On the Shadow Simplex method for curved polyhedra

Daniel Dadush[1] and Nicolai Hähnle[2]

1   **Centrum Wiskunde & Informatica, Netherlands**
     `dadush@cwi.nl`
2   **Universität Bonn, Germany**
     `haehnle@or.uni-bonn.de`

──── **Abstract** ────

We study the simplex method over polyhedra satisfying certain "discrete curvature" lower bounds, which enforce that the boundary always meets vertices at sharp angles. Motivated by linear programs with totally unimodular constraint matrices, recent results of Bonifas et al (SOCG 2012), Brunsch and Röglin (ICALP 2013), and Eisenbrand and Vempala (2014) have improved our understanding of such polyhedra.

We develop a new type of *dual* analysis of the shadow simplex method which provides a clean and powerful tool for improving all previously mentioned results. Our methods are inspired by the recent work of Bonifas and the first named author [4], who analyzed a remarkably similar process as part of an algorithm for the Closest Vector Problem with Preprocessing.

For our first result, we obtain a constructive diameter bound of $O(\frac{n^2}{\delta} \ln \frac{n}{\delta})$ for n-dimensional polyhedra with curvature parameter $\delta \in [0,1]$. For the class of polyhedra arising from totally unimodular constraint matrices, this implies a bound of $O(n^3 \ln n)$. For linear optimization, given an initial feasible vertex, we show that an optimal vertex can be found using an expected $O(\frac{n^3}{\delta} \ln \frac{n}{\delta})$ simplex pivots, each requiring $O(mn)$ time to compute. An initial feasible solution can be found using $O(\frac{mn^3}{\delta} \ln \frac{n}{\delta})$ pivot steps.

1   Introduction

The *simplex method* is one of the most important methods for solving linear programs (LP), that is, optimization problems of the form $\max\{\langle \mathbf{c}, \mathbf{x}\rangle : \mathbf{x} \in P\}$ where P is a polyhedron defined by linear constraints. Starting from an initial vertex $\mathbf{v}$, a simplex algorithm provides a rule for moving from vertex to vertex along edges of the graph or 1-skeleton of P until an optimal vertex $\mathbf{w}$ (or an unbounded ray) is found.

A long standing open question is whether there exists a polynomial-time simplex algorithm for LP. The first obstacle in proving the existence (or non-existence) of such a method is the following fundamental question:

▶ **Question 1.** Given any two vertices $\mathbf{v}, \mathbf{w}$ of a polyhedron P, what is the best possible bound on the length of the shortest path between them, as a function of the dimension n and the number of constraints m?

The polynomial Hirsch conjecture posits that the diameter of the graph of a polyhedron is bounded by a polynomial in m and n. The best known general upper bounds are however

much larger. Barnette [3] and Larman [14] proved a bound of $O(2^n m)$, and Todd [20] recently proved a bound of $(m - n)^{\log n}$, slightly improving an earlier bound of Kalai and Kleitman [12, 13]. The original Hirsch conjecture, which posited a bound of $m - n$, was recently disproved for polytopes (i.e. bounded polyhedra) by Santos [17, 15], who gave a lower bound of $(1 + \varepsilon)m$ (only slightly violating the conjectured bound).

Given the difficulty of the general question, much research has been aimed at bounding the diameter of special classes of polyhedra. For example, polynomial bounds have been given for 0/1 polytopes [16], transportation polytopes [2, 7, 9], and flag polytopes [1].

Another important class, which has recently received much attention and is directly related to this work, are polyhedra whose constraint matrices are "well-conditioned". Dyer and Frieze [10] showed that the diameter of totally unimodular polyhedra – i.e. having integer constraint matrices with all subdeterminants in $\{0, \pm 1\}$ – is bounded by $O(m^{16} n^3 (\log mn)^3))$. Their work also contains a polynomial time randomized simplex algorithm that solves linear programs over totally unimodular polyhedra.

The diameter bound of Dyer and Frieze was both generalized and improved in the work of Bonifas et al [5]. They showed that polyhedra with integer constraint matrices and all subdeterminants bounded by Δ have diameter $O(\Delta^2 n^4 \log(n\Delta))$ if they are unbounded and $O(\Delta^2 n^{3.5} \log(n\Delta))$ if they are bounded. Their proof used certain expansion properties of the polyhedral graph and was non-constructive.

In an attempt to make the bound of [5] constructive, Brunsch and Röglin [8] showed that given any two vertices $\mathbf{v}, \mathbf{w}$ on such a polyhedron P, a path between them of length $O(m\Delta^4 n^4)$ (note the dependence on m) can be constructed using the *shadow simplex method*. In fact, they give a more general bound based on the so-called δ-distance property of the constraint matrix, which measures how "well spread" the rows of the constraint matrix are[1]. Using this parameter they give a bound of $O(mn^2/\delta^2)$ on the length of the constructed path, and recover the previous bound by the relationship $\delta \geq 1/(n\Delta^2)$.

Most recently, Eisenbrand and Vempala [11] provided a different approach to making the Bonifas et al [5] result constructive, which more closely resembles the random walk approach of Dyer and Frieze and also extends to optimization. When the constraint matrix satisfies the δ-distance property, they show that given an initial vertex and objective, an optimal vertex can be computed using $\text{poly}(n, 1/\delta)$ random walk steps (no dependence on m). Furthermore, an initial feasible vertex can be computed using m calls to their optimization algorithm over subsets of the original constraints.

2   Results

Building and improving upon the works of Bonifas et al [5], Brunsch and Röglin [8] and Vempala and Eisenbrand [11], we give an improved (constructive) diameter bound and simplex algorithm for polyhedra satisfying the δ-distance and other related properties. We also make improvements in the treatment of unbounded polyhedra and degeneracy. All our results are based on a new variant and analysis of the *shadow simplex method*.

We now introduce the "discrete curvature measures" we use along with the corresponding results. We list these measures in order of increasing strength. In the next section, we shall explain our variant of the shadow simplex method and compare it with previous implementations.

---

[1] We note that this measure is already implicit in [5] and that the diameter bound factors through it.

Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \le \mathbf{b}\}$, $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^n$ be a pointed polyhedron (A has full column rank $\Leftrightarrow$ P has vertices). For a vertex $\mathbf{v}$ of P, the *normal cone* at $\mathbf{v}$ is $N_{\mathbf{v}} = \left\{ \sum_{i \in I_{\mathbf{v}}} \lambda_i \mathbf{a}_i : \lambda_i \ge 0, i \in I_{\mathbf{v}} \right\}$, where $I_{\mathbf{v}} = \{i \in [m] : \langle \mathbf{a}_i, \mathbf{v} \rangle = b_i\}$ is the set of tight constraints. Equivalently, $N_{\mathbf{v}}$ is the set of all linear objective functions whose maximum over P is attained at $\mathbf{v}$. $N_{\mathbf{v}}$ is simplicial (non-degenerate) if it is generated by a basis of A, that is, if exactly n linearly independent constraints of P are tight at $\mathbf{v}$. The *normal fan* of P is the collection of all the vertex normal cones, and the *support of the normal fan $N(P)$* is their union. A polyhedron is simple (or non-degenerate) if all its vertex normal cones are simplicial.

▶ **Definition 2** (τ-wide Polyhedra). We say that a cone C is τ-wide if it contains a Euclidean ball of radius τ centered on the unit sphere. We define a polyhedron P to have a τ-wide normal fan (or simply P to be τ-wide) if every vertex normal cone is τ-wide.
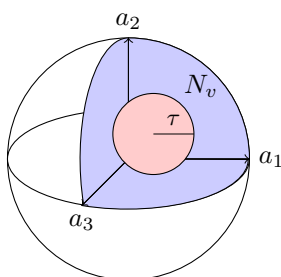
Roughly speaking, having a τ-wide normal fan enforces that facets always intersect at "sharp angles" (i.e. angle bounded away from π). In particular, for any vertex $\mathbf{v}$ of P, the angle between any two rays emanating from $\mathbf{v}$ and (non-trivially) passing through P is at most $\pi - 2\tau$. Hence one can interpret this condition as a discrete form of curvature for polyhedra. We now state our diameter bound for τ-wide polyhedra.

▶ **Theorem 3** (Diameter Bound, see Theorem 11). *Let $P \subseteq \mathbb{R}^n$ be an n-dimensional pointed polyhedron having a τ-wide normal fan. Then the graph of P has diameter bounded by $8n/\tau(1 + \ln(1/\tau))$. Furthermore, a path of this expected length can be constructed via the shadow simplex method.*

Restricting to n-dimensional polyhedra with subdeterminants bounded by Δ, using the relation $\tau \ge 1/(n\Delta)^2$ (see Lemma 35) we achieve a bound of $O(n^3 \Delta^2 \ln(n\Delta))$, improving on the existential bounds of Bonifas et al [5]. In contrast to [5], we note that our bound (and proof) is the same for polytopes and unbounded polyhedra.

While our bound is constructive – we follow a shadow simplex path – it is in general only efficiently implementable when the polyhedron is simple. In the presence of degeneracy, we note that computing a single edge of the path is essentially as hard as solving linear programming. Furthermore, standard techniques for removing degeneracy, such as the perturbation or lexicographic method, may unfortunately introduce a large number of extra simplex pivots.

Interestingly, our diameter bound can take advantage of degeneracy in situations where it makes the normal cones wider. While degeneracy does not occur for "generic polyhedra", it is very common for combinatorial polytopes. Furthermore, it can occur in ways that are useful to our diameter bound. For example, we remark that using degeneracy one can prove that the normal fan of the perfect matching polytope is $\Omega(1/\sqrt{|E|})$-wide, see Appendix 9.

To solve linear optimization problems via the shadow simplex method, we will need more than a wide normal fan. In fact, we will have different requirements for the two phases of

the simplex algorithm: Phase 1, which finds an initial feasible vertex, will require more than Phase 2, which finds an optimal vertex with respect to the objective starting from a feasible vertex.

▶ **Definition 4** (δ-distance property). A set of linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ satisfy the δ-distance property if for every $i \in [k]$, the vector $\mathbf{v}_i$ is at Euclidean distance at least $\delta\|\mathbf{v}_i\|$ from the span of $\{\mathbf{v}_j : j \in [k] \setminus \{i\}\}$.

For a polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$, we define P to satisfy the *local δ-distance property* if every *feasible basis* of A, i.e. the rows of A defining a vertex of P, satisfies the δ-distance property.

We say that a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ satisfy the *global δ-distance property* if every linearly independent subset satisfies the δ-distance property. We say that a matrix $A \in \mathbb{R}^{m \times n}$ satisfies the global δ-distance property if its row vectors do.

▶ **Lemma 5.** *Let* $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{S}^{n-1}$ *be a basis satisfying the δ-distance property. Then* $\mathrm{cone}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ *is δ/n-wide.*

Proof. See Appendix 10. ◀

The definitions differ in strength mainly based on the sets of bases to which they apply. The local δ-distance property is stronger than the τ-wide property for $\tau = \delta/n$, because it implies that *all triangulations* of the normal fan are τ-wide.[2] The global property is stronger than the local property since it applies also to infeasible bases, which allows one to control the geometry of polyhedra related to P, such as polyhedra obtained by removing a subset of constraints, which will be needed for Phase 1.

We now state our main result for Phase 2 simplex.

▶ **Theorem 6** (Optimization via Shadow Simplex, see Theorem 20). *Let* $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ *be an n-dimensional polytope with m constraints satisfying the local δ-distance property. Then, given an objective* $\mathbf{c} \in \mathbb{R}^n$ *and a vertex* $\mathbf{v}$ *of P, an optimal vertex can be computed using an expected* $O(n^3/\delta \ln(n/\delta))$ *shadow simplex pivots, where each pivot requires* $O(mn)$ *arithmetic operations.*

Our Phase 2 algorithm above is faster than the algorithms in [8, 11] and relies on a *weaker assumption* than [11]. The $\mathbf{v}, \mathbf{w}$ path finding algorithm of Brunsch and Röglin [8] is in fact a special case of the above, since we can choose $\mathbf{c}$ to be any objective maximized at $\mathbf{w}$. Comparing to the Phase 2 algorithm of Eisenbrand and Vempala [11], we require only the *local δ-distance property* instead of the global one. Whether one could rely only on the local property was left as open question in [11], which we resolve in the affirmative.

