

A geometric simplex method in infinite-dimensional spaces

Robert L. Smith* Christopher Thomas Ryan†

March 6, 2026

Abstract

We expand the basic geometric elements of the simplex method to linear programs in locally convex topological vector spaces and provide conditions under which the method converges in value to optimality. This setting generalizes many previous investigations of the simplex method, which are restricted to Hilbert spaces or otherwise specially structured instances. Our generality is obtained by avoiding the “algebraic” machinery of pivoting via column operations, which has required stronger topological conditions in establishing a connection between basic feasible solutions and extreme point structure. We show that our definition of polytopes captures optimization over the Hilbert cube, a quintessential object in infinite-dimensional spaces known for its surprisingly complicated properties. Moreover, all polytopes (under our definition) have exposed extreme points connected by edge paths.

1 Introduction

Infinite-dimensional linear programming is both classical and cutting-edge. Some of the very first linear programs studied were stated over infinite-dimensional spaces [27, 33], and there has been recent growing interest in several applied areas of mathematics that concern infinite-dimensional linear programs [12, 15, 22, 24, 25, 29, 30, 38].

Since at least [4], researchers have been interested in extensions of the simplex method to understand the structure and optimality of linear programs in topological vector spaces. Anderson and Nash [4] were interested in generalizing the *algebraic structure* underlying the simplex method, including notions of column bases, basic feasible solutions, and reduced costs. This program was championed over the years by several research groups (for example, [5, 6, 16, 20, 35, 40–42, 44]), including recent advancements [22, 41]. In this literature, the focus is largely on finding special types of linear programs where these algebraic concepts prove meaningful and amenable to analysis. In general, the dictionary between extreme points and basic feasible solutions that is foundational to the simplex method in finite dimensions (see Chapter 2 of [9]) breaks down in topological vector spaces (as detailed at length in [4]).

This paper pursues a different program. We are interested in generalizing the classical *geometric* understanding of the simplex method: start with an extreme point and iteratively follow the edges to adjacent extreme points until you reach an extreme point from which no adjacent edge direction improves the objective. We do not concern ourselves with finding an underlying algebraic description of extreme points. Despite alleviating ourselves of this concern, this new path is still fraught with complexities. It is not immediately obvious how to (i) define “edges”, (ii) ensure that

*Industrial and Operations Engineering, University of Michigan, E-mail: rlsmith@umich.edu

†UBC Sauder School of Business, University of British Columbia, E-mail: chris.ryan@sauder.ubc.ca

moving along the edges yields another extreme point, and (iii) that the process makes guaranteed progress toward optimality. We tackle these and other issues in this paper.

Our main result is to produce a geometric simplex algorithm with the following properties. Starting with an extreme point, it can identify a set of adjacent extreme points and define edges on which to reach them. We show ([Theorem 4](#)) that this algorithm converges to optimal value (that is, the objective values of successive iterates converge to the optimal value of the program). When a unique optimal solution exists, the iterates of the algorithm converge to the optimal solution ([Corollary 1](#)). Though the goal seems simple, a carefully constructed set of assumptions was required (see [Assumption 1–Assumption 9](#) throughout the paper). These assumptions will be motivated and justified as they are presented and we will not attempt to recapitulate that careful development here. A summarizing discussion of the assumptions can be found in [Section 6](#).

Our algorithm has important limitations. We do not guarantee that each iteration can be executed in finite time, let alone that an optimal solution can be found in finite time. However, it ensures that the simplex method has the necessary geometric structure of extreme points, edges, and improving directions so that the algorithm does not get “caught” in locally optimal solutions or be unable to determine where to go next. The assumptions all drive toward validating the basic mechanism of “moving from extreme point to extreme point while improving the objective” under minimal restrictiveness. Our goal is more conceptual than providing an implementable algorithm. We are interested in the general geometric structure of a compact convex set described by linear inequalities that make it amenable to a simplex method.

Our assumptions are broad enough to capture most of the well-known examples of when a simplex method “works” in the sense we put forward here, including the (corrected)¹ conditions of [\[22\]](#) that define a broad class of countably-infinite linear programs (including many instances of previous investigations). Our approach allows us to move beyond the sequence space setting of [\[22\]](#) and related papers, opening the door to problems defined over non-sequence spaces, including function and measure spaces.

However, we show that a very classical object — the Hilbert cube — fails the conditions of [\[22\]](#) but nonetheless satisfies our assumptions (see [Examples 2 to 8](#)). This demonstrates that our simplex method is broader than that explored in the previous literature. The fact that the Hilbert cube is problematic for previous methods is not too surprising, despite its simplicity, as it is a notoriously challenging example to understand and analyze [\[7\]](#).

Our assumptions provide a novel definition of what we will call a *polytope* for general topological vector spaces: a polytope is an object where a geometric simplex method can be used to optimize a linear functional over it, potentially requiring infinitely many iterations, under our assumptions. Why might there be interest in such a definition? The familiar definitions of a polytope do not extend to general topological spaces as we might hope. In finite-dimensional space, there are two alternative definitions of a polytope: (i) the convex hull of finitely many points or (ii) the bounded intersection of finitely many halfspaces. Both of these conceptions are flawed in infinite-dimensional space. For (i), the convex hull of finitely many points is only finite-dimensional, a trivial restriction in infinite-dimensional spaces. Indeed, several papers have successively executed simplex-like algorithms on infinite-dimensional objects [\[20, 22, 38, 40\]](#). The problem with (ii) is arguably even more severe: in infinite-dimensional space, the intersection of finitely many halfspaces is never bounded.

Other authors have proposed alternative definitions for polytopes, while there is generally no consensus on how they should be defined. In our search of the literature, we found at least twelve

¹We say “corrected” conditions because we found an error in the development of [\[22\]](#). We propose a fix to recover the results of that paper in [Appendix B](#).

alternative definitions of “polytope” in general spaces (and make no claims of exhaustion) [1, 3, 8, 11, 13, 23, 28, 31, 32, 34, 36]. For a comprehensive overview of this topic and related themes see [18, Chapter 6]. Several papers have attempted to reconcile this multitude of definitions, including [7, 14, 17]. [Section 5](#) provides additional discussion and shows that the Hilbert cube, arguably the simplest meaningful feasible region one can imagine, fails to meet the definition of many current notions of a polytope in infinite dimensions but nonetheless satisfies our assumptions.

The rest of the paper is organized as follows. In [Section 2](#), we provide the details of our basic settings and introduce the main notation. There, we introduce our object of study, which we denote by \mathcal{P} , the intersection of halfspaces in a topological vector space. [Section 3](#) elaborates the geometric properties of \mathcal{P} , including extreme points, edges, and support cones. [Section 4](#) presents our geometric simplex method building on these geometric structures. This section contains our main result ([Theorem 4](#)). Finally, in [Section 5](#), we briefly discuss the challenge of defining polytopes in topological vector spaces and how our simplex method approach enters the conversation. [Sections 4](#) and [5](#) both discuss the mysterious example of the Hilbert cube. [Section 6](#) concludes, including some additional discussion of our assumptions. The proofs of propositions and theorems are found in the main text; the proofs of technical lemmas are placed in the appendix.

2 Setup

We start with the basic setup of our problem and introduce all of the preliminary notation. For ease of reference, [Figure 1](#) captures much of the notation of this section in a finite-dimensional setting.

Let X be a real locally convex Hausdorff topological vector space [26, p. 53] with topology τ . Let X^* be the set of continuous linear functionals on X ; that is, functions $\phi : X \rightarrow \mathbb{R}$ that are continuous in the τ topology of X and the usual topology on the reals \mathbb{R} . A proper affine subspace that is maximal (when ordered by inclusion) is a *hyperplane* in X . From [26, p. 3], an affine subspace H is a closed hyperplane if and only if there is some non-zero continuous linear functional $\phi \in X^*$ and scalar $b \in \mathbb{R}$ such that $H = \{x \in X \mid \phi(x) = b\}$. We call $S = \{x \in X \mid \phi(x) \leq b\}$ the *halfspace* associated with hyperplane H . Since $\phi \in X^*$, S is closed (in the topology τ).

Given an index set A , for all $\alpha \in A$ define a hyperplane $H_\alpha \triangleq \{x \in X \mid \phi_\alpha(x) = b_\alpha\}$ where $\phi_\alpha \in X^*$ and $b_\alpha \in \mathbb{R}$. Let $\{S_\alpha \mid \alpha \in A\}$ be the associated collection of halfspaces (called *constraints*) of the form

$$S_\alpha \triangleq \{x \in X \mid \phi_\alpha(x) \leq b_\alpha\}.$$

Our main object of interest is the set

$$\mathcal{P} \triangleq \bigcap_{\alpha \in A} S_\alpha, \tag{1}$$

which denotes the intersection of all halfspaces in the family $\{S_\alpha \mid \alpha \in A\}$.² Clearly, \mathcal{P} is a closed convex set, since each S_α is closed and convex, and these properties are conserved under (arbitrary) intersections. Conversely, every closed convex set in X has the form (1) for some index set A and constraint set $\{S_\alpha \mid \alpha \in A\}$ [2, Corollary 5.83].

We say a constraint S_α is *active* (or *tight*) at a point $p \in \mathcal{P}$ if $p \in H_\alpha$. For a given $p \in \mathcal{P}$, let $A(p)$ denote the set of active indices of constraints at p . The set

$$\mathcal{H}(p) \triangleq \bigcap_{\alpha \in A(p)} H_\alpha = \{x \in X \mid \phi_\alpha(x) = b_\alpha, \alpha \in A(p)\} \tag{2}$$

²We do not call \mathcal{P} a polytope without further assumptions; this definition is more general than what we will call a polytope. Indeed, a disk in \mathbb{R}^2 satisfies this definition and is clearly not a polytope.

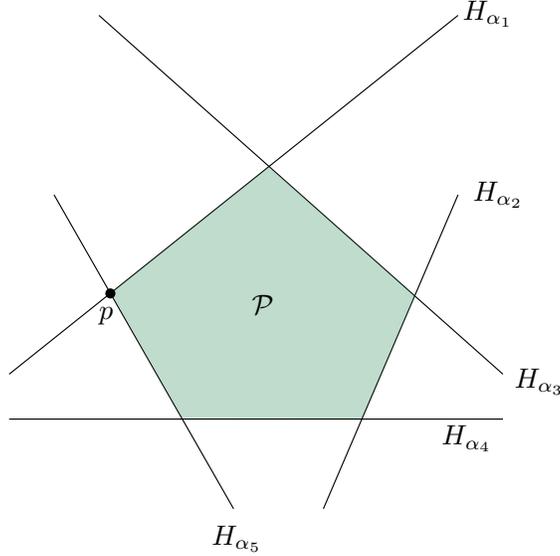


Figure 1: An illustration of the basic notions in the paper. Here $A = \{\alpha_1, \alpha_2, \dots, \alpha_5\}$, $A(p) = \{\alpha_1, \alpha_5\}$, $\mathcal{H}(p) = H_{\alpha_1} \cap H_{\alpha_5} = \{p\}$, and $s_{\alpha_1}(p) = s_{\alpha_5}(p) = 0$ with $s_{\alpha_j}(p) > 0$ for $j = 2, 3, 4$.

defines the intersection of all hyperplanes in the family $\{H_\alpha \mid \alpha \in A\}$ that pass through the point p . Note that $\mathcal{H}(p) = X$ when $A(p) = \emptyset$.

At a point $p \in \mathcal{P}$, for all $\alpha \in A$ we let

$$s_\alpha(p) \triangleq b_\alpha - \phi_\alpha(p) \quad (3)$$

denote the *slack* of constraint S_α at the point p . Constraints that are tight at p have a slack of 0.

To avoid trivialities throughout, we assume \mathcal{P} is not a singleton $\{p\}$.

3 Geometric properties

The geometry of the simplex method is built on the notions of extreme points and edges that connect adjacent extreme points. Readers who are familiar with these objects should take note of how our approach differs from the usual approaches in finite-dimensional spaces (like in [9]) by relying more exclusively on the algebraic structures of the underlying topological vector space X than other treatments.

3.1 Extreme point structure

Recall that an *extreme point* of a set K is an element $x \in K$ where there does not exist $y, z \in K$ such that $x \in (x', x'') \triangleq \{\lambda x' + (1 - \lambda)x'' \mid \lambda \in (0, 1)\}$ and $x' \neq x''$. Let E be the set of all extreme points of \mathcal{P} .

