Short-step Methods Are Not Strongly Polynomial-Time

Manru Zong* Yin Tat Lee† Man-Chung Yue‡

Abstract

Short-step methods are an important class of algorithms for solving convex constrained optimization problems. In this short paper, we show that under very mild assumptions on the self-concordant barrier and the width of the ℓ_2-neighbourhood, any short-step interior-point method is not strongly polynomial-time.

1 Introduction

An algorithm for solving linear programming problems is said to be strongly polynomial-time if the required number of arithmetic operations is a polynomial in the numbers of variables and constraints, independent of the bit-length for encoding the problem instance. A major open problem in optimization and computer science is whether there exists a strongly polynomial-time algorithm for solving linear programming problems. The purpose of this short note is to prove that short-step methods are not strongly polynomial-time.[1]

2 Preliminaries

Consider the linear programming problem

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & Ax \leq b, \end{array} \tag{P}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$. Denote the feasible region of problem (P) and its interior by $\mathcal{F}$ and $\mathcal{F}^\circ$, respectively. We assume that $\mathcal{F}$ is bounded and that $\mathcal{F}^\circ$ is non-empty. For simplicity, we denote the optimality gap of a feasible point $x \in \mathcal{F}$ by

$$\mathsf{gap}(x) = c^\top x - \min_{x' \in \mathcal{F}} c^\top x'.$$

The notion of self-concordant barriers plays vital role in the study of interior-point methods. Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a proper convex domain, *i.e.*, a convex set with non-empty

*The University of Hong Kong. E-mail: `zongmanru@connect.hku.hk`
†University of Washington. E-mail: `yintat@uw.edu`
‡Corresponding Author: The University of Hong Kong. E-mail: `mcyue@hku.hk`

[1] After we finished this work, we were informed that a similar result has been obtained independently by another group [2] around the same time, via a different approach.

interior and containing no 1-dimensional affine subspace. A function $\phi : \operatorname{int}(\mathcal{K}) \to \mathbb{R}$ is said to be a barrier on $\mathcal{K}$ if $\phi(x) \to +\infty \quad \text{as} \quad x \to \partial\mathcal{K}$. A three times continuously differentiable convex function ϕ is said to be self-concordant on $\mathcal{K}$ if for any $x \in \operatorname{int}(\mathcal{K})$ and $h \in \mathbb{R}^n$,

$$\left|\mathrm{D}^3\phi(x)[h,h,h]\right| \le 2\left(\mathrm{D}^2\phi(x)[h,h]\right)^{\frac{3}{2}}.$$

If ϕ additionally satisfies that for any $x \in \operatorname{int}(\mathcal{K})$ and $h \in \mathbb{R}^n$,

$$|\mathrm{D}\phi(x)[h]| \le \left(\nu\, \mathrm{D}^2\phi(x)[h,h]\right)^{\frac{1}{2}},$$

then ϕ is said to be ν-self-concordant on $\mathcal{K}$. This paper focuses on the proper convex domain $\mathcal{K} = \mathcal{F}$, which is a polytope defined by the linear inequalities $Ax \le b$. A standard self-concordant barrier on $\mathcal{F}$ is the logarithmic barrier

$$\phi_{\ln}(x) = -\sum_{i=1}^{m} \ln(b - Ax)_i.$$

Given a self-concordant barrier ϕ on $\mathcal{F}$, we consider the problem

$$\begin{aligned} \text{minimize} \quad & c^\top x + \mu\,\phi(x) \\ \text{subject to} \quad & Ax < b. \end{aligned} \tag{P$_\mu$}$$

From [3, Section 5.3.4], for any $\mu > 0$, there exists a unique minimizer $x^\phi(\mu)$ to problem (P$_\mu$). We call $x^\phi(\mu)$ the μ-analytic center of problem (P) associated with the self-concordant barrier ϕ. The central path of problem (P) associated with ϕ is then defined as (the image of) the curve $\mu \mapsto x^\phi(\mu)$ for $\mu > 0$. By [3, Theorem 5.3.10], $x^\phi(\mu)$ converges to an optimal solution to problem (P) as $\mu \to 0$.