A small technical caveat is that as stated, the algorithm requires knowledge of δ. Since $\delta \leq 1$, we can always guess a number $\delta' \leq \delta \leq 2\delta'$ by trying $O(\ln 1/\delta)$ different values, incurring an $O(\ln 1/\delta)$ factor increase in running time (overestimating δ only affects correctness, not runtime). For simplicity, we shall henceforth assume that δ is known.

A more important caveat is that the above algorithm requires that P be a polytope (i.e. bounded). This restriction is due to the fact that we can only generate the randomness required for our bounds efficiently (that is, without solving a general LP) when the support of the normal fan equals $\mathbb{R}^n$.

The unbounded setting can be reduced to the bounded setting, in the standard way, by adding one or more constraints to make P bounded while not cutting off any of its vertices.

---

[2] However, the τ-wide property is weaker even when all normal cones are simplicial: a 2-dimensional cone of inner angle close to π is almost 1-wide, but satisfies δ-distance only for δ close to 0.

▶ **Definition 7.** Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ be a pointed polyhedron. Then a polytope $P' = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}, A'\mathbf{x} \leq \mathbf{b}'\}$ is *LP equivalent* to P if every vertex $\mathbf{v} \in P$ satisfies $\langle \mathbf{a}'_i, \mathbf{v} \rangle < \mathbf{b}'_i$ for all i; in particular, x is a vertex of P'.

Given an optimal vertex $\mathbf{v}$ of P' as above, one can easily check whether $\mathbf{v}$ is a vertex of P. If it is not, the original LP must be unbounded. In general, however, adding constraints to P happens at the expense of a degraded δ. In particular, the standard reduction of adding a large box constraint can degrade δ arbitrarily, hence the constraints must be added with care. We state the guarantees we can achieve below.

▶ **Lemma 8** (Removing Unboundedness, see Appendix 7). *Let* $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ *be an* n-*dimensional pointed polyhedron with* m *constraints. Let* $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *denote the rows of* A *and* $b_{\max} = \max_{i \in [m]} |\mathbf{b}_i| / \|\mathbf{a}_i\|$.

1. *Assume that* P *satisfies the* local δ-distance property *and that* $I \subseteq [m]$, $|I| = n$, *indexes the rows of a feasible basis. Letting* $\mathbf{w} = -1/n \sum_{i \in I} \mathbf{a}_i / \|\mathbf{a}_i\|$, *we have that*

$$P' = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}, \quad \langle \mathbf{w}, \mathbf{x} \rangle \leq nb_{\max}/\delta\} \ ,$$

   is a polytope that is LP equivalent to P *and satisfies the* local $\delta^2/(2n)$-distance property.
2. *Assume that* A *satisfies the* global δ-distance property. *Then*

$$P' = \{\mathbf{x} \in \mathbb{R}^n : -n\|\mathbf{a}_i\| b_{\max}/\delta - 1 \leq \langle \mathbf{a}_i, \mathbf{x} \rangle \leq \mathbf{b}_i, \quad \forall i \in [m]\} \ ,$$

   is a polytope that is LP equivalent to P *and satisfies the* global δ-distance *property.*

Finally, we use standard techniques for reducing feasibility to phase 2 type optimization. As this generally requires pivoting over infeasible bases, we will require global instead of local properties here. Interestingly, for LPs with bounded subdeterminants, we get that the number of simplex pivots is completely independent of the number of constraints.

▶ **Theorem 9** (Feasibility via Shadow Simplex, see Appendix 8). *Let* $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq b\}$ *be an* n-*dimensional polyhedron whose constraint matrix has full column rank and satisfies the global* δ-distance property. *Then a feasible solution to* P *can be computed using an expected* $O(mn^3/\delta \ln(n/\delta))$ *shadow simplex pivots. Furthermore, if* A *is integral and has subdeterminants bounded by* Δ, *a feasible solution can be computed using an expected* $O(n^5\Delta^2 \ln(n\Delta))$ *shadow simplex pivots.*

Shadow Simplex Method

Our main technical contribution is a new analysis and variant of the *shadow simplex method*, which utilizes (rather unexpectedly) an approach developed in [4] for navigating over the Voronoi graph of a Euclidean lattice (see related work section).

The shadow simplex has been at the heart of many theoretical attempts to explain the surprising efficiency of the simplex method in practice. It has been shown to give polynomial bounds for the simplex method over random and smoothed linear programs [6, 19, 21]. As mentioned above, Brunsch and Röglin [8] already showed that it yields short paths for the polyhedra we consider here.

At a high level, the shadow simplex over a polyhedron P works as follows. Given an initial objective function $\mathbf{c}$, a vertex $\mathbf{v}$ of P which maximizes this objective, and a target objective function $\mathbf{d}$, the shadow simplex interpolates between the objective functions $\mathbf{c}$ and $\mathbf{d}$ and performs a pivot step whenever the optimal vertex changes (hence the alternative

name *parametric* simplex method referring to the parameterization $\mathbf{c}(\lambda) = (1 - \lambda)\mathbf{c} + \lambda\mathbf{d}$ of the objective function, where λ grows from 0 to 1 over the course of the algorithm).

Traditionally, this method is understood and analyzed with a primal interpretation: The polyhedron P is orthogonally projected onto the 2-dimensional plane spanned by $\mathbf{c}$ and $\mathbf{d}$ (hence the term "shadow"), and the algorithm is understood in terms of the boundary of the projection P'. The optimal vertices for $\mathbf{c}$ and $\mathbf{d}$ project to the boundary of P', and as long as $\mathbf{c}$ and $\mathbf{d}$ are in sufficiently general position, edges of P' lift to edges of P so that the boundary can be followed efficiently by an algorithm that performs simplex pivots in the original space. The number of pivot steps is then typically bounded in terms of the lengths of edges or in terms of angles between edges of P'.

Our analysis is substantially different and based on a *dual* perspective: The shadow simplex method follows the line segment $[\mathbf{c}, \mathbf{d}]$ through the normal fan of P, pivoting whenever the segment crosses into a different n-dimensional normal cone. We express the number of crossings, that is, the number of intersections between $[\mathbf{c}, \mathbf{d}]$ and the facets of the normal fan of P, in terms of certain surface area measures of translates of the normal fan. The bounds we obtain on the number of intersections are stated below.

▶ **Theorem 10** (Intersection bounds, see Lemmas 22 and 25). *Let $\mathcal{T} = (C_1, \ldots, C_k)$ be a partition of a cone Σ into polyhedral τ-wide cones. Let $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ and let $X \in \mathbb{R}^n$ be exponentially distributed on Σ.*

1. The expected number of facets hit by the shifted line segment $[\mathbf{c} + X, \mathbf{d} + X]$ satisfies

$$\mathbb{E}[|\partial\mathcal{T} \cap [\mathbf{c} + X, \mathbf{d} + X]|] \leq \frac{\|\mathbf{d} - \mathbf{c}\|}{\tau} \ .$$

2. Let $\alpha \in (0, 1)$. Then

$$\mathbb{E}[|\partial\mathcal{T} \cap [\mathbf{c} + \alpha X, \mathbf{c} + X]|] \leq \frac{2n}{\tau} \ln \frac{1}{\alpha} \ .$$

To achieve the above bounds, the main idea is to relate the probability that the above random line segments pass through a normal cone to the probability that the associated perturbation vector lands in the cone (or some joint shift). Under the τ-wideness condition, we can in fact uniformly upper bound these proportionality factors. Since the jointly shifted normal cones are all disjoint, we can deduce the desired bounds from the fact that the sum of their measures is ≤ 1.

We compose these bounds in a way that also departs from the classical template by using three consecutive shadow simplex paths instead of just one. For given vertices $\mathbf{v}$ and $\mathbf{w}$ we first pick objectives $\mathbf{c}$ and $\mathbf{d}$ that are "deep inside" the respective normal cones. From here, we sample an exponentially distributed perturbation vector X and traverse three paths through the normal fan in sequence:

$$\mathbf{c} \quad \xrightarrow{(a)} \quad \mathbf{c} + X \quad \xrightarrow{(b)} \quad \mathbf{d} + X \quad \xrightarrow{(c)} \quad \mathbf{d}$$

The perturbation X will be quite large and hence almost always large enough to push $\mathbf{c}$ and $\mathbf{d}$ away from their normal cones. Indeed, the high level intuition behind our path is that in order to avoid unusually long paths from $\mathbf{c}$ to $\mathbf{d}$, we first travel to a "random intermediate location".

We note that in phases (a) and (c), randomness is only used to perturb one of the objectives. As far as we are aware, this paper provides the first successful analysis of the shadow simplex path in this setting. Furthermore, this extension is crucial to achieving our

improved diameter bound. Previous algorithms were constrained to random perturbations that kept $\mathbf{c}$ and $\mathbf{d}$ inside their respective normal cones, making the amount of randomness they could take advantage of much smaller.

We now use the bounds from Theorem 10 to derive the diameter bound.

▶ **Theorem 11.** *Let $P \subseteq \mathbb{R}^n$ be a pointed full dimensional polyhedron with τ-wide outer normal cones. Then P has diameter bounded by $\frac{8n}{\tau}(1 + \ln 1/\tau)$.*

Proof. Let $\mathbf{v}_1, \mathbf{v}_2$ be vertices of P with outer normal cones $N_{\mathbf{v}_1}, N_{\mathbf{v}_2}$. Let $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{S}^{n-1}$ satisfy $\mathbf{c}_i + \tau \mathcal{B}_2^n \subseteq N_{\mathbf{v}_i}$, $i \in \{1, 2\}$. Let $\Sigma = N(P)$ denote the support of the normal fan of P, and let X be exponentially distributed over Σ.

We will construct a path from $\mathbf{v}_1$ to $\mathbf{v}_2$ by following the sequence of vertices optimizing the objectives in the segments $[s\mathbf{c}_1, s\mathbf{c}_1 + X]$, $[s\mathbf{c}_1 + X, s\mathbf{c}_2 + X]$, $[s\mathbf{c}_2 + X, s\mathbf{c}_2]$, where $s > 0$ is a scalar to be chosen later. We will condition on the event that $\|X\| \le 2n$. Since $\mathbb{E}[\|X\|] = n$ (see Lemma 13), by Markov's inequality this occurs with probablity at least $1/2$. Under this event, by τ-wideness, we will not pivot in the segments $[s\mathbf{c}_1, s\mathbf{c}_1 + \frac{s\tau}{2n}X]$ and $[s\mathbf{c}_2 + \frac{s\tau}{2n}X, s\mathbf{c}_2]$. Using Theorem 10, the number of pivots along the segments $[s\mathbf{c}_1 + \frac{s\tau}{2n}X, s\mathbf{c}_1 + X]$, $[s\mathbf{c}_1 + X, s\mathbf{c}_2 + X]$, $[s\mathbf{c}_2 + X, s\mathbf{c}_2 + \frac{s\tau}{2n}X]$, is bounded by

$$\frac{\left(\frac{s\|\mathbf{c}_2 - \mathbf{c}_1\|}{\tau} + \frac{4n}{\tau}\ln\left(\frac{2n}{s\tau}\right)\right)}{\Pr[\|X\| \le 2n]} \le 2\left(\frac{s\|\mathbf{c}_2 - \mathbf{c}_1\|}{\tau} + \frac{4n}{\tau}\ln\left(\frac{2n}{s\tau}\right)\right).$$

Setting $s = \frac{4n}{\|\mathbf{c}_2 - \mathbf{c}_1\|}$, the above bound becomes

$$\frac{8n}{\tau}\left(1 + \ln\left(\frac{\|\mathbf{c}_2 - \mathbf{c}_1\|}{2\tau}\right)\right) \le \frac{8n}{\tau}\left(1 + \ln\frac{1}{\tau}\right), \quad \text{as needed.} \qquad \blacktriangleleft$$

Related Work

In a surprising connection, we borrow techniques developed in a recent work of Bonifas and the first named author [4] for a totally different purpose, namely, for solving the Closest Vector Problem with Preprocessing on Euclidean lattices. In [4], a 3-step "perturbed" line path was analyzed to navigate over the Voronoi graph of the lattice, where lattice points are connected if their associated Voronoi cells touch in a facet.

