Title
Combinatorial and Machine Learning Problems Motivated by the Simplex Method

Author
Zhang, Zhenyang

Publication Date
2022

Peer reviewed|Thesis/dissertation

Combinatorial and Machine Learning Problems Motivated by the Simplex Method

By

ZHENYANG ZHANG
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

Luis Rademacher

Committee in Charge

2022

Contents

Zhenyang Zhang
June 2022
Mathematics

Abstract

This dissertation discusses several problems motivated by the simplex method, one of the most influential algorithms in optimization.

First, every generic linear functional f on a convex polytope P induces an orientation on the graph of P. We introduce the notions of f-arborescence and f-monotone path on P, as well as a natural graph structure on the vertex set of f-monotone paths on the resulting directed graphs. These combinatorial objects are proxies for pivot rules and simplex method pivot steps. We bound the number of f-arborescences, the number of f-monotone paths, and the diameter of the graph of f-monotone paths for polytopes P in terms of their dimension and number of vertices or facets. We also sample the distribution of lengths of monotone paths over different classes of random polytopes.

Second, inspired by the simplex and the criss-cross methods, we present an update on the search for bounds on the diameter of the cocircuit graph of an oriented matroid. We review the diameter problem and show the diameter bounds of general oriented matroids reduce to those of uniform oriented matroids. We give the latest exact bounds for oriented matroids of low rank and low corank, and for all oriented matroids with up to nine elements. For arbitrary oriented matroids, we present an improvement to a quadratic bound of Finschi. Our discussion highlights an old conjecture that states a linear bound for the diameter is possible. On the positive side, we show the conjecture is true for oriented matroids of low rank and low corank, and, verified with computers, for all oriented matroids with up to nine elements. On the negative side, our computer search showed two natural strengthenings of the main conjecture are false.

Finally, we discuss a data-driven, empirically-based framework to make algorithmic decisions or recommendations without expert knowledge. We improve the performance of the simplex method by selecting different pivot rules for different linear programs. We train machine learning methods to select the optimal pivot rule for given data without expert opinion. We use two types of techniques,

neural networks and boosted decision trees. Our selection framework recommends various pivot rules that improve overall total performance over just using a default fixed pivot rule. Here our recommendation system is best when using gradient boosted trees. Our data analysis also shows that the number of iterations by steepest-edge is no more than four percent from the optimal selection.

The thesis is structured as follows: In Chapter 1 we introduce the readers to the basic notions in Sections 1.1 and 1.2, and summarize our main results in Sections 1.3, 1.4 and 1.5. Section 2.1 through Section 2.4 prove our stated bounds on number of arborescences, monotone paths, and the diameter of flip graphs; Section 2.5 demonstrates the distributions related to length of monotone paths, and we conclude Chapter 2 with some open problems in Section 2.6. Chapter 3 proves the diameter conjecture in low rank and corank oriented matroids and shows the constructions of counterexamples. In Chapter 4 we show how we generate the training data and the comparison between different machine learning models to select pivot rules. In Appendix, we present the pseudocode for converting oriented matroids in different representations and for computing different features on directed polytope graphs.

Acknowledgments

I want to acknowledge a group of people that helped me all the way from my undergraduate years to finishing my thesis. Without your help and support I could not have made to this point.

First, I would like to thank my advisor, Jesús A. De Loera. You have been not only my teacher, but also my greatest cheerleader for the last four years. Thank you for being patient with me and informing all the opportunities you know on conferences, summer schools and internships. You taught me so many things, mathematical and more importantly how to treat math with passion and love. Thanks for all the help and support in this pivot step of my life path. For advisors, I could not think of anyone better.

Next, I would like to acknowledge all of my coauthors Ilan Adler, Christos A. Athanasiadis, Jesús A. De Loera and Steven Klee. This dissertation is based on studies in several papers I have coauthored with you.

My thanks also go to Professor Yan Zhang, David Aldous, and Fu Liu for inspiring me into my PhD research projects as well as teaching me the necessary knowledge that I utilize the most during my five years of study. Without you, this thesis will probably never be in this form.

A number of peers, whether here or there, have helped me both mathematically and personally. First and foremost, to my roommate Haolin Chen, thanks for all the help during my five years of PhD life. It would have been much harder, especially during the pandemic, if I were living alone. Yonggyu Lee, Jianping Pan and Stephen Sheng, thank you for all the helpful suggestions on choosing advisors and tips for research. Thank you, Xinyi Ni and Tongzhou Wang for the advice and support on coding. And thank you Joseph Pappe and Chengyang Wang for the helpful conversations when I was stuck in research.

Additionally, I am grateful to Limin Huang for his collaboration in an undergraduate research program.

My coauthors and I are grateful for the comments we received from anonymous referees as well as from Aviv Adler, Louis Billera, Lukas Finschi, Komei Fukuda, Kolja Knauer, Nati Linial, Francisco Santos, and Tamás Terlaky.

I was supported by NSF grants DMS-1818969 and NSF HDR TRIPODS grant CCF-1934568. We are grateful for the support. We are grateful for the wonderful working environments provided by UC Davis at which most of the research work was conducted. Part of my travel was supported by UC Davis Mathematics Travel Awards Program.

CHAPTER 1

Introduction

1.1. Preliminaries

1.1.1. Polytopes. We will briefly introduce definitions on polytopes. For more details, readers may refer to Barvinok [**Bar02**] and Chapters 1, 2 and 3 of Ziegler [**Zie95**]. A polytope is the generalization of polygons in higher dimensions. In order to define polytopes formally, we first introduce some concepts from affine linear algebra. A set S is *convex* if $\forall \mathbf{x}, \mathbf{y} \in S$, $\mathbf{z} = \lambda \mathbf{x} + (1 - \lambda)\,\mathbf{y} \in S$, $0 < \lambda < 1$. Let $a_1, a_2, \ldots, a_n \in \mathbb{R}$ and at least one of them is not zero. Let $c \in \mathbb{R}$ be a constant. A *hyperplane* is the set of all vectors $\mathbf{x} \in \mathbb{R}^n$ such that $a_1x_1 + a_2x_2 + \ldots + a_nx_n = c$. Geometrically, a hyperplane in $\mathbb{R}^n$ is an affine subspace with dimension $n-1$. A hyperplane in $\mathbb{R}^n$ separates $\mathbb{R}^n$ into two *halfspaces*, which are the sets $\{\mathbf{x} \in \mathbb{R}^n : a_1x_1 + a_2x_2 + \ldots + a_nx_n > c\}$ and $\{\mathbf{x} \in \mathbb{R}^n : a_1x_1 + a_2x_2 + \ldots + a_nx_n < c\}$. A hyperplane together with one of the halfspaces is called a *generalized halfspace.*

Formally speaking, a *polytope* P is the convex hull of a finite number of points in $\mathbb{R}^d$. A *polyhedron* P is the intersection of finitely many closed half-spaces in $\mathbb{R}^d$. Polytopes are bounded polyhedra by Weyl-Minkowski Theorem [**Bar02**]. A d-dimensional *simple* polytope is a d-dimension polytope where each vertex is included in at most d edges. Simple polytopes are corresponding to non-degenerate LP problems [**Kal97**]. A polytope is called *simplicial* if all its proper faces are simplices. Simple polytopes are the polar duals of simplicial polytopes.

A linear inequality $a_1x_1 + a_2x_2 + \ldots + a_nx_n \leq c$ is *valid* for a polytope P if it is satisfied for all $x \in P$. A *face* is any set of the form $F = P \cap \{x \in \mathbb{R}^d : a_1x_1 + a_2x_2 + \ldots + a_nx_n = c\}$ where $a_1x_1 + a_2x_2 + \ldots + a_nx_n \leq c$ is a valid inequality for P. The *dimension* of face F is the dimension of its affine hull. Note that both P and ϕ are faces of P (for inequality $\mathbf{0} \cdot \mathbf{x} \leq 0$ and $\mathbf{0} \cdot \mathbf{x} \leq 1$ respectively). Faces of dimensions 0, 1 and $dim(P) - 1$ are called *vertices, edges*, and *facets* respectively.

The vertices and edges of P form an undirected graph, which is closely related to Linear Programming and the simplex method. The *diameter* of a graph is the smallest number δ such that any two vertices can be connected by a path with at most δ edges. Denote $\Delta(d,n)$ the maximal diameter of the graph of an d-dimension polytope P with at most n facets. The original famous Hirsch Conjecture stated the following,

Conjecture 1.1.1. *for $n > d \geq 2$, $\Delta(n,d) \leq n-d$.*

The maximum *diameter* of a d-dimensional polytope with n facets is achieved by simple polytopes since all polytopes can be perturbed into simple polytopes with diameters at least as large [**KW67**]. Unfortunately, Conjecture 1.1.1 was disproved by F. Santos in [**San12**], by a counter-example with 40 facets in dimension 20. However, the question of whether $\Delta(n,d)$ can be bounded by a polynomial in n and d, or the polynomial Hirsch conjecture, remains open.

The best known general upper bound was $(n-d)^{\log d}$ by Todd [**Tod14**] and is improved by Sukegawa [**Suk16**]. Vershynin [**Ver09**] showed that every polyhedron can be perturbed by a small random amount so that the expected diameter of the perturbed polyhedron is bounded by a polynomial in d and $\log n$. For some special cases, researchers have proven the polynomial Hirsch conjecture to be true (see Naddef [**Nad89**] for 0/1 polytopes, Orlin [**Orl97**] for flow-polytopes, Brightwell et al. [**BvdHS06**] for transportation polytopes). Bonifas et al. [**Bon**] has shown that if the polytope can be written as $P = \{Ax \leq b\}, A \in \mathbb{Z}^{n\times d}$, its diameter is bounded by $O(\Delta d^4 \log d\Delta)$, where Δ is the largest absolute value of sub-determinants of A. See the survey by Kim and Santos [**KS10**] for more information on diameter bound.

1.1.2. Directed Polytope graphs. In this thesis we are interested in directed polytope graphs. Consider a d-dimensional convex polytope P in Euclidean space $\mathbb{R}^d$ and a generic linear functional f on P, meaning a linear functional on $\mathbb{R}^d$ which is nonconstant on every edge of P. The first part of the thesis investigates extremal enumerative problems about f-arborescences and f-monotone paths on the graph of P.

The functional f, which we think of as an objective function, induces an orientation on the graph of P which orients every edge in the direction of increasing objective value. Such orientations of polytope graphs are called *LP-admissible*; they are of great importance in the study of the simplex

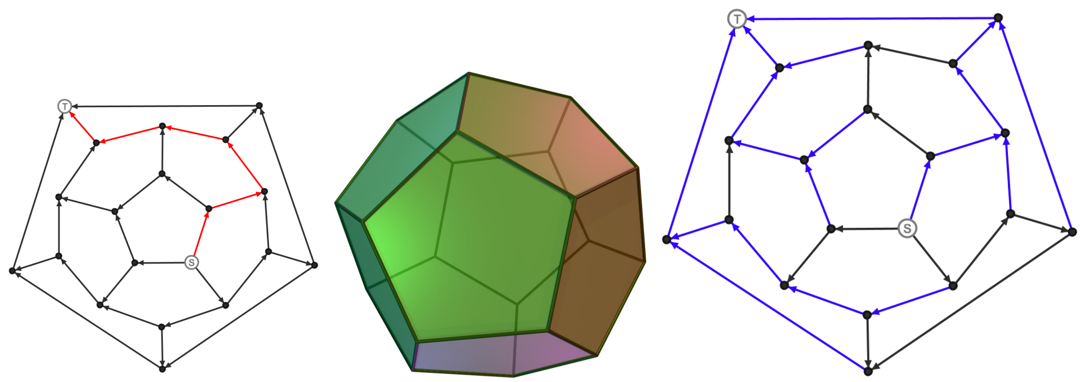

FIGURE 1.1. The regular dodecahedron (center), with examples of an f-monotone path (left) and an f-arborescence (right) for one of its LP-admissible orientations.

method for linear optimization (see [**Dev04, MK00**] and the references given there). The resulting directed graph, consisting of all vertices and oriented edges of P and denoted by $\omega(P, f)$, is acyclic and has a unique source and a unique sink on every face of P. An *f-monotone path* on P is any directed path in $\omega(P, f)$ having as initial and terminal vertex the unique source, say $v_{\min}$, and the unique sink, say $v_{\max}$, of $\omega(P, f)$ on P, respectively. An *f-arborescence* is any (necessarily acyclic) spanning subgraph $\mathcal{A}$ of the directed graph $\omega(P, f)$ such that for every vertex v of P there exists a unique directed path in $\mathcal{A}$ with initial vertex v and terminal vertex $v_{\max}$ (see Figure 1.1 for an example). As explained in the sequel, f-arborescences and f-monotone paths are important notions in geometric combinatorics and optimization. We denote *monotone height* the length of the longest f-monotone path. When the context is clear, we simply refer to them as arborescences and monotone paths.

The set of all f-monotone paths on P can be given a natural graph structure as follows. We say that two f-monotone paths on P differ by a *polygon flip* (also called *polygon move*, or simply *flip*) across a 2-dimensional face F if they agree on all edges not lying on F but follow the two different f-monotone paths on F, from the unique source to the unique sink of $\omega(P, f)$ on F. The *graph of f-monotone paths* (also called *flip graph*) on P is denoted by $G(P, f)$ and is defined as the simple (undirected) graph which has nodes all f-monotone paths on P and as edges all unordered pairs of such paths which differ by a polygon flip across a 2-dimensional face of P. The graph $G(P, f)$ is connected; its higher connectivity was studied in [**AER00**], where it was shown that $G(P, f)$

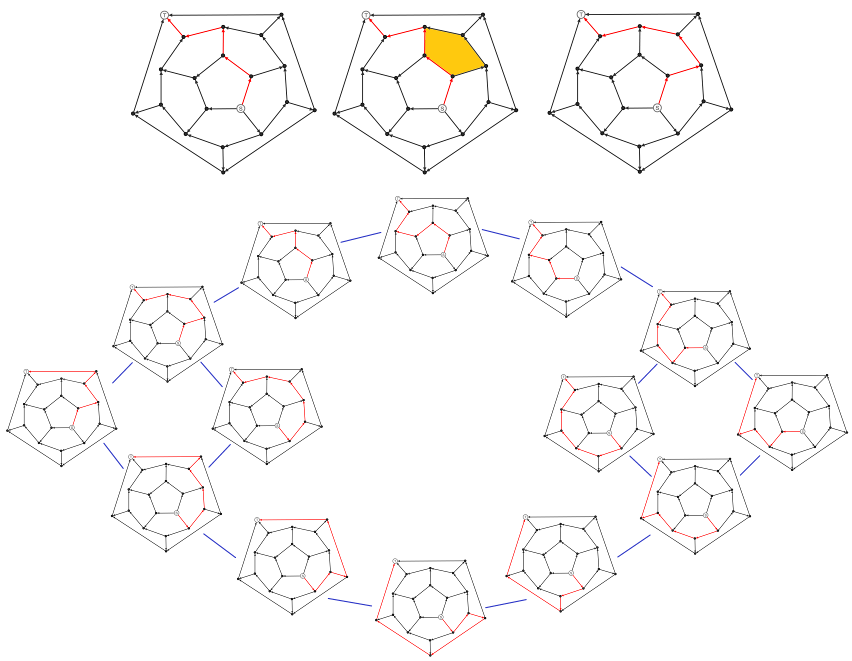

FIGURE 1.2. A polygon flip on the oriented dodecahedron and the resulting flip graph

is 2-connected for every polytope P of dimension $d \geq 3$ and $(d-1)$-connected for every simple polytope P of dimension d (see Figure 1.2 for an example).

Note that the pictures in both Figure 1.1 and Figure 1.2 are computed and plotted by a MATLAB software *PolyPathLab* we wrote. In Appendix, we will present the pseudocode of PolyPathLab for each feature (including generating random 3-dimensional polytope, computing diameter, monotone diameter, number of monotone paths and flip graphs of a polytope) and the simulated distributions of the monotone diameters of different types of polytopes.

1.1.3. Some special polytopes. Special classes of polytopes play an important role in this thesis, since they are optimal solutions of the extremal problems considered. A convenient way to encode the numbers of faces of each dimension of a simple or simplicial d-dimensional polytope P is provided by the *h-vector*, denoted as $h(P) = (h_0(P), h_1(P), \ldots, h_d(P))$; see pages 8, 59 and 248 of [**Zie95**] for details and more information. The h-vector of a simple polytope P has nonnegative integer coordinates which afford an elegant combinatorial interpretation: $h_k(P)$ equals the number of vertices of P of outdegree k in the directed graph $\omega(P, f)$, discussed in the introduction, for every generic linear functional f on P (see Sections 3.4 and 8.3 and Exercise 8.10 in [**Zie95**]); in particular, the multiset of such outdegrees is independent of f.

A polytope is called *2-neighborly* if every pair of vertices is connected by an edge. A d-dimensional simplicial polytope is called *neighborly* if any $\lfloor d/2 \rfloor$ or fewer of its vertices form the vertex set of a face. Neighborly polytopes other than simplices (cyclic polytopes being distinguished representatives) exist in dimensions four and higher. Their significance comes from the fact that they maximize the entries of the h-vector among all polytopes with given dimension and number of vertices (see pages 15-16, and 254-257 of [**Zie95**]); in particular, they maximize the numbers of faces of each dimension among such polytopes. The h-vector of a neighborly d-dimensional polytope P with n vertices is given by the formulas $h_k(P) = \binom{n-d+k-1}{k}$ for $0 \le k \le \lfloor d/2 \rfloor$ and $h_k(P) = h_{d-k}(P)$ for $0 \le k \le d$ (see Theorem 8.21 and Lemma 8.26 of [**?**]).

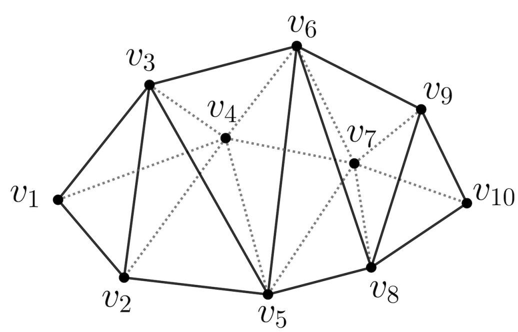

FIGURE 1.3. Example of the polytope $X(10)$

A *stacked polytope* is any simplicial polytope which can be obtained from a simplex by repeatedly glueing other simplices of the same dimension along common facets, so as to preserve convexity at each step. Equivalently, the boundary complex of a stacked polytope can be obtained combinatorially from that of a simplex by successive stellar subdivisions on facets. The h-vector of any stacked polytope P of dimension d with n vertices has the simple form $h(P) = (1, n-d, ..., n-d, 1)$ (see [**McM04**]). A fundamental result of Barnette [**Bar73**] states that among all simplicial polytopes with given dimension and number of vertices, the stacked polytopes have the fewest possible faces of each dimension. Moreover, as a consequence of the generalized lower bound theorem, stacked polytopes minimize the entries of the h-vector among all such polytopes (see [**Kal87, MN13**]).

Many different combinatorial types of stacked polytopes are possible. For each $n \geq 4$, we will consider a 3-dimensional stacked polytope of special type with n vertices, denoted by $X(n)$. This polytope comes together with a linear functional f which linearly orders its vertices as $f(v_1) < f(v_2) < \cdots < f(v_n)$. The associated triangulation comprises of all faces of the simplices with vertex sets $\{v_1, v_2, v_3, v_4\}$, $\{v_2, v_3, v_4, v_5\}, \ldots, \{v_{n-3}, v_{n-2}, v_{n-1}, v_n\}$, so the dual graph of this triangulation is a path (these dual graphs for general stacked polytopes are trees). The regularity of this triangulation easily implies that such polytope $X(n)$ and linear functional f exist for every $n \geq 4$. Figure 1.3 shows an example with $n = 10$.

A crucial property of $X(n)$ is that the directed graph $\omega(X(n), f)$ has as arcs the pairs (v_i, v_j) for $i, j \in \{1, 2, \ldots, n\}$ with $j \in \{i+1, i+2, i+3\}$. The following combinatorial lemma establishes the lower bound for the diameter of flip graphs, claimed in Theorem 1.3.2.

1.1.4. Oriented Matroids. Oriented matroids are combinatorial structures that generalize many types of objects, including point configurations, vector configurations, hyperplane arrangements, polyhedra, linear programs, and directed graphs. Oriented matroids have played a key role in combinatorics, geometry, and optimization (see the book by Björner et al. [**BLVS$^+$99**] for a complete treatise and Chapter 6 of Ziegler [**Zie95**] for a quick introduction we assume in the rest). In this thesis, we investigate a natural graph called the *cocircuit graph of an oriented matroid.*

We use standard notation about oriented matroids from Ziegler [**Zie95**] and the classic book of Björner et al. [**BLVS$^+$99**].

For the moment we will only provide the *cocircuit axioms* of oriented matroids. As with classical matroids, there are also several cryptomorphic definitions of oriented matroids; see [**BLVS$^+$99**] for more details. We will briefly introduce some of these details later. The cocircuits and covectors of an oriented matroid are special types of sign vectors that satisfy certain axioms:

Definition 1.1.2. An oriented matroid $\mathcal{M} = (E, \mathcal{C}^*)$ consists of a finite set E and a subset $\mathcal{C}^* \subseteq \{+, -, 0\}^E$, called signed cocircuits, that satisfy the following conditions.

(CC0) $\mathbf{0} \notin \mathcal{C}^*$;

(CC1) if $X \in \mathcal{C}^*$, then $-X \in \mathcal{C}^*$;

(CC2) for all $X, Y \in \mathcal{C}^*$, if $\mathrm{supp}(X) \subseteq \mathrm{supp}(Y)$, then $X = Y$ or $X = -Y$; and

(CC3) if $X, Y \in \mathcal{C}^*$, $X \neq -Y$, and $e \in S(X,Y)$, then there exists $Z \in \mathcal{C}^*$ such that $Z^+ \subseteq (X^+ \cup Y^+) \setminus \{e\}$ and $Z^- \subseteq (X^- \cup Y^-) \setminus \{e\}$.

A purely combinatorial description of oriented matroids can be given in terms of special *sign vectors*. If E is a finite set, we use $\{+,-,0\}^E$ to denote the set of all vectors of signs, with entries indexed by the elements of E. We will use capital letters $X, Y, Z, \ldots$ to represent elements of $\{+,-,0\}^E$ and subscripts X_e to reference the entry of X indexed by the element $e \in E$. We can always negate a sign vector: if $X = (X_e : e \in E)$, then $-X = (-X_e : e \in E)$.

The *positive*, *negative*, and *zero* parts of a sign vector $X \in \{+,-,0\}^E$ are defined respectively as $X^+ = \{e \in E : X_e = +\}$, $X^- = \{e \in E : X_e = -\}$, and $X^0 = \{e \in E : X_e = 0\}$. The *support* of X is defined as $\operatorname{supp}(X) = X^+ \cup X^-$. If X and Y are sign vectors, their *separating set* is $S(X,Y) = (X^+ \cap Y^-) \cup (X^- \cap Y^+)$, and their *composition* is the sign vector $X \circ Y$ whose entries are given by

$$(X \circ Y)_e = \begin{cases} X_e & \text{if } X_e \neq 0, \\ Y_e & \text{otherwise.} \end{cases}$$

Given an oriented matroid $\mathcal{M}$, we can consider the set $\mathcal{V}^* = \{X^0 \circ X^1 \circ \cdots \circ X^k \ : \ X^i \in \mathcal{C}^*(\mathcal{M})\}$ of all possible signed *covectors*, obtained by successively composing signed cocircuits. The set $\mathcal{V}^*$ has a natural poset structure, which we denote by $\mathcal{L}(\mathcal{V}^*)$ (in fact, $\mathcal{L}(\mathcal{V}^*)$ is a graded lattice). The order is obtained from the component-wise partial order on vectors in $\{+,-,0\}^E$ with $0 < +, -$. We will revisit this poset later in a geometric setting.

The *rank* of $\mathcal{M}$ is defined to be one less than the length of the longest chain of elements in the poset $\mathcal{L}(\mathcal{V}^*)$. Again, this is not the only way to define the rank of an oriented matroid [**BLVS$^+$99**]. We say an element of E is a *coloop* if it is not present in the support of any signed cocircuit. For brevity, signed cocircuits will also be called cocircuits. It is well known that every matroid has a *dual matroid*. In the case of oriented matroids, this concept is more delicate, but there is also a notion of duality. One can then talk about *circuits*, which are the cocircuits of the dual oriented matroid, and the related notions of *corank*, *loops*, etc. The corank of an oriented matroid on n elements of rank r is $n - r$.

The *cocircuit graph* of an oriented matroid $\mathcal{M}$ of rank r is the graph $G^*(\mathcal{M})$ whose vertices are the signed cocircuits of $\mathcal{M}$, with an edge connecting signed cocircuits X and Y if $|X^0 \cap Y^0| \geq r-2$ and $S(X,Y) = \emptyset$. An oriented matroid is *uniform* if $|X^0| = r-1$ for every cocircuit $X \in \mathcal{C}^*$. If X and Y are signed cocircuits in $\mathcal{M}$, we use $d_{\mathcal{M}}(X,Y)$ to denote the distance from X to Y in $G^*(\mathcal{M})$; that is, the length of the shortest path from X to Y in $G^*(\mathcal{M})$. We call a path P from X to Y *crabbed* (introduced in [**MBS06, KMBS14**]), if for every cocircuit $W \in P$, $W^+ \subseteq X^+ \cup Y^+$ and $W^- \subseteq X^- \cup Y^-$. The *diameter* of $G^*(\mathcal{M})$ is defined as $\mathrm{diam}(G^*(\mathcal{M})) = \max\{d_{\mathcal{M}}(X,Y) : X, Y \in \mathcal{C}^*(\mathcal{M})\}$.

An important family of oriented matroids, *realizable* oriented matroids, are given by *hyperplane arrangements*. In this case, the cocircuit graph is just the one-skeleton of the cell complex obtained by intersecting a central hyperplane arrangement with a unit sphere. In general, the cocircuit graph is the graph of a combinatorial manifold and it has a rich structure. Later we review this geometry in some detail.

For the ease of notation, let $OM(n,r)$ be the set of all oriented matroids of rank r whose ground set has cardinality n. Let $UOM(n,r)$ be the set of all uniform oriented matroids in $OM(n,r)$. Let $\Delta(n,r)$ denote the maximal diameter of $G^*(\mathcal{M})$ as $\mathcal{M}$ ranges over $OM(n,r)$.

General characterizations and properties of oriented matroids and their cocircuit graphs have been explored by several authors: While it is known that the cocircuit graph does not uniquely determine the oriented matroid (see [**CF93, CFGdO00**]), labeled cocircuit graphs can be characterized (see [**BFF01, CFGdO00**]). Other topics of research have been the connectivity of the cocircuit graph (see [**CF93, BLVS$^+$99**]) and how the cocircuit graph could define the entire oriented matroid and discussed the connectivity of the graph (see [**FGK$^+$11, KMBS14, MBS06**]). In this article we are interested instead in bounding the *diameter of the cocircuit graph* of an oriented matroid. We recall that the diameter of a graph is the largest distance between a pair of its vertices, where the distance between two vertices is the length of a shortest path connecting them. Oriented matroids are combinatorial structures that generalize many types of objects, including point configurations, vector configurations, hyperplane arrangements, polyhedra, linear programs, and directed graphs. Oriented matroids have played a key role in combinatorics, geometry, and optimization (see the book by Björner et al. [**BLVS$^+$99**] for a complete treatise and Chapter 6 of Ziegler [**Zie95**] for a quick introduction we assume in the rest).

We now introduce the geometric intuition that accompanies with the definitions of oriented matroids. Let $E = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \mathbb{R}^r$ be any set of vectors. For simplicity, we will assume that E spans $\mathbb{R}^r$. We will not make a distinction between E as a set of vectors or E as a matrix in $\mathbb{R}^{r\times n}$. In classical matroid theory, we consider the set of linear dependences among the vectors in E. In oriented matroid theory, we consider not only the set of linear dependences on E, but also the signs of the coefficients that make up these dependences. To any linear dependence $\sum_{i=1}^n z_i \, \mathbf{v}_i = \mathbf{0}$ we associate a *signed vector* $(\mathrm{sign}(z_i))_{i=1}^n$. The *sign* of a number $z \in \mathbb{R}$, denoted $\mathrm{sign}(z) \in \{+, -, 0\}$, encodes whether z is positive, negative, or equal to 0. If $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{R}^n$ is a vector, we use $\mathrm{sign}(\mathbf{z})$ to denote the vector of signs: $\mathrm{sign}(\mathbf{z}) := (\mathrm{sign}(z_i))_{i=1}^n \in \{+, 0, -\}^n$. We define the set of *signed vectors* on E as

$$\mathcal{V}(E) = \{\mathrm{sign}(\mathbf{z}) : \mathbf{z} \text{ is a linear dependence on } E\}.$$

In other words, $\mathcal{V}(E) = \{\mathrm{sign}(\mathbf{z}) : E\mathbf{z} = \mathbf{0}\}$.

Among all signed vectors determined by linear dependences on E, those with minimal (and nonempty) support under inclusion, are called the *signed circuits* of E. The set of such signed circuits is denoted $\mathcal{C}(E)$.

Dually, for any $\mathbf{c} \in \mathbb{R}^r$, we can consider the *signed covector* $\left(\mathrm{sign}(\mathbf{c}^T \, \mathbf{v}_i)\right)_{i=1}^n$. The set of all signed covectors on E is

$$\mathcal{V}^*(E) = \{\mathrm{sign}\left(\mathbf{c}^T E\right) : \mathbf{c} \in \mathbb{R}^r\}.$$

The set of signed covectors of minimal, nonempty support are called *signed cocircuits* and are denoted by $\mathcal{C}^*(E)$. It is important to note that if X is a cocircuit, then so is $-X$.

Summarily, to any collection of vectors $E \subseteq \mathbb{R}^r$, there are four sets of vectors that encode dependences among E. Those are the signed vectors $\mathcal{V}(E)$ arising from linear dependences, the signed circuits $\mathcal{C}(E)$ arising from minimal linear dependences, the signed covectors $\mathcal{V}^*(E)$ arising from valuations of linear functions, and signed cocircuits $\mathcal{C}^*(E)$ arising from linear valuations of minimal support. The first fundamental result in oriented matroid theory shows that any one of these sets is sufficient to determine the other three [**Zie95**, Corollary 6.9]. Any oriented matroid that arises from a collection of signed cocircuits in this way is called a *realizable* oriented matroid.

Now we are ready to motivate the definition of oriented matroids through a geometric model that proves to be more useful than the axiomatic definition. Let $E = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \mathbb{R}^r$ be a collection of vectors, and let $\mathcal{M}(E)$ be the oriented matroid determined by E. To each vector $\mathbf{v}_i$, there is an associated hyperplane $H_i := \{\mathbf{x} \in \mathbb{R}^r : \mathbf{x}^T \mathbf{v}_i = 0\}$. Each H_i is naturally oriented by taking $H_i^+ := \{\mathbf{x} \in \mathbb{R}^r : \mathbf{x}^T \mathbf{v}_i > 0\}$ and defining H_i^- analogously.

Therefore, the vectors in E determine a central hyperplane arrangement $\mathcal{H}$ in $\mathbb{R}^r$. Any vector $\mathbf{x} \in \mathbb{R}^r$ has an associated sign vector determined by its position relative to the hyperplanes in $\mathcal{H}$. These signs can be computed as $\operatorname{sign}(\mathbf{x}^T \mathbf{v}_i)$ for each i; in other words, by computing $\operatorname{sign}\left(\mathbf{x}^T E\right)$. Therefore, the signed covectors of $\mathcal{M}(E)$ are in bijection with the regions of the hyperplane arrangement $\mathcal{H}$.

Further, because $\operatorname{sign}\left(\mathbf{x}^T E\right) = \operatorname{sign}\left((c\mathbf{x})^T E\right)$ for any positive scalar c, no information from $\mathcal{H}$ is lost if we intersect $\mathcal{H}$ with the unit sphere $\mathbb{S}^{r-1}$, giving a collection of codimension-one spheres $\{s_i = H_i \cap \mathbb{S}^{r-1} : H_i \in \mathcal{H}\}$. This induces a cell decomposition of $\mathbb{S}^{r-1}$ whose nonempty faces correspond to covectors of $\mathcal{M}(E)$ and whose vertices correspond to cocircuits of $\mathcal{M}(E)$. The regions corresponding to covectors of maximal support are called *topes*. An example is illustrated in Figure 1.4. In that figure, the cocircuit X is encoded by the sign vector $(+, +, 0, -, 0)$. Similarly, the shaded region (a tope) corresponds to the covector $(+, +, +, -, +)$.

Not all matroids can be oriented. Determining whether a matroid is orientable is an NP-complete problem, even for fixed rank (see [**RG99**]). But, a topological model provides the "right" intuition for visualizing arbitrary oriented matroids. Every oriented matroid can be viewed as an arrangement of equators on a sphere, as in the realizable case, provided that one is allowed to slightly perturb the spheres determined by $H_i \cap \mathbb{S}^{r-1}$ in the following way.

