An inexact infeasible arc-search interior-point method for linear programming problems

Einosuke Iida*and Makoto Yamashita†

2024/09/08

Abstract

Inexact interior-point methods (IPMs) are a type of interior-point methods that inexactly solve the linear equation system for obtaining the search direction. On the other hand, arc-search IPMs approximate the central path with an ellipsoidal arc obtained by solving two linear equation systems in each iteration, while conventional line-search IPMs solve one linear system. Therefore, the improvement due to the inexact solutions of the linear equation systems can be more beneficial in arc-search IPMs than conventional IPMs. In this paper, we propose an inexact infeasible arc-search interior-point method. We establish that the proposed method is a polynomial-time algorithm through its convergence analysis. The numerical experiments for the large benchmark problems show that the proposed method using the conjugate gradient method as the inexact linear system solver can reduce both of the number of iterations and the computation time compared to the existing inexact IPM due to the reduction in computational complexity by the arc-search. Andmore, it can reduce the computation time compared to the existing exact IPMs because the dependence of the computational complexity on the dimension n of the coefficient matrix is smaller for the conjugate gradient method than for the Cholesky factorization.

Keywords: interior-point method, arc-search, inexact IPM, infeasible IPM, linear programming.

1 Introduction

Linear programming problems (LPs) have had an important role in both theoretical analysis and practical applications, and many methods have been studied for solving LPs efficiently. Since an interior-point method (IPM) was first proposed by Karmarkar [16], IPMs have been extended of optimization problems, for example, second-order cone programming and semidefinite programming. Many variations of the IPM have been

*Department of Mathematical and Computing Science, Tokyo Institute of Technology
†Department of Mathematical and Computing Science, Tokyo Institute of Technology.

proposed, such as the primal-dual IPM [18], Mehrotra's predictor-corrector method [21], and recently, two-dimensional search IPMs [30].

Inexact IPMs are one of such variations and they inexactly solve a linear equation system (LES) for obtaining the search direction in each iteration. An inexact IPM was first proposed for solving a constrained system of equations by Bellavia [3] and it has been extended for LPs [22, 1]. The inexact IPMs have recently gained much attention due to their relevance to quantum computing. Quantum linear system algorithms (QLSAs) have the potential to solve LESs fast; their complexity has a better dependence on the size of variables and the number of constraints but a worse one on other parameters compared to that on classical computers [8]. Recently, inexact IPMs using the QLSA called quantum interior-point methods are proposed in [17, 32].

On the other hand, studies to reduce the number of iterations in IPMs have also contributed to improving the numerical performance. The higher-order algorithms using second-order or higher derivatives in the framework of IPMs have been studied [24, 21, 11, 19, 10], but these sometimes have a worse polynomial bound, or the analysis of computational complexity is not simple. An arc-search IPM is the one of the higher-order algorithms originally proposed by Yang [34]. IPMs numerically trace a trajectory to an optimal solution called the central path. Standard IPMs find the next iterate on a straight line that approximates the central path by computing the search direction; such IPMs are called line-search IPMs in this paper. In contrast, arc-search IPMs employ an ellipsoidal arc for the approximation. Since the central path is generally a smooth curve, the ellipsoidal arc can approximate the central path better than the straight line, and a reduction in the number of iterations can be expected. Several studies [36, 39] found that the arc-search IPMs improve the iteration complexity from the line-search IPM in [31], and the numerical experiments in [35, 39] demonstrated that the number of iterations in solving LP is reduced compared to the existing methods.

Arc-search IPMs solve two LESs in each iteration for computing the search direction while line-search IPMs one LES, thus, the improvement due to solving LESs inexactly is expected to be more beneficial in arc-search IPMs than line-search IPMs. In fact, when the arc-search IPMs are extended to nonlinear programming problems [33] and convex optimization problems [38], the arc-search IPMs can reduce the computation time even if the computation of higher-order derivatives is omitted, i.e., the search direction is obtained inexactly.

In this paper, we propose a novel inexact infeasible arc-search interior-point method (II-arc) by integrating an inexact IPM and an arc-search IPM. We prove that the II-arc method achieves a better iteration complexity than the inexact infeasible line-search IPMs (II-lines) [22, 23]. We conduct the numerical experiments with the conjugate gradient (CG) method as an inexact linear equation solver for large benchmark problems in the Netlib collection [6]. The results show that the proposed method can reduce the number of iterations by twice and the computation time by 23% compared to II-line for almost half of the benchmark problems. Furthermore, the comparison with the IPMs for solving the LES exactly shows that the proposed method has an advantage in terms of the computation time, even if the number of iterations of II-arc is greater than it of the exact IPMs.

This paper is organized as follows. Section 2 introduces the standard form of LP problems and the formulas necessary for II-arc. In Section 3, we describe the proposed method, and in Section 4, we discuss the convergence and the polynomial iteration complexity. Section 5 provides the results of the numerical experiments and the discussion. Finally, Section 6 gives conclusions of this paper and discusses future directions.

1.1 Notations

We use x_i to denote the i-th element of a vector x. The Hadamard product of two vectors u and v is defined by $u \circ v$. The vector of all ones and the identity matrix are denoted by e and I, respectively. We use the capital character $X \in \mathbb{R}^{n\times n}$ as the diagonal matrix whose diagonal elements are taken from the vector $x \in \mathbb{R}^n$. For a set B, we denote the cardinality of the set by $|B|$. Given a matrix $A \in \mathbb{R}^{m\times n}$ and a set $B \subseteq \{1,\dots,n\}$, the matrix A_B is the submatrix consisting of the columns $\{A_i : i \in B\}$. Similarly, given a vector $v \in \mathbb{R}^n$ and a set $B \subseteq \{1,\dots,n\}$ where $|B| = m \le n$, the matrix $V_B \in \mathbb{R}^{m\times m}$ is the diagonal submatrix consisting of the elements $\{v_i : i \in B\}$. We use $\|x\|_2 = (\sum_i x_i^2)^{1/2}$, $\|x\|_\infty = \max_i |x_i|$ and $\|x\|_1 = \sum_i |x_i|$ for the Euclidean norm, the infinity norm and the ℓ_1 norm of a vector x, respectively. For simplicity, we denote $\|x\| = \|x\|_2$. For a matrix $A \in \mathbb{R}^{m\times n}$, $\|A\|$ denotes the operator norm associated with the Euclidian norm; $\|A\| = \max_{\|z\|=1} \|Az\|$.

2 Preliminaries

section_preliminaries

In this paper, we consider an LP in the standard form:

problem_main

$$\min_{x\in\mathbb{R}^n} c^\top x, \quad \text{s.t. } Ax = b, \quad x \ge 0, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$ with $m \le n$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are input data. The associated dual problem of (1) is

problem_dual

$$\max_{y\in\mathbb{R}^m, s\in\mathbb{R}^n} b^\top y, \quad \text{s.t. } A^\top y + s = c, \quad s \ge 0, \tag{2}$$

where y and s are the dual variable vector and the dual slack vector, respectively. Let $\mathcal{S}^*$ be the set of the optimal solutions of (1) and (2). When $(x^*, y^*, s^*) \in \mathcal{S}^*$, it is well-known that (x^*, y^*, s^*) satisfies the KKT conditions: KKT_conditions

$$\begin{aligned} Ax^* &= b && \text{(3a)}\\ A^\top y^* + s^* &= c && \text{(3b)}\\ (x^*, s^*) &\ge 0 && \text{(3c)}\\ x_i^* s_i^* &= 0, \quad i = 1,\dots,n. && \text{(3d)} \end{aligned}$$

We denote the primal and dual residuals in (1) and (2) as residuals_constraints

residual_main

$$r_b(x) = Ax - b \tag{4a}$$

residual_dual

$$r_c(y,s) = A^\top y + s - c, \tag{4b}$$

and define the duality measure as

def_mu

$$\mu = \frac{x^\top s}{n}. \tag{5}$$

Letting $\zeta \geq 0$, we define the set of ζ-optimal solutions as

def_zeta_optimal_SolSet

$$\mathcal{S}^*_\zeta = \left\{(x, y, s) \in \mathbb{R}^{2n+m} \mid (x, s) \geq 0,\ \mu \leq \zeta,\ \|(r_b(x), r_c(y, s))\| \leq \zeta\right\}. \tag{6}$$

From the KKT conditions (3), we know $\mathcal{S}^* \subset \mathcal{S}^*_\zeta$.

In this paper, we make the following assumptions for the primal-dual pair (1) and (2). These assumptions are common ones in the context of IPMs and are used in many papers (for example, see [31, 37]).

assumption_IPC

Assumption 2.1. *There exists an interior feasible solution* $(\bar{x}, \bar{y}, \bar{s})$ *such that*

$$A\bar{x} = b,\ A^\top \bar{y} + \bar{s} = c,\ \textit{and}\ (\bar{x}, \bar{s}) > 0.$$

assumption_full_row_rank

Assumption 2.2. *A is a full-row rank matrix, i.e.,* $\operatorname{rank}(A) = m$

Assumption 2.1 guarantees that the optimal set $\mathcal{S}^*$ is nonempty and bounded [31].

IPMs are iterative methods, so we denote the kth iteration by $(x^k, y^k, s^k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ and the initial point by (x^0, y^0, s^0). Without loss of generality, we assume that the initial point (x^0, y^0, s^0) is bounded. We denote the duality measure of kth iteration as $\mu_k = (x^k)^\top s^k / n$.

Given a strictly positive iteration (x^k, y^k, s^k) such that $(x^k, s^k) > 0$, the strategy of an infeasible IPM is to trace a smooth curve called an approximate central path:

def_ellipsoid

$$\mathcal{C} = \{(x(t), y(t), s(t)) \mid t \in (0, 1]\}, \tag{7}$$

where $(x(t), y(t), s(t))$ is the unique solution of the following system curve_to_optimal_solution

$$\begin{aligned}
&Ax(t) - b = t\ r_b(x^k), && \text{(8a)}\\
&A^\top y(t) + s(t) - c = t\ r_c(y^k, s^k), && \text{(8b)}\\
&x(t) \circ s(t) = t(x^k \circ s^k), && \text{(8c)}\\
&(x(t), s(t)) > 0. && \text{(8d)}
\end{aligned}$$

As $t \to 0$, $(x(t), y(t), s(t))$ converges to an optimal solution $(x^*, y^*, s^*) \in \mathcal{S}^*$.

Arc-search IPMs approximate $\mathcal{C}$ with an ellipsoidal arc. An ellipsoidal approximation of $(x(t), y(t), s(t))$ at (x^k, y^k, s^k) for an angle $\alpha \in [0, \pi/2]$ is obtained by $(x(\alpha), y(\alpha), s(\alpha))$ with the following [37, Theorem 5.1]: def_variable_alpha_original

$$\begin{aligned}
x(\alpha) &= x - \dot{x}\sin(\alpha) + \ddot{x}(1 - \cos(\alpha)), && \text{(9a)}\\
y(\alpha) &= y - \dot{y}\sin(\alpha) + \ddot{y}(1 - \cos(\alpha)), && \text{(9b)}\\
s(\alpha) &= s - \dot{s}\sin(\alpha) + \ddot{s}(1 - \cos(\alpha)). && \text{(9c)}
\end{aligned}$$

Here, $(\dot{x}, \dot{y}, \dot{s})$ and $(\ddot{x}, \ddot{y}, \ddot{s})$ are the first and second derivatives obtained by differentiating both sides of (8) by t, and they are computed as the solutions of the following LESs, respectively:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{s} \end{bmatrix} = \begin{bmatrix} r_b(x^k) \\ r_c(y^k, s^k) \\ x^k \circ s^k \end{bmatrix}, \tag{10}$$

first_derivative_original

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2\dot{x} \circ \dot{s} \end{bmatrix}. \tag{11}$$

second_derivative_original

Lastly, we define a neighborhood of the approximate central path [31, Chapter 6]:

$$\mathcal{N}(\gamma_1, \gamma_2) := \left\{ (x, y, s) \mid \begin{array}{l} (x, s) > 0,\ x_i s_i \geq \gamma_1 \mu \text{ for } i \in \{1, \ldots, n\}, \\ \|(r_b(x), r_c(y, s))\| \leq [\|(r_b(x^0), r_c(y^0, s^0))\|/\mu_0]\gamma_2 \mu \end{array} \right\}, \tag{12}$$

def_neighborhood

where $\gamma_1 \in (0, 1)$ and $\gamma_2 \geq 1$ are given parameters, and $\|(r_b(x), r_c(y, s))\|$ is the norm of the vertical concatenation of $r_b(x)$ and $r_c(y, s)$. This neighborhood will be used in the convergence analysis.