By [3, Theorem 5.1.6], $\nabla^2\phi(x)$ is positive definite for any $x \in \mathcal{F}^\circ$. This allows us to define a norm

$$\|h\|_x = \sqrt{h^\top \nabla^2\phi(x) h}, \quad h \in \mathbb{R}^n,$$

and its dual norm

$$\|h\|_x^* = \sqrt{h^\top \left(\nabla^2\phi(x)\right)^{-1} h}, \quad h \in \mathbb{R}^n.$$

For any $\theta \in (0,1)$ and $\mu > 0$, the ℓ_2-neighbourhood of problem (P) is defined as

$$\mathcal{N}_\theta^\phi(\mu) = \left\{x \in \mathcal{F}^\circ : \|c + \mu\,\nabla\phi(x)\|_x^* \le \theta\mu\right\}.$$

Note that $c + \mu\nabla\phi(x)$ is the gradient of the objective function of problem (P$_\mu$). The ℓ_2-neighbourhood is important and a natural choice for the design of path-following interior-point methods since the Newton's method applied to problem (P$_\mu$) converges quadratically to the optimal solution $x^\phi(\mu)$ [3, Theorem 5.2.2]. We write

$$\mathcal{N}_\theta^\phi = \bigcup_{\mu > 0} \mathcal{N}_\theta^\phi(\mu),$$

which is also called the ℓ_2-neighbourhood of problem (P). By the optimality condition of problem (P$_\mu$), we can see that $x^\phi(\mu) \in \mathcal{N}_\theta^\phi(\mu)$ for any $\theta \in (0,1)$ and $\mu > 0$.

A short-step method associated with ϕ is defined as an algorithm that generates a sequence of iterates $\{x^k\}_{k\geq 0}$ such that the polygonal (*i.e.*, continuous piecewise linear) curve formed using the sequence $\{x^k\}_{k\geq 0}$ is contained in the ℓ_2-neighbourhood $\mathcal{N}_\theta^\phi$ of problem (P) for some $\theta \in (0,1)$, where k is the iteration counter, or more precisely,

$$\bigcup_{k\geq 0} [x^k, x^{k+1}] \subseteq \mathcal{N}_\theta^\phi.$$

Another popular choice of the neighbourhood for path-following algorithms for solving linear programming problems is the so-called wide neighbourhood [4]:

$$\begin{aligned}\mathcal{W}_\theta(\mu) = \Big\{ x \in \mathcal{F}^\circ : \exists\, y \in \mathbb{R}^m_+ \text{ such that } A^\top y = -c,\ y^\top (b - Ax) = m\mu,\\ \text{and } y_i\,(b-Ax)_i \geq (1-\theta)\mu,\ i = 1,\ldots,m \Big\}.\end{aligned}$$

Similarly to the ℓ_2-neighbourhood, we write

$$\mathcal{W}_\theta = \bigcup_{\mu>0} \mathcal{W}_\theta(\mu).$$

3 Main Results

Our non-strong polynomiality result is based on the following family of linear programs introduced in [1]:

$$\begin{array}{rll} \text{minimize} & x_1 & \\ \text{subject to} & x_1 \leq t^2, & \\ & x_2 \leq t, & \\ & x_{2j+1} \leq t\, x_{2j-1}, \quad j = 1,\ldots,r-1, & \qquad \mathbf{LW}_r(t) \\ & x_{2j+1} \leq t\, x_{2j}, \quad j = 1,\ldots,r-1, & \\ & x_{2j+2} \leq t^{1-1/2^j}(x_{2j-1} + x_{2j}), \quad j = 1,\ldots,r-1, & \\ & x_{2r-1},\ x_{2r} \geq 0, & \end{array}$$

where $t > 1$ is a real number and $r \geq 1$ is an integer. The notation $\mathbf{LW}_r(t)$ follows from [1] and signifies that the central path of this linear program is long and winding.

To distinguish the properties of problem $\mathbf{LW}_r(t)$ from those of a general linear program, we introduce an extra subscript t to the notations. For instance, the feasible region, the μ-analytic center associated with a self-concordant barrier ϕ and the wide neighbourhood of problem $\mathbf{LW}_r(t)$ are denoted as $\mathcal{F}_t$, $x_t^\phi(\mu)$ and $\mathcal{W}_{\theta,t}$, respectively.

We can now present the main results of this paper, whose proofs are deferred to Section 4. The main contribution of this paper is the non-strong polynomiality of short-step methods.

Theorem 1 (Non-strong Polynomiality of Short-step Methods)**.** *Consider* $\mathbf{LW}_r(t)$ *for a sufficiently large* $t > 1$*. Let* $\theta \in (0, (\sqrt{69}-3)/10)$*,* ϕ *be a* ν*-self-concordant barrier on* $\mathcal{F}_t$

with ν independent of t and $\{x^k\}_{k\geq 0}$ be a sequence of iterates generated by a short-step method associated with ϕ. Suppose that

$$\mathsf{gap}(x^0) \geq \frac{(1+\beta_\theta)(3r+1)(1+\nu+2\sqrt{\nu})\sqrt{t}}{1-\beta_\theta} \quad \text{and} \quad \mathsf{gap}(x^K) \leq \frac{1-\beta_\theta}{2(1+\beta_\theta)(1+C_{3r+1})},$$

where $\beta_\theta \in (0,1)$ is a constant depending only on θ defined in (8). Then, $K \geq 2^{r-3}$.