Let Q be an equator of $\mathbb{S}^{r-1}$; that is, the intersection of $\mathbb{S}^{r-1}$ with some $(r-1)$-dimensional subspace of $\mathbb{R}^r$. If $\varphi : \mathbb{S}^{r-1} \to \mathbb{S}^{r-1}$ is a homeomorphism, then the image of the equator $\varphi(Q) \subseteq \mathbb{S}^{r-1}$ is called a *pseudosphere*. Because Q decomposes $\mathbb{S}^{r-1}$ into two pieces, so too does $\varphi(Q)$. Therefore, we may define an *oriented pseudosphere* to be a pseudosphere, s, together with a choice of a positive

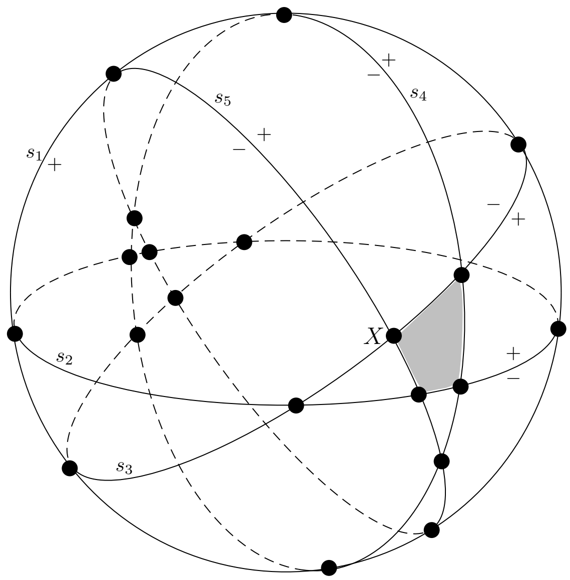

FIGURE 1.4. An oriented matroid arising from an arrangement of five hyperplanes.

side s^+ and negative side s^-. Now we may define an *arrangement* of pseudospheres in $\mathbb{S}^{r-1}$ to be a finite collection of pseudospheres $\mathcal{P} = \{s_e : e \in E\} \subseteq \mathbb{S}^{r-1}$ such that

(1) for any subset $A \subseteq E$, the set $S_A = \bigcap_{e \in A} s_e$ is a topological sphere, and
(2) if $S_A \not\subseteq s_e$ for $A \subseteq E$ and $e \in E$, then $S_A \cap s_e$ is a pseudosphere in S_A with two parts, $S_A \cap s_e^+$ and $S_A \cap s_e^-$.

A pseudosphere arrangement is *essential* if $\bigcap_{e \in E} s_e = \emptyset$. Any essential pseudosphere arrangement $\mathcal{P}$ induces a regular cell decomposition on $\mathbb{S}^{r-1}$. Because each pseudosphere in $\mathcal{P}$ has a positive and negative side, the cells of this decomposition are naturally indexed by sign vectors in $\{+,-,0\}^E$. We use $\Gamma(\mathcal{P})$ to denote the poset of such sign vectors, ordered by face containment. We have encountered this same (abstract) poset before as $\mathcal{L}(\mathcal{V}^*)$ in the introduction, the poset induced over the set of covectors $\mathcal{V}^*$ of an oriented matroid. As it turns out the following theorem of Folkman and Lawrence gives an exact correspondence between oriented matroids and pseudosphere arrangements. The same sets of sign vectors appear in both cases.

THEOREM 1.1.3. (Topological Representation Theorem [**FL78**])

Let $\mathcal{P}$ be an essential arrangement of pseudospheres in $\mathbb{S}^{r-1}$. Then $\Gamma(\mathcal{P}) \cup \{\mathbf{0}\}$ is the set of covectors of an oriented matroid of rank r. Conversely, if $\mathcal{V}^*$ is the set of covectors of a loopless *oriented matroid of rank r, then there exists an essential arrangement of pseudospheres $\mathcal{P}$ on $\mathbb{S}^{r-1}$ with* $\Gamma(\mathcal{P}) = \mathcal{V}^* \setminus \{\mathbf{0}\}$.

If $\mathcal{M}$ is an oriented matroid, the pseudosphere arrangement $\mathcal{P}$ guaranteed by the Topological Representation Theorem is called the *Folkman-Lawrence representation* of $\mathcal{M}$. Two elements $e, f \in E$ are *parallel* if $X_e = X_f$ for all $X \in \mathcal{V}^*$ or $X_e = -X_f$ for all $X \in \mathcal{V}^*$. Note that we can eliminate parallel elements without changing the pseudosphere arrangement $\mathcal{P}$.

REMARK 1.1.4. Let $\mathcal{M}$ be a uniform oriented matroid of rank r. If $A \subseteq E(\mathcal{M})$ is any set with $|A| \leq r-1$, then $S_A = \bigcap_{e \in A} s_e$ is an $(r-1-|A|)$-dimensional pseudosphere in the Folkman-Lawrence representation $\mathcal{P}(\mathcal{M})$.

Let $\mathcal{M}$ be an oriented matroid of rank r, and let $\mathcal{P}$ be the Folkman-Lawrence representation of $\mathcal{M}$. Then the underlying graph of $\mathcal{P}$ (as a cell complex) is the cocircuit graph $G^*(\mathcal{M})$. This provides a geometric model for visualizing cocircuit graphs of oriented matroids. A *coline* in $\mathcal{M}$ is a one-dimensional sphere in the Folkman-Lawrence representation of $\mathcal{M}$. In matroidal language, a coline is a covector that covers a cocircuit in the natural component-wise partial order where $0 < +, -$. For a uniform oriented matroid of rank r, a coline is a covector U with $|U^0| = r - 2$. Further, in a uniform oriented matroid, for each subset $S \in \binom{[n]}{r-2}$, there exists a coline U with $U^0 = S$. The graph of any coline is a simple cycle of length $2(n - r + 1)$.

The Folkman-Lawrence representation gives us a more concrete topological understanding of the following operations on oriented matroids. Let $\mathcal{M}$ be an oriented matroid on ground set E with signed covectors $\mathcal{V}^*(\mathcal{M})$, and let $A \subseteq E$. The *restriction* of a sign vector $X \in \{+, -, 0\}^E$ to A is the sign vector $X|_A \in \{+, -, 0\}^A$ defined by $(X|_A)_e = X_e$ for all $e \in A$. The *deletion* $\mathcal{M} \setminus A$ is the oriented matroid with covectors

$$\mathcal{V}^*(\mathcal{M} \setminus A) = \{X|_{E \setminus A} \ : \ X \in \mathcal{V}^*(\mathcal{M})\} \subseteq \{+, -, 0\}^{E \setminus A}.$$

The *contraction* $\mathcal{M} / A$ is the oriented matroid with covectors

$$\mathcal{V}^*(\mathcal{M} / A) = \{X|_{E \setminus A} \ : \ X \in \mathcal{V}^*(\mathcal{M}), A \subseteq X^0\} \subseteq \{+, -, 0\}^{E \setminus A}.$$

The fact that $\mathcal{M} \setminus A$ and $\mathcal{M} / A$ are oriented matroids is proved in [**BLVS$^+$99**, Lemma 4.1.8].

The deletion $\mathcal{M}\backslash A$ is also referred to as the restriction of $\mathcal{M}$ to $E\backslash A$. Geometrically, $\mathcal{M}\backslash A$ is the oriented matroid of the same rank as $\mathcal{M}$ obtained by removing pseudospheres $\{s_e : e \in A\}$. The contraction $\mathcal{M}/A$ is the oriented matroid obtained by intersecting S_A with $\{s_e : e \in E\backslash A\}$.

Note also that the pseudosphere arrangement of an oriented matroid of rank r lies on the sphere $\mathbb{S}^{r-1}$. The *topes* correspond to the regions, homeomorphic to balls of dimension $r-1$, that partition the sphere. For realizable oriented matroids coming from a hyperplane arrangement, topes are actual convex polytopes.

Given a tope $\mathcal{T}$ of an oriented matroid $\mathcal{M}$, we define its graph as the subgraph of $G^*(\mathcal{M})$ induced by the cociruits of $\mathcal{M}$ in $\mathcal{T}$. Next, we show the graph of a tope $\mathcal{T}$ in a uniform oriented matroid $\mathcal{M}$ of rank r on n elements, is isomorphic to a graph of an abstract polytope of dimension $r-1$ on n elements. Abstract polytopes, an abstraction of simple polytopes, were introduced by Adler and Dantzig [**AD74**] for the purpose of studying the diameter of their graphs. Abstract polytopes have been further generalized in recent years by several authors (see [**EHRR10, San13**] and references there for details).

Definition 1.1.5. Let T be a finite set. A family $\mathcal{A}$ of subsets of T (called *vertices*) forms a *d-dimensional abstract polytope on the ground set T* if the following three axioms are satisfied:

(i) Every vertex of $\mathcal{A}$ has cardinality d.

(ii) Any subset of $d-1$ elements of T is either contained in no vertices of $\mathcal{A}$ or in exactly two (called neighbors or adjacent vertices).

(iii) Given any pair of distinct vertices $X, Y \in \mathcal{A}$, there exists a sequence of vertices $X = Z_0, Z_1, \ldots, Z_k = Y$ in $\mathcal{A}$ such that

   (a) Z_i, Z_{i+1} are adjacent for all $i = 0, 1, \ldots, k-1$, and

   (b) $X \cap Y \subset Z_i$ for all $i = 0, 1, \ldots, k$.

The graph $G_{abs}(\mathcal{A})$ of an abstract polytope $\mathcal{A}$ is composed of nodes corresponding to its vertices, where two vertices are adjacent on the graph as specified in axiom (ii).

Consider a simple polytope $\mathcal{P}$ of dimension d which is the intersection of n facet-defining half-spaces. Then, indexing the n facets by $1, \ldots, n$, the family of all sets of indices that define a vertex of $\mathcal{P}$ is an abstract polytope of dimension d on the ground set $\{1, \ldots, n\}$. In particular, the three axioms of abstract polytopes state that the graph $G(\mathcal{P})$ associated with the vertices of $\mathcal{P}$ has the following three properties:

(i) $G(\mathcal{P})$ is regular of degree d (as all the hyperplanes corresponding to the half-spaces are in general position.)

(ii) All edges of $G(\mathcal{P})$ have two vertices as end points (as $\mathcal{P}$ is bounded).

(iii) For any two vertices X, Y that lie in a face F of $\mathcal{P}$, there exists a path between the nodes corresponding to X and Y on $G(\mathcal{P})$ composed entirely of nodes corresponding to vertices on F (as F is also a polytope.)

Interestingly, while the axioms of abstract polytopes represent only three basic properties related to graphs of simple polytopes, a substantial number of the results related to diameter of simple polytopes in [**KW67**] have been proved in [**AD74**] for abstract polytopes.

Next, we show that these properties are satisfied by the graph of topes of uniform oriented matroids. The graph of a tope, its connectivity, and the relation to pseudomanifolds has been studied in [**CF93**].

LEMMA 1.1.6. *Given a uniform oriented matroid $\mathcal{M} = (E, \mathcal{C}^*)$ of rank $r \geq 2$ and a tope $\mathcal{T}$ of $\mathcal{M}$, let*

$$\mathcal{C}_{\mathcal{T}} = \{X \in \mathcal{C}^* : X < \mathcal{T}\}, \text{ and } \mathcal{A} = \{X^0 : \; X \in \mathcal{C}_{\mathcal{T}}\}.$$

Then, $\mathcal{A}$ is a d-dimensional abstract polytope on the ground set E, where $d = r - 1$. Moreover, the graph $G(\mathcal{T})$ of $\mathcal{T}$ is isomorphic to the graph $G_{abs}(\mathcal{A})$ of $\mathcal{A}$.

PROOF. We show that $\mathcal{A}$ satisfies the three axioms of abstract polytopes:

(i) Axiom (i) holds because $\mathcal{M}$ is a uniform oriented matroid of rank r.

(ii) Let $E' \subset E$ such that $|E'| = d-1$, and assume that there exists $X \in \mathcal{C}_{\mathcal{T}}$ such that $E' \subset X^0$ (otherwise, no vertex of $\mathcal{A}$ contains E' and we are done). Let $U = \{W \in \mathcal{M}^* : E' \subset W^0\}$, then U is a coline of $\mathcal{M}$ whose graph is a simple cycle. Let Y_1, Y_2 be the two adjacent cocircuits to X in U. Then, there exists an element $e \in E \setminus E'$ such that $S(Y_1, Y_2) = e$ and $S(X, Y_i) = \emptyset$ $(i = 1, 2)$, implying that exactly one of Y_1, Y_2, say Y_1, is in $\mathcal{T}$. However, no other cocircuit in U is in $\mathcal{T}$. Suppose, to the contrary, that there exists $Z \in U$, distinct from X and Y_1 that belongs to $\mathcal{T}$. Then by definition

$$|X^0 \cap Y_1^0| = |X^0 \cap Z^0| = |Y_1^0 \cap Z^0| = d-1, \text{ and } S(X,Y_1) = S(X,Z) = S(Y_1,Z) = \emptyset.$$

This means that X, Y_1, and Z, are all adjacent on U. As the graph of U is a simple cycle of size $2(n-r+1)$, this leads to contradiction.

(iii) By [**FGK$^+$11**, Theorem 2.3], for any $X, Y \in \mathcal{C}^*$ there exists an (X,Y) crabbed path on $G^*(\mathcal{M})$. That is, there exists a path $X = Z_0, Z_1, \ldots, Z_k = Y$ on $G^*(\mathcal{M})$ such that $Z_i^+ \subseteq X^+ \cup Y^+$ and $Z_i^- \subseteq X^- \cup Y^-$ for all $0 \leq i \leq k$. This implies that if $X, Y \in \mathcal{C}_{\mathcal{T}}$, then for $i = 1, \ldots, k-1$, $Z_i \in \mathcal{C}_{\mathcal{T}}$ (as $Z_i < \mathcal{T}$), so $Z_i^0 \in \mathcal{A}$, and $X^0 \cap Y^0 \subseteq Z_i^0$. Now, let $G(\mathcal{T})$ be the graph of $\mathcal{T}$. Note that as $S(X,Y) = \emptyset$ for any $X, Y \in \mathcal{C}_{\mathcal{T}}$, X and Y share an edge on $G(\mathcal{T})$ if and only if $|X^0 \cap Y^0| = d-1$. However, the two vertices on $G_{abs}(\mathcal{A})$ corresponding to X^0, Y^0 are adjacent if and only if $|X^0 \cap Y^0| = d-1$. Thus, we conclude that $G(\mathcal{T})$ is isomorphic to $G_{abs}(\mathcal{A})$, so Axiom (iii) is satisfied.

Note that by the proof of part (iii) above we have that the graph $G(\mathcal{T})$ of $\mathcal{T}$ is isomorphic to the graph $G_{abs}(\mathcal{A})$ of $\mathcal{A}$. □

1.2. Linear Programming and the Simplex Method

This dissertation is mostly motivated by trying to understand the simplex method, and thus we introduce some standard definitions related to linear programming and the simplex algorithm. This introduction is meant to be brief, and we refer to textbooks (see [**Sch98, Dan98**]) for more extensive background knowledge. For the remainder of this section we let $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^n$ be given.

Linear programming (LP) is a method for maximizing or minimizing a linear objective function with respect to the constraints, which are a system of linear equations or inequalities. It can be written in the *canonical form* as:

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

.

Each constraint can be viewed as a generalized half-space, so the constraints of the LP form a convex polyhedron.

The *simplex method*, invented and developed by George Bernard Dantzig in 1947 (see the book [**Dan90**] for more information on the origins of of the simplex method), changed the field of optimization dramatically. When introduced, the simplex method became the basis of many further branches of LP, such as integer linear programming and nonlinear programming. Even today, the simplex method can outperform and compete with more recent algorithms [**Tod11**].

We now briefly recall how the simplex method works. We use the notation $[n] := \{1, 2, \ldots, n\}$ for any positive integer n. We say that $B \subseteq [n]$ with $|B| = m$ is a *basis* if and only if the columns of $\mathbf{A}_B$ are linearly independent, or equivalently $\mathbf{A}_B$ is non-singular. We say $\mathbf{x}_B$ a *basic feasible solution* with basis B if $\mathbf{A}\mathbf{x}_B = \mathbf{b}, \mathbf{x}_B \geq 0$ and for all $j \notin B$: $x_j = 0$. The vector of *reduced costs* for a basis B is defined as

$$\mathbf{z}^B = \mathbf{c} - \mathbf{A}\mathbf{A}_B^{-1}\mathbf{c}_B.$$

We say $j \in [n]$ is an *improving pivot* with respect to B if and only if $\mathbf{z}_j^B > 0$.

With the definition of an improving pivot, the simplex method can be summarized as a process of starting with a feasible basis B and updating with improving pivots until no such improving pivots exist. When performing the simplex method, we usually have more than one possible improving edges during one iteration. Thus, we need to use a *pivot rule* to tell us which edge the simplex method will go along. Note that when the LP problem is *degenerate*, the simplex method may run into a cycle, which means that the simplex method will revisit a vertex. However, if a “good” pivot rule is chosen, the simplex method does not cycle and terminates at the optimum. Examples of such rules include Bland’s rule [**Bla77**].

There are dozens of famous pivot rules [**TZ93, DL11**], but we present three basic pivot rules that our experiment consider:

(1) *Dantzig:* this rule was suggested by Dantzig [**Dan98**]. In every iteration Dantzig's rule picks the non-basic variable with the largest positive reduced cost to be the entering variable.
(2) *Greatest Improvement:* this rule picks the improving pivot that results in the largest increment of the objective function.
(3) *Steepest edge:* this rule performs the improving pivot with the largest rate of increment of objective function per distance traveled along the improving edge.

In the following example we briefly explain how different pivot rules choose different pivots using tableaux.

$$
\begin{array}{c}
\text{Variables} \\ z \\ w_1 \\ w_2 \\ w_3
\end{array}
\left[\begin{array}{c|ccccccc|c}
z & x_1 & x_2 & x_3 & w_1 & w_2 & w_3 & b \\
\hline
1 & -5 & -4 & -3 & 0 & 0 & 0 & 0 \\
\hline
0 & 2 & 3 & 1 & 1 & 0 & 0 & 5 \\
0 & 4 & 1 & 2 & 0 & 1 & 0 & 11 \\
0 & 3 & 4 & 2 & 0 & 0 & 1 & 8
\end{array}\right]
$$

We can see that x_1, x_2, x_3 have negative coefficients and are the potential entering variables. For Dantzig's rule, we pick x_1. For greatest improvement, we pick x_2 since the increment of objective function by each variable is $x_1 : 12.5, x_2 : \frac{20}{3}, x_3 : 12$. And for steepest edge, we pick x_3 since the rate of each variable is $x_1 : \frac{5}{\sqrt{30}}, x_2 : \frac{4}{\sqrt{27}}, x_3 : \frac{3}{\sqrt{10}}$.

We study the pivoting strategies for primal simplex algorithm implemented in DOcplex [**oCt**]. These include Dantzig's rule, hybrid (DOcplex's default), greatest improvement, steepest edge and devex. Hybrid is a pivot rule DOcplex implemented as default, which uses Dantzig's rule in the earlier iterations when there are a lot of choices of improving pivots and switch to steepest edge later. Devex is an approximate version of steepest edge developed by P. Harris [**Har73**]. DOcplex also implemented a steepest edge with slack initial norms, which is slightly cheaper in computation. But in our testing, it usually is not better than steepest edge. As a consequence it was not included in the algorithm portfolio.

1.3. Results in Directed Polytope Graphs

The main questions addressed in this section ask to determine:

- the minimum and maximum number of f-arborescences on P,
- the minimum and maximum number of f-monotone paths on P, and
- the minimum and maximum diameter of the graph $G(P, f)$,

where P ranges over all convex polytopes of given dimension and number of vertices and f ranges over all generic linear functionals on P. We will also consider these (or similar) questions when P is restricted to the important class of simple polytopes.

There are good reasons, from both a theoretical and an applied perspective, to study these problems. One motivation comes from the connection of f-arborescences and f-monotone paths to the behavior of the simplex method [**Sch86**]. The simplex method produces a partial f-monotone path, traversing $\omega(P, f)$ from an initial vertex to the optimal one. The simplex method has to make decisions to choose the improving arcs via a *pivot rule*. It is an open problem to find the longest possible simplex method paths and little is known about bounds (see [**BDLL21**] and references therein). Clearly, the lengths of f-monotone paths are of great interest, as they bound the number of steps in the simplex algorithm. A pivot rule gives a mapping from the set of instances of the algorithm to the set of f-arborescences of $\omega(P, f)$. Two pivot rules are equivalent if they always produce the same f-arborescence. Therefore, given P and f, there are only finitely many equivalence classes of pivot rules and counting f-arborescences is a proxy for the problem of counting pivot rules. See also [**BDLLS22**] for a recent found geometric structure on pivot rules.

Another motivation comes from enumerative and polyhedral combinatorics, especially from the theory of fiber polytopes [**BS92**]. The flip graph of f-monotone paths on P contains a well behaved subgraph, namely that induced on the set of *coherent* f-monotone paths (these are the monotone paths which come from the shadow vertex pivot rule [**DH16**]). This subgraph is isomorphic to the graph of a convex polytope of dimension $d - 1$, where $d = \dim(P)$, which is a fiber polytope known as a *monotone path polytope* [**BS92**, Section 5] [**BKS94**]. Monotone paths, monotone path polytopes and flip graphs of polytopes of combinatorial interest often have elegant combinatorial interpretations. For example, the monotone path polytope of a cube is a permutohedron [**BS92**, Example 5.4], while the flip graph of the latter encodes reduced decompositions of a certain permutation and the braid relations among them [**BLVS$^+$99**, Section 2.4]. More generally, monotone paths on zonotopes [**AS01, RR13**] correspond to certain galleries of chambers in a central hyperplane arrangement and the problem to estimate the diameter of the flip graph in this important special case has been intensely studied in [**Edm15, EJLM18, RR13**]. The diameter of flip graphs of fiber polytopes has also been studied in [**Pou14, Pou17**]. Moreover, certain zonotopes are in fact monotone path polytopes coming from projecting cyclic polytopes [**ADLRS00**, Section 3], or polytopes which look like piles of cubes [**Ath99**]. Monotone path polytopes are also related to fractional power series solutions of algebraic equations [**McD95**]. The combinatorial properties of f-monotone paths and flip graphs have thus been studied in comparison to those of coherent f-monotone paths, but also because of their own independent interest.

A special role in our results is played by a distinguished member $X(n)$ of the family of stacked 3-dimensional simplicial polytopes with n vertices. As it turns out, this polytope maximizes the number of both f-arborescences and f-monotone paths, and possibly the diameter of the flip graph too, in this dimension. We refer to Section 1.1.3 for a discussion of stacked polytopes and the precise definition of $X(n)$, which we always consider endowed with the specific LP-allowable orientation given there. We will typically denote by n (and sometimes by $n+1$) and m the number of vertices and facets of P, respectively. Let us also denote by

- $\tau(P, f)$ the number of f-arborescences on P,
- $\mu(P, f)$ the number of f-monotone paths on P,
- $\operatorname{diam}(G)$ the diameter of the graph $G = G(P, f)$.

Our first two main results provide a fairly complete description of tight bounds for the numbers of f-arborescences and f-monotone paths and the diameter of the graph of f-monotone paths on a 3-dimensional polytope with given number of vertices. The upper bound for the number of f-monotone paths involves the sequence of Tribonacci numbers (sequence A000073 in [**Slo**]), defined by the recurrence $T_0 = T_1 = 1$, $T_2 = 2$ and $T_n = T_{n-1} + T_{n-2} + T_{n-3}$ for $n \geq 3$.

THEOREM 1.3.1. *For $n \geq 4$,*

$$2(n-1) \;\leq\; \tau(P,f) \;\leq\; 2 \cdot 3^{n-3} \tag{1.3.1}$$

$$\left\lceil \frac{n}{2} \right\rceil + 2 \;\leq\; \mu(P,f) \;\leq\; T_{n-1} \tag{1.3.2}$$

for every 3-dimensional polytope P with n vertices and every generic linear functional f on P. The upper bound is achieved by the stacked polytope $X(n)$ in both situations.

The lower bound of (1.3.1) can be achieved by pyramids and that of (1.3.2) by prisms, when n is even, and by wedges of polygons over a vertex, when n is odd. In particular, prisms minimize the number of f-monotone paths over all simple 3-dimensional polytopes with given number of vertices. Moreover,

$$\tau(P,f) \;=\; 3 \cdot 2^{(n-2)/2} \;=\; 3 \cdot 2^{m-3}$$

for every 3-dimensional simple polytope P with n vertices and m facets.

THEOREM 1.3.2. *For every $n \geq 4$,*

$$\lceil \frac{(n-2)^2}{4} \rceil \leq \max \operatorname{diam} G(P,f) \leq (n-2) \lfloor \frac{n-1}{2} \rfloor, \tag{1.3.3}$$

where P ranges over all 3-dimensional polytopes with n vertices and f ranges over all generic linear functionals on P.

Our results are substantially weaker in dimensions $d \geq 4$ and leave plenty of room for further research. The upper bounds for the number of f-arborescences and the number of f-monotone paths are almost trivial, but are included here for the sake of completeness.

THEOREM 1.3.3. *(a) For $n > d \geq 4$,*

$$\tau(P,f) \leq (n-1)!$$

$$\mu(P,f) \leq 2^{n-2}$$

for every d-dimensional polytope P with n vertices and every generic linear functional f on P. These bounds are achieved by any 2-neighborly d-dimensional polytope with n vertices.

(b) For $m > d \geq 4$,

$$d \cdot ((d-1)!)^{m-d} \leq \tau(P,f) \leq \prod_{i=1}^{\lfloor \frac{d}{2} \rfloor} i^{\binom{m-d+i-1}{i}} \prod_{i=\lfloor \frac{d+1}{2} \rfloor}^{d} i^{\binom{m-i-1}{d-i}}$$

for every simple d-dimensional polytope P with m facets and every generic linear functional f on P. The lower and upper bounds are achieved by the polar duals of stacked simplicial polytopes and the polar duals of neighborly simplicial polytopes, respectively, of dimension d with m vertices.

The proofs of the results on f-arborescences, given in Section 2.2, rely on the fact that $\tau(P,f)$ is equal to the product of the outdegrees of the vertices of the directed graph $\omega(P,f)$ other than the sink (see Proposition 2.2.1). This has the curious consequence that $\tau(P,f)$ is independent of f for every simple polytope P. The proofs of the results on f-monotone paths and the diameter of flip graphs, given in Sections 2.3 and 2.4, respectively, use ideas from [**AER00**, Section 4] [**BKS94**], reviewed in Section 2.1, to construct $G(P,f)$ as an inverse limit in the category of graphs and simplicial maps.

1.4. Results in the Cocircuit Graphs of Oriented Matroids

The motivation for our investigations is again the complexity of the *simplex method* [**BT97, Sch86**] and of the *criss-cross method* [**FT97, FT99**]. Both algorithms are pivoting methods that jump from cocircuit to cocircuit along edges of the cocircuit graph. Bounds on the diameter are relevant for understanding their running time. The following conjecture is the oldest and the most ambitious open challenge about the diameter of oriented matroids today. It motivated a big part of this thesis.

CONJECTURE 1.4.1. *Let $\mathcal{M}$ be an oriented matroid of rank r on n elements, and let $G^*(\mathcal{M})$ be its cocircuit graph. Then* $\operatorname{diam}(G^*(\mathcal{M})) \leq n - r + 2$.

Prof. K. Fukuda (personal communication) kindly informed us that Conjecture 1.4.1 is an old folklore problem that goes back at least 25 years. We hope to revive interest in this conjecture with this article. Conjecture 1.4.1 bears a striking resemblance to the famous Hirsch conjecture for convex polytopes. Let $P \subseteq \mathbb{R}^d$ be a d-polytope defined by n hyperplane inequalities. Lifting P to $\mathbb{R}^{d+1}$ (and setting $r = d+1$) determines a central hyperplane arrangement in $\mathbb{R}^r$, one of whose cones is the nonnegative span of P. Therefore, P gives rise to an oriented matroid $\mathcal{M}$ whose cocircuit graph contains the graph of P as an induced subgraph (see Figure 1.5). Substituting $r = d + 1$ in

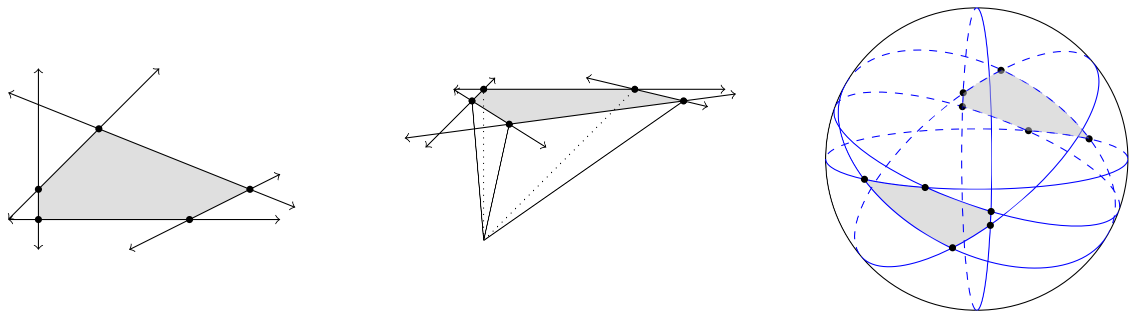

FIGURE 1.5. A polytope in $\mathbb{R}^2$ (left), its lifting to $\mathbb{R}^3$ (center), and the intersection with the resulting hyperplane arrangement on $\mathbb{S}^2$ (right).

Conjecture 1.4.1 gives an upper bound of $n - r + 2 = n - d + 1$, which differs from the conjectured Hirsch bound by 1. The reason for this is that each signed cocircuit X has an antipodal cocircuit $-X$. We will see later that when $\mathcal{M}$ is uniform, the distance between antipodal cocircuits is exactly $n - r + 2$.

Conjecture 1.4.1 has appeared in the literature in several forms. Babson, Finschi, and Fukuda [**BFF01**, Lemma 6] established Conjecture 1.4.1 for uniform oriented matroids of rank 2 and rank 3, showing further that only antipodal cocircuits can have distance $n-r+2$. Felsner et al. [**FGK$^+$11**, Lemma 4.1] also showed that the conjecture is true for uniform oriented matroids with rank at most 3 and stated again the famous Conjecture 1.4.1 in [**FGK$^+$11**, Question 4.2] with a strong emphasis on the important role of antipodal cocircuits. Finschi [**Fin01**, Open Problem 5] asked whether $\mathrm{diam}(G^*(\mathcal{M})) \leq c \cdot n$ for some constant c that is independent of n and r. Aside from the results of Babson, Finschi, and Fukuda in low rank, the most general progress that has been made on Conjecture 1.4.1 seems to have come from Finschi's Ph.D thesis.

THEOREM 1.4.2. (Finschi [**Fin01**, Proposition 2.6.1])

Let $\mathcal{M}$ be a uniform oriented matroid of rank r on n elements. Then

$$\mathrm{diam}(G^*(\mathcal{M})) \leq n - r + 2 + \sum_{k=1}^{\min(r-2,n-r)} \left(\left\lfloor \frac{n-r-k}{2} \right\rfloor + 1 \right).$$

The bound in Theorem 1.4.2 is tight when $r = 2$ or $r = n$, but in general it is not.

LEMMA 1.4.3. *Let $\mathcal{M}$ be an oriented matroid of rank r on n elements. Then there exists a uniform oriented matroid $\mathcal{M}'$ of rank r on n elements such that*

$$\mathrm{diam}(G^*(\mathcal{M})) \leq \mathrm{diam}(G^*(\mathcal{M}')).$$

Moreover, when $\mathcal{M}$ is realizable, then $\mathcal{M}'$ can be taken to be realizable as well.

Lemma 1.4.3, reduces Conjecture 1.4.1 to studying uniform oriented matroids. Therefore, for the purposes of studying Conjecture 1.4.1, it suffices to consider only uniform oriented matroids.

The following auxiliary lemma shows that the discrepancy between the diameter given in Conjecture 1.4.1 and the classical Hirsch bound cannot be improved. Essentially, Conjecture 1.4.1 cannot be improved because the distance between antipodal cocircuits is exactly $n - r + 2$.

LEMMA 1.4.4. *Let $\mathcal{M}$ be a uniform oriented matroid of rank r on n elements, and let $X, Y \in \mathcal{C}^*(\mathcal{M})$. Then*

$$d_{\mathcal{M}}(X,Y) \geq \begin{cases} |S(X,Y)| + |X^0 \setminus Y^0| & \text{if } X \neq -Y, \\ n - r + 2 & \text{if } X = -Y. \end{cases} \tag{1.4.1}$$

Moreover, if $|X^0 \setminus Y^0| \leq 1$, then the inequality (1.4.1) *holds with equality: $d_{\mathcal{M}}(X,Y) = 1 + |S(X,Y)|$, and in particular, when $X = -Y$, then $d_{\mathcal{M}}(X,Y) = n - r + 2$.*

We then move on to establish Conjecture 1.4.1 in low rank. Babson, Finschi, and Fukuda [**BFF01**, Lemma 6] and Felsner et al. [**FGK$^+$11**, Lemma 4.1] gave proofs of Conjecture 1.4.1 for $r \leq 3$. We present a new geometric proof of that same result (see Theorem 3.2.2). We explain why our method does not generalize for rank four matroids. Finally, we also settle the conjecture for oriented matroids of low corank (i.e. $n - r$). In summary, we have the following theorem:

THEOREM 1.4.5. *Let $\mathcal{M}$ be a uniform oriented matroid of rank r on n elements.*

a. If $n \leq 9$*, then* $\operatorname{diam}(G^*(\mathcal{M})) = n - r + 2$.

b. If $r \leq 3$*, then* $\operatorname{diam}(G^*(\mathcal{M})) = n - r + 2$.

c. If $n - r \leq 4$*, then* $\operatorname{diam}(G^*(\mathcal{M})) = n - r + 2$.

