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

Amitabh Basu
Kipp Martin
Christopher Thomas Ryan

March 17, 2014

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 the Slater constraint qualification, are reduced to guaranteeing that there are no dirty variables. This leads to completely new proofs of such sufficient conditions for zero duality.

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 constraint qualification* for convex programming. Slater's condition states that when there is feasible point that strictly satisfies all the convex constraints of the primal convex program (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. For surveys covering theory, applications and algorithms

for semi-infinite linear programming see Goberna and López [11], Hettich and Kortanek [12] and López [16]. In this paper we use the duality theory of semi-infinite linear programs 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 [2], Charnes, Cooper and Kortanek [4], Duffin and Karlovitz [6], Goberna and López [10], Karney [14], Kortanek [15], and Shapiro [20]). We provide an alternate and unifying approach to duality in semi-infinite linear programs.

We extend Fourier-Motzkin elimination (projection) [7, 18] 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, our approach does not rely on the theory of topological vector spaces or convex analysis. Instead, our approach combines simple aggregation of pairs of linear inequalities using nonnegative multipliers (an algebraic operation) with simple analysis on the reals, $\mathbb{R}$. 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 the Slater constraint qualification, are reduced to guaranteeing that there are no dirty variables in an appropriately defined semi-infinite linear program. To the author's knowledge, using Fourier-Motzkin elimination to study study semi-infinite linear programs is new. Blair [3] employed a Fourier-Motzkin elimination technique to extend a result by Jersoslow and Kortanek [13] on the feasibility of semi-infinite linear systems. However, their systems were over the ordered field $\mathbb{R}(M)$ obtained by appending the reals $\mathbb{R}$ with a transcendental number M larger than every real. By contrast, we consider systems over the reals $\mathbb{R}$. This allows us to study optimality and duality of semi-infinite linear programs.

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 can be leveraged to provide new proofs of some classical results in finite-dimensional conic linear programs and convex optimization. See Section A.3 and Section 5, respectively. Application of the results from Section 3 to the generalized Farkas' theorem for semi-infinite linear programs is given in Section 6. Additional sufficient conditions for zero duality gap in semi-infinite linear programs are in Section A.5 of the Electronic Companion.

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, 2^I denotes the power set of I and $\mathbb{R}^I$ 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 $i = h$, and $e^h(i) = 0$ if $i \neq h$. 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 an optimization problem $(*)$ is denoted by $v(*)$. If the objective of $(*)$ is a supremum and the problem is (i) unbounded then set $v(*) = \infty$ or (ii) infeasible then set $v(*) = -\infty$. Conversely, if the objective of $(*)$ is an infimum and the problem is (i) unbounded then set $v(*) = -\infty$ or (ii) infeasible then set $v(*) = \infty$.

Our results. The main topic of study is the semi-infinite program

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

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*

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

Our main results on this primal-dual pair are summarized in Table 1 (see page 21). These include a sufficient condition for primal solvability (Theorem 3.7) and characterizations of both dual solvability (Theorem 3.17) and zero duality gap (Theorem 3.18). 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 3.23). 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 3.23. 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).

The theory of tidy semi-infinite linear programs is leveraged to obtain new proofs of important duality results in conic and convex programming. In the main text of this manuscript (Section 5), we discuss convex programming. Section A.3 in the appendix (electronic companion) contains the discussion for conic programs. We consider the following general convex program

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

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

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

The following is a well-known key result in finite-dimensional convex programming.

Theorem 5.2 (Slater's theorem for convex programs). Suppose the convex program (CP) is feasible and bounded. Moreover, suppose there exists an $x^* \in \Omega$ 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 a $\lambda^* \geq 0$ such that $v(\text{LD}) = L(\lambda^*)$, i.e., the Lagrangian dual is solvable.

We provide a completely new proof of this classical result in Section 5. 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 can be used to show the tidiness of the system. By Theorem 3.24, this implies zero duality gap and dual solvability.

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. These results include the generalized Farkas' theorem for infinite systems of linear inequalities (see our Theorem 6.1 and Theorem 3.1 in Goberna and López [10]). Our proof does not rely on the theory of separating hyperplanes and thus does not mimic known proofs. Goberna and López use this 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. Other authors have also given characterizations of properties of primal-dual pairs of semi-infinite linear programs. A comparison is given in Section 3.4. Additional results on the finite approximability of semi-infinite linear programs are in Section A.5 of the Electronic Companion.

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 [18], and Williams [22]. In this section, Fourier-Motzkin 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.

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 $(x_1, \ldots, x_n) \in \mathbb{R}^n$ 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

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

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

4

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 $(x_2, \ldots, x_n) \in \mathbb{R}^{n-1}$ 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,

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

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 exist 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, $P(\Gamma; x_1) = FM(\Gamma; x_1)$. We show $FM(\Gamma; x_1)$ 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

5

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. $\square$

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

<div align="center">FOURIER-MOTZKIN ELIMINATION PROCEDURE</div>

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

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

with $\tilde{I} \subseteq 2^I$ and $\tilde{a}^k \in \mathbb{R}^{\tilde{I}}$. 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: $\mathcal{D} \leftarrow \{1, \ldots, n\}$, $\tilde{I} \leftarrow \{\{i\} \mid i \in I\}$, $\tilde{a}^k(\{i\}) \leftarrow a^k(i)$ for all $i \in I$ and $k \in \mathcal{D}$, $\tilde{b} \leftarrow b$, 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 $\tilde{I} \leftarrow \mathcal{H}_0(j) \cup \{p \cup q \mid p \in \mathcal{H}_+(j), q \in \mathcal{H}_-(j)\}$ 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

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

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

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

<div align="center">6</div>

(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

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

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 $d_i \leq d_j$ for $i \leq j$. 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 $\bar{d}_i \leq \bar{d}_j$ for $i \leq j$. In other words, $\ell - 1$ variables were eliminated and the remaining $n - \ell + 1$ variables indexed by the indices in $\mathcal{D}$ are not eliminated.

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

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

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

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

Remark 2.4. In the above procedure, $\tilde{I}$ is redefined in every iteration but remains a subset of 2^I; in particular, a family of finite subsets of I. The domain $\tilde{I}$ of the functions $\tilde{a}^k$ are redefined correspondingly. In contrast, the domain I of the functions u^h for $h \in \tilde{I}$ is unchanged throughout. The superscript h is the support of u^h. ◁

Examples 2.17, 3.8, 3.9, A.18 and 5.4 and Remark 3.14 below will illustrate various aspects of the Fourier-Motzkin elimination procedure.

Definition 2.5 (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 2.6 (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 2.7. For every semi-infinite linear system, there exists a permutation of the variables that puts it into canonical form. Moreover, if one applies 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. □

Remark 2.8. In light of Lemma 2.7, we may assume without loss of generality, 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. However, if a permutation reveals a dirty variable then at least one dirty variable will exist in every permutation. For details see Theorem A.1 in the Electronic Companion. For our purposes, the permutation of variables does not effect any of our results. Any permutation that puts the semi-infinite linear system into a canonical form will suffice. ◁

Definition 2.9. 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 2.10. 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 2.11 (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). Suppose M iterations of Step 2 of the Fourier-Motzkin elimination procedure yields the 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 $H_1 := \{h \in \tilde{I} : \tilde{a}^k(h) = 0 \text{ for all } k \in \{\ell, \ldots, n\}\}$ 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 2.12. 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 2.13. 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 2.14 (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 $\qquad$ (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 2.10, $P(\Gamma; x_1, \ldots, x_{\ell-1})$ 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. Since (i) and (2.8) are equivalent it remains to show (2.9) holds if and only if (ii) holds.

($\Longrightarrow$) 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 (ii) holds. Thus, there exists a $\delta \geq \max\{0, \sup_{h \in H_2} \tilde{b}(h) / \sum_{k=\ell}^n |\tilde{a}^k(h)|\}$. We show $x(\delta; \ell)$ satisfies (2.9). 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 $\delta \geq \sup_{h \in H_2} \tilde{b}(h) / \sum_{k=\ell}^n |\tilde{a}^k(h)|$.

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 2.11 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. $\qquad \square$

Remark 2.15. In the proof of Theorem 2.14 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 2.12 is feasible to (2.8)-(2.9) and thus can be extended to a feasible vector in Γ. This fact is used below. ◁

We next characterize the boundedness of the feasible set Γ.

Theorem 2.16 (Boundedness). If (2.1) defines a nonempty bounded set Γ then, after applying the Fourier-Motzkin elimination procedure, the resulting system (2.8)-(2.9) has $H_2 = \emptyset$.

9

Proof. The result follows from Theorem A.2 in the electronic companion, because in this case $\text{rec}(\Gamma) = \text{lin}(\Gamma) = \{0\}$. □

Example 2.17. The opposite implication in Theorem 2.16 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 2.18 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 2.18. Applying Fourier-Motzkin elimination to (2.1) gives (2.6). Let $\bar{u} \in \mathbb{R}_+^{(I)}$ such that $\langle a^k, \bar{u} \rangle = 0$ 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 $\bar{u} = \sum_{h \in \bar{I}} \lambda_h u^h$.

Proof. Proceed by induction on n. First prove the inductive step on n and then the $n = 1$ step. We 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. We invoke the induction hypothesis for the $n - 1$ variable system (2.13). For this to be valid, all the hypotheses for the $n-1$ system must hold. 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 index of 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 $\langle a^k, \bar{u} \rangle = 0$ 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 exist a finite index set $\{1, \ldots, t\} = \bar{I} \subseteq \hat{I}$ 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 $\hat{\alpha}_j > 0$ such that

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

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

By Theorem 2.10, 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). We claim that either (i) $\langle a^n, w^j \rangle = 0$ for $j = 1, \ldots, t$ or (ii) there exist $j^+, j^- \in \{1, \ldots, t\}$ such that $\langle a^n, w^{j^+} \rangle > 0$ and $\langle a^n, w^{j^-} \rangle < 0$. This follows since conditions (i) and (ii) are exhaustive, indeed $0 = \langle a^n, \bar{u} \rangle =$

10

$\sum_{j=1}^{t} \hat{\alpha}_j \langle a^n, \ w^j \rangle$ for $\hat{\alpha}_j > 0$ and so if $\langle a^n, w^j \rangle \geq 0$ for $j = 1, \dots, t$ (similiarly $\langle a^n, w^j \rangle \leq 0$ for $j = 1, \dots, t$) then $\langle a^n, w^j \rangle = 0$ for $j = 1, \dots, t$.

If (i) holds, and $\langle a^n, w^j \rangle = 0$ for $j = 1, \dots, t$, then $\langle a^k, w^j \rangle = 0$ for $j = 1, \dots, t$, $k = 1, \dots, n$; thus w^j for $j = 1, \dots, t$ 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 $k = 1, \dots, M$ 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 exist 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

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

Condition (2.17) follows since $\langle a^k, w^j \rangle = 0$ for $k = 1, \dots, M-1$ 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 $\sum_{j^- \in J^-} \theta_{j^+, j^-} = \alpha_{j^+}$ for $j^+ \in J^+$ and $\sum_{j^+ \in J^+} \theta_{j^+, j^-} = \alpha_{j^-}$ for $j^- \in J^-$. 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 case $n = 1$. By hypothesis, this forces $M = 1$, i.e., $\langle a^1, \bar{u} \rangle = 0$. If the coefficient of x_1 is zero for all constraints indexed by $\text{supp}(\bar{u})$, then the Fourier-Motzkin procedure initialization step gives multiplers $w^j = e^j$, $j \in \text{supp}(\bar{u})$. Then $\bar{u} = \sum_{j \in \text{supp}(\bar{u})} \bar{u}(j) w^j$.

