ON THE LENGTH OF MONOTONE PATHS IN POLYHEDRA

M. BLANCHARD*, J.A. DE LOERA†, AND Q. LOUVEAUX‡

Abstract. Motivated by the problem of bounding the number of iterations of the Simplex algorithm we investigate the possible lengths of monotone paths followed inside the oriented graphs of polyhedra (oriented by the objective function). We consider both the *shortest* and the *longest* monotone paths and estimate the *monotone diameter* and *height* of polyhedra. Our analysis applies to transportation polytopes, matroid polytopes, matching polytopes, shortest-path polytopes, and the TSP, among others.

We begin by showing that combinatorial cubes have monotone diameter and Bland simplex height upper bounded by their dimension and that in fact all monotone paths of zonotopes are no larger than the number of edge directions of the zonotope. We later use this to show that several polytopes have polynomial-size monotone diameter. In contrast, we show that for many well-known combinatorial polytopes, the height is at least exponential. Surprisingly, for some famous pivot rules, e.g., greatest improvement and steepest edge, these same polytopes have polynomial-size simplex paths.

Key words. Simplex method, diameter and height of polytopes, pivot rules, monotone paths, graphs of polyhedra, polyhedral combinatorics

AMS subject classifications. 52B12, 52B55, 90C05, 05C20 , 90C08

1. Introduction.

It is a famous open challenge to find a pivot rule that can make the Simplex method run in polynomial time for all linear programs or show that none exist (see e.g., [1, 6, 38] and the many references therein for a discussion of this famous algorithmic problem). In particular, such a pivot rule will take polynomially many monotonically-improving edge steps from any initial vertex. This paper discusses the possible lengths of the monotone paths followed by the Simplex method on several famous combinatorial polyhedra where computing monotone paths has nice combinatorial meaning.

We now introduce some basic terminology. In what follows we consider a polytope/polyhedron $P(A, b)$ in one of their canonical forms $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ or $\{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$. Here $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Objective function vectors will be typically denoted by $c \in \mathbb{R}^n$. $LP(A, b, c)$ will denote the (minimization) LP instance given by A, b, c.

Note that, each polyhedron $P(A, b)$ has a graph which is the 1-dimensional skeleton of faces of P. Given any A, b, c such that c is a nondegenerate linear objective function i.e., no two vertices have the same objective function, one obtains a natural directed acyclic graph on the vertices and edges of the polytope $P(A, b)$ by orienting each edge of the polytope $P(A, b)$ as per the objective value of the two endpoints. This will be denoted by $G(A, b, c)$. This kind of orientations of the graphs of $P(A, b)$ are called *LP-admissible*. We note that the resulting directed graph $G(A, b, c)$ is always acyclic, with a unique sink and source in each face (for more on LP-admissible orientations see [10, 24]).

We define a directed path Γ from node W to node Z inside the directed graph $G(A, b, c)$ as a subgraph of $G(A, b, c)$ having distinct nodes $v_0 = W, v_1, v_2, \ldots, v_n = Z$ and as arcs the pairs v_i, v_{i+1} for $i = 1$ to n. The length of Γ is n.

We introduce now the main combinatorial definitions and then give several remarks about these concepts. See Figure 1 for an example.

*Massachusetts Institute of Technology, Cambridge, MA, USA.
†University of California, Davis, CA, USA.
‡Université de Liège, Liège, Belgique

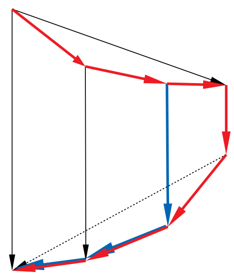

FIG. 1. *Two monotone paths on the directed graph $G(A, b, c)$ of the Klee−Minty cube. The longest monotone path in red gives the* height *and the blue monotone path gives the* monotone diameter *of this polytope.*

DEFINITION 1.1. *Let c be a generic linear objective function and π a pivot rule of the Simplex method.*

1. *A c-*monotone path *is a directed path in the LP-admissible oriented graph $G(A, b, c)$, that starts from some vertex to the optimal vertex. In particular, the path must satisfy if $c^T v_i > c^T v_{i+1}$ between consecutive nodes of the path. (note that we always consider the optimal vertex to be the terminal node of the path, but the paths do not necessarily start at a specific node).*
2. *From each vertex there is at least one shortest c-monotone path to the optimum. The c-*monotone diameter *is the maximum length of a shortest c-monotone path, the maximum being taken over all starting vertices.*
3. *The c-*height *is the length of the longest c-monotone path.*
4. *A c-π-*simplex path *is a c-monotone path in $G(A, b, c)$ following the pivot rule π. In this paper we will consider four popular pivot rules: Bland's pivot rule, Dantzig's pivot rule, greatest improvement pivot rule, and steepest edge pivot rule.*

We use these definitions to build our main concepts of interest.

DEFINITION 1.2.

1. *The* monotone diameter *of a polytope is the maximum c-monotone diameter, the maximum being taken over all objective functions c.*
2. *The* height *is the maximum c-height, the maximum being taken over all objective functions c.*
3. *The π-*simplex height *is the maximum length of a c-π-simplex path for the pivot rule π, the maximum being taken over all objective functions c.*

The study of the *undirected* diameter of the graph of polytopes is of course classical and related to the Hirsch conjecture (see e.g., [34], [9] and references), but the investigations of *directed monotone paths* are even more directly relevant to the Simplex method, and they have occupied researchers for some time too: In the 1960's Klee initiated the study of short/long monotone paths in his papers [21, 20, 22] where he proved bounds on the monotone diameter and height of simple polytopes. Later in the 1980's, in a remarkable *tour de force*, Todd [37] showed that the *monotone* Hirsch conjecture, saying that the monotone diameter is always less or equal to the number of facets minus the dimension, is false. In the 1990's Kalai [17] proved that for an n-dimensional polyhedron with m facets there is a subexponential upper bound on the *monotone* diameter of $m^{2\sqrt{m}}$ and Rispoli and collaborators wrote a series of papers about the monotone diameter of some specific combinatorial polytopes, such as the TSP [30, 31, 32]. Today, several papers continue the study of shortest monotone paths (see [8, 13, 29] and references therein).

The notion of height is useful to indicate the worst possible case of the Simplex method. In fact, long monotone paths have also been explored before: the *monotone upper bound problem* asks for the maximal number $M(n, m)$ of vertices on a strictly increasing edge-path on a simple n-dimensional polytope with m facets. This is the same as the largest height over all simple n-polytopes with m facets. It was conjectured that $M(n, m)$ is never more than the number of vertices of a dual-to-cyclic n-polytope with m facets, but Pfeifle and Ziegler proved it is strictly less than that in dimension six [29]. In our paper the reader can observe how the upper bound $M(n, m)$ is often too big for the specific polytopes we consider.

In the Simplex method a *pivot rule* is a method for selecting an improving neighboring extreme point. Each pivot rule will drive the algorithm to follow a different Simplex monotone path. Here we obtain a few results about the lengths of Simplex monotone paths. Today we know many pivot rules [36]. In this paper we will use four famous rules (here described in terms of tableau language, see Section 3.3 of [3]):

- **Dantzig's pivot rule:** The non-basic variable with the most negative reduced cost enters the basis.
- **Greatest improvement pivot rule:** The non-basic variable which provides largest improvement of the objective function enters the basis.
- **Steepest edge pivot rule:** The non-basic variable with the most negative reduced cost normalized by the length of the column enters the basis.
- **Bland's pivot rule:** Choose the entering basic variable x_j such that j is the smallest index with negative reduced cost. Also choose the leaving basic variable i with the smallest index (in case of ties in the ratio test).

At present, no pivot rule can guarantee a polynomial upper bound on the number of steps (see discussion and references in [6]). In fact, even for the four pivot rules above, there are exponentially-long c-π-simplex paths [14, 16, 22]. In contrast, we show that these four pivot rules behave nicely in some combinatorial polyhedra.

We wish to stress that the theory of computational complexity influences the geometry of monotone paths of polytopes. For instance, in [1] it was shown that there are Simplex pivoting rules for which it is PSPACE-complete to decide whether a particular basis will appear on the algorithm's path. This happens even for the Dantzig pivot rule [12]. Moreover, it was recently shown in [8] that it is NP-hard to compute the monotone diameter.

Finally, it is useful to note the key concepts we discuss satisfy the following relation:

$$\textit{(undirected) diameter} \leq \textit{monotone diameter} \leq \pi\textit{-simplex height} \leq \textit{height}.$$

The differences between these quantities can be rather dramatic. For example, for the Birkhoff polytope of $n \times n$ doubly-stochastic matrices, it is well-known that the undirected diameter is two, the monotone diameter is $\lfloor \frac{n}{2} \rfloor$, and the height is at least $O(n!)$. We now summarize our main results.

Our results. In Section 2 we show that combinatorial cubes have monotone diameter and Bland simplex height upper bounded by their dimension. Similarly, zonotopes have height never larger than the number of edge directions of the zonotope. In the following, for a polytope P we will denote by mono-diam(P) the monotone diameter of P.

THEOREM 1.3. *Let P be a convex polytope. Denote by $Z(P)$ the zonotope generated by the minimal set of vectors containing all directions of edges of P. Then, mono-diam$(P) \leq$ mono-diam$(Z(P))$ = number of different edge directions of P.*

This simple theorem has nice consequences. We can easily show that matroid polytopes, polymatroid polytopes, and some types of transportation polytopes have polynomial-size monotone diameter. Therefore, for polytopes such as the permutahedron or the spanning tree polytope there exist polynomial pivot rules for the Simplex method.

THEOREM 1.4.
1. *If P is a matroid polytope or a polymatroid polytope, then mono-diam$(P) \leq \binom{n}{2}$, where n is the number of elements of the matroid.*
2. *If P is a $k \times n$ transportation polytope, mono-diam$(P) \leq e \cdot k!\, n^k$. Therefore, the monotone diameter of $k \times n$ transportation polytopes for fixed k is polynomial in n.*

In Section 3 we show that many well-known combinatorial polytopes have exponentially-long monotone paths, and thus exponential height.

THEOREM 1.5. *The height of the matching, perfect matching, fractional matching and fractional perfect matching polytopes on the complete graph K_n is $> C \cdot \lfloor \frac{n}{2} - 1 \rfloor!$ for a universal constant $C > 0$.*

THEOREM 1.6. *The height of the perfect 2-matching polytope and the TSP with n nodes is $> C \cdot \phi^n$ for a universal constant $C > 0$ and $\phi = \frac{1+\sqrt{5}}{2}$ the golden ratio.*

THEOREM 1.7. *The height of the shortest path polytope on the complete graph K_n is $> \frac{C}{n^2} \sqrt[3]{n!}$ for some universal constant $C > 0$.*

In contrast, we prove that Bland's pivot rule, greatest improvement pivot rule, and steepest-edge pivot rule have polynomial-size simplex height for some combinatorial polytopes. Our discussion includes matching polytopes, fractional matching polytopes, and shortest-path polytopes.