Our first task is to relate the algebraic definition of extreme point to the halfspace structure that defines \mathcal{P} . This relationship is well-understood in finite-dimensional vector spaces X [9, Section 2.2]. The situation in more general X is more complex (see [4, Section 2.3] and [19]). One direction of the relationship is relatively straightforward:

Proposition 1. A point $p \in \mathcal{P}$ is an extreme point of \mathcal{P} if $\mathcal{H}(p) = \{p\}$.

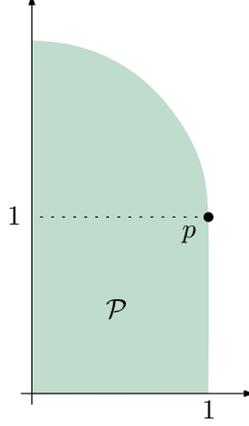


Figure 2: A counter-example to the converse of [Proposition 1](#).

Proof. Suppose $\mathcal{H}(p) = \{p\}$ and, by way of contradiction, that $p = \lambda x' + (1 - \lambda)x''$ for some $x', x'' \in \mathcal{P}$ and $\lambda \in (0, 1)$ with $x' \neq x''$.

Let's derive a contradiction to $x' \neq x''$. We do so by showing that $x', x'' \in \mathcal{H}(p)$ and using the fact that $\mathcal{H}(p)$ is the singleton $\{p\}$ and so $x', x'' \in \mathcal{H}(p)$ implies that $x' = x'' = p$.

It remains to show that $x', x'' \in \mathcal{H}(p)$. In particular, from [\(2\)](#) it suffices to show that $x', x'' \in H_\alpha$ for all $\alpha \in A(p)$. To see this, select an arbitrary $\alpha \in A(p)$. If both $x', x'' \in H_\alpha$ then we are done, and so may assume that one of x' or x'' is not in H_α . This violates the fact that $\alpha \in A(p)$. Indeed, since both x' and x'' are in S_α (due to the fact $x', x'' \in \mathcal{P}$), it must be that either $\phi_\alpha(x') < b_\alpha$ or $\phi_\alpha(x'') < b_\alpha$. But then from the linearity of ϕ_α we have, $\phi_\alpha(p) = \lambda\phi_\alpha(x') + (1 - \lambda)\phi_\alpha(x'') < b_\alpha$. But this contradicts the fact that $\alpha \in A(p)$; i.e., $\phi_\alpha(p) = b_\alpha$.

We conclude that $x', x'' \in H_\alpha$, completing the proof. \square

This leaves open the question of whether p being an extreme point implies $\mathcal{H}(p) = \{p\}$. The following example shows that this need not be the case.

Example 1. Consider the set \mathcal{P} shown in green in [Figure 2](#). Since it is a closed convex set, we know \mathcal{P} is of the form [\(1\)](#). However, notice that although p is an extreme point of \mathcal{P} , we do not have $\mathcal{H}(p) = \{p\}$. Indeed, $\mathcal{H}(p)$ lies along the line segment from $(1, 0)$ to p .³ \blacktriangleleft

To prove the converse, we make a few assumptions.

Assumption 1 (Compactness). The set \mathcal{P} is non-empty and compact (in the topology τ).

Assumption 2 (Strictly positive slacks). $\rho = \inf_{p \in E, \alpha \in A \setminus A(p)} s_\alpha(p) > 0$.

Assumption 3 (Uniformly bounded linear functionals). For all $p \in E$ and all $r \in \mathcal{H}(p) - p$, $\sup_{\alpha \in A \setminus A(p)} \phi_\alpha(r) < \infty$.

Example 2 (Hilbert cube). Consider the following example inspired by the development in [\[22\]](#). Let X be the Hilbert space defined in that paper, namely

$$X = \left\{ x = (x_j)_{j \geq 1} : \sum_{j=1}^{\infty} \delta_j^2 x_j^2 < \infty \right\}, \quad \langle x, y \rangle_X = \sum_{j=1}^{\infty} \delta_j^2 x_j y_j, \quad \|x\|_X^2 = \sum_{j=1}^{\infty} \delta_j^2 x_j^2,$$

³It is worth noting that p is known as a non-exposed extreme point, see [\[43, Chapter 12\]](#).

with weights (δ_j) chosen so that $\delta_j \in (0, 1)$ and $\sum_{j=1}^{\infty} \delta_j^2 < \infty$ (for instance, $\delta_j = \delta^j$ with $0 < \delta < 1$). Define the “Hilbert cube”

$$\mathcal{H} = \prod_{j=1}^{\infty} [0, 1] \subseteq X.^4$$

We may express \mathcal{H} in the form of (1) by setting up the constraints to express $0 \leq x_j \leq 1$ for all $j = 1, 2, \dots$. We can capture this by setting $A = \{1, 2, \dots\} \cup \{-1, -2, \dots\}$ where for $\alpha \in \{1, 2, \dots\}$ we set $\phi_\alpha(x) = \langle x, y^\alpha \rangle$ where $y_j^\alpha = 1$ for $j = \alpha$ and 0 otherwise with $b_\alpha = \delta_\alpha^2$. For $\alpha \in \{-1, -2, \dots\}$, set $\phi_\alpha(x) = \langle x, y^\alpha \rangle$ where $y_j^\alpha = -1$ for $j = \alpha$ and 0 otherwise with $b_\alpha = 0$.

We now show that the Hilbert cube satisfies **Assumptions 1 to 3**. Compactness follows by applying Tychonoff’s theorem, establishing **Assumption 1**.

As for **Assumption 2**, it is easy to see the extreme points of H have the form p such that $p_j = 0$ or 1 for all $j = 1, 2, \dots$. Observe that for all j such that $p_j = 0$, the constraint $-j \in A$ is tight and constraint $j \in A$ is slack with $s_j(p) = 1$. Also, for all j such that $p_j = 1$ we have constraint $j \in A$ is tight and constraint $-j \in A$ is slack with $s_{-j}(p) = 1$. In other words, $\inf_{\alpha \in A \setminus A(p)} s_\alpha(p) = 1$ for all $p \in E$ and so in **Assumption 2**, $\rho = 1 > 0$. This establishes **Assumption 2**.

Assumption 3 is nearly vacuous in this setting. It is easy to see that all extreme points p have $\mathcal{H}(p) = \{p\}$ and so the only possible value for r to consider in **Assumption 3** is $r = 0$. ◀

The example shown in **Figure 2** fails **Assumption 2** at p . Indeed, the constraints that define the “quarter disk” part of the figure about the line $y = 1$ have slacks that diminish to 0 at the point p .

Under these assumptions, we can establish the converse of **Proposition 1**. The argument requires the following two technical lemmas (with proofs relegated to the appendix). A *line* refers to a one-dimensional affine subspace. We refer to 0-dimensional affine subspaces (that is, singletons) as *trivial* affine subspaces.

Lemma 1. Suppose a line intersects a hyperplane but does not lie within that hyperplane. Then the line intersects the hyperplane at a single point.

Lemma 2. If S is a non-trivial affine subspace in X and H is a hyperplane intersecting S at the singleton point p , then S is a line through p .

Proposition 2. If **Assumptions 1 to 3** hold and p is an extreme point of \mathcal{P} then $\mathcal{H}(p) = \{p\}$.

Proof. We prove this by establishing the contrapositive. Suppose $\mathcal{H}(p) \neq \{p\}$ and so there exists a $q \in \mathcal{H}(p)$ with $q \neq p$. Let $r = q - p$. The goal is to show that p is not an extreme point.

The line $\ell(p, r) \triangleq \{p + \lambda r \mid \lambda \in \mathbb{R}\}$ lies in $\mathcal{H}(p)$ since $\mathcal{H}(p)$ is a non-trivial affine subspace containing p and q .

Let $A^+ \subseteq A \setminus A(p)$ be the set of indices of hyperplanes H_α (not in $A(p)$) that intersect $\ell(p)$ at $z_\alpha = \lambda_\alpha r + p$ for $\lambda > 0$. The set A^- is defined similarly but for $\lambda < 0$.

Note that A^+ and A^- may be finite or infinite but both must be nonempty. This follows by **Assumption 1** and [28, p. 736] that states that a compact convex set (namely, \mathcal{P}) in a linear topological space is linearly bounded, i.e. its intersection with any line is bounded. However, if (say) A^+ were empty then the ray $\{p + \lambda r \mid \lambda \geq 0\}$ would lie entirely in \mathcal{P} , contradicting the fact that $\mathcal{P} \cap \ell(p)$ is bounded.

We argue that this allows us to move from p in both directions along $\ell(p, r)$ and still remain in the set \mathcal{P} .

⁴For a similar definition of the Hilbert cube, see [2, Section 3.9].

To do this observe that we have for $\alpha \in A^+ \cup A^-$, $b_\alpha = \phi_\alpha(z_\alpha) = \phi_\alpha(\lambda_\alpha r + p)$ and by linearity of ϕ_α , $\lambda_\alpha = (b_\alpha - \phi_\alpha(p))/\phi_\alpha(r) = s_\alpha(p)/\phi_\alpha(r)$. Hence

$$\begin{aligned} \inf_{\alpha \in A^+} \lambda_\alpha &= \inf_{\alpha \in A^+} \frac{s_\alpha(p)}{\phi_\alpha(r)} \\ &\geq \inf_{\alpha \in A \setminus A(p)} \frac{s_\alpha(p)}{\phi_\alpha(r)} \\ &\geq \inf_{\alpha \in A \setminus A(p)} \frac{s_\alpha(p)}{\sup_{\alpha \in A \setminus A(p)} \phi_\alpha(r)} > 0 \end{aligned}$$

where the second inequality follows from [Assumptions 2](#) and [3](#). Similarly, $\inf_{\alpha \in A^-} -\lambda_\alpha > 0$.

Therefore, p is in the interior of the line segment joining $p^+ = \{\inf_{\alpha \in A^+} \lambda_\alpha\}r + p$ and $p^- = \{-\inf_{\alpha \in A^-} -\lambda_\alpha\}r + p$ which is in \mathcal{P} because p^+ and p^- are in \mathcal{P} as they are in $\mathcal{H}(p)$ as well and thereby satisfy the constraints passing through p and also satisfy inequality constraints not passing through p . Therefore the line segment joining p^+ and p^- is in \mathcal{P} by the convexity of \mathcal{P} . We have p is in the interior of this line segment in \mathcal{P} and so p is not an extreme point of \mathcal{P} . \square

From [Propositions 1](#) and [2](#), we have the following result:

Theorem 1. Suppose [Assumptions 1](#) to [3](#) hold. A point $p \in \mathcal{P}$ is an extreme point of \mathcal{P} if and only if the hyperplanes passing through p intersect in the singleton $\{p\}$, i.e. $\mathcal{H}(p) = \{p\}$.

3.2 Edge structure

In the development that follows we will work with a minimal set of constraints to define \mathcal{P} ; that is, we will assume that the set of constraints $\{S_\alpha \mid \alpha \in A\}$ is such that for every $\alpha' \in A$, $\bigcap_{\alpha \in A \setminus \{\alpha'\}} S_\alpha \not\subseteq \bigcap_{\alpha \in A} S_\alpha$. We call such a set of constraints *non-degenerate*. The following result shows that we can assume this without loss using Zorn's lemma.

Lemma 3 (Non-degeneracy). Let $\{S_\alpha \mid \alpha \in A\}$ be a set of constraints such that $\mathcal{P} = \bigcap_{\alpha \in A} S_\alpha$. There exists a subset of constraints $A^* \subseteq A$ such that $\mathcal{P} = \bigcap_{\alpha \in A^*} S_\alpha$ and the set of constraints $\{S_\alpha \mid \alpha \in A^*\}$ is non-degenerate.

Lemma 4. Suppose [Assumptions 1](#) to [3](#) hold. Then, for all extreme points $p \in E$ and $\alpha \in A(p)$, the set $\ell_\alpha(p) = \bigcap_{\alpha' \in A(p), \alpha' \neq \alpha} H_{\alpha'}$ is a line in X passing through p .