THEOREM 1.4.6. *Let* $\mathcal{M}$ *be an oriented matroid of rank* r *on* n *elements, and let* $X, Y \in \mathcal{C}^*(\mathcal{M})$ *with* $X \neq -Y$*. Then*

$$d_{\mathcal{M}}(X,Y) \leq n - r + 1 + \sum_{k=2}^{|X^0 \setminus Y^0|-1} \left(\left\lfloor \frac{n-r-k}{2} \right\rfloor + 1 \right). \tag{1.4.2}$$

In particular, when $r \geq 4$ *and* $n - r \geq 2$*,*

$$\operatorname{diam}(G^*(\mathcal{M})) \leq n - r + 1 + \sum_{k=2}^{\min(r-2,n-r)} \left(\left\lfloor \frac{n-r-k}{2} \right\rfloor + 1 \right). \tag{1.4.3}$$

This bound contrasts the best-known upper bounds on polytope diameters, which are linear in fixed dimension, but grow exponentially in the dimension (e.g., [**KK92**] and [**EHRR10**]). For a survey of the best bounds and more updates about diameters of polytopes see [**CS17, CS19, EHRR10, San13, Suk19**] and the references therein. To start, one may hope that $d_{\mathcal{M}}(X,Y) \leq n - r + 1$ when X and Y are not antipodal cocircuits. In fact, Finschi posed a similar question in his thesis [**Fin01**, Open Problem 2], as did Felsner et al. [**FGK$^+$11**, Question 4.2]. Here we answer this question. We show the answer is negative by considering Santos's counterexample to the Hirsch conjecture.

PROPOSITION 1.4.7. *There is a uniform oriented matroid* $\mathcal{M}$ *of rank 21 on 40 elements that has a pair of non-antipodal cocircuits* X *and* Y *such that* $d_{\mathcal{M}}(X,Y) \geq 21 = n - r + 2$.

It is not immediately clear whether bounds on the diameter of the cocircuit graph of a realizable oriented matroid imply bounds on polytope diameters. This possible connection has been discussed before. For example, a connection of the (original) Hirsch conjecture to Conjecture 1.4.1 was stated in Remark 4.3 of [**FGK$^+$11**]. Their proof of their remark shows that, thanks again to Santos's counterexample, there is an example of two cocircuits in the same tope that cannot be connected by a *crabbed* path of length $n - r + 1$. Their remark also suggests two natural strengthenings of

Conjecture 1.4.1, both of which would imply the polynomial Hirsch conjecture is true for convex polytopes:

First, if X and Y are vertices in a tope $\mathcal{T}$, does the shortest path from X to Y in the supergraph $G^*(\mathcal{M})$ of cocircuits leave the tope $\mathcal{T}$? The question is already interesting for a realizable $\mathcal{M}$ where a tope $\mathcal{T}$ corresponds to a polytope. If the shortest path between X, Y always stays in a tope containing both, then a quadratic bound on the diameter of polytopes follows from the quadratic bound for oriented matroids. This would prove the famous polynomial Hirsch Conjecture for those polytopes in the arrangement (recall the polynomial Hirsch conjecture states that the diameter of all convex polytopes is bounded by a polynomial in terms of the number of facets and the dimension, see [**San13**]).

Second, even more strongly, is there always a crabbed path from X to Y whose length is no bigger than the length of any path from X to Y in the entire cocircuit graph $\mathcal{M}$? Again, if this was true, the diameter computed over the topes that contain X, Y is always no larger than the diameter of the entire cocircuit graph. Unfortunately, we show the two strengthenings of Conjecture 1.4.1 are false. This is the content of Theorem 1.4.8. We used a computer search to find the counterexamples and to show they are smallest possible.

THEOREM 1.4.8. *There is a realizable rank 4 uniform oriented matroid $\mathcal{M}$ with 9 elements and a pair its cocircuits $X, Y \in \mathcal{C}^*(\mathcal{M})$, whose distance $d_{\mathcal{M}}(X, Y)$ is smaller than the length of any crabbed path from X to Y. We prove that no such example with fewer than 9 elements is possible. Moreover, by adding another element to $\mathcal{M}$, we construct a realizable rank 4 oriented matroid $\mathcal{M}'$ on 10 elements with two cocircuits X, Y inside a common tope $\mathcal{T}$, such that $d_{\mathcal{M}'}(X, Y) < d_{\mathcal{T}}(X, Y)$.*

1.5. Machine Learning for improving the simplex method

It is obvious that we are living in an era where huge amount of data have been continuously generated at increasing scales. Machine learning techniques have been widely adopted in a number of massive and complex data-intensive fields such as medicine, astronomy, biology, and so on, for these techniques provide possible solutions to mine the information hidden in the data. In this part of the thesis we explain how machine learning improves algorithms and in particular the simplex method.

What is the best way to select an algorithm? Two different algorithms for the same computational task have difference performances: one algorithm is better on some inputs, but worse on the others. Over the years there have been various theoretical frameworks answering this question. *Worst-case analysis* aims to find the extreme instances that strain the performance the most. *Average-case analysis* on the other hand assumes that input instances come from a fixed probability distribution, thus we can talk about average running time or average complexity. More recently, the *Smooth analysis* is a hybrid of the worst-case and average-case analysis of algorithms where one measures the maximum over inputs of the expected performance of an algorithm under small random perturbations of that input. The performance of many algorithms varies dramatically on the types of input one provides, thus the theoretical evaluations often say nothing useful for the non-expert user. How is a non-expert user supposed to make the right algorithmic choices when a large number of choices are possible? How can someone make reasonable consistent choices of parameters for tuning complicated algorithms?

In this thesis we consider the very famous simplex method. Researchers have found the worst-case behavior of the simplex algorithm is exponential, for most known deterministic pivot rules [**KM72, Jer73, GS79, AC09, Mur80, AZ99, Fri11**] and randomized pivot rules [**GK07, Kal97**]. On the other hand, under a specific probability distribution for input instances, the average running time of simplex algorithm is polynomial in terms of the input size [**Bor82**]. Similarly, the smooth analysis shows that the Simplex method is efficient [**DH18**]. Despite the theoretical success, neither of the three theoretical evaluations matches the *empirical performance* of the simplex method, which is known to be very fast in practice. Today the simplex method has been investigated and improved enormously from its original version [**Bix01**]. It is known that the running time or number of iterations for the Simplex method depends not just on the input data, but how we tune the algorithm itself. E.g., what choice of pivot rule shall we make? This is a question that has been answered by experts by fixing a default pivot rule, which often performs well, but may not be always the optimal choice.

In Chapter 4 we will discuss a pragmatic framework for empirical algorithm selection tuning and comparison. We demonstrate a *machine learning-based selection and tuning of algorithms*. Our framework is data-driven, empirically-based, and can help non-experts make reasonable consistent algorithmic decisions without prior knowledge of the algorithms. Users of algorithmic methods often have no knowledge of the worst-case examples, nor can they assume to know the exact distribution of their data. Users only have access to data sets. The simple principle we propose here is that, *if one has sufficiently many data instances, one can create a practical machine learning recommendation system to efficiently automate the selection of algorithms or their parameter configurations for concrete data sets, with the intention to speed up computation.*

We picked the case study of simplex method to illustrate the framework, but it would apply almost in the same way to other algorithms where the input is based on matrices. Algorithm selection has seen a strong surge in both practical and theoretical research and we only touch the fraction of the literature that we know deals with algorithm similar to our case studies (for much more we recommend [**LL00, YAKU19, GR17, BNVW17, LJD+18**] and the many references therein). Several authors have been directly concerned with algorithm selection and tuning for discrete algorithmic problems (see e.g., [**KDN+17, BDSV18, ADG+16**] and the many references therein). The papers [**BLP18, Smi99**] are great surveys of uses of learning in combinatorial optimization. In [**BS20**] the authors redefine mixed integer convex optimization problems as a multi-class classification problem where the machine learning predictor gives insights on the optimal solution. Dai et al. [**KDZ+17**] develop a method to learn heuristics over graph problems. Several authors have proposed ways to use machine learning to select the best branching rules (see [**ALW17, KDN+17**]). Machine learning methods have also been useful in aiding the selection of reformulations and decompositions for mixed-integer optimization [**BLZ18, KLP17**]. Some libraries organize data for various NP-hard tasks (where the aim is to predict how long an algorithm will take to solve concrete instances of NP-complete problems, or to choose best approximation schemes tailored by instances) [**NDSLB04, BKK+16, KHO17**]. In fact the approach we present here is a simplification of the empirical hardness model to predict the running time of algorithms applied to improve logic Satisfiability (SAT) solvers [**LHHX14, ELH18**]. There are also now a number of well-established software implementations for algorithm tuning (see [**ELH19, FKE+19**] and the many references therein).

The simplex method is widely used in solving linear programming (LP) problems. Geometrically, simplex algorithm starts on a vertex of the feasible region (which is a polytope), and generates a path via improving edges until optimum is reached. A pivot rule helps to decide which improving edge to pick if there are multiple choices. In this case, we are interested in applying different machine learning models to study and improve the choice among five pivoting rules for the simplex algorithm on linear programming based on features of different LP instances.

We demonstrate that the total performance of algorithms, when guided by Machine Learning (ML) decision-making, is clearly faster than using a single static choice for these algorithms. We implemented two ML methodologies, gradient boosting decision trees and neural networks. We tested two different schemes of predicting the fastest algorithm: *direct classification and run time prediction*. In direct classification a machine learning method is trained to predict which algorithm will run the fastest. In the run time prediction setting, a machine learning method is trained to estimate how long an algorithm will run on a particular instance, then we pick the algorithm that is expected to run the quickest. In addition, we tested different data representations and features. Our best performing model is able to choose the best pivot rule in 69.06% total instances, with 178.6 iterations on average, better than any single pivoting strategies we tested.

To end this section, we give a brief review of machine learning techniques that we used. For more detailed model definitions see The Elements of Statistical Learning [**HTF09**].

Generally machine learning is divided into supervised learning, unsupervised learning and reinforcement learning. *Supervised learning* is the most common subbranch of machine learning today. Supervised machine learning algorithms are designed to learn by example. When training a supervised learning algorithm, the training data will consist of inputs paired with the correct outputs. During training, the algorithm will search for patterns in the data that correlate with the desired outputs. After training, a supervised learning algorithm will take in new unseen inputs and will determine which label the new inputs will be classified as based on prior training data. The objective of a supervised learning model is to predict the correct label for newly presented input data. At its most basic form, a supervised learning algorithm can be written simply as $Y = f(x)$. Supervised learning can be split into two subcategories: classification and regression.

During training, a *classification* algorithm will be given data points with an assigned category. The job of a classification algorithm is to then take an input value and assign it a class, or category, that it fits into based on the training data provided. On the other hand, *regression* is a predictive statistical process where the model attempts to find the important relationship between dependent and independent variables. The goal of a regression algorithm is to predict a continuous number. A typical example of a classification task is to predict whether an incoming email is a spam or not. A regression task will be predicting sales of some products.

Decision trees create a model that predicts the label by evaluating a tree of if-then-else true/false feature questions, and estimating the minimum number of questions needed to assess the probability of making a correct decision. Decision trees can be used for classification to predict a category, or regression to predict a continuous numeric value. *Ensemble algorithms* combine multiple machine learning algorithms to obtain a better performing model. Similar to random forest, a *gradient boosting decision tree* is a decision tree ensemble learning algorithm. We use XGBoost [**xd21**] for our training.

The other training model we used in our case study is *neural networks* (or *multi-layer perceptrons*). The goal is not to create realistic models of the brain, but instead to develop robust algorithms and data structures that we can use to model difficult problems. Neural networks learn a mapping, and mathematically they are capable of learning most functions and thus being a universal approximation algorithm.

The building block for neural networks are called *neurons*. These are simple computational units that have weighted input signals and produce an output signal using an activation function. A neural network usually consists of at least three layers of neurons, an input layer, a hidden layer and an output layer. An *activation function* is a simple mapping of summed weighted input to the output of the neuron. It is called an activation function because it governs the threshold at which the neuron is activated and strength of the output signal. Non-linear functions such as the sigmoid function are often used as activation functions. Neural networks utilize back propagation, a supervised learning technique for training.

CHAPTER 2

Enumerative Problems on Directed Polytope Graphs

In this chapter we are going to prove the bounds on arborescences, monotone paths and diameter of monotone paths as stated in Theorem 1.3.1, 1.3.2 and 1.3.3. First, we prove a lemma for the diameter bound of stacked polytopes.

LEMMA 2.0.1. *The diameter of the graph of f-monotone paths on $X(n)$ is bounded below by $\lceil (n-2)^2/4 \rceil$ for every $n \geq 4$.*

PROOF. Let G be the graph of f-monotone paths on $X(n)$. Denoting f-monotone paths as sequences of vertices, we set

$$\gamma = \begin{cases} (v_1, v_3, v_5, \ldots, v_{n-1}, v_n), & \text{if } n \equiv 0 \pmod 2 \\ (v_1, v_2, v_4, \ldots, v_{n-3}, v_{n-1}, v_n), & \text{if } n \equiv 1 \pmod 4 \\ (v_1, v_3, v_5, \ldots, v_{n-2}, v_n), & \text{if } n \equiv 3 \pmod 4 \end{cases}$$

and $\delta = (v_1, v_2, v_3, \ldots, v_n)$. We claim that γ and δ are at a distance of $\lceil (n-2)^2/4 \rceil$ apart in G. Clearly, the lemma follows from the claim.

We only consider the case that n is even, the other two cases being similar. By passing to the complement of the set of vertices appearing on an f-monotone path on $X(n)$, such paths correspond bijectively to the subsets of $\{v_2, v_3, \ldots, v_{n-1}\}$ containing no three consecutive elements v_{k-1}, v_k, v_{k+1}. The subset which corresponds to γ, for instance, is $\{v_2, v_4, \ldots, v_{n-2}\}$ and the one which corresponds to δ is the empty set. The 2-dimensional faces of $X(n)$ have vertex sets $\{v_1, v_2, v_3\}$, $\{v_{n-2}, v_{n-1}, v_n\}$ and $\{v_{k-1}, v_k, v_{k+2}\}$ and $\{v_{k-1}, v_{k+1}, v_{k+2}\}$ for $2 \leq k \leq n-2$. From these facts it follows that polygon flips on f-monotone paths on $X(n)$ correspond to the following operations on the corresponding subsets:

- removal of v_2 or v_{n-1}, if present,
- inclusion of v_2, if absent and not both v_3 and v_4 are present,

- inclusion of v_{n-1}, if absent and not both v_{n-2} and v_{n-3} are present,
- removal or inclusion of one of v_k, v_{k+1}, if the other is present but v_{k-1} and v_{k+2} are absent.

Since the subsets which correspond to f-monotone paths on $X(n)$ contain no three consecutive elements, their maximal strings of consecutive elements are either singletons, or contain exactly two elements. Moreover, the strings cannot be merged with these operations, they cannot be removed except for $\{2\}$ and $\{n-1\}$, and each operation affects only one of them. To reach the empty set from $\{v_2, v_4, \ldots, v_{n-2}\}$, one needs to remove each of $v_2, v_4, \ldots, v_{n-2}$. Regardless of the order in which operations are performed, at least one is needed to remove v_2, at least three more are needed to remove v_{n-2}, at least five more are needed to remove v_4, and so on. For example, to remove v_{n-2} in at most three steps one needs to first include v_{n-1}, then remove v_{n-2} and finally remove v_{n-1} and to remove v_4 in at most five steps one needs to first include v_3, then remove v_4, include v_2, remove v_3 and finally remove v_2. This yields a distance of $1+3+5+\cdots+(n-3) = (n-2)^2/4$ between γ and δ in G. $\square$

Remark 2.0.2. Perhaps it is instructive to visualize the process of flipping γ to δ, described in the previous proof. The two f-monotone paths are shown on Figure 2.1 for $n = 10$ and the sequence of 2-dimensional faces (recording only vertex indices, for simplicity) across which the flips occur could be $\{1,2,3\}$, $\{7,9,10\}$, $\{7,8,10\}$, $\{8,9,10\}$, $\{2,3,5\}$, $\{2,4,5\}$, $\{1,2,4\}$, $\{1,3,4\}$, $\{1,2,3\}$, $\{5,7,8\}$, $\{5,6,8\}$, $\{6,8,9\}$, $\{6,7,9\}$, $\{7,9,10\}$, $\{7,8,10\}$ and $\{8,9,10\}$. $\square$

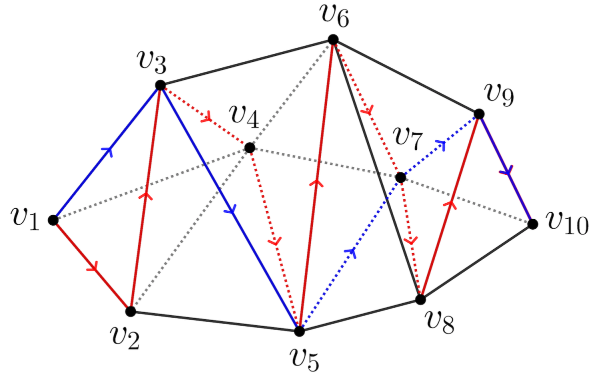

Figure 2.1. Two monotone paths on $X(10)$

Finally, we consider prisms and wedges of polygons. Given a $(d-1)$-dimensional polytope Q, the *prism* over Q is the d-dimensional polytope defined as the Cartesian product $Q \times [0,1]$. The *wedge* of Q over a face F of Q is the d-dimensional polytope W obtained combinatorially from the prism $Q \times [0,1]$ by collapsing the face $F \times [0,1]$ to $F \times 0$. Note that Q becomes a facet of W and that if F is a facet and Q is simple, then so is W. We will apply the wedge construction in the special cases that Q is a polygon and F is one of its vertices or edges.

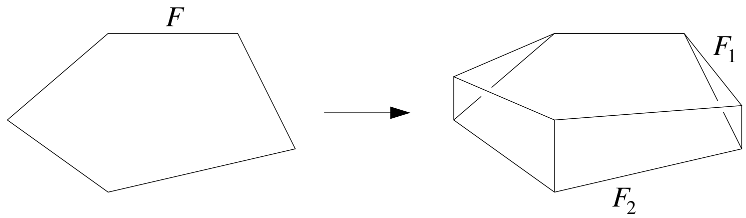

FIGURE 2.2. The wedge of a pentagon over an edge

2.1. The graph of f-monotone paths

Let P be a d-dimensional polytope and f be a generic linear functional on P. We will assume that f does not take the same value on any two distinct vertices of P.

To investigate the graph of f-monotone paths on P, we will describe another way to construct it from simpler graphs, arising in the fibers of the restriction of the projection map f on P. The technical device needed, which we now review, is the *inverse limit* in the category of graphs and simplicial maps. This concept was introduced in [**AER00**, Section 4] (with motivation coming from [**BKS94**]) to study the higher connectivity of $G(P, f)$; it leads to various more general graphs of partial f-monotone paths on P, a useful notion which allows for inductive arguments.

Let us linearly order the vertices $v_0, v_1, \ldots, v_n$ of P so that $f(v_0) < f(v_1) < \cdots < f(v_n)$. We recall that for every interior point t of the interval $f(P)$, the fiber $P(t) := f^{-1}(t) \cap P$ of the map $f : P \to \mathbb{R}$ is a $(d-1)$-dimensional polytope and thus it has a well defined graph. Setting $t_i = f(v_i)$ for $0 \le i \le n$, we may thus consider the graph G_i of $P(t_i)$ for $0 \le i \le n$ and the graph $G_{i,i+1}$ of $P(t)$ for some $t_i < t < t_{i+1}$, for $0 \le i \le n-1$ (the precise value of t being irrelevant because, by construction, the other choices of t in the same interval give a normally equivalent fiber $P(t)$); see

Figure 2.3 for an example. Considering these graphs as one-dimensional simplicial complexes, we have a diagram

$$G_{0,1} \xrightarrow{\alpha_1} G_1 \xleftarrow{\beta_1} G_{1,2} \xrightarrow{\alpha_2} G_2 \xleftarrow{\beta_2} \cdots \xleftarrow{\beta_{n-2}} G_{n-2,n-1} \xrightarrow{\alpha_{n-1}} G_{n-1} \xleftarrow{\beta_{n-1}} G_{n-1,n} \tag{2.1.1}$$

of graphs and simplicial maps for which $\alpha_i : G_{i-1,i} \to G_i$ and $\beta_i : G_{i,i+1} \to G_i$ result from the degeneration of the fiber $P(t)$ when t approaches t_i, with $t_{i-1} < t < t_i$ or $t_i < t < t_{i+1}$, respectively (recall that a simplicial map of one-dimensional simplicial complexes maps vertices to vertices and either maps edges linearly onto edges, or contracts them to vertices; in particular, such a map is determined by its images on vertices).

The inverse limit G of this diagram is defined as follows. The nodes are the sequences

$$(v_{0,1}, v_{1,2}, \ldots, v_{n-1,n}),$$

where $v_{i-1,i}$ is a vertex of $G_{i-1,i}$ for all $i \in [n]$ and $\alpha_i(v_{i-1,i}) = \beta_i(v_{i,i+1})$ for all $i \in [n-1]$. Two such sequences, say $(u_{0,1}, u_{1,2}, \ldots, u_{n-1,n})$ and $(v_{0,1}, v_{1,2}, \ldots, v_{n-1,n})$, are adjacent nodes in G if there exists a nonempty interval $\mathcal{I} \subseteq [n]$ such that:

- $u_{i-1,i}$ and $v_{i-1,i}$ are adjacent in $G_{i-1,i}$ for $i \in \mathcal{I}$,
- $u_{i-1,i} = v_{i-1,i}$ for $i \in [n] \smallsetminus \mathcal{I}$, and
- the edges $\{u_{i-1,i}, v_{i-1,i}\}$ and $\{u_{i,i+1}, v_{i,i+1}\}$ are mapped homeomorphically onto the same edge of G_i by α_i and β_i, respectively, whenever $i, i+1 \in \mathcal{I}$.

This construction associates an inverse limit graph to any diagram of graphs and simplicial maps (2.1.1). As explained in [**AER00**, Section 4] (see [**AER00**, Proposition 4.1]), the graph G is isomorphic to $G(P, f)$ when the diagram comes from a polytope P and linear functional f, as just described. The inverse limit of a subdiagram of (2.1.1) of the form

$$G_{k-1,k} \xrightarrow{\alpha_k} G_k \xleftarrow{\beta_k} G_{k,k+1} \xrightarrow{\alpha_{k+1}} \cdots \xleftarrow{\beta_{\ell-1}} G_{\ell-1,\ell} \xrightarrow{\alpha_\ell} G_\ell \xleftarrow{\beta_\ell} G_{\ell,\ell+1},$$

considered in Sections 2.3 and 2.4, has nodes which can be viewed as partial f-monotone paths on P, starting at the fiber $P(t)$ with $t_{k-1} < t < t_k$ and ending at $P(t')$ with $t_\ell < t' < t_{\ell+1}$, and adjacency given by a suitable extension of the notion of polygon flip, presented in the introduction.

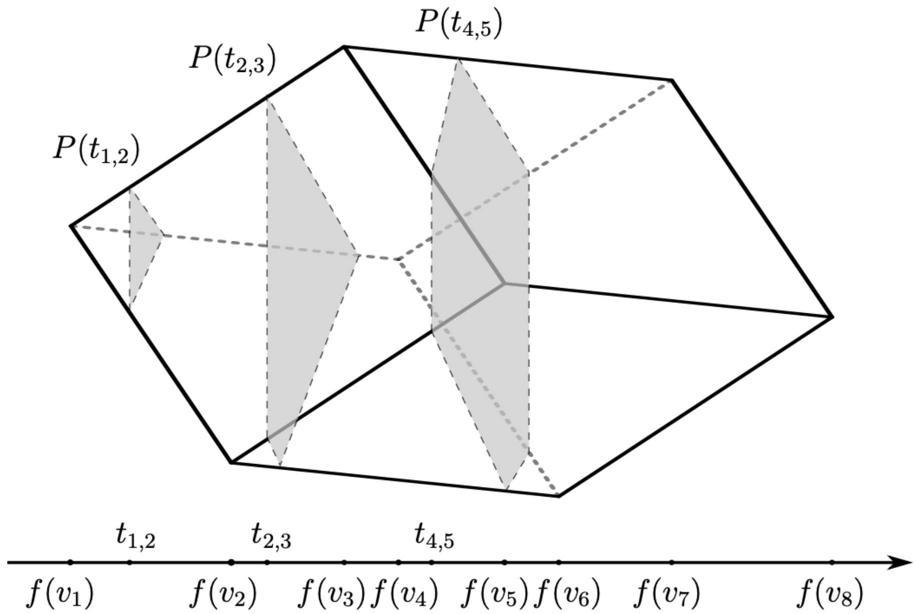

FIGURE 2.3. A combinatorial cube and some of its fibers

Now we are ready to prove bounds as stated in the Introduction.

2.2. On the number of arborescences

As explained in the introduction, we are interested in counting f-arborescences on a polytope P, meaning oriented spanning trees in the directed graph $\omega(P, f)$ which are rooted at the unique sink $v_{\max}$. Recall that $\tau(P, f)$ denotes the number of f-arborescences on P. The following statement provides an explicit product formula for this number.

PROPOSITION 2.2.1. *Given a d-dimensional polytope P and generic linear functional f, let $\mathrm{out}_f(v)$ denote the outdegree of the vertex v of P in the directed graph $\omega(P, f)$. Then,*

$$\tau(P, f) = \prod_{v \neq v_{\max}} \mathrm{out}_f(v),$$

where the product ranges over all vertices of P other than the sink $v_{\max}$. In particular, if P is simple, then

$$\tau(P, f) \;=\; \prod_{i=1}^{d} i^{h_i(P)}$$

is independent of f.

Proof. Since $\omega(P, f)$ is acyclic, an f-arborescence is uniquely determined by a choice of edge coming out of v for every vertex v of $\omega(P, f)$ other than the sink $v_{\max}$. Since there are exactly $\mathrm{out}_f(v)$ choices for every such v, the proof of the first formula follows. The second formula follows from the first and the combinatorial interpretation of the h-vector of a simple polytope P, mentioned in Section 1.1.3. $\square$

Remark 2.2.2. Since every edge of $\omega(P, f)$ has a unique initial vertex, the sum of the outdegrees $\mathrm{out}_f(v)$ of the vertices of P in the directed graph $\omega(P, f)$ is equal to the number of edges of P.

Corollary 2.2.3. *For $m > d \geq 4$, the maximum number of f-arborescences over all simple d-dimensional polytopes with m facets is achieved by the polar duals of neighborly polytopes and is given by the formula*

$$\max \tau(P, f) = \prod_{i=1}^{\lfloor \frac{d}{2} \rfloor} i^{\binom{m-d+i-1}{i}} \prod_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} (d-i)^{\binom{m-d+i-1}{i}}.$$

Similarly, the minimum number of f-arborescences in this situation is achieved by the polar duals of stacked polytopes and is given by the formula

$$\min \tau(P, f) = d \cdot ((d-1)!)^{m-d}\,.$$

For 3-dimensional simple polytopes P with m facets, $\tau(P, f) = 3 \cdot 2^{m-3}$.

Proof. The case $d \geq 4$ follows from the last sentence of Proposition 2.2.1, the upper and lower bound theorems for the h-vector of a simplicial polytope, discussed in Section 1.1.3, and the formulas for the h-vectors of d-dimensional neighborly and stacked simplicial polytope with m vertices given there. The case $d = 3$ follows again from the second formula of Proposition 2.2.1, since $h_0(P) = h_3(P) = 1$ and $h_1(P) = h_2(P) = m - 3$ for every 3-dimensional simple polytope P with m facets. $\square$

The following two statements apply to general polytopes. Combined with Corollary 2.2.3, they imply the results about f-arborescences stated in the introduction.

THEOREM 2.2.4. *For $n > d \geq 3$, the maximum number of f-arborescences over all d-dimensional polytopes with n vertices is achieved by the stacked polytope $X(n)$ for $d = 3$ and by any 2-neighborly polytope for $d \geq 4$. This number is equal to $2 \cdot 3^{n-3}$ and $(n-1)!$ in the two cases, respectively.*

PROOF. Let us order the vertices $v_1, v_2, \ldots, v_n$ of the d-dimensional polytope P so that $f(v_1) \leq f(v_2) \leq \cdots \leq f(v_n)$, where $v_n = v_{\max}$. Then, arcs of the directed graph $\omega(P, f)$ can only be pairs (v_i, v_j) with $i < j$ and hence $\mathrm{out}_f(v_i) \leq n - i$ for every $i \in [n]$. Thus, in view of Proposition 2.2.1, we get

$$\tau(P, f) \;=\; \prod_{i=1}^{n-1} \mathrm{out}_f(v_i) \;\leq\; \prod_{i=1}^{n-1} (n-i) \;=\; (n-1)!$$

and equality holds if and only if P is 2-neighborly.

Since no such polytopes other than simplices exist in dimension $d = 3$, this case has to be treated separately. Setting $d_i = \mathrm{out}_f(v_i)$ for $i \in [n-1]$, we have positive integers $d_1, d_2, \ldots, d_{n-1}$ such that $d_{n-1} = 1$ and $d_{n-2} \in \{1, 2\}$. Since P can have no more than $3n-6$ edges, we have $d_1 + d_2 + \cdots + d_{n-1} \leq 3n-6$ by Remark 2.2.2. It is an elementary fact that, under these assumptions, the product $\tau(P, f) = d_1 d_2 \cdots d_{n-1}$ is maximized when $d_{n-1} = 1$, $d_{n-2} = 2$ and $d_i = 3$ for every $i \in [n-3]$. Exactly that happens for the stacked polytope $X(n)$ and the proof follows. $\square$

THEOREM 2.2.5. *For all $n \geq 4$, the minimum number of f-arborescences over all 3-dimensional polytopes with n vertices is equal to $2(n-1)$. This is achieved by any pyramid P and any generic linear functional f which takes its minimum value on P at the apex.*

PROOF. As a simple application of Proposition 2.2.1, we have $\tau(P, f) = 2(n-1)$ for every pyramid P over an $(n-1)$-gon and every generic functional f which takes its minimum value on P at the apex.

We now consider any 3-dimensional polytope P with n vertices and any generic functional f on P. We need to show that $\tau(P, f) \geq 2(n-1)$. We may linearly order the vertices $v_1, v_2, \ldots, v_n$ of P in the order of decreasing outdegree in the directed graph $\omega(P, f)$ and denote by k the number of those vertices which have outdegree larger than one. Then, $k \geq 2$ and the respective outdegrees

$d_1, d_2, \ldots, d_n$ of $v_1, v_2, \ldots, v_n$ satisfy $d_1, d_2, \ldots, d_k \geq 2$, $d_n = 0$ and $d_i = 1$ for every other value of i. Letting $D_1, D_2, \ldots, D_n$ be the degrees of $v_1, v_2, \ldots, v_n$ in the undirected graph of P, respectively, we have $\tau(P, f) = d_1 d_2 \cdots d_k$ and

$$2 \cdot \sum_{i=1}^{n} d_i \;=\; \sum_{i=1}^{n} D_i$$

by Remark 2.2.2. Clearly, $D_i = d_i$ for one value of $i \in \{1, 2, \ldots, k\}$ (the one corresponding to the source vertex), $D_i \geq d_i + 1$ for every other such value and $D_i \geq 3$ for all $k < i \leq n$. These considerations result in the inequality $d_1 + d_2 + \cdots + d_k \geq n + 1$ and thus, it remains to show that $d_1 d_2 \cdots d_k \geq 2(n-1)$ for every $k \geq 2$ and all $d_1, d_2, \ldots, d_k \in \{2, 3, \ldots, n-1\}$ summing at least to $n+1$. Indeed, from the inequality $ab > (a-1)(b+1)$ for integers $a \leq b$, applied repeatedly when b is the largest of $d_1, d_2, \ldots, d_k$ and a is any other number from these larger than 2, we get

$$d_1 d_2 \cdots d_k \;\geq\; (d_1 + d_2 + \cdots + d_k - 2k + 2) \cdot 2^{k-1} \;\geq\; (n - 2k + 3) \cdot 2^{k-1}.$$

Applying repeatedly the fact that $2m \geq m + 2$ for $m \geq 2$, we conclude that $d_1 d_2 \cdots d_k \geq 2(n-1)$ and the proof follows. □

More generally, for any $d \geq 3$, the $(d-2)$-fold pyramid P over an $(n-d+2)$-gon has n vertices and dimension d. Moreover, if f is chosen so that every cone vertex has smaller objective value than any of the vertices of the $(n-d+2)$-gon, then the number of f-arborescences on P is equal to $2(n-1)(n-2)\cdots(n-d+2)$.

Question 2.2.6. What is the minimum number of f-arborescences over all d-dimensional polytopes with n vertices, for $d \geq 4$? Does it equal $2(n-1)(n-2)\cdots(n-d+2)$ for all $n > d \geq 4$?

