Pivot Rules for Circuit-Augmentation Algorithms in Linear Optimization

Jesús A. De Loera*, Sean Kafer⋆, Laura Sanità⋆

* University of California, Davis, USA. deloera@math.ucdavis.edu
⋆ Department of Combinatorics and Optimization, University of Waterloo, Canada.
{skafer, lsanita}@uwaterloo.ca

Abstract

Circuit-augmentation algorithms are generalizations of the Simplex method, where in each step one is allowed to move along a fixed set of directions, called circuits, that is a superset of the edges of a polytope. We show that in the circuit-augmentation framework the greatest-improvement and Dantzig pivot rules are NP-hard, already for 0/1-LPs. Differently, the steepest-descent pivot rule can be carried out in polynomial time in the 0/1 setting, and the number of circuit augmentations required to reach an optimal solution according to this rule is strongly-polynomial for 0/1-LPs.

The number of circuit augmentations has been of interest as a proxy for the number of steps in the Simplex method, and the circuit-diameter of polyhedra has been studied as a lower bound to the combinatorial diameter of polyhedra. Extending prior results, we show that for any polyhedron P the circuit-diameter is bounded by a polynomial in the input bit-size of P. This is in contrast with the best bounds for the combinatorial diameter of polyhedra.

Interestingly, we show that the circuit-augmentation framework can be exploited to make novel conclusions about the classical Simplex method itself: In particular, as a byproduct of our circuit results, we prove that (i) computing the shortest (monotone) path to an optimal solution on the 1-skeleton of a polytope is NP-hard, and hard to approximate within a factor better than 2, and (ii) for 0/1 polytopes, a monotone path of strongly-polynomial length can be constructed using steepest improving edges.

1 Introduction

Linear Programming (LP) is one of the most powerful mathematical tools for tackling optimization problems. While various algorithms have been proposed for solving LPs in the past decades, probably the most popular method remains the *Simplex method*, introduced by G. B. Dantzig in the 1940's. In what follows we assume the input LP is given in the general minimization format

$$\min\{\ \boldsymbol{c}^\mathsf{T}\boldsymbol{x}\ :\ A\boldsymbol{x} = \boldsymbol{b},\ B\boldsymbol{x} \leq \boldsymbol{d},\ \boldsymbol{x} \in \mathbb{R}^n\ \}. \tag{1}$$

Here the objective function is minimized over the polyhedron $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b},\ B\boldsymbol{x} \leq \boldsymbol{d}\}$ of feasible solutions. The 1-skeleton of the polyhedron $\mathcal{P}$ is the graph given by the 0-dimensional faces (vertices) and 1-dimensional faces (edges) of $\mathcal{P}$. Given an objective function to minimize on $\mathcal{P}$, a path on the 1-skeleton is called monotone if every vertex on the path has an associated objective function value larger than its subsequent one. The Simplex method starts with an initial feasible extreme point solution of the LP, and in each step, it moves along an improving edge to an adjacent

extreme point, until an optimal solution is found or unboundedness is detected. Effectively there is a *monotone* path from an initial extreme point solution of an LP to an optimal one.

Despite having been used and studied for more than 70 years, it is still unknown whether there is a rule for selecting an improving neighbor extreme point, called a *pivot rule*, that guarantees a polynomial upper bound on the number of steps performed by the Simplex method. In fact, a longstanding open question in the theory of optimization [43] states: *is there a version of the Simplex method, i.e., a choice of pivot rule, that can solve LPs in strongly-polynomial time?*

We contribute the following new results about the behavior of the Simplex method. First, we give a partial answer to the above question for linear programs whose feasible region $\mathcal{P}$ is a 0/1 polytope (in the rest of the paper we will often refer to them as *0/1-LPs*). We are able to exploit their *steepest edges*: At a given vertex $\boldsymbol{x}$, a steepest edge is given by the edge minimizing $\frac{\boldsymbol{c}^\mathsf{T}\boldsymbol{g}}{\|\boldsymbol{g}\|_1}$ among all the edges $\boldsymbol{g}$ incident at $\boldsymbol{x}$. A steepest-descent pivot rule asks to iteratively select a steepest edge as the improving direction.

Theorem 1. *Given an LP in the general form* (1) *whose feasible region $\mathcal{P}$ is a 0/1 polytope, then*

  (i) *One can construct a monotone path between any vertex and the optimum via steepest edges, whose length is strongly-polynomial in the input size of the LP.*

  (ii) *When the feasible region $\mathcal{P}$ is a non-degenerate 0/1 polytope, the Simplex method with a steepest-descent pivot rule reaches an optimal solution in strongly-polynomial time.*

In addition, we consider the following legitimate question: *can one hope to find a pivot rule that makes the Simplex method use the shortest monotone path?* We here show a negative answer.

Theorem 2. *Given an LP and an initial feasible solution, finding the shortest (monotone) path to an optimal solution is NP-hard. Furthermore, unless P=NP, it is hard to approximate within a factor strictly better than two.*

The above theorem implies that for *any* efficiently-computable pivoting rule, an edge-augmentation algorithm (like the Simplex method) cannot be guaranteed to reach an optimal solution via a minimum number of augmentations (i.e., non-degenerate pivots), unless P=NP. We remark that these hardness results require the use of degeneracy. The same theorem for non-degenerate polytopes would be a stronger result, and is left as an open question.

The methods we used to prove these two theorems are of independent interest. We arrived to the results above while thinking about a much more general family of algorithms that includes the Simplex method. *Circuit-augmentation* algorithms are extensions of the Simplex method where we have many more choices of improving directions available at each step—more than just the edges of the polyhedron. Our results are valid in this more general family of circuit-augmentation algorithms which we now introduce to the reader.

Given a polyhedron, its *circuits* are all potential edges that can arise by translating some of its facets. Circuits are important not just in the development of linear optimization [7, 38], but they appear very naturally in other areas of application where polyhedra need to be decomposed [35]. Formally:

Definition 1. *Given a polyhedron of the form $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b},\, B\boldsymbol{x} \le \boldsymbol{d}\}$, a non-zero vector $\boldsymbol{g} \in \mathbb{R}^n$ is a circuit if*

  (i) $\boldsymbol{g} \in \ker(A)$, *and*

  (ii) $B\boldsymbol{g}$ *is support-minimal in the collection* $\{B\boldsymbol{y} : \boldsymbol{y} \in \ker(A), \boldsymbol{y} \ne 0\}$.

Here $\ker(A)$ denotes the kernel of the matrix A. As we will see later, from a geometric perspective the circuits can capture the extreme rays of a cone containing any improving direction, and allow one to represent any augmentation as a sum of at most n circuit-augmentations.

To represent the circuits with a finite set, we can normalize them in various ways. Following [11, 12, 14, 18], we denote by $\mathcal{C}(A, B)$ the (finite) set of circuits with co-prime integer components.

Given an initial feasible point of an LP, a circuit-augmentation algorithm at each iteration moves maximally along an improving circuit-direction until an optimal solution is found (or unboundedness is detected). Circuits and circuit-augmentation algorithms have appeared in several papers and books on linear and integer optimization (see [8, 9, 10, 11, 12, 13, 14, 22, 25, 27, 36, 38, 42] and the many references therein). In particular, the authors of [14] considered linear programs in equality form and analyzed in detail three circuit-pivot rules that guarantee notable bounds on the number of steps performed by a circuit-augmentation algorithm to reach an optimal solution.

Given a feasible point $\boldsymbol{x} \in \mathcal{P}$, the proposed rules are as follows:

(i) *Greatest-improvement circuit-pivot rule:* select a circuit $\boldsymbol{g} \in \mathcal{C}(A, B)$ that maximizes the objective improvement $-\boldsymbol{c}^{\mathsf{T}}(\alpha \boldsymbol{g})$, among all circuits $\boldsymbol{g}$ and $\alpha \in \mathbb{R}_{>0}$ such that $\boldsymbol{x} + \alpha \boldsymbol{g} \in \mathcal{P}$.

(ii) *Dantzig circuit-pivot rule:* select a circuit $\boldsymbol{g} \in \mathcal{C}(A, B)$ that maximizes $-\boldsymbol{c}^{\mathsf{T}}\boldsymbol{g}$, among all circuits $\boldsymbol{g}$ such that $\boldsymbol{x} + \varepsilon \boldsymbol{g} \in \mathcal{P}$ for some $\varepsilon > 0$.

(iii) *Steepest-descent circuit-pivot rule:* select a circuit $\boldsymbol{g} \in \mathcal{C}(A, B)$ that maximizes $-\frac{\boldsymbol{c}^{\mathsf{T}}\boldsymbol{g}}{\|\boldsymbol{g}\|_1}$, among all circuits $\boldsymbol{g}$ such that $\boldsymbol{x} + \varepsilon \boldsymbol{g} \in \mathcal{P}$ for some $\varepsilon > 0$.

Note that these circuit-pivot rules are direct extensions to three famous pivot rules proposed for the Simplex method. Unfortunately, the Simplex method can require an exponential number of edge steps before reaching an optimal solution [23, 26, 31]. When all circuits are considered as possible directions to move, much better bounds can be given. Most notably, the greatest-improvement circuit-pivot rule guarantees a *polynomial* bound on the number of steps performed by a circuit-augmentation algorithm on LPs in equality form (see [14, 25, 36] and references therein). However, the set of circuits in general can have an exponential cardinality, and therefore selecting the best circuit according to the previously mentioned rules is not an easy optimization problem. Indeed, the central questions of this paper are the following:

- *How hard is it to carry out these three circuit-pivot rules over the exponentially-large set of circuits?*

- *Can we exploit (approximate) solutions to these circuit-pivot rules to design (strongly-) polynomial time augmentation algorithms?*

We now describe our results and how they lead to Theorem 1 and Theorem 2.

Hardness of circuit-augmentation. First we settle the computational complexity of the circuit-pivot rules (i) and (ii).

Theorem 3. *The greatest-improvement and Dantzig circuit-pivot rules are NP-hard.*

We prove this theorem by showing that computing a circuit, according to both the greatest-improvement and the Dantzig circuit-pivot rule, is already hard to do when $\mathcal{P}$ is a 0/1 polytope. In particular, we focus on the case when $\mathcal{P}$ is the matching polytope of a bipartite graph. We first characterize the circuits of the more general *fractional* matching polytope, i.e., the polytope given

3

by the standard LP-relaxation for the matching problem on general graphs, in Section 3.1. This builds on the known graphical characterization of adjacency given in [39, 6]. Then, we construct a reduction from the NP-hard Hamiltonian path problem in Section 3.2. The heart of the reduction yields the following interesting corollary.

Corollary 1. *Given a feasible extreme point solution of the bipartite matching polytope and an objective function, it is NP-hard to decide whether there is a neighbor extreme point that is optimal.*

With the above corollary, the hardness results stated in Theorem 2 can be easily derived. Even more, combining this corollary with the characterization of circuits for the fractional matching polytope, we can show that the hardness results of Theorem 2 hold more generally for *circuit*-paths, i.e., paths constructed by a circuit-augmentation algorithm (see Corollary 4).

Approximating pivoting rules. We next make a very useful observation: Any polynomial-time γ-approximation algorithm for the circuit-pivot rule optimization problems yields an increase of at most a γ-factor on the running time of the corresponding circuit-augmentation algorithm – this follows from an extension of the analysis given by [14]. This simple observation turns out to be quite powerful in combination with the greatest-improvement circuit-pivot rule, and it plays a key role in our subsequent results. We therefore formally state its main implication in the next lemma.

Lemma 1. *Consider an LP in the general form* (1). *Denote by δ the maximum absolute value of the determinant of any $n \times n$ submatrix of $\binom{A}{B}$. Let $\boldsymbol{x}_0$ be an initial feasible solution, and let $\gamma \geq 1$. Using a γ-approximate greatest-improvement circuit-pivot rule, we can reach an optimal solution $\boldsymbol{x}_{\min}$ of* (1) *with $O\big(n\gamma \log \big(\delta\, \boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min})\big)\big)$ augmentations.*

We exploit Lemma 1 in two ways. First, as an easy corollary, we get a polynomial bound on the *circuit-diameter* of any rational polyhedron, where the circuit-diameter is the maximum length of a shortest circuit-path between any two vertices. To the best of our knowledge, this has not been observed before. We recall that, as usual, the encoding length of a rational number $\frac{p}{q}$ is defined as $\lceil \log(p+1) \rceil + \lceil \log(q+1) \rceil + 1$.

