INDUSTRIAL AND SYSTEMS ENGINEERING

ISE

Ramification of Weighted Central Paths and Super-linear Interior-Point Methods for Semi-definite Optimization

Ali Mohammad-Nezhad[1], Pouya Sampourmahani[2], and Tamás Terlaky[2]

[1]Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, USA

[2]Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA, USA

Technical Report 26T-021

LEHIGH UNIVERSITY

arXiv:2611.07080v1 [math.OC] 4 Nov 2026

RAMIFICATION OF WEIGHTED CENTRAL PATHS AND SUPER-LINEAR INTERIOR-POINT METHODS FOR SEMI-DEFINITE OPTIMIZATION

ALI MOHAMMAD-NEZHAD, POUYA SAMPOURMAHANI, AND TAMÁS TERLAKY

Abstract. The local convergence of primal-dual interior-point methods for semi-definite optimization is closely tied to the regularity of the central path at its limiting point. Without strict complementarity, the central path can exhibit fractional power asymptotics and need not be differentiable in the ordinary barrier parameter.

We show that this obstruction can be removed simultaneously for a family of weighted AHO central paths. More precisely, there exists a common ramification exponent N such that all weighted paths in a fixed neighborhood of the identity become analytic at the limiting point after the reparametrization $\mu \mapsto \mu^N$. Using higher-order derivatives of these reparametrized paths, we construct a feasible pth-order predictor-corrector interior-point method.

Under a uniform predictor regularity condition, and without assuming strict complementarity, we prove that the normalized complementarity gaps converge with Q-order at least $(p+1)/\rho$, where ρ is the reparametrization exponent used by the method. We also establish super-linear convergence of the simultaneously vanishing primal-dual blocks, derive an explicit quantitative reduction bound throughout the predictor regularity region, and obtain a doubly logarithmic outer iteration bound once the iterates enter the local super-linear regime. Finally, we illustrate the predicted higher-order behavior on a collection of degenerate and non-degenerate semi-definite optimization problems.

Contents

1. Introduction 2
1.1. IPMs and central path 2
2. Main results 3
2.1. Ramification of weighted central paths 3
2.2. A higher-order IPM 4
2.3. Outline of the proofs 6
3. Background 7
3.1. Puiseux series 7
3.2. Semi-algebraic sets 8
3.3. Sign conditions, univariate representations, and Thom encodings 8
3.4. Cylindrical algebraic decomposition 9
3.5. Analytic reparametrization of the central path 9
4. Prior and related work 12
4.1. Central path 12
4.2. Weighted central paths and regularity 12
4.3. Local convergence of primal-dual IPMs 13
4.4. Algebraic methods and IPMs 15
5. Analytic reparametrization of weighted central paths 15
6. A higher-order IPM 20
6.1. AHO search directions 20
6.2. A higher-order IPM based on AHO search directions 22

2020 *Mathematics Subject Classification.* Primary 90C22, 90C51; Secondary 14P10.

Key words and phrases. Semi-definite optimization, interior-point methods, weighted central paths, semi-algebraic functions, Puiseux expansion, ramification index.

7. Proof of the main results 29
7.1. Local convergence 29
7.2. Iteration complexity 34
8. Numerical experiments 36
8.1. Initialization 37
8.2. Modification of the Newton systems 37
8.3. Experiments 38
8.4. Convergence of the sequence of complementarity gaps 46
9. Concluding remarks and future research 50
Acknowledgments 51
References 51

1. Introduction

Semi-definite optimization (SDO) is the minimization of a linear objective over the intersection of an affine subspace and the cone of positive semi-definite matrices. Let $\mathbb{S}^n$ be the vector space of symmetric matrices endowed with the inner product $\langle X, S\rangle := \mathrm{Tr}(XS)$. A pair of primal-dual SDO problems is defined as

$$\text{(P)} \qquad z_p^* := \inf_{X\in\mathbb{S}^n} \{\langle C, X\rangle \mid \langle A_i, X\rangle = b_i,\ i = 1,\dots,m,\ X \succeq 0\},$$

$$\text{(D)} \qquad z_d^* := \sup_{(y,S)\in\mathbb{R}^m\times\mathbb{S}^n} \left\{ b^T y \mid \sum_{i=1}^m y_i A_i + S = C,\ S \succeq 0,\ y \in \mathbb{R}^m \right\},$$

where $\succeq 0$ means positive semi-definite.

Notation 1. We use the notation $(\cdot,\cdot,\cdots,\cdot)$ to represent vectors of $\mathbb{R}^n$ or an arbitrary vector space. A primal-dual vector of SDO, identified by $(X, y, S) \in \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n$, is called an *optimal solution* if $X \in \mathcal{P}^*$ and $(y, S) \in \mathcal{D}^*$, where

$$\mathcal{P}^* := \left\{X \mid \langle A_i, X\rangle = b_i, \quad i = 1,\dots,m, \quad X \succeq 0, \quad \langle C, X\rangle = z_p^*\right\},$$

$$\mathcal{D}^* := \left\{(y, S) \mid \sum_{i=1}^m y_i A_i + S = C, \quad S \succeq 0,\ b^T y = z_d^*\right\}.$$

For a primal-dual feasible vector (X, y, S), we define the *normalized complementarity gap* by

$$\mu(X, S) := \frac{\langle X, S\rangle}{n}.$$

To avoid trivialities, we make the following assumptions from this point on, which in turn guarantee that $\mathcal{P}^*, \mathcal{D}^*$ are non-empty and compact, and $z_p^* = z_d^*$ (see e.g., [76, Corollary 4.2]).

Assumption 1. *The matrices $A_1,\dots,A_m$ are linearly independent in $\mathbb{S}^n$, and there exists a primal-dual feasible solution (X, y, S) with $X, S \succ 0$.*

SDO lies at the intersection of convex optimization, computational mathematics, and real algebraic geometry, with applications ranging from combinatorial optimization and control to polynomial optimization and sums of squares [39]. In this paper, our focus is on the interaction between the semi-algebraicity of weighted central paths and the local convergence of interior-point methods (IPMs).

1.1. IPMs and central path.

Analogous to linear optimization (LO), primal-dual path-following IPMs [1, 54] are among the most efficient algorithms to solve SDO numerically. The core idea, originating from [18], is to follow the *central path*. In this paper, we consider the following formulation of the central path (see also Remark 9): the central path of SDO is the map

$$\xi : (0,\infty) \longrightarrow \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n, \qquad \xi(\mu) := (X(\mu), y(\mu), S(\mu)),$$

where for each fixed $\mu > 0$, $(X(\mu), y(\mu), S(\mu))$ uniquely solves

$$\begin{aligned} &\langle A_i, X\rangle = b_i, \quad i = 1, \dots, m, \\ &\sum_{i=1}^{m} y_i A_i + S - C = 0, \\ &XS + SX - 2\mu I_n = 0,\ X, S \succ 0, \end{aligned} \tag{1.1}$$

in which the third equation is referred to as the centrality condition. It is easy to see that along the central path, the normalized complementarity gap coincides with the path parameter μ. Under Assumption 1, the central path exists and is real analytic; see [14, Theorems 3.1 and 3.3]. Further, ξ is a *semi-algebraic* function (see Section 3.2) and converges to an optimal solution (see Proposition 5.2 with $W = I_n$ or [8, 10]).

The machinery of primal-dual path-following `IPM`s involves generating a trajectory of solutions that approximately solve (1.1), while taking $\mu \downarrow 0$, towards an approximate optimal solution of `SDO`. The local behavior of primal-dual `IPM`s is closely tied to the regularity of the central path as $\mu \downarrow 0$. Under *strict complementarity* (see Definition 3.2), the central path admits an analytic extension to its limiting point, and this regularity underlies classical super-linear and quadratic convergence results. Without strict complementarity, however, the geometry changes qualitatively. The central path need not be differentiable at $\mu = 0$, and its local expansions may involve fractional powers of μ. Thus, the ordinary barrier parameter does not in general provide the higher-order derivatives needed to construct higher-order predictor directions.

The central idea of this paper is to overcome this obstruction by ramification. We prove that an entire family of *weighted `AHO` central paths* admits a common analytic reparametrization near the limiting point. The resulting higher-order derivatives are then used directly to construct and analyze a higher-order predictor-corrector `IPM`.

2. Main results

Our results connect three components: the ramified regularity of weighted central paths, a higher-order predictor-corrector construction, and its convergence without strict complementarity.

2.1. Ramification of weighted central paths.

The core of our higher-order `IPM` is a semi-algebraic description of weighted central paths, which allows for their analytic reparametrization. For the purpose of our algorithmic development, we consider a *feasible* parameterized family of *weighted central paths* defined by functions

$$\xi(\mu, W) := (X(\mu, W), y(\mu, W), S(\mu, W)) : (0, \infty) \times \mathbb{S}^n_{++} \to \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n,$$

where for each fixed $\mu > 0$ and $W \succ 0$, $(X(\mu, W), y(\mu, W), S(\mu, W))$ uniquely solves

$$\begin{aligned} &\langle A_i, X\rangle = b_i, \quad i = 1, \dots, m, \\ &\sum_{i=1}^{m} y_i A_i + S - C = 0, \\ &XS + SX - 2\mu W = 0,\ X, S \succ 0, \end{aligned} \tag{2.1}$$

where $\mathbb{S}^n_{++}$ denotes the interior of the positive semi-definite cone. For a fixed $W \succ 0$, we denote the corresponding weighted central path by

$$\xi_W(\mu) := \xi(\mu, W), \quad X_W(\mu) := X(\mu, W), \quad y_W(\mu) := y(\mu, W), \quad S_W(\mu) := S(\mu, W),$$

and with a fixed $\mu > 0$ $\big(X_W(\mu), y_W(\mu), S_W(\mu)\big)$ is called a *weighted central solution*.

When $W = I_n$, a weighted central path is equivalent to the central path defined in (1.1). Under Assumption 1, when $W \succ 0$, $\xi(\mu, W)$ exists and is analytic on $(0, \infty) \times \mathbb{S}^n_{++}$; see Proposition 5.1. Further, $\xi(\mu, W)$ is jointly semi-algebraic, and for a fixed $W \succ 0$, it converges to an optimal solution as $\mu \to 0$; see Proposition 5.2. More importantly, for the

purposes of this paper, we establish the existence of a minimal positive integer N such that the reparametrized path $\xi_W(\mu^N)$ is analytic at $\mu = 0$ over a bounded subset

$$\mathcal{W}_\beta := \{W \succ 0 \mid \|W - I_n\|_F \leq \beta\},$$

where $0 < \beta < 1$ is a fixed parameter. This analytic reparametrization arises naturally from the *Puiseux expansion* of weighted central paths, which we formally introduce in Section 3.1.

Notation 2. Throughout the paper, $\|\cdot\|_F$ denotes the Euclidean norm for vectors and the Frobenius norm for matrices.

Theorem 1. *For every fixed $0 < \beta < 1$, there exists a minimal positive integer $N = N(\beta)$ such that, for every $W \in \mathcal{W}_\beta$, the reparametrized path $\xi_W(\mu^N)$ is analytic at $\mu = 0$.*

Theorem 1 provides a common ramified reparametrization, but does not assert joint analyticity in (μ, W) or uniform bounds on derivatives as W varies. The latter requirement enters separately through the predictor regularity condition used in the convergence analysis of our higher-order `IPM`.

Remark 1 (Reparametrization exponent). The integer N in Theorem 1 is the smallest positive integer that provides a common analytic reparametrization of the weighted central paths. We call a positive integer ρ *admissible* on $\mathcal{W}_\beta$ if $\xi_W(\mu^\rho)$ is analytic at $\mu = 0$ for every $W \in \mathcal{W}_\beta$. In particular, $\rho = N$ is admissible, and every positive integer multiple of N is also admissible. In what follows, we use ρ to denote the reparametrization exponent employed by the higher-order `IPM`.

Definition 2.1. We refer to N as the *global ramification index* of weighted central paths on $\mathcal{W}_\beta$.

Remark 2. The proof of Theorem 1 can be modified to extend the result to all $W \succ 0$. However, doing so would require a substantially longer and more cumbersome statement, which is unnecessary for the algorithmic development in this paper.

2.2. **A higher-order `IPM`.** We develop a higher-order feasible predictor-corrector `IPM` based on the `AHO` (Alizadeh-Haeberly-Overton [3]) search directions. Our method exploits the analytic reparametrization $\mu \mapsto \mu^\rho$ of weighted central paths to construct higher-order predictor directions. Specifically, given an initial strictly feasible solution, the predictor directions are obtained from the pth-order Taylor approximation of the reparametrized weighted central path. We forward-track this path with respect to an outer neighborhood of the reparametrized central path. Upon reaching the boundary of the neighborhood, we apply centering to return to an inner neighborhood of the reparametrized central path.

Remark 3. Our higher-order `IPM` does not require the use of an admissible reparametrization exponent ρ. Indeed, any positive integer ρ for which the reparametrized weighted central paths possess derivatives through order p can be used to construct the pth-order predictor. Theorem 1 guarantees the existence of such a choice, since one may always take $\rho = N$, or any positive integer multiple of N.

Admissibility becomes important in our asymptotic convergence analysis. In particular, Theorems 2 and 3 require an admissible reparametrization exponent together with suitable uniform predictor regularity of the reparametrized weighted central paths.

We establish the super-linear convergence of the complementarity gaps $\langle X_k, S_k\rangle/n$ of Algorithm 1, without the strict complementarity condition.

Notation 3. For the kth iterate of Algorithm 1, we denote the normalized complementarity gap by

$$\mu_k := \frac{\langle X_k, S_k\rangle}{n}.$$

Theorem 2 (Super-linear convergence)**.** *Let $0 < \beta < \frac{1}{4}$, ρ be an admissible reparametrization exponent, and let $p \geq \rho$. Suppose, in addition, that the family of reparametrized weighted*

central paths is uniformly (p,ρ)-predictor regular on $\mathcal{W}_\beta$ (see Definition 7.1). Then the normalized complementarity gaps satisfy

$$\mu_{k+1} = O\Big(\mu_k^{(p+1)/\rho}\Big).$$

Consequently, the sequence $\{\mu_k\}$ converges with Q-order at least

$$\frac{p+1}{\rho} > 1.$$

In particular, Algorithm 1 converges Q-super-linearly in the normalized complementarity gap.

We also prove that the simultaneously vanishing blocks of X_k and S_k, in the sense of the *optimal partition* of SDO (see Definition 7.2), converge super-linearly, with a smaller order of convergence.

Theorem 3. *Under the assumptions of Theorem 2, let $Q := (Q_\mathcal{B}, Q_\mathcal{T}, Q_\mathcal{N})$ be an orthonormal basis partitioned according to $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$. Then the components $Q_\mathcal{N}^T X_k Q_\mathcal{N}$ and $Q_\mathcal{B}^T S_k Q_\mathcal{B}$ converge super-linearly with order $(p+1)/\rho$. Further, if p is large enough such that*

$$p > \rho(2^{n-1}-1) - 1,$$

then the components $Q_\mathcal{T}^T X_k Q_\mathcal{T}$ and $Q_\mathcal{T}^T S_k Q_\mathcal{T}$ converge super-linearly with order

$$(p+1)/(\rho(2^{n-1}-1)).$$

Finally, we quantify the local iteration complexity of Algorithm 1. Once the iterates enter the super-linear regime of Theorem 2, the remaining number of outer iterations required to attain a prescribed accuracy is doubly logarithmic in the target tolerance.

Theorem 4 (Outer iteration complexity)**.** *Let $0 < \beta < \frac{1}{4}$, ρ be an admissible reparametrization exponent, and let $p \geq \rho$. Suppose, in addition, that the family of reparametrized weighted central paths is uniformly (p,ρ)-predictor regular on $\mathcal{W}_\beta$ with constants $\tilde{t} > 0$ and $c > 0$. Assume that $0 < t_0 \leq \tilde{t}$, $\mu_0 = t_0^\rho$, and set*

$$q_0 := \frac{1}{1 + \left(\frac{2\beta}{c\, t_0^d}\right)^{1/\rho}} < 1,$$

where $d := p + 1 - \rho \geq 1$. Then for every $\eta \in (0, \mu_0)$, Algorithm 1 reaches an iterate satisfying $\mu_k \leq \eta$ after at most

$$K_g(\eta) := \left\lceil \frac{\log(\mu_0/\eta)}{\rho \log(1/q_0)} \right\rceil$$

outer iterations. Further, there exists $\mu_{\mathrm{loc}} > 0$ such that, once $\mu_k \leq \mu_{\mathrm{loc}}$, Algorithm 1 computes a ζ-optimal solution in

$$O\left(\log\log\frac{1}{\zeta}\right)$$

additional outer iterations. Consequently, taking $\eta = \mu_{\mathrm{loc}}$, for $0 < \zeta < \mu_{\mathrm{loc}}$, Algorithm 1 computes a ζ-optimal solution in

$$K_g(\mu_{\mathrm{loc}}) + O\left(\log\log\frac{1}{\zeta}\right)$$

outer iterations.

Remark 4 (Higher-order acceleration of a polynomial-time IPM). The higher-order method developed in this paper can also be used as a local acceleration phase for a standard polynomial-time predictor-corrector IPM for SDO. For example, one may first apply a polynomial-time method based on the AHO, or more generally the Monteiro-Zhang [47], search directions until the normalized complementarity gap reaches a threshold $\bar{\mu}$ for which

the local assumptions of Theorem 2 hold, and then switch to Algorithm 1. If the standard phase reaches such a threshold in

$$O\left(\sqrt{n}\log\frac{\mu_0}{\bar{\mu}}\right)$$

outer iterations, then the resulting two-phase method requires

$$O\left(\sqrt{n}\log\frac{\mu_0}{\bar{\mu}}+\log\log\frac{1}{\zeta}\right)$$

outer iterations to attain a ζ-optimal solution. This observation does not require modifying Algorithm 1 itself. The higher-order `IPM` is used only after the iterates have entered its local super-linear regime.

The rest of this paper is organized as follows. In Section 2.3, we briefly review the techniques used in the proofs. In Section 4, we review prior results on the analytic and algebraic properties of (weighted) central paths, super-linear convergence of `IPM`s, and real algebraic geometry methods in optimization. In Section 3, we provide the preliminaries from real algebraic geometry, Puiseux series, and the central path that will be needed for the theoretical results of our higher-order `IPM`. In Section 5, we extend the notion of analytic reparametrization to weighted central paths, and we also prove existence, convergence, analyticity, and semi-algebraicity of weighted central paths. We then prove Theorem 1. In Section 6, we present our higher-order `IPM`, and in Section 7, we provide the proofs of the main theoretical results, i.e., Theorems 2-4. In Section 8, we summarize the elements of our implementation, and we present our numerical experiments. Finally, in Section 9, we conclude the paper with directions for future research.

2.3. **Outline of the proofs.** In this section, we concisely review the ideas behind our proofs.

The proof of Theorem 1 utilizes tools from algorithmic real algebraic geometry [11] and the theory of algebraic curves [78]. More specifically, we apply parameterized bounded algebraic sampling [11, Algorithm 12.18] to obtain a parameterized real univariate representation containing the weighted central solutions. We then use cylindrical algebraic decomposition [11, Theorem 5.6] (see Definition 3.1) to partition the parameter space so that the relevant real roots are delineable and their Thom encodings remain fixed on each cell. After specializing the weight W, the Newton-Puiseux theorem (see Proposition 3.1) shows that the corresponding coordinates admit Puiseux expansions whose ramification indices are bounded by the degree of the univariate representation. This yields only finitely many possible ramification indices and hence a common reparametrization exponent for all $W \in \mathcal{W}_\beta$.

For the proof of Theorem 2, we use the pth-order Taylor approximation of the reparametrized weighted central paths (see (6.4) and (6.5)) together with the uniform (p,ρ)-predictor regularity condition to obtain a residual bound that is uniform over the relevant weights. When ρ is admissible and $p \geq \rho$, this bound yields a predictor parameter satisfying

$$t_+ = O\left(t_k^{(p+1)/\rho}\right),$$

and Lemma 6.3 then yields $\mu_{k+1} = t_+^\rho$, which gives the Q-super-linear convergence asserted in Theorem 2.

For Theorem 3, we apply an error bound result for linear matrix inequalities from [45, Theorem 4.3] to quantify the decay of the vanishing blocks associated with the optimal partition of `SDO`. Combining these estimates with Theorem 2 yields the stated block-wise super-linear convergence rates.

Finally, the proof of Theorem 4 has two parts. First, the quantitative predictor reduction of Proposition 7.1 yields a uniform contraction

$$\mu_{k+1} \leq q_0^\rho \mu_k$$

throughout the predictor regularity region, and hence an explicit bound on the number of outer iterations required to reach any fixed local threshold. Once the iterates enter the super-linear regime, Theorem 2 gives

$$\mu_{k+1} \leq C\mu_k^{(p+1)/\rho}.$$

After rescaling the normalized complementarity gaps to absorb the multiplicative constant, this recurrence yields doubly exponential decay with respect to the local outer iteration counter, and therefore the bound

$$O\left(\log\log\frac{1}{\zeta}\right).$$

3. Background

In this section, we recall some relevant notions from real algebraic geometry, the theory of algebraic curves, and the central path in accordance with [8, 10, 11, 13, 78]. We use standard terminology from field theory and commutative algebra; see, e.g., [38, 6].

Notation 4. In this paper, $\mathbb{R}$ and $\mathbb{C}$ denote the fields of real and complex numbers. Given a field $\mathbb{F}$, the ring of polynomials in $[X_1, \ldots, X_k]$ with coefficients in $\mathbb{F}$ is denoted by $\mathbb{F}[X_1, \ldots, X_k]$. Further, given a finite set $\mathcal{P}$ of polynomials in $\mathbb{F}[X_1, \ldots, X_k]$, the zero set of $\mathcal{P}$ in $\mathbb{F}^k$ is defined as

$$\mathbf{zero}(\mathcal{P}, \mathbb{F}^k) := \left\{ x \in \mathbb{F}^k \mid \bigwedge_{P\in\mathcal{P}} P(x) = 0 \right\}.$$

3.1. Puiseux series.

In algebraic geometry, roots of a polynomial $F(X, Y) \in \mathbb{C}[X, Y]$ in terms of X near $\mathbf{0}$ are expressed as Puiseux series. More specifically, in the proof of Theorem 1, each coordinate of the central path or a weighted central path is described by a Puiseux series.

A *Puiseux series* with coefficients in $\mathbb{C}$ (respectively $\mathbb{R}$) is a Laurent series with fractional exponents, i.e., a series of the form $\sum_{i=r}^{\infty} c_i\varepsilon^{i/q}$ where $c_i \in \mathbb{C}$ (resp. $c_i \in \mathbb{R}$), $i, r \in \mathbb{Z}$, and q is a minimal positive integer, which is called the *ramification index* of the Puiseux series. The set of Puiseux series in ε with coefficients in $\mathbb{C}$ (resp. $\mathbb{R}$) is a field, which we denote by $\mathbb{C}\langle\langle\varepsilon\rangle\rangle$ (resp. $\mathbb{R}\langle\langle\varepsilon\rangle\rangle$). We let $o(\cdot)$ denote the order of a Puiseux series, and it is defined as $o(\sum_{i=r}^{\infty} c_i\varepsilon^{i/q}) = r/q$ if $c_r \neq 0$, see [11, Section 2.6].

If $F(0, 0) = 0$ and $\partial F/\partial Y(0, 0) \neq 0$, then the Implicit Function Theorem [21, Page 19] provides a unique analytic function $Y = \psi(X)$ defined in a neighborhood of $X = 0$, such that $\psi(0) = 0$ and

$$F(X, \psi(X)) = 0.$$

If $\partial F/\partial Y(0, 0) = 0$, the Implicit Function Theorem no longer guarantees such an analytic parametrization, and fractional-power expansions may occur. In general, the Newton-Puiseux theorem [78, Theorem 3.1 of Chapter IV] shows that a Puiseux expansion of $F(X, Y) = 0$ always exists.