Otherwise, if variable x_1 has nonzero coefficients in the system indexed by $\text{supp}(\overline{u})$, 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 $\overline{u}_j$ 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 $\overline{u}$. $\qquad\square$

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

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

$$\text{s.t.} \quad -c_1 x_1 - c_2 x_2 - \cdots - c_n x_n + z \;\geq\; 0 \tag{3.2}$$

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

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. For this to make sense we assume without loss of generality that 0 is not an element of I.

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

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

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 $h \in I_2 \cup I_4$.

It is worth noting that the $\tilde{a}^k(h)$ and $\tilde{b}(h)$ in this section are different from those in Section 2. Indeed, by including the constraint (3.2) and enforcing the rule that a coefficient of 1 on z is maintained, the resulting output system will be different than if the Fourier-Motzkin elimination procedure was undertaken on (3.3) alone.

By Theorem 2.10, 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

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

A further step (Lemma 3.4) 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.

As mentioned in the introduction, other authors have made systematic study of semi-infinite programming duality using machinery other than Fourier-Motzkin (see, for instance, Goberna and López [10] and Kortanek [15]). We are not the first to provide characterizations of zero duality gap, dual solvability, etc. However, the content of our characterizations are new. We refer to the specifics of system (3.4), which has not previously appeared in the literature. A brief comparison of our results with those extant in the literature can be found in Section 3.4.

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 3.1 (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 2.14 to the dirty system (3.4) with $H_1 = I_1$ and $H_2 = I_2 \cup I_3 \cup I_4$. $\qquad\square$

Corollary 3.2 below states some consequences of primal feasibility for (SILP) which are useful later. The proof is analogous to the proof of Theorem 2.14. First introduce the function

$$\omega(\delta) := \sup_{h \in I_4} \left\{ \tilde{b}(h) - \delta \sum_{k=\ell}^{n} |\tilde{a}^k(h)| \right\} \tag{3.6}$$

that is used throughout the paper. Note ω can take values in the extended reals. If $I_4 = \emptyset$ then $\omega(\delta) = -\infty$. However, we show in the following corollary that if (SILP) is feasible then $\omega(\delta)$ cannot diverge to ∞. Observe ω is a nonincreasing function of δ since $\sum_{k=\ell}^{n} |\tilde{a}^k(h)| \geq 0$.

Corollary 3.2. If (SILP) is feasible then

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

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

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

(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 3.1. Condition iii) follows from the claim below and condition iv) of Theorem 3.1.

Claim 3.3. $\displaystyle\sup_{h \in I_4} \frac{\tilde{b}(h)}{\sum_{k=\ell}^{n} |\tilde{a}^k(h)| + 1} < \infty \iff \lim_{\delta \to \infty} \omega(\delta) < \infty$.

Proof of Claim. ($\Longrightarrow$) 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 a $\bar{\delta} < \infty$ such that $\bar{\delta} \geq \omega(\bar{\delta})$. Indeed, having $w(\delta) > \delta$ for all δ contradicts $\lim_{\delta \to \infty} \omega(\delta) < \infty$. 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 $h \in I_4$. Rearranging, $\bar{\delta} \geq \tilde{b}(h)/(\sum_{k=\ell}^n |\tilde{a}^k(h)| + 1)$ for $h \in I_4$ 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 3.1. 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 2.10. □

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 3.4. 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

$$z \geq \sup_{h \in I_3} \tilde{b}(h) \tag{3.7}$$

$$z \geq \lim_{\delta \to \infty} \omega(\delta). \tag{3.8}$$

Proof. Since (SILP) is feasible, conditions ii) and iii) in Corollary 3.2 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 3.2.

First, we suppose $\bar{z}$ satisfies (3.7)-(3.8) and show $\bar{z} \in \text{cl}(P(\Lambda; x_1, \ldots, x_n))$. Consider the following two exhaustive cases.

Case 1: $\bar{z} > \lim_{\delta \to \infty} \omega(\delta)$. There exists a $\hat{\delta} \in \mathbb{R}$ such that $\bar{z} > \omega(\hat{\delta})$. Choose $\bar{\delta} \geq \max\{0, \hat{\delta}, \delta_2\}$. By (3.7), $\bar{z} \geq \sup_{h \in I_3} \tilde{b}(h) = \delta_3$. 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 3.2. 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 (3.7), $z_m = \omega(\bar{\delta}_m) \geq \lim_{\delta \to \infty} \omega(\delta) = \bar{z} \geq \sup_{h \in I_3} \tilde{b}(h)$. Hence $z_m \geq \max\{\delta_3, \omega(\bar{\delta}_m)\}$ and by Corollary 3.2(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, we let $\bar{z} \in \text{cl}(P(\Lambda; x_1, \ldots, x_n))$ and show $\bar{z}$ satisfies (3.7) and (3.8). 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 $x^m = (x^m_\ell, \ldots, x^m_n)$ 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), $z_m \geq \sup_{h \in I_4}\{\tilde{b}(h) - \sum_{k=\ell}^n a^k(h)x^m_k\}$. Letting $\bar{\delta}_m = \max_{k=\ell,\ldots,n} |x^m_k|$ gives $z_m \geq \sup_{h \in I_4}\{\tilde{b}(h) - \sum_{k=\ell}^n a^k(h)x^m_k\} \geq \sup_{h \in I_4}\{\tilde{b}(h) - \bar{\delta}_m \sum_{k=\ell}^n |a^k(h)|\} = \omega(\bar{\delta}_m)$. 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 (3.7)-(3.8). $\qquad\square$

By Lemma 3.4, if (SILP) is feasible, then its optimal value is found by solving the optimization problem

$$\begin{aligned} \inf_z \quad & z \\ \text{s.t.} \quad & (3.7) - (3.8). \end{aligned} \tag{3.9}$$

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 3.5. 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 3.6 (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 3.5, $v(\text{SILP}) = -\infty$ 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 $I_3 = \emptyset$. $\qquad\square$

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 (3.9) even though an optimal solution to (SILP) does not exist (see for instance Example 3.8 below). This is due to the fact that (3.9) 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 3.7.

Theorem 3.7 (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 3.2 it follows that $\infty > \sup_{h \in I_3} \tilde{b}(h) > \lim_{\delta \to \infty} \omega(\delta)$. Moreover, by Lemma 3.5, $z^* = \sup_{h \in I_3} \tilde{b}(h)$. Let δ_2 be as defined in Corollary 3.2. Since $\omega(\delta)$ is a nonincreasing function, there exists a $\delta^* \geq \max\{0, \delta_2\}$ 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 3.2 and so $(x(\delta^*; \ell), z^*) \in P(\Lambda; x_1, \ldots, x_\ell)$, 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 $z^* = v(\text{SILP})$. $\qquad\square$

In light of the previous result, one may ask 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 3.8. Consider the following instance of (SILP)

$$
\begin{array}{rll}
\inf x_1 & & \\
x_1 + \frac{1}{t^2}x_2 & \geq & \frac{1}{t^2} + \frac{1}{t} \quad \text{for } t \geq 1 \\
x_1 & \geq & 0.
\end{array}
\tag{3.10}
$$

Applying Fourier-Motzkin elimination to

$$
\begin{array}{rll}
-x_1 \quad + z & \geq & 0 \\
x_1 + \frac{1}{t^2}x_2 & \geq & \frac{1}{t^2} + \frac{1}{t} \quad \text{for } t \geq 1 \\
x_1 & \geq & 0
\end{array}
\tag{3.11}
$$

yields (by eliminating x_1)

$$
\begin{array}{rll}
\frac{1}{t^2}x_2 + z & \geq & \frac{1}{t^2} + \frac{1}{t} \quad \text{for } t \geq 1 \\
z & \geq & 0.
\end{array}
\tag{3.12}
$$

The only I_3 constraint is $z \geq 0$ so $\sup_{h \in I_3} \tilde{b}(h) = 0$. Note 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 attained 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 3.5 the optimal value is 0.

However, for $z = 0$ the system (3.12) 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 (3.12) and the primal is not solvable. ◁

Example 3.9. Consider the following instance of (SILP)

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

Applying Fourier-Motzkin elimination (after introducing the $z - x_1$ constraint) to yields (after projecting out x_1)

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

Observe $I_3 = \{1\}$ and $\sup_{h \in I_3} \tilde{b}(h) = 0$. Note $\omega(\delta) = \sup \left\{ \tilde{b}(h) - \delta \sum_{h \in I_4} |\tilde{a}^k(h)| \ : \ h \in I_4 \right\}$ $= \sup \{ 0 - \delta/h \ : \ h = 3, 4, \ldots \} = 0$. Thus, $\lim_{\delta \to \infty} \omega(\delta) = 0 = \sup_{h \in I_3} \tilde{b}(h)$. By Lemma 3.5, 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 [4])

is the finite support (Haar) dual introduced in Section 1 and reproduced here for convenience.

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

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, $v(\text{SILP}) = v(\text{FDSILP})$.

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 3.10 gives basic properties of v^h that are used later.

Lemma 3.10. 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 part (iv) only. 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 2.10(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). $\square$

3.3.1 Dual feasibility

The next two subsections relate dual feasibility and boundedness to properties of the projected system (3.4). Theorem 2.18 and Lemma 3.10 play pivotal roles in the proofs.

Theorem 3.11 (Dual Feasibility). (FDSILP) is feasible if and only if $I_3 \neq \emptyset$.

Proof. ($\implies$) 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 2.18 to (3.2)-(3.3) with $M = n$, there exist a finite 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

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

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 3.10(ii) and (iv). Now, the 1 in the first component of $(1, \bar{v})$ 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 3.10(iv), v^h is a feasible solution to (FDSILP). $\qquad \square$

3.3.2 Dual boundedness

To characterize dual boundedess, first establish weak duality.

Lemma 3.12 (Weak Duality)**.** Suppose $\tilde{b}(h) \leq 0$ for all $h \in I_1$. If $\bar{v}$ is a feasible dual solution to problem (FDSILP) then

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

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

Proof. Applying Theorem 2.18 as in the proof of Theorem 3.11 implies there exists an index set $\bar{I} \subseteq I_1 \cup I_3$ such that $(1, \bar{v}) = \sum_{h \in \bar{I} \cap I_1} \lambda_h (0, v^h) + \sum_{h \in \bar{I} \cap I_3} \lambda_h (1, v^h)$. Reasoning about the components of $(1, \bar{v})$ separately gives,

$$\bar{v} = \sum_{h \in \bar{I} \cap I_1} \lambda_h v^h + \sum_{h \in \bar{I} \cap I_3} \lambda_h v^h \tag{3.13}$$

and $1 = \sum_{h \in \bar{I} \cap I_3} \lambda_h$. Lemma 3.10(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, (3.13) gives $\langle b, \bar{v} \rangle \leq \sum_{h \in \bar{I} \cap I_3} \lambda_h \langle b, v^h \rangle \leq \langle b, v^{\bar{h}} \rangle = \tilde{b}(\bar{h})$ 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 3.5. $\qquad \square$

Theorem 3.13 (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. We 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 3.12, there exist 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 an $h^* \in I_1$ such that $\tilde{b}(h^*) > 0$ and by Lemma 3.10(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 3.10(i), $\langle b, v^{h_m} \rangle = \tilde{b}(h_m) \to \infty$. By Lemma 3.10(iii), each v^{h_m} is a feasible solution to (FDSILP) and thus (FDSILP) is unbounded. $\qquad \square$

Remark 3.14. Observe that there are two distinct ways for a feasible (FDSILP) to be unbounded. The first is when there is a recession direction to the feasible region that drives the objective value to $+\infty$. From Lemma 3.10(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 3.13.