Corollary 2. *There exists a polynomial function $f(m, \alpha)$ that bounds above the circuit-diameter of any rational polyhedron $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}, B\boldsymbol{x} \leq \boldsymbol{d}\}$ with m row constraints and maximum encoding length among the coefficients in its description equal to α.*

Second, we prove that for an extreme point $\boldsymbol{x}$ of a 0/1-polytope $\mathcal{P}$, a polynomial approximation of a greatest-improvement circuit-pivot rule can be obtained by simply computing a steepest edge incident at $\boldsymbol{x}$.

Theorem 4. *Let $\boldsymbol{x}$ be an extreme point of a 0/1-LP with n variables. An augmentation along a steepest edge yields an n-approximation of an augmentation along a greatest-improvement circuit.*

We stress that the above theorem is not true for arbitrary (non 0/1) LPs, since in general augmenting along the edges incident at a vertex can be arbitrarily worse compared to augmenting along circuits, even in dimension two (see Figure 6 in Section 4). For 0/1-LPs instead, the above theorem can be combined with Lemma 1 and the seminal result of Frank & Tardos [19], to yield Theorem 1. Even more, we can show that the result of Theorem 1 holds more generally for *circuit*-augmentation algorithms (see Theorem 8).

Finally, as a byproduct of our result we also show (i) that a steepest-descent circuit-pivot rule for 0/1-LPs reduces to computing steepest edges, and this can be done in polynomial time (Lemma 5);

(ii) a new bound on the number of steps performed by a circuit-augmentation algorithm that uses a steepest-descent circuit-pivot rule, for general (non 0/1) bounded LPs (Theorem 7). In fact, the authors of [14], and later [12], gave bounds on the number of steps, which depend on the size of the set of circuits $\mathcal{C}(A, B)$ and the number of different values the objective function takes on that set. However, such bounds are a bit opaque, and often difficult to analyze. In Theorem 7 we get another type of bound of independent application. Our bound is more comparable to that obtained with the greatest-improvement circuit-pivot rule, since, unlike all previous bounds for steepest-descent, it depends more explicitly on the input description of the LP. Furthermore, it is valid for bounded LPs in general form, and not just LPs in equality form.

1.1 Relevant prior work and comparison with our results

We now give some extra context to our results, mentioning several important papers.

Pivoting rules. Several pivot rules have been proposed for the Simplex method, but since 1972 researchers have debunked most of them as candidates for good polynomial behavior. As of today, all popular pivoting rules are known to require an exponential number of steps to solve some concrete "twisted" linear programs (see [3, 4, 20, 23, 26, 31, 24, 43, 44, 46] and the many references there). Some ways to pivot over circuits that we do not consider here are presented e.g., in [34, 40, 28, 12]. In particular, the circuit-pivot rule in [34] also guarantees geometric convergence. We stress that in this article we consider only "pure" pivoting rules that cannot be adjusted along the way (just as the Simplex method requires). If one is allowed to adapt the pivot step along the way, the complexity may be different, but at the same time the algorithm would depart more substantially from the Simplex method. Similarly, we restrict ourselves to pivot rules that, while valid for circuits, relate to well-known Simplex pivot methods. Kitahara and Mizuno [29] provided a general upper bound for the number of different basic feasible solutions generated by the Simplex method for linear programming problems (LP) having optimal solutions. Their bound is polynomial on the number of constraints, the number of variables, and the ratio between two parameters: δ, the minimum value of a positive component among basic feasible solutions, and γ, the maximum value among all positive components. Unfortunately, their bound is not polynomial in the encoding length, and in fact, while γ can be efficiently computed, it has been shown that δ is NP-hard to determine (see [33]). Later in [30] the authors showed that the bounds are good for some special LP problems including those on 0/1 polytopes, those with totally unimodular matrices, and the Markov decision problems.

0/1 polytopes. Understanding monotone paths of 0/1 polytopes is particularly important in discrete optimization. Thus it is relevant to understand the performance of augmentation algorithms, and in particular of the Simplex method, in that family of polytopes. The authors of [32] showed that the monotone paths of some pivot rules are unfortunately exponentially long even in 0/1 polytopes. Later [29, 30] proved that the number of distinct basic feasible solutions generated by the Simplex method is strongly-polynomial for some 0/1-LPs of special format, namely, 0/1-LPs in standard equality form. For a comparison, our bound in Theorem 1 is weaker (though, still strongly-polynomial), but it is valid for *all* 0/1-LPs. An important remark is the following: The fact that paths of strongly-polynomial length on the 1-skeleton of 0/1 polytopes can be constructed from *any* augmentation oracle that outputs an improving edge-direction is already known (see [41], and [37] for a recent generalization to lattice polytopes). However, such arguments require some scaling/modifying of the objective function, and they are not realized with a "pure" pivoting rule.

In fact, the resulting path might not even be monotone with respect to the original objective function. In contrast, our result in Theorem 1 yields a monotone path of strongly-polynomial length that is realized by simply moving along steepest edges. In the context of the Simplex method, our Theorem 1 yields a strongly-polynomial bound on the number of distinct basic feasible solutions visited by the Simplex method, if implemented with a pivot rule that makes it move to an adjacent extreme point via a steepest edge. Note that our bound (as well as the results of [29, 30]) is not on the total number of *basis exchanges*, in presence of degeneracy. While computing a steepest edge can be done in polynomial time (as explained in the proof of Lemma 5), an interesting open problem is whether moving along steepest edges on 0/1-LPs can be realized efficiently via a "pure" pivot rule for the Simplex method in presence of degeneracy. In other words, can one guarantee that the number of different adjacent bases visited by the algorithm is strongly-polynomial (rather than the number of different extreme points, as we proved here)?

Hardness. Several authors arrived before us to the conclusion that pivot rules can be used to encode NP-hard problems [1, 16, 17]. For example, Fearnley and Savani [17] showed that it is PSPACE-complete to decide whether Dantzig's pivot rule ever chooses a specific variable to enter the basis. In contrast, Adler et al. [1] showed the simplex method with the shadow vertex pivot rule can decide this query in polynomial time. We stress that our hardness result in Theorem 2 does not follow from the hardness of computing the diameter of a polytope [21, 39]. In fact, the hardness results in [21] and [39] rely on the existence/non existence of vertices with a certain structure, and they do not provide a specific objective function to minimize over their polytopes.

We note that Theorem 2 and Corollary 1 can also be derived using the circulation polytope [5], instead of the bipartite matching polytope. In particular, the characterization of circuits for the circulation polytope becomes easier, since circuits can be associated with simple cycles. Hence the hardness results of the above oracles can be derived from the NP-hardness of finding a cycle that gives the biggest improvement in the cost, a result proven in [5]. However, our proof relies on the matching polytope for three reasons: First, bipartite matching is a well-studied combinatorial problem, it is quite interesting that the hardness resides already in such a simple and well-studied polyhedron. Second, if we had written our hardness result using the circulation polytope in [5], the presentation would not be any shorter than the proof written for the (integral) bipartite matching polytope. The extra material regarding the fractional matching polytope is only a couple of pages longer, and it does add new combinatorial properties to the literature. Third, we believe that the characterization of the circuits of the fractional matching polytope can be useful in determining the complexity of computing the circuit-diameter of a polytope. Interestingly, our results in Section 3.1 show that the circuits of the fractional matching polytope correspond to the actual edges of the polytope (i.e., the translation of any subset of the facets does not increase the initial set of edge-directions). However, we can construct examples showing that the circuit-diameter can be strictly smaller than the (combinatorial) diameter for the fractional matching polytope. As a consequence, the hardness result on the computation of its diameter given in [39] does not trivially extend to the circuit setting. We feel that the characterization of the circuits of the fractional matching polytope can still be exploited to resolve the complexity of the computation of its circuit-diameter, and hence can be useful in attacking the mentioned open question.

2 Preliminaries

We consider $\mathcal{P}$ to be a polyhedron of the form $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b},\ B\boldsymbol{x} \leq \boldsymbol{d}\}$ for integer matrices A and B of sizes $m_A \times n$ and $m_B \times n$ respectively, and integer vectors $\boldsymbol{b}$ and $\boldsymbol{d}$, and assume that we

wish to minimize a linear integral objective function $\mathbf{c}^{\mathsf{T}}\mathbf{x}$ over $\mathcal{P}$. We further assume that A has full row rank and that the rank of $\binom{A}{B}$ is n. For our purposes, we also assume $m_B \geq 1$ as otherwise $\mathcal{P}$ contains trivially only one point. As mentioned before, $\ker(A)$ denotes the kernel of the matrix A. For a matrix D and a subset T of row indices, we let D_T denote the submatrix of D given by the rows indexed by T. Furthermore, we let $\mathrm{rk}(D)$ denote its rank, and $\det(D)$ denote its determinant. For a vector $\mathbf{x}$, we let $\mathrm{supp}(\mathbf{x})$ be the support of the vector $\mathbf{x}$, and $\mathbf{x}(i)$ denote its i-th component.

Given a vertex $\bar{\mathbf{x}}$ of $\mathcal{P}$, we define the *feasible cone* at $\bar{\mathbf{x}}$ to be the set of all directions $\mathbf{z}$ such that $\bar{\mathbf{x}} + \varepsilon\mathbf{z} \in \mathcal{P}$ for some $\varepsilon > 0$. More formally, it is the set $\{\,\mathbf{z} \in \mathbb{R}^n : A\mathbf{z} = \mathbf{0}, B_{T(\bar{\mathbf{x}})}\mathbf{z} \leq \mathbf{0}\,\}$ where $T(\bar{\mathbf{x}})$ denotes the indices of the inequalities of $B\mathbf{x} \leq \mathbf{d}$ that are tight at $\bar{x}$. The extreme rays of the feasible cone at $\bar{\mathbf{x}}$ are the *edge-directions* at $\bar{\mathbf{x}}$.

A circuit-path is a finite sequence of feasible solutions $\mathbf{x}_1$, $\mathbf{x}_2$, $\ldots, \mathbf{x}_q$ satisfying $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i\mathbf{g}_i$, where $\mathbf{g}_i \in \mathcal{C}(A, B)$ and $\alpha_i \in \mathbb{R}_{>0}$ is such that $\mathbf{x}_i + \alpha_i\mathbf{g}_i \in \mathcal{P}$ but $\mathbf{x}_i + (\alpha_i + \varepsilon)\mathbf{g}_i \notin \mathcal{P}$ for all $\varepsilon > 0$ (i.e., the augmentation is maximal). Note that $\mathbf{x}_i$ is not necessarily a vertex of $\mathcal{P}$. A circuit-path is called monotone if each $\mathbf{g}_i$ satisfies $\mathbf{c}^{\mathsf{T}}\mathbf{g}_i < 0$ (i.e., it is an improving circuit).

A circuit-augmentation algorithm computes a monotone circuit-path starting at a given initial feasible solution, until an optimal solution is reached (or unboundedness is detected). The circuit $\mathbf{g}$ to use at each augmentation is usually chosen according to some *circuit-pivot rule*. As discussed before, in this paper we focus on three such rules, each of which gives rise to a corresponding optimization problem.

The optimization problem that arises when following the greatest-improvement circuit-pivot rule will be called $Great(\mathcal{P}, \mathbf{x}, \mathbf{c})$, and is as follows:

$$\max -\mathbf{c}^{\mathsf{T}}(\alpha\mathbf{g})$$
$$\text{s.t.}$$
$$\mathbf{g} \in \mathcal{C}(A, B),$$
$$\alpha > 0,$$
$$\mathbf{x} + \alpha\mathbf{g} \in \mathcal{P}.$$

The optimization problem that arises when following the Dantzig circuit-pivot rule will be called $Dan(\mathcal{P}, \mathbf{x}, \mathbf{c})$, and is as follows:

$$\max -\mathbf{c}^{\mathsf{T}}\mathbf{g}$$
$$\text{s.t.}$$
$$\mathbf{g} \in \mathcal{C}(A, B),$$
$$\mathbf{x} + \epsilon\mathbf{g} \in \mathcal{P} \quad \text{for some } \epsilon > 0.$$

The optimization problem that arises when following the steepest-descent circuit-pivot rule will be called $Steep(\mathcal{P}, \mathbf{x}, \mathbf{c})$, and is defined as follows:

$$\max -\frac{\mathbf{c}^{\mathsf{T}}\mathbf{g}}{\|\mathbf{g}\|_1}$$
$$\text{s.t.}$$
$$\mathbf{g} \in \mathcal{C}(A, B),$$
$$\mathbf{x} + \epsilon\mathbf{g} \in \mathcal{P} \quad \text{for some } \epsilon > 0.$$

A maximal augmentation given by an optimal solution to $Great(\mathcal{P}, \mathbf{x}, \mathbf{c})$ is called a greatest-improvement augmentation. A Dantzig augmentation and a steepest-descent augmentation are

7

defined similarly. In this work, we will only use maximal augmentations, and therefore will omit the word "maximal". We remark that, in the context of pivot rules for the Simplex method, the name 'steepest-descent' often refers to normalizations according to the 2-norm of a vector, rather than the 1-norm. Here we stick to this name as used previously in [14, 12], but at the end of the paper we note that this is not a major assumption.

