

# Circuit Diameter of Polyhedra is Strongly Polynomial

Bento Natura

Columbia University

## Abstract

We prove a strongly polynomial bound on the circuit diameter of polyhedra, resolving the circuit analogue of the polynomial Hirsch conjecture. Specifically, we show that the circuit diameter of a polyhedron  $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$  with  $A \in \mathbb{R}^{m \times n}$  is  $O(m^2 \log m)$ . Our construction yields monotone circuit walks, giving the same bound for the monotone circuit diameter.

The circuit diameter, introduced by Borgwardt, Finhold, and Hemmecke (SIDMA 2015), is a natural relaxation of the combinatorial diameter that allows steps along circuit directions rather than only along edges. All prior upper bounds on the circuit diameter were only weakly polynomial. Finding a circuit augmentation algorithm that matches this bound would yield a strongly polynomial time algorithm for linear programming, resolving Smale’s 9th problem.

## 1 Introduction

Linear programming is a cornerstone of modern optimization, with applications spanning operations research, economics, machine learning, and algorithm design. The simplex method, introduced by Dantzig in 1947, remains a workhorse algorithm for solving linear programs in practice, despite having exponential worst-case complexity. Understanding the geometry underlying the simplex method’s remarkable empirical performance has been a central pursuit in optimization theory for over seventy years.

At the heart of this question lies the *combinatorial diameter* of a polyhedron: the diameter of its vertex-edge graph, measuring the maximum number of edge steps needed to traverse between any pair of vertices. This quantity directly governs the worst-case iteration complexity of simplex-type algorithms. In 1957, Warren M. Hirsch conjectured that the combinatorial diameter of a  $d$ -dimensional polytope with  $f$  facets is at most  $f - d$ , a bound that would elegantly explain why the simplex method rarely requires many pivots in practice. Hirsch’s conjecture became one of the most celebrated open problems in polyhedral combinatorics, influencing decades of research in discrete geometry and optimization.

The conjecture remained open for over half a century. In a landmark result, Santos [San12] disproved it in 2012, constructing a counterexample with diameter exceeding the Hirsch bound. However, Santos’s counterexample still has polynomial diameter, leaving the more fundamental question unresolved: the *polynomial Hirsch conjecture*, which asks whether the combinatorial diameter of every polytope admits a polynomial bound  $\text{poly}(f, d)$ , remains the central open problem in this area. This question is not merely of geometric interest: a positive answer would fundamentally advance our understanding of why pivot algorithms perform well, while a negative answer would reveal inherent limitations of the simplex method and related approaches.

Current progress on the polynomial Hirsch conjecture remains limited. Kalai and Kleitman [Kal92, KK92] established the first quasipolynomial bound; subsequent improvements have refined the exponent [Suk17], but no polynomial bound is known for general polytopes. Polynomial bounds have been proven only in restricted cases: Dyer and Frieze [DF94] settled the conjecture for totally unimodular matrices, and polynomial bounds in terms of the maximum subdeterminant  $\Delta$  have been established for integer constraint matrices [BDSE<sup>+</sup>14, BR13, EV17, DH16]. After 65 years of effort, the polynomial Hirsch conjecture for general polyhedra remains tantalizingly open.

## 1.1 Circuit Diameter: A Natural Geometric Relaxation

Given the difficulty of the combinatorial diameter problem, it is natural to study relaxations that may be more tractable while still capturing essential geometric properties of polyhedra. Borgwardt, Finhold, and Hemmecke [BFH15] introduced the *circuit diameter* as such a relaxation, generalizing edge directions to *circuit directions*, which encompass all possible edge directions across all choices of the right-hand side vector.

To define circuit diameter precisely, consider a polyhedron in standard equality form

$$P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\} \quad (\text{P})$$

for  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ , with  $\text{rk}(A) = m$ . An *elementary vector* in  $\ker(A)$  is a support-minimal nonzero vector  $g \in \ker(A)$ : no  $h \in \ker(A) \setminus \{0\}$  satisfies  $\text{supp}(h) \subsetneq \text{supp}(g)$ . A *circuit* is the support of an elementary vector; these are precisely the circuits of the linear matroid defined over the columns of  $A$ . We denote by  $\mathcal{E}(A) \subseteq \ker(A)$  and  $\mathcal{C}(A) \subseteq 2^{[n]}$  the sets of elementary vectors and circuits, respectively. We remark that many papers on circuit diameter [BSY18, BDLF16, BFH15, BDLFM18, KPS19] refer to elementary vectors as circuits; we follow the traditional convention of [Ful67, Roc67, Lee89].

Elementary vectors naturally generalize edge directions. Every edge direction of  $P$  is an elementary vector, and conversely, the set  $\mathcal{E}(A)$  equals the set of all possible edge directions of polyhedra of the form (P) as  $b$  varies over  $\mathbb{R}^m$  [ST97]. Thus, elementary vectors capture the full geometric richness of the constraint matrix  $A$ , independent of any particular right-hand side.

A *circuit walk* is a sequence  $x^{(0)}, x^{(1)}, \dots, x^{(k)}$  in  $P$  where each step  $x^{(i+1)} = x^{(i)} + \alpha^{(i)} g^{(i)}$  uses an elementary vector  $g^{(i)} \in \mathcal{E}(A)$  with  $\alpha^{(i)} > 0$  chosen maximally:  $x^{(i)} + \alpha' g^{(i)} \notin P$  for any  $\alpha' > \alpha^{(i)}$ . The *circuit diameter* of  $P$  is the maximum, over all pairs of vertices  $x, y \in P$ , of the length of a shortest circuit walk from  $x$  to  $y$ . Note that circuit walks are non-reversible due to the maximal step requirement; this asymmetry is a key technical challenge.

Circuit directions arise naturally in classical combinatorial optimization algorithms. Many fundamental algorithms for network flows and transportation problems, including the network simplex method and minimum-cost flow algorithms, are circuit augmentation algorithms [Bla76, DLHL15]. Thus, understanding circuit walks is not merely a geometric abstraction, but directly relevant to algorithmic practice.

The *circuit Hirsch conjecture*, formulated in [BFH15], asserts that the circuit diameter of a  $d$ -dimensional polyhedron with  $f$  facets is at most  $f - d$ . For  $P$  in the form (P), where  $d = n - m$  and the number of facets is at most  $n$ , this conjectured bound is  $m$ . While this strong form of the conjecture remains open, our focus is on the weaker but more fundamental *polynomial circuit Hirsch conjecture*: whether circuit diameter is bounded by  $\text{poly}(m, n)$ .

**The circuit Hirsch conjecture as a central open problem.** Since its introduction, the circuit Hirsch conjecture has attracted significant attention in the polyhedral combinatorics and optimization communities, appearing prominently throughout the literature [BSY18, BBB25, DKNV24, BGK<sup>+</sup>25, BNS26, KPS19, NS24, Wu125, SY15, BFH16, BV22, ENV22]. Black, Borgwardt, and Brugger [BBB25] describe it as “the main open problem in this area”; Borgwardt, Grewe, Kafer, Lee, and Sanità [BGK<sup>+</sup>25] note that “a resolution of the circuit diameter conjecture would give insight to the reason the Hirsch conjecture does not hold”.

## 1.2 Our Results: The First Strongly Polynomial Bound

The main contribution of this paper is the first strongly polynomial bound on the circuit diameter of polyhedra, settling the polynomial circuit Hirsch conjecture.