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 semi-infinite linear program:

$$\inf x_1$$
$$\text{s.t. } x_1 \geq i \quad \text{for } i \in \mathbb{N}$$

with finite support dual

$$\sup \sum_{i \in \mathbb{N}} iv(i)$$
$$\text{s.t. } \sum_{i \in \mathbb{N}} v(i) = 1$$
$$v(i) \geq 0 \quad \text{for } i \in \mathbb{N}$$

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, $\sum_{i \in \mathbb{N}} ie^m(i) = m \to \infty$ as $m \to \infty$. Thus (FDSILP) is unbounded.

Fourier-Motzkin elimination can identify which of the conditions of Theorem 3.13 are violated and result in an unbounded problem. Applying Fourier-Motzkin elimination (after eliminating x_1) the system: $z \geq i$ 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 3.15. 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 3.12(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 3.10(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 3.16. 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 3.1(ii) $\tilde{b}(h) \leq 0$ for all $h \in I_1$. The result follows from Theorem 3.15. If (FDSILP) is feasible and bounded then by Theorem 3.13(i) $\tilde{b}(h) \leq 0$ for all $h \in I_1$. Once again, the result follows from Theorem 3.15. □

Theorem 3.17 (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 attained.

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 3.13(i), $\tilde{b}(h) \leq 0$ for all $h \in I_1$, establishing condition (i). Apply Lemma 3.12(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 3.10(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 3.10(i), the second equality holds from the arguments in the previous two sentences, and the third equality holds from Corollary 3.16. Thus, $\tilde{b}(h^*) = \sup_{h \in I_3} \tilde{b}(h)$, establishing condition (ii).

19

($\Longleftarrow$) 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 3.11. Thus, by Theorem 3.13 (FDSILP) is bounded. Since (FDSILP) is feasible and bounded, by Corollary 3.16 $\sup_{h \in I_3} \tilde{b}(h) = v$(FDSILP). Moreover, Lemma 3.10(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(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). $\qquad\square$

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(SILP) $= v$(FDSILP).

Theorem 3.18 (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, and $\qquad\qquad$ (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 3.16,

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

where the third equality holds by Lemma 3.5. Thus condition (ii) holds.

$\quad$($\Longleftarrow$) Now assume conditions (i) and (ii) hold. By (i) (SILP) is feasible. By Lemma 3.5, v(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 3.16 implies v(FDSILP) $= \sup_{h \in I_3} \tilde{b}(h)$. Thus, v(SILP) $= v$(FDSILP) and there is a zero duality gap. $\qquad\square$

$\quad$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 3.19.

Theorem 3.19 (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 attained 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 3.7. Since (SILP) is feasible, by Theorem 3.1(i), $\tilde{b}(h) \leq 0$ for all $h \in I_1$. Combined with condition (iii) dual solvability follows from Theorem 3.17.

$\quad$Conditions (i) and (ii) imply the sufficient conditions for zero duality gap given in Theorem 3.18 and the duality gap is zero.

$\quad$Conversely, suppose strong duality holds. Then there is a zero duality gap and so Theorem 3.18, (i) and (ii') hold. Theorem 3.17(ii) implies condition (iii). $\qquad\square$

Remark 3.20. Some authors define *strong duality* to mean zero duality gap and dual solvability, excluding the requirement of primal solvability. Under this definition, properties (i), (ii') and (iii) characterize strong duality. $\qquad\triangleleft$

The next two examples demonstrate how strong duality may either hold or not hold when $\sup_{h \in I_3} \tilde{b}(h) = \lim_{\delta \to \infty} \omega(\delta)$.

Example 3.21 (Example 3.8 revisited)**.** In this example the primal is feasible but not solvable, so strong duality fails. However, we showed that $\lim_{\delta \to \infty} \omega(\delta) = \sup_{h \in I_3} \tilde{b}(h) = 0$. ◁

Example 3.22 (Example 3.9 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 3.19 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 3.1)	I_1, I_2, I_3, I_4	Conditions i)-iv) of Theorem 3.1
Primal boundedness (Thm 3.6)	I_3, I_4	Primal feas. and ($I_3 \neq \emptyset$ OR $L > -\infty$)
Primal solvability* (Thm 3.7)	I_3, I_4	Primal feasible and $S > L$
Dual feasibility (Thm 3.11)	I_3	$I_3 \neq \emptyset$
Dual boundedness (Thm 3.13)	I_1, I_3	Dual feas., $\tilde{b}(h) \leq 0$ for all $h \in I_1$, $S < \infty$
Dual solvability (Thm 3.17)	I_1, I_3	$\tilde{b}(h) \leq 0$ for all $h \in I_1$, sup defining S attained
Zero duality gap (Thm 3.18)	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.

As discussed in the introduction, alternate characterizations of these properties have been obtained by other authors. These characterizations build on a different perspective of semi-infinite linear programming, typically based around topological conditions such as lower semicontinuity and closedness in the primal constraint space. They are not immediate consequences of our characterizations, or vice versa.

We invite the reader to compare our results with the following in the literature: primal feasibility (Table II of Kortanek [15], Theoerem 4.4 of Goberna and López [10]), primal boundedness (Table II of Kortanek [15], Theorem 9.3 of Goberna and López [10]), primal solvability (Theorem 7 of Kortanek [15], Table 8.1 of Goberna and López [10], Theorem 2.1 of Shapiro [20]), dual feasibility (Table II of Kortanek [15]), dual boundedness (Table II of Kortanek [15], Theorem 9.7 of Goberna and López [10]), dual solvability (Table 8.1 of Goberna and López [10], Theorem 2.3 of Shapiro [20]), zero duality gap (Table 8.1 of Goberna and López [10], Theorems 2.1 and 2.3 in Shapiro [20]). 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. Fortunately, tidiness is invariant under variable permutations and alternate orders of variable elimination in the Fourier-Motzkin elimination procedure. This follows from the comments in Remark 2.8 and Theorem A.1 in the Electronic Companion.

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 3.23 (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 3.7 implies that the primal is solvable. This establishes (i).

Since $I_3 \neq \emptyset$, (FDSILP) is feasible by Theorem 3.11. Since the primal is feasible, Theorem 3.1(i) and (ii) imply that the dual is bounded via Theorem 3.13. This establishes (ii).

Since the primal is feasible and $\sup_{h \in I_3} \tilde{b}(h) > \lim_{\delta \to \infty} \omega(\delta)$, Theorem 3.18 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. A similar result can be found in Goberna and López [10].

Theorem 3.24 (Bounded System). If there exists a $\gamma \in \mathbb{R}$ such that the system

$$\begin{aligned} -c_1 x_1 - c_2 x_2 - \cdots - c_n x_n &\geq -\gamma \\ a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n &\geq b(i) \quad \text{for } i \in I \end{aligned} \tag{3.14}$$

is feasible and bounded then (SILP) is feasible and tidy. In particular, if the set of solutions $(x_1, \ldots, x_n)$ that satisfy (3.14) 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 (3.14). Observe that the columns in systems (3.14) 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 2.16 guarantees that applying Fourier-Motzkin elimination to (3.14) 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 3.23 are met. Then by Theorem 3.23, (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 3.1). 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 [18]).

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

Theorem 3.25 (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. Note that conditions (i)-(iii) of Theorem 3.19 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 3.6. Thus condition (ii) in Theorem 3.19 holds. Finally, (iii) holds since I_3 is finite whenever I is finite. □

In Section A.2 of the Electronic Companion we illuminate further differences between semi-infinite linear programs and finite linear programs using the tool of Fourier-Motzkin elimination.

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

When I_3 is empty in (3.4), Theorem 3.11 implies that the finite support dual is infeasible. 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 $k = 1, \ldots, n$, $\tilde{a}^k(h_m)$ converge 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 [5].

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,

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

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 (3.6).

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 $\lim_{m \to \infty} \tilde{b}(h_m) = d$ and $\lim_{m \to \infty} \tilde{a}^k(h_m) = 0$ for all $k = \ell, \ldots, n$.

23

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 $\bar{I} = \{h \in \bar{I}, \tilde{b}(h) < d\}$. We consider two cases.

Case 1: $I_4 \setminus \bar{I} = \emptyset$. For any $m \in \mathbb{N}$, setting $\delta = m$, we have that $d \leq \sup\{\tilde{b}(h) - m \sum_{k=\ell}^{n} |\tilde{a}^k(h)| : h \in I_4\}$ and thus, there exists $h_m \in I_4$ such that $d - \frac{1}{m} < \tilde{b}(h_m) - m \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)|$. $I_4 \setminus \bar{I} = \emptyset$ implies $\tilde{b}(h) < d$ for all $h \in I_4$ and therefore we have

$$
\begin{aligned}
& d - \tfrac{1}{m} < d - m \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| \\
\Rightarrow \quad & \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| < \tfrac{1}{m^2}.
\end{aligned}
$$

This shows 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 $k = \ell, \ldots, n$.

Also,

$$
\begin{aligned}
& d - \tfrac{1}{m} < \tilde{b}(h_m) - m \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| \\
\Rightarrow \quad & d - \tfrac{1}{m} < \tilde{b}(h_m)
\end{aligned}
$$

since $m \sum_{k=\ell}^{n} |\tilde{a}^k(h_m)| \geq 0$. Since $\tilde{b}(h_m) < d$ we get $d - \frac{1}{m} < \tilde{b}(h_m) < d$. And so $\lim_{m\to\infty} \tilde{b}(h_m) = d$.

Case 2: $I_4 \setminus \bar{I} \neq \emptyset$. We show 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 we 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 $k = \ell, \ldots, n$. We 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 we 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

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

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. $\square$

24

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 3.5, $z^* = \max\{\sup\{\tilde{b}(h) : h \in I_3\}, \lim_{\delta \to \infty} \omega(\delta)\}$. If $z^* = \sup\{\tilde{b}(h) : h \in I_3\}$, then by Theorem 3.18, there is a zero duality gap, i.e., $z^* = d^*$ where d^* is the optimal value of (FDSILP). From Lemma 4.1, $\hat{d} \leq z^*$, so $z^* = d^*$ implies $\hat{d} \leq d^*$. By definition of limit value, $\hat{d} \geq d^*$. 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 3.10 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, $\hat{d} \geq z^*$. Again, from Lemma 4.1, $\hat{d} \leq z^*$, so $z^* = \hat{d}$. $\qquad\square$

5 Application: Convex programs

Recall the convex program (CP) and its Lagrangian dual (LD) introduced in Section 1. Construct the semi-infinite linear program

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

along with its 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

$$
\text{(CP-FDSILP)} \qquad \sup \sum_{x \in \Omega} u(x)f(x)
\tag{5.1}
$$

$$
\text{s.t.} \qquad \sum_{x \in \Omega} u(x) = 1
\tag{5.2}
$$

$$
-\sum_{x \in \Omega} u(x)g_i(x) + v_i = 0 \quad \text{for } i = 1, \ldots, p
\tag{5.3}
$$

$$
(u, v) \in \mathbb{R}_+^{(\Omega)} \times \mathbb{R}_+^p.
\tag{5.4}
$$

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 $v(\text{CP-SILP})$ is the optimal value of (5.1)-(5.4).

Remark 5.1. We show in the appendix (Theorems A.34 and A.35) the following holds:

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

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). ◁

We are now able to provide a new proof of a very well-known sufficient condition for zero duality gaps in convex programming.

Theorem 5.2 (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 a $x^* \in \Omega$ 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 a $\lambda^* \geq 0$ such that $v(\text{LD}) = L(\lambda^*)$, i.e., the Lagrangian dual is solvable.

Proof. Since $v(\text{CP}) < \infty$, 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 $v(\text{CP})$. Thus, we assume that (CP-SILP) is feasible : $\sigma = B$, $\lambda = 0$ where B is an upper bound on $f(x)$.