In the remainder of this section, we give a geometric interpretation of the lines $\ell_\alpha(p)$ for $\alpha \in A(p)$ in relation to the set \mathcal{P} . As we shall see, the lines “traces” out an edge $e_\alpha(p)$ of \mathcal{P} between p and another extreme point $q_\alpha(p)$ of \mathcal{P} . We call $q_\alpha(p)$ an *adjacent* extreme point of p . See [Figure 3](#) for an illustration of the idea of the more familiar setting of finite-dimensional space. Let's make all of this more precise.

We call the convex cone $\mathcal{S}(p) \triangleq \bigcap_{\alpha \in A(p)} S_\alpha$ the *support cone* at extreme point $p \in E$ and $\mathcal{R}(p) \triangleq \{d \in X \mid \phi_\alpha(p + \lambda d) \leq b_\alpha \text{ for all } \lambda \geq 0 \text{ and } \alpha \in A(p)\}$ the *recession cone* at extreme point $p \in E$. Note that $\mathcal{S}(p)$ has its apex at p , whereas $\mathcal{R}(p)$ has its apex at the origin 0.

Lemma 5. The recession cone $\mathcal{R}(p) = \{d \in X \mid \phi_\alpha(d) \leq 0 \text{ for all } \alpha \in A(p)\}$.

[Lemma 5](#) tells us that $\mathcal{R}(p) = \mathcal{S}(p) - p$; that is, $\mathcal{R}(p)$ is $\mathcal{S}(p)$ translated to the origin.

Note that every $d \in \mathcal{R}(p)$ with $d \neq 0$ generates a half-line or *ray* r of $\mathcal{R}(p)$, i.e. $r = \{\lambda d, \lambda \geq 0\}$. A ray r is an *extreme ray* of a convex cone C with vertex 0 if $x \in r, y \in C$ and $x - y \in C$ implies $y \in r$ [[39](#), p. 72]. A half-line $r \in X$ is a ray or extreme ray of a convex cone C with vertex p if $r - p$ is a ray or, respectively extreme ray, of $C - p$. Let $\eta_\alpha(p)$ for $\alpha \in A(p)$ be the ray in $\mathcal{R}(p)$ given by $\eta_\alpha \triangleq \ell_\alpha(p) \cap S_\alpha - p$. Note that $\eta_\alpha(p) \neq \{0\}$.

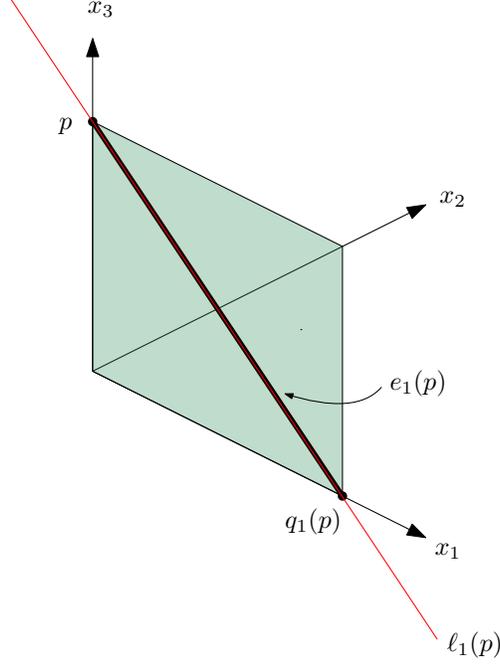


Figure 3: An illustration of extreme points, edges, and adjacent extreme points. Here $S_1 = \{x \in \mathbb{R}^3 \mid -x_1 \leq 0\}$, $S_2 = \{x \in \mathbb{R}^3 \mid -x_2 \leq 0\}$, $S_3 = \{x \in \mathbb{R}^3 \mid -x_3 \leq 0\}$, and $S_4 = \{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 \leq 1\}$.

Lemma 6. If **Assumptions 1 to 3** hold then rays $\eta_\alpha(p), \alpha \in A(p)$ are the extreme rays of $\mathcal{R}(p)$.

Since $\mathcal{R}(p)$ is $\mathcal{S}(p) - p$, we have the following theorem from **Lemma 6**.

Lemma 7. If **Assumptions 1 to 3** hold then rays $r_\alpha(p) = \eta_\alpha(p) + p, \alpha \in A(p)$ are the extreme rays of $\mathcal{S}(p)$.

Theorem 2 (Adjacent extreme points, $q_\alpha(p)$). Suppose **Assumptions 1 to 3** hold. For each $\alpha \in A(p)$, let $q_\alpha(p) \triangleq p + \lambda^* t$ where $\lambda^* = \min_{\lambda \geq 0} \{\lambda | p + \lambda r_\alpha(p) \in \cap_{\beta \in A'(p)} S_\beta\}$ where $A'(p) \triangleq \{\beta \in A \setminus A(p) \mid r_\alpha(p) \cap H_\beta \neq \emptyset\}$ and t is some point in the half-line $r_\alpha(p)$. Then $q_\alpha(p)$ is an extreme point of \mathcal{P} .

Proof. The set $A'(p)$ indexes the set of hyperplanes indexed by $A \setminus A(p)$ that intersect with $r_\alpha(p)$. These intersections are points by **Lemma 1**. Observe that $q_\alpha(p)$ is the closest point to p of intersections of hyperplanes in $A'(p)$ with $r_\alpha(p)$. This closest point $q_\alpha(p)$ is the intersection of $r_\alpha(p)$ with the boundary of $\cap_{\alpha \in A'(p)} S_\alpha$ —which being the boundary of the intersection of a collection of closed halfspaces (since $\phi_\alpha, \alpha \in A'(p)$ are continuous linear functions)—is closed. Note that there must be at least one point of intersection since \mathcal{P} is linearly bounded. This closest point must be distinct from p by **Assumptions 2 and 3** and the proof of **Proposition 2**. Note that $q_\alpha(p)$ is an extreme point of \mathcal{P} from **Theorem 1** since it is the unique point contained in the constraining hyperplanes passing through it. \square

The extreme points $\{q_\alpha(p) \mid \alpha \in A(p)\}$ are called the *adjacent* extreme points of p . The line segment $e_\alpha(p) = q_\alpha(p) - p$ joining $q_\alpha(p)$ and p is called an *edge direction* of \mathcal{P} for $\alpha \in A(p)$. The set $\mathcal{E}(p)$ of all edge directions out of p is given by $\mathcal{E}(p) = \cup_{\alpha \in A(p)} e_\alpha(p)$ and the set \mathcal{E} of all edges in \mathcal{P} is given by $\mathcal{E} = \cup_{p \in E} \mathcal{E}(p)$.

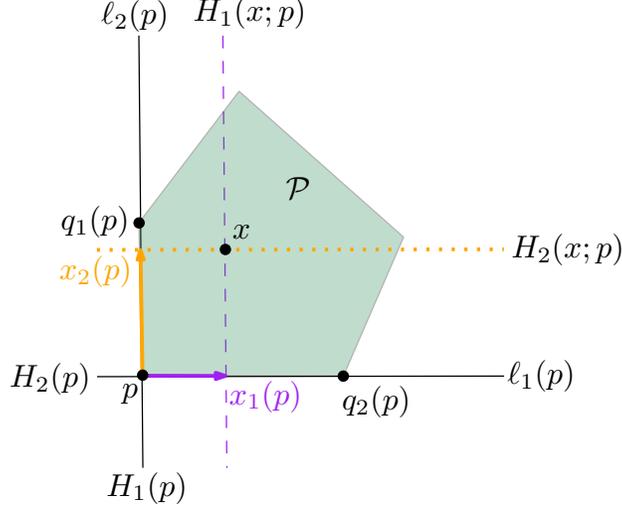


Figure 4: Visualizing the construction of the Schauder basis at an extreme point.

3.3 Edges from an extreme point form a Schauder basis

We now establish some key properties on the set of edges emanating from an extreme point that are essential in defining our simplex method. In particular, we will describe how these edge directions can be used to fashion a Schauder basis for the space and the Schauder basis coefficients are nonnegative when generating elements in the feasible region \mathcal{P} . This machinery is essential for showing the convergence of the simplex method we define in the next section. First, an assumption.

Assumption 4 (Countably many constraints). The set A is countable. In this case, we let k be the index of A and express $\mathcal{P} = \bigcap_{k=1}^{\infty} S_k$. In other words, from now, we use the index k instead of the index α .

Remark 1. One sufficient condition for [Assumption 4](#) to hold is that X is a separable Banach space [[10](#), Proposition 7.6.5]). In this case, even if \mathcal{P} was expressed with more than countably many constraints, there exists a countable set of constraints that also describes \mathcal{P} .

Remark 2. We slightly abuse notation in our indexing of the constraints $\phi_k(x)$. We will re-index the set of active constraints at p by $\phi_k(\cdot; p)$. In other words, we relabel the elements of $A(p)$ by $k = 1, 2, \dots$ and $\phi_k(\cdot; p)$ should be understood as the k active constraint at p .

Example 3 (Hilbert cube, continued). The index set $A = \{1, 2, \dots\} \cup \{-1, -2, \dots\}$ of the Hilbert cube is countable, and so satisfies [Assumption 4](#).

For the following development, refer to [Figure 4](#). Let p be an extreme point of a polytope \mathcal{P} . We begin by assuming $p = 0$. Fix $x \in \mathcal{P}$ and let $x_k(0)$ be the vector along the k th edge out of 0 given by

$$x_k(0) \triangleq \ell_k(0) \cap H_k(x; 0) \quad (4)$$

where $\ell_k(0)$ is the line defined in [Lemma 4](#) and $H_k(x; 0) = \{y \in X \mid \phi_k(y; 0) = \phi_k(x; 0)\}$ and $\phi_k(\cdot; 0)$.

We claim that

$$x_k(0) = \frac{\phi_k(x; 0)}{\phi_k(q_k(0); 0)} q_k(0) \quad (5)$$

by showing that the right-hand side of (5) is in $\ell_k(0) \cap H_k(x; 0)$ and is thus equal to $x_k(0)$ by the definition in (4). The right-hand side of (5) is clearly in $\ell_k(0)$. Moreover, note that

$$\phi_k \left(\frac{\phi_k(x; 0)}{\phi_k(q_k(0); 0)} q_k(0); 0 \right) = (\phi_k(x; 0) / \phi_k(x_k(0); 0)) \phi_k(x_k(0); 0) = \phi_k(x; 0)$$

and so the right-hand side of (5) is in $H_k(x; 0)$. This establishes (5).

Define $x^n(0) \triangleq \sum_{k=1}^n x_k(0)$.

Lemma 8. If Assumptions 1 to 4 hold then the sequence $x_n(0) \in \cap_{k=1}^n H_k(x; 0)$ for all $n = 1, 2, \dots$.

Lemma 9. Let $y \in X$. If $\cap_{k=1}^{\infty} H_k(y; 0) = \{y\}$ and $\{y^n\}$ is a convergent sequence with $y^n \in \cap_{k=1}^n H_k(y; 0)$ for $n = 1, 2, \dots$, then $y^n \rightarrow y$ as $n \rightarrow \infty$.

To ensure our Schauder basis construction converges, we need an additional assumption.

Assumption 5 (Compactness of the set of partial sums). For all $p \in \mathcal{P}$ and $x \in X$, the set $\{x^n(0)\}$ lies in a compact set (which may depend on x and p , but not on n).

Remark 3. A sufficient condition for Assumption 5 to hold is that the sequence $x^n(p), n = 1, 2, \dots$ be convergent since then we can set the containing compact set equal to the closure of $\{x^1(p), x^2(p), \dots\}$. This is a necessary as well as sufficient condition for Assumption 5 since $x^n(p), n = 1, 2, \dots$ is convergent under Assumption 5 as we show in Proposition 3 below. Also, since $\phi_k(x; p)$ and $\phi_k(q_k(p); p)$ are respectively bounded from above and below, a sufficient condition is that $\sum_{k=1}^{\infty} \|q_k(p)\|$ is absolutely convergent and X is some Banach space with norm $\|\cdot\|$.

◀

Remark 4. In the case where X is a reflexive Banach space (such as a Hilbert space or L_p space with $1 < p < \infty$), the condition of compactness can be relaxed to uniform boundedness. This yields the (weakly) convergent subsequence needed in the proof of Proposition 3 below, which carries through to the rest of the paper.