THEOREM 1.8. *The Dantzig simplex height and the greatest improvement simplex height are upper bounded by*
1. *$m\,[n \log(2n)]$ for the fractional perfect matching polytope on a graph with n nodes and m edges. For the complete graph K_n, we get a bound $\sim \frac{n^3}{2} \log n$.*
2. *$m\,[2n \log(2n)]$ for the fractional matching polytope on a graph with n nodes and m edges. For the complete graph K_n, we get a bound $\sim n^3 \log n$.*
3. *$n^2\,[n \log(2n-1)] \sim n^3 \log n$ for the Birkhoff polytope on the bipartite graph $K_{n,n}$.*
4. *$(n^2 - 2n + 1)\,[(n-1)\log(n-1)] \sim n^3 \log n$ for the shortest path polytope on n nodes.*

THEOREM 1.9. *The steepest-edge simplex height is upper bounded by*
1. *$m\left[2\sqrt{2} \cdot n\sqrt{n} \log(2n)\right]$ for the fractional perfect matching polytope on a graph with n nodes and m edges. For the complete graph K_n, we obtain the bound $\sim \sqrt{2} \cdot n^3 \sqrt{n} \log n$.*
2. *$m\left[4\sqrt{2} \cdot n\sqrt{n+1} \log(2n)\right]$ for the fractional matching polytope on a graph with n nodes and m edges. For the complete graph K_n, we obtain the bound $\sim 2\sqrt{2} \cdot n^3 \sqrt{n} \log n$.*
3. *$n^2\left[n\sqrt{\frac{n}{2}} \log(2n-1)\right] \sim n^3 \sqrt{\frac{n}{2}} \log n$ for the Birkhoff polytope on the bipartite graph $K_{n,n}$.*
4. *$(n^2 - 2n + 1)\left[(n-1)\sqrt{\frac{2n}{3}} \log(n-1)\right] \sim n^3 \sqrt{\frac{2n}{3}} \log n$ for the shortest path*

polytope on n nodes.

In Section 4 we revisit the problem of estimating the monotone diameter of transportation polytopes.

THEOREM 1.10. *A $2 \times n$ transportation polytope has monotone diameter $\leq n$. Therefore, $2 \times n$ transportation problems satisfy the monotone Hirsch conjecture.*

2. Monotone and Simplex paths on Cubes & Zonotopes. In this section we present several results about monotone paths and simplex paths on cubes and zonotopes. We will see that they have more general applicability.

We say that two polytopes P and Q are *combinatorially equivalent* if there is a bijection between their faces that preserves the inclusion relation. More precisely, the faces of a polytope P can be made into a lattice L(P), by using the order set by the containment of faces. In this way vertices of P are the atoms of the partial order. To be equivalent, $L(P) = L(Q)$ must be equal as partial orders. See [41] Section 2.2 for details. In what follows we will investigate polytopes that are combinatorially equivalent to hypercubes, which we simply call *combinatorial hypercubes*.

THEOREM 2.1. *Let $C \subset \mathbb{R}^n$ be a combinatorial hypercube. Then mono-diam$(C) = n$. Furthermore, there exists an ordering of the facets of C such that, using Bland's pivot rule with the corresponding ordering of columns, the simplex method leads to a Bland simplex height upper bounded by n.*

Proof. We first prove by induction on n that the monotone diameter is n. The same idea was used for the undirected diameter in [25]. For $n = 1$, the result is trivial. Assume now that the result is true for any combinatorial cube up to dimension $n-1$. Consider an arbitrary vertex $x \neq x^*$ of C. There must exist an improving edge going out of x. Consider a facet F containing x but that does not contain this edge. If $x^* \in F$, we are done by the induction hypothesis. Otherwise x^* belongs to the "opposite" facet. This is the facet of the polytope which does not have any vertex in common with the facet F. For example if C is the regular hypercube, this is the parallel facet to F. We take the improving edge to that opposite facet and apply the induction hypothesis. To conclude note that the monotone diameter is exactly n because there exists a vertex which needs at least n pivots to reach the optimal solution.

Now let us present good orderings of the facets for Bland's pivot rule. Denote by x^* the optimum vertex. We choose an ordering such that the first n facets satisfy $x^* \notin F_i$ for $1 \leq i \leq n$ and the last n facets satisfy $x^* \in F_i$ for $n+1 \leq i \leq 2n$. We will prove that Bland's rule with this ordering follows the path described above. More precisely, we prove that at each step, the index of the entering variable is in $\{1, \ldots, n\}$ while the index of the leaving variable is in $\{n+1, \ldots, 2n\}$ so inserted variables will never be removed from the basis.

Consider an arbitrary vertex $x \neq x^*$. Let $i_1, \ldots, i_k > n$ be the indices such that $x \in F_i$. Since x is not the optimum, there must exist an improving edge from x in the cube $F_{i_1} \cap \ldots \cap F_{i_k}$ of smaller dimension $n-k$. Consider the facet F_i of the n-dimensional cube such that $x \in F_i$ and that does not contain this improving edge. Note that $i \leq n$. Otherwise $x^* \in F_i$, therefore F_i is one of the facets $F_{i_1}, \ldots, F_{i_k}$ which is impossible because the improving edge is contained in their intersection.

The entering variable $\hat{i}$ chosen by Bland's pivot rule satisfies $\hat{i} \leq i \leq n$. Note that $x^* \notin F_{\hat{i}}$ so x^* is contained in the "opposite" facet which corresponds to the leaving variable. Therefore the index of the leaving variable is greater than n. Then, variables of index $\leq n$ cannot be removed from the basis. ∎

Note that these good orderings of the facets for Bland's pivot rule are very rare. Since the n facets not containing the optimum should have the first n indices in the ordering while the n facets containing the optimum should have the last n indices, there are $(n!)^2$ such good orderings among the $(2n)!$ possible orderings of the facets. Therefore, the proportion of good orderings for Bland's pivot rule is $\frac{1}{\binom{2n}{n}}$.

Next, we discuss the monotone diameter of another family of polytopes: zonotopes. A zonotope is the Minkowski sum of a set of edge directions. In the following, we will denote by $Z(v^1, \ldots, v^m)$ the zonotope resulting from the Minkowski sum of edge directions $v^1, \cdots, v^m$.

LEMMA 2.2. *Let $Z(v^1, \ldots, v^m) \subset \mathbb{R}^n$ be the zonotope generated by the edge directions $v^1, \cdots, v^m$. Assume any two directions v^i, v^j are non-colinear. The monotone diameter and the height of the zonotope are equal to m. In particular, the simplex height for any pivot rule is upper bounded by m.*

Proof. Let $c \in \mathbb{R}^n$. We define $J^+ = \{j \mid c^T v^j > 0\}$ and $J^- = \{j \mid c^T v^j < 0\}$. Observe that $x^* = \sum_{j \in J^-} v^j$. Consider a starting point $\hat{x}$. We can write it as $\hat{x} = \sum_{j \in S(\hat{x})} v^j$ for a certain subset $S(\hat{x}) \subset \{1, \ldots, m\}$.

Two adjacent vertices of the zonotope differ by $\pm v^i$ for some i. Then, the only edges a monotone path can use are $\varepsilon_i v^i$ where $\varepsilon_i = 1$ if $i \in J^-$ and $\varepsilon_i = -1$ if $i \in J^+$. Furthermore, the path can follow each of these directions at most once because once $\varepsilon_i v^i$ is added, this term cannot be removed by the other possible directions (two edge directions v^i, v^j are not colinear). Then, the length of the path is at most m.

Furthermore, since the admissible edges are of the form $\varepsilon_i v^i$, the point $\tilde{x} = \sum_{j \in J^+} v^j$ is at distance at least m from the optimum. Note that $\tilde{x}$ is a true vertex of the zonotope because it is the optimum for the cost function $-c$. ∎

LEMMA 2.3. *Let $Z(v^1, \ldots, v^m) \subset \mathbb{R}^n$ be a zonotope. Assume any pair of edge directions v^i, v^j are non-colinear. Then, $Z(v^1, \ldots, v^m)$ has at least $2m$ facets.*

Proof. Let $v^{i_1}, \ldots, v^{i_{n-1}}$ be a linearly independent subset of size $n-1$. Then $Z(v^1, \ldots, v^m)$ has two facets that are translates of $Z(v^{i_1}, \ldots, v^{i_{n-1}})$. This is because for an objective function $c \in \ker([v^{i_1}, \ldots, v^{i_{n-1}}])$ the optimum facet of $Z(v^1, \ldots, v^m)$ with respect to $\pm c$ are precisely these two facets. It suffices to show that there are $\geq m$ distinct subsets of this type.

Without loss of generality, let $v^1, \ldots, v^n$ be linearly independent. All subsets $S^i = \{v^1, \ldots, v^{i-1}, v^{i+1}, \ldots, v^n\}$, $i = 1, \ldots, n$ are n facet-inducing subsets as discussed in the previous paragraph. For any v^j with $j \geq n+1$, there exists $v \in \{v^1, \ldots, v^n\}$ such that $\{v^1, \ldots, v^n\} \setminus \{v\} \cup \{v^j\}$ is linearly independent (by the matroid axiom). Therefore drop v and add v^j, the corresponding subset gives two more facets. We get $m - n$ additional facet-inducing subsets like this. ∎

The following theorem shows that zonotopes satisfy the monotone Hirsch conjecture.

THEOREM 2.4. *For every edge directions $v^1, \ldots, v^m \in \mathbb{R}^n$,*

$$\textit{mono-diam}(Z(v^1, \ldots, v^m)) = m \leq \frac{|\textit{facets}|}{2} \leq |\textit{facets}| - n.$$

Proof. The first equality is given by Lemma 2.2, and the first inequality comes from Lemma 2.3. The second inequality is also a consequence of Lemma 2.3 because $m \geq n$. ∎

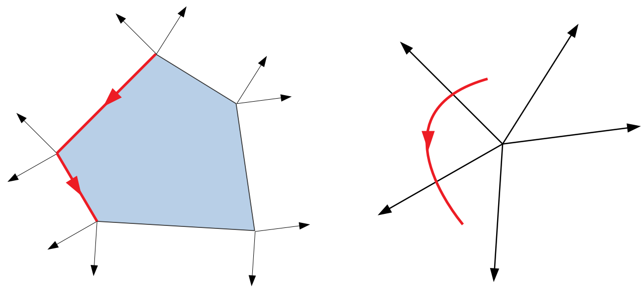

FIG. 2. *Left, a path on a polytope and right, the corresponding path on the normal fan.*

We will now use this result to upper bound the monotone diameter of general polytopes. For this, let us define the *normal cone* of a vertex v as the set of objective vectors c such that v is an optimal vertex for the corresponding objective function, and the *normal fan* as the collection of normal cones for all vertices of the polytope (see [4] and Figure 2 for an illustration). For two normal fans F_1, F_2 defined in the same space, we say that F_2 is a refinement of F_1 if the closure of any normal cone in F_1 can be obtained as the union of the closure of normal cones in F_2.

