Pivot Rules for Circuit-Augmentation Algorithms in Linear Optimization

Jesús A. De Loera∗, Sean Kafer⋆, Laura Sanitá⋆

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

Abstract

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

Interestingly, we show that this more general framework can be exploited also to make conclusions about the Simplex method itself. In particular, as a byproduct of our 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 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. The Simplex method is a perfect example of an augmentation algorithm: it starts with an initial feasible extreme point solution of the LP, and in each step it moves along an improving edge-direction to an adjacent extreme point, until an optimal solution is found or unboundedness is detected. 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 algorithm. The existence of a polynomial-time pivot rule for the Simplex method is a longstanding open question in the theory of optimization [30].

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 paper discusses several results about circuit-augmentation algorithms.

Given a polyhedron, its circuits are all potential edges that can arise by translating some of its facets. Formally:
Definition 1. Given a polyhedron of the form \( \mathcal{P} = \{ x \in \mathbb{R}^n : Ax = b, Bx \leq d \} \), a non-zero vector \( g \in \mathbb{R}^n \) is a circuit if

(i) \( g \in \ker(A) \), and

(ii) \( Bg \) is support-minimal in the collection \( \{ By : y \in \ker(A), y \neq 0 \} \).

Here \( \ker(A) \) denotes the kernel of the matrix \( A \). To represent the circuits with a finite set, we can normalize them in various ways. Following [7, 8, 10, 14], we denote by \( C(A, B) \) 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 [5, 6, 7, 8, 9, 10, 17, 19, 21, 25, 26, 29] and the many references therein). In particular, the authors of [10] considered in careful detail three pivot rules that guarantee notable bounds on the number of steps performed by a circuit-augmentation algorithm to reach an optimal solution. Specifically, consider now a general LP format where we wish to minimize an objective function \( c^\top x \) over \( \mathcal{P} \), i.e.,

\[
\min \left\{ c^\top x : Ax = b, Bx \leq d, x \in \mathbb{R}^n \right\}.
\]

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

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

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

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

Note that these circuit-pivot rules are similar to three famous pivot rules proposed for the Simplex method, for which it is known that the Simplex method can require an exponential number of steps before reaching an optimal solution [18, 20, 23]. When all circuits are considered as possible directions to move, finer bounds can be given. Most notably, the greatest-improvement pivot rule guarantees a \emph{polynomial} bound on the number of steps performed by a circuit-augmentation algorithm on LPs in equality form (see [10] 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 solve these three pivot rule optimization problems over the exponentially large set of circuits?
- Can we exploit (approximate) solutions to these pivot rules to design (strongly-) polynomial time augmentation algorithms?
- Can we exploit circuit-augmentation algorithms to analyze the Simplex method?
1.1 Our Contributions.

Hardness. First we settle the computational complexity of the pivot rules (i), (ii), and (iii) above.

**Theorem 1.** The greatest-improvement and Dantzig pivot rules are NP-hard. The steepest-descent pivot rule can be computed in polynomial time.

We prove the first part of the theorem by showing that computing a circuit according to both the greatest-improvement pivot rule and the Dantzig pivot rule is already hard to solve when $P$ is a 0/1 polytope. In particular, we focus on when $P$ 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 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 [27, 4]. Then, we construct a reduction from the NP-hard Hamiltonian path problem in Section 3.2.

We then show in Section 3.3 that the optimum of problem (iii) can be computed in polynomial time. The proof of this result follows the ideas of Borgwardt and Viss [8], who first proved that an extremely similar pivot rule to (iii) can be computed in polynomial time (they also called it a steepest-descent rule). Their pivot rule and our steepest-descent pivot rule are indeed very closely related (e.g., the two rules coincide when the matrix $B$ is the identity), but they are still quite distinct. As we see later ours has a nice advantage as it has applications to the analysis of the Simplex method. Similarly to [8], we show that the optimum of problem (iii) can be computed by solving an auxiliary LP whose size is polynomial in the size of the description of the polyhedron $P$. This is somewhat unsatisfactory of course, as one wishes to avoid solving a bigger LP to solve an LP, and it begs for an improvement. However, the polyhedral description of the set of circuits will allow us to derive some nice properties when applying circuit-augmentation algorithms on 0/1 polytopes.

**Hardness implications.** Interestingly, the proof of the first part of Theorem 1 can be used to derive additional hardness results for the computation of the shortest monotone path from an initial extreme point solution of an LP to an optimal one. To explain the result more formally, let us recall some more definitions. The 1-skeleton of $P$ is the graph given by the 0-dimensional faces (vertices) and 1-dimensional faces (edges) of $P$. Given an objective function to minimize on $P$, 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 finds an optimal solution of an LP from an initial one via monotone paths. A legitimate question is the following: can we hope to find a pivot rule that makes the Simplex method use shortest monotone paths? We show in Section 3.3 that the answer to this question is negative, unless $P$=NP.

**Theorem 2.** Given a feasible extreme point solution of an LP, 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 2.

Note that the above result is orthogonal to the NP-hardness results on the computation of the diameter of a polytope [16, 27]. In fact, the hardness results in [16] and [27] rely on the existence/non existence of vertices with a certain structure, and do not provide a specific objective function to minimize over their polytopes.

---

1 We note that the hardness results and subsequent implications can be also derived if instead of the bipartite matching polytope one considers the circulation polytope [3], where the characterization of circuits, and the correspondent hardness reduction, become easier. However, we believe that the characterization of the circuits of the fractional matching polytope is of independent interest.
We can also prove that the result of Theorem 2 holds for monotone circuit-paths, i.e., paths constructed by a circuit-augmentation algorithm (see Corollary 4). These hardness results show that:

**Corollary 1.** For any efficiently-computable pivoting rule, an augmentation algorithm that is only allowed to move along edge-directions (like the Simplex method) cannot reach an optimal solution via a minimum number of augmentations, unless \( P=NP \). This stays true even if the set of directions is enlarged to be the set of all circuits.

**Approximation.** We next note that any polynomial-time \( \gamma \)-approximation algorithm for the pivot rule optimization problems yields an increase of at most a \( \gamma \)-factor on the running time of the corresponding circuit-augmentation algorithm – this follows from an easy extension of the analysis given by [10]. This observation turns out to be quite powerful in combination with the greatest-improvement pivot rule, and it plays a key role in our subsequent results. We therefore formally state its main implication in the next lemma. There \( m_B \) refers to the number of rows of the matrix \( B \), and \( \delta \) is the maximum lowest common multiple of the determinants of any two \( n \times n \) submatrices of \( (A)^T \).

**Lemma 1.** Consider an LP of the form (1). Let \( x_0 \) be an initial feasible solution, and let \( \gamma \geq 1 \). Using a \( \gamma \)-approximate greatest-improvement pivot rule, we can reach an optimal solution \( x_{\min} \) of (1) with no more than \( 2m_B \gamma \log (\delta c^T (x_0 - x_{\min})) + n \) augmentations.