We now perform Fourier Motzkin on (CP-SILP) after reformulating as in Section 3:

$$
\begin{array}{rcll}
z \quad - \quad \sigma & \geq & 0 & \\
\sigma - \sum_{i=1}^{p} \lambda_i g_i(x) & \geq & f(x) & \forall x \in \Omega \\
\lambda_i & \geq & 0 & i = 1, \dots, p
\end{array}
$$

We first eliminate variable σ and end up in the following intermediate system during the Fourier-Motzkin elimination procedure:

$$
\begin{array}{rcll}
z - \sum_{i=1}^{p} \lambda_i g_i(x) & \geq & f(x) & \forall x \in \Omega \\
\lambda_i & \geq & 0 & i = 1, \dots, p
\end{array}
\tag{5.5}
$$

Claim 5.3. The variables $\lambda_1, \dots, \lambda_p$ remain clean as the Fourier Motzkin elimination procedure proceeds on (5.5).

Proof of Claim. We now track the intermediate inequalities produced by the Fourier Motzkin elimination procedure as we go through $\lambda_1, \dots, \lambda_p$. We claim that after processing variables $\lambda_1, \lambda_2, \dots \lambda_k$ where $1 \leq k \leq p$ we have the inequality $z - \sum_{i=k+1}^{p} \lambda_i g_i(x^*) \geq f(x^*)$ in the intermediate system of inequalities. We prove this by induction on k.

Consider $k = 1$ first. We have the constraint corresponding to x^*: $z - \sum_{i=1}^{p} \lambda_i g_i(x^*) \geq f(x^*)$ in (5.5). Since $g_1(x^*) > 0$ by hypothesis, the coefficient of λ_1 is negative in this constraint. Moreover, we have the constraint $\lambda_1 \geq 0$. We can multiply the constraint $\lambda_1 \geq 0$ by $g_1(x^*)$ and add to $z - \sum_{i=1}^{p} \lambda_i g_i(x^*) \geq f(x^*)$, resulting in the inequality $z - \sum_{i=2}^{p} \lambda_i g_i(x^*) \geq f(x^*)$. So the base case is done.

Now for the induction step for $k > 1$. By the induction hypothesis, we have the constraint $z - \sum_{i=k}^{p} \lambda_i g_i(x^*) \geq f(x^*)$ after processing $\lambda_1, \dots, \lambda_{k-1}$. Since $g_k(x^*) > 0$ the coefficient of λ_k is negative in this constraint. We also have the constraint $\lambda_k \geq 0$ in the intermediate system obtained after processing $\lambda_1, \dots, \lambda_{k-1}$. Multiplying the constraint $\lambda_k \geq 0$ by $g_k(x^*)$ and adding to $z - \sum_{i=k}^{p} \lambda_i g_i(x^*) \geq f(x^*)$, we obtain the constraint $z - \sum_{i=k+1}^{p} \lambda_i g_i(x^*) \geq f(x^*)$. Thus the induction is complete. †

By Claim 5.3, we have that all variables except z are clean throughout the Fourier-Motzkin elimination procedure. Since (CP-SILP) is feasible (by the discussion in the first paragraph of the proof), by Theorem 3.23 $v(\text{CP-SILP}) = v(\text{CP-FDSILP})$ and (CP-SILP) is solvable. By Remark 5.1 we have $v(\text{CP}) = v(\text{LD})$. Moreover, since (CP-SILP) is solvable, by Theorem A.34 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 5.4. Consider the convex optimization problem

$$
\begin{array}{rll}
\max_{x \in \mathbb{R}^n} & 0 & \\
\text{s.t.} & 1 - x_1^2 - x_2^2 \geq 0 & \\
& -1 + x_1 \quad\;\; \geq 0.
\end{array}
\tag{5.6}
$$

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

$$
\begin{array}{rll}
\inf & \sigma & \\
\text{s.t.} & \sigma + \lambda_1(x_1^2 + x_2^2 - 1) + \lambda_2(1 - x_1) \geq 0 & \text{for } x \in \mathbb{R}^n \\
& \lambda \geq 0.
\end{array}
\tag{5.7}
$$

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 is the zero function. Applying Fourier-Motzkin elimination to (5.7) 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 3.18 there is a zero duality gap between (5.7) and its finite support dual. By Theorem A.34 and A.35 this implies there is a zero duality gap between (5.6) and its Lagrangian dual. ◁

6 Application: Generalized Farkas' Theorem

In this section, Fourier-Motzkin elimination provides an alternate proof of the generalized Farkas' theorem, a well-known cornerstone result in the semi-infinite linear programming literature (see Goberna and López [10]). Consider a closed convex set given as the intersection of (possibly infinitely many) halfspaces

$$P = \{x \in \mathbb{R}^n \mid a^1(i)x_1 + \cdots + a^n(i)x_n \geq b(i) \text{ for } i \in I\}, \tag{6.1}$$

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.

In the theorem below, the difficulty is proving necessity of the conditions. We show how our Fourier-Motkzin approach can be used to prove necessity, as opposed to a separating hyperplane theorem, as was done in Goberna and López [10]. The sufficiency direction is identical to that of Theorem 3.1 in Goberna and López [10] and is omitted.

Theorem 6.1 (Generalized Farkas' Theorem, see Theorem 3.1 in Goberna and López [10]). 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. 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 (6.1) and obtain the system (2.8)-(2.9) with the corresponding index sets H_1 and H_2. Since $P = \emptyset$, by Theorem 2.14 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 2.10, 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 multiplers $\frac{u^{h^*}}{\langle b, u^{h^*} \rangle}$ 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 2.10, 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 $\mathrm{cone}\left(\left\{ \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix} ; i \in I \right\} \right)$ 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

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \quad & c^\top x \\ \text{s.t.} \quad & a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n \geq b(i), \quad \text{for } i \in I. \end{aligned} \tag{6.2}$$

If $P \neq \emptyset$, the semi-infinite linear program defined by (6.2) is feasible, i.e., $z^* < \infty$. Since $c^\top x \geq d$ is a consequence, (6.2) 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 3.5 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 3.10(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 $\mathrm{cone}\left(\left\{ \begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix} ; i \in I \right\} \right)$ where the multiplier for $\begin{bmatrix} 0_n \\ -1 \end{bmatrix}$ is $\tilde{b}(h^*) - (d - \epsilon)$. Since this is true for any $\epsilon > 0$, $\begin{bmatrix} c \\ d \end{bmatrix} \in \mathrm{cl}\left(\mathrm{cone}\left(\left\{ \begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix} ; i \in I \right\} \right) \right)$ 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 $k = \ell, \ldots, n$ and $\tilde{b}(h_m) \to z^*$. 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 3.10(v), for each $m \in \mathbb{N}$, $\tilde{\alpha}^m = \alpha^m - c$, for some $\alpha^m \in \mathrm{cone}(\{\alpha^i\}_{i \in I})$. Renaming $\tilde{b}(h_m) = b_m$, gives $\begin{bmatrix} \alpha^m \\ b_m \end{bmatrix} \in \mathrm{cone}\left(\left\{ \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix} ; i \in I \right\} \right)$ and $\begin{bmatrix} \tilde{\alpha}^m \\ b_m - z^* \end{bmatrix} = \begin{bmatrix} \alpha^m - c \\ b_m - d \end{bmatrix} + (z^* - d) \begin{bmatrix} 0_n \\ -1 \end{bmatrix}$. Since $\tilde{\alpha}^m \to 0$ and $b_m \to z^*$ as $m \to \infty$,

$$\begin{bmatrix} \alpha^m - c \\ b_m - d \end{bmatrix} + (z^* - d) \begin{bmatrix} 0_n \\ -1 \end{bmatrix} \to 0$$

$$\Rightarrow \quad \begin{bmatrix} \alpha^m \\ b_m \end{bmatrix} + (z^* - d) \begin{bmatrix} 0_n \\ -1 \end{bmatrix} - \begin{bmatrix} c \\ d \end{bmatrix} \to 0$$

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

Now $z^* \geq d$. Therefore $\begin{bmatrix} c \\ d \end{bmatrix} \in \mathrm{cl}\left(\mathrm{cone}\left(\left\{ \begin{bmatrix} 0_n \\ -1 \end{bmatrix}, \begin{bmatrix} \alpha^i \\ b(i) \end{bmatrix} ; i \in I \right\} \right) \right)$ and condition (i) of the theorem holds. $\qquad\square$

7   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. The second theme is that semi-infinite linear programming has implications for finite dimensional convex optimization.

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 the duality theory of finite-dimensional convex and conic programming imply simple conditions that ensure duality results. There is no need to appeal to advanced convex analysis or results from the theory of topological vector spaces.

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 [10]. Goberna and López's approach is topological and based on separating hyperplane theory, whereas our approach is based on the purely algebraic Fourier-Motzkin elimination procedure. Our proof of the generalized Farkas' theorem (see our Theorem 6.1 and Theorem 3.1 in Goberna and López [10]) 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 3.7) and zero duality gap (Theorem 3.18). However, tidiness is far from necessary, as demonstrated in Examples 3.9 and 5.4. 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.

This paper has not addressed the algorithmic aspects of Fourier-Motzkin elimination applied to semi-infinite linear programs. There is considerable work on computational approaches to solving semi-infinite linear programs, see for instance Glashoff and Gustavson [9] and Stein and Still [21]. Obviously, when applied to semi-infinite linear programs, Fourier-Motzkin elimination is not a finite process, so a direct comparison with existing computational methods will certainly prove unfavorable for our approach. 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.

Acknowledgements

References

[1] C.D. Aliprantis and K.C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Springer Verlag, second edition, 2006.

[2] E.J. Anderson and P. Nash. *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications.* Wiley, 1987.

[3] C.E. Blair. An extension of a theorem of Jeroslow and Kortanek. *Israel Journal of Mathematics*, 17(1):111–115, 1974.

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

[5] 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.

[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] K. Glashoff and S. Gustafson. *Linear Optimization and Approximation: An Introduction to the Theoretical Analysis and Numerical Treatment of Semi-infinite Programs*. Springer-verlag New York, 1983.

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

[11] M.A. Goberna and M.A. López. Linear semi-infinite programming theory: an updated survey. *European Journal of Operational Research*, 143(2):390–405, 2002.

[12] R. Hettich and K.O. Kortanek. Semi-infinite programming: theory, methods, and applications. *SIAM review*, 35(3):380–429, 1993.

[13] R.G. Jeroslow and K.O. Kortanek. On semi-infinite systems of linear inequalities. *Israel Journal of Mathematics*, 10(2):252–258, 1971.

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

[15] K.O. Kortanek. Classifying convex extremum problems over linear topologies having separation properties. *Journal of Mathematical Analysis and Applications*, 46(3):725–755, 1974.

[16] M.A. López. Stability in linear optimization and related topics. A personal tour. *Top*, 20(2):217–244, 2012.

[17] D.G. Luenberger. *Optimization by Vector Space Methods*. Wiley-Interscience, 1996.

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

[19] A. Shapiro. On duality theory of convex semi-infinite programming. *Optimization*, 54(6):535–543, 2005.

[20] A. Shapiro. Semi-infinite programming, duality, discretization and optimality conditions. *Optimization*, 58(2):133–161, 2009.

[21] O. Stein and G. Still. Solving semi-infinite optimization problems with interior point techniques. *SIAM Journal on Control and Optimization*, 42(3):769–788, 2003.

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

[23] Q. Zhang. Understanding linear semi-infinite programming via linear programming over cones. *Optimization*, 59(8):1247–1258, 2010.