In the current work, we show a strikingly close analogy between analyzing the number of intersections of a random straight line path with a Voronoi tiling of space and the intersections of a shadow simplex path with the normal fan of a polyhedron. This unexpected connection makes us hopeful that these ideas may have even broader applicability.

Organization

In section 3, we regroup all the necessary notation and definitions. In section 4, we give our shadow simplex based optimization algorithm. In section 5, we prove the intersection bounds for our shadow simplex method. In section 6, we show how to implement shadow simplex pivots and how to deal with degeneracy. In section 7, we give our reductions from unbounded δ-wide LPs to bounded ones. In section 8, we give our shadow simplex based algorithms for LP feasibility. In section 9, we give lower bounds on the width of the normal fan of the perfect matching polytope. Missing proofs can be found in Appendix 10.

3   Notation and definitions

For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we let $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{y}_i$ denote their inner product. We let $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ denote the Euclidean norm, $\mathcal{B}_2^n = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq 1\}$ the unit ball and $\mathbb{S}^{n-1} = \partial \mathcal{B}_2^n$ the unit sphere. We denote the linear span of a set $A \subseteq \mathbb{R}^n$ by $\mathrm{span}(A)$. We use the notation $\mathrm{I}[\mathbf{x} \in A]$ for the indicator of A, that is $\mathrm{I}[\mathbf{x} \in A]$ is 1 if $\mathbf{x} \in A$ and 0 otherwise. For a set of scalars $S \subseteq \mathbb{R}$, we write $SA = \{s\mathbf{a} : s \in S, \mathbf{a} \in A\}$. For two sets $A, B \subseteq \mathbb{R}^n$, we define their Minkowski sum $A + B = \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}$. We let $d(A, B) = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x} \in A, \mathbf{y} \in B\}$, denote the Euclidean distance between A and B. For vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ we write $[\mathbf{a}, \mathbf{b}]$ for the closed line segment and $[\mathbf{a}, \mathbf{b})$ for the half-open line segment from $\mathbf{a}$ to $\mathbf{b}$.

▶ **Definition 12** (Cone). A cone $\Sigma \subseteq \mathbb{R}^n$ satisfies the following three properties:

- $\mathbf{0} \in \Sigma$.
- $\mathbf{x} + \mathbf{y} \in \Sigma$ if $\mathbf{x}$ and $\mathbf{y}$ are in Σ.
- $\lambda \mathbf{x} \in \Sigma$ if $\mathbf{x} \in \Sigma$ and $\lambda \geq 0$.

For vectors $\mathbf{y}_1, \ldots, \mathbf{y}_k \in \mathbb{R}^n$, we define the closed cone they generate as

$$\mathrm{cone}(\mathbf{y}_1, \ldots, \mathbf{y}_k) = \left\{ \sum_{i=1}^{m} \lambda_i \mathbf{y}_i : \lambda_i \geq 0, i \in [m] \right\}.$$

A cone is polyhedral if it can be generated by a finite number of vectors, and is simplicial if the generators are linearly independent. By convention, we let $\mathrm{cone}(\emptyset) = \mathbf{0}$. A simplicial cone has the δ-distance property if its extreme rays satisfy the δ-distance property.[3]

For a convex set $K \subseteq \mathbb{R}^n$, a subset $F \subseteq K$ is a face of K, if for all $\mathbf{x}, \mathbf{y} \in K$, $\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \in F$, $\lambda \in [0, 1]$, implies that $\mathbf{x}, \mathbf{y} \in F$. For a simplicial cone C, we note that its faces are exactly all the subcones generated by any subset of the generators of C.

A set of cones $\mathcal{T} = \{C_1, \ldots, C_k\}$ is an n-dimensional *cone partition* if:

- Each $C_i \subseteq \mathbb{R}^n$, $i \in [k]$, is a closed n-dimensional cone.
- Any two cones C_i, C_j, $i \neq j$, meet in a shared face.
- The *support* of $\mathcal{T}$, $\mathrm{sup}(\mathcal{T}) \stackrel{\mathrm{def}}{=} \cup_{i \in [k]} C_i$, is a closed cone.

We say that F is a face of $\mathcal{T}$ if it is a face of one of its contained cones. A cone partition $\mathcal{T}$ is τ-wide if every C_i is τ-wide. It is simplicial if every C_i is simplicial. In this case, we also call $\mathcal{T}$ a *cone triangulation*. A cone triangulation satisfies the local δ-distance property if every C_i satisfies it. We define the boundary of $\mathcal{T}$, $\partial \mathcal{T} = \cup_{i=1}^{k} \partial C_i$. We say that a cone triangulation $\mathcal{T}$ *triangulates* a cone partition $\mathcal{P}$ if $\mathcal{T}$ and $\mathcal{P}$ have the same support and every cone $C \in \mathcal{T}$ is generated by a subset of the extreme rays of some cone of $\mathcal{P}$. This means that $\mathcal{T}$ partitions ("refines") every cone of $\mathcal{P}$ into simplicial cones.

3.1   Exponential distribution

We say that a random variable $X \in \mathbb{R}^n$ is exponentially distributed on a cone Σ if

$$\Pr[X \in S] = \int_S \zeta_\Sigma(\mathbf{x}) \mathrm{d}\mathbf{x},$$

---

[3] The δ-distance property is invariant under scaling, so the choice of generators of the extreme rays is irrelevant.

where $\zeta_\Sigma(\mathbf{x}) = c_\Sigma e^{-\|\mathbf{x}\|}\mathrm{I}[\mathbf{x} \in \Sigma]$. A standard computation, which we include for completeness, yields the normalizing constant and the expected norm.

▶ **Lemma 13.** *The normalizing constant $c_\Sigma^{-1} = n!\mathrm{vol}_n(\mathcal{B}_2^n \cap \Sigma)$. For X exponentially distributed on Σ, we have that $\mathbb{E}[\|X\|] = n$.*

Proof. See Appendix 10.   ◀

4 Optimization

While bounding the number of intersections of line segments $[\mathbf{c}, \mathbf{d}]$ with the facets of the normal fan of $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ is sufficient to obtain existential bounds on the diameter of P, we also need to be able to efficiently compute the corresponding pivots to obtain efficient algorithms. The following summarizes the required results, the technical details of which are found in Appendix 6.

▶ **Theorem 14** (Shadow simplex, see Theorem 29). *Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ be pointed, $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$, and B an optimal basis for $\mathbf{c}$. If every intersection of $[\mathbf{c}, \mathbf{d}]$ with a facet F of a cone spanned by a feasible basis of P lies in the relative interior of F, the Shadow Simplex can be used to compute an optimal basis for $\mathbf{d}$ in $O(mn^2 + Nmn)$ arithmetic operations, where N is the number of intersections of $[\mathbf{c}, \mathbf{d}]$ with some triangulation $\mathcal{T}$ of the normal fan of P, where $\mathcal{T}$ contains the cone spanned by the initial basis B.*

As explained in Section 2, we want to follow segments $[\mathbf{c}, \mathbf{c}+X]$, $[\mathbf{c}+X, \mathbf{d}+X]$, $[\mathbf{d}+X, \mathbf{d}]$ in the normal fan. Our intersection bounds from Theorem 10 are not quite sufficient to bound the number of steps on the first and last segments entirely. This is easily dealt with for the first segment, because we can control the initial objective function $\mathbf{c}$ so that it lies deep in the initial normal cone.

For the final segment, we follow the approach of Eisenbrand and Vempala [11], who showed that if A satisfies the *global δ-distance property*, then an optimal facet for $\mathbf{d}$ can be derived from a basis that is optimal for some $\tilde{\mathbf{d}}$ with $\|\mathbf{d} - \tilde{\mathbf{d}}\| \leq \frac{\delta}{n}$. Recursion can then be used on a problem of reduced dimension to move from $\tilde{\mathbf{d}}$ to $\mathbf{d}$. We strengthen their result (thereby answering a question left open by [11]) and show that the *local δ-distance property* is sufficient to get the same result as long as $\|\mathbf{d} - \tilde{\mathbf{d}}\| \leq \frac{\delta}{n^2}$.[4]

▶ **Definition 15.** Let F be a face of a cone triangulation $\mathcal{T}$ and let $\mathbf{x}$ be a vector in the support of $\mathcal{T}$. Let $G = \mathrm{cone}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$, $\|\mathbf{x}_i\| = 1$, be the minimal face of $\mathcal{T}$ that contains $\mathbf{x}$ and consider the unique conic combination $\mathbf{x} = \lambda_1\mathbf{x}_1 + \cdots + \lambda_k\mathbf{x}_k$. We define

$$\alpha_F(\mathbf{x}) := \sum_{i:\mathbf{x}_i \notin F} \lambda_i$$

In particular, $\alpha_F(\mathbf{x}) \geq 1$ if $\mathbf{x}$ is a unit vector and the minimal face containing it is disjoint from F, and $\alpha_F(\mathbf{x}) = 0$ if $\mathbf{x} \in F$.

▶ **Lemma 16.** *Let $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{S}^{n-1}$ be a set of vectors. Then the following are equivalent:*

1. $\mathbf{x}_1, \ldots, \mathbf{x}_m$ satisfy the δ-distance property.

---

[4] In the final bound, the loss of a factor n here disappears inside a logarithm.

2. $\forall\ I \subseteq [m]$ *for which* $\{\mathbf{x}_i : i \in I\}$ *are linearly independent and* $\forall\ (a_i \in \mathbb{R} : i \in I)$

$$\|\sum_{i \in I} a_i \mathbf{x}_i\| \geq \delta \max_{i \in I} |a_i| \ .$$

Proof. See Appendix 10. ◀

▶ **Lemma 17.** *Let F be a cone of an n-dimensional cone triangulation $\mathcal{T}$ satisfying the local δ-distance property. Let $\mathbf{x}$ be a point in the support of $\mathcal{T}$. Then $d(\mathbf{x}, F) \geq \alpha_F(\mathbf{x}) \cdot \frac{\delta}{n}$.*

Proof. Let $\mathbf{y} \in F$ be the (unique) point with $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, F)$. Note that by convexity, the segment $[\mathbf{x}, \mathbf{y}]$ is contained in the support of $\mathcal{T}$. By considering the cones of $\mathcal{T}$ that contain points on the segment $[\mathbf{x}, \mathbf{y}]$, we obtain a sequence of points

$$\mathbf{x} = \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_r = \mathbf{y}$$

on the segment $[\mathbf{x}, \mathbf{y}]$ and (full-dimensional) cones $G_1, \dots, G_r$ such that

$$G_i \cap [\mathbf{x}, \mathbf{y}] = [\mathbf{x}_{i-1}, \mathbf{x}_i].$$

Since $\alpha_F(\mathbf{y}) = 0$, the result of the lemma follows immediately from the claim that

$$d(\mathbf{x}_{i-1}, \mathbf{x}_i) \geq |\alpha_F(\mathbf{x}_{i-1}) - \alpha_F(\mathbf{x}_i)| \cdot \frac{\delta}{n},$$

which we will now prove.

Fix some $G_i = \text{cone}(\mathbf{y}_1, \dots, \mathbf{y}_n)$. By relabelling, we may assume that $\text{cone}(\mathbf{y}_1, \dots, \mathbf{y}_k) = G_i \cap F$ (since G_i and F are both faces of $\mathcal{T}$), for some $0 \leq k \leq n$ (if $k = 0$ then $G_i \cap F = \mathbf{0}$).

