Monotone Paths on Polytopes: Combinatorics and Optimization

By

Alexander E. Black

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Jesús A. De Loera (Chair)

Eric Babson

Raman Sanyal

Committee in Charge

2024

To Whiskers

Contents

Abstract

A foundational problem in optimization is that of **linear programming**, the problem of maximizing a linear function $\mathbf{c}^\intercal \mathbf{x}$ subject to linear inequality constraints $A\mathbf{x} \leq \mathbf{b}$. Geometrically, this problem corresponds to finding the highest point in some direction on a polytope. The simplex method, one of the oldest and most famous algorithms for linear programming, does this by following a path on the graph of the polytope such that at each step the linear objective function value increases. Such a path is called a **monotone path**. The simplex method could potentially follow any monotone path on the polytope. For implementation, one must specify a rule to choose such a path called a **pivot rule**. The goal of this thesis is to study the geometric combinatorics of monotone paths and pivot rules for linear programs and bound the performance of the simplex method.

We first show performance bounds for the simplex method in the case of 0/1-polytopes. In this case, we introduce two pivot rules guaranteed to follow paths of length at most the dimension of the polytope matching optimal bounds of Naddef. We extend these results to d-dimensional $(0, k)$-lattice polytopes finding a path following algorithm using the simplex method as a subroutine guaranteed to follow a path of length at most $2dk$ asymptotically matching the best known bounds on diameters of their graphs due to Kleinschmidt and Onn and improving upon the best known bounds in this framework due to Del Pia and Michini. It remains open whether there is a pivot rule for the simplex method that guarantees even a polynomial in d and k many steps.

In the remainder of this thesis, we study two constructions: the **monotone path polytope** and **pivot rule polytope**. The monotone path polytope, introduced by Billera and Sturmfels, is a geometric model for the spaces of all monotone paths on a polytope that could be chosen by a shadow pivot rule for a fixed linear program. We study these on examples of polytopes fundamental in algebraic combinatorics such as cross-polytopes and products of simplices. We also study how monotone path polytopes are affected by other constructions such as prisms and pyramids. In each case, we find rich combinatorics with surprising new interpretations in terms of the simplex method. The pivot rule polytope is an alteration of the monotone path polytope to classify a broad set of pivot rules we study in detail called **normalized weight pivot rules**, both of which are defined in this thesis. In particular, we show that if a polynomial time version of the simplex method exists, there must be a normalized weight pivot rule guaranteed to always choose polynomial length paths. Finally, we study this construction on examples finding new realizations of polytopes in algebraic combinatorics such as the associahedron.

Acknowledgments

First and foremost, I would like to thank Jesús De Loera and Raman Sanyal, since they played an advisory role for all of the work in this thesis. I have found this area of research enjoyable and fruitful, and their guidance has served as an excellent foundation for my career. Eric Babson also played an initial advisory role and deserves appreciation. Next are my collaborators on various related projects completed at the time of this thesis including both Raman and Jesús but also Niklas Lütjeharms, Sean Kafer, Laura Sanità, Steffen Borgwardt, Matthias Brugger, Igor Araujo, Amanda Burcroff, Yibo Gao, Robert Krueger, Alex McDonough, Kevin Liu, Garrett Nelson, Michael Wigal, Youngho Yoo, and Mei Yin. I would like to especially highlight the contributions of Paco Criado, Sean Kafer, Sophie Huiberts, and Sophie Rehberg, who each discussed results related to this work with me.

I would also like to thank the friends, family, and educators that made this possible and highlight especially Jordan Barrett, Elijah Barstis, and Emily Storck.

Work for this thesis was supported by the NSF Graduate Research Fellowship Program and NSF DMS-1818969.

CHAPTER 1

Introduction

1.1. Linear Programming and the Simplex Method

A **linear program** (LP) is a problem of the form $\max(\{\mathbf{c}^\intercal\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\})$. A **polyhedron** is the solution set to a system of linear inequalities $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ and is called a **polytope** if it is bounded. Linear programming is then the problem of finding the highest point in some direction on a polyhedron and reduces to the case of polytopes, which will be the focus of this work. Schrijver's encyclopedic work on combinatorial optimization [123] includes endless examples of problems that may be modeled directly by linear programming including shortest paths and network flow problems. Even when problems cannot be modeled directly by linear programs, they are often modeled by mixed integer programs (MIPs), problems of the form $\max(\mathbf{c}^\intercal\mathbf{x})$ such that $\mathbf{x} \in P \cap (\mathbb{R}^m \times \mathbb{Z}^n)$ for a polytope P. The algorithm used by standard MIP solvers such as Gurobi [80] is called branch and bound and solves MIPs by reducing them to solving multiple LPs. Even for optimization problems with nonlinear nonconvex objective functions, LP techniques still apply with many examples provided in [102]. Linear programming is also the fundamental tool for polyhedral computation [71], which is foundational for many other areas such as the study of Rectified Linear Unit (ReLU) neural networks [82]. LPs are ubiquitious, and it is a central problem in optimization to determine how fast one can solve them [126].

The first general LP algorithm is attributed to Fourier [68] in 1824, yet the field remains active after nearly 200 years as highlighted by the work of this thesis. The central problem in the area is Smale's 9th problem for the twentieth century [126], which asks whether there exists a strongly polynomial time algorithm for linear programming (i.e., that can solve an LP in time polynomial in m and n for an LP given by $\max(\mathbf{c}^\intercal\mathbf{x})$ such that $A\mathbf{x} \leq \mathbf{b}$ for an $m \times n$ matrix A). One of the most common algorithms for linear programming, the simplex method, remains a candidate solution to this problem. Like a polygon, a polytope comes together with a set of vertices and edges called its **graph**. The simplex method solves an LP by following a path on the graph of a polytope such that at each step the linear objective function value increases. Such a path is called a **monotone path**, and the lengths of such paths govern the run-time of the simplex method. There are many ways to choose monotone paths called **pivot rules**.

The goal of this thesis is to better understand pivot rules for the simplex method through studying the geometric combinatorics of monotone paths on polytopes. Our work is part of a long line of literature dating back to Dantzig in [49] with hundreds of papers on the topic having been written since then (see [20, 29, 31, 50, 83, 95, 127] and references therein for a sample). In this thesis, we will study the performance of the simplex method and the combinatorics of monotone paths.

In this chapter, we will provide some background on the simplex method and polyhedral geometry and state each of the main results of the thesis. In Chapter 2, we will study the behavior of the simplex method on 0/1-polytopes and lattice polytopes with the key new ingredient being a deterministic approach to shadow pivot rules. In Chapter 3, we will study the Baues poset and monotone path polytope, a CW-complex [16] and polyhedral construction introduced by Billera and Sturmfels in [17] respectively that encode the space of all monotone paths on a polytope. The core focus of that chapter will be on the rich combinatorics that arises from many examples. Finally, in Chapter 4, we will generalize the monotone path polytope construction to classify a broad set of pivot rules called **normalized weight pivot rules** and study highly structured examples of pivot rule polytopes.

There are different versions of the simplex method such as the revised simplex method and dual simplex method (see Chapters 3 and 4 of [15]). As a result, we will make explicit here what we mean when we say the simplex method and then present our results regarding its complexity. What we discuss here coincides with how the simplex method is generally covered in textbooks such as in [15, 106]. The simplex method was originally designed to solve linear programs in standard form:

$$\max(\{\mathbf{c}^\intercal \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}),$$

where A is an $m \times n$ matrix, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. Any linear program may be written in this form possibly by adding more variables. Namely, maximizing $\mathbf{c}^\intercal \mathbf{x}$ such that $A\mathbf{x} \leq \mathbf{b}$ is equivalent to maximizing $\mathbf{c}^\intercal \mathbf{x}$ such that $A\mathbf{x} + \mathbf{y} = \mathbf{b}$ for some $\mathbf{y} \geq \mathbf{0}$. That linear program is itself equivalent to maximizing $\mathbf{c}^\intercal(\mathbf{x}_+ - \mathbf{x}_-)$ subject to $A(\mathbf{x}_+ - \mathbf{x}_-) + \mathbf{y} = \mathbf{b}$, $\mathbf{x}_+, \mathbf{x}_-, \mathbf{y} \geq \mathbf{0}$. The final version of the LP is written in standard form.

Since the linear program is determined by the subspace $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$, one assumes that A is of full rank. Then, since A is full rank, given m linearly independent columns of A, there is always a unique linear combination that yields $\mathbf{b}$. Let B be a subset of indices corresponding to m linearly independent columns of A, and let A_B be the matrix obtained by restricting to those columns. Let $\mathbf{v}$ be given by

$$\mathbf{v}_i := \begin{cases} (A_B^{-1}\mathbf{b})_i \text{ if } i \in B \\ 0 \text{ otherwise} \end{cases}.$$

Then $\mathbf{v}$ is called a **basic solution**. In particular, note that

$$A\mathbf{v} = A_B(A_B^{-1}\mathbf{b}) = \mathbf{b}.$$

If $\mathbf{v} \geq \mathbf{0}$, then $\mathbf{v}$ is called a **basic feasible solution**. Basic feasible solutions have other geometric characterizations. Namely, a point in P is called a **vertex** if there exists a hyperplane H that intersects P at exactly $\mathbf{v}$.

LEMMA 1.1.0.1. *Let $P = \{\mathbf{x} \geq \mathbf{0} : A\mathbf{x} = \mathbf{b}\}$, and let $\mathbf{v} \in P$. Then the following are equivalent:*

- $\mathbf{v}$ *is a basic feasible solution,*
- $\mathbf{v}$ *is a vertex.*

PROOF. See the proof of Theorem 4.4.1 in [106]. □

Thus, instead of considering vertices, we will consider basic feasible solutions. In fact, we will only consider bases (i.e., subsets $B \subseteq [m]$ such that $|B| = n$, and the set of columns of A indexed by B are linearly independent) that correspond to basic feasible solutions. Note that a basis determines a unique basic feasible solution given by $A_B^{-1}\mathbf{v}$. However, many bases can potentially correspond to the same basic feasible solution. The simplex method solves a linear program by generating an initial basis B for a starting basic feasible solution. This step is often called **Phase** 1. Then it moves to a neighboring basis, where two bases B and B' are neighbors if $B' = (B \setminus \{i\}) \cup \{j\}$ for some $i \in B$ and $j \in B'$. Such a step is called a **pivot**. The neighboring basis either corresponds to the same basic feasible solution, and such a step is called a **degenerate pivot**, or it corresponds to a new basic feasible solution and is called a **non-degenerate pivot**. In order to perform pivots, the simplex method first generates the set of all neighboring bases to a fixed starting basis. To generate all neighbors, one applies the following lemma:

LEMMA 1.1.0.2. *Let $P = \{\mathbf{x} \geq \mathbf{0} : A\mathbf{x} = \mathbf{b}\}$ be a polyhedron, and let $\mathbf{v}$ be a basic feasible solution of P with basis B. Then given j not in the basis B, precisely one of the following will be true:*

- *There exists a maximal $\lambda \geq 0$ such that $\mathbf{v} + \lambda(e_j - A_B^{-1}A_j)$ is also a basic feasible solution.*
- *The polyhedron P is unbounded and contains $\mathbf{v} + \lambda(e_j - A_B^{-1}A_j)$ for all $\lambda \geq 0$.*

PROOF. To see this, note that for all $\lambda \geq 0$,

$$A(\mathbf{v} + \lambda(e_j - A_B^{-1}A_j) = A\mathbf{v} + \lambda(A_j - AA_B^{-1}A_j) = \mathbf{b} + \lambda(A_j - A_j) = \mathbf{b}.$$

Thus, $\mathbf{v} + \lambda(e_j - A_B^{-1}A_j) \in P$ so long as $\mathbf{v} + \lambda(e_j - A_B^{-1}A_j) \geq \mathbf{0}$. If $e_j - A_B^{-1}A_j \geq \mathbf{0}$, then this is true for all $\lambda \in \mathbb{R}$ and therefore P is unbounded and $\mathbf{v} + \lambda(e_j - A_B^{-1}A_j)$. Otherwise, some components of $e_j - A_B^{-1}A_j$ must be negative. Thus, there exists a maximal $\lambda \geq 0$ such that $\mathbf{v} + \lambda(e_j - A_B^{-1}A_j)$ is still non-negative as desired. □

The resulting basis after an exchange is of the form $(B \cup \{j\}) \setminus \{i\}$, where i may be chosen explicitly as any index such that $v_i - \lambda(A_B^{-1}A_j) = 0$. In exact terms,

$$\lambda = \min\left\{\frac{v_i}{(A_B^{-1}A_j)_i} : i \in [n] \setminus B, \mathbf{v}_i \geq 0 \text{ and } (A_B^{-1}A_j)_i \geq 0\right\},$$

where one defines $x/0 = \infty$ for all $x \in \mathbb{R}$. Note that if $\lambda = 0$, a pivot will occur, but the vertex does not change. Such a pivot is called a degenerate pivot. If $\lambda > 0$, then the resulting pivot is a non-degenerate pivot.

Note that any choice of j to add to the basis B and i such that $v_i - \lambda(A_B^{-1}A_j) = 0$ could potentially lead to a valid pivot for the simplex method. The only catch is the new basis must have large objective function value. Namely, if to maximize $\mathbf{c}^\intercal\mathbf{x}$, we only allow pivots such that:

$$\mathbf{c}^\intercal(\mathbf{v} + \lambda(e_j - A_B^{-1}A_j)) \geq \mathbf{c}^\intercal\mathbf{v}$$

or equivalently such that $\mathbf{c}^\intercal(e_j - A_B^{-1}A_j) = \mathbf{c}_j - \mathbf{c}^\intercal A_B^{-1}A_j > 0$. The value $\mathbf{c}_j - \mathbf{c}^\intercal A_B^{-1}A_j$ is commonly called the **reduced cost** in the literature such as in [15]. In particular, there is the following characterization of optimality used in the simplex method:

Lemma 1.1.0.3. *Consider an LP of the form* $\max(\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P\})$*, where* $P = \{\mathbf{x} \geq \mathbf{0} : A\mathbf{x} = \mathbf{b}\}$*. Then a basic feasible solution* $\mathbf{v}$ *with basis* B *is optimal if all reduced costs of all potential pivots from* B *are negative.*

Proof. See the discussion in Section 5.6 of [106]. □

Thus, either a vertex is optimal, or there exists a pivot with positive reduced cost. The idea behind the simplex method is to keep taking steps with positive reduced cost until none exist and one reaches the optimum. Note that the restriction that the reduced cost must be strictly positive does not uniquely determine the basis. The choice of which i to add to a basis is of fundamental importance for understanding the simplex method. In general, i is called the **entering variable** and j is called a **leaving variable** for an exchange of the form $(B \setminus \{j\}) \cup \{i\}$. A **pivot rule** is a method for choosing an entering and leaving variable at a given basis, and different pivot rules may lead to vastly different run-times for the simplex method. Though most pivot rules are solely stated for how to choose the entering variable. With this in place, we may formally state the simplex method. Note that our formulation of the simplex method only solves LPs for which we know an initial basic feasible solution already.

Algorithm 1 Simplex Method

Initialize basis feasible solution $\mathbf{v}$ with basis B
while $\exists$ an entering variable j with positive reduced cost **do**
  Select entering variable j using a pivot rule
  Compute the ratios $\frac{\mathbf{v}_i}{(A_B^{-1}A_j)_i}$ for each i such that $(A_B^{-1}A_j)_i < 0$
  Select pivot row i with the smallest non-negative ratio using a pivot rule and let λ be equal to that minimal ratio
  Replace $\mathbf{v}$ with $\mathbf{v} + \lambda(e_j - A_B^{-1}A_j)$ and B with $(B \cup \{j\}) \setminus \{i\}$
end while
Return $\mathbf{v}$

There are many possible pivot rules as we will discuss in Section 1.5 and Chapter 4. One example of such a pivot rule is called **Dantzig's pivot rule**, which is the standard first example in many textbooks [15, 49, 106]. Dantzig's pivot rule says to choose the entering variable with the most reduced cost and choose the leaving variable arbitrarily. Another natural choice of pivot rule is the **greatest improvement pivot rule**, which maximizes the objective function improvement over all improving neighbors. Another choice is the **random edge pivot rule**, which chooses an entering variable completely random. The one most often used in practice with some additional considerations is the **steepest edge pivot rule**, which maximizes the normalized reduced cost $\frac{\mathbf{c}_i - \mathbf{c}^\intercal A_B^{-1} A_j}{\|e_i - A_B^{-1} A_j\|_2}$. The pivot rule most often in theory due to their role in the random analysis of the simplex method [31, 127] are **shadow pivot rules**, which were originally introduced by Gass and Saaty in the context of parametric linear programming [74]. In terms of pivoting, one starts at an initial basis B and basic feasible solution $\mathbf{v}$ and then chooses an auxiliary vector $\mathbf{w}$ and then chooses a pivot by maximizing

$$\frac{\mathbf{c}^\intercal (e_i - A_B^{-1} A_j)}{\mathbf{w}^\intercal (e_i - A_B^{-1} A_j)},$$

for some $\mathbf{w} \in \mathbb{R}^n$. Shadow pivot rules get their name from a geometric interpretation. Namely, if one projects the feasible region $P = \{\mathbf{x} \geq \mathbf{0} : A\mathbf{x} = \mathbf{b}\}$ via the linear map $\pi(\mathbf{x}) = (\mathbf{w}^\intercal \mathbf{x}, \mathbf{c}^\intercal \mathbf{x})$, $\pi(P)$ is a polygon. Furthermore, the shadow simplex path is projected onto the boundary of that polygon. See Figure 1.1 for an example.

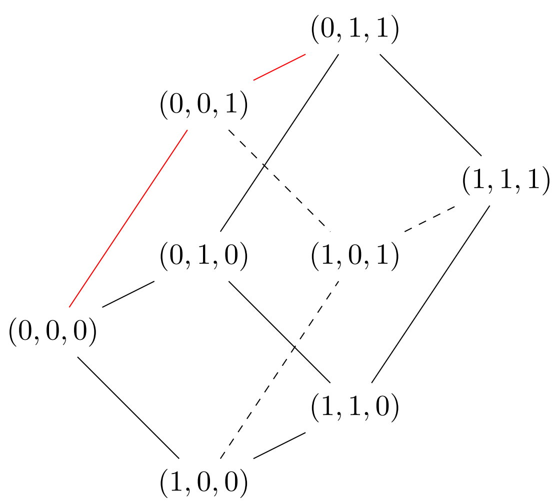

FIGURE 1.1. Pictured is an example of a shadow pivot rule for the simplex method on the cube $[0, 1]^n$ for maximizing $\mathbf{c} = (-2, 1, 3)$ starting at the vertex $(0, 0, 0)$. In particular, the cube is drawn in the plane using coordinates given by the shadow, so the boundary of the picture is exactly the image of the cube under the shadow projection $\pi(\mathbf{x}) = (\mathbf{w}^\intercal \mathbf{x}, \mathbf{c}^\intercal \mathbf{x})$. To see this, we first put the linear program in standard form, so that the linear program is the problem of maximizing $(-2, 1, 3, 0, 0, 0)^\intercal (\mathbf{z})$ subject to $\mathbf{z} \in \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{3+3} \geq \mathbf{0} : \mathbf{x} + \mathbf{y} = (1, 1, 1)\}$. Then starting at the vertex $(0, 0, 0, 1, 1, 1)$, we choose $\mathbf{w} = (1, 1, 1, 0, 0, 0)$ to obtain this shadow.

Note that it is not that case that all possible vectors $\mathbf{w}$ may be chosen. Namely, $\mathbf{w}$ must satisfy $\mathbf{w}_i = 0$ for all $i \in B$ and $\mathbf{w}_i > 0$ for all $i \notin B$. All of these pivot rules and essentially any other pivot rule that has been studied are known to have examples in which they run in exponential time [4, 61, 70, 95, 107, 110]. However, the following problem remains open:

Open Problem 1.1.1 ([49]). *Does there exist a pivot rule such that the simplex method runs in polynomial time?*

While this problem remains open and there are many negative results, there has been remarkable success in analyzing the performance of the simplex method in random models. In the 1970s Borgwardt analyzed the behavior of the simplex method for a random shadow pivot rule (i.e., a random choice of $\mathbf{w}$) on random linear programs taken from any spherically symmetric probability distribution and found that for an LP on a polytope given by $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ given by m inequalities in n variables, the expected total number of pivots taken is $\Theta(n^2 m^{\frac{1}{n-1}})$ [31]. More recently, in 2001, Spielman and Teng introduced the smoothed analysis of algorithms in [127] to study the expected performance of the simplex method with more recent improvements given in [48, 85, 137]. In this case, each entry of A and $\mathbf{b}$ is perturbed by a Gaussian with standard deviation σ, and the best known upper bounds on the expected number of steps in this framework is $O(\sigma^{-3/2} n^{13/4} \log^{7/4}(m))$ from [85]. Thus, in those two models of random linear programs, the simplex method performs in strongly polynomial time in expectation. One may also find weaker bounds using random shadows for fixed linear programs in terms of m and n alongside a measure of the size of the data measured by Δ, the maximum absolute value of a subdeterminant of the constraint matrix, or other similar parameters [35, 46] in which they find bounds of the form $\text{poly}(m, n, \Delta)$ for the expected performance of the shadow simplex algorithm. Though one should note that their bounds require either changing the linear program in [35] or generating multiple paths in [46].

In the deterministic setting, Kitahara and Mizuno proved in [93] that one may bound the performance of Dantzig's pivot rule in terms of the ratio of largest and smallest possible entries of any basic feasible solution, γ and δ, for an LP in standard form. In this setting, they achieved bounds of the form: $n\lceil m\frac{\gamma}{\delta}\log\left(m\frac{\gamma}{\delta}\right)\rceil$. The ratio $\frac{\gamma}{\delta}$ can be of exponential size, so these bounds are far from what is known in the random setting. Furthermore, the ratio $\frac{\gamma}{\delta}$ is NP-Hard to compute, since γ is polynomial time computable and δ is NP-Hard to compute [100]. These are, however, the best known bounds of their type.

The simplex method always follows a monotone path on the polytope given by its non-degenerate pivots. In particular, if a non-degenerate pivot occurs, that always corresponds to moving along an edge of the polytope. The run-time of the simplex method is then bounded by the number of pivots multiplied by the amount of time it takes to compute a single pivot. For our purposes, we are interested purely in bounding the number of non-degenerate pivots for certain pivot rules for the simplex method on certain classes of polytopes.

Even this problem is hard. Furthermore, the restriction to non-degenerate pivots here is partially justified by recent work of Kukharenko and Sanità [99] in which they show that there always exists a sequence of at most $m - n$ degenerate pivots necessary to compute some single nondegenerate pivot and provide an algorithm for finding that sequence of degenerate pivots given one knows an increasing direction. Our starting point for our work is the following 1997 question asked by Kortenkamp, Richter-Gebert, Sangarajan, and Ziegler on page 2 of [98]:

QUESTION 1.1.2. *Is there a pivot rule for the simplex method that guarantees a polynomial number of non-degenerate pivots on* $0/1$*-polytopes?*

1.2. Our Contributions: The Simplex Method for 0/1-Polytopes and Lattice Polytopes

All theorems in this section are proven in this thesis with earlier versions appearing in joint work with De Loera, Kafer, and Sanità [25]. A $0/1$-**polytope** is a polytope for which all vertices are in $\{0,1\}^n$, and the polynomial here is in terms of the number of facets and the dimension. One clear lower bound for the number of non-degenerate pivots needed to be taken by the simplex method is the **combinatorial diameter** of the polytopes, which is the maximal number of edges in a shortest path between two vertices in the graph of the polytope. Naddef showed in 1989 in [111] that the diameter of a 0/1-polytope is at most its dimension. However, this bound is not constructive. We resolve this question optimally with two new pivot rules for the simplex method.

THEOREM 1.2.1. *Given an LP on a d-dimensional* $0/1$*-polytope in* $\mathbb{R}^n$*, we have the following bounds:*

- *The slim shadow pivot rule solves the LP in at most n non-degenerate pivots.*
- *The ordered shadow pivot rule solves the LP in at most d non-degenerate pivots.*

Note that we do not bound degenerate steps. The purpose behind our result is to answer Question 1.1.2 and make Naddef's bounds constructive. For non-degenerate pivots, our bounds are optimal for the ordered shadow pivot rule and for the slim shadow pivot rule, when the polytope is full dimensional. At first glance, this may make one think that the ordered shadow pivot rule bounds are strictly stronger. However, the slim shadow pivot rule bounds improve with sparsity.

THEOREM 1.2.2. *Given an LP on a d-dimensional $0/1$-polytope in $\mathbb{R}^n$ such that the number of nonzero coordinates of the polytope is always $\min(k, n-k)$, the slim shadow pivot rule takes at most k non-degenerate pivots to solve the LP. If the number of nonzero coordinates is at most k, then it takes at most $2k$ non-degenerate pivots.*

Note that the bound of $\min(k, n-k)$ is optimal and attained for the hyper-simplex $\Delta(n, k)$, the convex hull of all $0/1$ vectors with support of cardinality k. For the case of nonzero coordinates being at most k, this improves upon a bound of Chubanov of $O(n \log(m))$, where m is the number of facets of the $0/1$ polytope, in Corollary 4.1 of [44]. By applying this theorem, we see that our results specialize to nearly optimal bounds for the **monotone diameter**, the worst-case number of edges needed to cross to reach a sink across all orientations of the graph induced by a linear function, for certain $0/1$-polytopes from combinatorial optimization. The monotone diameter similarly lower bounds the number of non-degenerate pivots the simplex method must take. We compare the slim shadow pivot rule's performance on various special instances:

- The rule yields at most n steps on both the asymmetric and symmetric traveling salesman polytopes for n vertex graphs. The respective optimal bounds are $\lfloor n/2 \rfloor$ and $\lfloor n/3 \rfloor$ [120].
- The rule yields at most n steps on the Birkhoff polytope for $n \times n$ matrices. The optimal bound is $\lfloor n/2 \rfloor$ [119].
- The rule yields at most $\lfloor n/2 \rfloor$ steps for the perfect matching polytope on the complete graph with n vertices. The optimal bound is $\lfloor n/4 \rfloor$ [119].
- The rule yields at most $\operatorname{rank}(\mathcal{M})$ for an independent set matroid polytope of a matroid $\mathcal{M}$ (starting at $\mathbf{0}$). As we saw this matches the optimal bound.

We were not the first to approach this problem, and there were various special cases resolved in other contexts. Namely, Chubanov studied two similar approaches to ours in [42, 44], but these were on generalizations of the simplex method in an oracle model. In our case, we study the simplex method as is applied in practice. We describe explicitly how one can implement each of our pivot rules on a computer. Del Pia and Michini also studied this problem in the more general setting of lattice polytopes and proved polynomial bounds, but they only do this by progressively changing the linear objective function [57]. We are able to follow a short path with a single run of the simplex method with precisely the same linear objective function they give us but with a particular choice of pivot rules. Finally, if the polytope is in standard-form and 0/1, polynomial bounds are implied by work of Kitahara and Mizuno in [92, 93]. However, for a general 0/1-polytope, converting it to standard form does not preserve being 0/1, and the slacks used in the bounds of Kitahara and Mizuno may blow up in size. Therefore, to our knowledge, no previous work had actually found a version of the simplex method guaranteed to take a polynomial number of non-degenerate steps to solve any LP on a 0/1-polytope. This concludes the results of [25] covered in this thesis.

The natural next step beyond 0/1-polytopes was to study $(0, k)$-lattice polytopes, polytopes whose vertices lie in $\{0, 1, \dots, k\}^n$. In this context, Kleinschmidt and Onn proved in [96] that their diameter is at most dk. In follow up work, Del Pia and Michini improved on these bounds in [56] and Deza and Pournin further improved on them in [59] without changing the asymptotics. In this case, the bound is conjecturally not tight. Though the lower bounds are of a similar order and have been studied in detail for the case of lattice zonotopes yielding surprising connections to number theory [58, 60]. Naturally, the next core question was whether one could match these bounds with a pivot rule for the simplex method. However, unlike for 0/1-polytopes, there is a major additional difficulty. While diameters of lattice polytopes are well known, the following remains only a conjecture:

CONJECTURE 1.2.1. *The monotone diameter of a d-dimensional $(0, k)$-lattice polytope is bounded by a polynomial in d and k.*

This question was the primary motivation behind my work in [22] in which I proved a series of partial results in the direction of this conjecture and made the bounds of Kleinschmidt and Onn constructive. The first step in this direction was to understand the case in which $k = 2$.

THEOREM 1.2.3. *The monotone diameter of a d-dimensional $(0, 2)$-lattice polytope $P \subset \mathbb{R}^n$ is at most $3d$. Furthermore, there is a pivot rule for the simplex method that is guaranteed to take at most $d + 2n$ non-degenerate pivots to solve an LP on P.*

For a general lattice polytope, I did find a bound on the performance of the simplex method so long as one allows one extra parameter.

THEOREM 1.2.4. *Let* $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ *be a* $(0,k)$*-lattice polytope defined by an integer matrix* A*. Then the simplex method with the slim shadow pivot rule is guaranteed to take at most* $n^2 k \|A\|_\infty$ *non-degenerate steps to solve any LP on* P*.*

This is also, to my knowledge, the best known general monotone diameter bound for $(0,k)$-lattice polytopes. The key parameter we want to remove or reduce dependency on is $\|A\|_\infty$. We do not achieve this for monotone diameters, but we are able to prove a partial result taking inspiration from Del Pia and Michini in [**57**]. Namely they show that on a lattice polytope, one may find a path of length at most $O(d^4 k \log(dk))$ between any pair of vertices by running the simplex method multiple times with different linear objective functions and warm starting.

THEOREM 1.2.5. *Let* $P \subseteq \mathbb{R}^n$ *be a* $(0,k)$*-lattice polytope. Let*

$$\mathbf{r} = (r, r^2, \ldots, r^n)$$

for r *sufficiently large. Then* P *has a unique* $\mathbf{r}$*-minimum* $\mathbf{v}$*. Furthermore, for any* $\mathbf{c} \in \mathbb{R}^n$*, there exists* r*, such that the shadow pivot rule for* $\mathbf{r}$ *to maximize* $\mathbf{c}$ *for the path starting at* $\mathbf{v}$ *takes at most* dk *non-degenerate steps.*

There is a notion of diameter lying between monotone diameter and combinatorial diameter called the **strict monotone diameter** as introduced in Chapter 3 of [**138**], which is the worst-case number of steps needed to walk from a minimum to a maximum across all orientations of the graph of a polytope induced by a linear objective function. As a corollary of Theorem 1.2.5, we have the following result:

COROLLARY 1.2.1. *The strict monotone diameter of a* d*-dimensional* $(0,k)$*-lattice polytope is at most* $2dk$*.*

Theorem 1.2.5 gives rise to the following algorithmic result.

COROLLARY 1.2.2. *Let* $P \subseteq \mathbb{R}^n$ *be a* $(0,k)$*-lattice polytope of dimension* d*, and let* $\mathbf{c} \in \mathbb{R}^n$*. Then one may solve the LP* $\max(\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P)$ *by first in Phase* 1 *finding an optimal basis for minimizing*

$(r, r^2, \ldots, r^n)$. *Then, for any $r \in \mathbb{R}$ sufficiently large, the simplex method with shadow pivot rule from auxiliary vector $\mathbf{r} = (r, r^2, \ldots, r^n)$, will solve the LP in at most dk non-degenerate steps.*

Note that this Phase 1 procedure may be done efficiently with the simplex method. This theorem may be understood as a way to make Kleinschmidt and Onn's original bound constructive and translated into the actual performance on an LP solver for lattice polytopes. Finally, there is an analogous class of polytopes to $(0, k)$-lattice polytopes called **$(\ell + 1)$-level polytopes**. These are polytopes for which each facet direction $\mathbf{w}$ takes on at most $\ell + 1$ distinct values on vertices of the polytope. For $\ell = 1$, these are all called 2-level polytopes or compressed polytopes [128]. Each 2-level polytope is affinely isomorphic to a 0/1-polytope, and these polytopes appear in many contexts in combinatorial optimization [7]. Then $(\ell + 1)$-level polytopes are to 2-level polytopes as $(0, k)$-lattice polytopes are to 0/1-polytopes. We prove the following:

THEOREM 1.2.6. *Let P be a d-dimensional $(\ell + 1)$-level polytope. Then the monotone diameter of P is at most $d\ell$, and this bound is achieved via a shadow pivot rule.*

1.3. Monotone Path Polytopes

For this section, we will try to understand monotone paths on polytopes from a combinatorial standpoint. We will start with a review of polytope theory and then move to a discussion of monotone paths.

We will recall the basics of polytopes at the level of Ziegler's book [138]. In particular, for our purposes we care most about Chapter 9 of [138] but must recall results from Chapters $2, 3$, and 7. A polytope has two equivalent definitions. The first is defined in terms of convex geometry. Recall that a set $C \subseteq \mathbb{R}^n$ is called **convex** if for all $x, y \in C$, the line segment from x to y is contained in C. That is $\lambda x + (1 - \lambda) y \in C$ for all $\lambda \in [0, 1]$. Note that the intersection of convex sets must still be convex, since for all $x, y \in C \cap D$ the line segment from x to y must also be in both C and D by definition for convex C and D. Hence, given any set S, one may take the intersection of all convex sets containing S to arrive at a unique smallest convex set containing S called the convex hull $\mathrm{conv}(S)$.

DEFINITION 1.3.1. *A **polytope** is the convex hull of a finite subset $S \subseteq \mathbb{R}^n$.*

That is the first definition. The second definition is grounded in linear algebra as opposed to convex geometry. Namely a **polyhedron** is the solution set to a system of linear inequalities $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$, where for vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{u} \leq \mathbf{v}$ if $\mathbf{u}_i \leq \mathbf{v}_i$ for all $i \in [n]$. For example, the first quadrant in $\mathbb{R}^2$ is a polyhedron defined by the system of inequalities $x \geq 0$ and $y \geq 0$. Recall that a set S is **bounded** if every point is within a fixed distance δ of the origin. The relation between polyhedra and polytopes is as follows:

THEOREM 1.3.2 (Weyl-Minkowski). *A set P is a polytope if and only if P is a bounded polyhedron.*

There are many proofs of this fundamental theorem. See [**138**] and [**13**] for two distinct approaches. For example, the hyper-cube $[0,1]^n$ is given by the system of inequalities $0 \leq \mathbf{x}_i \leq 1$ for all $i \in [n]$. It is also $\text{conv}(\{0,1\}^n)$.

The combinatorial structure of polytopes is fundamentally related to optimization. For a polytope P, a subset $F \subseteq P$ is a **face** if $F = \text{argmax}_{\mathbf{x} \in P}\, \mathbf{c}^\intercal \mathbf{x}$ for some $\mathbf{c} \in \mathbb{R}^n$. A geometric way of understanding this definition is by noticing that if take the hyper plane given by $\{\mathbf{x} : \mathbf{c}^\intercal \mathbf{x} = \max_{\mathbf{x} \in P}(\mathbf{c}^\intercal x)\}$, this will intersect P precisely at F, and all points in P lie on one side of this hyperplane. Such a hyperplane is called a **supporting hyperplane** of P. From the definition of a polytope as a bounded polyhedron, it is immediate that the intersection of a polytope with a hyperplane is still a polytope. Hence, each face of P is still a polytope. The set of faces of P with respect to set inclusion is called the **face lattice** of P. In Chapter 2 of [**138**], Ziegler proves that this is actually a lattice. In particular, it is a ranked poset, where the rank is given by the dimension of the faces, where the dimension of a polytope is the dimension of the smallest affine linear subspace containing it.

Consider again the case of a hyper-cube. Let $\mathbf{c} \in \mathbb{R}^n$, and suppose I want to maximize $\mathbf{c}^\intercal \mathbf{x}$ such that $\mathbf{x} \in [0,1]^n$. Then I hope to maximize $\sum_{i=1}^n \mathbf{c}_i \mathbf{x}_i$ such that $\mathbf{x} \in [0,1]^n$. Then there are three cases to consider. Suppose that $\mathbf{c}_i > 0$. Then, since $\mathbf{x}_i \leq 1$, $\mathbf{c}_i \mathbf{x}_i \leq \mathbf{c}_i$. Suppose that $\mathbf{c}_i < 0$. Then, since $\mathbf{x}_i \geq 0$, $\mathbf{c}_i \mathbf{x}_i \leq 0$. Suppose that $\mathbf{c}_i = 0$. Then $\mathbf{c}_i \mathbf{x}_i = 0$. Let S_+ be the set of i such that $\mathbf{c}_i > 0$, S_- be the set of i such that $\mathbf{c}_i < 0$, and let S_0 be the set of i such that $\mathbf{c}_i = 0$. Then we have for all choices of $\mathbf{x} \in [0,1]^n$,

$$\mathbf{c}^\intercal \mathbf{x} \leq \sum_{i \in S_+} \mathbf{c}_i + \sum_{j \in S_-} (0)\mathbf{c}_j + \sum_{k \in S_0} \mathbf{y}_k \mathbf{c}_k,$$

for all $\mathbf{y} \in [0,1]^{S_k}$. Furthermore, for all choices of $\mathbf{y} \in [0,1]^{S_k}$,

$$\mathbf{c}^{\mathsf{T}} \left(\sum_{i \in S_+} e_i + \sum_{k \in S_0} \mathbf{y}_k e_k \right) = \sum_{k \in S_+} \mathbf{c}_k.$$

Thus, the faces of the hyper-cube are in bijection with partitions of $[n]$ into three sets S_0, S_+, and S_-, where the coordinates in S_+ are all equal to 1, the coordinates in S_- are always equal to 0, and the coordinates in S_0 form the hyper-cube $[0,1]^{S_0}$. The ordering on the face lattice is given by $(S_0, S_+, S_-) \geq (T_0, T_+, T_-)$ if $S_0 \supseteq T_0$, $S_+ \subseteq T_+$, and $S_- \subseteq T_-$. The dimension of the face is given by $|S_0|$, since it is the image of a hyper-cube of dimension $|S_0|$ under an affine linear isomorphism. Throughout, we will also be considering operations on polytopes to create new polytopes from old ones. For example, the product of polytope $P \times Q = \{(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in P, \mathbf{y} \in Q\}$ is also a polytope. Namely, it is a polyhedron, since $\{(\mathbf{x}, \mathbf{y}) : A\mathbf{x} \leq \mathbf{b} \text{ and } A'\mathbf{y} \leq \mathbf{b}'\}$, where $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ and $Q = \{\mathbf{y} : A'\mathbf{y} \leq \mathbf{b}'\}$. Furthermore, it is bounded, since the product of bounded sets is bounded. The operations that will play a fundamental role throughout this work are the following:

Theorem 1.3.3 ([138]). *Let $P \subseteq \mathbb{R}^n$ be a polytope, $\pi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, and $H \subseteq \mathbb{R}^n$ be a hyperplane. Then $\pi(P)$ and $H \cap P$ are both polytopes. Furthermore, for each face F if $\pi(P)$, $\pi^{-1}(F)$ is a face of P, and the set of faces of $H \cap P$ is precisely the set $\{H \cap F : F \text{ is a face of } P\}$.*

Proof. See the proof of Lemma 7.10 in [138] for the case of projections, and the remaining results for sections follow from polar duality.

$\square$

For our purposes, we are primarily concerned with the zero dimensional faces and one dimensional faces of the polytope, called the vertices and edges of the polytope respectively. In particular, a point is a vertex of P if there exists $\mathbf{c} \in \mathbb{R}^n$ such that $\mathbf{c}^{\mathsf{T}}\mathbf{x}$ is uniquely maximized at $\mathbf{v}$. The set of vertices and edges of a polytope is called its **graph** or **one-skeleton**, where two vertices are adjacent if they are contained in the same edge. We will use the term graph throughout. It is also important to consider the codimension one faces called the **facets** of the polytope. The vertices and facets both yield minimal descriptions of the polytope:

Theorem 1.3.4. *Let P be a polytope. Let $V(P)$ be its set of vertices. Then $V(P)$ is the unique minimal set of points S in $\mathbb{R}^n$ such that $P = conv(S)$. Similarly, each facet F spans an affine* *hyperplane with a unique normal vector* $\mathbf{c}_F$ *such that* $P \subseteq \{\mathbf{x} : \mathbf{c}_F^\intercal \mathbf{x} \leq b_F\}$*. Then the unique smallest system of linear inequalities defining* P *is precisely given by* $\{\mathbf{c}_F^T\mathbf{x} \leq b_F : F$ *is a facet of* $P\}$*.*

Thus, the number of facets gives a measure of how big the polytope is. Note that the graph of P is always connected and, in fact, d-connected where d is the dimension of the polytope by Balinski's theorem. The combinatorial diameter of a polytope is the maximal length of a shortest path between vertices of the polytope. For example, the combinatorial diameter of a hyper-cube is its dimension. One of the biggest open problems in polytope theory is the following:

Conjecture 1.3.1 (Polynomial Hirsch Conjecture). *There exists a polynomial* $p(m, d)$ *such that the combinatorial diameter of a* d*-dimensional polytope with* m *facets is at most* $p(m, d)$*.*

Before the polynomial Hirsch conjecture, there was simply the Hirsch conjecture which asked for a bound of at most $m - d$ on the combinatorial diameters of polytopes. This bound stood for 50 years before being disproven by Santos in [**122**]. However, the best lower bounds remain on the order of $1.05(m - d)$ due to Santos' original construction and follow up work in [**108**]. The smallest counterexample known is in dimension 20 and has 40 facets with combinatorial diameter 21. The Hirsch conjecture holds for polygons, since the number of vertices equals the number of edges. It also holds in dimension 3 [**94**]. However, the Hirsch conjecture remains open in dimension 4. The best upper bounds for diameters of polytopes in terms of m and d are quasi-polynomial of the form $O(m^{1+\log(d)})$ due to Kalai and Kleitman [**87**] with mild improvements by Todd [**136**] and Sukegawa [**130**]. There is a linear bound in fixed dimension of the form $O(2^d n)$ due to Barnette [**12**] and Larman [**101**].

There is a substantial body work studying diameters of polytopes as discussed in [**91**] as well as Chapter 3 of [**138**] and Chapter 16 of [**79**]. The most recent of these surveys was in 2010, and there has continued to be activity in this area since then. For example, one line of research has been to study diameters of polytopes in terms of m, d, and the maximal absolute value of a subdeterminant Δ of the constraint matrix A for a polytope defined by $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$, where A is assumed to have integer coordinates. When $\Delta = 1$, the constraint matrix is called **totally unimodular**, and in that setting Dyer and Frieze showed in [**62**] a diameter bound of at most $O(m^{16}n^3 \log(mn)^3)$. The motivation for that setting comes from combinatorial optimization in which many polytopes have totally unimodular constraint matrices. Vastly improving and generalizing these bounds, Bonifas, Di Summa, Eisenbrand, Hähnle, and Niemeier [30] showed a bound of at most $O(\Delta^2 d^{3.5} \log(d\Delta))$. Dadush and Hähnle later improved these bounds in [47] to $O(d^3\Delta^2 \log(d\Delta))$ using similar techniques and bounds as in [35]. Similar bounds have been proven most recently in [112]. However, the Hirsch conjecture remains open even in the case of polytopes with totally unimodular constraint matrices. Though it is known in the important special case of network flow polytopes [32] and there has been recent work in this direction [33, 34].

The primary reason to study combinatorial diameters of polytopes is to better understand the simplex method for linear programming. In the previous section, we discussed the simplex method in detail. In this section, we primarily need to know the following: It is one of the most fundamental open problems in the theory of linear programming to find out whether there is a version of the simplex method that runs in polynomial time stated for example by Dantzig in 1963 in [49].