Proposition 3.1 (Theorems 3.2 and 4.1 and Section 4.2 of Chapter IV in [78])**.** *Let*

$$F(X, Y) = a_0 + a_1Y + \cdots + a_dY^d \in \mathbb{C}[X][Y],$$

where $a_d \neq 0$. There exist d (not necessarily distinct) Puiseux series $\psi_i(X) \in \mathbb{C}\langle\langle X\rangle\rangle$ for $i = 1, \ldots, d$ such that

$$F(X, Y) = a_d \prod_{i=1}^{d} (Y - \psi_i).$$

A subfield of $\mathbb{R}\langle\langle\varepsilon\rangle\rangle$, denoted by $\mathbb{R}\langle\varepsilon\rangle$, can be identified with the field of germs at 0^+ of continuous semi-algebraic functions [11, Theorem 3.14]. More precisely, the field $\mathbb{R}\langle\varepsilon\rangle$ consists of the elements of $\mathbb{R}\langle\langle\varepsilon\rangle\rangle$ that are algebraic over $\mathbb{R}(\varepsilon)$, the field of rational functions

in the indeterminate ε with coefficients in $\mathbb{R}$ [6, Page 370], and it is called the field of *algebraic Puiseux series* with coefficients in $\mathbb{R}$. It is the real closure [11, Page 43] of the ordered field $\mathbb{R}(\varepsilon)$ in which ε is positive and infinitesimal, i.e.,

$$0 < \varepsilon < r \qquad \text{for every } r \in \mathbb{R}_{>0}.$$

Thus, every element of $\mathbb{R}\langle\varepsilon\rangle$ is represented, for sufficiently small $t_0 > 0$, by a continuous semi-algebraic function

$$(0, t_0) \longrightarrow \mathbb{R}.$$

This interpretation makes $\mathbb{R}\langle\varepsilon\rangle$ particularly useful for studying germs of semi-algebraic functions, including the germs of weighted central paths at $\mu = 0$.

3.2. **Semi-algebraic sets.** Given the real field $\mathbb{R}$, a *quantifier-free formula* $\Phi(X_1, \ldots, X_k)$ with coefficients in $\mathbb{R}$ is a Boolean combination of atoms $P > 0$, $P = 0$, or $P < 0$ where $P \in \mathbb{R}[X_1, \ldots, X_k]$, and $\{X_1, \ldots, X_k\}$ are called the *free variables* of Φ. A quantified formula is given by

$$\Psi = (\mathrm{Q}_1 Y_1) \cdots (\mathrm{Q}_\ell Y_\ell)\ \Phi(X_1, \ldots, X_k, Y_1, \ldots, Y_\ell),$$

in which $\mathrm{Q}_i \in \{\forall, \exists\}$ are quantifiers and Φ is a quantifier-free formula with polynomials in

$$\mathbb{R}[X_1, \ldots, X_k, Y_1, \ldots, Y_\ell].$$

The set of all $(x_1, \ldots, x_k) \in \mathbb{R}^k$ satisfying Ψ is called the $\mathbb{R}$-realization of Ψ. A *semi-algebraic* subset of $\mathbb{R}^k$ is defined as the $\mathbb{R}$-realization of a quantifier-free formula. Alternatively, a semi-algebraic subset of $\mathbb{R}^k$ is a subset of the form

$$\bigcup_{i=1}^{s} \bigcap_{j=1}^{r_i} \left\{ x \in \mathbb{R}^k \mid P_{ij}(x) \leqq 0 \right\},$$

where $P_{ij} \in \mathbb{R}[X_1, \ldots, X_k]$. Given a semi-algebraic set $S \subset \mathbb{R}^k$, a function $f : S \to \mathbb{R}$ is called semi-algebraic if its graph is a semi-algebraic subset of $\mathbb{R}^{k+1}$.

The family of semi-algebraic subsets of $\mathbb{R}^k$ is closed under finite union, finite intersection, and complementation. Further, it is a classical result (see [75]) that every quantified formula is equivalent to a quantifier-free formula (see [11, Theorem 14.16] and [11, Algorithm 14.5]).

3.3. **Sign conditions, univariate representations, and Thom encodings.** In the proof of Theorem 1, we describe weighted central points symbolically, using their *real univariate representations*. We follow the representation used in [11].

Given a finite family $\mathcal{P} \subset \mathbb{R}[X_1, \ldots, X_k]$, a *sign condition* on $\mathcal{P}$ is an element of $\{-1, 0, 1\}^{\mathcal{P}}$, i.e., a mapping $\mathcal{P} \to \{-1, 0, 1\}$. Let $P \in \mathbb{R}[X]$ and

$$\mathrm{Der}(P) := \left\{ P, \frac{dP}{dX}, \frac{d^2P}{dX^2}, \ldots, \frac{d^{(\deg(P))}P}{dX^{(\deg(P))}} \right\}$$

denote a list of polynomials in which d^iP/dX^i for $i > 0$ is the formal ith-order derivative of P and $\deg(P)$ stands for the degree of P. A *Thom encoding* σ of real zero t_σ of P is a sign condition on $\mathrm{Der}(P)$ such that $\sigma(P) = 0$. The following proposition [11, Proposition 2.27] indicates that for any $P \in \mathbb{R}[X]$ and a root $x \in \mathbb{R}$ of P, its Thom encoding characterizes x.

Proposition 3.2 (Thom's Lemma)**.** *Let $P \in \mathbb{R}[X]$ be a univariate polynomial and $\sigma \in \{-1, 0, 1\}^{\mathrm{Der}(P)}$. Then the realization of the sign condition σ is either empty, a point, or an open interval.*

In particular, when $\sigma(P) = 0$, its realization is either empty or a single point.

A k-univariate representation is a $(k+2)$-tuple of polynomials $u = \big(f, g_0, \ldots, g_k\big) \in \mathbb{R}[T]^{k+2}$, where f and g_0 are coprime. A real k-univariate representation of an $x \in \mathbb{R}^k$ is a pair (u, σ)

of a k-univariate representation u and a Thom encoding σ of a real root t_σ of f such that

$$x = \left(\frac{g_1(t_\sigma)}{g_0(t_\sigma)}, \ldots, \frac{g_k(t_\sigma)}{g_0(t_\sigma)}\right) \in \mathbb{R}^k.$$

Given a finite family $\mathcal{P} \subset \mathbb{R}[X_1, \ldots, X_k]$ with $\mathbf{zero}(\mathcal{P}, \mathbb{R}^k)$ being non-empty and bounded, there exists an algorithm [11, Algorithm 12.16] to generate real k-univariate representations describing a set of points that meets every semi-algebraically connected component of $\mathbf{zero}(\mathcal{P}, \mathbb{R}^k)$. A parametric version of this algorithm [11, Algorithm 12.18] will be used in the proof of Theorem 1.

3.4. **Cylindrical algebraic decomposition.** The notion of cylindrical algebraic decomposition [13, 11] plays a key role in semi-algebraic geometry, and will be needed later in Section 5 for the proof of Theorem 1. In what follows, we include the definition from [11, Definition 5.1].

Definition 3.1 (Cell decomposition)**.** Fixing the standard basis of $\mathbb{R}^k$, we identify for each $i, 1 \leq i \leq k$, $\mathbb{R}^i$ with the span of the first i basis vectors. A cylindrical algebraic decomposition (or simply a cell decomposition) of $\mathbb{R}$ is a 1-tuple $(\mathcal{D}_1)$, where $\mathcal{D}_1$ is a finite set of subsets of $\mathbb{R}$, each element being a point or an open interval, which together gives a partition of $\mathbb{R}$. A cell decomposition of $\mathbb{R}^k$ is a k-tuple $(\mathcal{D}_1, \ldots, \mathcal{D}_k)$, where each $\mathcal{D}_i$ is a decomposition of $\mathbb{R}^i$, $(\mathcal{D}_1, \ldots, \mathcal{D}_{k-1})$ is a cell decomposition of $\mathbb{R}^{k-1}$, and $\mathcal{D}_k$ is a finite set of semi-algebraic subsets of $\mathbb{R}^k$ (called the cells of $\mathcal{D}_k$) giving a partition of $\mathbb{R}^k$ consisting of the following: for each $C \in \mathcal{D}_{k-1}$, there is a finite set of semi-algebraic continuous functions $f_{C,1}, \ldots, f_{C,N_C} : C \to \mathbb{R}$ such that $f_{C_1} < \cdots < f_{C,N_C}$, and each element of $\mathcal{D}_k$ is either the graph of a function $f_{C,i}$ or of the form

(a) $\{(x,t) \mid x \in C,\ t < f_{C,1}(x)\}$,
(b) $\{(x,t) \mid x \in C,\ f_{C,i}(x) < t < f_{C,i+1}(x)\}$,
(c) $\{(x,t) \mid x \in C,\ f_{C,N_C}(x) < t\}$,
(d) $\{(x,t) \mid x \in C\}$

(the last case arising is if the set of functions $\{f_{C,i} | 1 \leq i \leq N_C\}$ is empty).

We will say that the cell decomposition $(\mathcal{D}_1, \ldots, \mathcal{D}_k)$ is *adapted* to a semi-algebraic subset D of $\mathbb{R}^k$, if for each $C \in \mathcal{D}_k$, $C \cap D$ is either equal to C or empty. It is well-known that for every semi-algebraic subset D of $\mathbb{R}^k$ there exists a cell decomposition adapted to D [11, Theorem 5.6].

Given a finite family of polynomials $\mathcal{P} \subset \mathbb{R}[X_1, \ldots, X_k]$, a cylindrical algebraic decomposition is said to be *adapted to* $\mathcal{P}$ if every $P \in \mathcal{P}$ has constant sign on each cell of the decomposition. In particular, let $P \in \mathbb{R}[X_1, \ldots, X_k, T]$ and define

$$\mathrm{Der}_T(P) := \left\{P, \frac{\partial P}{\partial T}, \ldots, \frac{\partial^d P}{\partial T^d}\right\}, \qquad d = \deg_T(P).$$

A cylindrical algebraic decomposition adapted to $\mathrm{Der}_T(P)$ can be chosen so that, over each cell of the induced decomposition of the parameter space $\mathbb{R}^k$, the real roots of $P(x,T)$ are delineable. Consequently, the number and ordering of the real roots, as well as their Thom encodings, remain fixed as x varies within a cell; see [11, Chapter 5].

3.5. **Analytic reparametrization of the central path.** We first recall the strict complementarity condition.

Definition 3.2. A primal-dual optimal solution $(X^*, y^*, S^*) \in \mathcal{P}^* \times \mathcal{D}^*$ is called *strictly complementary* if

$$X^* + S^* \succ 0.$$

The strict complementarity condition holds if such an optimal solution exists.

Since a central solution, for every positive μ, is an isolated non-singular solution of (1.1) [14, Theorem 3.3], all higher-order derivatives of the central path exist, by the Implicit Function Theorem (see e.g., [13, Corollary 2.9.8]). More explicitly, the jth-order derivative $(X^{(j)}, y^{(j)}, S^{(j)})$ can be computed by solving

$$\begin{aligned} \langle A_i, X^{(j)} \rangle &= 0, \\ \sum_{i=1}^{m} y_i^{(j)} A_i + S^{(j)} &= 0, \\ X^{(j)}S(\mu) + X(\mu)S^{(j)} + S^{(j)}X(\mu) + S(\mu)X^{(j)} &= \\ \begin{cases} 2I_n & j = 1, \\ -\sum_{k=1}^{j-1} \binom{j}{k} \big(X^{(k)}S^{(j-k)} + S^{(k)}X^{(j-k)}\big), & j > 1. \end{cases} \end{aligned}$$

Under the strict complementarity condition, the central path admits an analytic extension to $\mu = 0$ [23, Theorem 1]. Consequently, its derivatives of every order have finite limits as $\mu \downarrow 0$. Without strict complementarity, this endpoint regularity may fail; in particular, the first-order derivative of the central path can become unbounded as $\mu \downarrow 0$ [20]. The following example, adopted from [10, Example 4], illustrates this phenomenon.

Example 1. Consider the minimization of a linear objective function over the *3-elliptope*, see Figure 1, as follows:

$$\inf_{(x,y,z)} \left\{ 4x - 4y - 2z \mid \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0 \right\},$$

which has a unique non-strictly complementary optimal solution

$$X^* = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \quad y^* = (-4,\ -1,\ -1)^T, \quad S^* = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{pmatrix}.$$

The central path equations can be equivalently written as

$$\underbrace{\begin{pmatrix} 1 & X_{12} & X_{13} \\ X_{12} & 1 & X_{23} \\ X_{13} & X_{23} & 1 \end{pmatrix}}_{X(\mu)} \underbrace{\begin{pmatrix} -y_1 & 2 & -2 \\ 2 & -y_2 & -1 \\ -2 & -1 & -y_3 \end{pmatrix}}_{S(\mu)} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix},$$

which, after eliminating y_1, y_2, and y_3, yields

$$\begin{cases} F(\mu, X_{12}) := 2X_{12}^3 + (2 - \frac{\mu}{2})X_{12}^2 - (\mu + 2)X_{12} - 2 = 0, \\ X_{13} = -X_{12}, \\ X_{23} = -2X_{12}^2 + \frac{\mu}{2}X_{12} + 1. \end{cases} \tag{3.1}$$

By Proposition 3.1, the first equation in (3.1) has a root corresponding to $X_{12}(\mu)$ that can be expressed as the Puiseux series

$$\begin{aligned} X_{12}(\mu) = -1 + \frac{\sqrt{8}}{8}\mu^{\frac{1}{2}} + \frac{1}{32}\mu - \frac{11\sqrt{8}}{2048}\mu^{\frac{3}{2}} - \frac{3}{512}\mu^2 - \frac{121\sqrt{8}}{1048576}\mu^{\frac{5}{2}} \\ + \frac{15}{32768}\mu^3 + \frac{19405\sqrt{8}}{268435456}\mu^{\frac{7}{2}} + \cdots \end{aligned} \tag{3.2}$$

with ramification index 2. All this implies the non-Lipschitzian convergence

$$\|X(\mu) - X^*\| = O(\sqrt{\mu}) \quad \text{and} \quad \|S(\mu) - S^*\| = O(\sqrt{\mu}).$$

It is clear from (3.2) that the first-order derivative of $X_{12}(\mu)$ does not exist at $\mu = 0$. However, the ramification index of $X_{12}(\mu)$ suggests the local parametrization $(\mu^2, \Phi(\mu))$ of the curve $F = 0$ around $\mu = 0$, where

$$\Phi(\mu) = -1 + \frac{\sqrt{8}}{8}\mu + \frac{1}{32}\mu^2 - \frac{11\sqrt{8}}{2048}\mu^3 - \frac{3}{512}\mu^4 - \frac{121\sqrt{8}}{1048576}\mu^5 + \frac{15}{32768}\mu^6 + \ldots,$$

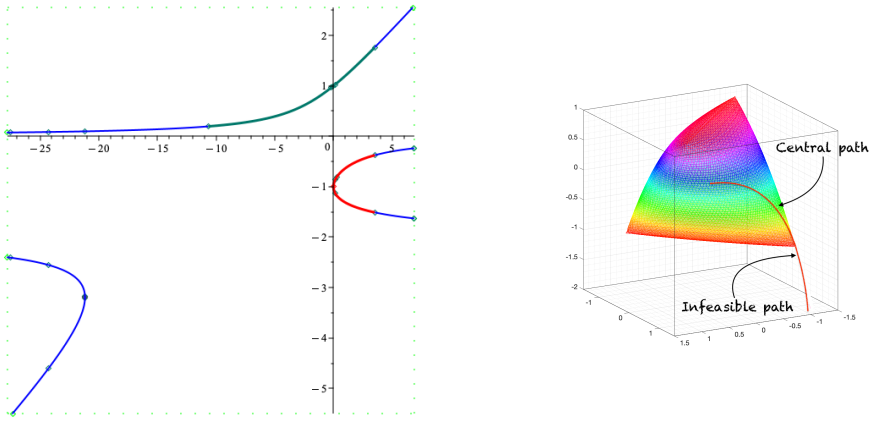


FIGURE 1. The central path is not analytic at $\mu = 0$.

see Figure 1. Thus, a reparametrization $\mu \mapsto \mu^\rho$ recovers the analyticity at $\mu = 0$, where ρ is a positive integer multiple of the ramification index of the Puiseux expansion.

The preceding example illustrates the general mechanism behind analytic reparametrization of the central path. This mechanism follows from the semi-algebraicity of the central path and the Newton-Puiseux theorem, as we briefly summarize below.

Notation 5. We identify a symmetric matrix X by a vector using the linear transformation

$$\operatorname{svec} : \mathbb{S}^n \to \mathbb{R}^{t(n)}, \quad t(n) := \binom{n+1}{2},$$

which multiplies the off-diagonal entries of a symmetric matrix by $\sqrt{2}$, and stacks the upper triangular part into a vector, i.e.,

$$\operatorname{svec}(X) := \big(X_{11}, \sqrt{2}X_{12}, \dots, \sqrt{2}X_{1n}, X_{22}, \sqrt{2}X_{23}, \dots, \sqrt{2}X_{2n}, \dots, X_{nn}\big)^T. \tag{3.3}$$

Notation 6. We use the notation $(\cdot;\cdot;\cdots;\cdot)$ for the concatenation of column vectors. Accordingly, using the identification (3.3), we identify the central path by $v(\mu) := (x(\mu); y(\mu); s(\mu))$. Furthermore, the coordinates of the central path are denoted by $v_i(\mu)$ for $i = 1, \dots, \bar{n}$, where

$$\bar{n} := m + 2t(n).$$

This notation will also be adopted for weighted central solutions.

Notation 7. Using the identification of Notation 6, the indeterminates of the polynomials defining (1.1) are denoted by $V_1, \dots, V_{\bar{n}}$.

3.5.1. *Puiseux reparametrization of the central path.* The central path is a semi-algebraic function (see Proposition 5.2). Hence, for each coordinate $v_i(\mu)$, the germ of v_i at 0^+ defines an element of the field $\mathbb{R}\langle\mu\rangle$ of algebraic Puiseux series. In particular, v_i is algebraic over $\mathbb{R}(\mu)$. Therefore, there exists a non-zero polynomial

$$P_i \in \mathbb{R}[\mu, V_i]$$

such that

$$P_i(\mu, v_i(\mu)) = 0, \qquad i = 1, \dots, \bar{n},$$

for all sufficiently small $\mu > 0$. By Proposition 3.1, the coordinate $v_i(\mu)$ admits a convergent Puiseux expansion of the form

$$v_i(\mu) = \sum_{j=r_i}^{\infty} c_{ij} \mu^{j/q_i},$$

where q_i is its ramification index. Since the central path converges as $\mu \downarrow 0$ (see Proposition 5.2), each $v_i(\mu)$ has a finite limit at $\mu = 0$, and therefore $r_i \geq 0$. Thus,

$$v_i(\mu) = \sum_{j=0}^{\infty} c_{ij} \mu^{j/q_i}.$$

Let q be the least common multiple of the ramification indices $q_1, \ldots, q_{\bar{n}}$. Then, for every positive integer ρ divisible by q, each coordinate

$$v_i(\mu^\rho)$$

is represented by a convergent power series in μ in a neighborhood of $\mu = 0$. Hence the reparametrized central path $v(\mu^\rho)$ admits an analytic extension to $\mu = 0$. In particular, under strict complementarity no ramification is needed: the analyticity of the central path at $\mu = 0$ implies that all coordinate ramification indices are equal to one, and hence $q = 1$.

4. Prior and related work

4.1. **Central path.** The endpoint regularity discussed in Section 3.5 has been studied extensively in the literature. Under strict complementarity, analyticity of the central path of SDO at $\mu = 0$ was established in [20, 23, 66]. This yields, in particular, the Lipschitzian estimates

$$\|X(\mu) - X^{**}\|_F = O(\mu), \qquad \|S(\mu) - S^{**}\|_F = O(\mu),$$

where (X^{**}, y^{**}, S^{**}) denotes the limit point of the central path. Without strict complementarity, substantially more complicated asymptotic behavior may occur, including fractional powers of μ.

The failure of analyticity has been extensively studied for LO and linear complementarity problems (LCP). Reparametrization provides a way to recover analyticity in this degenerate setting. For LCP problems, analytic extension can be recovered by the reparametrization $\mu \mapsto \mu^2$ [70, 71, 72]. This behavior is largely explained by the fact that variables along the central path have magnitudes of order $O(1)$, $O(\mu)$, or $O(\sqrt{\mu})$ [29]. For SDO, such a classification holds only under the strict complementarity condition. In its absence, however, solutions along the central path may exhibit blocks of order $O(\mu^d)$, where the exponent d can be exponentially large (see [45]). In [23], the author mentions that *"... some examples of SDPs show that the variables along the central path for SDP display much more complex behavior than those for LCP. A question now arises as to how the results of this paper may be used to construct higher-order methods for SDP ... the question that still remains open is how to define paths for SDP, so that their analyticity would serve for the construction of algorithms similarly as the analyticity of the weighted paths serves for LCP."*

For general SDO, Basu and Mohammad-Nezhad [10] showed that the central path admits an analytic reparametrization $\mu \mapsto \mu^\rho$ at $\mu = 0$. They characterized the minimal such exponent as the least common multiple of the ramification indices of the Puiseux expansions of the coordinate functions of the central path, and established the bound

$$\rho \leq 2^{O(m^2+n^2m+n^4)}.$$

In particular, the required ramification exponent may grow exponentially with the problem dimension.

4.2. **Weighted central paths and regularity.** Weighted central paths had already played an important role in the analysis of LO and LCP, where their structure is considerably simpler (see e.g., [22, 72, 49]). Their extension to SDO is more subtle because the notion of weighted centrality depends on the chosen symmetrization of the matrix complementarity equation. Instead, weighted paths in SDO are typically defined directly through a symmetrized complementarity equation

$$\Phi(X, S) = \mu W, \qquad W \succ 0,$$

where the choice of the interior-point map Φ determines the corresponding family of weighted paths [48, 52]. In particular, Monteiro and Pang [48] established existence and uniqueness properties for weighted trajectories associated with the AHO map

$$\Phi_{\texttt{AHO}}(X,S) = \frac{XS+SX}{2},$$

and Monteiro and Zanjácomo [52] subsequently developed a general framework for weighted paths associated with a broad class of interior-point maps. Alternative notions of weighted centers for SDO have also been studied; see, e.g., Sturm and Zhang [74].

Remark 5 (Comparison with previous weighted path regularity results). The limiting regularity of weighted central paths has been studied in several settings. Under strict complementarity, Preiß and Stoer [66] proved, for AHO weighted paths of monotone semi-definite LCPs, analytic dependence jointly on the path parameters, including the barrier parameter and the weight matrix. For SDO, Lu and Monteiro [42] established analyticity at $\mu = 0$, in the original barrier parameter, for each fixed $W \succ 0$, together with results on the limiting behavior and derivatives of the corresponding AHO weighted path. Related regularity results for other symmetrizations have also been obtained; see, e.g., [40].

Particularly relevant to Theorem 1 is [42, Proposition 3.5]. Without assuming strict complementarity, Lu and Monteiro state that, for each fixed $W \succ 0$, there exists an analytic scalar reparametrization $\nu = \nu(t)$, with $\nu(0) = 0$, such that the corresponding AHO weighted path

$$t \longmapsto \big(X(\nu(t),W), y(\nu(t),W), S(\nu(t),W)\big)$$

admits an analytic extension to $t = 0$. The reparametrization is thus obtained separately for each fixed W and is not further characterized.

Theorem 1 proceeds in a different direction. Without assuming strict complementarity, we show that the reparametrization can be chosen in the explicit power form $\mu \mapsto \mu^N$, where a single positive integer N works simultaneously for every $W \in \mathcal{W}_\beta$. Thus, in contrast with the fixed-W reparametrization above, the same ramification exponent applies to an entire family of weighted AHO paths. On the other hand, Theorem 1 does not assert joint analyticity in (μ, W) or uniform bounds on the derivatives as W varies. The latter uniformity is imposed separately through the predictor regularity condition in Definition 7.1.

The regularity of weighted central paths is closely connected with the local convergence analysis of primal-dual IPMs; in particular, the role played by the weight and by additional asymptotic centering conditions depends on the chosen symmetrization. We discuss these issues next.

4.3. **Local convergence of primal-dual IPMs.** Classical local convergence analyses of primal-dual IPMs are based on first-order search directions computed in neighborhoods of the central path. Under suitable regularity conditions, super-linear or quadratic convergence has been established for broad classes of primal-dual IPMs, including SDO [3, 31, 32, 35, 43, 61, 63, 64], semi-definite LCPs [35, 36, 41, 68], and LCPs [30, 33, 44, 50, 58, 59, 60, 62, 65, 73].

