CP})$ is the optimal value of (CP), $v(\text{LD})$ is the optimal value of (LD), $v(\text{CP-SILP})$ is the optimal value of (CP-SILP) and $v(\text{CP-FDSILP})$ is the optimal value of (6.1)-(6.4).

Theorem 6.1. $v(\text{LD}) = v(\text{CP-SILP})$. Moreover, (CP-SILP) is solvable if and only if there exists $\lambda^* \geq 0$ such that $L(\lambda^*) = \inf_{\lambda \geq 0} L(\lambda)$.

Proof. First show $v(\text{LD}) \geq v(\text{CP-SILP})$. If, for every $\lambda \geq 0$, $L(\lambda) = \infty$ then $v(\text{LD}) = \infty$ and the result is immediate. Else, consider any $\bar{\lambda} \geq 0$ such that $L(\bar{\lambda}) < \infty$. Set $\bar{\sigma} = L(\bar{\lambda})$. Then $(\bar{\sigma}, \bar{\lambda})$ is a feasible solution to (CP-SILP) with the same objective value as $L(\bar{\lambda})$. Thus, $L(\bar{\lambda}) \geq v(\text{CP-SILP})$. Since $\bar{\lambda} \geq 0$ was chosen arbitrarily, $\inf_{\lambda \geq 0} L(\lambda) \geq v(\text{CP-SILP})$.

Now show $v(\text{CP-SILP}) \geq v(\text{LD})$. If (CP-SILP) is infeasible then $v(\text{CP-SILP}) = \infty$ and the result is immediate. Otherwise, consider any feasible solution $(\bar{\sigma}, \bar{\lambda})$ to (CP-SILP). Then $\bar{\sigma} \geq L(\bar{\lambda})$ and thus $\bar{\sigma} \geq \inf_{\lambda \geq 0} L(\lambda)$. Since $\bar{\sigma}$ is the objective value of this feasible solution to (CP-SILP), the optimal value of (CP-SILP) is greater than or equal to $\inf_{\lambda \geq 0} L(\lambda)$.

The second part follows from very similar arguments. □

Theorem 6.2. $v(\text{CP}) = v(\text{CP-FDSILP})$.

Proof. First show $v(\text{CP}) \geq v(\text{CP-FDSILP})$. If (6.2)-(6.4) is infeasible, then $v(\text{CP-FDSILP}) = -\infty$ and the result is immediate. Assume (6.2)-(6.4) has feasible solution $(\bar{u}, \bar{v})$. Let $\bar{x} = \sum_{x \in \Omega} x \bar{u}(x)$. This sum is well-defined because $\bar{u}$ has finite support. Show $\bar{x}$ is feasible to (CP). First, since Ω is convex, by (6.2) $\bar{x} \in \Omega$. By (6.3), $-\sum_{x \in \Omega} \bar{u}(x) g_i(x) + \bar{v}_i = 0$ for all $i = 1, \ldots, p$. Since $\bar{v}_i \geq 0$, $\sum_{x \in \Omega} \bar{u}(x) g_i(x) \geq 0$. By (6.2) and concavity of g_i, $g_i(\bar{x}) = g_i(\sum_{x \in \Omega} x \bar{u}(x)) \geq \sum_{x \in \Omega} \bar{u}(x) g_i(x) \geq 0$ for all $i = 1, \ldots, p$. Thus the constraints of (CP) are satisfied. Since f is concave, $f(\bar{x}) = f(\sum_{x \in \Omega} x \bar{u}(x)) \geq \sum_{x \in \Omega} \bar{u}(x) f(x)$ which is the objective value of $(\bar{u}, \bar{v})$ in (6.1).

Now show that $v(\text{CP-FDSILP}) \geq v(\text{CP})$. If (CP) is infeasible, then $v(\text{CP}) = -\infty$ and the result is immediate. Otherwise, consider any feasible solution $\bar{x}$ to (CP). Let $\bar{u} \in \mathbb{R}_+^{(\Omega)}$ be defined by $\bar{u}(\bar{x}) = 1$ and $\bar{u}(x) = 0$ for all $x \neq \bar{x}$. Define $\bar{v} \in \mathbb{R}^p$ by $\bar{v}_i = g_i(\bar{x})$. Since $\bar{x}$ is feasible to (CP), $\bar{v} \in \mathbb{R}^p_+$. Thus, $(\bar{u}, \bar{v})$ is a feasible solution to (6.1). The objective value of $(\bar{u}, \bar{v})$ in (6.1) is $f(\bar{x})$ which is the objective value $\bar{x}$ in (CP). ☐

Remark 6.3. Theorems 6.1 and 6.2 imply

$$v(\text{CP-SILP}) = v(\text{LD}) \geq v(\text{CP}) = v(\text{CP-FDSILP})$$

where the inequality follows from weak duality of the Lagrangian dual (or the weak duality of semi-infinite linear programs as discussed in Section 3). ◁

Theorem 6.4 (Slater's theorem for convex programs)**.** Assume the convex program (CP) is feasible and bounded, i.e., $-\infty < v(\text{CP}) < \infty$ and there exists $x^* \in \Omega$ such that $g_i(x^*) > 0$ for all $i = 1, \ldots, p$. Then there is a zero duality gap between the convex program (CP) and its Lagrangian dual (LD) and there exists $\lambda^* \geq 0$ such that $v(\text{LD}) = L(\lambda^*)$, i.e., the Lagrangian dual is solvable.