We then prove that a circuit computed according to a steepest-descent pivot rule yields a \( \gamma \)-approximate solution to the greatest-improvement pivot rule, for a factor \( \gamma \) that depends on the 1-norm distance between some points of \( P \). Combining this with Lemma 1 we give a new bound on the number of steps performed by a circuit-augmentation algorithm that uses a steepest-descent pivot rule. In fact, the authors of [10], and later [8], gave bounds on the number of steps, which depended on the size of \( C(A,B) \) and the number of different values the objective function takes on that set. Here we get another type of bound of independent application:

**Theorem 3.** Let \( \omega_1 \) denote the minimum 1-norm distance from any extreme point \( v \) of \( P \) to any facet \( F \) of \( P \) such that \( v \notin F \). Let \( M_1 \) be the maximum 1-norm distance between any pair of extreme points of \( P \). Using a steepest-descent pivot rule, a circuit-augmentation algorithm reaches an optimal solution \( x_{\min} \) of (1) from any initial feasible solution \( x_0 \), performing

\[
O \left( nm_B \frac{M_1}{\omega_1} \log (\delta c^T (x_0 - x_{\min})) \right)
\]

augmentations.

**Implications on 0/1 polytopes.** The bound of Theorem 3 turns out to be polynomial for interesting classes of LPs, such as 0/1-LPs (i.e., LPs whose feasible region is a 0/1 polytope). Furthermore, it can be used to derive results on the performance of the Simplex method. We discuss these things next.

First, 0/1 polytopes are particularly important in optimization. Thus it is relevant to understand the performance of augmentation algorithms, and in particular of the Simplex method, in that family of polytopes [24].

Among others, the authors of [22] studied the number of distinct feasible solutions generated by the Simplex method, and proved that this number is strongly-polynomial for some 0/1-LPs of special format. With our approach, we can prove a weaker statement but that is valid for all 0/1-LPs.
Specifically, we first show that given a vertex $x$ of a 0/1 polytope $P$, there always exists an optimal solution to the steepest-descent pivot rule that is an edge-direction at $x$ (Lemma 5). Let us call such an optimal solution a steepest improving edge at $x$. Combining this with Theorem 3, we get the following.

**Theorem 4.** Given a problem of the form (1) whose feasible region $P$ is a 0/1 polytope, a circuit-augmentation algorithm with a steepest-descent 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 $P$.

An important remark is the following: The fact that paths of polynomial length on the 1-skeleton of 0/1 polytopes can be constructed from an augmentation oracle that outputs an improving edge-direction is already known [28]. What is not known is whether such paths can be realized via a simple pivot rule for the Simplex method. Our result tries to shed some new light on this matter. Indeed, the path constructed by our algorithm is the same path that the Simplex method would follow if implemented with a pivot rule that makes it move to an adjacent extreme point via a steepest improving edge. Note that this does not necessarily correspond to moving to an adjacent basis, because of degeneracy. While computing such an adjacent extreme point can be done in polynomial time (as explained in the proof of Lemma 5), translating this into a pivot rule for the Simplex method that is able to bypass degeneracy easily, remains an open question. Of course, if the polytope is non-degenerate then the solution is immediate. Therefore, a trivial consequence of the above theorem is:

**Corollary 2.** Given a problem of the form (1) whose feasible region $P$ is a non-degenerate 0/1 polytope, the Simplex method (with a steepest-descent pivot rule) reaches an optimal solution in strongly-polynomial time.

## 2 Preliminaries

We consider $P$ to be a polyhedron of the form $P = \{ x \in \mathbb{R}^n : Ax = b, Bx \leq d \}$ for integer matrices $A$ and $B$ of sizes $m_A \times n$ and $m_B \times n$ respectively, and integer vectors $b$ and $d$, and assume that we wish to minimize a linear objective function $c^T x$ over $P$. We further assume that $A$ has full row rank and that the rank of $\left( \begin{smallmatrix} A \\ B \end{smallmatrix} \right)$ is $n$. For our purposes, we also assume $m_B \geq 1$ as otherwise $P$ 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 $\text{rk}(D)$ denote its rank, and $\text{det}(D)$ denote its determinant. For a vector $x$, we let supp($x$) be the support of the vector $x$.

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

A circuit-path is a finite sequence of feasible solutions $x_1, x_2, \ldots, x_q$ satisfying $x_{i+1} = x_i + \alpha_i g_i$, where $g_i \in \mathcal{C}(A, B)$ and $\alpha_i \in \mathbb{R}_{>0}$ is such that $x_i + \alpha_i g_i \in P$ but $x_i + (\alpha_i + \varepsilon) g_i \notin P$ for all $\varepsilon > 0$ (i.e., the augmentation is maximal). Note that $x_i$ is not necessarily a vertex of $P$. A circuit-path is called monotone if each $g_i$ satisfies $c^T g_i < 0$ (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 $g$ 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 pivot rule will be called \( \text{Great}(\mathcal{P}, \mathbf{x}, \mathbf{c}) \), and is as follows:

\[
\begin{align*}
\max & \quad -\mathbf{c}^\top (\alpha \mathbf{g}) \\
\text{s.t.} & \quad \mathbf{g} \in \mathcal{C}(A, B), \\
& \quad \alpha > 0, \\
& \quad \mathbf{x} + \alpha \mathbf{g} \in \mathcal{P}.
\end{align*}
\]

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

\[
\begin{align*}
\max & \quad -\mathbf{c}^\top \mathbf{g} \\
\text{s.t.} & \quad \mathbf{g} \in \mathcal{C}(A, B), \\
& \quad \mathbf{x} + \epsilon \mathbf{g} \in \mathcal{P} \quad \text{for some } \epsilon > 0.
\end{align*}
\]

The optimization problem that arises when following the steepest-descent pivot rule\(^2\) will be called \( \text{Steep}(\mathcal{P}, \mathbf{x}, \mathbf{c}) \), and is as follows:

\[
\begin{align*}
\max & \quad -\frac{\mathbf{c}^\top \mathbf{g}}{||\mathbf{g}||_1} \\
\text{s.t.} & \quad \mathbf{g} \in \mathcal{C}(A, B), \\
& \quad \mathbf{x} + \epsilon \mathbf{g} \in \mathcal{P} \quad \text{for some } \epsilon > 0.
\end{align*}
\]

A maximal augmentation given by an optimal solution to \( \text{Great}(\mathcal{P}, \mathbf{x}, \mathbf{c}) \) is called a greatest-improvement augmentation. A Dantzig augmentation and a steepest-descent augmentation are defined similarly. In this work, we will only use maximal augmentations, and therefore will omit the word “maximal”.

## 3 Hardness of some Circuit-Pivot Rules

### 3.1 Key Tool: 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 \( \mathbf{x} \in \mathbb{R}^{E(G)} \), we let \( \mathbf{x}(X) \) denote \( \sum_{e \in X} \mathbf{x}(e) \).

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

\[
\begin{align*}
\mathbf{x} (\delta_G(v)) & \leq 1 \quad \text{for all } v \in V(G) \setminus L(G) \\
x & \geq 0.
\end{align*}
\]  

\(^2\)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. We here stick to this name as used previously in \([13, 8]\).
In this section, we fully characterize the circuits of $P_{\text{FMAT}}(G)$. We will prove that, if $x$ is a circuit of $P_{\text{FMAT}}(G)$, then $\text{supp}(x)$ induces a connected subgraph of $G$ that has a very special structure: namely, it belongs to one of the five classes of graphs ($E_1, E_2, E_3, E_4, E_5$) listed below.