For every $\mathbf{z} \in G_i$, the minimal cone containing $\mathbf{z}$ is a face of G_i. Therefore, using the unique conic combination $\mathbf{z} = \sum_{i=1}^n \lambda_i \mathbf{y}_i$, we have that $\alpha_F(\mathbf{z}) = \sum_{k<i\leq n} \lambda_i$.

Writing $\mathbf{x}_{i-1} = \sum_{i=1}^n a_i \mathbf{y}_i$ and $\mathbf{x}_i = \sum_{i=1}^n b_i \mathbf{y}_i$, by Lemma 16 we have that

$$d(\mathbf{x}_{i-1}, \mathbf{x}_i) \geq \delta \max_{1 \leq i \leq n} |a_i - b_i| \geq \delta \max_{k < i \leq n} |a_i - b_i| \geq \frac{\delta}{n} \sum_{k < i \leq n} |a_i - b_i|$$

$$\geq \frac{\delta}{n} |\alpha_F(\mathbf{x}_{i-1}) - \alpha_F(\mathbf{x}_i)|,$$

which completes the proof of the claim. ◀

▶ **Lemma 18.** *Let F be a cone of a triangulation $\mathcal{T}$ satisfying the local δ-distance property and let $\mathbf{x}$ be a point in the support of $\mathcal{T}$ with $d(\mathbf{x}, F) \leq \frac{\delta}{n^2}$. Let $G = \text{cone}(\mathbf{x}_1, \dots, \mathbf{x}_n)$, $\|\mathbf{x}_i\| = 1$, be a cone of $\mathcal{T}$ containing $\mathbf{x}$ and let*

$$\mathbf{x} = \lambda_1 \mathbf{x}_1 + \cdots + \lambda_n \mathbf{x}_n$$

be the corresponding conic combination. Then for every $i \in [n]$ with $\lambda_i > \frac{1}{n}$ one has $\mathbf{x}_i \in F$.

Proof. Suppose there is some i with $\lambda_i > \frac{1}{n}$ and $\mathbf{x}_i \notin F$. Then $\alpha_F(\mathbf{x}) > \frac{1}{n}$ and by Lemma 17 we get $d(\mathbf{x}, F) > \frac{\delta}{n^2}$, which is a contradiction. ◀

For the recursion on a facet, we let $\pi_i(\mathbf{x}) := \mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{a}_i \rangle}{\langle \mathbf{a}_i, \mathbf{a}_i \rangle} \mathbf{a}_i$ be the orthogonal projection onto the subspace orthogonal to $\mathbf{a}_i$ and we let F_i be the facet of P defined by $\langle \mathbf{a}_i, \mathbf{x} \rangle = b_i$.

▶ **Lemma 19.** *Let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ be linearly independent vectors that satisfy the δ-distance property and let π be the orthogonal projection onto the subspace orthogonal to $\mathbf{v}_k$. Then $\pi(\mathbf{v}_1), \dots, \pi(\mathbf{v}_{k-1})$ satisfy the δ-distance property.*

Proof. See Appendix 10. ◀

This Lemma, which was already used by [11], implies that if P satisfies the local δ-distance property then so does F_i, where the definition of local δ-distance is understood relative to the affine hull of F_i,[5] because the normal vectors of F_i arise from orthogonal projections of the normal vectors of P.

Input: polytope $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$, $\delta > 0$, feasible basis B, $\mathbf{d} \in \mathbb{R}^n$
Output: optimal basis $B \subset [m]$ for $\mathbf{d}$
$\mathbf{c} \leftarrow \sum_{i \in B} \frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}$, $\mathbf{d} \leftarrow 2\frac{\mathbf{d}}{\|\mathbf{d}\|}$
Sample $X \in \mathbb{R}^n$ from the exponential distribution conditioned on $\|X\| \leq 2n$
Follow segments $[\mathbf{c}, \mathbf{c} + X]$, $[\mathbf{c} + X, \mathbf{d} + X]$, $[\mathbf{d} + X, \mathbf{d} + \frac{\delta}{2n^3}X]$ using Shadow Simplex
Find λ_i such that $\tilde{\mathbf{d}} := \mathbf{d} + \frac{\delta}{2n^3}X = \sum_{i \in B} \lambda_i \frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}$ where B is the current basis
Choose $i^\star$ such that $\lambda_{i^\star} > \frac{1}{n}$
$B' \leftarrow$ optimal basis of $F_{i^\star}$ for $\pi_{i^\star}(\mathbf{d})$, obtained by recursion starting at $B \setminus \{i^\star\}$
return $B' \cup \{i^\star\}$

Algorithm 1: Optimization

▶ **Theorem 20.** *If P satisfies the local δ-distance property, then Algorithm 1 correctly computes an optimal basis for $\mathbf{d}$ using an expected $O(n^3/\delta \ln(n/\delta))$ shadow simplex pivots.*

Proof. For correctness, let $\mathcal{T}$ be some triangulation of the normal fan of P and let C be a cone in $\mathcal{T}$ that contains $\mathbf{d}$. We have $\|\frac{\delta}{2n^3}X\| \leq \frac{\delta}{n^2}$ and therefore $d(\tilde{\mathbf{d}}, C) \leq d(\tilde{\mathbf{d}}, \mathbf{d}) \leq \frac{\delta}{n^2}$. Furthermore, $\|\tilde{\mathbf{d}}\| \geq \|\mathbf{d}\| - \frac{\delta}{n^2} > 1$ implies that $\sum_{i \in B} \lambda_i > 1$ so that there is some i with $\lambda_i > \frac{1}{n}$. Applying Lemma 18 yields that $\mathbf{a}_{i^\star}$ is a generator of C, which means that $i^\star$ is contained in some optimal basis for $\mathbf{d}$. This implies that recursion on $F_{i^\star}$ yields the correct result.

In order to bound the number of pivots, let C be the cone of the initial basis and observe that $\mathbf{c} + \delta\mathcal{B}_2^n \subseteq C$ by the proof of Lemma 5. Hence the segment $[\mathbf{c} + \frac{\delta}{2n}X)$ does not cross a facet of the triangulation $\mathcal{T}_1$ of the normal fan that is implicitly used by the first leg of the shadow simplex path.

If X were exponentially distributed (without the conditioning on $\|X\| \leq 2n$), Theorem 10 would bound the expected number of pivot steps along the three segments by

$$\mathbb{E}[N] \leq \frac{2n^2}{\delta} \ln \frac{2n}{\delta} + \frac{n\|\mathbf{d} - \mathbf{c}\|}{\delta} + \frac{2n^2}{\delta} \ln \frac{2n^3}{\delta} \leq O(\frac{n^2}{\delta} \ln(\frac{n}{\delta}))$$

Since $\mathbb{E}[\|X\|] = n$ we have $\Pr[\|X\| \leq 2n] \geq \frac{1}{2}$ by Markov's inequality and therefore

$$\mathbb{E}[N \mid \|X\| \leq 2n] \leq 2\,\mathbb{E}[N] \leq O(\frac{n^2}{\delta} \ln(\frac{n}{\delta})).$$

The bound on the total expected number of pivot steps follows from the depth n of recursion. ◀

---

[5] Alternatively, one can apply a rotation and translation so that F_i lies in the subspace $\mathbb{R}^{n-1}$ spanned by the first $n-1$ coordinates. The rotation does not affect the δ-distance property, and we can then treat F_i as a polytope in $\mathbb{R}^{n-1}$.

5   Intersection Bounds and Diameter Bounds

▶ **Lemma 21.** *Let C be a polyhedral cone containing $\mathbf{u} + \tau \mathcal{B}_2^n$, where $\|\mathbf{u}\| = 1$. Let $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ and let $X \in \mathbb{R}^n$ be exponentially distributed on a full dimensional cone $\Sigma \ni \mathbf{u}$. Then the expected number of times the shifted line segment $[\mathbf{c} + X, \mathbf{d} + X]$ hits the boundary of C is at most*

$$\mathbb{E}[|\partial C \cap [\mathbf{c} + X, \mathbf{d} + X]|] \leq \frac{\|\mathbf{d} - \mathbf{c}\|}{\tau} \int_0^1 \int_{(C - ((1-\lambda)c + \lambda d)) \cap \Sigma} \zeta_\Sigma(\mathbf{x}) \mathrm{d}\mathbf{x} \mathrm{d}\lambda$$

Proof. Let F be a facet of C. Note that with probability 1, the line segment $[\mathbf{c} + X, \mathbf{d} + X]$ passes through F at most once. By linearity, we see that

$$\mathbb{E}[|\partial C \cap [\mathbf{c} + X, \mathbf{d} + X]|] = \sum_{F \text{ facet of } C} \Pr[(F \cap [\mathbf{c} + X, \mathbf{d} + X]) \neq \emptyset]. \tag{1}$$

We now bound the crossing probability for any facet F.

We first calculate the hitting probability as

$$\Pr[F \cap [\mathbf{c} + X, \mathbf{d} + X] \neq \emptyset] = \Pr[X \in -[\mathbf{c}, \mathbf{d}] + F]$$

$$= \int_{-[\mathbf{c}, \mathbf{d}] + F} \zeta_\Sigma(\mathbf{x}) \mathrm{d}\mathbf{x}$$

$$= |\langle \mathbf{n}, \mathbf{d} - \mathbf{c} \rangle| \int_0^1 \int_{F - ((1-\lambda)\mathbf{c} + \lambda \mathbf{d})} \zeta_\Sigma(\mathbf{x}) \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}\lambda$$

$$\leq \|\mathbf{d} - \mathbf{c}\| \int_0^1 \int_{(F - ((1-\lambda)\mathbf{c} + \lambda \mathbf{d})) \cap \Sigma} c_\Sigma e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}\lambda \tag{2}$$

where $\mathbf{n} \in \mathbb{R}^n$ is a unit normal vector to F. Bounding the hitting probability therefore boils down to bounding the measure of a shift of the facet F. Letting $h = |\langle \mathbf{n}, \mathbf{u} \rangle| \geq \tau$ (which holds by assumption on $\mathbf{u}$), for any shift $\mathbf{t} \in \mathbb{R}^n$ we have that

$$\int_{(F + \mathbf{t} + \mathrm{cone}(\mathbf{u})) \cap \Sigma} e^{-\|\mathbf{x}\|} \mathrm{d}\mathbf{x} \geq \int_{((F + \mathbf{t}) \cap \Sigma) + \mathrm{cone}(\mathbf{u})} e^{-\|\mathbf{x}\|} \mathrm{d}\mathbf{x} \quad (\text{ since } \mathbf{u} \in \Sigma)$$

$$= \int_0^\infty \int_{((F + \mathbf{t}) \cap \Sigma) + \frac{r}{h} \mathbf{u}} e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}r$$

$$= \int_0^\infty \int_{(F + \mathbf{t}) \cap \Sigma} e^{-\|\mathbf{x} + \frac{r}{h} \mathbf{u}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}r$$

$$\geq \int_0^\infty e^{-r/h} \mathrm{d}r \int_{(F + \mathbf{t}) \cap \Sigma} e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x})$$

$$\geq \tau \int_{(F + \mathbf{t}) \cap \Sigma} e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \tag{3}$$

The lemma now follows by combining (1),(2),(3), using the fact that the $F + \mathrm{cone}(\mathbf{u})$ partition the cone C up to sets of measure 0.   ◀

▶ **Lemma 22.** *Let $\mathcal{T} = (C_1, \ldots, C_k)$ be a partition of a cone Σ into polyhedral τ-wide cones. Let $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ and let $X \in \mathbb{R}^n$ be exponentially distributed on Σ. Then the expected number of facets hit by the shifted line segment $[\mathbf{c} + X, \mathbf{d} + X]$ satisfies*

$$\mathbb{E}[|\partial \mathcal{T} \cap [\mathbf{c} + X, \mathbf{d} + X]|] \leq \frac{\|\mathbf{d} - \mathbf{c}\|}{\tau} \ .$$

Proof. Using Lemma 21, we bound

$$\mathbb{E}[|\partial\mathcal{T} \cap [\mathbf{c} + X, \mathbf{d} + X]|] \leq \sum_{i=1}^{k} \mathbb{E}[|\partial C_i \cap [\mathbf{c} + X, \mathbf{d} + X]|]$$

$$\leq \sum_{i=1}^{k} \frac{\|\mathbf{d} - \mathbf{c}\|}{\tau} \int_0^1 \int_{(C_i - ((1-\lambda)\mathbf{c} + \lambda\mathbf{d})) \cap \Sigma} \zeta_\Sigma(\mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}\lambda$$

$$\leq \frac{\|\mathbf{d} - \mathbf{c}\|}{\tau} \int_0^1 \int_\Sigma \zeta_\Sigma(\mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}\lambda$$

$$\leq \frac{\|\mathbf{d} - \mathbf{c}\|}{\tau},$$

as needed.