We should emphasize that although the self-concordance parameter ν in Theorem 1 is assumed to be independent of t, it could possibly depend on the numbers of variables and constraints of problem $\mathbf{LW}_r(t)$. This subsumes almost all known barriers. Furthermore, we note that the requirement on the initial optimality gap is only very mild. Indeed, it is customary for interior-point methods to start at the ∞-analytic center $x^\phi(\infty)$. Using Lemma 2, we can check that $\mathsf{gap}(x^\phi(\infty)) = \Omega(\frac{t}{\nu})$, see also the proof of Lemma 4.

The cruxes of the proof of Theorem 1 are the following lemmas. The first shows a certain equivalence among the central paths of all self-concordant barriers. Geometrically speaking, it guarantees that when restricted to a constant-cost slice, all central paths are approximately equally close to (or away from) the boundary of the feasible region.

Lemma 1 (Equivalence of Central Paths)**.** *Consider the linear program* (P)*. Let ϕ and ψ be ν_ϕ- and ν_ψ-self-concordant barriers on the feasible region $\mathcal{F}$. Let $\mu, \eta > 0$ be such that $c^\top x^\phi(\mu) = c^\top x^\psi(\eta)$. Then, for any $i = 1, \ldots, m$,*

$$(1+\nu_\phi+2\sqrt{\nu_\phi})^{-1} \leq \frac{(b-Ax^\phi(\mu))_i}{(b-Ax^\psi(\eta))_i} \leq (1+\nu_\psi+2\sqrt{\nu_\psi}).$$

The second one bounds the optimality gap of analytic centers.

Lemma 2 (Optimality Gap of Analytic Center)**.** *Consider the linear program* (P)*. Let ϕ be a ν-self-concordant barrier on the feasible region $\mathcal{F}$. Then, for any $\mu > 0$,*

$$\min\left\{\frac{\mu}{2}, \frac{\rho\|c\|_2}{2\nu+4\sqrt{\nu}}\right\} \leq \mathsf{gap}(x^\phi(\mu)) \leq \mu\nu,$$

where $\rho > 0$ is the radius of the largest ball contained in $\mathcal{F}$.

It should be pointed out that the upper bound in Lemma 2 is not new and can be found in [3, Theorem 5.3.10]. The novelty of the lemma lies in the lower bound.

4 Proofs

Throughout this section, given a vector $\bar{x} \in \mathbb{R}^n$ and a positive definite matrix Q, we let $\mathcal{E}(Q,\bar{x}) = \{x \in \mathbb{R}^n : (x-\bar{x})^\top Q(x-\bar{x}) \leq 1\}$. Also, for any $\nu > 0$, we let $C_\nu = \nu + 2\sqrt{\nu}$.

We first collect a simple lemma about linear optimization over ellipsoids.

Lemma 3. *Let $\bar{x} \in \mathbb{R}^n$, $a \in \mathbb{R}^n$ and $Q \in \mathbb{R}^{n\times n}$ be a positive definite matrix. Then,*

$$\max_{x\in\mathcal{E}(Q,\bar{x})} a^\top x = a^\top\bar{x} + \sqrt{a^\top Q^{-1}a} \quad \text{and} \quad \min_{x\in\mathcal{E}(Q,\bar{x})} a^\top x = a^\top\bar{x} - \sqrt{a^\top Q^{-1}a}.$$

The proof of Lemma 3 uses only elementary optimality arguments and is thus omitted.

Proof of Lemma 1. Let $\mathcal{H} = \{x \in \mathbb{R}^n : c^\top x = c^\top x^\phi(\mu)\}$. By [3, Theorem 5.1.5], we have

$$\mathcal{E}(\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu)) \subseteq \mathcal{F} \quad \text{and} \quad \mathcal{E}(\nabla^2\psi(x^\psi(\eta)), x^\psi(\eta)) \subseteq \mathcal{F}.$$

Combining [3, Theorem 5.3.8] with these inclusions yields

$$\mathcal{E}(\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu)) \cap \mathcal{H} \subseteq \mathcal{F} \cap \mathcal{H} \subseteq \mathcal{E}(C_{\nu_\phi}^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu)) \cap \mathcal{H},$$

and

$$\mathcal{E}(\nabla^2\psi(x^\psi(\eta)), x^\psi(\eta)) \cap \mathcal{H} \subseteq \mathcal{F} \cap \mathcal{H} \subseteq \mathcal{E}(C_{\nu_\psi}^{-2}\nabla^2\psi(x^\psi(\eta)), x^\psi(\eta)) \cap \mathcal{H},$$

which implies, respectively,

$$x^\psi(\eta) \in \mathcal{E}(C_{\nu_\phi}^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu)) \cap \mathcal{H} \quad \text{and} \quad x^\phi(\mu) \in \mathcal{E}(C_{\nu_\psi}^{-2}\nabla^2\psi(x^\psi(\eta)), x^\psi(\eta)) \cap \mathcal{H}.$$

