Communications in Combinatorics and Optimization
Vol. ??, No. ?? (????), pp. 1-38
DOI:10.22049/CCO.2023.28256.1492

CCO
Commun. Comb. Optim.

Research Article

Algebraic-based primal interior-point algorithms for stochastic infinity norm optimization

Baha Alzalg[1,*], **Karima Tamsaouete**[1,2,†]

[1]Department of Mathematics, The University of Jordan, Amman, Jordan 11942
*b.alzalg@ju.edu.jo
†kry9180336@ju.edu.jo

[2]Department of Mathematics, M'Hamed Bougara University of Boumerdés, Algeria 22038

Received: 22 January 2023; Accepted: 28 May 2023
Published Online: 8 June 2023

Abstract: We study the two-stage stochastic infinity norm optimization problem with recourse based on the Jordan algebra. First, we explore and develop the Jordan algebra structure of the infinity norm cone, and utilize it to compute the derivatives of the barrier recourse functions. Then, we prove that the barrier recourse functions and the composite barrier functions for this optimization problem are self-concordant families with reference to barrier parameters. These findings are used to develop interior-point algorithms based on primal decomposition for this class of stochastic programming problems. Our complexity results for the short- and long-step algorithms show that the dominant complexity terms are linear in the rank of the underlying cone. Despite the asymmetry of the infinity norm cone, we also show that the obtained complexity results match (in terms of rank) the best known results in the literature for other well-studied stochastic symmetric cone programs. Finally, we demonstrate the efficiency of the proposed algorithm by presenting some numerical experiments on both stochastic uniform facility location problems and randomly-generated problems.

Keywords: Jordan algebras, Infinity norm optimization, Stochastic programming, Interior-point methods, Polynomial-time complexity

AMS Subject classification: 17C05, 90C15, 90C25, 90C51

1. Introduction

The core aim of this paper is to study, based on a Jordan algebraic treatment, the two-stage stochastic infinity norm programming (SINP for short) problem with K

* *Corresponding Author*

scenarios:

$$\begin{array}{llcll} \min & c^{\mathsf{T}} x + \sum_{k=1}^{K} \bar{\rho}^{(k)}(x) & \text{where} & \bar{\rho}^{(k)}(x) \triangleq \min & {d^{(k)}}^{\mathsf{T}} y^{(k)} \\ \text{s.t.} & Ax = b, & & \text{s.t.} & W^{(k)} y^{(k)} = q^{(k)} + T^{(k)} x, \; k = 1,2,\ldots,K, \\ & x \in \mathcal{I}^n; & & & y^{(k)} \in \mathcal{I}_+^m, \; k = 1,2,\ldots,K. \end{array} \tag{1}$$

Here, $\mathcal{I}^n$ is the nth-dimensional infinity norm cone of the first-stage decision variable $x \in \mathbb{R}^n$, and $\mathcal{I}_+^m$ is the mth-dimensional infinity norm cone of the second-stage decision variable $y^{(k)} \in \mathbb{R}^m$ for $k = 1,2,\ldots,K$. The function $\bar{\rho}^{(k)}(x)$ is called *the recourse function*. We also assumed that $A, W^{(k)}$ and $d^{(k)}$, $k = 1,2,\ldots,K$, have already absorbed the scenario probabilities.

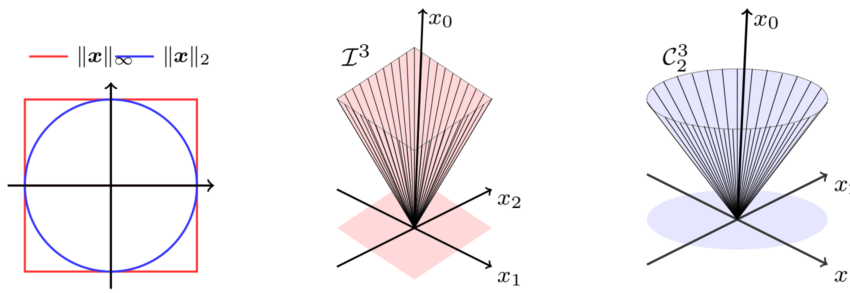


Figure 1: Graphs of the infinity norm cone (in light red) and the second-order cone (in light blue) in $\mathbb{R}^3$. The picture to the left shows the graphs of two-dimensional unit spheres in infinity and 2-norms.