For the furthermore, note that each intersection is overcounted twice in the summation above, since each facet belongs to exactly two cones in the partition. ◀

We will need the following simple lemma about the exponential distribution.

▶ **Lemma 23.** *Let Y be exponentially distributed on $\mathbb{R}_+$. Then for any $c \in \mathbb{R}$, $\mathbb{E}[|Y - c|] \geq |c|/2$.*

Proof. Since $Y \geq 0$, the inequality is trivial if $c \leq 0$. Hence we may assume that $c \geq 0$. Using integration by parts, we have that

$$\mathbb{E}[|Y - c|] = \int_0^c (c - x)e^{-x}\mathrm{d}x + \int_c^\infty (x - c)e^{-x}\mathrm{d}x$$

$$= (x - c)e^{-x}\big|_0^c - \int_0^c e^{-x}\mathrm{d}x + (c - x)e^{-x}\big|_c^\infty + \int_c^\infty e^{-x}\mathrm{d}x$$

$$= (x - c + 1)e^{-x}\big|_0^c + (c - x - 1)e^{-x}\big|_c^\infty = 2e^{-c} + c - 1.$$

We wish to show that $2e^{-c} + c - 1 \geq c/2$, hence it suffices to show $2e^{-c} + c/2 - 1 \geq 0$ for all $c \geq 0$. This function is minimized at $c = \ln 4$ where it achieves value $(\ln 4 - 1)/2 > 0$. ◀

While we could choose $\mathbf{c}$ and $\mathbf{d}$ such that $\mathbf{c} + X$ and $\mathbf{d} + X$ lie in the same cone with high probability, this would require us to choose $\|\mathbf{d} - \mathbf{c}\|$ quite large. Instead, we will bound the number of facets that are hit by the segment $[\mathbf{c}, \mathbf{c} + X]$.

▶ **Lemma 24.** *Let $C \subseteq \mathbb{R}^n$ be a polyhedral cone containing $\mathbf{u} + \tau\mathcal{B}_2^n$, where $\|\mathbf{u}\| = 1$. Let $\mathbf{c} \in \mathbb{R}^n$ and $X \in \mathbb{R}^n$ be exponentially distributed on a cone $\Sigma \ni \mathbf{u}$. Then for every $\alpha \in (0, 1)$ we have*

$$\mathbb{E}[|\partial C \cap [\mathbf{c} + \alpha X, \mathbf{c} + X]|] \leq \frac{2}{\tau} \int_1^{1/\alpha} \frac{1}{s} \int_{(C - s\mathbf{c}) \cap \Sigma} \|\mathbf{x}\|\zeta_\Sigma(\mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}s$$

Proof. As in the proof of Lemma 21, we will decompose the expectation over the facets of C, where we have

$$\mathbb{E}[|\partial C \cap [\mathbf{c} + \alpha X, \mathbf{c} + X]|] = \sum_{F \text{ facet of } C} \Pr[F \cap [\mathbf{c} + \alpha X, \mathbf{c} + X] \neq \emptyset] \qquad (4)$$

Take a facet F of C and let $\mathbf{n}$ denote a unit normal to F pointing in the direction of the cone (i.e., $\langle n, u \rangle > 0$).

$$\Pr[F \cap [\mathbf{c} + \alpha X, \mathbf{c} + X] \neq \emptyset] = \Pr[X \in [1, \frac{1}{\alpha}](F - \mathbf{c})]$$

$$= \int_1^{1/\alpha} \int_{(F-s\mathbf{c})\cap\Sigma} |\langle \mathbf{n}, \mathbf{c} \rangle| \zeta_\Sigma(\mathbf{x}) \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}s$$

$$= \int_1^{1/\alpha} \frac{1}{s} \int_{(F-s\mathbf{c})\cap\Sigma} |\langle \mathbf{n}, s\mathbf{c} \rangle| c_\Sigma e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}s. \qquad (5)$$

Again, we have to bound an integral over a shifted facet, similar to the proof of Lemma 21. Letting $h = |\langle \mathbf{n}, \mathbf{u} \rangle| \geq \tau$, we have that

$$\int_{(F+\mathbf{t}+\mathrm{cone}(\mathbf{u}))\cap\Sigma} \|\mathbf{x}\| e^{-\|\mathbf{x}\|} \mathrm{d}\mathbf{x} \geq \int_{((F+\mathbf{t})\cap\Sigma)+\mathrm{cone}(\mathbf{u})} \|\mathbf{x}\| e^{-\|\mathbf{x}\|} \mathrm{d}\mathbf{x} \quad (\text{since } \mathbf{u} \in \Sigma)$$

$$= \int_0^\infty \int_{((F+\mathbf{t})\cap\Sigma)+\frac{r}{h}\mathbf{u}} \|\mathbf{x}\| e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}r$$

$$= \int_0^\infty \int_{(F+\mathbf{t})\cap\Sigma} \|\mathbf{x} + \frac{r}{h}\mathbf{u}\| e^{-\|\mathbf{x}+\frac{r}{h}\mathbf{u}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}r$$

$$\geq \int_0^\infty \int_{(F+\mathbf{t})\cap\Sigma} |\langle \mathbf{n}, \mathbf{x} + \frac{r}{h}\mathbf{u} \rangle| e^{-r/h} e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \mathrm{d}r$$

$$= h^2 \int_0^\infty |\langle \mathbf{n}, \mathbf{t} \rangle / h + s| e^{-s} \mathrm{d}s \int_{(F+\mathbf{t})\cap\Sigma} e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x})$$

$$\geq \frac{h}{2} \int_{(F+\mathbf{t})\cap\Sigma} |\langle \mathbf{n}, \mathbf{t} \rangle| e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x}) \quad (\text{by Lemma 23})$$

$$\geq \frac{\tau}{2} \int_{(F+\mathbf{t})\cap\Sigma} |\langle \mathbf{n}, \mathbf{t} \rangle| e^{-\|\mathbf{x}\|} \mathrm{dvol}_{n-1}(\mathbf{x})$$

$$(6)$$

The Lemma now follows by combining (4),(5),(6). ◀

▶ **Lemma 25.** *Let $\mathcal{T} = (C_1, \ldots, C_k)$ be partition of a cone Σ into polyhedral τ-wide cones. Let $\mathbf{c} \in \mathbb{R}^n$ and $\alpha \in (0,1)$ be fixed and let $X \in \mathbb{R}^n$ be exponentially distributed over Σ. Then*

$$\mathbb{E}[|\partial\mathcal{T} \cap [\mathbf{c} + \alpha X, \mathbf{c} + X]|] \leq \frac{2n}{\tau} \ln \frac{1}{\alpha} .$$

Proof. By Lemmas 13 and 24, we have that

$$\mathbb{E}[|\partial\mathcal{T} \cap [\mathbf{c} + \alpha X, \mathbf{c} + X]|] \leq \sum_{i=1}^k \mathbb{E}[|\partial C_i \cap [\mathbf{c} + \alpha X, \mathbf{c} + X]|]$$

$$\leq \frac{2}{\tau} \sum_{i=1}^k \int_1^{1/\alpha} \frac{1}{s} \int_{(C_i-s\mathbf{c})\cap\Sigma} \|\mathbf{x}\| \zeta_\Sigma(\mathbf{x}) \mathrm{d}\mathbf{x} \mathrm{d}s$$

$$\leq \frac{2}{\tau} \int_1^{1/\alpha} \frac{1}{s} \int_\Sigma \|\mathbf{x}\| \zeta_\Sigma(\mathbf{x}) \mathrm{d}\mathbf{x} \mathrm{d}s$$

$$\leq \frac{2}{\tau} \int_1^{1/\alpha} \frac{1}{s} \mathbb{E}[\|X\|] \mathrm{d}s = \frac{2n}{\tau} \ln \frac{1}{\alpha} \qquad ◀$$

6   The shadow simplex method with symbolic perturbation

In this section, we will give a self-contained presentation of the algorithmic details of the shadow simplex method, including the details of coping with degenerate P using a perturbation of the right-hand sides $\mathbf{b}$. For clarity of presentation, we will first consider the case of simple polyhedra. We will also assume that $\mathbf{c}$ and $\mathbf{d}$ are in general position as made precise in the precondition of Algorithm 2.

Input: $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$, $\mathbf{c}, \mathbf{d} \in N(P)$, optimal basis $B \subset [m]$ for $\mathbf{c}$
Precondition: P is pointed and simple
Precondition: $[\mathbf{c}, \mathbf{d})$ intersects facets of normal cones only in their relative interiors
Output: optimal basis $B \subset [m]$ for $\mathbf{d}$
$\lambda \leftarrow 0$
Gauss elimination: $A \leftarrow AU$, $\mathbf{c} \leftarrow U^T\mathbf{c}$, $\mathbf{d} \leftarrow U^T\mathbf{d}$ so that $A_B = I_n$
loop
> $i^\star \leftarrow \arg\min \left\{ \frac{\mathbf{c}_i}{\mathbf{c}_i - \mathbf{d}_i} : i \in B, \mathbf{c}_i > \mathbf{d}_i \right\}$
> **if** $i^\star$ *undefined or* $\lambda^\star = \frac{\mathbf{c}_{i^\star}}{\mathbf{c}_{i^\star} - \mathbf{d}_{i^\star}} \geq 1$ **then return** B
> $j^\star \leftarrow \arg\min \left\{ \frac{\langle \mathbf{a}_j, \mathbf{b}_B \rangle - \mathbf{b}_j}{\mathbf{a}_{ji^\star}} : j \notin B, \mathbf{a}_{ji^\star} < 0 \right\}$
> $B \leftarrow B \setminus \{B_{i^\star}\} \cup \{j^\star\}$, $\lambda \leftarrow \lambda^\star$
> Gauss elimination: $A \leftarrow AU$, $\mathbf{c} \leftarrow U^T\mathbf{c}$, $\mathbf{d} \leftarrow U^T\mathbf{d}$ so that $A_B = I_n$
end

Algorithm 2: Shadow Simplex

▶ **Lemma 26.** *Let $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $U \in \mathbb{R}^{n \times n}$ invertible. Then a basis B is optimal for $\max \{\langle \mathbf{c}, \mathbf{x} \rangle : A\mathbf{x} \leq \mathbf{b}\}$ if and only if it is optimal for $\max \{\langle U^T\mathbf{c}, \mathbf{x} \rangle : AU\mathbf{x} \leq \mathbf{b}\}$.*

Proof. Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ be the rows of A. The basis B is optimal for the first problem if and only if $\mathbf{c} \in \text{cone} \{\mathbf{a}_i : i \in B\}$. This is equivalent to $U^T\mathbf{c} \in \text{cone} \{U^T\mathbf{a}_i : i \in B\}$. Since the $U^T\mathbf{a}_i$ are the rows of AU, this is equivalent to B being an optimal basis for the second problem. ◀

▶ **Theorem 27.** *Algorithm 2 is correct as specified and requires $O(mn^2 + Nmn)$ arithmetic operations, where N is the number of normal cone facets intersected by $[\mathbf{a}, \mathbf{b}]$.*

Proof. The initial Gauss elimination requires $O(mn^2)$ arithmetic operations. Each iteration is dominated by the computation of $j^\star$ and the rank-1 Gauss elimination update, both of which require $O(mn)$ arithmetic operations.