Proof. Write the semi-infinite linear program defined by system (CP-SILP) as in Section 3

$$\begin{array}{rcllll} z & - & \sigma & \geq & 0 & \\ & & \sigma - \sum_{i=1}^p \lambda_i g_i(x) & \geq & f(x) & \text{for } x \in \Omega \\ & & \lambda_i & \geq & 0 & \text{for } i = 1, \ldots, p. \end{array} \tag{6.5}$$

Without loss of generality, assume that $f(x)$ is bounded above on Ω. It is valid to replace the objective function $f(x)$, by the concave function $\tilde{f}(x) = \min\{f(x), B\}$, where B is an upper bound on $v(\text{CP})$. Such a bound exists by hypothesis. Therefore, $z = \sigma = B$, $\lambda = 0$ is a feasible solution to (6.5). Let γ be any value of z that is feasible in (6.5) and show

$$\begin{array}{rcllll} \gamma & - & \sigma & \geq & 0 & \\ & & \sigma - \sum_{i=1}^p \lambda_i g_i(x) & \geq & f(x) & \text{for } x \in \Omega \\ & & \lambda_i & \geq & 0 & \text{for } i = 1, \ldots, p \end{array} \tag{6.6}$$

is bounded. By hypothesis, (CP) has a Slater point x^*. Consider the sub-system of (6.6)

$$\begin{array}{rcllll} \gamma & - & \sigma & \geq & 0 & \\ & & \sigma - \sum_{i=1}^p \lambda_i g_i(x^*) & \geq & f(x^*) & \\ & & \lambda_i & \geq & 0 & \text{for } i = 1, \ldots, p. \end{array} \tag{6.7}$$

Since x^* is a Slater point, $g_i(x^*) > 0$, $i = 1, \ldots, p$; which together with $\sigma \leq \gamma$ implies

$$0 \leq \lambda_i \leq (\gamma - f(x^*))/g_i(x^*) \quad \text{for } i = 1, \ldots, p,$$

which in turn implies $f(x^*) \leq \sigma \leq \gamma$ and the set of feasible solutions to (6.7) is bounded. Since (6.7) is a sub-system of (6.6), the set of feasible solutions to (6.6) is bounded. Then by Theorem 3.25, $v(\text{CP-SILP}) = v(\text{CP-FDSILP})$ and (CP-SILP) is solvable. By Remark 6.3, $v(\text{CP}) = v(\text{LD})$. Moreover, since (CP-SILP) is solvable, by Theorem 6.1 there exists λ^* such that $v(\text{LD}) = L(\lambda^*)$. ☐

The following example demonstrates that it is possible to identify a zero duality gap with techniques of this paper, even when a Slater condition fails.

Example 6.5. Consider the convex optimization problem

$$\begin{array}{rl} \max_{x \in \mathbb{R}^n} & 0 \\ \text{s.t.} & 1 - x_1^2 - x_2^2 \geq 0 \\ & -1 + x_1 \qquad \geq 0. \end{array} \tag{6.8}$$

The feasible region is the singleton $\{(1,0)\}$ and so no Slater point exists, however there is a zero duality gap. For this instance, (CP-SILP) is

$$\begin{array}{llcll} \inf & \sigma & & & \\ \text{s.t.} & \sigma + \lambda_1(x_1^2 + x_2^2 - 1) + \lambda_2(1 - x_1) & \geq & 0 & \text{for } x \in \mathbb{R}^n \\ & \lambda & \geq & 0. & \end{array} \tag{6.9}$$

Setting $(\sigma, \lambda_1, \lambda_2) = (0,0,0)$ shows that this semi-infinite linear program (SILP) is feasible. Notice also that the right-hand function $b : \mathbb{R}^n \to \mathbb{R}$, is the zero function. Applying Fourier-Motzkin elimination to (6.9) gives $\tilde{b}(h) = 0$ for all h and this implies $\sup_{h \in I_3} \tilde{b}(h) = 0$. Also, for any $\delta \geq 0$,

$$\omega(\delta) = \sup_{h \in I_4} \left\{ \tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \right\} = \sup_{h \in I_4} \left\{ -\delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \right\} \leq 0.$$