For SDO, the local convergence analysis is considerably more delicate. Classical super-linear convergence results generally assume strict complementarity and exploit the optimal partition. However, the strict complementarity is not always the only hypothesis. For example, Kojima et al. [35] established super-linear convergence for a predictor-corrector method under strict complementarity, non-degeneracy, and an additional tangential condition requiring the iterates to approach the central path through a shrinking neighborhood. Potra and Sheng [61] subsequently showed that non-degeneracy can be removed, while retaining strict complementarity together with the shrinking-neighborhood condition, and later weakened the latter condition. Luo et al. [43] similarly established super-linear convergence of a symmetric primal-dual path-following method under strict complementarity together with asymptotic centering, and proposed a modification that enforces this behavior. The role of the additional centering conditions depends on the search direction. Kojima et al. [36]

proposed an AHO-based predictor-corrector method and claimed local quadratic convergence under strict complementarity. Lu and Monteiro [41] subsequently pointed out an error in that proof and showed that the original result is valid under an additional non-degeneracy assumption. They also introduced a slight modification of the corrector step and established local quadratic convergence of the resulting AHO-based method under strict complementarity, without imposing a separate tangential-convergence condition.

These additional asymptotic conditions can also be interpreted in terms of weighted central paths, and their role depends on the chosen symmetrization. For the scaling $X^{1/2}SX^{1/2}$, one classical sufficient condition is $W_k \to I$ [35], while weaker formulations (e.g., [61]) require only the appropriate off-diagonal block of W_k to vanish asymptotically [42].

The situation in the present paper is different. Our higher-order predictor is constructed from derivatives of weighted AHO paths. Because the weight varies between iterations, our local analysis requires uniform control of the higher-order predictors; this issue is discussed further in Remark 6.

More recent work continues to rely on strict complementarity, sometimes together with additional assumptions on the problem, initialization, or the behavior of the iterates. For example, implementable polynomial-time methods [69] with super-linear convergence have been obtained under strict complementarity by imposing non-degeneracy or neighborhood-narrowing conditions, or by considering particular homogeneous or feasibility formulations. To the best of our knowledge, however, super-linear convergence of a primal-dual IPM for general SDO has not been established without strict complementarity.

Remark 6. The distinction between strict complementarity and the additional regularity conditions used in super-linear convergence analysis is important for the present paper. Strict complementarity concerns the limiting primal-dual geometry and, in classical SDO analyses, provides strong regularity of the central or weighted central paths. Additional conditions such as non-degeneracy, shrinking neighborhoods, tangential convergence, or asymptotic conditions on the associated weights control how the algorithmic iterates approach these paths.

Our analysis removes the strict complementarity assumption, but replaces the classical endpoint regularity by an analytic reparametrization of weighted central paths. Since our method uses higher-order derivatives of weighted paths whose weights vary with the iterates, the local super-linear convergence result additionally requires a uniform predictor regularity condition. This condition is different from strict complementarity and from the classical tangential conditions, although they play analogous roles in controlling the local behavior of the algorithm.

Remark 7 (Relation to higher-order methods for LCP). Our predictor-corrector construction is inspired by the higher-order IPM framework of Stoer et al. [72] for LCPs. For sufficient LCPs, Stoer and Wechs [70] subsequently established polynomial iteration bounds for these higher-order predictor-corrector methods. Potra and Stoer [65] later developed a related class of super-linearly convergent polynomial-time higher-order IPMs for sufficient LCPs, obtaining arbitrarily high local convergence orders while retaining polynomial iteration complexity.

The semi-definite setting is substantially more delicate: weighted central paths depend on the matrix symmetrization, may fail to be analytic at the limiting point, and can have large ramification indices (see [10] or Section 8.3.2 of this paper).

At the level of the corrector step, our construction is closely related to the AHO centering mechanism of Kojima et al. [36]. In particular, the damping rule in (6.10) is based on the normalized product $\|\Delta X\|_F\|\Delta S\|_F$. However, our use of $\|\Delta X\|_F\|\Delta S\|_F$ is different from [36]: we do not require an asymptotic bound on $\|\Delta X\|_F\|\Delta S\|_F$. Instead, for each fixed centering parameter, continuity and compactness yield the finite bound needed to prove restoration of the inner neighborhood after finitely many damped corrector steps; see Lemma 6.4.

4.4. **Algebraic methods and IPMs.** Algebraic and semi-algebraic methods have been used both to study the structure of central paths and to compute exact solutions of SDO. This perspective is particularly useful because, even for rational input data, an exact optimal solution need not have rational coordinates (see e.g., [14, Page 17]). Thus, exact computation in this setting generally means representing the relevant coordinates as real algebraic numbers, for example through univariate representations and Thom encodings (see Section 3.3), rather than computing finite-precision approximations.

The exact computational complexity of SDO is considerably more subtle than its convexity might suggest. In the real-number model of computation [12], the SDO feasibility problem belongs to NP $\cap$ coNP; see, e.g., [7, 67]. In the classical bit model [19], even membership of general SDO feasibility in NP is not known [67]. Algorithmic real algebraic geometry nevertheless provides finite exact procedures for describing solutions of polynomial systems arising from SDO (see [11, Algorithm 14.9] or [8, Section 3.1]). Such procedures have singly exponential complexity in the relevant dimensions and degrees.

Algebraic geometry has also been used to quantify geometric and algebraic properties of central paths [4, 16, 17, 24, 25, 28, 34]. Complementary symbolic and numerical algebraic geometry approaches have been developed for computing solutions of SDO independently of IPMs [11, 25, 26, 27]. For example, Hauenstein et al. [25] interpret the central path as a bilinear homotopy and use analytic reparametrization to improve numerical endgames. Algebraic techniques have further provided quantitative bounds on the degree, ramification, and convergence rate of the central path of SDO

Proposition 4.1 (Theorem 1.1 in [8])**.** *Let (X^{**}, y^{**}, S^{**}) denote the limit point of the central path. Then, for all sufficiently small $\mu > 0$,*

$$\|X(\mu) - X^{**}\|_F = O\left(\mu^{1/\gamma}\right), \qquad \|S(\mu) - S^{**}\|_F = O\left(\mu^{1/\gamma}\right),$$

where

$$\gamma = 2^{O(m+n^2)}.$$

These results quantify aspects of the algebraic and geometric complexity of SDO, but such bounds do not by themselves translate into convergence order guarantees for IPMs. In particular, bounds on algebraic degree [8], error bound exponents [9], ramification indices [10], or the convergence rate of the central path [8] do not directly specify how an IPM should exploit this structure.

In contrast with these largely descriptive uses of algebraic geometry, the present work uses the ramification structure of weighted AHO paths directly in the construction and convergence analysis of a higher-order predictor-corrector IPM.

5. Analytic reparametrization of weighted central paths

The existence of a weighted central path for any $W \succ 0$ is well-known. For the sake of completeness, we include the proof here.

Notation 8. For ease of exposition, we define and use the following linear map:

$$\mathcal{A}(X) := \begin{bmatrix} \langle A_1, X\rangle \\ \vdots \\ \langle A_m, X\rangle \end{bmatrix}.$$

Proposition 5.1 (Existence and analyticity of weighted AHO central paths)**.** *For every $\mu > 0$ and every $W \in \mathbb{S}^n_{++}$, there exists a unique strictly primal-dual feasible solution*

$$(X(\mu, W), y(\mu, W), S(\mu, W))$$

satisfying (2.1). Further, the mapping $\xi(\mu, W)$ is well-defined and jointly analytic on $(0, \infty) \times \mathbb{S}^n_{++}$.

Proof. The proof follows from [48, Theorem 1(a)]. We define the affine map

$$F : \mathbb{S}^n \times \mathbb{S}^n \times \mathbb{R}^m \longrightarrow \mathbb{S}^n \times \mathbb{R}^m$$

by

$$F(X, S, y) := \big(\mathcal{A}^*(y) + S - C,\ \mathcal{A}(X) - b\big).$$

Since $A_1, \ldots, A_m$ are linearly independent, $\mathcal{A}^*$ is injective, and hence F is *y-injective* according to [48]. Furthermore, if $F(X, S, y) = F(\tilde{X}, \tilde{S}, \tilde{y})$, then

$$\mathcal{A}(X - \tilde{X}) = 0, \qquad S - \tilde{S} = -\mathcal{A}^*(y - \tilde{y}),$$

and therefore

$$\langle X - \tilde{X}, S - \tilde{S}\rangle = -\langle \mathcal{A}(X - \tilde{X}), y - \tilde{y}\rangle = 0.$$

Thus, F is (X, S)*-equilevel-monotone*. Since F is affine, y-injectivity also implies the *y-boundedness* condition required in [48, Theorem 1(a)]; see also [48, Lemma 3].

Now, consider the `AHO` interior-point map

$$\Phi_{\texttt{AHO}}(X, S) = \frac{XS + SX}{2}$$

and define

$$H := (\Phi_{\texttt{AHO}}, F), \quad \mathcal{U} := \left\{(X, S) \in \mathbb{S}^n_{++} \times \mathbb{S}^n_{++} \;\middle|\; XS + SX \in \mathbb{S}^n_{++}\right\}.$$

By Assumption 1, the central path exists [14, Theorem 3.1]. Hence, for any central point, we have

$$F(X, S, y) = 0, \qquad XS + SX = 2\mu I_n \succ 0,$$

so that $(X, S) \in \mathcal{U}$. All this implies $0 \in F(\mathcal{U} \times \mathbb{R}^m)$. It now follows from [48, Theorem 1(a)] that

$$H : \mathcal{U} \times \mathbb{R}^m \longrightarrow \mathbb{S}^n_{++} \times F(\mathcal{U} \times \mathbb{R}^m)$$

is a homeomorphism. Therefore, for every $\mu > 0$ and $W \in \mathbb{S}^n_{++}$, there exists a unique $(X(\mu, W), S(\mu, W), y(\mu, W))$ satisfying

$$H(X(\mu, W), S(\mu, W), y(\mu, W)) = (\mu W, 0),$$

which proves existence and uniqueness, and continuity of weighted central paths.

It remains to establish joint analyticity. The map H is polynomial, hence real analytic. Furthermore, the proof of [48, Lemma 4] shows that $DH(X, S, y)$ is an isomorphism at every $(X, S, y) \in \mathcal{U} \times \mathbb{R}^m$. Hence, by the analytic inverse function theorem (see e.g., [11, Proposition 3.24]), H^{-1} is real analytic locally throughout its image. Since H is globally one-to-one, these local analytic inverses agree on overlaps and therefore define a real-analytic inverse on

$$\mathbb{S}^n_{++} \times F(\mathcal{U} \times \mathbb{R}^m).$$

Finally, the map $(\mu, W) \longmapsto (\mu W, 0)$ is real analytic on $(0, \infty) \times \mathbb{S}^n_{++}$. Thus, $\xi(\mu, W) = H^{-1}(\mu W, 0)$ is jointly real analytic on $(0, \infty) \times \mathbb{S}^n_{++}$. □

For the purpose of local convergence of `IPM`s in this paper, we restrict weighted central paths to the bounded set

$$\Big\{(X, y, S) \mid \langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m,$$

$$\sum_{i=1}^m y_i A_i + S - C = 0, \quad XS + SX - 2\mu W = 0, \quad X, S \succ 0 \tag{5.1}$$

$$\text{for some } \|W - I_n\|_F \le \beta, \quad \mu \in (0, 1]\Big\},$$

where $\beta < 1$ is a fixed parameter (see (6.6)).

The semi-algebraicity of the weighted central paths together with their local boundedness implies that a weighted central path converges for every $W \succ 0$.

Proposition 5.2. *The mapping*

$$\xi : (0,\infty) \times \mathbb{S}^n_{++} \longrightarrow \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n$$

is jointly semi-algebraic. Moreover, for every fixed $W \succ 0$, the weighted central path $\xi_W(\mu)$ converges as $\mu \downarrow 0$.

Proof. By Proposition 5.1, for every $\mu > 0$ and $W \succ 0$ there exists a unique weighted central solution. Hence, the graph of ξ is precisely the semi-algebraic set

$$\left\{ (\mu, W, X, y, S) \;\middle|\; \begin{array}{l} \mu > 0, \quad W \succ 0, \quad X, S \succ 0, \\ \mathcal{A}(X) = b, \quad \mathcal{A}^*(y) + S = C, \\ XS + SX = 2\mu W \end{array} \right\}.$$

Therefore, ξ is a semi-algebraic mapping.

Now, we fix $W \succ 0$. We first show that $\xi_W(\mu)$ remains bounded for $0 < \mu \leq 1$. Let $(\bar{X}, \bar{y}, \bar{S})$ be any fixed strictly primal-dual feasible point, whose existence follows from Assumption 1. For every weighted central point $(X(\mu, W), y(\mu, W), S(\mu, W))$, primal-dual feasibility gives

$$\langle \bar{S}, X(\mu, W)\rangle + \langle \bar{X}, S(\mu, W)\rangle = \langle X(\mu, W), S(\mu, W) + \langle \bar{X}, \bar{S}\rangle.$$

Additionally, we have

$$\langle X(\mu, W), S(\mu, W)\rangle = \mu \ \mathrm{Tr}(W).$$

Consequently, we get

$$\langle \bar{S}, X(\mu, W)\rangle + \langle \bar{X}, S(\mu, W)\rangle = \mu \ \mathrm{Tr}(W) + \langle \bar{X}, \bar{S}\rangle,$$

which is uniformly bounded for $0 < \mu \leq 1$. Since $\bar{X}, \bar{S} \succ 0$ and $X(\mu, W), S(\mu, W) \succ 0$, it follows that $X(\mu, W)$ and $S(\mu, W)$ are uniformly bounded. Further, the linear independence of $A_1, \ldots, A_m$ implies that $y(\mu, W)$ is uniformly bounded as well. Thus, $\xi_W : (0, 1] \to \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n$ is a bounded semi-algebraic function, and By [11, Proposition 3.18], each of its coordinate functions has a finite limit as $\mu \downarrow 0$. □

From Proposition 5.2, each coordinate of ξ_W is described by Puiseux series $\sum_{i=r}^{\infty} c_i \mu^{i/q}$, when μ is sufficiently small, where $q \in \mathbb{Z}_{>0}$ is a constant integer, and c_i depends on W. Now, we prove Theorem 1.

Proof of Theorem 1. We show that there are only finitely many ramification indices for weighted central paths restricted to (5.1). The proof of Proposition 5.2 already establishes the uniform boundedness of weighted central paths restricted to (5.1).

We define $Q \in \mathbb{R}[\mu, W, V_1, \ldots, V_{\bar{n}}]$ as follows

$$Q := \|\mathcal{A}(X) - b\|^2 + \|\mathcal{A}^*(y) + S - C\|_F^2 + \|XS + SX - 2\mu W\|_F^2.$$

We choose $\varepsilon > 0$ sufficiently small so that

$$\|(X(\mu, W), y(\mu, W), S(\mu, W))\|_F < \frac{1}{\varepsilon}$$

for every $\mu \in (0, 1]$ and $W \in \mathcal{W}_\beta$, and we introduce an additional variable $V_{\bar{n}+1}$. We also define

$$\tilde{Q} := Q^2 + \left(\varepsilon^2 \sum_{i=1}^{\bar{n}+1} V_i^2 - 1 \right)^2.$$

The projection of

$$\mathbf{zero}\big(\tilde{Q}(\mu, W, \cdot), \mathbb{R}^{\bar{n}+1}\big)$$

onto the first $\bar{n}$ coordinates contains the weighted central point. Applying Parameterized Bounded Algebraic Sampling [11, Algorithm 12.18] to $\tilde{Q}$, with (μ, W) as parameters, produces a finite family

$$\mathcal{U} = \{u_1, \ldots, u_M\}$$

of parameterized univariate representations

$$u_a = (f_a, g_a) = \big(f_a, (g_{a,0}, g_{a,1}, \ldots, g_{a,\bar{n}+1})\big), \qquad a = 1, \ldots, M.$$

For every fixed (μ, W), the corresponding sample points meet every semi-algebraically connected component of

$$\mathbf{zero}\big(\tilde{Q}(\mu, W, \cdot), \mathbb{R}^{\bar{n}+1}\big).$$

Consequently, after projection onto the first $\bar{n}$ coordinates, the weighted central point is represented by at least one element $u_a \in \mathcal{U}$. More precisely, its coordinates have the form

$$\hat{v}_i(\mu, W) = \frac{g_{a,i}(\mu, W, t)}{g_{a,0}(\mu, W, t)}, \qquad i = 1, \ldots, \bar{n},$$

where t is a real root of $f_a(\mu, W, T)$ and $g_{a,0}(\mu, W, t) \neq 0$. Set

$$d := \max_{1 \le a \le M} \deg_T(f_a),$$

and consider the finite family

$$\mathcal{P} := \bigcup_{a=1}^{M} \mathrm{Der}_T(f_a),$$

where

$$\mathrm{Der}_T(f_a) := \left\{ f_a, \frac{\partial f_a}{\partial T}, \ldots, \frac{\partial^d f_a}{\partial^d T} \right\}.$$

Choose a cylindrical algebraic decomposition adapted to $\mathcal{P}$ (see Section 3.4). Over every cell of the induced decomposition of the (μ, W)-parameter space, the real roots of each f_a are delineable and their Thom encodings remain fixed. Refining the decomposition, if necessary, by the finitely many semi-algebraic conditions determining which of the above sample points is the weighted central point, we may assume that the relevant univariate representation and Thom encoding are fixed on each cell. Since we are interested only in sufficiently small positive values of μ, we refine the induced decomposition at $\mu = 0$. This yields a finite semi-algebraic partition

$$\mathcal{W}_\beta = \bigcup_{\ell=1}^{s} Z_\ell$$

with the following property. For each ℓ, there exist an index $a_\ell \in \{1, \ldots, M\}$ and a Thom encoding σ_ℓ such that, for every $W \in Z_\ell$, there exists $\varepsilon_W > 0$ for which

$$v_i(\mu, W) = \frac{g_{a_\ell,i}\big(\mu, W, t_{\sigma_\ell}(\mu, W)\big)}{g_{a_\ell,0}\big(\mu, W, t_{\sigma_\ell}(\mu, W)\big)}, \qquad i = 1, \ldots, \bar{n}, \tag{5.2}$$

for every $0 < \mu < \varepsilon_W$, where $t_{\sigma_\ell}(\mu, W)$ is the real root of $f_{a_\ell}(\mu, W, T)$ determined by σ_ℓ. Now, fix $W \in Z_\ell$. After specializing W, let

$$f_W(\mu, T) := f_{a_\ell}(\mu, W, T), \qquad d(W) := \deg_T(f_W).$$

The specialization associated with the above representation is non-zero, and $d(W) \le d$. Further, we have

$$f_W\big(\mu, t_{\sigma_\ell}(\mu, W)\big) = 0, \qquad f_W(\mu, T) \in \mathbb{R}[\mu, T] \subset \mathbb{C}[\mu, T].$$

By Proposition 3.1, $t_{\sigma_\ell}(\mu, W)$ admits a convergent Puiseux expansion with some ramification index $q(W)$ satisfying

$$q(W) \le d(W) \le d.$$

Notice that each coordinate in (5.2) is a rational function of $t_{\sigma_\ell}(\mu, W)$ with coefficients in $\mathbb{R}(\mu)$. Hence, all coordinates belong to the same finite Puiseux extension, and the ramification index of every coordinate of the weighted central path divides $q(W)$. Consequently, every coordinate ramification index is a positive integer not exceeding d. Let $\bar{q}$ be the least common multiple of $1, 2, \ldots, d$. Then it follows that $\bar{q}$ is a common multiple of the ramification indices of all coordinates of all weighted central paths with $W \in \mathcal{W}_\beta$. By Proposition 5.2, for every fixed W the weighted central path converges as $\mu \downarrow 0$. Hence, its coordinate Puiseux expansions contain no negative powers. Therefore,

$$X(\mu^{\bar{q}}, W), \quad y(\mu^{\bar{q}}, W), \quad S(\mu^{\bar{q}}, W)$$

are analytic at $\mu = 0$ for every $W \in \mathcal{W}_\beta$. All this implies that the set of positive integers q for which $\xi_W(\mu^q)$ is analytic at $\mu = 0$ for every $W \in \mathcal{W}_\beta$ is non-empty. We denote the smallest integer by N, which completes the proof. □

Example 2. Consider again the SDO problem in Example 1. We now introduce the one-parameter family of weights

$$W_\theta := \kappa_\theta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+\theta^2 & \theta \\ 0 & \theta & 1+\theta^2 \end{pmatrix}, \qquad \kappa_\theta := \frac{3}{3+2\theta^2}, \qquad \theta \in \mathbb{R}.$$

The matrix W_θ is positive definite for every $\theta \in \mathbb{R}$. The corresponding weighted central path is determined by

$$XS + SX = 2\mu W_\theta,$$

where

$$X = \begin{pmatrix} 1 & X_{12} & X_{13} \\ X_{12} & 1 & X_{23} \\ X_{13} & X_{23} & 1 \end{pmatrix}, \qquad S = \begin{pmatrix} -y_1 & 2 & -2 \\ 2 & -y_2 & -1 \\ -2 & -1 & -y_3 \end{pmatrix}.$$

The problem and W_θ are invariant under the orthogonal transformation

$$P = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

By uniqueness of the weighted central solution, $X_{13} = -X_{12}$ and $y_2 = y_3$. Set $x := X_{12}$, $z := X_{23}$, and $\nu := \kappa_\theta \mu$. Then

$$y_1 = 4x - \nu, \qquad y_2 = y_3 = 2x - z - (1+\theta^2)\nu,$$

and the remaining equations reduce to

$$\begin{cases} -6x^2 + xz + (\theta^2+2)\nu x + x - 2z + 4 = 0, \\ z^2 - 2xz + (1+\theta^2)\nu z - \theta\nu - 2x - 1 = 0. \end{cases}$$

For sufficiently small μ, the first equation gives

$$z = \frac{6x^2 - (\theta^2+2)\nu x - x - 4}{x - 2}.$$

Substitution into the second equation yields $F_\theta(\nu, x) = 0$, where

$$\begin{aligned} F_\theta(\nu, x) = {} & 24x^4 + \big(12 - (4\theta^2+14)\nu\big)x^3 \\ & + \Big((\theta^2+2)\nu^2 - (15\theta^2+\theta+17)\nu - 36\Big)x^2 \\ & + \Big((2\theta^4+6\theta^2+4)\nu^2 + (6\theta^2+4\theta+14)\nu - 12\Big)x \\ & + 12 + (8\theta^2 - 4\theta + 8)\nu. \end{aligned}$$

At $\nu = 0$,

$$F_\theta(0, x) = 12(x-1)(x+1)^2(2x-1),$$

thus the weighted central path is the branch converging to $x = -1$.

The corresponding Puiseux expansion is

$$X_{12}(\mu) = -1 + \sqrt{\frac{\theta^2+\theta+1}{8}}\,\nu^{1/2} + \frac{1+\theta-3\theta^2}{32}\,\nu + \frac{85\theta^4-254\theta^3+31\theta^2-294\theta-99}{4608\sqrt{2(\theta^2+\theta+1)}}\,\nu^{3/2} + O(\nu^2), \tag{5.3}$$

where $\nu = \kappa_\theta \mu$. The remaining coordinates are recovered from

$$X_{13}(\mu) = -X_{12}(\mu),$$
$$X_{23}(\mu) = \frac{6X_{12}(\mu)^2 - (\theta^2+2)\nu X_{12}(\mu) - X_{12}(\mu) - 4}{X_{12}(\mu) - 2},$$

and

$$y_1(\mu) = 4X_{12}(\mu) - \nu, \qquad y_2(\mu) = y_3(\mu) = 2X_{12}(\mu) - X_{23}(\mu) - (1+\theta^2)\nu.$$

Since

$$\theta^2+\theta+1 = \left(\theta+\frac{1}{2}\right)^2 + \frac{3}{4} > 0 \qquad \forall\,\theta \in \mathbb{R},$$

the coefficient of $\nu^{1/2}$ in (5.3) never vanishes. Therefore,

$$q(\theta) = 2, \qquad \forall\,\theta \in \mathbb{R}.$$

Consequently, for every interval $I \subseteq \mathbb{R}$, bounded or unbounded, the minimal common ramification index of the family $\{W_\theta : \theta \in I\}$ is

$$N_I = 2.$$

For $\theta = 0$, $W_0 = I_3$, and the expansion reduces to the ordinary central path expansion in Example 1.

Remark 8. For the SDO instance in Example 1, varying the positive definite weight cannot produce two different ramification indices on two parameter intervals. In fact, if $W = (w_{ij}) \succ 0$ is arbitrary, the branch converging to the non-strictly complementary optimal solution satisfies

$$X_{12}(\mu) = -1 + \frac{1}{4}\sqrt{w_{22}+2w_{23}+w_{33}}\,\mu^{1/2} + O(\mu).$$

Since

$$w_{22}+2w_{23}+w_{33} = (0,1,1)W(0,1,1)^T > 0,$$

the square-root term is present for every $W \succ 0$. Thus the ramification index is always equal to 2 for this particular SDO instance.