2.3. On the number of monotone paths

This section investigates the smallest and largest possible number of f-monotone paths on polytopes. For notational convenience, we let $v_0, v_1, \ldots, v_n$ be the vertices of a polytope P, linearly ordered so that $f(v_0) < f(v_1) < \cdots < f(v_n)$, as in Section 2.1. We recall that $\mu(P, f)$ denotes the number of f-monotone paths on P and that we refer to general directed paths in $\omega(P, f)$ as partial f-monotone paths, i.e., they may start or end at vertices other than $v_{\min}$ or $v_{\max}$.

The following formula is the key to most results in this section.

LEMMA 2.3.1. *The number of f-monotone paths on P can be expressed as*

$$\mu(P, f) \;=\; 1 + \sum_{k=0}^{n-1} (d_k - 1)\mu_k(P, f),$$

where $d_k = \mathrm{out}_f(v_k)$ is the outdegree of v_k in $\omega(P, f)$ and $\mu_k(P, f)$ stands for the number of partial f-monotone paths on P with initial vertex v_0 and terminal vertex v_k.

PROOF. Let $P(t) = f^{-1}(t) \cap P$ be the fibers of the map $f : P \to \mathbb{R}$, as in Section 2.1, and $t_i = f(v_i)$ for $0 \le i \le n$. For $0 \le k \le n-1$ let $\mathcal{H}_k(P, f)$ be the set of partial f-monotone paths on P having initial vertex v_0 and ending in the fiber $P(t)$ with $t_k < t < t_{k+1}$. Formally, these are essentially the nodes of the inverse limit of the part

$$G_{0,1} \xrightarrow{\alpha_1} G_1 \xleftarrow{\beta_1} G_{1,2} \xrightarrow{\alpha_2} G_2 \xleftarrow{\beta_2} \cdots \xrightarrow{\alpha_k} G_k \xleftarrow{\beta_k} G_{k,k+1}$$

of the diagram (2.1.1). Let $\eta_k(P, f)$ be the number of these partial f-monotone paths. We claim that

$$\eta_k(P, f) - \eta_{k-1}(P, f) \;=\; (d_k - 1)\mu_k(P, f) \tag{2.3.1}$$

for every $k \in [n-1]$. Since $\eta_0(P, f) = \mathrm{out}_f(v_0) = d_0$ and $\mu_0(P, f) = 1$, this implies that

$$\eta_k(P, f) \;=\; 1 + \sum_{i=0}^{k} (d_i - 1)\mu_i(P, f)$$

for $0 \le k \le n-1$. Since $\eta_{n-1}(P, f) = \mu_n(P, f) = \mu(P, f)$, the desired formula follows as the special case $k = n-1$ of this equation.

To verify (2.3.1), let $\varphi_k : \mathcal{H}_k(P, f) \to \mathcal{H}_{k-1}(P, f)$ be the natural map obtained by restriction of diagrams. More intuitively, $\varphi_k(\gamma)$ is obtained from $\gamma \in \mathcal{H}_k(P, f)$ by removing its last edge. Paths in $\mathcal{H}_{k-1}(P, f)$ and $\mathcal{H}_k(P, f)$ either pass through vertex v_k or not, depending on whether or not their last edge maps to v_k under the map α_k or β_k, respectively. Clearly, for every $\delta \in \mathcal{H}_{k-1}(P, f)$ which passes through v_k there are exactly d_k paths $\gamma \in \mathcal{H}_k(P, f)$ such that $\varphi_k(\gamma) = \delta$, obtained by choosing an edge of $\omega(P, f)$ coming out of v_k and attaching it to δ, while for every $\delta \in \mathcal{H}_{k-1}(P, f)$ which does not pass through v_k there is a unique path $\gamma \in \mathcal{H}_k(P, f)$ such that $\varphi_k(\gamma) = \delta$. These observations imply directly Equation (2.3.1) and the proof follows. □

Recall that the Tribonacci sequence (T_n) is defined by the recurrence relation $T_0 = T_1 = 1$, $T_2 = 2$ and $T_n = T_{n-1} + T_{n-2} + T_{n-3}$ for $n \geq 3$.

Theorem 2.3.2. *The maximum number of f-monotone paths over all 3-dimensional polytopes with $n+1$ vertices is equal to the nth Tribonacci number T_n for every $n \geq 3$. This is achieved by the stacked polytope $X(n)$.*

Proof. We proceed by induction on n. The result holds for $n = 3$, since there are exactly $T_3 = 4$ monotone paths on any 3-dimensional simplex. We assume that it holds for integers less than n and consider a 3-dimensional polytope P with $n+1$ vertices $v_0, v_1, \dots, v_n$, linearly ordered as in the beginning of this section by a generic functional f.

We wish to apply Lemma 2.3.1. Since partial f-monotone paths on P with initial vertex v_0 and terminal vertex v_k are f-monotone paths on the convex hull of $v_0, v_1, \dots, v_k$, we have $\mu_k(P, f) \leq T_k$ for $k \in \{3, 4, \dots, n-1\}$ by the induction hypothesis. Since this bound holds trivially for $k \in \{0, 1, 2\}$ as well, from Lemma 2.3.1 we get

$$\mu(P, f) \;\leq\; 1 + \sum_{k=0}^{n-1} (d_k - 1) T_k.$$

To bound the right-hand side, we note that

$$d_{n-k} + d_{n-k+1} + \cdots + d_{n-1} \leq 3k - 3$$

for $k \in \{2, 3, \dots, n-1\}$, since $d_{n-k} + d_{n-k+1} + \cdots + d_{n-1}$ is equal to the number of edges of P connecting vertices $v_{n-k}, v_{n-k+1}, \dots, v_n$ and hence to the number of edges of a planar simple graph with $k+1$ vertices. From these inequalities and the trivial one $d_{n-1} \leq 1$, and setting $T_{-1} := 0$, we get

$$\begin{aligned}\sum_{k=0}^{n-1} d_k T_k &= \sum_{k=1}^{n}(d_{n-1}+d_{n-2}+\cdots+d_{n-k})(T_{n-k}-T_{n-k-1}) \\ &\le (T_{n-1}-T_{n-2}) + (3k-3)\sum_{k=2}^{n}(T_{n-k}-T_{n-k-1}) \\ &= T_{n-1}+2T_{n-2}+3T_{n-3}+3T_{n-4}+\cdots+3T_0 \\ &= \sum_{k=1}^{n} T_k,\end{aligned}$$

where the last equality comes from summing the recurrence $T_k = T_{k-1}+T_{k-2}+T_{k-3}$ for $k \in [n]$. We conclude that

$$\mu(P,f) \le 1+\sum_{k=0}^{n-1}(d_k-1)T_k = 1+\sum_{k=0}^{n-1} d_k T_k - \sum_{k=0}^{n-1} T_k \le T_n.$$

This completes the induction.

Finally, it is straightforward to verify that the number of f-monotone paths on $X(n+1)$ satisfies the Tribonacci recurrence (or alternatively, that all inequalities hold as equalities in the previous argument) and is thus equal to T_n for every n. $\square$

Remark 2.3.3. The number of f-monotone paths on a polytope P with $n+1$ vertices is no larger than the number of subsets of its vertex set containing the source and the sink and hence at most 2^{n-1}. Equality holds exactly when P is 2-neighborly, meaning that the 1-skeleton of P is the complete graph on $n+1$ vertices, since then every such subset is the vertex set of an f-monotone path on P. As a result, the maximum number of f-monotone paths over all d-dimensional polytopes with $n+1$ vertices is equal to 2^{n-1} for all $n \ge d \ge 4$. $\square$

The following statement completes the proof of the results about the number of f-monotone paths, stated in the introduction.

Theorem 2.3.4. *The minimum number of f-monotone paths over all 3-dimensional polytopes with n vertices is equal to $\lceil n/2 \rceil + 2$. This is achieved by prisms, when n is even, and by wedges of polygons over a vertex, when n is odd.*

In particular, prisms minimize the number of f-monotone paths over all simple polytopes of dimension three with given number of vertices.

PROOF. Applying Lemma 2.3.1 and noting that $\mu_k(P,f) \geq 1$ for every k, we get

$$\mu(P,f) \;\geq\; 1 + \sum_{k=0}^{n-2} (d_k - 1) \;=\; \sum_{k=0}^{n-2} d_k - n + 2.$$

Since $\sum_{k=0}^{n-2} d_k$ is equal to the number of edges of P (see Remark 2.2.2), which is bounded below by $\lceil 3n/2 \rceil$, it follows that $\mu(P,f) \geq \lceil n/2 \rceil + 2$. It is straightforward to verify that prisms achieve the minimum when n is even and wedges of polygons over a vertex (obtained from prisms by identifying two vertices at different levels which are connected by an edge) achieve the minimum when n is odd. □

The lower bound for the number of f-monotone paths in any dimension, given in the following statement, is not expected to be tight.

PROPOSITION 2.3.5. *The number of f-monotone paths on any polytope of dimension d with n vertices is bounded below by $\lceil dn/2 \rceil - n + 2$.*

PROOF. Once again, this follows from the inequality $\sum_{k=0}^{n-2} d_k \geq \lceil dn/2 \rceil$ and Lemma 2.3.1. □

We end this section with a conjecture for the maximum number of monotone paths on simple 3-dimensional polytopes. The proposed maximum can be achieved by wedges of polygons over an edge whose vertices are the source and the sink, and all vertices of the polytope lie on a monotone path. We recall that the Fibonacci sequence (F_n) is defined by the recurrence $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$.

CONJECTURE 2.3.6. *We have $\mu(P,f) \leq F_{n+2} + 1$ for every simple 3-dimensional polytope P with $2n$ vertices.*

The argument in the proof of Theorem 2.3.2 yields the following weaker result.

PROPOSITION 2.3.7. *We have $\mu(P,f) \leq 2F_n$ for every 3-dimensional simple polytope P with $n+1$ vertices and every generic linear functional f on P.*

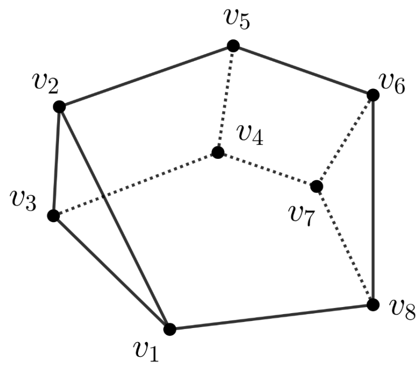

FIGURE 2.4. An example of a polytope on 8 vertices conjectured to be the maximizer of the number of monotone paths among simple polytopes. $f(v_1) < f(v_2) < \cdots < f(v_8)$.

PROOF. Let (a_n) be the sequence of numbers defined by the recurrence relation $a_0 = a_1 = 1$, $a_2 = 2$, $a_3 = 4$ and $a_n = a_{n-1} + a_{n-2}$ for $n \geq 4$. Note that $a_n = 2F_n$ for $n \geq 2$. We mimick the proof of Theorem 2.3.2 to show that $\mu(P, f) \leq a_n$ for all $n \geq 3$. For the inductive step, since P is simple, we have $d_0 = 3$, $d_1, d_2, \ldots, d_{n-2} \leq 2$ and $d_{n-1} = 1$ and compute that

$$\begin{aligned} \mu(P, f) \;&\leq\; 1 + \sum_{k=0}^{n-1} d_k a_k - \sum_{k=0}^{n-1} a_k \;\leq\; 1 + a_{n-2} + a_{n-3} + \ldots + a_1 + 2a_0 \\ &\leq\; a_{n-1} + a_{n-2} \;=\; a_n, \end{aligned}$$

since $a_{n-1} = 1 + a_{n-3} + \cdots + a_1 + 2a_0$. $\square$

2.4. On the diameter of monotone path graphs

The main goal of this section is to prove Theorem 1.3.2.

The lower bound of (1.3.3) for the maximum diameter follows from Lemma 2.0.1. The upper bound will be deduced from the following result. Clearly, given a polytope P and a generic linear functional f, every f-monotone path on P meets each of the fibers $f^{-1}(t) \cap P$, where $t \in f(P)$, in a unique point. For f-monotone paths γ and γ' on P, let us denote by $\nu(\gamma, \gamma')$ the number of connected components of the set of values $t \in f(P)$ for which γ and γ' disagree on $f^{-1}(t) \cap P$. For example, for the two monotone paths, say γ and γ', shown on Figure 2.1 we have $\nu(\gamma, \gamma') = 4$. Note that $\nu(\gamma, \gamma') = 0 \Leftrightarrow \gamma = \gamma'$.

THEOREM 2.4.1. *Let P be a 3-dimensional polytope and f be a generic linear functional on P. The distance between any two f-monotone paths γ and γ' in the graph $G = G(P, f)$ satisfies*

$$d_G(\gamma, \gamma') \leq \frac{\nu(\gamma, \gamma')}{2} \cdot f_2(P), \tag{2.4.1}$$

where $f_2(P)$ is the number of 2-dimensional faces of P.

REMARK 2.4.2. Theorem 2.4.1 gives a diameter bound for all three-dimensional polytopes. Cordovil and Moreira had studied bounds for three-dimensional zonotopes and rank-three oriented matroids [**CM93**], which they gave in terms of the dual pseudo-line arrangements.

We will first state a technical result (see Proposition 2.4.3) which constructs a walk in $G(P, f)$ between two monotone paths γ and γ' with the required properties from walks on the fibers, assuming that the latter satisfy certain necessary compatibility conditions. To allow for all possible ways that γ and γ' may intersect each other, we consider the following general situation. Let $\mathcal{F}$ be a connected polygonal complex in $\mathbb{R}^d$ having vertices $v_0, v_1, \ldots, v_n$ and $f : \mathbb{R}^d \to \mathbb{R}$ be a linear functional such that $f(v_0) < f(v_1) < \cdots < f(v_n)$. The graph of f-monotone paths on $\mathcal{F}$, denoted by $G(\mathcal{F}, f)$, having initial vertex v_0 and terminal vertex v_n, can be defined with adjacency given by polygon flips just as in the special case in which $\mathcal{F}$ is the 2-skeleton of a convex polytope (see Section 2.1). Alternatively, and in order to relate it to the graphs of the fibers of f, we may view $G(\mathcal{F}, f)$ as the inverse limit associated to a diagram

$$G_{0,1} \xrightarrow{\alpha_1} G_1 \xleftarrow{\beta_1} G_{1,2} \xrightarrow{\alpha_2} G_2 \xleftarrow{\beta_2} \cdots \xleftarrow{\beta_{n-2}} G_{n-2,n-1} \xrightarrow{\alpha_{n-1}} G_{n-1} \xleftarrow{\beta_{n-1}} G_{n-1,n} \tag{2.4.2}$$

of graphs and simplicial maps. This is defined as in Section 2.1 provided the fiber $f^{-1}(t) \cap P$ is replaced with $f^{-1}(t) \cap \|\mathcal{F}\|$, where $\|\mathcal{F}\|$ is the polyhedron (union of faces) of $\mathcal{F}$. Thus, the G_i and $G_{i,i+1}$ are graphs of (one-dimensional) fibers $f^{-1}(t) \cap \|\mathcal{F}\|$ and the α_i and β_i are natural degeneration maps.

Given an f-monotone path γ on $\mathcal{F}$ and $i \in [n]$, let us denote by $\pi_i(\gamma)$ the node of $G_{i-1,i}$ in which the union of the edges of γ intersects the corresponding fiber $f^{-1}(t) \cap \|\mathcal{F}\|$. Then, $\pi_i : G(\mathcal{F}, f) \to$

$G_{i-1,i}$ is a simplicial map. Given a walk $\mathcal{P}$ in a graph G, thought of as a sequence of edges, and a simplicial map $\varphi : G \to H$ of graphs, let us denote by $\varphi(\mathcal{P})$ the walk in H which is formed by the images of the edges of $\mathcal{P}$ under φ, disregarding those edges of $\mathcal{P}$ which are contracted to a node by φ.

PROPOSITION 2.4.3. *Let γ and δ be f-monotone paths on $\mathcal{F}$. Suppose that for every $i \in [n]$ there exists a walk $\mathcal{P}_i$ in $G_{i-1,i}$ with initial node $\pi_i(\gamma)$ and terminal node $\pi_i(\delta)$ which traverses each edge in $G_{i-1,i}$ exactly once, so that*

$$\alpha_i(\mathcal{P}_i) = \beta_i(\mathcal{P}_{i+1}) \tag{2.4.3}$$

for every $i \in [n-1]$. Then, there exists a walk $\mathcal{P}$ in $G(\mathcal{F}, f)$ with initial node γ and terminal node δ which traverses each 2-dimensional face of $\mathcal{F}$ exactly once, such that $\pi_i(\mathcal{P}) = \mathcal{P}_i$ for every $i \in [n]$.

We first illustrate the proposition with an important special case and then use it to prove Theorem 2.4.1.

EXAMPLE 2.4.4. To motivate the proof of Theorem 2.4.1, consider the special case in which the monotone paths γ and γ' do not have common vertices, other than those on which f attains its minimum and maximum value on P. Then, $\nu(\gamma, \gamma') = 1$ and the edges of γ and γ' form a simple cycle C which divides the boundary of P into two closed balls, say B^+ and B^-, having common boundary C. Let $\mathcal{F}^+$ and $\mathcal{F}^-$ be the two subcomplexes of the boundary complex of P which correspond to these balls. We wish to show that for each $\varepsilon \in \{+, -\}$, there exists a walk in $G(P, f)$ joining γ and γ' which traverses each 2-dimensional face of $\mathcal{F}^\varepsilon$ exactly once. This would imply the desired bound for $d_G(\gamma, \gamma')$. Such a walk must traverse every edge of each fiber $f^{-1}(t) \cap B^\varepsilon$ exactly once and thus induce walks on these fibers with the same property.

Let us consider the diagram (2.4.2) for the polygonal complex $\mathcal{F}^\varepsilon$. Clearly, the fiber $f^{-1}(t) \cap \partial P$ is the boundary of a polygon for every interior point t of the interval $f(P)$, where ∂P denotes the boundary of P. Since, by the f-monotonicity of γ and γ', this fiber intersects the cycle C, which is the boundary of the ball B^ε, in exactly two points, its intersection with B^ε must be homeomorphic to a line segment. Thus, all graphs appearing in the diagram (2.4.2) for $\mathcal{F}^\varepsilon$ are path graphs, where $G_{i-1,i}$ has endpoints $\pi_i(\gamma)$ and $\pi_i(\gamma')$ for every $i \in [n]$. As a result, there are unique walks $\mathcal{P}_i$, as in the statement of Proposition 2.4.3, where condition (2.4.3) holds by the degeneration of fibers. Thus, Proposition 2.4.3 implies the existence of a walk in $G(\mathcal{F}^\varepsilon, f)$ with initial node γ and terminal node γ' which traverses each 2-dimensional face of $\mathcal{F}^\varepsilon$ exactly once. $\square$

Proof of Theorem 2.4.1. We first observe that it suffices to prove the special case $\nu(\gamma, \gamma') = 1$. Indeed, given f-monotone paths γ and γ' on P and setting $\nu = \nu(\gamma, \gamma')$, it is straightforward to define f-monotone paths $\gamma = \gamma_0, \gamma_1, \ldots, \gamma_\nu = \gamma'$ on P satisfying $\nu(\gamma_{i-1}, \gamma_i) = 1$ for every $i \in [\nu - 1]$. Then, the triangle inequality and the special case imply that

$$d_G(\gamma, \gamma') \le \sum_{i=1}^{\nu} d_G(\gamma_{i-1}, \gamma_i) \le \nu \cdot \frac{f_2(P)}{2},$$

as claimed by (2.4.1).

So let γ, γ' be f-monotone paths on P such that $\nu(\gamma, \gamma') = 1$. Let u and v be their unique common vertices, satisfying $f(u) < f(v)$, for which γ and γ' disagree on each fiber $f^{-1}(t) \cap P$ with $f(u) < t < f(v)$ and agree on the other fibers; in the special case of Example 2.4.4, u and v are the unique vertices $v_{\min}$ and $v_{\max}$ on which f attains its minimum and maximum value on P, respectively. As in that special case, the edges of γ and γ' joining u and v form a simple cycle C which divides the 2-dimensional sphere ∂P into two closed 2-dimensional balls B^+ and B^- having common boundary C. Moreover, the f-monotonicity of γ and γ' implies that for each $\varepsilon \in \{+, -\}$ and every interior point t of the interval $f(B^\varepsilon)$, the fiber $f^{-1}(t) \cap B^\varepsilon$ is homeomorphic to a line segment or a circle. We wish to apply Proposition 2.4.3 to the subcomplex $\mathcal{F}^\varepsilon$ of the boundary complex of P which corresponds to B^ε.

We claim that there exist unique walks $\mathcal{P}_i$ satisfying the assumptions of the proposition. Indeed, according to our previous discussion, every graph $G_{i-1,i}$ appearing in the diagram (2.4.2) for $\mathcal{F}^\varepsilon$ is either a path graph, with endpoints $\pi_i(\gamma)$ and $\pi_i(\gamma')$, or a cycle. As a result, there exists a unique walk $\mathcal{P}_i$ in $G_{i-1,i}$ with initial node $\pi_i(\gamma)$ and terminal node $\pi_i(\gamma')$ which traverses each edge in $G_{i-1,i}$ exactly once, if the latter is a path graph, and exactly two such walks, corresponding to the two possible orientations of $G_{i-1,i}$, if the latter is a cycle. There are the following cases, illustrated in Example 2.4.5, to consider:

Case 1: The relative interior of B^{ε} contains neither $v_{\min}$ nor $v_{\max}$. Then, all the $G_{i-1,i}$ are path graphs and conditions (2.4.3) hold by degeneration of fibers, as in the special case $u = v_{\min}$ and $v = v_{\max}$ of Example 2.4.4.

Case 2: The relative interior of B^{ε} contains exactly one of $v_{\min}$ and $v_{\max}$, say $v_{\min}$. Then, the $G_{i-1,i}$ associated to fibers $f^{-1}(t) \cap B^{\varepsilon}$ with $t < f(u)$ are cycles and all others are path graphs which degenerate to cycles as the value of f approaches $f(u)$. Clearly, the cycles can be uniquely oriented, so that the resulting walks $\mathcal{P}_i$ satisfy conditions (2.4.3).

Case 3: The relative interior of B^{ε} contains both $v_{\min}$ and $v_{\max}$. Then, the $G_{i-1,i}$ associated to fibers $f^{-1}(t) \cap B^{\varepsilon}$ with $f(u) < t < f(v)$ are path graphs and the rest are cycles which can be uniquely oriented, so that the resulting walks $\mathcal{P}_i$ satisfy conditions (2.4.3).

Thus, Proposition 2.4.3 applies in all cases and we may conclude that $d_G(\gamma, \gamma') \leq f_2(\mathcal{F}^{\varepsilon})$ for each $\varepsilon \in \{+, -\}$. Hence,

$$d_G(\gamma, \gamma') \leq \frac{f_2(\mathcal{F}^{+}) + f_2(\mathcal{F}^{-})}{2} = f_2(P)/2$$

and the proof follows. □

Example 2.4.5. Let $P = X(10)$ be the stacked polytope shown in Figure 1.3. The following two situations illustrate the three cases within the proof of Theorem 2.4.1.

(a) Consider the f-monotone paths on P

$$\begin{aligned} \gamma &= (v_1, v_3, v_6, v_9, v_{10}), \\ \gamma' &= (v_1, v_3, v_5, v_8, v_9, v_{10}), \end{aligned}$$

presented as sequences of vertices. Then, the cycle C has edges with vertex sets $\{v_3, v_5\}$, $\{v_5, v_8\}$, $\{v_8, v_9\}$, $\{v_6, v_9\}$ and $\{v_3, v_6\}$, and one of the $\mathcal{F}^{\varepsilon}$ consists of the faces of the facets of P with vertex sets $\{v_3, v_5, v_6\}$, $\{v_5, v_6, v_8\}$ and $\{v_6, v_8, v_9\}$ and falls in the first case of the proof, while the other consists of the faces of the remaining thirteen facets of P and falls in the third case. Three flips are needed to reach γ' from γ across $\mathcal{F}^{\varepsilon}$ in the former case, and thirteen flips in the latter.

(b) Consider also the f-monotone paths

$$\begin{aligned} \gamma &= (v_1, v_3, v_6, v_9, v_{10}), \\ \gamma'' &= (v_1, v_3, v_4, v_5, v_8, v_9, v_{10}). \end{aligned}$$

Now C has six edges with vertex sets $\{v_3, v_4\}$, $\{v_4, v_5\}$, $\{v_5, v_8\}$, $\{v_8, v_9\}$, $\{v_6, v_9\}$ and $\{v_3, v_6\}$, and one of the $\mathcal{F}^\varepsilon$ consists of the faces of the facets of P with vertex sets $\{v_1, v_2, v_3\}$, $\{v_1, v_2, v_4\}$, $\{v_1, v_3, v_4\}$, $\{v_2, v_3, v_5\}$, $\{v_2, v_4, v_5\}$, $\{v_3, v_5, v_6\}$, $\{v_5, v_6, v_8\}$ and $\{v_6, v_8, v_9\}$, while the other consists of the faces of the remaining eight facets of P. Both fall in the second case of the proof. The fibers $f^{-1}(t) \cap B^\varepsilon$ are path graphs for $f(v_3) < t < f(v_9)$ in either case, and cycles for $t \leq f(v_3)$ or $t \geq f(v_9)$ in the two cases, respectively. $\square$

Proof of Theorem 1.3.2. As we have already mentioned, the lower bound of (1.3.3) follows from Lemma 2.0.1. The upper bound follows from Theorem 2.4.1 and the obvious inequalities $\nu(\gamma, \gamma') \leq \lfloor (n-1)/2 \rfloor$ and $f_2(P) \leq 2n - 4$. $\square$

QUESTION 2.4.6. What is the exact value of the maximum diameter in Theorem 1.3.2? In particular, is it equal to the lower bound given there for every n?

Proof of Proposition 2.4.3. Consider indices $0 < k \leq m \leq \ell < n$ and denote by K and L the graphs of partial f-monotone paths on $\mathcal{F}$ which arise as inverse limits of the subdiagrams

$$G_{k-1,k} \xrightarrow{\alpha_k} G_k \xleftarrow{\beta_k} G_{k,k+1} \xrightarrow{\alpha_{k+1}} \cdots \xrightarrow{\alpha_{m-1}} G_{m-1} \xleftarrow{\beta_{m-1}} G_{m-1,m} \tag{2.4.4}$$

and

$$G_{m,m+1} \xrightarrow{\alpha_{m+1}} G_{m+1} \xleftarrow{\beta_{m+1}} \cdots \xleftarrow{\beta_{\ell-1}} G_{\ell-1,\ell} \xrightarrow{\alpha_\ell} G_\ell \xleftarrow{\beta_\ell} G_{\ell,\ell+1} \tag{2.4.5}$$

of (2.4.2), respectively. Let us call a *polygon* any 2-dimensional face of $\mathcal{F}$ which intersects the fiber $f^{-1}(t) \cap \|\mathcal{F}\|$ for some $t_{k-1} < t < t_m$ in the case of (2.4.4) and any 2-dimensional face of $\mathcal{F}$ which intersects the fiber $f^{-1}(t) \cap \|\mathcal{F}\|$ for some $t_m < t < t_{\ell+1}$ in the case of (2.4.5). Thus, the polygons are exactly the 2-dimensional faces of $\mathcal{F}$ in the case of the entire diagram (2.4.2) and are in one-to-one correspondence with the edges of $G_{m-1,m}$ in the special case $k = m$ of (2.4.4). Define similarly the graph H of partial f-monotone paths on $\mathcal{F}$ and its polygons from the subdiagram

$$G_{k-1,k} \xrightarrow{\alpha_k} \cdots \xleftarrow{\beta_{m-1}} G_{m-1,m} \xrightarrow{\alpha_m} G_m \xleftarrow{\beta_m} G_{m,m+1} \xrightarrow{\alpha_{m+1}} \cdots \xleftarrow{\beta_\ell} G_{\ell,\ell+1} \tag{2.4.6}$$

of (2.4.2) and note that there are natural restriction maps $\pi_K : G(\mathcal{F}, f) \to K$, $\pi_L : G(\mathcal{F}, f) \to L$ and $\pi_H : G(\mathcal{F}, f) \to H$.

We assume that there exist a walk $\mathcal{Q}$ in K with initial node $\pi_K(\gamma)$ and terminal node $\pi_K(\delta)$ which traverses each polygon of (2.4.4) exactly once and a walk $\mathcal{R}$ in L with initial node $\pi_L(\gamma)$ and terminal node $\pi_L(\delta)$ which traverses each polygon of (2.4.5) exactly once, such that $\pi_i(\mathcal{Q}) = \mathcal{P}_i$ for $k \le i \le m$ and $\pi_i(\mathcal{R}) = \mathcal{P}_i$ for $m < i \le \ell + 1$. As a consequence, there exists a walk $\mathcal{P}$ in H with initial node $\pi_H(\gamma)$ and terminal node $\pi_H(\delta)$ which traverses each polygon of (2.4.6) exactly once, such that $\pi_i(\mathcal{P}) = \mathcal{P}_i$ for $k \le i \le \ell + 1$. The proposition then follows by applying the claim several times, for instance when $k = 1$ and $m = \ell$, for $m \in [n-1]$.

To prove the claim, we only need to patch $\mathcal{Q}$ and $\mathcal{R}$ along the walk $\alpha_m(\mathcal{P}_m) = \beta_m(\mathcal{P}_{m+1})$ in G_m. Any two nodes ζ of K and η of L produce by concatenation a node $\zeta * \eta$ of H, provided that the terminal edge of ζ and the initial edge of η have equal images under α_m and β_m, respectively. Let $\zeta_0, \zeta_1, \ldots, \zeta_q$ be the successive nodes of $\mathcal{Q}$ and $\eta_0, \eta_1, \ldots, \eta_r$ be the successive nodes of $\mathcal{R}$. By our assumptions, we have $\zeta_0 * \eta_0 = \pi_K(\gamma) * \pi_L(\gamma) = \pi_H(\gamma)$ and $\zeta_q * \eta_r = \pi_K(\delta) * \pi_L(\delta) = \pi_H(\delta)$. We define $\mathcal{P}$ to have nodes of the form $\zeta_i * \eta_j$, starting with $\zeta_0 * \eta_0$, so that the node immediately following $\zeta_i * \eta_j$ is

$$\begin{cases} \zeta_{i+1} * \eta_j, & \text{if well defined,} \\ \zeta_i * \eta_{j+1}, & \text{if well defined but } \zeta_{i+1} * \eta_j \text{ is not,} \\ \zeta_{i+1} * \eta_{j+1}, & \text{otherwise.} \end{cases} \tag{2.4.7}$$

We leave to the reader to verify that, because $\alpha_m(\mathcal{P}_m) = \beta_m(\mathcal{P}_{m+1})$, this is a well defined walk in H with initial node $\zeta_0 * \eta_0 = \pi_H(\gamma)$ and terminal node $\zeta_q * \eta_r = \pi_H(\delta)$. By construction, we have $\pi_i(\mathcal{P}) = \pi_i(\mathcal{Q})$ for $k \leq i \leq m$ and $\pi_i(\mathcal{P}) = \pi_i(\mathcal{R})$ for $m < i \leq \ell + 1$, and hence $\pi_i(\mathcal{P}) = \mathcal{P}_i$ for $k \leq i \leq \ell + 1$. Finally, we note that $\mathcal{P}$ traverses the polygons traversed by $\mathcal{Q}$ or $\mathcal{R}$ which do not intersect the fiber $f^{-1}(t_m) \cap \|\mathcal{F}\|$ by steps which move $\zeta_i * \eta_j$ to the first two paths shown in (2.4.7), respectively, each exactly once by our assumptions on $\mathcal{Q}$ and $\mathcal{R}$, and the 2-dimensional faces of $\mathcal{F}$ which intersect $f^{-1}(t_m) \cap \|\mathcal{F}\|$ by steps which move $\zeta_i * \eta_j$ to the third path shown in (2.4.7), each exactly once by our assumptions on $\mathcal{P}_m$ and $\mathcal{P}_{m+1}$, and that these are precisely the polygons of (2.4.6). $\square$

2.5. Distribution of Lengths of Monotone Paths

Monotone diameters and monotone heights are closely related to the number of iterations in the simplex method. In this section, we randomly sample paths from three different classes of polytopes: 3-dimensional simple polytopes, Birkhoff polytopes and Traveling Salesman polytopes. We present histograms of the distribution of certain statistics computed from monotone diameters and monotone heights of polytopes for these classes. The distributions might help us to establish some understanding on the behavior of the simplex method.

2.5.1. The Monotone Diameter and the Monotone Height of 3-dimensional Simple Polytopes. We use the random 3-dimensional polytope generator, which applies cutting plane method to generate 50,000 random 3-dimensional simple polytopes. The polytopes generated by this method will be roundish and will not have a similar behavior of polytopes of special design, like a Klee-Minty cube. By adjusting the number of cuts, we are able to sample polytopes with different numbers of facets, ranging from eight facets to 80 facets. We separate these polytopes into groups, and keep the groups that contain at least 30 different polytopes, which are groups from eight facets to 72 facets.