Remark 1.3.1. *If there is a version of the simplex method that runs in polynomial time, then the polynomial Hirsch conjecture must be true.*

In the context of our work, we are concerned with a variant of the polynomial Hirsch conjecture, which has received comparably less attention. Namely, there is an oriented notion of the combinatorial diameter, which is more closely to the application of linear programming. Namely, given a fixed choice of a vector $\mathbf{c} \in \mathbb{R}^n$, one can orient the graph of a polytope, where an edge $(\mathbf{u}, \mathbf{v})$ is oriented from $\mathbf{u}$ to $\mathbf{v}$ is $\mathbf{c}^\intercal \mathbf{u} < \mathbf{c}^\intercal \mathbf{v}$. As explained in Chapter 3 of [138], this oriented graph is always connected. The monotone diameter of a polytope is the worst case number of edges needed to walk to a sink in an oriented graph of the polytope across all orientations induced by some choice of $\mathbf{c}$.

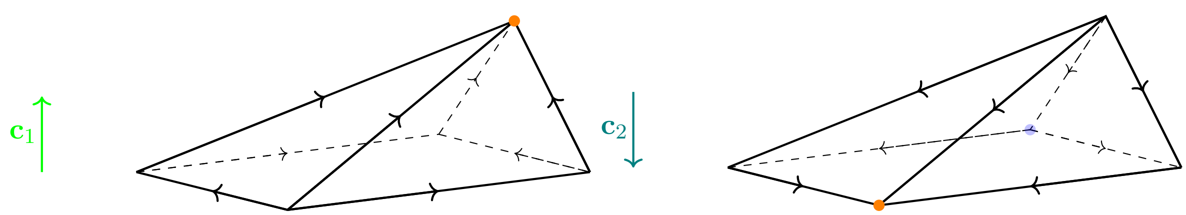

Figure 1.2. Pictured are two orientations of the square pyramid. For an orientation optimized at the pyramid point, the worst-case monotone distance to the optimum is always 1. However, on the right, an orientation is given for which the worst case distance to the optimum is 2. This showcases that the monotone diameter of the square pyramid is at least 2. To compute the monotone diameter, one must bound the worst-case number of steps needed to reach an optimum across all possible orientations induced by a linear function.

In general, the combinatorial diameter is a lower bound for the monotone diameter. They can be distinct. For example, the convex hull of the set of permutation matrices in $\mathbb{R}^{n^2}$, called the **Birkhoff polytope**, has combinatorial diameter two [11], but the monotone diameter is $n/2$ [119]. Little is known about monotone diameters. In particular, the natural analog to the polynomial Hirsch conjecture remains open:

CONJECTURE 1.3.2 (Polynomial Monotone Hirsch Conjecture). *There is a polynomial $p(m, d)$ such that any polytope with m facets and d dimensions has monotone diameter at most $p(m, d)$.*

The polynomial monotone Hirsch conjecture has the same motivation as the Hirsch conjecture:

REMARK 1.3.2. *If there is a version of the simplex method that runs in polynomial time, then the polynomial monotone Hirsch conjecture must be true.*

Before the polynomial monotone Hirsch conjecture, there was also the originally monotone Hirsch conjecture, which asked for a bound of at most $m - d$ for monotone diameters. Unlike for combinatorial diameters, the monotone Hirsch conjecture was disproven much earlier than the general Hirsch conjecture by Todd in 1980 in [135]. Besides Santos' counterexamples to the Hirsch conjecture, all known counterexamples to the monotone Hirsch conjecture have been constructed from Todd's 1980 example, and the best lower bound we have for the worst case monotone diameter of a polytope is $1.25(m - d)$. An advantage of Todd's counterexample is that it is remarkably simple. The counterexample is in dimension 4 with 8 facets and monotone diameter 5. Namely, it is given by $M_4 = \{\mathbf{x} : \mathbf{x} \geq \mathbf{0} \text{ and } A\mathbf{x} \leq \mathbf{b}\}$ with

$$A = \begin{pmatrix} 7 & 4 & 1 & 0 \\ 4 & 7 & 0 & 1 \\ 43 & 53 & 2 & 5 \\ 53 & 43 & 5 & 2 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 8 \\ 8 \end{pmatrix}.$$

I studied this example in the context of circuit augmentation schemes for linear programming with Steffen Borgwardt and Matthias Brugger in [23]. Along the way, we found there were precisely 7112 orientations of the graph of M_4 that are induced by a vector. Of these, only 5 contradict the conjectured bound meaning that, for all but 5 orientations, the worst case distance in the directed graph to a sink is 4. Todd only provided one orientation, but in [23], we described all 5 of them.

A path in an oriented graph of a polytope from an initial vertex to a sink for the orientation induced by a vector $\mathbf{c}$ is called a **$\mathbf{c}$-monotone path** or just a **monotone path**. The monotone diameter of a polytope is the longest length of a shortest $\mathbf{c}$-monotone path on a polytope across all choices of $\mathbf{c}$. There has not been any improvement on the lower bounds on the worst-case monotone diameters of polytopes in general in the 44 years since Todd's discovery of M_4. The primary motivation behind our work here is to develop new tools and examples to better understand monotone paths on polytopes. Namely, we fix a polytope P and a direction $\mathbf{c} \in \mathbb{R}^n$ and study all $\mathbf{c}$-monotone paths from a $\mathbf{c}$-minimizer to a $\mathbf{c}$-maximizer.

Example 1.3.5. *Consider all monotone paths for the orientation on the hyper-cube* $[0,1]^n$ *induced by* $\mathbf{c} = \sum_{i=1}^{n} e_i$. *As mentioned previously the vertices of a hyper-cube are* $\{0,1\}^n$. *Two vertices are adjacent if they are equal in all but one coordinate. In other words the vertices correspond to subsets* $S \subseteq [n]$, *and* S *is adjacent to* $S \cup \{x\}$ *and* $S \setminus \{y\}$ *for all* $x \in [n] \setminus S$ *and* $y \in S$. *For the orientation induced by* $\mathbf{c}$, *the directed edges from* S *precisely to* $S \cup \{x\}$ *for each* $x \in [n] \setminus S$. *In particular, the resulting graph is the Hasse diagram of the Boolean lattice (i.e., the poset of all subsets of* $[n]$ *with respect to inclusion). Then monotone paths are in bijection with sequences of subsets of the form:*

$$S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_{n-1} \subsetneq S_n$$

such that $|S_i \setminus S_{i-1}| = 1$ *for all* $i \in [n]$. *These are also called a* ***complete flag*** *of subsets. Complete flags of subsets are in bijection with permutations of* $[n]$ *via the map sending* $(S_0, S_1, \ldots, S_n)$ *to* $(S_1 \setminus S_0, S_2 \setminus S_1, \ldots, S_n \setminus S_{n-1})$. *Hence, the monotone paths on the hyper-cube encode exactly the permutations of* $[n]$. *See Figure 1.3 for an illustration.*

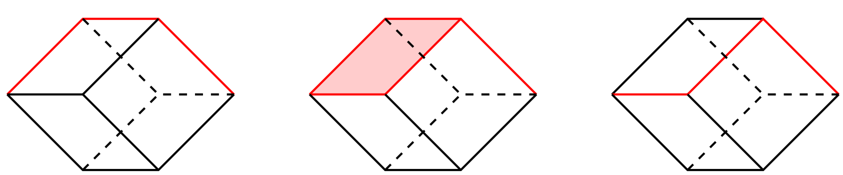

Figure 1.3. On the left is a monotone path on the cube given by $(0,0,0), (1,0,0), (1,1,0), (1,1,1)$, and on the right is the monotone path $(0,0,0), (0,1,0), (1,1,0), (1,1,1)$. Under the bijection, they correspond to $\emptyset \subsetneq \{1\} \subsetneq \{1,2\} \subsetneq \{1,2,3\}$ and $\emptyset \subsetneq \{2\} \subsetneq \{1,2\} \subsetneq \{1,2,3\}$ respectively. These two monotone paths are adjacent, and the cellular sting corresponding to the edge between them is drawn in the center and corrsponds to the flag of subsets $\emptyset \subseteq \{1,2\} \subseteq \{1,2,3\}$.

Our focus for studying monotone paths on polytopes is the **monotone path polytope**, a construction introduced by Billera and Sturmfels as part of their general theory of fiber polytopes in [18] with a textbook introduction given in Chapter 9 of [138]. The image of a polytope under a linear map is also a polytope. The fiber polytope is a way to associate a polytope to a linear map between polytopes $\pi : P \to Q$. For example, the fiber polytope of a projection of the standard simplex $\text{conv}(e_1, e_2, \ldots, e_n)$ is the **secondary polytope** introduced by Gelfand, Kapranov, and Zelevinsky in [75] for their study of Newton polytopes of multi-dimensional analogs of determinants and discriminants. Secondary polytopes have played a broad role. They for example appear in Gröbner basis theory [129] and tropical geometry in relation to tropical hypersurfaces [104]. See Chapter 5 of [55] for a detailed discussion of secondary polytopes and their many appearances. The monotone path polytope is the fiber polytope of a one-dimensional projection $\varphi : \mathbb{R}^n \to \mathbb{R}$ with $\varphi(\mathbf{x}) = \mathbf{c}^\intercal \mathbf{x}$ for some $\mathbf{c} \in \mathbb{R}^n$. However, we will not provide a further discussion of fiber polytopes in complete generality. Instead, our discussion of the monotone path polytope will be entirely self contained, and we will prove multiple equivalent formulations ourselves. The first step is to provide an explicit realization of the monotone path polytope. Recall from page 36 in [138] that the **Minkowski sum** of polytopes P and Q is $P+Q = \{x+y : x \in P, y \in Q\}$. The Minkowski sum of polytopes is always a polytope. Furthermore, let $R^{\mathbf{c}}$ denote the $\mathbf{c}$-maximizing face of a polytope R. Then $(P+Q)^{\mathbf{c}} = P^{\mathbf{c}} + Q^{\mathbf{c}}$ for all $\mathbf{c} \in \mathbb{R}^n$. In particular, the vertices of the Minkowski sum are precisely the sums of vertices of P and Q that are both maximal with respect to the same choice of $\mathbf{c}$. Using the Minkowski sum, we may define the monotone path polytope explicitly:

DEFINITION 1.3.6. *Given a polytope P and a vector $\mathbf{c} \in \mathbb{R}^n$, the monotone path polytope $\Sigma_{\mathbf{c}}(P)$ is given by*

$$\Sigma_{\mathbf{c}}(P) = \sum_{v \in V(P)} \frac{1}{|\{u \in V(P) : \mathbf{c}^\intercal \mathbf{u} = \mathbf{c}^\intercal \mathbf{v}\}|} P \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{c}^\intercal \mathbf{x} = \mathbf{c}^\intercal \mathbf{v}\}.$$

That is, it is the Minkowski sum of all the fibers containing vertices of P under the projection $\varphi(\mathbf{x}) = \mathbf{c}^\intercal \mathbf{x}$.

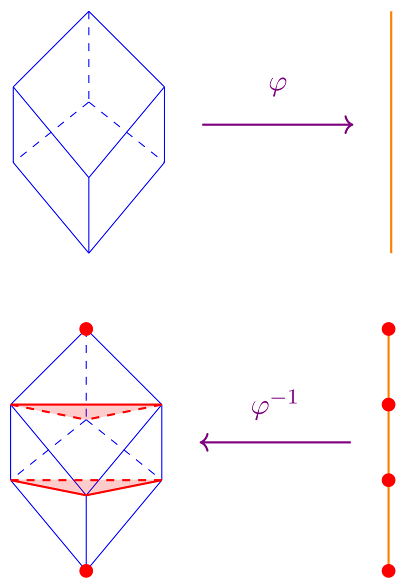

FIGURE 1.4. Pictured is the map from the cube to the interval given by $\varphi(\mathbf{x}) = (1,1,1)^{\intercal}\mathbf{x}$ together with each fiber of this map containing a vertex. Summing up all of the fibers yields the monotone path polytope.

See Figure 1.4 for an illustration of the summands in the definition. Neither the reason for this choice of fibers nor the connection to monotone paths are immediately obvious from this definition. The following is a key lemma for understanding why this choice of fibers.

LEMMA 1.3.6.1. *Let P be a polytope $\varphi : P \to \mathbb{R}$ be a linear map. Then let $P_0, P_1, \ldots, P_k$ be the set of fibers of φ that contain vertices of P. Let $\mathbf{x} \in \varphi(P)$. Then $\varphi^{-1}(x) = \lambda P_{i-1} + (1-\lambda)P_i$, where λ is chosen such that $x = \lambda\varphi(\mathbf{v}^{i-1}) + (1-\lambda)\varphi(\mathbf{v}^i)$ for some vertices $\mathbf{v}^{i-1}, \mathbf{v}^i$ of P contained in P_{i-1} and P_i respectively.*

PROOF. Let $x \in \varphi(P)$. Then if $x = \varphi(\mathbf{v})$ for some vertex v of P. Otherwise, there exists a unique index $i \in [k]$ such that $\varphi(\mathbf{v}^{i-1}) < x < \varphi(\mathbf{v}^i)$, where $\mathbf{v}^{i-1} \in P_{i-1}$ and $\mathbf{v}^i \in P_i$. Then $x = \lambda\varphi(\mathbf{v}^{i-1}) + (1-\lambda)\varphi(\mathbf{v}^i)$ for some $\lambda \in [0,1]$. We claim then that $\varphi^{-1}(x) = \lambda P_{i-1} + (1-\lambda)P_i$. Observe first that, by linearity, for all $\mathbf{a} \in P_{i-1}$ and $\mathbf{b} \in P_i$,

$$\varphi(\lambda\mathbf{a} + (1-\lambda)\mathbf{b}) = \lambda\varphi(\mathbf{a}) + (1-\lambda)\varphi(\mathbf{b}) = \lambda\varphi(\mathbf{v}^{i-1}) + (1-\lambda)\varphi(\mathbf{v}^i) = x.$$

Hence $\varphi(\lambda P_{i-1}+(1-\lambda)P_i) = \{x\}$, and by convexity $\lambda P_{i-1}+(1-\lambda)P_i \subseteq P$, so $\lambda P_{i-1}+(1-\lambda)P_i \subseteq \varphi^{-1}(x)$.

For the other direction, note first that by Theorem 1.3.4, since $\varphi^{-1}(x)$ is a polytope, it is determined by its vertices. Thus, it suffices to show that each vertex of $\varphi^{-1}(x)$ is contained in $\lambda P_{i-1}+(1-\lambda)P_i$. Note that $\varphi^{-1}(x) = P \cap H$, where H is the hyperplane given by $\{\mathbf{y} \in \mathbb{R}^n : \mathbf{c}^\intercal \mathbf{y} = x\}$. Hence, the faces of $P \cap H$ are precisely $\{H \cap F : F \text{ is a face of } P\}$ by Theorem 1.3.3. Note that H does not contain any vertex of P and therefore cannot contain any face of P, since each vertex of a face of P is a vertex of P (See Chapter 2 of [138]). Thus, each face H intersects of P must be intersected on its relative interior meaning that whenever $H \cap F \neq \emptyset$, $\dim(H \cap F) = \dim(F) - 1$. In particular, the vertices of $\varphi^{-1}(x)$ are precisely the intersection of H with an edge. Let $\mathbf{v}$ be one such vertex. For such an edge with endpoints $\mathbf{s}$ and $\mathbf{t}$, we must $\varphi(\mathbf{s}) < x < \varphi(\mathbf{t})$. In particular,

$$\varphi(\mathbf{s}) \leq \varphi(\mathbf{v}^{i-1}) < x < \varphi(\mathbf{v}^{i+1}) \leq \varphi(\mathbf{t}),$$

so the edge from $\mathbf{s}$ to $\mathbf{t}$ must intersect P_{i-1} and P_i meaning that $\mathbf{v}$ lies on a segment from P_{i-1} to P_i. Since $\varphi(\mathbf{v}) = x$, we must therefore have $\mathbf{v} \in \lambda P_{i-1} + (1-\lambda)P_i$ meaning that $\varphi^{-1}(x) \subseteq \lambda P_{i-1} + (1-\lambda)P_i$ as desired. $\square$

Thus, the fibers of the image of a polytope under a linear functional $\varphi : \mathbb{R}^n \to \mathbb{R}$ are precisely the interpolations of all fibers containing vertices. We have the following corollary:

Corollary 1.3.1. *Let $P \subseteq \mathbb{R}^n$ be a polytope and $\mathbf{c} \in \mathbb{R}^n$. Then a vertex of $\Sigma_{\mathbf{c}}(P)$ uniquely determines an element of each fiber of P under the map $\varphi(\mathbf{x}) = \mathbf{c}^\intercal \mathbf{x}$. In particular, the set of fibers is a piecewise linear curve starting at a $\mathbf{c}$-minimizing vertex of P and ending at a $\mathbf{c}$-maximizing vertex of P.*

Proof. Let $\mathbf{v}$ be a vertex of the monotone path polytope. Let $P_0, P_1, \ldots, P_k$ be the set of slices of P in the Minkowski sum decomposition of $\Sigma_{\mathbf{c}}(P)$. Then $\mathbf{v}$ has a unique decomposition into $\sum_{i=0}^{k} \mathbf{v}^i$, where each $\mathbf{v}^i \in P_i$ for all $i \in [0,k]$. Then for each fiber $\varphi^{-1}(x)$, the corresponding element of $\varphi^{-1}(x)$ is $\lambda \mathbf{v}^{i-1} + (1-\lambda)\mathbf{v}^i$, where $x = \lambda\varphi(v^{i-1}) + (1-\lambda)\varphi(\mathbf{v}^i)$, and $i \in [k]$ is chosen uniquely so that $\varphi(\mathbf{v}^{i-1}) \leq x \leq \varphi(\mathbf{v}^i)$. In particular, the set of all vertices coming from each fiber form exactly the piecewise linear curve by linearly interpolating from $\mathbf{v}^{i-1}$ to $\mathbf{v}^i$ for each $i \in [k]$. $\square$

While we have not proven it yet, this curve is always a monotone path on the polytope, which ise where the construction gets its name. See Figure 1.5 for an example on the cube.

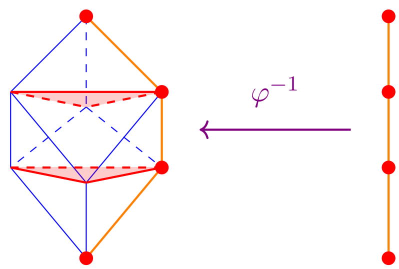

FIGURE 1.5. Pictured is each fiber of this map $\varphi : [0,1]^3 \to \mathbb{R}$ given by $\varphi(x) = (1,1,1)^\intercal \mathbf{x}$ containing a vertex. In the picture an example of a path is drawn by interpolating between vertices of the fibers.

In fact, while the resulting vertex is always a monotone path, not all monotone paths arise in this way. Those that do are called **coherent** paths. We call a path a $\mathbf{w}$-coherent $\mathbf{c}$-monotone path if corresponds to a $\mathbf{w}$-maximal vertex of the monotone path polytope $\Sigma_{\mathbf{c}}(P)$ for vectors $\mathbf{w}, \mathbf{c} \in \mathbb{R}^n$ and any polytope $P \subseteq \mathbb{R}^n$. There are multiple characterizations of coherence, which appear in differing references in the combinatorics literature [16, 18, 88, 138] and the optimization literature [31, 74]. While the following theorem is known from various sources, to my knowledge it has never appeared in print in a single location, and our proof is new.

THEOREM 1.3.7. *Let P be a polytope and $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n$. Then the following are equivalent for a $\mathbf{c}$-monotone path $\gamma \subseteq P$:*

(1) γ is $\mathbf{w}$-coherent.

(2) Shadow: $\gamma = \pi^{-1}(U)$, where $\pi(\mathbf{x}) = (\mathbf{c}^\intercal \mathbf{x}, \mathbf{w}^\intercal \mathbf{x})$ and U is the union of the set of all edges of $\pi(P)$ with outer normal vector with positive y coordinate.

(3) Slope: γ is built inductively start at $\mathbf{v}^0$, the $\mathbf{w}$-maximal vertex of the $\mathbf{c}$-minimal face of P and

$$\mathbf{v}^i = \operatorname*{argmax}_{\mathbf{u} \in N(v^{i-1})} \frac{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{v}^{i-1})}{\mathbf{c}^\intercal(\mathbf{u} - \mathbf{v}^{i-1})},$$

where $N(\mathbf{v}^{i-1})$ denotes the set of $\mathbf{c}$-increasing neighbors of $\mathbf{v}^{i-1}$.

(4) Parametric: $\gamma = \bigcup_{\lambda \in \mathbb{R}} P^{\lambda \mathbf{c} + \mathbf{w}}$, where $P^{\mathbf{t}}$ denotes the $\mathbf{t}$-maximal face of P for all $\mathbf{t} \in \mathbb{R}^n$.

We defer a detailed proof to Section A.2, but for now we add an informal explanation of the different formulation. For the shadow formulation, we define this set U, which is the **upper path** of a polygon and consists of all edges on the top of a polygon. Have an outer normal vector with positive y-coordinate is one way to determine whether a path is on the top of polygon. See Figure 1.6 for an example of a polygon together with its upper path.

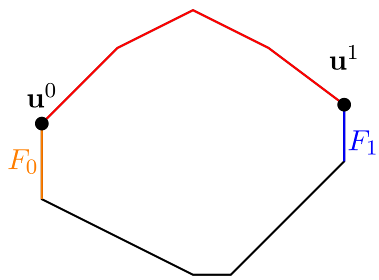

FIGURE 1.6. Pictured is the upper path of a polygon, which walks from the highest point $\mathbf{u}^0$ in its left-most face F_0 to the highest point $\mathbf{u}^1$ in its right-most face F_1.

What the shadow definition is saying is that a $\mathbf{w}$-coherent monotone path is the pre-image of the upper path in the image of P under the linear projection $\pi(\mathbf{x}) = (\mathbf{c}^\intercal \mathbf{x}, \mathbf{w}^\intercal \mathbf{x})$. The pre-image of this path will not necessarily be a monotone path, but when it is, that path is $\mathbf{w}$-coherent. In fact, this happens for a generic choice of $\mathbf{w}$. See Figure 1.7 for examples of projections of the tetrahedron for four different choices of $\mathbf{w}$ proving the coherence of all its monotone paths.

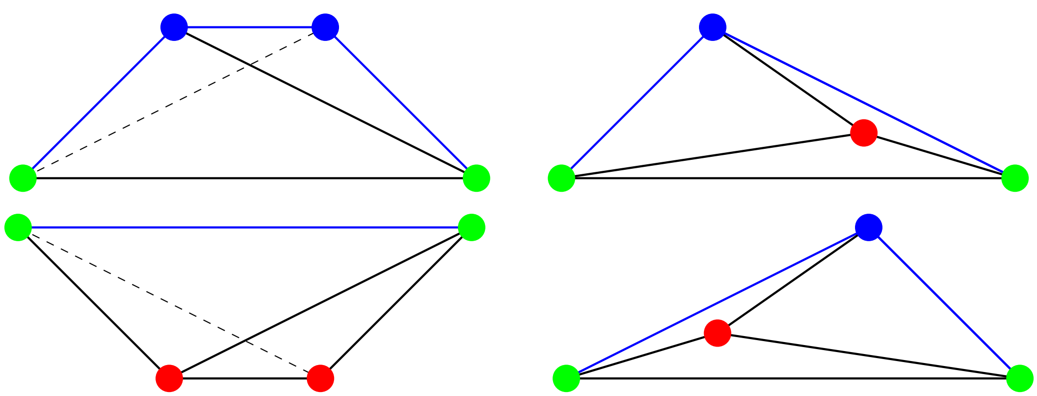

FIGURE 1.7. The four monotone paths on a tetrahedron $\operatorname{conv}(e_1, e_2, e_3, e_4)$ are depicted for $\mathbf{c} = (0, 1, 2, 3)$. The coherence of the top left and bottom left are shown using $\mathbf{w} = (0, 1, 1, 0)$ and $\mathbf{w} = (0, -1, -1, 0)$ respectively. The top right and bottom right are shown to be coherent using $\mathbf{w} = (0, 1, .3, 0)$ and $\mathbf{w} = (0, .3, 1, 0)$ respectively.

The slope formulation of coherence also can be seen in the shadow picture. Computing the ratio $\frac{\mathbf{w}^\intercal(u-\mathbf{v}^{i-1})}{\mathbf{c}^\intercal(u-\mathbf{v}^{\mathbf{i-1}})}$ corresponds to calculating the slope of the edge from $\mathbf{v}^{i-1}$ to $\mathbf{u}$ in the projection. The improving edge of maximal slope from a vertex on the upper path will always be the next edge in the upper path of the polygon. Hence, computing the maximal slope edge at each step gives precisely the pre-image of the upper path.

The parametric formulation also can be seen from the shadow picture. Namely, the face maximizing $\lambda\mathbf{c} + \mathbf{w}$ in P is the pre-image of the face of $\pi(P)$ maximizing $(\lambda, 1)^\intercal\mathbf{x}$, since $\pi(\mathbf{x}) = (\mathbf{c}^\intercal\mathbf{x}, \mathbf{w}^\intercal\mathbf{x})$. The set of faces of $\pi(P)$ maximizing $(\lambda, 1)^\intercal\mathbf{x}$ for some choice of $\lambda \in \mathbb{R}$ is exactly the upper path of $\pi(P)$ by definition. Thus, the final definition describes exactly the pre-image of the upper path.

This only describes the vertices of the monotone path polytope. For faces, there is a generalization of a monotone path called a **cellular string**, which is a sequence of faces $F_0, F_1, \ldots, F_k$ satisfying the following conditions:

- $F_{i-1}^{\mathbf{c}} = F_i^{-\mathbf{c}}$ for all $i \in [k]$,
- $F_i^{-\mathbf{c}} \neq F_i^{\mathbf{c}}$ for all $i \in [0, k]$,
- $F_0^{-\mathbf{c}} = P^{-\mathbf{c}}$ and $F_k^{\mathbf{c}} = P^{\mathbf{c}}$.

A cellular string is called **w-coherent** if it is the set of all $\mathbf{w}$-maxima of all fibers of $\varphi(\mathbf{x}) = \mathbf{c}^\intercal\mathbf{x}$. There is a nearly identical theorem for characterizing coherence of cellular strings.

THEOREM 1.3.8. *Let P be a polytope and $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n$. Then the following are equivalent for a $\mathbf{c}$-cellular string $\Gamma \subseteq P$:*

(1) Γ is $\mathbf{w}$-coherent.

(2) $\Gamma = \pi^{-1}(U)$, where $\pi(\mathbf{x}) = (\mathbf{c}^\intercal\mathbf{x}, \mathbf{w}^\intercal\mathbf{x})$ and U is the union of the set of all edges of $\pi(P)$ with outer normal vector with positive y coordinate.

(3) Γ is built inductively start at $\mathbf{F}^0$, the $\mathbf{w}$-maximal face of the $\mathbf{c}$-minimal face of P and

$$\mathbf{F}^i = \bigcup_{\mathbf{v}^{i-1} \in V(\mathbf{F}^{i-1})} \underset{\mathbf{u} \in N(v^{i-1})}{\operatorname{argmax}} \frac{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{v}^{i-1})}{\mathbf{c}^\intercal(\mathbf{u} - \mathbf{v}^{i-1})},$$

where $N(\mathbf{v}^{i-1})$ denotes the set of $\mathbf{c}$-increasing neighbors of $\mathbf{v}^{i-1}$.

(4) $\Gamma = \bigcup_{\lambda \in \mathbb{R}} P^{\lambda\mathbf{c}+\mathbf{w}}$, where $P^{\mathbf{t}}$ denotes the $\mathbf{t}$-maximal face of P for all $\mathbf{t} \in \mathbb{R}^n$.

In particular, Billera and Sturmfels proved the following, which follows as a corollary to our characterization:

THEOREM 1.3.9 ([17]). *The face lattice of the monotone path polytope is precisely the poset of coherent cellular strings.*

The poset of all cellular strings originally arose in work of Baues in topology [14], and the work of Billera, Kapranov, and Sturmfels resolved a conjecture of Baues about the topology of the space of cellular strings [16] using the monotone path polytope and the space of coherent cellular strings. Namely, the inclusion of the poset of coherent cellular strings into the poset of all cellular strings is a homotopy equivalence, so in that sense the coherent monotone paths capture exactly the topology of the space of all monotone paths. See Figure 1.8 for an example of the monotone path polytope of the octahedron together with the corresponding coherent cellular strings and coherent monotone paths.

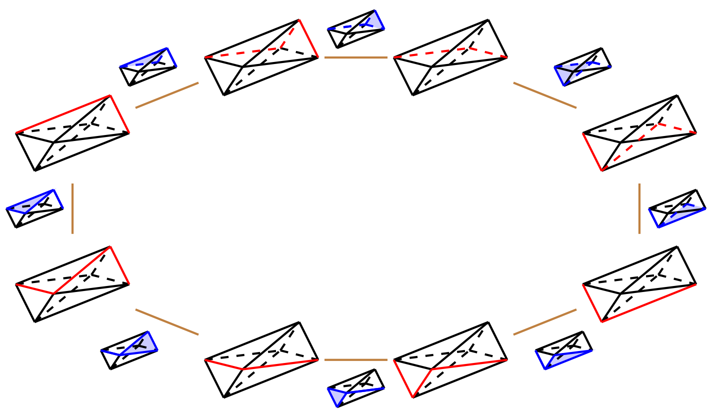

FIGURE 1.8. Pictured is $\Sigma_{(1,2,3)}(\diamond^3)$. The vertices are labeled by the corresponding coherent monotone paths and the edges are labeled by the corresponding coherent cellular strings. The two incoherent paths may be found in Figure 3.1.

Few enough examples of monotone path polytopes were known prior to this work that we can provide a fairly comprehensive survey of known results. In [17], Billera and Sturmfels introduced the monotone path polytope and computed the monotone path polytope of a simplex and hypercube. The monotone path polytope of a simplex appeared even earlier in Gelfand, Kapranov, and Zelevinsky's work on secondary polytopes [75], where they showed a certain monotone path polytope of a standard simplex is exactly the Newton polytope of the discriminant. Billera and Sturmfels also showed in [19] what happens when iterating the fiber polytope construction and computed a monotone path polytope of the monotone path polytope of a simplex. In [10], they computed general fiber polytopes of projections of cyclic polytopes, which included a discussion of the monotone path polytope. In contrast, while fiber polytopes of projections of hyper-simplices have been studied [114], computing their monotone path polytopes remains open in general. There are partial results characterizing coherence of monotone paths for the second hyper-simplex given in [118].

For zonotopes, there have been multiple notable works with a highlight being the thesis of Rob Edman [64] and associated paper [65] in which a complete characterization is given for zonotopes for which all cellular strings are coherent. The monotone path polytopes of zonotopes appeared more recently in relation to orderings of point configurations by vectors and allowable sequences [115], which turned out to have surprising connections to quantum physics [38, 39]. Also, in relation to zonotopes and hyperplane arrangements, Athanasiadis generalized results for monotone path polytopes of the hyper-cube to apply to what he calls piles of cubes originating from work of Ziegler [9]. In what remains, we will discuss in detail the results covered in this thesis.

1.4. Our Contributions: Many New Monotone Path Polytopes

Note that all theorems of this section originated from work towards this thesis. Billera and Sturmfels showed in their original paper on the topic that all cellular strings on the simplex and the hyper-cube are coherent for any generic linear functional on them. They showed the monotone path polytope for a simplex is **combinatorially equivalent** (i.e., has the same face lattice) to a hyper-cube, and the monotone path polytope of a hyper-cube is combinatorially equivalent to a permutahedron, the polytope with vertices given by $\{(\sigma(1), \sigma(2), \ldots, \sigma(n)) : \sigma \text{ is a permutation of } [n]\}$. Simplices and hyper-cubes are two of the three infinite families of **regular polytopes**, higher dimensional generalizations of the Platonic solids [45]. The only remaining infinite family is the cross-polytopes, the polytopes with vertices $\{\pm e_i : i \in [n]\}$. This was the starting point of my work on monotone path polytopes. In joint work with Jesús De Loera, we showed the following:

Theorem 1.4.1. *Let $\mathbf{c} \in \mathbb{R}^n$, and suppose that $0 < c_1 < \cdots < c_n$. Then the face lattice of $\Sigma_{\mathbf{c}}(\diamond^n)$ is the poset of intervals in the face lattice of $\diamond^{n-1}$ excluding the empty face. Furthermore, a cellular string is coherent on $\Sigma_{\mathbf{c}}(\diamond^n)$ if and only if it does not contain a pair of antipodes other than $\{\pm e_n\}$. In particular, not all cellular strings are coherent.*

The next example we studied were products of simplices in order to generalize the observations of Billera and Sturmfels. To do this, we try first to understand in general how coherence is affected by taking the product of a polytope with a simplex. In this case, it turns out the taking the product with a simplex preserves the property of having all cellular strings be coherent:

THEOREM 1.4.2. *Let $P \subset \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^m$. Suppose that all $\mathbf{c}$-cellular strings on P are coherent. Let $\mathbf{a} \in \mathbb{R}^n$ have distinct coordinates. Then all $(\mathbf{c}, \mathbf{a})$-cellular strings on $P \times \Delta_n$ are coherent.*

Hence, as a corollary, we have:

COROLLARY 1.4.1. *For any generic LP on a product of arbitrarily many simplices, all cellular strings are coherent.*

In fact, we prove more. On a polygon, all monotone paths are coherent for any generic linear function.

COROLLARY 1.4.2. *For any generic LP on a product of a polygon and arbitrarily many simplices, all cellular strings are coherent. Thus, there are infinitely many polytopes in every dimension for which all cellular strings are coherent.*

For the special case of a product of two simplices, we arrive at a nice characterization of the face lattice in analogy to the associahedron. Namely, consider a product of simplices $\Delta_m \times \Delta_n$, where Δ_k has vertices $e_1, e_2, \ldots, e_k$. We define the **colorful polygon** P_{m+n} associated to this product a polygon with m red vertices colored cyclically and then n blue vertices colored cyclically. In particular, we imagine the upper path of the polygon consists of m red vertices labeled from left to right $e_1, \ldots, e_m$ and n blue vertices on the lower path labeled $f_1, f_2, \ldots, f_n$ to represent the vertices of the two simplices. See Figure 3.3. We give a complete characterization of the face lattice in terms of certain subdivisions of P_{m+n}.

THEOREM 1.4.3. *Let $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{a} \in \mathbb{R}^n$ and assume the entries of $\mathbf{c}$ and $\mathbf{a}$ are strictly increasing. Then faces of $\Sigma_{(\mathbf{c},\mathbf{a})}(\Delta_m \times \Delta_n)$ are in bijection with subdivision of subpolygons of P_{m+n} such that each subpolygon contains the vertices corresponding to e_1, f_1, e_m, and f_n, and all internal edges are between vertices of different colors.*

A natural criticism of the monotone path polytope is that it only parametrizes monotone paths from a minimum to a maximum, while the simplex method may potentially choose a path starting from any vertex of the polytope. Define a **partial coherent monotone path** to be the restriction of a coherent monotone path to its final k steps. Then the set of partial coherent monotone paths model applying the shadow simplex method from any possible starting point. More generally, a **partial coherent cellular string** may be defined as the restriction of a coherent cellular string to its final k cells. It turns out that one can still find an accurate polyhedral model in this setting. Namely, given a polytope $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$, one can define its **pyramid** by $\text{pyr}(P) = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : A\mathbf{x} + \mathbf{b}y \leq \mathbf{b}, y \in [0,1]\}$. This new polytope has a facet isomorphic to P and otherwise has one new additional vertex e_{n+1} that is adjacent to all other vertices. We prove the following:

THEOREM 1.4.4. *Let $P \subset \mathbb{R}^n$ be a polytope, and let $\mathbf{c} \in \mathbb{R}^n$. Let $c^* = \min(\{\mathbf{c}^\intercal(\mathbf{x}) : \mathbf{x} \in P\})$. Then the set of coherent $(\mathbf{c}, -c^* + 1)$-cellular strings on pyr(P) is in bijection with $C \times \{0,1\}$, where C is the set of partial coherent cellular strings. Coherent $(\mathbf{c}, -c^* + 1)$-monotone paths on pyr(P) are in bijection with partial coherent monotone paths.*

We conclude this section with a discussion of coherence of monotone paths on 0/1-polytopes. In particular, we prove the following theorem:

THEOREM 1.4.5. *Let P be a 0/1-polytope, and suppose that all edge directions are of the form e_S for some $S \subseteq [n]$. Then all $(1, 1, \ldots, 1)$-cellular strings on P are coherent.*

This gives a sufficient condition for coherence of cellular strings for a large class of polytopes. In fact, for monotone paths, we may extend this result even further:

THEOREM 1.4.6. *Let P be a 0/1-polytope, and let $\mathbf{c} = (1, 1, \ldots, 1)$. Then if all edge directions of P are of the form e_S or are orthogonal to $\mathbf{c}$, all $\mathbf{c}$-monotone paths on P are coherent.*

One class of polytopes of recent interest due to their role in are Δ-matroids [66], a Type B generalization of matroids, which are precisely the 0/1-polytopes with edge directions of the form either $e_i - e_j$ or $e_i + e_j$. These satisfy the main condition of the lemma, so we must have

COROLLARY 1.4.3. *Let P be a Δ-matroid, and let $\mathbf{c} = (1, 1, \ldots, 1)$. Then all $\mathbf{c}$-monotone paths on P are coherent.*

Similarly matroid independence polytopes satisfy this condition.

Corollary 1.4.4. *Let P be matroid independence polytope, and let $\mathbf{c} = (1, 1, \ldots, 1)$. Then all $\mathbf{c}$-monotone paths on P are coherent.*

1.5. Pivot Rules for the Simplex Method

In the Section **??**, we defined a pivot rule for the simplex method as any method for choosing an entering and leaving variable at a given basis for the simplex method. However, we remained purposefully vague on what such a method could look like. There is not a universal consensus on what restrictions a pivot rule must satisfy or what exactly constitutes a simplex method with different interpretations given in each of [43, 57, 89] and plenty of others in the broader literature. Even when restricted to one version of the simplex method as we have presented here, there is a lack of clarity. For example, one could define a pivot rule as any method for selecting an entering variable in weakly polynomial time. However, since linear programming may be solved in weakly polynomial time via the ellipsoid method for example [90], one could then use a solution to the linear program as part of the pivot rule. Could such a method be called a pivot rule?

Instead of examining this questions, we will instead be specific about the kinds of pivot rules we consider. In particular, our approach will be geometric and combinatorial in order to avoid complexity considerations for example such as whether one allows a linear program solver as a subroutine. In this section, a **pivot rule** is going to be a $\mathcal{A}$ that takes in a linear program given as a pair $(P, \mathbf{c})$, where P is a polytope and $\mathbf{c} \in \mathbb{R}^n$. Then $f = \mathcal{A}(P, \mathbf{c})$ is a function with domain the vertices $V(P)$ of P, and $f(\mathbf{v})$ is a monotone path from $\mathbf{v}$ to a $\mathbf{c}$-maximum for all $\mathbf{v} \in V(P)$. Note that, in this section, we do not consider degenerate pivots for our pivot rules. We only consider edge steps, which correspond to non-degenerate pivots. In particular, for each pivot rule, one may compute its **footprint**, which is the union of $f(\mathbf{v})$ for all $\mathbf{v} \in V(P)$. We call a pivot rule **memory-less** if the underlying undirected graph of the footprint is a tree. In particular, a pivot rule is memory-less if the directed graph of the footprint is an **arborescence**.

Remark 1.5.1. *A pivot rule is memory-less for an LP (P, c) if its corresponding function f satisfies $f(\mathbf{u}) \subseteq f(\mathbf{v})$ for all vertices $\mathbf{u} \in f(\mathbf{v})$ and $\mathbf{v} \in V(P)$.*

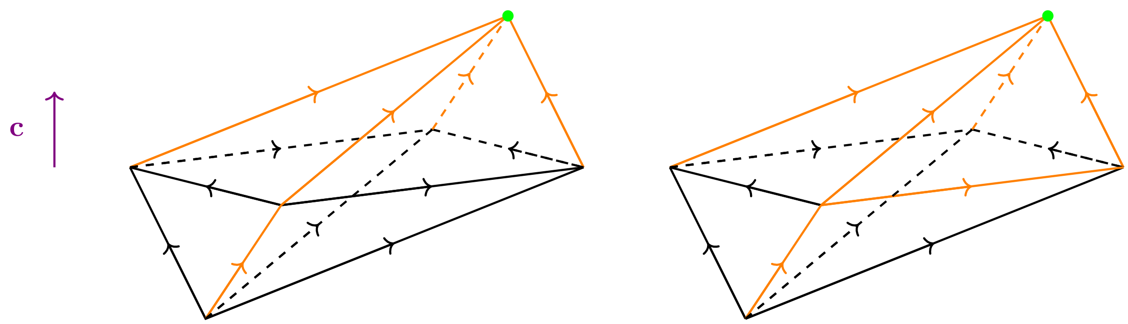

Figure 1.9. Pictured are footprints for two different pivot rules for a linear program on the octahedron. The one on the left is memoryless, while the right is not memoryless.

This remark is why we call these pivot rules memory-less. At a vertex they always choose the same path to an optimum regardless of the path taken to reach that vertex. Pivot rules that rely on memory cannot do that. One observation we can make about memory-less pivot rules is the following:

Proposition 1.5.1. *If there is a pivot rule that guarantees a polynomial run-time for the simplex method, there must be a memory-less pivot rule that always chooses a polynomial length monotone path from every starting vertex to an optimum.*

Proof. Fix any LP. Then take the footprint of a polynomial pivot rule, and compute a breadth first search tree from the set of optima to all remaining vertices in the directed graph induced by the LP. For each vertex in the polytope there will then be a unique directed path to the optimum. Taking the union of all such paths yields exactly the breadth first search tree. One may define a pivot rule by always a shortest path to the optimum in that breadth first search tree. Since the original pivot rule runs in polynomial time, each such path would be of polynomial length. Furthermore, each footprint for the pivot rule would be an arborescence, so the pivot rule is memory-less. $\square$

Thus, to study the worst-case behavior of the simplex method, it suffices to study memory-less pivot rules. Among the memory-less pivot rules, we will focus our attention on **normalized weight pivot rules** (NW-pivot rules), which are always given in the following manner. Let $\mathbf{v}^i$

$$\mathbf{v}^{i+1} = \operatorname*{argmax}_{\mathbf{u} \in N_{\mathbf{c}}(\mathbf{v}^i)} \frac{\mathbf{w}^{\intercal}(\mathbf{u} - \mathbf{v}^i)}{\eta(\mathbf{u} - \mathbf{v}^i)},$$

where $N_{\mathbf{c}}(\mathbf{v}^i)$ denotes the set of $\mathbf{c}$-improving neighbors of $\mathbf{v}^i$, $\mathbf{w} \in \mathbb{R}^n$ is called the **weight**, $\eta : \mathbb{R}^n \to \mathbb{R}_{>0}$ is called the **normalization**. The pair $(\mathbf{w}, \eta)$ then determines the pivot rule. See Table 1.1 for examples of pivot rules taken from [134] together with a determination of whether they are memory-less, normalized weight, or not. Note we only consider pivot rules for the primal simplex algorithm, which excludes, for example, criss-cross methods that allow for infeasible bases [72].