3   Hardness of some Circuit-Pivot Rules

3.1   The Circuits of the Fractional Matching Polytope

Let G be a simple connected graph with nodes $V(G)$ and edges $E(G)$. We assume $|V(G)| \geq 3$. Given $v \in V(G)$, we let $\delta_G(v)$ denote the edges of $E(G)$ incident with v. We call a node $v \in V(G)$ a *leaf* if $|\delta_G(v)| = 1$, and let $L(G)$ denote the set of leaf nodes of G. Furthermore, for $X \subseteq E$ and $\boldsymbol{x} \in \mathbb{R}^{E(G)}$, we let $\boldsymbol{x}(X)$ denote $\sum_{e \in X} \boldsymbol{x}(e)$.

Let $\mathcal{P}_{\mathrm{FMAT}}(G)$ denote the fractional matching polytope of G, which is defined by the following (minimal) linear system:

$$\boldsymbol{x}\left(\delta_G(v)\right) \leq 1, \qquad\qquad \text{for all } v \in V(G) \setminus L(G). \tag{2}$$

$$\boldsymbol{x} \geq \boldsymbol{0}. \tag{3}$$

In this section, we fully characterize the circuits of $\mathcal{P}_{\mathrm{FMAT}}(G)$. We will prove that, if $\boldsymbol{x}$ is a circuit of $\mathcal{P}_{\mathrm{FMAT}}(G)$, then $\mathrm{supp}(\boldsymbol{x})$ induces a connected subgraph of G that has a very special structure: namely, it belongs to one of the five classes of graphs $(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5,)$ listed below.

(i) Let $\mathcal{E}_1$ denote the set of all subgraphs $F \subseteq G$ such that F is an even cycle.

(ii) Let $\mathcal{E}_2$ denote the set of all subgraphs $F \subseteq G$ such that F is an odd cycle.

(iii) Let $\mathcal{E}_3$ denote the set of all subgraphs $F \subseteq G$ such that F is a simple path.

(iv) Let $\mathcal{E}_4$ denote the set of all subgraphs $F \subseteq G$ such that F is a connected graph satisfying $F = C \cup P$, where C and P are an odd cycle and a non-empty simple path, respectively, that intersect only at an endpoint of P. (See Figure 1).
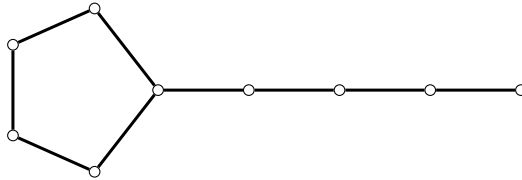

Figure 1: An Example of a subgraph belonging to $\mathcal{E}_4$

(v) Let $\mathcal{E}_5$ denote the set of all subgraphs $F \subseteq G$ such that F is a connected graph with $F = C_1 \cup P \cup C_2$, where C_1 and C_2 are odd cycles, and P is a (possibly empty) simple path satisfying the following: if P is non-empty, then C_1 and C_2 are node-disjoint and P intersects each C_i exactly at its endpoints (see Figure 2); if P is empty then C_1 and C_2 intersect only at one node v (see Figure 3).
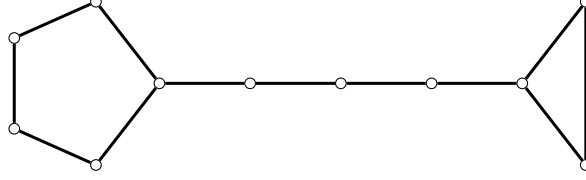
Figure 2: An Example of a subgraph belonging to $\mathcal{E}_5$ where P is non-empty.
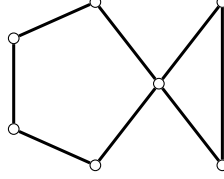

Figure 3: An Example of a subgraph belonging to $\mathcal{E}_5$ where P is empty.

We will associate a set of circuits to the subgraphs in the above families, by defining the following five sets of vectors. It is worth noticing that similar elementary moves appeared in [15] in applications of Gröbner bases in combinatorial optimization.

$$
\mathcal{C}_1 = \bigcup_{F \in \mathcal{E}_1} \left\{ \boldsymbol{g} \in \{-1,1\}^{E(G)} : \quad \begin{aligned} \boldsymbol{g}(e) &\neq 0 && \text{iff } e \in E(F) \\ \boldsymbol{g}(\delta_F(v)) &= 0 && \forall v \in V(F) \end{aligned} \right\},
$$

$$
\mathcal{C}_2 = \bigcup_{F \in \mathcal{E}_2} \left\{ \boldsymbol{g} \in \{-1,1\}^{E(G)} : \quad \begin{aligned} \boldsymbol{g}(e) &\neq 0 && \text{iff } e \in E(F) \\ \boldsymbol{g}(\delta_F(w)) &\neq 0 && \text{for one } w \in V(F) \\ \boldsymbol{g}(\delta_F(v)) &= 0 && \forall v \in V(F) \setminus \{w\} \end{aligned} \right\},
$$

$$
\mathcal{C}_3 = \bigcup_{F \in \mathcal{E}_3} \left\{ \boldsymbol{g} \in \{-1,1\}^{E(G)} : \quad \begin{aligned} \boldsymbol{g}(e) &\neq 0 && \text{iff } e \in E(F) \\ \boldsymbol{g}(\delta_F(v)) &= 0 && \forall v : |\delta_F(v)| = 2 \end{aligned} \right\},
$$

$$
\mathcal{C}_4 = \bigcup_{F = (P \cup C) \in \mathcal{E}_4} \left\{ \boldsymbol{g} \in \mathbb{Z}^{E(G)} : \quad \begin{aligned} \boldsymbol{g}(e) &\neq 0 && \text{iff } e \in E(F) \\ \boldsymbol{g}(\delta_F(v)) &= 0 && \forall v : |\delta_F(v)| \geq 2 \\ \boldsymbol{g}(e) &\in \{-1,1\} && \forall e \in E(C) \\ \boldsymbol{g}(e) &\in \{-2,2\} && \forall e \in E(P) \end{aligned} \right\},
$$

$$
\mathcal{C}_5 = \bigcup_{F = (C_1 \cup P \cup C_2) \in \mathcal{E}_5} \left\{ \boldsymbol{g} \in \mathbb{Z}^{E(G)} : \quad \begin{aligned} \boldsymbol{g}(e) &\neq 0 && \text{iff } e \in E(F) \\ \boldsymbol{g}(\delta_F(v)) &= 0 && \forall v \in V(F) \\ \boldsymbol{g}(e) &\in \{-1,1\} && \forall e \in E(C_1 \cup C_2) \\ \boldsymbol{g}(e) &\in \{-2,2\} && \forall e \in E(P) \end{aligned} \right\}.
$$

See Figure 4 for an example of a vector $\boldsymbol{g} \in \mathcal{C}_5$.

Let us denote by $\mathcal{C}(\mathcal{P}_{\text{FMAT}}(G))$ the set of circuits of $\mathcal{P}_{\text{FMAT}}(G)$ with co-prime integer components.

Lemma 2. $\mathcal{C}(\mathcal{P}_{\text{FMAT}}(G)) = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \cup \mathcal{C}_4 \cup \mathcal{C}_5$.
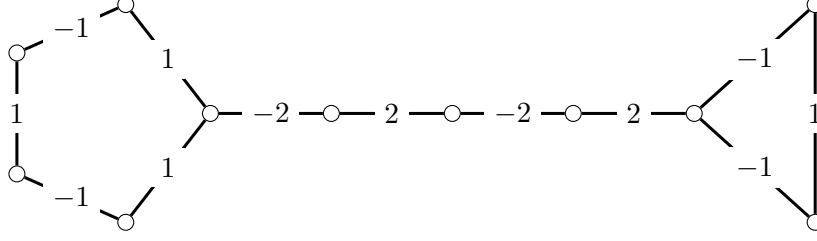
Figure 4: Example of a vector $\boldsymbol{g} \in \mathcal{C}_5$. Each edge e is labeled with $\boldsymbol{g}(e)$.

Proof. It is known that the vectors of $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_5$ correspond to edge-directions of $\mathcal{P}_{\mathrm{FMAT}}(G)$ (see e.g. [6, 39]), so it remains to be shown that all circuits belong to one of these sets.

Let B denote the constraint matrix corresponding to the inequality constraints (2). In what follows, the rows of B will be indexed by $V(G) \setminus L(G)$, and the columns of B will be indexed by $E(G)$. With this notation, we can treat $\mathrm{supp}(\mathsf{B}\boldsymbol{x})$ and $\mathrm{supp}(\boldsymbol{x})$ as a subset of $V(G)$ or $E(G)$, respectively. Let $\boldsymbol{g} \in \mathcal{C}(\mathcal{P}_{\mathrm{FMAT}}(G))$, and let $G(\boldsymbol{g})$ be the subgraph of G induced by the edges in $\mathrm{supp}(\boldsymbol{g})$.

First we note that $G(\boldsymbol{g})$ is connected. Otherwise, restricting $\boldsymbol{g}$ to the edges of any component of $G(\boldsymbol{g})$ gives a vector $\boldsymbol{f}$ with $\mathrm{supp}(\mathsf{B}\boldsymbol{f}) \subseteq \mathrm{supp}(\mathsf{B}\boldsymbol{g})$ and $\mathrm{supp}(\boldsymbol{f}) \subsetneq \mathrm{supp}(\boldsymbol{g})$, contradicting that $\boldsymbol{g}$ is a circuit.

Now, suppose that $G(\boldsymbol{g})$ contains no cycles. Let P be any edge-maximal path in $G(\boldsymbol{g})$, with endpoints u and w. Note that $\mathrm{supp}(\mathsf{B}\boldsymbol{g}) \supseteq \{u, w\} \setminus L(G)$. Let $\boldsymbol{f} \in \{-1, 1\}^{E(G)}$ be a vector that satisfies (i) $\boldsymbol{f}(e) \neq 0$ if and only if $e \in E(P)$, and (ii) $\boldsymbol{f}(\delta_P(v)) = 0 \; \forall v \neq u, w$. Note that $\boldsymbol{f} \in \mathcal{C}_3$. Then, $\mathrm{supp}(\mathsf{B}\boldsymbol{f}) = \{u, w\} \setminus L(G) \subseteq \mathrm{supp}(\mathsf{B}\boldsymbol{g})$, and $\mathrm{supp}(\boldsymbol{f}) \subseteq \mathrm{supp}(\boldsymbol{g})$. Therefore, it must be that the edges of $G(\boldsymbol{g})$ are exactly $E(P)$, and $\boldsymbol{g}(\delta_P(v)) = 0$ for all $v \in V(G) \setminus \{u, w\}$. Thus, $\boldsymbol{g} = \boldsymbol{f}$ or $\boldsymbol{g} = -\boldsymbol{f}$. In any case, $\boldsymbol{g} \in \mathcal{C}_3$.

Now, suppose that $G(\boldsymbol{g})$ contains an even cycle C. Let $\boldsymbol{f} \in \{-1, 1\}^{E(G)}$ be a vector that satisfies (i) $\boldsymbol{f}(e) \neq 0$ if and only if $e \in E(C)$, and (ii) $\boldsymbol{f}(\delta_C(v)) = 0 \; \forall v \in V(C)$. Note that $\boldsymbol{f} \in \mathcal{C}_1$. Then, $\mathrm{supp}(\mathsf{B}\boldsymbol{f}) = \emptyset \subseteq \mathrm{supp}(\mathsf{B}\boldsymbol{g})$, and $\mathrm{supp}(\boldsymbol{f}) \subseteq \mathrm{supp}(\boldsymbol{g})$. Therefore, it must be that the edges of $G(\boldsymbol{g})$ are exactly $E(C)$, and $\boldsymbol{g}(\delta_C(v)) = 0$ for all $v \in V(G)$. Thus, $\boldsymbol{g} = \boldsymbol{f}$ or $\boldsymbol{g} = -\boldsymbol{f}$. In any case, $\boldsymbol{g} \in \mathcal{C}_1$.