(i) Let $E_1$ denote the set of all subgraphs $F \subseteq G$ such that $F$ is an even cycle.
(ii) Let $E_2$ denote the set of all subgraphs $F \subseteq G$ such that $F$ is an odd cycle.
(iii) Let $E_3$ denote the set of all subgraphs $F \subseteq G$ such that $F$ is a simple path.
(iv) Let $E_4$ 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).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig1.png}
\caption{An Example of a subgraph belonging to $E_4$}
\end{figure}

(v) Let $E_5$ 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 3); if $P$ is empty then $C_1$ and $C_2$ intersect only at one node $v$ (see Figure 2).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2.png}
\caption{An Example of a subgraph belonging to $E_5$ where $P$ is non-empty.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig3.png}
\caption{An Example of a subgraph belonging to $E_5$ where $P$ is empty.}
\end{figure}

We will associate a set of circuits to the subgraphs in the above families, by defining the following 5 sets of vectors. It is worth noticing that similar elementary moves appeared in [11] in applications of Gröbner bases in combinatorial optimization.
First we note that \( C \) and let \( \text{supp}(B) \) can treat \( \text{supp}(g) \).

**Lemma 2.** \( \mathcal{C}(\mathcal{P}_{\text{FMAT}}(G)) = C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5. \)

**Proof.** It is known that the vectors of \( C_1 \cup \cdots \cup C_5 \) correspond to edge-directions of \( \mathcal{P}_{\text{FMAT}}(G) \) (see e.g. [4, 27]), so it remains to be shown that all circuits belong to one of these sets. Let \( B \) denote the constraint matrix corresponding to the 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 \( \text{supp}(Bx) \) and \( \text{supp}(x) \) as a subset of \( V(G) \) or \( E(G) \), respectively. Let \( g \in \mathcal{C}(\mathcal{P}_{\text{FMAT}}(G)) \), and let \( G(g) \) be the subgraph of \( G \) induced by the edges in \( \text{supp}(g) \).

First we note that \( G(g) \) is connected. Otherwise, restricting \( g \) to the edges of any component of \( G(g) \) gives a vector \( f \) with \( \text{supp}(Bf) \subseteq \text{supp}(Bg) \) and \( \text{supp}(f) \nsubseteq \text{supp}(g) \), contradicting that \( g \) is a circuit.

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.

**Figure 4:** Example of a vector \( g \in C_5 \). Each edge \( e \) is labeled with \( g_e \).
Now, suppose that $G(g)$ contains no cycles. Let $P$ be any edge-maximal path in $G(g)$, with endpoints $u$ and $w$. Note that $\text{supp}(Bg) \subseteq \{u, w\} \setminus L(G)$. Let $f \in \{-1, 1\}^{E(G)}$ be a vector that satisfies (i) $f_e \neq 0$ iff $e \in E(P)$, and (ii) $f(\delta_P(v)) = 0$ $\forall v \neq u, w$. Note that $f \in C_3$. Then, $\text{supp}(Bf) = \{u, w\} \setminus L(G) \subseteq \text{supp}(Bg)$, and $\text{supp}(f) \subseteq \text{supp}(g)$. Therefore, it must be that the edges of $G(g)$ are exactly $E(P)$, and $g(\delta_P(v)) = 0$ for all $v \in V(G) \setminus \{u, w\}$. Thus, $g = f$ or $g = -f$. In any case, $g \in C_3$.

Now, suppose that $G(g)$ contains an even cycle $C$. Let $f \in \{-1, 1\}^{E(G)}$ be a vector that satisfies (i) $f_e \neq 0$ iff $e \in E(C)$, and (ii) $f(\delta_C(v)) = 0$ $\forall v \in V(C)$. Note that $f \in C_1$. Then, $\text{supp}(Bf) = \emptyset \subseteq \text{supp}(Bg)$, and $\text{supp}(f) \subseteq \text{supp}(g)$. Therefore, it must be that the edges of $G(g)$ are exactly $E(C)$, and $g(\delta_C(v)) = 0$ for all $v \in V(G)$. Thus, $g = f$ or $g = -f$. In any case, $g \in C_1$.

We are left with the case where $G(g)$ 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 $G(g)$.

**Claim 1.** Any two odd cycles in $G(g)$ share at most one node.

**Proof.** Let $C, D \subseteq G(g)$ 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. □

Suppose that $G(g)$ contains at least two distinct odd cycles $C_1$ and $C_2$. Since $G(g)$ is connected, then either these two cycles share a node or there exists a simple path $P$ in $G(g)$ 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 $f \in \mathbb{Z}^{E(G)}$ be a vector that satisfies (i) $f_e \neq 0$ iff $e \in E(F)$, (ii) $f(\delta_F(v)) = 0$ $\forall v \in V(F)$, (iii) $f_e \in \{-1, 1\}$ for all $e \in E(C_1 \cup C_2)$, and (iv) $f_e \in \{-2, 2\}$ for all $e \in E(P)$. Note that $f \in C_5$. Then $\text{supp}(Bf) = \emptyset \subseteq \text{supp}(Bg)$, and $\text{supp}(f) \subseteq \text{supp}(g)$. Therefore, it must be that the edges of $G(g)$ are exactly $E(F)$, and $g(\delta_{C_i}(v)) = 0$ for all $v \in V(G)$. Thus, $g \in C_5$.

Finally, suppose that $G(g)$ contains exactly one odd cycle $C$. If there exists a node $w \in V(C)$ such that $g(\delta_C(w)) \neq 0$, then let $f \in \{-1, 1\}^{E(G)}$ be a vector that satisfies (i) $f_e \neq 0$ iff $e \in E(C)$, and (ii) $f(\delta_C(v)) = 0$ $\forall v \in V(C) \setminus \{w\}$. Note that $f \in C_2$. Then, $\text{supp}(Bf) = \{w\} \subseteq \text{supp}(Bg)$, and $\text{supp}(f) \subseteq \text{supp}(g)$. Therefore, it must be that the edges of $G(g)$ are exactly $E(C)$, and $g(\delta_C(v)) = 0$ for all $v \in V(G) \setminus \{w\}$. Thus, it must be that $g \in C_2$.

We are left with the case where $g(\delta_C(v)) = 0$ for all $v \in V(C)$. Note that this is not possible if $\text{supp}(g) = E(C)$, because $C$ is an odd cycle. Then let $P$ be any simple path in $G(g)$ 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 $f \in \mathbb{Z}^{E(G)}$ be a vector that satisfies (i) $f_e \neq 0$ iff $e \in E(F)$, (ii) $f(\delta_F(v)) = 0$ $\forall v \in V(F) \setminus \{w\}$, (iii) $f_e \in \{-1, 1\}$ for all $e \in E(C)$, and (iv) $f_e \in \{-2, 2\}$ for all $e \in E(P)$. Note that $f \in C_4$. Then $\text{supp}(Bf) = \{w\} \setminus L(G) \subseteq \text{supp}(Bg)$, and $\text{supp}(f) \subseteq \text{supp}(g)$. Therefore, it must be that the edges of $G(g)$ are exactly $E(F)$, and $g(\delta_F(v)) = 0$ for all $v \in V(G) \setminus \{w\}$. Thus, it must be that $g \in C_4$.