3 The proposed method

section_proposed_method

In this section, we propose the II-arc method. In the beginning, to guarantee the convergence of the proposed method, we introduce a perturbation into (10) as follows:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{s} \end{bmatrix} = \begin{bmatrix} r_b(x^k) \\ r_c(y^k, s^k) \\ x^k \circ s^k - \sigma \mu_k e \end{bmatrix}, \tag{13}$$

first_derivative_perturbed

where $\sigma \in (0, 1]$ is the constant called centering parameter. In the subsequent discussion, $(\dot{x}, \dot{y}, \dot{s})$ denote the solution of (13). The proposed method solves (13) and (11) inexactly in each iteration to obtain the ellipsoidal approximation.

Several approaches can be considered for solving the Newton system (13), such as the full Newton system and the Newton equation system (also known as the normal equation system, NES) [4]. The NES formula of (13) is

$$M^k \dot{y} = \rho_1^k, \tag{14}$$

first_derivative_NES

where

$$M^k = A(D^k)^2 A^\top, \tag{15a}$$

$$\begin{aligned} \rho_1^k &= A(D^k)^2 r_c(y^k, s^k) + r_b(x^k) - A(S^k)^{-1}(x^k \circ s^k - \sigma \mu_k e) \\ &= A(D^k)^2 A^\top y^k - A(D^k)^2 c + \sigma \mu_k A(S^k)^{-1} e + A x^k - b, \end{aligned} \tag{15b}$$

def_NES_rho_1

def_NES_constants

with $D^k = (X^k)^{\frac{1}{2}}(S^k)^{-\frac{1}{2}}$. When we solve the LES (14) exactly and obtain $\dot{y}$, we can compute the other components $\dot{x}$ and $\dot{s}$ of the solution in (13)

As discussed by Mohammadisiahroudi et al. [23], the iteration complexity of the II-line can be kept small by the modification to NES (14). This modified NES formula was examined for II-lines in [1, 25], it is called MNES. Since A is full row rank from Assumption 2.2, we can choose an arbitrary basis $\hat{B} \subset \{1, 2, \dots, n\}$ where $|\hat{B}| = m$ and $A_{\hat{B}} \in \mathbb{R}^{m\times m}$ is nonsingular. Now we can adapt (14) to

first_derivative_MNES

$$\hat{M}^k \dot{z} = \hat{\rho}_1^k, \tag{16}$$

where def_MNES_constants

def_MNES_coef_matrix

$$\hat{M}^k = (D_{\hat{B}}^k)^{-1} A_{\hat{B}}^{-1} M^k ((D_{\hat{B}}^k)^{-1} A_{\hat{B}}^{-1})^\top, \tag{17a}$$

$$\hat{\rho}_1^k = (D_{\hat{B}}^k)^{-1} A_{\hat{B}}^{-1} \rho_1^k, \tag{17b}$$

with $D_{\hat{B}}^k = (X_{\hat{B}}^k)^{\frac{1}{2}}(S_{\hat{B}}^k)^{-\frac{1}{2}}$. The inexact solution $\tilde{\dot{z}}$ of (16) satisfies

inexact_first_derivative_MNES

$$\hat{M}^k \tilde{\dot{z}} = \hat{\rho}_1^k + \hat{r}_1^k, \tag{18}$$

where $\hat{r}_1^k$ is the error of $\tilde{\dot{z}}$ defined as

$$\hat{r}_1^k := \hat{M}^k \tilde{\dot{z}} - \hat{\rho}_1^k = \hat{M}_k \left(\tilde{\dot{z}} - \dot{z}\right).$$

Then, we can obtain the first derivative $(\tilde{\dot{x}}, \tilde{\dot{y}}, \tilde{\dot{s}})$ from the inexact solution in (18) and the steps below: resolution_first_derivative_from_MNES

$$\tilde{\dot{y}} = \left(\left(D_{\hat{B}}^k\right)^{-1} A_{\hat{B}}^{-1}\right)^\top \tilde{\dot{z}} \tag{19a}$$

$$\tilde{\dot{s}} = r_c(y^k, s^k) - A^T \tilde{\dot{y}} \tag{19b}$$

$$v_1^k = \left(v_{\hat{B}}^k, v_{\hat{N}}^k\right) = \left(D_{\hat{B}}^k \hat{r}_1^k, 0\right) \tag{19c}$$

$$\tilde{\dot{x}} = x^k - (D^k)^2 \tilde{\dot{s}} - \sigma\mu_k (S^k)^{-1} e - v_1^k. \tag{19d}$$

We also apply the MNES formulation to the second derivative (11). Letting

$$\rho_2^k = 2A(S^k)^{-1}\tilde{\dot{x}} \circ \tilde{\dot{s}}, \qquad \hat{\rho}_2^k = (D_{\hat{B}}^k)^{-1} A_{\hat{B}}^{-1} \rho_2^k,$$

we have

second_derivative_MNES

$$\hat{M}^k \ddot{z} = \hat{\rho}_2^k \tag{20}$$

with the same definition of $\hat{M}^k$ as in (17a). We use $\tilde{\ddot{z}}$ to denote the inexact solution of (20), then we have

inexact_second_derivative_MNES

$$\hat{M}^k \tilde{\ddot{z}} = \hat{\rho}_2^k + \hat{r}_2^k, \tag{21}$$

where $\hat{r}_2^k$ is defined as $\hat{r}_2^k := \hat{M}_k\left(\tilde{\ddot{z}} - \ddot{z}\right)$. Similarly to (19), to obtain the inexact second derivative $(\tilde{\ddot{x}}, \tilde{\ddot{y}}, \tilde{\ddot{s}})$ from the inexact solution $\tilde{\ddot{z}}$ in (21), we compute as follows:

$$
\begin{aligned}
\tilde{\ddot{y}} &= \left(\left(D_{\hat{B}}^k\right)^{-1} A_{\hat{B}}^{-1}\right)^\top \tilde{\ddot{z}}, \\
\tilde{\ddot{s}} &= -A^T \tilde{\ddot{y}}, \\
v_2^k &= \left(v_{\hat{B}}^k, v_{\hat{N}}^k\right) = \left(D_{\hat{B}}^k \hat{r}_2^k, 0\right), \\
\tilde{\ddot{x}} &= -(D^k)^2 \tilde{\ddot{s}} - 2(S^k)^{-1} \tilde{\dot{x}} \circ \tilde{\dot{s}} - v_2^k.
\end{aligned}
$$

resolution_second_derivative_from_MNES

Using the derivatives obtained above, the next iteration will be found on the ellipsoidal arc with the following updated formula: def_variable_alpha_with_inexact_derivatives

$$
\begin{aligned}
x^k(\alpha) &= x^k - \tilde{\dot{x}} \sin(\alpha) + \tilde{\ddot{x}}(1 - \cos(\alpha)), && \text{(23a)} \\
y^k(\alpha) &= y^k - \tilde{\dot{y}} \sin(\alpha) + \tilde{\ddot{y}}(1 - \cos(\alpha)), && \text{(23b)} \\
s^k(\alpha) &= s^k - \tilde{\dot{s}} \sin(\alpha) + \tilde{\ddot{s}}(1 - \cos(\alpha)). && \text{(23c)}
\end{aligned}
$$

To give the framework of the proposed method, we prepare some functions below:

$$
\begin{aligned}
G_i^k(\alpha) &= x_i^k(\alpha) s_i^k(\alpha) - \gamma_1 \mu_k(\alpha) \text{ for } i \in \{1, \ldots, n\}, \\
g^k(\alpha) &= x^k(\alpha)^\top s^k(\alpha) - (1 - \sin(\alpha))(x^k)^\top s^k, \\
h^k(\alpha) &= (1 - (1 - \beta)\sin(\alpha))\,(x^k)^\top s^k - x^k(\alpha)^\top s^k(\alpha).
\end{aligned}
$$

Here, $h^k(\alpha) \geq 0$ corresponds to the Armijo condition with respect to the duality gap μ. In Section 4, we will show that the proposed algorithm converges to an optimal solution by selecting a step size α that satisfies the following conditions:

conditions_G_g_h_no_less_than_0

$$
G_i^k(\alpha) \geq 0 \text{ for } i \in \{1, \ldots, n\}, \quad g^k(\alpha) \geq 0, \quad h^k(\alpha) \geq 0. \tag{24}
$$

When (24) holds, the next lemma confirms that a next iteration point $(x^k(\alpha), y^k(\alpha), s^k(\alpha))$ is in the neighborhood $\mathcal{N}(\gamma_1, \gamma_2)$. This lemma can be proved in the same approach as Mohammadisiahroudi [23, Lemma 4.5] with Lemma 4.2 below.

lemma_in_neighborhood

Lemma 3.1. *Assume a step length $\alpha \in (0, \pi/2]$ satisfies $G_i^k(\alpha) \geq 0$ and $g^k(\alpha) \geq 0$. Then, $(x^k(\alpha), y^k(\alpha), s^k(\alpha)) \in \mathcal{N}(\gamma_1, \gamma_2)$.*

Lastly, we discuss the error range such that the inexact solutions still can make the proposed algorithm attain the polynomial iteration complexity. This accuracy will also be used for the convergence proof in Section 4. We assume the following inequality for the error $\hat{r}_1^k$ of (18) and $\hat{r}_2^k$ of (21):

def_upper_derivatives_residual_MNES

$$
\left\|\hat{r}_i^k\right\| \leq \eta \frac{\sqrt{\mu_k}}{\sqrt{n}}, \quad \forall i \in \{1, 2\} \tag{25}
$$

where $\eta \in [0, 1)$ is an enforcing parameter.

To prove the polynomial iteration complexity of the proposed algorithm in Proposition 4.1 below, we set the parameters so that parameter_conditions

parameter_condition_for_G_i
$$(1-\gamma_1)\sigma - (1+\gamma_1)\eta > 0, \tag{26a}$$
parameter_condition_beta_more_than_sigma_plus_eta
$$\beta > \sigma + \eta. \tag{26b}$$

We are now ready to give the framework of the proposed method (II-arc) as Algorithm 1.

Algorithm 1 The inexact infeasible arc-search interior-point method (II-arc)
algorithm_II_arc_IPM

Input: $\zeta > 0$, $\gamma_1 \in (0, 1)$, $\gamma_2 \geq 1$, σ, η, β satisfying (26) and an initial point $(x^0, y^0, s^0) \in \mathcal{N}(\gamma_1, \gamma_2)$ such that $x^0 > 0$ and $s^0 > 0$.