Then $\sup_{h \in I_3} \tilde{b}(h) \geq \lim_{\delta \to \infty} \omega(\delta)$ and by Theorem 3.21 there is a zero duality gap between (6.9) and its finite support dual. By Theorem 6.1 and 6.2 this implies there is a zero duality gap between (6.8) and its Lagrangian dual. ◁

7 Application: Generalized Farkas' Theorem

In this section, Fourier-Motzkin elimination is used to prove the generalized Farkas' theorem. Consider a closed convex set given as the intersection of (possibly infinitely many) halfspaces

$$P = \{x \in \mathbb{R}^n \mid a^1(i)x_1 + \cdots + a^n(i)x_n \geq b(i) \text{ for } i \in I\}, \tag{7.1}$$

where I is any index set, $a^1, \ldots, a^n$ and b are elements of $\mathbb{R}^I$. An inequality $c^\top x \geq d$ is a *consequence* of the system of inequalities $a^1(i)x_1 + \ldots a^n(i)x_n \geq b(i)$, $i \in I$ if $c^\top x \geq d$ for every $x \in P$. If $P = \emptyset$, then every inequality is a consequence the inequalities $a^1(i)x_1 + \ldots a^n(i)x_n \geq b(i)$, $i \in I$. Let α^i denote the vector in $\mathbb{R}^n$ given by $\alpha^i = (a^1(i), \ldots, a^n(i))^\top$. The notation 0_n is used to denote the n-dimensional vector of zeros.

Theorem 7.1 (Generalized Farkas' Theorem, see Theorem 3.1 in Goberna and López [9]). The inequality $c^\top x \geq d$ is a consequence of $(\alpha^i)^\top x \geq b(i)$ for all $i \in I$, if and only if at least one of the following holds:

(i)
$$\begin{bmatrix} c \\ d \end{bmatrix} \in \text{cl}\left(\text{cone}\left(\left\{\begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \; i \in I\right\}\right)\right)$$

(ii)
$$\begin{bmatrix} 0_n \\ 1 \end{bmatrix} \in \text{cl}\left(\text{cone}\left(\left\{\begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \; i \in I\right\}\right)\right).$$

Proof. First show that the condition is sufficient. Assume (i) holds. If $\begin{bmatrix} \gamma \\ \delta \end{bmatrix}$ is any vector in $\text{cone}\left(\left\{\begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \; i \in I\right\}\right)$, then $\gamma^\top x \geq \delta$ for every $x \in P$. Since $\begin{bmatrix} c \\ d \end{bmatrix} \in \text{cl}\left(\text{cone}\left(\left\{\begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \; i \in I\right\}\right)\right)$, there exists a sequence $\lim_{j \to \infty} \begin{bmatrix} \gamma^j \\ \delta^j \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$ such that each $\begin{bmatrix} \gamma^j \\ \delta^j \end{bmatrix}$ belongs to $\text{cone}\left(\left\{\begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \; i \in I\right\}\right)$. Thus, $(\gamma^j)^\top x - \delta^j \geq 0$

for every j and every $x \in P$ and $\lim_{j\to\infty}((\gamma^j)^\top x - \delta^j) = c^\top x - d \geq 0$ for every $x \in P$. Thus $c^\top x \geq d$ is a consequence. Now assume that (ii) above holds. Then, by the same reasoning, $0_n^\top x \geq 1$ is satisfied for all $x \in P$. This implies $P = \emptyset$ and then *any* inequality $c^\top x \geq d$ is a consequence.

For the other direction, assume $c^\top x \geq d$ is a consequence. There are two cases, depending on whether P is empty or not.

Case 1: $P = \emptyset$. Apply the Fourier-Motzkin elimination procedure to the constraints that define P in (7.1) and obtain the system (2.8)-(2.9) with the corresponding index sets H_1 and H_2. Since $P = \emptyset$, by Theorem 2.13 either $\tilde{b}(h) > 0$ for some $h^* \in H_1$, or $\sup\{\tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}^k(h)| : h \in H_2\} = \infty$. Consider these two cases in turn:

Case 1a: $\tilde{b}(h^*) > 0$ for some $h^* \in H_1$. By Theorem 2.9, there exists $u^{h^*} \in \mathbb{R}_+^{(I)}$ with finite support such that $\langle a^j, u^{h^*}\rangle = 0$ for all $j = 1, \ldots, n$ and $\langle b, u^{h^*}\rangle > 0$. Using the multipliers $\frac{u^{h^*}}{\langle b, u^{h^*}\rangle}$ for the constraints corresponding to the non-zero elements in u^{h^*} to aggregate constraints, gives $\begin{bmatrix} 0_n \\ 1 \end{bmatrix} \in \text{cone}\left(\left\{\begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \, i \in I\right\}\right)$. Condition (ii) in the statement of the theorem is satisfied.