6. A higher-order IPM

6.1. **AHO search directions.** Several primal-dual search directions have been developed for SDO, including Nesterov-Todd [55, 56], Alizadeh-Haeberly-Overton (AHO) [3], Kojima-Shindoh-Hara [37], or Monteiro-Zhang [53]). We focus on the AHO search direction, which is obtained by applying Newton's method to (1.1). Given an initial solution (X_0, y_0, S_0), this leads to the following linear system of equations

$$\begin{aligned} \langle A_i, \Delta X\rangle &= 0, \quad i = 1,\dots,m, \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S &= 0, \\ \Delta X S_0 + S_0 \Delta X + X_0 \Delta S + \Delta S X_0 &= 2\mu I_n - X_0 S_0 - S_0 X_0. \end{aligned} \tag{6.1}$$

When (X_0, y_0, S_0) with $X_0, S_0 \succ 0$ is sufficiently close to the central path [51, Theorem 2.1], the system (6.1) has a unique solution $(\Delta X, \Delta y, \Delta S)$.

Remark 9 (Semi-algebraic formulations). The centrality condition for the central path appears in several equivalent forms in the literature, for example,

$$XS = \mu I_n, \qquad X^{\frac{1}{2}} S X^{\frac{1}{2}} = \mu I_n, \qquad S^{\frac{1}{2}} X S^{\frac{1}{2}} = \mu I_n,$$

or, more generally, through symmetrized formulations of the form

$$PXSP^{-1} + P^{-T} S X P^T = 2\mu I_n,$$

where P is non-singular; see, e.g., [53]. If the positive-definiteness constraints $X, S \succ 0$ are omitted, then, for a fixed $\mu > 0$, these equations generally define different zero sets. In contrast, when restricted to $X, S \succ 0$, the corresponding centrality systems are equivalent and determine the same unique central solution for each $\mu > 0$. Consequently, they give rise to the same semi-algebraic central path function.

Analogously, AHO-type search directions can be obtained by the application of the Newton's method to the parameterized family of weighted central paths (2.1). The following proposition shows that the existence of such AHO-type directions still follows from [51, Theorem 2.1], when (X_0, y_0, S_0) with $X_0, S_0 \succ 0$ is sufficiently close to the central path.

Notation 9. We define the linear operator

$$\mathcal{J}_{X,S} : \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n \longrightarrow \mathbb{R}^m \times \mathbb{S}^n \times \mathbb{S}^n$$

by

$$\mathcal{J}_{X,S}(\Delta X, \Delta y, \Delta S) := \left(\mathcal{A}(\Delta X),\ \mathcal{A}^*(\Delta y) + \Delta S,\ \Delta X S + S \Delta X + X \Delta S + \Delta S X\right),$$

where

$$\mathcal{A}^*(\Delta y) := \sum_{i=1}^m \Delta y_i A_i$$

is the adjoint of the linear map $\mathcal{A}$.

Proposition 6.1 (Well-definedness of weighted AHO systems)**.** *Let $\mu > 0$, and let (X, y, S) be a strictly primal-dual feasible point with $X, S \succ 0$. Suppose that*

$$\left\| \frac{XS + SX}{2\mu} - I_n \right\|_F < \frac{1}{2}. \tag{6.2}$$

Then $\mathcal{J}_{X,S}$ is non-singular. Consequently, for every $H \in \mathbb{S}^n$, the linear system

$$\begin{aligned} \langle A_i, \Delta X \rangle &= 0, \quad i = 1, \ldots, m, \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S &= 0, \\ \Delta X S + S \Delta X + X \Delta S + \Delta S X &= H \end{aligned}$$

has a unique solution $(\Delta X, \Delta y, \Delta S)$. In particular, for every $W \succ 0$, the weighted AHO Newton system

$$\begin{aligned} \langle A_i, \Delta X \rangle &= 0, \quad i = 1, \ldots, m, \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S &= 0, \\ \Delta X S + S \Delta X + X \Delta S + \Delta S X &= 2\mu W - XS - SX \end{aligned}$$

has a unique solution whenever the neighborhood condition (6.2) holds.

Proof. Using the identity

$$\left\| \frac{XS + SX}{2\mu} - I_n \right\|_F^2 = \left\| \frac{S^{1/2} X S^{1/2}}{\mu} - I_n \right\|_F^2 + \frac{1}{4\mu^2} \|XS - SX\|_F^2,$$

we obtain

$$\left\| \frac{S^{1/2} X S^{1/2}}{\mu} - I_n \right\|_F \leq \left\| \frac{XS + SX}{2\mu} - I_n \right\|_F < \frac{1}{2}.$$

Hence, the sufficient condition of [51, Theorem 2.1] for existence and uniqueness of the AHO direction is satisfied. All this means that the coefficient matrix representing $\mathcal{J}_{X,S}$ is non-singular. Thus, for every $H \in \mathbb{S}^n$, the system

$$\mathcal{J}_{X,S}(\Delta X, \Delta y, \Delta S) = (0, 0, H)$$

has a unique solution. Taking $H = 2\mu W - XS - SX$ gives the weighted AHO Newton direction. □

6.2. **A higher-order IPM based on AHO search directions.** Our construction is inspired by the higher-order predictor-corrector IPM for LCP in [72] and by the AHO centering directions in [36].

Fix a positive integer reparametrization exponent ρ and a Taylor order p. We consider the family of reparametrized weighted AHO central paths

$$\begin{aligned} \langle A_i, X \rangle &= b_i, \quad i = 1, \dots, m, \\ \sum_{i=1}^{m} y_i A_i + S &= C, \\ XS + SX &= 2t^\rho W, \quad X, S \succ 0, \end{aligned} \tag{6.3}$$

where $W \in \mathcal{W}_\beta$. For a fixed $W \succ 0$, we denote the solution of (6.3) by

$$(X(t), y(t), S(t)) := (X(t^\rho, W), y(t^\rho, W), S(t^\rho, W)).$$

Throughout the construction below, we assume that the chosen reparametrization exponent ρ provides the derivatives required through order p at the current iterate. Admissibility of ρ is not required for the construction itself; it enters later in the convergence analysis (see Section 7.1).

Predictor step. The predictor step of our higher-order IPM involves computing the pth-order Taylor approximation of (6.3) at t_0 as follows

$$\begin{aligned} \big(\hat{X}_{t_0}(t), \hat{y}_{t_0}(t), \hat{S}_{t_0}(t)\big) &= \big(X(t_0), y(t_0), S(t_0)\big) \\ &\quad + \sum_{j=1}^{p} \frac{1}{j!} (t - t_0)^j \big(X^{(j)}(t_0), y^{(j)}(t_0), S^{(j)}(t_0)\big), \end{aligned} \tag{6.4}$$

where the higher-order derivatives are obtained by solving

(6.5)

$$\begin{aligned} \langle A_i, X^{(j)} \rangle &= 0, \quad i = 1, \dots, m, \\ \sum_{i=1}^{m} y_i^{(j)} A_i + S^{(j)} &= 0, \\ X^{(j)} S(t) + X(t) S^{(j)} + S^{(j)} X(t) + S(t) X^{(j)} &= \\ \begin{cases} 2\rho(\rho-1)\cdots(\rho-j+1) t^{\rho-j} W + \xi(j) & j \leq \rho, \\ \xi(j), & j > \rho, \end{cases} \end{aligned}$$

in which

$$\xi(j) := -\sum_{k=1}^{j-1} \binom{j}{k} \big(X^{(k)} S^{(j-k)} + S^{(k)} X^{(j-k)}\big).$$

We maintain feasibility and centrality by keeping the iterates in the following neighborhood of the central path:

$$\mathcal{N}_t(\beta) := \Big\{(X,y,S) \mid \vartheta_t(X,S) \le \beta\Big\}, \quad \vartheta_t(X,S) := \left\| \frac{XS+SX}{2t^\rho} - I_n \right\|_F, \qquad \beta < 1. \tag{6.6}$$

Lemma 6.1. *If $\big(\hat{X}_{t_0}(t), \hat{y}_{t_0}(t), \hat{S}_{t_0}(t)\big) \in \mathcal{N}_t(\beta)$ for all $t \in [\bar{t}, t_0]$, then $\big(\hat{X}_{t_0}(t), \hat{y}_{t_0}(t), \hat{S}_{t_0}(t)\big)$ is primal-dual feasible, and we have $\hat{X}_{t_0}(t), \hat{S}_{t_0}(t) \succ 0$ for all $t \in [\bar{t}, t_0]$.*

Proof. Since $\vartheta_t\big(\hat{X}_{t_0}(t), \hat{S}_{t_0}(t)\big) \le \beta < 1$, we have

$$\hat{X}_{t_0}(t)\hat{S}_{t_0}(t) + \hat{S}_{t_0}(t)\hat{X}_{t_0}(t) \succ 0$$

for every $t \in [\bar{t}, t_0]$. We first show that $\hat{X}_{t_0}(t)$ and $\hat{S}_{t_0}(t)$ are non-singular throughout this interval. Indeed, if $\hat{X}_{t_0}(t)v = 0$ for some $v \neq 0$, then

$$v^T \Big(\hat{X}_{t_0}(t)\hat{S}_{t_0}(t) + \hat{S}_{t_0}(t)\hat{X}_{t_0}(t)\Big) v = 0,$$

contradicting the positive definiteness above. Hence $\hat{X}_{t_0}(t)$ is non-singular, and the same argument applies to $\hat{S}_{t_0}(t)$. Since $\hat{X}_{t_0}(t)$ and $\hat{S}_{t_0}(t)$ are continuous symmetric matrix-valued functions of t, and

$$\hat{X}_{t_0}(t_0) = X(t_0) \succ 0, \quad \hat{S}_{t_0}(t_0) = S(t_0) \succ 0,$$

their inertia cannot change on $[\bar{t}, t_0]$ without one of them becoming singular. Therefore,

$$\hat{X}_{t_0}(t) \succ 0, \quad \hat{S}_{t_0}(t) \succ 0, \qquad t \in [\bar{t}, t_0].$$

Finally, primal-dual feasibility follows directly from (6.4)-(6.5). □

To ensure adequate reduction in the duality gap, the iterates from the predictor step stay in the wider neighborhood $\mathcal{N}(2\beta)$, where the next t is obtained from

$$t_+ := \inf_t \Big\{\bar{t} > 0 \mid \vartheta_t(\hat{X}_{t_0}(t), \hat{S}_{t_0}(t)) \le 2\beta, \quad \forall\, \bar{t} \le t \le t_0\Big\}. \tag{6.7}$$

Remark 10. By Proposition 6.1, when $\beta < \frac{1}{4}$, the `AHO` direction is well-defined throughout the larger predictor neighborhood $\mathcal{N}(2\beta)$.

Remark 11 (Numerical computation of t_+). The convergence analysis of our `IPM` uses the exact boundary parameter t_+ defined by (6.7). However, in our numerical implementation, we approximate t_+ by a geometric line search. Starting from t_0, we successively reduce the path parameter by a fixed factor $\alpha_{\text{ls}} \in (0,1)$ and retain the last accepted value before the next trial point leaves $\mathcal{N}(2\beta)$. We denote this numerical approximation by t_+^{ls}. In our experiments, we take $\alpha_{\text{ls}} = 0.9999$ (see Section 8).

Corrector step. The corrector step of our higher-order `IPM` regains centrality (which is classically an important factor for the super-linear convergence of `IPM`s). This step, starting from

$$(X_0^c, y_0^c, S_0^c) := (\hat{X}_{t_0}(t_+), \hat{y}_{t_0}(t_+), \hat{S}_{t_0}(t_+)),$$

performs multiple damped Newton corrections as follows

$$X_{r+1}^c(\alpha) := X_r^c + \alpha\Delta X_r, \quad y_{r+1}^c(\alpha) := y_r^c + \alpha\Delta y_r, \quad S_{r+1}^c(\alpha) := S_r^c + \alpha\Delta S_r, \tag{6.8}$$

where $(\Delta X_r, \Delta y_r, \Delta S_r)$ is the solution of

$$\begin{aligned} \langle A_i, \Delta X_r\rangle &= 0, \quad i = 1,\dots,m, \\ \sum_{i=1}^m (\Delta y_r)_i A_i + \Delta S_r &= 0, \qquad\qquad r = 0,1,2,\dots \\ \Delta X_r S_r^c + S_r^c \Delta X_r + X_r^c \Delta S_r + \Delta S_r X_r^c &= 2t_+^\rho I - X_r^c S_r^c - S_r^c X_r^c, \end{aligned} \tag{6.9}$$

and $X^c_{r+1} := X^c_{r+1}(\alpha_r)$, $y^c_{r+1} := y^c_{r+1}(\alpha_r)$, $S^c_{r+1} := S^c_{r+1}(\alpha_r)$, where

$$\alpha_r = \begin{cases} 1 & \delta_r \le \gamma_r/2 \\ \dfrac{\gamma_r}{2\delta_r} & \delta_r > \gamma_r/2 \end{cases}, \quad \gamma_r := \vartheta_{t_+}(X^c_r, S^c_r), \quad \delta_r := \frac{\|\Delta X_r\|_F \|\Delta S_r\|_F}{t_+^\rho}. \tag{6.10}$$

The corrector steps are continued until $\vartheta_{t_+}(X^c_r, S^c_r) \le \beta$. We call the new corrected point $(X^+, y^+, S^+) := (X^c_r, y^c_r, S^c_r)$. In what follows (Lemmas 6.2-6.4), we show that the iterate (X^+, y^+, S^+) is strictly feasible, and can be obtained using finitely many corrector steps.

The following lemma shows that each corrector step strictly reduces the centrality while keeping the strict feasibility of the iterate.

Lemma 6.2. *Recall $t_+ > 0$ from (6.7). Suppose that, at the rth corrector step, (X^c_r, y^c_r, S^c_r) is strictly primal-dual feasible and*

$$\beta < \gamma_r = \vartheta_{t_+}(X^c_r, S^c_r) \le 2\beta.$$

Let $(\Delta X_r, \Delta y_r, \Delta S_r)$ be the solution of (6.9), and let δ_r and α_r be defined as in (6.10). Then

$$(X^c_{r+1}, y^c_{r+1}, S^c_{r+1}) := \big(X^c_{r+1}(\alpha_r), y^c_{r+1}(\alpha_r), S^c_{r+1}(\alpha_r)\big)$$

is strictly primal-dual feasible and

$$\gamma_{r+1} := \vartheta_{t_+}(X^c_{r+1}, S^c_{r+1}) \le \begin{cases} \delta_r \le \dfrac{\gamma_r}{2}, & \text{if } \delta_r \le \dfrac{\gamma_r}{2}, \\ \gamma_r\left(1 - \dfrac{\gamma_r}{4\delta_r}\right) < \gamma_r, & \text{if } \delta_r > \dfrac{\gamma_r}{2}. \end{cases}$$

Further, we have

$$X^c_{r+1}(\bar\alpha) \succ 0, \qquad S^c_{r+1}(\bar\alpha) \succ 0 \qquad \text{for every } \bar\alpha \in [0, \alpha_r].$$

In particular,

$$\gamma_{r+1} < \gamma_r \le 2\beta,$$

and thus every accepted corrector iterate remains in $\mathcal{N}_{t_+}(2\beta)$.

Proof. Set $\tau := t_+^\rho$ and define

$$E_r := \frac{X^c_r S^c_r + S^c_r X^c_r}{2\tau} - I_n,$$

so that

$$\|E_r\|_F = \gamma_r.$$

For $\bar\alpha \in [0, 1]$, define

$$E_{r+1}(\bar\alpha) := \frac{X^c_{r+1}(\bar\alpha) S^c_{r+1}(\bar\alpha) + S^c_{r+1}(\bar\alpha) X^c_{r+1}(\bar\alpha)}{2\tau} - I_n.$$

Using (6.8) and the third equation of (6.9), we obtain

$$E_{r+1}(\bar\alpha) = (1 - \bar\alpha) E_r + \frac{\bar\alpha^2}{2\tau}\left(\Delta X_r \Delta S_r + \Delta S_r \Delta X_r\right).$$

Therefore,

$$\begin{aligned} \|E_{r+1}(\bar\alpha)\|_F &\le (1 - \bar\alpha)\gamma_r + \frac{\bar\alpha^2}{2\tau}\|\Delta X_r \Delta S_r + \Delta S_r \Delta X_r\|_F \\ &\le (1 - \bar\alpha)\gamma_r + \bar\alpha^2 \delta_r. \end{aligned} \tag{6.11}$$

If $\delta_r \le \gamma_r/2$, then $\alpha_r = 1$, and hence

$$\gamma_{r+1} \le \delta_r \le \frac{\gamma_r}{2}.$$

If $\delta_r > \gamma_r/2$, then

$$\alpha_r = \frac{\gamma_r}{2\delta_r} < 1,$$

and consequently

$$\begin{aligned}\gamma_{r+1} &\le \left(1 - \frac{\gamma_r}{2\delta_r}\right)\gamma_r + \frac{\gamma_r^2}{4\delta_r^2}\delta_r \\ &= \gamma_r\left(1 - \frac{\gamma_r}{4\delta_r}\right) < \gamma_r.\end{aligned}$$

Now, it remains to prove that positive definiteness is preserved along the accepted corrector segment. We define

$$q_r(\bar{\alpha}) := (1-\bar{\alpha})\gamma_r + \bar{\alpha}^2\delta_r.$$

By the definition of α_r,

$$2\bar{\alpha}\delta_r \le \gamma_r \qquad \text{for every } \bar{\alpha} \in [0, \alpha_r].$$

Thus, we have

$$q_r'(\bar{\alpha}) = -\gamma_r + 2\bar{\alpha}\delta_r \le 0,$$

and (6.11) yields

$$\|E_{r+1}(\bar{\alpha})\|_F \le q_r(\bar{\alpha}) \le q_r(0) = \gamma_r \le 2\beta < 1.$$

Therefore, we get

$$I_n + E_{r+1}(\bar{\alpha}) \succ 0,$$

and hence

$$X_{r+1}^c(\bar{\alpha})S_{r+1}^c(\bar{\alpha}) + S_{r+1}^c(\bar{\alpha})X_{r+1}^c(\bar{\alpha}) \succ 0.$$

Since $X_r^c, S_r^c \succ 0$, analogous to the continuity argument of Lemma 6.1, neither $X_{r+1}^c(\bar{\alpha})$ nor $S_{r+1}^c(\bar{\alpha})$ can become singular for $\bar{\alpha} \in [0, \alpha_r]$. Hence

$$X_{r+1}^c(\bar{\alpha}) \succ 0, \qquad S_{r+1}^c(\bar{\alpha}) \succ 0$$

throughout the accepted segment. Finally, the first two equations of (6.9) imply that the primal and dual affine constraints are preserved. Thus, $(X_{r+1}^c, y_{r+1}^c, S_{r+1}^c)$ is strictly primal-dual feasible. □

Next, we show that the damped corrector steps preserve the complementarity gap.

Lemma 6.3. *Let $(X(t), y(t), S(t))$ be the reparametrized weighted central path through the current corrected iterate at t_0, with weight*

$$W := \frac{X(t_0)S(t_0) + S(t_0)X(t_0)}{2t_0^\rho}.$$

Assume that

$$\frac{\langle X(t_0), S(t_0)\rangle}{n} = t_0^\rho$$

and that $p \ge \rho$. Let $(\hat{X}_{t_0}(t), \hat{y}_{t_0}(t), \hat{S}_{t_0}(t))$ be the pth-order Taylor predictor defined in (6.4). Then

$$\frac{\langle \hat{X}_{t_0}(t), \hat{S}_{t_0}(t)\rangle}{n} = t^\rho. \tag{6.12}$$

In particular, for $(X_0^c, y_0^c, S_0^c) = \big(\hat{X}_{t_0}(t_+), \hat{y}_{t_0}(t_+), \hat{S}_{t_0}(t_+)\big)$, we have

$$\frac{\langle X_0^c, S_0^c\rangle}{n} = t_+^\rho.$$

Further, let (X_r^c, y_r^c, S_r^c) *be any corrector iterate generated by* (6.8)-(6.10)*. Then we have*

$$\frac{\langle X_r^c, S_r^c\rangle}{n} = t_+^\rho \qquad \textit{for every corrector iterate } r \geq 0.$$

Proof. By the definition of W,

$$\operatorname{Tr}(W) = \frac{\operatorname{Tr}\big(X(t_0)S(t_0) + S(t_0)X(t_0)\big)}{2t_0^\rho} = \frac{\langle X(t_0), S(t_0)\rangle}{t_0^\rho} = n.$$

Hence, along the weighted central path ξ_W, the centrality condition

$$X(t)S(t) + S(t)X(t) = 2t^\rho W$$

implies

$$\langle X(t), S(t)\rangle = t^\rho \operatorname{Tr}(W) = nt^\rho.$$

The Taylor predictor is primal-dual feasible, since the zeroth-order term is feasible and all higher-order derivatives satisfy the homogeneous primal and dual feasibility equations. Therefore,

$$\langle \hat{X}_{t_0}(t), \hat{S}_{t_0}(t)\rangle = \langle C, \hat{X}_{t_0}(t)\rangle - b^T \hat{y}_{t_0}(t).$$

Since the right-hand side is a linear functional of the predictor variables, it is precisely the pth-order Taylor polynomial at t_0 of

$$\langle C, X(t)\rangle - b^T y(t) = \langle X(t), S(t)\rangle = nt^\rho.$$

Because $p \geq \rho$, the pth-order Taylor polynomial of nt^ρ is the polynomial itself. Thus,

$$\langle \hat{X}_{t_0}(t), \hat{S}_{t_0}(t)\rangle = nt^\rho,$$

which proves (6.12).

We now consider one damped corrector step. Set

$$\tau := t_+^\rho, \qquad \nu_r := \frac{\langle X_r^c, S_r^c\rangle}{n}.$$

Taking the trace of the third equation in (6.9) gives

$$\langle X_r^c, \Delta S_r\rangle + \langle S_r^c, \Delta X_r\rangle = n(\tau - \nu_r).$$

Furthermore, the first two equations in (6.9) imply

$$\mathcal{A}(\Delta X_r) = 0, \qquad \Delta S_r = -\mathcal{A}^*(\Delta y_r),$$

and hence

$$\langle \Delta X_r, \Delta S_r\rangle = -\langle \Delta X_r, \mathcal{A}^*(\Delta y_r)\rangle = -\langle \mathcal{A}(\Delta X_r), \Delta y_r\rangle = 0.$$

Using

$$X_{r+1}^c = X_r^c + \alpha_r \Delta X_r, \qquad S_{r+1}^c = S_r^c + \alpha_r \Delta S_r,$$

we obtain

$$\begin{aligned}\frac{\langle X_{r+1}^c, S_{r+1}^c\rangle}{n} &= \nu_r + \frac{\alpha_r}{n}\left(\langle X_r^c, \Delta S_r\rangle + \langle S_r^c, \Delta X_r\rangle\right) + \frac{\alpha_r^2}{n}\langle \Delta X_r, \Delta S_r\rangle \\ &= \nu_r + \alpha_r(\tau - \nu_r) \\ &= (1-\alpha_r)\nu_r + \alpha_r \tau.\end{aligned}$$

Therefore, if $\nu_r = \tau$, then

$$\frac{\langle X_{r+1}^c, S_{r+1}^c\rangle}{n} = \tau = t_+^\rho.$$

Since the predictor point satisfies

$$\frac{\langle X_0^c, S_0^c\rangle}{n} = t_+^\rho,$$

the conclusion follows by induction on r. □

Now, we show that only finitely many damped corrector steps are needed to ensure the desired centrality tolerance.