Output: ζ-optimal solution (x^k, y^k, s^k)

$k \leftarrow 0$
while $(x^k, y^k, s^k) \notin S_\zeta$ **do** line_algo_II_arc_search_checking_stop
  $\mu_k \leftarrow (x^k)^\top s^k/n$
  Calculate $(\tilde{\dot{x}}, \tilde{\dot{y}}, \tilde{\dot{s}})$ by solving (16) inexactly satisfying (25).
  Calculate $(\tilde{\ddot{x}}, \tilde{\ddot{y}}, \tilde{\ddot{s}})$ by solving (20) inexactly satisfying (25). line_algo_II_arc_search_calculate_se
  $\alpha_k \leftarrow \max\{\alpha \in (0, \pi/2] \mid \alpha \text{ satisfies } (24)\}$ line_algo_II_arc_search_decide_step_size
  Set $(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k(\alpha_k), y^k(\alpha_k), s^k(\alpha_k))$ by (23).
  $k \leftarrow k + 1$
end while

4 Theoretical proof

section_theoretical_proof

In this section, we prove the convergence of Algorithm 1 and its polynomial iteration complexity. Our analysis is close to Mohammadisiahroudi et al. [23], but it also employs properties of arc-search IPMs.

First, we evaluate the constraint residuals (4). From (18) and (19), the residual appears only in the last equation as a term $S^k v_1^k$, as the following lemma shows.

lemma_inexact_solution_MNES_conditions

Lemma 4.1. *For the inexact first derivative* $(\tilde{\dot{x}}, \tilde{\dot{y}}, \tilde{\dot{s}})$ *of* (8) *obtained by the inexact solution of* (16) *and the steps in* (19), *we have*

inexact_first_derivative_MNES_main_residual
$$A\tilde{\dot{x}} = r_b(x^k), \tag{27a}$$
inexact_first_derivative_MNES_dual_residual
$$A^\top \tilde{\dot{y}} + \tilde{\dot{s}} = r_c(y^k, s^k), \tag{27b}$$
inexact_first_derivative_MNES_duality
$$S^k \tilde{\dot{x}} + X^k \tilde{\dot{s}} = X^k s^k - \sigma \mu_k e - S^k v_1^k. \tag{27c}$$

Lemma 4.1 can be proved from (16) and (19) in the same way as Mohammadisiahroudi [23, Lemma 4.1], thus we omit the proof. As in Lemma 4.1, $(\tilde{\dot{x}}, \tilde{\dot{y}}, \tilde{\dot{s}})$ obtained by (21) and (22) satisfies indexact_second_derivative_conditions

inexact_second_derivative_main_residual
$$A\tilde{\ddot{x}} = 0, \tag{28a}$$
inexact_second_derivative_dual_residual
$$A^\top \tilde{\ddot{y}} + \tilde{\ddot{s}} = 0, \tag{28b}$$
inexact_second_derivative_duality
$$S^k \tilde{\ddot{x}} + X^k \tilde{\ddot{s}} = -2\tilde{\dot{x}} \circ \tilde{\dot{s}} - S^k v_2^k. \tag{28c}$$

Therefore, the following lemma holds from (27a), (27b), (28a) and (28b) due to (23).

lemma_decrease_constraint_residuals

Lemma 4.2 ([37, Lemma 7.2])**.** *For each iteration k, the following relations hold.*

$$r_b(x^{k+1}) = r_b(x^k)\left(1 - \sin(\alpha_k)\right),$$
$$r_c(y^{k+1}, s^{k+1}) = r_c(y^k, s^k)\left(1 - \sin(\alpha_k)\right).$$

For the following discussions, we introduce the following notation:

$$\nu_k = \prod_{i=0}^{k-1}(1 - \sin(\alpha_i)).$$

From Lemma 4.2, we can obtain residuals_decreasing

$$r_b(x^k) = \nu_k r_b(x^0) \tag{29a}$$
$$r_c(y^k, s^k) = \nu_k r_c(y^0, s^0) \tag{29b}$$

In the next proposition, we prove the existence of the lower bound of the step size α_k to guarantee that Algorithm 1 is well defined.

proposition_lower_bound_of_step_size

Proposition 4.1. *Let $\{(x^k, y^k, s^k)\}$ be the sequence generated by Algorithm 1. Then, there exists $\hat{\alpha} > 0$ satisfying (24) for any $\alpha_k \in (0, \hat{\alpha}]$ and*

$$\sin(\hat{\alpha}) = \frac{C}{n^{1.5}},$$

where C is a positive constant.

The proof of Proposition 4.1 will be given later. For this proof, we first evaluate x^k and s^k with the ℓ_1 norm.

lemma_upper_nu_x_s

Lemma 4.3. *There is a positive constant C_1 such that*

upper_bound_norm_x_s
$$\nu_k \left\| (x^k, s^k) \right\|_1 \le C_1 n \mu_k. \tag{30}$$

The proof below is based on [31, Lemma 6.3].

Proof. From the definition of $\mathcal{N}(\gamma_1, \gamma_2)$ in (12) and $\gamma_2 \geq 1$, we know

$$\frac{\left\|(r_b(x^k), r_c(y^k, s^k))\right\|}{\mu_k} \leq \gamma_2 \frac{\left\|(r_b(x^0), r_c(y^0, s^0))\right\|}{\mu_0} \leq \frac{\left\|(r_b(x^0), r_c(y^0, s^0))\right\|}{\mu^0},$$

which implies

$$\mu_k \geq \frac{\left\|(r_b(x^k), r_c(y^k, s^k))\right\|}{\left\|(r_b(x^0), r_c(y^0, s^0))\right\|}\mu_0 = \nu_k \mu_0 \tag{31}$$

mu_decreasing_lower_bound

from (29). When we set

$$(\bar{x}, \bar{y}, \bar{s}) = \nu_k(x^0, y^0, s^0) + (1-\nu_k)(x^*, y^*, s^*) - (x^k, y^k, s^k),$$

we have $A\bar{x} = 0$ and $A^\top \bar{y} + \bar{s} = 0$ from (29) and (3), then

$$\begin{aligned} 0 &= \bar{x}^\top \bar{s} \\ &= (\nu_k x^0 + (1-\nu_k)x^* - x^k)^\top(\nu_k s^0 + (1-\nu_k)s^* - s^k) \\ &= \nu_k^2 (x^0)^\top s^0 + \nu_k(1-\nu_k)\left((x^0)^\top s^* + (x^*)^\top s^0\right) + (x^k)^\top s^k + (1-\nu_k)^2 (x^*)^\top s^* \\ &\quad - \left(\nu_k((x^0)^\top s^k + (s^0)^\top x^k) + (1-\nu_k)((x^k)^\top s^* + (s^k)^\top x^*)\right) \end{aligned}$$

is satisfied. Since all the components of x^k, s^k, x^*, s^* are nonnegative, we have $((x^k)^\top s^* + (s^k)^\top x^*) \geq 0$. In addition, we have $(x^*)^\top s^* = 0$ from (3). By using these and rearranging, we obtain

$$\begin{aligned} \nu_k((x^0)^\top s^k + (s^0)^\top x^k) &\leq \nu_k^2 (x^0)^\top s^0 + \nu_k(1-\nu_k)\left((x^0)^\top s^* + (x^*)^\top s^0\right) + (x^k)^\top s^k \\ [\because (5)] \quad &= \nu_k^2 n\mu_0 + \nu_k(1-\nu_k)\left((x^0)^\top s^* + (x^*)^\top s^0\right) + n\mu_k \\ [\because (31)] \quad &\leq \nu_k n\mu_k + \frac{\mu_k}{\mu_0}(1-\nu_k)\left((x^0)^\top s^* + (x^*)^\top s^0\right) + n\mu_k \\ [\because \nu_k \in [0,1]] \quad &\leq 2n\mu_k + \frac{\mu_k}{\mu_0}\left((x^0)^\top s^* + (x^*)^\top s^0\right). \end{aligned} \tag{32}$$

x_s_upper

Defining a constant ξ by

$$\xi = \min_{i=1,2,\ldots,n} \min(x_i^0, s_i^0) > 0, \tag{33}$$

def_xi

we have $(x^0)^\top s^k + (s^0)^\top x^k \geq \xi\left\|(x^k, s^k)\right\|_1$. Therefore, from (32), we obtain

$$\nu_k\left\|(x^k, s^k)\right\|_1 \leq \xi^{-1}\left(2 + \frac{(x^0)^\top s^* + (x^*)^\top s^0}{(x^0)^\top s^0}\right) n\mu_k.$$

We complete this proof by setting

$$C_1 = \xi^{-1}\left(2 + \frac{(x^0)^\top s^* + (x^*)^\top s^0}{(x^0)^\top s^0}\right) \tag{34}$$

def_C1

in (30), where C_1 is independent of n. □

Next, we prove upper bounds of the terms related to $\tilde{\dot{x}}, \tilde{\dot{s}}, \tilde{\ddot{x}}, \tilde{\ddot{s}}$. From (25), the following lemma gives an upper bound of (27c) and (28c):

lemma_upper_derivatives_residual

Lemma 4.4 ([23, Lemma 4.2])**.** *For the derivatives* $(\tilde{\dot{x}}, \tilde{\dot{y}}, \tilde{\dot{s}})$ *and* $(\tilde{\ddot{x}}, \tilde{\ddot{y}}, \tilde{\ddot{s}})$*, when the residuals* $\hat{r}_i^k$ *satisfy* (25)*, it holds that*

upper_residual_term_MNES

$$\left\| S^k v_i^k \right\|_\infty \le \eta\mu_k. \tag{35}$$

Then, the following lemma holds similarly to [31, Lemma 6.5] and [23, Lemma 4.6].

Lemma 4.5. *There is a positive constant* C_2 *such that*

$$\max\left\{ \left\| (D^k)^{-1}\tilde{\dot{x}} \right\|, \left\| D^k \tilde{\dot{s}} \right\| \right\} \le C_2 n\sqrt{\mu_k}$$

lemma_first_derivative_upper

Proof. Let

$$(\bar{x}, \bar{y}, \bar{s}) = (\tilde{\dot{x}}, \tilde{\dot{y}}, \tilde{\dot{s}}) - \nu_k(x^0, y^0, s^0) + \nu_k(x^*, y^*, s^*).$$

From (27a), (27b), (29) and (3), we have $A\bar{x} = 0$ and $A^\top\bar{y} + \bar{s} = 0$, therefore, $\bar{x}^\top\bar{s} = 0$. Thus, we obtain

eq_norm_D_inv_bar_x_plus_D_bar_s

$$\left\| (D^k)^{-1}\bar{x} + D^k\bar{s} \right\|^2 = \left\| (D^k)^{-1}(\tilde{\dot{x}} - \nu_k(x^0 - x^*)) \right\|^2 + \left\| D^k(\tilde{\dot{s}} - \nu_k(s^0 - s^*)) \right\|^2. \tag{36}$$

From (27c), it holds that

$$\begin{aligned} S^k\bar{x} + X^k\bar{s} &= (S^k\tilde{\dot{x}} + X^k\tilde{\dot{s}}) - \nu_k S^k(x^0 - x^*) - \nu_k X^k(s^0 - s^*) \\ &= (X^k s^k - \sigma\mu_k e - S^k v_1^k) - \nu_k S^k(x^0 - x^*) - \nu_k X^k(s^0 - s^*). \end{aligned}$$

Consequently, we verify

eq_D_inv_bar_x_plus_D_bar_s

$$(D^k)^{-1}\bar{x} + D^k\bar{s} = (X^k S^k)^{-\frac{1}{2}}(X^k s^k - \sigma\mu_k e - S^k v_1^k) - \nu_k(D^k)^{-1}(x^0 - x^*) - \nu_k D^k(s^0 - s^*). \tag{37}$$