The following example exhibits another approach to establishing Assumption 5.

Example 4 (Hilbert cube, continued). We claim that \mathcal{H} satisfies Assumption 5. Fix the extreme point $p = 0$ (the argument is analogous for other extreme points) and note that $q_k(0) = e_k$ (the k th unit vector). Accordingly, for any $x = (x_j) \in \mathcal{H}$, $x_k(0) = x_k e_k$ and so the associated k th partial sum

$$x^n(0) = (x_1, \dots, x_n, 0, 0, \dots).$$

Each $x^n(0) \in \mathcal{H}$ since all coordinates lie in $[0, 1]$. Moreover,

$$\|x - x^n(0)\|_X^2 = \sum_{k=n+1}^{\infty} \delta_k^2 x_k^2 \longrightarrow 0 \quad \text{as } n \rightarrow \infty,$$

because $x \in X$ implies $\sum_{k=1}^{\infty} \delta_k^2 x_k^2 < \infty$. Therefore $x^n(0) \rightarrow x$, so the set $\{x_n(0) : n \geq 1\}$ is closed and thus compact since it lies in the compact set \mathcal{H} . This establishes Assumption 5. ◀

Proposition 3 (Schauder basis formed at origin). Suppose Assumptions 1 to 5 hold. Let $x \in \mathcal{P}$ and extreme point $p = 0$. The sequence $x^n(0) \rightarrow x$ as $n \rightarrow \infty$, that is, $x = \sum_{k=1}^{\infty} \frac{\phi_k(x; 0)}{\phi_k(q_k(0); 0)} q_k(0)$ and $\{q_k(0), k = 1, 2, \dots\}$ is a Schauder basis for the elements of \mathcal{P} .

Proof. Recall $x^n(0) = \sum_{k=1}^n x_k(0)$. Suppose the sequence $(x^n(0))$ does not converge to x . Then there exists an open neighborhood U of x such that there is a subsequence $(x^{n_\ell}(0))$ entirely outside U . By the compactness condition in [Assumption 1](#), this sequence must have an accumulation point \bar{x} . Because the complement of U is closed, $\bar{x} \notin U$, and therefore $\bar{x} \neq x$. However, because $x^{n_\ell}(0) \in \bigcap_{k=1}^{n_\ell} H_k(x; 0)$, and the hyperplanes are closed, it follows from [Lemma 9](#) that any accumulation point of these partial sums must be x . This contradicts $\bar{x} \neq x$. Hence, $x^n(0) \rightarrow x$ as $n \rightarrow \infty$. \square

Theorem 3 (Schauder basis formed at a given extreme point). Suppose [Assumptions 1](#) to [5](#) hold. Let $x \in \mathcal{P}$ and $p \in E$. Then

$$x = \sum_{k=1}^{\infty} \frac{\phi_k(x-p; p)}{\phi_k(q_k(p)-p; p)} (q_k(p) - p) + p$$

where the coefficients $\frac{\phi_k(x-p; p)}{\phi_k(q_k(p)-p; p)}$ on the basis vectors $q_k(p) - p$ are nonnegative.

Proof. Since topological vector spaces are invariant under translations, we can replace x by $x - p$ and substitute in [Proposition 3](#) above. \square

4 A geometric simplex algorithm

Consider the following linear program

$$\begin{aligned} \min \quad & c(x) \\ \text{s.t.} \quad & a_k(x) \leq b_k, \text{ for } k = 1, 2, \dots \end{aligned} \tag{6}$$

where $c : X \rightarrow \mathbb{R}$ is a continuous linear functional. We assume that the feasible region $\mathcal{P} = \{x : a_k(x) \leq b_k, \text{ for } k = 1, 2, \dots\}$ satisfies our assumptions up until this point in the paper. Note that an optimal solution x^* is attained since c is continuous and \mathcal{P} is compact by [Assumption 1](#).

What follows is a simplex algorithm for solving for x^* . It successively moves from an extreme point along a steepest descent edge to its adjacent extreme point. Let's formalize the notion of steepest descent edge. Consider the lines $\ell_k(p), k = 1, 2, \dots$ with $p \in E$ as one-dimensional vector spaces of X . Let $\|\cdot\|_{\ell_k(p)}$ be a norm for each vector space $\ell_k(p)$. Note that the topology on each space $\ell_k(p)$ is homeomorphic to that of the real line and not necessarily in agreement with the relative topology of X when restricted to $\ell_k(p)$.

We have that $(c(q_k(p)) - c(p)) / \|q_k(p) - p\|_{\ell_k(p)}$ is the rate of cost decrease (or increase) when moving along edge k out of p . Note that $\|q_k(p) - p\|_{\ell_k(p)} > 0$ since $q_k(p) \neq p$ for all $p \in E$. The *steepest descent edge* out of p would be an edge k that minimizes $(c(q_k(p)) - c(p)) / \|q_k(p) - p\|_{\ell_k(p)}$. We next assume that this minimum exists.

Assumption 6 (Steepest descent edge exists). For each extreme point p the following infimum is attained:

$$\gamma(p) := \min_{k=1,2,\dots} \frac{c(q_k(p)) - c(p)}{\|q_k(p) - p\|_{\ell_k(p)}}.$$

We call a resulting edge a *steepest descent edge* emanating from p .

Remark 5. A sufficient condition for [Assumption 6](#) is that for every extreme point p , the set of adjacent extreme points to p (and including p) are a closed set, since then by Weierstrass extreme value theorem, the minimization problem defining $\gamma(p)$ has an optimal solution.

Example 5 (Hilbert cube, continued). The extreme points of \mathcal{H} are the $\{0, 1\}$ -valued sequences (each belongs to X because $\sum_j \delta_j^2 < \infty$). Fix an extreme point $p \in \mathcal{P}$ and, for each $k \in \mathbb{N}$, let $q_k(p)$ be the adjacent extreme point obtained by flipping the k th coordinate of p from 0 to 1 or 1 to 0. Then $\|q_k(p) - p\|_X = \delta_k \rightarrow 0$ as $k \rightarrow \infty$. Hence $q_k(p) \rightarrow p$ in \mathcal{H} . This implies the set $E(p)$ is closed for all $p \in E$. Thus, via [Remark 5](#), [Assumption 6](#) is satisfied. ◀

It is important for the geometric simplex algorithm to select edges of steepest descent in order to reach an optimal solution (or converge in value to optimality) as we show in [Theorem 4](#). The issue is illustrated in the binary tree example of [21, p. 869-7]. There, the authors illustrate that if a simplex method takes a non-steepest descent choice initially, it can thereafter make steepest descent choices and not converge in value to optimality. This is because it can get “stuck” on an infinitely improving path that converges in value to a suboptimal level.

We are now ready to state the “geometric” simplex algorithm, the core mechanism studied in this paper:

Simplex Algorithm

1. Choose an initial extreme point $p \in \mathcal{P}$ and set $n = 0$ and $p^0 = p$.
2. **while** $\gamma(p^n) < 0$ **do**

Find the adjacent extreme point p^{n+1} along a steepest descent edge from p^n , that is,

$$\text{Set } p^{n+1} = \arg \min_{k=1,2,\dots} \frac{c(q_k(p^n)) - c(p^n)}{\|q_k(p^n) - p^n\|_{\ell_k(p^n)}}.$$

Set $n = n + 1$.

3. **end while**

Remark 6. Note that the Simplex Algorithm moves or pivots from p^n to p^{n+1} by removing exactly one of the hyperplanes passing through p^n and adding exactly one of the hyperplanes not passing through p^n during each iteration n .

We turn now to establishing that there is a strictly positive minimum cost reduction in each iteration of the simplex algorithm. This requires the following assumption.

Assumption 7 (Edge lengths bounded away from zero and infinity). $\nu = \inf_{p \in E, k \in A(p)} \|e_k(p)\|_{\ell_k(p)} > 0$ and $D \triangleq \sup_{p \in E, k \in A(p)} \|e_k(p)\|_{\ell_k(p)} < \infty$.

[Assumption 7](#) ensures that each edge direction has a minimal geometric “thickness,” preventing degenerate cases where the distance between adjacent extreme points tends to zero. Without this, the simplex path could accumulate infinitely many pivot steps in a bounded region without improving the objective sufficiently to guarantee convergence in value.

Example 6 (Hilbert cube, continued). Following [Example 5](#), we know every extreme point p has the form $p_i = 0$ or 1 for all i , and the set of adjacent extreme points to p has elements $q_k(p)$ where the k th coordinate of p is “flipped” from 0 to 1 or 1 to 0. The edge $e_k(p)$ is thus the unit vector with a 1 in the k th position and 0 otherwise, and $\ell_k(p)$ is the k th “coordinate” axis. Thus, $\|\cdot\|_{\ell_k(p)}$ is the usual topology on the real line. Thus, $\|e_k(p)\|_{\ell_k(p)} = 1$ for all k and all p and so $\nu = 1$. ◀

Lemma 10. If [Assumptions 1](#) to [7](#) hold then either the Simplex Algorithm terminates at some iteration n or $\lim_{n \rightarrow \infty} \gamma(p^n) = 0$.

For the remainder of the argument, we need another uniformity property to complement [Assumption 3](#).

Assumption 8 (Constraint functions uniformly bounded over extreme points). $\xi \triangleq \sup_{p \in E, \alpha \in A} |\phi_\alpha(p)| < \infty$.

Note that by Bauer's Maximum Principle [2, Theorem 7.69], this implies that the supremum over the whole set \mathcal{P} is also bounded by ξ , since an optimal solution exists among the extreme points of \mathcal{P} when optimizing over any continuous linear functional. [Assumption 8](#) also prevents pathological scaling of the constraint system that could destroy the uniformity of convergence arguments later in [Theorem 4](#).

Example 7 (Hilbert cube, continued). Recall from [Example 2](#) that $\phi_k(x) = \langle x, y^k \rangle$ where $y^k = \pm 1$ for all $k \in A$. As we have also observed, every extreme p has the form $p_j = 0$ or 1 for all j . Thus, we can see that $\phi_k(p) \leq 1$ for all $p \in E$. In particular $\xi = 1 < \infty$ and so [Assumption 8](#) is satisfied.

◀

In order to assure convergence in value of the Simplex Algorithm, we also need an assumption on the objective function:

Assumption 9 (Uniform convergence of edge costs). Let $c(q_k(p) - p)$ be the cost of edge $e_k(p)$ for $p \in E$ and $k = 1, 2, \dots$. We assume $\sum_{k=1}^{\infty} |c(q_k(p) - p)|$ converges uniformly over $p \in E$.

Example 8 (Hilbert cube, continued). Let $c : X \rightarrow \mathbb{R}$ be a continuous linear functional, represented by some $h = (h_j) \in X$ via the Riesz map: $c(x) = \langle h, x \rangle_X = \sum_{j \geq 1} \delta_j^2 h_j x_j$. For an extreme point p of H and its adjacent extreme point $q_k(p)$ (obtained by flipping the k th coordinate), the edge is $q_k(p) - p = \pm e_k$, so $c(q_k(p) - p) = \langle h, e_k \rangle_X = \pm \delta_k^2 h_k$ independent of p up to sign. Therefore $\sum_{k=1}^{\infty} |c(q_k(p) - p)| = \sum_{k=1}^{\infty} \delta_k^2 |h_k|$ and this series converges *uniformly in* $p \in E$ (indeed, the right-hand side does not depend on p) whenever $\sum_{k \geq 1} \delta_k^2 |h_k| < \infty$. Hence, under the mild additional summability condition $\sum_k \delta_k^2 |h_k| < \infty$ on the Riesz representative h of c , [Assumption 9](#) holds.

In finite dimensions, the simplex method can optimize *any* linear functional over a polyhedron. For a simplex method to work in infinite spaces, we need both properties on \mathcal{P} and c , since a linear functional is not nearly as well-behaved in such spaces as in finite-dimensional space. [Assumption 9](#) aims to offer a minimal restriction on c with respect to the set of extreme points E .

Remark 7. Note that $|c(q_k(p) - p)| \leq C\alpha^k$ for all $p \in E$ and $k = 1, 2, \dots$ for some $0 < \alpha < 1$ and constant $C > 0$ is a sufficient condition for [Assumption 9](#) to hold.