**Theorem 1.1.** *The circuit diameter of a polyhedron of the form (P) with  $A \in \mathbb{R}^{m \times n}$  is  $O(m^2 \log m)$ .*

This result represents a significant departure from all prior work on circuit diameter. Previous upper bounds fall into two categories: (i) bounds derived by analyzing the iteration complexity of specific circuit augmentation algorithms, which inherit dependencies on numerical properties of the input, and (ii) bounds obtained by designing circuit augmentation schemes that imitate known algorithms, yielding circuit diameter bounds that approximately match the iteration counts of those algorithms. Both approaches inherit the limitations of these algorithms: they are either strongly polynomial only on highly structured polytopes, depend on encoding size, or depend on condition numbers of the constraint matrix.

In contrast, our bound of  $O(m^2 \log m)$  depends *only* on  $m$ , the number of constraints. This is the first unconditional geometric result on circuit diameter: it reveals an intrinsic property of the polyhedron’s combinatorial structure, independent of how the constraint matrix is represented numerically. Our result conclusively establishes that short circuit walks exist between any pair of vertices, with length governed solely by the dimension of the ambient space.

**Strongly polynomial versus weakly polynomial complexity.** The distinction between strongly and weakly polynomial bounds is fundamental in optimization. A bound is *weakly polynomial* if it depends on the bit-complexity of the input (e.g., logarithms of numerical values or condition numbers), and *strongly polynomial* if it depends only on combinatorial parameters such as the dimensions of the constraint matrix. Strongly polynomial bounds reveal that a problem’s complexity is governed by its combinatorial structure rather than numerical artifacts of representation.

In linear programming, this distinction is central. Both Khachiyan’s ellipsoid method [Kha79] and interior point methods [Kar84] only run in weakly polynomial time. Finding a strongly polynomial algorithm for LP remains a major open problem, listed by Smale [Sma98] as one of the key mathematical challenges for the 21st century (Smale’s 9th problem). Our result on circuit diameter is, to our knowledge, the first strongly polynomial bound on any diameter measure for general polyhedra.

**Existence versus Computation.** Our construction is algorithmic: given two vertices, a short circuit walk between them can be computed in strongly polynomial time. However, this does not yield a strongly polynomial algorithm for linear programming, since computing an optimal vertex is itself the problem we wish to solve. The challenge is to find short circuit walks toward an optimum without knowing the target vertex in advance. Our result shows that no geometric obstruction prevents such walks from existing; the barrier is purely algorithmic.

Beyond settling the polynomial circuit Hirsch conjecture, our result suggests that polyhedra possess fundamentally good geometric structure. The focus shifts toward understanding what structure might enable efficient navigation without knowing the target: perhaps randomized or approximate methods that find near-optimal walks with high probability, or augmentation rules guided by the geometric insights underlying our proof.

### 1.3 Prior Work on Circuit Diameter

We now survey related work on circuit diameter bounds, organizing by methodology.

**Circuit augmentation algorithms and condition number bounds.** The *circuit imbalance*  $\kappa(\mathbf{A})$  (defined in Section 2) measures the maximum ratio between entries of an elementary vector. This parameter equals 1 if and only if  $\mathbf{A}$  admits a totally unimodular representation [Cam65, ENV22]. Ekbani et al. [ENV22] showed that a natural extension of the Goldberg–Tarjan minimum-mean cycle canceling algorithm [GT89] yields a steepest-descent circuit augmentation algorithm with iteration complexity  $O(n^2 m \kappa(\mathbf{A}) \log(\kappa(\mathbf{A}) + n))$ ; see also Gauthier and Desrosiers [GD21].

Dadush et al. [DKNV24] improved this to  $O(m^2 \log(m + \kappa(\mathbf{A})))$ , giving the first bound with only logarithmic dependence on  $\kappa(\mathbf{A})$ . Their key innovation is a “shoot towards the optimum” scheme: rather than greedily improving the objective, the algorithm moves directly toward a known optimal vertex  $x^*$ , using circuit directions that make geometric progress toward  $x^*$ . This more global strategy reduces dependence on local condition numbers, but still requires logarithmic dependence on the circuit imbalance.

**Interior point methods and straight-line complexity.** Allamigeon et al. [ADL<sup>+</sup>25] introduced *straight-line complexity* (SLC), a geometric measure that quantifies the curvature of trajectories in logarithmic coordinates. They show that this measure provides, up to polynomial factors in  $n$ , an upper bound on the number of iterations their path-following interior-point methods require. Dadush, Kober, and Koh [DKK25] further established that SLC also bounds circuit augmentation complexity: by tracing the central path and decomposing it into “polarized segments”, they obtain bounds of the form  $O(n^3 \sum_{i \in [n]} \text{SLC}_{1/2}(x_i^m))$ .

While SLC is a natural geometric quantity, it can be exponential in the problem dimension [ABGJ18], so this bound is not strongly polynomial.

**Hardness of approximation.** Black, Nöbel, and Steiner [BNS26] showed that approximating the monotone circuit diameter (a variant where steps must decrease the objective) within a factor of  $O(m^{1-\varepsilon})$  is NP-hard for any  $\varepsilon > 0$ . This suggests that efficiently computing optimal circuit walks is computationally intractable, consistent with the Existence versus Optimization gap discussed above.

**Special polytopes.** Constant and linear circuit diameter bounds have been established for specific polytopes. Borgwardt, Finhold, and Hemmecke [BFH15] proved such bounds for dual transportation polyhedra. Kafer, Pashkovich, and Sanità [KPS19] extended these results to matching polytopes, traveling salesman polytopes, and fractional stable set polytopes, using problem-specific structure. Borgwardt, De Loera, and Finhold [BDLF16] introduced several variants of circuit diameter and explored their relationships.

**Comparison of bounds.** The following table summarizes key circuit diameter bounds for general polyhedra, highlighting their dependencies:

Reference	Bound	Depends on
De Loera et al. [LKS22]	$\text{poly}(m, n, \text{size}(\mathbf{A}, b, c))$	$\text{size}(\mathbf{A}, b, c)$
Ekbatalani et al. [ENV22]	$O(n^2 m \kappa(\mathbf{A}) \log(\kappa(\mathbf{A}) + n))$	$\kappa(\mathbf{A})$
Dadush et al. [DKNV24]	$O(m^2 \log(\kappa(\mathbf{A}) + m))$	$\kappa(\mathbf{A})$
Dadush et al. [DKK25]	$O(n^3 \sum_i \text{SLC}_{1/2}(x_i^m))$	SLC
<b>This paper</b>	$O(m^2 \log m)$	-

Note that  $\text{SLC}_{1/2}$  is upper bounded by  $\text{poly}(n) \log(n + \kappa(\mathbf{A}))$  as shown in [DKN<sup>+</sup>24].

Our work is the first to eliminate all dependence on numerical properties, achieving a purely combinatorial bound.

## 1.4 Paper Organization

The remainder of this paper is organized as follows. Section 2 establishes notation and reviews necessary background on elementary vectors, circuits, and conformal decompositions. Section 3 contains our main result.