In all the above cases, $g \in C_1 \cup \cdots \cup C_5$, as desired. □
3.2 Hardness Reduction

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

**Theorem 5.** Solving the optimization problem $\text{Dan}(\mathcal{P}, x, c)$ 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 $c$, 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 $\text{Dan}(\mathcal{P}_{\text{FMAT}}(H), x^M, c)$ (where $x^M$ 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 $v_av_b \in E$. For all arcs $uv \in F$, with $u, v \neq t$, we add an edge $u bv_a \in E$. 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 $u bt \in E$. 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 $H$.](image)

Now we define the cost function $c$. We set $c(v_av_b) = 0$ for all $v \in N \setminus \{t\}$, $c(s's_a) = -W = -c(tt')$ such that $W \in \mathbb{Z}$, $W \gg |E|$, and let all other edges have cost -1. Finally, we let $M = \{v_av_b : v \in N \setminus \{t\}\} \cup \{tt'\}$ 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 $g$ to $\text{Dan}(\mathcal{P}_{\text{FMAT}}(H), x^M, c)$ with objective function value at least $-c^tg = 2W + n - 1$.

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

$$P' = (s's_a, s_as_b, s_bv^1_a, v^1_av^1_b, v^1_bv^2_a, v^2_av^2_b, \ldots, v^{k-1}_av^{k-1}_b, v^{k-1}_bv^k_a, v^k_av^k_b, v^k_btt').$$

Let $g$ be defined as

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

Then $g \in \mathcal{C}_3$, and is therefore a circuit of $\mathcal{P}_{\text{FMAT}}(H)$. Note that $x^M + g \in \mathcal{P}_{\text{FMAT}}(H)$, and $-c^tg = 2W + n - 1$. Thus, $g$ is a feasible solution to $\text{Dan}(\mathcal{P}_{\text{FMAT}}(H), x^M, c)$ with the claimed objective function value.
Now suppose that there is a solution \( g \) to \( \text{Dan}(\mathcal{P}_{\text{FMAT}}(H), \chi^M, c) \), with objective function value at least \( 2W + n - 1 \). First, we argue that the support of \( g \) is indeed an \( M \)-alternating path with endpoints \( s' \) and \( t' \).

Note that, by construction, \( H \) is bipartite, so \( g \in C_1 \cup C_3 \). In either case, \( g \in \{1, 0, -1\}^E \). By our choice of \( W \), since \( -c^t g \geq 2W + n - 1 \), it must be that \( g(s's_a) = 1 \) and \( g(tt') = -1 \). Then, since \( s' \) and \( t' \) are not in any cycles of \( H \), necessarily \( g \in C_3 \) and its support is an \( s', t' \)-path. It follows that \( g \) has at most \(|V|-1\) non-zero entries. Two of the non-zero entries are \( g(s's_a) \) and \( g(tt') \), and of those that remain, exactly half have value 1. Thus,

\[
-c^t g \leq 2W + \frac{1}{2}(|V| - 3) = 2W + \frac{1}{2}(2n + 1) - 3 = 2W + n - 1.
\]

It is clear that the above inequality holds tight only if \( g(e) = 1 \) for \( \frac{1}{2}(|V| - 3) \) edges of \( E \setminus \{s's_a, tt'\} \), all of which have \( c(e) = -1 \), and \( c(f) = 0 \) for all edges \( f \) such that \( g(f) = -1 \). Since the number of edges \( e \) with \( g(e) = 1 \) equals the number of edges \( f \) with \( g(f) = -1 \), we have that \(|\text{supp}(g)| = |V| - 1 \), and therefore \( \text{supp}(g) \) 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 \((v_av_b)\) (for \( v \in N \)), we obtain a path that naturally corresponds to a directed Hamiltonian \( s', t' \)-path in \( D \).

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

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

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

\[
\max\{-c^t(\alpha y) : \chi^M + \alpha y \in \mathcal{P}_{\text{FMAT}}(H), \alpha > 0\} = -c^t y.
\]

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

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 \( \mathcal{P}_{\text{FMAT}}(H) \) is integral.

### 3.3 Hardness implications

Here we observe that the reductions in the previous section have interesting hardness implications. As a first corollary, we get that given a vertex \( x \) of \( \mathcal{P}_{\text{FMAT}}(H) \), computing the best neighbor extreme point of \( x \) is NP-hard. Here the best neighbor extreme point is an extreme point that minimizes the objective function value among all vertices \( \bar{x} \) that are adjacent to \( x \).

**Corollary 3.** Given a feasible extreme point solution of a 0/1 polytope, 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 \( \text{Great}(\mathcal{P}_{\text{FMAT}}(H), \chi^M, c) \) is a circuit \( g \) that corresponds to an edge-direction at \( \chi^M \). As a consequence, if we consider the LP obtained by minimizing \( c^t x \) over \( \mathcal{P}_{\text{FMAT}}(H) \), and take \( \chi^M \) 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.
In addition, we can now prove Theorem 2 that we restate for convenience.

**Theorem 2** Given a feasible extreme point solution of an LP, 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 2.

*Proof.* Once again, consider the hardness reduction used in the proof of Theorem 6 and the LP obtained by minimizing $c^T x$ over $\mathcal{P}_\text{FMAT}(H)$. 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 $\chi^M$ 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 $\chi^M$ 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 $\mathcal{P}_\text{FMAT}(H)$, 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.

As mentioned in the introduction, our result implies that for any efficiently-computable pivot- ing rule, the Simplex method cannot 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 [13, 12, 1].

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

**Corollary 4.** Given an initial feasible solution of an LP, 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 2.

### 3.4 Efficiency of steepest-descent

Despite the news so far, we can at least guarantee that the steepest-edge pivot rule can be computed in polynomial time. Our idea follows the work of Borgwardt and Viss [8] (see also [17]).

The steepest-descent pivot rule presented in [10] is stated only for LPs in standard equality form. While in this paper we extend their definition to general form LPs directly, the authors of [8] instead give an alternate, generalized definition of the steepest-descent pivot rule in order to extend it to general form LPs. One of the main results of their paper is to show that their generalized steepest-descent pivot rule can be computed in polynomial time. They accomplish this by showing that, given an LP of the form (11), the circuits of $\mathcal{P}$ satisfying $\|B g\|_1 = 1$ appear as vertices in the polytope

$$\mathcal{P}_{A,B} = \{(x, y^+, y^-) \in \mathbb{R}^{n+2m} : Ax = 0, Bx = y^+ - y^-, \|y^+\|_1 + \|y^-\|_1 = 1, y^+, y^- \geq 0\}.$$ 

This polyhedral model allows for the efficient computation of a close relative of our steepest-descent pivot rule. It can be computed by simply solving a linear program over $\mathcal{P}_{A,B}$ intersected with the feasible cone at a current vertex.

In this section, we mirror their technique to show that the definition of steepest-descent considered in this paper can also be computed in polynomial time in a similar way. This requires only a very slight modification to the polyhedral model and proofs presented in [8]. For the sake of completeness, we give a proof in the Appendix.

**Theorem 7.** The optimization problem $\text{Steep}(\mathcal{P}, x, c)$ can be solved in polynomial time.
Proof sketch. Consider the polytope $P_{A,B} = \{(x^+, x^-, y^+, y^-) \in \mathbb{R}^{2n+2m_B} : A(x^+ - x^-) = 0, Bx = y^+ - y^-, ||x^+||_1 + ||x^-||_1 = 1, x^+, x^- \geq 0, y^+, y^- \geq 0\}$. If we ignore the equality constraint $||x^+||_1 + ||x^-||_1 = 1$, we obtain a cone whose extreme rays are minimal support solutions of the above system. Thus, these rays include all the circuits of the system (1). By the specific normalization we took, the optimization problem $Steep(P, x, c)$ can be solved via a linear program over $P_{A,B}$ intersected with the feasible cone at a current vertex. □

We remark that while the above result does give a guarantee on the efficiency of computing a steepest-descent pivot rule, it is not effective in practice, for our purposes. In particular, if we wish to solve an LP via an augmentation procedure based on a steepest-descent pivot rule, using this method would require solving a different larger LP at each step of the augmentation procedure.