Then, we apply a fixed set of 26 objective functions on each polytope to obtain the characteristics. These objective functions come from the set $F = \{[x_1, x_2, x_3]$ s.t. $x_1 = 0, \pm 1$; $x_2 = 0, \pm 1$; $x_3 = 0, \pm 1\} \setminus [0, 0, 0]$ with a little perturbation to reduce the possible degeneracy.

In this sampling, we are interested in two ratios: the *monotone diameter ratio* (monotone diameter divided by number of facets) and the *monotone height ratio* (monotone height divided by number of facets). Figure 2.5, 2.6 and 2.7 are the results for the monotone diameter ratio, and each subgraph is a normalized histogram, representing the ratio for the polytopes with the same number of facets. The x-axis of each subgraph represents the monotone diameter ratio, and the y-axis shows the percentage of occurrence.

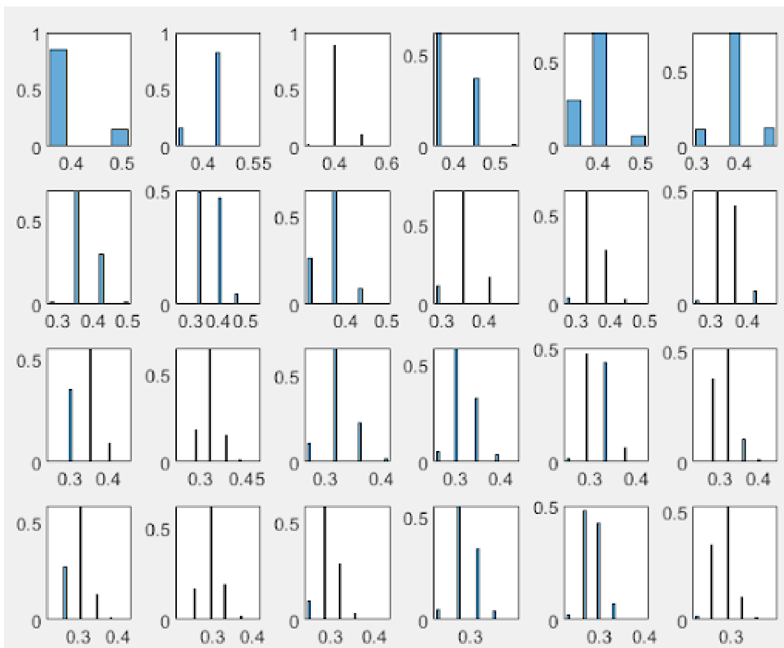

FIGURE 2.5. The monotone diameter ratio for polytopes with 8-31 facets.

From these graphs, we can see that each subgraph looks like a skewed normal distribution. Hence, we observe that, in general, as the number of facets increases, the mean of the monotone diameter ratio decreases. From this observation, we find it to be interesting to put all these ratios into one normalized histogram. We decide to give an equal weight to each number of facets from eight to 72. Figure 2.12 shows the histogram of the occurrence of each ratio:

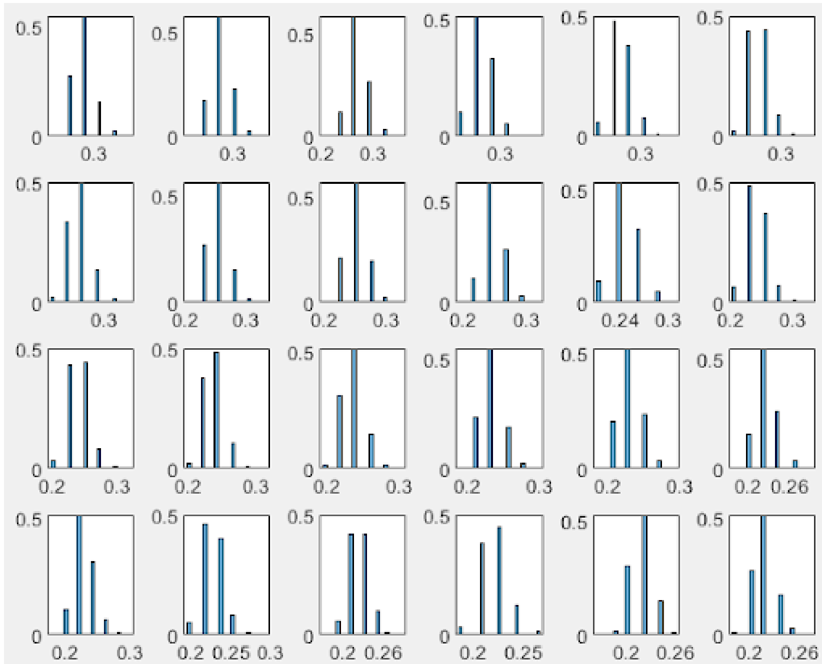

FIGURE 2.6. The monotone diameter ratio for polytopes with 32-55 facets.

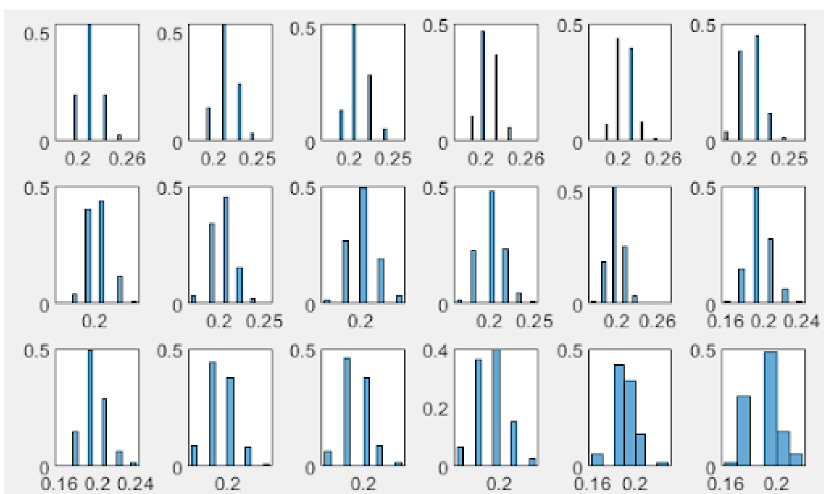

FIGURE 2.7. The monotone diameter ratio for polytopes with 56-72 facets.

If our observation is correct that the mean of the monotone diameter ratio decreases as the number of facets increases, we will see more bars on the left side of the 0.2 on the x-axis, and the ratio may approach zero or a small number as the number of facets increases.

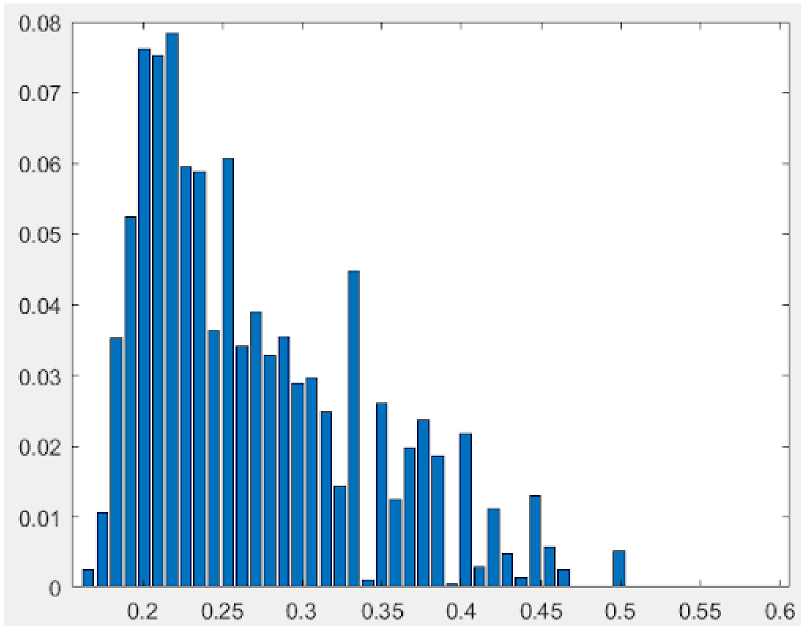

FIGURE 2.8. Weighted monotone diameter histogram.

The monotone height ratio also follows a similar pattern, and the results of the monotone height ratio are shown in the Figure 2.9, 2.10 and 2.11. The x-axis represents the monotone height ratio, and the y-axis represents the percentage of occurrence. Similarly, each subgraph looks like skewed normal distribution, and the mean is decreasing as the number of facets increases:

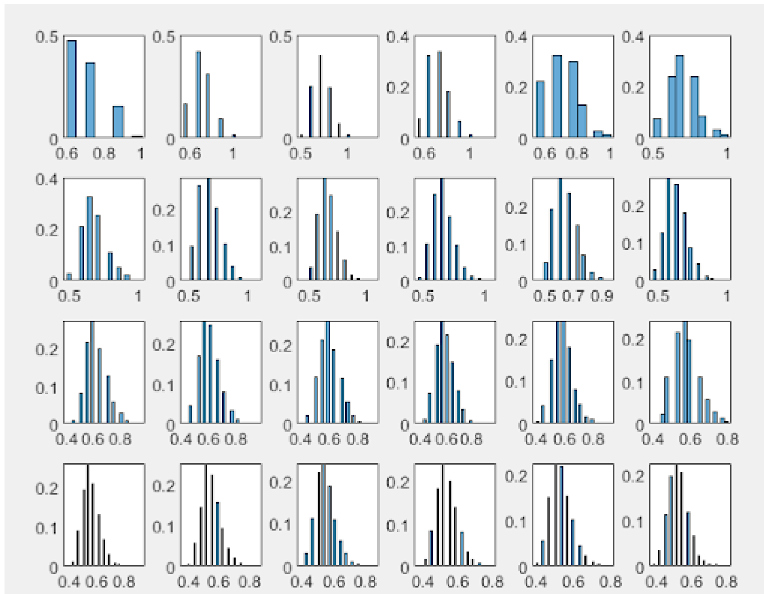

FIGURE 2.9. The monotone height ratio for polytopes with 8-31 facets.

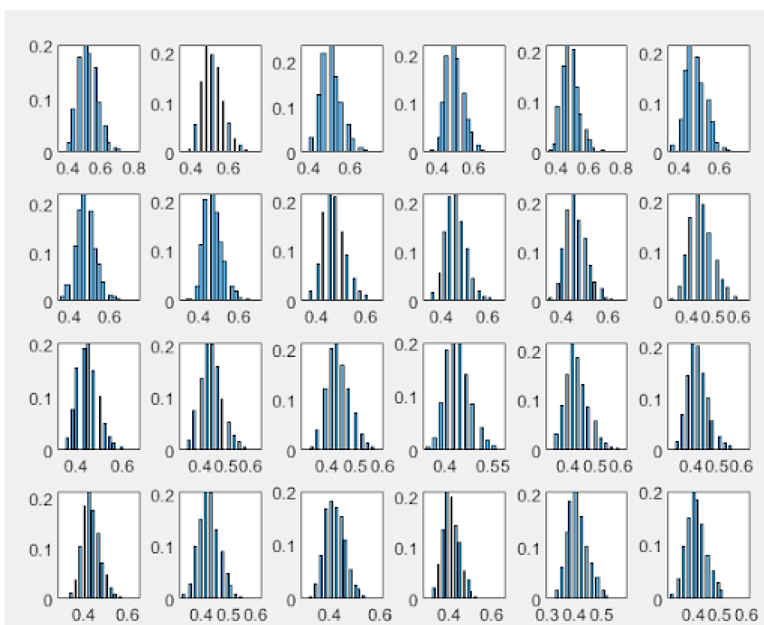

FIGURE 2.10. The monotone height ratio for polytopes with 32-55 facets.

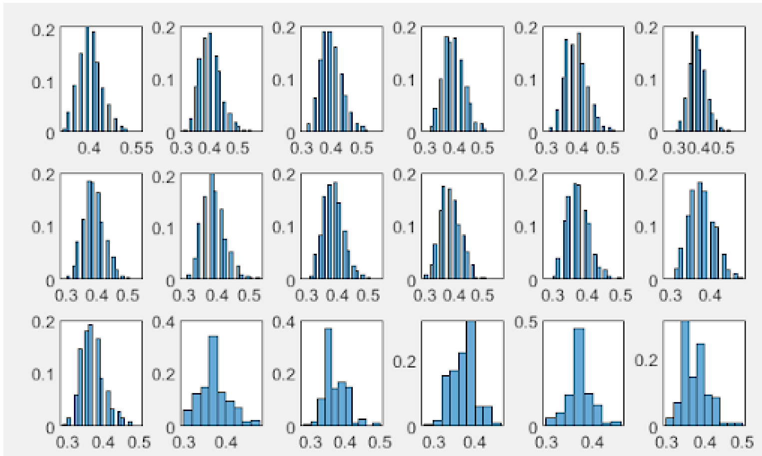

FIGURE 2.11. The monotone height ratio for polytopes with 56-72 facets.

We also aggregate the ratios into one normalized histogram presented below, and it seems to follow the pattern we observed in the monotone diameter ratio:

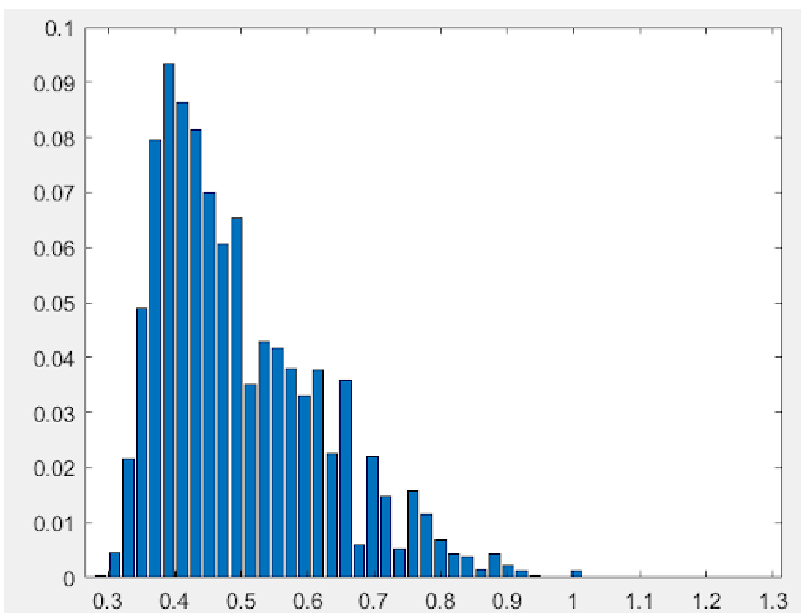

FIGURE 2.12. Weighted monotone diameter histogram.

2.5.2. The Monotone Height of Birkhoff Polytope. B_n is a convex polytope that comes from an $n \times n$ doubly stochastic matrix, a matrix with non-negative real entries and every column sum and every row sum equal to 1. It is also called the Birkhoff Polytope [**DLLY08**]. We want to explore some features of the Birkhoff Polytope, with a focus on the monotone height. We generated B_3, B_4, B_5, and B_6 based on the idea in [**DLLY08**]. Then, we applied 500 different objective functions on B_3, B_4, and B_5, as well as 150 different objective functions on B_6. The following table is what we get:

Birkhoff Polytope	# of runs	# of facets	# of vertices	diameter	monotone diameter	monotone height	# of paths	# of arborescences
B_3	500	9	6	1	1	5	16	120
B_4	500	16	24	2	2	15 to 23	188340 to 2812400	0.75e21 to 1.83e21
B_5	500	25	120	2	2	80 to 107	9.26e25 to 3.79e30	0.16e180 to 4.36e180
B_6	150	36	720	2	2 to 3	453 to 514	2.13e138 to 5.99e150	Out of bound

Among these numbers, we are interested in the distribution of the monotone height, so we ran more experiments on B_4, B_5, and B_6. The distributions are shownin Figure 2.13, 2.14 and 2.15 respectively. The x-axis represents the monotone height, and the y-axis represents the number of occurrences of each monotone height:

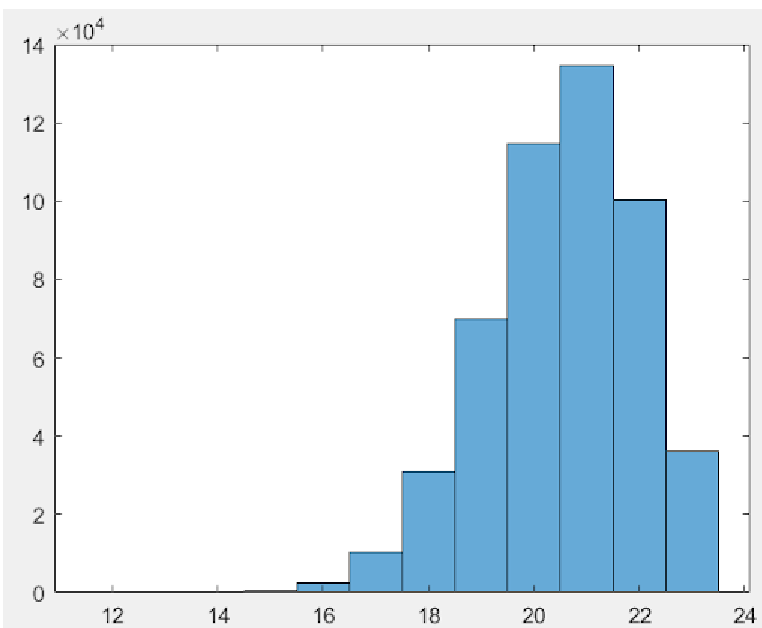

FIGURE 2.13. The monotone height distribution of B_4.

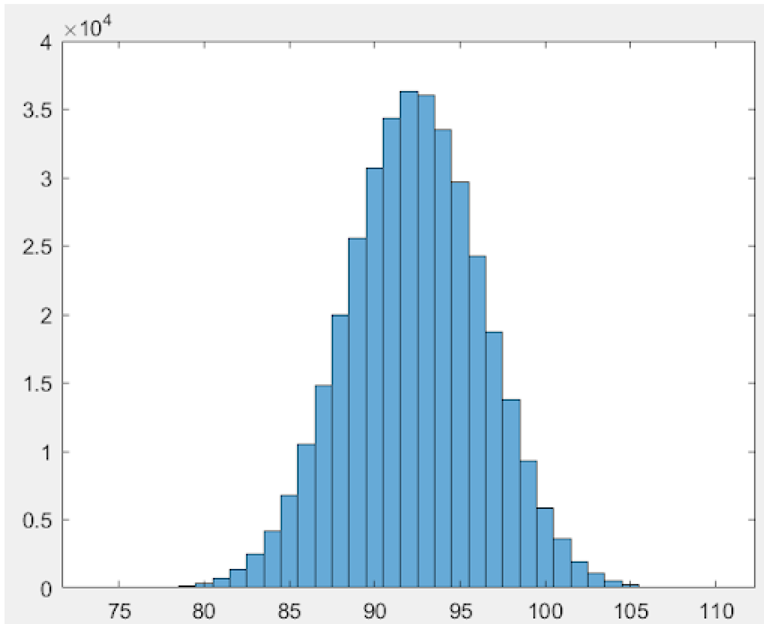

FIGURE 2.14. The monotone height distribution of B_5.

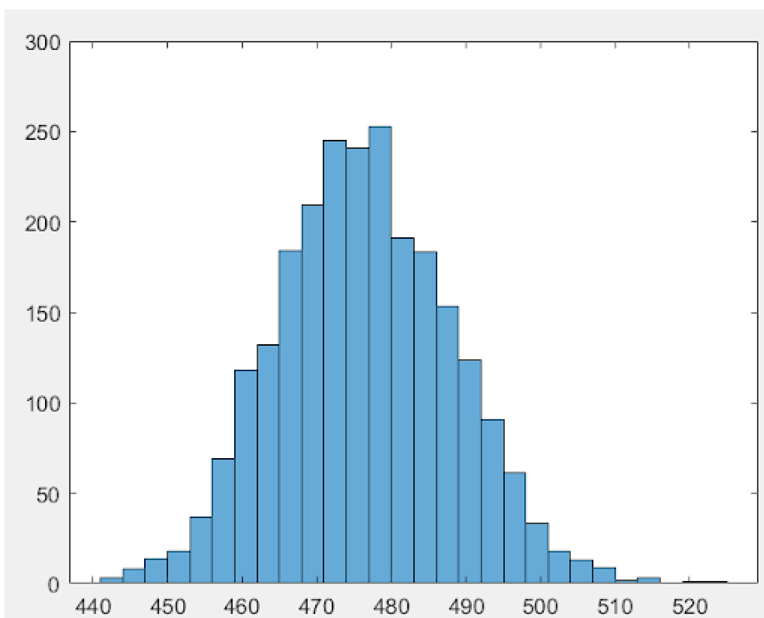

FIGURE 2.15. The monotone height distribution of B_6.

We also tried to compute the features for B_7. Unfortunately, B_7 requires too much computation power. Instead of calculating the exact monotone height, we choose to calculate the monotone height divided by ten to speed up the process. Below is what we get for 60 runs. The numbers in x-axis times ten is roughly the monotone height:

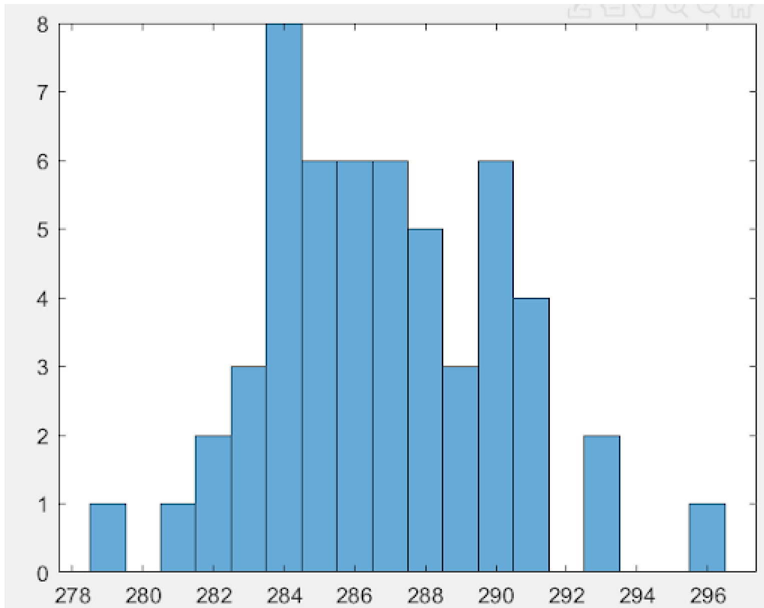

FIGURE 2.16. The monotone height distribution of B_7.

From these figures, we can see that the monotone height of B_n based on randomly generated objective functions is normally distributed with a mean is growing exponentially in terms of n. In particular, there can be a monotone path traveling through all the vertices of the directed polyhedral graph coming from B_4 and special objective functions.

2.5.3. The Monotone Height of the Traveling Salesman Polytope. Traveling salesman problem (TSP) is an NP-hard problem, and there is no known efficient algorithm for solving the TSP. The setting of the problem is that a salesman wants to minimize the cost to travel all n cities and return to the starting city, without going into the same city twice. In the traveling salesman polytope (TSP polytope for short), each vertex represents a solution to the TSP (a Hamiltonian cycle in a complete graph of n nodes), and each entry in the objective function represents the cost to travel between a pair of cities. By varying the objective functions, we obtain TSPs with different costs between cities.

We are interested in the distribution of the monotone height of the TSP polytope. However, the computation power needed for the n-city TSP polytopes is exponential, and we could only compute the features for the 5-city and the 6-city TSP polytopes, as well as the distribution of the monotone height of the 6-city and the 7-city TSP polytope. The following table shows the features for the 5-city and the 6-city TSP polytopes:

TSP Polytope	# of runs	# of facets	# of vertices	diameter	monotone diameter	monotone height	# of paths	# of arborescences
5-city	500	20	12	2	2	10	682	1.81e7
6-city	500	100	60	2	2	39 to 55	7.69e12 to 2.08e15	9.15e71 to 3.49e72

From the features, we can see that the 5-city TSP polytope is not very interesting. Therefore, we run more experiments to compute the monotone height for the 6-city and the 7-city TSP polytopes. The histogram of the distribution of the monotone height are shown below in Figure 2.17 and Figure 2.18, with the x-axis representing the monotone height and the y-axis representing the number of occurrences of each monotone height.

One thing interesting for the 6-city TSP polytope is that the maximum of the monotone height of the 6-city TSP polytope is 58, which means that, starting from one vertex, the simplex method needs to go through all except one vertex (59 in total except the starting vertex) to reach the optimum, which is very inefficient. The minimum of the monotone height is 32, and we do not know if it could be lower. The mean is 46.58, the standard deviation is 2.69, and the median is 47. The minimum for the monotone height for the 7-city TSP polytope in 8800 runs is 193, and the max is 266. The mean is 223.3, the standard deviation is 9.43, and the median is 223.

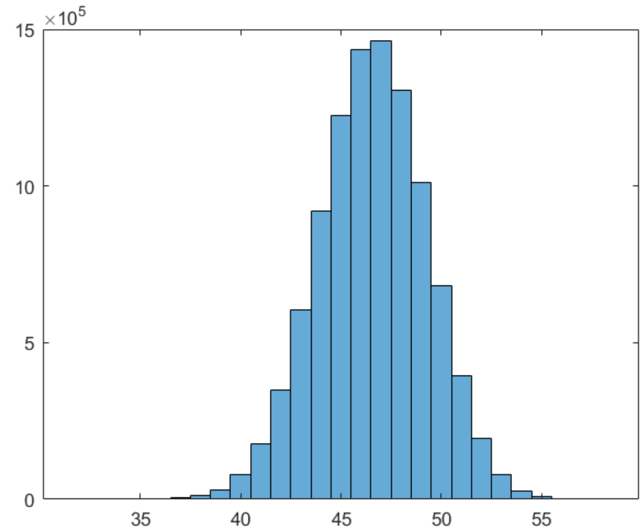

FIGURE 2.17. The distribution of the monotone height for 6-city TSP polytope using 10,000,000 random objective functions.

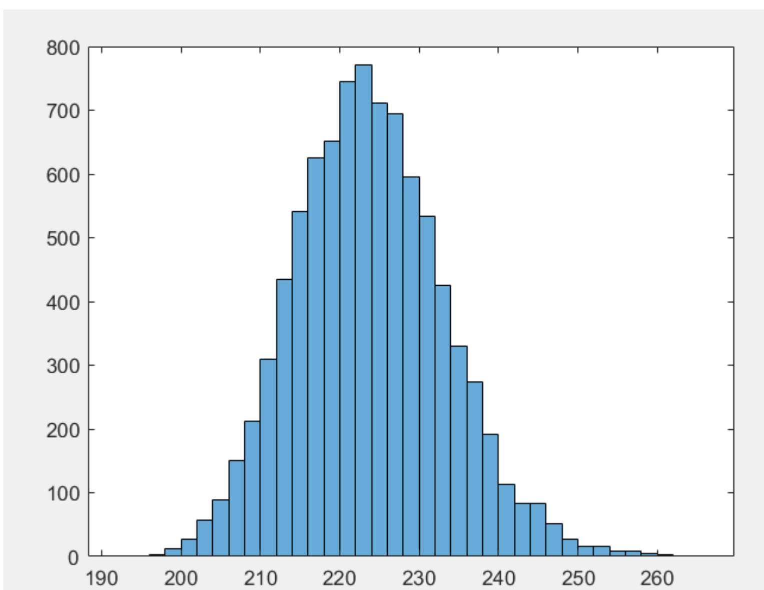

FIGURE 2.18. The distribution of the monotone height for 7-city TSP polytope using 8,800 random objective functions.

2.6. Open Problems for Directed Polytope Graphs

Based on our sampled distributions for monotone heights of Birkhoff polytopes and TSP polytopes, it is natural to ask why all these distributions look unimodal. *Is it possible that for certain classes of polytopes, the monotone heights will follow a normal distribution?* For general 3-dimensional polytopes, it is less obvious that the monotone diameter ratio comes from uniform distribution. However, it will be an interesting question to consider under what conditions will monotone diameter ratios follow a bimodal distribution or even a multimodal distribution?

Apart from the distribution of monotone paths, the following tables summarize our extremal results on monotone paths and arborescences and indicate the problems which remain open.

# of arborescences		all polytopes	simple polytopes
$d = 3$	upper bound	Theorem 2.2.4	Corollary 2.2.3
	lower bound	Theorem 2.2.5	
$d \geq 4$	upper bound	Theorem 2.2.4	Corollary 2.2.3
	lower bound	Question 2.2.6	Corollary 2.2.3

Table 2.1. Summary for f-arborescences

# of monotone paths		all polytopes	simple polytopes
$d = 3$	upper bound	Theorem 2.3.2	Conjecture 2.3.6, Proposition 2.3.7
	lower bound	Theorem 2.3.4	
$d \geq 4$	upper bound	Remark 2.3.3	open
	lower bound	Proposition 2.3.5	

Table 2.2. Summary for f-monotone paths

diameter of flip graph		all polytopes	simple polytopes
$d = 3$	upper bound	Theorem 1.3.2, Question 2.4.6	open
	lower bound	Conjecture 2.6.1	
$d \geq 4$	upper bound	open	open
	lower bound	open	open

Table 2.3. Summary for the diameter of flip graphs

We have no reason to doubt that Question 2.2.6 on the minimum number of f-arborescences in dimensions $d \geq 4$ and Question 2.4.6 on the maximum diameter of flip graphs in dimension 3 have positive answers. For the minimum diameter of flip graphs, we expect that the diameter of $G(P, f)$ is bounded below by the integral part of half the number of facets for every 3-dimensional polytope P. In particular, we expect that the following conjecture is true.

Conjecture 2.6.1. *The minimum diameter of $G(P, f)$, when P ranges over all 3-dimensional polytopes with n vertices and f ranges over all generic linear functionals on P, is equal to $\lfloor (n+5)/4 \rfloor$ for every $n \geq 4$. This can be achieved by simple polytopes for every even n.*

The outdegrees of the vertices of $\omega(P, f)$ play an important role in the proofs of Theorems 1.3.1 and 1.3.3. It seems a very interesting problem to characterize, or at least obtain significant information about, the possible multisets of these outdegrees when P ranges over all polytopes of given dimension and number of vertices and f ranges over all generic linear functionals on P. Finally, it would be interesting to address the questions raised in this thesis for *coherent* f-monotone paths as well. Their number typically grows much slower than the total number of f-monotone paths [**ADLRS00**].

CHAPTER 3

Diameters of Cocircuit Graphs of Oriented Matroids

In this chapter we will first prove Lemma 1.4.3 and Lemma 1.4.4 in Section 3.1. Then we discuss diameter for oriented matroids of small rank and corank as stated in Theorem 1.3.5 in Section 3.2. We finish by showing the connection between the diameter bound for oriented matroids and the diameter bound for polytopes in Section 3.3.

3.1. Reductions and Lower Bounds

Klee and Walkup [**KW67**] showed that the maximal diameter among all d-dimensional polytopes with n facets is achieved by a simple polytope. Their argument was straightforward: if P is a d-polytope with n facets that is not simple, then slightly perturbing the facets of P will produce a simple polytope whose diameter is at least as large as that of P. Our goal in this section is to prove an analogous result for oriented matroids. First we require some definitions, see [**BLVS$^+$99**, Section 7.1 and 7.2] for more details.

Let $\mathcal{M}$ be an oriented matroid on ground set E. An *extension* of $\mathcal{M}$ is an oriented matroid $\widetilde{\mathcal{M}}$ on a ground set $\widetilde{E}$ that contains E, such that the restriction of $\widetilde{\mathcal{M}}$ to E is $\mathcal{M}$. We say $\widetilde{\mathcal{M}}$ is a *single element extension* if $|\widetilde{E}\backslash E| = 1$. For any single element extension $\widetilde{\mathcal{M}}$, there is a unique way to extend cocircuits of $\mathcal{M}$ to cocircuits of $\widetilde{\mathcal{M}}$. Specifically, there is a function

$$\sigma : \mathcal{C}^*(\mathcal{M}) \to \{+, -, 0\}$$

such that $\sigma(-Y) = -\sigma(Y)$ for all $Y \in \mathcal{C}^*(\mathcal{M})$ and

$$\{(Y, \sigma(Y)) : \ Y \in \mathcal{C}^*(\mathcal{M})\} \subseteq \mathcal{C}^*(\widetilde{\mathcal{M}}).$$

That is, $(Y, \sigma(Y))$ is a cocircuit of $\widetilde{\mathcal{M}}$ for every cocircuit Y of $\mathcal{M}$. The functions $\sigma : \mathcal{C}^* \to \{+, -, 0\}$ that correspond to single element extensions are called *localizations*. Furthermore, $\widetilde{\mathcal{M}}$ is uniquely determined by σ, with

$$\mathcal{C}^*(\widetilde{\mathcal{M}}) = \{(Y, \sigma(Y)) :\ Y \in \mathcal{C}^*(\mathcal{M})\} \cup$$
$$\{(Y^1 \circ Y^2, 0) :\ Y^1, Y^2 \in \mathcal{C}^*(\mathcal{M}), \sigma(Y^1) = -\sigma(Y^2) \neq 0, S(Y^1, Y^2) = \emptyset, \rho(Y^1 \circ Y^2) = 2\}.$$

Here ρ is the rank function and $\circ$ is the composition of covectors.

Now we are ready to define the perturbation map on non-uniform oriented matroids.