For any vector $a \in \mathbb{R}^d$,

inequality_norms

$$\|a\|_1 \le \sqrt{n}\|a\| \le n\|a\|_\infty \tag{38}$$

holds from [37, Lemma 3.1]. From (36), (37), (38) and Lemma 4.4, we obtain

upper_sum_of_norm_of_D_inv_bar_x_plus_D_bar_s

$$\begin{aligned} &\left\| (D^k)^{-1}(\tilde{\dot{x}} - \nu_k(x^0 - x^*)) \right\|^2 + \left\| D^k(\tilde{\dot{s}} - \nu_k(s^0 - s^*)) \right\|^2 \\ &= \left\| (X^k S^k)^{-\frac{1}{2}}(X^k s^k - \sigma\mu_k e - S^k v_1^k) - \nu_k(D^k)^{-1}(x^0 - x^*) - \nu_k D^k(s^0 - s^*) \right\|^2 \\ &\le \left\{ \left\| X^k S^k \right\|^{-\frac{1}{2}} \left(\left\| X^k s^k - \sigma\mu_k e \right\| + \left\| S^k v_1^k \right\| \right) + \nu_k \left\| (D^k)^{-1}(x^0 - x^*) \right\| + \nu_k \left\| D^k(s^0 - s^*) \right\| \right\}^2 \\ &\le \left\{ \left\| X^k S^k \right\|^{-\frac{1}{2}} \left(\left\| X^k s^k - \sigma\mu_k e \right\| + \sqrt{n}\eta\mu_k \right) + \nu_k \left(\left\| (D^k)^{-1}(x^0 - x^*) \right\| + \left\| D^k(s^0 - s^*) \right\| \right) \right\}^2. \end{aligned} \tag{39}$$

In addition, $x_i^k s_i^k \geq \gamma\mu_k$ in (12) implies

upper_x_s_half_inverse

$$\left\|X^k S^k\right\|^{-\frac{1}{2}} \leq \frac{1}{\sqrt{\gamma_1 \mu_k}}. \tag{40}$$

From (30) and (40), we have

upper_xs_XS_half_inv

$$\nu_k \left\|(x^k, s^k)\right\|_1 \left\|(XS)^{-1/2}\right\| \leq \frac{C_1 n \sqrt{\mu_k}}{\sqrt{\gamma_1}}. \tag{41}$$

According to the derivation in [31, Lemma 6.5], we have

upper_X_s

$$\left\|X^k s^k - \sigma\mu_k e\right\| \leq n\mu_k, \tag{42}$$

upper_nu_k_norm

$$\begin{aligned} &\nu_k \left(\left\|(D^k)^{-1}(x^0 - x^*)\right\| + \left\|D^k(s^0 - s^*)\right\|\right) \\ &\leq \nu_k \left\|(x^k, s^k)\right\|_1 \left\|(XS)^{-1/2}\right\| \max\left\{\left\|x^0 - x^*\right\|, \left\|s^0 - s^*\right\|\right\}. \end{aligned} \tag{43}$$

Therefore, from (43) and (41), we obtain

upper_sum_of_norm_of_D_xs_0_minus_xs_star

$$\begin{aligned} &\nu_k \left(\left\|(D^k)^{-1}(x^0 - x^*)\right\| + \left\|D^k(s^0 - s^*)\right\|\right) \\ &\leq \frac{C_1}{\sqrt{\gamma_1}} n\sqrt{\mu_k} \max\left\{\left\|x^0 - x^*\right\|, \left\|s^0 - s^*\right\|\right\}. \end{aligned} \tag{44}$$

Therefore, we have

$$\begin{aligned} \left\|(D^k)^{-1}\tilde{x}\right\| &\leq \left\|(D^k)^{-1}(\tilde{x} - \nu_k(x^0 - x^*))\right\| + \nu_k \left\|(D^k)^{-1}(x^0 - x^*)\right\| \\ [\because (39)] \quad &\leq \left\|X^k S^k\right\|^{-\frac{1}{2}} \left(\left\|X^k s^k - \sigma\mu_k e\right\| + \sqrt{n}\eta\mu_k\right) \\ &\quad + 2\nu_k \left(\left\|(D^k)^{-1}(x^0 - x^*)\right\| + \left\|D^k(s^0 - s^*)\right\|\right) \\ [\because (40), (42)] \quad &\leq \frac{\sqrt{\mu_k}}{\sqrt{\gamma_1}} \left(n + \sqrt{n}\eta\right) + 2\nu_k \left(\left\|(D^k)^{-1}(x^0 - x^*)\right\| + \left\|D^k(s^0 - s^*)\right\|\right) \\ [\because (44)] \quad &\leq \frac{\sqrt{\mu_k}}{\sqrt{\gamma_1}} \left(n + \sqrt{n}\eta\right) + \frac{2C_1 n\sqrt{\mu_k}}{\sqrt{\gamma_1}} \max\left\{\left\|x^0 - x^*\right\|, \left\|s^0 - s^*\right\|\right\} \\ &\leq \frac{1}{\sqrt{\gamma_1}} \left(1 + \eta + 2C_1 \max\left\{\left\|x^0 - x^*\right\|, \left\|s^0 - s^*\right\|\right\}\right) n\sqrt{\mu_k}. \end{aligned}$$

Since the optimal set is bounded from Assumption 2.1 and the initial point is bounded,

def_C2

$$C_2 := \gamma_1^{-1/2} \left(1 + \eta + 2C_1 \max\left\{\left\|x^0 - x^*\right\|, \left\|s^0 - s^*\right\|\right\}\right) \tag{45}$$

is also bounded, and we can prove this lemma by setting this C_2. We can similarly show $\tilde{s} \leq C_2 n\sqrt{\mu_k}$. □

From Lemma 4.5,

upper_first_derivative_Hadamard

$$\left\|\dot{\tilde{x}} \circ \dot{\tilde{s}}\right\| \leq \left\|(D^k)^{-1}\dot{\tilde{x}}\right\|\left\|D^k\dot{\tilde{s}}\right\| \leq C_2^2 n^2 \mu_k. \tag{46}$$

Similarly, we evaluate the terms related to $G_i^k(\alpha)$, $g^k(\alpha)$ and $h^k(\alpha)$.

lemma_upper_of_first_and_second_derivatives

Lemma 4.6. *There are positive constants C_3 and C_4 such that*

$$\begin{aligned} \left\|\ddot{\tilde{x}} \circ \ddot{\tilde{s}}\right\| &\leq C_3 n^4 \mu_k, \\ \max\left\{\left\|(D^k)^{-1}\ddot{\tilde{x}}\right\|, \left\|D^k\ddot{\tilde{s}}\right\|\right\} &\leq C_4 n^2 \sqrt{\mu_k}, \\ \max\left\{\left\|\ddot{\tilde{x}} \circ \dot{\tilde{s}}\right\|, \left\|\dot{\tilde{x}} \circ \ddot{\tilde{s}}\right\|\right\} &\leq C_2 C_4 n^3 \mu_k. \end{aligned}$$

Proof. When $u^\top v \geq 0$ for any vector pairs of u, v, the inequality

$$\|u \circ v\| \leq 2^{-\frac{3}{2}}\|u+v\|^2$$

holds from [31, Lemma 5.3], so the following is satisfied:

$$\left\|\ddot{\tilde{x}} \circ \ddot{\tilde{s}}\right\| = \left\|(D^k)^{-1}\ddot{\tilde{x}} \circ D^k\ddot{\tilde{s}}\right\| \leq 2^{-\frac{3}{2}}\left\|(D^k)^{-1}\ddot{\tilde{x}} + D^k\ddot{\tilde{s}}\right\|^2.$$

From $(D^k)^{-1}\ddot{\tilde{x}} + D^k\ddot{\tilde{s}} = (X^k S^k)^{-1/2}(S^k\ddot{\tilde{x}} + X^k\ddot{\tilde{s}})$,

upper_norm_D_inv_ddot_x_plus_D_ddot_s

$$\begin{aligned} \left\|(D^k)^{-1}\ddot{\tilde{x}} + D^k\ddot{\tilde{s}}\right\| &\leq \left\|X^k S^k\right\|^{-\frac{1}{2}}\left\|S^k\ddot{\tilde{x}} + X^k\ddot{\tilde{s}}\right\| \\ [\because (28c)] \quad &\leq \left\|X^k S^k\right\|^{-\frac{1}{2}}\left(2\left\|\dot{\tilde{x}} \circ \dot{\tilde{s}}\right\| + \left\|S^k v_2^k\right\|\right) \\ [\because (40), (46), (35), (38)] \quad &\leq \frac{1}{\sqrt{\gamma_1 \mu_k}}\left(2C_2^2 n^2 \mu_k + \sqrt{n}\eta\mu_k\right) \\ &\leq \frac{\sqrt{\mu_k}}{\sqrt{\gamma_1}}(2C_2^2 n^2 + \sqrt{n}\eta). \end{aligned} \tag{47}$$

From the above, we can obtain

$$\left\|\ddot{\tilde{x}} \circ \ddot{\tilde{s}}\right\| \leq 2^{-\frac{3}{2}}\frac{\mu_k}{\gamma_1}(2C_2^2 n^2 + \sqrt{n}\eta)^2 \leq \frac{(2C_2^2+\eta)^2}{2^{\frac{3}{2}}\gamma_1} n^4 \mu_k =: C_3 n^4 \mu_k.$$

From (28a) and (28b), we know

second_derivative_x_s_zero_inner_product

$$\ddot{\tilde{x}}^\top \ddot{\tilde{s}} = 0, \tag{48}$$

then (47) leads to

$$\begin{aligned} \max\left\{\left\|(D^k)^{-1}\ddot{\tilde{x}}\right\|^2, \left\|D^k\ddot{\tilde{s}}\right\|^2\right\} &\leq \left\|(D^k)^{-1}\ddot{\tilde{x}} + D^k\ddot{\tilde{s}}\right\|^2 \\ &\leq \frac{\mu_k}{\gamma_1}(2C_2^2 n^2 + \sqrt{n}\eta)^2 \\ &\leq \frac{\mu_k}{\gamma_1}(2C_2^2+\eta)^2 n^4 =: C_4^2 n^4 \mu_k, \end{aligned}$$

$$\left\|\ddot{\tilde{x}} \circ \dot{\tilde{s}}\right\| \le \left\|(D^k)^{-1}\ddot{\tilde{x}}\right\|\left\|D^k \dot{\tilde{s}}\right\| \le C_4 n^2 \sqrt{\mu_k} C_2 n \sqrt{\mu_k} = C_2 C_4 n^3 \mu_k.$$

We can show the boundedness of $\left\|\dot{\tilde{x}} \circ \ddot{\tilde{s}}\right\|$ similarly. □

Using these lemmas, we are ready to prove Proposition 4.1.