Case 1b: $\sup_{h\in H_2} \tilde{b}(h)/\sum_{k=\ell}^n |\tilde{a}^k(h)| = \infty$. This implies that there is a sequence $h_m \in H_2$, $m = 1, 2, \ldots$ such that $\tilde{b}(h_m)/\sum_{k=\ell}^n |\tilde{a}^k(h_m)| > m$. This implies $\tilde{b}(h_m) > 0$ for all m. Rearranging the terms, gives

$$\lim_{m\to\infty} \frac{\sum_{k=\ell}^n |\tilde{a}^k(h_m)|}{\tilde{b}(h_m)} = 0.$$

The above limit implies

$$\lim_{m\to\infty} \frac{\tilde{a}^k(h_m)}{\tilde{b}(h_m)} = 0$$

for $k = \ell, \ell + 1, \ldots, n$. By Theorem 2.9, there exists $u^{h_m} \in \mathbb{R}_+^{(I)}$ with finite support such that $\langle a^j, u^{h_m}\rangle = 0$ for $j = 1, \ldots, \ell - 1$, $\langle a^j, u^{h_m}\rangle = \tilde{a}^j(h_m)$ for $j = \ell, \ldots, n$ and $\langle b, u^{h_m}\rangle = \tilde{b}(h_m)$. Since $\tilde{b}(h_m) > 0$, $\langle a^j, \frac{u^{h_m}}{\tilde{b}(h_m)}\rangle = 0$ for all $j = 1, \ldots, \ell - 1$, $\langle a^j, \frac{u^{h_m}}{\tilde{b}(h_m)}\rangle = \frac{\tilde{a}^j(h_m)}{\tilde{b}(h_m)}$ for $j = \ell, \ldots, n$ and $\langle b, \frac{u^{h_m}}{\tilde{b}(h_m)}\rangle = 1$. Since $\lim_{m\to\infty} \frac{\tilde{a}^j(h_m)}{\tilde{b}(h_m)} = 0$ for $j = 1, \ldots, n$, this gives a sequence of points in $\text{cone}\left(\left\{\begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \, i \in I\right\}\right)$ that converges to $\begin{bmatrix} 0_n \\ 1 \end{bmatrix}$ and condition (ii) holds.

Case 2: $P \neq \emptyset$. Consider the semi-infinite linear program

$$\begin{array}{lll} \inf_{x\in\mathbb{R}^n} & c^\top x & \\ \text{s.t.} & a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n \geq b(i), & \text{for } i \in I. \end{array} \tag{7.2}$$

If $P \neq \emptyset$, the semi-infinite linear program defined by (7.2) is feasible, i.e., $z^* < \infty$. Since $c^\top x \geq d$ is a consequence, (7.2) is bounded, i.e., $z^* \geq d > -\infty$. Reformulate as in (3.1)-(3.3) and apply Fourier-Motzkin elimination and obtain the system (3.4) with the corresponding index sets I_1, I_2, I_3 and I_4. Then by Lemma 3.8 the primal optimal value is

$$z^* = \max\{\sup_{h\in I_3} \tilde{b}(h), \lim_{\delta\to\infty} \omega(\delta)\}.$$

Again consider two cases :

Case 2a: $z^* = \sup_{h\in I_3} \tilde{b}(h)$. This implies that for any fixed $\epsilon > 0$ there is an $h^* \in I_3$ such that $\tilde{b}(h^*) \geq z^* - \epsilon \geq d - \epsilon$. Since $h^* \in I_3$, Lemma 3.13(iv) implies that there exists $v^{h^*} \in \mathbb{R}^{(I)}$ such that $\langle a^j, v^{h^*}\rangle = c_j$ and $\tilde{b}(h^*) = \langle b, v^{h^*}\rangle \geq d - \epsilon$. Thus, $\begin{bmatrix} c \\ d - \epsilon \end{bmatrix}$ is in

$\text{cone}\left(\left\{\begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \ i \in I\right\}\right)$ where the multiplier for $\begin{bmatrix} 0_n \\ -1 \end{bmatrix}$ is $\tilde{b}(h^*) - (d - \epsilon)$. Since this is true for any $\epsilon > 0$,

$$\begin{bmatrix} c \\ d \end{bmatrix} \in \text{cl}\left(\text{cone}\left(\left\{\begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \ i \in I\right\}\right)\right)$$

and condition (i) of the theorem holds.