Lemma 6.4 (Finite restoration of centrality)**.** *Recall $t_+ > 0$ from (6.7), assume that $p \geq \rho$, and let*

$$(X_0^c, y_0^c, S_0^c) = \big(\hat{X}_{t_0}(t_+), \hat{y}_{t_0}(t_+), \hat{S}_{t_0}(t_+)\big)$$

be the predictor point, where $\vartheta_{t_+}(X_0^c, S_0^c) \leq 2\beta$, and $\beta < \frac{1}{4}$. Then the damped corrector iteration (6.8)-(6.10) reaches $\mathcal{N}_{t_+}(\beta)$ after finitely many inner iterations, i.e., there exists $r_* < \infty$ such that

$$\gamma_{r_*} = \vartheta_{t_+}(X_{r^*}^c, S_{r^*}^c) \leq \beta.$$

Proof. Set $\tau := t_+^\rho$. From Lemma 6.3 we have

$$\frac{\langle X_r^c, S_r^c \rangle}{n} = \tau \qquad \text{for every corrector iterate } r.$$

Further, Lemma 6.2 implies that every accepted corrector iterate is strictly primal-dual feasible, remains in $\mathcal{N}_{t_+}(2\beta)$, and satisfies $\gamma_{r+1} < \gamma_r$ whenever $\gamma_r > \beta$. Now, it remains to show that the strict decrease cannot continue indefinitely while $\gamma_r > \beta$. To this end, consider

$$\mathcal{K}_\tau := \left\{ (X, y, S) \;\middle|\; \begin{array}{l} \langle A_i, X \rangle = b_i, \quad i = 1, \ldots, m, \\ \sum_{i=1}^m y_i A_i + S = C, \\ X, S \succ 0, \quad \langle X, S \rangle = n\tau, \\ \beta \leq \vartheta_{t_+}(X, S) \leq 2\beta \end{array} \right\}.$$

We first show that $\mathcal{K}_\tau$ is compact. Let $(\bar{X}, \bar{y}, \bar{S})$ be any fixed strictly primal-dual feasible point (see Assumption 1). For every primal-dual feasible $(X, y, S) \in \kappa_\tau$, we have

$$\langle \bar{S}, X \rangle + \langle \bar{X}, S \rangle = \langle X, S \rangle + \langle \bar{X}, \bar{S} \rangle = n\tau + \langle \bar{X}, \bar{S} \rangle,$$

which implies

$$\lambda_{\min}(\bar{S}) \operatorname{Tr}(X) + \lambda_{\min}(\bar{X}) \operatorname{Tr}(S) \leq n\tau + \langle \bar{X}, \bar{S} \rangle.$$

Since $\bar{X}, \bar{S} \succ 0$ and $X, S \succeq 0$, this yields uniform Frobenius norm bounds on X and S. Furthermore, since $A_1, \ldots, A_m$ are linearly independent, the boundedness of S implies the boundedness of y.

To prove closedness of $\mathcal{K}_\tau$, let a sequence in $\mathcal{K}_\tau$ converge to (X, y, S). The affine feasibility equations and $\langle X, S \rangle = n\tau$ pass to the limit, and $X, S \succeq 0$. Furthermore, $\beta \leq \vartheta_{t_+}(X, S) \leq 2\beta < \frac{1}{2}$, and hence $XS + SX \succ 0$. This also implies $X, S \succ 0$, because otherwise there would exist $v \neq 0$ with $Xv = 0$, which would give $v^T(XS + SX)v = 0$. Hence $(X, y, S) \in \mathcal{K}_\tau$, and thus $\mathcal{K}_\tau$ is compact.

By Proposition 6.1, for each $(X, y, S) \in \mathcal{K}_\tau$, the `AHO` centering system (6.9) is non-singular, because $\mathcal{K}_\tau \subseteq \mathcal{N}_{t_+}(2\beta)$. Therefore, by the Implicit Function Theorem, its solution $(\Delta X, \Delta y, \Delta S)$ depends continuously on (X, y, S), that also implies

$$\delta(X, y, S) := \frac{\|\Delta X\|_F \|\Delta S\|_F}{\tau}$$

is continuous on $\mathcal{K}_\tau$. Therefore, there exists a finite $D_\tau \geq 0$ such that

$$\delta(X, y, S) \leq D_\tau \qquad \forall (X, y, S) \in \mathcal{K}_\tau.$$

If $D_\tau = 0$, then Lemma 6.2 gives $\gamma_{r+1} = 0$ whenever $\gamma_r > \beta$, and the corrector phase terminates. Now, suppose that $D_\tau > 0$. As long as $\gamma_r > \beta$, Lemma 6.2 yields either

$$\gamma_{r+1} \leq \frac{\gamma_r}{2},$$

or

$$\gamma_{r+1} \leq \gamma_r \left(1 - \frac{\gamma_r}{4\delta_r}\right) \leq \gamma_r \left(1 - \frac{\beta}{4D_\tau}\right),$$

which can be written as $\gamma_{r+1} \leq q_\tau \gamma_r$, with

$$q_\tau := \max\left\{\frac{1}{2}, 1 - \frac{\beta}{4D_\tau}\right\} < 1.$$

Since $\gamma_0 \leq 2\beta$, it follows that $\gamma_r \leq 2\beta q_\tau^r$, and thus there exists a finite r_* such that $2q_\tau^{r_*} \leq 1$, and then $\gamma_{r_*} \leq \beta$. This completes the proof. □

As a result of Lemmas 6.1-6.4, the corrected point

$$(X^+, y^+, S^+) := (X^c_{r_*}, y^c_{r_*}, S^c_{r_*})$$

is strictly primal-dual feasible, and satisfies

$$(X^+, y^+, S^+) \in \mathcal{N}_{t_+}(\beta), \quad \frac{\langle X^+, S^+\rangle}{n} = t_+^\rho.$$

Further, the new iterate (X^+, y^+, S^+) also leads to a positive definite weight matrix

$$W_+ := \left(X^+S^+ + S^+X^+\right)/2t_+^\rho,$$

which defines a new weighted central path passing through the current solution.

Numerical algorithm. Algorithm 1 is a numerical realization, in which the higher-order directions and the corrector systems are computed using numerical linear algebra. For the theoretical analysis, t_+ denotes the exact boundary parameter defined by (6.7); in the numerical implementation of Section 8, it is approximated by the geometric line search described in Remark 11.

Algorithm 1 Higher-order `IPM` for `SDO`

Input: The degree ρ of reparametrization, the order p of approximation, the neighborhood parameter $0 < \beta < \frac{1}{4}$, the target tolerance $\zeta > 0$, an initial solution (X_0, y_0, S_0) with $t_0^\rho = \langle X_0, S_0\rangle/n$ such that $\vartheta_{t_0}(X_0, S_0) \leq \beta$.
Initialize: Set $k = 0$
while $\frac{\langle X_k, S_k\rangle}{n} > \zeta$ **do**
  $t_k \leftarrow \left(\frac{\langle X_k, S_k\rangle}{n}\right)^{1/\rho}$
  $W_k \leftarrow \frac{X_kS_k + S_kX_k}{2t_k^\rho}$
  for $j = 1, \ldots, p$ **do**
    $(X^{(j)}, y^{(j)}, S^{(j)}) \leftarrow$ Solve (6.5) with $t = t_k$
  end for
  $t_+ \leftarrow$ Determine the boundary parameter from (6.7)
  Predictor: $\left(\hat{X}_{t_k}(t_+), \hat{y}_{t_k}(t_+), \hat{S}_{t_k}(t_+)\right) \leftarrow$ Compute (6.4) at $t = t_+$
  $(X_0^c, y_0^c, S_0^c) \leftarrow \left(\hat{X}_{t_k}(t_+), \hat{y}_{t_k}(t_+), \hat{S}_{t_k}(t_+)\right)$
  Set $r = 0$
  while $\vartheta_{t_+}(X_r^c, S_r^c) > \beta$ **do**
    Corrector: $(X_{r+1}^c, y_{r+1}^c, S_{r+1}^c) \leftarrow$ Solve (6.9) starting from (X_r^c, y_r^c, S_r^c) at $t = t_+$ and then compute (6.8) using (6.10)
    $r \leftarrow r + 1$
  end while
  $(X_{k+1}, y_{k+1}, S_{k+1}) \leftarrow (X_r^c, y_r^c, S_r^c)$
  $k \leftarrow k + 1$
end while
Return: (X_k, y_k, S_k)

In the following sections, we analyze the local convergence and iteration complexity of Algorithm 1.

7. Proof of the main results

In this section, we prove Theorems 2-4, i.e., the local convergence and the iteration complexity of Algorithm 1.

7.1. **Local convergence.**

The higher-order predictor is constructed from weighted central paths whose weight varies with the iterates. Accordingly, point-wise differentiability of each weighted path is not sufficient for the convergence analysis; a uniform control of the predictor residual along the relevant family of weights is required. We formalize this property next.

Definition 7.1 (Uniform predictor regularity)**.** Let $\rho \in \mathbb{Z}_{>0}$ and $p \geq 1$, and let $\mathcal{W} \subseteq \mathbb{S}^n_{++}$. We say that the family of reparametrized weighted central paths is *uniformly* (p,ρ)*-predictor regular on* $\mathcal{W}$ if there exist constants $\tilde{t} > 0$ and $c > 0$ such that, for every $W \in \mathcal{W}$, every $0 < t_0 \leq \tilde{t}$, and every $0 < t \leq t_0$, the pth-order Taylor predictor satisfies

$$\left\| \hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) - 2t^{\rho}W \right\|_F \leq c|t - t_0|^{p+1}.$$

We now prove the super-linear convergence by deriving bounds on a feasible solution of (6.7).

Proof of Theorem 2. Let

$$h := t - t_0, \qquad X^{(0)} := X(t_0), \qquad S^{(0)} := S(t_0).$$

By the Taylor expansion (6.4), we have

$$\hat{X}(t) = \sum_{j=0}^{p} \frac{h^j}{j!} X^{(j)}(t_0), \qquad \hat{S}(t) = \sum_{j=0}^{p} \frac{h^j}{j!} S^{(j)}(t_0).$$

Therefore,

$$\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) = \sum_{r=0}^{2p} h^r \sum_{\ell=\max\{0,r-p\}}^{\min\{p,r\}} \frac{X^{(\ell)}S^{(r-\ell)} + S^{(\ell)}X^{(r-\ell)}}{\ell!(r-\ell)!}, \tag{7.1}$$

where all the derivatives on the right-hand side are evaluated at t_0. For every $0 \leq r \leq p$, the coefficient of h^r in (7.1) is

$$\frac{1}{r!} \sum_{\ell=0}^{r} \binom{r}{\ell} \left(X^{(\ell)}S^{(r-\ell)} + S^{(\ell)}X^{(r-\ell)} \right).$$

By (6.5), this is precisely

$$\frac{1}{r!} \frac{d^r}{dt^r} \left(2t^{\rho}W\right) \Big|_{t=t_0}.$$

Hence, when $p \geq \rho$,

$$\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) = 2t^{\rho}W + R_p(t), \tag{7.2}$$

where

$$R_p(t) := \sum_{r=p+1}^{2p} h^r \sum_{\ell=r-p}^{p} \frac{X^{(\ell)}S^{(r-\ell)} + S^{(\ell)}X^{(r-\ell)}}{\ell!(r-\ell)!}.$$

Here we used the fact that, since $p \geq \rho$, the pth-order Taylor expansion of the polynomial $2t^{\rho}W$ about t_0 is exact.

Now, we restrict the reparametrized weighted central paths to $(0, \tilde{t})$ for some $\tilde{t} > 0$. By the assumption, there exist a constant $c > 0$, independent of t and t_0, such that

$$\|R_p(t)\| \leq c|t - t_0|^{p+1}, \qquad 0 < t \leq t_0 \leq \tilde{t}.$$

It follows from (7.2) that

$$\left\| \frac{\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t)}{2t^\rho} - W \right\| \leq \frac{c|t - t_0|^{p+1}}{2t^\rho}.$$

Further,

$$\begin{aligned} \left\| \frac{\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t)}{2t^\rho} - I_n \right\| &\leq \left\| \frac{\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t)}{2t^\rho} - W \right\| + \|W - I_n\| \\ &\leq \left\| \frac{\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t)}{2t^\rho} - W \right\| + \beta, \end{aligned} \tag{7.3}$$

where the last inequality follows from $(X(t_0), y(t_0), S(t_0)) \in \mathcal{N}(\beta)$. Since the right-hand side of (7.3) is a monotonically decreasing continuous function of t over $(0, t_0]$ taking values in $[0, \infty)$, by the Intermediate Value Theorem there exists a $\bar{t} < t_0$ such that

$$\begin{aligned} &\frac{(t_0 - \bar{t})^{p+1}}{2\bar{t}^\rho} c = \beta, \\ &\frac{(t_0 - t)^{p+1}}{2t^\rho} c \leq \beta, \quad \forall t \in [\bar{t}, t_0]. \end{aligned}$$

Then it follows from (6.7) and (7.3) that $\bar{t}$ is a feasible solution for (6.7), and thus $t_+ \leq \bar{t}$. Moreover, from

$$\frac{c(t_0 - \bar{t})^{p+1}}{2\bar{t}^\rho} = \beta$$

we obtain

$$\frac{2\beta\bar{t}^\rho}{ct_0^{p+1}} = \left(\frac{t_0 - \bar{t}}{t_0} \right)^{p+1} \leq 1,$$

and therefore

$$t_+ \leq \bar{t} \leq \left(\frac{c}{2\beta} \right)^{1/\rho} t_0^{(p+1)/\rho}.$$

Since $p \geq \rho$, we have $(p+1)/\rho > 1$. In particular, for

$$t_0 \leq \left(\frac{2\beta}{c} \right)^{1/(p+1-\rho)},$$

the predictor step yields a strict reduction in t, and

$$t_+ = O\left(t_0^{(p+1)/\rho} \right).$$

By definition, $\mu_k = t_0^\rho$ and by Lemma 6.3, $\mu_{k+1} = t_+^\rho$. Hence

$$\mu_{k+1} = O\left((\mu_k)^{(p+1)/\rho} \right),$$

which proves Q-super-linear convergence of the complementarity gaps with order

$$\gamma = \frac{p+1}{\rho}.$$

□

The predictor regularity condition in Theorem 2 is formulated directly in terms of the Taylor residuals of the weighted central paths. We next record several sufficient conditions under which this assumption is automatically satisfied. We begin with a standard joint-smoothness condition, and then specialize it to the strictly complementary setting and to semi-algebraic families of weighted paths.

Corollary 1 (Joint smoothness)**.** *Suppose that $p \geq \rho$, and for some $\tilde{t} > 0$, the mapping*

$$(t, W) \longmapsto \xi_W(t^\rho)$$

extends to a $\mathcal{C}^p$-smooth mapping on $[0, \tilde{t}] \times \mathcal{W}_\beta$, where $\mathcal{W}_\beta$ is compact. Then the family is uniformly (p, ρ)-predictor regular on $\mathcal{W}_\beta$.

Proof. Since $\mathcal{W}_\beta$ is compact and the reparametrized weighted central path is $\mathcal{C}^p$-smooth on $[0, \tilde{t}] \times \mathcal{W}_\beta$, there exists a constant $M > 0$ such that

$$\max_{0 \leq j \leq p} \left\{ \|X^{(j)}(t, W)\|, \|S^{(j)}(t, W)\| \right\} \leq M$$

for all $(t, W) \in [0, \tilde{t}] \times \mathcal{W}_\beta$. Let $h := t - t_0$. By definition,

$$\hat{X}(t) = \sum_{j=0}^{p} \frac{h^j}{j!} X^{(j)}(t_0, W), \qquad \hat{S}(t) = \sum_{j=0}^{p} \frac{h^j}{j!} S^{(j)}(t_0, W).$$

Therefore,

$$\begin{aligned} &\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) \\ &\quad = \sum_{r=0}^{2p} h^r \sum_{\substack{0 \leq j \leq p \\ 0 \leq r-j \leq p}} \frac{X^{(j)}(t_0, W)S^{(r-j)}(t_0, W) + S^{(j)}(t_0, W)X^{(r-j)}(t_0, W)}{j!(r-j)!}. \end{aligned}$$

For $0 \leq r \leq p$, the inner sum is precisely

$$\frac{1}{r!} \frac{d^r}{dt^r} \left(X(t, W)S(t, W) + S(t, W)X(t, W) \right) \Big|_{t=t_0}.$$

Since the weighted central path satisfies

$$X(t, W)S(t, W) + S(t, W)X(t, W) = 2t^\rho W,$$

we obtain

$$\sum_{r=0}^{p} \frac{h^r}{r!} \frac{d^r}{dt^r} \left(2t^\rho W \right) \Big|_{t=t_0} = 2t^\rho W,$$

where the last equality follows from $p \geq \rho$. Consequently,

$$\begin{aligned} &\hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) - 2t^\rho W \\ &\quad = \sum_{r=p+1}^{2p} h^r \sum_{j=r-p}^{p} \frac{X^{(j)}(t_0, W)S^{(r-j)}(t_0, W) + S^{(j)}(t_0, W)X^{(r-j)}(t_0, W)}{j!(r-j)!}. \end{aligned}$$

Without loss of generality, we may take $\tilde{t} \leq 1$. Since $0 < t \leq t_0 \leq \tilde{t}$, we have $|h| \leq 1$, and hence $|h|^r \leq |h|^{p+1}$ for every $r \geq p + 1$. Using the sub-multiplicativity of the matrix norm and the uniform bound above,

$$\begin{aligned} &\left\| \hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) - 2t^\rho W \right\| \\ &\qquad \leq 2M^2 |h|^{p+1} \sum_{r=p+1}^{2p} \sum_{j=r-p}^{p} \frac{1}{j!(r-j)!}. \end{aligned}$$

Thus, defining

$$c := 2M^2 \sum_{r=p+1}^{2p} \sum_{j=r-p}^{p} \frac{1}{j!(r-j)!},$$

we obtain

$$\left\| \hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) - 2t^\rho W \right\| \leq c|t - t_0|^{p+1},$$

where c is independent of t_0 and W. □

The preceding criterion immediately recovers the classical strictly complementary regime.

Corollary 2 (Strict complementarity)**.** *Suppose that the strict complementarity condition holds. Then, for every $p \geq 1$, the* ***AHO*** *weighted central paths are uniformly $(p, 1)$-predictor regular on $\mathcal{W}_\beta$. Consequently,*

$$\mu_{k+1} = O(\mu_k^{p+1}).$$

In particular, choosing $p = 1$ yields quadratic convergence of the normalized complementarity gaps.

Proof. Under strict complementarity, the AHO weighted central paths depend analytically jointly on the path parameter and the weight matrix (see [66]). Hence, after possibly restricting the path parameter to a sufficiently small interval, the mapping

$$(\mu, W) \longmapsto \xi_W(\mu)$$

is jointly $\mathcal{C}^p$-smooth on the compact set $[0, \tilde{\mu}] \times \mathcal{W}_\beta$ for every $p \geq 1$. Corollary 1 therefore implies uniform $(p, 1)$-predictor regularity. The result now follows from Theorem 2 with $\rho = 1$. □

We next give a variant adapted to the semi-algebraic structure of the weighted path family established earlier in the paper.

Corollary 3 (Cell-wise algebraic regularity)**.** *Let $\mathcal{W} \subset \mathbb{S}^n_{++}$ be compact and suppose that $\mathcal{W}$ admits a finite semi-algebraic partition*

$$\mathcal{W} = C_1 \cup \cdots \cup C_r$$

such that, for every i, the mapping

$$(t, W) \longmapsto \xi_W(t^\rho), \qquad (t, W) \in (0, \tilde{t}] \times C_i,$$

admits a $\mathcal{C}^p$ extension to an open neighborhood of

$$[0, \tilde{t}] \times \overline{C_i}.$$

Then the family is uniformly (p, ρ)-predictor regular on $\mathcal{W}$.

Proof. For each i, the set $[0, \tilde{t}] \times \overline{C_i}$ is compact. Hence, by the assumed $\mathcal{C}^p$ extension, all derivatives up to order p of the reparametrized weighted central path are uniformly bounded on this set. Therefore, by the proof of Corollary 1, there exists $c_i > 0$ such that

$$\left\| \hat{X}(t)\hat{S}(t) + \hat{S}(t)\hat{X}(t) - 2t^\rho W \right\| \leq c_i |t - t_0|^{p+1}$$

for every $W \in C_i$ and $0 < t \leq t_0 \leq \tilde{t}$. Since the partition is finite, taking $c := \max_{1 \leq i \leq r} c_i$ proves uniform (p, ρ)-predictor regularity on $\mathcal{W}$. □

Remark 12 (Non-admissible reparametrizations). Analyticity is stronger than what is needed to construct a predictor of a fixed finite order. To see this, consider a Puiseux term $c\mu^\alpha$ appearing in a coordinate of a weighted central path. Under the reparametrization $\mu = t^\rho$, this term becomes $ct^{\rho\alpha}$. If ρ is admissible, all such exponents are non-negative integers, and derivatives of arbitrary order are available at $t = 0$. If ρ is not admissible, some exponent $\rho\alpha$ remains non-integral, and the reparametrized path generally possesses only finitely many bounded derivatives at $t = 0$. A pth-order predictor may still be constructed as long as the derivatives through order p exist and satisfy the uniform bounds required in the preceding analysis.

There is, however, an additional restriction for super-linear convergence. The proof of Theorem 2 requires $p \geq \rho$, and yields the convergence order $(p + 1)/\rho$. Thus, a non-admissible reparametrization can provide the same super-linear guarantee only if it retains at least ρ bounded derivatives. This need not occur. For example, if a relevant coordinate contains a non-zero Puiseux term $c\mu^{1/N}$, then after reparametrization it becomes $ct^{\rho/N}$. If ρ is not a multiple of N, this term provides only $\lfloor \rho/N \rfloor$ bounded derivatives. Since, for

$N > 1$, $\lfloor \frac{\rho}{N} \rfloor < \rho$, the requirements for constructing a pth-order predictor and for obtaining the super-linear convergence guarantee cannot simultaneously be satisfied in this case.

Therefore, an admissible reparametrization removes the finite-differentiability obstruction altogether. Among all admissible choices, $\rho = N$ is the smallest and, for a fixed Taylor order p, gives the largest guaranteed convergence order.

We now prove super-linear convergence of simultaneously vanishing blocks of X_k and S_k with respect to the optimal partition of SDO. To this end, we need to define the notion of the optimal partition (see e.g., [14, Lemma 2.3] or [45]).

Definition 7.2. Let $(X^*, y^*, S^*) \in \mathrm{ri}(\mathcal{P}^* \times \mathcal{D}^*)$, where $\mathrm{ri}(\cdot)$ denotes the relative interior of a convex set, and let $\mathcal{B} := \mathcal{R}(X^*)$, $\mathcal{N} := \mathcal{R}(S^*)$, and $\mathcal{T} := \big(\mathcal{R}(X^*) + \mathcal{R}(S^*)\big)^{\perp}$, where $\mathcal{R}(\cdot)$ is the column space of a matrix. Then the 3-tuple $\big(\mathcal{B}, \mathcal{T}, \mathcal{N}\big)$ is called the *optimal partition* of (P) − (D).

The optimal partition $\big(\mathcal{B}, \mathcal{T}, \mathcal{N}\big)$ is independent of the choice of (X^*, y^*, S^*) [14, Lemma 2.3(i)]. By the complementarity condition, X^* and S^* commute, and thus they have a common eigenvector basis Q, whose columns can be partitioned into $Q = [Q_{\mathcal{B}}, Q_{\mathcal{T}}, Q_{\mathcal{N}}]$ according to the partitions $\mathcal{B}$, $\mathcal{T}$, and $\mathcal{N}$.

Proof of Theorem 3. An orthogonal transformation of $(X^*, y^*, S^*) \in \mathrm{ri}(\mathcal{P}^* \times \mathcal{D}^*)$ with respect to Q reveals the optimal partition as

$$Q^T X^* Q = \begin{pmatrix} U_{X^*} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Q^T S^* Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & U_{S^*} \end{pmatrix},$$

where $U_{X^*}, U_{S^*} \succ 0$. Now, it follows from [45, Theorem 4.3] that when k is large enough, there exist positive constants κ, κ' independent of t_{k+1} (but dependent on n and β) and a positive exponent γ such that

$$\kappa' t_{k+1}^{(1-\gamma)\rho} \le \|Q_{\mathcal{T}}^T X_{k+1} Q_{\mathcal{T}}\| \le \kappa t_{k+1}^{\gamma\rho},$$

where

$$\frac{1}{2^{n-1}} \le \gamma \le \frac{1}{2}.$$

Further, using [45, Theorem 4.3] again and uniform boundedness, there exist positive constants η, η' independent of t_{k+1} (dependent on n and a condition number defined in [45, Page 232]) such that

$$\eta' t_{k+1}^{\rho} \le \|Q_{\mathcal{N}}^T X_{k+1} Q_{\mathcal{N}}\| \le \eta t_{k+1}^{\rho}.$$

Now, for sufficiently large k, it follows from Theorem 2 that

$$\begin{aligned} \|Q_{\mathcal{T}}^T X_{k+1} Q_{\mathcal{T}}\| \le \kappa t_{k+1}^{\gamma\rho} &\le \kappa \left(\frac{c}{2\beta}\right)^{\gamma} t_k^{(p+1)\gamma} \\ &\le \kappa \left(\frac{c}{2\beta}\right)^{\gamma} \left(\frac{1}{\kappa'}\right)^{\frac{(p+1)\gamma}{\rho(1-\gamma)}} \|Q_{\mathcal{T}}^T X_k Q_{\mathcal{T}}\|^{\frac{(p+1)\gamma}{\rho(1-\gamma)}}, \\ &\le \kappa \left(\frac{c}{2\beta}\right)^{\gamma} \left(\frac{1}{\kappa'}\right)^{\frac{(p+1)\gamma}{\rho(1-\gamma)}} \|Q_{\mathcal{T}}^T X_k Q_{\mathcal{T}}\|^{\frac{p+1}{\rho(2^{n-1}-1)}}, \end{aligned}$$

where the last inequality follows from $\|Q_{\mathcal{T}}^T X_k Q_{\mathcal{T}}\| \to 0$ and $\gamma/(1-\gamma) \ge 1/(2^{n-1}-1)$. Analogously, we can derive

$$\begin{aligned} \|Q_{\mathcal{N}}^T X_{k+1} Q_{\mathcal{N}}\| \le \eta t_{k+1}^{\rho} &\le \eta \left(\frac{c}{2\beta}\right) t_k^{(p+1)} \\ &\le \eta \left(\frac{c}{2\beta}\right) \left(\frac{1}{\eta'}\right)^{\frac{p+1}{\rho}} \|Q_{\mathcal{N}}^T X_k Q_{\mathcal{N}}\|^{\frac{p+1}{\rho}}. \end{aligned}$$

The proof is similar for $Q_{\mathcal{B}}^T S_k Q_{\mathcal{B}}$ and $Q_{\mathcal{T}}^T S_k Q_{\mathcal{T}}$. □

7.2. **Iteration complexity.** We first derive a quantitative reduction bound for the normalized complementarity gap on any region where the uniform (p,ρ)-predictor regularity estimate holds. We then combine this bound with the super-linear recurrence of Theorem 2 to obtain a doubly logarithmic local outer iteration bound.