A Electronic Companion

A.1 Invariance of cleanliness under permutations

In this section of the Electronic Companion we provide a geometric interpretation of a *clean system*. Recall a clean system is one where all of the variables are projected out, that is, there are no dirty variables. The key results are Theorem A.1, Theorem A.2, and Theorem A.3. By Theorem A.1, if there is a variable permutation that results in a clean system, then every variable permutation results in a clean system. This is a very useful result. It tells us that if Fourier-Motzkin elimination applied to (SILP) results in a dirty variable, then there is no permutation that could ever make the elimination process find a clean system. Hence there is no need to ever search for such a permutation, it does not exist. Furthermore, by Theorem A.2, if Fourier-Motzkin elimination does result in a clean system, under any permutation, then we know the recession cone of the (SILP) feasible region is equal to the lineality space of the (SILP). Hence dirty variables are always the result of the geometric property that the recession cone is not equal to the lineality space. Finally, in Theorem A.3 we give a necessary and sufficient condition for the Fourier-Motzkin elimination procedure to conclude that the feasible region of (SILP) is bounded.

Theorem A.1. If there exists a permutation of the variables that results in a clean system using Fourier-Motzkin elimination, then every variable permutation results in a clean system.

Proof. By Proposition A.17, if there exists a permutation of the variables that results in a clean system when the Fourier-Motzkin procedure is applied, then every permutation of the variables results in a clean system. □

Recall Γ is the feasible region of the semi-infinite linear system of (SILP). The *recession cone* of Γ is denoted by $\mathrm{rec}(\Gamma)$ and *lineality space* of Γ is denoted by $\mathrm{lin}(\Gamma)$, respectively.

Theorem A.2. Every permutation of the variables results in a clean system using Fourier-Motzkin elimination if and only if $\mathrm{rec}(\Gamma) = \mathrm{lin}(\Gamma)$.

Proof. The logic is as follows.

1. By Definition A.4, there exists a conic index set for (SILP) if and only if $\mathrm{rec}(\Gamma) \neq \mathrm{lin}(\Gamma)$.

2. By Proposition A.16, if (SILP) contains a conic index set, then the Fourier Motzkin elimination procedure will terminate with at least one dirty variable regardless of the variable permutation used in the elimination procedure. By Corollary A.11 if there is a permutation of the variables that results in a dirty variable then (SILP) has a conic index set. Hence (SILP) has a conic index set if and only if there is permutation of the variables that results in a dirty variable using Fourier-Motzkin elimination.

3. By Theorem A.1 there is permutation of the variables that results in a dirty variable using Fourier-Motzkin elimination if and only if there is no permutation of the variable that results in a clean system. Then by item 2., (SILP) contains a conic index set if and only if there is no permutation of the variable that results in a clean system.

4. Items 1. and 3. imply there is no permutation of the variables that results in a clean system if and only if $\mathrm{rec}(\Gamma) \neq \mathrm{lin}(\Gamma)$.

The contrapositive of item 4. gives our result. □

Theorem A.3. If (SILP) is feasible, then the feasible region of (SILP) is bounded if and only if, for every variable permutation, application of the Fourier-Motzkin elimination procedure (see Section 2) to (SILP) results in both $\mathcal{H}_+(j)$ and $\mathcal{H}_-(j)$ nonempty at each iteration of Step 2b.

Proof. Assume without loss the variable permutation is $\{1, 2, \ldots, n\}$ and that at each iteration of step 2b of the Fourier-Motzkin elimination procedure, both $\mathcal{H}_+(j)$ and $\mathcal{H}_-(j)$ are not empty. Show that this implies (SILP) is bounded. Since both $\mathcal{H}_+(j)$ and $\mathcal{H}_-(j)$ are not empty

$$x_j \geq \frac{\tilde{b}(p)}{\tilde{a}^j(p)} - \sum_{k=j+1}^n \frac{\tilde{a}^k(p)}{\tilde{a}^j(p)} x_k, \quad \forall p \in \mathcal{H}_+(j)$$

$$x_j \leq \frac{\tilde{b}(q)}{\tilde{a}^j(q)} - \sum_{k=j+1}^n \frac{\tilde{a}^k(q)}{\tilde{a}^j(q)} x_k, \quad \forall q \in \mathcal{H}_-(j).$$

Therefore x_j has an upper bound and a lower bound if the variables $x_{j+1}, \ldots, x_n$ are bounded. When $j = n$,

$$x_n \geq \sup\{\frac{\tilde{b}(p)}{\tilde{a}^j(p)} : p \in \mathcal{H}_+(n)\}$$

$$x_n \leq \inf\{\frac{\tilde{b}(q)}{\tilde{a}^j(q)} : q \in \mathcal{H}_-(n)\}.$$

Therefore variable x_n has a lower bound and an upper bound. Then it follows from a simple recursive argument that variables $x_{n-1}, \ldots, x_1$ are bounded and the feasible region of (SILP) is bounded.

Now assume the feasible region of (SILP) is bounded. Then there cannot exists a conic index set nor a lineality index set. Then by Corollary A.11 there cannot be a dirty variable, i.e. the case where $\mathcal{H}_+(j)$ or $\mathcal{H}_-(j)$ is empty, but not both empty. By Corollary A.14 there is never a variable j with both $\mathcal{H}_+(j)$ and $\mathcal{H}_-(j)$ empty. Then at each iteration of step 2b of the Fourier-Motzkin elimination procedure, both $\mathcal{H}_+(j)$ and $\mathcal{H}_-(j)$ are not empty. $\square$

The results used in the proofs of Theorem A.1, Theorem A.2, and Theorem A.3 are in Section A.1.2. Basic definitions used in these theorems are in Section A.1.1.

A.1.1 Basic Definitions

Definition A.4. An index set $J_C = \{k_1, k_2, \ldots, k_m\} \subseteq \{1, \ldots, n\}$ is a *conic index set* if and only if there exist nonzero $\alpha_{k_1}, \alpha_{k_2}, \cdots \alpha_{k_m}$ such that for every feasible solution $(\overline{x}_1, \ldots, \overline{x}_n)$ to (SILP), the vector $(\hat{x}_1, \ldots, \hat{x}_n)$ defined by

$$\hat{x}_k = \overline{x}_k, \quad \forall k \notin J_C, \qquad \hat{x}_k = \overline{x}_k + r\alpha_k, \quad \forall k \in J_C$$

is feasible for all $r > 0$, but the vector $(\tilde{x}_1, \ldots, \tilde{x}_n)$ defined by

$$\tilde{x}_k = \overline{x}_k, \quad \forall k \notin J_C, \qquad \tilde{x}_k = \overline{x}_k - r\alpha_k, \quad \forall k \in J_C.$$

is infeasible for a sufficiently large $r > 0$.

Remark A.5. If $J_C = \{k_1, k_2, \ldots, k_m\} \subseteq \{1, \ldots, n\}$ is a *conic index set* in Definition A.4 then $y = (y_1, \ldots, y_n)$ defined by

$$y_k = 0, \quad \forall k \notin J_C, \qquad y_k = \alpha_k, \quad \forall k \in J_C$$

is an element of $\text{rec}(\Gamma)$ since for any feasible $\overline{x}$, $\overline{x} + ry \in \Gamma$ for all $r > 0$. However, for sufficiently large r, $\overline{x} - ry \notin \Gamma$ so $y \notin \text{lin}(\Gamma)$. Likewise each element in $\text{rec}(\Gamma) \setminus \text{lin}(\Gamma)$ corresponds to a conic index set. $\triangleleft$

Definition A.6. An index set $J_L = \{k_1, k_2, \ldots, k_m\} \subseteq \{1, \ldots, n\}$ is a *lineality index set* if and only if there exist nonzero $\alpha_{k_1}, \alpha_{k_2}, \cdots \alpha_{k_m}$ such that for every feasible solution $(\overline{x}_1, \ldots, \overline{x}_n)$ to (SILP), the vector $(\hat{x}_1, \ldots, \hat{x}_n)$ defined by

$$\hat{x}_k = \overline{x}_k, \quad \forall k \notin J_L, \qquad \hat{x}_k = \overline{x}_k + r\alpha_k, \quad \forall k \in J_L$$

is feasible for all $r > 0$, and the vector $(\tilde{x}_1, \ldots, \tilde{x}_n)$ defined by

$$\tilde{x}_k = \overline{x}_k, \quad \forall k \notin J_L, \qquad \tilde{x}_k = \overline{x}_k - r\alpha_k, \quad \forall k \in J_L.$$

is also feasible for all $r > 0$.

A.1.2 Clean Systems are Permutation Independent

In this Section we assume that $(SILP)$ is feasible. Also assume that the FM procedure has eliminated variables $x_1, \ldots, x_\ell$ and the system of inequalities describing $P(\Gamma; x_1, \ldots, x_\ell)$ is

$$\tilde{a}^{\ell+1}(i)x_{\ell+1} + \tilde{a}^{\ell+2}(i)x_{\ell+2} + \cdots + \tilde{a}^n x_n \geq \tilde{b}(i), \quad i \in \tilde{I}. \tag{A.1}$$

We use the notation $J_C(\ell+1) \subseteq \{\ell+1, \ldots, n\}$ to denote a conic index set with respect to the system (A.1).

Remark A.7. If $(\overline{x}_1, \ldots, \overline{x}_n)$ is a feasible solution to (SILP), then by Theorem 2.2, $(\overline{x}_{\ell+1}, \ldots, \overline{x}_n)$ is a feasible solution to (A.1). Then by definition of conic index set, for all $r > 0$, $(\hat{x}_{\ell+1}, \ldots, \hat{x}_n)$ is also feasible to (A.1) where

$$\hat{x}_k = \overline{x}_k, \quad \forall k \notin J_C(\ell+1) \text{ and } k > \ell, \qquad \hat{x}_k = \overline{x}_k + r\alpha_k, \quad \forall k \in J_C(\ell+1).$$

Since $(\hat{x}_{\ell+1}, \ldots, \hat{x}_n)$ is also feasible to (A.1) for all $r > 0$ it follows that

$$\sum_{k \in J_C(\ell+1)} r\alpha_k \tilde{a}^k(i) = r \sum_{k \in J_C(\ell+1)} \alpha_k \tilde{a}^k(i) \geq 0, \quad i \in \tilde{I}.$$

This implies

$$\sum_{k \in J_C(\ell+1)} \alpha_k \tilde{a}^k(i) \geq 0, \quad i \in \tilde{I}. \tag{A.2}$$

Lemma A.8. (Conic Index Set Extension) Assume that the Fourier-Motzkin procedure has eliminated variables $x_1, \ldots, x_\ell$ producing the system (A.1) that describes $P(\Gamma; x_1, \ldots, x_\ell)$. If $J_C(\ell+1) \subseteq \{\ell+1, \ldots, n\}$ is a conic index set of $P(\Gamma; x_1, \ldots, x_\ell)$, then there is a conic index set $J_C(\ell)$ of $P(\Gamma; x_1, \ldots, x_{\ell-1})$ such that $J_C(\ell) = J_C(\ell+1) \cup \{\ell\}$ or $J_C(\ell) = J_C(\ell+1)$.