LEMMA 2.5. *(Gritzmann-Sturmfels, Proposition 2.1.8. in [15]) Let $P \subset \mathbb{R}^n$ be a polytope and let E be a finite set of vectors containing all edge directions of P, that is, a maximal set of non-colinear edges of P. The normal fan of the zonotope generated by E is a refinement of the normal fan of P. In particular, the diameter of $Z(E)$ upper bounds the diameter of P.*

Proof of Theorem 1.3. The first inequality of the theorem uses Lemma 2.5 by viewing a path on the graph of the polytope as a sequence of normal fans where consecutive normal fans share a facet. The normal to this shared facet is the direction of the corresponding edge between the two vertices on the graph of the polytope. Therefore any monotone path p on the zonotope $Z(P)$ for the linear function c leads to a path $\tilde{p}$ with smaller or equal length on the original polytope P. Further, $\tilde{p}$ is still monotone for c because the directions of the edges of $\tilde{p}$ are contained in the directions of the edges of p according to the hypothesis of Theorem 1.3. From this result one can show that the monotone diameter on $Z(P)$ upper bounds the monotone diameter on P using a simple comparison: Let v be the vertex of P such that the length of a shortest $c-$monotone path from v to the optimum of P is the monotone diameter of P. Denote by L such a path of length $mono - diam(P)$. Now let L' be the shortest monotone path on $Z(P)$ from an equivalent vertex to v — such that its cone in the normal fan of $Z(P)$ is included in the cone of v in the normal fan of P – to the optimum in $Z(P)$. By definition, L' is shorter than the monotone diameter of $Z(P)$. As shown above, from L', one can construct a shorter monotone path $\tilde{L}$ from v to the optimum of P. By definition, L is shorter than $\tilde{L}$, which in turn is shorter L'. Therefore, $mono - diam(P) \leq mono - diam(Z(P))$. ∎

Note that we obtained the result mono-diam$(P) \leq$ mono-diam$(Z(P))$ by showing that from a monotone path on $Z(P)$ we can construct a shorter monotone path on P. Unfortunately, this does not prove that height$(P) \leq$ height$(Z(P))$ because we would have to show that from a monotone path on P one can construct a longer path on $Z(P)$ that is still monotone.

We can now apply Theorem 1.3 to several polytopes. The essential message is that if the set of edge-directions is "small" or polynomially bounded, then we can obtain an upper bound on the monotone diameter using the above result. While we show some nice situations below, in most cases this is not useful (see [26] where a lower bound on the number of edge directions is discussed).

Proof of Theorem 1.4. To prove both statements, we will use Theorem 1.3 to upper bound the monotone diameter of the considered polytope by the number of edge directions of the polytope.

1. Let E be the finite set of n elements defining a matroid polytope or a polymatroid polytope. We know from Theorem 5.1 in [39] that edges of the polymatroid are of the form $e_i - e_j$ for $e_i, e_j \in E$. Therefore the number of edge directions (the sign does not count here) is upper bounded by $\binom{n}{2}$.
2. Edges on transportation polytopes are alternating sign cycles on the bipartite graph. Since there are k supply nodes, the length of the cycle is $2p$ for $2 \leq p \leq k$. The number of such cycles of length $2p$ is $\frac{1}{p}\frac{n!k!}{(n-p)!(k-p)!}$. Then, the number of different edge directions is upper bounded by

$$\sum_{p=2}^{k} \frac{1}{p} \frac{n^p k!}{(k-p)!} \leq n^k k! \sum_{p=2}^{k} \frac{1}{(k-p)! n^{k-p}} \leq e^{1/n} n^k k!$$

and the proof follows. ∎

Finally, the transportation polytopes family in Theorem 1.4 is naturally generalized by N-fold linear programs, see Chapter 4 in [7]. In that case the defining matrix A has a very specific shape as multiple copies of smaller matrices. We omit details.

3. Monotone and Simplex paths on $0/1$ and $0/\frac{1}{2}/1$ polyhedra. In this section, we give results on the height, monotone diameter, and the simplex height of some well-known polytopes. It should be noted that the recent paper [8] provides a new general polynomial upper bound for the diameter of $0/1-$polytopes which is independent of the polyhedral representation. However, in this section we look at specific families and thus we can obtain more precise polynomial bounds.

As shown in Section 2, for some particular polytopes (e.g., zonotopes) the height is polynomially bounded, thus it gives a polynomial upper bound for the Simplex algorithm for *any* pivot rule. However, it turns out that, for many polytopes of interest and for some well-known $0/1$ and $0/\frac{1}{2}/1$ polyhedra, monotone paths can be very long.

Let us recall the definitions and basic properties of the combinatorial polytopes we will consider in this section. A *matching* in a graph $G = (V, E)$ is a subset of edges $M \subset E$ such that every vertex is incident to at most one edge of M. A matching is *perfect* if every vertex meets exactly one edge of M. The matching polytope (M) of G is defined as the convex hull of the $0/1$ incidence vectors of matchings i.e.,

$$\mathrm{M(G)} = \mathrm{conv}\{\chi^M : M \text{ is a matching of G}\}.$$

The perfect matching polytope (PM) of G is the convex hull of the incidence vectors of the perfect matchings. Note that the perfect matching polytope on the complete bipartite graph is the Birkhoff polytope.

$$\mathrm{PM(G)} = \mathrm{conv}\{\chi^M : M \text{ is a perfect matching of G}\}.$$

For these two polytopes, two matchings are adjacent if and only if the union of

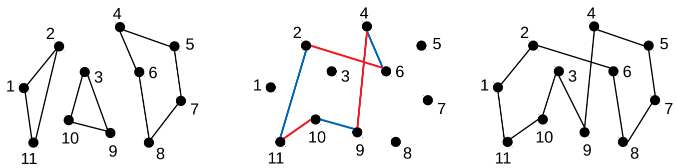

FIG. 3. *Left and right, two adjacent perfect 2-matchings. In the middle, the corresponding alternating cycle.*

their support graph contains a unique cycle (see Lemma 1 in [30]). A set of inequalities describing these polytopes is given by the Edmond's matching theorem [11].

We also consider the relaxations of these polytopes obtained by omitting the *odd cycle inequalities*. The fractional matching polytope (FM) of G is defined by

$$\mathrm{FM(G)} = \{x \in \mathbb{R}^E(G) : \ x_e \geq 0 \ \forall e \in E(G), \ x(\delta(v)) \leq 1 \ \forall v \in V(G)\},$$

where $E(G), V(G)$ denote, respectively, the sets of edges and vertices of the graph G. Similarly, the fractional perfect matching (FPM) is described by

$$\mathrm{FPM(G)} = \{x \in \mathbb{R}^E(G) : \ x_e \geq 0 \ \forall e \in E(G), \ x(\delta(v)) = 1 \ \forall v \in V(G)\}.$$

The adjacency of these fractional polytopes is given in Theorem 25 of [2]. In the following we will only use the fact that the graph of M(G) and PM(G) are, respectively, a subgraph of FM(G) and FPM(G).

A 2-perfect matching of G is a subset of edges M such that every vertex is incident to exactly 2 edges in M. Note that a 2-perfect matching is the union of disjoint cycles. The perfect 2-matching polytope (P2M) of G is defined as a 0/1 polytope as follows,

$$\mathrm{P2M(G)} = \mathrm{conv}\{\chi^M : M \text{ is a perfect 2-matching of G}\}.$$

Two 2-perfect matchings are adjacent if and only if the symmetric difference of their support graphs contains a unique alternating cycle (see Lemma 1 in [31] and Figure 3 for an illustration).

In the following, if the graph is not specified we will consider the complete graph K_n.

The traveling salesman polytope (TSP) on K_n is the convex hull of tours i.e., cycles of length n. In the following, we will also use the term n-tours when needed, to clarify the number of nodes in the considered graph. The TSP graph is therefore a proper subgraph of the perfect 2-matching polytope of K_n (see [33]).

Finally, the shortest path polytope on n nodes is defined as the convex hull of paths from say node 1 to node n without cycles. A system of equations and inequalities describing the shortest path polytope is given by

$$\left\{ x \geq 0 : \sum_{j=2}^{n} x_{1,j} = 1, \sum_{j \neq i} x_{i,j} - \sum_{j \neq i} x_{j,i} = 0, \sum_{j \neq i} x_{i,j} \leq 1, \ 2 \leq i \leq n-1 \right\},$$

where $x = (x_{i,j})_{1 \leq i \leq n-1, \ 2 \leq j \leq n}$ (see [30] for an equivalent system). Two paths are adjacent if and only if the union of their support graph contains a unique cycle (see Lemma 2 in [30]).

Here we collect some of the results about the height of polyhedra. We first state a result from Pak [28] that we will use as a lemma for our next results, in which he shows that the height of the Birkhoff polytope is exponential.

LEMMA 3.1 (Pak, Theorem 1.4. in [28]). *There exists a linear function ϕ with a decreasing sequence of vertices of length $> C \cdot n!$ on the Birkhoff polytope on the bipartite graph $K_{n,n}$ for a universal constant $C > 0$.*

Note that the graph of a proper face F of a polytope P is a proper subgraph of the graph of P. Therefore the height of a polytope is greater or equal to the height of any of its faces. Indeed, let c be a cost function in F. For P take that same cost function parallel to F and denote it by $\tilde{c}$. Then any c-monotone path in F is a $\tilde{c}$-monotone path in P.

Proof of Theorem 1.5. We show that the Birkhoff polytope is a face of each of the considered polytopes. We will denote by $x_{i,j}$ the component corresponding to the edge between nodes i and j for a vertex x in the polytope. Note that graphs are non-oriented here.

Define $E_1 := \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$ and $E_2 := \{\lfloor \frac{n}{2} \rfloor + 1, \dots, 2\lfloor \frac{n}{2} \rfloor\}$. For the matching polytope and the fractional matching polytope, the corresponding face can be described by the several equalities $x_{i,j} = 0$ for $(i,j) \notin E_1 \times E_2 \cup E_2 \times E_1$ and $x(\delta(i)) = 1$ for $i \in E_1 \cup E_2$. In both the matching and fractional matching polytopes, these equalities describe the set of perfect matchings on the bipartite graph between E_1 and E_2 i.e., vertices of these facets are in exact correspondence with the vertices of the K_{E_1,E_2} Birkhoff polytope. Furthermore, the adjacency between the perfect matchings of these faces is exactly the same as in the Birkhoff polytope. Hence the corresponding face is equivalent to the Birkhoff polytope with $2 \times \lfloor \frac{n}{2} \rfloor$ nodes. The height of these polytopes is therefore greater than the lower bound for the height of the Birkhoff polytope $C\lfloor \frac{n}{2} \rfloor!$ given in Lemma 3.1.