## 2 Preliminaries

Let  $[n] = \{1, 2, \dots, n\}$ . For  $\alpha \in \mathbb{R}$ , we denote  $\alpha^+ = \max\{0, \alpha\}$  and  $\alpha^- = \max\{0, -\alpha\}$ . For a vector  $z \in \mathbb{R}^n$ , we define  $z^+, z^- \in \mathbb{R}^n$  as  $(z^+)_i = (z_i)^+$ ,  $(z^-)_i = (z_i)^-$  for  $i \in [n]$ . For  $z \in \mathbb{R}^n$ , we let  $\text{supp}(z) = \{i \in [n] : z_i \neq 0\}$  denote its support, and  $1/z \in (\mathbb{R} \cup \{\infty\})^n$  denote the vector  $(1/z_i)_{i \in [n]}$ . Throughout, we use the convention that  $0/0 = 0$  for convenience. We use  $\|\cdot\|_p$  to denote the  $\ell_p$ -norm. We denote by  $\mathbb{R}_+^n$  the non-negative orthant  $\{x \in \mathbb{R}^n : x \geq 0\}$ .

For technical reasons we assume that  $m \geq 2$ . For  $m = 1$ , our main statements are all trivial. Note that every circuit has size at most  $m + 1$  since  $\text{rk}(\mathbf{A}) \leq m$ . The *circuit imbalance measure* of  $\mathbf{A}$  is defined as

$$\kappa(\mathbf{A}) := \max_{g \in \mathcal{E}(\mathbf{A})} \left\{ \frac{|g_i|}{|g_j|} : i, j \in \text{supp}(g) \right\}.$$

For  $P$  as in (P),  $x \in P$  and an elementary vector  $g \in \mathcal{E}(\mathbf{A}) \setminus \mathbb{R}_+^n$ , we let  $\text{aug}_P(x, g) := x + \alpha g$  where  $\alpha = \max\{\tilde{\alpha} : x + \tilde{\alpha} g \in P\}$ .

**Definition 2.1** ([DLHK12]). We say that  $x, y \in \mathbb{R}^n$  are *sign-compatible* if  $x_i y_i \geq 0$  for all  $i \in [n]$ . We write  $x \sqsubseteq y$  if they are sign-compatible and further  $|x_i| \leq |y_i|$  for all  $i \in [n]$ . For  $x \in \ker(\mathbf{A})$ , a *conformal circuit decomposition* of  $x$  is a set of elementary vectors  $h^{(1)}, h^{(2)}, \dots, h^{(k)}$  in  $\ker(\mathbf{A})$  such that  $x = \sum_{j=1}^k h^{(j)}$ ,  $k \leq n - m$ , and  $h^{(j)} \sqsubseteq x$  for all  $j \in [k]$ .

The following lemma shows that every vector in a linear space has a conformal circuit decomposition. It is a simple corollary of the Minkowski–Weyl and Carathéodory theorems.

**Lemma 2.2** ([DKNV24]). *For a matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , every  $x \in \ker(\mathbf{A})$  has a conformal circuit decomposition  $x = \sum_{j=1}^k h^{(j)}$  such that  $k \leq \min\{\dim(\ker(\mathbf{A})), |\text{supp}(x)|\}$ .*

**Lemma 2.3** ([DNV20]). *A conformal circuit decomposition for a vector  $x \in \ker(\mathbf{A})$  can be found in strongly polynomial time.*

### 3 A strongly polynomial bound

In this section we prove our main result, a strongly polynomial bound on the circuit diameter of polyhedra defined by a system of linear equations and non-negativity constraints.

**Theorem 3.1.** *Let  $P = \{x : \mathbf{A}x = b, x \geq 0\}$  be a polyhedron with  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Then, for any feasible  $x \in P$  and any vertex  $v \in P$  there exists a circuit walk from  $x$  to  $v$  of length  $n + O(m^2 \log(m))$ . Furthermore, such a circuit walk can be constructed in strongly polynomial time.*

The following corollaries follow directly from Theorem 3.1.

**Corollary 3.2.** *Let  $P = \{x : \mathbf{A}x = b, x \geq 0\}$  be a polyhedron with  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Then, its circuit diameter is upper bounded by  $O(m^2 \log(m))$ .*

*Proof.* Any two vertices  $u, v \in P$  correspond to basic feasible solutions of (P) and hence have support of at most  $m$  variables each. Reducing the system  $\mathbf{A}$  to the columns indexed by  $\text{supp}(u) \cup \text{supp}(v)$ , we get a polyhedron in at most  $2m$  variables. By Theorem 3.1, there exists a circuit walk from  $u$  to  $v$  of length at most  $2m + O(m^2 \log(m)) = O(m^2 \log(m))$ , as desired.  $\square$

**Corollary 3.3.** *Let  $P = \{x : \mathbf{A}x = b, x \geq 0\}$  be a polyhedron with  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Then, its circuit diameter can be approximated up to a factor  $O(m\sqrt{\log(m)})$  in strongly polynomial time. Furthermore, a circuit walk of length  $O(m^2 \log(m))$  between any two vertices can be constructed in strongly polynomial time.*

*Proof.* The second part of the corollary is trivial from Corollary 3.2. The first part is achieved by guessing a circuit diameter of  $O(m\sqrt{\log(m)})$ .  $\square$

Another important corollary concerns the running time to find a short circuit walk to an optimal vertex of a linear program. As our algorithm can find a circuit walk of length  $O(m^2 \log(m))$  between any two given vertices, we can use any algorithm to solve (P) optimally for some objective vector  $c \in \mathbb{R}^n$  to find an optimal solution and then apply our algorithm.

**Corollary 3.4.** *Given a polyhedron  $P = \{x : \mathbf{A}x = b, x \geq 0\}$  with  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and an objective  $c \in \mathbb{R}^n$ . If  $\mathcal{T}(P)$  is the time to solve (P) optimally with objective  $c$ , then in time  $\mathcal{T}(P) + \text{poly}(m, n)$  we can find a circuit walk from any given initial feasible solution to an optimal vertex of  $P$  of length  $O(m^2 \log(m))$ .*

#### 3.1 Algorithmic Overview

We now outline the main ideas underlying our proof, deferring technical details to Section 3.2. Our proof constructs an explicit circuit walk from any feasible point to a target vertex  $x^*$  (with basis  $B$  and non-basic indices  $N = [n] \setminus B$ ) in  $n + O(m^2 \log m)$  steps. The algorithm operates in two phases.

**Phase 1: Support reduction.** The first phase is straightforward: we reduce the number of nonzero non-basic coordinates. Since any  $m + 1$  variables contain a circuit of  $\ker(\mathbf{A})$ , whenever  $|\text{supp}(x_N)| \geq m + 1$  we can find an elementary vector  $g$  supported entirely within  $\text{supp}(x_N)$  and augment along it, zeroing at least one coordinate. After at most  $n - 2m$  steps, we achieve  $|\text{supp}(x_N)| \leq m$ , giving at most  $2m$  nonzero coordinates total. This phase requires no sophisticated analysis and contributes only the additive  $n$  term to our bound.



**Phase 2: Trapped variables and amortized progress.** The main contribution lies in Phase 2, where we achieve the  $O(m^2 \log m)$  bound through a novel combination of *trapped variables* and *elimination steps*.

After Phase 1, we maintain a *trapped set*  $T \subseteq B$  of basic indices satisfying the invariant

$$x_t \leq m \cdot x_t^* \quad \text{for all } t \in T.$$

This invariant says that trapped coordinates are within a factor of  $m$  of their target values. Crucially, once an index becomes trapped, it remains trapped throughout the algorithm. Progress in Phase 2 is measured by two events: either the trapped set  $T$  grows, or a non-basic coordinate is zeroed. Since  $|T| \leq m$  and  $|\text{supp}(x_N)| \leq m$  after Phase 1, at most  $2m$  such progress events can occur.