Table 1.1. Pivot Rules from [134]

Pivot Rule	Memory-Less	Normalized Weight
Greatest Improvement	Yes	Yes
Steepest Edge	Yes	Yes
Roos' Minimial Volume Simplex Pivot Rule	Yes	No
Dual Interior Primal Simplex	Yes	No
Bland's Minimal Index Pivot Rule	Yes	No
Edmonds Fukuda Pivot Rule	Yes	No
Dantzig's Pivot Rule	Yes	No
Bland's Recursive Pivot Rule	No	No
Zadeh's Pivot Rule	No	No
Shadow Pivot Rules	No	No
Random Edge Pivot Rule	No	No

Dantzig's pivot rule is memory-less, since the decision of an entering variable is made purely locally in terms of the current basis. However, Dantzig's pivot rule is only defined in our sense for **simple polytopes**, polytopes for which each vertex corresponds to a unique basis. Otherwise, due to degenerate steps, the choice of outgoing edge taken by Dantzig's pivot rule may depend on the initial basis. Even on simple polytopes, Dantzig's pivot rule is not a normalized weight pivot rule, since the normalization changes at each step.

1.6. Our Contributions: Pivot Rule Polytopes

The definitions given for memoryless pivot rules and normalized weight pivot rules are nonstandard and earlier versions are given in joint work with De Loera, Lütjeharms, and Sanyal [26]. All theorems here are proven in this thesis. While some memory-less pivot rules are not explicitly given by a normalized weight pivot rule, it turns out that generically this is always the case:

Theorem 1.6.1. *Given a linear program $(P, \mathbf{c})$, there is a small perturbation P' of P for which every memory-less pivot rule is given by some normalized weight pivot rule.*

A key question we will be interested in understanding is how a normalized weight pivot rule varies for a fixed linear program as one changes the weight but fixes the normalization. This allows one to take pivot rules that are not typically parametric such as the steepest edge pivot rule and make them parametric. In particular, it allows one to perturb a pivot rule. Here we show the following:

THEOREM 1.6.2. *Fix a linear program* $(P, \mathbf{c})$ *and a normalization* η. *Then the set of all weights* $\mathbf{w}$ *that induces the same arborescence for the normalized weight pivot rule with normalization* η *and weight* $\mathbf{w}$ *forms a polyhedral cone. The set of all such cones forms a polyhedral fan and is the normal fan of a polytope.*

We call the polytope with this normal fan the **pivot rule polytope**. In fact, the polytope is constructed explicitly using Minkowski sums as follows:

DEFINITION 1.6.3. *Given a polytope* P *and* $\mathbf{c} \in \mathbb{R}^n$, *we define the pivot rule polytope for the pivot rule with normalization* η *by*

$$\Pi_{\eta,\mathbf{c}}(P) = \sum_{v \in V(P)} conv\left(\left\{\frac{\mathbf{u}-\mathbf{v}}{\eta(\mathbf{u}-\mathbf{v})} : \mathbf{u} \in N_{\mathbf{c}}(\mathbf{v})\right\}\right),$$

where $N_{\mathbf{c}}(\mathbf{v})$ *is the set of* $\mathbf{c}$*-improving neighbors of* $\mathbf{v}$.

Our construction is an alteration of Billera and Sturmfels fiber polytope construction [18] for the monotone path polytope. In fact, we define a **max-slope pivot rule** to be any pivot rule with normalization given by the linear objective function $\eta(\mathbf{x}) = \mathbf{c}^\intercal \mathbf{x}$. In particular, $\mathbf{w}$ may be any weight. Then the max-slope pivot rule polytope is given as follows:

$$\Pi_{\mathbf{c}}(P) = \sum_{v \in V(P)} \mathrm{conv}\left(\left\{\frac{\mathbf{u}-\mathbf{v}}{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})} : \mathbf{u} \in N_{\mathbf{c}}(\mathbf{v})\right\}\right).$$

See Figure 1.10 for an example of the max-slope pivot rule polytope and an illustration of this Minkowski sum decomposition.

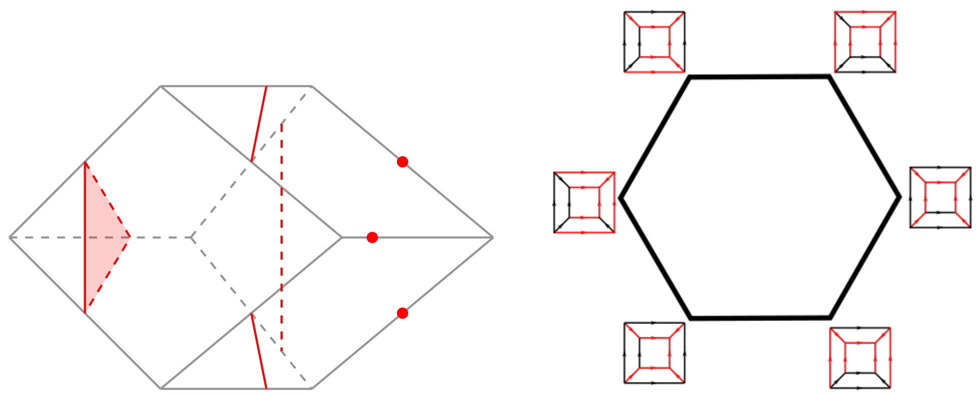

FIGURE 1.10. Pictured is the max-slope pivot rule polytope of the hyper-cube together with a depiction of the Minkowski sum decomposition. Compare the Minkowski sum decomposition to that of the monotone path polytope in Figure 1.5.

The max-slope pivot rule polytope is related directly to the monotone path polytope. In fact, the max-slope pivot rule is defined to be a memory-less version of shadow pivot rules. Its name come from the interpretation of the ratio $\frac{\mathbf{w}^{\intercal}(\mathbf{u}-\mathbf{v})}{\mathbf{c}^{\intercal}(\mathbf{u}-\mathbf{v})}$ as being the the slope of the edge from $\mathbf{v}$ to $\mathbf{u}$ in the projection $\pi(\mathbf{x}) = (\mathbf{c}^{\intercal}\mathbf{x}, \mathbf{w}^{\intercal}\mathbf{x})$. The way to relate the two is through their normal fans. A complete fan F_1 **refines** another complete fan F_2 if each cone in F_1 is contained in a cone of F_2. When a normal fan of a polytope P_1 is refined by a normal fan of a polytope P_2, we say that P_1 is a **weak Minkowski summand** of P_2 (See Chapter 15 of [79] and in particular Theorem 15.1.2). If two polytopes have the same normal fan, we say they are normally equivalent.

THEOREM 1.6.4. *Let $(P, \mathbf{c})$ be a fixed linear program, and suppose that the $\mathbf{c}$-minimum on P is unique. Then the monotone path polytope $\Sigma_{\mathbf{c}}(P)$ is a weak Minkowksi summand of the pivot rule polytope $\Pi_{\mathbf{c}}(P)$. If P is a zonotope, then they are normally equivalent.*

In particular, for a hyper-cube the max-slope pivot rule polytope and monotone path polytope are normally equivalent, so we have the following corollary:

COROLLARY 1.6.1. *The max-slope pivot rule polytope of a hyper-cube for a generic choice of $\mathbf{c}$ is combinatorially equivalent to the permutahedron.*

Our final main results of this section are on an interesting class of examples. While the following theorem is stated in an alternative form in [26], a complete proof first appears in this thesis and arises from joint work with Lütjeharms and Sanyal.

Theorem 1.6.5. *The max-slope pivot rule polytope of a simplex is an associahedron.*

An **associahedron** is any polytope whose face lattice is given by the poset of all subdivisions of an n-gon. It was originally described as a lattice, and Tamari asked in his 1951 PhD thesis [132] whether this lattice can arise as the face lattice of a polytope. There are many positive answers to this question, and the history and diversity of realizations of the associahedron is covered in the following survey [40]. There are many different realizations of this face lattice using various geometric tools such as the theory of generalized permutahedra and cluster algebras [41,75,84,117]. Ours adds a new realization to the list with a distinction that our realization is coming from optimization and the study of the simplex method. Associahedra also appear in applications in both theoretical computer science [124] and in fundamental physics through scattering theory [8]. In this introduction, we stated the key contributions in this thesis. In the next chapters, we present the details and proofs surrounding each of the results.

CHAPTER 2

The Simplex Method

2.1. On the Simplex Method for 0/1-Polytopes

The section expands and updates joint work with Jesús De Loera, Sean Kafer, and Laura Sanità in [25]. Our goal here is to devise a simplex method for 0/1-polytopes such that the total number of non-degenerate pivots is at most n, the number of variables used to describe the 0/1-polytope. In essence, we argue that the complexity of the simplex method for 0/1-polytopes reduces entirely to degeneracy. To do this, we consider linear programs of the following form:

$$\begin{aligned} &\max(\mathbf{c}^\intercal \mathbf{x}) \\ &\text{s.t. } A\mathbf{x} \leq \mathbf{b} \\ &0 \leq \mathbf{x}_i \leq 1 \text{ for all } i \in [n], \end{aligned}$$

where the set of vertices of the resulting polytope have coordinates that are in $\{0,1\}^n$. Note that the inequalities $0 \leq \mathbf{x}_i \leq 1$ may be redundant, but we do not remove them for designing our simplex method. Namely, in standard form, our linear program looks like the following:

$$\begin{aligned} &\max(\mathbf{c}^\intercal \mathbf{x}) \\ &\text{s.t. } A\mathbf{x} + \mathbf{y} = \mathbf{b} \\ &\mathbf{x}_i + \mathbf{z}_i = 1 \text{ for all } i \in [n] \\ &\mathbf{x}, \mathbf{y}, \mathbf{z} \geq 0, \end{aligned}$$

where any basic feasible solution $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathbf{R}^{n+m+n}$ must satisfy $\mathbf{x} \in \{0,1\}^n$. For phase 1 of the simplex method, we first produce an arbitrary basis $B \subseteq [n+m+n]$ using a standard black box phase 1 algorithm for the simplex method. From this basis, one may read off an initial vertex $\mathbf{v} \in \{0,1\}^n$. Let $[n] = S_0 \cup S_1$, where S_0 and S_1 are the set of coordinates of $\mathbf{v}$ equal to 0 and 1 respectively. Then $[n+m+n] \setminus S_0 \cup (S_1+m)$ is a basis for $\mathbf{v}$, where $S_1+m = \{i+m : i \in S_1\}$. Thus, in phase 1 of the simplex method, we will generate a basis of that form. That is, the set of non-basic variable will be $S_0 \cup (S_1+m)$.

It remains to describe phase 2 of the simplex method. For this, we provide two different options. The first option is the simpler of the two:

Definition 2.1.1. *Let B be the initial basis generated by phase 1. Define $\mathbf{w} = \sum_{i \notin B} e_i$. Then the* ***slim shadow pivot rule*** *says to choose a new basis B' from a basis B by*

$$B' = (B \setminus \{j\}) \cup \operatorname*{argmax}_{k \in [n+m+n] \setminus B} \frac{\mathbf{c}^\intercal (e_k - A_B^{-1} A_k)}{\mathbf{w}^\intercal (e_k - A_B^{-1} A_k)}.$$

The second option is like the first but with a lexicographic ordering.

Definition 2.1.2. *Let B be the initial basis generated by phase 1. Choose any bijection $\sigma : [n+m+n] \setminus B \to [n]$. Define $\mathbf{w} = \sum_{i \notin B} r^{\sigma(i)} e_i$ for any $r > \max(2\|\mathbf{c}\|_1^2 n, \|\mathbf{c}\|_1 + 2)$. Then the* ***ordered shadow pivot rule*** *says to choose a new basis B' from a basis B by*

$$B' = (B \setminus \{j\}) \cup \operatorname*{argmax}_{k \in [n+m+n] \setminus B} \frac{\mathbf{c}^\intercal (e_k - A_B^{-1} A_k)}{\mathbf{w}^\intercal (e_k - A_B^{-1} A_k)}.$$

Our argument comes down to the following. We first show that if a non-degenerate step is taken by this method, it is given by maximizing

$$\frac{\mathbf{c}^\intercal (\mathbf{u} - \mathbf{v})}{\mathbf{w}^\intercal (\mathbf{u} - \mathbf{v})}$$

over all improving neighbors $\mathbf{u}$ of the current vertex $\mathbf{v}$. Since our rule for leaving variable prevents cycling, we will always take a non-degenerate step after finitely many degenerate steps. Then we argue that the path generated by this set of non-degenerate steps must be short.

Recall that a basis B corresponds to a set of n linearly independent inequalities that are tight at a given $\mathbf{v}$. The **basic cone** C_B is the cone given by the intersection of the inequalities from that basis. We define the **normal cone** to a basis N_B to be the polar dual cone to C_B. Then $N_B = \{\mathbf{y} : \mathbf{y}^\intercal \mathbf{x} \leq \mathbf{0} \text{ for all } \mathbf{y} \in C_B\}$. The normal cone to a basis is distinct from the normal cone $N_\mathbf{v}$ to a vertex $\mathbf{v}$, which is the set of vectors $\mathbf{c}$ for which $\mathbf{v}$ is $\mathbf{c}$-maximal. However, by construction, if $\mathbf{v}$ is the vertex corresponding to a basis B, then $N_B \subseteq N_\mathbf{v}$.

Lemma 2.1.2.1. *Let* $C = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$ *be a polyhedral cone. Let* $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n \setminus \mathbf{0}$*. Let* $\lambda \in \mathbb{R}$*, and suppose that* $\mathbf{w} \in C$*. Then the maximal* λ *such that* $-\mathbf{w} + \lambda\mathbf{c} \in C$ *is*

$$\lambda^* = \max\left\{\frac{\mathbf{c}^\intercal A_i}{\mathbf{w}^\intercal A_i} : A_i > 0\right\}$$

Proof. Consider the line $-\mathbf{w} + \lambda\mathbf{c}$. For this line to leave $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$, it must pass through one of its bounding hyperplanes given by $\mathbf{a}_i^\intercal\mathbf{x} = 0$. If $\mathbf{a}_i^\intercal\mathbf{c} < 0$ then $\mathbf{a}_i^\intercal(\lambda\mathbf{c} + -\mathbf{w}) < \mathbf{a}_i^\intercal(-\mathbf{w}) \leq 0$ for all $\lambda > 0$. For all such hyper-planes, adding a positive multiple of $\mathbf{c}$ only moves further away from the hyper-plane. In contrast, for all $\mathbf{a}_i^\intercal\mathbf{c} > 0$ adding a positive multiple of $\mathbf{c}$ moves $-\mathbf{w}$ closer to violating the inequalities defining C. In particular, observe that $\mathbf{a}_i^\intercal(-\mathbf{w} + \lambda\mathbf{c}) = 0$ precisely when $\lambda = \frac{\mathbf{w}^\intercal\mathbf{a}_i}{\mathbf{c}^\intercal\mathbf{a}_i}$. The first violation of a hyper-plane by moving along the line is given by minimizing λ, which is equivalent to maximizing $\frac{\mathbf{c}^\intercal\mathbf{a}_i}{\mathbf{w}^\intercal\mathbf{a}_i}$.

$\square$

This observation is standard and gives rise to the dual perspective on the shadow simplex method as one would find in [31]. Here we use it to understand how to deal with degeneracy. See Figure 2.1 for an illustration of the argument.

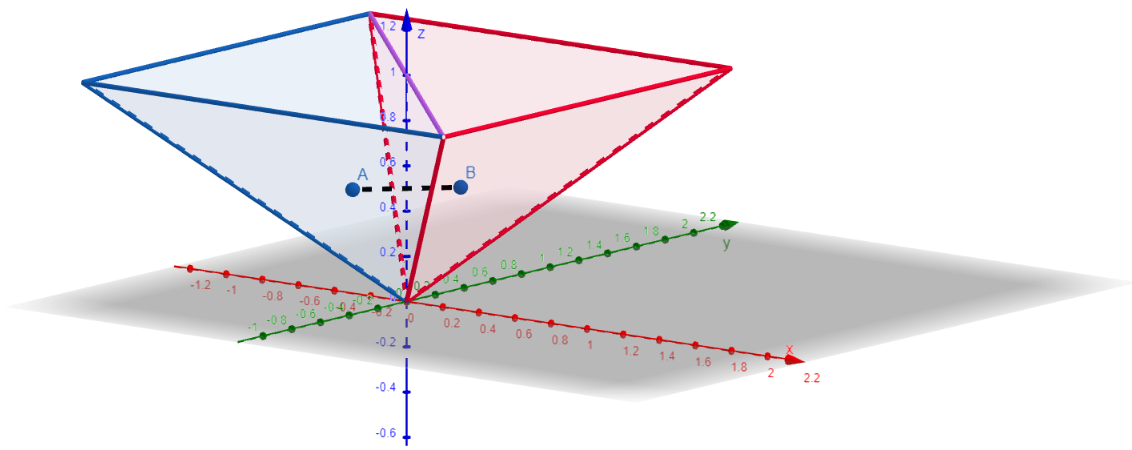

FIGURE 2.1. Consider the square pyramid $\operatorname{conv}((-1,-1,0),(-1,1,0),(1,-1,0),(1,1,0),(0,0,1))$. Then the vertex $(0,0,1)$ is degenerate. Two degenerate bases for the vertex $(0,0,1)$ are depicted in the image through their basic normal cones C_1 and C_2 generated by the rays $\{(-1,-1,1),(-1,1,1),(1,-1,1)\}$ and $\{(1,1,1),(-1,1,1),(1,-1,1)\}$ respectively. The cone C_1 contains the point $A=(-1/6,-1/6,1/2)$, and the cone C_2 contains the point $(1/6,1/6,1/2)$. Starting from the cone C_1, the shadow pivot rule with auxiliary vector $\mathbf{w}=(1/6,1/6,-1/2)$ for maximizing $(1,1,0)$ would walk from cone C_1 to cone C_2 as described in Lemma 2.1.2.2. The image was made using Geogebra (geogebra.org).

LEMMA 2.1.2.2. *Let N_B and $N_{B'}$ be normal cones for adjacent bases B and B' for a linear program $\{\mathbf{x}\geq\mathbf{0} : A\mathbf{x}=\mathbf{b}\}$. Then C_B and $C_{B'}$ share a facet. Furthermore, if B' is the basis chosen by the shadow pivot rule for $\mathbf{w}$ choosing the shadow and maximizing $\mathbf{c}$ with $-\mathbf{w}+\lambda\mathbf{c}\in C_B$, then $-\mathbf{w}+\lambda'\mathbf{c}\in C_{B'}$ for some $\lambda'>\lambda$.*

PROOF. Note that $N_B=\ker(A)^\perp+\{\sum_{i\in B^C}\lambda_i e_i : \lambda_i\geq 0 \text{ for all } i\}$. Then $N_B\cap N_{B'}=\{\sum_{i\in B^C\cap(B')^C}\lambda e_i : \lambda_i\geq 0 \text{ for all } i\}+\ker(A)^\perp$. Hence, the two normal cones for adjacent bases share a maximal dimensional face. The normal to that face is precisely the edge direction for exchanging from one basis to the other. This face is the last cone that is intersected by the line $-\mathbf{w}+\lambda\mathbf{c}$ by Lemma 2.1.2.1. Hence, after the exchange the line still intersects the cone as desired. □

This argument tells us that there is a measure of progress made at each pivot step. Namely each consecutive basic cone contains a point further along the line $\mathbf{w} + \lambda\mathbf{c}$. Hence, as a corollary, we have the following:

Corollary 2.1.1. *The simplex method with both the slim shadow pivot rule and the ordered shadow pivot rule will not cycle when solving an LP on a* $0/1$*-polytope.*

We already have this observation for free from our rule for entering variable being lexicographic. This observation tells us that any rule for entering variable suffices. To finish dealing with degeneracy, we require that if a non-degenerate pivot occurs, it is a pivot of the desired type.

Lemma 2.1.2.3. *Let B be a basis for a linear program* $\max(\mathbf{c}^\intercal\mathbf{x})$ *such that* $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq 0$. *Suppose that N_B intersects the line $\mathbf{w} + \lambda\mathbf{c}$ for $\mathbf{w} \in \mathbb{R}^n$. Let B' be the basis chosen for B by the shadow for $\mathbf{w}$, and suppose that the step followed is non-degenerate. Then*

$$\mathbf{v}_{B'} = \operatorname*{argmax}_{\mathbf{u} \in N_{\mathbf{c}}(\mathbf{v}_B)} \frac{\mathbf{c}^\intercal(\mathbf{u} - \mathbf{v}_B)}{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{v}_B)},$$

where $\mathbf{v}_B$ is the vertex associated to a basis B and $N_{\mathbf{c}}(\mathbf{v}_B)$ is the set of $\mathbf{c}$-improving neighbors of $\mathbf{v}_B$.

Proof. In Lemma 2.1.2.2, we showed that the set of bases in the shadow simplex path always intersect the line $-\mathbf{w} + \lambda\mathbf{c}$. In fact, we argued that the normal cones for consecutive bases must intersect consecutive parts of the line meaning that the union of the two pieces of the line in connected normal cones is connected.

The same argument holds for vertices. The vertex for

$$\operatorname*{argmax}_{\mathbf{u} \in N_{\mathbf{c}}(\mathbf{v}_B)} \frac{\mathbf{c}^\intercal(\mathbf{u} - \mathbf{v}_B)}{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{v}_B)}$$

is exactly the vertex $\mathbf{v}$ with normal cone that intersects the line $-\mathbf{w} + \lambda\mathbf{c}$ after it intersects $\mathbf{v}_B$. Hence, the normal cones for B' must intersect the normal cone for $\mathbf{v}$ meaning that $\mathbf{v} = \mathbf{v}_{B'}$ as desired. ☐

It remains to argue that the total number of non-degenerate steps taken must be at most n for the slim shadow pivot rule and at most d, the dimension of P, for the ordered shadow pivot rule. For bounding the total number of non-degenerate steps, we rely on the following lemma.

LEMMA 2.1.2.4. *Consider the LP* $\max(\mathbf{c}^\intercal \mathbf{x})$ *such that* $\mathbf{x} \in P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\} \subseteq \mathbb{R}^n$ *for* P *a polytope. Let* $\mathbf{w} \in \mathbb{R}^n$*, and suppose that* $\mathbf{v}$ *is the* $\mathbf{c}$*-maximum of the* $\mathbf{w}$*-minimal face of* P*. Then any path generated by the* $\mathbf{w}$*-shadow rule is both* $\mathbf{c}$*-increasing and* $\mathbf{w}$*-increasing.*

PROOF. Consider the projection $\pi(P)$ under the map $\pi(\mathbf{x}) = (\mathbf{w}^\intercal \mathbf{x}, \mathbf{c}^\intercal \mathbf{x})$. Then, since $\mathbf{v}$ is a $\mathbf{c}$-maximum of the $\mathbf{w}$-minimal face of P, $\pi(\mathbf{v})$ is the e_2-maximum of the e_1-minimal face of the polygon $\pi(P)$. The $\mathbf{c}$-maximizing face F of P is mapped onto the e_2-maximizing face $\pi(F)$ of $\pi(P)$. The path followed by the shadow rule is given by starting at the vertex $\mathbf{v}$ and choosing the neighbor maximizing $\frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{\mathbf{w}^\intercal(\mathbf{u}-\mathbf{v})}$. In the shadow this corresponds to choosing an improving edge of maximal slope. In particular, the image of the path followed under $\pi(P)$ is contained in the path on the top part of the polygon $\pi(P)$. By convexity, each edge of that path in $\pi(P)$ is of positive slope until reaching an e_2-maximum meaning that it is both e_1-improving and e_2-improving until reaching e_2-maximizer. Hence, the pre-image of the path in P consists of edges that are both $\mathbf{c}$-improving and $\mathbf{w}$-improving. □

As a corollary, we immediately have our bound for the slim shadow pivot rule.

COROLLARY 2.1.2. *The slim shadow pivot rule solves an LP on a* $0/1$*-polytope by following a path of length at most* n*.*

PROOF. By Lemma 2.1.2.4, the path must be strictly $\mathbf{w}$-improving for $\mathbf{w} = \sum_{i \in N} e_i$, where N is the complement of the initial basis. By construction and assumption that P is a $0/1$-polytope, for all $i \in N$ and each vertex $\mathbf{v}$ of P, $\mathbf{v}_i \in \{0, 1\}$. Hence, $0 \leq \mathbf{w}^\intercal \mathbf{v} \leq n$ for each vertex $\mathbf{v}$ in P. Hence, since the path is $\mathbf{w}$-improving the total length of the path is at most n. □

By applying precisely the same argument, for $0/1$-polytopes with bounded size of support for vertices and fixed size of support, we arrive at the proof of Theorem 1.2.2. For the ordered shadow pivot rule, the same tool is used, but the argument is a bit more involved. We will prove a more general statement.

LEMMA 2.1.2.5. *Let* $P \subseteq \mathbb{R}^n$ *be a* d*-dimensional polytope. Suppose that* $V(P) \subseteq \mathbb{Z}^n$ *and let*

$$k = \max_{\mathbf{u},\mathbf{v} \in V(P)} \|\mathbf{u} - \mathbf{v}\|_\infty.$$

For any $\mathbf{c} \in \mathbb{Z}^n$ *and for any*

$$r > \max(2\|\mathbf{c}\|_1^2 k^3 n, \|\mathbf{c}\|_1 k^2 + 1 + k)$$

the $(r, r^2, \ldots, r^n)$ *shadow path for maximizing* $\mathbf{c}$ *from the* $(r, r^2, \ldots, r^n)$*-minimal vertex of* P *is of length at most*

$$d\left(\max_{i\in[n]}\{\textit{Length of any } e_i\textit{-monotone path}\}\right).$$

Proof. Consider the chain of faces $P = F_{n+1} \supseteq F_n \supseteq \cdots \supseteq F_1$, where F_i is defined inductively the e_i-minimal face of F_{i+1}. Then F_1 is the vertex $\mathbf{v}$ minimizing $\mathbf{r} = (r, r^2, \ldots, r^n)$ for all $r > 0$ with r sufficiently large. We consider the shadow path for the rule given by maximizing $\frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}-\mathbf{v})}$. To prove the result, we show that starting at a $\mathbf{c}$-maximum of F_ℓ, this pivot rule follows an e_ℓ-monotone path to the $\mathbf{c}$-maximum of $F_{\ell+1}$. This suffices to prove the theorem, since the initial vertex is the $\mathbf{c}$-maximum of F_1, and at most d of the F_i are distinct leading the total length of the path to be at most the product of d and the maximal length of an e_ℓ monotone path.

If $F_\ell = F_{\ell+1}$, then we are done, so suppose that $F_\ell \neq F_{\ell+1}$. Then let $\mathbf{v}$ be a $\mathbf{c}$-maximum of F_ℓ. Then, since F_ℓ is the e_ℓ-minimal face of F_ℓ, a shadow path for e_ℓ from $\mathbf{v}$ to the $\mathbf{c}$-maximum of $F_{\ell+1}$ is e_ℓ-monotone. Thus, it suffices to show that an $\mathbf{r}$ shadow path restricted to $F_{\ell+1}$ is an e_ℓ-shadow path. To do this, let $\mathbf{u}$ be on that shadow path. Our goal is to show that

$$\operatorname*{argmax}_{\mathbf{u}'\in N_{\mathbf{c}}(\mathbf{u})} \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{u})}{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{u})} = \operatorname*{argmax}_{\mathbf{u}'\in N_{\mathbf{c}}(\mathbf{u})\cap F_{\ell+1}} \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{u})}{e_\ell^\intercal(\mathbf{u}'-\mathbf{u})}.$$

for all r sufficiently large.

Let $\mathbf{y} \in N_{\mathbf{c}}(\mathbf{u}) - F_{\ell+1}$. Then, in particular, $\mathbf{y} \notin F_{\ell+1}$, so $\mathbf{y}_j > \mathbf{u}_j$ for some $j \geq \ell + 1$.

Let $\mathbf{z} \in N_{\mathbf{c}}(\mathbf{u}) \cap F_{\ell+1}$. Then $\mathbf{z}_j - \mathbf{u}_j = 0$ for all $j \geq k+1$. Furthermore, since $\mathbf{u}$ is on the (e_k, $\mathbf{c}$)-shadow of $F_{\ell+1}$, each $\mathbf{c}$-improving neighbor of $\mathbf{u}$ in $F_{\ell+1}$ must also be e_ℓ-improving. Thus, $\mathbf{z}$ must also satisfy $\mathbf{z}_\ell - \mathbf{u}_\ell = e_\ell^\intercal(\mathbf{z} - \mathbf{u}) > 0$.

It remains to find effective bounds on how large r needs to be. Fix a vertex $\mathbf{v}$ on the e_ℓ-shadow path on $F_{\ell+1}$. Let $\mathbf{u} \in N_{\mathbf{c}}(\mathbf{v}) - F_{\ell+1}$. Then we have for some $t > \ell$, since we assume $\mathbf{c} \in \mathbb{Z}^n$,

$$\begin{aligned}\frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}-\mathbf{v})} &\leq \frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{r^t - \sum_{i=1}^{t-1} \|\mathbf{u}-\mathbf{v}\|_\infty r^i} \\ &\leq \frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{r^t - 1 - \|\mathbf{u}-\mathbf{v}\|_\infty \frac{r^t-1}{r-1}} \\ &= \left(\frac{1}{r^t-1}\right) \frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{1 - \frac{\|\mathbf{u}-\mathbf{v}\|_\infty}{r-1}} \\ &\leq \left(\frac{1}{r^t-1}\right) \frac{\|\mathbf{c}\|_1 \|\mathbf{u}-\mathbf{v}\|_\infty}{1 - \frac{\|\mathbf{u}-\mathbf{v}\|_\infty}{r-1}} \\ &\leq \left(\frac{1}{r^t-1}\right) \frac{\|\mathbf{c}\|_1 k}{1 - \frac{k}{r-1}}.\end{aligned}$$

Similarly, for $\mathbf{u}' \in N_{\mathbf{c}}(\mathbf{v}) \cap F_{\ell+1}$, since $\mathbf{c}, \mathbf{u}', \mathbf{v} \in \mathbb{Z}^n$,

$$\begin{aligned}\frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})} &\geq \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\sum_{i=1}^{\ell} r^i \|\mathbf{u}'-\mathbf{v}\|_\infty} \\ &= \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\|\mathbf{u}'-\mathbf{v}\|_\infty \frac{r^{\ell+1}-1}{r-1}} \\ &= \left(\frac{r-1}{r^{\ell+1}-1}\right) \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\|\mathbf{u}'-\mathbf{v}\|_\infty} \\ &\geq \left(\frac{r-1}{r^{\ell+1}-1}\right) \frac{1}{\|\mathbf{u}'-\mathbf{v}\|_\infty} \\ &\geq \left(\frac{r-1}{r^{\ell+1}-1}\right) \frac{1}{k}.\end{aligned}$$

It follows that $\frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}-\mathbf{v})} \leq \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})}$ if

$$\left(\frac{1}{r^t-1}\right) \frac{\|\mathbf{c}\|_1 k}{1 - \frac{k}{r-1}} \leq \left(\frac{r-1}{r^{k+1}-1}\right) \frac{1}{k},$$

which is true if and only if

$$\frac{r^{\ell+1}-1}{(r^t-1)} \left(\frac{1}{r-1-k}\right) \leq \frac{1}{\|\mathbf{c}\|_1 k^2}.$$

Suppose that $r > \|\mathbf{c}\|_1 k^2 + 1 + k$. Then, since $\ell + 1 \leq t$,

$$\frac{r^{\ell+1}-1}{(r^t-1)} \left(\frac{1}{r-1-k}\right) \leq \left(\frac{1}{r-1-k}\right) < \frac{1}{\|c\|_1 k^2}.$$

Hence, for all $r > \|\mathbf{c}\|_1 k^2 + 1 + k$,

$$\underset{\mathbf{u}\in N_{\mathbf{c}}(\mathbf{v})-F_{\ell+1}}{\operatorname{argmax}} \frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}-\mathbf{v})} < \underset{\mathbf{u}'\in N_{\mathbf{c}}(\mathbf{v})\cap F_{\ell+1}}{\operatorname{argmax}} \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})}.$$

Then to prove the result, it remains to show for r sufficiently large that

$$\underset{\mathbf{u}'\in N_{\mathbf{c}}(\mathbf{v})\cap F_{\ell+1}}{\operatorname{argmax}} \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})} = \underset{\mathbf{u}'\in N_{\mathbf{c}}(\mathbf{v})\cap F_{\ell+1}}{\operatorname{argmax}} \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{e_\ell^\intercal(\mathbf{u}'-\mathbf{v})}.$$

Furthermore, observe that for any $\mathbf{u}' \in N_{\mathbf{c}}(\mathbf{v}) \cap F_{\ell+1}$,

$$\left(\frac{1}{r^\ell}\right) \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{e}_\ell^\intercal(\mathbf{u}'-\mathbf{v}) + \frac{n\|\mathbf{u}-\mathbf{v}\|_\infty}{r}} \le \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})} \le \left(\frac{1}{r^\ell}\right) \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{e}_\ell^\intercal(\mathbf{u}'-\mathbf{v}) - \frac{n\|\mathbf{u}'-\mathbf{v}\|_\infty}{r}}.$$

In particular, by manipulating these inequalities:

$$\begin{aligned}
\frac{e_k^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} - \frac{n\|\mathbf{u}'-\mathbf{v}\|_\infty}{r} &\le \frac{e_k^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} - \frac{n\|\mathbf{u}'-\mathbf{v}\|_\infty}{r\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} \\
&\le \frac{1}{r^\ell}\frac{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} \\
&\le \frac{e_\ell^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} + \frac{(k-1)\|\mathbf{u}'-\mathbf{v}\|_\infty}{r\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} \\
&\le \frac{e_k^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} + \frac{n|\mathbf{u}'-\mathbf{v}\|_\infty}{r}.
\end{aligned}$$

Hence, we have that

$$\left|\frac{e_k^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} - \frac{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})}{r^k\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}\right| \le \frac{n\|\mathbf{u}'-\mathbf{v}\|_\infty}{r} \le \frac{nk}{r}$$

Let $\mathbf{a}, \mathbf{b} \in N_{\mathbf{c}}(\mathbf{v})$. Suppose that $\dfrac{e_k^\intercal(\mathbf{a}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{a}-\mathbf{v})} > \dfrac{e_k^\intercal(\mathbf{b}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{b}-\mathbf{v})}$. Our goal is to prove that:

$$\frac{\mathbf{r}^\intercal(\mathbf{a}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{a}-\mathbf{v})} > \frac{\mathbf{r}^\intercal(\mathbf{b}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{b}-\mathbf{v})}.$$

Note that

$$\frac{e_k^\intercal(\mathbf{a}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{a}-\mathbf{v}])} - \frac{e_k^\intercal(\mathbf{b}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{b}-\mathbf{v}])} \ge \frac{1}{\mathbf{c}^\intercal(\mathbf{a}-\mathbf{v})\mathbf{c}^\intercal(\mathbf{b}-\mathbf{v})} \ge \frac{1}{\|\mathbf{c}\|_1^2\|\mathbf{a}-\mathbf{v}\|_\infty\|\mathbf{b}-\mathbf{v}\|_\infty} \ge \frac{1}{\|\mathbf{c}\|_1^2 k^2}.$$

Suppose that $r > 4\|\mathbf{c}\|_1^2 k^3 n$. Then

$$\left|\frac{e_\ell^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})} - \frac{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})}{r^\ell\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}\right| < \frac{nk}{r} < \frac{1}{2\|\mathbf{c}\|_1^2 k^2}.$$

It follows that

$$\frac{r^\intercal(\mathbf{a}-\mathbf{v})}{r^\ell\mathbf{c}^\intercal(\mathbf{a}-\mathbf{v})} - \frac{r^\intercal(\mathbf{b}-\mathbf{v})}{r^\ell\mathbf{c}^\intercal(\mathbf{b}-\mathbf{v})} > \frac{e_\ell^\intercal(\mathbf{a}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{a}-\mathbf{v})} - \frac{e_\ell^\intercal(\mathbf{b}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{b}-\mathbf{v})} - \frac{1}{\|\mathbf{c}\|_1^2 k^2} > 0,$$

as desired. Hence, for $r > \max(2\|c\|_1^2 k^3 n, \|c\|_1 k^2 + 1 + k)$, we have that

$$\underset{\mathbf{u}'\in N_{\mathbf{c}}(\mathbf{v})}{\operatorname{argmax}} \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{\mathbf{r}^\intercal(\mathbf{u}'-\mathbf{v})} = \underset{\mathbf{u}'\in N_{\mathbf{c}}(\mathbf{v})\cap F_{\ell+1}}{\operatorname{argmax}} \frac{\mathbf{c}^\intercal(\mathbf{u}'-\mathbf{v})}{e_\ell^\intercal(\mathbf{u}'-\mathbf{v})}.$$

□

The performance bound for the ordered shadow pivot rule follows immediately from Lemma 2.1.2.5, which completes the proof of Theorem 1.2.1.

2.2. Small Shadows of Lattice Polytopes

In contrast to the previous section, where we explicitly worked out the details of how exactly to implement our ideas with the simplex method, here we will focus on the combinatorial and geometric aspect of our results. Namely, we will show that short paths exist and may be found using shadow rules. We will not discuss the implementation. However, one may implement each of our ideas using similar techniques to the previous chapter.

Our first step of interest is to try to go beyond 0/1-polytopes instead to $\{0,1,2\}$ lattice polytopes. In this case, the arguments do not follow in the same way. The key difference is that normal cones for any vertex of a 0/1-polytope will always contain an orthant. This is no longer the case for half integral polytopes.

Proof of Theorem 1.2.3. Let P be a $(0,2)$-lattice polytope. Up to an affine transformation, we may assume the polytope has vertices in $\{-1,0,1\}^n$ and that it is full dimensional. We proceed by induction. Let $M(k)$ be the worst-case monotone diameter of a $\{-1,0,1\}$-polytope in which the maximal size of a support of a vector is k. Then $M(1) = 1$. Then, in particular, the worst-case monotone diameter of a d-dimensional $\{-1,0,1\}$-polytope is at most $M(d)$, since any $\{-1,0,1\}$-polytope has a full dimensional representation that is also $\{-1,0,1\}$.

Let $\mathbf{v}$ be an initial vertex. Suppose that $\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{v}$ is $\mathbf{v}$-maximal, since $\mathbf{v}^\intercal\mathbf{v} \geq \mathbf{v}^\intercal\mathbf{u}$ for all $\mathbf{u} \in \{-1,0,1\}^n$ by the Cauchy-Schwarz inequality. The $\mathbf{v}$-maximal face consists of all vectors $\mathbf{u} \in P$ such that $\mathbf{u}|_i = \mathbf{v}_i$ for all $i \in \operatorname{supp}(\mathbf{v})$. Hence, the $\mathbf{v}$-maximal face is linearly isomorphic to a $\{-1, 0, 1\}$-polytope for which the maximal support of a vector is $d - |\text{supp}(\mathbf{v})|$. Thus, in at most $M(d - |\text{supp}(\mathbf{v})|)$ steps one can reach a $\mathbf{c}$-maximum of the $\mathbf{v}$-maximal face. By Lemma 2.1.2.4, the shadow path from such a $\mathbf{c}$-maximum is $\mathbf{v}$-improving. Hence, its length is at most the number of different values $\mathbf{v}^\intercal \mathbf{x}$ can take on P, which is $2|\text{supp}(\mathbf{v})|$. Hence, the length of a path from $\mathbf{v}$ to the optimum is at most $M(d - |\text{supp}(\mathbf{v})|) + 2|\text{supp}(\mathbf{v})|$. Suppose instead that $\mathbf{v} = \mathbf{0}$. Then $\mathbf{v}$ can reach a nonzero vertex by taking at most 1 step. Thus, we have

$$M(d) \leq \max(M(d-k) + 2k + 1).$$

By induction, we may assume that $M(j) \leq 3j$ for all $1 \leq j < d$. Then we have

$$M(d) \leq \max_{j \in [d]}(M(d-j) + 2j + 1)) \leq \max_{j \in [d]}(3(d-j) + 2j + 1) = \max_{j \in [d]}(3d - j + 1) = 3d.$$

Note that each step of this path may be computed by either taking an arbitrary improving neighbor over applying a shadow rule and it therefore may be made constructive. $\square$

For $(m+1)$-level polytopes, we may more directly apply the bound from Lemma 2.1.2.4:

Proof of Theorem 1.2.6. Let P be a $(\ell+1)$-level polytope. Assume without loss of generality that P is full dimensional. Let $\mathbf{v}$ be a vertex of P, and let $\mathbf{c} \in \mathbb{R}^n$. Let $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ be a set of n inequalities defining a basis at $\mathbf{v}$. Apply the affine transformation to P taking each $\mathbf{a}_i \to e_i$ to arrive at P'. Then for P', the vertex $\mathbf{v}$ is e_i-minimal for all $i \in [n]$. In particular, it $\mathbf{r} = \sum_{i=1}^n r^i e_i$-minimal for all $i \in [n]$ and $r > 0$. Then by Lemma 2.1.2.5 and our assumption that P is an $(\ell+1)$-level polytope, the shadow pivot rule for all sufficiently large $r > 0$ is of length at most $d\ell$.

$\square$

Next we turn to general lattice polytopes to prove Theorem 1.2.4. The key tool here is the natural choice of deterministic shadow rule that is the proper generalization of the slim shadow pivot rule. Namely, start with a $(0, k)$-lattice polytope $P \subseteq \mathbb{R}^n$ and put that lattice polytope in standard form $P = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{m+n} : A\mathbf{x} + \mathbf{y} = \mathbf{b}, \mathbf{x}, \mathbf{y} \geq \mathbf{0}\}$. In particular, $V(P) \subseteq \{0, k\}^n \times \mathbb{R}^m$. Assume our representation of P is irredundant. Then the pivot rule we use here is exactly the slim shadow pivot rule. However, in contrast to the 0/1-case, we only require our phase 1 procedure to start at a basis for which all nonbasic variables are contained in $[n+1, m+n]$.

Proof of Theorem 1.2.4. By Lemma 2.1.2.4, the path followed must be monotone with respect to $\mathbf{w}$. Let $(\mathbf{x}, \mathbf{y}) \in P$. Let $\mathbf{a}_1, \ldots, \mathbf{a}_m$ denote the rows of A. Then we have

$$\mathbf{w}^\intercal(\mathbf{x}, \mathbf{y}) = \sum_{i \in N \cap [n]} \mathbf{x}_i + \sum_{j \in N \cap [n+1, m+n]} \mathbf{y}_j = \sum_{i \in N} \mathbf{x}_i + \sum_{j \in N} b_j - \mathbf{a}_j^\intercal \mathbf{x}.$$

By Lemma 2.1.2.4, the path is $\mathbf{w}$-monotone, so the length is upper bounded by the number of different values $\mathbf{w}$ can take one. Note that

$$-|\{i : A_{ij} < 0\}| k \|A\|_\infty \leq \mathbf{a}_j^\intercal \mathbf{x} \leq |\{i : A_{ij} > 0\}| k \|A\|_\infty.$$

Since A is integral, this yields a bound of $n^2 k \|A\|_\infty$. $\square$

For the constructive diameter bound, we do not follow a single monotone path. Instead Lemma 2.1.2.5 guarantees that we can follow two shadow paths to reach the optimum. Namely, consider the following scheme for solving a linear program $\max(\mathbf{c}^\intercal \mathbf{x})$ such that $\mathbf{x} \in P$ for a lattice polytope P. Start at a basis B and choose a shadow $\mathbf{a} = \sum_{i \in [m+n] \setminus B} e_i$. Then choose $\mathbf{r}$ in terms of $\mathbf{a}$ in order to satisfy the assumptions of Lemma 2.1.2.5. Then the $\mathbf{a}$-shadow path followed to minimize $\mathbf{r}$ is precisely the shadow path from Lemma 2.1.2.5 but followed in reverse. Then to optimize $\mathbf{c}$, follow the $\mathbf{r}$-shadow path from the $\mathbf{r}$-minimizer. This yields a proof of Theorem 1.2.5.