Proof. By hypothesis, variable ℓ can be eliminated so $\mathcal{H}_+(\ell)$ is not empty and $\mathcal{H}_-(\ell)$ is not empty (if both $\mathcal{H}_+(\ell)$ and $\mathcal{H}_-(\ell)$ are empty we have a zero column and it follows immediately that $J_C(\ell) = J_C(\ell+1)$ is a conic index set for $P(\Gamma; x_1, \ldots, x_{\ell-1})$). Assume prior to elimination variable x_ℓ the system is

$$a^\ell(i)x_\ell + a^{\ell+1}(i)x_{\ell+1} + \cdots + a^n x_n \geq b(i), \quad i \in I. \tag{A.3}$$

Now show there is a well-defined α_ℓ so we extend the conic index set $J_C(\ell+1)$ to include variable x_ℓ. When projecting out variable x_ℓ the $\tilde{a}(i)$ and $\tilde{b}(i)$ in (A.1) are generated from the $a(i)$ and $b(i)$ in (A.3). For the feasible $(\overline{x}_1, \ldots, \overline{x}_n)$,

$$\sum_{k=\ell+1}^{n} a^{\ell+1}(i)\overline{x}_k \;\geq\; b(i), \quad i \in \mathcal{H}_0(\ell). \tag{A.4}$$

$$\frac{b(p)}{a^\ell(p)} - \sum_{k=\ell+1}^{n} \frac{a^k(p)}{a^\ell(p)}\overline{x}_k \;\leq\; \frac{\tilde{b}(q)}{a^\ell(q)} - \sum_{k=\ell+1}^{n} \frac{a^k(q)}{a^\ell(q)}\overline{x}_k,$$
$$\forall p \in \mathcal{H}_+(\ell), \; \forall q \in \mathcal{H}_-(\ell) \tag{A.5}$$

and

$$\overline{x}_\ell \geq \frac{b(p)}{a^\ell(p)} - \sum_{k=\ell+1}^{n} \frac{a^k(p)}{a^\ell(p)} \overline{x}_k, \quad \forall p \in \mathcal{H}_+(\ell) \tag{A.6}$$

$$\overline{x}_\ell \leq \frac{b(q)}{a^\ell(q)} - \sum_{k=\ell+1}^{n} \frac{a^k(q)}{a^\ell(q)} \overline{x}_k, \quad \forall q \in \mathcal{H}_-(\ell). \tag{A.7}$$

If there exists an α_ℓ that satisfies

$$\alpha_\ell \geq - \sum_{k \in J_C(\ell+1)} \alpha_k \frac{a^k(p)}{a^\ell(p)}, \quad \forall p \in \mathcal{H}_+(\ell) \tag{A.8}$$

$$\alpha_\ell \leq - \sum_{k \in J_C(\ell+1)} \alpha_k \frac{a^k(q)}{a^\ell(q)}, \quad \forall q \in \mathcal{H}_-(\ell) \tag{A.9}$$

then $r > 0$ gives

$$r\alpha_\ell \geq -r \sum_{k \in J_C(\ell+1)} \alpha_k \frac{a^k(p)}{a^\ell(p)}, \quad \forall p \in \mathcal{H}_+(\ell) \tag{A.10}$$

$$r\alpha_\ell \leq -r \sum_{k \in J_C(\ell+1)} \alpha_k \frac{a^k(q)}{a^\ell(q)}, \quad \forall q \in \mathcal{H}_-(\ell). \tag{A.11}$$

Claim: The system is (A.8)-(A.9) is consistent. Multiply (A.9) by -1 and apply Fourier-Motzkin elimination. This yields (A.2) and the fact that (A.2) is nonnegative implies (A.8)-(A.9) is consistent. †

Combining (A.10)-(A.11) with (A.6)-(A.7)

$$\overline{x}_\ell + r\alpha_\ell \geq \frac{b(p)}{a^\ell(p)} - \sum_{k=\ell+1}^{n} \frac{a^k(p)}{a^\ell(p)} \overline{x}_k - r \sum_{k \in J_C(\ell+1)} \alpha_k \frac{a^k(p)}{a^\ell(p)}, \forall p \in \mathcal{H}_+(\ell) \tag{A.12}$$

$$\overline{x}_\ell + r\alpha_\ell \leq \frac{b(q)}{a^\ell(q)} - \sum_{k=\ell+1}^{n} \frac{a^k(q)}{a^\ell(q)} \overline{x}_k - r \sum_{k \in J_C(\ell+1)} \alpha_k \frac{a^t(q)}{a^\ell(q)}, \forall q \in \mathcal{H}_-(\ell) \tag{A.13}$$

If there is an $\alpha_\ell = 0$ that is a solution to (A.8)-(A.9) set $J_C(\ell) = J_C(\ell+1)$. Otherwise, if all solutions to (A.8)-(A.9) are nonzero, pick a nonzero α_ℓ and set $J_C(\ell) = \{\ell\} \cup J_C(\ell+1)$. In either case, $J_C(\ell)$ is a conic index set for $P(\Gamma; x_1, \ldots, x_{\ell-1})$.

$\square$

Example A.9 (Example Illustrating Lemma A.8)**.** Consider the system

$$-\frac{2}{3}x_1 - x_2 \geq b_1$$
$$-\frac{1}{2}x_1 - x_2 \geq b_2$$
$$-x_1 - x_2 \geq b_3$$
$$x_1 + 3x_2 \geq b_4$$

In reference back to Lemma A.8, $\ell = 1$. Project out x_1 and get

$$\frac{3}{2}x_2 \geq \frac{3}{2}b_1 + b_4$$
$$x_2 \geq 2b_2 + b_4$$
$$2x_2 \geq b_3 + b_4.$$

Observe $J_C(2) = \{2\}$ and the inequalities corresponding to (A.8)-(A.9) are

$$
\begin{aligned}
\alpha_1 &\leq -\frac{3}{2} \\
\alpha_1 &\leq -2 \\
\alpha_1 &\leq -1 \\
\alpha_1 &\geq -3.
\end{aligned}
$$

Then $\{1,2\}$ is a conic index set and if $(\bar{x}_1, \bar{x}_2)$ is feasible then $(\bar{x}_1 + \alpha_1 r, \bar{x}_2 + r)$ is feasible for all $r > 0$ when $-3 \leq \alpha_1 \leq -2$ and $\alpha_2 = 1$.   ◁

Proposition A.10 (Conic Index Set Extension). Assume that $(SILP)$ is feasible and the Fourier-Motzkin procedure has eliminated variables $x_1, \ldots, x_\ell$ and the system of inequalities describing $P(\Gamma; x_1, \ldots, x_\ell)$ is (A.1). If $J_C(\ell + 1) \subseteq \{\ell + 1, \ldots, n\}$ is a conic index set of $P(\Gamma; x_1, \ldots, x_\ell)$, then there is a conic index set J_C of (SILP) such that $J_C(\ell + 1) \subseteq J_C$.

Proof. Use Lemma A.10 and generate conic index set $J_C(\ell)$ on $P(\Gamma; x_1, \ldots, x_{\ell-1})$ from conic index set $J_C(\ell + 1)$ on $P(\Gamma; x_1, \ldots, x_\ell)$. Repeat this process using Lemma A.10 and successively generate conic index set $J_C(k)$ from $J_C(k + 1)$ for $k = 1, \ldots, \ell$, terminating with $J_C = J_C(1)$ a conic index set of (SILP).   □

Corollary A.11. (Dirty Variable in Conic Index Set) Assume that the Fourier-Motzkin procedure has eliminated variables $x_1, \ldots, x_\ell$ producing the system (A.1) that describes $P(\Gamma; x_1, \ldots, x_\ell)$. If variable x_t for $t > \ell$ is dirty, then there is a conic index set J_C for (SILP) and $t \in J_C$.

Proof. If variable x_t is dirty, set $J_C(\ell + 1) = \{t\}$ and observe that $J_C(\ell + 1)$ is a conic index set for the projected space $P(\Gamma; x_1, \ldots, x_\ell)$. Then by Proposition A.10 there is a conic index set J_C for (SILP) with $t \in J_C$.   □

In what follows Lemma A.12 replicates Lemma A.8 for lineality index sets instead of conic index sets. Proposition A.13 replicates Proposition A.10 for lineality index sets instead of conic index sets. Corollary A.14 replicates Corollary A.11 for a variable in the projected system with all zero coefficients instead of all nonnegative or all nonpositive coefficients.

Lemma A.12. (Lineality Index Set Extension) Assume that the Fourier-Motzkin procedure has eliminated variables $x_1, \ldots, x_\ell$ producing the system (A.1) that describes $P(\Gamma; x_1, \ldots, x_\ell)$. If $J_L(\ell + 1) \subseteq \{\ell + 1, \ldots, n\}$ is a lineality index set of $P(\Gamma; x_1, \ldots, x_\ell)$, then there is a lineality index set $J_L(\ell)$ of $P(\Gamma; x_1, \ldots, x_{\ell-1})$ such that $J_L(\ell) = J_L(\ell + 1) \cup \{\ell\}$ or $J_L(\ell) = J_L(\ell + 1)$.

Proof. Observe that in case of a lineality variable, instead of a conic variable, the system (A.2)

$$
\sum_{k \in J_C(\ell+1)} \alpha_k \tilde{a}^k(i) \geq 0, \quad i \in \tilde{I}
$$

used in the proof of Lemma A.8 becomes

$$
\sum_{k \in J_L(\ell+1)} \alpha_k \tilde{a}^k(i) = 0, \quad i \in \tilde{I}
$$

since for a lineality variable r is both positive and negative. The implication of equality is that when replicating the proof of Lemma A.8 we can multiply (A.8)-(A.9) by positive and negative r and still guarantee the existence of an α_ℓ solution. Since multiplying by both r and $-r$ is valid when calculating α_ℓ it follows that $J_L(\ell) = J_L(\ell + 1)$ is a lineality index set.   □

35

Proposition A.13. (Lineality Index Set Extension) Assume that the Fourier-Motzkin procedure has eliminated variables $x_1, \ldots, x_\ell$ producing the system (A.1) that describes $P(\Gamma; x_1, \ldots, x_\ell)$. If $J_L(\ell + 1) \subseteq \{\ell + 1, \ldots, n\}$ is a lineality index set of $P(\Gamma; x_1, \ldots, x_\ell)$, then there is a lineality index set J_L of (SILP) such that $J_L(\ell + 1) \subseteq J_L$.

Proof. Replicate the proof of Proposition A.10. $\qquad\square$

Corollary A.14. (A Zero Variable is in a Lineality Index Set) Assume that the Fourier-Motzkin procedure has eliminated variables $x_1, \ldots, x_\ell$ producing the system (A.1) that describes $P(\Gamma; x_1, \ldots, x_\ell)$. If x_t for $t > \ell$ is a zero variable, i.e. $\tilde{a}^t(i) = 0$ for all $i \in \tilde{I}$, then there is a lineality index set J_L for (SILP) and $t \in J_L$.

Proof. Replicate the proof of Corollary A.11. $\qquad\square$

The following lemma is critical in proving our main result. The basic idea is that if there is an index set J that "behaves" like a conic index set, but projection of the variables $x_1, \ldots, x_\ell$ results in the remaining variables in the index set having zero coefficients in the projected system, then J must actually be a lineality index set. The proof relies heavily on the ideas used in the proof of Lemma A.12.

Lemma A.15. (Lineality Implication) Assume that the Fourier-Motzkin procedure has eliminated variables $x_1, \ldots, x_\ell$ producing the system (A.1) that describes $P(\Gamma; x_1, \ldots, x_\ell)$. Further assume $J = \{k_1, \ldots, k_m\} \subseteq \{1, \ldots, n\}$ is an index set with associated nonzero $\alpha_{k_1}, \ldots, \alpha_{k_m}$ such that

1. the set $J \cap \{\ell + 1, \ldots, n\}$ is not empty and $J \cap \{\ell + 1, \ldots, n\}$ is a lineality index set for $P(\Gamma; x_1, \ldots, x_\ell)$ based on the nonzero α_{k_m} where $k_m > \ell$, and

2. for every feasible solution $(\overline{x}_1, \ldots, \overline{x}_n)$ to (SILP), $(\hat{x}_1, \ldots, \hat{x}_n)$ where

$$\hat{x}_k = \overline{x}_k, \quad \forall k \notin J, \qquad \hat{x}_k = \overline{x}_k + r\alpha_k, \quad \forall k \in J$$

   is feasible for all $r > 0$.