Proposition 7.1 (Predictor reduction)**.** *Suppose that the family of reparametrized weighted central paths is uniformly (p,ρ)-predictor regular on $\mathcal{W}_\beta$ with constants $\tilde{t} > 0$ and $c > 0$. Let $p \geq \rho$, and consider an iterate with $0 < t_k \leq \tilde{t}$. Let $\bar{t}_k \in (0, t_k)$ be the unique solution of*

$$\frac{c(t_k - \bar{t}_k)^{p+1}}{2\bar{t}_k^{\rho}} = \beta,$$

as in the proof of Theorem 2. If $\alpha_k := \bar{t}_k / t_k$, then

$$\alpha_k \leq \frac{1}{1 + \left(\frac{2\beta}{c\, t_k^{p+1-\rho}}\right)^{1/\rho}}. \tag{7.4}$$

Further, we have

$$\mu_{k+1} \leq \left[1 + \left(\frac{2\beta}{c\, t_k^{p+1-\rho}}\right)^{1/\rho}\right]^{-\rho} \mu_k.$$

Proof. Apply the argument in the proof of Theorem 2 with $t_0 = t_k$ and with the weight associated with the current iterate. In particular, inequality (7.3) shows that, if $\bar{t}_k \in (0, t_k)$ satisfies

$$\frac{c(t_k - \bar{t}_k)^{p+1}}{2\bar{t}_k^{\rho}} = \beta,$$

then every $t \in [\bar{t}_k, t_k]$ is feasible for the predictor problem (6.7). Since the left-hand side is continuous and strictly decreasing from $+\infty$ to 0 as $\bar{t}_k$ increases from 0 to t_k, such $\bar{t}_k$ is unique. Hence $t_+ \leq \bar{t}_k$. Now, set

$$\alpha_k := \frac{\bar{t}_k}{t_k} \in (0, 1).$$

The defining equation for $\bar{t}_k$ gives

$$\frac{(1-\alpha_k)^{p+1}}{\alpha_k^{\rho}} = \frac{2\beta}{c\, t_k^{p+1-\rho}}.$$

Since $p \geq \rho$,

$$\left(\frac{1-\alpha_k}{\alpha_k}\right)^{\rho} \geq \frac{2\beta}{c\, t_k^{p+1-\rho}},$$

and therefore

$$\alpha_k \leq \frac{1}{1 + \left(\frac{2\beta}{c\, t_k^{p+1-\rho}}\right)^{1/\rho}}.$$

Finally, Lemma 6.3 gives $\mu_{k+1} = t_+^{\rho}$, while by definition $\mu_k = t_k^{\rho}$. Thus,

$$\mu_{k+1} = t_+^{\rho} \leq \bar{t}_k^{\rho} = \alpha_k^{\rho} \mu_k,$$

and substituting the preceding bound on α_k yields the claimed estimate. □

Remark 13 (Lower-order predictors). Theorem 2 requires $p \geq \rho$ and hence does not provide a super-linear convergence guarantee when $p < \rho$. We leave the behavior of the method in this regime for future investigation.

Remark 14 (Global reduction versus polynomial complexity). Proposition 7.1 shows that, whenever the uniform (p,ρ)-predictor regularity estimate is valid, Algorithm 1 makes a quantitative reduction in the normalized complementarity gap. In particular, the reduction becomes increasingly aggressive as $t_k \downarrow 0$.

This estimate, however, does not by itself yield the conventional polynomial iteration complexity of path-following `IPM`s. Indeed, the bound on the predictor reduction depends on the constant c in the bound (7.4). Consequently, the estimate does not provide a uniform reduction of the form

$$\mu_{k+1} \le \left(1 - \frac{\delta}{\sqrt{n}}\right) \mu_k$$

with a constant $\delta > 0$ independent of the problem data. Thus, Proposition 7.1 should be interpreted as a quantitative progress estimate for the higher-order predictor rather than as a polynomial outer iteration complexity bound.

We now turn to the local regime. When $p \ge \rho$, Theorem 2 gives the stronger recurrence

$$\mu_{k+1} = O\left(\mu_k^{(p+1)/\rho}\right),$$

from which we obtain a doubly logarithmic local outer iteration bound.

Proof of Theorem 4. Recall $d = p + 1 - \rho \ge 1$. By Proposition 7.1,

$$\alpha_k \le \frac{1}{1 + \left(\frac{2\beta}{c t_k^d}\right)^{1/\rho}}.$$

Since $t_k \le t_0$ and $d \ge 1$,

$$\frac{2\beta}{c\, t_k^d} \ge \frac{2\beta}{c\, t_0^d},$$

and therefore $\alpha_k \le q_0$. Now, using Proposition 7.1 we get

$$\mu_{k+1} \le \alpha_k^\rho \mu_k \le q_0^\rho \mu_k.$$

Inductively, we get $\mu_k \le q_0^{\rho k} \mu_0$. Hence, $\mu_k \le \eta$ whenever

$$q_0^{\rho k} \mu_0 \le \eta,$$

which gives

$$k \ge \frac{\log(\mu_0/\eta)}{\rho \log(1/q_0)}.$$

For the local phase, Theorem 2 gives

$$\mu_{k+1} \le C \mu_k^q, \qquad q := \frac{p+1}{\rho} > 1,$$

for all sufficiently small μ_k. Choose $\mu_{\mathrm{loc}} > 0$ sufficiently small so that $C\mu_{\mathrm{loc}}^{q-1} < 1$. Setting $v_k := C^{1/(q-1)} \mu_k$ gives

$$v_{k+1} \le v_k^q.$$

Thus, the normalized complementarity gap decreases doubly exponentially with the local outer iteration counter, and the number of additional outer iterations required to reach $\mu_k \le \zeta$ is

$$O\left(\log\log \frac{1}{\zeta}\right).$$

□

We can also comment on the costs of predictor and corrector steps per iteration.

Notation 10. The symmetric Kronecker product of any two square matrices K_1 and K_2 is defined as a mapping

$$(K_1 \otimes_s K_2)\,\mathrm{svec}(H) := \frac{1}{2}\,\mathrm{svec}\left(K_2 H K_1^T + K_1 H K_2^T\right),$$

where H is any symmetric matrix (see e.g., [14] for more details).

Remark 15 (Cost of higher-order directions). At a fixed outer iterate, all derivative systems in (6.5) have the same coefficient matrix and differ only in their right-hand sides. More specifically, to obtain higher-order derivatives, we iteratively solve the system (6.5) at each iteration of `IPM`, where the coefficient matrix is fixed in each iteration. The full Newton system is given by

$$\begin{bmatrix} 0 & \mathbf{A}^T & \mathcal{I} \\ \mathbf{A} & 0 & 0 \\ S \otimes_s I & 0 & X \otimes_s I \end{bmatrix} \begin{bmatrix} \mathrm{svec}(X^{(j)}) \\ y^{(j)} \\ \mathrm{svec}(S^{(j)}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mathrm{svec}(f(W, \xi(j))) \end{bmatrix} \qquad \forall j = 1, \ldots, p, \tag{7.5}$$

where $\mathbf{A} = [\mathrm{svec}(A_1), \ldots, \mathrm{svec}(A_m)]^T$, and

$$f(W, \xi(j)) = \begin{cases} 2\rho(\rho-1)\cdots(\rho-j+1)t^{\rho-j}W + \xi(j) & j \le \rho, \\ \xi(j) & j > \rho. \end{cases}$$

As an alternative, we can build the normal equations system:

$$\left(\mathbf{A}\mathbf{E}^{-1}\mathbf{F}\mathbf{A}^T\right) y^{(j)} = -\mathbf{A}\mathbf{E}^{-1}\,\mathrm{svec}(f(W, \xi(j))), \qquad \forall j = 1, \ldots, p, \tag{7.6}$$

where $\mathbf{E} = S \otimes_s I$ and $\mathbf{F} = X \otimes_s I$ (notice that both mappings $\mathbf{E}$ and $\mathbf{F}$ are invertible at the interior solution (X, y, S)). Given that $\mathbf{A}\mathbf{E}^{-1}\mathbf{F}\mathbf{A}^T$ is positive definite, we can perform a one-time factorization to solve (7.6). We then perform forward-backward substitution for computing higher-order derivatives.

Thus, the additional cost of increasing the predictor order is primarily the cost of these additional right-hand sides and back-solves, rather than p independent factorizations. Each corrector step similarly requires the solution of one `AHO` Newton system. We do not, however, obtain here a uniform bound on the number of inner corrector steps per outer iteration.

8. Numerical experiments

In this section, we elaborate on the implementation details of Algorithm 1. After fixing (ρ, p) at some pre-determined pair of integers, we initiate Algorithm 1 with a feasible solution $(X_0, y_0, S_0) \in \mathcal{N}(\beta)$ with the predictor-neighborhood parameter $\beta = 0.5$. The value $\beta = 0.5$ used in the numerical experiments is larger than the sufficient neighborhood parameter required by the theoretical well-definedness argument in Proposition 6.1. This is an implementation choice intended to permit larger predictor steps.

Predictor step. The predictor step consists of computing the higher-order derivatives of a reparametrized weighted central path. To that end, we solve (7.5) directly. Once the higher-order information is computed, we construct the pth-order Taylor approximation of the weighted central path passing through the current point (X_k, y_k, S_k).

The theoretical algorithm uses the boundary parameter t_+ defined by (6.7). Rather than solving this one-dimensional semi-algebraic problem exactly, in our implementation we approximate t_+ by a geometric line search. Starting from $t = t_k$, we successively reduce t according to

$$t \leftarrow \alpha_{\mathrm{ls}} t, \qquad \alpha_{\mathrm{ls}} = 0.9999,$$

and retain the last accepted value before the next trial point leaves $\mathcal{N}(2\beta)$. Since we use $\beta = 0.5$ in the numerical experiments, we additionally check the semi-definite constraints during the forward-tracking line search. In particular, a trial point is accepted only if it remains in $\mathcal{N}(2\beta)$ and its primal and dual matrices remain positive definite. Thus, the line search is prevented from overshooting the positive semi-definite cone.

Corrector step. At the point obtained from the predictor step, we perform, possibly multiple, damped centering steps toward the central path according to (6.8)-(6.10). For this purpose, we set the centrality tolerance, which determines the desired proximity to the central path, to 10^{-1}. Upon completion of the centering step, the resulting interior solution is used as the base point for constructing the Taylor approximation in the subsequent iteration. Although the implementation allows for damped corrector steps, all centering steps taken in the experiments reported below were full steps ($\alpha = 1$); no damping was required.

The above procedure is repeated until the complementarity gap falls below the prescribed tolerance $\zeta = 10^{-10}$.

To evaluate the performance of our higher-order `IPM` with respect to the convergence order, we compared it to MOSEK[1] and SDPA[2] through CVXPY[3], and SeDuMi[4] and SDPT3[5]. While MOSEK serves as our primary benchmark, we occasionally include additional solvers for further comparison. MOSEK employs a homogeneous self-dual (HSD) embedding together with a Mehrotra-type predictor-corrector scheme. SeDuMi likewise uses a predictor-corrector approach. SDPT3 provides both an HSD-embedded path-following method (SDPT3-HSD) and an infeasible path-following method (SDPT3), while SDPA is based on a Mehrotra-type predictor-corrector primal-dual `IPM`. Unless otherwise stated, the gap values reported for SDPT3 in the figures correspond to the original problem formulation, irrespective of whether the infeasible or HSD variant is used. Our primary objective is to compare the asymptotic complementarity gap reduction and outer iteration counts of these established solvers with those of our higher-order `IPM`, rather than wall-clock performance. A higher-order iteration generally requires more linear algebra than a conventional first-order predictor-corrector iteration. Our algorithm is implemented in Python, and the source code is available at `https://github.com/psm-optimizes/HOIPM4SDO.git`.

8.1. **Initialization.** Solvers based on HSD embedding are initialized with a point (X_0, y_0, S_0) such that $\frac{\langle X_0, S_0\rangle}{n} = 1$. To facilitate a fair comparison of complementarity gap reduction, we initialize our `IPM` with a point having the same initial gap. Moreover, we set the optimality tolerance to 10^{-10} for both our higher-order `IPM` and all competing solvers.

The figures in this section illustrate the convergence history of our algorithm, either in isolation or in comparison with other solvers. In each figure, the horizontal axis (*Iteration*) represents the iteration count, while the vertical axis (*Normalized Complementarity Gap (log)*) displays the normalized complementarity gap $\frac{\langle X_k, S_k\rangle}{n}$ at iteration k on a logarithmic scale.

For the prescribed maximum order p of the Taylor approximation, we choose $p \in \{\rho, 2\rho, 3\rho\}$, unless stated otherwise. In the numerical implementation, the computation of higher-order derivatives may terminate before reaching p if the scaled Taylor coefficient at a given order falls below the threshold 10^{-16}. Thus, p denotes the prescribed maximum Taylor order, while the effective order used at a particular iteration may be smaller.

8.2. **Modification of the Newton systems.** Due to the structure of the right-hand side (RHS) in the linear systems used to compute higher-order derivatives, which contains the factor $\rho(\rho-1)\cdots(\rho-j+1)$, the right-hand side might have a significantly large value, even for moderately large reparametrization exponents. To improve the numerical accuracy of the linear solves, we employ two scaling strategies. First, we scale the system with t_0^j for computing the jth-order derivative. As our `IPM` starts at a complementarity gap of 1 and proceeds downward, this scaling helps reduce the magnitude of the RHS. Second, we further

[1] `https://www.mosek.com/`
[2] `https://sdpa.sourceforge.net/`
[3] `https://www.cvxpy.org/`
[4] `https://sedumi.ie.lehigh.edu/`
[5] `https://www.math.cmu.edu/~reha/sdpt3.html`

scale the system using the $1/j!$ coefficient from the Taylor expansion. For convenience, let

$$(\bar{X}^{(j)}, \bar{y}^{(j)}, \bar{S}^{(j)}) := \frac{t_0^j}{j!}(X^{(j)}, y^{(j)}, S^{(j)}).$$

Specifically, instead of (7.5), we solve

$$\begin{bmatrix} 0 & \mathbf{A}^T & I \\ \mathbf{A} & 0 & 0 \\ \mathbf{E} & 0 & \mathbf{F} \end{bmatrix} \begin{bmatrix} \operatorname{svec}(\bar{X}^{(j)}) \\ \bar{y}^{(j)} \\ \operatorname{svec}(\bar{S}^{(j)}) \end{bmatrix} = \frac{t_0^j}{j!} \begin{bmatrix} 0 \\ 0 \\ \operatorname{svec}(f(W, \xi^{(j)})) \end{bmatrix}, \qquad j = 1, \dots, p.$$

Analogously, instead of the normal equations system (7.6), we solve

$$\left(\mathbf{A}\mathbf{E}^{-1}\mathbf{F}\mathbf{A}^T\right) \bar{y}^{(j)} = -\frac{t_0^j}{j!}\mathbf{A}\mathbf{E}^{-1} \operatorname{svec}(f(W, \xi^{(j)})), \qquad j = 1, \dots, p.$$

After the scaled systems are solved, the resulting quantities are multiplied by $j!/t_0^j$ to recover the true derivatives. Together, these modifications significantly improve the numerical precision of the computed derivatives by reducing the magnitude of the right-hand sides of the linear systems.

8.3. **Experiments.** We evaluate our higher-order `IPM` on five sets of benchmark problems. In Section 8.3.1, we consider the 3-elliptope problem, whose central path has ramification index 2. In Section 8.3.2, we solve three instances from [14, Example 3.3, p. 58], for which a lower bound on the global ramification index is known. We then turn to problem classes for which no prior information on the ramification index is available. In Section 8.3.3, we study a set of generated problems that fail to satisfy the strict complementarity condition. We next consider problems with a positive duality gap. Since the homogeneous self-dual embedding of such problems fails strict complementarity [15], Section 8.3.4 reports results for solving their self-dual embedding counterparts. Finally, in Section 8.3.5, we evaluate our higher-order `IPM` on problems that satisfy strict complementarity and therefore do not require reparametrization.

For our experiments, the two main input parameters are the candidate reparametrization exponent ρ and the Taylor order p. When ρ is admissible and the uniform (p, ρ)-predictor regularity assumption of Theorem 2 holds, choosing $p \geq \rho$ yields a Q-order of at least $(p+1)/\rho > 1$. Accordingly, we take

$$p \in \{\rho, 2\rho, 3\rho\}.$$

When an admissible reparametrization exponent is not known, we test the candidate values $\rho = 2, 3, 4$, motivated by the heuristic choices suggested in [25, Remark 2]. For several of the tested choices of (ρ, p), we observe behavior consistent with super-linear convergence. This speed-up substantially reduces the number of outer iterations required to reach the prescribed tolerance relative to the other solvers considered in our study.

8.3.1. *Experiments on the 3-elliptope problem.* We recall Example 1, whose ordinary central path has ramification index 2. Motivated by this value, we compare $\rho = 1$ (no reparametrization) with the candidate reparametrization exponent $\rho = 2$. As shown in Figure 2, reparametrization is essential for improving the convergence rate: when $\rho = 1$, increasing the order of the Taylor approximation may reduce the number of iterations, but it does not improve the asymptotic convergence rate. In contrast, when $\rho = 2$, higher-order Taylor approximations lead to both a faster convergence rate and a smaller number of iterations. Figure 3 further demonstrates a comparison between our higher-order `IPM` with $\rho = 2$ versus several solvers. We observe that our higher-order `IPM` requires fewer outer iterations and exhibits a steeper reduction in the normalized complementarity gap than the other solvers considered on the 3-elliptope problem. Particularly, with $p = 8$, the prescribed tolerance is reached in two outer iterations. As discussed in Remark 15, in addition to the full Newton system, one can exploit the normal equations system, and our algorithm requires one factorization per iteration, followed by p back-solves to compute the higher-order

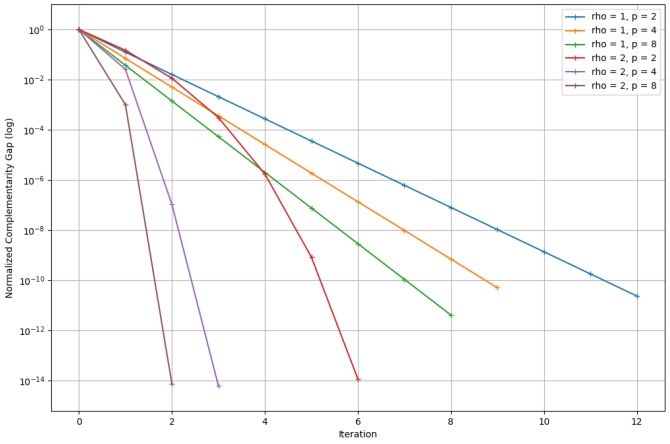


FIGURE 2. Impact of higher-order derivatives for solving the 3-elliptope problem with and without a reparametrization

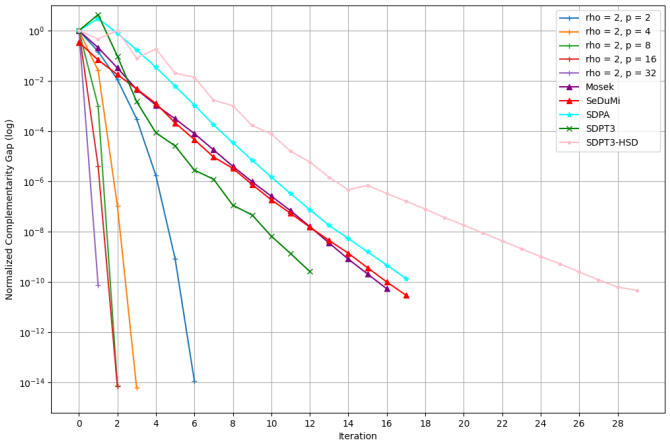


FIGURE 3. Increasing the order of the Taylor approximation helps reduce the number of iterations

derivatives. This strategy, which uses one factorization followed by multiple back-solves, has also been advocated as a means of reducing the number of `IPM` iterations [5, Page 31]. Although our preference is to keep the order of the Taylor approximation small, allowing more solves per iteration can further decrease the total number of iterations. In fact, as illustrated in Figure 3, increasing p sufficiently enables the prescribed tolerance to be reached in a single outer iteration.

Finally, recall the Puiseux expansion (3.2) of the central path near $\mu = 0$. We note that with a reparametrization $\mu \mapsto \mu^3$, we get

$$X_{12}(\mu) = -1 + \frac{\sqrt{8}}{8}\mu^{3/2} + \frac{1}{32}\mu^3 - \frac{11\sqrt{8}}{2048}\mu^{9/2} - \frac{3}{512}\mu^6 - \frac{121\sqrt{8}}{1048576}\mu^{15/2} + \dots,$$

whose second derivative does not exist at $\mu = 0$. Consequently, choosing $\rho = 3$ restores first-order differentiability of the central path, but does not recover analyticity at $\mu = 0$.

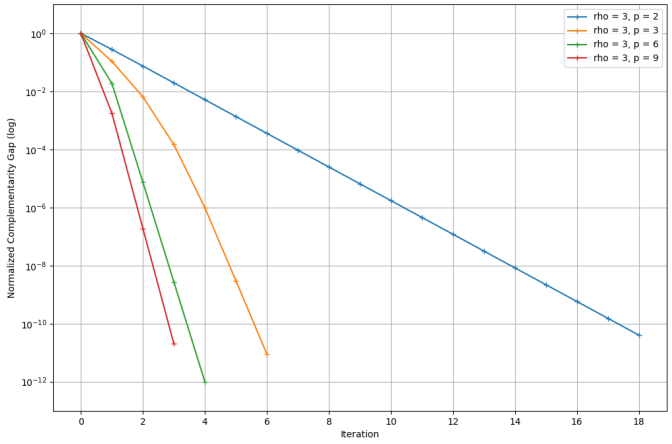


FIGURE 4. Using $\rho = 3$ for the 3-elliptope problem

Figure 4 shows that we get speed-ups even with $\rho = 3$. Still, as we approach optimality, the diagnostic corresponding to $\omega = 1.05$ eventually ceases to decrease, so the reported iterates do not provide evidence of a Q-order exceeding 1.05. We examine this behavior in more detail in Section 8.4.

8.3.2. *Experiments on a scalable instance.* In this section, we consider the following `SDO` problem (from [14]) in dual form:

$$\max \left\{ -y_n \mid S = \begin{pmatrix} 1 & y_1 & y_2 & \dots & y_{n-1} \\ y_1 & y_2 & 0 & \dots & 0 \\ y_2 & 0 & y_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ y_{n-1} & 0 & \dots & 0 & y_n \end{pmatrix} \succeq 0 \right\},$$

which has a unique optimal solution $y_i^* = 0$ for $i = 1, \dots, n$. It is easy to check that $y_2(\mu) = c\mu^{2^{-(n-2)}}$ for some constant c independent of μ, which yields the lower bound 2^{n-2} on the ramification index of the central path. Since the central path corresponds to $W = I$, this also gives the lower bound $N \geq 2^{n-2}$ on the global ramification index of the weighted family. We begin with the case $n = 3$, for which the above bound gives $N \geq 2$. Motivated by this lower bound, we first test the candidate reparametrization exponent $\rho = 2$. Figure 5 illustrates the effect of the reparametrization with $\rho = 2$ for several orders of Taylor approximation. Remarkably, orders $p = 2$ and $p = 3$ reach the prescribed tolerance in a single outer iteration. In contrast, when $p = 1$, our `IPM` exhibits a convergence rate comparable to that of the other solvers, namely, linear convergence.