CHAPTER 3

Monotone Path Polytopes

For results in the following chapter, it is recommended that one reads Appendix A first, since it covers all of the relevant definitions and lemmas necessary for studying monotone path polytopes.

3.1. Cross-polytopes

The section expands on my first paper with my advisor Jesús De Loera in which we studied the monotone paths on the cross-polytope in depth [24]. Here we provide new simplified proofs of the main theorems in that work together with some new observations. Having studied monotone path polytopes for three more years since those results, here I am able to offer cleaner statements and cleaner proofs. A remarkable distinction between the cross-polytope and other polytopes for which we can compute the monotone path polytope is that not all monotone paths are coherent. See for example Figure 3.1. However, we can easily characterize precisely which cellular strings are coherent and them provide a complete combinatorial description of the face lattice of the monotone path polytope.

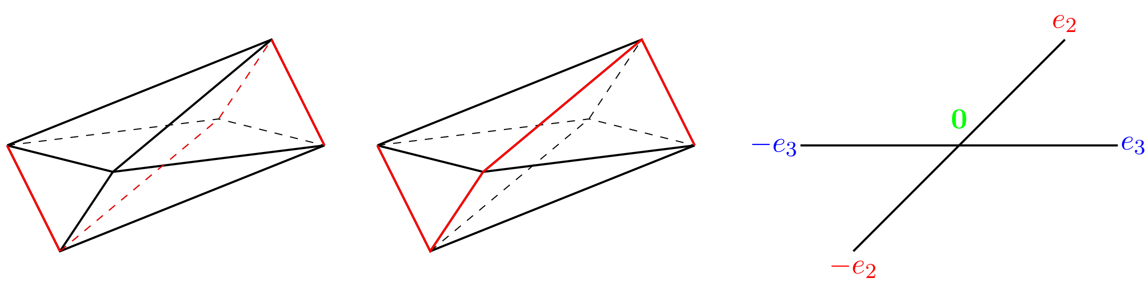

FIGURE 3.1. There are precisely two incoherent monotone paths on the octahedron, which are picture in the left and middle of the figure. The obstruction is pictured on the right, since both paths contain $-e_3, -e_2, e_2$, and e_3.

First recall that the cross-polytope $\diamond^n = \text{conv}(\{\pm e_i : i \in [n]\}) \subset \mathbb{R}^n$. A subset of vertices is a face of $\diamond^n$ if and only if it does not contain a pair of antipodes. In particular, faces of the cross-polytope are in bijection with **signed subsets**, that is a subset $S \subseteq [n]$ together with a map $\varepsilon : S \to \{\pm 1\}$. The key observation is the following:

Lemma 3.1.0.1. *Consider the cross-polytope* $\diamond^n$*, and* $\mathbf{c} \in \mathbb{R}^n$ *such that* $0 < c_1 < c_2 < \cdots < c_n$*. Then a cellular string* $\mathcal{C}$ *on* $\diamond^n$ *is not coherent if* $V(\mathcal{C})$ *contains* $\{\pm e_i\}$ *for some* $1 \leq i < n$*.*

Proof. Suppose for the sake of contradiction that $\mathcal{C}$ is coherent and contains $\{\pm e_i\}$. Then there exists $\mathbf{w} \in \mathbb{R}^n$ such that $\mathcal{C} = \pi^{-1}(U_{\mathbf{w}})$, where $\pi(\mathbf{x}) = (\mathbf{c}^\intercal \mathbf{x}, \mathbf{w}^\intercal \mathbf{x})$ and $U_{\mathbf{w}}$ is the upper path of $\pi(\diamond^n)$. Then $\pi(e_i)$ is on the upper path. Note that $\pi(\diamond^n)$ must be centrally symmetric, so $-\pi(e_i) = \pi(-e_i)$ is on the lower path. However, by assumption, $\pi(-e_i)$ is also on the upper path meaning that $\pi(-e_i)$ is on both the upper and lower path. This is only possible if $-e_i$ is $\mathbf{c}$-minimal or $\mathbf{c}$-maximal, so $i = n$, a contradiction. $\square$

The key observation is that this is actually the only obstruction to coherence, which allows us to prove our main theorem.

Theorem 3.1.1. *Let* $\mathbf{c} \in \mathbb{R}^n$*, and suppose that* $0 < c_1 < \cdots < c_n$*. Then the face lattice of* $\Sigma_{\mathbf{c}}(\diamond^n)$ *is the poset of intervals in the face lattice of* $\diamond^{n-1}$ *excluding the empty face. Furthermore, a cellular string is coherent on* $\Sigma_{\mathbf{c}}(\diamond^n)$ *if and only if it does not contain a pair of antipodes other than* $\{\pm e_n\}$*.*

Proof. We start by characterizing the cellular strings not containing a pair of antipodes other than $\{\pm e_n\}$. Then we will show the lattice of cellular strings with this property is isomorphic to the set of intervals in the face lattice of $\diamond^{n-1}$ with respect to inclusion. Finally, we will prove coherence. Note that any cellular string is determined by its endpoints and its set of vertices by Lemma A.4.0.1. By our assumption that they do not contain a pair of interior antipodes, the endpoints and vertices here are $E \subseteq V \subseteq \{\pm e_i : 1 \leq i < n\}$ together with $\{\pm e_n\}$. Note that E and V do not contain antipodes and therefore both correspond to faces of $\diamond^{n-1}$ and $E \subseteq V$. The map then taking a cellular string to an interval in the face lattice of $\diamond^{n-1}$ is the map taking the cellular string to the interval $[E, V]$. The inverse of this map is taking any pair $[E, V]$ to the cellular string with cells given by

$$\operatorname{conv}(\{\mathbf{x} \in V : \mathbf{c}^\intercal \mathbf{y}_{i-1} \leq \mathbf{c}^\intercal \mathbf{x} \leq \mathbf{c}^\intercal \mathbf{y}_i\}),$$

for each $i \in [k]$, where $\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_k$ is the sequence of endpoints in $\mathbf{c}$-increasing order. It follows from Lemma A.4.0.1, that the ordering in the set of cellular strings satisfying this property is precisely inclusion of intervals. Hence, this subposet of cellular strings is precisely the poset of intervals in the face lattice of $\diamond^{n-1}$.

It remains to show that any such cellular string is coherent. To do this, first choose $\mathbf{w}_i = 0$ for all i such that $e_i \notin V$ and $-e_i \notin V$, where V denotes the set of vertices contained in the cellular string other than $\{\pm e_n\}$. This ensures that each vertex not contained in $\pm V$ lies on the segment from $\pi(-e_n)$ to $\pi(e_n)$. Then note the remaining elements of V are the vertices of a simplex, so by applying the same proof as in the simplex case, by linear independence we may force them to all be at positive height and to yield the desired coherent cellular string for the upper path. All elements of $-V$ or in the complement of V will lie below the segment from $\pi(-e_n)$ to $\pi(e_n)$ and therefore not be contained in the upper path. Hence, the pre-image of the upper path is precisely the desired cellular string. Therefore, by Lemma 3.1.0.1, we have a cellular string is coherent if and only if it does not contain a pair of antipodes other than $\{\pm e_n\}$. $\square$

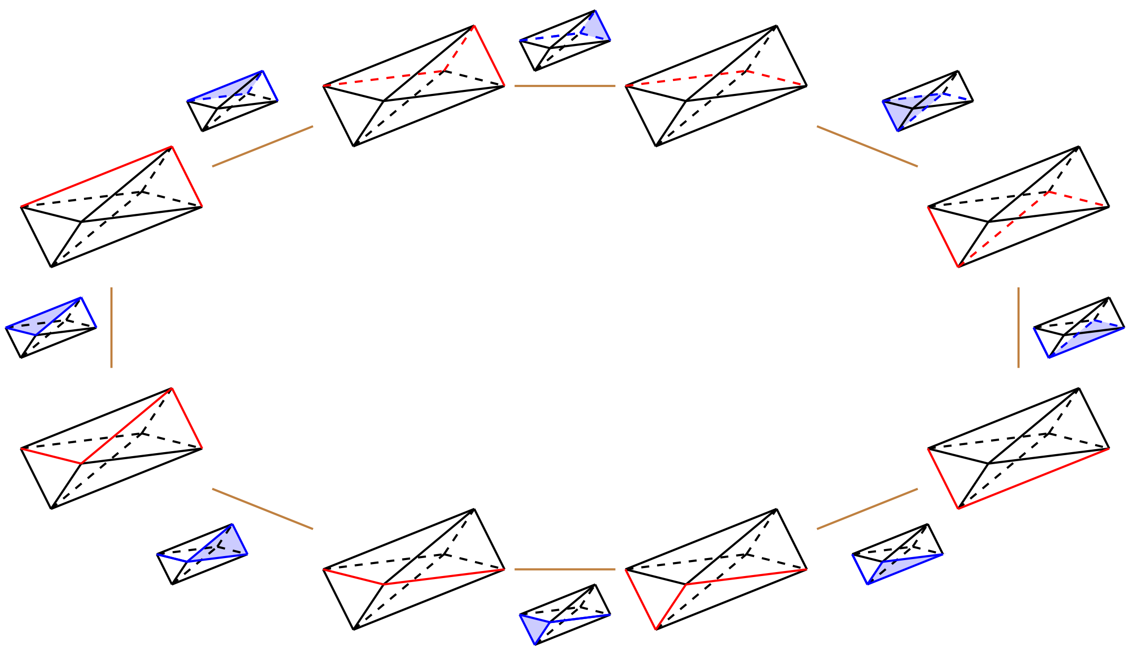

FIGURE 3.2. Pictured is $\Sigma_{(1,2,3)}(\diamond^3)$. The vertices are labeled by the corresponding coherent monotone paths and the edges are labeled by the corresponding coherent cellular strings. The two incoherent paths may be found in Figure 3.1.

Given a lattice the lattice of intervals in that poset under reverse inclusion is called the **antiprism** of that lattice [21]. For both the simplex and cross-polytope, the monotone path polytope is the polar dual of the anti-prism of a lower dimensional copy of the polytope. Let P^* denote the polar dual of a polytope P.

THEOREM 3.1.2. *Let* $\mathbf{c} \in \mathbb{R}^n$ *and suppose that* $0 < c_1 < \cdots < c_n$. *Then* $\Sigma_{\mathbf{c}}(\diamond^n)$ *is combinatorially equivalent to* $(\diamond^{n-1} + [-1,1]^{n-1})^*$

PROOF. To compute the vertices of $[0,1]^n + \diamond^n$, it suffices to find the $\mathbf{c}$-maximizers for generic $\mathbf{c}$ on $[0,1]^n + \diamond^n$. Let $\mathbf{c} \in \mathbb{R}^n$. Then the $\mathbf{c}$-maximizer on $\diamond^n$ is the vertex e_m such that $m = \operatorname{argmax}_{i \in [n] \cup [-n]} c_i$. By genericity, we may assume that m is unique. Analogously, the max on $[-1,1]^n$ returns the subset $\sum_{i \in R_+} e_i - \sum_{j \in R_-} e_j$ such that $R_+ = \{i : c_i > 0\}$ and $R_- = \{j : c_j < 0\}$, and we may assume by genericity of φ that $R_+ \cup R_- = [n]$. The sum of the two maximizers is then the corresponding vertex of $[-1,1]^n + \diamond^n$. The resulting vertex is then $e_m + \sum_{i=1}^n e_{\text{sign}(c_i)i} = 2e_m + \sum_{i \neq m} e_{\text{sign}(c_i)i}$.
Thus, $([-1,1]^n + \diamond^n)$ has vertex set $B_n\,(2e_1 + \sum_{i=2}^n e_i)$, where B_n is the set of signed permutation matrices. Let $S \in \{+,-,0\}^n \setminus \{\mathbf{0}\}$. Then S induces a partition of $[n]$ into $S_+ \cup S_- \cup S_0$, and there is a naturally associated linear functional φ_S given by $\sum_{i \in S_+} e_i^T - \sum_{j \in S_-} e_j^T$. The vertices maximized by φ_S are precisely

$$\{\text{sign}(k)e_k : k \in S_+ \cup S_-\} + \{\sum_{a \in A} \varepsilon(a)e_a : A \in S_0, \varepsilon : A \to \{\pm 1\}\}$$
$$+ \sum_{k \in S_+ \cup S_-} \text{sign}(k)e_k.$$

These vectors span the affine hyperplane given by

$$\sum_{i \in S_+} x_i - \sum_{j \in S_-} x_j = |S_+| + |S_-| + 1.$$

Thus, each sign vector φ_S corresponds to a facet of the resulting polytope. Let φ be a linear functional. Then any vertex maximized by that linear functional would also be maximized by the linear functional with the same sign pattern. Hence, the sign vectors induce all of the facets of $[-1,1]^n + \diamond^n$, which gives us a polyhedral formulation of this polytope.
Namely, for any partition $S_+ \cup S_- \cup S_0$ of n, we must have

$$\frac{1}{|S_+| + |S_-| + 1}\left(\sum_{i \in S_+} x_i - \sum_{j \in S_-} x_j\right) \leq 1.$$

These are precisely the relations given by maximizing the sign functionals. Observe that if a sign vector contains 0's, then any choice of $\pm e_i$ for $i \in S_0$ is allowable for a vertex in that facet.

Furthermore, if two facets have distinct positive sets or negative sets, then they cannot intersect, since that imposes the sign of e_i for any vertex in the set. It follows that two facets intersect if and only if their corresponding sign vectors are comparable. In particular, m-dimensional faces correspond exactly to intervals of sign vectors in the poset of length $n-m$. It follows then that the face lattice of this polytope is isomorphic to the lattice of intervals of the sign poset under reverse inclusion. Hence, we have $([-1,1]^{n-1}+\diamond^{n-1})^{\Delta}$ and $\Sigma_{\mathbf{c}}(\diamond^n)$ are combinatorially equivalent. $\square$

It is worth comparing the coherent monotone paths to the space of all monotone paths.

PROPOSITION 3.1.1. *The number of* $\mathbf{c}$*-monotone paths on* $\diamond^n$ *for* $\mathbf{c} \in \mathbb{R}^n$ *such that* $0 < c_1 < \cdots < c_n$ *is* $\frac{2}{3}(4^{n-1}-1)$.

PROOF. A monotone path corresponds to a subsequence $s_1, s_2, \ldots, s_m$ of

$$(-e_{n-1}, -e_{n-2}, \ldots, -e_1, e_1, e_2, \ldots, e_{n-1})$$

such that $s_k + s_{k+1} \neq 0$, since all vertices are connected to all vertices other than their antipodes. There are $2^{2(n-1)}-1$ non-empty subsets of $\{e_i : i \in \{\pm 1, \pm 2, \ldots, \pm n-1\}$. Observe that for $0 \leq k \leq n-2$, precisely 2^{2k} of those subsets contain $\{-e_{n-1-k}, e_{n-1-k}\}$ but are disjoint from $\{e_j : |j| < |n-1-k|\}$. Hence, the resulting number of possible sequences is

$$2^{2(n-1)} - 1 - \sum_{k=0}^{n-2} 2^{2k} = 2^{2(n-1)} - 1 - \frac{2^{2(n-1)}-1}{3} = \frac{2^{2n-1}-2}{3} = \frac{2}{3}(4^{n-1}-1).$$

$\square$

See OEIS Sequence A020988 [125], for other interpretations of the sequence of numbers of monotone paths. Note that, by Theorem 3.1.1, coherent monotone paths are in bijection with nonempty signed subsets $\{0,-1,1\}^n \setminus \mathbf{0}$. Hence, the total number of coherent monotone paths is $3^{n-1}-1$. We have the following corollary then:

COROLLARY 3.1.1. *The probability that a random monotone path is coherent tends to* 0 *at a rate* $\Theta((3/4)^n)$.

Thus, while all monotone paths are coherent on the simplex, the probability that a random monotone path on the cross-polytope is coherent tends to 0 exponentially fast. There is more we can show, and more is said in my joint paper with Jesús De Loera [24], but the remaining results all follow from our characterization of precisely which cellular strings are coherent. Namely, we can compute an explicit vertex and inequality description for a particular realization of the polytope, the f-vector, and extend some of our observations to centrally symmetric polytopes in general.

3.2. Products with Simplices

This is the first section on monotone path polytopes that has not appeared in any published form. Results on monotone path polytopes of products of two simplices will appear in a different form in joint work with Sophie Rehberg and Raman Sanyal. The goal of this section is to prove two main theorems. The first goal is to understand how the monotone path polytope plays with a basic construction: The prism. Namely, we show to compute the set of coherent cellular strings on the prism of a polytope from the set of coherent cellular strings on a polytope for corresponding choices of orientations. Furthermore, the prism operation is, by definition, the product with an edge. We prove something more general and study the product of a polytope with a simplex. In particular, we show the following:

Theorem 3.2.1. *Let $P \subset \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^m$. Suppose that all $\mathbf{c}$-cellular strings on P are coherent. Let $\mathbf{a} \in \mathbb{R}^n$ have distinct coordinates. Then all $(\mathbf{c}, \mathbf{a})$-cellular strings on $P \times \Delta_n$ are coherent.*

The second main goal is a complete characterization of the face lattice of a monotone path polytope of the product of two simplices. A characterization is easily found for a product of arbitrary many simplices. However, the picture for the product of two simplices is particularly nice. We already have as a corollary of the previous theorem the following:

Corollary 3.2.1. *For generic $\mathbf{a} \in \mathbb{R}^{n_1+n_2+\cdots+n_k}$, all cellular string on the product of simplices $\Delta_{n_1} \times \cdots \times \Delta_{n_k}$ are coherent.*

It is worth contrasting the statement with a different important problem concerning products of simplices. Namely, in my advisor Jesús De Loera's thesis [53], he showed that there exist non-regular triangulations of products of simplices $\Delta_m \times \Delta_n$ so long as m and n are sufficiently large. Regularity of triangulations is analogous to coherence of monotone paths. Both may be understood in terms of projections and the fiber polytope construction. However, for triangulations, understanding coherence on products of simplices is complicated. In fact, the corresponding object to the Baues poset for triangulations need not even be connected [103]. However, the analogous question for monotone paths is always positive. All cellular strings are coherent, and here we prove even more.

To prove the Theorem 3.2.1, we prove something stronger. The start is an understanding of cellular strings on products of polytopes. The following lemma from faces of products of polytopes being products of faces of each individual polytope.

LEMMA 3.2.1.1. *Let $P \subset \mathbb{R}^m$ and $Q \subset \mathbb{R}^n$ be polytopes, and let $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{a} \in \mathbb{R}^n$. Then $(\mathbf{c}, \mathbf{a})$-cellular strings on $P \times Q$ may be written as*

$$F_1 \times G_1, F_2 \times G_2, \dots, F_k \times G_k,$$

where $\{F_i : i \in [k]\}$ and $\{G_i : i \in [k]\}$ are $\mathbf{c}$-cellular strings on P and $\mathbf{a}$-cellular strings on Q respectively with F_i and G_i not necessarily distinct such that $F_i \neq F_{i+1}$ or $G_i \neq G_{i+1}$ for all $i \in [k-1]$. Furthermore, any sequence of that form yields a cellular string.

PROOF. It is clear that each cellular string would have to cells consisting of products of faces. The key additional condition here is for a face $F \times G$ of $P \times Q$, the $(\mathbf{c}, \mathbf{a})$-maximal face of a face $F \times G$ is the the product of the $\mathbf{c}$-maximal face of F and $\mathbf{a}$-maximal face of G. For any cell $F' \times G'$ that could be the next step of the string, if $F' \neq F$, the $\mathbf{c}$-minimum of F' must be the $\mathbf{c}$-maximum of F and an analogous condition holds if $G' \neq G$. Hence, the set of faces from P and Q that appear as a factor of some face in the cellular string must form cellular strings on each of P and Q.

To show that any sequence yields a cellular string of that form apply Lemma A.5.0.1 and note that the $(\mathbf{c}, \mathbf{a})$-maximum of $F \times G$ is the $(\mathbf{c}, \mathbf{a})$-minimum of $F' \times G$, $F \times G'$, and $F' \times G'$ for any F' or G' that could extend the corresponding string in each piece. $\square$

Another interpretation of this lemma may be stated directly in terms of shuffle permutations. Namely, given two ordered sets $A = \{a_1 < a_2 < \cdots < a_m\}$ and $\{b_1 < b_2 < \cdots < b_n\}$, a shuffle permutation is a total ordering of $A \cup B$. We are interested in more generally quotients of shuffle permutations. This is an ordering an $A \cup B$, which is total except we are allowed to say that pairs in A and B are equal. We will call such an order a **shuffle order**. For example, $a_1 < a_2 = b_1 < b_2$ is a shuffle order. However, $a_1 = a_2 = b_1 < b_2$ is not a shuffle order, since it identifies two distinct a_i. We may reinterpret products of cellular strings in terms of shuffle orders.

Lemma 3.2.1.2. *Let $P \subset \mathbb{R}^m$ and $Q \subset \mathbb{R}^n$ be polytopes, and let $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{a} \in \mathbb{R}^n$. Then $(\mathbf{c}, \mathbf{a})$-cellular strings on $P \times Q$ are in bijection with shuffle orders of pairs of $\mathbf{c}$-cellular strings on P and $\mathbf{a}$-cellular strings on Q.*

Proof. Consider a cellular string $F_1 \times G_1, F_2 \times G_2, \dots, F_k \times G_k$. There is a total ordering on $\{F_i\}$ and $\{G_i\}$ given by the restriction of the cellular string to each factor as demonstrated in Lemma 3.2.1.1. To define the corresponding shuffle order we say that $F_i < G_i$, $F_i > G_i$, or $F_i = G_i$, if $F_{i-1} = F_i$, $G_{i-1} = G_i$, or neither respectively. That is we record the new cell or pair of new cells at each step. Clearly this yields a shuffle order, and from any shuffle order one can recover the ordering on the cells. Furthermore, any shuffle order yields a squence by Lemma 3.2.1.1 due to the equivalence with the representation in the trichotomy. $\square$

For coherent cellular strings, we may be explicit about how exactly to obtain the cellular string on the product assuming we know the vector that realizes the cellular string.

Lemma 3.2.1.3. *Let $P \subset \mathbb{R}^m$ and $Q \subset \mathbb{R}^n$ be polytopes, and let $\mathbf{c}, \mathbf{w} \in \mathbb{R}^m$ and $\mathbf{a}, \mathbf{z} \in \mathbb{R}^n$. Let $F_1, F_2, \dots, F_k$ and $G_1, G_2, \dots, G_\ell$ be the $\mathbf{w}$-coherent $\mathbf{c}$-cellular string on P and $\mathbf{z}$-coherent $\mathbf{a}$-cellular string on Q respectively. Then the $(\mathbf{w}, \mathbf{z})$-coherent $(\mathbf{c}, \mathbf{a})$-cellular string $P \times Q$ is given by taking the cellular string induces by the shuffle order of F_i, G_j induced by taking the slope each face.*

Proof. This is immediate from the slope maximizing definition of coherence. Namely, at each step the next face in the string is the face spanned by the set of all slope maximizing edges. In the product the set of edges incident to a given vertex $(\mathbf{u}, \mathbf{v})$ is the union of edges incident to $\mathbf{u}$ and $\mathbf{v}$. The slopes of those edges are precisely those from each factor, since they lie entirely within each factor. It follows then that the set of maximal slope edges at $(\mathbf{u}, \mathbf{v})$ are either given by the maximal slope edges of $\mathbf{u}$, the maximal slope edges of $\mathbf{v}$, or a mix if the maximal slope edges happen to have the same slope. This is precisely the trichotomy used in Lemma 3.2.1.2. $\square$

Finally, our statement is not for general products, so it must use some structural advantage from taking the product with a simplex. For this, we have the following lemma,

Lemma 3.2.1.4. *Let $\mathbf{c} \in \mathbb{R}^n$ and suppose that all coordinates of $\mathbf{c}$ are distinct. Furthermore, let $F_0, F_1, \dots, F_k$ be a coherent $\mathbf{c}$-cellular string on Δ_n. Then for any $a_0 > a_1 > \cdots > a_k$, there exists*

$\mathbf{w} \in \mathbb{R}^n$ such that the slope of $\pi(F_i)$ is a_i for all $i \in [0,k]$, and $\mathbf{w}$ certifies the coherence the cellular string.

PROOF. Let $e_1, e_{i_1}, e_{i_2}, \ldots, e_{i_k}, e_n$ denote the endpoints of the cellular string such that e_{i_j} is the $\mathbf{c}$-minimum of F_j for all $j \in [k]$. Then, to choose $\mathbf{w}$, force $w_1 = 0$. Then for each $e_j \in F_0$, choose $\mathbf{w}_j$ so that (c_j, w_j) lies on the line of slope a_0 starting at $(c_1, 0)$. Do the same inductively for each remaining cell for the corresponding choice of a_i. Then for all remaining e_i not contained in the string, choose $\mathbf{w}_i$ so that they lie below the line from $(c_1, 0)$ to (c_n, w_n). Then, by construction, the pre-image of the upper path of the resulting polygon is the coherent cellular string and the slopes are as desired. $\square$

This statement tells us we have almost complete control over the sizes of slopes of edges in the shadow of a simplex. The only restriction is convexity, which forces consecutive slopes to decrease in the path. This is not in general going to be true for polytopes. In particular, this requires the edges in each coherent monotone path are linearly independent, which is why the result extends to the hyper-cube and cross-polytope but not necessarily to, say, the permutahedron or many other zonotopes in general. Then our main theorem for products with a simplex is as follows:

THEOREM 3.2.2. *Let $P \subset \mathbb{R}^m$ be a polytope, let $\mathbf{c} \in \mathbb{R}^m$, and let $\mathbf{a} \in \mathbb{R}^n$ with each coordinate of $\mathbf{a}$ distinct. Then the set of coherent $(\mathbf{c}, \mathbf{a})$-cellular strings on $P \times \Delta_n$ are given by*

$$F_1 \times G_1, F_1, \times G_2, \ldots, F_1 \times G_{k_1}, F_2 \times G_{k_1}, F_2 \times G_{k_1+1} \ldots, F_k \times G_{k_\ell},$$

where the set of F_i is a coherent $\mathbf{c}$-cellular string on P, the G_i form a coherent $\mathbf{a}$-cellular string on Δ_n, and the F_i are not necessarily distinct.

PROOF. Choose any $\mathbf{w}$ that certifies F_i. Then choose $b_0 > b_1 > \cdots > b_{k_\ell}$ in terms of the slopes of each F_i freely to ensure the desired shuffle order, which completes the proof. $\square$

Note that Theorem 3.2.1 follows as an immediate corollary to Theorem 3.2.2. Before understanding the product of simplices, we may start by taking a closer look at what monotone paths and coherent monotone paths look like for the case of a prism.

Proposition 3.2.1. *Let $P \subset \mathbb{R}^n$ be a polytope and let $\mathbf{c} \in \mathbb{R}^n$. Then the set of $(\mathbf{c}, 1)$ (coherent) monotone paths on $P \times [0,1]$, are in bijection with*

$$\{\gamma \times (V(\gamma) \cup \{\emptyset\}) : \gamma \text{ is a } \mathbf{c} - \text{monotone path}\}.$$

Proof. (Coherent) monotone paths on $P \times [0,1]$ are in bijection with shuffles of monotone paths on P and $[0,1]$ by Theorem 3.2.2. Then monotone paths are in bijection with a pair of path $\gamma = v_1, v_2, \ldots, v_k$ together with a choice of where to shuffle e_{n+1} into the sequence. By recording the last vertex before e_{n+1} and the path γ, we obtain the desired bijection. ☐

This yields an interesting observation about the possible distributions of monotone paths by length. Namely, we may define the **monotone path polynomial** by $f_{P,c}(x) = \sum_{k \geq 1} a_k x^k$, where a_k is the number of $\mathbf{c}$-monotone paths on a polytope of length k.

Corollary 3.2.2. *Let $P \subset \mathbb{R}^n$ be a polytope and let $\mathbf{c} \in \mathbb{R}^n$. Then the (coherent) monotone path polynomial satisfies:*

$$f_{P \times [0,1], (\mathbf{c},1)}(x) = x \left(\frac{d}{dx} \left(x f_{P,\mathbf{c}}(x) \right) \right).$$

Proof. From Proposition 3.2.1, we have the number of monotone paths on the prism of length $k+1$ is precisely the number of monotone paths of length k multiplied by $(k+1)$. Then the remainder of the proof follows from the generating function identity

$$x \frac{d}{dx} \left(x \sum_{k \geq 1} a_k x^k \right) = x \sum_{k \geq 1} (k+1) a_k x^k = \sum_{k \geq 0} (k+1) a_k x^{k+1}.$$

☐

There is a further generalization of this result to a product with a simplex.

Proposition 3.2.2. *Let $P \subset \mathbb{R}^m$ be a polytope and let $\mathbf{c} \in \mathbb{R}^m$. Then, if the coherent monotone path polynomial $f_P = \sum_{k \geq 1} a_k x^k$, we have*

$$f_{P \times \Delta_n, (\mathbf{c}, 1, 2, \ldots, n)} = \sum_{k \geq 1} \sum_{i=1}^{n-1} \binom{n-2}{i-1} \binom{i+k}{k} a_k x^{k+i}.$$

Proof. Fixing a monotone path of a given length k on P and a monotone path of a given length i in Δ_n, one finds the number of shuffles of the two paths together is precisely $\binom{k+\ell}{k}$ by a standard stars and bars argument. The number of monotone paths of length ℓ on Δ_n is precisely $\binom{n-2}{\ell-1}$, which yields the result. $\square$

3.2.1. Products of Simplices. Now we focus on the example of a product of two simplices. Consider standard simplices Δ_m and Δ_n and their product $\Delta_m \times \Delta_n$. We define the **colorful polygon** P_{m+n} associated to this product a polygon with m red vertices colored cyclically and then n blue vertices colored cyclically. In particular, we imagine the upper path of the polygon consists of m red vertices labeled from left to right $e_1, \dots, e_m$ and n blue vertices on the lower path labeled $f_1, f_2, \dots, f_n$ to represent the vertices of the two simplices. See Figure 3.3 for an example.

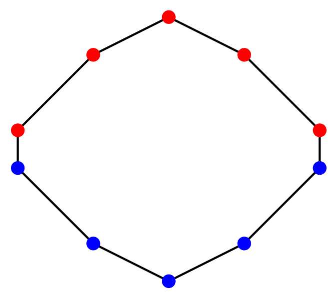

FIGURE 3.3. Pictured is the colorful polygon P_{5+5}.

Our goal is to define a bijection between cellular strings of on $\Delta_m \times \Delta_n$ and certain subdivisions of the polygon. We first show this for subdivisions of subpolygons. Namely, we have the following:

THEOREM 3.2.3. *Let $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{a} \in \mathbb{R}^n$ and assume the entries of $\mathbf{c}$ and $\mathbf{a}$ are strictly increasing. Then faces of $\Sigma_{(\mathbf{c},\mathbf{a})}(\Delta_m \times \Delta_n)$ are in bijection with subdivision of subpolygons of P_{m+n} such that each subpolygon contains the vertices corresponding to e_1, f_1, e_m, and f_n, and all internal edges are between vertices of different colors.*

PROOF. By Corollary 3.2.1, all cellular strings of $\Delta_m \times \Delta_n$ are coherent, so it suffices to show a bijection to the set of cellular strings. From Lemma A.5.0.1, cellular strings are determined uniquely by their endpoints and the set of vertices contained in each string. We define the map to subdivisions as follows. Namely, let E be the set of endpoints of a cellular string and V be the set of vertices of that appear in each cell of the string. Then we take the subpolygon with vertices given by the set of e_i and f_j that appears as a factor of some element of V, and we draw an edge from e_i to f_j whenever $e_i \times f_j$ is an endpoint of the cellular string.

To start, we need to show this map actually yields a subdivision of the subpolygon of the described type. Note that $e_1 \times f_1$ and $e_m \times f_n$ are endpoints of the string, since they are minimal and maximal with respect to the orientation induced by $(\mathbf{c}, \mathbf{a})$. Hence, the subpolygon contains e_1, f_1, e_m, and f_n. Furthermore, by construction, edges are always from e_i to f_j and thus, by construction, have different colors. It remains to show that this yields a subdivision, it suffices to show that no two edges cross. Note that, by construction, the edge from e_i to f_j and e_k to f_ℓ intersect if and only if $i < k$ and $j > \ell$ or $i < k$ and $j > \ell$.

Neither of these outcomes are possible due to monotonicity. Namely, faces of $\Delta_m \times \Delta_n$ are always of the $\text{conv}(\{e_i \times f_j : i \in S, j \in T\})$ for some $S \subseteq [m]$ and $T \subseteq [n]$. Since the coordinates of $\mathbf{c}$ and $\mathbf{a}$ are strictly increasing by hypothesis, the endpoints of a face are $e_{\min(S)} \times f_{\min(T)}$ and $e_{\max(S)} \times f_{\max(T)}$ respectively. Hence, if $e_i \times f_j$ and $e_k \times f_\ell$ are both endpoints $e_i \times e_j$ appearing first, then $i \leq j$ and $k \leq \ell$. Hence, the edges are do not cross, so the map yields a subdivision of a subpolygon.

For injectivity, note that the endpoints of the cellular string may be read off from the edges. Then for each cell of the subdivision of the subpolygon, the corresponding cell in the cellular string is given by

$$\text{conv}(\{e_i \times f_j : e_i, f_j \text{ are vertices of the cell}\}).$$

This is the inverse map and also may be defined for any subdivision of a subpolygon of the desired type. Hence, they are in bijection. □

Note that the monotone paths have another interpretation in terms of lattice paths. Namely, by recording the pair of indices (i, j), we see that monotone paths are in bijection with lattice paths from $(1, 1)$ to (m, n) by taking steps to the right or upward. Note that the length of the step may vary. We can have $(1, 1)$ move to $(3, 1)$ and then to $(4, 1)$. This is also considered distinct from $(1, 1)$ to $(4, 1)$ directly. These are also called **rook walks** due to a relation to the way the rook moves on a chessboard. There is an implicit bijection between these rook walks and monotone paths. From this theorem, we may immediately describe monotone paths in terms of subdivisions. Namely, they are precisely the subdivisions that are triangulations.

Corollary 3.2.3. *Let $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{a} \in \mathbb{R}^n$ and assume the entries of $\mathbf{c}$ and $\mathbf{a}$ are strictly increasing. Then faces of $\Sigma_{(\mathbf{c},\mathbf{a})}(\Delta_m \times \Delta_n)$ are in bijection with triangulations of subpolygons of P_{m+n} such that each subpolygon contains the vertices corresponding to e_1, f_1, e_m, and f_n, and all internal edges* *are between vertices of different colors. In particular, the number of vertices is*

$$\sum_{i=0}^{m-2}\sum_{j=0}^{n-2}\binom{m-2}{i}\binom{n-2}{j}\binom{i+j+2}{i+1}.$$

Proof. Following the bijection in Theorem 3.2.3, cells in the subdivision are in bijection with cells on the cellular string on the simplex. There the cells are of the form $\operatorname{conv}(\{e_i \times f_j : e_i, f_j \in C\})$ for each cell C in the subdivision of the subpolygon P_{m+n}. For a triangle, the vertices of the cell are either of the form $\{e_i, e_j, f_k\}$ or $\{e_i, f_j, f_k\}$ for some choice i, j, k, which is true if and only if the corresponding cell in P_{m+n} is an edge. Hence, a subdivision consists only of triangles if and only if the corresponding cellular string consists only of edges and is therefore a monotone path. For the formula, note that for a subpolygon with i vertices on top in addition to e_1 and e_m and j vertices on bottom in addition to f_1 and f_n, there are precisely $\binom{m-2}{i}$ and $\binom{n-2}{j}$ choices. Then for a fixed subpolygon of that type, the number of vertices is precisely the number of colorful triangulations of that polygon. These are in bijection with lattice paths from $(0,0)$ to $(i+1, j+1)$. Thus, there are precisely $\binom{i+j+2}{i+1}$ of them, which yields the desired formula. $\square$

See Figure 3.4 for an example of the bijection between rook walks and triangulations. For a fixed subpolygon, the set of colorful triangulations are in bijection with staircase triangulations of products of simplices, which correspond to the vertices of a different polytope called the **shuffle polytope** introduced by Gelfand, Kapranov, and Zelevinsky in Chapter 12, Section 2 of [75]. They are certain canonical triangulations of products of simplices that are always regular.

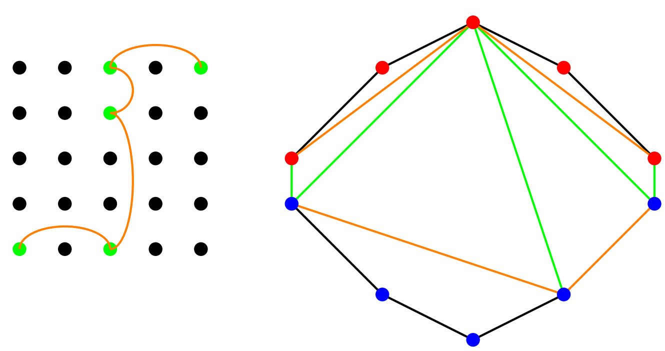

Figure 3.4. On the left is the rook walk from the bottom left node $(1,1)$ to $(5,5)$ for the monotone path $e_1 \times f_1, e_3 \times f_1, e_3 \times f_4, e_3 \times f_5, e_5 \times f_5$ on $\Delta_5 \times \Delta_5$. On the right is the colorful triangulation of a subpolygon for the same monotone path.

While we have a description of the faces, we have not yet described the ordering on the face lattice. We capture that in the following Corollary:

Corollary 3.2.4. *Let* $\mathbf{c} \in \mathbb{R}^m$ *and* $\mathbf{a} \in \mathbb{R}^n$ *and assume the entries of* $\mathbf{c}$ *and* $\mathbf{a}$ *are strictly increasing. Then the order on the face lattice of* $\Sigma_{(\mathbf{c},\mathbf{a})}(\Delta_m \times \Delta_n)$ *for subdivisions of subpolygons of* P_{m+n} *is as follows. Indicate a subdivision by its internal edges* E *and the set of vertices of the subpolygon* V*. Then for two subdivision* (E_1, V_1) *and* (E_2, V_2)*, we have* $(E_1, V_1) \leq (E_2, V_2)$ *if and only if* $E_1 \subseteq E_2$ *and* $V_1 \supseteq V_2$*.*

Proof. Since the set of internal edges of a subdivision correspond to endpoints of cells in cellular strings and the set of vertices that appear correspond to the set of vertices that appear in the cellular string, this result follows immediately from Lemma A.4.0.1. □

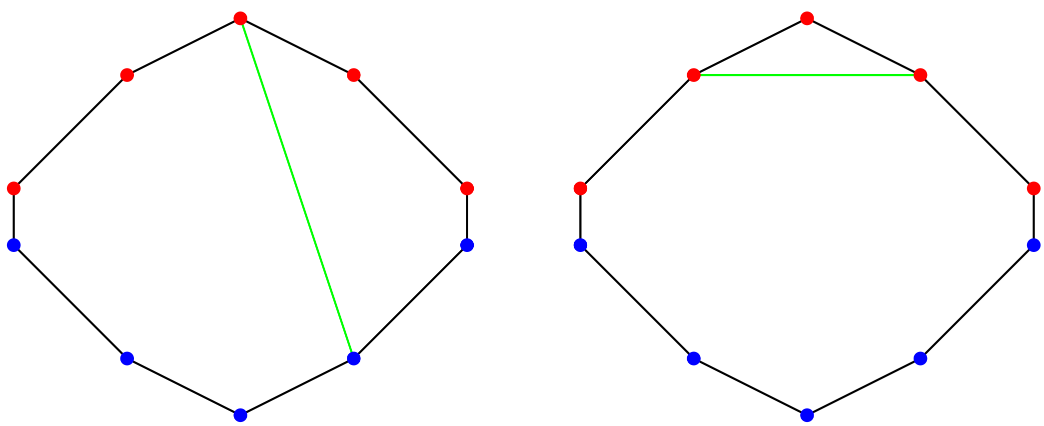

Figure 3.5. These subdivisions correspond to the two types of facets of the MPP of $\Delta_4 \times \Delta_4$. The left

From this corollary, we see that given a fixed subpolygon, the set of all faces contained in it are precisely the colorful subdivisions of subpolygons of that subpolygon. This observation allows us to also describe the facets of the polytope. See Figure 3.5 for pictures of the two cases of facets.

Corollary 3.2.5. *Let* $\mathbf{c} \in \mathbb{R}^m$ *and* $\mathbf{a} \in \mathbb{R}^n$ *and assume the entries of* $\mathbf{c}$ *and* $\mathbf{a}$ *are strictly increasing. Then the facets of* $\Sigma_{(\mathbf{c},\mathbf{a})}(\Delta_m \times \Delta_n)$ *are in bijection with* $\{e_i : 2 \leq i \leq m-1\}, \{f_j : j \leq i \leq n-1\}$*, and* $\{e_i \times f_j : (i,j) \notin \{(1,1),(m,n)\}$*. In particular, it has* $m + n + mn - 6$ *facets.*

Proof. For $\{e_i\}$ consider the subdivision of P_{m+n} given by taking the subpolygon including all vertices other than e_i. No edges may be removed from this subdivision to make it larger, and similarly no vertices may be added to make it larger without it being the trivial subdivision corresponding to a the face that is the whole polytope. Hence, it is a facet. There is analogous correspondence between $\{f_j\}$ and facets by excluding f_j. For $\{e_i \times f_j\}$ consider the subdivision that is the whole polygon but with a single edge from e_i to f_j. The same reasoning yields that this must be a facet.

To show all that facets arise in this way, note that any subdivision must either exclude a vertex or include one internal edge. By Corollary 3.2.4, we therefore have that any subdivision is less than or equal to one of this time in the partial order induced by the face lattice. Hence, these must be all of the facets, since no facet of a polytope is contained in another. □

Note that the number of vertices of the monotone path polytope of the product of simplices is far larger than the number of facets even though for the product of simplices, the numbers are the same. This is interesting from the standpoint of understanding the use of this construction. In optimization, we are interested in polytopes with few facets but many vertices, since then the inequality description of that polytope provides a compact way of representing that set of vertices by a set of inequalities that allows for optimizing over that large set efficiently via linear programming algorithms. The monotone path polytope being able to sometimes provide us with new polytopes with this property shows that this construction could give us new, interesting examples to study from an optimization standpoint. In optimization, we are also interested in non-degeneracy of polytopes. This also holds in this case:

COROLLARY 3.2.6. *Let* $\mathbf{c} \in \mathbb{R}^m$ *and* $\mathbf{a} \in \mathbb{R}^n$ *and assume the entries of* $\mathbf{c}$ *and* $\mathbf{a}$ *are strictly increasing. Then* $\Sigma_{(\mathbf{c},\mathbf{a})}(\Delta_m \times \Delta_n)$ *is a simple polytope.*

PROOF. By Corollary 3.2.3, the vertices of the monotone path polytope correspond to colorful triangulations of subpolygons of P_{m+n}. Note that $\Delta_m \times \Delta_n$ is $(m-1)+(n-1)$ dimensional, and so the monotone path polytope $(m-1)(n-1)-1$ dimensional. To show simplicity, it suffices to show that each vertex is contained in precisely $(m-1)+(n-1)-1 = m+n-3$ facets. By using the facet description from Corollary 3.2.5, we see that the number of facets containing a vertex is the number of internal edges plus the number of vertices excluded. For a triangulation of a subpolygon with i additional vertices on top and j additional vertices on bottom, this is precisely $m-2-i+n-2-j = m+n-i-j-4$. The number of internal edges of a triangulation of a convex polygon is always the number of vertices minus 3, which in this case yields $i+2+j+2-3 = i+j+1$. Adding these together yields $m+n-3$ as desired. $\square$

Thus, this is a simple polytope with faces corresponding to subdivisions of an n-gon and vertices corresponding to certain triangulations of n-gon, many of the properties of the associahedron. However, to our knowledge this polytope has not appeared previously in the literature. It is perhaps combinatorially equivalent to some member of a generalized class of associahedra. We leave this as an open problem:

Open Problem 3.2.4. *Is the monotone path polytope of a product of simplices combinatorially equivalent to the graph associahedron for some graph?*

The graph of the associahedron is of particular interest due to its appearance in many contexts including in the cluster algebra for the Grassmannian $Gr(2, n)$ [67] and in computer science through rotation of binary trees [124]. The graph of the monotone path polytope of the product of simplices is analogous.