Theorem 5, Theorem 6 and Theorem 7 yield a proof of Theorem 1.

4 Approximation of Circuit-Pivot Rules

The main goal of this section is to prove Lemma 1 (in Section 4.1) and Theorem 3 (in Section 4.2). We start with the following formal definition of approximate greatest-improvement augmentations.

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

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

As mentioned in the introduction, we define

$$\delta := \max \left\{ \text{lcm} \left( |\det \left( \begin{array}{c} A \\ D_1 \end{array} \right)|, |\det \left( \begin{array}{c} A \\ D_2 \end{array} \right)| \right) \right\},$$

where $\text{lcm}(a, b)$ denotes the least common multiple of $a$ and $b$, and where the max is taken over all pairs of $n \times n$ submatrices $(A_D)$ of $(A_B)$ such that $(A_D)$ has rank $n$.

Furthermore, we let $\omega_1$ be the minimum 1-norm distance from any extreme point to any facet not containing it. Formally, let $\text{vert}(\mathcal{P})$ be the set of vertices of $\mathcal{P}$. For a given $v \in \text{vert}(\mathcal{P})$, let $\mathcal{F}(v)$ be the set of feasible points of $\mathcal{P}$ that lie on any facet $F$ of $\mathcal{P}$ with $v \notin F$.

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

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

$$M_1 := \max_{v_1, v_2 \in \text{vert}(\mathcal{P})} \|v_1 - v_2\|_1.$$

4.1 Approximate greatest-improvement augmentations

Let us recall the statement of Lemma 1.

Lemma 1. Consider an LP of the form (1). Let $x_0$ be an initial feasible solution, and let $\gamma \geq 1$. Using a $\gamma$-approximate greatest-improvement pivot rule, we can reach an optimal solution $x_{\min}$ of (1) with no more than $2m_B \gamma \log \left( \delta c^T (x_0 - x_{\min}) \right) + n$ augmentations.
The proof of Lemma 1 closely mimicks the arguments used in [10]. However, since the authors of [10] 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 [8]).

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

**Proposition 1** (see Proposition 1.4 in [14]). Let $v \in \ker(A) \setminus \{0\}$. Then we can express $v$ as $v = \sum_{i=1}^{k} \alpha_i g^i$ such that for all $1 \leq i \leq k$

- $g^i \in \mathcal{C}(A, B)$,
- $g^i$ and $v$ are sign-compatible with respect to $B$ and $\text{supp}(Bg^i) \subseteq \text{supp}(Bv)$,
- $\alpha_i \in \mathbb{R}_{\geq 0}$,
- and $k \leq m_B$.

Let $x_{\text{max}}$ 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 [10] based on well-known estimates of [2]:

**Lemma 3** (see Lemma 1 in [10]). Let $\epsilon > 0$ be given. Let $c$ be an integer vector. Define $f_{\text{min}} := c^\top x_{\text{min}}$, $f_{\text{max}} := c^\top x_{\text{max}}$. Suppose that $f^k = c^\top x_k$ 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 $k$,

$$(f^k - f^{k+1}) \geq \beta (f^k - f_{\text{min}}).$$

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

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

**Lemma 4.** Let $\tilde{x}$ 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 $c^\top \hat{x} \leq c^\top \tilde{x}$.

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

We can iterate this process, and note that the number of linearly independent rows of $(B^\top_{B^\top})$ increases by one at each step. Therefore, after at most $n - \text{rk}(A)$ 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 $\epsilon$. By the sign-compatible representation property of circuits though, at any non-optimal point $\tilde{x}$, there always exists an improving direction that is a circuit.

We can now give a proof of Lemma 1.
Proof of Lemma 7. By the sign-compatible representation property of the circuits,

\[ x_{\text{min}} - x_k = \sum_{i=1}^{p} \alpha_i g^i \]

where \( g^i \in \mathcal{C}(A, B) \) and \( p \leq m_B \).

We then have

\[ 0 > c^T(x_{\text{min}} - x_k) = c^T \sum_{i=1}^{p} \alpha_i g^i = \sum_{i=1}^{p} \alpha_i c^T g^i \geq -m_B \Delta, \]

where \( \Delta > 0 \) is the largest value of \( -\alpha c^T z \) over all \( z \in \mathcal{C}(A, B) \) and \( \alpha > 0 \) for which \( x_k + \alpha z \) is feasible. Equivalently, we get

\[ \Delta \geq \frac{c^T(x_k - x_{\text{min}})}{m_B}. \]

Now let \( \alpha z \) be a \( \gamma \)-approximate greatest-improvement augmentation applied to \( x_k \), leading to \( x_{k+1} := x_k + \alpha z \). Since \( -\alpha c^T z \geq \frac{1}{\gamma} \Delta \), we get

\[ c^T(x_k - x_{k+1}) = -\alpha c^T z \geq \frac{1}{\gamma} \Delta \geq \frac{c^T(x_k - x_{\text{min}})}{\gamma m_B}. \]

Thus, we have at least a factor of \( \beta = \frac{1}{\gamma m_B} \) of objective function value decrease at each augmentation. Applying Lemma 3 with \( \epsilon = 1/\delta \) then yields a solution \( \bar{x} \) with \( c^T(\bar{x} - x_{\text{min}}) < 1/\delta \), obtained within \( 2m_B \gamma \log(\delta c^T(x_0 - x_{\text{min}})) \) augmentations.

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

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

\[ \begin{pmatrix} A \\ B_{T_1} \end{pmatrix} x = \begin{pmatrix} b \\ d_{T_1} \end{pmatrix}, \]

and \( x_{\text{min}} \) is the unique solution to

\[ \begin{pmatrix} A \\ B_{T_2} \end{pmatrix} x = \begin{pmatrix} b \\ d_{T_2} \end{pmatrix}. \]

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 \( x' \) are integer multiples of \( \frac{1}{\delta_1} \) and the entries of \( x_{\text{min}} \) are integer multiples of \( \frac{1}{\delta_2} \). Then by letting \( \delta = \text{lcm}(\delta_1, \delta_2) \), we have that the entries of \( (x' - x_{\text{min}}) \) are integer multiples of \( \frac{1}{\delta} \). Since \( c \) is an integer vector, we have that \( c^T(x' - x_{\text{min}}) \geq \frac{1}{\delta} \), and by the definition of \( \delta \), we have that \( \frac{1}{\delta} \geq \frac{1}{\gamma} \). This is a contradiction to the fact that \( c^T(x' - x_{\text{min}}) < 1/\delta \).

Note that the above proof also establishes that the result obtained by [10] regarding the number greatest-improvement augmentations needed to solve an equality form LP extends to the general-form LP (trivially by taking \( \gamma = 1 \)).