Between progress events, the algorithm performs two types of steps:

**Norm-reduction steps.** These steps are standard and have, e.g., been used already in [DKNV24] to make geometric progress towards a target vertex: We decompose the direction  $x^* - x$  into a conformal sum of elementary vectors  $g^{(1)}, \dots, g^{(k)}$  with  $k \leq m$  (by Lemma 2.2, the number of circuits in a conformal decomposition can be upper bounded by the kernel dimension, which is at most  $m$  after Phase 1). We then select the circuit  $g^{(j^*)}$  that maximizes progress on a weighted  $\ell_1$ -norm over  $N$ , and augment along it.

By conformality, each  $g_N^{(j)} \leq 0$  (since  $x_N^* = 0$  and  $x_N \geq 0$ ), so these steps always decrease non-basic coordinates. Crucially, the step size  $\alpha$  is at most  $m$ , which will follow from the greedy circuit selection. This bounded step size will preserve the trapped invariant: for any  $t \in T$ , conformality ensures  $|g_t| \leq |x_t^* - x_t|$ , so the change in  $x_t$  is at most  $m \cdot |x_t^* - x_t| \leq m \cdot x_t^*$ , keeping  $x_t \leq mx_t^*$ .

Each norm-reduction step decreases the weighted  $\ell_1$ -norm by a  $(1 - 1/m)$  factor. After  $O(m \log m)$  consecutive norm-reduction steps, the non-basic coordinates shrink below a threshold  $\tau = 1/\text{poly}(m)$  relative to a reference point  $x^{(r)}$ , triggering an elimination step.

**Elimination steps.** When  $\|x_N/x_N^{(r)}\|_\infty \leq \tau$ , we perform an elimination step designed to force progress. The idea is to extrapolate beyond the current point in the direction it has been moving. Define

$$y = x + \frac{\varrho}{1 - \varrho}(x - x^{(r)}),$$

where  $\varrho = \max_{j \in N} x_j/x_j^{(r)} \leq \tau$  measures how much non-basic coordinates have shrunk. By construction,  $y_N \leq 0$ : we have extrapolated past zero on all non-basic coordinates.

However, moving directly towards  $y$  via a circuit direction in the decomposition of  $y - x$  might zero a trapped coordinate before any non-basic coordinate, hindering the progress we seek. To prevent this, we pull the target slightly from  $y$  toward  $x^*$  by setting

$$z = y + \lambda(x^* - y)$$

with  $\lambda = \Theta(1/\text{poly}(m))$ . This perturbation is carefully calibrated: for trapped coordinates  $t$  with  $x_t \ll x_t^*$ , it ensures  $z_t \geq x_t$ , preventing conformal steps toward  $z$  from decreasing such coordinates. Conversely, for trapped coordinates  $t$  with  $x_t$  close to  $x_t^*$ , the construction ensures that the relative change from  $x$  toward  $z$  for any coordinate in  $N$  dominates that of  $t$ .

We then decompose  $z - x$  conformally and augment along the circuit that makes most progress toward zeroing some coordinate in  $N$ . Crucially, due to the relative change on coordinates  $N$  dominating the relative change on coordinates in  $T$ , we will observe that either some non-basic or non-trapped basic coordinate must hit zero first, guaranteeing progress: either  $|T|$  increases or  $|\text{supp}(x_N)|$  decreases.

**A strongly polynomial bound.** The crucial property of our augmentation scheme is that our parameter  $\tau$  that we require to ensure combinatorial progress is only polynomially small and, unlike prior works, does not depend on condition numbers. Previous approaches could not achieve this because they analyzed specific algorithmic paths guided by numerical properties (objective gradients, circuit imbalances, or curvature measures). Our algorithm, by contrast, is designed so that *any* conformal decomposition suffices: we select

circuits greedily for concreteness, but the analysis only requires that the decomposition exists and has size at most  $m$ . This allows us to bound the walk length using purely combinatorial arguments.

While *trapped variables* appeared already in [DKNV24], the main novelty of our work is the introduction of *elimination steps* and the use of carefully chosen auxiliary points which force combinatorial progress after only  $O(m \log m)$  norm-reduction steps, eliminating all dependence on condition numbers.

### 3.2 Formal Algorithm and Proof of Theorem 3.1

In this section we present the full algorithm and prove Theorem 3.1. The full algorithm is given in Algorithm 1.

---

#### Algorithm 1: Circuit Augmentation Algorithm for Strongly Polynomial Diameter Bound

---

**Input:** Feasible solution  $x^{(0)}$  to (P) with  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , target vertex  $x^*$  with basis  $B$ , threshold value  $\tau = (2m)^{-3}$ ,  $\lambda = (2m)^{-2}$ .

**Output:** Sequence of circuit augmentations from  $x^{(0)}$  to  $x^*$

```

1  $N \leftarrow [n] \setminus B$ ; // Non-basic variables
2  $T^{(-1)} \leftarrow \emptyset$ ; // Trapped variables
3 // Phase 1: Reduce support on non-basic variables
4  $i \leftarrow 0$ ;
5 while  $|\text{supp}(x_N^{(i)})| \geq m + 1$  do
6   Select elementary vector  $g \in \mathcal{E}(\mathbf{A})$  with  $\text{supp}(g) \subseteq \text{supp}(x_N^{(i)})$  and  $\text{supp}(g^-) \neq \emptyset$ ;
7    $x^{(i+1)} \leftarrow \text{aug}_P(x^{(i)}, g)$ ;
8    $i \leftarrow i + 1$ ;
9 // Phase 2: Main circuit augmentation loop
10  $r \leftarrow i$ ; // Reference index (reset when  $T$  grows)
11 while  $x^{(i)} \neq x^*$  do
12   // Update trapped set
13    $T^{(i)} \leftarrow \{j \in B : x_j^{(i)} \leq mx_j^*\}$ ;
14   if  $T^{(i)} \neq T^{(i-1)}$  then
15      $r \leftarrow i$ ; // Reset reference point
16   // Norm-reduction step: shrink non-basic coordinates
17   if  $\|x_N^{(i)} / x_N^{(r)}\|_\infty > \tau$  then
18      $(g^{(1)}, \dots, g^{(k)}) \leftarrow \text{conformal decomposition of } x^* - x^{(i)}$ ;
19      $j^* \leftarrow \arg \max_{j \in [k]} \left\| \frac{g_N^{(j)}}{x_N^{(r)}} \right\|_1$ ; // Best progress on weighted  $\ell_1$ -norm
20      $x^{(i+1)} \leftarrow \text{aug}_P(x^{(i)}, g^{(j^*)})$ ;
21   else
22     // Elimination step: force a desired coordinate to zero
23      $q \leftarrow \arg \max_{j \in N} \frac{x_j^{(i)}}{x_j^{(r)}}$ ;
24      $\varrho \leftarrow \frac{x_q^{(i)}}{x_q^{(r)}}$ ; // Shrinkage ratio; note  $\varrho \leq \tau$ 
25      $y \leftarrow x^{(i)} + \frac{\varrho}{1-\varrho}(x^{(i)} - x^{(r)})$ ; // Extrapolate past zero on  $N$ 
26      $z \leftarrow y + \lambda(x^* - y)$ ; // Pull toward  $x^*$  to protect small trapped coords
27      $(g^{(1)}, \dots, g^{(k)}) \leftarrow \text{conformal decomposition of } z - x^{(i)}$ ;
28      $j^* \leftarrow \arg \max_{j \in [k]} -g_q^{(j)}$ ; // Circuit with most progress on  $q$ 
29      $x^{(i+1)} \leftarrow \text{aug}_P(x^{(i)}, g^{(j^*)})$ ;
30    $i \leftarrow i + 1$ ;
31 return Circuit walk  $(x^{(0)}, x^{(1)}, \dots, x^{(i)})$ 

```

---

*Proof of Theorem 3.1.* We prove Theorem 3.1 by analyzing Algorithm 1. We begin with Phase 1, which reduces the support on the non-basic variables, and then analyze Phase 2.