We are left with the case where $G(\boldsymbol{g})$ contains at least one cycle, but it does not contain any even cycle. In this case, first we state an easy claim that gives some more structure for the graph $G(\boldsymbol{g})$.

Claim 1. *Under the assumption that $G(\boldsymbol{g})$ contains at least one cycle, but it does not contain an even cycle, any two odd cycles in $G(\boldsymbol{g})$ must share at most one node.*

Proof. Let $C, D \subseteq G(\boldsymbol{g})$ be two odd cycles, and suppose for the sake of contradiction that $|V(C) \cap V(D)| \geq 2$. Then C can be written as the union of two edge-disjoint paths $C_1 \cup C_2$ where C_1 is some sub-path of C such that $V(C_1) \cap V(D) = \{u, v\}$ where u and v are the endpoints of C_1, and $E(C_1) \cap E(D) = \emptyset$. Since D is a cycle, we can decompose D into two sub-paths D_1 and D_2 each with endpoints u and v. Since $|E(D)|$ is odd, for *exactly* one $i \in \{1, 2\}$, $|E(D_i)|$ is even. Note that since $V(C_1) \cap V(D_i) = \{u, v\}$, $C_1 \cup D_i$ is a cycle for all $i \in \{1, 2\}$, and therefore there exists $i \in \{1, 2\}$ such that $C_1 \cup D_i$ is an even cycle, a contradiction with the assumption. $\square$

Suppose that $G(\boldsymbol{g})$ contains at least two distinct odd cycles C_1 and C_2. Since $G(\boldsymbol{g})$ is connected, then either these two cycles share a node or there exists a simple path P in $G(\boldsymbol{g})$ connecting them. In particular, we can choose P so that $E(P) \cap E(C_i) = \emptyset$ for $i \in \{1, 2\}$. Let $F = C_1 \cup P \cup C_2$ (where $E(P) = \emptyset$ if C_1 and C_2 share a node). Let $\boldsymbol{f} \in \mathbb{Z}^{E(G)}$ be a vector that satisfies (i) $\boldsymbol{f}(e) \neq 0$

if and only if $e \in E(F)$, (ii) $\boldsymbol{f}(\delta_F(v)) = 0 \ \forall v \in V(F)$, (iii) $\boldsymbol{f}(e) \in \{-1, 1\}$ for all $e \in E(C_1 \cup C_2)$, and (iv) $\boldsymbol{f}(e) \in \{-2, 2\}$ for all $e \in E(P)$. Note that $\boldsymbol{f} \in \mathcal{C}_5$. Then $\mathrm{supp}(\mathsf{B}\boldsymbol{f}) = \emptyset \subseteq \mathrm{supp}(\mathsf{B}\boldsymbol{g})$, and $\mathrm{supp}(\boldsymbol{f}) \subseteq \mathrm{supp}(\boldsymbol{g})$. Therefore, it must be that the edges of $G(\boldsymbol{g})$ are exactly $E(F)$, and $\boldsymbol{g}(\delta_G(v)) = 0$ for all $v \in V(G)$. Thus, $\boldsymbol{g} \in \mathcal{C}_5$.

Finally, suppose that $G(\boldsymbol{g})$ contains exactly one odd cycle C. If there exists a node $w \in V(C)$ such that $\boldsymbol{g}(\delta_G(w)) \neq 0$, then let $\boldsymbol{f} \in \{-1, 1\}^{E(G)}$ be a vector that satisfies (i) $\boldsymbol{f}(e) \neq 0$ if and only if $e \in E(C)$, and (ii) $\boldsymbol{f}(\delta_C(v)) = 0 \ \forall v \in V(C) \setminus \{w\}$. Note that $\boldsymbol{f} \in \mathcal{C}_2$. Then, $\mathrm{supp}(\mathsf{B}\boldsymbol{f}) = \{w\} \subseteq \mathrm{supp}(\mathsf{B}\boldsymbol{g})$, and $\mathrm{supp}(\boldsymbol{f}) \subseteq \mathrm{supp}(\boldsymbol{g})$. Therefore, it must be that the edges of $G(\boldsymbol{g})$ are exactly $E(C)$, and $\boldsymbol{g}(\delta_C(v)) = 0$ for all $v \in V(G) \setminus \{w\}$. Thus, it must be that $\boldsymbol{g} \in \mathcal{C}_2$.

We are left with the case where $\boldsymbol{g}(\delta_G(v)) = 0$ for all $v \in V(C)$. Note that this is not possible if $\mathrm{supp}(\boldsymbol{g}) = E(C)$, because C is an odd cycle. Then let P be any simple path in $G(\boldsymbol{g})$ which is inclusion-wise maximal subject to the condition that $E(P) \cap E(C) = \emptyset$ and $|V(P) \cap V(C)| = \{u\}$, where u is an endpoint of P. Let $F = C \cup P$, and let $w \in V(G)$ be the unique node such that $|\delta_F(w)| = 1$. Let $\boldsymbol{f} \in \mathbb{Z}^{E(G)}$ be a vector that satisfies (i) $\boldsymbol{f}(e) \neq 0$ if and only if $e \in E(F)$, (ii) $\boldsymbol{f}(\delta_F(v)) = 0 \ \forall v \in V(F) \setminus \{w\}$, (iii) $\boldsymbol{f}(e) \in \{-1, 1\}$ for all $e \in E(C)$, and (iv) $\boldsymbol{f}(e) \in \{-2, 2\}$ for all $e \in E(P)$. Note that $\boldsymbol{f} \in \mathcal{C}_4$. Then $\mathrm{supp}(\mathsf{B}\boldsymbol{f}) = \{w\} \setminus L(G) \subseteq \mathrm{supp}(\mathsf{B}\boldsymbol{g})$, and $\mathrm{supp}(\boldsymbol{f}) \subseteq \mathrm{supp}(\boldsymbol{g})$. Therefore, it must be that the edges of $G(\boldsymbol{g})$ are exactly $E(F)$, and $\boldsymbol{g}(\delta_F(v)) = 0$ for all $v \in V(G) \setminus \{w\}$. Thus, it must be that $\boldsymbol{g} \in \mathcal{C}_4$.

In all the above cases, $\boldsymbol{g} \in \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_5$, as desired. $\qquad\square$

3.2 Hardness Reduction

We will start by proving hardness for the Dantzig circuit-pivot rule.

Theorem 5. *Solving the optimization problem $Dan(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$ is NP-hard.*

Proof. We will prove this via reduction from the directed Hamiltonian path problem. Let $D = (N, F)$ be a directed graph with $n = |N|$, and let $s, t \in N$ be two given nodes. We will construct a suitable auxiliary undirected graph H, cost function $\boldsymbol{c}$, and a matching M in H, such that the following holds: D contains a directed Hamiltonian s, t-path if and only if an optimal solution to $Dan(\mathcal{P}_{\mathrm{FMAT}}(H), \chi^M, \boldsymbol{c})$ (where χ^M is the characteristic vector of M) attains a certain objective function value.

We start by constructing $H = (V, E)$. For each node $v \in N \setminus \{t\}$ we create two copies v_a and v_b in V. For all $v \in N \setminus \{t\}$, we let $v_a v_b \in E$. For all arcs $uv \in F$, with $u, v \neq t$, we add an edge $u_b v_a \in E$. That is, every in-arc at a node v corresponds to an edge incident with v_a, and every out-arc at v corresponds to an edge incident with v_b. We add t in V, and for all arcs $ut \in F$, we have that $u_b t \in E$. Finally, we add nodes s' and t', where $s's_a \in E$ and $tt' \in E$ (see Figure 5).
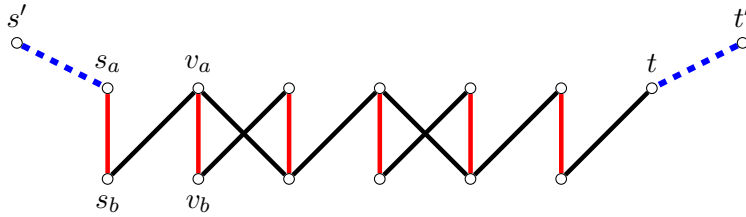

Figure 5: An example of the auxiliary graph H.

Now we define the cost function $\boldsymbol{c}$. We set $\boldsymbol{c}(v_a v_b) = 0$ for all $v \in N \setminus \{t\}$, $\boldsymbol{c}(s's_a) = -W = -\boldsymbol{c}(tt')$ such that $W \in \mathbb{Z}$, $W \gg |E|$, and let all other edges have cost -1. Finally, we let

$$M = \{v_a v_b : v \in N \setminus \{t\}\} \cup \{tt'\}$$

be a matching in H. We claim that there exists a directed Hamiltonian s, t-path in D if and only if there is a solution $\boldsymbol{g}$ to $Dan(\mathcal{P}_{\text{FMAT}}(H), \chi^M, \boldsymbol{c})$ with objective function value at least $-\boldsymbol{c}^\mathsf{T} \boldsymbol{g} = 2W + n - 1$.

($\Rightarrow$) Suppose that there exists a directed Hamiltonian s, t-path $P = (sv^1, v^1 v^2, \cdots, v^{k-1} v^k, v^k t)$ in D. Then, P can be naturally associated to an M-alternating path P' in H with endpoints s' and t', as follows:

$$P' = (s' s_a, s_a s_b, s_b v_a^1, v_a^1 v_b^1, v_b^1 v_a^2, v_a^2 v_b^2, \cdots, v_a^{k-1} v_b^{k-1}, v_b^{k-1} v_a^k, v_a^k v_b^k, v_b^k t, tt').$$

Let $\boldsymbol{g}$ be defined as

$$\boldsymbol{g}(e) := \begin{cases} 1 & \text{if } e \in E(P') \setminus M, \\ -1 & \text{if } e \in M, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\boldsymbol{g} \in \mathcal{C}_3$, and is therefore a circuit of $\mathcal{P}_{\text{FMAT}}(H)$. Note that $\chi^M + \boldsymbol{g} \in \mathcal{P}_{\text{FMAT}}(H)$, and $-\boldsymbol{c}^\mathsf{T} \boldsymbol{g} = 2W + n - 1$. Thus, $\boldsymbol{g}$ is a feasible solution to $Dan(\mathcal{P}_{\text{FMAT}}(H), \chi^M, \boldsymbol{c})$ with the claimed objective function value.

($\Leftarrow$) Now suppose that there is a solution $\boldsymbol{g}$ to $Dan(\mathcal{P}_{\text{FMAT}}(H), \chi^M, \boldsymbol{c})$, with objective function value at least $2W + n - 1$. First, we argue that the support of $\boldsymbol{g}$ is indeed an M-alternating path with endpoints s' and t'.

Note that, by construction, H is bipartite, so $\boldsymbol{g} \in \mathcal{C}_1 \cup \mathcal{C}_3$. In either case, $\boldsymbol{g} \in \{1, 0, -1\}^E$. By our choice of W, since $-\boldsymbol{c}^\mathsf{T} \boldsymbol{g} \geq 2W + n - 1$, it must be that $\boldsymbol{g}(s's_a) = 1$ and $\boldsymbol{g}(tt') = -1$. Then, since s' and t' are not in any cycles of H, necessarily $\boldsymbol{g} \in \mathcal{C}_3$ and its support is an s', t'-path. It follows that $\boldsymbol{g}$ has at most $|V| - 1$ non-zero entries. Two of the non-zero entries are $\boldsymbol{g}(s's_a)$ and $\boldsymbol{g}(tt')$, and of those that remain, exactly half have value 1. Thus,

$$-c^\mathsf{T} \boldsymbol{g} \leq 2W + \frac{1}{2}(|V| - 3) = 2W + \frac{1}{2}((2n + 1) - 3) = 2W + n - 1.$$

It is clear that the above inequality holds tight only if $\boldsymbol{g}(e) = 1$ for $\frac{1}{2}(|V| - 3)$ edges of $E \setminus \{s's_a, tt'\}$, all of which have $\boldsymbol{c}(e) = -1$, and $\boldsymbol{c}(f) = 0$ for all edges f such that $\boldsymbol{g}(f) = -1$. Since the number of edges e with $\boldsymbol{g}(e) = 1$ equals the number of edges f with $\boldsymbol{g}(f) = -1$, we have that $|\operatorname{supp}(\boldsymbol{g})| = |V| - 1$, and therefore $\operatorname{supp}(\boldsymbol{g})$ is a path P' spanning H. Furthermore, all edges of M are in $E(P')$. By removing the first and the last edge of P', and by contracting all edges of M that are the form $(v_a v_b)$ (for $v \in N$), we obtain a path that naturally corresponds to a directed Hamiltonian s', t'-path in D. $\square$

Note that the above proof immediately yields the following theorem as a corollary.