A similar approximation argument can be used to recover the bound obtained in [10] on the number of Dantzig-descent augmentations needed to solve an LP. In fact, the proof they give can be interpreted (with slight modification) as showing that at an extreme point, under the coprime integer scaling of circuits, a Dantzig-descent augmentation is an approximate greatest-improvement augmentation. Since this is less relevant for our subsequent results, we omit the details.
4.2 Steepest-descent 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 an LP. We restate Theorem 3 for convenience.

Theorem 3. Let $\omega_1$ denote the minimum 1-norm distance from any extreme point $v \in P$ to any facet $F$ of $P$ such that $v \notin F$. Let $M_1$ be the maximum 1-norm distance between any pair of extreme points of $P$. Using a steepest-descent pivot rule, we can reach an optimal solution $x_{\text{min}}$ of (1) from any initial feasible solution $x_0$, performing

$$O \left( nm_B \frac{M_1}{\omega_1} \log \left( \delta c^T(x_0 - x_{\text{min}}) \right) \right)$$

augmentations.

Proof. First, we can apply Lemma 4 to move from $x_0$ to an extreme point $x'$ of the LP in at most $n$ steps.

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

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

Since the augmentation $\alpha z$ is maximal, we have that at the point $x' + \alpha z$, there exists some facet of our feasible region which contains $x' + \alpha z$ but not $x'$. Then $\omega_1 \leq \|(x' + \alpha z) - x'\|_1 = \alpha \|z\|_1$. Since $\|z\|_1 = 1$, it follows that $\alpha \geq \omega_1$. Since $x' + \alpha^* z^*$ is feasible, we have that $\|(x' + \alpha^* z^*) - x'\|_1$ is at least the maximum 1-norm distance from $x'$ to any other feasible point. As above, it follows that $\alpha^* \leq M_1$.

Given these bounds on $\alpha$ and $\alpha^*$, it follows that

$$-\alpha c^T z \geq \left( \frac{\omega_1}{M_1} \right) (-\alpha^* c^T z^*).$$

Now let $\bar{x} = x' + \alpha z$. By Lemma 4, an extreme point solution $\bar{x}$ can be found from $\bar{x}$ in at most $n - 1$ additional augmentations (e.g., using again steepest-descent augmentations, but on a sequence of face-restricted LPs) with $c^T \bar{x} \leq c^T \bar{x}$. Then we have that $\bar{x} = \bar{x}^*$ is an $(\frac{\omega_1}{M_1})$-approximate greatest-improvement augmentation at $x'$, and since $\bar{x}$ is also an extreme point, we can continue to apply this procedure. Since it takes at most $n$ steepest-descent augmentations to find such an $(\frac{\omega_1}{M_1})$-approximate greatest-improvement augmentation, it follows from Lemma 1 that from an initial solution $x_0$, we can reach $x_{\text{min}}$ in $O \left( m_B n M_1 \omega_1 \log \left( \delta c^T(x_0 - x_{\text{min}}) \right) \right)$ steepest-descent augmentations.

The example given by Figure 6 shows that the approximation factor of $\frac{\omega_1}{M_1}$ used in the above proof is tight.
Figure 6: An example where the objective function improvement of taking a steepest-descent direction at $x$ is $\frac{M_1}{\omega_1}$ times the objective function improvement of taking a greatest-improvement direction. This polygon has vertices $x = (0, 0), y_1 = (0, 1), y_2 = (2, 2), (2, 1), \text{and} (1, 0)$. One can check that at $x$, $y_1$ is a steepest-descent augmentation, $y_2$ is a greatest-improvement augmentation, and $c^T y_1 = \frac{1}{4} c^T y_2 = \frac{\omega_2}{M_1} c^T y_2$.

5 Implications for 0/1 Polytopes

In this section, we consider the implications that Theorem 3 has in the case of 0/1-polytopes. In particular, we will prove Theorem 4, and hence Corollary 2. We start with the following lemma:

**Lemma 5.** Consider a problem of the form (1) whose feasible region $P$ is a 0/1 polytope, and let $x$ be a non-optimal vertex of $P$. Then, the optimal solution to Steep$(P, x, c)$ corresponds to an edge-direction at $x$.

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

$$\begin{align*}
\max & -c^T z \\
\text{s.t.} & \|z\|_1 \leq 1 \\
& x + \varepsilon z \in P \\ & \text{for some } \varepsilon > 0
\end{align*}$$

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

Let $P_Q$ denote the feasible region of $Q$. Note that $P_Q$ is the feasible cone at $x$ in $P$—given by the constraint (3)—intersected with an $n$-dimensional orthoplex (or cross-polytope)—given by the constraint (4). The constraint (4) can be modeled using the linear constraints

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

It follows that $P_Q$ is a polytope, and therefore $Q$ is a feasible bounded LP. There exists an optimal vertex $y$ of $P_Q$ which is determined by $n$ linearly independent constraints of $P_Q$.

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

17
As a consequence of this, we have that the optimal solution of \( Q \) corresponds to an edge-direction of \( P \) incident with \( x \). It follows that the optimal solution of \( \text{Steep}(P, x, c) \) is an edge-direction of \( P \) incident with \( x \).

We will rely on the following result of Frank and Tardos [15].

**Lemma 6** ([15]). Let \( w \in \mathbb{R}^n \) be a rational vector, and \( \alpha \) be a positive integer. Define \( N := (n + 1)!2^{m_3} + 1 \). Then one can compute an integral vector \( w' \in \mathbb{Z}^n \) satisfying:

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

(b) Consider any rational LP of the form \( \max \{ w^T x : A'x \leq b', x \in \mathbb{R}^n \} \), where the encoding length\(^3\) of any entry of \( A' \) is at most \( \alpha \). Then, \( x \in \mathbb{R}^n \) is an optimal solution to that LP if and only if it is an optimal solution to \( \max \{ w'^T x : A'x \leq b', x \in \mathbb{R}^n \} \).

We are now ready to prove Theorem 4, which we restate for convenience.

**Theorem 4**. Given a problem of the form (1) whose feasible region \( P \) is a 0/1 polytope, a circuit-augmentation algorithm with a steepest-descent 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 \( P \).

*Proof.* Let us call \((LP_1)\) the given LP problem of the form (1) whose feasible region is \( P \). Since \( P \) 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 \( \leq \frac{n^{n/2}}{2} \) [31]. Apply Lemma 5 to the LP obtained from \((LP_1)\) after changing the objective function to \( \max w^T x \), with \( w := -c \). Set \( c' := -w' \). Finally, let \((LP_2) := \min \{ c'^T x : Ax = b, Bx \leq d, x \in \mathbb{R}^n \} \).

Let \( x_0 \) and \( x_{\min} \) be respectively the initial solution and the optimal solution. By performing at most \( n \) additional augmentations, we can assume \( x_0 \) is an extreme point.

First, we will show that Theorem 4 holds for \((LP_2)\). Then, we will show that a circuit-augmentation algorithm traverses the same edge-walk when solving \((LP_2)\) and \((LP_1)\), if one uses the steepest-edge pivot rule. This will prove the statement.