**Phase 1: Support Reduction (Lines 5 to 8).** Initially, if  $|\text{supp}(x_N^{(0)})| \geq m+1$ , we perform circuit augmentations to reduce the support. Since any set of  $m+1$  variables in  $\ker(\mathbf{A})$  contains a circuit, we can always find an elementary vector  $g \in \mathcal{E}(\mathbf{A})$  with  $\text{supp}(g) \subseteq \text{supp}(x_N^{(0)})$  and  $\text{supp}(g^-) \neq \emptyset$ . Each augmentation reduces  $|\text{supp}(x_N^{(i)})|$  by at least one. This phase terminates after at most  $n - 2m$  iterations with  $|\text{supp}(x_N^{(i)})| \leq m$ , giving us at most  $2m$  non-zero variables in total.

**Reducing the norm on variables in  $N$  (Lines 17 to 20).** We maintain a set  $T \subseteq B$  of *trapped* variables satisfying  $x_t^{(i)} \leq mx_t^*$  for all  $t \in T$ .

Let us first show that the set of trapped variables will only increase in the norm-reduction steps in Line 20 and that furthermore, these steps make significant progress in reducing the weighted  $\ell_1$  norm on  $N$ . Claim 3.5 is analogous to progress guarantees obtained in prior work on circuit augmentation algorithms [DKNV24].

**Claim 3.5.** *If the augmentation in iteration  $i$  happens in Line 20, then  $T^{(i+1)} \supseteq T^{(i)}$  and furthermore, if the current reference index is  $r$ , then*

$$\left\| \frac{x_N^{(i+1)}}{x_N^{(r)}} \right\|_1 \leq \left(1 - \frac{1}{m}\right) \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1.$$

*Proof of Claim.* Let  $t \in T^{(i)}$  be any trapped variable at iteration  $i$ . Let  $(g^{(1)}, \dots, g^{(k)})$  be the conformal decomposition of  $x^* - x^{(i)}$ , and let  $j^* \in \{1, \dots, k\}$  with  $g^{(j^*)}$  being the selected circuit for augmentation in Line 20. By the induction hypothesis, we have  $x_t^{(i)} \leq mx_t^*$ . By definition  $x^{(i+1)} = x^{(i)} + \alpha g^{(j^*)}$  for some  $\alpha > 0$ . First, note that  $g_N^{(j^*)} \leq \mathbf{0}_N$  by conformality and  $\alpha \leq m$  as we have that

$$\begin{aligned} \left\| \frac{x_N^{(i+1)}}{x_N^{(r)}} \right\|_1 &= \left\| \frac{x_N^{(i)}}{x_N^{(r)}} + \alpha \frac{g_N^{(j^*)}}{x_N^{(r)}} \right\|_1 = \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1 - \alpha \left\| \frac{g_N^{(j^*)}}{x_N^{(r)}} \right\|_1 \\ &\leq \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1 - \frac{\alpha}{k} \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1 = \left(1 - \frac{\alpha}{k}\right) \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1. \end{aligned} \tag{1}$$

where the last inequality follows by the choice of  $g^{(j^*)}$  maximizing the weighted  $\ell_1$  norm reduction on  $N$ . The second equality used the fact that  $g_N^{(j^*)} \leq \mathbf{0}_N$ . Using the further fact that  $k \leq \dim(\ker(\mathbf{A}_{\text{supp}(x^{(i)})})) \leq 2m - \text{rk}(\mathbf{A}_B) = 2m - m = m$  by Lemma 2.2 and the fact that the left hand side of (1) is nonnegative gives  $\alpha \leq m$ . From here, note that we have for any  $t \in T$  with  $x_t^{(i)} \geq x_t^*$  that  $g_t^{(j^*)} \leq 0$  by conformality. Hence  $x_t^{(i+1)} \leq x_t^{(i)} \leq mx_t^*$ . For  $t \in T$  with  $x_t^{(i)} \leq x_t^*$ , we have by conformality that  $0 \leq g_t^{(j^*)} \leq x_t^* - x_t^{(i)}$ , giving

$$x_t^{(i+1)} = x_t^{(i)} + \alpha g_t^{(j^*)} \leq x_t^{(i)} + \alpha(x_t^* - x_t^{(i)}) \leq x_t^{(i)} + m(x_t^* - x_t^{(i)}) \leq mx_t^*,$$

which proves the first part of the claim. For the second part of the claim, note that we also have that  $\alpha \geq 1$ , again by conformality as we have that

$$\alpha = \min \left\{ -\frac{x_\ell^{(i)}}{g_\ell^{(j^*)}} : \ell \in \text{supp}([g^{(j^*)}]^-) \right\} \geq \min \left\{ \frac{x_\ell^{(i)}}{x_\ell^{(i)} - x_\ell^*} : \ell \in \text{supp}([x^* - x^{(i)}]^-) \right\} \geq 1. \tag{2}$$

Therefore, we have with (1) that

$$\left\| \frac{x_N^{(i+1)}}{x_N^{(r)}} \right\|_1 \leq \left(1 - \frac{\alpha}{k}\right) \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1 \leq \left(1 - \frac{1}{m}\right) \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1,$$

which proves the claim. ■



A direct consequence of Claim 3.5 is the following.

**Claim 3.6.** *Within  $m \log(m/\tau)$  many consecutive executions of Line 20 either the set of trapped variables  $T$  gets extended or the if condition  $\|x_N^{(i)}/x_N^{(r)}\|_\infty > \tau$  in Line 17 becomes false.*

*Proof of Claim.* Let  $i$  be an iterate and let  $p = i + m \cdot \log(m/\tau)$  be such that in each iteration  $j \in \{i, \dots, p-1\}$  the augmentation is performed in Line 20. If  $T$  has not been extended in any of these iterations, then the reference index  $r \leq i$  remains unchanged. By Claim 3.5, we have that

$$\begin{aligned} \left\| \frac{x_N^{(p)}}{x_N^{(r)}} \right\|_\infty &\leq \left\| \frac{x_N^{(p)}}{x_N^{(r)}} \right\|_1 \leq \left(1 - \frac{1}{m}\right)^{p-i} \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1 = \left(1 - \frac{1}{m}\right)^{m \log(m/\tau)} \left\| \frac{x_N^{(i)}}{x_N^{(r)}} \right\|_1 \leq \left(1 - \frac{1}{m}\right)^{m \log(m/\tau)} \cdot m \\ &\leq e^{-\log(m/\tau)} \cdot m \leq \frac{\tau}{m} \cdot m = \tau, \end{aligned}$$

as desired. ■

We now claim that a single execution of the elimination step in Line 29 either sets a variable in  $N$  to zero or extends  $T^{(r)} (= T^{(i)})$ . From now on, we let  $T := T^{(r)}$  for simplicity.