Theorem 4 (Convergence in value). If [Assumptions 1 to 9](#) hold then either the Simplex Algorithm terminates with an optimal solution or the infinite sequence of extreme points generated converges in value to an optimal solution; i.e. $\lim_{n \rightarrow \infty} c(p^n) = c^*$, where $c^* = c(x^*)$ and x^* is an optimal solution of (6).

Proof. Suppose the Simplex Algorithm terminates at iteration n . Then $\gamma(p^n) = \min_{k=1,2,\dots} (c(q_k(p^n)) - c(p^n)) / \|q_k(p^n) - p^n\| \geq 0$ so that $c(q_k(p^n)) - c(p^n) \geq 0$ for all $k = 1, 2, \dots$. From [Theorem 3](#), for any $x \in \mathcal{P}$, $x \neq p^n$, we have

$$x = \sum_{k=1}^{\infty} \frac{\phi_k(x - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (q_k(p^n) - p^n) + p^n$$

and since c is a continuous function,

$$c(x) - c(p^n) = \sum_{k=1}^{\infty} \frac{\phi_k(x - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (c(q_k(p^n)) - c(p^n)). \quad (7)$$

But $\phi_k(x - p^n; p^n) \leq 0$ and $\phi_k(q_k(p^n) - p^n; p^n) < 0$ since x and $q_k(p^n)$ are in \mathcal{P} and $q_k(p^n)$ is not equal to p^n . Hence $c(x) - c(p^n) \geq 0$ so that $c(p^n) \leq c(x)$ and therefore p^n is optimal.

Now, suppose the Simplex Algorithm does not terminate. Consider $x^* \in \mathcal{P}$. We have

$$c(x^*) - c(p^n) = \sum_{k=1}^{\infty} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (c(q_k(p^n)) - c(p^n)).$$

Let $\epsilon > 0$. We have by [Assumption 8](#)

$$|\phi_k(x^* - p^n; p^n)| \leq 2\xi < \infty$$

using the triangle equality.

Moreover, $\phi_k(q_k(p^n) - p^n; p^n) = \phi_k(q_k(p^n); p^n) - \phi_k(p^n; p^n) = \phi_k(q_k(p^n); p^n) - b_k(p^n) = -s_k(q_k(p^n); p^n)$. Thus $s_k(q_k(p^n)) = -\phi_k(q_k(p^n) - p^n; p^n) > 0$. But $s_k(q_k(p^n)) \geq \inf_{p \in E, \alpha \in A \setminus A(p)} s_k(p) = \rho > 0$ by [Assumption 2](#). Hence $|\phi_k(q_k(p^n) - p^n; p^n)| \geq \rho > 0$. By [Assumption 9](#), for all $\epsilon' > 0$ there is a $K'(\epsilon')$ such that $\sum_{k=1}^{\infty} |c(e_k(p))| < \epsilon$ for all $p \in E$.

Then

$$\begin{aligned} & \sum_{k=K'(\epsilon')}^{\infty} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (c(q_k(p^n)) - c(p^n)) \\ & \leq \sum_{k=K'(\epsilon')}^{\infty} \frac{|\phi_k(x^* - p^n; p^n)|}{|\phi_k(q_k(p^n) - p^n; p^n)|} |c(q_k(p^n)) - c(p^n)| \\ & \leq \frac{\xi D}{\rho} \sum_{k=K'(\epsilon')}^{\infty} |c(q_k(p^n)) - c(p^n)| \\ & = \frac{2\xi D}{\rho} \sum_{k=K'(\epsilon')}^{\infty} |c(e_k(p^n))| \\ & < \frac{2\xi D}{\rho} \epsilon'. \end{aligned}$$

Set $\epsilon' = \rho\epsilon/2\xi D$, $K(\epsilon) = K'(\epsilon')$ and we have

$$\sum_{k=K(\epsilon)}^{\infty} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (c(q_k(p^n)) - c(p^n)) < \epsilon$$

for all $\epsilon > 0$ and $p^n, n = 1, 2, \dots$. That is

$$\sum_{k=1}^{\infty} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (c(q_k(p^n)) - c(p^n))$$

converges uniformly over $p^n, n = 1, 2, \dots$

Then for any $\epsilon > 0$, there is a $K(\epsilon)$ independent of p^n such that

$$-\epsilon < c(x^*) - c(p^n) - \sum_{k=1}^{K(\epsilon)} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (c(q_k(p^n)) - c(p^n)) < \epsilon$$

so

$$c(x^*) - c(p^n) > \sum_{k=1}^{K(\epsilon)} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (c(q_k(p^n)) - c(p^n)) - \epsilon \quad (8)$$

for all $p^n, n = 1, 2, \dots$

We have by [Lemma 10](#), $\lim_{n \rightarrow \infty} \gamma(p^n) = 0$ while $\gamma(p^n) < 0$ for all n . Hence for all $\epsilon' > 0$, there is a $N(\epsilon') > 0$ such that $0 > \gamma(p^n) > -\epsilon'$ for all $n > N(\epsilon')$. Hence for $n > N(\epsilon')$,

$$0 > \min_{k=1,2,\dots} (c(q_k(p^n)) - c(p^n)) / \|q_k(p^n) - p^n\|_{\ell_k(p^n)} > -\epsilon'$$

so that

$$0 > (c(q_k(p^n)) - c(p^n)) / \|q_k(p^n) - p^n\|_{\ell_k(p^n)} > -\epsilon'$$

or

$$c(q_k(p^n)) - c(p^n) > -\|q_k(p^n) - p^n\|_{\ell_k(p^n)} \epsilon' > -D\epsilon' \quad (9)$$

for all $k = 1, 2, \dots$ and $n > N(\epsilon')$.

From [\(8\)](#) and [\(9\)](#),

$$c(x^*) - c(p^n) > \sum_{k=1}^{K(\epsilon)} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} (-D\epsilon') - \epsilon \quad (10)$$

for all $n > N(\epsilon')$ since $\phi_k(x^* - p^n; p^n) \leq 0$ and $\phi_k(q_k(p^n) - p^n; p^n) < 0$ because $x^* \in \mathcal{P}$ and $q_n(p^n) \in \mathcal{P}$ and $q_n(p^n) \neq p^n$.

Hence from [\(10\)](#),

$$c(x^*) - c(p^n) > -D\epsilon' \sum_{k=1}^{K(\epsilon)} \frac{\phi_k(x^* - p^n; p^n)}{\phi_k(q_k(p^n) - p^n; p^n)} - \epsilon$$

for all $n > N(\epsilon')$.

Let $\bar{\phi}(\epsilon, n, x^*) = \max_{k=1,2,\dots,K(\epsilon)} \phi_k(x^* - p^n; p^n)$ and $\underline{\phi}(\epsilon, n, x^*) = \min_{k=1,2,\dots,K(\epsilon)} \phi_k(q_k(p^n) - p^n; p^n)$.

Then

$$c(x^*) - c(p^n) > -D\epsilon' K(\epsilon) \frac{\bar{\phi}(\epsilon, n, x^*)}{\underline{\phi}(\epsilon, n, x^*)} - \epsilon$$

for all $n > N(\epsilon')$. Set $\epsilon' = (\epsilon \bar{\phi}(\epsilon, n, x^*)) / (2DK(\epsilon)\underline{\phi}(\epsilon, n, x^*))$. Then $c(x^*) - c(p^n) > -\epsilon/2 - \epsilon/2 = -\epsilon$ for all $n > N(\epsilon')$. Since $\epsilon > 0$ was arbitrary, $\lim_{n \rightarrow \infty} c(p^n) = c(x^*)$. \square

Note that since \mathcal{P} is compact and the sequence $(p^n)_{n=1,2,\dots}$ is contained in \mathcal{P} , the sequence must have an accumulation point $p \in P$. By [Theorem 4](#) and the continuity of c , we have $c^* = \lim = c(p^{n_\delta})$ where n_δ is a converging subnet, so that p is an optimal solution. Below we show that when the optimal solution is unique, we obtain a sequence of pivots p^n that converges to x^* .

Corollary 1. Let [Assumptions 1 to 9](#) hold and suppose the optimal solution x^* is unique. Either the Simplex Algorithm terminates with the optimal solution x^* or the infinite sequence of extreme points generated converges to the optimal solution x^* , i.e. $\lim_{n \rightarrow \infty} p^n = x^*$.

Proof. From [Theorem 4](#), either the simplex algorithm terminates in an optimal solution or $\lim_{n \rightarrow \infty} c(p^n) = c^*$ where $c^* = c(x^*)$. Suppose $p^n \rightarrow x^*$ as $n \rightarrow \infty$. Then there is an open neighborhood U containing x^* such that a subsequence $(p^{i_j})_{j=1,2,\dots}$ lies entirely outside of U . But \mathcal{P} is compact, the closed subset $\mathcal{P} \setminus U$ is also compact. Therefore, the sequence $(p^{i_j})_{j=1,2,\dots}$ has an accumulation point $x' \in \mathcal{P} \setminus U$. By the continuity of c and [Theorem 4](#), evaluating c along the subnet converging to x' yields $c(x') = c^*$, making x' an optimal solution. This contradicts the assumption that x^* is uniquely optimal, since $x' \notin U$ ensures $x \neq x'$. \square

5 Defining polytopes

As mentioned in the introduction, there is a long history of attempting to define what we mean by a “polytope” in general topological space. As outlined, the usual notions from the finite-dimensional setting fail somewhat spectacularly. This gave rise to a long succession of attempts to define polytopes in infinite dimensions in a more workable manner [[1](#), [3](#), [8](#), [11](#), [13](#), [23](#), [28](#), [31](#), [32](#), [34](#), [36](#)]. For example, Phelps [[34](#)] was interested in the Choquet “simplex” and attempted definitions of a polytope that captured this “simplex”. However, a stunning set of results by [[7](#)] showed that nearly all of these attempted definitions excluded something quite simple: infinite-dimensional cubes. This is especially surprising from an optimization perspective, since a cube corresponds to “box constraints” in linear programming. One would hope that any definition of polytope captures such a foundational setting. We have shown in [Examples 2 to 8](#) that the Hilbert cube satisfies the assumptions of this paper’s setting. We are thus the first paper that we are aware of to construct a simplex method that is amenable to the Hilbert cube.

For concreteness, let us illustrate how a cube can present issues for an existing definition of a polytope in the literature. Consider the definition of Klee in [[28](#)], which is arguably the easiest to state. Klee defines a polytope to be any closed convex set whose intersection with every finite-dimensional affine subspace (also called a “flat” in Klee’s language) is a finite-dimensional polytope (in its usual definition) in that flat. The following example shows that the Hilbert cube \mathcal{H} defined in [Example 2](#) fails this definition in a spectacular fashion.

Example 9 (Hilbert cube is not a Klee polytope). We show that the Hilbert cube \mathcal{H} defined in [Example 2](#) is *not* a polytope in Klee’s sense by showing that a section of \mathcal{H} with a 2-dimensional flat is not a finite-dimensional polytope in that flat. In fact, we can cut the Hilbert cube with a 2-dimensional flat and reveal a disc!

Let $p, u, v \in \mathbb{R}^{\mathbb{N}}$ where u and v are linearly independent (that is, there does not exist a $c \in \mathbb{R}$ such that $v = cu$). Then the set

$$F_{uv}(p) = \{p + \alpha u + \beta v \mid \alpha, \beta \in \mathbb{R}\}$$

is a 2-dimensional flat of $\mathbb{R}^{\mathbb{N}}$. The intersection $\mathcal{H} \cap F_{uv}(p)$ is an (at most 2-dimensional) section of \mathcal{H} given by

$$\mathcal{H} \cap F_{uv}(p) = \{p + \alpha u + \beta v \mid 0 \leq p_i + \alpha u_i + \beta v_i \leq 1, \text{ for all } \alpha, \beta \in \mathbb{R} \text{ and } i = 1, 2, \dots\}.$$

Re-expressing the above set through its linearly isomorphic representation in the two-dimensional subspace V of $\mathbb{R}^{\mathbb{N}}$ spanned by $\{u, v\}$ yields:

$$S_{uv}(p) = \{(\alpha, \beta) \in \mathbb{R}^2 \mid 0 \leq p_i + u_i \alpha + v_i \beta \leq 1, \text{ for all } i = 1, 2, \dots\}. \quad (11)$$

The goal for the rest of the proof is to show that u , v and p can be chosen so that $S_{uv}(p)$ is the unit disk centered at the origin in V . Note that $S_{uv}(p)$ is a countable intersection of halfspaces in \mathbb{R}^2 .