We next consider the cases $n = 4, 5$, for which the above lower bound gives $N \geq 4$ and $N \geq 8$, respectively. Motivated by these bounds, we test $\rho = 4$ for $n = 4$ and $\rho = 8$ for $n = 5$. For $n = 4$, the fourth-order Taylor approximation reaches the prescribed tolerance in a single outer iteration (see Figure 6), while $p = 2, 3$ exhibit nearly sub-linear and linear convergence rates, respectively. For $n = 5$, Figure 7 shows that the prescribed tolerance is reached in a single outer iteration for $p = 8, 9$. We further observed, in additional experiments not shown here, that the tested choices $\rho < 7$ did not exhibit super-linear convergence.

8.3.3. *Random problems failing strict complementarity.* In this section, we extend our experiments to randomly generated instances that violate the strict complementarity condition. The problems are generated using [46, Algorithm 11], which constructs `SDO` instances using

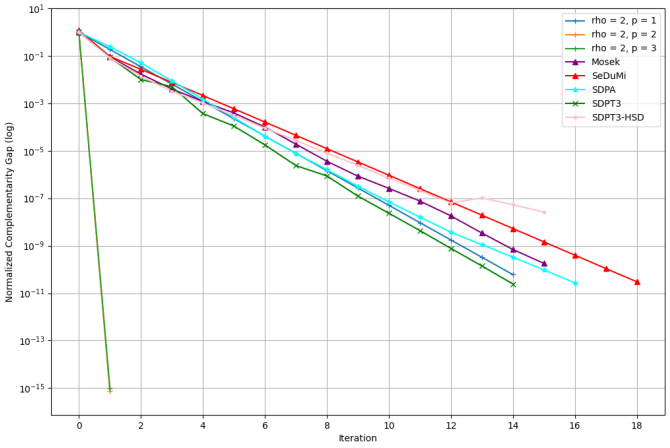


FIGURE 5. Impact of reparametrization for the scalable problem of size $n = 3$

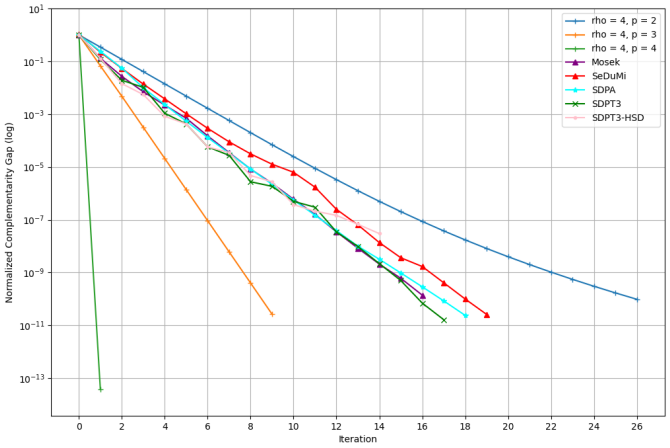


FIGURE 6. Impact of reparametrization for the scalable problem of size $n = 4$

a known strictly feasible solution and an optimal solution. We restrict our attention to problems whose partition $\mathcal{T}$ is non-empty (see Definition 7.2) and additionally enforce the primal and dual non-degeneracy conditions of [2]. We generate a set of `SDO` problems that satisfy both primal and dual non-degeneracy conditions, implying that primal and dual optimal solutions are unique (see [2, Theorems 7 and 10]), while failing the strict complementarity condition, i.e., $X^* + S^* \not\succ 0$. The optimal partition configurations for all test instances are summarized in Table 1, where $|\cdot|$ means the dimension of the corresponding subspace of the optimal partition.

Since the global ramification index is unavailable for these instances, we test several candidate reparametrization exponents. From this point onward, we restrict our comparisons to MOSEK, owing to its consistently strong performance in the preceding experiments. We also tested SDPA and observed that its performance is at best similar to that of MOSEK.

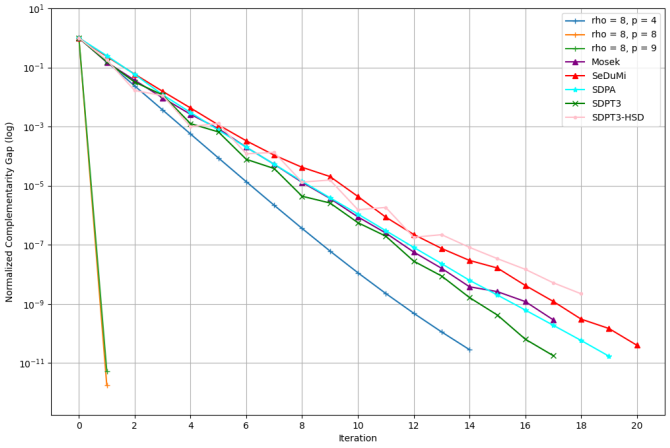


FIGURE 7. Impact of reparametrization for the scalable problem of size $n = 5$

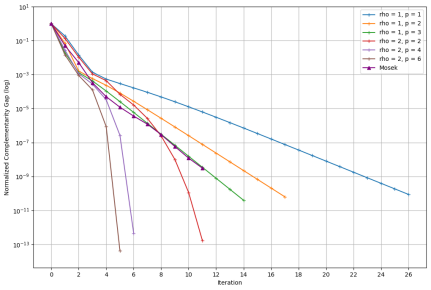
The convergence histories of our `IPM` on six non-degenerate instances are shown in Figures 8 and 9. For each instance, we first compare the performance of the method without reparametrization against the reparametrized variant with $\rho = 2$ for several preselected values of p. Consistent with the previous examples, increasing the order of the Taylor approximation in the absence of reparametrization does not improve the asymptotic convergence rate; it only reduces the total number of iterations required. In contrast, with $\rho = 2$ we observe behavior consistent with super-linear convergence. As the iterates approach optimality, the rate of decrease of the complementarity gap accelerates. We also compared the performance of reparametrizations with $\rho = 2$ and $\rho = 3$. While the case $\rho = 2$ exhibits the expected behavior that higher-order approximations lead to steeper convergence, the same trend is not observed for $\rho = 3$. In particular, increasing the approximation order does not yield comparable improvements in the convergence speed.

It is worth emphasizing that our `IPM` preserves its super-linear convergence behavior up to the target complementarity gap of $\langle X, S\rangle/n = 10^{-10}$. Moreover, in several instances, our method attains a smaller final normalized complementarity gap than MOSEK under comparable tolerance settings.

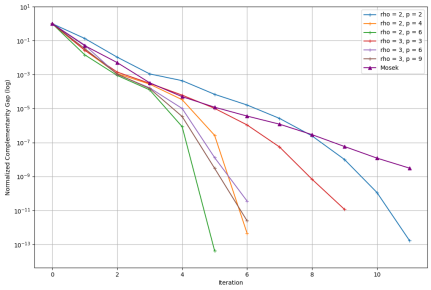
TABLE 1. Random instances with primal and dual non-degeneracy

Instance	m	n	$\lvert\mathcal{B}\rvert$	$\lvert\mathcal{N}\rvert$	$\lvert\mathcal{T}\rvert$	$\lVert X^*\rVert$	$\lVert S^*\rVert$
ND-1	3	4	1	2	1	6.45E-01	3.08E+00
ND-2	4	6	1	4	1	6.93E-01	3.33E+00
ND-3	5	6	1	4	1	2.64E-01	3.51E+00
ND-4	6	7	1	4	2	8.02E-01	7.27E+00
ND-5	6	8	1	5	2	4.05E-01	6.75E+00
ND-6	7	10	1	7	2	2.92E-01	6.00E+00

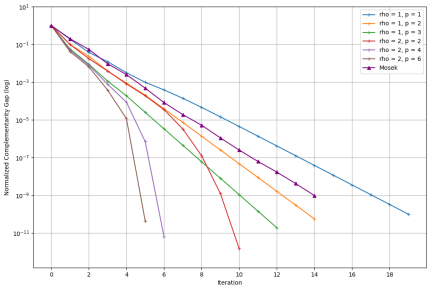
8.3.4. *Problems with positive duality gaps.* In this section, we consider problems with a positive duality gap (PDG). It is known that, when formulated in homogeneous self-dual (HSD) embedding form, such problems fail to satisfy the strict complementarity condition [77].



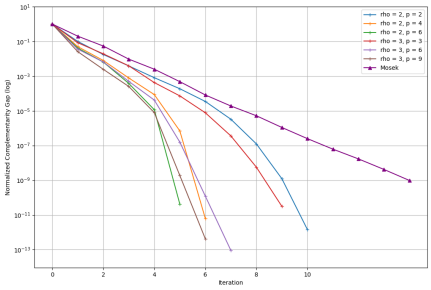
ND-1: $\rho = 1$ vs 2



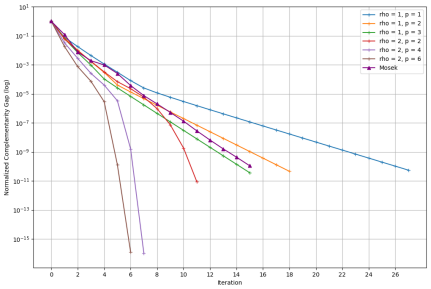
ND-1: $\rho = 2$ vs 3



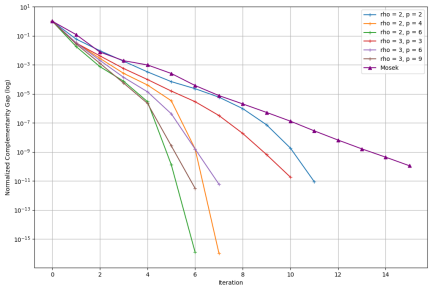
ND-2: $\rho = 1$ vs 2



ND-2: $\rho = 2$ vs 3



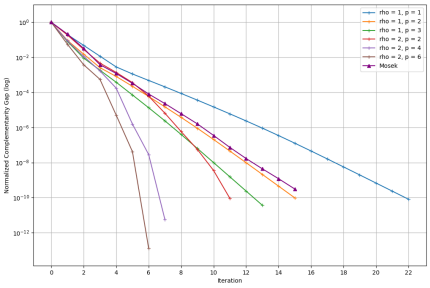
ND-3: $\rho = 1$ vs 2



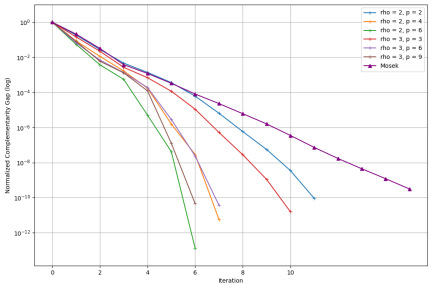
ND-3: $\rho = 2$ vs 3

FIGURE 8. Convergence of random instances with primal-dual non-degeneracy (ND1-ND3)

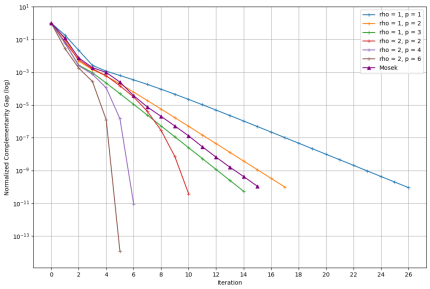
We consider a collection of PDG instances from [57], which are described by the author as clean examples with a finite duality gap. Each instance is first transformed into its HSD embedding, and the resulting embedded problem is then solved using our higher-order `IPM`. Table 2 summarizes the dimensions of the original PDG problems, (m, n), together with the dimensions of their corresponding embedded formulations, $(m_{\text{HSD}}, n_{\text{HSD}})$. The convergence histories of our `IPM` on several embedded PDG instances are presented in Figure 10.



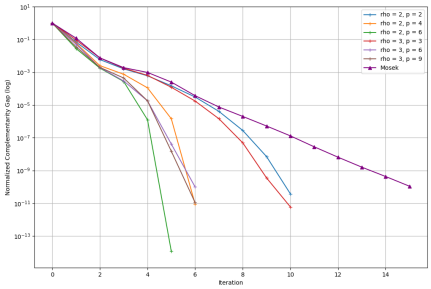
ND-4: $\rho = 1$ vs 2



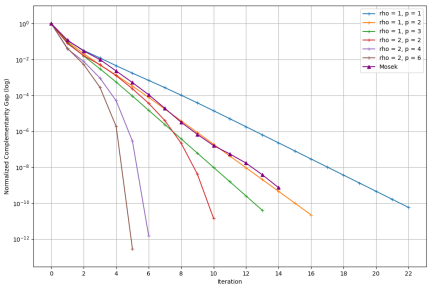
ND-4: $\rho = 2$ vs 3



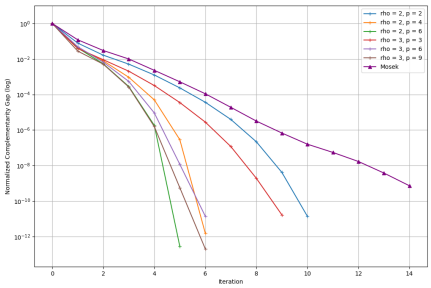
ND-5: $\rho = 1$ vs 2



ND-5: $\rho = 2$ vs 3



ND-6: $\rho = 1$ vs 2

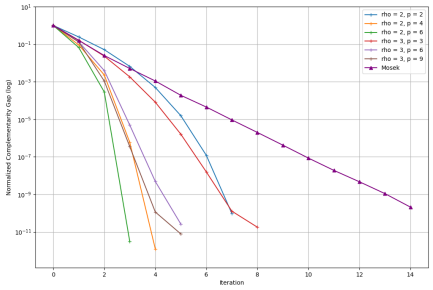


ND-6: $\rho = 2$ vs 3

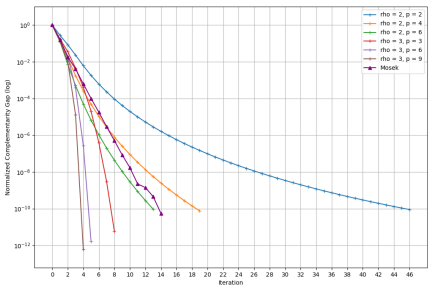
FIGURE 9. Convergence of random instances with primal-dual non-degeneracy (ND4-ND6)

TABLE 2. Problems with positive duality gap

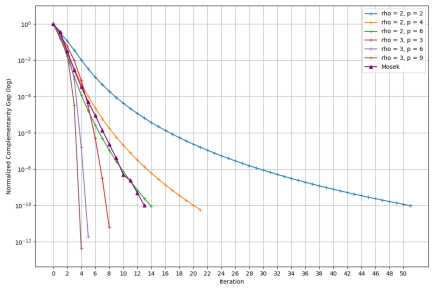
Instance	m	n	m_{HSD}	n_{HSD}
1	2	3	12	18
2	3	4	18	24
3	4	5	25	30
4	5	6	33	36
5	6	7	42	42
6	7	8	52	48



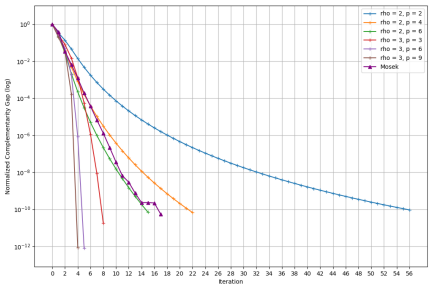
PDG: Instance 1



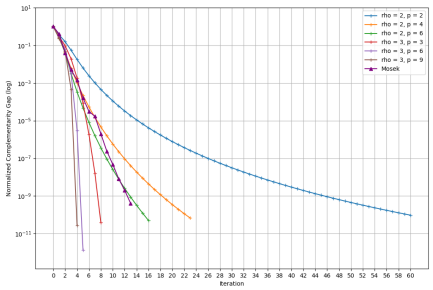
PDG: Instance 2



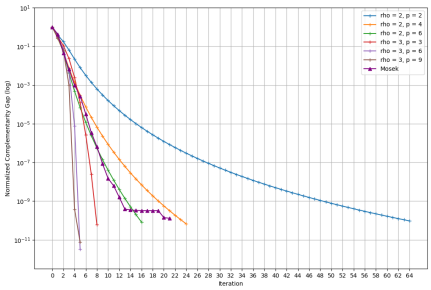
PDG: Instance 3



PDG: Instance 4



PDG: Instance 5



PDG: Instance 6

FIGURE 10. Convergence on PDG Problems

Among the tested reparametrization exponents, $\rho = 3$ is the smallest value for which we observe sustained super-linear behavior on all but Instance 1. Under the choice $\rho = 3$, our method requires fewer outer iterations than MOSEK and exhibits a steeper asymptotic reduction in the normalized complementarity gap on these instances. We further observe that our algorithm converges to a primal-dual optimal solution that fails the strict complementarity condition and differs from the solution returned by MOSEK. This discrepancy is not unexpected, as the embedded problems generally do not possess a unique optimal solution. Rather, they admit a set of optimal solutions, all of which fail to satisfy strict complementarity.

8.3.5. *Experiments on instances satisfying strict complementarity.* We conclude by testing our `IPM` on instances for which strict complementarity is known to hold. Since the strict complementarity condition is satisfied, no reparametrization is required, and thus $\rho = 1$ is sufficient. Nevertheless, for comparison purposes, we also include the cases $\rho = 2, 3$.

Random problems with dual non-degeneracy and strict complementarity. In this section, we consider randomly generated problems that satisfy the dual non-degeneracy condition of [2] (which implies the uniqueness of the primal optimal solution [2, Theorem 10]) and the strict complementarity condition, but they violate the primal non-degeneracy condition (see Table 3). Similar to the problem set considered in Section 8.3.3, these instances are generated using the procedure described in [46, Algorithm 11].

The convergence histories of our `IPM` on three such instances are presented in Figure 11. The super-linear convergence of predictor-corrector `IPM`s under strict complementarity is well established (see Section 4.3), and our numerical results are consistent with this theory. In particular, our `IPM` converges substantially faster than in instances whose unique optimal solution does not satisfy strict complementarity. The convergence profiles exhibit a super-linear rate across all instances and for a wider range of choices of ρ. Furthermore, although the target complementarity gap was set to 10^{-10}, the algorithm attained solutions with significantly smaller complementarity gaps in all tested instances.

TABLE 3. Random instances with dual non-degeneracy and strict complementarity

Instance	m	n	$\lvert\mathcal{B}\rvert$	$\lvert\mathcal{N}\rvert$	$\lVert X^*\rVert$	$\lVert S^*\rVert$
D-1	6	5	1	4	2.18E-01	1.13E+00
D-2	8	8	1	7	6.48E-01	3.89E+00
D-3	10	9	1	8	4.14E-01	1.11E+00

Random problems with strict complementarity. Next, we consider two benchmark instances: one from the literature [25]:

$$\inf\left\{2x_{11} - 2x_{12} \mid \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & (1-2x_{11})/3 \end{bmatrix} \succeq 0\right\}, \tag{8.1}$$

and another generated using the `SDO` generator of [46, Algorithm 11]. We generate a small instance where $n = 5$, $m = 5$, $|\mathcal{B}| = 2$, $|\mathcal{N}| = 3$, $|\mathcal{T}| = 0$ (recall that $|\mathcal{T}| = 0$ implies the strict complementarity), which also satisfies the primal-dual non-degeneracy conditions of [2].

Figures 12 and 13 present the numerical results for the benchmark problem (8.1) and the random problem. Across all experiments, we observe a super-linear rate of convergence for every choice of p. Moreover, as expected, increasing the order of the Taylor approximation leads to greater reductions in the complementarity gap.

8.4. **Convergence of the sequence of complementarity gaps.** In this section, we examine the convergence behavior of the sequence of complementarity gaps in greater detail. Under the assumptions of Theorem 2, the complementarity gaps satisfy

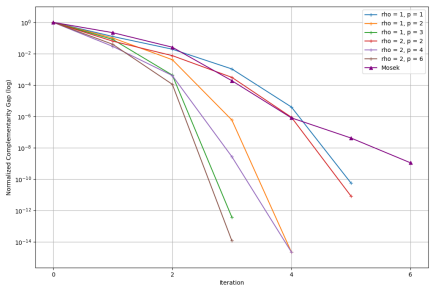
$$\mu_k = O\left(\mu_{k-1}^{(p+1)/\rho}\right).$$

Hence, for any $\omega < (p+1)/\rho$, the ratio

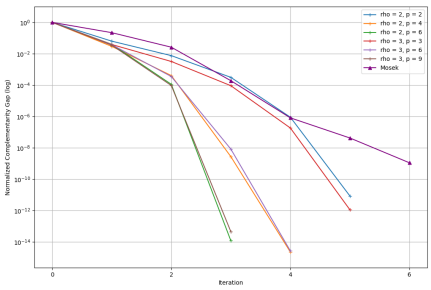
$$\frac{\mu_k}{\mu_{k-1}^{\omega}}$$

tends to zero in the local regime. We take $\omega = 1.05$ throughout this section as a numerical diagnostic of convergence behavior.

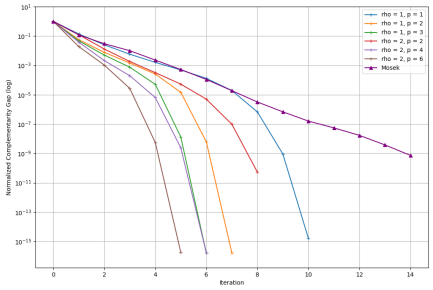
As before, we begin with the 3-elliptope problem. Previously, we demonstrated the promising performance obtained with $\rho = 2$. For several choices of p, Table 4 reports the convergence history of our `IPM`.