Definition 3.1.1. [**BLVS$^+$99**, Theorem 7.3.1] Let $\mathcal{M}$ be an oriented matroid of rank $r \geq 2$ on E. If $f \in E$ is not a coloop, then $\mathcal{M}$ is a single element extension of a rank r oriented matroid $\mathcal{M}_0 := \mathcal{M} \setminus f$, with localization σ_f. Let $\overline{W} \in \mathcal{C}^*(\mathcal{M}_0)$ be a cocircuit with $\sigma_f(\overline{W}) = 0$, meaning $W = (\overline{W}, 0)$ is a cocircuit of $\mathcal{M}$. Then the local perturbation $\mathcal{M}'$ of $\mathcal{M}$ can be defined as a single element extension of $\mathcal{M}_0$ with localization

$$\sigma_{LP}(\overline{Y}) = \begin{cases} + & \text{if } \overline{Y} = \overline{W}, \\ - & \text{if } \overline{Y} = -\overline{W}, \\ \sigma_f(\overline{Y}) & \text{otherwise.} \end{cases}$$

We can now reduce the general diameter problem to the case of uniform oriented matroids, as promised by Lemma 1.4.3.

Proof. (of Lemma 1.4.3)

Let $\mathcal{M}$ be a non-uniform oriented matroid. We may assume without loss of generality that, $\mathcal{M}$ does not have any loops, coloops or parallel elements since removing them will not affect the cocircuit graph of $\mathcal{M}$. Note that there exists $W \in \mathcal{C}^*(\mathcal{M})$ with $|W^0| > r - 1$. Pick an arbitrary $f \in W^0$. Let $\mathcal{M}_0 := \mathcal{M} \setminus f$ and let $\mathcal{M}'$ be the perturbed oriented matroid defined in Definition 3.1.1. We will show $\operatorname{diam}(\mathcal{M}) \leq \operatorname{diam}(\mathcal{M}')$. In addition, if $\mathcal{M}$ is realizable, then we will show the perturbed $\mathcal{M}'$ can also be made realizable. From this, it will follow that for all n and r, the optimal bound $\Delta(n, r)$ is achieved by a uniform oriented matroid.

Denote by $\{X^1, X^2, \ldots, X^k\} = \{X \in \mathcal{C}^*(\mathcal{M}_0) : \sigma_f(X) = -,\ S(\overline{W}, X) = \emptyset,\ \rho(\overline{W}, X) = 2\}$. Note that $X^1, \ldots, X^k$ are exactly the cocircuits that are adjacent to $\overline{W}$ in $G^*(\mathcal{M}_0)$ before the extension with $\sigma_f(X^i) = -$. Let $Z^i = (X^i \circ \overline{W}, 0)$. After the perturbation by σ_{LP}, W is mapped

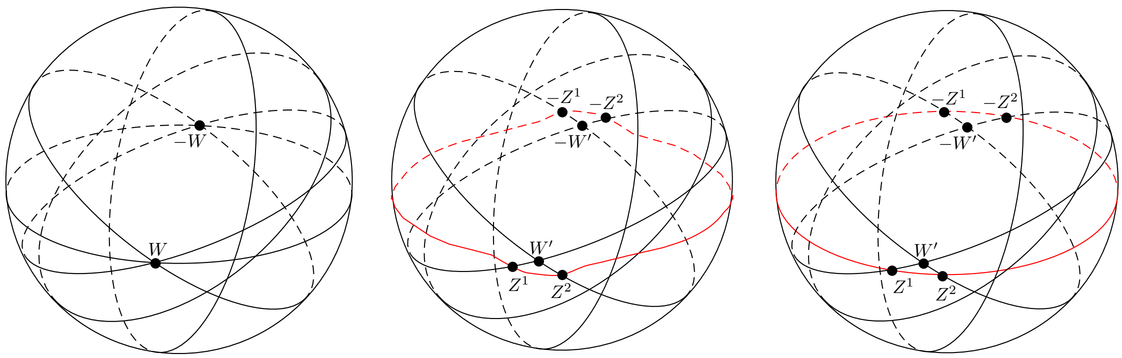

FIGURE 3.1. A non-uniform oriented matroid (left), a local perturbation (center), and a realizable local perturbation (right).

to $W' = (\overline{W}, +)$. Since σ_{LP} and σ_f only differ on $\pm\overline{W}$, it follows that $\pm Z^1, \dots, \pm Z^k$ are all the cocircuits created by this perturbation. After the perturbation, each edge of the form $\{W, X^i\}$ in $G^*(\mathcal{M})$ is subdivided into two edges $\{W, Z^i\}$ and $\{Z^i, X^i\}$ (similarly $\{-W, -X^i\}$ is subdivided into $\{-W', Z^i\}$ and $\{-Z^i, -X^i\}$).

Now let $X, Y \in \mathcal{C}^*(\mathcal{M})$ be any two cocircuits of $\mathcal{M}$ such that $X, Y \in \mathcal{C}^*(\mathcal{M}')$ (X, Y could be $\pm W$, in this case we just consider $\pm W'$ in $\mathcal{M}'$). Take a minimal path between X and Y on $G^*(\mathcal{M}')$, and replace any elements of $\{\pm W', \pm Z^1, \dots, \pm Z^k\}$ with $\pm W$ respectively. This gives us a path (potentially having repeated elements and not necessarily shortest) between X and Y in $\mathcal{M}$. Now if we pick $X, Y \in \mathcal{C}^*(\mathcal{M})$ that realize the diameter of $\mathcal{M}$, since $d_{\mathcal{M}}(X, Y) \leq d_{\mathcal{M}'}(X, Y)$, we have $\operatorname{diam}(\mathcal{M}) = d_{\mathcal{M}}(X, Y) \leq d_{\mathcal{M}'}(X, Y) \leq \operatorname{diam}(\mathcal{M}')$.

Now suppose $\mathcal{M}$ is realizable. Let $\mathcal{H} = \{H_1, \dots, H_n\}$ be the hyperplane arrangement corresponding to $\mathcal{M}$ (with f corresponding to H_n). Let $H_i = \{\mathbf{x} : \mathbf{x}^T \mathbf{v}_i = 0\}$, and $\mathbf{w}$ be the vector realizing W. Note that we have $\mathbf{w}^T \mathbf{v}_n = 0$ since the last entry of W is 0. Consider $\mathbf{y}$, the minimizer of $\mathbf{x}^T \mathbf{v}_n$ over all cocircuits of $\mathcal{M}$ subject to $\mathbf{x}^T \mathbf{v}_n > 0$. Now we replace H_n by $H'_n = \{\mathbf{x} : \mathbf{x}^T((1-\epsilon)\mathbf{v}_n + \epsilon\mathbf{y}) = 0\}$, in which the choice of ϵ will be made later. Note that,

$$\mathbf{x}^T((1-\epsilon)\mathbf{v}_n + \epsilon\mathbf{y}) = \mathbf{x}^T\mathbf{v}_n - \epsilon\mathbf{x}^T\mathbf{v}_n + \epsilon\mathbf{x}^T\mathbf{y}\,.$$

We first pick the sign of ϵ so that $\epsilon\,\mathbf{w}^T\mathbf{y} > 0$; as a result, $\mathbf{w} \in H_n'^{+}$ and $-\mathbf{w} \in H_n'^{-}$. Then we take $|\epsilon|$ small enough such that $|\mathbf{x}^T\mathbf{v}_n| > |\epsilon(\mathbf{x}^T\mathbf{v}_n - \mathbf{x}^T\mathbf{w}')|$ for all $\mathbf{x}$ vectors that realize a cocircuit in $\mathcal{M}$ (this choice of ϵ exists since the number of cocircuits is finite and we may scale the vector). The construction ensures that all cocircuits, except those that lie on H_n with degeneracy, will have the same sign as defined in Definition 3.1.1. As a result $\mathcal{H}' = \{H_1, \ldots, H_{n-1}, H_n'\}$ corresponds to some realizable oriented matroid $\mathcal{M}'$ after some local perturbations (the composition of perturbation maps on all cocircuits with degeneracy on H_n (including W) as defined in Definition 3.1.1).

To conclude, we have decreased the number of pairs of (W, f) with $|W^0| > r - 1$ and $W_f = 0$ without decreasing the diameter. By continuing this procedure, we will eventually obtain an oriented matroid in which no such pair of (W, f) can be found, or equivalently $|X^0| = r - 1$ for all $X \in \mathcal{C}^*(\mathcal{M})$. Hence $\Delta(n, r)$ will be achieved by a uniform oriented matroid. $\square$

Hence it suffices to study uniform oriented matroids for the purpose of bounding $\Delta(n, r)$. The bound in Conjecture 1.4.1 can be rewritten as $\Delta(n, r) \leq n - (r - 1) + 1$. For polytopes, $n - (r - 1) + 1 = n - d + 1$. It may seem mysterious that the bound here is one more than the Hirsch bound, so we will pause for a moment to discuss this. We begin by proving Lemma 1.4.4 from the Introduction.

Proof. (of Lemma 1.4.4)

Recall that if cocircuits Z and W are adjacent in $G^*(\mathcal{M})$, then there are elements $e \in Z^0 \setminus W^0$ and $e' \in W^0 \setminus Z^0$ such that $Z^0 = (W^0 \setminus \{e'\}) \cup \{e\}$. In other words, when we move from Z to W, we see $Z_e = 0$ change to become $W_e \neq 0$ and $Z_{e'} \neq 0$ change to become $W_{e'} = 0$. Therefore, we will say that each edge in $G^*(\mathcal{M})$ encodes two "basic transformations", which are changes to the cocircuit that transform a nonzero entry into a zero entry or vice versa.

Now we consider the differences in the sign patterns of X and Y. For each $e \in S(X, Y)$ we require two basic transformations to move from X to Y: one to transform X_e to 0, and another to transform 0 to $-X_e = Y_e$. For each $e \in X^0 \setminus Y^0$, we require one basic transformation to transform 0 to Y_e. Similarly, for each $e \in Y^0 \setminus X^0$, we require one basic transformation to transform X_e to 0. Therefore, moving from X to Y requires at least $2|S(X, Y)| + |X^0 \setminus Y^0| + |Y^0 \setminus X^0| = 2|S(X, Y)| + 2|X^0 \setminus Y^0|$ basic transformations. Thus $d_{\mathcal{M}}(X, Y) \geq |S(X, Y)| + |X^0 \setminus Y^0|$.

Now we examine the case where $X = -Y$ more closely. In this case, $S(X, Y) = \operatorname{supp}(X)$ and $X^0 = Y^0$. Pick a shortest path from X to Y in $G^*(\mathcal{M})$ and let Z be the neighbor of X on this path. Then $|S(Y,Z)| = n-r$ and $|Z^0 \setminus Y^0| = 1$, so $d_{\mathcal{M}}(Y,Z) \geq n-r+1$ by the above argument. Therefore, $d_{\mathcal{M}}(X,Y) = 1 + d_{\mathcal{M}}(Y,Z) \geq n-r+2$.

Next, consider the case $|X^0 \setminus Y^0| \leq 1$. We show that the equality holds for expression (1.4.1).

Let $A \subseteq X^0 \cap Y^0$ have cardinality $r-2$. If $|X^0 \setminus Y^0| = 1$, then $A = X^0 \cap Y^0$; otherwise, $X = -Y$ and we can pick $r-2$ elements arbitrarily from $X^0 = Y^0$. Let $\{s_e : e \in E\}$ be the pseudospheres in the Folkman-Lawrence representation of $\mathcal{M}$ and let $S_A = \bigcap_{e \in A} s_e$. Because $\mathcal{M}$ is uniform, we know $S_A \approx \mathbb{S}^1$.

We saw above that in general $d_{\mathcal{M}}(X,Y) \geq 1 + |S(X,Y)|$. On the other hand, the elements of $S(X,Y)$ are in bijective correspondence with cocircuits along the shortest path from X to Y in S_A. Indeed, if Z is such a cocircuit, then Z and $-Z$ are antipodal vertices on S_A, so they constitute a 0-dimensional pseudosphere whose positive side contains one of X or Y and whose negative side contains the other. Thus the distance from X to Y on S_A is exactly $1 + |S(X,Y)|$. This proves $d_{\mathcal{M}}(X,Y) \leq 1 + |S(X,Y)|$. □

3.2. Results for small oriented matroids

3.2.1. Computer-based results for oriented matroids with few elements.

Finschi and Fukuda [**FF01a**] computed the exact number of isomorphism classes of uniform oriented matroids, and gave a representative of each isomorphism class, when $n \leq 9$ and in small rank/corank when $n = 10$. We established Conjecture 1.4.1 for all of these examples using computers.

	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
$r=2$	1	1	1	1	1	1	1	1	1
$r=3$		1	1	1	4	11	135	4382	312356
$r=4$			1	1	1	11	2628	9276595	unknown
$r=5$				1	1	1	135	9276595	unknown
$r=6$					1	1	1	4382	unknown
$r=7$						1	1	1	312356
$r=8$							1	1	1
$r=9$								1	1
$r=10$									1

TABLE 3.1. Number of uniform oriented matroids for $n \leq 10$.

Each isomorphism class is encoded by its chirotope representation. Chirotopes, or basis orientations, are one of the equivalent axiomatic systems for oriented matroids (see [**BLVS$^+$99**, Section 3] for more details). For a given oriented matroid on ground set E, the chirotope defines a mapping $\chi : E^r \to \{-, 0, +\}$. For a realizable oriented matroid with vector configuration $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$,

$$\chi(\lambda_1, \ldots, \lambda_r) = \mathrm{sign}(\det(\mathbf{v}_{\lambda_1}, \mathbf{v}_{\lambda_2}, \ldots, \mathbf{v}_{\lambda_r})).$$

The data can be found on Finschi and Fukuda's Homepage of Oriented Matroids [**FF01b**]. Given a chirotope map χ of an oriented matroid of rank r on $E = \{1, 2, \ldots, n\}$, we can generate the cocircuits by computing the set $\mathcal{C}^*(\chi) = \{(\chi(\lambda, 1), \chi(\lambda, 2), \ldots, \chi(\lambda, n)) : \lambda \in E^{r-1}\}$. Since $\mathcal{M}$ is uniform, we add an edge between $X, Y \in \mathcal{C}^*(\mathcal{M})$ if and only if $|X^0 \cap Y^0| = r - 2$ and $|S(X,Y)| = 0$. For $n = 9$, $r = 5$ and $n = 10$, $r = 7$, the chirotope maps are missing in the original dataset. However we can look at their duals ($n = 9$, $r = 4$ and $n = 10$, $r = 3$) and consider the set of circuits instead. See Appendix for the pseudocode of computing the set of cocircuits and circuits.

After finding all the cocircuits and edges, we used the Python NetworkX package [**Dev18**] to construct the cocircuit graph. This package has a method for computing the diameter of a graph, and also for determining the distance between any pairs of vertices. Table 3.1 shows the number of isomorphism classes (up to reorientation) of uniform oriented matroids of cardinality n and rank r. We used a MacBook Pro with quad-core 2.2GHz Intel i7 processor, as well as UC Davis Math servers to construct the cocircuit graphs and compute their diameters. When $n = 9$, $r = 4, 5$ the algorithm takes the longest to terminate. On average, each instance of an oriented matroid takes about 0.36 seconds to compute, resulting in around 38.7 days to complete the checking of all oriented matroids of cardinality nine and rank four.

We investigate other interesting questions such as whether the shortest path between two cocircuits on the same tope stays on the tope. Our code is available on Github.[1] Based on our explicit computations, we derive the following theorem for small matroids, as promised in the introduction.

Theorem 3.2.1. *Let $r \leq n \leq 9$ and $\mathcal{M} \in UOM(n, r)$, then* $\mathrm{diam}(G^*(\mathcal{M})) = n - r + 2$. *Moreover, if $X, Y \in \mathcal{C}^*(\mathcal{M})$ with $X \neq -Y$ and $n \leq 9$, then $d_{\mathcal{M}}(X, Y) \leq n - r + 1$.*

[1] `https://github.com/zzy1995/OrientedMatroid`

3.2.2. Results in low rank.

As a next step, we explore Conjecture 1.4.1 in low rank. If $\mathcal{M} \in UOM(n, 2)$, then the cocircuit graph $G^*(\mathcal{M})$ is a cycle on $2n$ vertices, so its diameter is $n = n - r + 2$. Thus Conjecture 1.4.1 holds trivially when $r = 2$. Now we move on to study uniform oriented matroids of rank three.

THEOREM 3.2.2. *Let $\mathcal{M} \in UOM(n, 3)$, then* $\mathrm{diam}(G^*(\mathcal{M})) = n - r + 2 = n - 1$.

PROOF. Let $\mathcal{M} \in UOM(n, 3)$ and $X, Y \in \mathcal{C}^*(\mathcal{M})$. If $X = -Y$, then $d_{\mathcal{M}}(X, Y) = n - r + 2$ by Lemma 1.4.4. If $|X^0 \setminus Y^0| = 1$, then $d_{\mathcal{M}}(X, Y) \leq n - r + 1$ by Lemma 1.4.4. So we only need to consider the case that $|X^0 \setminus Y^0| \geq 2$. But $|X^0| = |Y^0| = r - 1 = 2$, so this means $X^0 \cap Y^0 = \emptyset$.

Identify the elements of $E(\mathcal{M})$ with $\{1, 2, \ldots, n\}$. Let $\mathcal{P}(\mathcal{M})$ be the Folkman-Lawrence representation of $\mathcal{M}$ with pseudospheres $\{s_1, \ldots, s_n\}$.

Without loss of generality we can assume $X^0 = \{1, 2\}$ and $Y^0 = \{3, 4\}$. Let $\mathcal{M}'$ denote the restriction of $\mathcal{M}$ to $\{1, 2, 3, 4\} \subseteq E$. The Folkman-Lawrence representation of $\mathcal{M}'$ is obtained from $\mathcal{P}(\mathcal{M})$ by removing s_i for all $i > 4$. Up to relabeling and reorientation, there is only one uniform oriented matroid of rank three on four elements. We can further assume $X_3 = X_4 = Y_1 = Y_2 = +$. In particular, there are cocircuits W and Z such that $W^0 = \{1, 3\}$, $Z^0 = \{2, 4\}$, and $W_2 = W_4 = Z_1 = Z_3 = +$. Consider the region, $D = s_1^+ \cap s_2^+ \cap s_3^+ \cap s_4^+ \subseteq \mathcal{P}(\mathcal{M})$. This is the quadrilateral region bounded by cocircuits X, Y, Z, and W in Figure 3.2.

We claim that for each $i > 4$, the pseudosphere s_i can intersect the boundary of D in at most two points. Indeed, suppose s_i intersects the boundary of D at a point $p_0 \in s_j$ for some $j \in \{1, 2, 3, 4\}$. Because $\mathcal{M}$ is uniform, $p_0 \notin \{X, Y, Z, W\}$, so s_j is unique. Let $\varphi_i : [0, 1] \to \mathcal{P}(\mathcal{M})$ be a parametrization of s_i. We can assume $\varphi_i(0) = p_0$ and $\varphi_i(t)$ passes into the interior of D for sufficiently small $t > 0$. Let t_1 be the next time when $\varphi_i(t_1)$ is on the boundary of D. Assume $\varphi_i(t_1) \in s_k$. Once again, s_k is unique because $\mathcal{M}$ is uniform. Further, $k \neq j$ because otherwise s_j would intersect s_i in at least four points: $\varphi_i(0)$, $\varphi_i(t_1)$, and their antipodes.

When $t > 0$ is sufficiently small, $\varphi_i(t) \in s_j^+ \cap s_k^+$. When $t > t_1$ and $t - t_1$ is sufficiently small, $\varphi_i(t) \in s_j^+ \cap s_k^-$. By the definition of a pseudosphere arrangement, the image of φ_i cannot cross back into s_k^+ before it crosses into s_j^-. However, any other points where the image of φ_i could

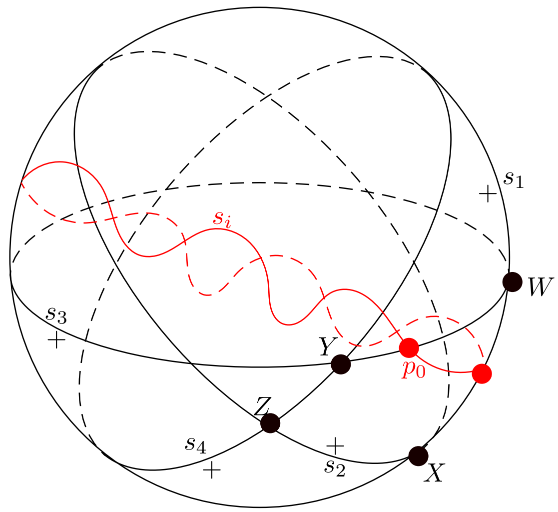

FIGURE 3.2. The unique rank-3 pseudosphere arrangement with four pseudolines.

intersect the boundary of D lie in $s_j^+ \cap s_k^+$. Thus $\varphi_i(0)$ and $\varphi_i(t_1)$ are the only points of intersection of s_i with the boundary of D.

Now we consider two paths from X to Y in $G^*(\mathcal{M})$. The first path P_W travels from X to W along s_1, then from W to Y along s_3. The second path P_Z travels from X to Z along s_2, then from Z to Y along s_4. Let $\ell(P_W)$ and $\ell(P_Z)$ denote the lengths of these paths. Initially, in $\mathcal{M}'$, $\ell(P_W) = \ell(P_Z) = 2$.

For each $i > 4$, the pseudosphere s_i meets the boundary of D in at most two points. This means $\ell(P_W) + \ell(P_Z)$ increases by at most two when we add s_i back into $\mathcal{P}(\mathcal{M})$. Thus, in $\mathcal{M}$,

$$\ell(P_W) + \ell(P_Z) \leq 4 + 2(n-4) = 2n - 4.$$

By the pigeonhole principle, either $\ell(P_W) \leq n-2$ or $\ell(P_Z) \leq n-2$, so $d_{\mathcal{M}}(X, Y) \leq n-2$. $\square$

COROLLARY 3.2.3. *Let $r \geq 3$ and $\mathcal{M} \in UOM(n, r)$. If $X, Y \in \mathcal{C}^*(\mathcal{M})$ and $|X^0 \setminus Y^0| = 2$, then $d_{\mathcal{M}}(X, Y) \leq n - r + 1$.*

PROOF. Let $A = X^0 \cap Y^0$. Let $\{s_e : e \in E(\mathcal{M})\}$ be the pseudospheres in the Folkman-Lawrence representation of $\mathcal{M}$ and let $S_A = \bigcap_{e \in A} s_e$. Because $\mathcal{M}$ is uniform, $|A| = r-3$ and hence $S_A \approx \mathbb{S}^2$ is the Folkman-Lawrence representation of the uniform oriented matroid $\mathcal{M}/A \in UOM(n-r+3, 3)$.

Both X and Y are cocircuits on S_A and clearly $X \neq -Y$, so by Theorem 3.2.2,

$$d_{\mathcal{M}}(X,Y) \leq d_{\mathcal{M}/A}(X,Y) \leq (n-r+3) - 2 = n-r+1.$$

□

Recall that in the proof of Theorem 3.2.2 for oriented matroids of rank three, the two cocircuits we choose lie on four different hyperplanes, and they form a combinatorial square. Each additional hyperplane will intersect the square twice, which implies that one of the two paths will at most increase by one. Santos (personal communication) has pointed out that this cannot be directly extended to establish Conjecture 1.4.1 in rank four. For a realizable uniform oriented matroid of rank four, six hyperplanes will enclose a combinatorial cube. For concreteness, we can consider the cube with $-1 \leq x_i \leq 1$ for all $i = 1, 2, 3$.

Figure 3.3 illustrates three edge-disjoint paths, colored red, green, and blue, from $(-1,-1,-1)$ to $(1,1,1)$. Here, $(-1,-1,-1)$ is the vertex incident to the three dotted edges, and $(1,1,1)$ is its polar opposite. The three images show slices of the cube by hyperplanes $x_i + x_j = (2-\varepsilon_k)x_k$ for all choices of $\{i,j,k\} = \{1,2,3\}$ and with ε_1, ε_2, and ε_3 all distinct. Each plane intersects two edges incident to $(-1,-1,-1)$ and two edges incident to $(1,1,1)$, and hence increases the total length of all three paths by at least four. If each of the remaining $n-6$ hyperplanes has one of the three illustrated types (with the ε_k generic) then the total length of the red, blue, and green paths will be at least $4(n-6)+9$. If there are approximately $\frac{n-6}{3}$ hyperplanes of each type, then each of the red, green, and blue paths will have length at least $\left\lfloor \frac{4}{3}n \right\rfloor - 5$.

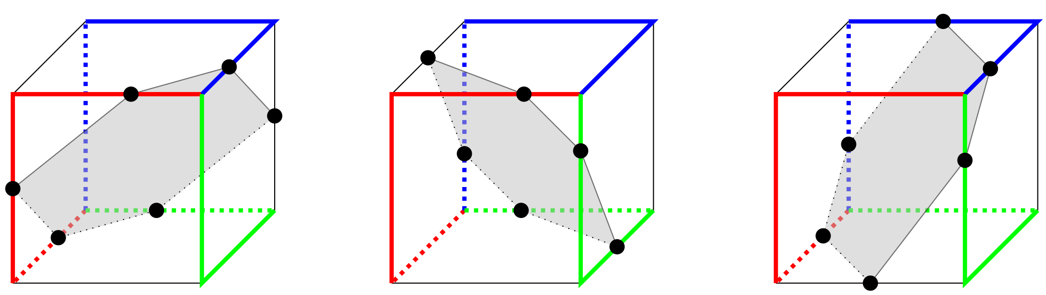

FIGURE 3.3. Hyperplanes $x_i + x_j = (2-\varepsilon_k)x_k$ slicing the ± 1 cube for $\{i,j,k\} = \{1,2,3\}$.

3.2.3. Results in low corank.

Recall that the corank of an oriented matroid of rank r on n elements is equal to $n-r$.

LEMMA 3.2.4. *Let $\mathcal{M} \in UOM(n,r)$ with $n-r=k$ for $k \geq 0$. Then*

$$\operatorname{diam}(G^*(\mathcal{M})) \leq \max\{\operatorname{diam}(G^*(\mathcal{M}')) : \mathcal{M}' \in UOM(r'+k, r'),\ 2 \leq r' \leq k+2\}.$$

PROOF. Let $\mathcal{M}$ be a uniform oriented matroid of corank k, and let $X, Y \in \mathcal{C}^*(\mathcal{M})$ such that $\operatorname{diam}(G^*(\mathcal{M})) = d_{\mathcal{M}}(X,Y)$. If $Y = -X$, we are done, since by, Lemma 1.4.4 the diameter of any uniform oriented matroid of corank k is at least $k+2$, and $d_{\mathcal{M}}(X,-X) = k+2$. So we assume that $Y \neq -X$.

Consider the contraction $\mathcal{M}' = \mathcal{M}/(X^0 \cap Y^0)$, and let X' and Y' be the images of X and Y under this contraction. Let $r' = \operatorname{rank}(\mathcal{M}')$ and $n' = |E(\mathcal{M}')|$. We know that $\mathcal{M}'$ is uniform because $\mathcal{M}$ is. Note that $(X')^0 \cap (Y')^0 = \emptyset$ by construction, so $\operatorname{supp}(X') \cup \operatorname{supp}(Y') = E(\mathcal{M}')$. In addition, since $\mathcal{M}'$ is uniform, $|\operatorname{supp}(X')| = |\operatorname{supp}(Y')| = k+1$. This shows $|E(\mathcal{M}')| \leq 2(k+1)$. Further, $\operatorname{supp}(X') \neq \operatorname{supp}(Y')$ because $Y \neq -X$, so $|\operatorname{supp}(X') \cup \operatorname{supp}(Y')| \geq k+2$, which implies $2 \leq r' \leq k+1$, as $|E(\mathcal{M}')| = r'+k$.

Then, as $X', Y' \in \mathcal{C}^*(\mathcal{M}')$ and $G^*(\mathcal{M}')$ is a subgraph of $G^*(\mathcal{M})$, we have that $\operatorname{diam}(G^*(\mathcal{M})) = d_{\mathcal{M}}(X,Y) \leq d_{\mathcal{M}'}(X',Y') \leq \operatorname{diam}(G^*(\mathcal{M}'))$. Thus, we conclude that for every matroid $\mathcal{M}$ of corank k, there exists a matroid $\mathcal{M}' \in UOM(r'+k, r')$, where $2 \leq r' \leq k+2$, such that $\operatorname{diam}(G^*(\mathcal{M})) \leq \operatorname{diam}(G^*(\mathcal{M}'))$. $\square$

THEOREM 3.2.5. *Let $\mathcal{M} \in UOM(n,r)$ with $n-r \leq 4$. Then* $\operatorname{diam}(G^*(\mathcal{M})) = n-r+2$.

PROOF. If $n-r \leq 3$ the theorem follows directly from Lemma 3.2.4 and Theorem 3.2.1.

When $n-r=4$, by Lemma 3.2.4 we have $\Delta(r+4,r) \leq \max_{2 \leq r' \leq 6}\{\Delta(r'+4,r')\}$. However, by Theorem 3.2.1, for $2 \leq r' \leq 5$, $\max\{\Delta(r'+4,r')\} \leq r'+4-r'+2 = 6$. So we only need to consider $\mathcal{M} \in UOM(10,6)$. Let $X, Y \in \mathcal{C}^*(\mathcal{M})$ be such that $\operatorname{diam}(G^*(\mathcal{M})) = d_{\mathcal{M}}(X,Y)$. If $Y = -X$, the result holds by Lemma 1.4.4. If $X^0 \cap Y^0 \neq \emptyset$, then as in Theorem 3.2.4, the contraction $\mathcal{M}' = \mathcal{M}/(X^0 \cap Y^0)$ satisfies $d_{\mathcal{M}}(X,Y) \leq \operatorname{diam}(\mathcal{M}')$. Since $|E(\mathcal{M}')| \leq 9$, the result holds by Theorem 3.2.1. So we may assume that $X^0 \cap Y^0 = \emptyset$.

Define $\mathcal{T} = X \circ Y$. Then, by Lemma 1.1.6 the graph $G(\mathcal{T})$ of $\mathcal{T}$ is isomorphic to the graph $G_A(\mathcal{A})$ of $\mathcal{A}$, where $\mathcal{A}$ is the abstract polytope on the covector of $\mathcal{T}$ with dimension 5 on 10 elements.

However, by [**AD74**, Theorem 7.1] the diameter of $G_A(\mathcal{A})$ is 5, implying that $d_{\mathcal{M}}(X,Y) = 5$. Noting that $d_{\mathcal{M}}(X,-X) = 6$, we conclude that $\mathrm{diam}(G^*(\mathcal{M})) = 6$ which completes the proof. □

Note that while the theorems about coranks in this subsection are for uniform oriented matroids, they are valid for general oriented matroids due to Lemma 1.4.3. Now we are ready to combine all the results in this section to prove Theorem 1.4.5.

Proof. (of Theorem 1.4.5)

The proof of part (a) for small oriented matroids is in Theorem 3.2.1. The proof of part (b) for rank three oriented matroids is in Theorem 3.2.2. The proof of part (c) for oriented matroids of corank no more than four is in Theorem 3.2.5.

□

3.3. An Improved Quadratic Diameter Bound

Proof. (of Theorem 1.4.6)

By Lemma 1.4.3, it suffices to consider the case that $\mathcal{M}$ is uniform. We prove the claim by induction on $|X^0 \backslash Y^0|$. If $|X^0 \backslash Y^0| = 1$, then $d_{\mathcal{M}}(X,Y) \leq n-r+1$ by Lemma 1.4.4. If $|X^0 \backslash Y^0| = 2$, then $d_{\mathcal{M}}(X,Y) \leq n-r+1$ by Corollary 3.2.3.

Now we move on to the inductive step. Suppose $|X^0 \backslash Y^0| = \ell \geq 3$. Pick any element $e \in Y^0 \backslash X^0$, and consider the coline U, with $U^0 = Y^0 \setminus \{e\}$. Note that $|U^0 \setminus X^0| = \ell - 1$.

Now we look more carefully at the coline U, which is a cycle on $2(n-r+2)$ cocircuits. We distinguish ℓ pairs of these cocircuits. For each element $f \in X^0 \setminus U^0$, there is a cocircuit Z^f with $(Z^f)^0 = U^0 \cup \{f\}$. Because $|X^0 \setminus U^0| = \ell$, there are ℓ such pairs of antipodal cocircuits, which we denote as $\pm Z^1, \ldots, \pm Z^\ell$ for simplicity.

The cocircuits Y and $-Y$ are antipodal on U, and hence partition U into two halves, each of which contains $n-r+1$ cocircuits. Assume without loss of generality that $Z^1, \ldots, Z^\ell$ all lie on one half of the coline (as it is partitioned by Y and $-Y$), and further that $Z^1, \ldots, Z^\ell$ are ordered by their distance from Y, with Z^1 closest to Y and Z^ℓ farthest.

Because there are $(n-r+2)-(\ell+1) = n-r-\ell+1$ remaining pairs of antipodal circuits on U, and at most one element from each pair can lie on the arc from Z^1 to $-Z^\ell$ that contains Y, it follows that there exists a path of length at most $\lfloor \frac{n-r-\ell+1}{2} \rfloor + 1$ from Y to one of Z^1 or $-Z^\ell$ along U. For simplicity, let Z denote whichever of Z^1 and $-Z^\ell$ is closer to Y along U.