Proof of Proposition 4.1. Firstly, we derive the equations necessary for the proofs. We have the following simple identity:

sin_cos_1

$$-2(1-\cos(\alpha)) + \sin^2(\alpha) = -(1-\cos(\alpha))^2. \tag{49}$$

Therefore, we can obtain

$$\begin{aligned}
x^k(\alpha) \circ s^k(\alpha) =& \left(x^k - \dot{\tilde{x}}\sin(\alpha) + \ddot{\tilde{x}}(1-\cos(\alpha))\right) \circ \left(s^k - \dot{\tilde{s}}\sin(\alpha) + \ddot{\tilde{s}}(1-\cos(\alpha))\right) \\
=& \; x^k \circ s^k - \left(x^k \circ \dot{\tilde{s}} + \dot{\tilde{x}} \circ s^k\right)\sin(\alpha) + \left(x^k \circ \ddot{\tilde{s}} + \ddot{\tilde{x}} \circ s^k\right)(1-\cos(\alpha)) \\
& + \dot{\tilde{x}} \circ \dot{\tilde{s}} \sin^2(\alpha) - \left(\dot{\tilde{x}} \circ \ddot{\tilde{s}} + \ddot{\tilde{x}} \circ \dot{\tilde{s}}\right)\sin(\alpha)(1-\cos(\alpha)) + \ddot{\tilde{x}} \circ \ddot{\tilde{s}}(1-\cos(\alpha))^2 \\
[\because (27c),(28c)] \quad =& \; x^k \circ s^k - (x^k \circ s^k - \sigma\mu_k e - S^k v_1^k)\sin(\alpha) + \left(-2\dot{\tilde{x}} \circ \dot{\tilde{s}} - S^k v_2^k\right)(1-\cos(\alpha)) \\
& + \dot{\tilde{x}} \circ \dot{\tilde{s}} \sin^2(\alpha) - \left(\dot{\tilde{x}} \circ \ddot{\tilde{s}} + \ddot{\tilde{x}} \circ \dot{\tilde{s}}\right)\sin(\alpha)(1-\cos(\alpha)) + \ddot{\tilde{x}} \circ \ddot{\tilde{s}}(1-\cos(\alpha))^2 \\
[\because (49)] \quad =& \; x^k \circ s^k(1-\sin(\alpha)) + \sigma\mu_k \sin(\alpha) e \\
& + \left(\ddot{\tilde{x}} \circ \ddot{\tilde{s}} - \dot{\tilde{x}} \circ \dot{\tilde{s}}\right)(1-\cos(\alpha))^2 - \left(\dot{\tilde{x}} \circ \ddot{\tilde{s}} + \ddot{\tilde{x}} \circ \dot{\tilde{s}}\right)\sin(\alpha)(1-\cos(\alpha)) \\
& + S^k v_1^k \sin(\alpha) - S^k v_2^k (1-\cos(\alpha))
\end{aligned} \tag{50}$$

x_s_alpha_Hadamard

and

$$\begin{aligned}
x^k(\alpha)^\top s^k(\alpha) =& \left(x^k - \dot{\tilde{x}}\sin(\alpha) + \ddot{\tilde{x}}(1-\cos(\alpha))\right)^\top \left(s^k - \dot{\tilde{s}}\sin(\alpha) + \ddot{\tilde{s}}(1-\cos(\alpha))\right) \\
[\because (50),(5),(48)] \quad =& \; (x^k)^\top s^k \left((1-\sin(\alpha)) + \sigma\sin(\alpha)\right) \\
& - \dot{\tilde{x}}^\top \dot{\tilde{s}}(1-\cos(\alpha))^2 - \left(\dot{\tilde{x}}^\top \ddot{\tilde{s}} + \ddot{\tilde{x}}^\top \dot{\tilde{s}}\right)\sin(\alpha)(1-\cos(\alpha)) \\
& + \sin(\alpha)\sum_{i=1}^n [S^k v_1^k]_i - (1-\cos(\alpha))\sum_{i=1}^n [S^k v_2^k]_i.
\end{aligned} \tag{51}$$

x_s_alpha_inner_product

From Lemmas 4.5 and 4.6 and the Cauchy-Schwartz inequality, we know

$$\left|\dot{\tilde{x}}_i \dot{\tilde{s}}_i\right|, \; \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| \le \left\|(D^k)^{-1}\dot{\tilde{x}}\right\|\left\|D^k \dot{\tilde{s}}\right\| \le C_2^2 n^2 \mu_k \tag{52a}$$

$$\left|\ddot{\tilde{x}}_i \dot{\tilde{s}}_i\right|, \; \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right| \le \left\|(D^k)^{-1}\ddot{\tilde{x}}\right\|\left\|D^k \dot{\tilde{s}}\right\| \le C_2 C_4 n^3 \mu_k \tag{52b}$$

$$\left|\dot{\tilde{x}}_i \ddot{\tilde{s}}_i\right|, \; \left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| \le \left\|(D^k)^{-1}\dot{\tilde{x}}\right\|\left\|D^k \ddot{\tilde{s}}\right\| \le C_2 C_4 n^3 \mu_k \tag{52c}$$

$$\left|\ddot{\tilde{x}}_i \ddot{\tilde{s}}_i\right| \le \left\|(D^k)^{-1}\ddot{\tilde{x}}\right\|\left\|D^k \ddot{\tilde{s}}\right\| \le C_4^2 n^4 \mu_k \tag{52d}$$

upper_product_of_ddot_x_and_ddot_s_element_wise uppers_product_of_derivatives

Here, $\left|\ddot{\tilde{x}}^\top \ddot{\tilde{s}}\right| = 0$ holds due to (48). Furthermore, we have

sin_square_more_than_one_minus_cos

$$\sin^2(\alpha) = 1 - \cos^2(\alpha) \geq 1 - \cos(\alpha) \tag{53}$$

from $\alpha \in (0, \pi/2]$.

We prove that the step size α satisfying $g^k(\alpha) \geq 0$ is bounded away from zero. From (51),

lower_x_s_alpha_inner_product

$$\begin{aligned}
x^k(\alpha)^\top s^k(\alpha) &\geq (x^k)^\top s^k \left((1 - \sin(\alpha)) + \sigma \sin(\alpha)\right) \\
&\quad - \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| (1 - \cos(\alpha))^2 - \left(\left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| + \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right|\right) \sin(\alpha)(1 - \cos(\alpha)) \\
&\quad - \left\|S^k v_1^k\right\|_1 \sin(\alpha) - \left\|S^k v_2^k\right\|_1 (1 - \cos(\alpha)) \\
[\because (38), (35)] \quad &\geq (x^k)^\top s^k \left((1 - \sin(\alpha)) + \sigma \sin(\alpha)\right) \\
&\quad - \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| (1 - \cos(\alpha))^2 - \left(\left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| + \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right|\right) \sin(\alpha)(1 - \cos(\alpha)) \\
&\quad - \eta n \mu_k (\sin(\alpha) + 1 - \cos(\alpha)).
\end{aligned} \tag{54}$$

Therefore,

$$\begin{aligned}
g^k(\alpha) &= x^k(\alpha)^\top s^k(\alpha) - (1 - \sin(\alpha))(x^k)^\top s^k \\
[\because (54)] \quad &\geq \sigma (x^k)^\top s^k \sin(\alpha) - \eta n \mu_k \left(\sin(\alpha) + 1 - \cos(\alpha)\right) \\
&\quad - \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| (1 - \cos(\alpha))^2 - \left(\left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| + \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right|\right) \sin(\alpha)(1 - \cos(\alpha)) \\
[\because (5), (53)] \quad &\geq \sigma n \mu_k \sin(\alpha) - \eta n \mu_k \left(\sin(\alpha) + \sin^2(\alpha)\right) \\
&\quad - \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| \sin^4(\alpha) - \left(\left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| + \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right|\right) \sin^3(\alpha) \\
[\because (52)] \quad &\geq n \mu_k \sin(\alpha) \left((\sigma - \eta) - \eta \sin(\alpha) - C_2^2 n \sin^3(\alpha) - 2 C_2 C_4 n^2 \sin^2(\alpha)\right).
\end{aligned}$$

Since $\left(-\eta \sin(\alpha) - C_2^2 n \sin^3(\alpha) - 2 C_2 C_4 n^2 \sin^2(\alpha)\right)$ is monotonically decreasing and $\sigma > \eta$ holds from (26a) and $\gamma_1 \in (0, 1)$, there exists the step size $\hat{\alpha}_1 \in (0, \pi/2]$ satisfying the last formula of the right-hand side is no less than 0. When

$$\sin(\hat{\alpha}_1) \leq \frac{\sigma - \eta}{2n} \frac{1}{\max\left\{\eta, C_2^{\frac{2}{3}}, \sqrt{2 C_2 C_4}\right\}},$$

from $0 < \sigma - \eta < \sigma \leq 1$,

$$\begin{aligned}
&(\sigma - \eta) - \eta \sin(\hat{\alpha}_1) - C_2^2 n \sin^3(\hat{\alpha}_1) - 2 C_2 C_4 n^2 \sin^2(\hat{\alpha}_1) \\
&\geq (\sigma - \eta) - \frac{\sigma - \eta}{2n} - \frac{(\sigma - \eta)^3}{8n^2} - \frac{(\sigma - \eta)^2}{4} \\
&\geq (\sigma - \eta) \left(1 - \frac{1}{2} - \frac{1}{8} - \frac{1}{4}\right) \quad \geq 0.
\end{aligned}$$

Therefore, $g^k(\alpha) \geq 0$ is satisfied for any $\alpha \in (0, \hat{\alpha}_1]$.

Next, we consider the range of α such that $G_i^k(\alpha) \geq 0$. From (52), upper_derivatives_element_wise_mu

$$\left|\tilde{\dot{x}}_i\tilde{\dot{s}}_i - \frac{\gamma_1}{n}\tilde{\dot{x}}^\top\tilde{\dot{s}}\right| \leq \left(1+\frac{\gamma_1}{n}\right)C_2^2n^2\mu_k \leq 2C_2^2n^2\mu_k \tag{55a}$$

$$\left|\tilde{\ddot{x}}_i\tilde{\dot{s}}_i - \frac{\gamma_1}{n}\tilde{\ddot{x}}^\top\tilde{\dot{s}}\right|, \left|\tilde{\dot{x}}_i\tilde{\ddot{s}}_i - \frac{\gamma_1}{n}\tilde{\dot{x}}^\top\tilde{\ddot{s}}\right| \leq 2C_2C_4n^3\mu_k \tag{55b}$$

is satisfied. Therefore, we have

$$\begin{aligned}
& G_i^k(\alpha) = x_i^k(\alpha)s_i^k(\alpha) - \gamma_1\mu_k(\alpha) \\
[\because (50),(5),(51)] \quad & \geq x_i^ks_i^k(1-\sin(\alpha)) + \sigma\mu_k\sin(\alpha) \\
& \quad + \left(\tilde{\ddot{x}}_i\tilde{\ddot{s}}_i - \tilde{\dot{x}}_i\tilde{\dot{s}}_i\right)(1-\cos(\alpha))^2 - \left(\tilde{\dot{x}}_i\tilde{\ddot{s}}_i + \tilde{\ddot{x}}_i\tilde{\dot{s}}_i\right)\sin(\alpha)(1-\cos(\alpha)) \\
& \quad - \left\|S^kv_1^k\right\|_\infty\sin(\alpha) - \left\|S^kv_2^k\right\|_\infty(1-\cos(\alpha)) \\
& \quad - \frac{\gamma_1}{n}\Bigg(n\mu_k\left((1-\sin(\alpha)) + \sigma\sin(\alpha)\right) \\
& \quad - \tilde{\dot{x}}^\top\tilde{\dot{s}}(1-\cos(\alpha))^2 - \left(\tilde{\dot{x}}^\top\tilde{\ddot{s}} + \tilde{\ddot{x}}^\top\tilde{\dot{s}}\right)\sin(\alpha)(1-\cos(\alpha))) \\
& \quad + \left\|S^kv_1^k\right\|_1\sin(\alpha) + \left\|S^kv_2^k\right\|_1(1-\cos(\alpha))\Bigg) \\
[\because (12),(35),(38)] \quad & \geq (1-\gamma_1)\sigma\mu_k\sin(\alpha) - (1+\gamma_1)\eta\mu_k(\sin(\alpha)+1-\cos(\alpha)) \\
& \quad + \tilde{\ddot{x}}_i\tilde{\ddot{s}}_i(1-\cos(\alpha))^2 - \left(\tilde{\dot{x}}_i\tilde{\dot{s}}_i - \frac{\gamma_1}{n}\tilde{\dot{x}}^\top\tilde{\dot{s}}\right)(1-\cos(\alpha))^2 \\
& \quad - \left(\tilde{\dot{x}}_i\tilde{\ddot{s}}_i - \frac{\gamma_1}{n}\tilde{\dot{x}}^\top\tilde{\ddot{s}} + \tilde{\ddot{x}}_i\tilde{\dot{s}}_i - \frac{\gamma_1}{n}\tilde{\ddot{x}}^\top\tilde{\dot{s}}\right)\sin(\alpha)(1-\cos(\alpha)) \\
[\because (53),(52\text{d}),(55)] \quad & \geq \mu_k\sin(\alpha)\Bigg((1-\gamma_1)\sigma - (1+\gamma_1)\eta - (1+\gamma_1)\eta\sin(\alpha) \\
& \quad - (C_4^2n^4 + 2C_2^2n^2)\sin^3(\alpha) - 4C_2C_4n^3\sin^2(\alpha)\Bigg).
\end{aligned}$$