For the perfect matching polytope we can simply take the equalities $x_{i,j} = 0$ for $(i,j) \notin E_1 \times E_2 \cup E_2 \times E_1$. The other equalities of the form $x(\delta(i))$ are already satisfied. We get the same lower bound $C\lfloor \frac{n}{2} \rfloor!$ for the height.

The same argument holds for the fractional perfect matching polytope when n is even. However, when n is odd, matchings on K_{E_1,E_2} are not vertices of the polytope anymore. In this case, we can restrict to the face $x_{n-2,n-1} = x_{n-1,n} = x_{n-2,n} = 1/2$ and use the same arguments as above with $E_1 := \{1, \dots, \frac{n-3}{2}\}$ and $E_2 := \{\frac{n-1}{2}, \dots, n-3\}$. We finally get the lower bound $C\lfloor \frac{n}{2} - 1 \rfloor!$ for the height. ◻

Proof of Theorem 1.6. Recall that tours, which are the vertices of the TSP, are also vertices of the perfect 2-matching polytope. If two tours are adjacent on the perfect 2-matching polytope, then they are also adjacent in TSP (see [33]). Therefore it suffices to prove that there exists a long monotone path on the perfect 2-matching polytope going only through tours.

Denote by $x_{i,j}$ the component corresponding to the edge between nodes i and j for a vertex x in the polytope. Consider the following linear function:

$$\psi = x_{1,2} + \alpha x_{1,3} + \ldots + \alpha^{n-2} x_{1,n} + \alpha^{n-1} x_{2,3} + \alpha^{n} x_{2,4} + \ldots + \alpha^{\frac{n(n-1)}{2}-1} x_{n-1,n}$$

for $0 < \alpha < 1/2$ such that the linear order on the perfect 2-matching polytope, or on the TSP, is the lexicographic order on the edges with the following order: $\{1,2\}, \{1,3\}, \dots, \{1,n\}, \{2,3\}, \dots, \{n-1,n\}$.

Denote by $x^* = (1, n, 2, n-2, 4, \cdots, n-3, 3, n-1)$ the optimum for TSP (see Figure 5e). The initial tour is going to be the cycle $x^0 = (1, 2, \dots, n)$. We will construct

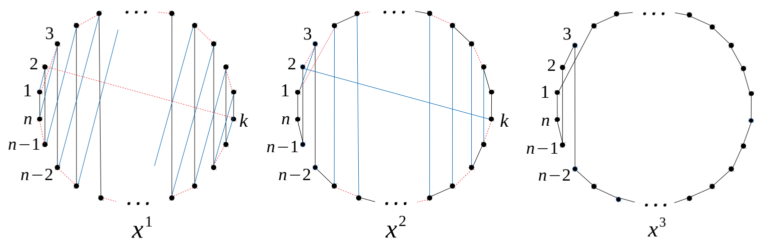

FIG. 4. *Step 2 of the monotone path on TSP. Edges in blue are the edges going to be deleted and the dashed red edges are going to be inserted. Since they form an alternating cycle, these tours are adjacent.*

by induction a monotone path with exponential length starting in x^0 and ending in x^*. We denote by L_n the length of this monotone path. For $n \geq 4$, assume that we have constructed these long monotone paths for $k = 4, \cdots, n-1$. Let us now construct the path of length L_n.

Step 1: We first restrict to $x_{1,2} = 1$. We can get to the optimum x^1 of this facet (see Figure 5a) in at least L_{n-1} steps. Indeed if x is a $(n-1)$-tour in the long path for $n-1$ nodes, define $\tilde{x}$ a n-tour by dividing node 1 into two nodes 1 and 2. The indices of the other nodes should be shifted by one accordingly. Recall that two 2-matchings are adjacent if and only if the symmetric difference of their edges defines a unique alternating cycle. Let x_1 and x_2 be two adjacent tours in the $(n-1)$-perfect 2-matching. Then either $\tilde{x}_1$ and $\tilde{x}_2$ are adjacent in the n-perfect 2-matching or $\tilde{x}_1$ and $\hat{x}_2$ are adjacent where $\hat{x}_2$ is the same tour as $\tilde{x}_2$ except the two nodes coming from the division of node 1 have been switched. We can therefore construct a path of length L_{n-1} corresponding to the same path for $(n-1)$-tours. We then get from the corresponding end point to the optimum x^1 of the facet $x_{1,2} = 1$. These two tours might be distinct if we have to switch the two nodes coming from the division of node 1, which takes at most one step.

Step 2: We now get in two improving steps to the tour $x^3 = (1, 4, 5, \ldots, n-2, 3, 2, n-1, n)$ (see Figure 4). The edges of the current vertex x^1 are $x^1_{i,n+1-i} = 1$ for all i, $x^1_{1,2} = x^1_{i,n+3-i} = 1$ for $i \geq 3$. We now get to the tour $x^2 = (2, n-1, n, 1, 3, n-2, n-3, 4, 5, \ldots, k)$ which uses all the edges of the form $x_{i,n+1-i}$. Here $k = \frac{n}{2} + 1$ if $n \equiv 0 \mod 4$, $k = \frac{n}{2}$ if $n \equiv 2 \mod 4$ and $k = \frac{n+1}{2}$ otherwise. This is an adjacent node because the symmetric difference of the graphs of the two tours has a unique alternating cycle $(2, 1, 3, n, n-1, 4, 5, n-2, n-3, 6, 7, \ldots k)$. The precise end of this alternating cycle depends on $n \mod 4$. If $n \equiv 0 \mod 4$, the ending of this cycle is $(\ldots, k-2, k-1, k+2, k+1, k)$. If $n \equiv 2 \mod 4$, it is $(\ldots, k-2, k-1, k+4, k+3, k)$. For $n \equiv 1 \mod 4$, it is $(\ldots, k-2, k+4, k+3, k-1, k, k+2, k+1, k)$ and for $n \equiv 3 \mod 4$ it is $(\ldots, k+2, k+1, k+3, k+2, k)$. Because $x^2_{1,2} = 0$, this is an improving step for the lexicographic order on the edges. Now use the alternating cycle $(2, 3, 1, 4, n-3, n-4, 5, 6, n-5, n-6, \ldots, k)$ to get to the neighbor tour $x^3 = (1, 4, 5, \ldots, n-2, 3, 2, n-1, n)$. More precisely, the ending of the alternating cycle is $(\ldots, k-3, k-2, k+1, k)$ if $n \equiv 0 \mod 4$, $(\ldots, k-2, k+3, k+2, k-1, k)$ if $n \equiv 2 \mod 4$, $(\ldots, k-2, k-1, k+1, k)$ if $n \equiv 1 \mod 4$ and $(\ldots, k-2, k+2, k+1, k-1, k)$ if $n \equiv 3 \mod 4$. This is also an improving step because $x^3_{1,2} = x^3_{1,3} = 0$.

Now we fix $x_{n-1,2} = x_{2,3} = x_{3,n-2} = 1$. We get to optimal tour of this facet (see Figure 5b) in at least L_{n-3} steps, similarly to the technique used for the L_{n-1} long

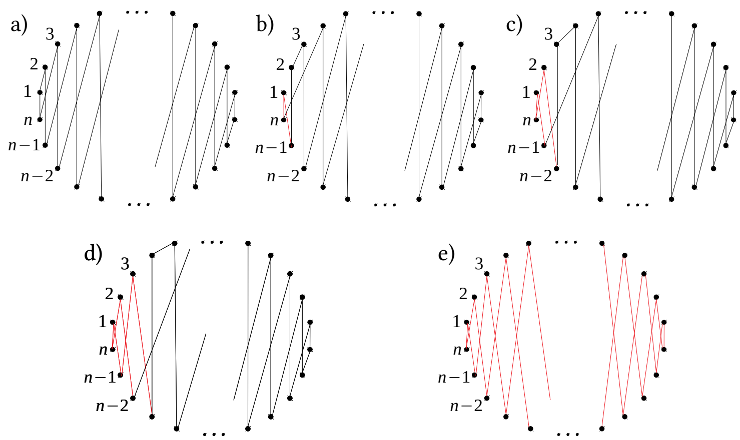

FIG. 5. *Main steps of the monotone path. Once the red edges that belong to the optimum e) are inserted they will never be deleted.*

step: in the n-tour, nodes $n-1, 2, 3$ and $n-2$ will be merged together to obtain a $(n-3)$-tour.

Step 3: Note that now $\{1, n\}$ and $\{1, n-1\}$ are edges that will never be removed so we are restricted to the facet $x_{1,n} = x_{1,n-1} = 1$. Merging together nodes $1, n, n-1$, the resulting $(n-2)$-tour is exactly the tour given at the end of step 1 for $n-2$ nodes. Apply Step 2 again to get to the next tour in at least $2 + L_{n-5}$ steps which is the optimum of the facet $x_{1,n} = x_{1,n-1} = x_{n-2,3} = x_{3,4} = x_{4,n-3} = 1$ (see Figure 5c). Now, $\{2, n\}$ and $\{2, n-2\}$ are edges that will never be removed so we are restricted to the facet $x_{2,n} = x_{2,n-2} = 1$. With the same arguments, we progressively reconstruct the edges of the optimum x^* in at least $2 + L_{n-7} + 2 + L_{n-9} + \ldots$ steps.

Together, we have $L_n \geq L_{n-1} + 2 + L_{n-3} + 2 + L_{n-5} + \ldots + 2 + L_k$ with $k = 4$ if n is odd and $k = 5$ otherwise. Define $\tilde{L}_n$ by $\tilde{L}_4 = 1$, $\tilde{L}_5 = 3$ and $\tilde{L}_n = \tilde{L}_{n-1} + 2 + \tilde{L}_{n-3} + 2 + \tilde{L}_{n-5} + \ldots + 2 + \tilde{L}_k$. Then $L_n \geq \tilde{L}_n$ because $L_4 \geq 1$ and $L_5 \geq 3$. Furthermore, $\tilde{L}_n = \tilde{L}_{n-1} + \tilde{L}_{n-2} + 2$ therefore note that $\tilde{L}_n + 2 = F_n$ is the Fibonacci sequence. Then $L_n \geq \tilde{L}_n \geq \frac{\phi^n}{\sqrt{5}} - 3$. ∎

Proof of Theorem 1.7. Recall that the vertices of the shortest path polytope are the paths from node say 1 to n and that two paths from 1 to n are adjacent if and only if the union of their graphs contains a unique cycle. Denote by $x_{i,j}$ the coordinate of the edge going from node $i \neq n$ to $j \neq 1$ in a vertex x of the polytope. Similarly to the cost function used in Theorem 1.6 we use the linear function

$$\psi = x_{1,2} + \alpha x_{1,3} + \ldots + \alpha^{n-2} x_{1,n} + \alpha^{n-1} x_{2,3} + \ldots + \alpha^{2n-4} x_{2,n} + \ldots + \alpha^{n^2-3n+2} x_{n-1,n},$$

so that the linear order is the lexicographic order on the edges $\{1, 2\}, \{1, 3\}, \ldots, \{1, n\}, \{2, 3\}, \ldots, \{2, n\}, \{3, 2\}, \{3, 4\}, \ldots, \{3, n\}, \ldots, \{n-1, n\}$ with a chosen small enough $\alpha > 0$. We start from the path $1, 2, \ldots, n$ which is the maximum value vertex for ψ. Denote by L_n the length of the monotone path we will construct here by induction.