Case 2b: $z^* = \lim_{\delta\to\infty} \omega(\delta)$. Since $-\infty < z^* < \infty$, by Lemma 4.2, there exists a subsequence of indices $h_m, m = 1, 2, \ldots$ such that $h_m \in I_4$, $\tilde{a}^k(h_m) \to 0$ for all $k = \ell, \ldots, n$, $\tilde{b}(h_m) \to z^*$ and $\tilde{b}(h_m) \geq z^*$ for all $m \in \mathbb{N}$. Let $\tilde{\alpha}^m \in \mathbb{R}^n$ be defined by $(\tilde{\alpha}^m)_k = 0$ for $k = 1, \ldots, \ell - 1$ and $(\tilde{\alpha}^m)_k = \tilde{a}^k(h_m)$ for $k = \ell, \ldots, n$. By Lemma 3.13(v), for each $m \in \mathbb{N}$, $\tilde{\alpha}^m = \alpha^m - c$, for some $\alpha^m \in \text{cone}(\{\alpha^i\}_{i\in I})$. Renaming $\tilde{b}(h_m) = b_m$, gives

$$\begin{bmatrix} \alpha^m \\ b_m \end{bmatrix} \in \text{cone}\left(\left\{\begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \ i \in I\right\}\right)$$

and

$$\begin{bmatrix} \tilde{\alpha}^m \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha^m - c \\ b_m - d \end{bmatrix} + (b_m - d)\begin{bmatrix} 0_n \\ -1 \end{bmatrix}.$$

Since $\tilde{\alpha}^m \to 0$ as $m \to \infty$,

$$\begin{aligned} & \begin{bmatrix} \alpha^m - c \\ b_m - d \end{bmatrix} + (b_m - d)\begin{bmatrix} 0_n \\ -1 \end{bmatrix} \to 0 \\ \Rightarrow \ & \begin{bmatrix} \alpha^m \\ b_m \end{bmatrix} + (b_m - d)\begin{bmatrix} 0_n \\ -1 \end{bmatrix} - \begin{bmatrix} c \\ d \end{bmatrix} \to 0 \\ \Rightarrow \ & \begin{bmatrix} \alpha^m \\ b_m \end{bmatrix} + (b_m - d)\begin{bmatrix} 0_n \\ -1 \end{bmatrix} \to \begin{bmatrix} c \\ d \end{bmatrix}. \end{aligned}$$

Now $b_m \geq d$ because $b_m = \tilde{b}(h_m) \geq z^* \geq d$. Therefore

$$\begin{bmatrix} c \\ d \end{bmatrix} \in \text{cl}\left(\text{cone}\left(\left\{\begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix}; \ i \in I\right\}\right)\right)$$

and condition (i) of the theorem holds. □

8 Application: Further results for semi-infinite linear programs

8.1 Additional sufficient conditions for zero duality gap

By looking at the recession cone of (3.18) it is possible gain further insights and discover useful sufficient conditions for zero duality gaps in general semi-infinite linear programs. We show results first discovered by Karney [10] follow directly and easily from our methods. The recession cone of (3.18) is defined by the system

$$\begin{aligned} -c_1x_1 - c_2x_2 - \cdots - c_nx_n &\geq 0 && (8.1) \\ a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n &\geq 0 \quad \text{for } i \in I. && (8.2) \end{aligned}$$

Applying Fourier-Motzkin elimination to (8.1)-(8.2) gives

$$\begin{aligned} 0 &\geq 0 \quad \text{for } h \in H_1 && (8.3) \\ \tilde{a}^\ell(h)x_\ell + \tilde{a}^{\ell+1}(h)x_{\ell+1} + \cdots + \tilde{a}^n(h)x_n &\geq 0 \quad \text{for } h \in H_2. && (8.4) \end{aligned}$$

Following the notation conventions of Karney [10], K denotes the recession cone of (SILP), given by the inequalities (8.2) and N denotes the null space of the objective function vector c.

Lemma 8.1. If H_2 is nonempty in (8.4), then there exists a ray $r \in \mathbb{R}^n$ satisfying (8.1)-(8.2) with at least one of the inequalities in (8.1)-(8.2) satisfied strictly.

Proof. If H_2 is nonempty, there is a $k \geq \ell$ such that $\tilde{a}^k(h)$ is nonzero for at least one $h \in H_2$. Since x_k is a dirty variable, the nonzero $\tilde{a}^k(h)$ are of the same sign for all $h \in H_2$. If the $\tilde{a}^k(h)$ are all nonnegative, then set $x_k = 1$ and $x_i = 0$ for $i \neq k$; if the $\tilde{a}^\ell(h)$ are all nonpositive, then set $x_k = -1$ and $x_i = 0$ for $i \neq k$. This solution to (8.3)-(8.4) satisfies at least one of the inequalities in (8.3)-(8.4) strictly. Since this is the projection of some r satisfying (8.1)-(8.2), this r must satisfy at least one inequality in (8.1)-(8.2) strictly, since all inequalities in (8.3)-(8.4) are conic combinations of inequalities in (8.1)-(8.2). $\square$

Theorem 8.2. If (SILP) is feasible and $K \cap N$ is a subspace, then $v(\text{SILP}) = v(\text{FDSILP})$.