In summary, we have shown that there exists a cocircuit Z whose distance to Y is at most $\lfloor \frac{n-r-\ell+1}{2} \rfloor + 1$ with $|X^0 \setminus Z^0| = \ell - 1$. Because $\ell - 1 \neq 0$, we know $Z \neq -X$ as well. The result now follows by induction, and after reindexing with $k = \ell - 1$ we have

$$d_{\mathcal{M}}(X,Y) \leq n - r + 1 + \sum_{k=2}^{|X^0 \setminus Y^0|-1} \left(\left\lfloor \frac{n-r-k}{2} \right\rfloor + 1 \right).$$

To get Eq. (1.4.3), note that $|X^0 \setminus Y^0| \leq \min(r-1, n-r+1)$, because $|X^0 \setminus Y^0| \leq |X^0| = r-1$ and $|X^0 \setminus Y^0| \leq |E \setminus Y^0| = n - r + 1$. So, when $r \geq 4$ and $n - r \geq 2$,

$$\operatorname{diam}(G^*(\mathcal{M})) \leq n - r + 1 + \sum_{k=2}^{\min(r-2,n-r)} \left(\left\lfloor \frac{n-r-k}{2} \right\rfloor + 1 \right).$$

$\square$

3.4. Similarities to the diameter of polytopes problem and two conjectures

One could hope that $d_{\mathcal{M}}(X,Y) \leq n-r+1$ provided $X, Y \in \mathcal{C}^*(\mathcal{M})$ are not antipodal cocircuits. However, this is not the case. Matschke, Santos, and Weibel [**MSW15**] built on the methodology of Santos's original non-Hirsch polytope [**San12**] to construct a simple polytope $P_{20,40}$ of dimension 20 with 40 facets which has diameter 21. Let $\mathcal{M}_{20,40}$ be the oriented matroid obtained by lifting $P_{20,40}$ into $\mathbb{R}^{21}$ and intersecting its hyperplane arrangement with the unit sphere. Since $P_{20,40}$ is simple, $\mathcal{M}_{20,40}$ is uniform, and one of its topes is $P_{20,40}$. We will show that the oriented matroid $\mathcal{M}_{20,40} \in UOM(40, 21)$ has a pair of non-antipodal cocircuits X and Y such that $d_{\mathcal{M}_{20,40}}(X,Y) \geq 21 = n - r + 2$.

Proof. (of Proposition 1.4.7) Let X, Y be the pair of cocircuits that are of distance 21 in $P_{20,40}$. Let $E = \{1, \ldots, 40\}$. After reorientation and relabeling, we may assume that $X^0 = \{1, 2, \ldots, 20\}$, $X^+ = \{21, \ldots, 40\}$ and $Y^0 = \{21, \ldots, 40\}$, $Y^+ = \{1, \ldots, 20\}$.

Consider a shortest path, γ, from X to Y in $\mathcal{M}_{20,40}$. If each cocircuit on γ belongs to the tope $P_{20,40}$, then its length is 21. So we may suppose instead that γ contains a cocircuit Z that does not belong to $P_{20,40}$. This means $Z^- \neq \emptyset$.

Recall the notion of a “basic transformation” from the proof of Lemma 1.4.4. Each edge in the cocircuit graph accounts for two basic transformations, which change some entry on a cocircuit from $+/-$ to 0 or from 0 to $+/-$.

Let $i \in Z^-$. If $X_i = +$ and $Y_i = 0$, then walking from X to Y via Z requires at least $20+19+3 = 42$ basic transformations. This is because each $j \in X^0$ requires one basic transformation to become an element of Y^+; each $j \in X^+ \setminus \{i\}$ requires one basic transformation to become an element of Y^0, and $i \in X^+$ requires two basic transformations to become an element of Z^- and one additional transformation to subsequently become an element of Y^0. Similarly, if $X_i = 0$ and $Y_i = +$, then walking from X to Y via Z also requires at least 42 basic transformations. This tells us $d_{\mathcal{M}_{20,40}}(X, Y) \geq 21 = n - r + 2$. $\square$

We now prove Theorem 1.4.8. Let $\mathcal{M}$ be a uniform oriented matroid. We say a path $X^1, X^2, \ldots, X^k$ in the cocircuit graph $G^*(\mathcal{M})$ *stays on a tope* $\mathcal{T}$ if each cocircuit X^i is a vertex of $\mathcal{T}$.

Proof. (of Theorem 1.4.8) We used computers to look over all (chirotopes) oriented matroids with $n \leq 10$ computed by Finschi and Fukuda [**FF01a**] (we used them already earlier in the thesis). We used the Python package NetworkX [**Dev18**] to find all shortest paths between a given pair of cocircuits and verify that one of the shortest paths between them is a crabbed path. We also checked when shortest paths stay on a common tope. We found that for all $\mathcal{M} \in UOM(n, r)$ with $n \leq 8$, there exists a crabbed path from X to Y whose length is no bigger than the length of any path from X to Y in the entire cocircuit graph $\mathcal{M}$. But we eventually found a smallest counterexample in our search. This is an oriented matroid with 9 elements in rank 4. The chirotope mapping of this counterexample is

$++++++++++++++++++++++++++++++++++--++++++++$

$+++++++-++--+++++-+---+----++-++-+-+++-+--+-$

$-+++----+--++--++++----+--++--++---+-$

shown here ordered by reverse lexicographical order. That is, the first two and the last two signs in the list (shown in red there) are $\chi(1,2,3,4,5) = +, \chi(1,2,3,4,6) = +, \chi(4,6,7,8,9) = +, \chi(5,6,7,8,9) = -$.

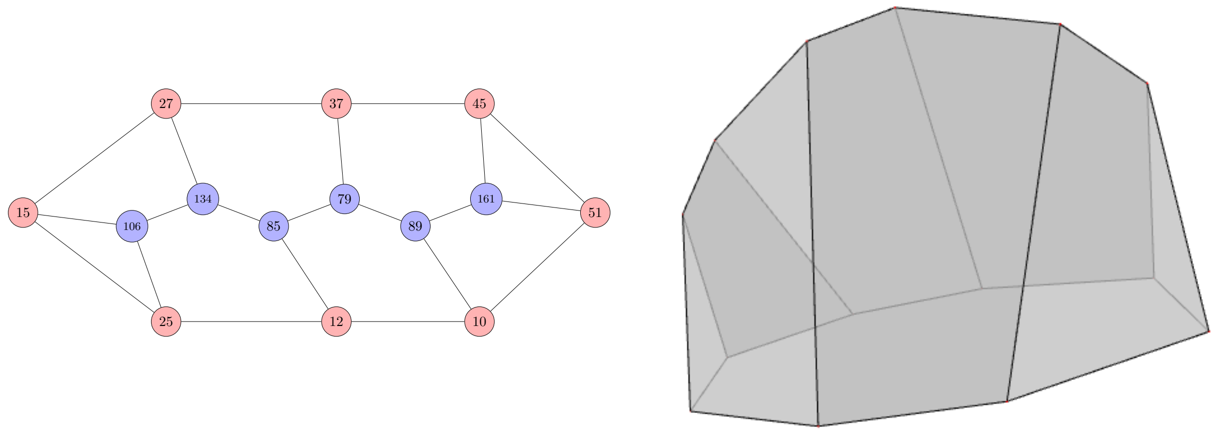

FIGURE 3.4. The subgraph induced by the tope containing X_{12} and X_{37} (left) and a realization of the tope as a 3-polytope (right).

Cocircuit $X_{12} = (0, 0, -, -, -, -, -, -, 0)$ and cocircuit $X_{37} = (0, -, -, -, 0, 0, -, -, +)$ lie on the same tope. But as shown in Figure 3.4, the shortest path between X_{12} and X_{37} goes through cocircuits X_{85} and X_{79}, where

$$X_{85} = (+, 0, -, -, 0, -, -, -, 0), X_{79} = (+, 0, -, -, 0, 0, -, -, +).$$

Note that $X_{85}^+ = \{1\} \nsubseteq X_{12}^+ \cup X_{37}^+$, and thus the shortest path is shorter than any crabbed path.

Next, we take a hyperplane arrangement that realizes the tope in Figure 3.4 as a polytope and add a tenth hyperplane that cuts through (among others) the edge between X_{37} and X_{79}, X_{85} and X_{12}, as shown in Figure 3.5. By lifting all these hyperplanes (see Figure 1.5 for intuition of what is happening, we go from three to four dimensions), we obtain the central hyperplane arrangement that yields a (realizable) uniform oriented matroid $\mathcal{M}'$ of rank 4 with 10 elements. Below are the explicit equations of these ten hyperplanes of the arrangement:

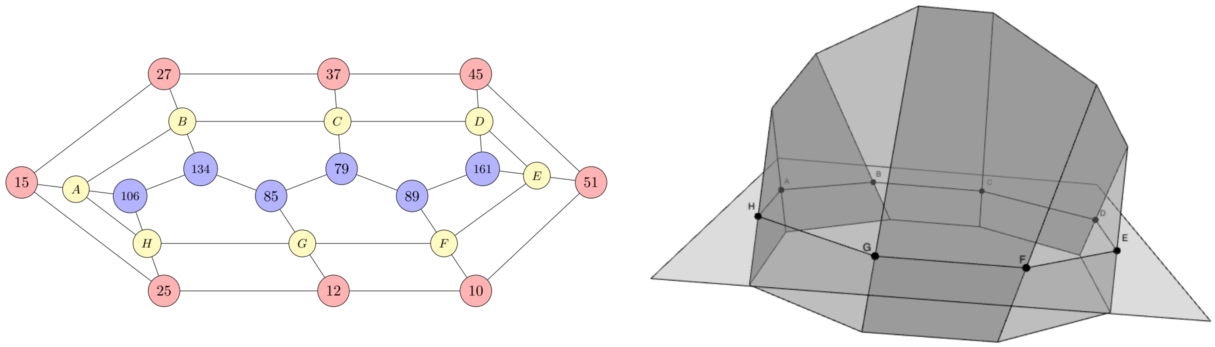

FIGURE 3.5. The subgraph after the tenth hyperplane is added, creating cocircuits $X'_A, \dots, X'_H$ (left) and a realization of the tope and the hyperplane (right). The relevant shortest path is $X'_{37} \to X'_C \to X'_{79} \to X'_{85} \to X'_G$.

$$H_0 : -8x_1 - 15.99x_2 - 9x_3 + 160z = 0,$$

$$H_1 : -56x_1 + 112x_2 - 39x_3 + 672z = 0,$$

$$H_2 : 56x_1 - 112x_2 - 39x_3 + 448z = 0,$$

$$H_3 : 8x_1 + 16x_2 - 9x_3 = 0,$$

$$H_4 : -2x_2 - x_3 + 12z = 0,$$

$$H_5 : 280x_1 - 31x_3 = 0,$$

$$H_6 : x_3 = 0,$$

$$H_7 : 2x_2 - x_3 + 4z = 0,$$

$$H_8 : -280x_1 - 31x_3 + 3360z = 0,$$

$$H_9 : x_1 + 2x_2 + 100x_3 - 300z = 0.$$

After constructing the cocircuit graph of $\mathcal{M}'$, we find that the path $X'_{37} \to X'_C \to X'_{79} \to X'_{85} \to X'_G$, going from X'_{37} to X'_G, leaves the tope they share. Their common tope is composed by the red and yellow vertices in Figure 3.5 (these are points with indices A to H). The path we proposed is shorter than any path from X'_{37} to X'_G staying on their common tope. This shows two cocircuits on a common tope while the shortest path between them leaves the tope. This completes the proof of the second part of the theorem.

□

For polytopes , it is natural to ask the following question: if two vertices lie on a common facet, does there exist a shortest path between them that stays within that facet? One can show that this property implies the *non-revisiting path property* [**KK87**], and therefore implies the (linear) Hirsch conjecture. The linear Hirsch conjecture was disproved by Santos, thus we know the polytope version of must be false starting in dimension 20. But Aviv Adler (personal communication) pointed out to us that already in three dimensions it is possible to have two vertices on a common facet while the shortest path between them leaves the facet. Our Theorem 1.4.8 demonstrates this fails for oriented matroids too and here we provided the smallest counterexample.

CHAPTER 4

Machine Learning to Improve the Simplex Method

In this chapter we are going to present the case study on how machine learning and the simplex method are related to each other. We will present several trained machine learning models on choosing pivot rules for the simplex method.

4.1. Data Generation

The existing libraries (MIPLIB 2017 [**mip18**], netlib [**Don97**] etc.) of linear programming or integer programming are too small for our training purpose. Hence we generated our own data for training and testing. We adapted the algorithms introduced by Bowly et al [**BSMBM20**]. Their method involves generating constraint matrix $\mathbf{A}$, and a solution pair (α, β). They used $\mathbf{A}, \alpha, \beta$ to generate the final linear problem maximizing $\mathbf{c}^T\mathbf{x}$ subject to $\mathbf{A}\mathbf{x} \leq \mathbf{b}$. For simplicity, we replaced the generation of variable constraint graph by generating Erdos-Renyi (ER) random graphs. For training and validation set, we generated 24634 instances of linear programming problems with number of constraints ranging from 120 to 200 and number of variables ranging from 50 to 100. For testing, we generate 7279 more instances. Note that these linear programs will most likely be characterized as "easy" problems by MIPLIB 2017. For the ER random graphs, the parameter p was drawn from $\mathcal{U}\{0.2, 0.8\}$. For other hyperparameters in generating the LP instances, we draw the coefficient mean μ_A from normal distribution $\mathcal{N}(0, 1)$, coefficient standard deviation σ_A from uniform distribution $\mathcal{U}\{1, 10\}$, primal versus slack basis γ from $\mathcal{U}\{0.2, 0.8\}$, fractional primal λ from $\mathcal{N}(0, 1)$ and Beta fraction $a = 0.5$.

After generating the LP instances, we solve our LP problems using primal simplex solver in DOcplex with default initialization. We store the number of iterations for each instance using different pivot rules. Note that the LP instances we generate may have degeneracy, and empirically there is a high likelihood of degeneracy where the constraint matrix is low-density.

4.2. Feature selection

We have two different ways of choosing features for the linear programming instances. The first method we use is a bag-of-features, where we add features based on heuristics from previous studies on the simplex method. Apart from m, n the number of constraints and the number of variable, we add three sets of features: variable constraint graph features, coefficient values, and normalized coefficients. Variable constraint graph features include the minimum, maximum, mean, and standard deviation of the degree sequences of variable nodes and constraint nodes. Coefficient values include the statistics of the coefficient matrix $\mathbf{A}$, the constraint vector $\mathbf{b}$, and the objective function $\mathbf{c}$ (i.e. he minimum, maximum, mean, standard deviation, norm of the vector, and the smallest non-zero absolute value). Finally, normalized coefficients are the statistics of row and column normalized coefficients ($\{\frac{\mathbf{A}_{ij}}{\mathbf{b}_j} | \mathbf{b}_j \neq 0\}$ and $\{\frac{\mathbf{A}_{ij}}{\mathbf{c}_j} | \mathbf{c}_j \neq 0\}$) and degree normalized coefficients ($\{\frac{\mathbf{b}_j}{deg(u_j)}\}$ and $\{\frac{\mathbf{c}_i}{deg(v_i)}\}$).

The other way we have implemented features related to the coefficient matrix $\mathbf{A}$, is the Truncated Singular Value Decomposition (SVD), which is a method of dimension reduction [**MRS08**]. The truncated SVD of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ returns three matrices U, Σ, V such that:

$$\mathbf{A} \approx U \Sigma V$$

where $U \in \mathbb{R}^{m \times k}, \Sigma \in \mathbb{R}^{k \times k}$, and $V \in \mathbb{R}^{k \times n}$, where k is the number of top singular values to keep. Multiplying U by Σ allows for the computation of an $m \times k$ matrix. Applying this procedure again to $(U\Sigma)^{\mathrm{T}}$ will then compute a $k \times k$ matrix with similar features to the original matrix $\mathbf{A}$. We choose $k = 20$ in this experiment for the best performance. We still include the features of statistics of the constraint vector $\mathbf{b}$ and objective function $\mathbf{c}$.

4.3. Experiments

We train four models to choose which pivoting strategies will perform the best on each LP instance. Two models use the bag of features that we choose for LP problems, and the other two use the SVD to replace the features of the coefficient matrix $\mathbf{A}$.

4.3.1. Gradient Boosting Decision Tree. We train two gradient boosting decision tree to predict the best pivot rule for each LP instance. The first one is an empirical hardness model,

Hyperparameters	Dantzig	Hybrid	Devex	Steepest	Greatest
learning rate	0.1	0.1	0.1	0.1	0.05
# estimators	271	137	173	173	371
max depth	5	6	4	6	6
min child weight	6	6	5	4	1
γ	0	0	0	0	0.3
subsample ratio	1	0.8	0.8	0.9	0.8
column subsample	1	1	1	0.8	0.9
regularization α	100	10	1e-5	100	1e-5

TABLE 4.1. Hyperparameters for each regressor.

that is, for all five pivot rules, we use regression on the features we selected to predict number of iterations that the solver will take using certain pivot rule. The second model is a GBDT classifier using truncated SVD as features.

Bag-of-features GBDT The first model is an empirical hardness model, where we use GBDT to do regression and predict the number of iterations each pivot rule would cost. Table 4.1 shows the hyperparameters for different pivot rules. This model results in a 67.78% accuracy on the test set with 178.5934 iterations on average.

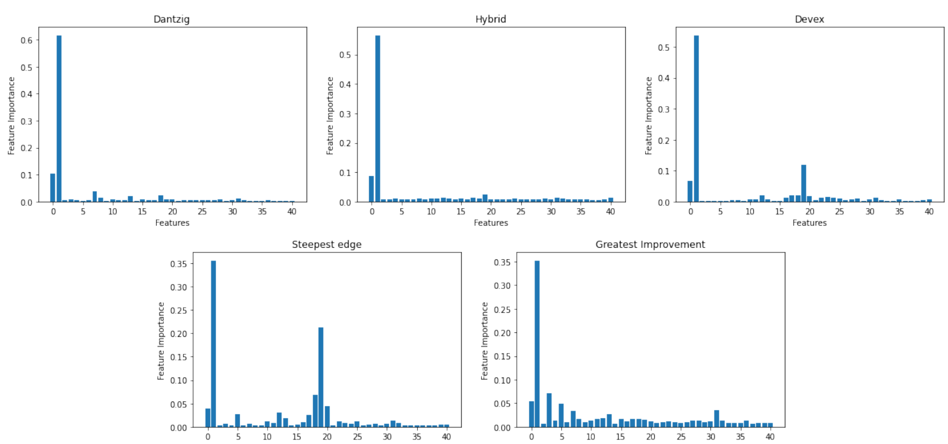

FIGURE 4.1. The gain of features for GBDT regressors.

Figure 4.1 shows the gain of features for GBDT regressors. We can see that apart from number of constraints and number of variables, some of the common features that are important are: maximum number of constraint degree, max and mean of variable degree, min and mean of coefficient matrix **A**, min, mean, norm and standard deviation of objective function **c** etc. One could take the subset of important features to train smaller models, which makes the training much faster, but the accuracy will drop to 66.44% with 178.7804 iterations on average.

GBDT classifier The other GBDT uses the truncated SVD with $k = 20$ as part of the features while keeping the features of constraint vector and objective function. This random forest contains 102 trees with minimum child weight of 5, maximum depth of 5, learning rate of 0.1, subsample and column subsample by tree ratio of 0.8. This model results in a 67.15% accuracy on the test set with 179.0714 iterations on average. The feature importance is shown in Figure 4.2.

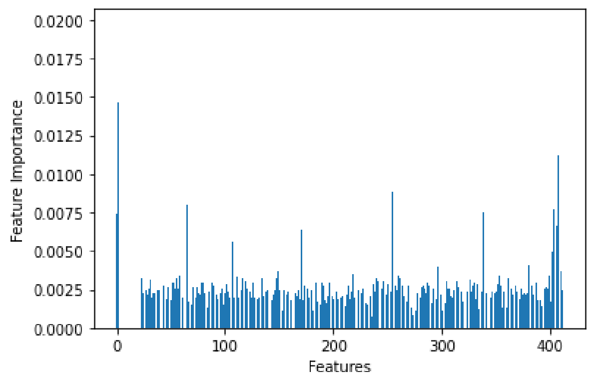

FIGURE 4.2. The gain of features for GBDT classifier.

As we can see, number of variables and constraints (the first and second feature), as well as features related to constraint vector and objective function are of great importance. Meanwhile, the diagonal entries of the SVD matrix have a relatively high importance.

4.3.2. Neural Networks. We train two models to classify which pivoting strategies will perform the best on each LP instance. The first model uses the bag of features that we choose for LP problems, and the second model uses the truncated SVD matrix to replace the features of the coefficient matrix **A**.

Bag-of-features Neural Network We first train a neural network using features of LP instances we pre-selected. The architecture of the network consists of four hidden layers of ReLU activation function with 64 neurons. Each hidden layer has a dropout of 0.1. The output layer contains five neurons with the softmax activation function. We train the neural network to minimize the categorical cross-entropy loss with the RMSProp optimizer with a learning rate of 0.01 and momentum of 0.2. We train with a batch size of 64 for 50 epochs. This model results in a

62.2% accuracy on the test set with 179.279 iterations on average. Figure 4.3 plots the accuracy and loss during each epoch.

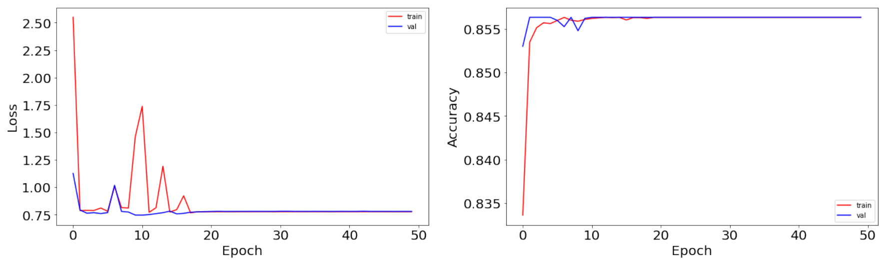

FIGURE 4.3. The accuracy and loss of the bag of features classification against number of epochs

Truncated SVD Neural Network We then train a neural network using truncated SVD matrices as features for coefficient matrix $\mathbf{A}$ while keeping the features of constraint vector and objective function. The architecture consists of four layers of 512 hidden units with ReLU activation function. The output layer contains 5 neurons with the softmax activation function. We train the neural network to minimize the categorical cross-entropy loss with the ADAM optimizer with a learning rate of 0.001. We train with a batch size of 64 for 100 epochs. This model results in a 72.78% accuracy on the test set with 179.18 iterations on average. Figure 4.4 plots the accuracy and loss during each epoch. Here we summarize the performance of our models. Table 4.2 shows

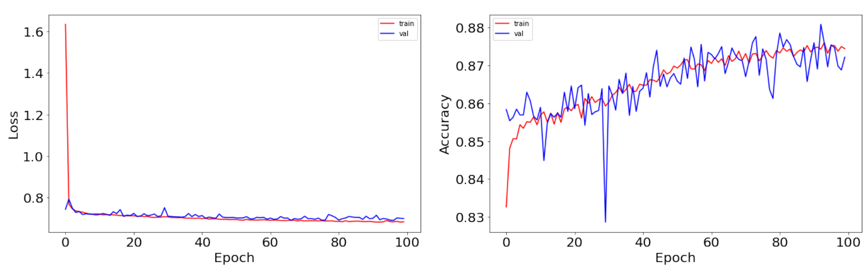

FIGURE 4.4. The accuracy and loss of the SVD20 classification against number of epochs

the average number of iterations (from the most to the fewest) if we use certain pivot rule or follow our models to solve the LP instances in the test set. It also demonstrates the prediction accuracy of our models. Table 4.3 shows the instance-wise comparison between our model recommendations with the most popular steepest edge pivot rule. We can see that the best performance of our four models is 69.06% of the number of iterations steepest edge will take. And they vary on the worst case behavior, with our GBDT regressor being the most consistent: their worst case will only cost 174.19% of what steepest edge will perform.

Classifier	Average iterations on test set	Accuracy
Greatest Improvement	326.1419	-
Dantzig	319.7501	-
Devex	262.2335	-
Hybrid	217.2856	-
Steepest edge	179.4161	-
Bag-of-features NN	179.279	62.2%
SVD-20 NN	179.18	72.78%
XGBClassifier	179.0714	70.15%
XGBRegressor	178.5934	67.78%
Best in theory	173.1783	100%

Table 4.2. Summary of average number of iterations and accuracy of each model.

Classifier	Best	Worst
Bag-of-features NN	69.06%	258.91%
SVD-20 NN	69.06%	210.25%
XGBClassifier	69.06%	318.06%
XGBRegressor	69.06%	174.19%

Table 4.3. Comparison between our models and steepest edge on test set per instance.

5

Appendix: Pseudocode for PolyPathLab and Oriented Matroid

5.1. PolyPathLab

We developed *PolyPathLab*, a MATLAB-based package, that takes in polytopes in cdd/cdd+ format and computes the following features:

- the *diameter*
- the *monotone diameter*
- the *monotone height*
- the *number of monotone paths*
- the *number of directed arborescences*
- characteristics related to four famous pivot rules: Dantzig's rule, greatest descent, steepest edge, and Bland's rule
- the *flip graph* and the *diameter of the flip graph*

To run *PolyPathLab*, users need to know how to use cdd/cdd+. The process of using *PolyPathLab* is:

(1) Create an inequality file or a vertices file for cdd/cdd+.
(2) Run cdd/cdd+ to obtain the files needed for the input of *PolyPathLab*, including *.ine, *.ead, *.ext files. We will call these the polytope files.
(3) Create an objective function file *.txt for the input of *PolyPathLab*.
(4) Put the polytope files with the correct file-naming convention and the objective function file into the input folder under the package directory.
(5) Run "main_general.m" or "flip_graph.m" based on what users want to compute and follow the instructions to set up the settings. Collect the outputs.

We will use the dodecahedron and a fixed objective function f for examples we use in this chapter. The dodecahedron has 20 vertices, 12 facets, and 30 edges. Each facet of the dodecahedron is a regular pentagon.

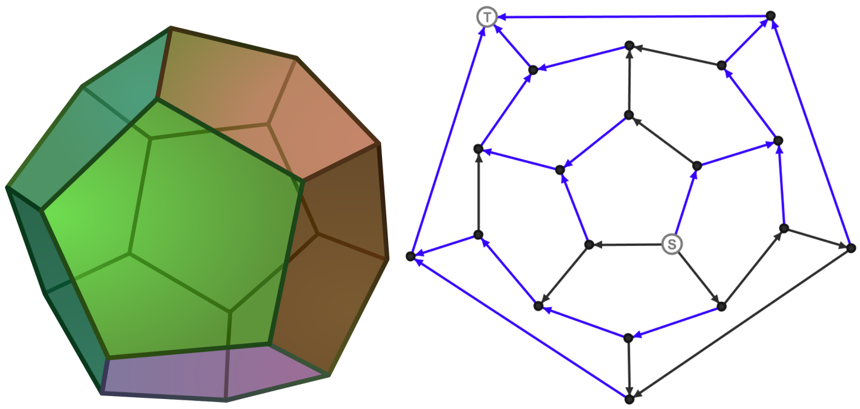

FIGURE 5.1. A picture of the dodecahedron [**Wik20**] and the directed 1-skeleton of the dodecahedron.

Besides computation of the above features, *PolyPathLab* has two functions, "generate_3d_CuttingPlane.m" and "generate_3d_PointOnSphere.m," to generate random $3d$ polytopes, which we use to test some properties of $3d$ polytopes. The functions are described below.

5.1.1. Random 3-dimensional Polytope Generator. We have two ways to generate random $3d$ polytopes: random cutting planes and random points on a sphere. The polytopes created by the cutting plane method are simple polytopes, while the points on sphere method may generate non-simple polytopes when the number of vertices is large.

For the cutting plane method, we first construct a random cuboid that is bigger than a $2 \times 2 \times 2$ cube in any direction. We then generate cutting planes one to two unit lengths away from the center of the cuboid. This ensures that we get a polytope with moderate size which increases the possibility of having effective cutting planes (cutting planes that produces one of the facets of the polytope). If we do not have a limit on the distance, there might be some "deep" cuts that are very close to the center of the cuboid and make many other cuts ineffective.

For the points on sphere method, we first generate a sphere of radius two, then select points at random. The these points form a polytope.

Here is the code to generate a random point on a sphere of radius two:

```
%n is the number of vertices of the polytope
%set the radius to 2
r = 2;

%create random theta and phi
theta = 2 * pi * (2 * rand(n, 1) - ones(n, 1));
phi = pi * ( 2 * rand(n, 1) - ones(n, 1));

%obtain the coordinates of the vertices
x = r .* cos(theta) .* sin(phi);
y = r .* sin(theta) .* sin(phi);
z = r .* cos(phi);

vertices = [x,y,z];
```

This code is a usage of spherical coordinate system. If we have $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$, we can obtain all the possible points on the sphere by fixing the radius r to equal the radius of the sphere. If we randomize θ and ϕ, we can obtain all the possible points on the sphere, thus randomizing the coordinates of the vertices.

5.1.2. Diameter. The diameter of a polytope shows us how efficient is the simplex method on the polytope. There are two ways to get the diameter of a polytope. First, MATLAB has a function called "distances" which measures the distances between any two pair of vertices [**MAT**]. For a polytope P with n vertices, it outputs an $n \times n$ matrix where each entry corresponds to the distance between the vertices of the row and column of the entry. The maximum number of the matrix produced by the distance function is the diameter according to definition. Second, we can multiply the matrix of $n \times n$ identity matrix I_n by adjacency matrix A plus an $n \times n$ identity matrix I_n, or $(A + I_n)$. Repeat multiplying the current matrix by $(A + I_n)$ until all of its entries become bigger than zero. Then, we record the number of times we have multiplied. The idea is that, every time the new matrix has an entry that turns from 0 to some positive number, the vertices represented by the row and column of the entry are connected by a path. For example, if $A^k_{i,j} = 1$, there is a path from i to j with length k. By adding I_n to A and raise $(A + I_n)$ to the power k, we allows the vertices to have self-loops so that if there is a zero in the ij entry of the current matrix, it must be that there is no path from i to j with length k. If the matrix has only non-zero entries, all the vertices are connected to each other with paths that have lengths less than the number of times we multiply.

In our program, we implemented the second method, matrix multiplication, to keep it consistent with the methods we used for the monotone diameter and the monotone height. Here is the pseudocode that computes the diameter.

```
%A is the n-by-n adjacency matrix. I is the n-by-n identity matrix.
%R is the matrix that records the entries

%initialize R as the identity matrix

R = I;

 diameter = 0;

%multiply until all the entries are greater than 1

 while ismember(0, R)  %if there is 0 in R

     R = R * (A + I); %multiply R by (A + I)

     diameter = diameter + 1; %record the increase in the diameter
```

end

For the dodecahedron, which has 20 vertices, A is the matrix shown in the next page:

$$A = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0
\end{bmatrix}.$$

If we multiply the identity matrix by $(A+I_{20})^4$, we obtain a matrix that only has a few zeros; if we multiply the current matrix by $(A+I_{20})$ again, we have a non-zero matrix. Thus, the diameter of the dodecahedron is five, meaning that any vertex is connected to any other vertices in a path with length less than five.

5.1.3. Monotone Diameter. Recall that the monotone diameter of DG is the maximum monotone distance between any two vertices in DG that are connected by monotone paths. The monotone diameter is not the same as the monotone distance between the source and the sink. Figure 5.2 is an example where the source s does not have a longer monotone distance to the sink

t than some other vertex i. The monotone distance from s to t is two, but the monotone distance from i to t is four.

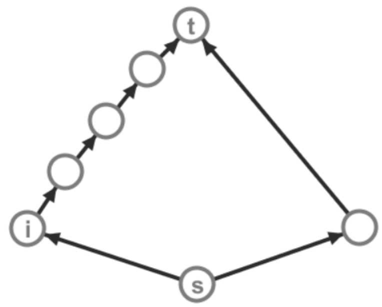

FIGURE 5.2. A directed graph where the monotone distance between the source s and the sink t is smaller than the monotone distance between i and t.

To get the monotone distance for vertex i, we multiply I_n by the matrix $(D + I_n)$, the directed adjacency matrix with diagonals being all 1. We repeat multiplying the current matrix by $(D+I_n)$ until the column of the sink t has non-zero values on i-th row for the first time. The number of multiplications k_i means that there is at least one monotone path from i to t with length k_i. We take the maximum k_i among all vertices to get the monotone diameter. In other words, we multiply the starting matrix I_n again and again by $(D+I_n)$ and record the number of times of multiplication k when the column of the sink has all non-zero values.

Here is the pseudocode for computing the monotone diameter:

```
%D is the directed adjacency matrix; I is the identity matrix;
%R is the current matrix; t is the index of the sink

%initialize R as the identity matrix

R = I;

monotone_diameter = 0;

%multiply until all the entries in column t are nonzero
```

```
while ismember(0, A2(:, t)) == 1 %if there is a zero in column t
    R = R * (D + I);
    mono_diameter = mono_diameter + 1;
end
```

For the same dodecahedron with an objective function f, the directed adjacency matrix is shown below:

$$D = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}$$

D is calculated from the adjacency matrix A and the objective function f by comparing the objective values of the two connected vertices. If i is connected with j, $A_{i,j} = A_{j,i} = 1$. Then, we compare the objective value of i and that of j. If the objective value of i is less than that of j, $D_{i,j} = 1$ and $D_{j,i} = 0$.