**Elimination step analysis** Let  $i > r$  be such that the condition  $\|x_N^{(i)}/x_N^{(r)}\|_\infty \leq \tau$  holds. Let  $q \in N$  be the index maximizing  $x_q^{(i)}/x_q^{(r)}$  and let  $\varrho := x_q^{(i)}/x_q^{(r)} \leq \tau$ . As in the algorithm, we define the vectors

$$y := x^{(i)} + \frac{x_q^{(i)}}{x_q^{(r)} - x_q^{(i)}}(x^{(i)} - x^{(r)}) = x^{(i)} + \frac{\varrho}{1 - \varrho}(x^{(i)} - x^{(r)}), \quad z := y + \lambda(x^* - y).$$

Note that  $y_q = 0$ . By construction, we have  $y_N \leq \mathbf{0}_N$  as for any  $j \in N$ , we get that

$$(1 - \varrho)y_j = (1 - \varrho)x_j^{(i)} + \varrho(x_j^{(i)} - x_j^{(r)}) = x_j^{(i)} - \varrho x_j^{(r)} \leq x_j^{(i)} - \frac{x_j^{(i)}}{x_j^{(r)}} x_j^{(r)} = 0.$$

For all  $t \in T$ , we have

$$\begin{aligned} z_t - x_t^{(i)} &= \lambda(x_t^* - y_t) + \frac{\varrho}{1 - \varrho}(x_t^{(i)} - x_t^{(r)}) \\ &= \lambda(x_t^* - x_t^{(i)}) + \lambda(x_t^{(i)} - y_t) + \frac{\varrho}{1 - \varrho}(x_t^{(i)} - x_t^{(r)}) \\ &= \lambda(x_t^* - x_t^{(i)}) + \frac{(1 - \lambda)\varrho}{1 - \varrho}(x_t^{(i)} - x_t^{(r)}). \end{aligned}$$

Now, by Claim 3.5 we have that  $x_t^{(i)}, x_t^{(r)} \leq m x_t^*$ . Therefore, we have that

$$\begin{aligned} \frac{|z_t - x_t^{(i)}|}{x_t^*} &= \frac{\left| \lambda(x_t^* - x_t^{(i)}) + \frac{(1 - \lambda)\varrho}{1 - \varrho}(x_t^{(i)} - x_t^{(r)}) \right|}{x_t^*} \\ &\leq \lambda \frac{|x_t^* - x_t^{(i)}|}{x_t^*} + \frac{(1 - \lambda)\varrho}{1 - \varrho} \frac{|x_t^{(i)} - x_t^{(r)}|}{x_t^*} \\ &\leq \left( \lambda + \frac{(1 - \lambda)\varrho}{1 - \varrho} \right) m. \end{aligned} \tag{3}$$

For  $t \in T$  with  $x_t^{(i)} \leq \frac{1}{2}x_t^*$ , we furthermore have

$$\begin{aligned}
z_t - x_t^{(i)} &= \lambda(x_t^* - x_t^{(i)}) + \frac{(1-\lambda)\varrho}{1-\varrho}(x_t^{(i)} - x_t^{(r)}) \\
&\geq \lambda(x_t^* - x_t^{(i)}) - \frac{(1-\lambda)\varrho}{1-\varrho} |x_t^{(i)} - x_t^{(r)}| \\
&\geq \lambda \frac{x_t^*}{2} - \frac{(1-\lambda)\varrho}{1-\varrho} \cdot mx_t^* \\
&= \left( \frac{\lambda}{2} - \frac{(1-\lambda)\varrho}{1-\varrho} m \right) x_t^* \\
&\geq \left( \frac{\lambda}{2} - \varrho m \right) x_t^* \\
&\geq 0.
\end{aligned}$$

Observe that  $\lambda = (2m)^{-2} \leq 1$ , which implies that  $z$  is a convex combination of  $y$  and  $x^*$ . Consequently,  $z_N \leq \mathbf{0}_N$ . So, in contrast, for  $j \in N$ , we have

$$\frac{|z_j - x_j^{(i)}|}{x_j^{(i)}} \geq \frac{|0 - x_j^{(i)}|}{x_j^{(i)}} \geq 1$$

as  $z_j \leq 0$  and  $x_j^{(i)} \geq 0$ .

We now show that when we decompose  $z - x^{(i)}$  conformally (Line 27) and select  $g^{(j^*)}$  maximizing  $-g_q^{(j)}$  (Line 28), the augmentation (Line 29) will zero out a coordinate in  $N$  or  $B \setminus T$  before any coordinate in  $T$ .

The coordinate that gets zeroed out first is

$$\arg \max_{j: g_j^{(j^*)} < 0} \frac{-g_j^{(j^*)}}{x_j^{(i)}}.$$

By conformality, we have for all  $t \in T$  with  $x_t^{(i)} \leq \frac{1}{2}x_t^*$  that  $g_t^{(j^*)} \geq 0$ . Hence, such coordinates cannot be zeroed out in the conformal augmentation. For  $t \in T$  with  $x_t^{(i)} \geq \frac{1}{2}x_t^*$  we have that

$$\begin{aligned}
\frac{-g_t^{(j^*)}}{x_t^{(i)}} &\leq \frac{|z_t - x_t^{(i)}|}{x_t^{(i)}} \\
&\leq \frac{2|z_t - x_t^{(i)}|}{x_t^*} \\
&\leq 2 \left( \lambda + \frac{(1-\lambda)\varrho}{1-\varrho} \right) m && \text{(by (3))} \\
&\leq 2 \left( \lambda + \frac{(1-\lambda)\tau}{1-\tau} \right) m && \text{(as } \varrho \leq \tau) \\
&= 2 \left( (2m)^{-2} + \frac{(1-(2m)^{-2}) \cdot (2m)^{-3}}{1-(2m)^{-3}} \right) m \\
&< \frac{1}{m}.
\end{aligned}$$

On the other hand, we have that  $z_q = 0$  and so  $\sum_j g_q^{(j)} = z_q - x_q^{(i)} = -x_q^{(i)}$  and so

$$\max_{j \in N} \frac{-g_j^{(j^*)}}{x_j^{(i)}} \geq \frac{-g_q^{(j^*)}}{x_q^{(i)}} \geq \frac{1}{k} \sum_{j=1}^k \frac{-g_q^{(j)}}{x_q^{(i)}} = \frac{1}{k} \geq \frac{1}{m}.$$

and so the two inequalities above yield

$$\max_{j \in N} \frac{-g_j^{(j^*)}}{x_j^{(i)}} > \max_{t \in T} \frac{-g_t^{(j^*)}}{x_t^{(i)}}. \quad (4)$$

Therefore, the augmentation in Line 29 cannot set a variable in  $T$  to zero. As a consequence, either a variable in  $N$  is set to zero, or a variable in  $B \setminus T$  is set to zero. In the former case, this variable will remain zero forever, while in the latter case, the variable that is set to zero will now satisfy  $x_j^{(i+1)} = 0 \leq mx_j^*$  and so  $T$  gets extended.

With this we proved the main part of the theorem. To show that the algorithm is correct we still have to show that after this augmentation, variables that originated in  $T$  remain trapped.

**Claim 3.7.** *Let  $i$  be an iterate where the augmentation is performed in Line 29 and let  $x^{(i+1)}$  be the resulting point. Then,  $T^{(i)} \subseteq T^{(i+1)}$ .*

*Proof of Claim.* For any  $t \in T$  with  $x_t^{(i)} \geq z_t$ , we have by conformality that  $g_t^{(j^*)} \leq 0$  and so  $x_t^{(i+1)} \leq x_t^{(i)} \leq mx_t^*$ .