Proof. *Case 1: H_2 in (8.4) is empty.* Observe that the columns in systems (8.1)-(8.2) and (3.2)-(3.3) are identical for variables $x_1, \ldots, x_n$. This means if x_k is eliminated when Fourier-Motzkin elimination is applied to one system, it is eliminated in the other system. Since H_2 in (8.4) is empty, (SILP) is tidy. Then by Theorem 3.24, $v(\text{SILP}) = v(\text{FDSILP})$.

Case 2: H_2 in (8.4) is not empty. If H_2 is not empty, by Lemma 8.1, there exists r satisfying (8.1)-(8.2) such that at least one of the inequalities in (8.1)-(8.2) is satisfied strictly. If $c^T r < 0$ and $r \in K$, then $v(\text{SILP}) = -\infty$. Therefore (FDSILP) is infeasible by weak duality and $v(\text{SILP}) = v(\text{FDSILP}) = -\infty$. If $c^T r = 0$ then the constraint (8.1) is tight at r and so $r \in N$. Then $r \in K \cap N$ which is a subspace by hypothesis. Then $-r \in K \cap N$. This implies $r \in K \cap -K$. But this means that r satisfies all inequalities in (8.2) at equality and this contradicts the fact established for this case that at least one inequality in (8.1)-(8.2) is strict. $\square$

8.2 Finite approximation results

Consider an instance of (SILP) and the corresponding finite support dual (FDSILP). For any subset $J \subseteq I$, define SILP(J) as the semi-infinite linear program with only the constraints indexed by J and the same objective function, and $v(J)$ the optimal value of SILP(J). For example, if J is a finite subset of I, SILP(J) is a finite linear program.

Theorem 8.3. If (SILP) is feasible, then $v(\text{FDSILP}) = \sup\{v(J) : J \text{ is a finite subset of } I\}$.

Proof. First show $v(\text{FDSILP}) \leq \sup\{v(J) : J \text{ is a finite subset of } I\}$. By hypothesis, (SILP) is feasible and this implies by Corollary 3.19 that $v(\text{FDSILP}) = \sup_{h \in I_3} \tilde{b}(h)$. Therefore, for every $\epsilon > 0$, there exists $h^* \in I_3$ such that $v(\text{FDSILP}) - \epsilon \leq \tilde{b}(h^*)$. By Lemma 3.13(iv), there exists $v^{h^*} \in \mathbb{R}^{(I)}$ with support J^* such that $\tilde{b}(h^*) = \langle b, v^{h^*} \rangle = \sum_{i \in J^*} b(i) v^{h^*}(i)$, and $\sum_{i \in J^*} a^k(i) v^{h^*}(i) = c_k$. Since (SILP) is feasible, SILP(J^*) is feasible; let $\bar{x}$ be *any* feasible solution to this finite LP. Thus,

$$
\begin{aligned}
c^T \bar{x} &= \textstyle\sum_{k=1}^n c_k \bar{x}_k \\
&= \textstyle\sum_{k=1}^n (\sum_{i \in J^*} a^k(i) v^{h^*}(i)) \bar{x}_k \\
&= \textstyle\sum_{i \in J^*} (\sum_{k=1}^n a^k(i) \bar{x}_k) v^{h^*}(i) \\
&\geq \textstyle\sum_{i \in J^*} b(i) v^{h^*}(i) \\
&= \tilde{b}(h^*).
\end{aligned}
$$

Since this holds for any feasible solution to SILP(J^*), $v(J^*) \geq \tilde{b}(h^*) \geq v(\text{FDSILP}) - \epsilon$. Thus, for every $\epsilon > 0$, there exists a finite $J^* \subseteq I$ such that $v(J^*) \geq v(\text{FDSILP}) - \epsilon$. Hence, $v(\text{FDSILP}) \leq \sup\{v(J) : J \text{ is a finite subset of } I\}$.

Next show that $v(\text{FDSILP}) \geq \sup\{v(J) : J \text{ is a finite subset of } I\}$. Consider any finite $J^* \subseteq I$. It suffices to show that $v(\text{FDSILP}) \geq v(J^*)$. If $v(J^*) = -\infty$, then the result is immediate. So assume $v(J^*) > -\infty$. Then SILP(J^*) is bounded. Since (SILP) is feasible by hypothesis, SILP(J^*) is also feasible. Then by Theorem 3.26, there exists $v^* \in \mathbb{R}^{J^*}$ such that $\sum_{i \in J^*} b(i)v^*(i) = v(J^*)$ and $\sum_{i \in J^*} a^k(i)v^*(i) = c_k$. Define $\bar{v} \in \mathbb{R}^{(I)}$ by $\bar{v}(i) = v^*(i)$ for $i \in J^*$ and $\bar{v}(i) = 0$ for $i \notin J^*$. Thus, $\bar{v}$ is a feasible solution to (FDSILP) with objective value $v(J^*)$. Therefore, $v(\text{FDSILP}) \geq v(J^*)$. □

Theorem 8.3 is used to prove a series of results by Karney [10]. Consider a semi-infinite linear program with countably many constraints, i.e., $I = \mathbb{N}$. For every $n \in \mathbb{N}$, let P_n denote the finite linear program formed using the constraints indexed by $\{1, \ldots, n\}$ and the same objective function. Let $v(P_n)$ denote its optimal value.