Corollary 3.2.7. *Let $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{a} \in \mathbb{R}^n$ and assume the entries of $\mathbf{c}$ and $\mathbf{a}$ are strictly increasing. Then two vertices of $\Sigma_{(\mathbf{c},\mathbf{a})}(\Delta_m \times \Delta_n)$ are adjacent if and only if the corresponding triangulations differ by a flip or one is obtained from the other by adding a new vertex and internal edge to the vertex to the subpolygon.*

Proof. By Corollary 3.2.3, all vertices correspond to colorful triangulations of subpolygons. Then by Corollary 3.2.4, one may create an edge by either removing an internal edge of the triangulation or adding a new vertex to the subpolygon. These two possibilities yield precisely the two desired cases. $\square$

See Figure 3.6 for pictures of the two types of flips.

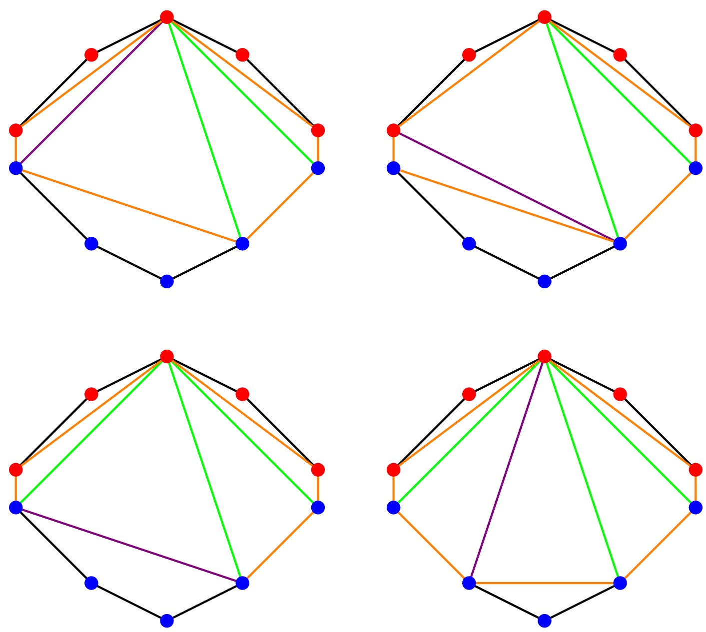

FIGURE 3.6. The two notions of adjacency for the monotone path polytope of a product of simplices are shown. The top shows a flip, and the bottom shows adding a new vertex.

There is a generalization of this picture for products of many simplices, but they are best done using the formulation in terms of lattice paths. In that case, the monotone path become lattice paths on a higher dimensional grid, and the cellular strings correspond to steps that are formed by taking unions of boxes in the grid. This can all be described explicitly but loses the nice quality of being interpreted in terms of certain subdivisions of a polygon, and we do not describe it here.

3.3. Pyramids of Polytopes

As in the previous section, the work in this section is new and has not appeared in any publication. The focus here is to address the following concern: When running the simplex method to solve an optimization problem, we do not necessarily know we starting at the minimum and walking to the maximum. There are cases for which this is true such as for optimization a non-negative linear function on an anti-blocking polytope such as the matching polytope of a bipartite graph, where we may always assume we are starting at the vertex $\mathbf{0}$. However, in general, this is rare and not representative of how the simplex method works. The main theorem of this section shows that this is actually not a problem.

For this, recall that the pyramid of P is defined by $\operatorname{pyr}(P) = \operatorname{conv}(\{\mathbf{0} \cup P \times \{1\}\}) \subset \mathbb{R}^n$. We define a **partial (coherent) cellular string** to be a sequence of faces $F_0, F_1, \ldots, F_k$ of P such that there exist faces $G_0, G_1, \ldots, G_\ell$ such that $G_0, G_1, \ldots, G_\ell, F_1, \ldots, F_k$ forms a (coherent) cellular string.

THEOREM 3.3.1. *Let $P \subset \mathbb{R}^n$ be a polytope, and let $\mathbf{c} \in \mathbb{R}^n$. Let $c^* = \min(\{\mathbf{c}^\intercal(\mathbf{x}) : \mathbf{x} \in P\})$. Then the set of coherent $(\mathbf{c}, -c^* + 1)$-cellular strings on pyr(P) is in bijection with $C \times \{0, 1\}$, where C is the set of partial coherent cellular strings. Coherent $(\mathbf{c}, -c^* + 1)$-monotone paths on pyr(P) are in bijection with partial coherent monotone paths.*

To prove this, we start with the case in which cellular strings are not coherent.

LEMMA 3.3.1.1. *Let $P \subset \mathbb{R}^n$ be a polytope, and let $\mathbf{c} \in \mathbb{R}^n$. Let $c^* = \min(\{\mathbf{c}^\intercal(\mathbf{x}) : \mathbf{x} \in P\})$. Then the set of $(\mathbf{c}, -c^* + 1)$-cellular strings on pyr(P) is in bijection with $C \times \{0, 1\}$, where C is the set of partial cellular strings. The $(\mathbf{c}, -c^* + 1)$-monotone paths on pyr(P) are in bijection with partial monotone paths.*

PROOF. By our choice of $(\mathbf{c}, -c^* + 1)$, $(\mathbf{c}, -c^* + 1)^\intercal(\mathbf{x}, 1) = \mathbf{c}^\intercal\mathbf{x} - c^* + 1 \geq 1 > 0$ for all $\mathbf{x} \in P$. It follows that $\mathbf{0}$ is the unique $(\mathbf{c}, -c^* + 1)$-minimum on $\operatorname{pyr}(P)$. Thus, the starting endpoint of every cellular string $\mathbf{0}$. Then the first cell is any face containing $\mathbf{0}$. For a pyramid, the set of faces containing $\mathbf{0}$ is the set of pyramids of each face of P. Hence, the first cell is $\operatorname{pyr}(F)$ for some face F of P. Then there are two cases. If F is orthogonal to $\mathbf{c}$, then $F \times \{1\}$ is an endpoint and we take any partial cellular string on P starting with $F \times \{1\}$ as its endpoint. If F is not orthogonal to $\mathbf{c}$, we take any partial cellular string on P with F as its start. Thus, for a fixed partial cellular string $F_0, F_1, \ldots, F_k$, there are always precisely two cellular strings on $\operatorname{pyr}(P)$ that correspond to it. Namely, either $\operatorname{pyr}(F_0), F_1, \ldots, F_k$ or $\operatorname{pyr}(E_0), F_0, F_1, \ldots, F_k$, where E_0 is the $\mathbf{c}$-minimum of F_0. This yields the desired bijection.

For monotone paths, note that $\operatorname{pyr}(F_0)$ would be a two dimensional face, since F_0 is an edge. Hence, there is a unique monotone on $\operatorname{pyr}(P)$ corresponding to each partial monotone path on P. □

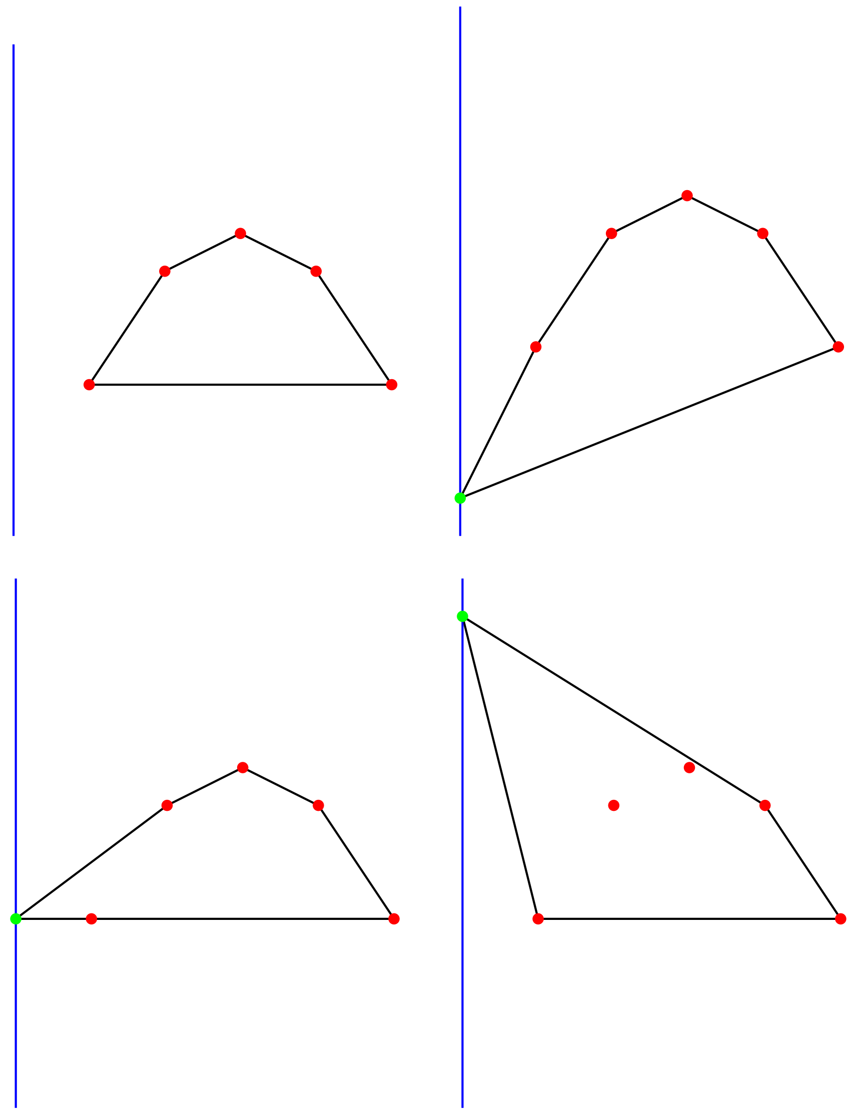

FIGURE 3.7. When projecting the pyramid of a polytope, one has complete control over the height of the additional point. The image shows various shadows of a pyramid of a polytope by changing the heights of the pyramid point. This illustrates the key idea behind the proof of Theorem 1.4.4.

To prove the theorem, it comes to down to showing for each coherent cellular string, that the resulting restrictions are coherent.

PROOF OF THEOREM 3.3.1. Let $F_0, F_1, \ldots, F_k$ be a partial coherent cellular string. Then by definition, there is a coherent cellular string $G_0, G_1, \ldots, G_\ell, F_0, F_1, \ldots, F_k$ that extends this cellular string. By the proof of Lemma 3.3.1.1, it suffices to show that $\operatorname{pyr}(F_0), F_1, \ldots, F_k$ and

$\mathrm{pyr}(E_0), F_0, F_1, \ldots, F_k$ are both coherent. To do this let $\mathbf{w} \in \mathbb{R}^n$ certify coherence of

$$G_0, G_1, \ldots, G_\ell, F_0, F_1, \ldots, F_k.$$

Define $\mathbf{w}_\lambda = (\mathbf{w}, \lambda)$ for each $\lambda \in \mathbb{R}$. Consider $\pi_\lambda(\mathbf{x}) = ((\mathbf{c}, -\alpha + 1)^\intercal(\mathbf{x}), \mathbf{w}_\lambda^\intercal \mathbf{x})$. Then, by the choice of $\mathbf{w}$, we have

$$\pi_\lambda(P \times \{1\}) = \{(\mathbf{c}^\intercal \mathbf{x} - c^* + 1, \mathbf{w}^\intercal \mathbf{x} + \lambda) : \mathbf{x} \in P\}.$$

Hence, the upper path on $\pi(P \times \{1\})$ is a translate of the upper path certifying coherence of the coherent cellular string $G_0, G_1, \ldots, G_\ell, F_0, \ldots, F_k$. Then we have

$$\pi_\lambda(\mathrm{pyr}(P)) = \mathrm{conv}(\{(0,0)\} \cup \{(\mathbf{c}^\intercal \mathbf{x} - c^* + 1, \mathbf{w}^\intercal \mathbf{x} + \lambda) : \mathbf{x} \in P\}).$$

One may visualize this parametrically as fixing $(0,0)$ and moving the polygon vertically uniformly up or down based on the choice of λ before taking the convex hull. Note that there is a unique choice of λ such that the line going the $\pi_0(F_0) + (0, \lambda)$ intersects $(0,0)$. Note that this line is supporting for $\pi_\lambda(P \times \{1\})$, since it contains the edge from the upper path corresponding to $\pi_\lambda(F_0)$. Furthermore, $(0,0)$ lies on this line, so it continues to be supporting for $\pi_\lambda(\mathrm{pyr}(P))$ and defines an edge going through $(0,0)$ and $\pi_\lambda(F_0)$. This edge must be the first edge of the upper path in the resulting polygon, while no other edges change. Hence, by construction, the resulting coherent cellular string is $\mathrm{pyr}(F_0), F_1, \ldots, F_k$.

To obtain $\mathrm{pyr}(E_0), F_0, F_1, \ldots, F_k$, add ε to λ so small that the slope from $(0,0)$ to $\pi(E_0)$ is greater than the slope of the line through $\pi_\lambda(F_0)$ and less than the slope of the line through $\pi_\lambda(G_\ell)$. Such a choice exists since the slopes of $\pi_\lambda(F_0)$ and $\pi_\lambda(G_\ell)$ are independent of λ and the slope from $(0,0)$ to $\pi(E_0)$ varies continuously. Then the line from $(0,0)$ to $\pi_{\lambda+\varepsilon}(E_0)$, intersects $\pi_{\lambda+\varepsilon}(G_\ell)$ at $\pi_{\lambda+\varepsilon}(E_0)$ and lies strictly above it prior to that point due to having a smaller slope. The line also lies strictly above the line through $\pi_{\lambda+\varepsilon}(F_0)$ after $\pi_{\lambda+\varepsilon}(E_0)$ by similar reasoning. This line is therefore supporting, and by similar reasoning to the previous case, this ensures that $\mathrm{pyr}(E_0), F_0, F_1, \ldots, F_k$ is the resulting coherent cellular string. Therefore, each partial cellular string that extends to a coherent cellular string is coherent for both possible realizations in the pyramid.

To show that these are all coherent cellular strings, note that we $\pi_\lambda(\mathrm{pyr}(P))$ for all choices of $\mathbf{w}$ and λ is precisely the set of all possible of shadows for $\mathrm{pyr}(P)$ for $(\mathbf{c}, -c^* + 1)$. The max-slope neighbor of $(0,0)$ will always lie on the upper path of $\pi_\lambda(P \times \{1\})$, since for any point on the interior of the projection, increasing its height will always yield a point of higher slope form $(0, 0)$. Then the max-slope neighbor must already be a vertex of the upper path. The max-slope neighbor of that vertex will not change after restricting. Hence, the resulting upper path will consist of an edge from $(0, 0)$ and then the last k edges of the original upper path induced by $\mathbf{w}$. For any such upper path, the pre-image is a coherent cellular string on $\mathrm{pyr}(P)$ of the desired form. Hence, the set of coherent cellular strings are precisely the partial coherent cellular strings together with the choice of taking the pyramid of the first endpoint or the first whole cell as desired. $\square$

3.4. 0/1-Polytopes

The results in this section have not appeared elsewhere in the literature. In the previous sections, we studied classes of polytopes and provided fairly complete descriptions of their monotone path polytopes either through an explicit description of their face lattices or the normal fans or sometimes even both. For 0/1-polytopes such a comprehensive description is likely out of reach. A **0/1-polytope** is any polytope whose vertices are in $\{0, 1\}^n$. These polytopes can be very complicated in general. Even determining whether two vertices are adjacent on the traveling salesman polytope, the convex hull of all indicator vectors of Hamiltonian cycles on a complete graph, is NP-complete [116]. If adjacency is hard to detect, then describing all possible monotone paths is even harder.

The focus of this section is instead to provide some sufficient conditions that make dealing with coherence tractable for large classes of 0/1-polytopes. The primary observation to take advantage of is that every 0/1 polytope is a subpolytope of a hyper-cube, and on a hyper-cube coherent monotone paths are easy to describe completely. In particular, all monotone paths are coherent. The following lemma gives a first idea of what this insight buys us.

LEMMA 3.4.0.1. *Let $P \subset \mathbb{R}^n$ be a polytope, and let Q be a subpolytope of P. Let $\mathbf{c} \in \mathbb{R}^n$. Let $F_0, F_1, \ldots, F_k$ be a $\mathbf{w}$-coherent $\mathbf{c}$-cellular string on P. Suppose that the each endpoint of the cellular for each cell F_ℓ for $\ell \in [i, j]$ is contained in Q, and suppose that the $\mathbf{c}$-minimum on Q is the left endpoint of F_i and the $\mathbf{c}$-maximum is the right endpoint of F_j. Then $F_i \cap Q, F_{i+1} \cap Q, \ldots, F_j \cap Q$ is a $\mathbf{w}$-coherent $\mathbf{c}$-cellular string on P.*

PROOF. Note that the set of faces and endpoints of the $\mathbf{w}$-coherent $\mathbf{c}$-cellular string on Q are precisely the set of faces that maximize $\mathbf{w} + \lambda\mathbf{c}$ for some $\lambda \in \mathbb{R}$. For any fixed λ, the $\mathbf{w} + \lambda\mathbf{c}$-maximizing face of Q is the $\mathbf{w} + \lambda\mathbf{C}$-maximizing face of P intersected with Q so long as that intersection is nonempty. Since each endpoint of the cellular string on P is contained in Q, that intersection is always nonempty for each face and endpoint of the cellular string. Hence, the resulting $\mathbf{w}$-coherent $\mathbf{c}$-cellular string on Q is must contain $F_i \cap Q, F_{i+1} \cap Q, \ldots, F_j \cap Q$. Since the left endpoint of F_i is a $\mathbf{c}$-minimum on Q and the right endpoint of F_j is a $\mathbf{c}$-maximum on Q, this coherent cellular string spans the full length of Q and therefore no further cells can be added. $\square$

To understand what this lemma means for 0/1-polytopes, we have the following immediate corollary:

Corollary 3.4.1. *Let P be a 0/1-polytope. Then if a monotone path on P is also a monotone path on $[0,1]^n$, then it is coherent on P.*

For 0/1-polytopes, we can say something even stronger by taking advantage of additional structure. Namely, we have the following useful technical lemma that is the core observation of this section. For each $S \subseteq T \subseteq [n]$, let $C_{S,T}$ denote the face of $[0,1]^n$ consisting all e_K such that $S \subseteq K \subseteq T$. Then we have the following:

Lemma 3.4.0.2. *Let $P \subset \mathbb{R}^n$ be a 0/1-polytope, and let $\mathbf{c} = (1, 1, \ldots, 1)$. Let $F_0, F_1, \ldots, F_k$ be a $\mathbf{c}$-cellular string with endpoints $e_{S_1}, e_{S_2}, \ldots, e_{S_{k+1}}$. Suppose that $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_{k+1}$ and each $F_i \subseteq C_{S_i,S_{i+1}} \cap P$. Then the cellular string is coherent.*

Proof. Let $C_{S,T}$ be the face of the cube given by all subsets in the interval $[S,T]$ for any $S \subseteq T \subseteq [n]$. Since the S_i form a flag, we have, by Lemma 3.4.0.1,

$$C_{S_1,S_2} \cap P, C_{S_2,S_3} \cap P, \ldots, C_{S_k,S_{k+1}} \cap P$$

is a coherent cellular string on P with the same endpoints as each F_i. Note that each F_i is contained in and therefore a face of $C_{S_i,S_{i+1}}$ by hypothesis. Thus, there exists a support vector $\mathbf{w}^i$ for F_i such that $\mathrm{supp}(\mathbf{w}) \subseteq S_{i+1} \setminus S_i$. For any $\varepsilon > 0$, the maximum slope face for $\mathbf{w} + \varepsilon \mathbf{w}^i$ of F_i starting at E_i is precisely F_i. Similarly, for any $\varepsilon > 0$ it does not change the slope of any other F_j, since $\mathbf{w}^i$ returns 0 for any difference of vectors on that face by our assumption about its support. However, we also need to require that ε is chosen so small that the slope of any other edge that is not maximal continues to not be maximal. Adding such a perturbation for each face in the path to yield a vector $\mathbf{w} + \sum_{i=1}^{n} \varepsilon_i \mathbf{w}^i$ yields a vector proving the cellular string is coherent. $\square$

This lemma is a nice sufficient condition that applies to understand coherence on a broad class of 0/1-polytopes. Namely, Matsui and Tamura studied monotone diameters of 0/1-polytopes of the form $\{\mathbf{x} \in [0,1]^n : A\mathbf{x} = \mathbf{b}\}$ [109]. An example of one such polytope is the Birkhoff polytope. There they showed that for any 0/1-polytope P of that form, any $\mathbf{v} \in V(P)$, and $\mathbf{c} \in \mathbb{R}^n$, there is a special monotone path from $\mathbf{v}$ to the $\mathbf{c}$-optimum such that at each step the set of coordinates that $\mathbf{v}$ and the $\mathbf{c}$-maximum $\mathbf{v}^*$ have in common strictly increases. They called such a path a **monotone vertex sequence**. Up to linear transformation, one may assume that $\mathbf{v}^* = (1, 1, \ldots, 1)$. In that case the monotone path is a flag of subsets on the cube. Therefore, by Lemma 3.4.0.2, the monotone path is always coherent. Hence, any 0/1-polytope from a large class contains a monotone path that may be proven to coherent in this manner.

We are also generally interested in polytopes for which all cellular strings are coherent, since then understanding the face lattice of the monotone path polytope reduces to the combinatorial question of understanding the Baues poset of cellular strings. We show that all cellular strings are coherent for a large class of 0/1-polytopes vastly generalizing the observation that all cellular strings on hyper-cubes are coherent.

Corollary 3.4.2. *Let P be a 0/1-polytope, and suppose that all edge directions are of the form e_S for some $S \subseteq [n]$. Then all $(1, 1, \ldots, 1)$-cellular strings on P are coherent.*

Proof. Since all edge directions are of the form e_S, the endpoints of the cells in any cellular string have to form a flag. Furthermore, every vertex in each cell must be able to be reached by a path from the minimum of the flag to the maximum and therefore must both contain the minimum and be contained in the maximum set of each cell. It follows then that each cell must be contained in the corresponding face $C_{S,T}$ of the cube. Therefore, by Lemma 3.4.0.2, all cellular strings are coherent. □

There is another interesting class of polytopes where we may prove a slightly weaker statement.

Lemma 3.4.0.3. *Let P be a 0/1-polytope, and let $\mathbf{c} = (1, 1, \ldots, 1)$. Then if all edge directions of P are of the form e_S or are orthogonal to $\mathbf{c}$, all $\mathbf{c}$-monotone paths on P are coherent.*

Proof. Any $\mathbf{c}$-monotone path must be strictly monotone and therefore all edges are of the form e_S, since the all remaining edges are orthogonal to $\mathbf{c}$. Hence, any $\mathbf{c}$-monotone path must be coherent by applying the same argument as in Lemma 3.4.0.2. □

CHAPTER 4

Pivot Rule Polytopes

4.1. The Pivot Rule Polytope Construction

Recall from Section 1.6 that the Pivot Rule Polytope for a linear program $\max(\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P\})$ and fixed normalization η is given as follows:

$$\Pi_{\eta,\mathbf{c}}(P) = \sum_{v \in V(P)} \operatorname{conv}\left(\left\{\frac{\mathbf{u}-\mathbf{v}}{\eta(\mathbf{u}-\mathbf{v})} : \mathbf{u} \in N_{\mathbf{c}}(\mathbf{v})\right\}\right),$$

where $N_{\mathbf{c}}(\mathbf{v})$ denotes the set of $\mathbf{c}$-improving neighbors of $\mathbf{v}$. The proof of Theorem 1.6.2 follows from unpacking this definition:

Proof. Proof of Theorem 1.6.2 To prove the characterization of the normal fan, it suffices to show for generic $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ that $\mathbf{a}$ and $\mathbf{b}$ are maximized at the same vertex of $\Pi_{\eta,\mathbf{c}}(P)$ if and only if they induce the same footprint for the normalized weight pivot rules with normalization η and weights $\mathbf{a}$ and $\mathbf{b}$ respectively.

From the general theory of Minkowski sums as in Chapter 7 of [138], the $\mathbf{c}$-maximal face of a Minkowski sum of polytopes corresponds exactly to the sum of all simultaneously $\mathbf{c}$-maximal faces on each of the summands. If the maximizer is a vertex, this decomposition is unique. In our case, then a maximizer of $\Pi_{\eta,\mathbf{c}}(P)$ corresponds uniquely to a choice of maximizer

$$\operatorname{conv}\left(\left\{\frac{\mathbf{u}-\mathbf{v}}{\eta(\mathbf{u}-\mathbf{v})} : \mathbf{u} \in N_{\mathbf{c}}(\mathbf{v})\right\}\right)$$

for all $\mathbf{v} \in V(P)$. In particular, a vertex of $\Pi_{\eta,\mathbf{c}}(P)$ maximizing $\mathbf{a}$ corresponds to maximizing

$$\operatorname*{argmax}_{\mathbf{u} N_{\mathbf{c}}(\mathbf{v})} \frac{\mathbf{a}^\intercal(\mathbf{u}-\mathbf{v})}{\eta(\mathbf{u}-\mathbf{v}}.$$

Therefore $\mathbf{a}$ and $\mathbf{b}$ are maximized at the same vertex of $\Pi_{\eta,\mathbf{c}}(P)$ if and only if

$$\operatorname*{argmax}_{\mathbf{u} N_{\mathbf{c}}(\mathbf{v})} \frac{\mathbf{a}^\intercal(\mathbf{u}-\mathbf{v})}{\eta(\mathbf{u}-\mathbf{v}} = \operatorname*{argmax}_{\mathbf{u} N_{\mathbf{c}}(\mathbf{v})} \frac{\mathbf{b}^\intercal(\mathbf{u}-\mathbf{v})}{\eta(\mathbf{u}-\mathbf{v}}$$

for all each vertex $\mathbf{v}$ of P. Hence, two vertices are co-maximal if and only if they induces the same footprint for their corresponding normalized weight pivot rules as desired. $\square$

The proof given here was simple with the main complexity being finding the right construction of the polytope. In particular, the key observation is that to generalize the monotone path polytope, one can consider arborescences instead of monotone paths.

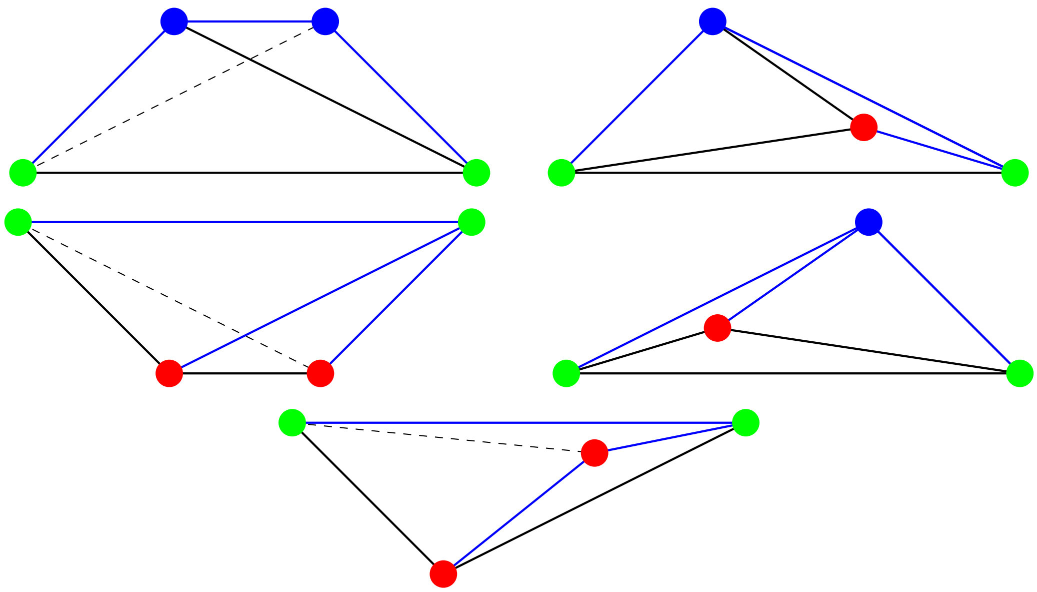

FIGURE 4.1. The five max-slope arborescences on a tetrahedron $\mathrm{conv}(e_1, e_2, e_3, e_4)$ are depicted for $\mathbf{c} = (0, 1, 2, 3)$. The coherence of the top left and middle left are shown using $\mathbf{w} = (0, 1, 1, 0)$ and $\mathbf{w} = (0, -1, -1, 0)$ respectively. The top right and middle right are shown to be coherent using $\mathbf{w} = (0, 1, .3, 0)$ and $\mathbf{w} = (0, .3, 1, 0)$ respectively. The bottom uses $\mathbf{w} = (0, -1, -.2, 0)$. Compare to Figure 1.7.

It remains open whether there is a similar construction that only records the monotone paths. Such a construction could still be possible. Namely, we make the following observation:

PROPOSITION 4.1.1. *Consider a linear program* $\max(\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P\})$ *for a polytope* P*, and suppose that* $\mathbf{c}$ *is generic. Fix a normalization* η*. Let* C *denote the set all vectors* $\mathbf{w} \in \mathbb{R}^n$ *that induce the same monotone path from the* $\mathbf{c}$*-minimum on* P *to the* $\mathbf{c}$*-maximum on* P *for the normalized weight pivot rule with normalization* η *and weight* $\mathbf{w}$*. Then* C *is an open polyhedral cone.*

PROOF. To prove this, let $\mathbf{v}^0, \mathbf{v}^1, \ldots, \mathbf{v}^k$ be a monotone path from the $\mathbf{c}$-minimum to the $\mathbf{c}$-maximum on P. Then this path is chosen by the normalized weight pivot rule for normalization η

and weight $\mathbf{w}$ if and only if

$$\frac{\mathbf{w}^\intercal(\mathbf{v}^i - \mathbf{v}^{\mathbf{i-1}})}{\eta(\mathbf{v}^i - \mathbf{v}^{i-1})} > \frac{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{v}^i)}{\eta(\mathbf{u} - \mathbf{v}^{i-1})}$$

for all $\mathbf{u} \in N_{\mathbf{c}}(\mathbf{v}^{i-1})$ and $i \in [k]$. Since η and the polytope are fixed, each inequality here is linear, so $\mathbf{w}$ chooses a given monotone path if and only if it satisfies a fixed linear inequalities. Hence, the set of all such $\mathbf{w}$ is an open polyhedral cone. $\square$

We leave open whether the set of all such cones forms a fan and whether that fan is the normal fan of some polytope. There is at least one special case where this results. Namely, recall that the max-slope pivot rule polytope is given by

$$\Pi_{\mathbf{c}}(P) = \sum_{v \in V(P)} \operatorname{conv}\left(\left\{\frac{\mathbf{u} - \mathbf{v}}{\mathbf{c}^\intercal(\mathbf{u} - \mathbf{v})} : \mathbf{u} \in N_{\mathbf{c}}(\mathbf{v})\right\}\right).$$

Then we have the following observation:

LEMMA 4.1.0.1. *Consider a linear program* $\max(\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P\})$ *for a polytope* P. *Let* C *denote the set of all* $\mathbf{w} \in \mathbb{R}^n$ *that induce the same monotone path from the* $\mathbf{w}$*-maximum of the* $\mathbf{c}$*-minimal face of* P. *Then* C *is a normal cone in the monotone path polytope* $\Sigma_{\mathbf{c}}(P)$.

PROOF. This follows immediately from the definition of coherence in terms of slopes in Theorem 1.3.7. $\square$

Thus, in the case of the max-slope pivot rule, the set of cones for each monotone path does form the normal fan of a polytope. Using this observation, we may prove Theorem 1.6.4.

PROOF. Proof of Theorem 1.6.4 To show that $\Sigma_{\mathbf{c}}(P)$ is a weak Minkowski summand of $\Pi_{\mathbf{c}}(P)$, it suffices to show that for generic $\mathbf{w} \in \mathbb{R}^n$, the $\mathbf{w}$-maximum on $\Pi_{\mathbf{c}}(P)$ determines the $\mathbf{w}$-maximum on $\Sigma_{\mathbf{c}}(P)$. Consider the max-slope arborescence for a weight $\mathbf{w}$. Then in the arborescence there is a unique $\mathbf{c}$-monotone path from the unique $\mathbf{c}$-minimum of P. From Lemma 4.1.0.1, this monotone path is precisely the $\mathbf{w}$-coherent $\mathbf{c}$-monotone path. Hence, from the max-slope arborescence corresponding to $\mathbf{w}$, one can completely recover the corresponding coherent monotone path as desired.

Suppose that P is a zonotope. Then, as noted in Chapter 7 of [138], the normal fan of P is a hyper-plane arrangement. Let $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^k$ denote the normals of the arrangement and choose them so that $-\mathbf{c}^\intercal \mathbf{v}^i < 0$ for all $i \in [k]$, which must be possible by our assumption that P has a unique **c**-minimum. As discussed in Chapter 7 of [138] and in standard oriented matroid theory, the regions of the hyperplane arrangement are given by a vector $\varepsilon \in \{-1,1\}^k$, where the region corresponding to ε is precisely $C_\varepsilon = \{\mathbf{x} \in \mathbb{R}^n : \varepsilon_i(\mathbf{v}^i)^\intercal \mathbf{x} \geq \mathbf{0}$. The minimal vertex corresponds to $\varepsilon = (-1,-1,\ldots,-1)$ and the maximal region correspond to $\varepsilon = (1,1,\ldots,1)$.

Two regions are adjacent if they differ by a flip of a single sign, and a monotone path is a sequence $(-1,\ldots,-1) = \varepsilon^0, \varepsilon^1, \ldots, \varepsilon^k = (1,1,\ldots,1)$ such that at each step, the total number of positive coordinates of ε^i increases. A sign flip in each coordinate occurs precisely once, and flipping the sign in coordinate i corresponds to following an edge parallel to $\mathbf{v}^i$. Thus, a monotone path corresponds to a total ordering of $\mathbf{v}^{\sigma(1)}, \ldots, \mathbf{v}^{\sigma(k)}$ for $\sigma : [k] \to [k]$. If the path is **w**-coherent, by convexity, the slopes of consecutive edges in the path are decreasing. Hence, a coherent monotone path determines a total ordering on the slopes of all edge directions. A total ordering on the slopes all edge directions determines a max-slope arborescence. Therefore, for P a zonotope, we also have $\Pi_{\mathbf{c}}(P)$ is a weak Minkowski summand of $\Sigma_{\mathbf{c}}(P)$. Hence, in that case, $\Sigma_{\mathbf{c}}(P)$ and $\Pi_{\mathbf{c}}(P)$ are normally equivalent. $\square$

The assumption that there is a unique **c**-minimum is necessary. Namely, consider following degenerate example: $P = \operatorname{conv}(e_1, e_2, e_3, 2e_1, 2e_2, 2e_3)$. Then for the orientation $(1,1,1)$, by definition, the monotone path polytope is precisely

$$\Sigma_{(1,1,1)}(P) = \operatorname{conv}(e_1,e_2,e_3) + \operatorname{conv}(2e_1,2e_2,2e_3) = \operatorname{conv}(3e_1,3e_2,3e_3).$$

However, the max-slope pivot rule polytope by definition is:

$$\Pi_{(1,1,1)}(P) = \operatorname{conv}(e_1) + \operatorname{conv}(e_2) + \operatorname{conv}(e_3) = \{(1,1,1)\}.$$

Hence, in that case $\Sigma_{(1,1,1)}(P)$ is not a weak Minkowski summand of $\Pi_{(1,1,1)}(P)$. Note also that here the dimensions of the two do not even coincide. In general, however, they do.

THEOREM 4.1.1. *Let P be a d-dimensional polytope in $\mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^n$, and suppose that $\mathbf{c}$ has a unique minimum on P. Then $\Pi_{\mathbf{c}}(P)$ has dimension $d-1$ if $d \geq 1$.*

PROOF. First recall that Billera and Sturmfels showed in [17] that the dimension of $\Sigma_{\mathbf{c}}(P)$ is always $d-1$ so long as **c** is nonconstant on P. Since P has a unique **c**-minimum, **c** is nonconstant on P, so $\Sigma_{\mathbf{c}}(P)$ is of dimension $d-1$. Furthermore, by Theorem 1.6.4, $\Sigma_{\mathbf{c}}(P)$ is a weak Minkowski summand of $\Pi_{\mathbf{c}}(P)$, so $\Pi_{\mathbf{c}}(P)$ has dimension at least the dimension of $\Sigma_{\mathbf{c}}(P)$ and therefore has dimension at least $d-1$. Let $\mathbf{v}^*$ be a vertex of $\Pi_{\mathbf{c}}(P)$. Then, by definition,

$$\mathbf{v}^* = \sum_{v \in V(P) \setminus V(F)} \frac{\mathbf{u}-\mathbf{v}}{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})},$$

where F is the $\mathbf{c}$-maximal face of P. Hence, we must have,

$$\mathbf{c}^\intercal \mathbf{v}^* = \mathbf{c}^\intercal \sum_{v \in V(P) \setminus V(F)} \frac{\mathbf{u}-\mathbf{v}}{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})} = \sum_{v \in V(P) \setminus V(F)} \frac{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})} = |V(P) \setminus V(F)|.$$

Hence, $\Pi_{\mathbf{c}}(P)$ is contained in the hyperplane $\mathbf{c}^\intercal \mathbf{x} = |V(P) \setminus V(F)|$. Furthermore, suppose that the smallest plane containing P is given by $A\mathbf{x} = \mathbf{b}$. Then we must have

$$A\mathbf{v} = \sum_{v \in V(P) \setminus V(F)} \frac{A(\mathbf{u}-\mathbf{v})}{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})} = \sum_{v \in V(P)} \frac{\mathbf{b}-\mathbf{b}}{\mathbf{c}^\intercal(\mathbf{u}-\mathbf{v})} = \mathbf{0}.$$

Note that $\mathbf{c}$ is not normal to that plane, so the codimension of $\Pi_{\mathbf{c}}(P)$ is at most $n-d+1$ meaning that $\Pi_{\mathbf{c}}(P)$ has dimension $d-1$. ☐

From our description so far, we understand only the vertices of the pivot rule polytope, and this is already enough to prove many structural features. The faces of the polytope have an alternative description. Namely, faces occur exactly when there is a tie in which two or more improving neighbors all have maximal slope. The set of vertices of the face are precisely the set of possible methods of breaking ties via perturbing $\mathbf{w}$. To encode this, we just take the set of neighbors that are tied and treat that as an outgoing neighbor. We call this a **multi-arborescence**. To be formal an arborescence for us a map $\mathcal{A} : V(P) \setminus V(P^{\mathbf{c}}) \to V(P)$, where $P^{\mathbf{c}}$ is the $\mathbf{c}$-maximal face of P, such that $\mathcal{A}(\mathbf{v}) \in N_{\mathbf{c}}(\mathbf{v})$ for all $\mathbf{v} \in V(P)$. A multi-arborescence has the same definition except $\mathcal{A} : V(P) \setminus V(P^{\mathbf{c}}) \to 2^{V(P)}$ such that $\mathcal{A}(\mathbf{v}) \subseteq N_{\mathbf{c}}(\mathbf{v})$.

PROPOSITION 4.1.2. *The face lattice of the pivot rule polytope $\Pi_{\eta,\mathbf{c}}(P)$ for a polytope $P \subseteq \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^n$, and a normalization η is the poset of $\mathbf{c}$-multi-arborescences, where $\mathcal{A} \leq \mathcal{A}'$ if $\mathcal{A}(\mathbf{v}) \subseteq \mathcal{A}'(\mathbf{v})$ for all $\mathbf{v} \in V(P) \setminus V(F)$ with F the $\mathbf{c}$-maximal face of P.*

PROOF. This follows immediately from unpacking what it means to be the face of a Minkowski sum of polytopes. ☐

4.2. The Associahedron as a Pivot Rule Polytope

To prove Theorem 1.6.5, we will rely on an interpretation of the max-slope pivot rule polytope in terms of collisions of particles moving on a line. Suppose we have particles racing on a track at constant speeds: $\mathbf{c}_1 > \mathbf{c}_2 > \cdots > \mathbf{c}_n$ starting at points $\mathbf{w}_1 \leq \mathbf{w}_2 \leq \cdots \leq \mathbf{w}_n$. For $i < j$, the speed of particle i is greater than the speed of particle j. Furthermore, particle i starts behind particle j. Hence, eventually, at time t_{ij}, particle i will catch up to particle j. To model this, we see the following:

$$\begin{aligned}
\mathbf{w}_i + \mathbf{c}_i t_{ij} &= \mathbf{w}_j + \mathbf{c}_j t_{ij} \\
(\mathbf{c}_i - \mathbf{c}_j) t_{ij} &= \mathbf{w}_j - \mathbf{w}_i \\
t_{ij} &= \frac{\mathbf{w}_j - \mathbf{w}_i}{\mathbf{c}_i - \mathbf{c}_j} \\
&= -\frac{\mathbf{w}_j - \mathbf{w}_i}{\mathbf{c}_j - \mathbf{c}_i}
\end{aligned}$$

Note that $t_{ij} \geq 0$, since $\mathbf{w}_i - \mathbf{w}_j \leq 0$ and $\mathbf{c}_j - \mathbf{c}_i < 0$ by assumption for all $j > i$. Note that maximizing $-t_{ij}$ is equivalent to minimizing t_{ij}. Thus, computing a max-slope arborescence on the standard simplex for the orientation $\mathbf{c}_1 > \mathbf{c}_2 > \cdots > \mathbf{c}_n$ with auxiliary vector $\mathbf{w}_1 \leq \mathbf{w}_2 \leq \cdots \leq \mathbf{w}_n$ corresponds to recording for each particle, the first particle that catches up to it.