Step 1: Fix the edge $x_{1,2} = 1$. This facet corresponds to the shortest path polytope on the complete graph K_{n-1} with nodes $2, 3, \ldots, n$. The objective function ψ is still the same lexicographic order on the edges of K_{n-1}. Then, by induction, we

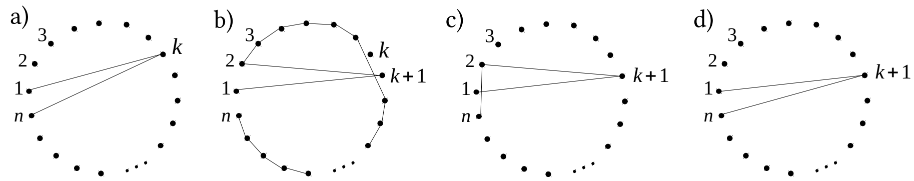

FIG. 6. *Step 3 of the long monotone path on the shortest path polytope. The length of the path from a) to d) is* $L_{n-3}+2$.

can get to path $1, 2, n$ in L_{n-1} monotone steps.

Step 2: We now get to the path $1, 3, 4, \dots, n$ which is a decreasing neighbor because we do not use the edge $\{1, 2\}$ anymore. Similarly to Step 1, we get to path $1, 3, n$ in L_{n-2} monotone steps.

Step 3: We are now going to go from path $1, 3, n$ to $1, 4, n$, then to $1, 5, n$ etc... to $1, n-1, n$. Figure 6 shows how to go from the path $1, k, n$ to the path $1, k+1, n$ where $3 \le k \le n-1$. From the path $1, k, n$ (see Figure Figure 6 a) we first get to the decreasing neighbor $1, k+1, 2, 3, \dots, k-1, k+2, k+3, \dots, n$ (see Figure 6 b). Fixing edges $x_{1,k+1} = x_{k+1,2} = 1$, this facet is equivalent to the shortest path on the complete graph K_{n-3} with nodes $2, 3, \dots, k-1, k+2, k+3, \dots, n$, starting in 2 and ending in n. We therefore get to path $1, k+1, 2, n$ (see Figure 6 c) in L_{n-3} steps and then to path $1, k+1, n$ (see Figure 6 d) in an improving step. We can repeat this operation $n-4$ times until we reach path $1, n-1, n$. We finally get to path $1, n$ in one improving step. All together we get

$$L_n = L_{n-1} + L_{n-2} + (n-4)L_{n-3} + 2(n-3) \ge (n-2)L_{n-3}.$$

Therefore $L_{3k+2} \ge 3^k \cdot k!$ and $L_{3k+1}, L_{3k} \ge 3^{k-1} \cdot (k-1)!$ where $3^k \cdot k! \sim \frac{\tilde{C}}{k^{1/3}} \sqrt[3]{(3k)!}$ for some constant $\tilde{C}$. The result follows. ∎

Although the height of all the combinatorial polytopes above is exponential, several authors have shown that their monotone diameter can be short. For example Rispoli [30] showed that the monotone diameter of the Birkhoff polytope of vertices in $\mathcal{S}_n$ is $\lfloor \frac{n}{2} \rfloor$. Furthermore, he also proved that several matching polytopes [31], the shortest path polytope [30] and the TSP [32] have linear monotone diameter.

We now give estimates for their simplex height for some specific pivot rules. For this we use an analysis of the number of basic feasible solutions (BFS) generated by the algorithm. The ideas we use are inspired from the work of Kitahara, Mizuno and co-authors (see [18], [19] and [35]).

Consider the following linear program in standard form for a bounded polytope:

$$\begin{aligned} \text{(3.1)} \qquad & \min \quad c^T x \\ & \text{s.t.} \quad Ax = b, \quad x \ge 0 \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $m < n$ and A is a matrix with full row rank.

For a given BFS x, let B and N denote the submatrices of A corresponding to basic and non-basic columns respectively. We split the objective function vector c and the variables x accordingly,

$$c = \begin{bmatrix} c_B \\ c_N \end{bmatrix}, \quad x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}, \quad x_B = B^{-1}b, \quad x_N = 0.$$

Definition 3.2. *Define γ and δ respectively as the maximum and the minimum among the positive coordinates of all BFS. We also denote by ν and μ respectively the maximum and minimum among the Euclidean length of all possible edges.*

In the paper [19] Kitahara and Mizuno proved that for Dantzig's pivot rule and the greatest (descent) improvement pivot rule, the number of steps is upper bounded by $n\left[\frac{m\gamma}{\delta}\log\left(m\frac{\gamma}{\delta}\right)\right]$ iterations. In [18] Kitahara, Matsui and Mizuno improved that result and obtained the following upper bound:

$$(n-m)\left[\min\{m,n-m\}\frac{\gamma}{\delta}\log\left(\min\{m,n-m\}\frac{\gamma}{\delta}\right)\right].$$

Tano, Miyashiro and Kitahara [35] then showed that the number of different BFS for the generalized $p-$norm steepest edge rule is upper bounded by

$$(n-m)\left[m^{1+1/p}\frac{\gamma^2}{\delta^2}\log\left(m\frac{\gamma}{\delta}\right)\right].$$

Next we derive another new upper bound for the *steepest edge* pivot rule ($p=2$) which later will be applicable to the polytopes of our interest. See Theorem 3.7. We remark that the resulting bounds are in general still exponential in the bit-size of the input, and that the constants are complicated to compute. For example, δ is NP-hard to compute in general (see [23]).

Consider now a single step of steepest edge pivoting rule for the Simplex method. To simplify the argument, we assume that the current basis consists of the first m columns. If column q ($q>m$) is entering the basis and the column p is leaving the basis, then the next BFS $\bar{x}$ we encounter would be of the form

$$\bar{x}=x+\theta\eta_q$$

where θ is the step-size, and η_N^q is the pivot direction from the set of edge directions $\eta_N=[\eta_N^{m+1},...,\eta_N^n]$,

$$\eta_N^q=\begin{bmatrix}-B^{-1}N\\ I\end{bmatrix}e_{q-m}.$$

Let $\bar{c}_N$ denote the reduced cost vector for non-basic variables, so

$$\bar{c}_N^q=c^T\eta_N^q,\quad \bar{c}_N^T=c^T\eta_N=c_N^T-c_B^TB^{-1}N. \tag{3.2}$$

Denote by ζ_N^q the Euclidean norm of $q-$th edge direction, and W_N a diagonal matrix whose diagonal elements are ζ_N^q.

$$\zeta_N^q=\|\eta_N^q\|_2,\quad W_N=diag(\zeta_N^{m+1},...,\zeta_N^n).$$

In the steepest edge Simplex algorithm, we determine our pivoting column by minimizing the normalized reduced cost i.e., choosing $\hat{q}$ such that

$$\hat{q}=\arg\min\bar{c}_N^q/\zeta_N^q.$$

Set $\Lambda=-\bar{c}_N^{\hat{q}}/\zeta_N^{\hat{q}}>0$. With all the notations above, Problem Section 3 can be rewritten as

$$\min_{x_N}\quad c_B^TB^{-1}b+\bar{c}_N^TW_N^{-T}W_Nx_N. \tag{3.3}$$

$$\text{s.t.} \quad x_B = B^{-1}b - B^{-1}Nx_n,$$
$$x_B \geq 0, \quad x_N \geq 0.$$

Note that $W_N^{-1}\bar{c}_N$ is the normalized reduced cost vector.

Lemma 1 of [19] gives an upper bound on the distance between the current objective value and the optimal value. The following lemma is an extension for the steepest edge pivoting rule.

LEMMA 3.3. *Assume z^* is the optimal value and $x^{(t)}$ the BFS generated at the $t-th$ iteration, with the corresponding basic and non-basic columns $B^{(t)}, N^{(t)}$. Then we have*

$$z^* \geq c^T x^{(t)} - \Lambda^{(t)} m\nu \frac{\gamma}{\delta}.$$

Proof. The proof of this lemma comes from modifications of the techniques used in [19] to extend the results to the steepest edge pivoting rule. We decompose the optimal value z^* with the current basis.

$$\begin{aligned} z^* &= c^T x^* \\ &= c^T x^{(t)} + \bar{c}_{N^{(t)}}^T x_{N^{(t)}}^* \\ &= c^T x^{(t)} + \bar{c}_{N^{(t)}}^T W_{N^{(t)}}^{-T} W_{N^{(t)}} x_{N^{(t)}}^*. \end{aligned}$$

Using the definition of $\Lambda^{(t)}$ we get

$$\begin{aligned} z^* &\geq c^T x^{(t)} - \Lambda^{(t)} e^T W_{N^{(t)}} x_{N^{(t)}}^* \\ &\geq c^T x^{(t)} - \Lambda^{(t)} \left(e^T W_{N^{(t)}} e\right) \gamma \\ &\geq c^T x^{(t)} - \Lambda^{(t)} m \frac{\nu}{\delta} \gamma, \end{aligned}$$

where the last inequality results from the definition of ν. ∎

The following theorem shows the decreasing rate of the gap between the optimal value and the objective value at iteration t.

THEOREM 3.4. *For the steepest edge pivoting rule, if the t-th iterate $x^{(t)}$ is not optimal then*

$$\frac{c^T x^{(t+1)} - z^*}{c^T x^{(t)} - z^*} \leq 1 - \frac{\mu\delta^2}{m\nu\gamma^2}.$$

Proof.

$$\begin{aligned} c^T x^{(t)} - c^T x^{(t+1)} &= \Lambda^{(t)} \zeta_{N^{(t)}}^{\hat{q}^{(t)}} x_{\hat{q}^{(t)}}^{(t+1)} \\ &\geq \Lambda^{(t)} \frac{\mu}{\gamma} \delta \\ &\geq \frac{\mu\delta^2}{m\nu\gamma^2} (c^T x^{(t)} - z^*). \end{aligned}$$

This theorem is an analog of Theorem 1 in [19] for steepest edge pivoting rules and uses similar proof techniques. The last inequality follows from Lemma 3.3. Rearranging the terms gives us the desired result. ∎

Lemma 2 in the original paper [19] does not depend on pivoting rules, so it can be applied directly here.