We can derive the same discussion as $g^k(\alpha)$ using (26a). When

$$\sin(\hat{\alpha}_2) \leq \frac{(1-\gamma_1)\sigma - (1+\gamma_1)\eta}{2n^{\frac{3}{2}}}\frac{1}{\max\left\{(1+\gamma_1)\eta, (C_4^2+2C_2^2)^{\frac{1}{3}}, 2\sqrt{C_2C_4}\right\}},$$

from $0 < (1-\gamma_1)\sigma - (1+\gamma_1)\eta < \sigma \leq 1$,

$$\begin{aligned}
& (1-\gamma_1)\sigma - (1+\gamma_1)\eta - (1+\gamma_1)\eta\sin(\hat{\alpha}_2) - (C_4^2n^4+2C_2^2n^2)\sin^3(\hat{\alpha}_2) - 4C_2C_4n^3\sin^2(\hat{\alpha}_2) \\
& \quad \geq ((1-\gamma_1)\sigma - (1+\gamma_1)\eta)\left(1 - \frac{1}{2n^{\frac{3}{2}}} - \frac{1}{2^3n^{\frac{1}{2}}} - \frac{1}{2^2}\right) \\
& \quad \geq ((1-\gamma_1)\sigma - (1+\gamma_1)\eta)\left(1 - \frac{1}{2} - \frac{1}{8} - \frac{1}{4}\right) \\
& \quad \geq 0.
\end{aligned}$$

Therefore, $G_i^k(\alpha) \geq 0$ is satisfied for $\alpha \in (0, \hat{\alpha}_2]$.

Lastly, we consider $h^k(\alpha) \geq 0$. Similarly to the derivation of (54), we can obtain the following:

$$\begin{aligned} x^k(\alpha)^\top s^k(\alpha) \leq\ & (x^k)^\top s^k \left((1-\sin(\alpha)) + \sigma \sin(\alpha)\right) \\ & + \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| (1-\cos(\alpha))^2 + \left(\left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| + \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right|\right) \sin(\alpha)(1-\cos(\alpha)) \\ & + \eta n \mu_k (\sin(\alpha) + 1 - \cos(\alpha)), \end{aligned} \tag{56}$$

upper_x_s_alpha_inner_product

Therefore,

$$\begin{aligned} h^k(\alpha) =\ & (1-(1-\beta)\sin(\alpha))\,(x^k)^\top s^k - x^k(\alpha)^\top s^k(\alpha) \\ [\because (56)] \quad \geq\ & (x^k)^\top s^k \left(\beta \sin(\alpha) - \sigma \sin(\alpha)\right) - \eta n \mu_k (\sin(\alpha) + 1 - \cos(\alpha)) \\ & - \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| (1-\cos(\alpha))^2 - \left(\left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| + \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right|\right) \sin(\alpha)(1-\cos(\alpha)) \\ [\because (5)] \quad =\ & n\mu_k \left(\beta \sin(\alpha) - \sigma \sin(\alpha) - \eta(\sin(\alpha) + 1 - \cos(\alpha))\right) \\ & - \left|\dot{\tilde{x}}^\top \dot{\tilde{s}}\right| (1-\cos(\alpha))^2 - \left(\left|\dot{\tilde{x}}^\top \ddot{\tilde{s}}\right| + \left|\ddot{\tilde{x}}^\top \dot{\tilde{s}}\right|\right) \sin(\alpha)(1-\cos(\alpha)) \\ [\because (52)] \quad \geq\ & n\mu_k \left((\beta - \sigma - \eta)\sin(\alpha) - \eta(1-\cos(\alpha))\right) \\ & - C_2^2 n^2 \mu_k (1-\cos(\alpha))^2 - 2C_2C_4 n^3 \mu_k \sin(\alpha)(1-\cos(\alpha)) \\ [\because (53)] \quad \geq\ & n\mu_k \sin(\alpha) \left((\beta - \sigma - \eta) - \eta \sin(\alpha) - C_2^2 n \sin^3(\alpha) - 2C_2C_4 n^2 \sin^2(\alpha)\right). \end{aligned}$$

The last coefficient on the right-hand side is cubic for $\sin(\alpha)$ and monotonically decreasing for α. Therefore, it is possible to take a step size $\hat{\alpha}_3$ satisfying $h^k(\hat{\alpha}_3) \geq 0$ from (26b). When

$$\sin(\hat{\alpha}_3) \leq \frac{\beta - \sigma - \eta}{2n} \frac{1}{\max\left\{\eta, C_2^{\frac{2}{3}}, \sqrt{2C_2C_4}\right\}},$$

from $0 < \beta - \sigma - \eta < \beta < 1$, we know

$$\begin{aligned} & (\beta - \sigma - \eta) - \eta \sin(\hat{\alpha}_3) - C_2^2 n \sin^3(\hat{\alpha}_3) - 2C_2C_4 n^2 \sin^2(\hat{\alpha}_3) \\ & \geq (\beta - \sigma - \eta) - \frac{\beta - \sigma - \eta}{2n} - \frac{(\beta - \sigma - \eta)^3}{8n^2} - \frac{(\beta - \sigma - \eta)^2}{4} \\ & > (\beta - \sigma - \eta)\left(1 - \frac{1}{2} - \frac{1}{8} - \frac{1}{4}\right) \\ & = \frac{\beta - \sigma - \eta}{8} > 0. \end{aligned}$$

Therefore, $g^k(\alpha) \geq 0$ is satisfied for $\alpha \in (0, \hat{\alpha}_3]$.

From the above discussions, when $\hat{\alpha}$ is taken such that

$$\sin(\hat{\alpha}) = \frac{1}{n^{\frac{3}{2}}} \frac{\min\left\{(1-\gamma_1)\sigma - (1+\gamma_1)\eta, \beta - \sigma - \eta\right\}}{2\max\left\{(1+\gamma_1)\eta, (C_4^2 + 2C_2^2)^{\frac{1}{3}}, 2\sqrt{C_2C_4}\right\}}, \tag{57}$$

def_min_step_size

$g^k(\alpha), G_i^k(\alpha), h^k(\alpha) \geq 0$ are satisfied for all k and $\alpha \in (0, \hat{\alpha}]$. □

Since $\hat{\alpha}$ defined in (57) can satisfy the conditions in line 6 of Algorithm 1, we can find the step length $\alpha_k \geq \hat{\alpha} > 0$. Therefore, Algorithm 1 is well-defined. From $h^k(\alpha_k) \geq 0$ for all k,

mu-decrement

$$\begin{aligned} h^k(\alpha_k) \geq 0 \Rightarrow x^k(\alpha_k)^\top s^k(\alpha_k) &\leq (1-(1-\beta)\sin(\alpha_k))(x^k)^\top s^k \\ &\leq (1-(1-\beta)\sin(\hat{\alpha}))(x^k)^\top s^k \\ &\leq (1-(1-\beta)\sin(\hat{\alpha}))^k (x^0)^\top s^0. \end{aligned} \tag{58}$$

Due to (29), it also holds that

constraints_residual_decrement

$$\|(r_b(x^k), r_c(y^k, s^k))\| \leq (1-\sin(\hat{\alpha}))^k \|(r_b(x^0), r_c(y^0, s^0))\|. \tag{59}$$

We can prove the polynomial complexity of the proposed method based on the following theorem.

polynomiality_by_mu

Theorem 4.1 ([37, Theorem 1.4])**.** *Suppose that an algorithm for solving* (3) *generates a sequence of iterations that satisfies*

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n^\omega}\right)\mu_k, \quad k = 0, 1, 2, \ldots,$$

for some positive constants δ and ω. Then there exists an index K with

$$K = \mathcal{O}(n^\omega \log(\mu_0/\zeta))$$

such that

$$\mu_k \leq \zeta \text{ for } \forall k \geq K.$$

Applying (58), (12), $(x^k, y^k, s^k) \in \mathcal{N}(\gamma_1, \gamma_2)$, (59) and a result that $\sin(\hat{\alpha})$ is propositional to $n^{-1.5}$ in (57) to this theorem, we can obtain the following theorem.

main-theorem

Theorem 4.2. *Algorithm 1 generates a ζ-optimal solution in at most*

$$\mathcal{O}\left(n^{1.5}\log\left(\frac{\max\{\mu_0, \|r_b(x^0), r_c(y^0, s^0)\|\}}{\zeta}\right)\right)$$

iterations.

In the case that the input data is integral, Al-Jeiroudi et al. [1] and Mohammadisiahroudi et al. [23] analyze that the iteration complexity of II-line is $\mathcal{O}(n^2 L)$, where L is the binary length of the input data denoted by

$$L = mn + m + n + \sum_{i,j} \lceil \log(|a_{ij}| + 1) \rceil + \sum_i \lceil \log(|c_i| + 1) \rceil + \sum_j \lceil \log(|b_j| + 1) \rceil.$$

Theorem 4.2 indicates that II-arc can reduce the iteration complexity from n^2 to $n^{1.5}$, by a factor of $n^{0.5}$. This reduction is mainly brought by the ellipsoidal approximation in the arc-search method.

5 Numerical experiments

section_numerical_experiments

In this section, we describe the implementation and the numerical experiments of the proposed method. The experiments were conducted on a Linux server with Opteron 4386 (3.10GHz), 16 cores, and 128GB RAM, and the methods were implemented with Python 3.10.9.

5.1 Implementation details

We describe the implementation details before discussing the results.

5.1.1 Parameter settings

In these numerical experiments, we set

$$\sigma = 0.4, \quad \eta = 0.3, \quad \gamma_1 = 0.1, \quad \gamma_2 = 1, \quad \beta = 0.9.$$

These parameters satisfy (26), and we use the same parameters for II-line as well.

5.1.2 Solving LESs

To solve the LESs inexactly, we employ the conjugate gradient (CG) method in Scipy package. Although we examined other iterative solvers than CG, the preliminary experiments showed that CG was the fastest inexact solver in II-arc.

The proposed method uses the MNES formulation in Section 3, but preliminary experiments showed that MNES lacks numerical stability. Specifically, CG did not converge to a certain accuracy even when a preconditioner was employed, and the search direction did not satisfy (24). A possible cause is that the condition number of the coefficient matrix $\hat{M}^k$ for MNES is extremely worse than that for NES; it is known that the condition number of MNES can grow up to the square of that of NES [26].