Lemma 4.2.0.1. *Let $\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_n)$ such that $\mathbf{c}_1 > \mathbf{c}_2 > \cdots > \mathbf{c}_n$. Then every normal cone of $\Pi_{\mathbf{c}}(\Delta^n)$ contains a vector $\mathbf{w}$ such that $\mathbf{w}_1 \leq \mathbf{w}_2 \leq \cdots \leq \mathbf{w}_n$.*

Proof. Let C be a normal cone of the max-slope pivot rule polytope, and let $\mathbf{x} \in C$. Note that $\mathbf{c}$ is in the lineality space of the max-slope pivot rule polytope, so $\mathbf{x} + \lambda\mathbf{c} \in C$ for any $\lambda \in \mathbb{R}$. Since $\mathbf{c}_1 > \mathbf{c}_2 > \cdots > \mathbf{c}_n$, there exists $\lambda > 0$ such that $\mathbf{x}_1 - \lambda\mathbf{c}_1 \geq \mathbf{x}_2 - \lambda\mathbf{c}_2 \geq \cdots \geq \mathbf{x}_n - \lambda\mathbf{c}_n$. Hence, for such a choice of λ, $\mathbf{x} - \lambda\mathbf{c} \in C$ and $\mathbf{x} - \lambda\mathbf{c}$ has decreasing coordinates as desired. $\square$

Therefore, without loss of generality, we may always assume that $\mathbf{w}_1 \leq \mathbf{w}_2 \leq \cdots \leq \mathbf{w}_n$. Hence, the particle model is sufficient to completely describe the combinatorics of the pivot rule polytope of a simplex. A first observation that will be useful for understanding particle collisions is the following:

Lemma 4.2.0.2. *Consider a particle system with n particles at initial positions $\mathbf{w}_1 < \mathbf{w}_2 < \cdots < \mathbf{w}_n$ all moving at speed $\mathbf{c}_1 > \mathbf{c}_2 > \cdots > \mathbf{c}_n$. For each $1 \leq a < b \leq n$, let t_{ab} denotes the time of collision* *between particles a and b. Then for any $1 \leq i < j < k \leq n$,*

$$\min(t_{ij}, t_{jk}) \leq t_{ik} \leq \max(t_{ij}, t_{jk})$$

with equality if and only if $t_{ij} = t_{jk} = t_{ik}$.

Proof. Suppose first that the collisions times are all distinct. For particle i to pass particle k, j must either no longer be between them meaning j has already passed k or already been passed by i. Hence, $t_{ik} > \min(t_{ij}, t_{jk})$. Thus, either i passes j first or j passes k first. If i passes j first, then since i is faster than j, i will pass k before j does. Hence, if $t_{ij} < t_{ik}$, then $t_{ik} < t_{jk}$. By similar reasoning, it $t_{jk} < t_{ik}$, then $t_{ik} < t_{ij}$. Therefore,

$$\min(t_{ij}, t_{jk}) < t_{ik} < \max(t_{ij}, t_{jk}).$$

If $t_{ik} = \min(t_{ij}, t_{jk})$, then $t_{ik} = t_{ij}$ or $t_{ik} = t_{jk}$ meaning that i collides with j and k at the same time or k collides with i and j at the same time. In either case, all three particles collide simultaneously. □

Lemma 4.2.0.3. *Consider, for an interval $[a, b]$ with $1 \leq a < b \leq n$ with $[a, b] \neq [1, n]$, the face $F_{[a,b]}$ of $\Pi_{\mathbf{c}}(\Delta_n)$ for $\mathbf{c}$ such that $\mathbf{c}_1 > \mathbf{c}_2 > \cdots > \mathbf{c}_n$ induced by the vector $\mathbf{w}_i = -\mathbf{c}_i$ if $i \notin [a, b]$ and $-\mathbf{c}_b$ otherwise. Then the following hold:*

(a) A vertex $\mathbf{v}$ is in $F_{[a,b]}$ if and only if the arborescence of $\mathbf{v}$ satisfies $A(x) \in [a, b]$ if $x \in (a, b]$ and $A(x) \notin (a, b]$ otherwise.

(b) For each $[a, b]$, $F_{[a,b]}$ is a facet of $\Pi_{\mathbf{c}}(\Delta_n)$. Furthermore, each choice of $F_{[a,b]}$ is distinct.

(c) All facets are of the form $F_{[a,b]}$.

Proof. To prove (a), we compute the collisions that occur for this particle system explicitly. Then we unpack what this means in terms of the max-slope multi-arborescence. Note that at time $t = 0$, by construction, the coordinates of $\mathbf{w}$ are weakly increasing. Hence, since $\mathbf{c}$ is strictly decreasing, $\mathbf{w} + t\mathbf{c}$ has strictly increasing coordinates for all $t < 0$ meaning that the first collision that occurs is at time $t = 0$.

At time $t = 0$, $w_i = w_j$ for all $i, j \in [a, b]$ meaning that the particles in $[a, b]$ all collide simultaneously at time 0, so the set of particles that catch up to each element i of $(a, b]$ first is precisely $[a, i)$.

There are no other collisions outside of $[a, b]$, since the remaining initial positions are $-\mathbf{c}_i$, which is strictly increasing.

It remains to find which particles collide with those in $[1, a] \cup (b, n]$ first to describe the corresponding face of the polytope. Note that for all $t > 0$, $-\mathbf{c}_a + t\mathbf{c}_i < -\mathbf{c}_a + t\mathbf{c}_a$ for all $i \in (a, b]$, so particle a will catch up to any particle in $(b, n]$ before any particle in $(a, b]$. Thus, the first particle to catch up to a particle in $[1, a] \cup (b, n]$ must be a particle in $[1, a] \cup (b, n]$. At time $t = 1$, all particles in $[1, a] \cup (b, n]$ collide simultaneously, since, by construction, their positions will all be 0. Hence, the first collision for any $i \in [1, a] \cup (b, n]$ is always precisely $[1, i] \cap [1, a] \cup (b, n]$. Thus, the multiarborescence corresponding to the face F^w is given by $A(i) = [1, i] \cap [a, b]$ for $i \in (a, b]$ and $[1, i] \cap ([1, a] \cup (b, n])$ for $i \notin [a, b)$. The characterization of inclusion in F^w given by (a) then follows immediately.

To prove (b), it suffices to show that w is the unique choice of support vector up to non-negative rescaling and addition of a vector in the lineality space of the max-slope pivot rule polytope. Since the standard simplex has $\mathbf{1} = (1, 1, \ldots, 1)$ in its lineality space, the lineality space of the max-slope pivot rule polytope is $\operatorname{span}(\mathbf{1}, c)$.

Note that for a vector w to induce the arborescence as desired, there must be times $t_1 < t_2$ such that $w_i + t_1 c_i = w_j + t_1 c_j$ for all $i, j \in [a, b]$ and $w_k + t_2 c_k = w_\ell + t_2 c_\ell$ for all $k, \ell \in [a, b)^C$. Note that by adding a multiple of c to w and rescaling, we may without loss of generality assume that $t_1 = 0$ and $t_2 = 1$. Furthermore, by adding multiples of the all ones vector, we may assume that $w_i = -c_b$ for all $i \in [a, b]$. Then at time 1, $w_b + c_b = 0$. Thus, $w_i + c_i = 0$ for all $i \notin [a, b)$ meaning that $w_i = -c_i$ for all $i \notin [a, b)$. Hence, w is uniquely determined up to rescaling and addition of a vector from the lineality space of the normal cone it must lie in. Hence, F_w is a facet.

To prove (c), it suffices to show that any proper face of P is contained in $F_{[a,b]}$ for some choice of $[a, b]$. Let $\mathbf{w} \in \mathbb{R}^n$ with $\mathbf{w} \notin \operatorname{span}(\mathbf{c}, \mathbf{1})$, so $\mathbf{w}$ is maximized at a proper face of $\Pi_{\mathbf{c}}(\Delta_{[n]})$. Consider the first collision of particles of $\mathbf{w}$, and suppose without loss of generality that it occurs at time $t = 0$. Suppose during the first collision particles i and k collide. By Lemma 4.2.0.2, $t_{ik} \geq \min(t_{ij}, t_{jk})$ for all j such that $i < j < k$. Since t_{ik} is the first collision, $t_{ik} \leq \min(t_{ij}, t_{jk}) \leq t_{ik}$. Hence, $t_{ik} = \min(t_{ij}, t_{jk})$, so by Lemma 4.2.0.2, all particles in $[i, k]$ collide simultaneously at time $t = 0$. Assume that $[i, k]$ is a maximal interval with respect to inclusion of particles that all collide at time $t = 0$. Then for all $\ell \in (i, k]$, $\mathcal{A}(\ell) = [i, \ell)$. For all $j \notin [i, k]$, if one of the first particles to catch up to

j is in $[i, k]$ it must be i, since after time $t = 0$, i is ahead of all of other particles in $[i, k]$. Hence, in the multi-arborescence corresponding to $\mathbf{w}$, $\mathbf{A}(\ell) \subseteq [i, k]$ for all $\ell \in (i, k]$ and $\mathcal{A}(\ell) \subseteq [1, i] \cup [k+1, n]$ otherwise meaning that $\mathbf{A}$ is a multi-arborescence contained in $F_{[i,k]}$. Hence, any face of P must be contained in a face of the form $F_{[a,b]}$ meaning that the set of $F_{[a,b]}$ is precisely the set of faces of $\Pi_{\mathbf{c}}(P)$. $\square$

Now that we understand facets and vertex-facet incidences, to describe the face lattice, it suffices to characterize the vertices.

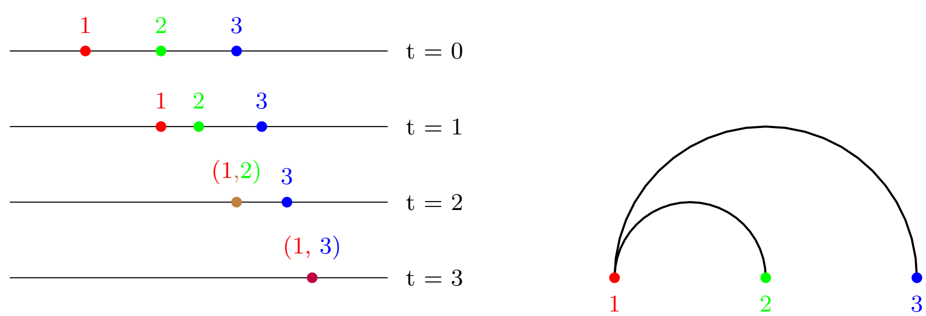

FIGURE 4.2. On the left is an illustration of the collision scheme for $\mathbf{w} = (1, 2, 3)$ and $\mathbf{c} = (1, 1/2, 1/3)$. To the right is an illustration of the corresponding non-crossing arborescence.

.

LEMMA 4.2.0.4. *An arborescence A corresponds to a vertex of the max-slope pivot rule polytope of a simplex if and only if whenever $A(k) = i$ and $i < j < k$, we have $A(j) \geq i$. We call such an arborescence* ***non-crossing****.*

PROOF. Suppose first that $\mathcal{A}$ is an arborescence corresponding to a vertex of the max-slope pivot rule polytope. Suppose for the sake of contradiction that there exists $i < j < k < \ell$ for $i, j, k, \ell \in [n]$ such that $\mathcal{A}(k) = i$ and $\mathcal{A}(\ell) = j$. By Lemma 4.2.0.2,

$$\min(t_{ij}, t_{jk}) < t_{ik} < \max(t_{ij}, t_{ik}).$$

Since i collides with k first among all particles less than k, we must then have,

$$t_{ij} < t_{ik} < t_{jk}.$$

By similar reasoning,

$$t_{jk} < t_{j\ell} < t_{k\ell}.$$

Then $t_{ij} < t_{j\ell}$. But then, since i is faster than j, i has to pass ℓ before j does. In particular, by Lemma 4.2.0.2, $t_{ij} < t_{i\ell} < t_{j\ell}$, a contradiction to our assumption that j passes ℓ before any other particle does.

For the other direction, proceed by induction. Certainly all arborescences of the 1-simplex are non-crossing and are realized by some weights. Thus, the base case holds. Now suppose it is true for all $1 \leq k < n$. Then consider a fixed non-crossing arborescence A. Let $i^* = \max(A^{-1}(1))$. Then, by the non-crossing condition, the interval $[0, i^*)$ is mapped to $[0, i^*]$ and the interval $[i^* + 1, n)$ is mapped to $[i^* + 1, n]$. Then we may realize these arborscences by vectors $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{i^*})$ and $(\mathbf{w}_{i^*+1}, \ldots, \mathbf{w}_n)$ respectively. Then to glue the arborescences together we consider the vector

$$(\mathbf{w}_1 - \lambda, \mathbf{w}_2 - \lambda, \ldots, \mathbf{w}_{i^*} - \lambda, \mathbf{w}_{i^*+1}, \ldots, \mathbf{w}_n)$$

for λ strictly larger than the time it takes for $\mathbf{w}_1$ to catch up to $\mathbf{w}_{i^*}$ and the time it takes for $\mathbf{w}_{i^*}$ to catch up to $\mathbf{w}_n$. Then all collisions between between particles within $[0, i^*]$ and within $[i^*+1, n]$ will occur before i^* collides with any particle ahead of it. Then i^* collides with n to finish. Therefore, by induction, we are done. $\square$

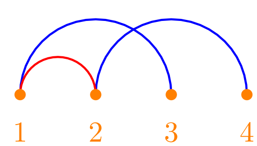

FIGURE 4.3. Pictured is an example of an arborescence together with a crossing highlighted in blue. In particular, it is not a max-slope arborescence of the simplex.

Now given a vertex, we may be more explicit about how to compute the set of facets containing it. We say an interval $[a, b]$ **coarsens** an arborescence $\mathcal{A}$ if $\mathcal{A}(i) \in [a, b]$ for all $i \in (a, b]$ and $\mathcal{A}(j) \in [n] \setminus (a, b]$ for all $j \in [n] \setminus (a, b]$. By Lemma 4.2.0.3, an interval coarsens a non-crossing arborescence if and only if the corresponding facet contains it. This notion of coarsening gives rise to subarborescences.

LEMMA 4.2.0.5. *Let $\mathcal{A}$ be a non-crossing arborescence on $[1, n]$. Then if $[a, b]$ coarsens $\mathcal{A}$, then $\mathcal{A}$ restricted to $[a, b]$ is still a non-crossing arborescence on $[a, b]$. Furthermore, each interval coarsening $\mathcal{A}|_{[a,b]}$ coarsens $\mathcal{A}$.*

PROOF. Let $j \in (a, b]$. Then $\mathcal{A}(j) \in [a, b]$ by definition of coarsening. Hence, $\mathcal{A}|_{[a,b]}$ is still an arborescence. Furthermore, $\mathcal{A}|_{[a,b]}$ must still be non-crossing, since any crossing on $\mathcal{A}|_{[a,b]}$ must be a crossing on $\mathcal{A}$.

Let $[c, d]$ be interval coarsening the restricted arborescence $\mathcal{A}|_{[a,b]}$. Then for all $i \in (c, d]$, $\mathcal{A}(i) \in [c, d]$ and for all $j \in [a, b] \setminus (c, d]$, $\mathcal{A}(j) \in [a, b] \setminus (c, d]$. Then to show that $[c, d]$ also coarsens $\mathcal{A}$, it suffices to show that $\mathcal{A}(k) \notin (c, d]$ for all $k > b$. Let $k > b$. Since $[a, b]$ coarsens $\mathcal{A}$, $\mathcal{A}(k) \notin (a, b] \supseteq (c, d]$. Hence, each interval coarsening $[a, b]$ also coarsens $\mathcal{A}$. $\square$

Now we may give a more explicit explanation of precisely which facets contain a given vertex.

LEMMA 4.2.0.6. *Let $\mathcal{A}$ be a non-crossing arborescence. Let $i^* = \max(\mathcal{A}^{-1}(1))$. Then the set of intervals that coarsen $\mathcal{A}$ are precisely $[1, i^* - 1]$ if $i^* > 2$, $[i^*, n]$ if $i^* < n$, and all intervals that coarsen the restricted arborescences $\mathcal{A}|_{[1,i^*-1]}$ and $\mathcal{A}|_{[i^*,n]}$.*

PROOF. Let $j \in (i^*, n]$. Then $\mathcal{A}(j) \neq 1$, since $i^* = \max(\mathcal{A}^{-1}(1))$. Furthermore, since $\mathcal{A}(1) = i^*$, by Lemma 4.2.0.4, $\mathcal{A}(j) \geq i^*$. Hence, $\mathcal{A}((i^*, n]) \subseteq [i^*, n]$. Furthermore, since $1 = \mathcal{A}(i^*)$ $\mathcal{A}([i^*, n]) = \{1\} \cup [i^*, n]$. Hence, $\mathcal{A}$ is contained in the facets $[1, i^* - 1]$ if $i^* > 2$ or $[i^*, n]$ if $i^* < n$. Then, by Lemma 4.2.0.5, $\mathcal{A}$ is also coarsened by all the arborescences that coarsen $\mathcal{A}|_{[1,i^*-1]}$ and $\mathcal{A}|_{[i^*,n]}$. It suffices to show that any other interval that coarsen $\mathcal{A}$ must be in this set.

Let $[a, b]$ be interval that coarsens $\mathcal{A}$. Then are three cases to consider: $[a, b] \subseteq [1, i^* - 1]$, $[a, b] \subseteq [i^*, n]$, or $a \in [1, i^* - 1]$ and $b \in [i^*, n]$.

Suppose that $[a, b]$ is properly contained in $[1, i^* - 1]$. Then $\mathcal{A}((a, b]) \subseteq [a, b]$ and

$$\mathcal{A}([1, i^* - 1] \setminus (a, b]) \subseteq [1, i^* - 1] \cap \mathcal{A}([n] \setminus (a, b]) \subseteq [1, i^*] \cap [n] \setminus (a, b]) = [1, i^* - 1] \setminus (a, b].$$

Hence, $[a, b]$ also coarsens $\mathcal{A}|_{[1,i^*-1]}$.

Suppose instead that $[a, b]$ is properly contained in $[i^*, n]$. Then, since $[a, b]$ coarsens $\mathcal{A}$, $\mathcal{A}((a, b]) \subseteq [a, b]$. Similarly, $\mathcal{A}([n] \setminus (a, b]) \subseteq [1, a] \cup [b + 1, n]$. Since $\mathcal{A}(1) = i^*$ and $a \geq i^*$, by Lemma 4.2.0.4, $\mathcal{A}((i^*, n] \setminus (a, b])) \subseteq [i^*, n]$. Hence, $[a, b]$ also coarsens $[i^*, n]$.

Finally, let $a \in [1, i^* - 1]$ and $b \in [i^*, n]$. Suppose that $[a, b]$ coarsens $\mathcal{A}$. Then $\mathcal{A}((a, b]) \subseteq [a, b]$. Then, in particular, there exists $k \in [n]$ such that $b, \mathcal{A}(b), \mathcal{A}^2(b), \ldots, \mathcal{A}^k(b) = a$. Since $a \in [1, i^* - 1]$ and $b \in [i^*, n]$, there exists some $i < k$ such $\mathcal{A}^{i-1}(b) \in [i^*, n]$ and $\mathcal{A}^i(b) \in [1, i^* - 1]$. Since $\mathcal{A}([i^*, n]) \subseteq \{1\} \cup [i^*, n]$, $\mathcal{A}^i(b) = 1$. Hence, $a = 1$.

Since $[a, b]$ is a proper interval, $b < n$, so in particular, $n \notin [a, b]$. Then consider the sequence $n, \mathcal{A}(n), \mathcal{A}^2(n), \ldots, \mathcal{A}^\ell(n) = 1$. Note that, by the non-crossing condition in Lemma 4.2.0.4, $\mathcal{A}^{\ell-1}(n) = i^* \in (a, b]$. Hence, there there exists some $j \in [\ell]$ such that $\mathcal{A}^{j-1}(\ell) \notin (a, b]$ and $\mathcal{A}^j(\ell) \in (a, b]$. But then $[a, b]$ does not coarsen $\mathcal{A}$, a contraditionc. Hence, any such choice of $[a, b]$ does not coarsen $\mathcal{A}$. Therefore, the set of $[a, b]$ that coarsen $\mathcal{A}$ are of the desired type.

□

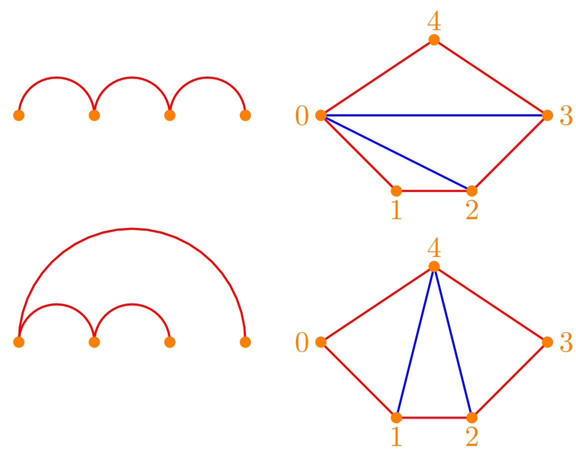

FIGURE 4.4. Pictured are two non-crossing arborescences together for the tetrahedron with the corresponding triangulations of an pentagon.

From what we have shown thus far, we have a complete description of the face lattice of the pivot rule polytope. Our goal is to compare to the face lattice of the associahedron. There are many different interpretations of the associahedron face lattice. Here we rely on the interpretation in terms of triangulations of an n-gon as one may find in Chapter 0 of [138] for example. Namely, the vertices of the associahedron n dimensional associahedron correspond to triangulations of an $(n+2)$-gon using only diagonals for additional edges. Facets correspond to diagonals of the $(n + 2)$-gon, and a vertex is contained in a facet if the corresponding triangulation contains the corresponding diagonal as one of its additional edges.

THEOREM 4.2.1. *The max-slope pivot rule polytope of a simplex $\Pi_{\mathbf{c}}(\Delta^n)$ is combinatorially equivalent to the associahedron for triangulations of an $n+1$-gon.*

PROOF. To prove combinatorial equivalence, it suffices to define bijections from vertices to vertices and facets to facets that is inclusion preserving as noted by in Chapter 2 of [138]. We consider the $(n+1)$-gon with vertices labeled $(0, 1, \ldots, n)$ and choose the associahedron corresponding to all subdivisions of this $(n+1)$-gon.

Facets of the associahedron correspond to diagonals in the polygon. In particular, they correspond to pairs (i, j) such that $\min(i - j \mod (n+1), j - i \mod (n+1)) \geq 2$. By Lemma 4.2.0.3, facets of $\Pi_{\mathbf{c}}(\Delta^n)$ correspond exactly to intervals $[a, b]$ such that $1 \leq a < b \leq n$ and $[a, b] \neq [1, n]$. Note that these two sets of facets have the same cardinality, since the number of diagonals of an $(n+1)$-gon is $\binom{n+1}{2} - (n+1) = \binom{n}{2} - 1$, the number of intervals other than $[1, n]$.

For our bijection, the facet corresponding to the interval $F_{[a,b]}$ is mapped to the edge $(a, b+1 \mod (n+1))$. For $b < n$, $b+1 \mod (n+1) = b+1$. Since $b > a$, $b+1-a \geq 2$. Otherwise, for $b < n$,

$$a - (b+1) \mod (n+1) \geq 1 - (b+1) \mod (n+1) \geq 2.$$

For $b = n$, the edge is $(a, 0)$. For $a - 0 < 2$, since $a \geq 1$, $a = 1$ and $b = n$, but the interval $[1, n]$ is not in the domain. Hence, the map is well-defined. Furthermore, it is injective with inverse given an edge (a, b) with $a < b$, the pre-image is to $(a, b-1)$ if $a \neq 0$ and (b, n) otherwise. Since the two sets of facets, have the same cardinality, this map is therefore a bijection.

The theorem clearly holds if $n = 2$ and $n = 3$. Given a non-crossing arborescence $\mathcal{A}$, we construct a triangulation of the $(n+1)$-gon inductively as follows. Suppose there is a bijection for $1 \leq k < n$. Let $i^* = \max(\mathcal{A}^{-1}(1))$. Then add the edges $(0, i^*)$ and $(1, i^*)$ if they do not already appear in the polygon. By Lemma 4.2.0.5, the $\mathcal{A}$ restricted to $[1, i^* - 1]$ and to $[i^*, n]$ yields non-crossing arborescences. By induction, the pair of arborescences on $[1, i^* - 1]$ and $[i^*, n]$ are in bijection to a pair of triangulations of an i^*-gon and an $n - i^* + 1$-gon respectively. The $i^* - 1$-gon has vertices $[1, i^*]$ and the $n - i^* + 1$-gon $\{0\} \cup [i^*, n]$. Hence, there is a well-defined map. From the triangulation, one can find the unique triangle containing the edge $(0, 1)$ to find the $\max(\mathcal{A}^{-1}(1))$ and then reconstruct the remaining arborescences by induction. Hence, the map is injective. To show surjectivity, pick any triangulation of an n-gon and find the unique third vertex i^* of the triangle containing $(0, 1)$. Then define $\mathcal{A}(i^*) = 1$ and compute arborescences of $[1, i^* - 1]$ and $[i^*, n]$ by induction. Such an arborescence is non-crossing. Hence, the vertices are in bijection.

Finally, by Lemma 4.2.0.6, every interval that coarsens an arborescence $\mathcal{A}$ is of the form $[1, i^* - 1]$, $[i^*, n]$, or coarsens a restricted arborescence of $\mathcal{A}|_{[1,i^*-1]}$ and $\mathcal{A}_{[i^*,n]}$. Since the triangulation contains the triangle $(0, 1, i^*)$, $\mathcal{A}$ is contained in the facets corresponding to $(0, i^*)$ and $(1, i^*)$ so long as those edges correspond to diagonals of the $(n+1)$-gon. The remaining facets correspond to the edges in the triangulations of $[1, i^*]$ and $\{0\} \cup [i^*, n]$. By the inductive construction, these are precisely the edges in the triangulation mapped to by $\mathcal{A}|_{[1,i^*-1]}$ and $\mathcal{A}|_{[i^*,n]}$ respectively. Furthermore, by induction, we may assume that the diagonals in each triangulation correspond exactly to the intervals that coarsen each arborescence. Hence, the bijections on facets and vertices preserves vertex-facet incidences and is therefore a lattice isomorphism.

□

Note that this bijection is similar to the bijection for the realization of the associahedron using nestohedra in [117]. In Chapter 5, we will discuss some prospective examples, where computing pivot rule polytopes would be both interesting and feasible.

CHAPTER 5

Future Directions

In the process of working on this thesis, I have found many problems in the study of monotone paths both for the sake of better understanding the simplex method and combinatorial algorithms for linear programming as well as for uncovering rich combinatorial structure. In this section, I will mention some of the many open problems that I would like to see a solution. The biggest general problem of interest is the following:

Open Problem 5.0.1 ([126]). *Is there a strongly polynomial time algorithm for linear programming?*

This problem is affectionately called the **Holy Grail** by the linear programming community and is Smale's 9th problem for the 21st century. A related and equally important problem going back to Dantzig is the following:

Open Problem 5.0.2 ([49]). *Is there a polynomial time version of the simplex method?*

Finally, the combinatorial version of this problem is the following:

Open Problem 5.0.3 (Polynomial Hirsch Conjecture). *Does there exist a polynomial bound on the diameters of polytopes in terms of m and n, where m is the number of facets of the polytope and n is its dimension?*

These three problems are the central motivating problems of this thesis, and any partial progress on these problems would be of great interest to the linear programming and polytopes community. From the combinatorics side, I am generally interested in understanding structural results concerning monotone path polytopes. The following result is of interest for example:

Open Problem 5.0.4. *What is the maximal number of vertices of a monotone path polytope of a polytope with n vertices or m facets?*

Such a result would yield an estimate on the maximal number of distinct shadows of a fixed polytope. Out of doing so, as noticed here, one should find rich connections to many other areas of combinatorics as well as insights into the behavior of the simplex method.

5.1. Pivot Rules for the Simplex Method

Even if the polynomial Hirsch conjecture is proven, finding a polynomial time algorithm for the simplex may still be out of reach. Namely, finding short paths on polytopes is NP-Hard even to approximate within a constant factor [36, 37, 54, 121]. From a positive perspective, Kaibel and Kukharenko showed that if one can find a pivot rule for the simplex method guaranteed to approximate a shortest path within a polynomial factor of the diameter, then one can solve linear programs in polynomial time [86].

Open Problem 5.1.1. *Is it NP-Hard to compute a polynomial length path between a pair of vertices of a polytope?*

A particularly interesting case of this question is for combinatorial cubes (i.e., polytopes combinatorially equivalent to a hyper-cube). It is open whether there is a strongly polynomial time algorithm for combinatorial cubes. In fact, they are the key tool for finding lower bounds for many known pivot rules for the simplex method [3,95] and for interior point methods [1,2]. However, the monotone diameter of a combinatorial cube is at most its dimension with a bound always realized by some version of Bland's pivot rule [28].

Open Problem 5.1.2. *Is there a strongly polynomial time simplex method for linear programming on combinatorial cubes? In fact, is it NP-Hard to find a path of length at most d between a pair of vertices of a combinatorial cube when given as a system of linear inequalities?*

Such a result is related to work on representing general linear programs in terms of unique sink orientations on graphs of hyper-cubes [73, 131]. While a pivot rule may be defined as any way to choose an entering variable in polynomial time. However, this leaves room for one to use a linear programming solver as part of our pivot rule. A hope would be that a pivot rule could avoid using a linear programming solver as a subroutine and should be a simple rule easily computed in terms of a current vertex.

Open Problem 5.1.3. *For any continuous function $\eta : \mathbb{R}^n \to \mathbb{R}$ such that $\eta(\mathbf{x}) = \mathbf{0}$ if and only if $\mathbf{x} = \mathbf{0}$, and $\mathbf{w} \in \mathbb{R}^n$, define the $(\mathbf{w}, \eta)$-pivot rule by*

$$\underset{\mathbf{u} \in N_{\mathbf{c}(\mathbf{v})}}{\operatorname{argmax}} \left\{ \frac{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{v})}{\eta(\mathbf{u} - \mathbf{v})} \right\},$$

where $N_{\mathbf{c}}(\mathbf{v})$ is the set of $\mathbf{c}$-improving neighbors of $\mathbf{v}$. Is there some choice of $\mathbf{w}$ and η that guarantees a polynomial run-time for every LP?

This essentially asks whether we need to choose the normalization and weight in terms of the linear program. For shadow rules, the method of choosing a shadow makes a difference in the run-time. In fact, Murty showed in [110] that some polytopes have exponentially large shadows. By applying an affine transformation, one may make any particular choice of shadow is chosen with arbitrarily high probability. However, affine transformations do not change the space of all shadows of a polytope.

Open Problem 5.1.4. *Does every polytope have a polynomial size shadow? Equivalently, for any polytope, does there exist an affine transformation for which a randomly chosen is of expected polynomial size?*

Other rules such as the steepest edge rule are also not preserved under affine transformation. While there are exponential examples known for fixed polytopes such as the Goldfarb-Sit cube from [76], no such exponential example is know for which one can not fix the behavior with a suitable affine transformation.

Open Problem 5.1.5. *For any linear program, is there an affine transformation of the underlying polytope for which the steepest edge rule will perform in strongly polynomial time?*

A related question to this one would be to describe the space of possible pivot rules that arise from applying an affine transformation of the steepest edge rule. Then one could construct an example that fails for any pivot rule in that class instead of describing all affine transformations.

It would furthermore be useful to have a similar tool for the monotone path polytope for every pivot rule. The pivot rule polytope plays that role to a degree. However, Proposition 4.1.1 argues that to any monotone path chosen by a normalized weight pivot rule, there is a cone of weights that will choose the same path.

Open Problem 5.1.6. *Does the set of cones from Proposition 4.1.1 always form the normal fan of a polytope?*

5.2. Monotone Diameters of Polytopes

The Hirsch conjecture that the worst-case diameter of a polytope with m facets in n dimensions is $m-n$ stood for 50 years until being disproven by Santos in 2010 in [122]. The same conjecture for monotone diameters lasted much less time. Namely, in 1980, Todd showed the following:

Theorem 5.2.1 ([135]). *The worst-case monotone diameter of a polytope with m facets in n dimensions is at least $1.25(m-n)$.*

He did this by exhibiting an example of a polytope in 4 dimensions with 8 facets but monotone diameter 5. Now in the 44 years since then, there has been no improvement on the lower bounds due to Todd. One fundamental open problem I am interested in is improving the lower bounds due to Todd:

Open Problem 5.2.2. *Improve on Todd's lower bounds for the worst-case monotone diameters of polytopes with m facets in n dimensions.*

Diameters of polytopes are also studied in terms of different parameters. For example, one studies diameters of polytopes of the form $\{\mathbf{x} : A\mathbf{x} \le \mathbf{b}\}$ with $A \in \mathbb{Z}^{m\times n}$ and the maximum absolute value of any subdeterminant of A is bounded by Δ. In this context, the best known diameter bounds are in [47] of the form $O(n^3 \Delta \ln(n\Delta))$. However, for monotone diameters, no similar result is known:

Open Problem 5.2.3. *Can one find a polynomial bound poly(n,Δ) on the monotone diameters of n-dimensional polytopes of the form $\{\mathbf{x} : A\mathbf{x} \le \mathbf{b}\}$ with A integral and such that the maximal subdeterminant is bounded by Δ?*

Note this problem is even open for the case in which $\Delta = 1$. Similarly, for $(0,k)$-lattice polytopes (i.e., polytopes with vertices contained in $\{0,k\}^n$), their diameters are known to be $O(dk)$ due to Kleinschmidt and Onn [96]. However, no similar result is known for monotone diameters:

Open Problem 5.2.4. *Can one bound the monotone diameters of d-dimensional $(0,k)$ lattice polytopes by a polynomial in d and k?*

One distinction between diameters of polytopes and monotone diameters is that knowing the graph of the polytope is insufficient to compute the monotone diameter, since one must take into account the different possible orientations of the graph. In particular, computing the diameter of a graph is in P. However, the following remains open:

OPEN PROBLEM 5.2.5. *Is it NP-Hard to compute the monotone diameter of a polytope in terms of its number of vertices and facets?*

In [121], Sanità showed that computing the diameter of a polytope is NP-Hard in the special case of fractional matching polytopes with the input being an inequality description of the polytope. In very recent follow up work, Nöbel and Steiner extended this result to monotone diameters [113], but their input is an inequality description. In contrast, Open Problem 5.2.5 suggests the monotone diameter is NP-Hard to compute even when the graph of the polytope is known. One challenge for dealing with this problem is that there are few families of polytopes for which the monotone diameter has been explicitly computed.

OPEN PROBLEM 5.2.6. *Compute monotone diameters of more examples of polytopes.*

5.3. Monotone Paths on Generalized Permutahedra

In [16, 18, 19], there is a discussion of how we know the monotone path polytope of a simplex is a cube, the monotone path polytope of a cube is a permutahedron, but we have no purely combinatorial description for coherent monotone paths on the permutahedron. For generic $\mathbf{c} \in \mathbb{R}^n$, the digraph of the permutahedron induced by $\mathbf{c}$ is precisely the Hasse diagram for the weak Bruhat order on the permutation group S_n, the partial order on permutations with cover relations given by $\sigma \leq \sigma'$ if $\sigma' = \sigma(i, i+1)$ and the number of inversions of σ' is greater than the number of inversion of σ. The Edelman-Green Bijection then says the following:

PROPOSITION 5.3.1 ([63]). *For generic* $\mathbf{c}$*, the set of monotone paths on the permutahedron* Π_n *are in bijection with Standard Young Tableaux of staircase shape* $(n-1, \dots, 1)$.

These monotone paths appeared in the works of Goodman and Pollack under the name **allowable sequences** [77, 78] of permutations, which are precisely maximal chains in the weak Bruhat order. There they introduce the notion of **stretchable allowable sequences**, which are sequences of permutations that arise via the following geometric method. Namely, consider a point cloud of n points in $\mathbb{R}^2$. Then the set of all orderings of those points induced by a vector $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2) \in \mathbb{R}^2$ such that $\mathbf{c}_2 \geq 0$ is an allowable sequence with the identity permutation for the orientation $(1, 0)$.

Proposition 5.3.2. *A monotone path on the permutahedron is coherent if and only if in the corresponding allowable sequence is stretchable. In particular, a* $\mathbf{w}$*-coherent* $\mathbf{c}$*-monotone path corresponds exactly to the allowable sequence realized by the set of points* $(\mathbf{w}_1, \mathbf{c}_1), \ldots, (\mathbf{w}_n, \mathbf{c}_n)$*.*

The connection between allowable sequences and coherent monotone paths was made originally in [**115**]. Goodman and Pollack did not provide a combinatorial characterization of when an allowable sequence is stretchable.

Open Problem 5.3.1. *Find a combinatorial description of the coherent monotone paths on the permutahedron.*

In particular, the permutahedron is a zonotope, so coherent monotone paths correspond to total orderings of the edge directions of the zonotope. A total ordering corresponds to a coherent path if and only if it is realized by taking the slopes of the edges [**65**]. Then testing whether a monotone path on the permutahedron is coherent corresponds to a feasibility problem for a cone defined by $\binom{n}{2}$ inequalities. A combinatorial description of the coherent paths could correspond to an efficient algorithm for solving such a linear program.

The connection appears also in the context of probabilistic combinatorics in connection to statistical mechanics, where monotone paths on permutahedra have yet another name: **sorting networks** [**5**, **6**, **51**, **52**]. There they proved a limit shape theorem that, in more precise terms, says that random monotone paths on the permutahedron behave like coherent monotone paths. I pose the following open problem:

Open Problem 5.3.2. *Generalize the results on random sorting networks to monotone paths on a broader class of polytopes.*

It would in particular be interesting to do this for generalized permutahedra, since each monotone path on a generalized permutahedron extends to a monotone path on the permutahedron. The case of the associahedron is of particular interest. In that case, one could take an orientation as in Example 5.5 of [**81**] for which the directed graph of the associahedron is the well studied Tamari lattice [133]. Is there a limit shape theorem one can state for maximal chains in the Tamari lattice that is compatible with the limit shape theorem for maximal chains in the weak Bruhat order?

For the permutahedron, we can also provide a general descriptions of what exactly coherence means in terms of a generalization of the particle model we introduce for pivot rule polytopes. Namely, consider a choice of n particles all moving on a line at speeds $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$ starting at initial point $\mathbf{w} = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$. Suppose that $\mathbf{c}_1 > \mathbf{c}_2 > \cdots > \mathbf{c}_n$. Then one can consider the positions of the particles $\mathbf{w}_i + \lambda \mathbf{c}_i$ overtime. For the permutahedron, a $\mathbf{w}$-coherent $\mathbf{c}$-monotone path corresponds to recording the set of all total orderings of those particles. This is because the path corresponds to the set of cones in the normal fan of the permutahedron intersected by the line $\mathbf{w} + \lambda \mathbf{c}$. The slopes record the times in which particle i passes particle j.

Recall that a **generalized permutahedron** is any weak Minkowski summand of a permutahedron. In particular, it is any polytope whose normal fan is refined by the normal fan of the permutahedron or equivalently, for which a total ordering on the coordinates of $\mathbf{c}$ will always a determine a unique $\mathbf{c}$-maximal vertex. Then the set of vertices of a generalized permutahedron correspond to collections of permutations. In the particle model, then a coherent monotone path corresponds to seeing how particles move on a line and recording some information about the orderings of the particles. On the simplex, for example, we may think of the particles as racers on a track. Then the monotone path records the set of particles which are in first place at some point in the race. For a hyper-simplex $\Delta(n, k)$, one instead records which particles are in the top k places. For matroid base polytopes, it records which basis is maximal. Thus, the monotone path polytopes of generalized permutahedra have a fundamental representation in terms of evolving particle systems. In fact, this is true for pivot rule polytopes of generalized permutahedra as well and the key tool we leveraged to compute the pivot rule polytope of a simplex.

Open Problem 5.3.3. *Compute more monotone path polytopes and max-slope pivot rule polytopes of generalized permutahedra. Find general expressions for the vertices of these polytopes in terms of sorting networks.*

A particularly interesting orientation for permutahedra is the degenerate orientation $e_{[k]} = \sum_{i=1}^{k} e_i \in \mathbb{R}^n$ for $k < n$. Note that

$$e_{[k]}^{\intercal} e_i - e_j = \begin{cases} 1 \text{ if } i \in [k], j \notin [k] \\ 0 \text{ if } i, j \in [k] \\ -1 \text{ if } i \notin [k], j \in [k] \end{cases}.$$

In particular, every $e_{[k]}$-monotone path is of the same length, since $e_{[k]}^{\intercal}(\mathbf{u}-\mathbf{v}) = 1$ for all improving neighbors. Then monotone paths correspond exactly to total orderings $e_a - e_b$ with $a \in [k]$ and $b \notin [k]$ induced by a generic choice of $\mathbf{w} \in \mathbb{R}^n$. Hence, we have as shown in joint work with Raman Sanyal:

Theorem 5.3.4 ([27]). *The monotone path polytope of the permutahedron $\Sigma_{e_{[k]}}(\Pi_n)$ has normal fan given by the hyperplane arrangement with normals*

$$\{e_a + e_b - e_c - e_d : a, c \in [k], b, d \in [\ell]\}.$$

This arrangement has appeared in many other settings including quantum computing [38, 97] deconvolution algorithms [105], and sorting [69]. Understanding the extreme rays of this arrangement is of particular interest in physics and corresponds precisely to understand the facets of the monotone path polytope of the permutahedron for this orientation. Optimization on this polytope corresponds to sorting a list of numbers $X + Y = \{x + y : x \in X, y \in Y\}$. Whether one can sort a list of that form using comparison sort in time $O(n^2)$ for two lists of length of n is a nearly 50 year old open problem going back to Fredman's work proving linear lower bounds [69]. Better understanding the behavior of coherent monotone paths on the permutahedron and on generalized permutahedra for this orientation could yield insight into this problem.

Open Problem 5.3.5. *Describe monotone path polytopes for generalized permutahedra for the orientation $e_{[k]}$.*

Matroid base polytopes would be an interesting starting point for this problem.

5.4. Generalities Surrounding Monotone Path Polytopes

Monotone path polytopes are generally easier to compute if one knows that all monotone paths are coherent, for then understanding monotone paths reduces to understanding the oriented graph of the polytope. Likewise, for the simplex method, one may treat every monotone path as a choice of shadow, so one could lower bound the length of a shortest monotone path by lower bounding the length of a smallest shadow.

Open Problem 5.4.1. *Characterize when all monotone paths are coherent and bound diameters of monotone path polytopes. Is the polynomial Hirsch conjecture true under that assumption?*

Computing monotone path polytopes is quite challenging. In general, for a polytope with $|V(P)|$ vertices, it requires computing the Minkowski sum of $|V(P)|$ slices of the polytope. A tool to potentially compute new monotone path polytopes comes from their interpretation in the context of toric geometry [88]. Namely, the monotone path polytope may be realized as a quotient of a toric variety by a one dimensional subtorus. One way in which such a quotient could arise would be if one took two varieties V_1 and V_2 each equipped with a torus action and defined a rational map between them via $\alpha : V_1 \to V_2$ that is torus equivariant. Then for a toric variety $T \subseteq V_1$, $\alpha(T)$ is a toric variety contained in V_2. Then $\alpha(T)$ may be realized as a toric variety of T. To my knowledge, this observation has not been used to compute new monotone path polytopes.

Open Problem 5.4.2. *Compute new monotone path polytopes or more generally fiber polytopes using toric quotients.*

While any polytope can arise as the monotone path polytope of a polytope if we allow for degenerate orientations. This is no longer the case if one only allows for generic orientations $\mathbf{c}$.

Open Problem 5.4.3. *Describe the space of all polytopes that arise as monotone path polytopes for generic choices of linear functions.*

Other generalities about monotone path polytopes are unclear. For example, the monotone path polytopes of simplices, cubes, and products of simplices are all simple. It is unclear when this condition should occur but could be of interest to use the monotone path polytope as a construction tool.

Open Problem 5.4.4. *Find nontrivial sufficient conditions that guarantee a monotone path polytope will be simple.*

In this final section, we covered a sample of questions that could follow up on our results. There are many other directions. The simplex method is a remarkable algorithm of deep practical importance in optimization, and much remains to be understood to have a complete theory of its behavior. So far our current theory of the simplex method is unable to provide a robust explanation of what we see in practice. Furthermore, the underlying structure of the simplex method gives a new perspective on combinatorial objects open to tools from optimization.