Using Lemma 3, for any $i = 1, \ldots, m$,

$$\max_{x' \in \mathcal{E}(C_{\nu_\phi}^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu))} (b - Ax')_i = (b - Ax^\phi(\mu))_i + C_{\nu_\phi}\sqrt{a_i^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} a_i}, \tag{1}$$

and

$$\min_{x' \in \mathcal{E}(C_{\nu_\phi}^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu))} (b - Ax')_i = (b - Ax^\phi(\mu))_i - C_{\nu_\phi}\sqrt{a_i^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} a_i}, \tag{2}$$

where $a_i^\top$ is the i-row of A. Also, since $Ax' \le b$ for any $x' \in \mathcal{E}(\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu)) \subseteq \mathcal{F}$, Lemma 3 implies that

$$0 \le \min_{x' \in \mathcal{E}(\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu))} (b - Ax')_i = (b - Ax^\phi(\mu))_i - \sqrt{a_i^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} a_i}.$$

It follows from (1), (2) and $x^\psi(\eta) \in \mathcal{E}(C_{\nu_\phi}^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu))$ that

$$\left|(b - Ax^\phi(\mu))_i - (b - Ax^\psi(\eta))_i\right| \le C_{\nu_\phi}\sqrt{a_i^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} a_i} \le C_{\nu_\phi}(b - Ax^\phi(\mu))_i.$$

Using the same arguments, we also obtain

$$\left|(b - Ax^\phi(\mu))_i - (b - Ax^\psi(\eta))_i\right| \le C_{\nu_\psi}\sqrt{a_i^\top \left(\nabla^2\psi(x^\psi(\eta))\right)^{-1} a_i} \le C_{\nu_\psi}(b - Ax^\psi(\eta))_i.$$

Thus, for any $i = 1, \ldots, m$,

$$(b - Ax^\psi(\eta))_i \le (1 + C_{\nu_\phi})(b - Ax^\phi(\mu))_i \quad \text{and} \quad (b - Ax^\phi(\mu))_i \le (1 + C_{\nu_\psi})(b - Ax^\psi(\eta))_i,$$

which completes the proof. $\square$

Proof of Lemma 2. The upper bound follows directly from [3, Theorem 5.3.10]. Therefore, we prove only the lower bound. Two cases are discussed separately:

$$\begin{aligned} &\nabla\phi(x^\phi(\mu)))^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} \nabla\phi(x^\phi(\mu))) \le \frac{1}{4} \\ \text{and}\quad &\nabla\phi(x^\phi(\mu)))^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} \nabla\phi(x^\phi(\mu))) > \frac{1}{4}. \end{aligned}$$

We first consider the case of

$$\nabla\phi(x^\phi(\mu)))^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} \nabla\phi(x^\phi(\mu))) \le \frac{1}{4}.$$

By [3, Lemma 5.1.5 and Theorem 5.2.1], we have

$$\frac{\|x^\phi(\mu) - x^\phi(\infty)\|^2_{x^\phi(\infty)}}{1 + \frac{2}{3}\|x^\phi(\mu) - x^\phi(\infty)\|_{x^\phi(\infty)}} \le \frac{(\frac{1}{4})^2}{1 - \frac{1}{4}} = \frac{1}{12},$$

where $x^\phi(\infty)$ is the ∞-analytic center, *i.e.*, the unique optimal solution to the problem

$$\begin{array}{ll} \text{minimize} & \phi(x) \\ \text{subject to} & Ax < b. \end{array}$$

Solving the quadratic inequality, we get

$$\|x^\phi(\mu) - x^\phi(\infty)\|_{x^\phi(\infty)} \le \frac{1}{2}. \tag{3}$$

Using Lemma 3 and inequality (3), we have that

$$\begin{aligned} \mathsf{gap}(x^\phi(\mu)) &\ge \min_{x \in \mathcal{E}(16\nabla^2\phi(x^\phi(\infty)), x^\phi(\infty))} \mathsf{gap}(x) \\ &= \mathsf{gap}(x^\phi(\infty)) - \frac{1}{2}\sqrt{c^\top \left(\nabla^2\phi(x^\phi(\infty))\right)^{-1} c}. \end{aligned} \tag{4}$$

Next, since $\mathcal{E}(\nabla^2\phi(x^\phi(\infty)), x^\phi(\infty)) \subseteq \mathcal{F}$, Lemma 3 implies

$$0 \le \min_{x \in \mathcal{E}(\nabla^2\phi(x^\phi(\infty)), x^\phi(\infty))} \mathsf{gap}(x) = \mathsf{gap}(x^\phi(\infty)) - \sqrt{c^\top \left(\nabla^2\phi(x^\phi(\infty))\right)^{-1} c}. \tag{5}$$