Theorem 6. *Solving the optimization problem $Great(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$ is NP-hard.*

Proof. The proof is identical to that of Theorem 5, we only need to replace $Dan(\mathcal{P}_{\text{FMAT}}(H), \chi^M, \boldsymbol{c})$ with $Great(\mathcal{P}_{\text{FMAT}}(H), \chi^M, \boldsymbol{c})$. This is because for any circuit $\boldsymbol{y} \in \mathcal{C}(\mathcal{P}_{\text{FMAT}}(H))$, we have $\chi^M + 1\boldsymbol{y} \in \mathcal{P}$, and $\chi^M + \alpha \boldsymbol{y} \notin \mathcal{P}$ for any $\alpha > 1$. Therefore, for all $\boldsymbol{y} \in \mathcal{C}(\mathcal{P}_{\text{FMAT}}(H))$ such that $-\boldsymbol{c}^\mathsf{T} \boldsymbol{y} > 0$, we have

$$\max\{-\boldsymbol{c}^\mathsf{T}(\alpha \boldsymbol{y}) : \chi^M + \alpha \boldsymbol{y} \in \mathcal{P}_{\text{FMAT}}(H), \alpha > 0\} = -\boldsymbol{c}^\mathsf{T} \boldsymbol{y}.$$

It is not difficult to see that this implies the result. $\square$

We highlight that these hardness results hold indeed for 0/1 polytopes. In fact, since by our construction the graph H is bipartite, the polytope $\mathcal{P}_{\mathrm{FMAT}}(H)$ is integral.

3.3 Hardness implications

Here we prove that the reductions in the previous section have interesting hardness implications for the Simplex method.

Corollary 3. *Given a feasible extreme point solution of a 0/1-LP, computing the best neighbor extreme point is NP-hard.*

Proof. Consider again the hardness reduction used in the proof of Theorem 6, and note that the optimal solution of $Great(\mathcal{P}_{\mathrm{FMAT}}(H), \chi^M, \boldsymbol{c})$ is a circuit $\boldsymbol{g}$ that corresponds to an edge-direction at χ^M. As a consequence, if we consider the LP obtained by minimizing $\boldsymbol{c}^\intercal \boldsymbol{x}$ over $\mathcal{P}_{\mathrm{FMAT}}(H)$, and take χ^M as an initial vertex solution, there is a neighbor optimal solution of objective function value $-W - n + 1$ (which is the minimum possible value) if and only if the initial directed graph has a Hamiltonian path. The result follows. $\qquad\square$

Note that the above proof also yields a proof for Corollary 1. We can now prove Theorem 2, that we restate for convenience.

Theorem 2. *Given an LP and an initial feasible solution, finding the shortest (monotone) path to an optimal solution is NP-hard. Furthermore, unless P=NP, it is hard to approximate within a factor strictly better than two.*

Proof. Once again, consider the hardness reduction used in the proof of Theorem 6, and the LP obtained by minimizing $\boldsymbol{c}^\intercal \boldsymbol{x}$ over $\mathcal{P}_{\mathrm{FMAT}}(H)$. In order for a Hamiltonian path to exist on D, the optimal solution of this LP must have objective function value $-W - n + 1$, so without loss of generality, we can assume that this is the case. Take χ^M as the initial vertex solution. Under the latter assumption, as noted in the proof of the previous corollary, there is a neighbor optimal solution to χ^M if and only if D has a Hamiltonian path. This implies the following: (i) if D has a Hamiltonian path, then there is a shortest (monotone) path to an optimal solution on the 1-skeleton of $\mathcal{P}_{\mathrm{FMAT}}(H)$, that consists of one edge; (ii) if D does not have a Hamiltonian path, then any shortest (monotone) path to an optimal solution has at least two edges. The result follows. $\quad\square$

As mentioned in the introduction, our result implies that for any efficiently-computable pivoting rule, the Simplex method cannot be guaranteed to reach an optimal solution via a minimum number of non-degenerate pivots, unless P=NP. In a way, this result is similar in spirit to some hardness results proven about the vertices that the Simplex method can visit during its execution [17, 16, 1].

The latter hardness result also holds for circuit-paths, via the exact same argument.

Corollary 4. *Given an LP and an initial feasible solution, finding the shortest (monotone) circuit-path to an optimal solution is NP-hard. Furthermore, unless P=NP, it is hard to approximate within a factor strictly better than two.*

Note that Theorem 5 and Theorem 6 yield a proof of Theorem 3.

4 Approximation of Circuit-Pivot Rules

We start with the following formal definition of approximate greatest-improvement augmentations.

Definition 2. *Let $\gamma \geq 1$, $\boldsymbol{x} \in \mathcal{P}$, and $\alpha^* \boldsymbol{g}^*$ be a greatest-improvement augmentation at $\boldsymbol{x}$. We say that an augmentation $\alpha \boldsymbol{g}$ is a γ-approximate greatest-improvement augmentation at $\boldsymbol{x}$, if*

$$c^{\mathsf{T}}\boldsymbol{x} - c^{\mathsf{T}}(\boldsymbol{x} + \alpha \boldsymbol{g}) \geq \frac{1}{\gamma}\Big(c^{\mathsf{T}}\boldsymbol{x} - c^{\mathsf{T}}(\boldsymbol{x} + \alpha^* \boldsymbol{g}^*)\Big).$$

As mentioned in the introduction, we define

$$\delta := \max \left\{ \left| \det \begin{pmatrix} A \\ D \end{pmatrix} \right| \right\},$$

where the max is taken over all $n \times n$ submatrix $\begin{pmatrix} A \\ D \end{pmatrix}$ of $\begin{pmatrix} A \\ B \end{pmatrix}$ such that $\begin{pmatrix} A \\ D \end{pmatrix}$ has rank n.
We also define

$$\bar{\delta} := \max \left\{ \left| \det \begin{pmatrix} \bar{A} \\ \bar{D} \end{pmatrix} \right| \right\},$$

where the max is taken over all $n \times n$ submatrix $\begin{pmatrix} \bar{A} \\ \bar{D} \end{pmatrix}$ of $\begin{pmatrix} A|\boldsymbol{b} \\ B|\boldsymbol{d} \end{pmatrix}$ such that $\begin{pmatrix} \bar{A} \\ \bar{D} \end{pmatrix}$ has rank n.

Furthermore, we let ω_1 be the minimum 1-norm distance from any extreme point to any facet not containing it. Formally, let $vert(\mathcal{P})$ be the set of vertices of $\mathcal{P}$. For a given $\boldsymbol{v} \in vert(\mathcal{P})$, let $\mathcal{F}(\boldsymbol{v})$ be the set of feasible points of $\mathcal{P}$ that lie on any facet F of $\mathcal{P}$ with $\boldsymbol{v} \notin F$.

$$\omega_1 := \min_{\boldsymbol{v} \in vert(\mathcal{P}), \, \boldsymbol{f} \in \mathcal{F}(\boldsymbol{v})} \|\boldsymbol{v} - \boldsymbol{f}\|_1.$$

Finally, we let M_1 be the maximum 1-norm distance between any pair of extreme points, i.e.

$$M_1 := \max_{\boldsymbol{v}_1, \boldsymbol{v}_2 \in vert(\mathcal{P})} \|\boldsymbol{v}_1 - \boldsymbol{v}_2\|_1.$$

4.1 Approximate greatest-improvement augmentations

Let us recall the statement of Lemma 1:

Lemma 1. *Consider an LP in the general form (1). Denote by δ the maximum absolute value of the determinant of any $n \times n$ submatrix of $\begin{pmatrix} A \\ B \end{pmatrix}$. Let $\boldsymbol{x}_0$ be an initial feasible solution, and let $\gamma \geq 1$. Using a γ-approximate greatest-improvement circuit-pivot rule, we can reach an optimal solution $\boldsymbol{x}_{\min}$ of (1) with with $O\big(n\gamma \log \big(\delta \, \boldsymbol{c}^{\mathsf{T}}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min})\big)\big)$ augmentations.*

The proof of Lemma 1 closely mimics the arguments used in [14]. However, since the authors of [14] consider LPs in equality form, for the sake of completeness we will re-state (and sometimes reprove) some of their lemmas for our more general setting. Indeed, when working with circuits, converting an LP to equality form by adding slack variables cannot be done without loss of generality, since this operation might increase the number of circuits (see [12]).

The first proposition that we state is the *sign-compatible representation property* of circuits. We say two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ are *sign-compatible with respect to B* if the i-th components of the vectors $(B\boldsymbol{v})$ and $(B\boldsymbol{w})$ satisfy $(B\boldsymbol{v})(i) \cdot (B\boldsymbol{w})(i) \geq 0$ for all $1 \leq i \leq m_B$. The representation property is as follows:

Proposition 1 (see Proposition 1.4 in [18]). *Let $v \in \ker(A) \setminus \{\mathbf{0}\}$. Then we can express v as $v = \sum_{i=1}^{k} \alpha_i g^i$ such that for all $1 \le i \le k$*

- $g^i \in \mathcal{C}(A, B)$,

- g^i *and* v *are sign-compatible with respect to* B *and* $\operatorname{supp}(Bg^i) \subseteq \operatorname{supp}(Bv)$,

- $\alpha_i \in \mathbb{R}_{\ge 0}$,

- *and* $k \le n$.

Let $x_{\max}$ be a maximizer of the LP problem (1), i.e., an optimal solution of the LP obtained from (1) by multiplying the objective function by -1. We will use the following lemma from [14] based on well-known estimates of [2]:

Lemma 3 (see Lemma 1 in [14]). *Let $\epsilon > 0$ be given. Let c be an integer vector. Define $f^{\min} := c^{\mathsf{T}} x_{\min}$, $f^{\max} := c^{\mathsf{T}} x_{\max}$. Suppose that $f^k = c^{\mathsf{T}} x_k$ is the objective function value of the solution x_k at the k-th iteration of an augmentation algorithm. Furthermore, suppose that the algorithm guarantees that for every augmentation k,*

$$(f^k - f^{k+1}) \ge \beta(f^k - f^{\min}).$$

Then the algorithm reaches a solution with $f^k - f^{\min} < \epsilon$ in no more than $2 \log\left((f^{\max} - f^{\min})/\epsilon\right)/\beta$ augmentations.

We now state the following easy lemma, that we reprove for completeness.

Lemma 4. *Let $\bar{x}$ be any feasible solution of the LP problem (1). Then with a sequence of at most n maximal augmentations, we can reach an extreme point $\hat{x}$ of (1) such that $c^{\mathsf{T}} \hat{x} \le c^{\mathsf{T}} \bar{x}$.*

Proof. Let $T = \{\, i : B_i \bar{x} = d(i) \,\}$. If $\bar{x}$ is not a vertex, then we can select any direction $g \in \ker\binom{A}{B_T}$ such that $c^{\mathsf{T}} g \le 0$, and such that for some $\epsilon > 0$, $\tilde{x} := \bar{x} + \epsilon g$ satisfies $B_i \tilde{x} \le d(i)$ for all $i \notin T$. We then use g to perform a maximal step αg at $\bar{x}$. Since the step is maximal, there exists an index $i \notin T$ such that $B_i(\bar{x} + \alpha g) = d(i)$. This enables us to grow the set T at the new feasible solution. Furthermore, $c^{\mathsf{T}}(\bar{x} + \alpha g) \le c^{\mathsf{T}} \bar{x}$.

We can iterate this process, and note that the number of linearly independent rows of $\binom{A}{B_T}$ increases by one at each step. Therefore, after at most $n - \operatorname{rk}(A)$ iterations we arrive at a vertex $\hat{x}$.

Note that the above argument does not require the use of circuits, but it requires only that the selected directions are improving with respect to c. By the sign-compatible representation property of circuits though, at any non-optimal point $\bar{x}$, there always exists an improving direction that is a circuit. $\qquad\square$

We can now give a proof of Lemma 1.