APPENDIX A

Background on Monotone Path Polytopes

This appendix is meant to be a self contained introduction to monotone paths on polytopes and monotone path polytopes that relies on polyhedral theory at the level of Chapters $0, 1, 2$, and 7 of [138]. In pursuit of this goal, we will also recover some of what is written in the introduction. While there are many new theorems, we spend time and space on slowly recovering some well known results for the sake of illustration.

To start, let $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ be a polytope, and consider the linear program $\max(\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P)$ for some $\mathbf{c} \in \mathbb{R}^n$. We say that a vertex $\mathbf{v}$ of P is a $\mathbf{c}$-minimum or $\mathbf{c}$-maximum if $\mathbf{v}$ minimizes or maximizes $\mathbf{c}^\intercal \mathbf{x}$ respectively.

Definition A.0.1. *A $\mathbf{c}$-**monotone path** for $\mathbf{c} \in \mathbb{R}^n$ is a sequence of vertices $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_k$ of a polytope $P \subset \mathbb{R}^n$ satisfying*

- *For all $i \in [k]$, $\mathbf{v}_{i-1}$ is adjacent to $\mathbf{v}_i$,*
- *For all $i \in [k]$, $\mathbf{c}^\intercal \mathbf{v}_{i-1} < \mathbf{c}^\intercal \mathbf{v}_i$,*
- *$\mathbf{v}_0$ is $\mathbf{c}$-minimal,*
- *$\mathbf{v}_k$ is $\mathbf{c}$-maximal.*

To illustrate some of these properties, we will look at the example of a simplex. Namely, consider the standard simplex $\operatorname{conv}(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ and the linear function $\mathbf{c} = (1, 2, \ldots, n)$. Note that the graph of the simplex is the complete graph. Then the $\mathbf{c}$-monotone paths are precisely sequences of basis vectors such that $e_1, e_{i_0}, e_{i_1}, \ldots, e_{i_k}, e_n$, such that $1 < i_{j-1} < i_j < n$ for all $j \in [k]$. In particular, they are in bijection with subsets of $[n-2]$. See Figure A.1 for a visual depiction of all monotone paths on the tetrahedron.

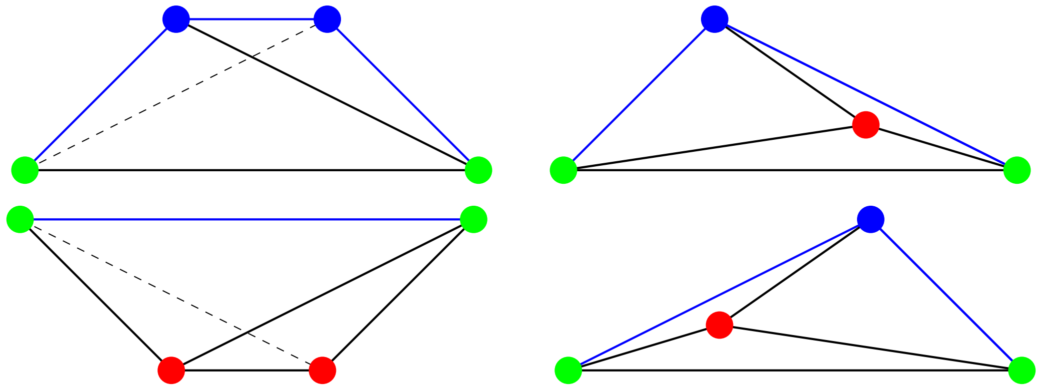

FIGURE A.1. The four monotone paths on a tetrahedron $\mathrm{conv}(e_1, e_2, e_3, e_4)$ are depicted for $\mathbf{c} = (0, 1, 2, 3)$. The coherence of the top left and bottom left are shown using $\mathbf{w} = (0, 1, 1, 0)$ and $\mathbf{w} = (0, -1, -1, 0)$ respectively. The top right and bottom right are shown to be coherent using $\mathbf{w} = (0, 1, .3, 0)$ and $\mathbf{w} = (0, .3, 1, 0)$ respectively.

A.1. Coherent Monotone Paths

We will focus on studying a special subset of monotone paths called **coherent monotone paths**, which correspond to the vertices of the **monotone path polytope**. There are multiple equivalent definitions of coherent monotone paths. We use this terminology due to its historical introduction in Billera and Sturmfels original introduction of monotone path polytopes in [18]. However, the name **shadow path** is more evocative and used in the optimization literature.

DEFINITION A.1.1. *Let $P \subset \mathbb{R}^n$ be a polytope, $\mathbf{c} \in \mathbb{R}^n$, and $\gamma = (\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k)$ be a $\mathbf{c}$-monotone path. Let $\pi : \mathbb{R}^n \to \mathbb{R}^2$ be defined by $\pi(\mathbf{x}) = (\mathbf{c}^\intercal \mathbf{x}, \mathbf{w}^\intercal \mathbf{x})$. Then $\pi(P)$ is a polygon, and let $U_{\mathbf{w}} = \{\mathbf{y} \in \pi(P) : \mathbf{y} + (0, \lambda) \notin \pi(P) \text{ for all } \lambda > 0\}$. Then γ is $\mathbf{w}$-coherent if $\gamma = \pi^{-1}(U_{\mathbf{w}})$ and γ is coherent if there exists some $\mathbf{w} \in \mathbb{R}^n$ such that $\gamma = \pi^{-1}(U_{\mathbf{w}})$.*

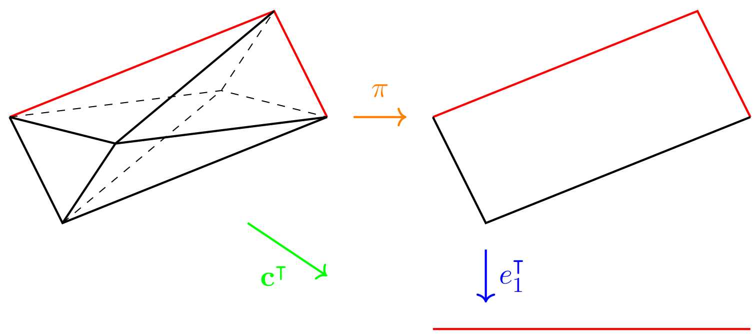

FIGURE A.2. Depicted is a description of coherence in terms of a commutative diagram. Namely, the path depicted is coherent, since it arises as the pre-image of the upper path on a polygon $\pi(P)$ such that $e_1^\intercal \pi = \mathbf{c}^\intercal$. The coherent monotone path, the upper on the polygon, and the interval from $\mathbf{c}^\intercal P$ are all highlighted in red to emphasize the correspondence between them.

The formal definition may be hard to parse as written. I encourage the reader to look at Figure A.2 for a labeled example. The intuition is as follows. Namely, on a polytope, there are many possible of monotone paths. However, on a polygon, the graph is a cycle meaning that there are always exactly two monotone paths one could take to reach the optimum. If we choose the linear function $\varphi(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1$, then one path goes along the top of the polygon Q and the other goes along the bottom. The upper path is given set theoretically as $\mathbf{U} = \{\mathbf{y} \in Q : \mathbf{y} + (0, \lambda) \notin Q\}$.
Then given a polytope P, I can project P via $\pi(P)$, where $\pi(\mathbf{x}) = (\mathbf{c}^\intercal \mathbf{x}, \mathbf{w}^\intercal \mathbf{x})$. The resulting polygon has an upper path $U_{\mathbf{w}}$. For a generic choice of $\mathbf{w}$, the pre-image of each edge of $\pi(P)$ is an edge in P. Then $\pi^{-1}(U_{\mathbf{w}})$ is a path. In fact, $U_{\mathbf{w}}$ is a $(1,0)$-monotone path. Since $\mathbf{c}^\intercal \mathbf{x} = (1,0)\pi(\mathbf{x})$, $\pi^{-1}(U_{\mathbf{w}})$ must be $\mathbf{c}$-monotone. Therefore, for a generic choice of $\mathbf{w}$, $\pi^{-1}(U_{\mathbf{w}})$ is a monotone path. The coherent monotone paths are precisely the paths that arise in this way across all possible choices of $\mathbf{w}$.
To further understand the definition, consider the case of a simplex. As stated previously, for the orientation $\mathbf{c} = (1, 2, \ldots, n)$ a monotone path on Δ_n is any path of the form $e_1, e_{i_1}, \ldots, e_{i_k}, e_n$ such that consecutive indices increase. Then for any choice of $\mathbf{w}$, we have $\pi(e_i) = (i, \mathbf{w}_i)$ meaning that choosing $\mathbf{w}$ effectively corresponds to choosing the heights of the respective points.

PROPOSITION A.1.1 ([17]). *All monotone paths on Δ_n are coherent.*

PROOF. Let $S \subseteq [2, n-1]$ be the subset of indices corresponds to points on a monotone path on the simplex. Define $\mathbf{w}_i = -1$ for all $i \in [2, n-1] \setminus S$ and $\mathbf{w}_1 = \mathbf{w}_n = 0$. Then $\pi(e_1) = (1, 0)$,

$\pi(e_n) = (n, 0)$, and $\pi(e_i) = (i, -1)$ for all $i \in [2, n-1] \setminus S$. Then for all such that i,

$$\pi(e_i) + (0,1) = (i,-1) + (0,1) = (i,0) \in \operatorname{conv}((1,0),(n,0)) = \operatorname{conv}(\pi(e_1), \pi(e_n)).$$

Hence, $\pi(e_i) \notin U_{\mathbf{w}}$ by definition.

For $i \in S$, we we will rely on the function $f : \mathbf{R} \to \mathbb{R}$ defined by $f(x) = -(1-x)(n-x)$. Note that f satisfies $f(1) = f(n) = 0$, $f(x) > 0$ for $x \in (1, n)$, and $f''(x) = -2 < 0$ for all $x \in \mathbb{R}$. In particular, f is concave, so $\{(i, f(i)) : i \in S \cup \{1, n\}\}$ is the set of vertices of a concave polygon. Note that, since f is concave,

$$C = \operatorname{conv}(\{(x, f(x)) : x \in \mathbb{R}\}) = \{(x, y) \in \mathbb{R} : y \leq f(x)\}.$$

Define $w_i = f(i)$ for $i \in S$. Then $\mathbf{w}_j \leq f(j)$ for all $j \in [n]$ meaning that $\pi(P) \subseteq C$. Hence, for all $i \in S$ and $\lambda > 0$, $(i, f(i)) + \lambda(0,1) \notin S$. Therefore, the vertices of the upper hull of P are precisely e_1, e_n, and $\{e_i : i \in S\}$. By construction, no 3 vertices in the upper hull are collinear. Hence, the pre-image of an edge in the upper path must always be an edge in P. Hence, the pre-image of $\pi(P)$ is precisely P, so the path defined by S is $\mathbf{w}$-coherent. $\square$

From this example and the definition it is not immediately clear that there exist polytopes for which not all monotone paths are coherent. In Section 3.1, we will study monotone path polytopes of cross-polytopes and note that even on the octahedron not all monotone paths are coherent.

I was very explicit in the proof of Proposition A.1.1, but using another definition, we can have more freedom for a less constructive proof. The next definition is particularly important from an optimization standpoint. It is nice that there are paths that come naturally from geometry of projections. However, it is not clear at all how an optimizer could compute each successive step of the path from the shadow definition of coherence.

Definition A.1.2. *Let $P \subset \mathbb{R}^n$ be a polytope, and let $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n$. Let $\mathbf{v}_0$ be a $\mathbf{c}$-minimum that is $\mathbf{w}$-maximal amongst all $\mathbf{c}$-minima. Then construct a path via*

$$\mathbf{v}_{i+1} = \underset{u \in N_c(\mathbf{v}_i)}{\operatorname{argmax}} \left(\frac{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{v}_i)}{\mathbf{c}^\intercal(\mathbf{u} - \mathbf{v}_i)} \right),$$

where $N_{\mathbf{c}}(\mathbf{v}_i)$ is the set of $\mathbf{c}$-improving neighbors of $\mathbf{v}_i$. If the choice of $\mathbf{v}_0$ and argmax at each step are both unique, we say the path $\gamma = \mathbf{v}_0, \ldots, \mathbf{v}_k$ is $\mathbf{w}$-coherent. We call a path coherent if there exists such a $\mathbf{w}$.

This equivalent definition is the one actually used to give an implementation for finding coherent monotone paths using the shadow simplex method, since one can compute the ratio for each neighbor to find the neighbor to choose. The intuition for the definition is grounded in interpreting what this optimization is doing. Namely, in the case of the simplex, we see that

$$\left(\frac{\mathbf{w}^{\intercal}(\mathbf{u} - \mathbf{v}_i)}{\mathbf{c}^{\intercal}(\mathbf{u} - \mathbf{v}_i)}\right) = \left(\frac{\mathbf{w}^{\intercal}(e_j - e_i)}{\mathbf{c}^{\intercal}(e_j - e_i)}\right) = \frac{w_j - w_i}{c_j - c_i}.$$

In the projection, $\mathbf{c}_i$ is the x-coordinate of $\pi(e_i)$ and $\mathbf{w}_i$ is the y-coordinate of $\pi(e_i)$. Then the ratio $\frac{\mathbf{w}_j - \mathbf{w}_i}{\mathbf{c}_j - \mathbf{c}_i}$ is the slope from $\pi(e_i)$ to $\pi(e_j)$. This is true in general. Namely, the optimization at each step is find the $\mathbf{c}$-improving neighbor of maximal slope in the projection.

The reason for the connection to slopes is the following algorithm for finding the upper path of a point cloud. This algorithm is similar to the gift wrapping algorithm for finding the convex hull of a point set in 2-D. Namely, start at the left most vertex. If there are multiple, the one of greatest height is the only one in the upper path. Take that one and call it $\mathbf{v}_0$. Then to find the next point adjacent to $\mathbf{v}_0$ on the upper hull, check all points to the right of $\mathbf{v}_0$ and find the one that yields the greatest slope $\mathbf{v}_1$. Then, if that point is unique, all remaining points will lie under the line from $\mathbf{v}_0$ to $\mathbf{v}_1$. Hence, that line is actually supporting line for the segment from $\mathbf{v}_0$ to $\mathbf{v}_1$, so $\mathbf{v}_0$ and $\mathbf{v}_1$ are adjacent and both vertices. Continuing this process, one buildings the entirely upper path. Therefore, if a path is constructed in the manner given by Definition A.1.2, it must be coherent in the sense of Definition A.1.1.

This gives an easier way to prove coherence on the simplex for a generic choice of $\mathbf{c}$. Without loss of generality, we may assume that $\mathbf{c}_1 < \mathbf{c}_2 < \cdots < \mathbf{c}_n$. Then the orientation induced on Δ_n is the same as for $(1, 2, \ldots, n)$, so the set of monotone paths does not change. Now to prove that the monotone paths are coherent, it remains to choose for each subset S of $[2, n-1]$, some $\mathbf{w}_S$ in $\mathbb{R}^n$ satisfying the desired system. To do this, I let $S = \{i_1 < i_2 < \cdots < i_k\}$. I can choose $\mathbf{w}_1$ and $\mathbf{w}_{i_1}$ arbitrarily. Then I choose values of $\mathbf{w}_i$ for $1 < i < i_1$ that force $(\mathbf{c}_i, \mathbf{w}_i)$ to be below the $\mathbf{w}_{i_1}$. This is easy to do just by making the heights small. Then for $\mathbf{w}_{i_2}$, I just choose it so that the slope from $\mathbf{w}_{i_1}$ to $\mathbf{w}_{i_2}$ is less than the slope from $\mathbf{w}_1$ to $\mathbf{w}_{i_1}$ and similarly force all points in between to lie below the resulting line. Doing this inductively builds precisely the necessary w, which proves coherence in general.

The slope condition also allows for a simple expression for the set of $\mathbf{w}$ for which a fixed monotone path is $\mathbf{w}$-coherent. Namely, for $S = \{1 = i_0, i_1, i_2, \ldots, i_k, i_{k+1} = n\}$, we have two types of inequalities:

- $\frac{\mathbf{w}_{i_j} - \mathbf{w}_{i_{j-1}}}{\mathbf{c}_{i_j} - \mathbf{c}_{i_{j-1}}} < \frac{\mathbf{w}_{i_{j+1}} - \mathbf{w}_{i_j}}{\mathbf{c}_{i_{j+1}} - \mathbf{c}_{i_j}}$ for $1 \leq j \leq k$
- $\frac{\mathbf{w}_{i} - \mathbf{w}_{i_j}}{\mathbf{c}_{i} - \mathbf{c}_{i_j}} < \frac{\mathbf{w}_{i_{j+1}} - \mathbf{w}_{i_j}}{\mathbf{c}_{i_{j+1}} - \mathbf{c}_{i_j}}$ for $i_j < i \leq i_{j+1}$ and $0 \leq j \leq k$.

Note that for fixed $\mathbf{c}$, these are linear inequalities, so these describe a polyhedral cone. Later we will observe inequalities of this form define exactly the normal cones for each vertex in the monotone path polytope of the simplex.

The final definition is also called the dual framework for coherent monotone paths, since it has a natural interpretation in terms of the normal fan of the polytope. Namely, we have the following definition for coherence:

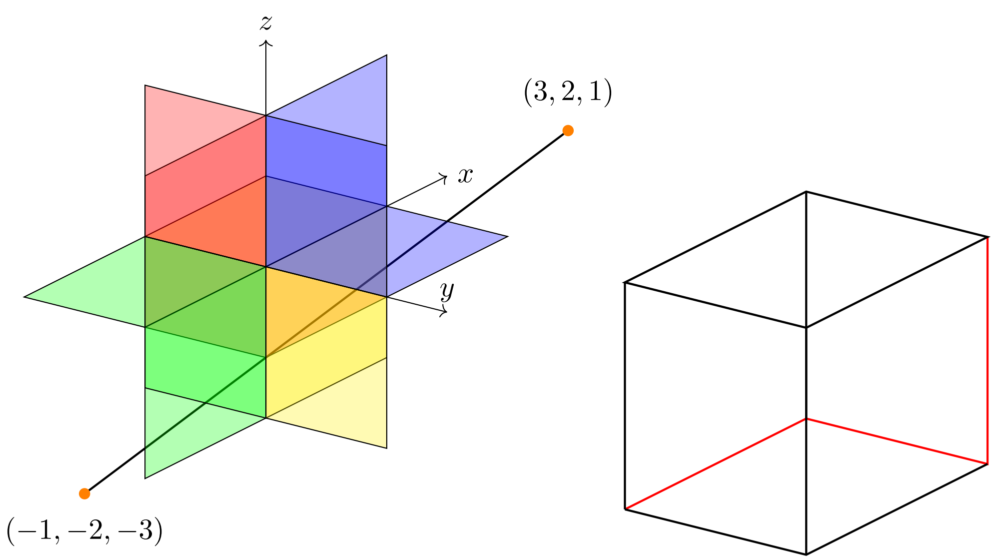

FIGURE A.3. Consider the $\mathbf{w}$-coherent $\mathbf{c}$-monotone path on the cube for $\mathbf{w} = (-1, -2, -3)$ and $\mathbf{c} = (1, 1, 1)$. The normal fan of the cube is given by the set of coordinate hyper-planes. Then the coherent monotone path is given by the vertices corresponding to the cones $(-, -, -)$, $(+, -, -)$, $(+, +, -)$, and $(+, +, +)$, which correspond to the monotone path for the vertices $(0, 0, 0), (1, 0, 0), (1, 1, 0), (1, 1, 1)$.

DEFINITION A.1.3. *Let $P \subset \mathbb{R}^n$ be a polytope, and let $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n$. Define*

$$P_{\mathbf{w}} = \bigcup_{\lambda \in \mathbb{R}} \operatorname*{argmax}_{\mathbf{x} \in P}((\mathbf{w} + \lambda c)^{\intercal} \mathbf{x}).$$

Then in $P_{\mathbf{w}}$ is a path, that monotone path is $\mathbf{w}$-coherent. Furthermore, a monotone path is coherent if only if it is of the form $P_{\mathbf{w}}$ for some $\mathbf{w}$.

See Figure A.3 for an example of an application of this definition. To unpack this definition, note that $\operatorname{argmax}_{\mathbf{x} \in P}((\mathbf{w} + \lambda c)^{\intercal} \mathbf{x})$ is always a face of P for each $\lambda \in \mathbb{R}$. For a generic choice of $\mathbf{w}$ this will always be a vertex or an edge. This can be seen in the normal fan of the polytope P. Namely, $\operatorname{argmax}_{\mathbf{x} \in P}((\mathbf{w} + \lambda c)^{\intercal} \mathbf{x})$ records exactly the smallest cone in the normal fan that contains $\mathbf{w} + \lambda c$. Then $P_{\mathbf{w}}$ records the set of all cones in the normal fan that meet the line $\{\mathbf{w} + \lambda \mathbf{c} : \lambda \in \mathbb{R}\}$. For a generic choice of $\mathbf{w}$, such a line only meets codimension 0 and 1 cones and therefore corresponds to a union of vertices and edges. Furthermore, if one traces along this line and records the cones as λ increases from $-\infty$ to ∞, consecutive cones must be adjacent, and hence this forms a path. We claim the resulting path must be the coherent monotone path one would find by projecting with $\pi(\mathbf{x}) = (\mathbf{c}^{\intercal}\mathbf{x}, \mathbf{w}^{\intercal}\mathbf{x})$. This observation is visualized in Figure A.4. To see this, note that $\mathbf{w} + \lambda \mathbf{c}$ after projecting is the vector $(\lambda, 1)$, and the $(\lambda, 1)$-maximizing face on $\pi(P)$ is an upper face, since it has a normal with positive y-coordinate. The pre-image of that face is the $\mathbf{w} + \lambda \mathbf{c}$ maximizer on P. Hence, the faces maximizing $\mathbf{w} + \lambda \mathbf{c}$ are precisely the faces that arise as pre-images of the vertices and edges in the upper path on the projection $\pi(P)$. This is why the definitions coincide. The main advantage with this definition in the case of the standard simplex is that it gives an interesting interpretation of the problem. Namely, given a vector $\mathbf{w} = (w_1, w_2, \ldots, w_n)$, it would record the maximal coordinates of $(w_1 + c_1 t, w_2 + c_2 t, \ldots, w_n + c_n t)$ for each $t \in \mathbb{R}$, where we can think of t is some unit of time. Namely, we can imagine a set of particles in some initial position $(w_1, w_2, \ldots, w_n)$ and ask which is the further to the right at all times t as t evolves both forward and backward and they move at speeds c_i. This interpretation is nonobvious from the slope definition. The key insight is that two particles share a position at a point in time if and only if

$$w_i + c_i t_{ij} = w_j + c_j t_{ij},$$

so we have

$$t_{ij} = \frac{w_j - w_i}{c_i - c_j} = -\frac{w_j - w_i}{c_j - c_i}$$

Then maximizing slope corresponds to minimizing time. There is a physical intuition for understanding what this problem models. Namely, if $c_1 < c_2 < \cdots < c_n$ as is true without loss of generality, then particle 1 is the slowest. This means that at any point in sufficiently large negative time $w_1 - \lambda c_1$ is maximal. Maximizing slope from the point (w_1, c_1) corresponds to checking for the first time the lead of 1 is overtaken by i_1. No particle between 1 and i_1 can then take the lead after, since i_1 is faster than those particles. They have no chance to catch up. Then i_1 is eventually overtaken by i_2 for some $i_2 > i_1$ and so on. All monotone paths being coherent means that the restriction that no one slower ever overtakes someone faster is the only restriction on what possible patterns of people in the lead could arise.

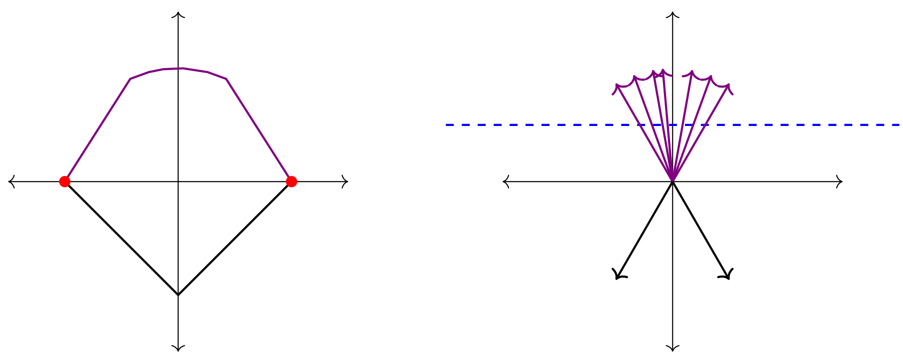

FIGURE A.4. Depicted is the upper path on a polygon together with the corresponding sequence of cones in the normal cone. One can see that the precisely the cones intersected by the line $(0,1) + \lambda(-1,0)$.

A.2. The Monotone Path Polytope

Fix a polytope $P \subset \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^n$. Then for a generic choice of $\mathbf{w}$, there is a $\mathbf{w}$-coherent $\mathbf{c}$-monotone path $\gamma_{\mathbf{w}}$ corresponding to $\mathbf{w}$. One can then define the set

$$C_{\mathbf{w}} = \{\mathbf{w}' : \gamma_{\mathbf{w}'} = \gamma_{\mathbf{w}}\}.$$

To arrive at the **shadow fan**, we first the prove the following:

LEMMA A.2.0.1. *Let $P \subseteq \mathbb{R}^n$ and $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n$ with $\mathbf{w}$ chosen generically. Then $C_{\mathbf{w}}$ is an open polyhedral cone.*

Proof. We can prove this using the slope maximizing condition of a coherent monotone path. Namely, let $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k$ be the vertices of the path. Then the max-slope rule will choose this path starting from $\mathbf{v}_0$ if and only if for all $i \in [k]$

$$\frac{\mathbf{w}^\intercal(\mathbf{v}_i - \mathbf{v}_{i-1})}{\mathbf{c}^\intercal(\mathbf{v}_i - \mathbf{v}_{i-1})} > \frac{\mathbf{w}^\intercal(\mathbf{u} - \mathbf{u})}{\mathbf{c}^\intercal(\mathbf{v}_i - \mathbf{v}_{i-1})}$$

for all $\mathbf{u} \neq \mathbf{v}_i$ in $N_c(\mathbf{v}_{i-1})$, the set of $\mathbf{c}$ improving neighbors of $\mathbf{v}_{i-1}$. The additional restriction for a coherent monotone path is that $\mathbf{v}_0$ must be the $\mathbf{w}$-maximimum of the $\mathbf{c}$-minimal face. This condition is equivalent to requiring that $\mathbf{w}$ lies in the normal cone of $\mathbf{v}_0$ in the normal fan of $\mathbf{c}$-minimal face, which is another polyhedral cone. An intersection of open polyhedral cones is an open polyhedral cone, so $C_{\mathbf{w}}$ is an open polyhedral cone. □

We define the **shadow fan** to be the polyhedral fan with open cones given by $\{C_{\mathbf{w}} : \mathbf{w} \in \mathbb{R}^n \setminus S\}$, where S is the boundary of each cone for generic $\mathbf{w}$. It is not immediately obvious that this is a polyhedral fan. To prove this, we will show it is the normal fan of a polytope.

Definition A.2.1. *Let $P \subseteq \mathbb{R}^n$ be a polytope and $\mathbf{c} \in \mathbb{R}^n$. Then the* ***monotone path polytope*** *$\Sigma_c(P)$ is defined up to normal equivalence by*

$$\sum_{v \in V(P)} \varphi^{-1}(\varphi(v)),$$

where $\varphi(x) = \mathbf{c}^\intercal(\mathbf{x})$ and $V(P)$ is the set of vertices of P.

Note the definition is only up to normal equivalence. In the context of this thesis, we care primarily about characterizing which monotone paths are coherent, since that tells us about the shadow simplex method. However, there are canonical representatives such that, for example, the monotone path polytope of a lattice polytope is always a lattice polytope, which is useful for understanding the toric interpretation of the monotone path polytope construction in [88]. However, the main initial complaint that one would have about this definition is that it is not obvious the vertices correspond to monotone paths or even how that could happen. The key observation is the following:

Theorem A.2.2. *Let $P \subset \mathbb{R}^n$ be a polytope and let $\mathbf{c} \in \mathbb{R}^n$. Then the normal fan of the monotone path polytope $\Sigma_{\mathbf{c}}(P)$ is the shadow fan.*

Proof. To prove this, it suffices to show that for each vertex $\mathbf{v} \in V(\Sigma_c(P))$, the set of $\mathbf{w}$ for which $\mathbf{v}$ is uniquely optimal is the shadow cone $C_{\mathbf{w}}$. Note that the $\mathbf{w}$-maximum of a Minkowksi sum is precisely determined by the set of $\mathbf{w}$-maxima from each summand. In particular, it suffices to understand $\varphi^{-1}(\varphi(\mathbf{v}))$ for each $\mathbf{v}$. If $\mathbf{v}$ is $\mathbf{c}$-minimal, then $\varphi^{-1}(\varphi(\mathbf{v}))$ is the $\mathbf{c}$-minimal face of P. Then $\mathbf{w}$-maximum of the $\mathbf{c}$-minimal face is exactly the vertex chosen as the start of the $\mathbf{w}$-coherent $\mathbf{c}$-monotone path. Hence, the first step of the path determines the maximum of that face.

Consider more generally the set of vertices $V(\gamma_{\mathbf{w}})$ of the coherent monotone path. Let $\mathbf{v}$ be one such vertex. Then, by the dual Definition A.1.3, there must exist some choice of λ such that $\mathbf{v}$ is a unique $\mathbf{w} + \lambda\mathbf{c}$-maximizer on P. Then, by restriction, $\mathbf{v}$ will be the unique $\mathbf{w} + \lambda\mathbf{c}$-maximizer on $\varphi^{-1}(\varphi(\mathbf{v}))$. However, note that $\varphi(x) = \mathbf{c}^\intercal(\mathbf{x})$, so $\mathbf{c}^\intercal(\mathbf{x})$ is constant on $\varphi^{-1}(\varphi(\mathbf{v}))$. Hence, $\mathbf{v}$ is also the unique $\mathbf{w}$-maximizer on $\varphi^{-1}(\varphi(\mathbf{v}))$. Thus, $V(\gamma_{\mathbf{w}})$ is a subset of the set of $\mathbf{w}$-maximizers of the slices that are vertices.

Suppose instead that $\mathbf{v}$ is the $\mathbf{w}$-maximizer of $\varphi^{-1}(\varphi(v))$. Then consider the shadow $\pi(P)$ with $\pi(\mathbf{x}) = (\mathbf{c}^\intercal\mathbf{x}, \mathbf{w}^\intercal\mathbf{x})$. Since $\mathbf{v}$ is $\mathbf{w}$-maximal on the slice of fixed φ-value, $\pi(\mathbf{v})$ is $(0,1)$-maximal only the slice of $\pi(P)$ with fixed first coordinate. Hence, $\pi(\mathbf{v})$ must lie on the upper path of the shadow and is therefore either a vertex of the shadow or lies on the interior of an edge. Since $\mathbf{w}$ is chosen generically, $\pi(\mathbf{v})$ cannot be on the interior of an edge. Hence, $\pi(\mathbf{v})$ is a vertex of the shadow and $\mathbf{v}$ is a vertex of the $\mathbf{w}$-coherent $\mathbf{c}$-monotone path.

Therefore, $V(\gamma_{\mathbf{w}})$ is precisely the set of vertices $\mathbf{v}$ that are equal to the $\mathbf{w}$-maximum of $\varphi^{-1}(\varphi(\mathbf{w}))$. This tells us that knowing which vertex of $\Sigma_c(P)$ is $\mathbf{w}$-maximal determines the shadow cone that $\mathbf{w}$ is in. That is, geometrically, the normal cone for that vertex is contained in the shadow cone. It remains to show the reverse that the shadow cone is contained in the normal cone.

Let $\mathbf{u}$ be a vertex that is not on the coherent monotone path. Then if $\varphi(\mathbf{u}) = \varphi(\mathbf{v})$ for some $\mathbf{v} \in V(\gamma_{\mathbf{w}})$, $\varphi^{-1}(\varphi(\mathbf{u})) = \varphi^{-1}(\varphi(\mathbf{v}))$, so the $\mathbf{w}$-maximum is determined by $\gamma_{\mathbf{w}}$ as desired. Otherwise, there exist consecutive $\mathbf{u}_1$ and $\mathbf{u}_2$ in $V(\gamma_{\mathbf{w}})$ such that

$$\mathbf{c}^\intercal\mathbf{u}_1 < \mathbf{c}^\intercal\mathbf{v} < \mathbf{c}^\intercal\mathbf{u}_2.$$

Then from the dual formulation of coherence in Definition A.1.3, there exists some $\lambda \in \mathbb{R}$ such that the $\mathbf{w} + \lambda\mathbf{c}$-maximal face of P is the edge $E = \operatorname{conv}(\mathbf{u}_1, \mathbf{u}_2)$. Then the $\mathbf{w} + \lambda\mathbf{c}$-maximal face of $\varphi^{-1}(\varphi(\mathbf{v}))$ is the intersection $E \cap \varphi^{-1}(\varphi(\mathbf{v}))$, which is a vertex $\mathbf{v}^*$ of $\varphi^{-1}(\varphi(\mathbf{v}))$, since E

is strictly $\mathbf{c}$-improving. Therefore, the vertices are determined precisely by the coherent monotone path. Hence, the normal cone is contained in the shadow cone meaning that the shadow cone and normal cone coincide. Hence, the normal fan of $\Sigma_{\mathbf{c}}(P)$ is the shadow fan.

$\square$

This proof is new in the sense that the standard proof appeals to monotone path polytopes being a special case of fiber polytopes in [18]. As a corollary, we have the following:

COROLLARY A.2.1. *The shadow fan is a polyhedral fan, and the vertices of the monotone path polytope correspond to coherent monotone paths.*

An evocative example is the hyper-cube $[0,1]^n$ and the vector $\mathbf{c} = (1,1,\ldots,1)$. Then $\varphi(\mathbf{x}) = \sum_{i=1}^n x_i$. Let $e_S = \sum_{s\in S} e_s$. Then $\varphi(e_S) = |S|$, and

$$\varphi^{-1}(\varphi(e_S)) = \{\mathbf{x} \in [0,1]^n : \sum_{i=1}^n \mathbf{x}_i = |S|\} = \Delta(n,|S|),$$

where $\Delta(n,|S|)$ is the hyper-simplex $\operatorname{conv}(\{e_T : |T| = |S|\})$. Up to normal equivalence, we then have

$$\Sigma_{\mathbf{c}}([0,1]^n) = \sum_{k=0}^n \Delta(n,k) = \Pi_n,$$

where Π_n denotes the permutahedron given by $\operatorname{conv}(\{(\sigma(1),\sigma(2),\ldots,\sigma(n)) : \sigma \in S_n\})$ for S_n the group of permutations of $[n]$. In this case, we are able to compute the monotone path polytope explicitly directly from the definition, though that is quite rare. Fibers are generally hard to understand.

A.3. The Baues Poset and Cellular Strings

We now know the normal fan of the monotone path polytope, but thus far we have only interpreted the vertices. The question here is motivated by understanding what happens to the shadow. Namely, let P be a polytope, and let $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n$. Then consider the shadow map π defined by $\pi(\mathbf{x}) = (\mathbf{c}^\intercal \mathbf{x}, \mathbf{w}^\intercal \mathbf{x})$. Consider the upper path $U_{\mathbf{w}}$ of $\pi(P)$. Then $\mathcal{C} = \pi^{-1}(U_{\mathbf{w}})$ is the union of the pre-images of the edges in the upper path of $\pi(P)$ and therefore a union of faces. We have the following axiomatic definition:

Definition A.3.1. *Let P be a polytope, and let $\mathbf{c} \in \mathbb{R}^n$ be a linear function. Then a* $\mathbf{c}$-***cellular string*** *on P is a sequence of faces $\mathcal{F} = (F_0, F_1, \ldots, F_k)$ such that:*

(1) F_0 and F_k contain a $\mathbf{c}$-minimizer and maximizer on P respectively,

(2) $F_{i-1}^{\mathbf{c}} = F_i^{-\mathbf{c}}$ for all $i \in [k]$, where for a polytope Q and vector $\mathbf{a}$, we let $Q^{\mathbf{a}}$ denotes the $\mathbf{a}$-maximal face of Q,

(3) $\mathbf{c}$ is non-constant on each face F_i.

A cellular string is called ***coherent*** *if it is of the form $\pi^{-1}(U_{\mathbf{w}})$ for some choice of $\mathbf{w}$.*

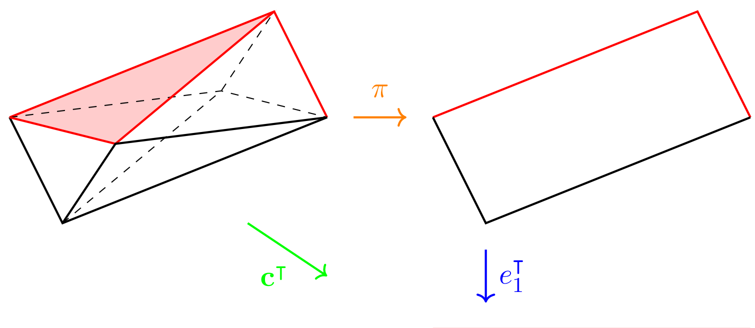

Figure A.5. Depicted is a description of coherence for cellular strings in terms of a commutative diagram. Namely, the path depicted is coherent, since it arises as the pre-image of the upper path on a polygon $\pi(P)$ such that $e_1^\intercal \pi = \mathbf{c}^\intercal$. The coherent cellular string, the upper on the polygon, and the interval from $\mathbf{c}^\intercal P$ are all highlighted in red to emphasize the correspondence between them. Note that the coherent monotone path from Figure A.2 is a vertex of the face given by the cellular string depicted here.

If the F_i are restricted to be edges, then this is equivalent to the definition of monotone path. It turns out that $\pi^{-1}(U_{\mathbf{w}})$ is always a cellular string with the cells F_i being the pre-images of consecutive edges in the path. Since the first and last edge of the upper path contain a leftmost and rightmost vertex, the pre-image faces F_0 and F_k contain $\mathbf{c}$-minimal and $\mathbf{c}$-maximal vertices respectively. The $\mathbf{c}$-maximal face of F_{i-1} is the pre-image of the rightmost vertex of the $i-1$st edge of the upper path, which is the same as leftmost vertex of the ith edge of the upper path, which yields the second property. The third and final property arises from each edge in the upper path needing to increase from left to right. This shows that all choices of $\mathbf{w}$ yield a coherent cellular string.

There are equivalent definitions in this case as well, which come down to relaxing the condition that $\mathbf{w}$ be generic. To start, the $\mathbf{w}$-maximizing face of the $\mathbf{c}$-minimal face may no longer be a vertex.

Starting from that face, maximizing slope amongst all edges exiting that face will potentially lead to a multitude of edges spanning an affine subspace. The intersection of that affine subspace and the polytope will be a face, which is the face in the cellular string. One can continue this to generate each face in the cellular string. A cellular string is coherent if and only if it arises in that way.

There is also the dual formulation for coherence of a cellular string. Namely, consider the set of faces obtained by $\{P^{\mathbf{c}+\lambda\mathbf{w}} : \lambda \in \mathbb{R}\}$, where $P^{\mathbf{y}}$ is the $\mathbf{y}$-maximal face of P. The set of faces obtained in this way will be the faces of the coherent cellular string $F_0, F_1, \ldots, F_k$ together with the endpoints $E_0, E_1, \ldots, E_k, E_{k+1}$, where E_i is the $\mathbf{c}$-minimal faces of F_i and the $\mathbf{c}$-maximal face of F_{i+1} for each i. This and the previous alternative definition are proven in precisely the same manner of proving equivalence of definitions of coherent monotone paths.

There is a poset on all cellular strings called the **Baues poset**, where for cellular strings $\mathcal{F}, \mathcal{G}$, we have $\mathcal{F} \leq \mathcal{G}$ if $\bigcup_{F\in\mathcal{F}} F \subseteq \bigcup_{G\in\mathcal{G}} G$. This poset is, in fact, a CW-complex that was studied from the perspective of combinatorial topology. An important theorem about this poset is that it is homotopy equivalent to an $(n-1)$-sphere [16]. One way to observe this is to note that this poset deformation retracts onto the induced poset of coherent cellular strings and apply the following theorem:

THEOREM A.3.2. *The face lattice of the monotone path polytope is the corresponding poset of coherent cellular strings.*

PROOF. To do this, we will describe the normal fan in more detail. Namely, let $P \subset \mathbb{R}^n$ be a polytope, and let $\mathbf{c}, \mathbf{w} \in \mathbb{R}^n$ with $\mathbf{c} \neq \mathbf{0}$. Consider the set $C_{\mathbf{w}}$ of all $\mathbf{w}'$ that induce the same coherent cellular string as $\mathbf{w}$. By using the slope formulation of coherence, one can show this is a polyhedral cone. Then, by a precisely analogous argument to that of showing that coherent monotone paths correspond to vertices, one can argue that this polyhedral cone is the normal cone of the face $\sigma_{\mathbf{c}}(P)^{\mathbf{w}}$. Namely, $\sigma_{\mathbf{c}}(P)$ may be decomposed into a sum of fibers $\varphi^{-1}(\varphi(v))$, and the $\mathbf{w}$ optimum of each fiber is the coherent cellular string intersected with it. Knowing all these intersections uniquely determines the coherent cellular string. $\square$

From this, we not only know the monotone path polytope but have a complete description. In what remains of this chapter, we will use what we have shown here to compute many examples of monotone path polytopes.

A.4. Simplices

In this section, we will describe the combinatorics of the monotone path polytope of the simplex. Note, however, that we are by no means the first to do so. This was originally done by Gelfand, Kapranov, and Zelevinsky in their remarkable tour de force Discriminants, Resultants, and Multi-dimensional Determinants [75]. In their context, they consider the monotone path polytope of the standard simplex for the orientation $\mathbf{c} = (1, 2, \ldots, n)$. In that case, the monotone path polytope is the Newton polytope of the discriminant, and they describe its combinatorics explicitly. Billera and Sturmfels coined the term monotone path polytope in [18] and also provided a proof. However, to our knowledge, the proofs we provide here are new.

Consider the simplex $\Delta_n = \operatorname{conv}(\{e_i : i \in [n]\})$, and let $\mathbf{c} \in \mathbb{R}^n$. We will consider only generic $\mathbf{c}$ in that sense that we assume all coordinates of $\mathbf{c}$ are distinct. Up to relabeling and symmetry, we may assume that $c_1 < c_2 < \cdots < c_n$. In the previous section, we used the simplex as an example throughout and showed that all monotone paths on the simplex are coherent, so the vertices correspond to subsets of $[2, n-1]$. We rely on the following lemma to completely characterize the face lattice:

Lemma A.4.0.1. *Let $P \subset \mathbb{R}^n$ be a polytope, and let $\mathbf{c} \in \mathbb{R}^n$. Suppose that $\mathbf{c}$ is vertex generic meaning that $\mathbf{c}^\intercal \mathbf{v} \neq \mathbf{c}^\intercal \mathbf{u}$ for any pair of vertices $\mathbf{u}, \mathbf{v} \in P$. Consider two cellular strings $\mathcal{F}$ and $\mathcal{G}$. Let $V(\mathcal{C})$ denote the set of vertices that appear in a cellular string $\mathcal{C}$, and let $E(\mathcal{C})$ denote the set of endpoints of the string. Then $\mathcal{F} \leq \mathcal{G}$ if and only if $E(\mathcal{F}) \supseteq E(\mathcal{G})$ and $V(\mathcal{F}) \subseteq V(\mathcal{G})$.*

Proof. Suppose first that $\mathcal{F} \leq \mathcal{G}$. Then, since $\mathcal{F}$ is a subset of $\mathcal{G}$, each vertex of $\mathcal{F}$ appears as a vertex in $\mathcal{G}$ meaning that $V(\mathcal{F}) \subseteq V(\mathcal{G})$. Arguing for endpoints is more subtle.