Many authors studied deterministic conic optimization problems including those over the infinity norm cone (see, for example, [12, 14, 15, 17–19, 22, 23]). Despite the need for studying the optimization problems in stochastic environments, there are no algorithmic methods to specifically solve infinity norm optimization problems in the stochastic setting. Taking this literature gap into account, we study in this paper two-stage stochastic optimization problems over the infinity norm cone (also called the infinity-order cone), which is defined as

$$\mathcal{I}^n \triangleq \left\{ x \triangleq \begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} \in \mathbb{R} \times \mathbb{R}^{n-1} : x_0 \geq \|\bar{x}\|_\infty \right\}, \quad \text{where } \|\bar{x}\|_\infty \triangleq \max_{1 \leq i \leq n-1} |x_i|,$$

and $\bar{x} \triangleq (x_1, x_2, \ldots, x_{n-1})^{\mathsf{T}} \in \mathbb{R}^{n-1}$. The dual cone of $\mathcal{I}^n$ is the nth-dimensional first-order cone, which is defined as

$$\mathcal{C}_1^n \triangleq \left\{ x \triangleq \begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} \in \mathbb{R} \times \mathbb{R}^{n-1} : x_0 \geq \|\bar{x}\|_1 \right\}, \quad \text{where } \|x\|_1 = \sum_{i=1}^{n-1} |x_i|.$$

The cone $\mathcal{C}_1^n$ is a special case of the p-th order cone of order n, which is defined as

$$\mathcal{C}_p^n \triangleq \left\{ x \triangleq \begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} \in \mathbb{R} \times \mathbb{R}^{n-1} : x_0 \geq \|\bar{x}\|_p \right\}, \quad p \geq 1, \quad \text{where } \|\bar{x}\|_p \triangleq \left(\sum_{i=1}^{n-1} |x_i|^p \right)^{1/p}.$$

Note that when $p = 2$, $\mathcal{C}_p^n$ reduces to the well-studied second-order cone $\mathcal{C}_2^n$. Like any p-th order cone $\mathcal{C}_p^n$, the infinity norm cone $\mathcal{I}^n$ is solid (i.e., its interior, $\text{int}(\mathcal{I}^n)$, is nonempty), pointed (i.e., $\mathcal{I}^n \cap -\mathcal{I}^n = \{0\}$), closed convex cone in $\mathbb{R}^n$ (see Figure 1). Unlike the second-order cone $\mathcal{C}_2^n$, the infinity norm cone is non-self-dual and hence is asymmetric.

Benders' decomposition has long been employed in the development of solution methodologies for both two-stage stochastic linear and nonlinear programs [1–3, 5, 6, 8–10, 13, 16, 20, 24, 25]. The L-shaped method, for example, uses this strategy to construct cuts by taking into account subgradients of the recourse function. Later on in the last two decades, decomposition interior-point algorithms have been developed to find solution methodologies for different classes of two-stage stochastic conic programs. These algorithms can be summarized as follows. Zhao [24] derived logarithmic barrier interior-point methods for solving two-satge stochastic linear programming using Benders' decomposition. Alzalg [1] (see also [2, 5, 9]) derived decomposition-based interior-point methods for two-stage stochastic second-order cone programming by generalizing the work of Zhao [24]. Mehrotra and Özevin [20] (see also Ariyawansa and Zhu [13]) generalized the work of Zhao [24] for two-stage stochastic semidefinite programming. The work of Alzalg and Ariyawansa [8] generalizes the results in [1, 20, 24] to derive logarithmic barrier decomposition-based interior-point algorithms for stochastic programming on all symmetric cones. Finally, Chen and Mehrotra [16] (see also Zhao [25]) derived a prototype interior-point algorithm for stochastic convex programming.