The first insight is that the unit disk can also be expressed as the countable intersection of halfspaces. Let C be the unit circle and D be the unit disk in \mathbb{R}^2 and let

$$E = \{(\alpha, \beta) \in \mathbb{R}^2 : r\alpha + q\beta \leq 1 \text{ for all } (r, q) \in C \cap \mathbb{Q}^2\},$$

where \mathbb{Q} is set of rational numbers. We claim that $D = E$. We first show that $D \subseteq E$. Let T be any tangent line of C in \mathbb{R}^2 . By the convexity of D , it is clear that D lies on entirely one side of T . Let $(r, q) \in C \cap \mathbb{Q}^2$ and note that the set $\{(\alpha, \beta) \mid r\alpha + q\beta = 1\}$ is the tangent line to C through (r, q) . It is clear that $D \subset \{(\alpha, \beta) \mid r\alpha + q\beta \leq 1\}$. Taking the intersection over all $(r, q) \in C \cap \mathbb{Q}^2$ yields $D \subseteq E$.

Conversely, let $(a, b) \notin D$. Let (\bar{a}, \bar{b}) be the point in D with the smallest distance to (a, b) , where the distance derives from the Euclidean norm in \mathbb{R}^2 . The tangent plane $\{(\alpha, \beta) \mid \bar{a}\alpha + \bar{b}\beta = 1\}$ separates D from (a, b) . In particular, we know $\bar{a}a + \bar{b}b > 1$. Hence, there exist rationals $(\bar{r}, \bar{q}) \in C \cap \mathbb{Q}^2$ sufficiently close to (\bar{a}, \bar{b}) so that it remains that $\bar{r}a + \bar{q}b > 1$. Hence $(a, b) \notin E$. That is, $E \subseteq D$. This implies $E = D$.

Now, let \bar{u}, \bar{v} be chosen such that $\bar{u}_i = r/2$ and $\bar{v}_i = q/2$ where (r, q) span the countably many points $C \cap \mathbb{Q}^2$. Choose \bar{p} such that $\bar{p}_i = 1/2$ for all $i \in \mathbb{N}$. We claim that $S_{\bar{u}\bar{v}}(\bar{p}) = E$. First we show $S_{\bar{u}\bar{v}}(\bar{p}) \subseteq E$. According to (11) we have

$$S_{\bar{u}\bar{v}}(\bar{p}) = \{(\alpha, \beta) \in \mathbb{R}^2 \mid -1 \leq r\alpha + q\beta \leq 1, \text{ for all } (r, q) \in C \cap \mathbb{Q}^2\}$$

after some rearranging. Hence it clearly follows that $S_{\bar{u}\bar{v}}(\bar{p}) \subseteq E$ since $S_{\bar{u}\bar{v}}(\bar{p})$ includes all the inequalities defining E (indeed all the $r\alpha + q\beta \leq 1$ constraints). Conversely, we note that $E \subseteq S_{\bar{u}\bar{v}}(\bar{p})$. This also follows immediately since if $r\alpha + q\beta \leq 1$ is an inequality defining E then so is $-r\alpha - q\beta \leq 1$ since if $(r, q) \in C \cap \mathbb{Q}^2$ then $(-r, -q) \in C \cap \mathbb{Q}^2$. Rewriting the latter inequality yields $r\alpha + q\beta \geq -1$. This argument for all $(r, q) \in C \cap \mathbb{Q}^2$ recovers all of the constraints of $S_{\bar{u}\bar{v}}(\bar{p})$. Hence, $E \subseteq S_{\bar{u}\bar{v}}(\bar{p})$ and conclude $E = S_{\bar{u}\bar{v}}(\bar{p})$.

Putting things together, our work shows $D = S_{\bar{u}\bar{v}}(\bar{p})$. This implies that $\mathcal{H} \cap F_{\bar{u}\bar{v}}(\bar{p})$ is linearly isomorphic to a unit disk. Clearly, the unit disk cannot be expressed as the intersection of finitely many halfspaces. Hence, \mathcal{H} is not a Klee polyhedron. ◀

Remark 8. Note that if we replace the unit disk D in the above example with *any* finite-dimensional compact convex set, then a similar construction can be made and p, u, v chosen so that D is an affine section of \mathcal{H} . This follows from [2, Corollary 7.49] that every finite-dimensional compact convex set is the intersection of countably many closed halfspaces. ◀

We here put forward our definition of a polytope in infinite dimensional space—an object that can be traversed by the Simplex Algorithm defined in this paper. Such polytopes may hopefully enjoy the simplex method efficiency enjoyed by finite-dimensional linear programs. This offers a kind of “optimizer’s perspective” on the geometric question of defining a polytope—a polytope is something that we have an algorithmic procedure to traverse extreme points to optimize a linear functional. This definition includes the Hilbert cube, but also many others objects that satisfy [Assumptions 1 to 8](#).

These assumptions recover key geometric properties that are essential for the simplex viewpoint, as we show in the following result. We will say two extreme points p and p' of \mathcal{P} are *connected by a path* if there exists a finite or infinite sequence of adjacent edges that define a path of extreme points (p, p^1, p^2, \dots) such that p^n converges to p' in the limit. We say \mathcal{P} is *path-connected* if all of its extreme points are connected by a path.

Proposition 4. If a \mathcal{P} defined in (1) satisfies **Assumptions 1 to 8** then it is path-connected and all its extreme points are exposed.

Proof. We first show all extreme points are exposed. Fix an extreme point $e \in E$ and we show that the linear functional

$$K_e(x) = \sum_{i \in A(e)} s_i(x)(1/2)^i,$$

where $s_i(x) = b_i - \phi_i(x)$, exposes the point p .

First, we claim that K_e is a continuous linear functional for every $e \in E$. First, recall that $\xi = \sup_{i \in A, e \in E} s_i(e) < \infty$ by **Assumption 8** and note that s_i are continuous convex functions. By the Bauer Maximum Principle [2, Theorem 7.69] a continuous convex function on a compact convex set attains its maximum at an extreme point, so we conclude that $\sup_{i \in A, p \in \mathcal{P}} s_i(p) \leq \xi < \infty$. Second, the Weierstrass M-Test provides that a series of functions converges uniformly whenever its terms are bounded by terms of a second series which is absolutely convergent. We have $s_i(p)(1/2)^i \leq \xi(1/2)^i$ for all $i \in A$ and $\sum_{i \in A(e)} \xi(1/2)^i \leq 2\xi < \infty$ so that the series $\sum_{i \in A(e)} s_i(p)(1/2)^i$ is uniformly convergent. Finally, by the Uniform Limit Theorem, the limit of a uniformly convergent sequence of continuous functions is continuous.

Second, we show that K_e is uniquely optimized at e over \mathcal{P} . Note that $s_i(p) \geq 0$ for all $i \in A(e)$ since $p \in \mathcal{P}$. Moreover $s_i(p) = 0$ for all $i \in A(e)$ implies $p = e$ since then $0 = b_i - \phi_i(p)$ or $\phi_i(p) = b_i$ for all $i \in A(e)$ so that $p = e$ uniquely since $\mathcal{H}(e) = \{e\}$ by **Proposition 2**. Since $s_i(p) = 0$ for all $i \in A(e)$ is optimal for the problem

$$\min_{p \in \mathcal{P}} K_e(p), \tag{12}$$

extreme point e is the unique optimal solution for $\min_{p \in \mathcal{P}} K_e(p)$ so that $\min_{p \in \mathcal{P}} K_e(p) = K_e(e)$. Hence all extreme points are exposed points.

Finally, we show every pair of extreme points is connected by a path of edges. The simplex method applied to the minimization problem (12) will by **Corollary 1** generate a sequence of edge connected extreme points p^n terminating in e or converging to e as $n \rightarrow \infty$ regardless of starting extreme point e' since (12) has the unique solution e . We conclude that there is an edge path connecting any two extreme points e and e' in polytope \mathcal{P} . \square

6 Conclusion with a discussion of the assumptions

In this paper we have studied a geometric version of the simplex method for topological vector spaces. We imposed a collection of assumptions to ensure that the geometric mechanics of the simplex method—moving from extreme points to extreme points along “improving” edges—converges in objective value to optimality. We show that our conditions generalize those found in the literature and cover the Hilbert cube, an interesting example that has troubled earlier definitions of the notion of a polytope (even though it is arguably the simplest possible feasible region of a linear program in our context).

A natural concern with any infinite-dimensional extension of the simplex method is the number and strength of structural assumptions required to recover the familiar geometric picture from finite dimensions. We emphasize that the assumptions introduced in this paper are not designed to be minimal. Rather, they serve to isolate the precise points at which infinite dimensionality obstructs the classical simplex argument and to make those obstructions explicit. In finite-dimensional linear programming, properties such as isolation of extreme points, nondegenerate edges, finite basis

representations, and assured attainment of steepest descent directions are guaranteed by finite-dimensional polyhedrality. In general locally convex spaces, these properties must be imposed directly.

Viewed in this light, our framework should be understood as a structural existence result: it identifies a broad class of closed convex sets—extending beyond previously studied settings—for which a geometric simplex method is well-defined in principle and converges in value. Several of the assumptions jointly rule out distinct forms of geometric or analytic degeneracy, and we do not claim their logical independence. However, they are consistent, since, for example, the Hilbert cube satisfies all assumptions. Still, a number of the assumptions imposed in this paper reflect genuine obstructions that arise when attempting to extend the geometric simplex method beyond finite dimensions, rather than mere technical conveniences.

Some form of compactness or effective boundedness ([Assumption 1](#)) is unavoidable if one wishes to guarantee the existence of extreme points and attainment of optimal solutions for continuous linear objectives. In general locally convex spaces, closed convex sets may lack extreme points altogether or admit minimizing sequences with no convergent subsequence, rendering any simplex-type method ill-posed at the outset (see, e.g., [4]).

Likewise, uniform separation of active and inactive constraints at extreme points ([Assumption 2](#)) is necessary to exclude extreme points with vanishing slack, a phenomenon closely related to non-exposed extreme points studied by [28, 34]. In the presence of diminishing slack, extreme points fail to behave as isolated geometric vertices, and no meaningful notion of an edge direction along which to move can be defined.

Several further assumptions rule out failure modes that are specific to infinite dimensions and have no finite-dimensional analogue. Uniform boundedness of constraint functionals ([Assumptions 3 and 8](#)) prevents pathological scaling in which constraint normals grow without bound along geometrically relevant directions, destabilizing the construction of edges and descent rates. Similarly, the requirement that edge lengths be bounded away from zero ([Assumption 7](#)) excludes situations in which sequences of adjacent extreme points accumulate, allowing infinitely many pivots within a compact region with negligible objective improvement. Finally, uniform summability of edge costs ([Assumption 9](#)) reflects the necessity of controlling how infinitely many local descent directions aggregate; without such control, infinitesimal but nonsummable improvements can defeat convergence in value.

Taken together, these considerations suggest that while individual assumptions may admit reformulation or weakening, any geometric simplex method that retains the classical interpretation of moving along edges between adjacent extreme points must confront—and in some form exclude—the same underlying obstructions identified here. Determining which conditions can be weakened, replaced, or derived from more intrinsic geometric properties is an important direction for future work. We hope that by making these requirements explicit, the present work provides a useful organizing framework for understanding when simplex-type methods can meaningfully extend beyond finite dimensions.

Additionally, future work may extend this geometric perspective in several directions. There is scope for a theory of degeneracy, cycling, and perturbation analysis in this broader setting. One might also investigate what further assumptions guarantee that each iteration of the Simplex Method can be executed in finite time.