Note that, since $\mathbf{c}$ is vertex generic, all endpoints are vertices. Let $\mathbf{v} \in E(\mathcal{G})$. Note that any cellular string must contain a point of each objective function value. In particular, $\mathcal{F}$ has to contain a point with the same objective function value as $\mathbf{v}$. The unique point with objective function value $\mathbf{c}^\intercal \mathbf{v}$ in $\mathcal{G}$ is $\mathbf{v}$. Hence, since $\mathcal{F}$ is a subset of $\mathcal{G}$, we must have that $\mathbf{v}$ is the unique point with objective function value $\mathbf{c}^\intercal \mathbf{v}$ in $\mathcal{F}$. It remains to show that $\mathbf{v} \in E(\mathcal{F})$. Since $\mathbf{c}$ is constant on $\mathbf{v}$, by definition of a cellular string, $\mathbf{v}$ must be contained in a face in the string. Since $\mathbf{v}$ is the unique point with its objective function value, by convexity, it must either be a maximum or minimum of that face and therefore an endpoint of the cellular string. Hence, $\mathbf{v} \in E(\mathcal{F})$ meaning that $E(\mathcal{F}) \supseteq E(\mathcal{G})$, which finishes the proof of the first direction.

Suppose instead that $E(\mathcal{F}) \supseteq E(\mathcal{G})$ and $V(\mathcal{F}) \subseteq V(\mathcal{G})$. Consider two consecutive endpoints $\mathbf{x}$ and $\mathbf{y}$ in $E(\mathcal{G})$. Then let F be the face with minimum $\mathbf{x}$ and maximum $\mathbf{y}$, and V denote the set of vertices of $V(\mathcal{F})$ that have objective function value between $\mathbf{x}$ and $\mathbf{y}$. Then $V \subseteq F$, since F includes all vertices with objective function values between $\mathbf{x}$ and $\mathbf{y}$, and all cells from $\mathbf{x}$ to $\mathbf{y}$ in $\mathcal{F}$ must therefore be contained in F. Hence, each of those cells is contained in $\mathcal{G}$. Doing this for each consecutive endpoints shows eahc cell of $\mathcal{F}$ must be a subset of a cell in $\mathcal{G}$ meaning that $\mathcal{F} \leq \mathcal{G}$ as desired.

$\square$

Note that the consideration of endpoints here is necessary. For example on the tetrahedron, the $\{e_1, e_2, e_3\}$ and $\{e_3, e_4\}$ is a cellular string with the same set of vertices as $\{e_1, e_2\}$, $\{e_2, e_3\}$, and $\{e_3, e_4\}$, but they are nonetheless distinct due to differing endpoints. Note also that the endpoints do not determine the cellular string. Namely, $\{e_1, e_2, e_3\}, \{e_3, e_4\}$ has the same endpoints as $\{e_1, e_3\}, \{e_3, e_4\}$. The previous result shows that the two pieces of data determine both the cellular string and inclusion relations between cellular strings so long as $\mathbf{c}$ is generic. We may take advantage of this to find a combinatorial descriptions of the cellular strings of the simplex.

Theorem A.4.1 ([18]). *For the simplex Δ_n and any $\mathbf{c} \in \mathbb{R}^n$ such that $c_1 < c_2 < \cdots < c_n$, we have $\Sigma_{\mathbf{c}}(\Delta_n)$ is combinatorially equivalent to a hyper-cube. Furthermore, all cellular strings are coherent.*

Proof. To prove this, it suffices to characterize the cellular strings combinatorially and prove directly they are coherent. Note first that any subset of vertices of the simplex is a face. To determine the cellular strings, it then suffices to understand how to glue the cellular strings together. First of all, recall the **endpoints** of a cellular string are the $\mathbf{c}$-minimal and $\mathbf{c}$-maximal faces of each face in the string. Since $\mathbf{c}$ is generic, the endpoints are always going to be vertices of Δ_n.

For the sake of our analysis, we will determine a cellular string by its set of endpoints E together with the set of vertices that appear in some cell of the string V. Note that both sets always contain e_1 and e_n, so we will consider the pair $E' = E\backslash\{e_1, e_n\}$ and $V' = V\backslash\{e_1, e_n\}$. Then, by construction $E' \subseteq V'$. Then we make the following claims.

(1) (E', V') uniquely determines the cellular string and gives a bijection between cellular strings and pairs of subsets (S, T) such that $S \subseteq T \subseteq [2, n-1]$

(2) We have $(E', V') \leq (X, Y)$ if and only if $X \subseteq E'$ and $Y \supseteq V'$.

Given a pair of subsets (E', V') with $E' \subseteq V'$, we may write the set of elements of E' in order, $e_{i_1}, e_{i_2}, \ldots, e_{i_k}$. Then the cellular string is given by

$$\bigcup_{j=0}^{k} \operatorname{conv}(\{e_\ell : \ell \in V', i_j \leq \ell \leq i_{j+1}\}),$$

where we define $i_0 = 0$ and $i_{k+1} = n$. The inverse of this map is given by taking the set of endpoints and vertices contained in a given cellular string and it is therefore a bijection. The inclusion relation follows immediately from Lemma A.4.0.1. Note then that the face lattice is precisely the set of intervals in the Boolean poset with respect to inclusion, which is the face lattice of the hyper-cube. Hence, the poset of all cellular strings is isomorphic to the face lattice of the hyper-cube. Note that using poset topology machinery and the work of Billera and Sturmfels, this observation is already enough to complete the proof that all cellular strings are coherent. However, in this work, I prefer to prove things like this directly.

Let $\mathcal{C}$ be a cellular string. To prove coherence, we need to find a choice of $\mathbf{w} \in \mathbb{R}^n$ such that $\pi^{-1}(U_{\mathbf{w}}) = \mathcal{C}$, where $\pi(\mathbf{x}) = (\mathbf{c}^\intercal \mathbf{x}, \mathbf{w}^\intercal \mathbf{x})$ and $U_{\mathbf{w})}$ is the upper hull of $\pi(P)$. To do this, first let E denote the endpoints of the cellular string. For choosing a projection π, note that the x-coordinates in the image are fixed, and by linear endpoints, we may choose the heights of the points freely. In particular, let $e_1 = e_{i_0}, e_{i_1}, \ldots, e_{i_k}, e_{i_{k+1}} = e_n$ be the endpoints in the cellular string. I may choose the height of the leftmost endpoints e_1 and e_{i_1} freely. Then I choose the height of $e_{i_{j+1}}$ such that the slope from $\pi(e_{i_j})$ to $\pi(e_{i_{j+1}})$ is less than the slope from $\pi(e_{i_{j-1}})$ to $\pi(e_{i_j})$ for all remaining choices of j. Then for the remaining vertices of each cell, I choose heights to force that they lie on the line segment between each endpoint of that cell. For any remaining vertex $\mathbf{v}$, there exist endpoints $\mathbf{c}^\intercal e_{i_j} < \mathbf{c}^\intercal \mathbf{v} < \mathbf{c}^\intercal e_{i_{j+1}}$ since $\mathbf{c}$ is vertex generic. We may choose the height of $\mathbf{v}$ to force that it lies strictly below the line segment from $\pi(e_{i_j})$ to $\pi(e_{i_{j+1}})$. By construction, $\mathcal{C} = \pi^{-1}(U_{\mathbf{w}})$ for such a choice of $\mathbf{w}$. Therefore, all cellular strings are coherent as desired. ☐

All of this was proven for the standard simplex, but one can actually show this holds for any simplex and any vertex generic $\mathbf{c}$. To do this, we will rely on a few lemmas. The first explains one reason why the monotone path polytope is also called a **fiber polytope** for one-dimensional linear projections.

LEMMA A.4.1.1 ([18]). *Let $P \subseteq \mathbb{R}^n$ be a polytope, let $\mathbf{c} \in \mathbb{R}^n$, and define $\varphi(\mathbf{x}) = \mathbf{c}^\intercal \mathbf{x}$. Then every fiber of φ is a weak Minkowski summand of $\Sigma_{\mathbf{c}}(P)$.*

PROOF. Consider a fiber $\varphi^{-1}(x)$. If $x = \varphi(\mathbf{v})$ for a vertex of P, then we are done as it is a Minkowski summand. Suppose otherwise. First, to show it is a weak Minkowski summand, it suffices to show for each $\mathbf{w} \in \mathbb{R}^n$, the $\mathbf{w}$-maximum of $\Sigma_{\mathbf{c}}(P)$ determines the $\mathbf{w}$-maximum of $\varphi^{-1}(x)$. Note that $x = \varphi(y)$ for some y in a unique cell F_i of the coherent cellular string determined by $\mathbf{w}$. Since F_i is a cell of the coherent cellular string, F_i must be $\mathbf{w}+\lambda\mathbf{c}$-maximal for some $\lambda \in \mathbb{R}$. Hence, $F_i \cap \varphi^{-1}(x)$ must be a $\mathbf{w}+\lambda\mathbf{c}$-maximal face $\varphi^{-1}(x)$. Since φ is constant on $\varphi^{-1}(x)$ by definition, $F_i \cap \varphi^{-1}(x)$ must also be $\mathbf{w}$-maximal. Hence, the coherent cellular string for $\mathbf{w}$ on P determines the $\mathbf{w}$-maximal face of $\varphi^{-1}(x)$ meaning that $\varphi^{-1}(x)$ is a weak Minkowski summand of $\Sigma_{\mathbf{c}}(P)$. ☐

In general a fiber polytope $\Sigma_\rho(P)$ comes from a linear projection $\rho : \mathbb{R}^n \to \mathbb{R}^m$, and the normal fan of the fiber polytope is the set of the refine of the fans of all fibers of ρ. It turns out the normal fans of fibers only come in finitely many types. By taking a representative of each type and adding them up, we see that the resulting refinement is the normal fan of a polytope that we call the fiber polytope. This all generalizes what happens for the monotone path polytope, and there are many analogous statements in that setting. There is a second lemma we rely on that is also true for fiber polytopes. Our statement is only for monotone path polytopes, but generalizes to arbitrary fiber polytopes.

LEMMA A.4.1.2 ([18]). *Let $P \subseteq \mathbb{R}^n$ be a polytope, $\mathbf{c} \in \mathbb{R}^m$, and consider a linear map $\rho : \mathbb{R}^n \to \mathbb{R}^m$. Then*

$$\rho(\Sigma_{\mathbf{c}'}(P)) = \Sigma_{\mathbf{c}}(\rho(P)),$$

where $\mathbf{c}' = \mathbf{c}^\intercal \rho$ and equality is normal equivalence. If ρ is a linear isomorphism of polytopes, then $\Sigma_{\mathbf{c}'}(P)$ is linearly isomorphic to $\Sigma_{\mathbf{c}}(\rho(P))$.

PROOF. Let $\varphi(\mathbf{x}) = \mathbf{c}^\intercal \mathbf{x}$ and $\varphi'(\mathbf{x}) = \mathbf{c}'^\intercal \mathbf{x}$. Then, by construction of our choice of $\mathbf{c}'$, each fiber $\varphi'^{-1}(\varphi'(\mathbf{v}))$ for a vertex $\mathbf{v} \in P$ satisfies

$$\rho(\varphi'^{-1}(\varphi'(\mathbf{v}))) = \varphi^{-1}(\varphi'(\mathbf{v})).$$

Thus, ρ maps fibers to fibers. Furthermore, each vertex of $\rho(P)$ has a pre-image in P, so for each $\mathbf{v}$, $\varphi^{-1}(\varphi(\mathbf{v}))$ is of the form $\rho(\varphi^{-1}(\varphi'(\rho^{-1}(\mathbf{v})))$. That is the map on fibers covers over fibers that sum to the monotone path polytope. From the Minkowski sum decomposition definition of the monotone path polytope, it follows that $\Sigma_{\mathbf{c}}(\rho(P))$ is a Minkowski summand of $\rho(\Sigma_{\mathbf{c}'}(P))$. To show normal equivalence, it suffices to show that $\rho(\Sigma_{\mathbf{c}'}(P))$ is a weak Minkowski summand of $\Sigma_{\mathbf{c}}(\rho(P))$. Equivalently, it suffices to show each Minkowski summand in the Minkowski sum decomposition of $\rho(\Sigma_{\mathbf{c}'}(P))$ is a weak Minkowski summand of $\Sigma_{\mathbf{c}}(P)$. Note first that $\rho(\varphi'^{-1}(\varphi'(\mathbf{v}))$ is a fiber of φ, so by Lemma A.4.1.1, we are done.

In the case of a linear isomorphism, the restriction of the map to fibers at vertices is also a linear isomorphism, ρ induces a linear isomorphism on the Minkowski sums of fibers at vertices and therefore the monotone path polytopes as desired. □

By applying linear transformations to our previous result, we see that

Corollary A.4.1. *For a vertex generic choice of* $\mathbf{c} \in \mathbb{R}^n$ *and any simplex* $P \subseteq \mathbb{R}^n$, $\Sigma_{\mathbf{c}}(P)$ *is always a combinatorial cube.*

A.5. Hyper-cubes and Zonotopes

The work in this section is expository based on what was already known for general fiber polytopes in Billera and Sturmfels' original paper [18]. Monotone path polytopes of zonotopes were studied further under the name **Sweep Polytopes** and in somewhat different framework in [115]. Zonotopes for which all cellular string are coherent were characterized completely in [65], and monotone path polytopes of zonotopes were studied in great detail in the PhD thesis of Rob Edman [64].

We start with the hyper-cube. In Section A.2, we computed the mpp of a hyper-cube for one particular orientation and showed it was normally equivalent to the permutahedron. From the definition of cellular strings, we have a more practical way to describe the space of cellular strings.

Lemma A.5.0.1. *Let* $P \subset \mathbb{R}^n$ *be a polytope, and let* $\mathbf{c} \in \mathbb{R}^n$ *be an edge generic vector. That is for all* $u, v \in V(P)$ *such that* u *and* v *are adjacent,* $\mathbf{c}^\intercal(\mathbf{u} - \mathbf{v}) \neq 0$. *Define a directed graph* $\Gamma(P)$ *on* $V(P)$, *where there is a directed edge from* $\mathbf{u}$ *to* $\mathbf{v}$ *if there exists non-vertex face* F *such that* $\mathbf{u}$ *is the* $\mathbf{c}$*-minimum of* F *and* $\mathbf{v}$ *is the* $\mathbf{c}$*-maximum of* F. *Then we have*

- *The set of endpoints of cellular strings on $\Gamma(P)$ are in bijection with paths in $\Gamma(P)$ from the* $\mathbf{c}$*-minimum to the* $\mathbf{c}$*-maximum*
- *The set of cellular strings is in bijection with a path* $\mathbf{v}_1, \dots, \mathbf{v}_k$ *together with a face certifying adjacency for each consecutive pair of vertices* $\mathbf{v}_i$, $\mathbf{v}_{i+1}$.

PROOF. This is immediate from the definition of cellular strings. Namely, consider any sequence of vertices $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k$ such that $\mathbf{v}_0$ is $\mathbf{c}$-minimal, $\mathbf{v}_k$ is $\mathbf{c}$-maximal, and there is a directed edge from $\mathbf{v}_i$ to $\mathbf{v}_{i+1}$ in $\Gamma(P)$. Then there must exist a sequence of face $F_1, F_2, \dots, F_k$ such that $\mathbf{v}_{i-1}$ is the $\mathbf{c}$-minimum of F_i and $\mathbf{v}_i$ is the $\mathbf{c}$-maximuum of F_i. By construction, this sequence of faces is a cellular string with the desired endpoints. Furthermore, all cellular strings arise in this way. $\square$

We start by extending the observation that $\Sigma_{(1,1,\dots,1)}([0,1]^n)$ is an $(n-1)$-permutahedron to any generic orientation.

THEOREM A.5.1 ([18]). *Let $\mathbf{c} \in \mathbb{R}^n$ and suppose that $\mathbf{c}$ is chosen generically. Then $\Sigma_{\mathbf{c}}([0,1]^n)$ is combinatorially equivalent to the permutahedron, and all cellular strings are coherent.*

PROOF. Note that, up to symmetries of the cube, we may without loss of generality assume that $\mathbf{c} > \mathbf{0}$ by Lemma A.4.1.2. Each vertex of the cube may be written as $e_S = \sum_{i \in S} e_i$ for some $S \subseteq [n]$. Note that faces of the cube are in bijection with intervals in the boolean lattice $[S,T]$, where the face is given by $\{e_R : R \in [S,T]\}$. For the orientation $\mathbf{c}$, the $\mathbf{c}$-minimum and $\mathbf{c}$-maximum of such a face are e_S and e_T respectively. Then, by Lemma A.5.0.1, sets of endpoints of cellular strings are in bijection with flags of proper nonempty subsets (i.e., a sequence of subsets of $S_i \subsetneq S_1 \subsetneq \cdots \subsetneq S_k$). Furthermore, the $\mathbf{c}$-minimum and $\mathbf{c}$-maximum uniquely determine the face of the cube, so the endpoints determine the cellular string entirely.

To prove coherence, start with a flag $\emptyset = S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_k \subsetneq S_{k+1} = [n]$. Then choose $\mathbf{w}_i = \frac{k+1-i}{c_i}$ for all $i \in S_i \setminus S_{i-1}$. The coherent cellular string starts at $\mathbf{0}$. The set of improving neighbors with maximal slope at $\mathbf{0}$ is precisely the $\{j \in S_1 \setminus S_0\}$. Hence, the first cell taken is the one spanned by all those neighbors, which is the face corresponding to $[S_0, S_1]$. The same principle continues for all remaining steps yielding a realization of the desired cellular string. Hence, all cellular strings are coherent.

Since all cellular strings are coherent, and the choice of cellular strings does not depend on the choice of $\mathbf{c}$, for all $\mathbf{c}$, the face lattice is isomorphic to the poset of coherent cellular strings for $\mathbf{c} = (1, 1, \ldots, 1)$, which we showed in Section A.2 is the face lattice of the permutahedron. $\square$

For the hyper-cube and more generally zonotopes, we may also find an explicit description of the normal fan of the corresponding monotone path polytope. Namely recall that a **zonotope** is, by definition, the image of the hyper-cube under a linear map (up to affine transformation). We may write the hyper-cube $\sum_{i=1}^{n}[\mathbf{0}, e_i]$, where $[\mathbf{x}, \mathbf{y}]$ is shorthand for the $\mathrm{conv}(\mathbf{x}, \mathbf{y})$ for vectors $\mathbf{x}, \mathbf{y} \in \mathbf{R}^m$. Then a zonotope may be written as $Z = \sum_{i=1}^{n}[\mathbf{0}, \mathbf{v}_i]$. Then choose $\mathbf{c} \in \mathbb{R}^n$ such that $\mathbf{c}^\intercal \mathbf{v}_i > 0$ for all $i \in [n]$. This choice is universal in the sense that any generic orientation of a zonotope may arise in this way by possibly transforming the zonotope by replacing some $\mathbf{v}_i$ with $-\mathbf{v}_i$. For such a $\mathbf{c}$, $\mathbf{0}$ and $\sum_{i=1}^{n} \mathbf{v}_i$ are the $\mathbf{c}$-minimum and maximum respectively. Furthermore, up to normal equivalence, we may assume that no pair of $\mathbf{v}_i$ are parallel meaning that monotone paths correspond to ordered decompositions of $\sum_{i=1}^{n} \mathbf{v}_i$. That is a choice of $\sigma \in S_n$ such that the kth step of the path of the path is $\sum_{i=1}^{k} \mathbf{v}_{\sigma(i)}$.

For $e_1, e_2, \ldots, e_n$, any choice of σ yields a valid path on the boundary of the polytope. In general, the situation is more complicated. Some permutations correspond to paths, while others do not. However, if there is a choice of $\mathbf{w}$ such that

$$\frac{\mathbf{w}^\intercal \mathbf{v}_{\sigma(1)}}{\mathbf{c}^\intercal \mathbf{v}_{\sigma(1)}} > \frac{\mathbf{w}^\intercal \mathbf{v}_{\sigma(2)}}{\mathbf{c}^\intercal \mathbf{v}_{\sigma(2)}} > \cdots > \frac{\mathbf{w}^\intercal \mathbf{v}_{\sigma(n)}}{\mathbf{c}^\intercal \mathbf{v}_{\sigma(n)}},$$

then by the max-slop formulation for coherent monotone paths, σ induces a coherent monotone path. Furthermore, this is an if and only if, since by convexity, consecutive edges in the upper path in a project must always have decreasing slopes. This gives us a description of the normal fan:

Theorem A.5.2. *Let $Z = Z(\mathbf{v}_1, \ldots, \mathbf{v}_k) \subset \mathbb{R}^n$ be a zonotope, and let $\mathbf{c}$ be such that $\mathbf{c}^\intercal(\mathbf{v}_i) > 0$ for all i. Then we have the normal fan of $\Sigma_{\mathbf{c}}(Z)$ is the hyper-plane arrangement given by*

$$\mathcal{H} = \bigcup_{1 \leq i < j \leq k} \left\{ \mathbf{x} \in \mathbb{R}^n : \left(\frac{\mathbf{v}_i}{\mathbf{c}^\intercal \mathbf{v}_i} - \frac{\mathbf{v}_j}{\mathbf{c}^\intercal \mathbf{v}_j} \right)^\intercal \mathbf{x} = 0 \right\}.$$

Proof. The regions of this arrangement correspond to indicating which side of a hyperplane a vector $\mathbf{w} \in \mathbb{R}^n$ lies on. Each hyper-plane here separates space by whether the difference of slope between two $\mathbf{v}_i$ is positive or negative. Equivalently, each hyperplane determines which of the slopes is larger between a pair of $\mathbf{v}_i$. Hence, a region is determined precisely by the total orderings on the slopes of the $\mathbf{v}_i$. A total ordering of such a slope is also precisely what determines a coherent monotone path meaning that that the hyper-plane arrangement is the normal fan $\Sigma_{\mathbf{c}}(Z)$. $\square$

Zonotopes generalize all of what happens for the hyper-cube. However, it is rare for all cellular strings to be coherent on a zonotope. In many of the remaining examples we study here, we will prove that the phenomenon of all cellular strings being coherent extends to a broad class of other polytopes that do generalize the cube.

Bibliography

[1] X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig, *Log-barrier interior point methods are not strongly polynomial*, SIAM Journal on Applied Algebra and Geometry, 2 (2018), pp. 140–178.

[2] X. Allamigeon, S. Gaubert, and N. Vandame, *No self-concordant barrier interior point method is strongly polynomial*, in Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing, 2022, pp. 515–528.

[3] N. Amenta and G. Ziegler, *Shadows and slices of polytopes*, in Proceedings of the twelfth annual symposium on computational geometry, 1996, pp. 10–19.

[4] ———, *Deformed products and maximal shadows of polytopes*, Contemporary Mathematics, 223 (1999), pp. 57–90.

[5] O. Angel, V. Gorin, and A. Holroyd, *A pattern theorem for random sorting networks*, Electronic Journal of Probability, 17 (2012), pp. 1–16.

[6] O. Angel, A. Holroyd, D. Romik, and B. Virág, *Random sorting networks*, Advances in Mathematics, 215 (2007), pp. 839–868.

[7] M. Aprile, A. Cevallos, and Y. Faenza, *On 2-level polytopes arising in combinatorial settings*, SIAM Journal on Discrete Mathematics, 32 (2018), pp. 1857–1886.

[8] N. Arkani-Hamed, S. He, and T. Lam, *Stringy canonical forms*, Journal of High Energy Physics, 2021 (2021), pp. 1–62.

[9] C. Athanasiadis, *Piles of cubes, monotone path polytopes, and hyperplane arrangements*, Discrete & Computational Geometry, 21 (1999), pp. 117–130.

[10] C. Athanasiadis, J. De Loera, V. Reiner, and F. Santos, *Fiber polytopes for the projections between cyclic polytopes*, European Journal of Combinatorics, 21 (2000), pp. 19–47.

[11] M. Balinski and A. Russakoff, *On the assignment polytope*, Siam Review, 16 (1974), pp. 516–525.

[12] D. Barnette, *An upper bound for the diameter of a polytope*, Discrete Mathematics, 10 (1974), pp. 9–13.

[13] A. Barvinok, *A course in convexity*, vol. 54, American Mathematical Soc., 2002.

[14] H. Baues, *Geometry of loop spaces and the cobar construction*, vol. 230, American Mathematical Soc., 1980.

[15] D. Bertsimas and J. Tsitsiklis, *Introduction to linear optimization*, vol. 6, Athena scientific Belmont, MA, 1997.

[16] L. Billera, M. Kapranov, and B. Sturmfels, *Cellular strings on polytopes*, Proc. Amer. Math. Soc., 122 (1994), pp. 549–555.

[17] L. Billera and B. Sturmfels, *Fiber polytopes*, Ann. of Math. (2), 135 (1992), pp. 527–549.

[18] ———, *Fiber polytopes*, Annals of Mathematics, (1992), pp. 527–549.

[19] ———, *Iterated fiber polytopes*, Mathematika, 41 (1994), pp. 348–363.

[20] R. Bixby, *Solving real-world linear programs: A decade and more of progress*, Operations research, 50 (2002), pp. 3–15.

[21] A. Björner, *The antiprism fan of a convex polytope*, in Amer. Math. Soc, vol. 18, 1997, p. 1.

[22] A. Black, *Small shadows of lattice polytopes*, in Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), SIAM, 2023, pp. 1669–1679.

[23] A. Black, S. Borgwardt, and M. Brugger, *Short circuit walks on hirsch counterexamples*, arXiv preprint arXiv:2302.03977, (2023).

[24] A. Black and J. De Loera, *Monotone paths on cross-polytopes*, Discrete & Computational Geometry, 70 (2023), pp. 1245–1265.

[25] A. Black, J. De Loera, S. Kafer, and L. Sanità, *On the simplex method for 0/1-polytopes*, Mathematics of Operations Research, To Appear (2023).

[26] A. Black, J. De Loera, N. Lütjeharms, and R. Sanyal, *The polyhedral geometry of pivot rules and monotone paths*, SIAM Journal on Applied Algebra and Geometry, 7 (2023), pp. 623–650.

[27] A. Black and R. Sanyal, *Flag polymatroids*, arXiv preprint arXiv:2207.12221, (2022).

[28] M. Blanchard, J. De Loera, and Q. Louveaux, *On the length of monotone paths in polyhedra*, SIAM Journal on Discrete Mathematics, 35 (2021), pp. 1746–1768.

[29] R. Bland, *New finite pivoting rules for the simplex method*, Mathematics of operations Research, 2 (1977), pp. 103–107.

[30] N. Bonifas, M. Di Summa, F. Eisenbrand, N. Hähnle, and M. Niemeier, *On sub-determinants and the diameter of polyhedra*, Discrete & Computational Geometry, 52 (2014), pp. 102–115.

[31] K. Borgwardt, *The Simplex Method: a probabilistic analysis*, vol. 1, Springer Science & Business Media, 2012.

[32] S. Borgwardt, J. De Loera, and E. Finhold, *The diameters of network-flow polytopes satisfy the Hirsch conjecture*, Mathematical Programming, 171 (2018), pp. 283–309.

[33] S. Borgwardt, W. Grewe, and J. Lee, *On the diameter of a 2-sum of polyhedra*, arXiv preprint arXiv:2311.02047, (2023).

[34] ———, *On the combinatorial diameters of parallel and series connections*, SIAM Journal on Discrete Mathematics, 38 (2024), pp. 485–503.

[35] T. Brunsch and H. Röglin, *Finding short paths on polytopes by the shadow vertex algorithm*, in Automata, Languages, and Programming: 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part I 40, Springer, 2013, pp. 279–290.

[36] J. Cardinal and R. Steiner, *Inapproximability of shortest paths on perfect matching polytopes*, in International Conference on Integer Programming and Combinatorial Optimization, Springer, 2023, pp. 72–86.

[37] J. Cardinal and R. Steiner, *Shortest paths on polymatroids and hypergraphic polytopes*, arXiv preprint arXiv:2311.00779, (2023).

[38] F. Castillo and J.-P. Labbé, *Lineup polytopes of product of simplices*, arXiv preprint arXiv:2306.00082, (2023).

[39] F. Castillo, J.-P. Labbé, J. Liebert, A. Padrol, E. Philippe, and C. Schilling, *An effective solution to convex 1-body n-representability*, in Annales Henri Poincaré, Springer, 2023, pp. 1–81.

[40] C. Ceballos and G. Ziegler, *Realizing the associahedron: Mysteries and questions*, in Associahedra, Tamari Lattices and Related Structures: Tamari Memorial Festschrift, Springer, 2012, pp. 119–127.

[41] F. Chapoton, S. Fomin, and A. Zelevinsky, *Polytopal realizations of generalized associahedra*, Canadian Mathematical Bulletin, 45 (2002), pp. 537–566.

[42] S. Chubanov, *A generalized simplex method for integer problems given by verification oracles*, SIAM J. Optim., 31 (2021), pp. 686–701.

[43] ———, *A generalized simplex method for integer problems given by verification oracles*, SIAM Journal on Optimization, 31 (2021), pp. 686–701.

[44] ———, *A scaling algorithm for optimizing arbitrary functions over vertices of polytopes*, Mathematical Programming, 190 (2021), pp. 89–102.

[45] H. S. M. Coxeter, *Regular polytopes*, Dover Publications, Inc., 1973.

[46] D. Dadush and N. Hähnle, *On the shadow simplex method for curved polyhedra*, Discrete & Computational Geometry, 56 (2016), pp. 882–909.

[47] D. Dadush and N. Hähnle, *On the shadow simplex method for curved polyhedra*, Discrete & Computational Geometry, 56 (2016), pp. 882–909.

[48] D. Dadush and S. Huiberts, *A friendly smoothed analysis of the simplex method*, SIAM Journal on Computing, 49 (2019), pp. 18–449.

[49] G. Dantzig, *Linear programming and extensions*, Princeton university press, 1963.

[50] G. Dantzig, *Origins of the simplex method*, in A history of scientific computing, 1990, pp. 141–151.

[51] D. Dauvergne, *The Archimedean limit of random sorting networks*, Journal of the American Mathematical Society, 35 (2022), pp. 1215–1267.

[52] D. Dauvergne and B. Virág, *Circular support in random sorting networks*, Transactions of the American Mathematical Society, 373 (2020), pp. 1529–1553.

[53] J. De Loera, *Triangulations of polytopes and computational algebra*, Cornell University, 1995.

[54] J. De Loera, S. Kafer, and L. Sanità, *Pivot rules for circuit-augmentation algorithms in linear optimization*, SIAM Journal on Optimization, 32 (2022), pp. 2156–2179.

[55] J. De Loera, J. Rambau, and F. Santos, *Triangulations: structures for algorithms and applications*, vol. 25, Springer Science & Business Media, 2010.

[56] A. Del Pia and C. Michini, *On the diameter of lattice polytopes*, Discrete & Computational Geometry, 55 (2016), pp. 681–687.

[57] ———, *Short simplex paths in lattice polytopes*, Discrete & Computational Geometry, (2022), pp. 1–22.

[58] A. Deza, G. Manoussakis, and S. Onn, *Primitive zonotopes*, Discrete & Computational Geometry, 60 (2018), pp. 27–39.

[59] A. Deza and L. Pournin, *Improved bounds on the diameter of lattice polytopes*, Acta Mathematica Hungarica, 154 (2018), pp. 457–469.

[60] A. Deza and L. Pournin, *Primitive point packing*, arXiv:2006.14228, (2020).

[61] Y. Disser, O. Friedmann, and A. V. Hopp, *An exponential lower bound for zadeh's pivot rule*, Mathematical Programming, 199 (2023), pp. 865–936.

[62] M. Dyer and A. Frieze, *Random walks, totally unimodular matrices, and a randomised dual simplex algorithm*, Mathematical Programming, 64 (1994), pp. 1–16.

[63] P. Edelman and C. Greene, *Balanced tableaux*, Advances in Mathematics, 63 (1987), pp. 42–99.

[64] R. Edman, *Diameter and coherence of monotone path graphs in low corank*, PhD thesis, University of Minnesota, 2015.

[65] R. Edman, P. Jiradilok, G. Liu, and T. McConville, *Zonotopes whose cellular strings are all coherent*, European Journal of Combinatorics, 96 (2021), p. 103352.

[66] C. Eur, A. Fink, M. Larson, and H. Spink, *Signed permutohedra, delta-matroids, and beyond*, arXiv preprint arXiv:2209.06752, (2022).

[67] S. Fomin and A. Zelevinsky, *Cluster algebras i: foundations*, Journal of the American mathematical society, 15 (2002), pp. 497–529.

[68] J. Fourier, *Histoire de l'académie, partie mathématique (1824)*, Mémoires de l'Académie des sciences de l'Institut de France, 7 (1827).

[69] M. Fredman, *How good is the information theory bound in sorting?*, Theoretical Computer Science, 1 (1976), pp. 355–361.

[70] O. Friedmann, T. D. Hansen, and U. Zwick, *Subexponential lower bounds for randomized pivoting rules for the simplex algorithm*, in Proceedings of the forty-third annual ACM symposium on Theory of computing, 2011, pp. 283–292.

[71] K. Fukuda, *Polyhedral computation*, (2020).

[72] K. Fukuda and T. Terlaky, *Criss-cross methods: A fresh view on pivot algorithms*, Mathematical programming, 79 (1997), pp. 369–395.

[73] B. Gärtner and I. Schurr, *Linear programming and unique sink orientations*, in Symposium on Discrete Algorithms: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm, vol. 22, 2006, pp. 749–757.

[74] S. Gass and T. Saaty, *The computational algorithm for the parametric objective function*, Naval Res. Logist. Quart., 2 (1955), pp. 39–45.

[75] I. Gelfand, M. Kapranov, and A. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Springer Science & Business Media, 2008.

[76] D. Goldfarb and W. Sit, *Worst case behavior of the steepest edge simplex method*, Discrete Applied Mathematics, 1 (1979), pp. 277–285.

[77] J. Goodman and R. Pollack, *Upper bounds for configurations and polytopes in* $\mathbb{R}^d$, Discrete & Computational Geometry, 1 (1986), pp. 219–227.

[78] ———, *Allowable sequences and order types in discrete and computational geometry*, in New trends in discrete and computational geometry, Springer, 1993, pp. 103–134.

[79] B. Grünbaum, *Convex polytopes*, vol. 16, Springer, 2003.

[80] Gurobi Optimization, LLC, *Gurobi Optimizer Reference Manual*, 2023.

[81] P. Hersh, *Posets arising as 1-skeleta of simple polytopes, the nonrevisiting path conjecture, and poset topology*, Discrete & Computational Geometry, (2023), pp. 1–30.

[82] C. Hertrich, A. Basu, M. Di Summa, and M. Skutella, *Towards lower bounds on the depth of relu neural networks*, Advances in Neural Information Processing Systems, 34 (2021), pp. 3336–3348.

[83] A. Hoffman, *Cycling in the simplex algorithm*, Selected Papers of Alan Hoffman with Commentary, (2003), p. 177.

[84] C. Hohlweg and C. Lange, *Realizations of the associahedron and cyclohedron*, Discrete & Computational Geometry, 37 (2007), pp. 517–543.

[85] S. Huiberts, Y. T. Lee, and X. Zhang, *Upper and lower bounds on the smoothed complexity of the simplex method*, in Proceedings of the 55th Annual ACM Symposium on Theory of Computing, 2023, pp. 1904–1917.

[86] V. Kaibel and K. Kukharenko, *Polytope extensions with linear diameters*, arXiv preprint arXiv:2307.05246, (2023).

[87] G. Kalai and D. J. Kleitman, *A quasi-polynomial bound for the diameter of graphs of polyhedra*, Bulletin of the American Mathematical Society, 26 (1992), pp. 315–316.

[88] M. Kapranov, B. Sturmfels, and A. Zelevinsky, *Quotients of toric varieties*, Mathematische Annalen, 290 (1991), pp. 643–655.

[89] J. Kelner and D. Spielman, *A randomized polynomial-time simplex algorithm for linear programming*, in Proceedings of the thirty-eighth annual ACM symposium on Theory of computing, 2006, pp. 51–60.

[90] L. Khachiyan, *A polynomial algorithm in linear programming*, in Doklady Akademii Nauk, vol. 244, Russian Academy of Sciences, 1979, pp. 1093–1096.

[91] E. Kim and F. Santos, *An update on the Hirsch conjecture*, Jahresbericht der Deutschen Mathematiker-Vereinigung, 112 (2010), pp. 73–98.

[92] T. Kitahara and S. Mizuno, *Klee-Minty's LP and upper bounds for Dantzig's simplex method*, Operations Research Letters, 39 (2011), pp. 88–91.

[93] ———, *A bound for the number of different basic solutions generated by the simplex method*, Mathematical Programming, 137 (2013), pp. 579–586.

[94] V. Klee, *Diameters of polyhedral graphs*, Canadian Journal of Mathematics, 16 (1964), pp. 602–614.

[95] V. Klee and G. J. Minty, *How good is the simplex algorithm*, Inequalities, 3 (1972), pp. 159–175.

[96] P. Kleinschmidt and S. Onn, *On the diameter of convex polytopes.*, Discret. Math., 102 (1992), pp. 75–77.

[97] A. Klyachko, *Quantum marginal problem and n-representability*, in Journal of Physics: Conference Series, vol. 36, IOP Publishing, 2006, p. 014.

[98] U. Kortenkamp, J. Richter-Gebert, A. Sarangarajan, and G. Ziegler, *Extremal properties of 0/1-polytopes*, Discrete & Computational Geometry, 17 (1997), pp. 439–448.

[99] K. Kukharenko and L. Sanità, *On the number of degenerate simplex pivots*, in International Conference on Integer Programming and Combinatorial Optimization, Springer, 2024, p. To appear.

[100] T. Kuno, Y. Sano, and T. Tsuruda, *Computing kitahara–mizuno's bound on the number of basic feasible solutions generated with the simplex algorithm*, Optimization letters, 12 (2018), pp. 933–943.

[101] D. Larman, *Paths on polytopes*, Proceedings of the London Mathematical Society, 3 (1970), pp. 161–178.

[102] J. Lee and S. Leyffer, *Mixed integer nonlinear programming*, vol. 154, Springer Science & Business Media, 2011.

[103] G. Liu, *A zonotope and a product of two simplices with disconnected flip graphs*, Discrete & Computational Geometry, 59 (2018), pp. 810–842.

[104] D. Maclagan and B. Sturmfels, *Introduction to tropical geometry*, vol. 161, American Mathematical Society, 2021.

[105] C. Mallows and R. Vanderbei, *Which Young tableaux can represent an outer sum?*, Journal of Integer Sequences, 18 (2015), p. 15.9.1.

[106] J. Matoušek and B. Gärtner, *Understanding and using linear programming*, vol. 1, Springer, 2007.

[107] J. Matoušek and T. Szabó, *Random edge can be exponential on abstract cubes*, Advances in Mathematics, 204 (2006), pp. 262–277.

[108] B. Matschke, F. Santos, and C. Weibel, *The width of five-dimensional prismatoids*, Proceedings of the London Mathematical Society, 110 (2015), pp. 647–672.

[109] T. Matsui and S. Tamura, *Adjacency on combinatorial polyhedra*, Discrete Applied Mathematics, 56 (1995), pp. 311–321.

[110] K. Murty, *Computational complexity of parametric linear programming*, Mathematical programming, 19 (1980), pp. 213–219.

[111] D. Naddef, *The Hirsch conjecture is true for (0, 1)-polytopes*, Mathematical Programming, 45 (1989), pp. 109–110.

[112] H. Narayanan, R. Shah, and N. Srivastava, *A spectral approach to polytope diameter*, in 13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA, M. Braverman, ed., vol. 215 of LIPIcs, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[113] C. Nöbel and R. Steiner, *Hardness of circuit and monotone diameters of polytopes*, arXiv preprint arXiv:2404.04158, (2024).

[114] J. Olarte and F. Santos, *Hypersimplicial subdivisions*, Selecta Mathematica, 28 (2022), p. 4.

[115] A. Padrol and E. Philippe, *Sweeps, polytopes, oriented matroids, and allowable graphs of permutations*, Combinatorica, (2023), pp. 1–61.

[116] C. Papadimitriou, *The adjacency relation on the traveling salesman polytope is np-complete*, Mathematical Programming, 14 (1978), pp. 312–324.

[117] A. Postnikov, *Permutohedra, associahedra, and beyond*, International Mathematics Research Notices, 2009 (2009), pp. 1026–1106.

[118] G. Poullot, *Geometric combinatorics of paths and deformations of convex polytopes*, PhD thesis, Sorbonne Université, 2023.

[119] F. Rispoli, *The monotonic diameter of the perfect matching and shortest path polytopes*, Operations Research Letters, 12 (1992), pp. 23–27.

[120] ———, *The monotonic diameter of traveling salesman polytopes*, Operations Research Letters, 22 (1998), pp. 69–73.

[121] L. Sanità, *The diameter of the fractional matching polytope and its hardness implications*, in 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 2018, pp. 910–921.

[122] F. Santos, *A counterexample to the Hirsch conjecture*, Annals of mathematics, (2012), pp. 383–412.

[123] A. Schrijver, *Combinatorial optimization: polyhedra and efficiency*, vol. 24, Springer, 2003.

[124] D. Sleator, R. Tarjan, and W. Thurston, *Rotation distance, triangulations, and hyperbolic geometry*, in Proceedings of the eighteenth annual ACM symposium on Theory of computing, 1986, pp. 122–135.

[125] N. Sloane and T. O. F. Inc., *The on-line encyclopedia of integer sequences*, 2020.

[126] S. Smale, *Mathematical problems for the next century*, Mathematical Intelligencer, 20 (1998), pp. 7–15.

[127] D. Spielman and S. Teng, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, Journal of the ACM (JACM), 51 (2004), pp. 385–463.

[128] R. P. Stanley, *Decompositions of rational convex polytopes*, Ann. Discrete Math, 6 (1980), pp. 333–342.

[129] B. Sturmfels, *Grobner bases and convex polytopes*, vol. 8, American Mathematical Soc., 1996.

[130] N. Sukegawa, *Improving bounds on the diameter of a polyhedron in high dimensions*, Discrete Mathematics, 340 (2017), pp. 2134–2142.

[131] T. Szabó and E. Welzl, *Unique sink orientations of cubes*, in Proceedings 42nd IEEE Symposium on Foundations of Computer Science, IEEE, 2001, pp. 547–555.

[132] D. Tamari, *Monoides préordonnés et chaines de Malcev*, PhD thesis, Paris, 1951.

[133] ———, *The algebra of bracketings and their enumeration*, Nieuw Arch. Wisk, 3 (1962), pp. 131–146.

[134] T. Terlaky and S. Zhang, *Pivot rules for linear programming: A survey on recent theoretical developments*, Annals OR, 46-47 (1993), pp. 203–233.

[135] M. Todd, *The monotonic bounded Hirsch conjecture is false for dimension at least 4*, Mathematics of Operations Research, 5 (1980), pp. 599–601.

[136] ———, *An improved Kalai–Kleitman bound for the diameter of a polyhedron*, SIAM Journal on Discrete Mathematics, 28 (2014), pp. 1944–1947.

[137] R. Vershynin, *Beyond Hirsch conjecture: walks on random polytopes and smoothed complexity of the simplex method*, SIAM J. Comput., 39 (2009), pp. 646–678.

[138] G. Ziegler, *Lectures on polytopes*, vol. 152, Springer Science & Business Media, 2012.