We will show the invariant that B is an optimal basis for $\max \{\langle \mathbf{c}_\lambda, \mathbf{x} \rangle : A\mathbf{x} \leq \mathbf{u}\}$, where $\mathbf{c}_\lambda = (1 - \lambda)\mathbf{c} + \lambda\mathbf{d}$. The invariant initially holds by definition of the input and remains unchanged by the Gauss elimination steps due to Lemma 26.

A typical iteration is illustrated in Figure 1. The vertex corresponding to B is $\mathbf{b}_B$, its normal cone is the positive orthant $\mathbb{R}^n_+$. The invariant implies that $\mathbf{c}_\lambda \in \mathbb{R}^n_+$. If $\mathbf{d} - \mathbf{c} \geq 0$, the algorithm returns B. Indeed, B is an optimal basis for $\mathbf{d}$ in this case because $\mathbf{d} \in \mathbb{R}^n_+$.

Otherwise, the ray from $\mathbf{c}_\lambda$ through $\mathbf{d}$ eventually leaves the positive orthant, and in fact $\mathbf{c}_{\lambda^\star}$ is the last point contained in $\mathbb{R}^n_+$. If $\lambda^\star \geq 1$ we know that $\mathbf{d} = \mathbf{c}_1 \in \mathbb{R}^n_+$ and so the current basis is optimal for $\mathbf{d}$. The index $i^\star$ is the index into the basis whose contribution to
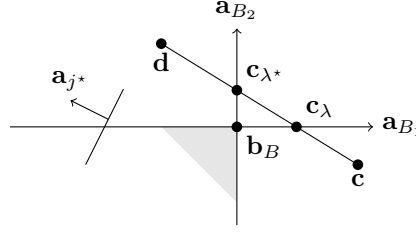
Figure 1 One iteration of the Shadow Simplex

the conic combination describing $\mathbf{c}_{\lambda^\star}$ is 0.[6]

Letting $B' = B \setminus \{B_{i^\star}\} \cup \{j^\star\}$, where $B_{i^\star}$ is the $i^\star$-th index in the basis, it is trivially true that $\mathbf{c}_{\lambda^\star} \in \mathrm{cone}A_{B'}^T$, so it only remains to show that B' is a feasible basis. The edge of the polyhedron described by the constraints $B \setminus \{B_{i^\star}\}$ is contained in the ray starting at the vertex $\mathbf{b}_B$ in direction $-\mathbf{e}_{i^\star}$. This ray can only be cut off by constraints $\mathbf{a}_j\mathbf{x} \leq \mathbf{b}_j$ with $\mathbf{a}_{ji^\star} < 0$. At least one such constraint must exist by the condition that $\mathbf{d}$ lies in the support of the normal fan, i.e. the corresponding linear program is bounded. The fraction $\frac{\langle \mathbf{a}_j, \mathbf{b}_B \rangle - \mathbf{b}_j}{\mathbf{a}_{ji^\star}}$ is the Euclidean distance from $\mathbf{b}_B$ to the intersection of the constraint $\langle \mathbf{a}_j, \mathbf{x} \rangle \leq \mathbf{b}_j$ with the ray, and so B' is feasible.[7] This completes the proof of the invariant and thus the proof of correctness.

If the initial objective $\mathbf{c}$ lies in the facet of a normal cone we may get $\lambda^\star = \lambda$ in the first iteration. Nevertheless, due to the precondition on $[\mathbf{c}, \mathbf{d}]$, every computed value $\lambda^\star$ is distinct and, except for the last one, corresponds to one intersection point of $[\mathbf{c}, \mathbf{d}]$ with a facet of a normal cone. Therefore, the number of iterations is bounded by $N + 1$.   ◄

If the input polyhedron were non-simple, the proof of correctness and termination would still work given $\mathbf{c}$ in sufficiently general position. However, the $\lambda^\star$ would then correspond to points of intersection between $[\mathbf{c}, \mathbf{d}]$ and cones corresponding to bases. Those cones are merely subsets of normal cones, and they may not be mutually consistent with a single triangulation of the normal fan. For this reason, we consider a perturbed polyhedron

$$P_\varepsilon := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b} + \gamma(\varepsilon)\}$$

where $\gamma(\varepsilon) := (\varepsilon, \varepsilon^2, \ldots, \varepsilon^m)$ for $\varepsilon > 0$.

▶ **Lemma 28.** *If $\varepsilon > 0$ is sufficiently small, one has that*

1. *every feasible basis B for P_ε is also feasible for P,*
2. *the normal fan of P_ε triangulates the normal fan of P, and*
3. *if $B = \{m - n + 1, \ldots, m\}$ is a feasible basis for P, then it is also feasible for P_ε.*

Proof. If P_ε is not simple, there is some basis B and a constraint $j \notin B$ such that

$$\left\langle \mathbf{a}_j, A_B^{-1}(\mathbf{b}_B + \gamma(\varepsilon)_B) \right\rangle = \mathbf{b}_j + \varepsilon^j$$

In other words, ε is a root of a non-zero polynomial. Since there are only finitely many pairs of B and $j \notin B$, we have that P_ε is simple for sufficiently small $\varepsilon > 0$. In particular, the

---

[6] This is just a different way of saying that $(\mathbf{c}_{\lambda^\star})_{i^\star} = 0$. Due to the precondition on $[\mathbf{c}, \mathbf{d}]$ this index is uniquely defined when $\lambda^\star < 1$.

[7] Note that the minimum is positive and unique because the underlying polyhedron is simple.

normal fan of P_ε is a triangulation. For the second claim, it remains to show that for all vertices $\mathbf{x} \in P_\varepsilon$ one has $N_\mathbf{x} \subseteq N_\mathbf{y}$ for some vertex $\mathbf{y} \in P$. In fact, this is implied by the first claim, which we show next.

Let B be an infeasible basis for P, i.e. there is some $j \notin B$ such that

$$\left\langle \mathbf{a}_j, A_B^{-1}(\mathbf{b}_B + \gamma(\varepsilon)_B) \right\rangle > \mathbf{b}_j + \varepsilon^j$$

holds for $\varepsilon = 0$. Since both sides of the inequality are continuous as functions in ε, strict inequality also holds for all sufficiently small $\varepsilon > 0$. There are only finitely many infeasible bases for P, so we have that all of them are infeasible for P_ε when $\varepsilon > 0$ is sufficiently small. This implies the first claim.

Finally, if $B = \{m - n + 1, \dots, m\}$ is feasible for P, we have that

$$\left\langle \mathbf{a}_j, A_B^{-1}(\mathbf{b}_B + \gamma(\varepsilon)_B) \right\rangle - \mathbf{b}_j - \varepsilon^j \le 0$$

holds for all $j \le m-n$ when $\varepsilon = 0$. The left hand side is a polynomial in ε whose lowest-degree non-constant monomial is $-\varepsilon^j$. This implies that the inequality also holds when $\varepsilon > 0$ is sufficiently small, hence B is feasible for P_ε. ◄

Without explicit bounds on the coefficients describing P, we cannot give a quantitative bound for ε. We avoid the need for such a bound by applying the perturbation symbolically. Since the right-hand sides $\mathbf{b}$ never appear in divisors, we can perform related computations in the polynomial ring $R := \mathbb{R}[\varepsilon]$. The order $\le$ on $\mathbb{R}$ naturally extends to a lexicographic order on R such that $a \le b$ holds for $a, b \in R$ if and only if $a(\varepsilon) \le b(\varepsilon)$ holds over the reals for all sufficiently small $\varepsilon > 0$.

▶ **Theorem 29.** *Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \le \mathbf{b}\}$ be pointed, $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$, and B an optimal basis for $\mathbf{c}$. If every intersection of $[\mathbf{c}, \mathbf{d})$ with a facet F of a cone spanned by a feasible basis of P lies in the relative interior of F, the Shadow Simplex can be used to compute an optimal basis for $\mathbf{d}$ in $O(mn^2 + Nmn)$ arithmetic operations, where N is the number of intersections of $[\mathbf{c}, \mathbf{d}]$ with some triangulation of the normal fan of P that contains the cone spanned by the initial basis B.*

Proof. Rearrange the rows of $A\mathbf{x} \le \mathbf{b}$ so that $B = \{m - n + 1, \dots, m\}$ and apply Lemma 28. This gives us all preconditions of the Shadow Simplex algorithm, and the proof of Theorem 27 applies with a single caveat: the computation of $j^\star$ involves computations and comparisons of terms of the form $\frac{\langle \mathbf{a}_j, \mathbf{b}_B + \gamma(\varepsilon)_B \rangle - \mathbf{b}_j - \varepsilon^j}{\mathbf{a}_{ji^\star}} \in R$. The resulting polynomials contain at most $n + 2$ monomials. By storing them as sparse sorted vectors, we can compute each term and compare it to the previous best in time $O(n)$, so that the computation of $j^\star$ still requires only $O(mn)$ arithmetic operations. ◄

7 Removing unboundedness

In this section, we discuss two approaches to add constraints to a pointed polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \le \mathbf{b}\}$ to construct an LP equivalent *polytope P'*. This requires the additional constraints to be strictly valid for all vertices of P.

▶ **Lemma 30.** *Let $M \in \mathbb{R}^{n \times n}$ be a matrix whose rows are $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{S}^{n-1}$. Then $\mathbf{v}_1, \dots, \mathbf{v}_n$ satisfy the δ-distance property if and only if the columns $\mathbf{u}_1, \dots, \mathbf{u}_n$ of M^{-1} satisfy $\|\mathbf{u}_j\| \le 1/\delta$ for all j.*

Proof. Let $H_i = \operatorname{span}\{\mathbf{v}_j : j \neq i\}$. Then

$$1 = \langle \mathbf{v}_i, \mathbf{u}_i \rangle = \|\mathbf{u}_i\| d(\mathbf{v}_i, H_i) \iff d(\mathbf{v}_i, H_i) = \frac{1}{\|\mathbf{u}_i\|}$$

implies the statement. ◀

▶ **Lemma 31.** *Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ satisfy the local δ-distance property and let* $b_{\max} = \max\left\{\frac{|\mathbf{b}_i|}{\|\mathbf{a}_i\|} : i \in [m]\right\}$. *Then every vertex $\mathbf{x} \in P$ satisfies $\|\mathbf{x}\| \leq \frac{nb_{\max}}{\delta}$.*

Proof. We may assume without loss of generality that $\|\mathbf{a}_i\| = 1$ for all $i \in [m]$. Let $\mathbf{x}$ be a vertex and let B be a basis for $\mathbf{x}$. Let $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^n$ be the columns of A_B^{-1}. The triangle inequality and Lemma 30 imply $\|\mathbf{x}\| = \|A_B^{-1}\mathbf{b}_B\| \leq \sum_{i \in B} \|\mathbf{u}_i\| b_{\max} \leq nb_{\max}/\delta$. ◀

Proof of Lemma 8. **1.** Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ satisfy the local δ-distance property. For some feasible basis I we let

$$\mathbf{w} = -\frac{1}{n} \sum_{i \in I} \frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}$$

and $P' = \left\{\mathbf{x} \in P : \langle \mathbf{w}, \mathbf{x} \rangle \leq \frac{nb_{\max}}{\delta}\right\}$. The normal fan of P' covers $\mathbb{R}^n$, because $\mathbf{0}$ lies in the interior of $\operatorname{conv}(\{\mathbf{a}_i : i \in I\} \cup \{\mathbf{w}\})$, and so P' is a polytope. Since $\|\mathbf{w}\| < 1$, we have $\langle \mathbf{w}, \mathbf{x} \rangle < \|\mathbf{x}\| \leq \frac{nb_{\max}}{\delta}$ for every vertex $\mathbf{x} \in P$ by Lemma 31, so P' is LP equivalent to P. Every vertex of P' is either a vertex of P or the intersection of an unbounded ray of P with the new constraint. Consider a feasible basis B of P'. Either B is already feasible for P, in which case it satisfies the δ-distance property. Otherwise, it is of the form $\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}, \mathbf{w}$, where $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is a feasible basis for P, after a suitable renumbering of indices.