Recall that the steepest-descent pivot rule selects at each step an improving circuit \( g \) that minimizes \( \frac{c'^T g}{\|g\|} \). Since the feasible region of \((LP_2)\) 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 3 by

\[
\mathcal{O} \left( n! m_B \frac{M_1}{\omega_1} \log (\delta c'^T (x_0 - x_{\min})) \right).
\]

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

Since \( P \) is a 0/1 polytope, we immediately have that \( M_1 \leq n \). We now show that \( \omega_1 \geq 1 \). Let \( v \) be any extreme point of \( P \) and let \( F \) be any facet of \( P \) which does not contain \( v \). By reflecting and translating \( P \), we may assume without loss of generality that \( v = 0 \) (Note that these operations do not change the 1-norm distance between any pair of points in \( P \)). It therefore suffices to show that for any facet \( F \) not containing \( 0 \), \( \|y\|_1 \geq 1 \) for all \( y \in F \). Since all points in \( F \) have non-negative

\[^3\text{The encoding length of a rational number } \frac{p}{q} \text{ is defined as } [\log(p + 1)] + [\log(q + 1)] + 1\]
coordinates, the minimum value of $\|y\|_1$ over all $y \in F$ is equal to the optimal solution to the following LP:

$$\begin{align*}
\min & \quad 1^\top y \\
\text{s.t.} & \quad y \in F.
\end{align*}$$

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

Finally, we address the term $\log \left( c'(x_0 - x_{\min}) \right)$. Since $x_0$ and $x_{\min}$ are both in $\{0, 1\}^n$, we have that $\log \left( c'(x_0 - x_{\min}) \right) \leq \log(\|c'\|_1) \leq \log(n\|c'\|_\infty)$, which is polynomial in $n$ due to Lemma 6(a). Therefore, the number of augmentations required to solve $(LP2)$ is polynomial in $n$ and $m_B$, and hence 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 $\text{Steep}(\mathcal{P}, x, c)$, defined in Section 3. By Lemma 5, the edge-direction $g$ selected by our algorithm applied to $(LP2)$ is an optimal solution to the LP describing $\text{Steep}(\mathcal{P}, x, c')$. Note that the maximum absolute value of a matrix-coefficient of this LP is also at most $\frac{n^{n/2}}{2}$. Therefore, due to Lemma 6(b), $g$ is an optimal solution to $\text{Steep}(\mathcal{P}, x, c')$ if and only if it is an optimal solution to $\text{Steep}(\mathcal{P}, x, c)$. Therefore, the circuit-augmentation algorithm implemented according to the steepest-descent pivot rule, performs the exact same pivots for the objective functions $c'$ and $c$. \hfill \square

Acknowledgements. J.A. De Loera was partially supported by NSF grant DMS-1818969. S. Kafer and L. Sanità were supported by the NSERC Discovery Grant Program and an Early Researcher Award by the Province of Ontario.

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] F. Barahona and . Tardos. Note on Weintraubs minimum-cost circulation algorithm. SIAM Journal on Computing, 18(3):579–583, 1989.

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

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

[6] S. Borgwardt, J. A. De Loera, and E. Finhold. Edges vs circuits: a hierarchy of diameters in polyhedra. Advances in Geometry, 16, 09 2014.
[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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

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

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

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

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

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

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

[21] 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.

[22] 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.
Appendix

We here give a more detailed proof of Theorem 7, mirroring the results in [8]. We first prove a version of Theorem 3 from [8]. Recall that if \( g \) is a circuit of \( P \), then \( g = \alpha g' \) for some \( g' \in \mathcal{C}(A,B) \) and \( \alpha \in \mathbb{R} \).

**Theorem 8.** Given a pointed polyhedron \( P = \{ x \in \mathbb{R}^n : Ax = b, Bx \leq d \} \), the pointed cone

\[
C_{A,B} = \{ (x^+, x^-, y^+, y^-) \in \mathbb{R}^{2n+2m_B} : A(x^+ - x^-) = 0, \\
B(x^+ - x^-) = (y^+ - y^-), \\
x^+, x^-, y^+, y^- \geq 0 \}
\]

has extreme rays generated by the following sets:

1. The set

\[
S := \{ (g^+, g^-, y^+, y^-) : g^+ - g^= g \in \mathcal{C}(A,B), \\
g^+_i = \max \{ g_i, 0 \} , g^-_i = \max \{ -g_i, 0 \} , \\
y^+_i = \max \{ (Bg)_i, 0 \} , y^-_i = \max \{ -(Bg)_i, 0 \} \},
\]

which gives the circuits of \( P \).
2. A subset of $T_x := \{(x^+, x^-, 0, 0) : x^+ = x^- = 1 \text{ for some } i \leq n, x^+_j = x^-_j = 0 \text{ for } j \neq i\}$, which has size at most $n$.

3. A subset of $T_y := \{(0, 0, y^+, y^-) : y^+_i = y^-_i = 1 \text{ for some } i \leq m_B, y^+_j = y^-_j = 0 \text{ for } j \neq i\}$, which has size at most $m_B$.

**Proof.** Since $P$ is pointed, we have that $\text{rank}(A_B^t) = n$ and $C_{A,B}$ is a pointed cone. We consider a canonical representation $C_{A,B} = \{ r \in \mathbb{R}^{2n+2m_B} : Mr \geq 0 \}$, where

$$M = \begin{pmatrix} A & -A & 0 & 0 \\ -A & A & 0 & 0 \\ B & -B & -I & I \\ -B & B & I & -I \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \quad \text{and} \quad r = \begin{pmatrix} x^+ \\ x^- \\ y^+ \\ y^- \end{pmatrix}.$$

Given $r \in C_{A,B}$, let $Z(r)$ denote the set of row indices $i$ of $M$ such that $M_i r = 0$. As in [S], we rely on the following characterization of the extreme rays of a pointed cone: a point $r \in C_{A,B}$ is an extreme ray of $C_{A,B}$ if for all nonzero $r' \in C_{A,B}$ with $Z(r) \subseteq Z(r')$, $r'$ satisfies $r' = \alpha r$ for some $\alpha > 0$. For any $r \in C_{A,B}$, the only inequalities of the system $Mr \geq 0$ which may not be satisfied with equality are those corresponding to the constraints $x^+, x^-, y^+, y^- \geq 0$. It follows that $r \in C_{A,B}$ is an extreme ray of $C_{A,B}$ if for all nonzero $r' = (x'^+, x'^-, y'^+, y'^-) \in C_{A,B}$ with $\text{supp}(r') \subseteq \text{supp}(r)$, $r'$ satisfies $r' = \alpha r$ for some $\alpha > 0$. Using this characterization, we will first show that the vectors of $S$ are extreme rays of $C_{A,B}$, and then we will show that all extreme rays have positive scalar multiples in $S \cup T_x \cup T_y$.

To show the former, we first note that $S \subseteq C_{A,B}$. Now, let $r = (g^+, g^-, y^+, y^-) \in S$ and let $r' = (x'^+, x'^-, y'^+, y'^-) \in S$ be any point in $C_{A,B}$ with $\text{supp}(r') \subseteq \text{supp}(r)$. Note that by definition, $\text{supp}(g^+) \cap \text{supp}(g^-) = \emptyset$ and $\text{supp}(y^+) \cap \text{supp}(y^-) = \emptyset$. Then since $\text{supp}(r') \subseteq \text{supp}(r)$, we have that $\text{supp}(x'^+) \cap \text{supp}(x'^-) = \emptyset$ and $\text{supp}(y'^+) \cap \text{supp}(y'^-) = \emptyset$. Let $g = g^+ - g^-$ and $x' = x'^+ - x'^-$. Then $\text{supp}(Bx') = \text{supp}(y'^+ - y'^-) = \text{supp}(y'^+) \cup \text{supp}(y'^-) \subseteq \text{supp}(y^+) \cup \text{supp}(y^-) = \text{supp}(y^+) - y^- = \text{supp}(B_g)$, so $\text{supp}(Bx') \subseteq \text{supp}(B_g)$.