Then J is a lineality index set with associated nonzero $\alpha_{k_1}, \ldots, \alpha_{k_m}$.

Proof. The essence of the proof is to follow the proof of Proposition A.13 and start with the lineality index set for $P(\Gamma; x_1, \ldots, x_\ell)$ and recurse back and construct the entire set J. However, for this to work, it is necessary to construct the given $\alpha_{k_1}, \ldots, \alpha_{k_m}$ in the hypothesis. By hypothesis we know $J \cap \{\ell + 1, \ldots, n\}$ is a lineality index set for $P(\Gamma; x_1, \ldots, x_\ell)$ based on the nonzero α_{k_m} where $k_m > \ell$. Hence we need to match the α_{k_m} for all $k_m \leq \ell$. Consider an arbitrary $k_h \in J$ with $k_h \leq \ell$. Assume at this point we have a match for the lineality set $J(k_h + 1)$. It suffices to show

$$\alpha_{k_h} \quad \geq \quad - \sum_{k \in J(k_h+1)} \alpha_k \frac{a^k(p)}{a^{k_h}(p)}, \quad \forall p \in \mathcal{H}_+(k_h) \tag{A.14}$$

$$\alpha_{k_h} \quad \leq \quad - \sum_{k \in J(k_h+1)} \alpha_k \frac{a^k(q)}{a^{k_h}(q)}, \quad \forall q \in \mathcal{H}_-(k_h) \tag{A.15}$$

Assume without loss (A.14) is violated (a similar argument is valid if (A.15) is violated). Then there is a $\hat{p} \in \mathcal{H}_+(k_h)$ such that

$$\alpha_{k_h} \quad < \quad - \sum_{k \in J(k_h+1)} \alpha_k \frac{a^k(\hat{p})}{a^{k_h}(\hat{p})}. \tag{A.16}$$

But

$$\overline{x}_{k_h} + r\alpha_{k_h} \quad \geq \quad \frac{b(p)}{a^{k_h}(p)} - \sum_{k=k_h+1}^{n} \frac{a^k(p)}{a^{k_h}(p)}\overline{x}_k - r \sum_{k \in J(k_h+1)} \alpha_k \frac{a^k(p)}{a^{k_h}(p)} \qquad (A.17)$$

must hold for all $p \in \mathcal{H}_+(\ell)$ and feasible $\hat{x}$ by part 2. of the hypothesis for this lemma. But (A.16) implies for sufficiently large r that (A.17) will be violated. Hence it is possible in the backward recursion to generate the exact α_k. By a similar argument, $j \notin J$ implies that $\alpha_j = 0$ must be in the interval defined by (A.14)-(A.15).

Also, as in the proof of Lemma A.12, if α_{k_h} satisfies (A.14)-(A.15), then α_{k_h} satisfies

$$\alpha_{k_h} \quad \leq \quad - \sum_{k \in J(k_h+1)} \alpha_k \frac{a^k(p)}{a^{k_h}(p)}, \quad \forall p \in \mathcal{H}_+(k_h)$$

$$\alpha_{k_h} \quad \geq \quad - \sum_{k \in J(k_h+1)} \alpha_k \frac{a^k(q)}{a^{k_h}(q)}, \quad \forall q \in \mathcal{H}_-(k_h)$$

since $J(k_h + 1)$ is a lineality set and this implies

$$\sum_{k \in J(k_h+1)} \alpha_k \tilde{a}^k(i) = 0, \quad i \in \tilde{I}$$

and adding variable k_h to $J(k_h + 1)$ results in a new lineality index set.

Finally, given that the recursion began with an index set of lineality variables, this is maintained at each step and the lemma is proved. $\qquad \square$

Proposition A.16. If (SILP) contains a conic index set, then the Fourier-Motzkin elimination procedure will terminate with at least one dirty variable regardless of the variable permutation used in the elimination procedure.

Proof. Assume an arbitrary variable permutation. By hypothesis there is a conic index set $J_C = \{j_1, \ldots, j_m\}$. The order of the index set is irrelevant, so assume without loss that for the given variable permutation $j_1 < j_2 < \cdots < j_m$. Apply Fourier-Motzkin elimination and attempt to eliminate the variables (reindexed to reflect the selected variable permutation) $1, \ldots, \ell$ where $\ell = j_{m-1}$. If a dirty variable is discovered prior to eliminating variable x_ℓ, we are done, there exists a dirty variable. Therefore, we can assume variables $1, \ldots, \ell$ where $\ell = j_{m-1}$ are eliminated and the projected system is

$$\tilde{a}^{\ell+1}(i)x_{\ell+1} + \tilde{a}^{\ell+2}(i)x_{\ell+2} + \cdots + \tilde{a}^n x_n \geq \tilde{b}(i), \quad i \in \tilde{I}$$

where $j_m \geq \ell + 1$. There are two cases to consider.

Case 1: Eliminating a variable in the index set $\{1, \ldots, \ell\}$ does not result in j_m indexing a zero column in the projected system. We show variable j_m is dirty in the projected system. This follows because the projected system does not include variables $j_1, \ldots, j_{m-1}$. Variable j_m indexes the only variable in the conic index set J_C that remains in the projected system and the not all $\tilde{a}^{j_m}(i)$ are zero. Remark A.7 implies $r\alpha_{j_m}\tilde{a}^{j_m}(i) \geq 0$ for all $i \in \tilde{I}$. If $\alpha_{j_m} > 0$ then $\tilde{a}^{j_m}(i) \geq 0$ for all $i \in \tilde{I}$. If $\alpha_{j_m} < 0$ then $\tilde{a}^{j_m}(i) \leq 0$ for all $i \in \tilde{I}$. In either case, column j_m is dirty since not all coefficients are equal to zero.

Case 2: Eliminating a variable in the index set $\{1, \ldots, \ell\}$ does result in j_m indexing a zero column in the projected system. Then set $J \cap \{\ell + 1, \ldots, n\} = \{j_m\}$ is a lineality index set in $P(\Gamma; x_1, \ldots, x_\ell)$. Then by Lemma A.15 the index set J_C must be lineality index set which contradicts the hypotheses. Therefore Case 2 cannot occur. $\qquad \square$

Proposition A.17. If there exists a permutation of the variables that results in a clean system when the Fourier-Motzkin procedure is applied to the variables, then every permutation of the variables results in a clean system.

Proof. If the Fourier-Motzkin procedure for some permutation results in a clean system then there are no conic index sets by the contrapositive of Proposition A.16. If there are no conic index sets, then applying the Fourier-Motzkin procedure to any permutation of the variables cannot result in dirty variables since a dirty variable implies the existence of a conic index set by Corollary A.11. □

A.2 Differences between semi-infinite linear programming and finite linear programming

In this section we show how the Fourier-Motzkin elimination procedure can be used to reveal 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 A.18. Consider the following problem

$$\inf x_1$$
$$\frac{1}{i}x_2 \geq 1 \quad \text{for } i = 1, 2, \ldots$$
$$x_1 \geq 0.$$

This problem is infeasible since for any x_2, there exists a 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.

$$\frac{1}{i}x_2 \geq 1 \quad \text{for } i = 1, 2, \ldots$$
$$z \geq 0,$$

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 attained and the dual is solvable. ◁

Example A.19 (Example 3.8 revisited)**.** In Example 3.8, 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 (3.10). This is because $\omega(\delta)$ still equals $\frac{1}{4(\delta-1)}$ and $\lim_{\delta \to \infty} \omega(\delta) = 0$. Then by Lemma 3.5, 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 3.11 because in this case $I_3 = \emptyset$. ◁

A.3 Application: Conic linear programs

We consider conic programming with the following primal problem:

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

where X is a finite dimensional vector space and Y an arbitrary vector space, $A : X \to Y$ 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

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

In this section of the electronic companion, 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 A.31 (Zero duality gap via an interior point)**.** Let Y be finite-dimensional, and let P be reflexive. Assume the primal conic program (ConLP) is feasible. Suppose there exists a $\psi^* \in \text{int}(P')$ with $A'(\psi^*) = \phi$. 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. We believe our results add fresh insight to the literature on connections between conic programming and semi-infinite linear programming (see, for instance Zhang [23]).

A.3.1 Preliminaries

For the linear map A defined on X, let $\ker(A) = \{x \in X : A(x) = 0\}$ denote the kernel of T. The *algebraic adjoint* $A' : Y' \to X'$ of A, where Y' and X' are the algebraic dual vector spaces of X and Y respectively, is the mapping $A'(\psi) = \psi \circ A$ (the algebraic adjoint has been defined elsewhere, see for instance Chapter 6 of [1]).

The following results demonstrate how, without loss of generality, we may assume A' is surjective when (ConLP) is feasible and bounded.

Lemma A.20. 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'. We 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 $\psi \in Y'$. We now 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 $A(x) = 0$. Note 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, $\langle x, \phi \rangle = \langle x, A'(\psi) \rangle = \langle A(x), \psi \rangle = \langle 0, \psi \rangle = 0$. $\qquad\square$

Remark A.21. Note that a related result to Lemma A.20 for topological adjoints is well-known in the functional analysis literature (see for instance Theorem 2 on page 156 of [17]). This more familiar result requires that $Im(A)$ be a closed set in Y. This requirement does not fit our setting since we assume no topology on Y. In contrast, for algebraic adjoints, no assumption on $Im(A)$ is necessary. Indeed, *any* extension of the linear functional in the forward direction of the above proof suffices, it need not be continuous in a given topology. $\qquad\triangleleft$

Lemma A.22. 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 $\lambda \geq 0$. 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 A.23. 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

$$\inf_{x \in \ker(A)^\perp} \quad \langle x, \bar{\phi} \rangle \tag{A.18}$$
$$\text{s.t.} \quad \bar{A}(x) \succeq_P d.$$

If (ConLP) is feasible and bounded, the optimal value of (ConLP) equals the optimal value of (A.18). Moreover, if O is the set of optimal primal solutions for (ConLP), and $\overline{O}$ is the set of optimal primal solutions for (A.18), then $O = \overline{O} + \ker(A)$.

Proof. Since (ConLP) is feasible and bounded, $\ker(A) \subseteq \ker(\phi)$ by Lemma A.22. 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 (A.18) with the same objective value as $\langle x, \phi \rangle$. □

Remark A.24. By Lemma A.23, when (ConLP) is feasible and bounded, it suffices to consider a restricted optimization problem like (A.18). 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 A.20. ◁

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 the linear map A is surjective. As discussed above (Lemmas A.20–A.23 and Remark A.24) this can be assumed without loss of generality when (ConLP) is feasible and bounded.

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 condition that P is reflexive naturally holds in many important special cases of conic programming. Once such case is when Y is finite dimensional 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.

The primal semi-infinite linear program is

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

where $n = \dim(X)$, and we choose a basis $e^1, \ldots, e^n \in X$ to view X as isomorphic to $\mathbb{R}^n$, 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 $b(\psi) := \langle b, \psi \rangle$. The finite support dual of (ConSILP) is

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

The close connection of this primal-dual pair to the conic pair (ConLP)–(ConLPD) is shown in Theorem A.25 and Theorem A.26 below. Theorem A.25 shows that (ConLP) and (ConSILP) are *equivalent*, meaning their respective feasible sets are isomorphic under an isomorphism which preserves objective values. In particular, this means $v(\text{ConLP}) = v(\text{ConSILP})$. Similarly, Theorem A.26 shows that (ConLPD) and (ConFDSILP) are equivalent. In particular this means, $v(\text{ConLPD}) = v(\text{ConFDSILP})$.