In this case, cone $\{\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}\}$ is on the boundary of the support of the normal fan of P, and so by the proof of Lemma 5 we have $d(\mathbf{w}, H) \geq \delta/n$, where $H = \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}\}$. Using an orthogonal transformation, we may assume without loss of generality that $\mathbf{a}_{1n} = \cdots = \mathbf{a}_{(n-1),n} = 0$ and so the matrix M whose rows are the basis vectors normalized to unit length is of the form

$$M = \begin{pmatrix} A' & 0 \\ \mathbf{w}'^T & h \end{pmatrix} \in \mathbb{R}^{n \times n}$$

where $h = d(\mathbf{w}/\|\mathbf{w}\|, H) \geq \delta/n$ and $\|\mathbf{w}'\| < 1$. We compute

$$M^{-1} = \begin{pmatrix} A'^{-1} & 0 \\ -\mathbf{w}'^T A'^{-1}/h & 1/h \end{pmatrix}$$

By Lemma 30, it is sufficient to show that the norms of the columns of M^{-1} are bounded by $2n/\delta^2$. This is immediate for the last column. Let $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ be the columns of A'^{-1}. We have $\|\mathbf{u}_i\| \leq 1/\delta$ by Lemma 30. Furthermore, the i-th entry of the last row of M^{-1} is bounded in absolute value by

$$\left|\frac{\langle \mathbf{w}', \mathbf{u}_i \rangle}{h}\right| \leq \frac{\|\mathbf{u}_i\| n}{\delta} \leq \frac{n}{\delta^2}$$

By the triangle inequality, the norms of the first $n-1$ columns of M^{-1} are bounded by $1/\delta + n/\delta^2 \leq 2n/\delta^2$. This completes the proof that P' satisfies the local $\delta^2/(2n)$-distance property.

2. Now suppose that A satisfies the global δ-distance property. For every vertex $\mathbf{x} \in P$, we have

$$\langle -\mathbf{a}_i, \mathbf{x} \rangle \leq \|\mathbf{a}_i\| \|\mathbf{x}\| \leq \frac{n \|\mathbf{a}_i\| b_{\max}}{\delta}$$

by Lemma 31, and so the polytope

$$P' = \left\{ \mathbf{x} \in \mathbb{R}^n : -\frac{n \|\mathbf{a}_i\| b_{\max}}{\delta} - 1 \leq \langle \mathbf{a}_i, \mathbf{x} \rangle \leq \mathbf{b}_i \quad \forall i \in [m] \right\}$$

is LP equivalent to P. Furthermore, every (not necessarily feasible) basis of the constraint matrix of P' is equal to a basis of A up to sign changes, which do not affect the δ-distance property. Therefore, P' satisfies the global δ-distance property. ◀

8 Feasibility

▶ **Lemma 32.** *Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ where $A \in \mathbb{Z}^{m \times n}$ is an integral matrix of full column rank with subdeterminants bounded by Δ. One can compute a feasible basis of P or decide infeasibility using an expected $O(n^5 \Delta^2 \ln n\Delta)$ shadow simplex pivots.*

Proof. Consider the linear program

$$\min s$$
$$\langle \mathbf{a}_i, \mathbf{x} \rangle - s \leq \mathbf{b}_i \qquad\qquad \forall i \in [m]$$
$$s \geq 0$$

The $(m+1) \times (n+1)$-constraint matrix is integral of full column rank and has $n \times n$-subdeterminants bounded by $n\Delta$. Therefore, it satisfies the global δ-distance property with $\delta = \frac{1}{n^2\Delta^2}$ by Lemma 35. The point $(\mathbf{0}, -\min(\{0\} \cup \{\mathbf{b}_i : i \in [m]\}))$ is feasible, and so a feasible basis can be found using a standard ray-casting procedure. Lemma 8 implies that we can construct an LP equivalent polytope with the same parameter δ, which we can then optimize in $O(n^5 \Delta^2 \ln n\Delta)$ shadow simplex pivots by Theorem 6. If the optimal solution we found satisfies $s = 0$, we can read off a feasible basis of P; otherwise, we know that P is empty. ◀

▶ **Lemma 33.** *Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ where $A \in \mathbb{R}^{m \times n}$ satisfies the global δ-distance property. One can compute a feasible basis of P or decide infeasibility using an expected $O(mn^3/\delta \ln n/\delta)$ shadow simplex pivots.*

Proof. We proceed by iteratively adding the constraints of P, one at a time. First, observe that we can find a basis B of A efficiently using Gauss elimination, which gives us the unique feasible basis of $P_B = \{\mathbf{x} \in \mathbb{R}^n : A_B\mathbf{x} \leq \mathbf{b}_B\}$.

Now suppose we already found a feasible basis B of $P_I = \{\mathbf{x} \in \mathbb{R}^n : A_I\mathbf{x} \leq \mathbf{b}_I\}$, $I \subsetneq [m]$. Let $i \notin I$. Since P_I satisfies the global δ-distance property, we can combine Lemma 8 with Theorem 6 to solve the linear program

$$\min \{\langle \mathbf{a}_i, \mathbf{x} \rangle : \mathbf{x} \in P_I\} = \gamma$$

using an expected $O(n^3/\delta \ln n/\delta)$ shadow simplex pivots. If $\gamma > \mathbf{b}_i$, this implies that P_{I+i} is empty and therefore P is empty. Otherwise, the solution of the linear program yields a point $\mathbf{x} \in P_{I+i}$ (if $\gamma = -\infty$, we find a suitable point on an unbounded ray) which we can round to a feasible basis of P_{I+i} using a standard ray-casting procedure if necessary.

Applying this procedure iteratively until $I = [m]$, we use an expected $O(mn^3/\delta \ln n/\delta)$ shadow simplex pivots to obtain a feasible basis of P. Observe that the time required for those pivots dominates the time required for any intermediate ray-casting. ◀
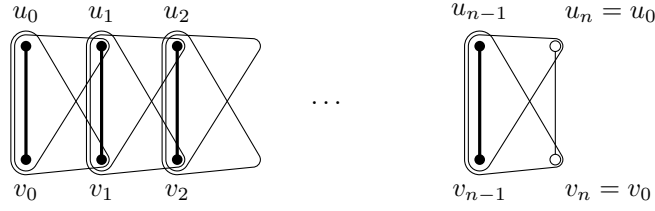
■ **Figure 2** A perfect matching and a selection of tight odd sets.

9   The perfect matching polytope

The perfect matching polytope $P_G \subset \mathbb{R}^E$ of an undirected graph $G = (V, E)$, $|V| = 2n$, is the convex hull of the characteristic vectors χ_M of perfect matchings $M \subseteq E$ (see chapter 25 of [18] for a collection of fundamental results on P_G). It is described by the system of inequalities

$$\sum_{e \in \delta(v)} x(e) \leq 1 \qquad\qquad \forall v \in V$$

$$\sum_{e \in \delta(U)} x(e) \geq 1 \qquad\qquad \forall U \subseteq V, |U| \text{ odd}$$

$$x(e) \geq 0 \qquad\qquad \forall e \in E$$

We will also use the fact that two vertices χ_M and χ_N of P_G are adjacent if and only if $M \triangle N$ is a cycle.

We will show that even though this polytope has an exponential number of facets, so that its normal fan contains an exponential number of extreme rays, every normal cone is rather wide. This is due to the high level of degeneracy of the polytope. While a much better bound on the diameter of P_G follows directly from the combinatorial adjacency structure noted above, it is interesting to see that some of our techniques can be applied to P_G. As far as we know, this is the first example of a combinatorial polytope that satisfies this kind of "discrete curvature bound" without having a constraint matrix with small subdeterminants.

Since P_G is not full-dimensional, there are two different but essentially equivalent definitions for the normal cones of P_G. One can treat P_G as a polytope in the ambient space $\mathbb{R}^E$, keeping the definition of the normal cone N_v of a vertex $\mathbf{v} \in P_G$ as the set of objective functions $\mathbf{c} \in \mathbb{R}^E$ that are maximized at $\mathbf{v}$. Hence N_v is not pointed, and its lineality space $L^\perp$ is the set of vectors that are orthogonal to the affine hull $\operatorname{aff}(P_G)$ of P_G. Alternatively, one may treat P_G as a full-dimensional polytope within $\operatorname{aff}(P_G)$, in which case the normal cone N'_v is simply the restriction of N_v to the linear space L of vectors parallel to $\operatorname{aff}(P_G)$. The choice of definition does not affect τ-width: given a ball $\mathbf{w} + \tau B_2^E \subseteq N_v$, $\|\mathbf{w}\| = 1$, the orthogonal projection of the ball onto L is $\mathbf{w}' + \tau B_2^L \subseteq N'_v$ with $\|\mathbf{w}'\| \leq 1$.

▶ **Theorem 34.** P_G is τ-wide for $\tau = 1/(3\sqrt{|E|})$.

Proof. Let $\chi_M \in P_G$ be a vertex and let N_{χ_M} be its normal cone. Let us label the vertices of G such that

$$M = \{u_0 v_0, u_1 v_1, \ldots, u_{n-1} v_{n-1}\}$$

For an odd set $U \subseteq V$ of vertices, let $\mathbf{a}_U$ be the corresponding row of the constraint matrix in normal form, i.e. $\mathbf{a}_U = -\chi_{\delta(U)}$. We consider the following 3-element sets, see Figure 2:

$$\mathcal{U} := \left\{ \{u_k, v_k, w\} : 0 \leq k \leq n-1, w \in \left\{u_{k+1 \pmod n}, v_{k+1 \pmod n}\right\} \right\}$$

Since the corresponding constraints are tight at χ_M, we have that

$$\mathbf{w} := \sum_{U \in \mathcal{U}} \mathbf{a}_U \in N_{\chi_M}$$

Note that for $e \in M$, we have $\mathbf{w}(e) = -2$, and for $e \in E \setminus M$, we have $\mathbf{w}(e) \in \{-4, -6\}$ (every vertex is contained in exactly 3 of the sets in $\mathcal{U}$, and there can be at most one set in $\mathcal{U}$ that contains both endpoints of $e \notin M$). In particular, $\|\mathbf{w}\| \leq 6\sqrt{|E|}$.

Every facet F of N_{χ_M} corresponds to an edge from χ_M to some other vertex χ_N. We know that $M \triangle N$ is a cycle C. The direction $\mathbf{v}$ of the edge from χ_N to χ_M, which is (inner) normal to F, satisfies

$$\mathbf{v}(e) = \begin{cases} 1 & e \in C \cap M \\ -1 & e \in C \setminus M \\ 0 & e \notin C \end{cases}$$

We compute

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{e \in C \cap M} \mathbf{w}(e) - \sum_{e \in C \setminus M} \mathbf{w}(e) \geq -2|C \cap M| + 4|C \setminus M| = |C|$$

where we use the fact that C alternates between edges in M and edges in N for the last equation. Let H be the affine span of F. We obtain

$$d(\mathbf{w}, H) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|} \geq \frac{|C|}{\|\mathbf{v}\|} = \sqrt{|C|} \geq 2$$

In other words, N_{χ_M} contains a ball of radius 2 around w, and so N_{χ_M} is τ-wide for $\tau = 2/\|\mathbf{w}\| \geq 1/(3\sqrt{|E|})$. ◀

────── **References** ──────

1   Karim Alexander Adiprasito and Bruno Benedetti. The Hirsch conjecture holds for normal flag complexes. Arxiv Report 1303.3598, 2014.
2   M. L. Balinski. The Hirsch conjecture for dual transportation polyhedra. *Math. Oper. Res.*, 9(4):629–633, 1984.
3   David Barnette. An upper bound for the diameter of a polytope. *Discrete Math.*, 10:9–13, 1974.
4   Nicolas Bonifas and Daniel Dadush. Short paths on the voronoi graph and the closest vector problem with preprocessing. In *To appear in SODA*, 2015.
5   Nicolas Bonifas, Marco Di Summa, Friedrich Eisenbrand, Nicolai Hähnle, and Martin Niemeier. On sub-determinants and the diameter of polyhedra. *Discrete Comput. Geom.*, 52(1):102–115, 2014. Preliminary version in SOCG 12.
6   Karl-Heinz Borgwardt. *The simplex method: A probabilistic analysis*, volume 1 of *Algorithms and Combinatorics: Study and Research Texts*. Springer-Verlag, Berlin, 1987.