Combining the inequalities (4) and (5), we obtain

$$\mathsf{gap}(x^\phi(\mu)) \ge \frac{1}{2}\sqrt{c^\top \left(\nabla^2\phi(x^\phi(\infty))\right)^{-1} c}. \tag{6}$$

On the other hand, by supposition, the feasible region $\mathcal{F}$ contains a ball of radius ρ. By [3, Theorem 5.3.9], it is contained in the ellipsoid $\mathcal{E}(C_\nu^{-2}\nabla^2\phi(x^\phi(\infty)), x^\phi(\infty))$. Hence,

$$\mathcal{E}(\rho^{-2}I, x^\phi(\infty)) \subseteq \mathcal{E}(C_\nu^{-2}\nabla^2\phi(x^\phi(\infty)), x^\phi(\infty)),$$

where the left-hand side is a ball with center $x^\phi(\infty)$ of radius ρ. Using Lemma 3, we get

$$\begin{aligned} \mathsf{gap}(x^\phi(\infty)) + \rho\|c\|_2 &= \max_{x\in\mathcal{E}(\rho^{-2}I, x^\phi(\infty))} \mathsf{gap}(x) \le \max_{x\in\mathcal{E}(C_\nu^{-2}\nabla^2\phi(x^\phi(\infty)), x^\phi(\infty))} \mathsf{gap}(x) \\ &= \mathsf{gap}(x^\phi(\infty)) + (\nu + 2\sqrt{\nu})\sqrt{c^\top \left(\nabla^2\phi(x^\phi(\infty))\right)^{-1} c}, \end{aligned}$$

which, upon substitution into (6), yields

$$\mathsf{gap}(x^\phi(\mu)) \ge \frac{\rho\|c\|_2}{2(\nu + 2\sqrt{\nu})}.$$

Then, we consider the case of

$$\nabla\phi(x^\phi(\mu)))^\top \nabla^2\phi(x^\phi(\mu))\nabla\phi(x^\phi(\mu))) > \frac{1}{4}.$$

Using the optimality condition of problem (P_μ), we get $c + \mu\nabla\phi(x^\phi(\mu)) = 0$. Therefore,

$$c^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} c > \frac{1}{4}\mu^2. \tag{7}$$

Since $\mathcal{E}(\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu)) \subseteq \mathcal{F}$, by Lemma 3,

$$0 \le \min_{x\in\mathcal{E}(\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu))} = \mathsf{gap}(x^\phi(\mu)) - \sqrt{c^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} c},$$

which, together with (7), yields

$$\mathsf{gap}(x^\phi(\mu)) > \frac{\mu}{2}.$$

This completes the proof. □

We then prove a proposition that compares the slack of the μ-analytic center with that of any point in its ℓ_2-neighbourhood.

Proposition 1. *Consider the linear program* (P). *Let $\theta \in (0, (\sqrt{69} - 3)/10)$, $\mu > 0$, ϕ be a self-concordant function on $\mathcal{F}$ and $x \in \mathcal{N}_\theta^\phi(\mu)$. Then, for any $i = 1, \ldots, m$,*

$$(1 - \beta_\theta)(b - Ax^\phi(\mu))_i \le (b - Ax)_i \le (1 + \beta_\theta)(b - Ax^\phi(\mu))_i,$$

where $\beta_\theta \in (0, 1)$ is a constant depending only on θ defined in (8).

Proof. Let $x \in \mathcal{N}_\theta^\phi(\mu)$. By [3, Lemma 5.1.5 and Theorem 5.2.1] and the definition of ℓ_2-neighbourhood,

$$\frac{\|x - x^\phi(\mu)\|_{x^\phi(\mu)}^2}{1 + \frac{2}{3}\|x - x^\phi(\mu)\|_{x^\phi(\mu)}} \le \frac{\theta^2}{1 - \theta}.$$

Solving the quadratic inequality yields $\|x - x^\phi(\mu)\|_{x^\phi(\mu)} \le \beta_\theta$, where

$$\beta_\theta = \frac{1}{3}\left(\frac{\theta^2}{1-\theta} + \sqrt{\frac{\theta^4}{(1-\theta)^2} + \frac{9\theta^2}{1-\theta}}\right). \tag{8}$$

We note that $\beta_\theta < 1$ whenever $\theta \in (0, (\sqrt{69}-3)/10)$. Using Lemma 3, we get

$$\max_{x' \in \mathcal{E}(\beta_\theta^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu))} (b - Ax')_i = (b - Ax^\phi(\mu))_i + \beta_\theta \sqrt{a_i^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} a_i},$$

and

$$\min_{x' \in \mathcal{E}(\beta_\theta^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu))} (b - Ax')_i = (b - Ax^\phi(\mu))_i - \beta_\theta \sqrt{a_i^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} a_i} \geq 0,$$

where the inequality follows from that $Ax' \leq b$ for any $x' \in \mathcal{E}(\beta_\theta^{-2}\nabla^2\phi(x^\phi(\mu)), x^\phi(\mu)) \subseteq \mathcal{F}$. Therefore, we have

$$\left|(b - Ax)_i - (b - Ax^\phi(\mu))_i\right| \leq \beta_\theta \sqrt{a_i^\top \left(\nabla^2\phi(x^\phi(\mu))\right)^{-1} a_i} \leq \beta_\theta (b - Ax^\phi(\mu))_i.$$

This completes the proof. $\square$

The next proposition compares the optimality gap of the μ-analytic center with that of any point in its ℓ_2-neighbourhood.