Therefore, in the numerical experiments, we choose the NES formulations (14) and

second_derivative_NES

$$M^k \ddot{y} = \rho_2^k, \tag{60}$$

instead of the MNES (16) and (20), respectively. The inexact solution of (14) satisfies

inexact_first_derivative_NES

$$M^k \tilde{\dot{y}} = \rho_1^k + r_1^k, \tag{61}$$

where the error r_1^k is defined as $r_1^k := M^k \tilde{\dot{y}} - \rho_1^k = M^k \left(\tilde{\dot{y}} - \dot{y} \right)$, and that of (60) satisfies

inexact_second_derivative_NES

$$M^k \tilde{\ddot{y}} = \rho_2^k + r_2^k, \tag{62}$$

where the error r_2^k is defined similar to r_1^k. As for the solution accuracy, we set the following threshold as in (25):

def_upper_derivatives_residual_NES

$$\left\| r_i^k \right\| \leq \eta \frac{\sqrt{\mu_k}}{\sqrt{n}} \quad \forall i \in \{1, 2\}. \tag{63}$$

When we solve (14) and (60) by CG, we use the inverse matrix of the diagonal components of M^k as the preconditioner matrix to speed up its convergence of CG [12]. We adopt this preconditioner because it is simpler than the other methods, such as the controlled Cholesky Factorization preconditioner [5], the splitting preconditioner [27], and the hybrid of these [2], and we checked its convergence in a preliminary test.

The coefficient matrix M^k has to be a symmetric positive definite matrix when solving (14) and (60) in II-arc and [23, (NES)] in II-line by CG of Scipy. Though this condition should hold theoretically from Assumption 2.2 and $x^k, s^k > 0$, M^k may not be positive definite due to numerical errors. Therefore, when the CG method fails to satisfy (63), we replace M^k with $M^k + 10^{-3}I$, as indicated in [20].

5.1.3 The modification of $(\tilde{\ddot{x}}, \tilde{\ddot{y}}, \tilde{\ddot{s}})$

If $\left\|-2\tilde{\dot{x}} \circ \tilde{\dot{s}}\right\|_\infty \le \eta\mu_k$ is satisfied, (28) and (35) can hold with $(\tilde{\ddot{x}}, \tilde{\ddot{y}}, \tilde{\ddot{s}}) = (0, 0, 0)$. Therefore, to shorten the computation time, we skip solving (60) and set $(\tilde{\ddot{x}}, \tilde{\ddot{y}}, \tilde{\ddot{s}}) = (0, 0, 0)$. In this case, (23) can be interpreted as a line-search method.

Furthermore, when the inexact solution of (60) satisfies $\left\|M_2^k \tilde{\ddot{y}} - \rho_2^k\right\| > \left\|\rho_2^k\right\|$, $\tilde{\ddot{y}}$ is replaced with a zero vector as in [22] to avoid a large error.

5.1.4 Step size

In line 6 of Algorithm 1 and [23, Algorithm 1, Line 9], since it is difficult to obtain the solution of (24) analytically, Armijo's rule [31] is employed to determine an actual step size α_k.

5.1.5 Stopping criteria

section_stopping_criteria

The algorithms are designed to terminate when $(x^k, y^k, s^k) \in \mathcal{S}^*_\zeta$ is satisfied. The condition $\mu_k \le \zeta$, however, does not consider the magnitude of the data, thus it is not practical especially when the magnitude of the optimal values is relatively large.

Therefore, in addition to condition $\mu_k \le \zeta$ (where $\zeta = 10^{-2}$), as in [33], we terminate the algorithms when the following condition is met:

$$\max\left\{\frac{\left\|r_b(x^k)\right\|}{\max\{1, \|b\|\}}, \frac{\left\|r_c(y^k, s^k)\right\|}{\max\{1, \|c\|\}}, \frac{\mu_k}{\max\left\{1, \|c^\top x^k\|, \|b^\top y^k\|\right\}}\right\} < \epsilon, \qquad (64)$$

condition_solved

where we set the threshold $\epsilon = 10^{-7}$.

In addition, we stop the algorithm prematurely when the step size α_k diminishes as $\alpha_k < 10^{-7}$.

5.2 Test problems

section_test_problems

The CG or other iterative solvers are often employed when the matrix related to the normal equation is very large and makes the Cholesky factorization impractical. In this context, we use the largest problems in the NETLIB collection [6]; QAP15 and the fifteen Kennington problems [7] except KEN-18[1]. We applied the same preprocessing as in [15, Section 5.1] to the problems, e.g., removing redundant rows of the constraint matrix A.

5.3 Numerical Results

section_numerical_results

We report numerical results as follows. In Section 5.3.1, we compare II-arc and the inexact infeasible line-search IPM [23, Algorithm 1] (II-line), and show II-arc can solve the large problems with less iterations and computation time. In Section 5.3.2, we compare II-arc and the existing exact infeasible IPMs. This result indicates that the proposed method requires more iterations but less computation time.

The detailed numerical results of the all methods are reported in Appendix A.

5.3.1 Comparison with the Inexact line-search

section_comparison_II_line_IPM

We compare II-arc with II-line by solving the benchmark problems using CG in this section. We set the initial point as $(x^0, y^0, s^0) = 10^4(e, 0, e)$ that always satisfies $(x^0, y^0, s^0) \in \mathcal{N}(\gamma_1, \gamma_2)$.

Firstly, Figure 1 shows a performance profile [9, 13] on the numbers of iterations of II-arc and II-line. The figures on the performance profile in this section was generated with a Julia package [28].

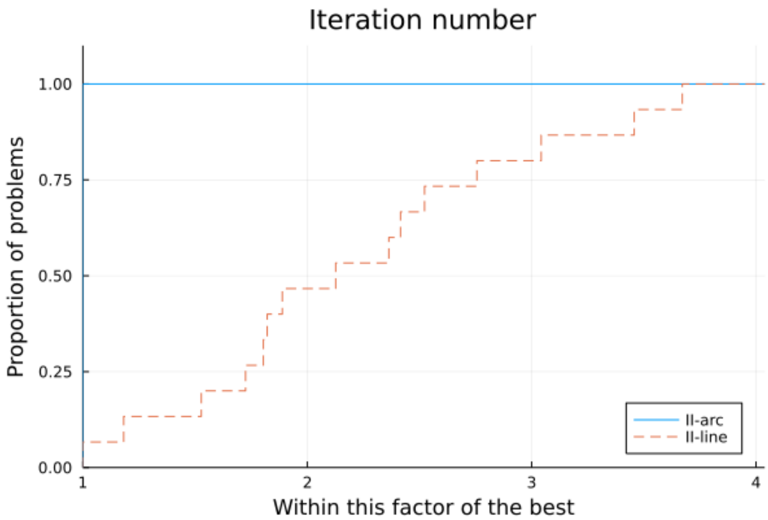

Figure 1: Performance profile of the number of iterations with II-arc and II-line fig_comparison_iter_num_with_existing_method

We observe from Figure 1 that II-arc demands fewer iterations than II-line in all problems. For more than half of the test problems, II-line required more than twice as many iterations as II-arc. Therefore, these results indicate that the number of iterations

[1]The size of KEN-18 (n = 255248 and m = 205676) was so large that all of the methods in this section exceeded the time limit of 36000 seconds.

can be reduced by approximating the central path with the ellipsoidal arc, when the LESs for the search direction are solved inexactly.

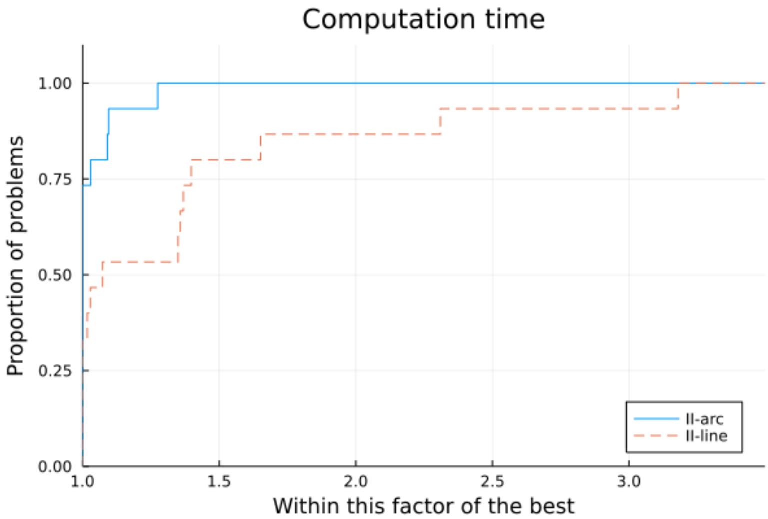

Figure 2: Performance profile of the computation time with II-arc and II-line. fig_comparison_calc_time_with_existing_method

Next, Figure 2 shows a performance profile on the computation time. The computation time of II-arc is shorter than that of II-line. These results show that even though II-arc requires an additional LES (62) to be solved, II-arc can solve the large problems faster than II-line due to the reduction in the number of iterations.

5.3.2 Comparison with the existing exact IPMs

section_comparison_Exact_IPM

Next, we compare II-arc and the exact infeasible IPMs; the arc-search IPM [35] (EI-arc) and the Mehrotra-type line-search IPM [21] (EI-line). We employ Scipy's Cholesky factorization to solve the LES exactly. We exclude KEN-18, OSA-60 and PDS-20 from the comparison, since the computation exceeded the time limit of 36000 seconds due to the Cholesky factorization for the extremely large LESs. For the initial points, the II-arc method uses the same initial points as in Section 5.3.1. On the other hand, since $(x^0, y^0, s^0) \in \mathcal{N}(\gamma_1, \gamma_2)$ is not required for EI-arc and EI-line, these use the same method as Yang [35, Section 4.1]. Therefore, EI-arc and EI-line generate initial point candidates using the Mehrotra method [21] and the Lusting one [19], and select the one.

Figure 3 shows the performance profile for the number of iterations. This figure shows that II-arc is inferior to the exact methods. If the exact search direction can be calculated, it can be inferred that the number of iterations can be reduced.

Next, Figure 4 shows the performance profile of the computation time. This figure shows that II-arc has an advantage in terms of computation time in spite of a larger number of iterations. When solving the LESs (14) and (60) for the search direction, the Cholesky factorization requires $\mathcal{O}(n^3)$ of the computational complexity, whereas CG requires $\mathcal{O}(nd\sqrt{\kappa}\log(1/\varepsilon))$ [29], where d being the maximum number of non-zero elements

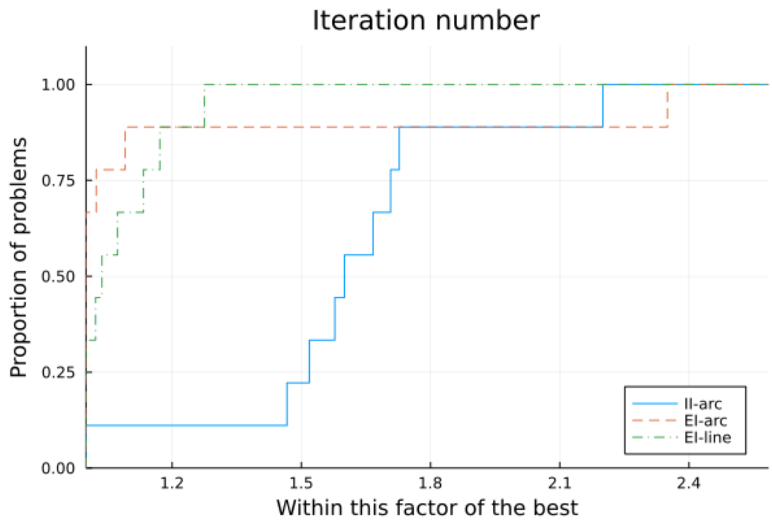

Figure 3: Performance profile of the iteration number with II-arc, EI-arc and EI-line.
fig_comparison_inexact_and_exact_in_iter_num

in any row or column of M^k, κ the condition number of M^k, and ε the error allowed (we set ε satisfying (63) in II-arc and II-line). It is known that κ increases as the iterations proceed in the IPM [12], but CG still can be faster than the Cholesky factorization if n is remarkably large. For instance, when CRE-B was solved, the overall II-arc method took 255.85 seconds, of which the CG consumed 247.24 seconds (96.63%). In contrast, the EI-arc method took 657.61 seconds and the Cholesky factorization occupies 639.02 seconds (97.17%). Therefore, the time required to find the search direction per iteration can be shorter in the inexact IPM than that in the exact one, and as a result, the entire computation time can be reduced.