Corollary 8.4. If (SILP) is feasible with $I = \mathbb{N}$, then $\lim_{n\to\infty} v(P_n) = v(\text{FDSILP})$.

Proof. Since $\{1, \ldots, n\}$ is a finite subset of I, $v(P_n) \leq \sup\{v(J) : J \text{ is a finite subset of } I\} = v(\text{FDSILP}) < \infty$ where the equality follows from Theorem 8.3 and the "$<$" follows from weak duality since (SILP) is feasible. Since $v(P_n)$ is a nondecreasing sequence of real numbers bounded above, $\lim_{n\to\infty} v(P_n)$ exists and $\lim_{n\to\infty} v(P_n) \leq v(\text{FDSILP})$. Next prove that $\lim_{n\to\infty} v(P_n) \geq v(\text{FDSILP})$. Observe that for any finite subset $J^* \subseteq I$ there exists $n^* \in \mathbb{N}$ such that $J^* \subseteq \{1, \ldots, n^*\}$ and this implies $v(P_{n^*}) \geq v(J^*)$. Thus, $\lim_{n\to\infty} v(P_n) \geq \sup\{v(J) : J \text{ is a finite subset of } I\} = \text{v(FDSILP)}$ where the equality follows from Theorem 8.3. □

Corollary 8.5 (Karney [10] Theorem 2.1)**.** If the feasible region of (SILP) with $I = \mathbb{N}$ is nonempty and bounded, then $\lim_{n\to\infty} v(P_n) = v(\text{SILP})$.

Proof. This follows from Theorem 3.25 and Corollary 8.4. □

Corollary 8.6 (Karney [10] Theorem 2.4)**.** If (SILP) with $I = \mathbb{N}$ is feasible and the zero vector is the unique solution to the system (8.1)-(8.2), then $\lim_{n\to\infty} v(P_n) = v(\text{SILP})$.

Proof. If the zero vector is the unique solution to the system (8.1)-(8.2), then the recession cone of (3.18) is $\{0\}$ and (3.18) is bounded for any value of γ such that (3.18) is feasible (such a γ exists because (SILP) is feasible). The result then follows from Theorem 3.25 and Corollary 8.4. □

Corollary 8.7 (Karney [10] Theorem 2.5)**.** Assume (SILP) with $I = \mathbb{N}$ is feasible and let r be a ray satisfying (8.1)-(8.2). If r is not an element of the null space N, then $\lim_{n\to\infty} v(P_n) = v(\text{SILP}) = -\infty$.

Proof. If $r \in K$ and $r \notin N$, then $c^T r < 0$. This implies $v(\text{SILP}) = -\infty$ and (FDSILP) is infeasible by weak duality. Then $v(\text{SILP}) = v(\text{FDSILP}) = -\infty$ and the result follows from Corollary 8.4. □

Corollary 8.8 (Karney [10] Theorem 2.6)**.** If (SILP) is feasible and $K \cap N$ is a linear subspace, then $\lim_{n\to\infty} v(P_n) = v(\text{SILP})$.

Proof. This follows from Theorem 8.2 and Corollary 8.4. □

9 Conclusion

This paper explores two related themes. The first is how the powerful extension of Fourier-Motzkin elimination to semi-infinite systems of linear inequalities is used to prove and provide insights about duality theory for semi-infinite linear programs. This was the topic of Section 3 where projection is used to characterize feasibility, boundedness, solvability and the duality gap between primal and dual semi-infinite linear programs. In Sections 7 and 8 we established connections between our approach and the duality theories of Goberna and Lopez [9] and Karney [10] respectively. In particular, we prove the generalized Farkas' theorem in Section 7, a result which forms the basis of most of Goberna and Lopez's arguments. Hence, in principle, their work can be obtained from the Fourier-Motzkin approach. In Section 8, we explicitly show how Karney's results follow directly from the Fourier-Motzkin approach. This underscores our claim that projection is a universal and unifying approach to the study of semi-infinite linear programming.

The second theme is that semi-infinite linear programming has implications for finite dimensional convex optimization. Sections 5 and 6 illustrate how both well-known and lesser-known duality results in conic and convex programming are special cases of semi-infinite linear programming duality.