For  $t \in T$  with  $x_t^{(i)} \geq \frac{3}{2}x_t^*$ , we get that

$$z_t - x_t^{(i)} = \lambda(x_t^* - x_t^{(i)}) + \frac{(1-\lambda)\varrho}{1-\varrho}(x_t^{(i)} - x_t^{(r)}) \leq -\frac{\lambda}{2}x_t^* + \frac{(1-\lambda)\varrho}{1-\varrho}mx_t^* \leq \left(-\frac{\lambda}{2} + \varrho m\right)x_t^* \leq \left(-\frac{\lambda}{2} + \tau m\right)x_t^* = 0.$$

and so  $g_t^{(j^*)} \leq 0$ .

It therefore remains to consider the case  $x_t^{(i)} \leq \frac{3}{2}x_t^*$  and  $z_t \geq x_t^{(i)}$ . Now, note that similar to the proof of Claim 3.5, we have that the step size  $\alpha$  in the augmentation  $x^{(i+1)} = x^{(i)} + \alpha g^{(j^*)}$  satisfies  $\alpha \leq m$ . This can be observed by noting that  $-g_q^{(j^*)} = \max_{j \in [k]} -g_q^{(j)} \geq \frac{1}{k} \sum_{j=1}^k -g_q^{(j)} = \frac{1}{k}x_q^{(i)} \geq \frac{1}{m}x_q^{(i)}$  and so

$$\alpha = \min \left\{ -\frac{x_j^{(i)}}{g_j^{(j^*)}} : j \in \text{supp}([g^{(j^*)}]^-) \right\} \leq -\frac{x_q^{(i)}}{g_q^{(j^*)}} \leq m. \quad (5)$$

But then, we have with (3) that

$$\begin{aligned} x_t^{(i+1)} &= x_t^{(i)} + \alpha g_t^{(j^*)} \\ &\leq x_t^{(i)} + \alpha(z_t - x_t^{(i)}) && \text{(by conformality)} \\ &\leq x_t^{(i)} + m \left( \lambda + \frac{(1-\lambda)\varrho}{1-\varrho} \right) mx_t^* && \text{(by } \alpha \leq m \text{ and (3))} \\ &\leq \frac{3}{2} \left( 1 + m^2 \cdot \left( \lambda + \frac{(1-\lambda)\varrho}{1-\varrho} \right) \right) x_t^* && \text{(by } x_t^{(i)} \leq \frac{3}{2}x_t^*) \\ &\leq \frac{3}{2} (1 + m^2 \cdot (\lambda + \varrho)) x_t^* && \text{(by } \lambda > \varrho) \\ &\leq \frac{3}{2} (1 + m^2 \cdot ((2m)^{-2} + (2m)^{-3})) x_t^* \\ &\leq mx_t^*, \end{aligned}$$

where the last inequality follows by the assumption that  $m \geq 2$ . This proves the claim.  $\blacksquare$

It remains to analyze the total number of circuit augmentations of the algorithm. By Claim 3.5, we perform at most  $m \log(m/\tau)$  consecutive norm-reduction iterations before either  $T$  gets extended or we reduced the relative norm on  $N$  below  $\tau$  so that the elimination step is called, which either extends  $T$  or sets a variable in  $N$  to zero. Since  $|T| \leq m$  and  $|N| \leq m$  after Phase 1, there are at most  $2m$  such progress events, so the total number of augmentations after Phase 1 is at most  $m \cdot m \log(m/\tau) = O(m^2 \log(m))$ .  $\square$

### 3.3 Monotone Diameter

A circuit walk is *monotone* with respect to an objective  $c \in \mathbb{R}^n$  if the objective value does not increase along the walk. The *monotone circuit diameter* measures the worst-case length of shortest monotone walks.

**Definition 3.8.** Let  $P = \{x : Ax = b, x \geq 0\} \subseteq \mathbb{R}^n$  be a polyhedron. A circuit walk  $(x^{(0)}, x^{(1)}, \dots, x^{(k)})$  is *monotone with respect to*  $c \in \mathbb{R}^n$  if  $c^\top x^{(i+1)} \leq c^\top x^{(i)}$  for all  $i \in \{0, \dots, k-1\}$ . The *monotone circuit diameter* of  $P$  is the maximum, over all vertices  $u, v \in P$  and objectives  $c \in \mathbb{R}^n$  with  $v \in \arg \min_{x \in P} c^\top x$ , of the length of the shortest monotone circuit walk from  $u$  to  $v$ .

Since every circuit walk is trivially monotone for  $c = 0$ , the circuit diameter is a lower bound on the monotone circuit diameter. We show that Algorithm 1 produces monotone walks, yielding a strongly polynomial bound on the monotone circuit diameter as well.

The key observation is that monotonicity can be characterized in terms of the non-basic coordinates. Let  $v = x^*$  be the target vertex with basis  $B$  and non-basic indices  $N = [n] \setminus B$ . By complementary slackness, any objective  $c$  minimized at  $x^*$  satisfies  $c = A^\top y + s$  for some  $y \in \mathbb{R}^m$  and  $s \in \mathbb{R}^n$  with  $s_B = 0$  and  $s_N \geq 0$ . For any circuit step  $x^{(i)} \rightarrow x^{(i+1)} = x^{(i)} + \alpha g$  with  $g \in \ker(A)$ , we have

$$c^\top (x^{(i+1)} - x^{(i)}) = \alpha \cdot c^\top g = \alpha \cdot s^\top g = \alpha \sum_{j \in N} s_j g_j.$$

Since  $\alpha > 0$  and  $s_N \geq 0$ , the step is monotone if  $g_N \leq 0$ . Thus, an algorithm that only uses circuit directions  $g$  with  $g_N \leq 0$  is monotone with respect to all objectives minimized at  $x^*$ .

Algorithm 1 satisfies this property. In the norm-reduction steps, the conformal decomposition of  $x^* - x^{(i)}$  yields circuits  $g^{(j)} \subseteq x^* - x^{(i)}$ . Since  $x_N^* = 0$  and  $x_N^{(i)} \geq 0$ , we have  $(x^* - x^{(i)})_N \leq 0$ , and conformality implies  $g_N^{(j)} \leq 0$ . In the elimination steps, the target  $z$  satisfies  $z_N \leq 0$  by construction, and since  $x_N^{(i)} \geq 0$ , we again have  $(z - x^{(i)})_N \leq 0$ , so the conformal decomposition yields circuits with  $g_N^{(j)} \leq 0$ .

**Corollary 3.9.** Let  $P = \{x : Ax = b, x \geq 0\} \subseteq \mathbb{R}^n$  be a polyhedron with  $A \in \mathbb{R}^{m \times n}$ . Then, the monotone circuit diameter of  $P$  is bounded by  $O(m^2 \log(m))$ .

We complement this with a hardness result of Black, Nöbel, and Steiner [BNS26], showing that computing the monotone circuit diameter exactly is hard.

**Theorem 3.10** ([BNS26, Corollary 1.7]). For every  $\varepsilon > 0$  and  $m \geq 2$ , the following problem is NP-hard: Given a polytope  $P = \{x : Ax = b, x \geq 0\} \subseteq \mathbb{R}^n$  with  $A \in \mathbb{R}^{m \times n}$ , two vertices  $u, v \in P$ , and an objective  $c$  minimized at  $v$ , compute a monotone circuit walk from  $u$  to  $v$  approximating the minimum possible length of such a walk to within a factor of  $O(m^{1-\varepsilon})$ .