6 Conclusion

section_conclusion

In this work, we proposed an inexact infeasible arc-search interior-point method (II-arc) for solving LPs. In particular, by formulating MNES and setting the parameters appropriately, we showed that the proposed method achieves polynomial iteration complexity that is smaller than the II-line by a factor of $n^{0.5}$. In the numerical experiments for the largest problems in the NETLIB collection with CG as the solver for the LESs, the II-arc outperformed the II-line in terms of both the number of iterations and the computation time due to the reduction in the computational complexity by the arc-search. Additionally, solving the LESs inexactly resulted in a reduction of the computation time compared to the existing exact IPMs for the large problems because the computational complexity of CG is less dependent on the problem size n than that of the Cholesky factorization.

As a future direction, we can consider the following:

- utilizing QLSA, such as the Harrow-Hassidim-Lloyd algorithm [14], to solve the

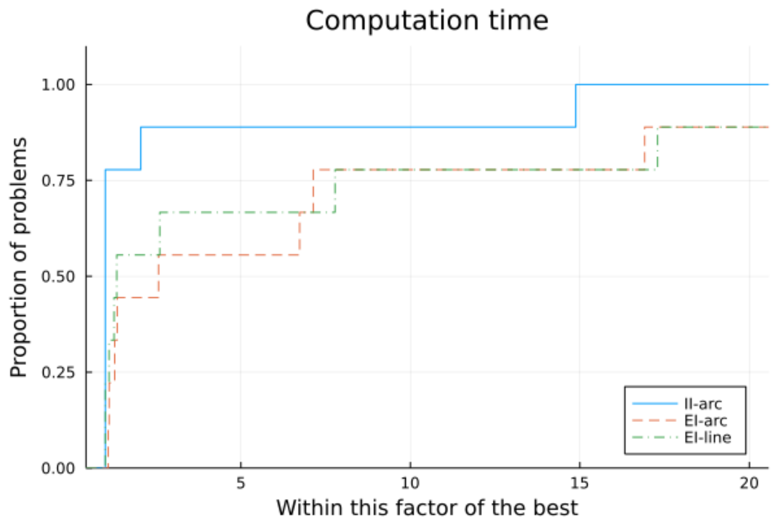

Figure 4: Performance profile of the computation time with II-arc, EI-arc and EI-line. fig_comparison_inexact_and_exact_in_calc_time

LESs more quickly,

- combining Nesterov's restarting strategy as in [15] to shorten the entire computation time,
- exploring hybrid methods to improve the efficiency of inexact solutions, such as Bartmeyer et al. [2],
- extending the approach to other optimization problems, such as quadratic programming problems.

References

[1] G. Al-Jeiroudi and J. Gondzio. Convergence analysis of the inexact infeasible interior-point method for linear optimization. *Journal of Optimization Theory and Applications*, 141:231–247, 2009.

[2] P. M. Bartmeyer, S. Bocanegra, and A. R. L. Oliveira. Switching preconditioners using a hybrid approach for linear systems arising from interior point methods for linear programming. *Numerical Algorithms*, 86:397–424, 2021.

[3] S. Bellavia. Inexact interior-point method. *Journal of Optimization Theory and Applications*, 96:109–121, 1998.

[4] S. Bellavia and S. Pieraccini. Convergence analysis of an inexact infeasible interior point method for semidefinite programming. *Computational Optimization and Applications*, 29:289–313, 2004.

[5] S. Bocanegra, F. F. Campos, and A. R. Oliveira. Using a hybrid preconditioner for solving large-scale linear systems arising from interior point methods. *Computational Optimization and Applications*, 36:149–164, 2007.

[6] S. Browne, J. Dongarra, E. Grosse, and T. Rowan. The Netlib mathematical software repository. *D-lib Magazine*, 1(9), 1995.

[7] W. J. Carolan, J. E. Hill, J. L. Kennington, S. Niemi, and S. J. Wichmann. An empirical evaluation of the korbx® algorithms for military airlift applications. *Operations Research*, 38(2):240–248, 1990.

[8] P. A. M. Casares and M. A. Martin-Delgado. A quantum interior-point predictor-corrector algorithm for linear programming. *Journal of Physics A: Mathematical and Theoretical J. Phys. A: Math. Theor*, 53:30, 2020.

[9] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical programming*, 91:201–213, 2002.

[10] T. A. Espaas and V. S. Vassiliadis. An interior point framework employing higher-order derivatives of central path-like trajectories: Application to convex quadratic programming. *Computers & Chemical Engineering*, 158:107638, 2022.

[11] J. Gondzio. Multiple centrality corrections in a primal-dual method for linear programming. *Computational Optimization and Applications*, 6(2):137–156, 1996.

[12] J. Gondzio. Interior point methods 25 years later. *European Journal of Operational Research*, 218(3):587–601, 2012.

[13] N. Gould and J. Scott. A note on performance profiles for benchmarking software. *ACM Transactions on Mathematical Software (TOMS)*, 43(2):1–5, 2016.

[14] A. W. Harrow, A. Hassidim, and S. Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103(15):150502, 2009.

[15] E. Iida and M. Yamashita. An infeasible interior-point arc-search method with Nesterov's restarting strategy for linear programming problems. *Computational Optimization and Applications*, pages 1–34, 2024.

[16] N. Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 302–311, 1984.

[17] I. Kerenidis and A. Prakash. A quantum interior point method for LPs and SDPs. *ACM Transactions on Quantum Computing*, 1(1):1–32, 2020.

[18] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point algorithm for linear programming. In *Progress in Mathematical Programming*, pages 29–47. Springer, New York, 1989.

[19] I. J. Lustig, R. E. Marsten, and D. F. Shanno. On implementing mehrotra's predictor–corrector interior-point method for linear programming. *SIAM journal on Optimization*, 2(3):435–449, 1992.

[20] A. Malyshev, R. Quirynen, and A. Knyazev. Preconditioning of conjugate gradient iterations in interior point mpc method. *IFAC-PapersOnLine*, 51(20):394–399, 2018.

[21] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, 2:575–601, 1992.

[22] S. Mizuno and F. Jarre. Global and polynomial-time convergence of an infeasible-interior-point algorithm using inexact computation. *Mathematical Programming*, 84(1), 1999.

[23] M. Mohammadisiahroudi, R. Fakhimi, and T. Terlaky. Efficient use of quantum linear system algorithms in interior point methods for linear optimization. *arXiv preprint arXiv:2205.01220*, 2022.

[24] R. D. Monteiro, I. Adler, and M. G. Resende. A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension. *Mathematics of Operations Research*, 15(2):191–214, 1990.

[25] R. D. Monteiro and J. W. O'Neal. Convergence analysis of a long-step primal-dual infeasible interior-point lp algorithm based on iterative linear solvers. *Georgia Institute of Technology*, 2003.

[26] R. D. Monteiro, J. W. O'Neal, and T. Tsuchiya. Uniform boundedness of a preconditioned normal matrix used in interior-point methods. *SIAM Journal on Optimization*, 15(1):96–100, 2004.

[27] A. R. Oliveira and D. C. Sorensen. A new class of preconditioners for large-scale linear systems from interior point methods for linear programming. *Linear Algebra and its applications*, 394:1–24, 2005.

[28] D. Orban and contributors. BenchmarkProfiles.jl: A Simple Julia Package to Plot Performance and Data Profiles. `https://github.com/JuliaSmoothOptimizers/BenchmarkProfiles.jl`, February 2019.

[29] Y. Saad. *Iterative methods for sparse linear systems.* SIAM, PA, 2003.

[30] F. Vitor and T. Easton. Projected orthogonal vectors in two-dimensional search interior point algorithms for linear programming. *Computational Optimization and Applications*, 83(1):211–246, 2022.

[31] S. J. Wright. *Primal-dual interior-point methods.* SIAM, PA, 1997.

[32] Z. Wu, M. Mohammadisiahroudi, B. Augustino, X. Yang, and T. Terlaky. An inexact feasible quantum interior point method for linearly constrained quadratic optimization. 2022.

[33] M. Yamashita, E. Iida, and Y. Yang. An infeasible interior-point arc-search algorithm for nonlinear constrained optimization. *Numerical Algorithms*, 2021.

[34] Y. Yang. A polynomial arc-search interior-point algorithm for convex quadratic programming. *European Journal of Operational Research*, 215(1):25–38, 2011.

[35] Y. Yang. CurveLP-A MATLAB implementation of an infeasible interior-point algorithm for linear programming. *Numerical Algorithms*, 74:967–996, 4 2017.

[36] Y. Yang. Two computationally efficient polynomial-iteration infeasible interior-point algorithms for linear programming. *Numerical Algorithms*, 79(3):957–992, 2018.

[37] Y. Yang. *Arc-search techniques for interior-point methods*. CRC Press, FL, 2020.

[38] Y. Yang. A polynomial time infeasible interior-point arc-search algorithm for convex optimization. *Optimization and Engineering*, 24(2):885–914, 2023.

[39] Y. Yang and M. Yamashita. An arc-search $O(nL)$ infeasible-interior-point algorithm for linear programming. *Optimization Letters*, 12(4):781–798, 2018.

A Details on numerical results

section_appendix

Table 1 reports the numerical results in Section 5.3. The first column of the table is the problem name, and the second and the third are the variable size n and the number of constraints m, respectively, after preprocessing denoted in Section 5.2. The fourth to last columns report the number of iterations and the computation time (in seconds). The underlined results indicate the best results among the four methods. A mark '-' indicates the algorithms stop before reaching the optimality, since the step size α_k diminishes prematurely. In columns of EI-arc and EI-line, '*' means that these methods exceeded the time limit of 36000 seconds.

Table 1: Numerical results on the proposed method and the existing methods table_results_for_comparison

problem	n	m	II-arc		II-line		EI-arc		EI-line	
			Itr.	Time	Itr.	Time	Itr.	Time	Itr.	Time
CRE-A	6997	3299	41	37.18	113	39.88	27	41.43	28	41.28
CRE-B	36382	5336	70	255.85	257	357.5	41	657.61	44	666.94
CRE-C	5684	2647	44	50.45	111	51.86	30	26.72	30	24.66
CRE-D	28601	4102	70	175.99	242	240.82	43	237.77	42	221.57
KEN-07	5127	3951	33	23.66	39	21.61	15	2.03	17	1.59
KEN-11	32996	26341	36	2239.44	55	2276.8	-	-	-	-
KEN-13	72784	58757	46	16841.77	83	16367.94	-	-	-	-
OSA-07	25067	1118	40	4.45	69	6.04	94	31.75	51	5.95
OSA-14	54797	2337	45	10.29	85	16.99	-	-	60	14.85
OSA-30	104374	4350	44	18.87	104	43.58	-	-	66	39.44
OSA-60	243246	10280	47	52.13	143	165.76	*	*	*	*
PDS-06	36920	17604	56	128.64	102	117.98	35	866.18	41	1000.53
PDS-10	63905	30773	71	420.27	151	420.29	45	7107.41	46	7267.52
PDS-20	139330	65437	89	17269.62	215	23297.11	*	*	*	*
QAP15	22275	6330	19	2.46	19	1.93	12	580.75	11	534.94