In the case when $x' \neq 0$, the proof of [S] can be followed verbatim to conclude that $x' = \alpha g$ and $(y'^+, y'^-) = \alpha(y^+, y^-)$ for some $\alpha > 0$. It remains to show in this case that $(x'^+, x'^-) = \alpha(g^+, g^-)$. We have that $\text{supp}(x') = \text{supp}(x'^+) \cup \text{supp}(x'^-) \subseteq \text{supp}(g^+) \cup \text{supp}(g^-) = \text{supp}(g) = \text{supp}(x')$, so equality holds throughout. Therefore, $\text{supp}(x'^+) = \text{supp}(g^+)$, $\text{supp}(x'^-) = \text{supp}(g^-)$, and so $(x'^+, x'^-) = \alpha(g^+, g^-)$, as desired.

Now, suppose instead that $x' = 0$. Then $x'^+ = x'^-$. However, as observed above, $\text{supp}(x'^+) \cap \text{supp}(x'^-) = \emptyset$, so in fact $x'^+ = x'^- = 0$. Furthermore, $Bx' = 0$, and so similarly $y'^+ = y'^- = 0$, and so $r' = 0$, a contradiction. Therefore, we have that $r' = \alpha r$ for some $\alpha > 0$, so $r$ is an extreme ray of $C_{A,B}$.

We now show that every extreme ray of $C_{A,B}$ is in $S \cup T_x \cup T_y$. Let $r = (x^+, x^-, y^+, y^-) \in S$ be an extreme ray of $C_{A,B}$. It is shown in the proof in [S] that for any index $i$ such that $y^+_i = y^-_i$ are both positive, $y^+_i = y^-_i$. The same proof shows that for any index $j$ such that $x^+_j$ and $x^-_j$ are both positive, then $x^+_j = x^-_j$. Now, suppose there exists an index $i$ such that $y^+_i = y^-_i > 0$. Let $y'^+$ be defined by $y'^+_i = y^+_i$, $y'^+_j = 0$ for all $j \neq i$, and let $y'^-_i = y^-_i$. Then the point $r' = (0, 0, y'^+, y'^-)$ is in $C_{A,B}$ and $\text{supp}(r') \subseteq \text{supp}(r)$. Then since $r$ is an extreme ray, $r' = \alpha r$ for some $\alpha > 0$, and
since \( y_i^+ = y_i^- \), we have \( \alpha = 1 \). Therefore, in this case \( r = r' \) and so by the definition of \( r' \), \( r \) has a positive scalar multiple in \( T_y \). The same proof shows that if there exits any index \( j \) such that \( x_j^+ = x_j^- > 0 \), then \( r \) has a positive scalar multiple in \( T_x \).

Then we may assume that for each index \( i \leq m_B \), at most one of \( y_i^+ \) and \( y_i^- \) is non-zero, and for each index \( j \leq n \), at most one of \( x_j^+ \) and \( x_j^- \) is non-zero. Let \( x = x^+ - x^- \). Then we may also assume \( x \neq 0 \), as otherwise \( x^+ = x^- \), \( y^+ = y^- \) and since \( r \neq 0 \), we are in one of the cases considered above. From here, the proof in [S] can be followed verbatim to conclude that \( x \) is a circuit of \( \mathcal{P} \) with \( y_i^+ = \max \{ (Bx)_i, 0 \} \) and \( y_i^- = \max \{ -(Bx)_i, 0 \} \) for all \( i \leq m_B \). Since \( x = x^+ - x^- \) and \( \text{supp}(x^+) \cap \text{supp}(x^-) = \emptyset \), it follows that \( x_j^+ = \max \{ x_j, 0 \} \) and \( x_j^- = \max \{ -x_j, 0 \} \) for all \( j \leq n \). Therefore \( r \) has a positive scalar multiple in \( S \), as desired.

**Theorem 9.** Given a pointed polyhedron \( \mathcal{P} = \{ x \in \mathbb{R}^n : Ax = b, Bx \leq d \} \), the set of vertices of the polytope

\[
P_{A,B} = \{ (x^+, x^-, y^+, y^-) \in \mathbb{R}^{2n+2m_B} : A(x^+ - x^-) = 0, B(x^+ - x^-) = (y^+ - y^-), \|x^+\|_1 + \|x^-\|_1 = 1, x^+, x^-, y^+, y^- \geq 0 \}
\]

is \( S' \cup T'_x \) where \( S' \) consists the scaled extreme rays from \( S \) of \( C_{A,B} \), and \( T'_x \) consists of the scaled extreme rays from \( T_x \) of \( C_{A,B} \).

**Proof.** Since \( x^+, x^- \geq 0 \), we have that the equality \( \|x^+\|_1 + \|x^-\|_1 = 1 \) corresponds to the hyperplane

\[
\sum_{j=1}^n x_j^+ + \sum_{j=1}^n x_j^- = 1.
\]

Furthermore, each extreme ray of \( C_{A,B} \) which is parallel to a vector in \( S \cup T_x \) intersects this hyperplane exactly once. The convex hull of these intersection points gives the polytope \( P_{A,B} \).

It remains to show how we can use this polyhedral model to compute a steepest-descent pivot at any feasible point \( x_0 \) of \( \mathcal{P} \). This is done, as in [S], by optimizing over a face of \( P_{A,B} \).

**Corollary 5.** Given a pointed polyhedron \( \mathcal{P} = \{ x \in \mathbb{R}^n : Ax = b, Bx \leq d \} \) and a point \( x_0 \in \mathcal{P} \), the set of vertices of the polytope

\[
P_{A,B,x_0} = \{ (x^+, x^-, y^+, y^-) \in P_{A,B} : y_i^+ = 0 \text{ for each } i \leq m_B \text{ such that } (Bx_0)_i = d_i \}
\]

is \( S'' \cup T''_x \) where \( S'' \) is a subset of the vertices \( S' \) of \( P_{A,B} \) which correspond to circuits which are strictly feasible at \( x_0 \), and \( T''_x \) is a subset of the vertices from \( T'_x \) of \( P_{A,B} \).

Given a vertex \( (g^+, g^-, y^+, y^-) \in S'' \) of \( P_{A,B,x_0} \) corresponding to the circuit \( g = g^+ - g^- \), the constraint \( \|g^+\|_1 + \|g^-\|_1 = 1 \) implies that \( \|g\|_1 = 1 \). This implies the following corollary:

**Corollary 6.** Consider an LP of the form [1] and a non-optimal solution \( x_0 \). If \( (g^+, g^-, y^+, y^-) \) is an optimal extreme-point solution to

\[
\begin{align*}
\max -c^T(g^+ - g^-) \\
\text{s.t.} \\
(g^+, g^-, y^+, y^-) \in P_{A,B,x_0},
\end{align*}
\]

23
then \( g = g^+ - g^- \), when scaled to have coprime integer components, is an optimal solution to \( \text{Steep}(\mathcal{P}, \mathbf{x}_0, c) \).