The sink of this example is vertex 10. If we raise $(D + I_{20})$ to the fourth power, there are zeros in the tenth column; if we raise $(D + I_{20})$ to the fifth power, there is no zero in the tenth column, which means that there is at least one path from any vertex to the sink with the length less than five. Thus, the monotone diameter of this dodecahedron with the objective function f is five.

5.1.4. Monotone Height. Recall that the monotone height of DG is the length of the longest monotone path on DG. Equivalently, monotone height is the maximum length of the monotone path from the source s to the sink t. We can prove this equivalence by contradiction. Let i be a vertex that is not the source or the sink. Assume the maximum length of the monotone path from i to t is k, which is longer than the maximum length of the monotone path from s to t. There is a monotone path of length bigger than one from s to i because of the property of DG. Then, we can build a new monotone path from s to i to t, which has the length at least $(k+1)$. This is a contradiction to our assumption.

To get the monotone height, we raise D to power k, where k is the *first time* that the matrix becomes a zero matrix. For all i and j, $D^k_{i,j} = 0$, meaning that there is no monotone path with length k between any pair of vertices. Thus, the monotone height is equal to $(k-1)$ because the maximum length of the monotone paths will be $(k-1)$.

```
%D is the directed adjacency matrix; R is the recording matrix;
%I is the identity matrix

R = I;

mono_height = -1;
%starts from -1 since when the matrix becomes all zero, we have no monotone path

%Multiply R by D until all the entries are zero

while ~ismember(0, R) %if the current matrix is all zero, stop

    mono_height = mono_height + 1;

    R = R * D;
```

end

Using the directed adjacency matrix D from the same dodecahedron and the same objective function f, we see that, for D^7, there are still ones in the matrix; for D^8, the matrix becomes a zero matrix for the first time. Thus, there is at least one path with length equals seven. Therefore, The monotone height of this dodecahedron with the same objective function f is seven.

5.1.5. Number of Monotone Paths. The number of monotone paths (from the source to the sink) tells us how many ways can the source s go to the sink t. The number of monotone paths is recorded when *PolyPathLab* computes the monotone height. Every time when the recording matrix is multiplied with the directed adjacency matrix, the entry st of the resulting matrix tells us how many monotone paths with the current length from s to t. When the matrix becomes all zero, we sum the values to get the number of total monotone paths.

For the dodecahedron and the same objective function, the recorded values in entry st are: $0, 0, 0, 0, 6, 6, 2$. Summing these values, we get that the number of monotone paths is 14. There are only seven values because the monotone height is seven, so anything after the last value will be zero.

5.1.6. Number of Arborescences. Because the pivot rule picks one improving edge from every vertex of DG that is not the sink, it creates a directed tree, which is a subgraph of DG. Furthermore, there is a monotone path from every non-sink vertex to the sink, so the subgraph is an arborescence. We say two pivot rules are equivalent on DG if they produce the same arborescence. Then, the number of arborescence is the number of the equivalence groups of pivot rules, and we can estimate the number of pivot rules by computing the number of arborescences [**ADLZ21**]. Figure 5.3 shows one arborescence of the dodecahedron with objective function f.

To compute the number of arborescences, we multiply the number of outgoing edges of each vertex except the sink (which has 0 outgoing edges) because changing the outgoing edge from a vertex will give us a different arborescence.

For the dodecahedron and the objective function f, if we count the outgoing edges of each vertex and put the numbers into a vector, we get:

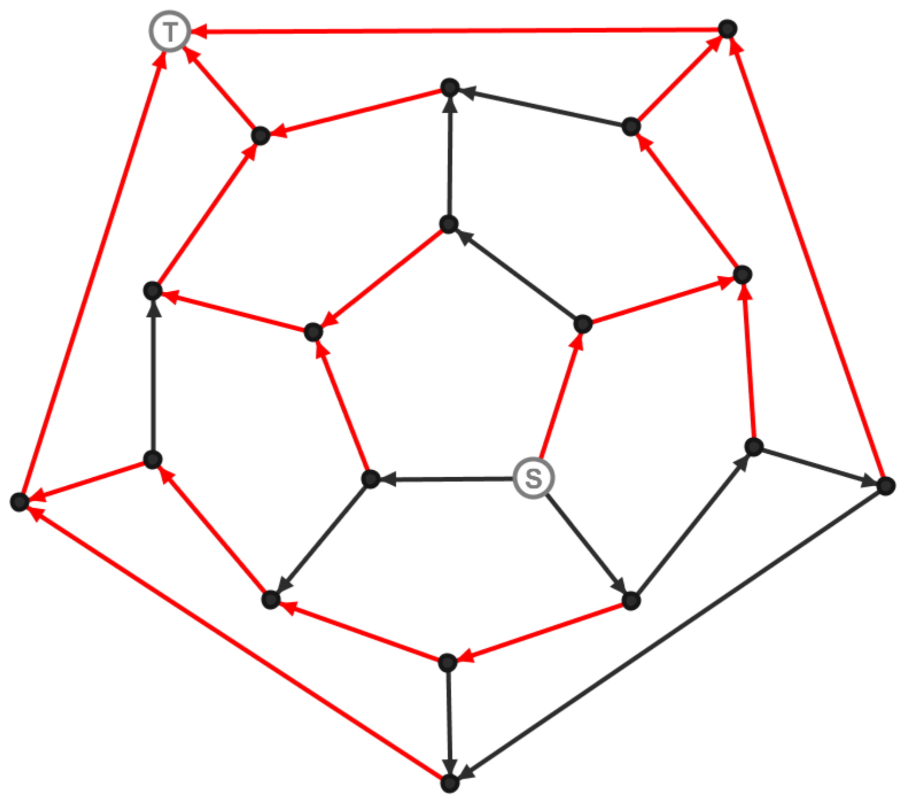

FIGURE 5.3. An arborescence of the dodecahedron. From each vertex, the pivot rule picks only one outgoing edge.

$$\begin{bmatrix} 2 & 2 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 0 & 1 & 1 & 2 & 2 & 1 & 3 & 2 & 1 & 2 & 2 \end{bmatrix}$$

If we multiply the nonzero entries, we get 1536, which is the number of arborescences of the dodecahedron with the objective function f.

5.1.7. Characteristics Related to Pivot Rules. In our program, we include four pivot rules: Dantzig's rule, Greatest Descent, Steepest Edge, and Bland's rule. Based on these four pivot rules, we collect characteristics related to the arborescences outputted by these pivot rules to see the performance of each pivot rule on the polytope P with the objective function f. The characteristics include: the average length of the monotone paths, the standard deviation of the length of the monotone paths of the arborescence, the standard deviation of the indegrees of the arborescence, the length of the monotone path between the source and the sink, the maximum length of the monotone paths of the arborescence, and the number of leaves in the arborescence.

We did not include the average indegree of the arborescence for a polytope with n vertices because it is always $\frac{n-1}{n}$.

5.1.8. Flip Graph. One interesting combinatorial feature of a polytope with an objective function is the flip graph. The flip graph tells us how are the monotone paths related to each other. Since the flip graph is 2-connected, we are also interested in the diameter of the flip graph because the diameter tells us how many flips do the furthest pair of monotone paths differ.

To compute the flip graph, first, *PolyPathLab* uses a code called *getpaths* (see [**Ans**]) to obtain all the monotone paths. Then, to find out if two monotone paths, p_i and p_j, are connected by a flip, we find out the first vertex u and the last vertex v that the two monotone paths differ from each other. Collecting all the vertices in between, we obtain a set of vertices between $u-1$ and $v+1$ for p_i and p_j.

If p_i and p_j differ by a polygon flip, then this set of vertices belong to the same 2-dimensional face. In order to compare to the faces, We put the coordinates of the vertices into a matrix V with each row being a vertex in the set. Then, we compute AV^T, with A being the inequalities in the form $Ax \leq b$. If there is a zero in the entry ij of the resulting matrix, vertex j is on the facet represented by the i-th row of A; if the i-th row is a zero vector, all the vertices between $u-1$ and $v+1$ are on the facet represented by the i-th row of A. For dimension d, the intersection of $d-2$ hyperplanes will provide us a 2-dimensional face. Therefore, if there are $d-2$ rows that are zero vectors, we know p_i and p_j differ by a polygon flip.

5.2. Oriented Matroid

We wrote code in Python to construct cocircuits graphs and run experiments on different oriented matroids to look for new conjectures. The code has the following features:

- compute the *cocircuits* given the chirotope of the oriented matroid
- compute the *cocircuits* given the hyperplane arrangement for a realizable oriented matroid
- compute the *circuits* given the chirotope of the oriented matroid
- construct the *circuit/cocircuit graphs* given the set of circuits/cocircuits
- compute the *diameter* of the graph and the pair of vertices whose distance is equal to the diameter

- check whether the shortest path between two cocircuits is a crabbed path.

The code is available at the Github directory at `https://github.com/zzy1995/OrientedMatroid`. To run the code, users only need to run "python3 OM.py" plus the text files containing oriented matroid in chirotope representations in the terminal. If the user wishes to explore on single instances, play around with the code or investigate certain conjectures, then running "python3 -i OM.py" would bring the user into python interaction command lines.

We illustrate how we implement each functions below:

5.2.1. Circuits and Cocircuits from chirotopes. Recall that the oriented matroids on Finschi and Fukuda's homepage [**FF01b**] are stored in the representation of chirotopes. In order to compute the cocircuit graph, we need to obtain the set of cocicruits. Recall that we can generate the cocircuits by computing the set $\mathcal{C}^*(\chi) = \{(\chi(\lambda, 1), \chi(\lambda, 2), \ldots, \chi(\lambda, n)) : \lambda \in E^{r-1}\}$.

Here we attach the pseudocode for computing the set of cocircuits and the set of circuits from the chirotope of the oriented matroid. See the functions *circuits* and *cocircuits* in the source code for more details on the implementation.

Algorithm 1 Construct cocircuits set given the chirotope map

Input Cardinality n, rank r of $\mathcal{M}$ and χ the chirotope map
Output A list containing all cocircuits $\mathcal{C}^*(\mathcal{M})$

```
for A ⊆ [n] and |A| = r − 1 do
  Initialize v = 0 ∈ ℝ^n
  Sort and vectorize A to λ
  for i = 1 to n do
    if i ∉ A then
      v[i] ← χ(i, λ)
    end if
  end for
  Add ± v to the set of cocircuits
end for
```

5.2.2. Cocircuits from hyperplane arrangement. In the findings of counter-examples, we often want to realize an oriented matroid directly from a hyperplane arrangement. In our representation, the central hyperplane arrangement is $Hx = 0$ where $H \in \mathbb{R}^{n\times r}$. The intersection of every $r - 1$ hyperplanes shall give a pair of cocircuits. Here we attach the pseudocode for

Algorithm 2 Construct circuits set given the chirotope map

Input Cardinality n, rank r of $\mathcal{M}$ and χ the chirotope map
Output a list containing all circuits $\mathcal{C}(\mathcal{M})$

for $A \subseteq [n]$ and $|A| = r - 1$ **do**
  Initialize $\mathbf{v} = 0 \in \mathbb{R}^n$
  Sort and vectorize A to λ
  for $i = 1$ to n **do**
    if $i \in A$ **then**
      if $i = \min A$ **then**
        $\mathbf{v}[i] \leftarrow 1$
      else
        $\mathbf{v}[i] \leftarrow -\chi(\min(A), \lambda) \times \chi(i, \lambda)$
      end if
    end if
  end for
  Add $\pm\,\mathbf{v}$ to the set of circuits
end for

computing the set of cocircuits with this strategy. See the function *cocircuits_from_arrangement* for more details on the implementation.

Algorithm 3 Construct cocircuits set given the hyperplane arrangement

Input Cardinality n, rank r of $\mathcal{M}$ and the matrix H for the central hyperplane arrangement $Hx = 0$
Output A list containing all cocircuits $\mathcal{C}^*(\mathcal{M})$

for $A \subseteq [n]$ and $|A| = r - 1$ **do**
  Initialize H_A containing rows of H as in A
  Compute $\mathbf{v}_A$ in the null space of H_A
  Add $\pm\,\mathrm{sign}(H_A \cdot \mathbf{v}_A$ to the set of cocircuits
end for

5.2.3. Constructing graphs. Recall that given two cocircuit signed vectors X, Y, X and Y are connected in cocircuit graph if $|X^0 \cap Y^0| \geq r - 2$ and $X \neq \pm Y$. Thus, once we have the set of cocircuits, we are able to construct the cocircuit graph. We utilize python package NetworkX for constructing the graph. We first add cocircuits as vertices of the graph, then add edges based on this condition. See functions *make_graph* and *construct_graph* for detailed implementation.

5.2.4. Diameter. Given the cocircuit graph in NetworkX, we are able to compute the diameter of the graph, as well as finding which pairs of cocircuits have the distance of the diameter. See functions *find_diameter* and *find_pairs* for detailed implementation. There are also other functions built in NetworkX to compute things such as shortest path between two cocircuits.

5.2.5. Crabbed paths. Recall that a path P from X to Y *crabbed* if for every cocircuit $W \in P$, $W^+ \subseteq X^+ \cup Y^+$ and $W^- \subseteq X^- \cup Y^-$. The *diameter* of $G^*(\mathcal{M})$ is defined as $\mathrm{diam}(G^*(\mathcal{M})) = \max\{d_{\mathcal{M}}(X, Y) : X, Y \in \mathcal{C}^*(\mathcal{M})\}$. For any pairs of cocircuits X and Y, we use the function all_shortest_paths from NetworkX to find all shortest path between them. We then check if at least one of the paths is crabbed. See function *is_crabbed*, *check_crabbed_graph* and *check_crabbed_file* for detailed implementation.

Bibliography

[AC09] D. Avis and V. Chvatal, *Notes on Bland's pivoting rule*, vol. 8, pp. 24–34, Springer, 03 2009.

[AD74] I. Adler and G. Dantzig, *Maximum diameter of abstract polytopes*, Mathematical Programming Study **1** (1974), 20–40.

[ADG+16] M. Andrychowicz, M. Denil, S. Gomez, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, and N. De Freitas, *Learning to learn by gradient descent by gradient descent*, Advances in Neural Information Processing Systems, 2016, pp. 3981–3989.

[ADLRS00] C. A. Athanasiadis, J. A. De Loera, V. Reiner, and F. Santos, *Fiber polytopes for the projections between cyclic polytopes*, European J. Combin. **21** (2000), 19–47.

[ADLZ21] C. A. Athanasiadis, J. A. De Loera, and Z. Zhang, *Enumerative problems for arborescences and monotone paths on polytope graphs*, Journal of Graph Theory **99** (2021).

[AER00] C. A. Athanasiadis, P. Edelman, and V. Reiner, *Monotone paths on polytopes*, Math. Z. **235** (2000), 549–555.

[ALW17] A. M. Alvarez, Q. Louveaux, and L. Wehenkel, *A machine learning-based approximation of strong branching*, INFORMS J. Comput. **29** (2017), no. 1, 185–195.

[Ans] M. Answers, *Calculating all paths from a given node in a digraph*, `https://www.mathworks.com/matlabcentral/answers/417396-calculating-all-paths-from-a-given-node-in-a-digraph`.

[AS01] C. A. Athanasiadis and F. Santos, *Monotone paths on zonotopes and oriented matroids*, Canad. J. Math. **53** (2001), no. 6, 1121–1140.

[Ath99] C. A. Athanasiadis, *Piles of cubes, monotone path polytopes, and hyperplane arrangements*, Discrete Comput. Geom. **21** (1999), no. 1, 117–130.

[AZ99] N. Amenta and G. M. Ziegler, *Deformed products and maximal shadows of polytopes*, Advances in Discrete and Computational Geometry (1999).

[Bar73] D. Barnette, *A proof of the lower bound conjecture for convex polytopes*, Pacific J. Math. **46** (1973), 349–354.

[Bar02] A. I. Barvinok, *A course in convexity*, Graduate studies in mathematics, vol. 54, American Mathematical Society, 2002.

[BDLL21] M. Blanchard, J. A. De Loera, and Q. Louveaux, *On the length of monotone paths in polyhedra*, SIAM J. Discret. Math. **35** (2021), no. 3, 1746–1768.

[BDLLS22] A. E. Black, J. A. De Loera, N. Lütjeharms, and R. Sanyal, *The polyhedral geometry of pivot rules and monotone paths*, 2022.

[BDSV18] M. Balcan, T. Dick, T. Sandholm, and E. Vitercik, *Learning to branch*, Int. Conf. Mach. Learn., 2018, pp. 353–362.

[BFF01] E. Babson, L. Finschi, and K. Fukuda, *Cocircuit graphs and efficient orientation reconstruction in oriented matroids*, European J. Combin. **22** (2001), no. 5, 587–600, Combinatorial geometries (Luminy, 1999).

[Bix01] R. Bixby, *Solving real-world linear programs: A decade and more of progress*, Operations Research **50** (2001).

[BKK+16] B. Bischl, P. Kerschke, L. Kotthoff, M. T. Lindauer, Y. Malitsky, A. Fréchette, H. H. Hoos, F. Hutter, K. Leyton-Brown, K. Tierney, and J. Vanschoren, *ASlib: A benchmark library for algorithm selection*, Artif. Intell. **237** (2016), 41–58.

[BKS94] L. Billera, M. M. Kapranov, and B. Sturmfels, *Cellular strings on polytopes*, Proc. Am. Math. Soc. **122** (1994), no. 2, 549–555.

[Bla77] R. G. Bland, *New finite pivoting rules for the simplex method*, Mathematics of Operations Research **2** (1977), no. 2, 103–107.

[BLP18] Y. Bengio, A. Lodi, and A. Prouvost, *Machine learning for combinatorial optimization: a methodological tour d'horizon*, CoRR **abs/1811.06128** (2018), 1811.06128.

[BLVS+99] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. M. Ziegler, *Oriented matroids*, second ed., Encyclopedia of Mathematics and its Applications, vol. 46, Cambridge University Press, Cambridge, 1999.

[BLZ18] P. Bonami, A. Lodi, and G. Zarpellon, *Learning a classification of mixed-integer quadratic programming problems*, Integration of Constraint Programming, Artificial Intelligence, and Operations Research - 15th International Conference, CPAIOR 2018, Delft, The Netherlands, June 26-29, 2018, Proceedings (W. J. van Hoeve, ed.), Lecture Notes in Computer Science, vol. 10848, Springer, 2018, pp. 595–604.

[BNVW17] M. Balcan, V. Nagarajan, E. Vitercik, and C. White, *Learning-theoretic foundations of algorithm configuration for combinatorial partitioning problems*, Proc. Conf. Learn. Th., 2017, pp. 213–274.

[Bon]

[Bor82] K. H. Borgwardt, *The average number of pivot steps required by the simplex-method is polynomial*, Zeitschrift for Operations Research **26** (1982), 157–177.

[BS92] L. Billera and B. Sturmfels, *Fiber polytopes*, Ann. of Math. **135** (1992), no. 3, 527–549.

[BS20] D. Bertsimas and B. Stellato, *The voice of optimization*, Machine Learning (2020), 1–29.

[BSMBM20] S. Bowly, K. Smith-Miles, D. Baatar, and H. Mittelmann, *Generation techniques for linear programming instances with controllable properties*, Mathematical Programming Computation **12** (2020), 389–415.

[BT97] D. Bertsimas and J. Tsitsiklis, *Introduction to linear optimization*, Athena Scientific Series in Optimization and Neural Computation, 1997.

[BvdHS06] G. Brightwell, J. van den Heuvel, and L. Stougie, *A linear bound on the diameter of the transportation polytope**, Combinatorica **26** (2006), 133–139.

[CF93] R. Cordovil and K. Fukuda, *Oriented matroids and combinatorial manifolds*, European Journal of Combinatorics **14** (1993), no. 1, 9–15.

[CFGdO00] R. Cordovil, K. Fukuda, and A. Guedes de Oliveira, *On the cocircuit graph of an oriented matroid*, Discrete and Computational Geometry **24** (2000), 257–266.

[CM93] R. Cordovil and M. Moreira, *A homotopy theorem on oriented matroids*, Discrete Mathematics **111** (1993), no. 1, 131 – 136.

[CS17] F. Criado and F. Santos, *The maximum diameter of pure simplicial complexes and pseudo-manifolds*, Discrete & Computational Geometry **58** (2017), no. 3, 643–649.

[CS19] ———, *Topological prismatoids and small simplicial spheres of large diameter*, (2019).

[Dan90] G. B. Dantzig, *Origins of the simplex method*, p. 141–151, Association for Computing Machinery, New York, NY, USA, 1990.

[Dan98] ———, *Linear programming and extensions*, Landmarks in Physics and Mathematics, Princeton University Press, 1998.

[Dev04] M. Develin, *LP-orientations of cubes and crosspolytopes*, Adv. Geom. **4** (2004), no. 4, 459–468.

[Dev18] N. Developers, *Networkx*, September 2018.

[DH16] D. Dadush and N. Hähnle, *On the shadow simplex method for curved polyhedra*, Discrete Comput. Geom. **56** (2016), no. 4, 882–909.

[DH18] D. Dadush and S. Huiberts, *A friendly smoothed analysis of the simplex method*, Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing - STOC 2018 (2018).

[DL11] J. A. De Loera, *New insights into the complexity and geometry of linear optimization*, Optima 87 (2011).

[DLLY08] J. A. De Loera, F. Liu, and R. Yoshida, *A generating function for all semi-magic squares and the volume of the Birkhoff polytope*, Journal of Algebraic Combinatorics **30** (2008), no. 1, 113–139.

[Don97] J. Dongarra, *Netlib*, 1997, http://netlib.org.

[Edm15] R. Edman, *Diameter and coherence of monotone path graphs in low corank*, Ph.D. thesis, May 2015.

[EHRR10] F. Eisenbrand, N. Hähnle, A. Razborov, and T. Rothvoß, *Diameter of polyhedra: Limits of abstraction*, Math. Oper. Res. **35** (2010), no. 4, 786–794.

[EJLM18] R. Edman, P. Jiradilok, G. Liu, and T. McConville, *Zonotopes whose cellular strings are all coherent*, 2018, 1801.09140.

[ELH18] K. Eggensperger, M. Lindauer, and F. Hutter, *Neural networks for predicting algorithm runtime distributions*, Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden (J. Lang, ed.), ijcai.org, 2018, pp. 1442–1448.

[ELH19] ———, *Pitfalls and best practices in algorithm configuration*, J. Artif. Intell. Res. **64** (2019), 861–893.

[FF01a] L. Finschi and K. Fukuda, *Combinatorial generation of small point configurations and hyperplane arrangements*, CCCG '01 (2001), 97–100.

[FF01b] ———, *Homepage of oriented matroids*, November 2001.

[FGK+11] S. Felsner, R. Gómez, K. Knauer, J. J. Montellano-Ballesteros, and R. Strausz, *Cubic time recognition of cocircuit graphs of uniform oriented matroids*, European J. Combin. **32** (2011), no. 1, 60–66.

[Fin01] L. Finschi, *A graph theoretical approach for reconstruction and generation of oriented matroids*, Ph.D. thesis, ETH Zürich, 2001.

[FKE+19] M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum, and F. Hutter, *Auto-sklearn: Efficient and robust automated machine learning*, Automated Machine Learning - Methods, Systems, Challenges (F. Hutter, L. Kotthoff, and J. Vanschoren, eds.), The Springer Series on Challenges in Machine Learning, Springer, 2019, pp. 113–134.

[FL78] J. Folkman and J. Lawrence, *Oriented matroids*, J. Combin. Theory Ser. B **25** (1978), no. 2, 199–236.

[Fri11] O. Friedmann, *A subexponential lower bound for Zadeh's pivoting rule for solving linear programs and games*, Integer Programming and Combinatoral Optimization (Berlin, Heidelberg) (O. Günlük and G. J. Woeginger, eds.), Springer Berlin Heidelberg, 2011, pp. 192–206.

[FT97] K. Fukuda and T. Terlaky, *Criss-cross methods: A fresh view on pivot algorithms*, Mathematical Programming: Series A and B **Volume 79 Issue 1-3** (1997), 369–395.

[FT99] ———, *On the existence of a short admissible pivot sequence for feasibility and linear optimization problems*, Pure Mathematics and Applications **vol 10(4)** (1999), 431–447.

[GK07] B. Gärtner and V. Kaibel, *Two new bounds for the random-edge simplex-algorithm*, SIAM Journal on Discrete Mathematics **21** (2007), no. 1, 178–190.

[GR17] R. Gupta and T. Roughgarden, *A PAC approach to application-specific algorithm selection*, SIAM J. Comput. **46** (2017), no. 3, 992–1017.

[GS79] D. Goldfarb and W. Y. Sit, *Worst case behavior of the steepest edge simplex method*, Discrete Applied Mathematics **1** (1979), no. 4, 277 – 285.

[Har73] P. M. J. Harris, *Pivot selection methods of the devex LP code*, Mathematical Programming **5** (1973), no. 1, 1–28.

[HTF09] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*, Springer New York, 2009.

[Jer73] R. Jeroslow, *The simplex algorithm with the pivot rule of maximizing criterion improvement*, Discrete Mathematics **4** (1973), no. 4, 367 – 377.

[Kal87] G. Kalai, *Rigidity and the lower bound theorem. I*, Invent. Math. **88** (1987), no. 1, 125–151.

[Kal97] ———, *Linear programming, the simplex algorithm and simple polytopes*, Mathematical Programming **79** (1997), no. 1, 217–233.

[KDN+17] E. B. Khalil, B. Dilkina, G. L. Nemhauser, S. Ahmed, and Y. Shao, *Learning to run heuristics in tree search*, Proc. Int. Joint Conf. Artif., 2017, pp. 659–666.

[KDZ+17] E. B. Khalil, H. Dai, Y. Zhang, B. Dilkina, and L. Song, *Learning combinatorial optimization algorithms over graphs*, Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA (I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, and R. Garnett, eds.), 2017, pp. 6348–6358.

[KHO17] L. Kotthoff, B. Hurley, and B. O'Sullivan, *The ICON challenge on algorithm selection*, AI Magazine **38** (2017), no. 2, 91–93.

[KK87] V. Klee and P. Kleinschmidt, *The d-step conjecture and its relatives*, Math. Oper. Res. **12** (1987), no. 4, 718–755.

[KK92] G. Kalai and D. J. Kleitman, *A quasi-polynomial bound for the diameter of graphs of polyhedra*, Bull. Amer. Math. Soc. (N.S.) **26** (1992), no. 2, 315–316.

[KLP17] M. Kruber, M. E. Lübbecke, and A. Parmentier, *Learning when to use a decomposition*, Integration of AI and OR Techniques in Constraint Programming - 14th International Conference, CPAIOR 2017, Padua, Italy, June 5-8, 2017, Proceedings (D. Salvagnin and M. Lombardi, eds.), Lecture Notes in Computer Science, vol. 10335, Springer, 2017, pp. 202–210.

[KM72] V. Klee and G. J. Minty, *How good is the simplex algorithm*, Inequalities III (1972).

[KMBS14] K. Knauer, J. Montellano-Ballesteros, and R. Strausz, *A graph-theoretical axiomatization of oriented matroids*, European Journal of Combinatorics **35** (2014), 388–391, Selected Papers of EuroComb'11.

[KS10] E. D. Kim and F. Santos, *An update on the Hirsch conjecture*, Jahresbericht der Deutschen Mathematiker-Vereinigung **112** (2010), no. 2, 73–98.

[KW67] V. Klee and D. W. Walkup, *The d-step conjecture for polyhedra of dimension $d < 6$*, Acta Math. **117** (1967), 53–78.

[LHHX14] K. Leyton-Brown, H. H. Hoos, F. Hutter, and L. Xu, *Understanding the empirical hardness of* NP-*complete problems*, Commun. ACM **57** (2014), no. 5, 98–107.

[LJD+18] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, *Hyperband: A novel bandit-based approach to hyperparameter optimization*, Journal of Machine Learning Research **18** (2018), no. 185, 1–52.

[LL00] M. G. Lagoudakis and M. L. Littman, *Algorithm selection using reinforcement learning*, Int. Conf. Mach. Learn., 2000, pp. 511–518.

[MAT] MATLAB, *Shortest path distances of all node pairs - MATLAB*, `https://www.mathworks.com/help/matlab/ref/graph.distances.html`.

[MBS06] J. Montellano-Ballesteros and R. Strausz, *A characterization of cocircuit graphs of uniform oriented matroids*, J. Combin. Theory Ser. B **96** (2006), no. 4, 445–454.

[McD95] J. McDonald, *Fiber polytopes and fractional power series*, J. Pure Appl. Algebra **104** (1995), no. 2, 213–233.

[McM04] P. McMullen, *Triangulations of simplicial polytopes*, Beiträge Algebra Geom. **45** (2004), no. 1, 37–46.

[mip18] *MIPLIB 2017*, 2018, http://miplib.zib.de.

[MK00] J. Mihalisin and V. Klee, *Convex and linear orientations of polytopal graphs*, Discrete Comput. Geom. **24** (2000), no. 2-3, 421–435.

[MN13] S. Murai and E. Nevo, *On the generalized lower bound conjecture for polytopes and spheres*, Acta Math. **210** (2013), no. 1, 185–202.

[MRS08] C. Manning, P. Raghavan, and H. Schütze, *Introduction to information retrieval*, Cambridge University Press, 2008.

[MSW15] B. Matschke, F. Santos, and C. Weibel, *The width of five-dimensional prismatoids*, Proc. Lond. Math. Soc. (3) **110** (2015), no. 3, 647–672.

[Mur80] K. G. Murty, *Computational complexity of parametric linear programming*, Mathematical Programming **19** (1980), no. 1, 213–219.

[Nad89] D. Naddef, *The Hirsch conjecture is true for (0, 1)-polytopes*, Math. Program. **45** (1989), no. 1–3, 109–110.

[NDSLB04] E. Nudelman, A. Devkar, Y. Shoham, and K. Leyton-Brown, *Understanding random SAT: Beyond the clauses-to-variables ratio*, Lect. Notes Comput. SC, 2004, pp. 438–452.

[oCt] T. I. D. O. on Cloud team, *Docplex*.

[Orl97] J. B. Orlin, *A polynomial time primal network simplex algorithm for minimum cost flows*, Mathematical Programming **78** (1997), no. 2, 109–129.

[Pou14] L. Pournin, *The diameter of associahedra*, Adv. Math. **259** (2014), 13–42.

[Pou17] ———, *The asymptotic diameter of cyclohedra*, Israel J. Math. **219** (2017), no. 2, 609–635.

[RG99] J. Richter-Gebert, *Testing orientability for matroids is NP-complete*, Advances in Applied Mathematics **23** (1999), 78–90.

[RR13] V. Reiner and Y. Roichman, *Diameter of graphs or reduced words and galleries*, Trans. Amer. Math. Soc **365** (2013), no. 5, 2779–2802.

[San12] F. Santos, *A counterexample to the Hirsch conjecture*, Ann. of Math. (2) **176** (2012), no. 1, 383–412.

[San13] ———, *Recent progress on the combinatorial diameter of polytopes and simplicial complexes*, TOP **21** (2013), no. 3, 426–460.

[Sch86] A. Schrijver, *Theory of linear and integer programming*, John Wiley and Sons, Inc. New York, NY, USA, 1986.

[Sch98] ———, *Theory of linear and integer programming*, Wiley Series in Discrete Mathematics & Optimization, Wiley, 1998.

[Slo] N. Sloane, *The on-line encyclopedia of integer sequences.*

[Smi99] K. A. Smith, *Neural networks for combinatorial optimization: A review of more than a decade of research*, INFORMS Journal on Computing **11** (1999), no. 1, 15–34, https://doi.org/10.1287/ijoc.11.1.15.

[Suk16] N. Sukegawa, *Improving bounds on the diameter of a polyhedron in high dimensions*, Discrete Mathematics **340** (2016).

[Suk19] N. Sukegawa, *An asymptotically improved upper bound on the diameter of polyhedra*, Discret. Comput. Geom. **62** (2019), no. 3, 690–699.

[Tod11] M. J. Todd, *The basic George B. Dantzig, by Richard W. Cottle*, Bulletin of the American Mathematical Society **48** (2011), no. 1, 123–129.

[Tod14] M. J. Todd, *An improved Kalai-Kleitman bound for the diameter of a polyhedron*, SIAM J. Discrete Math. **28** (2014), no. 4, 1944–1947.

[TZ93] T. Terlaky and S. Zhang, *Pivot rules for linear programming: A survey on recent theoretical developments*, Annals of Operations Research **46-47** (1993), 203–233.

[Ver09] R. Vershynin, *Beyond Hirsch conjecture: Walks on random polytopes and smoothed complexity of the simplex method*, SIAM Journal on Computing **39** (2009), no. 2, 646–678, https://doi.org/10.1137/070683386.

[Wik20] Wikipedia, *Archimedean solid*, `https://en.wikipedia.org/wiki/Archimedean_solid`, Apr 2020.

[xd21] xgboost developers, *Xgboost documentation*, 2021.

[YAKU19] C. Yang, Y. Akimoto, D. W. Kim, and M. Udell, *Oboe: Collaborative filtering for automl model selection*, Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2019).

[Zie95] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995.