7   Graham Brightwell, Jan van den Heuvel, and Leen Stougie. A linear bound on the diameter of the transportation polytope. *Combinatorica*, 26(2):133–139, 2006.

8   Tobias Brunsch and Heiko Röglin. Finding short paths on polytopes by the shadow vertex algorithm. In *Automata, languages, and programming. Part I*, volume 7965 of *Lecture Notes in Comput. Sci.*, pages 279–290. Springer, Heidelberg, 2013.

9   Jesús A. De Loera, Edward D. Kim, Shmuel Onn, and Francisco Santos. Graphs of transportation polytopes. *J. Combin. Theory Ser. A*, 116(8):1306–1325, 2009.

10   Martin Dyer and Alan Frieze. Random walks, totally unimodular matrices, and a randomised dual simplex algorithm. *Math. Programming*, 64(1, Ser. A):1–16, 1994.

11   Friedrich Eisenbrand and Santosh Vempala. Geometric random edge. Arxiv Report 1404.1568, 2014.

12   Gil Kalai. The diameter of graphs of convex polytopes and f-vector theory. In *Applied geometry and discrete mathematics*, volume 4 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 387–411. Amer. Math. Soc., Providence, RI, 1991.

13   Gil Kalai and Daniel J. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bull. Amer. Math. Soc. (N.S.)*, 26(2):315–316, 1992.

14   D. G. Larman. Paths of polytopes. *Proc. London Math. Soc. (3)*, 20:161–178, 1970.

15   Benjamin Matschke, Francisco Santos, and Christophe Weibel. The width of 5-dimensional prismatoids. Arxiv Report 1202.4701, 2013.

16   Denis Naddef. The Hirsch conjecture is true for $(0,1)$-polytopes. *Math. Programming*, 45(1, (Ser. B)):109–110, 1989.

17   Francisco Santos. A counterexample to the Hirsch conjecture. *Ann. of Math. (2)*, 176(1):383–412, 2012.

18   Alexander Schrijver. *Combinatorial Optimization: Polyhedra and Efficiency*, volume 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003.

19   Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463 (electronic), 2004.

20   Michael J. Todd. An improved Kalai-Kleitman bound for the diameter of a polyhedron. Arxiv Report 1402.3579, 2014.

21   Roman Vershynin. Beyond Hirsch conjecture: walks on random polytopes and smoothed complexity of the simplex method. *SIAM J. Comput.*, 39(2):646–678, 2009.

10   **Additional proofs**

Proof of Lemma 5. Let $\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*$ be the dual basis satisfying $\langle \mathbf{v}_i, \mathbf{v}_j^* \rangle = 1$ if $i = j$ and 0 otherwise. By definition of the δ-distance property

$$\langle \mathbf{v}_i, \mathbf{v}_i^*/\|\mathbf{v}_i^*\| \rangle \geq \delta \Rightarrow \|\mathbf{v}_i^*\| \leq 1/\delta.$$

Note that $\mathbf{x} \in \text{cone}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ iff $\langle \mathbf{v}_i^*, \mathbf{x} \rangle \geq 0$, for all $i \in [n]$. Let $\bar{\mathbf{v}} = \sum_{i=1}^n \mathbf{v}_i/n$, and note that $\|\bar{\mathbf{v}}\| \leq 1$, since it is an average of unit vectors.

We will show that $\bar{\mathbf{v}} + \frac{\delta}{n}\mathcal{B}_2^n \subseteq \text{cone}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, which suffices to prove the lemma. Take any vector $\mathbf{e}$, $\|\mathbf{e}\| \leq \delta/n$. Then for any $i \in [m]$, note that

$$\langle \mathbf{v}_i^*, \bar{\mathbf{v}} + \mathbf{e} \rangle = 1/n + \langle \mathbf{v}_i^*, \mathbf{e} \rangle \geq 1/n - \|\mathbf{v}_i^*\|\|\mathbf{e}\| \geq 0.$$

Hence $\bar{\mathbf{v}} + \mathbf{e} \in \text{cone}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, as needed.   ◀

Proof of Lemma 13. For the first part,

$$c_\Sigma^{-1} = \int_\Sigma e^{-\|\mathbf{x}\|} d\mathbf{x} = \int_\Sigma \int_{\|\mathbf{x}\|}^\infty e^{-t} dt d\mathbf{x}$$

$$= \int_0^\infty e^{-t} \int_\Sigma \mathrm{I}[\|\mathbf{x}\| \le t] d\mathbf{x} dt = \int_0^\infty e^{-t} t^n \mathrm{vol}_n(\mathcal{B}_2^n \cap \Sigma) dt = n! \mathrm{vol}_n(\mathcal{B}_2^n \cap \Sigma).$$

For the expected norm,

$$\mathbb{E}[\|X\|] = c_\Sigma \int_\Sigma \|\mathbf{x}\| e^{-\|\mathbf{x}\|} d\mathbf{x} = c_\Sigma \int_\Sigma \|\mathbf{x}\| \int_{\|\mathbf{x}\|}^\infty e^{-t} dt d\mathbf{x}$$

$$= c_\Sigma \int_0^\infty e^{-t} \int_{t(\mathcal{B}_2^n \cap \Sigma)} \|\mathbf{x}\| d\mathbf{x} dt = c_\Sigma \int_0^\infty e^{-t} t^{n+1} dt \int_{\mathcal{B}_2^n \cap \Sigma} \|\mathbf{x}\| d\mathbf{x}$$

$$= c_\Sigma(n+1)! \int_0^1 (1-s^n) \mathrm{vol}_n(\mathcal{B}_2^n \cap \Sigma) ds$$

$$= c_\Sigma(n+1)! \mathrm{vol}_n(\mathcal{B}_2^n \cap \Sigma) \frac{n}{n+1} = n. \qquad \blacktriangleleft$$

Proof of Lemma 16. **(1) $\Rightarrow$ (2)** Let $I \subseteq [m]$ for which $\{\mathbf{x}_i : i \in I\}$ are linearly independent, and examine the linear combination $\sum_{i \in I} a_i \mathbf{x}_i$. Letting $j = \arg\max_{i \in I} |a_i|$, we have that

$$\|a_j \mathbf{x}_j + \sum_{i \in I \setminus \{j\}} a_i \mathbf{x}_i\| \ge d(a_j \mathbf{x}_j, \mathrm{span}(\{\mathbf{x}_i : i \in I \setminus \{j\}\}))$$

$$= |a_j| d(\mathbf{x}_j, \mathrm{span}(\{\mathbf{x}_i : i \in I \setminus \{j\}\})) \ge \delta |a_j|, \quad (\text{ by property (1) })$$

as needed.

(2) $\Rightarrow$ (1) Take $i \in [m]$ and $J \subseteq [m]$, such that $\mathbf{x}_i \notin \mathrm{span}(\{\mathbf{x}_j : j \in J\})$. Since we need only prove a lower bound on $d(\mathbf{x}_i, \mathrm{span}(\{\mathbf{x}_j : j \in J\}))$, we may clearly assume that $\{\mathbf{x}_j : j \in \{i\} \cup J\}$ are linearly independent. Given this, we have that

$$d(\mathbf{x}_i, \mathrm{span}(\{\mathbf{x}_j : j \in J\})) = \min\left\{ \mathbf{x}_i - \sum_{j \in J} a_j \mathbf{x}_j : a_j \in \mathbb{R}, j \in J \right\}$$

$$\ge \min\left\{ \delta \max\left\{ 1, \max_{j \in J} |a_j| \right\} : a_j \in \mathbb{R}, j \in J \right\} \quad (\text{ by property (2) })$$

$$\ge \delta, \quad \text{as needed.} \qquad \blacktriangleleft$$

Proof of Lemma 19. Let $i \in \{1, \ldots, k-1\}$. Let $S_i = \mathrm{span}\{\pi(\mathbf{v}_j) : j \ne i, k\}$. First, observe

$$\mathrm{span}(\{\mathbf{v}_k\} \cup S_i) = \mathrm{span}(\{\mathbf{v}_j : j \ne i\}).$$

Assume that $d(\pi(\mathbf{v}_i), S_i) < \delta \|\pi(\mathbf{v}_i)\|$. So there exists some $\mathbf{x} \in S_i$ such that $d(\pi(\mathbf{v}_i), \mathbf{x}) < \delta \|\pi(\mathbf{v}_i)\|$. Since we can write $\mathbf{v}_i = \pi(\mathbf{v}_i) + \lambda \mathbf{v}_k$ for some $\lambda \in \mathbb{R}$, we get

$$d(\mathbf{v}_i, \mathrm{span}(\{\mathbf{v}_j : j \ne i\})) \le d(\mathbf{v}_i, \mathbf{x} + \lambda \mathbf{v}_k) = d(\pi(\mathbf{v}_i), \mathbf{x}) < \delta \|\pi(\mathbf{v}_i)\| \le \delta \|\mathbf{v}_i\|,$$

which contradicts the δ-distance property of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. So we must in fact have $d(\pi(\mathbf{v}_i), S_i) \ge \delta \|\pi(\mathbf{v}_i)\|$ for all $i \in \{1, \ldots, k-1\}$, which completes the proof. $\blacktriangleleft$

The following result is already implicit in [5]. We provide it here for completeness.

▶ **Lemma 35.** *Let $A \in \mathbb{R}^{m \times n}$ be an integral matrix whose entries are bounded by Δ_1 and whose $(n-1) \times (n-1)$ subdeterminants are bounded by Δ_{n-1} in absolute value. Then A satisfies the global δ-distance property with $\delta = 1/(n\Delta_1\Delta_{n-1})$. Furthermore, any polyhedron with constraint matrix A is τ-wide with $\tau = 1/(n^2\Delta_1\Delta_{n-1})$.*

Proof. It is sufficient to consider the case where $A \in \mathbb{Z}^{n \times n}$ is invertible. Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the rows of A and let $H_i = \text{span}\{\mathbf{a}_j : j \neq i\}$. The vector $\mathbf{u}_i$ satisfying $A\mathbf{u}_i = |\det(A)|\mathbf{e}_i$ is a normal vector of H_i with $\mathbf{u}_i \in \mathbb{Z}^n$ and $\|\mathbf{u}_i\|_\infty \leq \Delta_{n-1}$ by Cramer's rule. We can compute

$$d(\mathbf{a}_i/\|\mathbf{a}_i\|, H_i) = \frac{\langle \mathbf{a}_i, \mathbf{u}_i \rangle}{\|\mathbf{a}_i\|\|\mathbf{u}_i\|} \geq \frac{1}{(\sqrt{n}\Delta_1)(\sqrt{n}\Delta_{n-1})} = \frac{1}{n\Delta_1\Delta_{n-1}}$$

using the fact that $\|\mathbf{a}_i\|_\infty \leq \Delta_1$.

The "furthermore" part of the statement of the Lemma follows from Lemma 5. ◀