Proposition 2. *Consider the linear program* (P). *Let $\theta \in (0, (\sqrt{69}-3)/10)$, $\mu > 0$, ϕ be a self-concordant function on $\mathcal{F}$ and $x \in \mathcal{N}_\theta^\phi(\mu)$. Then,*

$$(1 - \beta_\theta)\mathsf{gap}(x^\phi(\mu)) \leq \mathsf{gap}(x) \leq (1 + \beta_\theta)\mathsf{gap}(x^\phi(\mu)).$$

The proof of Proposition 2 uses the same arguments as in the proof of Proposition 1 and is therefore omitted.

We will need to compare the μ-analytic center associated with ϕ and the η-analytic center associated with the logarithmic barrier $\phi_{\ln}$ along the constant-cost slices of $\mathcal{F}_t$.

Lemma 4. *Consider the linear program $\mathbf{LW}_r(t)$. For any self-concordant barrier ϕ on $\mathcal{F}_t$ and $\mu > 0$ satisfying*

$$\mathsf{gap}(x^\phi(\mu)) < \frac{t}{4(3r+1) + 8\sqrt{3r+1}},$$

there exists a unique $\eta(\mu) > 0$ such that

$$\frac{\eta(\mu)}{2} \leq \mathsf{gap}(x^\phi(\mu)) = \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu))) \leq (3r+1)\eta(\mu).$$

Proof. We first prove the existence of $\eta(\mu)$ satisfying $\mathsf{gap}(x^\phi(\mu)) = \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu)))$. Since $\mu \mapsto \mathsf{gap}(x^\phi(\mu))$ and $\eta \mapsto \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu)))$ are both continuous and strictly increasing functions attaining the minimum value 0 at 0, it suffices to show that for any $\mu > 0$ with

$$\mathsf{gap}(x^\phi(\mu)) < \frac{t}{4(3r+1) + 8\sqrt{3r+1}},$$

there exists $\eta > 0$ such that $\mathsf{gap}(x^\phi(\mu)) \leq \mathsf{gap}(x^{\phi_{\ln}}(\eta))$. To show this, we note that the feasible region $\mathcal{F}_t$ contains a hypercube of side length t, which in turn contains a ball of radius $\frac{t}{2}$. Also, the cost vector of problem $\mathbf{LW}_r(t)$ is $c_t = (1, 0, \ldots, 0)^\top$ and hence $\|c_t\|_2 = 1$. Furthermore, by [3, Section 5.3], $\phi_{\ln}$ is $(3r+1)$-self-concordant on $\mathcal{F}_t$. Lemma 2 then implies that for a large enough $\eta > 0$,

$$\mathsf{gap}(x^\phi(\mu)) < \min\left\{\frac{t}{4(3r+1) + 8\sqrt{3r+1}}, \frac{\eta}{2}\right\} \leq \mathsf{gap}(x^{\phi_{\ln}}(\eta)).$$

Therefore, there exists a unique $\eta(\mu) > 0$ with $\mathsf{gap}(x^\phi(\mu)) = \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu)))$.

It remains to prove the inequalities for $\mathsf{gap}(x^\phi(\mu)) = \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu)))$. By Lemma 2,

$$\min\left\{\frac{t}{4(3r+1) + 8\sqrt{3r+1}}, \frac{\eta(\mu)}{2}\right\} \leq \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu))) \leq \eta(\mu)(3r+1).$$

Since

$$\mathsf{gap}(x^\phi(\mu)) = \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu))) < \frac{t}{4(3r+1) + 8\sqrt{3r+1}},$$

we get

$$\frac{\eta(\mu)}{2} \leq \mathsf{gap}(x^\phi(\mu)) = \mathsf{gap}(x^{\phi_{\ln}}(\eta(\mu))) \leq \eta(\mu)(3r+1).$$

This completes the proof. □

We will also need the following remarkable result from [1, Theorem 29].