A.3.2 Equivalent semi-infinite linear programming formulations of conic programs

The following two theorems show the equivalence of (ConLP) and (ConLPD) with their semi-infinite programming formulations given in (ConSILP) and (ConFDSILP).

Theorem A.25 (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). We show F is isomorphic to $\hat{F}$ under the basis $e^1, \ldots, e^n$. First we 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 we show that if $(x_1, \ldots, x_n) \in \hat{F}$, then $x = x_1 e^1 + \ldots + x^n e^n \in F$. We establish 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 $(x_1, \ldots, x_n) \notin \hat{F}$. $\qquad \square$

Theorem A.26 (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. $\qquad\square$

A.3.3 Zero duality gap via boundedness

This result is known in the literature (see for instance Shapiro [19]), but we show it as an immediate consequence of Theorem 3.24 based on Fourier-Motzkin elimination techniques.

Theorem A.27 (Zero duality gap via boundedness)**.** If P is reflexive and there exists a scalar γ such the set $\{x : A(x) \succeq_P d$ and $\langle x, \phi \rangle \leq \gamma\}$ is nonempty and bounded, then there is no duality gap for the primal-dual pair (ConLP)-(ConLPD).

Remark A.28. The above result shows that semi-definite programs (SDPs) and copositive programs with nonempty, bounded feasible regions have zero duality gap.

A.3.4 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 A.29 (Regular duality for conic programs)**.** 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 A.25, 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 A.26, every feasible sequence for (ConLPD) maps to a feasible sequence for (ConFDSILP). Similarly, every feasible sequence for (ConFDSILP) maps to a feasible sequence for (ConLPD). Thus, the limit value $\hat{d}$ of (ConLPD) is equal z^*, the limit value of (ConFDSILP). $\qquad\square$

A.3.5 Zero duality gap via an interior point condition

The main result of this section demonstrates how the Fourier-Motzkin elimination procedure can be used to establish a "Slater-like" theorem for conic programs. The result is well known. Alternate proofs can be found in the conic programming literature (see for instance Chapter 4 of [8]). The novelty here is the new proof using projection techniques.

For this section, we impose the condition that Y is also finite-dimensional (along with X). As in the discussion after the definition of (ConSILP), we identify X and 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 A.30. Let Y be finite-dimensional, and let P be reflexive. Assume $A' : Y' \to X'$ is surjective and there exists $\psi^* \in \mathrm{int}(P')$ with $A'(\psi^*) = c$. Then there exists $\epsilon > 0$ and such that for all $\bar{c} \in B(c, \epsilon)$, there exists a $\bar{\psi} \in P'$ such that $\bar{c}^\top x \geq \langle d, \bar{\psi} \rangle$ 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 $\psi^* \in \mathrm{int}(P')$ 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 A.31 (Zero duality gap via an interior point)**.** Let Y be finite-dimensional, and let P be reflexive. If the primal conic program (ConLP) is feasible and there exists $\psi^* \in \mathrm{int}(P')$ with $A'(\psi^*) = \phi$, then there is a zero duality gap for the primal dual pair (ConLP)-(ConLPD). Moreover, the primal is solvable.

Proof. By hypothesis, there exists $\psi^* \in \mathrm{int}(P')$ 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 A.24, it is valid to assume A' is surjective.

Claim A.32. The variables $x_1, \ldots, x_n$ remain clean when Fourier-Motzkin elimination is applied to (ConSILP).

Proof of Claim. 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 A.30, there exists $\epsilon > 0$ such that every $\bar{c} \in B(c, \epsilon)$ gives a constraint $\bar{c}^\top x \geq b'$ in (ConSILP) where $b' = \langle d, \bar{\psi} \rangle$ with $A'(\bar{\psi}) = \bar{c}$. 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 $\langle d, \psi_+^j \rangle$ and $\langle d, \psi_-^j \rangle$ respectively with $A'(\psi_+^j) = c + \delta e^j$ and $A'(\psi_-^j) = c - \delta e^j$).

Case 1: $c_j = 0$ for all $j = 1, \ldots, n$. 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 $j \in \{1, \ldots, n\}$. Relabel the variables such that $j = 1$ and $c_1 \neq 0$. Note that coefficient of x_1 in $-c^\top x + z \geq 0$, has opposite sign to the coefficient of x_1 in each pair of constraints $(c + \frac{\epsilon}{2} e^k)^\top x \geq b^k j_+$ and $(c - \frac{\epsilon}{2} e^k)^\top x \geq b_-^k$ for $j = 2, \ldots, n$. Clearly x_1 is clean, and when x_1 is eliminated the constraint $-c^\top x + z \geq 0$ is aggregated with the constraints $(c + \frac{\epsilon}{2} e^k)^\top x \geq b_+^k$ and $(c - \frac{\epsilon}{2} e^k)^\top x \geq b_-^k$, for each $k = 2, \ldots, n$. This leaves the constraints $\frac{\epsilon}{2} x_k + z \geq b_+^k$ and $-\frac{\epsilon}{2} x_j + z \geq b_-^k$ in the system for $k = 2, \ldots, n$, after x_1 is eliminated. As in Case 1, these constraints remain in the system variable until x_k 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 3.23, there is a zero duality gap between the pair (ConSILP)-(ConFDSILP), and (ConSILP) is solvable. This implies that there is zero duality gap for the pair (ConLP)-(ConLPD), and the primal (ConLP) is solvable. □

Remark A.33. 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 A.31, 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 in this case. ◁

A.4 Convex programs

Theorem A.34. $v(\text{LD}) = v(\text{CP-SILP})$. 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 we 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 we 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 A.35. $v(\text{CP}) = v(\text{CP-FDSILP})$.

Proof. First we show $v(\text{CP}) \geq v(\text{CP-FDSILP})$. If (5.2)-(5.4) is infeasible, then $v(\text{CP-FDSILP}) = -\infty$ and the result is immediate. Assume (5.2)-(5.4) 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. Note that $\bar{x}$ is feasible to (CP). First, since Ω is convex, by (5.2) $\bar{x} \in \Omega$. By (5.3), $-\sum_{x \in \Omega} \bar{u}(x)g_i(x) + \bar{v}_i = 0$ 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 (5.2) 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 it follows that $f(\bar{x}) = f(\sum_{x \in \Omega} x\bar{u}(x)) \geq \sum_{x \in \Omega} \bar{u}(x)f(x)$ and $\sum_{x \in \Omega} \bar{u}(x)f(x)$ is the objective value of $(\bar{u}, \bar{v})$ in (5.1). This implies $v(\text{CP}) \geq v(\text{CP-FDSILP})$.

Now we 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 (5.1). The objective value of $(\bar{u}, \bar{v})$ in (5.1) is $f(\bar{x})$ which is the objective value $\bar{x}$ in (CP). □

A.5 Additional sufficient conditions for zero duality gap

By looking at the recession cone of (3.14) 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 [14] follow directly and easily from our methods. The recession cone of (3.14) is defined by the system

$$-c_1 x_1 - c_2 x_2 - \cdots - c_n x_n \quad \geq \quad 0 \tag{A.19}$$
$$a^1(i)x_1 + a^2(i)x_2 + \cdots + a^n(i)x_n \quad \geq \quad 0 \quad \text{for } i \in I. \tag{A.20}$$

Applying Fourier-Motzkin elimination to (A.19)-(A.20) gives

$$0 \quad \geq \quad 0 \quad \text{for } h \in H_1 \tag{A.21}$$
$$\tilde{a}^\ell(h)x_\ell + \tilde{a}^{\ell+1}(h)x_{\ell+1} + \cdots + \tilde{a}^n(h)x_n \quad \geq \quad 0 \quad \text{for } h \in H_2. \tag{A.22}$$

Following the notation of Karney [14], K denotes the recession cone of (SILP), given by the inequalities (A.20) and N denotes the null space of the objective function vector c.

Lemma A.36. If H_2 is nonempty in (A.22), then there exists a ray $r \in \mathbb{R}^n$ satisfying (A.19)-(A.20) with at least one of the inequalities in (A.19)-(A.20) 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 (A.21)-(A.22) satisfies at least one of the inequalities in (A.21)-(A.22) strictly. Since this is the projection of some r satisfying (A.19)-(A.20), this r must satisfy at least one inequality in (A.19)-(A.20) strictly, since all inequalities in (A.21)-(A.22) are conic combinations of inequalities in (A.19)-(A.20). $\qquad\square$

Theorem A.37. If (SILP) is feasible and $K \cap N$ is a subspace, then $v(\text{SILP}) = v(\text{FDSILP})$.

Proof. Case 1: H_2 in (A.22) is empty. Observe that the columns in systems (A.19)-(A.20) 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 (A.22) is empty, (SILP) is tidy. Then by Theorem 3.23, $v(\text{SILP}) = v(\text{FDSILP})$.

Case 2: H_2 in (A.22) is not empty. If H_2 is not empty, by Lemma A.36, there exists a r satisfying (A.19)-(A.20) such that at least one of the inequalities in (A.19)-(A.20) 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 (A.19) 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 (A.20) at equality and this contradicts the fact established for this case that at least one inequality in (A.19)-(A.20) is strict. $\qquad\square$

A.5.1 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 A.38. If (SILP) is feasible, then $v(\text{FDSILP}) = \sup\{v(J) : J \text{ is a finite subset of } I\}$.

Proof. We first show that $v(\text{FDSILP}) \leq \sup\{v(J) : J \text{ is a finite subset of } I\}$. By hypothesis, (SILP) is feasible and this implies by Corollary 3.16 that $v(\text{FDSILP}) = \sup_{h \in I_3} \tilde{b}(h)$. Therefore, for every $\epsilon > 0$, there exists a $h^* \in I_3$ such that $v(\text{FDSILP}) - \epsilon \leq \tilde{b}(h^*)$. By Lemma 3.10(iv), there exists a $v^{h^*} \in \mathbb{R}^{(I)}$ 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 we 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 3.25, there exists a $v^* \in \mathbb{R}^{J^*}$ 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, $v(\text{FDSILP}) \geq v(J^*)$. $\qquad\square$

Theorem A.38 is used to prove a series of results by Karney [14]. 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 A.39. 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, $v(P_n) \leq \sup\{v(J) : J$ is a finite subset of $I\} = v(\text{FDSILP}) < \infty$ where the equality follows from Theorem A.38 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 an $n^* \in \mathbb{N}$ such that $J^* \subseteq \{1, \ldots, n^*\}$ and this implies $v(P_{n^*}) \geq v(J^*)$. Thus, $\lim_{n \to \infty} v(P_n) \geq \sup\{v(J) : J$ is a finite subset of $I\} = v(\text{FDSILP})$ where the equality follows from Theorem A.38. $\qquad\square$

Corollary A.40 (Karney [14] 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 3.24 and Corollary A.39. $\qquad\square$

Corollary A.41 (Karney [14] Theorem 2.4). If (SILP) with $I = \mathbb{N}$ is feasible and the zero vector is the unique solution to the system (A.19)-(A.20), then $\lim_{n \to \infty} v(P_n) = v(\text{SILP})$.

Proof. If the zero vector is the unique solution to the system (A.19)-(A.20), then the recession cone of (3.14) is $\{0\}$ and (3.14) is bounded for any value of γ such that (3.14) is feasible (such a γ exists because (SILP) is feasible). The result then follows from Theorem 3.24 and Corollary A.39. $\qquad\square$

Corollary A.42 (Karney [14] Theorem 2.5). Assume (SILP) with $I = \mathbb{N}$ is feasible and let r be a ray satisfying (A.19)-(A.20). 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 A.39. $\qquad\square$

Corollary A.43 (Karney [14] 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 A.37 and Corollary A.39. $\qquad\square$