LEMMA 3.5. *(Kitahara and Mizuno, Lemma 2 in [19]) If $x^{(t)}$ is not optimal, then there exists $\bar{j} \in B^t$, such that $x^{(t)} > 0$, and for any k, $x^{(k)}$ satisfies*

$$x_{\bar{j}}^{(k)} \leq \frac{m(c^T x^{(k)} - z^*)}{c^T x^{(t)} - z^*} x_{\bar{j}}^{(t)}.$$

Combining the results from Theorem 3.4 and Lemma 3.5, we have the following lemma.

LEMMA 3.6. *If $x^{(t)}$ is not an optimal solution, then there exists $\bar{j} \in B^t$, such that $x_{\bar{j}}^{(t)} > 0$ and becomes zero and stays zero after $\left[\frac{m\gamma^2\nu}{\delta^2\mu} \log\left(m\frac{\gamma}{\delta}\right)\right]$ iterations.*

Proof.

$$x_{\bar{j}}^{(t+k)} \leq m\left(1 - \frac{\mu\delta^2}{m\nu\gamma^2}\right)^k x_{\bar{j}}^{(t)} \leq m\gamma\left(1 - \frac{\mu\delta^2}{m\nu\gamma^2}\right)^k \leq m\gamma \exp\left(-\frac{k\mu\delta^2}{m\nu\gamma^2}\right).$$

Therefore, if $k > \left[\frac{m\gamma^2\nu}{\delta^2\mu} log\left(m\frac{\gamma}{\delta}\right)\right]$, we would have $x_{\bar{j}}^{(t+k)} < \delta$. By the definition of δ, the lemma follows. ∎

The event described in Lemma 3.6 can happen at most once for each variable. Since we have in total n variables, we have the following theorem.

THEOREM 3.7. *The steepest-edge simplex height for the Problem (3.1) is upper bounded by*

$$n\left[\frac{m\gamma^2\nu}{\delta^2\mu} \log\left(m\frac{\gamma}{\delta}\right)\right]. \tag{3.4}$$

In other words, the steepest edge algorithm reaches the optimal solution in at most $n\left[\frac{m\gamma^2\nu}{\delta^2\mu} \log\left(m\frac{\gamma}{\delta}\right)\right]$ non-degenerate pivots.

As a remark, we will now show that from Theorem 3.7 we can derive similar but weaker upper bounds to those given by Tano, Miyashiro and Kitahara [35] for steepest edge. We give an upper bound in terms of the sub-determinants of the input matrix A. In the following, we will denote by Δ and λ respectively the maximum and minimum absolute value of non-zero determinants over the $m \times m$ sub-matrices of A.

LEMMA 3.8. *For any $m \times m$ sub-matrix B of A and any column A_k of the matrix A, $\|B^{-1}A_k\|_2 \leq \sqrt{m}\frac{\Delta}{\lambda}$.*

Proof. By Cramer's rule, the j-th entry of $B^{-1}A_k$ is given by $\frac{\det(B_j)}{\det(B)}$ for any $j \in \{1, \dots, m\}$, where B_j is the matrix obtained by replacing the j-th column of B by A_k. Since A_k is also a column of A, B_j is an $m \times m$ submatrix of A. Thus, $|\frac{\det(B_j)}{\det(B)}| \leq \frac{\Delta}{\lambda}$. The bound follows. ∎

Remark 3.9. The steepest-edge simplex height for the Problem (3.1) is upper bounded by

$$n\left[m\sqrt{2m}\frac{\gamma^3\Delta}{\delta^3\lambda} \log\left(m\frac{\gamma}{\delta}\right)\right]. \tag{3.5}$$

Proof. Let x_B be the vertex corresponding to a basis B, and a neighbor $\tilde{x}$. Denote by $\hat{q}$ the entering variable to get from x_B to $\tilde{x}$. Then $\tilde{x} - x_B = -\tilde{x}_{\hat{q}} A_B^{-1} A_{\hat{q}}$ where $A_{\hat{q}}$ is the $\hat{q}$-th column of A and A_B is the $m \times m$ submatrix of A of columns in the basis B. Then,

$$\|\tilde{x} - x_B\|_2 = \tilde{x}_{\hat{q}} \sqrt{1 + \|A_B^{-1} A_{\hat{q}}\|_2^2}.$$

By Lemma 3.8, $\nu \leq \gamma\sqrt{1 + m\left(\frac{\Delta}{\lambda}\right)^2}$ and $\mu \geq \delta$. The proof follows from the upper bound given in Theorem 3.7. ∎

When the matrix A is totally unimodular, Remark 3.9 gives an upper bound for the number of different BFS of $n\left[m\sqrt{2m}\frac{\gamma^3}{\delta^3}\log\left(m\frac{\gamma}{\delta}\right)\right]$ for the steepest edge rule. In this case we get a very similar bound to that given by Tano, Miyashiro and Kitahara [35]. In addition, when b is integral, Kitahara and Mizuno [19] derived from their result the upper bound $n[m\|b\|_1 \log(m\|b\|_1)]$ on the number of different BFS generated by the simplex method with Dantzig's rule or the greatest improvement rule. Here we improve this result for different polytopes of interest and give the corresponding explicit polynomial upper bounds.

COROLLARY 3.10. *The Dantzig simplex height and the greatest improvement simplex height for a transportation problem written as $Ax = b$, $x \geq 0$ are upper bounded by*

$$n\left[\|b\|_1 \log\left(m\|b\|_\infty\right)\right] \tag{3.6}$$

and more precisely by $n[S\log(m\|b\|_\infty)]$ where S is the total supply, equal to the total demand in the transportation problem. In other words, at most $n[S\log(m\|b\|_\infty)]$ different BFS are generated by the Dantzig algorithm or the greatest improvement algorithm.

Proof. We slightly change the proof of the result given by Kitahara and Mizuno [19].

$$\begin{aligned} z^* &= c^T x^* \\ &= c^T x^{(t)} + \bar{c}_{N^{(t)}}^T x_{N^{(t)}}^* \\ &\geq c^T x^{(t)} - \Delta^{(t)} \|x_{N^{(t)}}^*\|_1 \end{aligned}$$

where $\Delta^{(t)} = -\min \bar{c}_N^q$. If $x_{i,j}$ is the value for the edge from supply node i to demand node j, $\|x_{N^{(t)}}^*\|_1 \leq \|x^*\|_1 \leq \sum_{i,j} x_{i,j}^* = S$ the total supply (or total demand). Similarly to the proof of Theorem 3.4, we use the above inequality to find

$$\begin{aligned} c^T x^{(t)} - c^T x^{(t+1)} &= \Delta^{(t)} x_{\hat{q}^{(t)}}^{(t+1)} \\ &\geq \Delta^{(t)}\delta \\ &\geq \frac{\delta}{S}(c^T x^{(t)} - z^*). \end{aligned}$$

Therefore $c^T x^{(t+1)} - z^* \leq \left(1 - \frac{\delta}{S}\right)(c^T x^{(t)} - z^*)$. Using Lemma 3.5, we get

$$x_{\bar{j}}^{(t+k)} \leq m\left(1 - \frac{\delta}{S}\right)^k x_{\bar{j}}^{(t)} \leq m\gamma\left(1 - \frac{\delta}{S}\right)^k \leq m\gamma e^{-\frac{k\delta}{S}}.$$

The number of different BFS is then at most $n[\frac{S}{\delta}\log(m\frac{\gamma}{\delta})]$. As noted in [19], since A is a totally unimodular matrix, δ is a positive integer, so $\delta \geq 1$. Denote by s_i and v_j the supply and demand at supply node i and demand node j respectively. Then $\gamma = \max x_{i,j} \leq \min(\max_i s_i, \max_j d_j) \leq \|b\|_\infty$. The proof follows. ∎

Note that in the proof of Corollary 3.10, instead of replacing $\|x^*\|_1$ by $m\gamma$, we kept $\|x^*\|_1$. If we do the same in the proof of Theorem 3.7, we obtain additional upper bounds for the number of generated BFS for several pivot rules in the following lemma.

LEMMA 3.11.
1. *The Dantzig simplex height and the greatest improvement simplex height for Problem (3.1) are upper bounded by* $n\left[\frac{\|x^*\|_1}{\delta}\log\left(m\frac{\gamma}{\delta}\right)\right]$.
2. *The steepest edge simplex height for Problem (3.1) is upper bounded by*

$$\left[\frac{\|x^*\|_1\gamma\nu}{\delta^2\mu}\log\left(m\frac{\gamma}{\delta}\right)\right].$$

We are now ready to use Lemma 3.11 to prove our upper bounds on several combinatorial polytopes.

Proof of Theorem 1.8 and Theorem 1.9. We prove the two theorems in parallel, as we only need to apply two different estimations to the same polytope for each item of the same index as listed in the theorems.

1. The fractional perfect matching polytope is a $0/\frac{1}{2}/1$ polytope so $\gamma = 1$ and $\delta = 1/2$. Furthermore, $x \in FPM$ is a vertex if and only if it is the union of a perfect matching $\mathcal{M}_x$ given by the edges $\{e \in E, x_e = 1\}$ and a collection $\mathcal{C}_x$ of disjoint cycles of odd length given by the edges $\{e \in E, x_e = 1/2\}$. Then $\|x\|_1 = \frac{k_1+k_2}{2}$ where k_1 is the number of nodes in the odd length cycles and k_2 the number of nodes in the matching $\mathcal{M}_x$. Therefore $\|x^*\|_1 = \frac{|V|}{2}$. Now let us give bounds for μ and ν. For two vertices x_1 and x_2 and any edge $e \in E$, $|(x_1 - x_2)_e| \leq 1$. Then, $\|x_1 - x_2\|_2^2 \leq \|x_1 - x_2\|_1 \leq |V|$ so $\nu \leq \sqrt{|V|}$. Furthermore $\mu \geq \sqrt{2}\cdot\delta$. Indeed, two adjacent vertices differ at least by δ for the entering variable and exiting variable coordinates. Thus, $\nu \geq \sqrt{2}/2$.
2. The fractional matching polytope is still a half integral polytope so $\gamma = 1$ and $\delta = 1/2$. Vertices are still the union of a perfect matching on $\mathcal{M}_x$ given by the edges $\{e \in E, x_e = 1\}$ and disjoint odd-length cycles $\mathcal{C}_x$ given by the edges $\{e \in E, x_e = 1/2\}$. We have to add the n slack variables s_i for the inequality at each node so $\|x\|_1 = |\mathcal{M}_x|/2 + |\mathcal{C}_x|/2 + |V - (\mathcal{M}_x \cup \mathcal{C}_x)|$ where the last term comes from the slack variables. Then, $\|x^*\|_1 \leq |V|$. Note that two adjacent vertices differ by at most $m+1$ coordinates, corresponding to the basis variables and the entering variable. Therefore, $\nu \leq \sqrt{m+1}\gamma$. Therefore, $\nu \leq \sqrt{|V|+1}$. Finally, the same arguments as above give $\mu \geq \sqrt{2}/2$.