Theorem 2. *Let $\omega \in (0,1)$. Consider the linear program $\mathbf{LW}_r(t)$ for*

$$t > \left(\frac{2(5r-1)(10r-1)^4((10r-2)!)^8}{1-\omega}\right)^{2^{r+2}}.$$

Suppose that

$$[x^0, x^1] \cup [x^1, x^2] \cup \cdots \cup [x^{K-1}, x^K] \subseteq \mathcal{W}_{\omega,t},$$

with $x^0 \in \mathcal{W}_{\omega,t}(\eta_0)$ and $x^K \in \mathcal{W}_{\omega,t}(\eta_K)$ for some $\eta_0 \geq \sqrt{t}$ and $\eta_K \leq 1$. Then, $K \geq 2^{r-3}$.

Some remarks are in order. First, the original statement of [1, Theorem 29] concerns iterates of a longer vector consisting of primal, dual and slack variables. But Theorem 2 concerns only the primal variables and hence is seemingly stronger than [1, Theorem 29], because, for a polygonal curve in $\mathbb{R}^N$ with K pieces, its projection to a lower dimensional space could possibly have less than K pieces in general. However, upon a close inspection of the proof of [1, Theorem 29] and [1, Section 6.2], we find that Theorem 2 holds for the linear program $\mathbf{LW}_r(t)$. Using the language of [1], the reason is that for the tropical central path of the linear program $\mathbf{LW}_r(t)$, even the projection onto the primal variables is already a polygonal curve with a huge number of pieces. Second, in the original statement of [1, Theorem 29], the requirement on η_0 is that $\eta_0 \geq t^2$, and the conclusion is that $K \geq 2^{r-1}$. By considering a shorter part of the tropical central path of $\mathbf{LW}_r(t)$, we find that replacing the requirement on η_0 with the weaker bound of $\eta_0 \geq \sqrt{t}$ leads the weaker conclusion of $K \geq 2^{r-3}$, see [1, Section 4.3].

Proof of Theorem 1. Let

$$\omega = 1 - \frac{(1-\beta_\theta)}{(1+\beta_\theta)(1+C_\nu)(1+C_{3r+1})}$$

and

$$t > \left(\frac{2(5r-1)(10r-1)^4((10r-2)!)^8}{1-\omega} \right)^{2^{r+2}}.$$

Note that $\omega \in (0,1)$.

Reduction

Without loss of generality, we can assume that for any $x \in [x^0, x^1] \cup \cdots \cup [x^{K-1}, x^K]$,

$$\mathsf{gap}(x^0) = \frac{(1+\beta_\theta)(3r+1)(1+C_\nu)\sqrt{t}}{1-\beta_\theta} \geq \mathsf{gap}(x) = \frac{1-\beta_\theta}{2(1+\beta_\theta)(1+C_{3r+1})} = \mathsf{gap}(x^K). \quad (9)$$

Indeed, if this is not the case, because of the continuity of gap and the assumption that

$$\mathsf{gap}(x^0) \geq \frac{(1+\beta_\theta)(3r+1)(1+C_\nu)\sqrt{t}}{1-\beta_\theta} \quad \text{and} \quad \mathsf{gap}(x^K) \leq \frac{1-\beta_\theta}{2(1+\beta_\theta)(1+C_{3r+1})},$$

we can choose a sub-curve (connected subset) of $[x^0, x^1] \cup \cdots \cup [x^{K-1}, x^K]$ that satisfies all the assumptions of Theorem 1 as well as assumption (9).

Our goal is to show that $[x^0, x^1] \cup \cdots \cup [x^{K-1}, x^K] \subseteq \mathcal{W}_{\omega,t}$ and that $x^0 \in \mathcal{W}_{\omega,t}(\eta_0)$ and $x^K \in \mathcal{W}_{\omega,t}(\eta_K)$, for some $\eta_0 \geq \sqrt{t}$ and $\eta_K \leq 1$. If this can be proved, the desired conclusion would then follow from Theorem 2.

Proving $[x^0, x^1] \cup \cdots \cup [x^{K-1}, x^K] \subseteq \mathcal{W}_{\omega,t}$

Let $x \in [x^0, x^1] \cup \cdots \cup [x^{K-1}, x^K]$. Then, $x \in \mathcal{N}^{\phi}_{\theta,t}(\mu)$ for some $\mu > 0$. By assumption (9) and Proposition 2, we have

$$\mathsf{gap}(x^\phi(\mu)) \leq \frac{(1+\beta_\theta)^2(3r+1)(1+C_\nu)\sqrt{t}}{1-\beta_\theta} < \frac{t}{4(3r+1) + 8\sqrt{3r+1}}.$$

Using Lemma 4, there exists a unique $\eta(\mu) > 0$ such that

$$c_t^\top x_t^\phi(\mu) = c_t^\top x_t^{\phi_{\ln}}(\eta(\mu)).$$

By Proposition 1 and Lemma 1, for any $i = 1, \ldots, 3r+1$,

$$\frac{(b_t - A_t x)_i}{(b_t - A_t x_t^{\phi_{\ln}}(\eta(\mu)))_i} = \frac{(b_t - A_t x)_i}{(b_t - A_t x_t^{\phi}(\mu))_i} \cdot \frac{(b_t - A_t x_t^{\phi}(\mu))_i}{(b_t - A_t x_t^{\phi_{\ln}}(\eta(\mu)))_i} \geq \frac{(1-\beta_\theta)}{(1+C_\nu)}, \quad (10)$$