To analyze the proposed algorithm, we develop a novel Jordan algebra associated with the undelying cone and discuss its characteristics in great detail. We exploit this algebra to derive a logarithmic barrier primal interior-point algorithm for the two-stage stochastic infinity norm programming (SINP) problem via a utilization of the work of Chen and Mehrotra [16] for stochastic convex programming. While the explicit expressions for the derivatives of the barrier function in [16] are not available, the merit of this work is sufficiently evinced by explicitly computing such derivatives. These derivatives are used to prove the self-concordance properties (see Nesterov and Nemirovskii [21]) of the barrier recourse function that guarantee nice performance of Newton's method used for the proposed algorithms. These findings are used to develop short- and long-step interior-point decomposition algorithms for the two-stage SINP problem.

We will see that, for a two-stage stochastic program with K number of realizations over infinity norm cones with ranks $\mathcal{O}(n + Km)$, the short-step algorithm restores the proximity condition in one step, while the long-step algorithm may perform several inner iterations. Let ϵ be the desired accuracy of the final solution, we will also see that we need at most $\mathcal{O}((n + Km)^{1/2}\ln(\mu^0/\epsilon))$ outer iterations in the short-step algorithm to follow the central path from a starting value of the barrier parameter μ^0 to the terminating value ϵ, and we need at most $\mathcal{O}((n+Km)\ln(\mu^0/\epsilon))$ outer iterations in the long-step algorithm for this recentering. We will see that the above complexity results agree in terms of rank the best known results in the literature for two-stage stochastic linear programming in [24], two-stage stochastic second-order cone programming in

[1], and two-stage stochastic semidefinite programming in [20]. This agreement is in spite of the fact that the infinity norm cone is asymmetric.
The following is how the paper is structured. In Section 2, we study and establish algebraic structure of the infinity norm cone. In Section 3, we introduce the barrier function associated with the infinity norm cone, compute its derivatives, and prove its self-concordance complexity. Section 4 is devoted to explicitly computing the derivatives of the composite barrier function and establishing its self-concordance analytical properties. In Section 5, we state path-following interior-point algorithms for solving our problem and present their complexity results. We present numerical experiments to show the efficiency of the proposed algorithms in Section 6. Sections 7 draws some closing conclusions. The proofs of the complexity results are given in Appendix A.

2. The algebraic structure of the cone

In this section, we dive into the algebraic structure of the infinity norm cone. We will see that the algebra that we construct and associate with this cone is a Jordan algebra. To review some preliminaries of Jordan algebras, see [4, Section 2] and [11, Appendix 2].
Let $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors in $\mathbb{R}^n$, we write

$$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = (\boldsymbol{x}^\mathsf{T}, \boldsymbol{y}^\mathsf{T})^\mathsf{T} = (\boldsymbol{x}; \boldsymbol{y}),$$

where "," is used to adjoin vectors and matrices in a row, and ";" is used to adjoin them in a column. For each vector $\boldsymbol{x} \in \mathbb{R}^n$ indexed from 0, we denote $\bar{\boldsymbol{x}}$ the sub-vector comprising entries 1 through $n-1$; therefore $\boldsymbol{x} = (x_0; \bar{\boldsymbol{x}}) \in \mathbb{R} \times \mathbb{R}^{n-1}$.
By $\mathcal{E}^n$ we mean the nth-dimensional real vector space $\mathbb{R} \times \mathbb{R}^{n-1}$. We use I_n to denote the identity matrix of order n and $I_n^{\langle i \rangle}$ to denote a matrix of order n such that all its entries are zeros except the (i,i)th-entry which is a one. By $\boldsymbol{e}_n \triangleq (1; 0)$ we mean the *identity vector* of $\mathcal{E}^n$, by $\boldsymbol{u}_n^{\langle i \rangle}$ we mean the vector in $\mathcal{E}^n$ such that all its entries are zeros except the ith-entry, which is a one for $i = 1, 2, \ldots, n-1$, by O we mean the zero matrix of appropriate size, and by $\mathbf{0}$ we mean the zero vector of appropriate dimension.