A Proofs of lemmas

A.1 Proof of Lemma 1

Let line $\ell(p, r) \triangleq \{x \in X \mid x = p + \lambda r, \lambda \in \mathbb{R}\}$ where $r, p \in X$ and let hyperplane $H = \{x \in X \mid \phi(x) = b\}$ where ϕ is a linear functional on X and $b \in R$. Suppose $\ell(p, r)$ intersects H , i.e., there is a $\lambda_0 \in \mathbb{R}$ such that $\phi(p + \lambda_0 r) = b$. Suppose further that $\ell(p, r)$ does not lie in H . Then $\phi(r) \neq 0$ for otherwise $b = \phi(p + \lambda_0 r) = \phi(p) + \lambda_0 \phi(r) = \phi(p)$ in which case $\phi(p + \lambda r) = \phi(p) + \lambda \phi(r) = b$ for all $\lambda \in \mathbb{R}$ and $\ell(p, r)$ would lie in H . Now suppose $\ell(p, r)$ meets H in more than one point. Then there are λ_1 and λ_2 , $\lambda_1 \neq \lambda_2$ such that $\phi(p + \lambda_1 r) = \phi(p + \lambda_2 r)$ so that $\phi(p) + \lambda_1 \phi(r) = \phi(p) + \lambda_2 \phi(r)$ or $\lambda_1 \phi(r) = \lambda_2 \phi(r)$ or $\lambda_1/\lambda_2 = 1$ since $\phi(r) \neq 0$. Contradiction since $\lambda_1 \neq \lambda_2$. Hence $\ell(p, r)$ must meet H in exactly one point. \square

A.2 Proof of Lemma 2

Let S be a non-trivial affine subspace in X and let H be a hyperplane intersecting S in a singleton point p . Let q be another point in S distinct from p which we know exists since S is non-trivial. Then the line $\ell(p, q - p) = \{p + \lambda(q - p) \mid \lambda \in R\} \subseteq S$ since S is an affine subspace. Suppose there is a point $o \in S$ with $o \notin \ell(p, q - p)$. Consider $\ell(p, q - p)$ translated to pass through o , i.e. $\ell(o, q - p) = \{o + \lambda(q - p) \mid \lambda \in \mathbb{R}\}$. Note that $\ell(o, q - p) \subset S$ since it is a translate of a set in S . By Lemma 1, $\ell(o, q - p)$ intersects H in a single point p' .

Claim: $\ell(o, q - p) \cap \ell(p, q - p) = \emptyset$. Proof: If not, there is a $y \in X$ with $y = o + \lambda'(q - p) = p + \lambda''(q - p)$ or $o - p = (\lambda' - \lambda'')(q - p)$ in which case $o = p + (\lambda' - \lambda'')(q - p)$ so that $o \in \ell(p, q - p)$. Contradiction.

Hence $p' \notin \ell(p, q - p)$ so that $p' \neq p$. Contradiction since H intersects S in the unique point p . Hence $S = \ell(p, q - p)$ is a line through p . \square

A.3 Proof of Lemma 3

We prove the set A^* exists using Zorn's lemma. Let $\mathcal{G} := \{K \subseteq A : R(K) \triangleq \bigcap_{\alpha \in A \setminus K} S_\alpha = \mathcal{P}\}$. We claim \mathcal{G} has a maximal element K^* . Indeed, let $\mathcal{C} \subseteq \mathcal{G}$ be an arbitrary chain in \mathcal{G} and set $K^{\mathcal{C}} \triangleq \bigcup_{K \in \mathcal{C}} K$. Then

$$R(K^{\mathcal{C}}) = \bigcap_{\alpha \in A \setminus K^{\mathcal{C}}} S_\alpha = \bigcap_{K \in \mathcal{C}} \left(\bigcap_{\alpha \in A \setminus K} S_\alpha \right) = \bigcap_{K \in \mathcal{C}} R(K) = \bigcap_{K \in \mathcal{C}} \mathcal{P} = \mathcal{P}.$$

Hence $K^{\mathcal{C}}$ is an upper bound on chain \mathcal{C} in the inclusion order. This implies that every chain in \mathcal{G} has an upper bound and so there exists an inclusion-maximal element K^* of \mathcal{G} by Zorn's lemma.

Finally, we claim $A^* = A \setminus K^*$ has the desired property. Indeed, since $K^* \in \mathcal{G}$ we have $\bigcap_{\alpha \in A^*} S_\alpha = R(K^*) = \mathcal{P}$. Next, suppose, by way of contradiction that the set of constraints $\{S_\alpha \mid \alpha \in A^*\}$ is degenerate; that is, there exists a $\beta \in A^*$ such that $\bigcap_{\alpha \in A^* \setminus \{\beta\}} S_\alpha = \mathcal{P}$. But then this implies $K^* \cup \{\beta\} \in \mathcal{G}$, contradicting the maximality of K^* . \square

A.4 Proof of Lemma 4

The set $\ell_\alpha(p)$ is an intersection of affine subspaces containing p and is thus an affine subspace containing p . By nondegeneracy (Lemma 3), there is a point $p' \in \ell_\alpha(p)$ with $p' \neq p$ so that $\ell_\alpha(p)$ is a non-trivial affine subspace. Also, H_α intersects $\ell_\alpha(p)$ at the singleton point p by Theorem 1. Hence, by Lemma 2, $\ell_\alpha(p)$ is a line through p . \square

A.5 Proof of Lemma 5

Let $d \in \mathcal{R}(p)$. Then $b_\alpha \geq \phi_\alpha(p + \lambda d) = \phi_\alpha(p) + \lambda \phi_\alpha(d) = b_\alpha + \lambda \phi_\alpha(d)$ for all $\lambda \geq 0$ and $\alpha \in A(p)$ since $\phi_\alpha(p) = b_\alpha$ for $\alpha \in A(p)$. Hence $\phi_\alpha(d) \leq 0$ for all $\alpha \in A(p)$.

For the reverse, let $d \in \{d \in X \mid \phi_\alpha(d) \leq 0 \text{ for all } \alpha \in A(p)\}$. Then $\phi_\alpha(p + \lambda d) = \phi_\alpha(p) + \lambda \phi_\alpha(d) = b_\alpha + \lambda \phi_\alpha(d) \leq b_\alpha$ for all $\lambda \geq 0$ and $\alpha \in A(p)$. \square

A.6 Proof of Lemma 6

Note that $d \in \eta_\alpha(p) \Leftrightarrow d+p \in \ell_\alpha(p) \cap S_\alpha \Leftrightarrow d \in \ell_\alpha(p) - p$ and $d \in S_\alpha - p \Leftrightarrow d \in \bigcap_{\alpha' \in A(p), \alpha' \neq \alpha} H_{\alpha'} - p$ and $\phi_\alpha(d) \leq 0 \Leftrightarrow \phi_{\alpha'}(d) = 0$ for $\alpha' \in A(p), \alpha' \neq \alpha$ and $\phi_\alpha(d) \leq 0$ since subtracting p translates p in $\ell_\alpha(p)$ and S_α to the origin.

First we will show that $\eta_\alpha(p), \alpha \in A(p)$ is an extreme ray of $\mathcal{R}(p)$. Let $x \in \eta_\alpha(p), y \in \mathcal{R}(p)$ and $x - y \in \mathcal{R}(p)$. Then a) $\phi_{\alpha'}(x) = 0$ for $\alpha' \neq \alpha$, b) $\phi_\alpha(x) \leq 0$, c) $\phi_{\alpha'}(y) \leq 0$ and $\phi_{\alpha'}(x - y) \leq 0$ for all $\alpha' \in A(p)$. From c), we have d) $\phi_{\alpha'}(x) \leq \phi_{\alpha'}(y)$ for all $\alpha' \in A(p)$. From a) and d), we have e) $\phi_{\alpha'}(y) = 0$ for all $\alpha' \neq \alpha$. From c) and e), we conclude $y \in \eta_\alpha(p)$ and hence $\eta_\alpha(p)$ is an extreme ray of $\mathcal{R}(p)$.

We turn now to showing that every extreme ray of $\mathcal{R}(p)$ is equal to $\eta_\alpha(p)$ for some $\alpha \in A(p)$. Let η be an extreme ray of $\mathcal{R}(p)$ and suppose $\eta \neq \eta_\alpha(p)$ for any $\alpha \in A(p)$. Let $d \in \eta$. Since $d \in \mathcal{R}(p)$ and $d \neq 0, \phi_{\alpha'}(d) \leq 0$ for all $\alpha \in A(p)$ while $\phi_{\alpha_1} < 0$ for some $\alpha_1 \in A(p)$ since $d = 0$ uniquely satisfies $\phi_{\alpha'}(d) = 0, \alpha' \in A(p)$. Also since $\eta \neq \eta_\alpha(p)$ for any $\alpha \in A(p), \phi_{\alpha_2}(d) < 0$ for some $\alpha_2 \neq \alpha_1$ in $A(p)$. Let $d' \in \delta \eta_{\alpha_1}(p) + (1 - \delta) \eta_{\alpha_2}(p)$ and choose $\delta \in (0, 1)$ so that $d' \notin \eta$ and $d' \neq 0$. Note that $d' \in \mathcal{R}(p)$ by convexity of $\mathcal{R}(p)$. We want to choose $\lambda \in (0, 1)$ so that

$$d - \lambda d' \in \mathcal{R}(p) \quad (13)$$

since then we have $d \in \eta, d - \lambda d' \in \mathcal{R}(p)$, and $d - (d - \lambda d') = \lambda d' \in \mathcal{R}(p)$ while $d - \lambda d' \notin \eta$ so that η is not an extreme ray.

We have (13) holds if $\phi_{\alpha'}(d - \lambda d') \leq 0$ for all $\alpha' \in \mathcal{R}(p)$. Now $d' \in \delta \eta_{\alpha_1}(p) + (1 - \delta) \eta_{\alpha_2}(p)$ means $d' = \delta d'_1 + (1 - \delta) d'_2$ where $d'_1 \in \eta_{\alpha_1}(p)$ and $d'_2 \in \eta_{\alpha_2}(p)$. Then $\phi_{\alpha'}(d - \lambda d') = \phi_{\alpha'}(d) - \lambda(\delta \phi_{\alpha'}(d'_1) - (1 - \delta) \phi_{\alpha'}(d'_2))$.

We have for $\alpha' = \alpha_1, \phi_{\alpha'}(d) < 0, \phi_{\alpha'}(d'_1) < 0, \phi_{\alpha'}(d'_2) = 0$, for $\alpha' = \alpha_2, \phi_{\alpha'}(d) < 0, \phi_{\alpha'}(d'_1) = 0, \phi_{\alpha'}(d'_2) < 0$ and for $\alpha' \neq \alpha_1, \alpha' \neq \alpha_2, \phi_{\alpha'}(d) \leq 0, \phi_{\alpha'}(d'_1) = 0, \phi_{\alpha'}(d'_2) = 0$. Hence (13) holds for $\lambda < \min(\phi_{\alpha_1}(d)/\delta \phi_{\alpha_1}(d'_1), \phi_{\alpha_2}(d)/(1 - \delta) \phi_{\alpha_2}(d'_2))$. Therefore, η is not an extreme ray. Contradiction. \square

A.7 Proof of Lemma 8

Let $x^n(0) = \sum_{k=1}^n x_k(0)$. Let $i \in \{1, 2, \dots, n\}$. Note that for $i \neq k$,

$$\phi_i(x_k(0); 0) = (\phi_k(x; 0)/\phi_k(q_k(0); 0)) \phi_i(q_k(0); 0) = (\phi_k(x; 0)/\phi_k(q_k(0); 0)) \phi_i(q_k(0); 0) = 0$$

since $\phi_i(q_k(0); 0) = 0$. Hence $\phi_i(x^n(0); 0) = \phi_i(\sum_{k=1}^n x_k(0); 0) = \phi_i(x_i(0); 0) = \phi_i(x; 0)$ since $r_i(0) \in H_i(x; 0)$. Hence $x^n(0) \in H_i(x; 0)$ for all $i = 1, 2, \dots, n$. \square

A.8 Proof of Lemma 9

Let $z = \lim_{n \rightarrow \infty} y_n$. Suppose $z \notin H_k(y; 0)$ for some k . Put an open set U around z not intersecting closed set $H_k(y; 0)$. Let N be large enough so that $y_n \in U$ for all $n \geq N$. Set $M = \max\{k, N\}$. Then $y_M \in U$ and therefore $y_M \notin H_k(y; 0)$ while $y_M \in \bigcap_{n=1}^M H_k(y; 0)$. Contradiction. Hence $z \in H_k(y; 0)$ for all $k = 1, 2, \dots$. Then $z \in \bigcap_{k=1}^\infty H_k(y; 0) = \{y\}$ so that $z = y$, noting that $\bigcap_{k=1}^\infty H_k(y; 0) = \{y\}$ since $\bigcap_{k=1}^\infty H_k(0) = \{0\}$. \square

A.9 Proof of Lemma 10

Suppose the simplex algorithm does not terminate at any finite iteration n . Then $\gamma(p^n) < 0$ for all n . Suppose $\gamma(p^n) \not\rightarrow 0$ as $n \rightarrow \infty$. Then there is an $\epsilon > 0$ and a subsequence n_1, n_2, \dots such that $\gamma(p^{n_k}) < -\epsilon$ for $k = 1, 2, \dots$. But then $c(p^{n_k+1}) \leq c(p) - k\epsilon\nu$ for all k and therefore $c(p^{n_k+1}) < c(x^*)$ for $k > (c(p) - c(x^*))/\epsilon\nu$. Contradiction since x^* is a least cost point in \mathcal{P} . \square

B Technical issue in [22] and proposed fix

There is a technical issue with the results in [22]. In particular, no CILP defined in [22] satisfies both of that paper’s assumptions [22, (A3)] and [22, (A5)]. To see this, note that any doubly-infinite matrix A satisfying [22, (A3)] is a Hilbert-Schmidt operator, since the sums of the squares of the entries of any such A is finite.⁵ It is well known that a Hilbert-Schmidt operator is compact and no compact operator can be onto between two infinite-dimensional spaces (see, for instance, [37, Theorem 4.25] and its surrounding discussion), thus violating [22, (A5)].