Proof of Lemma 1. Let x_k be the solution at the k-th iteration of an augmentation algorithm. By the sign-compatible representation property of the circuits,

$$x_{\min} - x_k = \sum_{i=1}^{p} \alpha_i g^i$$

where $g^i \in \mathcal{C}(A, B)$ and $p \le n$. Note that, as a consequence of Proposition 1, for any i we have that $\alpha_i \ge 0$ and $x_k + \alpha_i g^i$ is feasible. More precisely, $A(x_k + \alpha_i g^i) = A x_k + \alpha_i A g^i = A x_k = b$.

Furthermore, we know $B\boldsymbol{x}_k \leq \boldsymbol{d}$ and $B(\boldsymbol{x}_{\min}) = B(\boldsymbol{x}_k + \sum_{i=1}^{p} \alpha_i \boldsymbol{g}^i) \leq \boldsymbol{d}$. This and the sign-compatibility of $\boldsymbol{g}^i$ and $(\boldsymbol{x}_{\min} - \boldsymbol{x}_k)$ implies that $B(\boldsymbol{x}_k + \alpha_i \boldsymbol{g}^i) \leq \boldsymbol{d}$ for all $i = 1 \ldots p$.
We then have

$$0 > \boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_{\min} - \boldsymbol{x}_k) = \boldsymbol{c}^\mathsf{T} \sum_{i=1}^{p} \alpha_i \boldsymbol{g}^i = \sum_{i=1}^{p} \alpha_i \boldsymbol{c}^\mathsf{T} \boldsymbol{g}^i \geq -n\Delta,$$

where $\Delta > 0$ is the largest value of $-\alpha \boldsymbol{c}^\mathsf{T} \boldsymbol{z}$ over all $\boldsymbol{z} \in \mathcal{C}(A, B)$ and $\alpha > 0$ for which $\boldsymbol{x}_k + \alpha \boldsymbol{z}$ is feasible. Those circuit vectors and coefficients include the $\boldsymbol{g}^i, \alpha_i$ above in particular. Equivalently, we get

$$\Delta \geq \frac{\boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_k - \boldsymbol{x}_{\min})}{n}.$$

Now let $\alpha \boldsymbol{z}$ be a γ-approximate greatest-improvement augmentation applied to $\boldsymbol{x}_k$, leading to $\boldsymbol{x}_{k+1} := \boldsymbol{x}_k + \alpha \boldsymbol{z}$. Since $-\alpha \boldsymbol{c}^\mathsf{T} \boldsymbol{z} \geq \frac{1}{\gamma}\Delta$, we get

$$\boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_k - \boldsymbol{x}_{k+1}) = -\alpha \boldsymbol{c}^\mathsf{T} \boldsymbol{z} \geq \frac{1}{\gamma}\Delta \geq \frac{\boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_k - \boldsymbol{x}_{\min})}{\gamma n}.$$

Thus, we have at least a factor of $\beta = \frac{1}{\gamma n}$ of objective function value decrease at each augmentation. Applying Lemma 3 with $\epsilon = 1/\delta^2$ then yields a solution $\bar{\boldsymbol{x}}$ with $\boldsymbol{c}^\mathsf{T}(\bar{\boldsymbol{x}} - \boldsymbol{x}_{\min}) < 1/\delta^2$, obtained within at most $4n\gamma \log(\delta \, \boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min}))$ augmentations.

By Lemma 4, a vertex solution $\boldsymbol{x}'$ with $\boldsymbol{c}^\mathsf{T} \boldsymbol{x}' \leq \boldsymbol{c}^\mathsf{T} \bar{\boldsymbol{x}}$ can be reached from $\bar{\boldsymbol{x}}$ in at most n additional augmentations. It remains to prove that $\boldsymbol{x}'$ is optimal.

Suppose $\boldsymbol{x}'$ is a non-optimal vertex. There exist subsets T_1 and T_2 of $\{1, \ldots, m_B\}$ such that $\boldsymbol{x}'$ is the unique solution to

$$\begin{pmatrix} A \\ B_{T_1} \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{d}(T_1) \end{pmatrix},$$

and $\boldsymbol{x}_{\min}$ is the unique solution to

$$\begin{pmatrix} A \\ B_{T_2} \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{d}(T_2) \end{pmatrix}.$$

Let $\delta_1 = |\det \begin{pmatrix} A \\ B_{T_1} \end{pmatrix}|$ and $\delta_2 = |\det \begin{pmatrix} A \\ B_{T_2} \end{pmatrix}|$. By Cramer's rule, the entries of $\boldsymbol{x}'$ are integer multiples of $\frac{1}{\delta_1}$ and the entries of $\boldsymbol{x}_{\min}$ are integer multiples of $\frac{1}{\delta_2}$. Then, by letting $\delta' = \text{lcm}(\delta_1, \delta_2)$, we have that the entries of $(\boldsymbol{x}' - \boldsymbol{x}_{\min})$ are integer multiples of $\frac{1}{\delta'}$. Since $\boldsymbol{c}$ is an integer vector, we have that $\boldsymbol{c}^\mathsf{T}(\boldsymbol{x}' - \boldsymbol{x}_{\min}) \geq \frac{1}{\delta'}$, and by the definition of δ, we have that $\frac{1}{\delta'} \geq \frac{1}{\delta^2}$. This is a contradiction to the fact that $\boldsymbol{c}^\mathsf{T}(\boldsymbol{x}' - \boldsymbol{x}_{\min}) < 1/\delta^2$. $\qquad\square$

Note that the above proof also establishes that the result obtained by [14] regarding the number greatest-improvement augmentations needed to solve an equality form LP extends to the general-form LP (trivially by taking $\gamma = 1$). Therefore, an easy corollary of the above result shows a (weakly) polynomial bound on the circuit-diameter of a rational polyhedron. We recall the statement of Corollary 2:

Corollary 2. *There exists a polynomial function $f(m, \alpha)$ that bounds above the circuit-diameter of any rational polyhedron $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}, B\boldsymbol{x} \leq \boldsymbol{d}\}$ with m row constraints and maximum encoding length among the coefficients in its description equal to α.*

Proof. Consider a polyhedron $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}, B\boldsymbol{x} \leq \boldsymbol{d}\}$, and without loss of generality assume $m = m_A + m_B \geq n$. Recall that, by possibly scaling, we can assume that all coefficients in its description (i.e., all entries of $A, B, \boldsymbol{b}, \boldsymbol{d}$) are integers. Let $\boldsymbol{x}$ and $\bar{\boldsymbol{x}}$ be two extreme points of $\mathcal{P}$. Let $T := \{i : B_i\bar{\boldsymbol{x}} = \boldsymbol{d}(i)\}$. Let $\boldsymbol{c}^\mathsf{T}$ be the vector obtained by adding the rows of A and B_T, multiplied by -1. By construction, using this vector as an objective function, $\bar{\boldsymbol{x}}$ is the unique optimal solution to the LP problem (1). Lemma 1 shows that we can reach $\bar{\boldsymbol{x}}$ from $\boldsymbol{x}$ with $O(n \log (\delta \, \boldsymbol{c}^\mathsf{T}(\boldsymbol{x} - \bar{\boldsymbol{x}})))$ augmentations. Note that $\boldsymbol{c}^\mathsf{T}(\boldsymbol{x} - \bar{\boldsymbol{x}}) \leq ||\boldsymbol{c}||_\infty ||\boldsymbol{x} - \bar{\boldsymbol{x}}||_1$. The result then follows by observing that $\log(||\boldsymbol{c}||_\infty) = O(\alpha + \log m)$, $\log(||\boldsymbol{x} - \bar{\boldsymbol{x}}||_1) \leq \log(2n\bar{\delta})$ (using Cramer's rule), and $\log(\delta) \leq \log(\bar{\delta}) = O(n(\alpha + \log n))$. $\qquad\square$

4.2 Steepest-descent circuit-pivot rule

Using the approximation result developed in the previous section, we here give a new bound on the number of steepest-descent augmentations needed to solve a bounded LP.

Theorem 7. *Let ω_1 denote the minimum 1-norm distance from any extreme point $\boldsymbol{v} \in \mathcal{P}$ to any facet F of $\mathcal{P}$ such that $\boldsymbol{v} \notin F$. Let M_1 be the maximum 1-norm distance between any pair of extreme points of $\mathcal{P}$. Using a steepest-descent circuit-pivot rule, a circuit-augmentation algorithm reaches an optimal solution $\boldsymbol{x}_{\min}$ of a bounded LP (1) from any initial feasible solution $\boldsymbol{x}_0$, performing*

$$\mathcal{O}\left(n^2 \frac{M_1}{\omega_1} \log \left(\delta \, \boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min})\right)\right)$$

augmentations.

Proof. First, we can apply Lemma 4 to move from $\boldsymbol{x}_0$ to an extreme point $\boldsymbol{x}'$ of the LP in at most n steps.

Let $\hat{\boldsymbol{z}}$ be an optimal solution to $Steep(\mathcal{P}, \boldsymbol{x}', \boldsymbol{c})$, and let $\boldsymbol{z} := \frac{1}{||\hat{\boldsymbol{z}}||_1}\hat{\boldsymbol{z}}$. Note that $\boldsymbol{z}$ is a circuit of $\mathcal{P}$, being a rescaling of $\hat{\boldsymbol{z}} \in \mathcal{C}(A, B)$. Let $\alpha\boldsymbol{z}$ be a steepest-descent augmentation at $\boldsymbol{x}'$. Similarly, let $\hat{\boldsymbol{z}}^*$ be an optimal solution to $Great(\mathcal{P}, \boldsymbol{x}', \boldsymbol{c})$, let $\boldsymbol{z}^* := \frac{1}{||\hat{\boldsymbol{z}}^*||_1}\hat{\boldsymbol{z}}^*$ and let $\alpha^*\boldsymbol{z}^*$ be a greatest-improvement augmentation at $\boldsymbol{x}'$. Then we have that $-(\boldsymbol{c}^\mathsf{T}\hat{\boldsymbol{z}})/||\hat{\boldsymbol{z}}||_1 \geq -(\boldsymbol{c}^\mathsf{T}\hat{\boldsymbol{z}}^*)/||\hat{\boldsymbol{z}}^*||_1$, and so $-\boldsymbol{c}^\mathsf{T}\boldsymbol{z} \geq -\boldsymbol{c}^\mathsf{T}\boldsymbol{z}^*$. Therefore

$$-\alpha\boldsymbol{c}^\mathsf{T}\boldsymbol{z} \geq -\alpha\boldsymbol{c}^\mathsf{T}\boldsymbol{z}^* = \left(\frac{\alpha}{\alpha^*}\right)(-\alpha^*\boldsymbol{c}^\mathsf{T}\boldsymbol{z}^*).$$

Since the augmentation $\alpha\boldsymbol{z}$ is maximal, we have that at the point $\boldsymbol{x}' + \alpha\boldsymbol{z}$, there exists some facet of our feasible region which contains $\boldsymbol{x}' + \alpha\boldsymbol{z}$ but not $\boldsymbol{x}'$. Then $\omega_1 \leq ||(\boldsymbol{x}' + \alpha\boldsymbol{z}) - \boldsymbol{x}'||_1 = \alpha||\boldsymbol{z}||_1$. Since $||\boldsymbol{z}||_1 = 1$, it follows that $\alpha \geq \omega_1$. Since $\boldsymbol{x}' + \alpha^*\boldsymbol{z}^*$ is feasible, we have that $||(\boldsymbol{x}' + \alpha^*\boldsymbol{z}^*) - \boldsymbol{x}'||_1$ is at most the maximum 1-norm distance from $\boldsymbol{x}'$ to any other feasible point. As above, it follows that α^* is at most the maximum 1-norm distance from $\boldsymbol{x}'$ to any other feasible point. Since the function $f(\boldsymbol{y}) = ||\boldsymbol{y} - \boldsymbol{x}'||_1$ is convex, this maximum is achieved at an extreme point. It follows that $\alpha^* \leq M_1$.