   The next polytopes are 0/1 polytopes, therefore $\gamma = \delta = 1$.
3. The Birkhoff polytope has exactly n positive edges then $\|x\|_1 = n$ for any permutation x. Two vertices x, y are adjacent on this polytope if the symmetric difference of their edges form a single alternating cycle of norm $\sqrt{l}$ where l is its length. Because the cycle is alternating, we have $4 \leq l \leq 2n$ and then $\mu = 2$, $\nu = \sqrt{2n}$.
4. For the shortest path polytope, there are $n^2 - 3n + 3$ variables and $n - 2$ slack variables for each node of indices 2 to n. A path of length l is represented by a vertex x where the positive slack variables are the variables for the nodes

which are not visited by the path. Then $\|x\|_1 = l + (n-1-l) = n-1$. Two paths are adjacent if the union of their edges contains a unique cycle. The norm of the corresponding direction is at least $\sqrt{l'}$ where l' is the length of this cycle and at most $\sqrt{2l'}$ where we consider the l' possibly affected slack variables. Therefore $\mu \geq \sqrt{3}$ and $\nu \leq \sqrt{2n}$. ∎

4. Monotone paths on Transportation polytopes.

Exponentially-long simplex paths can be found even for very simple linear programs given by network flow problems using Dantzig's pivot rule [40]. Nevertheless, Orlin showed that for certain pivot rules, the network Simplex method runs in a polynomial number of pivots [27]. Here we try to look at the special case of transportation polytopes and improve the bound.

In the paper [5], Borgwardt, De Loera and Finhold proved that the undirected diameter of $m \times n$ transportation polytopes is upper bounded by the Hirsch bound $m + n - 1$. In this section we study the monotone diameter of this polytope. From any degenerate transportation we can derive a non-degenerate transportation polytope with greater or equal monotone diameter by perturbing the original polytope. We will therefore assume non-degeneracy in this section. Recall that for a non-degenerate transportation polytope P, $x \in P$ is a vertex if and only if its support forms a spanning tree on the bipartite graph $K_{m,n}$ given by the m supply nodes and the n demand nodes (see references in [5]). For a vertex x we will write $s \sim d$ when supply node s and demand node d are adjacent in the support graph of x.

LEMMA 4.1. *Let x^* be the optimum of a $n \times m$ transportation polytope for a given linear functional c. Denote by $c_{v,w}$ the cost of the edge between vertex v and w. Let $s_1, s_2, \dots, s_k$ be $k \geq 2$ supply nodes and $d_1, d_2, \dots, d_k$ demand nodes. If $s_1 \sim d_1, s_2 \sim d_2, \dots, s_k \sim d_k$ in x^* then $c_{s_1,d_1} - c_{d_1,s_2} + c_{s_2,d_2} - c_{d_2,s_3} + \dots + c_{s_k,d_k} - c_{d_k,s_1} < 0$.*

Therefore, an edge between two vertices of the transportation polytope following the cycle $s_1 d_1 s_2 d_2 \dots s_k d_k$ is an improving edge for the linear functional.

Proof. Let s and d be respectively a supply and demand node which are not adjacent in x^*. Let $s = x^0, x^1, x^2, \dots, x^l = d$ be the path from s to d in x^*. By optimality of x^*, entering the edge (s,d) into the spanning tree associated to x^* will increase the cost function. In other words, the reduced cost of the variable (s,d) is positive i.e., $\tilde{C}_{s,d} := c_{s,d} - c_{x^0,x^1} + c_{x^1 x^2} - \dots + c_{x^{l-2},x^{l-1}} - c_{x^{l-1},x^l} > 0$, which gives us an inequality on the alternating cycle $s = x^0, x^1, x^2, \dots, x^l = d$.

We will add k inequalities of this type to obtain the desired inequality. More precisely, we will add the inequality resulting from the cycle given by adding the edge (s_2, d_1) to x^*, the cycle given by the edge (s_3, d_2), etc... and the cycle given by (s_1, d_k). We prove by induction on k that in the resulting sum $\tilde{C}_{s_2,d_1} + \tilde{C}_{s_3,d_2} + \dots + \tilde{C}_{s_1,d_k}$, terms cancel out to leave out $-(c_{s_1,d_1} - c_{d_1,s_2} + c_{s_2,d_2} - c_{d_2,s_3} + \dots + c_{s_k,d_k} - c_{d_k,s_1})$, which will then be positive.

Denote by T the smallest subtree of the support spanning tree of x^* containing the edges $(s_1,d_1), (s_2,d_2), \dots, (s_k,d_k)$. Without loss of generality, assume (s_1,d_1) is a leaf in T. We are going to merge together $\tilde{C}_{s_2,d_1}$ and $\tilde{C}_{s_1,d_k}$. The term $-c_{s_1,d_1}$ appears exactly once in their sum, say in $\tilde{C}_{s_1,d_k}$. We can therefore write the two paths in x^* from d_1 to s_2 and s_1 to d_k by $d_1 v^1 v^2 \dots v^l p^1 p^2 \dots p^{r-1} p^r = s_2$ and $s_1 d_1 v^1 v^2 \dots v^l q^1 q^2 \dots q^{t-1} q^t = d_k$ where $p^1 \neq q^1$. Note that the path in x^* from d_k to s_2 is exactly $q^t q^{t-1} \dots q^1 v^l p^1 p^2 \dots p^r$. Then the terms from the path $d_1 v^1 v^2 \dots v^l$ cancel to give $\tilde{C}_{s_1,q^t} + \tilde{C}_{d_1,p^r} = c_{s_2,d_1} + c_{s_1,d_k} - c_{s_1,d_1} - c_{s_2,d_k} + \tilde{C}_{s_2,d_k}$.

If $k = 2$, the above calculations directly give the desired result $\tilde{C}_{s_2 \sim d_1} + \tilde{C}_{s_1 \sim d_2} =$

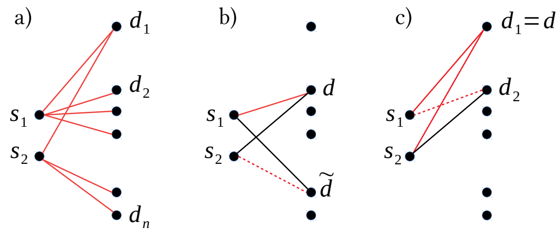

FIG. 7. *Illustration of the choice of entering variable in dashed lines when D_1 and D_2 non empty. Edges belonging to the optimum tree a) are in red.*

$c_{s_2,d_1} + c_{s_1,d_2} - c_{s_1,d_1} - c_{s_2,d_2}$. Otherwise, we use the induction on $\tilde{C}_{s_3 \sim d_2} + \tilde{C}_{s_4 \sim d_3} + \ldots + \tilde{C}_{s_2 \sim d_k}$ and the result follows. ∎

We now consider the case of a $2 \times n$ transportation polytope. We denote the supply and demand nodes respectively by s_1, s_2 and $d_1, \ldots, d_n$. Consider a vertex of the $2 \times n$ transportation polytope. Assuming that the transportation polytope is non-degenerate, we can partition the demand nodes in the following way:

- the set D_1 of demand nodes that are leaves adjacent to supply node s_1 only.
- the set D_2 of demand nodes that are leaves adjacent to supply node s_2 only.
- the last demand node adjacent to s_1 and s_2.

Proof of Theorem 1.10. We will show that from any vertex we can get to the optimum x^* in at most n steps using only edges of the type given by Lemma 4.1.

Without loss of generality, assume d_1 is adjacent to the two supply nodes in x^*, $D_1 = \{2, \ldots, k\}$ and $D_1 = \{k+1, \ldots, n\}$. We work by induction on $n \geq 1$. The result is true for $n = 1$ and the monotone diameter is even $0 = n - 1$ so now assume $n > 1$. Let x be the initial vertex of the transportation polytope. If any node $d \in D_1$ is a leaf incident to s_1 in x, likewise in x^*, we may remove this node and set the supply of s_1 to $S - D$ where S and D are respectively the supply at s_1, and the demand at d. The new problem is non-degenerate with $n - 1$ demand nodes so the induction gives the desired result. The result similarly holds if a node in D_2 is a leaf adjacent to supply node 2.

We therefore assume that all nodes in D_1 are adjacent to supply node 2 and all nodes in D_1 are adjacent to supply node 1 in x. Let d the demand node adjacent to both supply nodes in x.

Case 1: $d \neq d_1$
We are in fact going to prove that only $n-1$ steps are necessary to get to the optimum.

If D_1 and D_2 are not empty (see Figure 7b), without loss of generality, assume $d \in D_1$ and let $\tilde{d} \in D_2$. We make the edge $(s_2, \tilde{d})$ enter the basis. The corresponding cycle in x is $s_2 d s_1 \tilde{d}$ with $(s_2 \tilde{d})$ and (s_1, d) being two edges present in the optimum x^*. By Lemma 4.1, this pivot reduces the cost function. Denote by x^2 the resulting vertex. The demand node of the edge which has been deleted, either (s_2, d) or $(s_1, \tilde{d})$ is now a leaf in x^2 adjacent to the same supply node as in x^*. Similarly to above, we can delete this demand node and we get the result by induction.

Otherwise, without loss of generality we assume D_2 empty and $D_1 = \{2, \ldots, n\}$ (see Figure 8). But s_2 is a leaf adjacent to d_1 in x^* so the demand at d_1 is greater to the supply at s_2. Then, in an admissible tree, d_1 cannot be a leaf adjacent to s_2. Since $d \neq d_1$, d_1 is a leaf and it has to be adjacent to s_1 in x. We make the variable

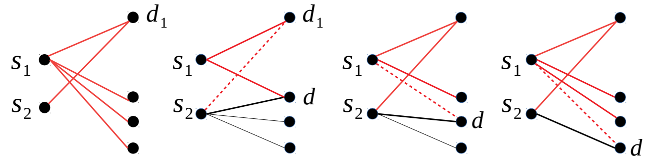

FIG. 8. *Illustration of the choice of entering variable in dashed lines when D_2 null. Edges belonging to the optimum tree on the left are in red.*

(s_2, d_1) enter the basis. The corresponding cycle is $s_2 d s_1 d_1$ and (s_1, d) and (s_2, d) are present edges in the optimum x^*. By Lemma 4.1 this pivot is increasing. Denote by x^2 the new spanning tree. The potential leaving variables are only (s_1, d_1) and (s_2, d), but it cannot be (s_1, d_1). Otherwise d_1 would be a leaf adjacent to s_2 in x^2. Therefore, (s_2, d) has been deleted and d is now a leaf adjacent to the correct supply node in x^2. Thus, we can delete the demand leaf d.

In x^2, d_1 is now adjacent to both supply nodes and all other demand nodes are adjacent to s_2. We enter the variable (s_1, d_2) into the basis. The corresponding cycle $s_1 d_1 s_2 d_2$ is improving since (s_1, d_2) and (s_2, d_1) are in x^*. Similarly to above, (s_1, d_1) cannot be the leaving variable, otherwise d_1 would become a leaf adjacent to s_2. Therefore, in the new spanning tree x^2, d_2 is a leaf adjacent to the correct supply node so we can delete it.