## References

- [ABGJ18] X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig. Log-barrier interior point methods are not strongly polynomial. *SIAM Journal on Applied Algebra and Geometry*, 2(1):140–178, 2018.
- [ADL<sup>+</sup>25] X. Allamigeon, D. Dadush, G. Loho, B. Natura, and L. A. Végh. Interior point methods are not worse than simplex. *SIAM Journal on Computing*, 54(5):FOCS22–178–FOCS22–264, 2025.
- [BBB25] A. E. Black, S. Borgwardt, and M. Brugger. On the circuit diameter conjecture for counterexamples to the hirsch conjecture. *Discrete Optimization*, 58:100910, 2025.
- [BDLF16] S. Borgwardt, J. A. De Loera, and E. Finhold. Edges versus circuits: a hierarchy of diameters in polyhedra. *Advances in Geometry*, 16(4):511–530, 2016.
- [BDLFM18] S. Borgwardt, J. A. De Loera, E. Finhold, and J. Miller. The hierarchy of circuit diameters and transportation polytopes. *Discrete Applied Mathematics*, 240:8–24, 2018.
- [BDSE<sup>+</sup>14] N. Bonifas, M. Di Summa, F. Eisenbrand, N. Hähnle, and M. Niemeier. On sub-determinants and the diameter of polyhedra. *Discrete & Computational Geometry*, 52(1):102–115, 2014.

- [BFH15] 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.
- [BFH16] S. Borgwardt, E. Finhold, and R. Hemmecke. Quadratic diameter bounds for dual network flow polyhedra. *Mathematical Programming*, 159(1–2):237–251, 2016. arXiv:1408.4184.
- [BGK<sup>+</sup>25] S. Borgwardt, P. Grewe, S. Kafer, J. Lee, and L. Sanità. On the hardness of short and sign-compatible circuit walks. *Discrete Applied Mathematics*, 367, 2025. arXiv:2402.01066.
- [Bla76] R. G. Bland. On the generality of network flow theory. Presented at the ORSA/TIMS Joint National Meeting, Miami, FL, 1976.
- [BNS26] A. E. Black, C. Nöbel, and R. Steiner. *Short circuit walks in fixed dimension*, page 563–573. Society for Industrial and Applied Mathematics, January 2026.
- [BR13] T. Brunsch and H. Röglin. Finding short paths on polytopes by the shadow vertex algorithm. In *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 279–290. Springer, 2013.
- [BSY18] S. Borgwardt, T. Stephen, and T. Yusun. On the circuit diameter conjecture. *Discrete & Computational Geometry*, 60(3):558–587, April 2018.
- [BV22] S. Borgwardt and C. Viss. Circuit walks in integral polyhedra. *Discrete Optimization*, 44:100566, 2022. arXiv:1712.01933.
- [Cam65] P. Camion. Characterization of totally unimodular matrices. *Proceedings of the American Mathematical Society*, 16(5):1068–1068, May 1965.
- [DF94] M. Dyer and A. Frieze. Random walks, totally unimodular matrices, and a randomised dual simplex algorithm. *Mathematical Programming*, 64(1):1–16, 1994.
- [DH16] D. Dadush and N. Hähnle. On the shadow simplex method for curved polyhedra. *Discrete & Computational Geometry*, 56(4):882–909, 2016.
- [DKK25] D. Dadush, S. Kober, and Z. K. Koh. On circuit diameter and straight line complexity. Unpublished manuscript, 2025.
- [DKN<sup>+</sup>24] D. Dadush, Z. K. Koh, B. Natura, N. Olver, and L. A. Végh. A strongly polynomial algorithm for linear programs with at most two non-zero entries per row or column. In *Proceedings of the 56th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 1916–1925, 2024.
- [DKNV24] D. Dadush, Z. K. Koh, B. Natura, and L. A. Végh. On circuit diameter bounds via circuit imbalances. *Mathematical Programming*, 206(1–2):631–662, July 2024.
- [DLHK12] J. A. De Loera, R. Hemmecke, and M. Köppe. *Algebraic and Geometric Ideas in the Theory of Discrete Optimization*. Society for Industrial and Applied Mathematics, USA, 2012.
- [DLHL15] 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(4):2494–2511, 2015.
- [DNV20] D. Dadush, B. Natura, and L. A. Végh. Revisiting Tardos’s framework for linear programming: Faster exact solutions using approximate solvers. In *Proc. 61st IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 931–942, 2020.
- [ENV22] F. Ekbatalani, B. Natura, and L. A. Végh. Circuit imbalance measures and linear programming. In *Surveys in Combinatorics 2022*, London Mathematical Society Lecture Note Series, pages 64–114. Cambridge University Press, 2022.
- [EV17] F. Eisenbrand and S. Vempala. Geometric random edge. *Mathematical Programming*, 164(1-2):325–339, 2017.

- [Ful67] D. R. Fulkerson. *Networks, Frames, Blocking Systems*. RAND Corporation, Santa Monica, CA, 1967.
- [GD21] J. B. Gauthier and J. Desrosiers. The minimum mean cycle-canceling algorithm for linear programs. *European Journal of Operational Research*, 2021.
- [GT89] A. V. Goldberg and R. E. Tarjan. Finding minimum-cost circulations by canceling negative cycles. *Journal of the ACM (JACM)*, 36(4):873–886, 1989.
- [Kal92] G. Kalai. A subexponential randomized simplex algorithm. In *Proceedings of the 24th annual ACM Symposium on Theory of Computing*, pages 475–482, 1992.
- [Kar84] N. Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the 16th Annual ACM Symposium on Theory of Computing (STOC)*, pages 302–311, 1984.
- [Kha79] L. G. Khachiyan. A polynomial algorithm in linear programming. In *Doklady Akademii Nauk SSSR*, volume 244, pages 1093–1096, 1979.
- [KK92] G. Kalai and D. J. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bulletin of the American Mathematical Society*, 26(2):315–316, 1992.
- [KPS19] 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.
- [Lee89] J. Lee. Subspaces with well-scaled frames. *Linear Algebra and its Applications*, 114:21–56, 1989.
- [LKS22] J. A. D. Loera, S. Kafer, and L. Sanità. Pivot rules for circuit-augmentation algorithms in linear optimization. *SIAM Journal on Optimization*, 32(3):2156–2179, 2022. arXiv:1909.12863.
- [NS24] C. Nöbel and R. Steiner. Complexity of polytope diameters via perfect matchings. *arXiv preprint*, 2024.
- [Roc67] R. T. Rockafellar. The elementary vectors of a subspace of  $\mathbb{R}^n$ . In R. C. Bose and T. A. Dowling, editors, *Combinatorial Mathematics and Its Applications, Proceedings of the Chapel Hill Conference*, pages 104–127, 1967.
- [San12] F. Santos. A counterexample to the Hirsch conjecture. *Annals of Mathematics*, pages 383–412, 2012.
- [Sma98] S. Smale. Mathematical problems for the next century. *The Mathematical Intelligencer*, 20:7–15, 1998.
- [ST97] B. Sturmfels and R. R. Thomas. Variation of cost functions in integer programming. *Mathematical Programming*, 77(2):357–387, 1997.
- [Suk17] N. Sukegawa. Improving bounds on the diameter of a polyhedron in high dimensions. *Discrete Mathematics*, 340(9):2134–2142, 2017.
- [SY15] T. Stephen and T. Yusun. The circuit diameter of the Klee–Walkup polyhedron. *Electronic Notes in Discrete Mathematics*, 49:505–512, 2015. arXiv:1503.05252.
- [Wul25] L. Wulf. Computing the polytope diameter is even harder than NP-hard. *arXiv preprint*, 2025.