Given these bounds on α and α^*, it follows that

$$-\alpha\boldsymbol{c}^\mathsf{T}\boldsymbol{z} \geq \left(\frac{\omega_1}{M_1}\right)(-\alpha^*\boldsymbol{c}^\mathsf{T}\boldsymbol{z}^*) \tag{$\star$}$$

Now let $\bar{\boldsymbol{x}} = \boldsymbol{x}' + \alpha\boldsymbol{z}$. By Lemma 4, an extreme point solution $\hat{\boldsymbol{x}}$ can be found from $\bar{\boldsymbol{x}}$ in at most $n-1$ additional augmentations (e.g., using again steepest-descent augmentations, but on a sequence of face-restricted LPs) with $\boldsymbol{c}^\mathsf{T}\hat{\boldsymbol{x}} \leq \boldsymbol{c}^\mathsf{T}\bar{\boldsymbol{x}}$. Then we have that $\hat{\boldsymbol{x}} - \bar{\boldsymbol{x}}$ is an $\left(\frac{\omega_1}{M_1}\right)$-approximate greatest-improvement augmentation at $\boldsymbol{x}'$, and since $\hat{\boldsymbol{x}}$ is also an extreme point, we can continue to apply
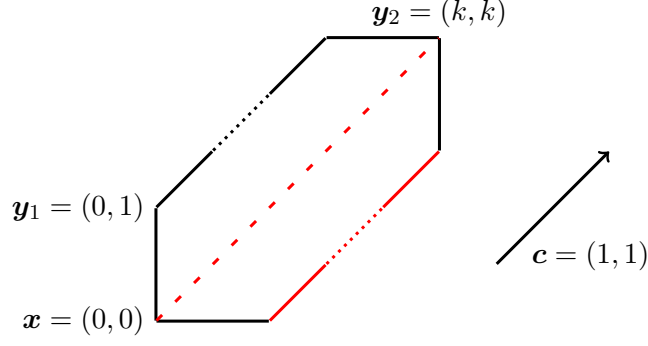
17

Figure 6: This gives a family of examples (parameterized by k) where i) moving along the edges incident at a vertex yields an arbitrarily bad approximation of moving along the greatest-improvement circuit, and ii) a steepest-descent augmentation at $\boldsymbol{x}$ is a (tight) $\frac{M_1}{\omega_1}$-approximate greatest-improvement augmentation. This polygon has vertices $\boldsymbol{x} = (0,0), \boldsymbol{y}_1 = (0,1), \boldsymbol{y}_2 = (k,k), (k, k-1), (k-1, k)$ and $(1,0)$. One can check that at $\boldsymbol{x}$, $\boldsymbol{y}_1$ is both a steepest-descent augmentation as well as a steepest edge, while $\boldsymbol{y}_2$ is a greatest-improvement augmentation. We have that $\boldsymbol{c}^\mathsf{T}\boldsymbol{y}_1 = \frac{1}{2k}\boldsymbol{c}^\mathsf{T}\boldsymbol{y}_2 = \frac{\omega_1}{M_1}\boldsymbol{c}^\mathsf{T}\boldsymbol{y}_2$.

this procedure. Since it takes at most n steepest-descent augmentations to find such an $\left(\frac{\omega_1}{M_1}\right)$-approximate greatest-improvement augmentation, it follows from Lemma 1 that from an initial solution $\boldsymbol{x}_0$, we can reach $\boldsymbol{x}_{\min}$ in $\mathcal{O}\left(n^2 \frac{M_1}{\omega_1} \log\left(\delta \boldsymbol{c}^\mathsf{T}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min})\right)\right)$ steepest-descent augmentations. □

We note that the inequality $(\star)$ yields the following corollary.

Corollary 5. *Let $\boldsymbol{x}$ be an extreme point of a bounded LP. A steepest-descent augmentation is a $\left(\frac{M_1}{\omega_1}\right)$-approximate greatest-improvement augmentation.*

The example given by Figure 6 shows that the approximation factor $\frac{M_1}{\omega_1}$ can be tight.

5 Implications for 0/1 Polytopes

In this section, we explore the implications that Theorem 7 has in the case of 0/1 polytopes, eventually proving Theorem 4 and Theorem 1. We start with the following lemma:

Lemma 5. *Consider a problem in the general form (1) whose feasible region $\mathcal{P}$ is a 0/1 polytope, and let $\boldsymbol{x}$ be a non-optimal vertex of $\mathcal{P}$. Then, the optimal solution to $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$ corresponds to an edge-direction at $\boldsymbol{x}$, and can be computed in polynomial time.*

Proof. Consider the optimal objective function value of $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$. It is not difficult to see that this value is bounded above by the optimal objective function value of the following optimization problem $\mathcal{Q}$:

$$\max -\boldsymbol{c}^\mathsf{T}\boldsymbol{z}$$
$$\text{s.t.}$$
$$\|\boldsymbol{z}\|_1 \leq 1 \tag{4}$$
$$\boldsymbol{x} + \varepsilon\boldsymbol{z} \in \mathcal{P} \qquad \text{for some } \varepsilon > 0 \tag{5}$$

18

over all $\boldsymbol{z} \in \mathbb{R}^n$. This is true since if $\boldsymbol{g} \in \mathcal{C}(A, B)$ is a feasible solution to $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$, then $\frac{\boldsymbol{g}}{\|\boldsymbol{g}\|_1}$ is a feasible solution of $\mathcal{Q}$ with the same objective function value.

Let $\mathcal{P}_{\mathcal{Q}}$ denote the feasible region of $\mathcal{Q}$. Note that $\mathcal{P}_{\mathcal{Q}}$ is the feasible cone at $\boldsymbol{x}$ in $\mathcal{P}$—given by the constraint (5)—intersected with an n-dimensional cross-polytope—given by the constraint (4). The constraint (4) can be modeled using the linear constraints

$$\boldsymbol{v}^{\mathsf{T}}\boldsymbol{z} \leq 1 \ \text{ for all } \boldsymbol{v} \in \{\, 1, -1 \,\}^n. \tag{6}$$

It follows that $\mathcal{P}_{\mathcal{Q}}$ is a polytope, and therefore $\mathcal{Q}$ is a feasible bounded LP. There exists an optimal vertex $\boldsymbol{y}$ of $\mathcal{P}_{\mathcal{Q}}$ which is determined by n linearly independent constraints of $\mathcal{P}_{\mathcal{Q}}$.

Since $\boldsymbol{x} \in \{\, 0, 1 \,\}^n$ and $\mathcal{P}$ is a 0/1 polytope, each entry of $\boldsymbol{x}$ is either equal to its upper bound or its lower bound. Thus, the feasible cone at $\boldsymbol{x}$ lies within a single orthant of $\mathbb{R}^n$. This implies that among all the linear constraints that model $\|\boldsymbol{z}\|_1 \leq 1$, only one is facet defining. Therefore, $\boldsymbol{y}$ is contained in at least $n-1$ facets corresponding to inequalities that describe the feasible cone at $\boldsymbol{x}$. Since $\boldsymbol{x}$ is not optimal, $\boldsymbol{y} \neq \boldsymbol{0}$.

As a consequence of this, we have that the optimal solution of $\mathcal{Q}$ corresponds to an edge-direction of $\mathcal{P}$ incident with $\boldsymbol{x}$. It follows that the optimal solution of $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$ is an edge-direction of $\mathcal{P}$ incident with $\boldsymbol{x}$, and it can be computed in polynomial time. $\qquad\square$

We note that it was shown in [12] that for a slightly different definition of steepest-descent, a steepest-descent circuit-augmentation can be computed in polynomial time for all LPs. Despite similarities in the two versions of the steepest-descent pivot rule, the technique they employ cannot be straightforwardly applied to work for the definition of steepest-descent used here, and our computability result in Lemma 5 is not implied by theirs.

We can now prove Theorem 4, which we restate for convenience.

Theorem 4. *Let $\boldsymbol{x}$ be an extreme point of a 0/1-LP with n variables. An augmentation along a steepest edge yields an n-approximation of an augmentation along a greatest-improvement circuit.*

Proof. We combine Corollary 5 and Lemma 5. Since $\mathcal{P}$ is a 0/1 polytope, we immediately have that $M_1 \leq n$. We now show that $\omega_1 \geq 1$. Let $\boldsymbol{v}$ be any extreme point of $\mathcal{P}$ and let F be any facet of $\mathcal{P}$ which does not contain $\boldsymbol{v}$. By reflecting and translating $\mathcal{P}$, we may assume without loss of generality that $\boldsymbol{v} = \boldsymbol{0}$ (Note that these operations do not change the 1-norm distance between any pair of points in $\mathcal{P}$). It therefore suffices to show that for any facet F not containing $\boldsymbol{0}$, $\|\boldsymbol{y}\|_1 \geq 1$ for all $\boldsymbol{y} \in F$. Since all points in F have non-negative coordinates, the minimum value of $\|\boldsymbol{y}\|_1$ over all $\boldsymbol{y} \in F$ is equal to the optimal solution to the following LP:

$$\min \boldsymbol{1}^{\mathsf{T}}\boldsymbol{y}$$
$$\text{s.t.}$$
$$\boldsymbol{y} \in F.$$

There exists an optimal solution $\boldsymbol{y}^*$ to this LP which is an extreme point solution. Since $\boldsymbol{y}^*$ is an extreme point of F, it is also an extreme point of $\mathcal{P}$, and since $\boldsymbol{y}^* \in F$, it is an extreme point not equal to $\boldsymbol{0}$. Therefore, $\boldsymbol{y}^*$ has at least one coordinate equal to 1, and so $\|\boldsymbol{y}^*\|_1 \geq 1$, as desired. Therefore, $\frac{\omega_1}{M_1} \geq \frac{1}{n}$. $\qquad\square$

We will rely next on the following result of Frank and Tardos [19].

Lemma 6 ([19]). *Let $\boldsymbol{w} \in \mathbb{R}^n$ be a rational vector, and α be a positive integer. Define $N := (n+1)! 2^{n\alpha} + 1$. Then one can compute an integral vector $\boldsymbol{w}' \in \mathbb{Z}^n$ satisfying:*

19

(a) $\|\boldsymbol{w}'\|_\infty \le 2^{4n^3} N^{n(n+2)}$;

(b) *Consider any rational LP of the form* $\max\{\ \boldsymbol{w}^\mathsf{T}\boldsymbol{x}\ :\ A'\boldsymbol{x} \le \boldsymbol{b}',\ \boldsymbol{x} \in \mathbb{R}^n\ \}$, *where the encoding length of any entry of A' is at most α. Then, $\boldsymbol{x} \in \mathbb{R}^n$ is an optimal solution to that LP if and only if it is an optimal solution to* $\max\{\ \boldsymbol{w}'^\mathsf{T}\boldsymbol{x}\ :\ A'\boldsymbol{x} \le \boldsymbol{b}',\ \boldsymbol{x} \in \mathbb{R}^n\ \}$.

We are now ready to prove the following theorem.

Theorem 8. *Given a problem in the general form (1) whose feasible region $\mathcal{P}$ is a 0/1 polytope, a circuit-augmentation algorithm with a steepest-descent circuit-pivot rule reaches an optimal solution performing a strongly-polynomial number of augmentations.*
Furthermore, if the initial solution is a vertex, the algorithm follows a path on the 1-skeleton of $\mathcal{P}$.

Proof. Let us call (LP_1) the given LP problem of the form (1) whose feasible region is $\mathcal{P}$. Since $\mathcal{P}$ is a 0/1 polytope, for the sake of the analysis we can assume that the maximum absolute value of any element in A and B is $\le \frac{n^{n/2}}{2^n}$ [45]. Apply Lemma 6 to the LP obtained from $(LP1)$ after changing the objective function to $\max \boldsymbol{w}^\mathsf{T}\boldsymbol{x}$, with $\boldsymbol{w} := -\boldsymbol{c}$. Set $\boldsymbol{c}' := -\boldsymbol{w}'$. Finally, let $(LP2) := \min\{\ \boldsymbol{c}'^\mathsf{T}\boldsymbol{x}\ :\ A\boldsymbol{x} = \boldsymbol{b},\ B\boldsymbol{x} \le \boldsymbol{d},\ \boldsymbol{x} \in \mathbb{R}^n\ \}$.

Let $\boldsymbol{x}_0$ and $\boldsymbol{x}_{\min}$ be respectively the initial solution and the optimal solution. By performing at most n additional augmentations, we can assume $\boldsymbol{x}_0$ is an extreme point.

First, we will show that Theorem 8 holds for $(LP2)$. Then, we will show that a circuit-augmentation algorithm traverses the same edge-walk when solving $(LP2)$ and $(LP1)$, if one uses the steepest-descent circuit-pivot rule. This will prove the statement.