Note that in all pivot steps considered here we deleted a demand node. In the new spanning tree, either d_1 is a leaf or D_1 or D_2 are null which are the cases we handled. The induction therefore holds and we can get to $n' = 1$ in at most $n - 1$ steps. For $n' = 1$ there is only one spanning tree which is the optimum.

Case 2: $d = d_1$
We have already considered the case where D_1 or D_2 are empty. Now assume this is not the case. Therefore $d_2 \in D_1$ and d_2 is a leaf adjacent to s_2 in x (see Figure 7c).

We make the edge (s_1, d_2) enter the basis. The corresponding cycle is $s_1 d_1 s_2 d_2$. This is an improving cycle according to Lemma 4.1 given that edges (s_1, d_2) and (s_2, d_1) are present in x^*. Denote by x^2 the new vertex of the polytope. Either edge (s_1, d_1) or (s_2, d_2) has been removed. If (s_2, d_2) was removed, d_2 is a leaf in x^2 adjacent to s in x^2, likewise in x^*. Removing node d_2 therefore gives the result by induction. Otherwise, (s_1, d_1) has been removed so in x^2, the demand node adjacent to both supply nodes is now $d_2 \neq d_1$ and we use case 1.

We proved that the monotone diameter is $\leq n$. The bound n can be attained potentially if there exists at least one vertex with $d = d_1$ and D_1, D_2 non empty. This can only happen if $n \geq 3$, otherwise the monotone diameter is $n - 1$. ∎

CONJECTURE 4.2. *The monotone diameter of $m \times n$ transportation polytopes is linear in m and n.*

Acknowledgments. The authors are grateful for the useful comments of the referees and review team of SIAM Journal of Discrete Mathematics. We are also grateful to Amitabh Basu, Laurent Poirrier, Francisco Santos, Laura Sanità, Sean Kafer and Xiaotie Chen for several remarks and comments that were useful to the creation of this manuscript. We are particularly grateful to Dr. Tomonari Kitahara who found a small mistake in our original calculations. The first and second author were partially supported by NSF grant DMS-1818969. The first author was also supported by École Polytechnique.

REFERENCES

[1] I. Adler, C. Papadimitriou, and A. Rubinstein, *On simplex pivoting rules and complexity theory*, in Integer Programming and Combinatorial Optimization 2014, IPCO'14, Springer International Publishing, Cham, 2014, pp. 13–24.

[2] R. E. Behrend, *Fractional perfect b-matching polytopes I: General theory*, Linear Algebra and its Applications, 439 (2013), pp. 3822–3858.

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

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

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

[6] J. A. De Loera, *New insights into the complexity and geometry of linear optimization*, Optima, 87 (2011), pp. 1–11, http://www.mathopt.org/Optima-Issues/optima87.pdf.

[7] J. A. De Loera, R. Hemmecke, and M. Köppe, *Algebraic and geometric ideas in the theory of discrete optimization*, vol. 14 of MOS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2013.

[8] J. A. De Loera, S. Kafer, and L. Sanità, *Pivot rules for circuit-augmentation algorithms in linear optimization*, (2019). preprint available at arXiv:1909.12863.

[9] A. Del Pia and C. Michini, *Short simplex paths in lattice polytopes*, arXiv e-prints, (2020), arXiv:1912.05712v2, p. arXiv:1912.05712v2, https://arxiv.org/abs/1912.05712v2.

[10] M. Develin, *LP-orientations of cubes and crosspolytopes*, Advances in Geometry, 4 (2004), pp. 459–468.

[11] J. Edmonds, *Maximum matching and a polyhedron with 0,1-vertices*, Journal of Research of the National Bureau of Standards, 69B (1965), pp. 125–130.

[12] J. Fearnley and R. Savani, *The complexity of the simplex method*, in Symposium on Theory of Computing 2015, STOC'15, ACM, New York, 2015, pp. 201–208.

[13] M. J. Gallagher and W. D. Morris, *A proof of the strict monotone 5-step conjecture*, arXiv e-prints, (2018), arXiv:1806.03403, p. arXiv:1806.03403, https://arxiv.org/abs/1806.03403.

[14] D. Goldfarb and W. Y. Sit, *Worst case behavior of the steepest edge simplex method*, Discrete Applied Mathematics, 1 (1979), pp. 277 – 285, https://doi.org/https://doi.org/10.1016/0166-218X(79)90004-0, http://www.sciencedirect.com/science/article/pii/0166218X79900040.

[15] P. Gritzmann and B. Sturmfels, *Minkowski addition of polytopes: computation complexity and applications to Gröbner bases*, SIAM Journal on Discrete Mathematics, 6 (1993), pp. 246–269.

[16] R. Jeroslow, *The simplex algorithm with the pivot rule of maximizing criterion improvement*, Discrete Mathematics, 4 (1973), pp. 367 – 377, https://doi.org/https://doi.org/10.1016/0012-365X(73)90171-4, http://www.sciencedirect.com/science/article/pii/0012365X73901714.

[17] G. Kalai, *Upper bounds for the diameter and height of graphs of convex polyhedra*, Discrete Comput. Geom., 8 (1992), pp. 363–372, https://doi.org/10.1007/BF02293053, https://doi.org/10.1007/BF02293053.

[18] T. Kitahara, T. Matsui, and S. Mizuno, *On the number of solutions generated by Dantzig's simplex method for LP with bounded variables*, Pacific journal of optimization, 8 (2012), pp. 447–455.

[19] T. Kitahara and S. Mizuno, *A bound for the number of different basic solutions generated by the simplex method*, Mathematical Programming, 137 (2013), pp. 579–586, https://doi.org/10.1007/s10107-011-0482-y, https://doi.org/10.1007/s10107-011-0482-y.

[20] V. Klee, *A class of linear programming problems requiring a large number of iterations*, Numerische Mathematik, 7 (1965), pp. 313–321, https://doi.org/10.1007/BF01436525, https://doi.org/10.1007/BF01436525.

[21] V. Klee, *Heights of convex polytopes*, Journal of Mathematical Analysis and Applications, 11 (1965), pp. 176 – 190, https://doi.org/https://doi.org/10.1016/0022-247X(65)90077-6, http://www.sciencedirect.com/science/article/pii/0022247X65900776.

[22] V. Klee and G. J. Minty, *How good is the simplex algorithm?*, in Inequalities–III, Proceedings Third Symposium, 1969, Academic Press, New York, 1972, pp. 159–175.

[23] T. Kuno, Y. Sano, and T. Tsuruda, *Computing Kitahara–Mizuno's bound on the number of basic feasible solutions generated with the simplex algorithm*, Optimization Letters, 12 (2018), pp. 933–943, https://doi.org/10.1007/s11590-018-1276-4, https://doi.org/10.1007/

s11590-018-1276-4.

[24] J. Mihalisin and V. Klee, *Convex and linear orientations of polytopal graphs*, Discrete & Computational Geometry, 24 (2000), pp. 421–436, https://doi.org/10.1007/s004540010046, https://doi.org/10.1007/s004540010046.

[25] D. J. Nadeff, *The Hirsch conjecture is true for (0, 1)-polytopes*, Mathematical Programming, 45 (1989), pp. 109–110.

[26] S. Onn and R. Pinchasi, *A note on the minimum number of edge-directions of a convex polytope*, Journal on Combinatorial Theory, Series A, 107 (2004), pp. 147–151, https://doi.org/10.1016/j.jcta.2004.03.008, https://doi.org/10.1016/j.jcta.2004.03.008.

[27] J. B. Orlin, *A polynomial time primal network simplex algorithm for minimum cost flows*, Mathematical Programming, 77 (1997), pp. 109–129, https://doi.org/10.1007/BF02614365.

[28] I. Pak, *Four questions on Birkhoff polytope*, Annals of Combinatorics, 4 (2000), pp. 83–90.

[29] J. Pfeifle and G. M. Ziegler, *On the monotone upper bound problem*, Experimental Mathematics, 13 (2004), pp. 1–11, https://doi.org/10.1080/10586458.2004.10504519, https://doi.org/10.1080/10586458.2004.10504519, https://arxiv.org/abs/https://doi.org/10.1080/10586458.2004.10504519.

[30] F. J. Rispoli, *The monotonic diameter of the perfect matching and shortest path polytopes*, Operations Research Letters, 12 (1992), pp. 23–27.

[31] F. J. Rispoli, *The monotonic diameter of the perfect 2-matching polytope*, SIAM Journal on Optimization, 4 (1994), pp. 455–460.

[32] F. J. Rispoli, *The monotonic diameter of traveling salesman polytopes*, Operations Research Letters, 22 (1998), pp. 69–73.

[33] F. J. Rispoli and S. Cosares, *A bound of 4 for the diameter of the symmetric traveling salesman polytope*, SIAM Journal on Discrete Mathematics, 11 (1998), pp. 373–380.

[34] F. Santos, *Recent progress on the combinatorial diameter of polyhedra and simplicial complexes*, in Symposium on Computational Geometry 2013, SoCG'13, 2013, pp. 165–166, https://doi.org/10.1145/2462356.2462408, https://doi.org/10.1145/2462356.2462408.

[35] M. Tano, R. Miyashiro, and T. Kitahara, *Steepest-edge rule and its number of simplex iterations for a nondegenerate lp*, Operations Research Letters, 47 (2019), pp. 151 – 156, https://doi.org/https://doi.org/10.1016/j.orl.2019.02.003, http://www.sciencedirect.com/science/article/pii/S0167637718304346.

[36] T. Terlaky and S. Zhang, *Pivot rules for linear programming: A survey on recent theoretical developments*, Annals OR, 46-47 (1993), pp. 203–233, https://doi.org/10.1007/BF02096264, https://doi.org/10.1007/BF02096264.

[37] M. J. Todd, *The monotonic bounded Hirsch conjecture is false for dimension at least 4*, Mathematics of Operations Research, 5 (1980), pp. 599–601, https://doi.org/10.1287/moor.5.4.599.

[38] M. J. Todd, *The many facets of linear programming*, Mathematical Programming Series B, 91 (2002), pp. 417–436, https://doi.org/10.1007/s101070100261, http://dx.doi.org/10.1007/s101070100261.

[39] D. M. Topkis, *Adjacency on polymatroids*, Mathematical Programming, 30 (1984), pp. 229–237, https://doi.org/10.1007/BF02591887.

[40] N. Zadeh, *A bad network problem for the simplex method and other minimum cost flow algorithms*, Mathematical Programming, 5 (1973), pp. 255–266, https://doi.org/10.1007/BF01580132.

[41] G. M. Ziegler, *Lectures on polytopes*, Springer-Verlag New York, NY, USA, 1995.