We introduce the following matrices in $\mathbb{R}^{n \times n}$:

$$J_n \triangleq \begin{bmatrix} \frac{1}{n-1} & \mathbf{0}^\mathsf{T} \\ \mathbf{0} & I_{n-1} \end{bmatrix}, J_n^{\langle i \rangle} \triangleq \begin{bmatrix} \frac{1}{n-1} & \mathbf{0}^\mathsf{T} \\ \mathbf{0} & I_{n-1}^{\langle i \rangle} \end{bmatrix}, R_n \triangleq \begin{bmatrix} 1 & \mathbf{0}^\mathsf{T} \\ \mathbf{0} & -I_{n-1} \end{bmatrix}, R_n^{\langle i \rangle} \triangleq \begin{bmatrix} 1 & \mathbf{0}^\mathsf{T} \\ \mathbf{0} & -I_{n-1}^{\langle i \rangle} \end{bmatrix}, i = 1, 2, \ldots, n-1.$$

Note that $I_n = \sum_{i=1}^{n-1} J_n^{\langle i \rangle}$ and $R_n = J_n \sum_{i=1}^{n-1} R_n^{\langle i \rangle}$.
For each $\boldsymbol{x} \in \mathcal{E}^n$, we associate the vectors $\boldsymbol{x}^{\langle 1 \rangle}, \boldsymbol{x}^{\langle 2 \rangle}, \ldots, \boldsymbol{x}^{\langle n-1 \rangle}$, where each vector

$\boldsymbol{x}^{\langle i\rangle} \in \mathcal{E}^n$ is defined as

$$\boldsymbol{x}^{\langle i\rangle} \triangleq \begin{bmatrix} x_0 \\ \mathbf{0} \\ x_i \\ \mathbf{0} \end{bmatrix} \begin{matrix} \\ \longleftarrow \text{the zero vector in } \mathbb{R}^{i-1}, \\ \longleftarrow (i+1)\text{th-entry}, \\ \longleftarrow \text{the zero vector in } \mathbb{R}^{n-i-1}, \end{matrix}$$

for $i = 1, 2, \ldots, n-1$. Note that each vector $x \in \mathcal{E}^n$ can be uniquely written as

$$\boldsymbol{x} = J_n \sum_{i=1}^{n-1} \boldsymbol{x}^{\langle i\rangle} = \sum_{i=1}^{n-1} J_n^{\langle i\rangle} \boldsymbol{x}^{\langle i\rangle}.$$

The *spectral decomposition* of each vector $\boldsymbol{x}^{\langle i\rangle} \in \mathcal{E}^n$ is defined as

$$\boldsymbol{x}^{\langle i\rangle} = \underbrace{(x_0 + x_i)}_{\lambda^+(\boldsymbol{x}^{\langle i\rangle})} \underbrace{\left(\frac{1}{2}\right) \begin{bmatrix} 1 \\ \boldsymbol{u}_{n-1}^{\langle i\rangle} \end{bmatrix}}_{\boldsymbol{c}^+(\boldsymbol{x}^{\langle i\rangle})} + \underbrace{(x_0 - x_i)}_{\lambda^-(\boldsymbol{x}^{\langle i\rangle})} \underbrace{\left(\frac{1}{2}\right) \begin{bmatrix} 1 \\ -\boldsymbol{u}_{n-1}^{\langle i\rangle} \end{bmatrix}}_{\boldsymbol{c}^-(\boldsymbol{x}^{\langle i\rangle})}.$$