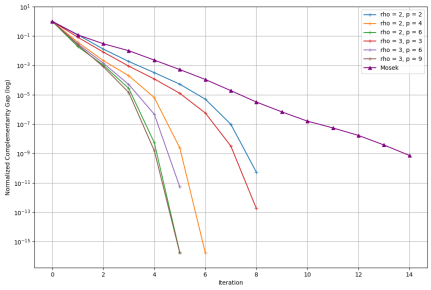
D-1: $\rho = 1$ vs 2



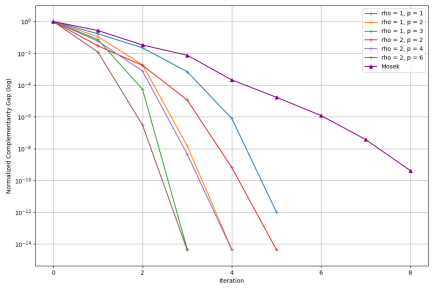
D-1: $\rho = 2$ vs 3



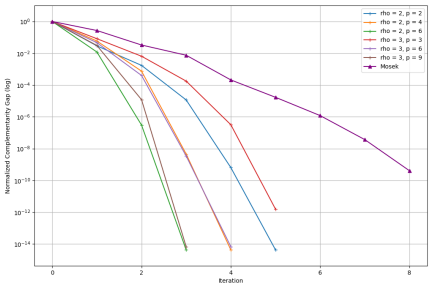
D-2: $\rho = 1$ vs 2



D-2: $\rho = 2$ vs 3



D-3: $\rho = 1$ vs 2



D-3: $\rho = 2$ vs 3

FIGURE 11. Convergence of random instances with dual non-degeneracy and strict complementarity

TABLE 4. Convergence history for the 3-elliptope problem (Figure 2)

	$(\rho, p) = (2, 2)$		$(\rho, p) = (2, 4)$		$(\rho, p) = (2, 8)$	
Iter. (k)	$\langle X_k, S_k \rangle / n$	Ratio	$\langle X_k, S_k \rangle / n$	Ratio	$\langle X_k, S_k \rangle / n$	Ratio
0	1.00E+00	N/A	1.00E+00	N/A	1.00E+00	N/A
1	1.49E-01	1.49E-01	2.58E-02	2.58E-02	1.00E-03	1.00E-03
2	1.14E-02	8.46E-02	1.08E-07	5.01E-06	7.51E-15	1.06E-11
3	3.09E-04	3.38E-02	6.29E-15	1.30E-07	-	-
4	1.73E-06	8.41E-03	-	-	-	-
5	8.48E-10	9.51E-04	-	-	-	-
6	1.17E-14	3.93E-05	-	-	-	-

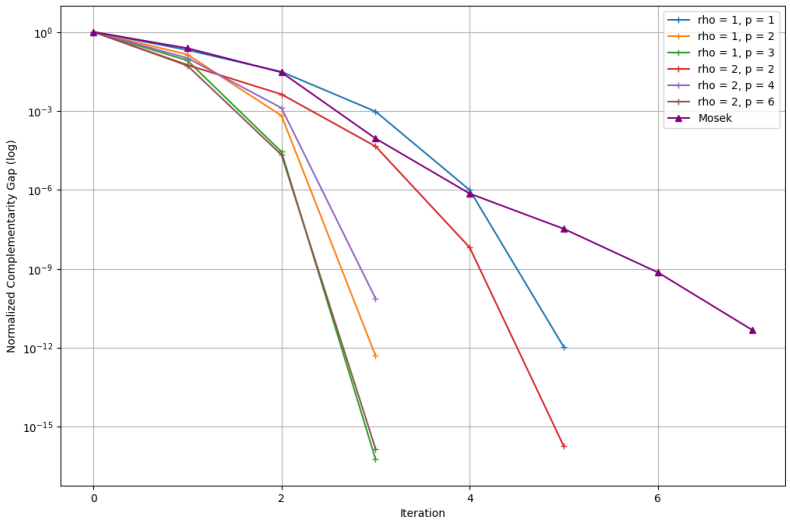


FIGURE 12. Convergence history on the benchmark problem (8.1)

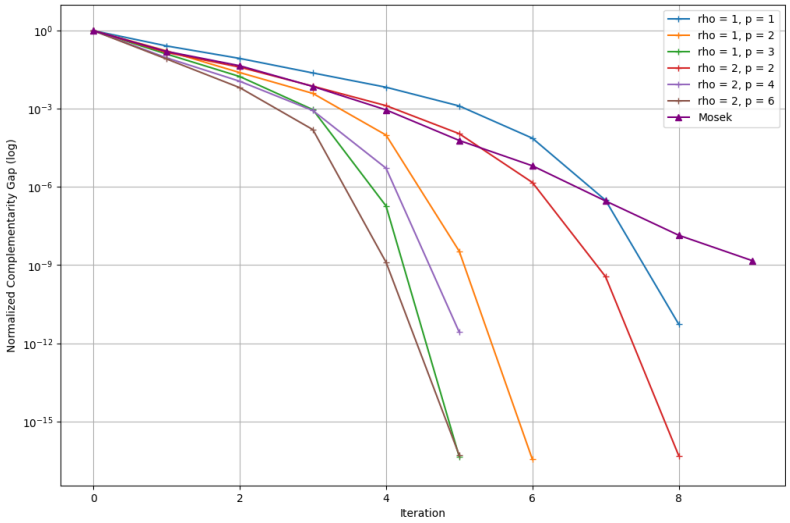


FIGURE 13. Convergence history on the random problem with strict complementarity

For all choices of p, the reported values of $\frac{\mu_k}{\mu_{k-1}^{1.05}}$ decrease rapidly toward zero. This provides numerical evidence of a Q-order exceeding 1.05 over the reported iterates. Moreover, larger values of p lead to a more rapid decrease in this ratio, consistent with the higher convergence orders predicted by Theorem 2.

Earlier, we noted that using $\rho = 3$ for the 3-elliptope problem may reduce the number of iterations, although it is not expected to preserve super-linear convergence near optimality. To investigate this behavior, Table 5 presents the convergence history obtained with $\rho = 3$. As the iterates approach optimality, the ratio eventually ceases to decrease, so the reported iterates do not provide evidence of a Q-order exceeding 1.05 in the asymptotic regime.

TABLE 5. Convergence history for the 3-elliptope problem (Figure 4)

	$(\rho,p)=(3,3)$		$(\rho,p)=(3,6)$		$(\rho,p)=(3,9)$	
Iter. (k)	$\langle X_k,S_k\rangle/n$	Ratio	$\langle X_k,S_k\rangle/n$	Ratio	$\langle X_k,S_k\rangle/n$	Ratio
0	1.00E+00	N/A	1.00E+00	N/A	1.00E+00	N/A
1	1.09E-01	1.09E-01	1.89E-02	1.89E-02	1.77E-03	1.77E-03
2	6.52E-03	6.72E-02	7.47E-06	4.81E-04	1.90E-07	1.47E-04
3	1.47E-04	2.90E-02	2.71E-09	6.54E-04	2.02E-11	2.31E-04
4	9.80E-07	1.04E-02	9.81E-13	9.72E-04	-	-
5	3.00E-09	6.11E-03	-	-	-	-
6	8.89E-12	7.92E-03	-	-	-	-

Next, we consider the random instance ND-5. Table 6 presents the convergence history of our `IPM` for this instance. After an initial transient phase, the ratios eventually decrease for all three choices of p, with a substantially faster decrease for the higher-order predictors. This behavior is consistent with the onset of the local super-linear regime.

TABLE 6. Convergence history for the random instance ND-5 (Figure 9)

	$(\rho,p)=(2,2)$		$(\rho,p)=(2,4)$		$(\rho,p)=(2,6)$	
Iter. (k)	$\langle X_k,S_k\rangle/n$	Ratio	$\langle X_k,S_k\rangle/n$	Ratio	$\langle X_k,S_k\rangle/n$	Ratio
0	1.00E+00	N/A	1.00E+00	N/A	1.00E+00	N/A
1	6.97E-02	6.97E-02	5.94E-02	5.94E-02	2.78E-02	2.78E-02
2	5.91E-03	9.69E-02	2.52E-03	4.89E-02	1.78E-03	7.64E-02
3	1.60E-03	3.50E-01	7.50E-04	4.02E-01	2.72E-04	2.10E-01
4	6.03E-04	5.20E-01	1.11E-04	2.13E-01	1.25E-06	6.94E-03
5	1.48E-04	3.55E-01	1.44E-06	2.04E-02	1.16E-14	1.84E-08
6	3.06E-05	3.22E-01	9.05E-12	1.23E-05	-	-
7	4.06E-06	2.23E-01	-	-	-	-
8	2.78E-07	1.28E-01	-	-	-	-
9	7.06E-09	5.40E-02	-	-	-	-
10	3.77E-11	1.37E-02	-	-	-	-

Finally, we consider one of the PDG problems and examine the convergence history of our `IPM` applied to its homogeneous self-dual embedding. Specifically, we revisit PDG Instance 3, whose embedded problem has dimensions $m_{HSD}=25$ and $n_{HSD}=30$. As in the previous experiments, we set $\rho=3$, which was observed to perform well for this instance. Table 7 shows that, for all three choices of p, the reported ratios eventually decrease toward zero. This provides numerical evidence of a Q-order exceeding 1.05 over the reported iterates, with the decrease becoming more pronounced as p increases.

TABLE 7. Convergence history for the embedding of PDG Instance 3 (Figure 10)

	$(\rho,p)=(3,3)$		$(\rho,p)=(3,6)$		$(\rho,p)=(3,9)$	
Iter. (k)	$\langle X_k,S_k\rangle/n$	Ratio	$\langle X_k,S_k\rangle/n$	Ratio	$\langle X_k,S_k\rangle/n$	Ratio
0	1.00E+00	N/A	1.00E+00	N/A	1.00E+00	N/A
1	2.19E-01	2.19E-01	1.95E-01	1.95E-01	1.54E-01	1.54E-01
2	6.06E-02	2.99E-01	3.44E-02	1.92E-01	2.18E-02	1.56E-01
3	9.76E-03	1.85E-01	9.07E-04	3.12E-02	3.35E-05	1.87E-03
4	7.66E-04	9.89E-02	1.59E-07	2.48E-04	4.20E-13	2.10E-08
5	2.88E-05	5.39E-02	1.86E-12	2.57E-05	-	-
6	4.94E-07	2.89E-02	-	-	-	-
7	3.10E-09	1.30E-02	-	-	-	-
8	6.09E-12	5.23E-03	-	-	-	-

We also monitored the number of corrector steps and the norms of the higher-order derivatives throughout these experiments. In most runs, a single full centering step was sufficient to return the iterate to the 10^{-1}-neighborhood. The main systematic exception was the 3-elliptope with $\rho = 3$. Since $W = I_n$ belongs to $\mathcal{W}_\beta$ and the ordinary central path is not analytic after the reparametrization $\mu \mapsto \mu^3$, this value of ρ is not admissible on $\mathcal{W}_\beta$. In this case, as the iterations progressed, two full centering steps were required. For ND-5, two centering steps were needed only at isolated iterations, while one full step sufficed otherwise.

The norms of the computed higher-order derivatives did not exhibit systematic growth as the iterates approached optimality. Although they were not necessarily monotone and could become relatively large at intermediate iterations, they generally decreased in the local regime. This behavior suggests that the scaling strategy of Section 8.2 is effective in controlling the higher-order systems, while quantitative uniform bounds on these derivatives remain an interesting subject for further study.

9. Concluding remarks and future research

In this paper, we developed a higher-order feasible predictor-corrector `IPM` for `SDO` based on the `AHO` search directions. The method exploits the semi-algebraic structure of weighted central paths and their ramified analytic reparametrizations near the limiting point. This makes it possible to construct higher-order predictors even when strict complementarity fails. Under a uniform predictor regularity condition controlling these local models as the weight varies, we prove super-linear convergence of the normalized complementarity gap without assuming strict complementarity. We also establish super-linear convergence of the vanishing blocks of the primal-dual iterates, derive a quantitative reduction estimate throughout the predictor-regularity region, and obtain a doubly logarithmic outer iteration bound for the resulting predictor-corrector `IPM`.

We evaluated the proposed algorithm on five problem sets with diverse characteristics, including four that violate the strict complementarity condition. For suitable choices of (ρ, p), the numerical results exhibit behavior consistent with super-linear convergence across all test sets, including instances without strict complementarity.

Local convergence and global iteration complexity. We establish super-linear convergence of the complementarity gap and an explicit quantitative reduction estimate when $p \geq \rho$. It would be of considerable interest to further investigate the convergence behavior of Algorithm 1 in the regime $p < \rho$. Another important question is whether the higher-order method itself admits a polynomial global iteration complexity bound in the standard input-size model. Our present global estimate depends on the predictor regularity constant c; establishing polynomially controlled bounds for the normalized higher-order `AHO` systems may provide a route toward such a result.

Corrector complexity and higher-order conditioning. In all of our numerical experiments, the centering phase required only full `AHO` steps, and in most iterations a single full step was sufficient. It would be interesting to identify conditions under which one full centering step is guaranteed locally, or more generally to obtain a uniform bound on the number of corrector steps required after a higher-order predictor. The experiments also show that the norms of higher-order derivatives can vary substantially with the iteration, the weight, and the derivative order. Developing quantitative uniform bounds for these derivatives, and relating such bounds to the predictor regularity condition, could lead both to sharper convergence results and to more robust numerical implementations.

Approximation of ρ. We did not address the problem of estimating an admissible reparametrization exponent ρ for the family of weighted central paths. For the ordinary central path, an exact symbolic procedure for computing a feasible reparametrization exponent was developed in [10]. In our numerical experiments, we instead use heuristic candidate values motivated by [25]. Developing reliable numerical techniques for estimating and certifying an admissible exponent for the weighted family remains an interesting direction for future research.

Acknowledgments

The authors were supported by NSF grants CCF-2128527, CCF-1910441 and CCF-2128702 during the preparation of the initial version of this paper.

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Program.*, 95(1):3–51, 2003. 2
[2] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Complementarity and nondegeneracy in semidefinite programming. *Math. Program.*, 77(1):111–128, 1997. 41, 46
[3] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM J. Optim.*, 8(3):746–768, 1998. 4, 13, 20
[4] X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig. Log-barrier interior point methods are not strongly polynomial. *SIAM J. Appl. Algebra Geom.*, 2(1):140–178, 2018. 15
[5] E. D. Andersen, J. Gondzio, C. Mészáros, and X. Xu. Implementation of interior-point methods for large scale linear programs. In *Interior Point Methods of Mathematical Programming*, volume 5 of *Appl. Optim.*, pages 189–252. Kluwer Acad. Publ., Dordrecht, 1996. 39
[6] M. Artin. *Algebra*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1991. 7, 8
[7] S. Basu. Algorithms in real algebraic geometry: a survey. In *Real Algebraic Geometry*, volume 51 of *Panor. Synthèses*, pages 107–153. Soc. Math. France, Paris, 2017. 15
[8] S. Basu and A. Mohammad-Nezhad. On the central path of semidefinite optimization: degree and worst-case convergence rate. *SIAM J. Appl. Algebra Geom.*, 6(2):299–318, 2022. 3, 7, 15
[9] S. Basu and A. Mohammad-Nezhad. Improved effective Łojasiewicz inequality and applications. *Forum Math. Sigma*, 12:Paper No. e115, 27, 2024. 15
[10] S. Basu and A. Mohammad-Nezhad. On the complexity of analyticity in semi-definite optimization. *Adv. in Appl. Math.*, 156:Paper No. 102670, 35, 2024. 3, 7, 10, 12, 14, 15, 51
[11] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, second edition, 2006. 6, 7, 8, 9, 15, 16, 17, 18
[12] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation*. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp. 15
[13] J. Bochnak, M. Coste, and M.-F. c. Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors. 7, 9, 10
[14] E. de Klerk. *Aspects of Semidefinite Programming: Interior Point algorithms and Selected Applications*, volume 65 of *Applied Optimization*. Kluwer Academic Publishers, Dordrecht, 2002. 3, 10, 15, 16, 33, 36, 38, 40
[15] E. de Klerk, C. Roos, and T. Terlaky. Infeasible-start semidefinite programming algorithms via self-dual embeddings. In *Topics in Semidefinite and Interior-point Methods*, volume 18 of *Fields Inst. Commun.*, pages 215–236. Amer. Math. Soc., Providence, RI, 1998. 38
[16] J. A. De Loera, B. Sturmfels, and C. Vinzant. The central curve in linear programming. *Found. Comput. Math.*, 12(4):509–540, 2012. 15

[17] J.-P. Dedieu, G. Malajovich, and M. Shub. On the curvature of the central path of linear programming theory. *Found. Comput. Math.*, 5(2):145–171, 2005. 15

[18] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, volume 4 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990. 2

[19] M. R. Garey and D. S. Johnson. *Computers and intractability*. A Series of Books in the Mathematical Sciences. W. H. Freeman and Co., San Francisco, CA, 1979. A guide to the theory of NP-completeness. 15

[20] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programming. *SIAM J. Optim.*, 8(4):871–886, 1998. 10, 12

[21] P. Griffiths and J. Harris. *Principles of algebraic geometry*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1994. Reprint of the 1978 original. 7

[22] O. Güler. Limiting behavior of weighted central paths in linear programming. *Math. Program.*, 65(3):347–363, 1994. 12

[23] M. Halická. Analyticity of the central path at the boundary point in semidefinite programming. *European J. Oper. Res.*, 143(2):311–324, 2002. 10, 12

[24] M. Halická, E. de Klerk, and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM J. Optim.*, 12(4):1090–1099, 2002. 15

[25] J. D. Hauenstein, A. C. Liddell, Jr., S. McPherson, and Y. Zhang. Numerical algebraic geometry and semidefinite programming. *Results Appl. Math.*, 11:100166, 20, 2021. 15, 38, 46, 51

[26] D. Henrion, S. Naldi, and M. Safey El Din. Exact algorithms for linear matrix inequalities. *SIAM J. Optim.*, 26(4):2512–2539, 2016. 15

[27] D. Henrion, S. Naldi, and M. Safey El Din. Exact algorithms for semidefinite programs with degenerate feasible set. *J. Symbolic Comput.*, 104:942–959, 2021. 15

[28] S. Hoşten, I. Shankar, and A. Torres. The degree of the central curve in semidefinite, linear, and quadratic programming. *Matematiche (Catania)*, 76(2):483–499, 2021. 15

[29] T. Illés, J. Peng, C. Roos, and T. Terlaky. A strongly polynomial rounding procedure yielding a maximally complementary solution for $P_*(\kappa)$ linear complementarity problems. *SIAM J. Optim.*, 11(2):320–340, 2000. 12

[30] J. Ji, F. A. Potra, and S. Huang. Predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence. *J. Optim. Theory Appl.*, 85(1):187–199, 1995. 13

[31] J. Ji, F. A. Potra, and R. Sheng. On a general class of interior-point algorithms for semidefinite programming with polynomial complexity and superlinear convergence. *Methods Appl. Anal.*, 6(4):549–569, 1999. 13

[32] J. Ji, F. A. Potra, and R. Sheng. On the local convergence of a predictor-corrector method for semidefinite programming. *SIAM J. Optim.*, 10(1):195–210, 1999. 13

[33] J. Ji, F. A. Potra, R. A. Tapia, and Y. Zhang. An interior-point method with polynomial complexity and superlinear convergence for linear complementarity problems. Technical report, Department of Computational and Applied Mathematics, Rice University, Houston, Texas (TR91-23), 1991. 13

[34] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A unified approach to interior point algorithms for linear complementarity problems*, volume 538 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1991. 15

[35] M. Kojima, M. Shida, and S. Shindoh. Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. *Math. Program.*, 80(2):129–160, 1998. 13, 14

[36] M. Kojima, M. Shida, and S. Shindoh. A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberly-Overton search direction. *SIAM J. Optim.*, 9(2):444–465, 1999. 13, 14, 22

[37] M. Kojima, S. Shindoh, and S. Hara. Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM J. Optim.*, 7(1):86–125, 1997. 20

[38] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002. 7

[39] J. B. Lasserre. *An Introduction to Polynomial and Semi-algebraic Optimization*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2015. 2

[40] Z. Lu and R. D. C. Monteiro. Error bounds and limiting behavior of weighted paths associated with the SDP map $X^{1/2}SX^{1/2}$. *SIAM J. Optim.*, 15(2):348–374, 2005. 13

[41] Z. Lu and R. D. C. Monteiro. A note on the local convergence of a predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem based on the Alizadeh-Haeberly-Overton search direction. *SIAM J. Optim.*, 15(4):1147–1154, 2005. 13, 14

[42] Z. Lu and R. D. C. Monteiro. Limiting behavior of the Alizadeh-Haeberly-Overton weighted paths in semidefinite programming. *Optim. Methods Softw.*, 22(5):849–870, 2007. 13, 14

[43] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming. *SIAM J. Optim.*, 8(1):59–81, 1998. 13

[44] S. Mizuno. A superlinearly convergent infeasible-interior-point algorithm for geometrical LCPs without a strictly complementary condition. *Math. Oper. Res.*, 21(2):382–400, 1996. 13

[45] A. Mohammad-Nezhad and T. Terlaky. On the identification of the optimal partition for semidefinite optimization. *INFOR Inf. Syst. Oper. Res.*, 58(2):225–263, 2020. 6, 12, 33

[46] M. Mohammadisiahroudi, R. Fakhimi, B. Augustino, and T. Terlaky. Generating linear, semidefinite, and second-order cone optimization problems for numerical experiments. *Optim. Methods Softw.*, 39(4):725–755, 2024. 40, 46

[47] R. D. C. Monteiro. Polynomial convergence of primal-dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions. *SIAM J. Optim.*, 8(3):797–812, 1998. 5

[48] R. D. C. Monteiro and J.-S. Pang. On two interior-point mappings for nonlinear semidefinite complementarity problems. *Math. Oper. Res.*, 23(1):39–60, 1998. 13, 16

[49] R. D. C. Monteiro and T. Tsuchiya. Limiting behavior of the derivatives of certain trajectories associated with a monotone horizontal linear complementarity problem. *Math. Oper. Res.*, 21(4):793–814, 1996. 12

[50] R. D. C. Monteiro and S. J. Wright. Local convergence of interior-point algorithms for degenerate monotone LCP. *Comput. Optim. Appl.*, 3(2):131–155, 1994. 13

[51] R. D. C. Monteiro and P. R. Zanjácomo. A note on the existence of the Alizadeh-Haeberly-Overton direction for semidefinite programming. *Math. Program.*, 78(3):393–396, 1997. 20, 21, 22

[52] R. D. C. Monteiro and P. R. Zanjácomo. General interior-point maps and existence of weighted paths for nonlinear semidefinite complementarity problems. *Math. Oper. Res.*, 25(3):381–399, 2000. 13

[53] R. D. C. Monteiro and Y. Zhang. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Math. Programming*, 81(3):281–299, 1998. 20, 21

[54] Y. Nesterov and A. Nemirovskii. *Interior-point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. 2

[55] Y. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22(1):1–42, 1997. 20

[56] Y. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.*, 8(2):324–364, 1998. 20

[57] G. Pataki. On positive duality gaps in semidefinite programming. *arXiv preprint arXiv:1812.11796*, 2018. 43

[58] F. A. Potra. Q-superlinear convergence of the iterates in primal-dual interior-point methods. *Math. Program.*, 91(1):99–115, 2001. 13

[59] F. A. Potra. A superlinearly convergent predictor-corrector method for degenerate LCP in a wide neighborhood of the central path with $O(\sqrt{n}L)$-iteration complexity. *Math. Program.*, 100(2):317–337, 2004. 13

[60] F. A. Potra and R. Sheng. A path following method for LCP with superlinearly convergent iteration sequence. *Ann. Oper. Res.*, 81:97–114, 1998. 13

[61] F. A. Potra and R. Sheng. Superlinear convergence of interior-point algorithms for semidefinite programming. *J. Optim. Theory Appl.*, 99(1):103–119, 1998. 13, 14

[62] F. A. Potra and R. Sheng. Superlinearly convergent infeasible-interior-point algorithm for degenerate LCP. *J. Optim. Theory Appl.*, 97(2):249–269, 1998. 13

[63] F. A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. *SIAM J. Optim.*, 8(4):1007–1028, 1998. 13

[64] F. A. Potra and R. Sheng. Superlinear convergence of a predictor-corrector method for semidefinite programming without shrinking central path neighborhood. *Bull. Math. Soc. Sci. Math. Roumanie (N.S.)*, 43(91)(2):107–124, 2000. 13

[65] F. A. Potra and J. Stoer. On a class of superlinearly convergent polynomial time interior point methods for sufficient LCP. *SIAM J. Optim.*, 20(3):1333–1363, 2010. 13, 14

[66] M. Preiß and J. Stoer. Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems. *Math. Program.*, 99(3):499–520, 2004. 12, 13, 32

[67] M. V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Math. Program.*, 77(2):129–162, 1997. 15

[68] C.-K. Sim. Superlinear convergence of an infeasible predictor-corrector path-following interior point algorithm for a semidefinite linear complementarity problem using the Helmberg-Kojima-Monteiro direction. *SIAM J. Optim.*, 21(1):102–126, 2011. 13

[69] C.-K. Sim. Superlinear convergence of an interior point algorithm on linear semi-definite feasibility problems. *Optim. Methods Softw.*, 40(2):287–307, 2025. 14

[70] J. Stoer and M. Wechs. Infeasible-interior-point paths for sufficient linear complementarity problems and their analyticity. *Math. Program.*, 83(3):407–423, 1998. 12, 14

[71] J. Stoer and M. Wechs. On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems. *Numer. Math.*, 81(4):631–645, 1999. 12

[72] J. Stoer, M. Wechs, and S. Mizuno. High order infeasible-interior-point methods for solving sufficient linear complementarity problems. *Math. Oper. Res.*, 23(4):832–862, 1998. 12, 14, 22

[73] J. F. Sturm. Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists. *Math. Oper. Res.*, 24(1):72–94, 1999. 13

[74] J. F. Sturm and S. Zhang. On weighted centers for semidefinite programming. *European J. Oper. Res.*, 126(2):391–407, 2000. 13

[75] A. Tarski. A decision method for elementary algebra and geometry. In *Quantifier Elimination and Cylindrical Algebraic Decomposition (Linz, 1993)*, Texts Monogr. Symbol. Comput., pages 24–84. Springer, Vienna, 1998. 8

[76] M. J. Todd. Semidefinite optimization. *Acta Numer.*, 10:515–560, 2001. 2

[77] L. Tunçel and H. Wolkowicz. Strong duality and minimal representations for cone optimization. *Comput. Optim. Appl.*, 53(2):619–648, 2012. 42

[78] R. J. Walker. *Algebraic Curves*. Springer-Verlag, New York-Heidelberg, 1978. Reprint of the 1950 edition. 6, 7

Department of Statistics and OR, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A.

Email address: alimn@unc.edu

Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015, U.S.A.

Email address: pos220@lehigh.edu

Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015, U.S.A.

Email address: terlaky@lehigh.edu