Recall that the steepest-descent circuit-pivot rule selects at each step an improving circuit $\boldsymbol{g}$ that minimizes $\frac{\boldsymbol{c}^\mathsf{T}\boldsymbol{g}}{\|\boldsymbol{g}\|_1}$. Since the feasible region of $(LP2)$ is a 0/1 polytope, we can apply Lemma 5. Therefore, each augmentation corresponds to moving from an extreme point to an adjacent extreme point. Furthermore, the total number of augmentations can be bounded via Theorem 4 and Lemma 1 by

$$\mathcal{O}\left(n^2 \log\left(\delta\,\boldsymbol{c}'^\mathsf{T}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min})\right)\right).$$

Since the maximum absolute value of any entry of A and B is at most $\frac{n^{n/2}}{2^n}$, we have that the maximum absolute value of the determinant of any $n \times n$ submatrix is at most $\left(\frac{n^{n/2}}{2^n}\right)^n n!$. By the definition of δ, we have that $\delta \le \left(\left(\frac{n^{n/2}}{2^n}\right)^n n!\right)$, and so $\log(\delta)$ is polynomial in n.

Finally, we address the term $\log\left(\boldsymbol{c}'^\mathsf{T}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min})\right)$. Since $\boldsymbol{x}_0$ and $\boldsymbol{x}_{\min}$ are both in $\{0,1\}^n$, we have that $\log\left(\boldsymbol{c}'^\mathsf{T}(\boldsymbol{x}_0 - \boldsymbol{x}_{\min})\right) \le \log(\|\boldsymbol{c}'\|_1) \le \log(n\|\boldsymbol{c}'\|_\infty)$, which is polynomial in n due to Lemma 6(a). Therefore, the number of augmentations required to solve $(LP2)$ is strongly-polynomial in the input size.

To finish our proof, it remains to show that when the circuit-augmentation algorithm is applied to $(LP1)$, it performs the same edge-walk as it does when it is applied to $(LP2)$. To see this, we will rely on the polyhedral characterization of the problem $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$, used in the proof of Lemma 5. As explained there, the edge-direction $\boldsymbol{g}$ selected by our algorithm applied to $(LP2)$ is an optimal solution to the LP $\max\{-\boldsymbol{c}'^\mathsf{T}\boldsymbol{x}\ :\ \boldsymbol{x} \in \mathcal{P}_\mathcal{Q}\}$, which describes $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c}')$. Note that the maximum absolute value of a matrix-coefficient of this LP is also at most $\frac{n^{n/2}}{2^n}$. Therefore, due to Lemma 6(b), $\boldsymbol{g}$ is an optimal solution to $\max\{-\boldsymbol{c}'^\mathsf{T}\boldsymbol{x}\ :\ \boldsymbol{x} \in \mathcal{P}_\mathcal{Q}\}$ (i.e., $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c}')$) if and only if it is an optimal solution to $\max\{-\boldsymbol{c}^\mathsf{T}\boldsymbol{x}\ :\ \boldsymbol{x} \in \mathcal{P}_\mathcal{Q}\}$ (i.e., $Steep(\mathcal{P}, \boldsymbol{x}, \boldsymbol{c})$). Therefore, the circuit-augmentation algorithm implemented according to the steepest-descent circuit-pivot rule, performs the exact same pivots for the objective functions $\boldsymbol{c}'$ and $\boldsymbol{c}$. $\qquad\square$

It can be readily seen that Theorem 8 implies Part (i) of Theorem 1. This gives a strongly-polynomial bound on the number of *distinct* basic feasible solutions visited with a steepest-descent pivoting rule. In the context of the Simplex method, moving to a neighboring vertex along a steepest edge might require several *degenerate* basis exchanges. Having a 'pure' pivoting rule that implies a strongly-polynomial bound on the total number of basis exchanges remains an open question, but if the polytope is non-degenerate, then the two concepts coincide, hence we obtain Part (ii) of Theorem 1.

We conclude our paper with the following remark. As already mentioned, in the context of pivot rules for the Simplex method, the name *steepest-descent* often refers to normalizations according to the 2-norm of a vector, rather than the 1-norm. Since for any vector $\boldsymbol{g} \in \mathbb{R}^n$, we have $\sqrt{(n)}||\boldsymbol{g}||_2 \geq ||\boldsymbol{g}||_1 \geq ||\boldsymbol{g}||_2$, it is not difficult to note that the result of Theorem 8 still holds if we normalize according to the 2-norm.

References

[1] I. Adler, C. Papadimitriou, and A. Rubinstein. On simplex pivoting rules and complexity theory. In *Integer programming and combinatorial optimization*, volume 8494 of *Lecture Notes in Comput. Sci.*, pages 13–24. Springer, Cham, 2014.

[2] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network flows.* Prentice Hall Inc., Englewood Cliffs, NJ, 1993.

[3] N. Amenta and G. M. Ziegler. Deformed products and maximal shadows of polytopes. In *Advances in discrete and computational geometry (South Hadley, MA, 1996)*, volume 223 of *Contemp. Math.*, pages 57–90. Amer. Math. Soc., Providence, RI, 1999.

[4] D. Avis and O. Friedmann. An exponential lower bound for Cunningham's rule. *Mathematical Programming*, 161(1):271–305, Jan 2017.

[5] F. Barahona and E. Tardos. Note on Weintraub's minimum-cost circulation algorithm. *SIAM Journal on Computing*, 18(3):579–583, 1989.

[6] R. E. Behrend. Fractional perfect b-matching polytopes. i: General theory. *Linear Algebra and its Applications*, 439, 01 2013.

[7] R.G. Bland. *Complementary orthogonal subspaces of n-dimensional Euclidean space and orientability of matroids.* 1974. Thesis (Ph.D.)–Cornell University.

[8] R.G. Bland. On the generality of network flow theory. Presented at the ORSA/TIMS Joint National Meeting, Miami, Florida, 1976.

[9] R.G. Bland and J. Edmonds. Fast primal algorithms for totally unimodular linear programming. Presented at XI-th International Symposium on Mathematical Programming, Bonn, West Germany, 1982.

[10] S. Borgwardt, J. A. De Loera, and E. Finhold. Edges vs circuits: a hierarchy of diameters in polyhedra. *Advances in Geometry*, 16, 09 2014.

[11] S. Borgwardt, E. Finhold, and R. Hemmecke. On the circuit diameter of dual transportation polyhedra. *SIAM Journal on Discrete Mathematics*, 29(1):113–121, 2015.

[12] S. Borgwardt and C. Viss. A Polyhedral Model for Enumeration and Optimization over the Set of Circuits. *arXiv e-prints*, page arXiv:1811.00444, Nov 2018.

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

[14] J. A. De Loera, R. Hemmecke, and J. Lee. On augmentation algorithms for linear and integer-linear programming: From Edmonds–Karp to Bland and beyond. *SIAM Journal on Optimization*, 25:2494–2511, 2015.

[15] J. A. De Loera, B. Sturmfels, and R. R. Thomas. Gröbner bases and triangulations of the second hypersimplex. *Combinatorica*, 15(3):409–424, Sep 1995.

[16] Y. Disser and M. Skutella. The simplex algorithm is NP-mighty. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 858–872. SIAM, Philadelphia, PA, 2015.

[17] J. Fearnley and R. Savani. The complexity of the simplex method. In *STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing*, pages 201–208. ACM, New York, 2015.

[18] E. Finhold. *Circuit diameters and their application to transportation problems.* PhD thesis, Technische Univ. München, 2015.

[19] A. Frank and E. Tardos. An application of simultaneous Diophantine approximation in combinatorial optimization. *Combinatorica*, 7(1):49–65, 1987.

[20] O. Friedmann, T. D. Hansen, and U. Zwick. Subexponential lower bounds for randomized pivoting rules for the simplex algorithm. In *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC'11, San Jose, CA, USA*, pages 283–292, 2011.

[21] A. M. Frieze and S. H. Teng. On the complexity of computing the diameter of a polytope. *Comput. Complexity*, 4:207–219, 1994.

[22] J. B. Gauthier, J. Desrosiers, and M. Lübbecke. Decomposition theorems for linear programs. *Operations Research Letters*, 42(8):553 – 557, 2014.

[23] D. Goldfarb and W. Y. Sit. Worst case behavior of the steepest edge simplex method. *Discrete Applied Mathematics*, 1(4):277 – 285, 1979.

[24] T. Hansen and U. Zwick. An improved version of the random-facet pivoting rule for the simplex algorithm. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 209–218, 2015.

[25] R. Hemmecke, S. Onn, and R. Weismantel. A polynomial oracle-time algorithm for convex integer minimization. *Mathematical Programming*, 126:97–117, 2011.

[26] R.G. Jeroslow. The simplex algorithm with the pivot rule of maximizing criterion improvement. *Discrete Mathematics*, 4(4):367 – 377, 1973.

[27] S. Kafer, K. Pashkovich, and L. Sanità. On the circuit diameter of some combinatorial polytopes. *SIAM Journal on Discrete Mathematics*, 33(1):1–25, 2019.

[28] A. V. Karzanov and S. T. McCormick. Polynomial methods for separable convex optimization in unimodular linear spaces with applications. *SIAM J. Comput.*, 26(4):1245–1275, 1997.

[29] T. Kitahara and S. Mizuno. A bound for the number of different basic solutions generated by the simplex method. *Mathematical Programming*, 137(1-2, Ser. A):579–586, 2013.

[30] T. Kitahara and S. Mizuno. On the number of solutions generated by the simplex method for LP. In *Optimization and Control Techniques and Applications*, pages 75–90, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

[31] V. Klee and G. J. Minty. How good is the simplex algorithm? In *Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)*, pages 159–175. Academic Press, New York, 1972.

[32] U. H. Kortenkamp, J. Richter-Gebert, A. Sarangarajan, and G. M. Ziegler. Extremal properties of 0/1-polytopes. *Discrete & Computational Geometry*, 17(4):439–448, Jun 1997.

[33] 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(5):933–943, 2018.

[34] S. T. McCormick and A. Shioura. Minimum ratio canceling is oracle polynomial for linear programming, but not strongly polynomial, even for networks. *Operations Research Letters*, 27(5):199–207, 2000.

[35] S. Müller and G. Regensburger. Elementary vectors and conformal sums in polyhedral geometry and their relevance for metabolic pathway analysis. *Frontiers in Genetics*, 7:90, 2016.

[36] S. Onn. *Non-Linear Discrete Optimization*. Zurich Lectures in Advanced Mathematics. European Mathematical Society, 2010.

[37] A. Del Pia and C. Michini. Short simplex paths in lattice polytopes. *arXiv e-prints*, abs/1912.05712, 2019.

[38] R.T. Rockafellar. The elementary vectors of a subspace of R^N. In *Combinatorial Mathematics and its Applications (Proc. Conf., Univ. North Carolina, Chapel Hill, N.C., 1967)*, pages 104–127. Univ. North Carolina Press, Chapel Hill, N.C., 1969.

[39] L. Sanità. The diameter of the fractional matching polytope and its hardness implications. In *FOCS 2018 (59th Annual IEEE Symposium on Foundations of Computer Science)*, pages 910–921, 10 2018.

[40] A. S. Schulz and R. Weismantel. The complexity of generic primal algorithms for solving general integer programs. *Mathematics of Operations Research*, 27:681–692, 2002.

[41] A. S. Schulz, R. Weismantel, and G. M. Ziegler. 0/1-integer programming: Optimization and augmentation are equivalent. In Paul Spirakis, editor, *Algorithms — ESA '95*, pages 473–483, Berlin, Heidelberg, 1995. Springer Berlin Heidelberg.

[42] B. Sturmfels. *Gröbner bases and convex polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1996.

[43] T. Terlaky and S. Zhang. Pivot rules for linear programming: A survey on recent theoretical developments. *Annals OR*, 46-47(1):203–233, 1993.

[44] N. Zadeh. What is the worst case behavior of the simplex algorithm? In *Polyhedral computation*, volume 48 of *CRM Proc. Lecture Notes*, pages 131–143. Amer. Math. Soc., Providence, RI, 2009.

[45] G. M. Ziegler. Lectures on 0/1 polytopes. In *Polytopes – Combinatorics and Computation. DMV Seminar*, volume 29. Birkhaüser, Basel, 2000.

[46] G. M. Ziegler. Typical and extremal linear programs. In *The sharpest cut*, MPS/SIAM Ser. Optim., pages 217–230. SIAM, Philadelphia, PA, 2004.