From the definition of the logarithmic barrier $\phi_{\ln}$ and the optimality conditions of problem (P_μ), there exists $y \in \mathbb{R}^{3r+1}_+$ such that $A_t^\top y = -c_t$ and

$$y_i(b_t - A_t x_t^{\phi_{\ln}}(\eta(\mu)))_i = \eta(\mu), \quad i = 1, \ldots, 3r+1.$$

Averaging these $3r+1$ equalities and then using Proposition 1, Lemma 1 and the fact that $\phi_{\ln}$ is $(3r+1)$-self-concordant on $\mathcal{F}_t$, we get

$$\begin{aligned}\eta(\mu) &= \frac{1}{3r+1}\sum_{i=1}^{3r+1} y_i(b_t - A_t x_t^{\phi_{\ln}}(\eta(\mu)))_i \\ &\geq \frac{(1+C_{3r+1})^{-1}}{3r+1}\sum_{i=1}^{3r+1} y_i(b_t - A_t x_t^{\phi}(\mu))_i \\ &\geq (1+C_{3r+1})^{-1}(1+\beta_\theta)^{-1}\frac{y^\top(b_t - A_t x)}{3r+1}.\end{aligned} \tag{11}$$

Hence, using (10) and (11), for any $i = 1, \ldots, 3r+1$,

$$\begin{aligned}y_i(b_t - A_t x)_i &\geq \frac{(1-\beta_\theta)}{(1+C_\nu)}\cdot y_i(b_t - A_t x_t^{\phi_{\ln}}(\eta(\mu)))_i = \frac{(1-\beta_\theta)}{(1+C_\nu)}\cdot\eta(\mu) \\ &\geq \frac{(1-\beta_\theta)}{(1+\beta_\theta)(1+C_\nu)(1+C_{3r+1})}\cdot\frac{y^\top(b_t - A_t x)}{3r+1},\end{aligned} \tag{12}$$

which implies that $x \in \mathcal{W}_{\omega,t}(\eta_x)$ with

$$\eta_x = \frac{y^\top(b_t - A_t x)}{3r+1}. \tag{13}$$

Proving $x^0 \in \mathcal{W}_{\omega,t}(\eta_0)$ and $x^K \in \mathcal{W}_{\omega,t}(\eta_K)$ for some $\eta_0 \geq \sqrt{t}$ and $\eta_K \leq 1$

Using definition (13), inequality (11), Lemma 4 and Proposition 2, we get

$$\begin{aligned}\eta_{x^K} &= \frac{y^\top(b_t - A_t x^K)}{3r+1} \leq (1+\beta_\theta)(1+C_{3r+1})\eta(\mu_K) \leq 2(1+\beta_\theta)(1+C_{3r+1})\mathsf{gap}(x^\phi(\mu_K)) \\ &\leq \frac{2(1+\beta_\theta)(1+C_{3r+1})}{1-\beta_\theta}\mathsf{gap}(x^K) \;\leq 1,\end{aligned}$$

where $x^K \in \mathcal{N}^\phi_{\theta,t}(\mu_K)$. Similarly, using definition (13), inequality (12), Lemma 4 and Proposition 2, we get

$$\begin{aligned}\eta_{x^0} &= \frac{y^\top(b_t - A_t x^0)}{3r+1} \geq \frac{1-\beta_\theta}{1+C_\nu}\eta(\mu^0) \geq \frac{1-\beta_\theta}{(3r+1)(1+C_\nu)}\mathsf{gap}(x^\phi(\mu_0)) \\ &\geq \frac{1-\beta_\theta}{(1+\beta_\theta)(3r+1)(1+C_\nu)}\mathsf{gap}(x^0) \geq \sqrt{t},\end{aligned}$$

where $x^0 \in \mathcal{N}^\phi_{\theta,t}(\mu_0)$ Taking $\eta_0 = \eta_{x^0}$ and $\eta_K = \eta_{x^0}$, the proof is completed. □

Acknowledgments

The authors thank Defeng Sun, Kim-Chuan Toh and Stephen Wright for their helpful discussions.

Statements and Declarations

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] 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.

[2] X. Allamigeon, S. Gaubert, and N. Vandame. No self-concordant barrier interior point method is strongly polynomial. *arXiv preprint arXiv:2201.02186*, 2022.

[3] Y. Nesterov. *Lectures on Convex Optimization.* Springer, 2018.

[4] S. J. Wright. *Primal-dual Interior-point Methods.* SIAM, 1997.