The connection between semi-infinite linear programming and convex optimization is made clear by the method of projection. Fourier-Motzkin elimination is purely algebraic. It is simply the aggregation of pairs of linear inequalities using nonnegative multipliers. The key insight is that topological conditions common in duality theory of finite-dimensional convex optimization are implied by the algebraic conditions for solvability and duality in semi-infinite linear programming via projection.

Both themes, and the connections between them, deserve further exploration. Regarding the first, it might be fruitful to further explore the connections between our characterization of zero duality gap and the characterization presented in Theorem 8.2 of Goberna and López [9]. Goberna and López's approach is topological and based on separating hyperplane theory, whereas our approach is purely algebraic. Our proof of the generalized Farkas' theorem (see our Theorem 7.1 and Theorem 3.1 in Goberna and López [9]) provides a useful starting point for further exploration.

Regarding the second theme, there are at least two avenues for further research. First, all the duality results for finite-dimensional convex optimization considered here were derived by showing the associated semi-infinite linear program was tidy. Recall that when (SILP) is tidy, $\lim_{\delta\to\infty}\omega(\delta) = -\infty$. This condition (along with primal feasibility) suffices to establish primal solvability (Theorem 3.10) and zero duality gap (Theorem 3.21). However, tidiness is far from necessary, as demonstrated in Examples 3.12 and 6.5. Exploring how to translate more subtle sufficient conditions for zero duality gap arising from finite values for $\lim_{\delta\to\infty}\omega(\delta)$ into the language of finite dimensional convex optimization could prove fruitful.

A second avenue is to examine algorithmic approaches to solving convex programs from the viewpoint of semi-infinite programming. As an example, Wolfe [13] proposed a method for solving (CP) using column generation. See also Dantzig [3]. The restricted master of Dantzig and Wolfe corresponds to a modified (CP-FDSILP) containing a *finite* subset of columns. At each step of the Dantzig and Wolfe algorithm, a column with a negative reduced cost is added to the current restricted master. This results in a new restricted master with one more column. Dantzig [3] proves that the optimal value of the restricted master converges to the optimal value of the convex program (CP) as the number of iterations (columns in the restricted master) goes to infinity. In other words, once the number of columns becomes infinite, it is possible to recover the optimal value of the convex program. It would be an interesting project to see if an alternate proof of Dantzig's result, and possibly further insight into his algorithm, derive from a deeper understanding of the semi-infinite linear programs (CP-SILP) and (CP-FDSILP) associated with the convex program (CP).

This paper has not addressed the algorithmic aspects of Fourier-Motzkin elimination applied to semi-infinite linear programs. Obviously, when applied to semi-infinite linear programs, Fourier-Motzkin elimination is not a finite process. However, if the functions $b, a^k \in \mathbb{R}^I$ for $k = 1, \ldots, n$ could be characterized in a reasonably simple format, then symbolic elimination might be possible. This is another avenue of research.

References

[1] E.J. Anderson and P. Nash. *Linear programming in infinite-dimensional spaces: theory and applications.* Wiley, 1987.

[2] A. Charnes, WW Cooper, and K. Kortanek. Duality in semi-infinite programs and some works of Haar and Carathéodory. *Management Science*, 9(2):209–228, 1963.

[3] G. B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, NJ, 1963.

[4] R.J. Duffin. Infinite programs. In H. W. Kuhn and A. W. Tucker, editors, *Linear Inequalities and Related Systems*, pages 157–170. Princeton University Press, Princeton, NJ, 1956.

[5] R.J. Duffin. Clark's theorem on linear programs holds for convex programs. *Proceedings of the National Academy of Sciences*, 75(4):1624–1626, 1978.

[6] R.J. Duffin and L.A. Karlovitz. An infinite linear program with a duality gap. *Management Science*, 12:122–134, 1965.

[7] J. B. J. Fourier. Solution d'une question particulière du calcul des inégalités. *Oeuvres II Paris*, pages 317–328, 1826.

[8] B. Gärtner and J. Matousék. *Approximation Algorithms and Semi-Definite Programming.* Springer-Verlag, 2012.

[9] M.A. Goberna and M.A. López. *Linear semi-infinite optimization.* John Wiley & Sons Chichester, 1998.

[10] D. F. Karney. Duality gaps in semi-infinite linear programming – an approximation problem. *Mathematical Programming*, 20:129–143, 1981.

[11] T. S. Motzkin. *Beitrage zur Theorie der Linearen Ungleichungen.* PhD thesis, University of Besel, Jerusalem, 1936.

[12] H. P. Williams. Fourier's method of linear programming and its dual. *The American Mathematical Monthly*, 93:681–695, 1986.

[13] P. Wolfe. Methods of nonlinear programming. In J. Abadie, editor, *Nonlinear Programming*, pages 100–142. John Wiley and Sons, New York, NY, 1967.