There is a challenge in weakening [22, (A3)] to recover [22, (A5)]: it can easily compromise [22, Lemma 8.2], which argues that convergence of bases in the product discrete topology implies convergence in the operator norm. In order to recover [22, (A5)] by weakening [22, (A3)], one needs to strengthen the conditions leading up to [22, Lemma 8.2]. One combination that works is the following:

(A3′) The doubly infinite matrix A defines a bounded (continuous) linear operator from the Hilbert space H to the Hilbert space Y .

(A8′) The set of bases \mathcal{B} is closed in the product discrete topology (column-wise convergence). Furthermore, for any sequence of bases $\{B^n\} \subseteq \mathcal{B}$ converging to $B \in \mathcal{B}$ in this topology, the sequence of operators converges in the operator norm topology (i.e., $\|B^n - B\|_{op} \rightarrow 0$).

(A3′) is a weakening of [22, (A3)], but (A8′) is a strengthening of [22, (A8)]. One can check (A3′) is no longer inconsistent with [22, (A5)] and (A8′) asserts the structure that the old [22, (A3)] and [22, (A8)] was leveraged to deliver [22, (Lemma 8.2)]. Fortunately, the motivating network flow example from [38] in [22, Section 9] satisfies (A3′) and (A8′).

Acknowledgements

The authors would like to thank Google’s Gemini, xAI’s Grok, and OpenAI’s ChatGPT for technical assistance in the preparation of this manuscript, including assistance in verifying proofs and identifying grammar issues and typos.

⁵To see that an operator A satisfying [22, (A3)] is a Hilbert-Schmidt operator, let X be the Hilbert space of sequences $x = (x_j)_{j \geq 1}$ with norm $\|x\|_X^2 = \sum_{j \geq 1} \delta^{2j} |x_j|^2$, and let Y be defined analogously with weight β^i on the i th coordinate, where $0 < \beta, \delta < 1$. Define the isometric isomorphisms $T_\delta : X \rightarrow \ell^2$ and $T_\beta : Y \rightarrow \ell^2$ by $(T_\delta x)_j = \delta^j x_j$ and $(T_\beta y)_i = \beta^i y_i$. Conjugating the operator $A = (a_{ij})$ yields $A^0 := T_\beta A T_\delta^{-1} : \ell^2 \rightarrow \ell^2$ with matrix entries $m_{ij} = \beta^i a_{ij} / \delta^j$. Under [22, (A3)], there exist $\bar{a} > 0$ and $\alpha \in (0, \delta)$ such that $|a_{ij}| \leq \bar{a} \alpha^j$ for all i, j , and hence $|m_{ij}| \leq \bar{a} \beta^i (\alpha/\delta)^j$. It follows that

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |m_{ij}|^2 \leq \bar{a}^2 \left(\sum_{i=1}^{\infty} \beta^{2i} \right) \left(\sum_{j=1}^{\infty} (\alpha/\delta)^{2j} \right) < \infty,$$

since $\beta < 1$ and $\alpha < \delta$. Therefore, A^0 is a Hilbert-Schmidt operator on ℓ^2 , and by invariance of the Hilbert-Schmidt class under conjugation by isometries, A is Hilbert-Schmidt as an operator from X to Y .

References

- [1] Erik M Alfsen and Tore Nordseth. Vertices of choquet simplexes. *Mathematica Scandinavica*, 23(1): 171–176, 1969. [3](#), [16](#)
- [2] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker’s Guide*. Springer, 3rd edition, 2006. [3](#), [6](#), [13](#), [17](#), [18](#)
- [3] Dan Amir and Frank Deutsch. Suns, moons, and quasi-polyhedra. *Journal of Approximation Theory*, 6(2):176–201, 1972. [3](#), [16](#)
- [4] Edward J. Anderson and Peter Nash. *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications*. Wiley, 1987. [1](#), [4](#), [19](#)
- [5] E.J. Anderson and A.S. Lewis. An extension of the simplex algorithm for semi-infinite linear programming. *Math. Prog.*, 44(1-3):247–269, 1989. [1](#)
- [6] E.J. Anderson, M.A. Goberna, and M.A. López. Simplex-like trajectories on quasi-polyhedral sets. *Mathematics of Operations Research*, 26(1):147–162, 2001. [1](#)
- [7] Thomas Armstrong. Polyhedrality of infinite dimensional cubes. *Pacific Journal of Mathematics*, 70(2):297–307, 1977. [2](#), [3](#), [16](#)
- [8] Andrée Bastiani. Cones convexes et pyramides convexes. *Ann. Inst. Fourier*, 9:249–292, 1959. [3](#), [16](#)
- [9] Dimostri Bertsimas and John N. Tsitsiklis. *Introduction to Linear Optimization*. Athena, 1997. [1](#), [4](#)
- [10] Jonathan M Borwein, Jon D Vanderwerff, et al. *Convex functions: constructions, characterizations and counterexamples*, volume 109. Cambridge University Press Cambridge, 2010. [9](#)
- [11] Bruno Brosowski and Frank Deutsch. On some geometric properties of suns. *Journal of Approximation Theory*, 10(3):245–267, 1974. [3](#), [16](#)
- [12] Suresh Chand, Vernon Ning Hsu, and Suresh Sethi. Forecast, solution, and rolling horizons in operations management problems: A classified bibliography. *Manufacturing & Service Operations Management*, 4(1):25–43, 2002. [1](#)
- [13] Roland Durier and Christian Michelot. Sets of efficient points in a normed space. *Journal of Mathematical Analysis and Applications*, 117(2):506–528, 1986. [3](#), [16](#)
- [14] Roland Durier and Pier Luigi Papini. Polyhedral norms in an infinite dimensional space. *The Rocky Mountain Journal of Mathematics*, pages 863–875, 1993. [3](#)
- [15] Piotr Dworzak and Anton Kolotilin. The persuasion duality. *Theoretical Economics*, 19(4):1701–1755, 2024. [1](#)
- [16] Michael C Ferris and Andrew B Philpott. An interior point algorithm for semi-infinite linear programming. *Mathematical Programming*, 43:257–276, 1989. [1](#)
- [17] Vladimir P Fonf and Libor Veselý. Infinite-dimensional polyhedrality. *Canadian Journal of Mathematics*, 56(3):472–494, 2004. [3](#)
- [18] VP Fonf, J Lindenstrauss, and RR Phelps. Infinite dimensional convexity. In *Handbook of the Geometry of Banach Spaces*, volume 1, pages 599–670. Elsevier, 2001. [3](#)
- [19] Archis Ghate and Robert L. Smith. Characterizing extreme points as basic feasible solutions in infinite linear programs. *Operations Research Letters*, 37(1):7–10, 2009. [4](#)
- [20] Archis Ghate and Robert L. Smith. A linear programming approach to non stationary infinite-horizon Markov decision processes. *Operations Research*, 61:413–425, 2013. [1](#), [2](#)

- [21] Archis Ghate, Dushyant Sharma, and Robert L. Smith. A shadow simplex method for infinite linear programs. *Operations Research*, 58(4):865–877, 2010. 12
- [22] Archis Ghate, Christopher T Ryan, and Robert L Smith. A simplex method for countably infinite linear programs. *SIAM Journal on Optimization*, 31(4):3157–3183, 2021. 1, 2, 5, 22
- [23] Alan Gleit and Robert McGuigan. A note on polyhedral Banach spaces. *Proceedings of the American Mathematical Society*, pages 398–404, 1972. 3, 16
- [24] Wenxuan Guo, YoonHaeng Hur, Tengyuan Liang, and Christopher Thomas Ryan. Online learning to transport via the minimal selection principle. In *Conference on Learning Theory*, pages 4085–4109. PMLR, 2022. 1
- [25] Onesimo Hernández-Lerma and Jean B. Lasserre. *Discrete-time Markov Control Processes: Basic Optimality Criteria*, volume 30. Springer, 2012. 1
- [26] Richard B Holmes. *Geometric Functional Analysis and its Applications*, volume 24. Springer Science & Business Media, 2012. 3
- [27] Leonid Kantorovitch. On the translocation of masses. *Management Science*, 5(1):1–4, 1958. 1
- [28] VL Klee. Strict separation of convex sets. *Proceedings of the American Mathematical Society*, 7(4):735–737, 1956. 3, 6, 16, 19
- [29] Andreas Kleiner and Alejandro Manelli. Strong duality in monopoly pricing. *Econometrica*, 87(4):1391–1396, 2019. 1
- [30] Arkadii V Kryazhimskii and Andrzej Ruszczyński. Constraint aggregation in infinite-dimensional spaces and applications. *Mathematics of Operations Research*, 26(4):769–795, 2001. 1
- [31] Ka-sing Lau. Infinite dimensional polytopes. *Mathematica Scandinavica*, 32(2):193–213, 1974. 3, 16
- [32] PH Maserick. Convex polytopes in linear spaces. *Illinois Journal of Mathematics*, 9(4):623–635, 1965. 3, 16
- [33] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Mem. Math. Phys. Acad. Royale Sci.*, pages 666–704, 1781. 1
- [34] Robert R Phelps. Infinite dimensional compact convex polytopes. *Mathematica Scandinavica*, 24(1):5–26, 1969. 3, 16, 19
- [35] Malcolm C. Pullan. An algorithm for a class of continuous linear programs. *SIAM Journal on Control and Optimization*, 31(6):1558–1577, 1993. 1
- [36] M Rajagopalan and AK Roy. Maximal core representing measures and generalized polytopes. *The Quarterly Journal of Mathematics*, 25(1):257–271, 1974. 3, 16
- [37] Walter Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw–Hill, New York, 2 edition, 1991. ISBN 0-07-054236-8. 22
- [38] Christopher Thomas Ryan, Robert L Smith, and Marina A Epelman. A simplex method for uncapacitated pure-supply infinite network flow problems. *SIAM Journal on Optimization*, 28(3):2022–2048, 2018. 1, 2, 22
- [39] H.H. Schaefer. *Topological Vector Spaces*. Springer, 1971. 7
- [40] Thomas C. Sharkey and H. Edwin Romeijn. A simplex algorithm for minimum-cost network-flow problems in infinite networks. *Networks*, 52(1):14–31, 2008. 1, 2

- [41] Evgeny Shindin and Gideon Weiss. A simplex-type algorithm for continuous linear programs with constant coefficients. *Mathematical Programming*, 180:157–201, 2020. [1](#)
- [42] Robert L Smith and Rachel Q Zhang. Infinite horizon production planning in time-varying systems with convex production and inventory costs. *Management Science*, 44(9):1313–1320, 1998. [1](#)
- [43] Valeriu Soltan. *Lectures on Convex Sets*. World Scientific, 2019. [5](#)
- [44] Gideon Weiss. A simplex based algorithm to solve separated continuous linear programs. *Mathematical Programming*, 115(1):151–198, 2008. [1](#)