We call $\lambda^{\mp}(\boldsymbol{x}^{\langle i\rangle})$ and $\boldsymbol{c}^{\mp}(\boldsymbol{x}^{\langle i\rangle})$ the *eigenvalues* and *eigenvectors* of $\boldsymbol{x}^{\langle i\rangle}$, respectively. The *trace* and *determinant* of $\boldsymbol{x}^{\langle i\rangle}$ are defined respectively as

$$\operatorname{trace}\left(\boldsymbol{x}^{\langle i\rangle}\right) \triangleq \lambda^+\left(\boldsymbol{x}^{\langle i\rangle}\right) + \lambda^-\left(\boldsymbol{x}^{\langle i\rangle}\right) = 2x_0, \quad \text{and} \quad \det\left(\boldsymbol{x}^{\langle i\rangle}\right) \triangleq \lambda^+\left(\boldsymbol{x}^{\langle i\rangle}\right)\lambda^-\left(\boldsymbol{x}^{\langle i\rangle}\right) = x_0^2 - x_i^2.$$

Note that any $\boldsymbol{x} \in \mathcal{E}^n$ can be decomposed as

$$\boldsymbol{x} = \begin{bmatrix} x_0 \\ \bar{\boldsymbol{x}} \end{bmatrix} = J_n \sum_{i=1}^{n-1} \left(\overbrace{(x_0 + x_i)}^{\lambda^+(\boldsymbol{x}^{\langle i\rangle})} \overbrace{\left(\frac{1}{2}\right) \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}^{\boldsymbol{c}^+(\boldsymbol{x}^{\langle i\rangle})} + \overbrace{(x_0 - x_i)}^{\lambda^-(\boldsymbol{x}^{\langle i\rangle})} \overbrace{\left(\frac{1}{2}\right) \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}^{\boldsymbol{c}^-(\boldsymbol{x}^{\langle i\rangle})} \right) \longleftarrow \text{the } (i+1)\text{th-entry}.$$

This leads us to define the *spectral decomposition* of each vector $\boldsymbol{x} \in \mathcal{E}^n$ as

$$x = J_n \sum_{i=1}^{n-1} \left(\underbrace{(x_0 + x_i)}_{\lambda_i^+(\boldsymbol{x})} \underbrace{\left(\frac{1}{2}\right) \begin{bmatrix} 1 \\ \boldsymbol{u}_{n-1}^{\langle i\rangle} \end{bmatrix}}_{\boldsymbol{c}_i^+(\boldsymbol{x})} + \underbrace{(x_0 - x_i)}_{\lambda_i^-(\boldsymbol{x})} \underbrace{\left(\frac{1}{2}\right) \begin{bmatrix} 1 \\ -\boldsymbol{u}_{n-1}^{\langle i\rangle} \end{bmatrix}}_{\boldsymbol{c}_i^-(\boldsymbol{x})} \right).$$

We call $\lambda_i^{\mp}(\boldsymbol{x})$ and $\boldsymbol{c}_i^{\mp}(\boldsymbol{x})$ the *eigenvalues* and *eigenvectors* of $\boldsymbol{x}$, respectively. Note that $\lambda_i^{\mp}(\boldsymbol{x}) = \lambda^{\mp}(\boldsymbol{x}^{\langle i \rangle})$ and that $\boldsymbol{c}_i^{\mp}(\boldsymbol{x}) = \boldsymbol{c}^{\mp}(\boldsymbol{x}^{\langle i \rangle})$ for all $i = 1, 2, \ldots, n-1$. We also call $\mathsf{rk}(\mathcal{I}^n) \triangleq 2(n-1)$ the *rank* of the cone $\mathcal{I}^n$.
The *determinant* and *trace* of $\boldsymbol{x}$ are defined as

$$\det(\boldsymbol{x}) \triangleq \prod_{i=1}^{n-1} \lambda_i^{\mp}(\boldsymbol{x}) = \prod_{i=1}^{n-1} \det\left(\boldsymbol{x}^{\langle i \rangle}\right) = \prod_{i=1}^{n-1} \left(x_0^2 - x_i^2\right),$$

and

$$\operatorname{trace}(\boldsymbol{x}) \triangleq \sum_{i=1}^{n-1} \lambda_i^{\mp}(\boldsymbol{x}) = \sum_{i=1}^{n-1} \operatorname{trace}\left(\boldsymbol{x}^{\langle i \rangle}\right) = 2(n-1)x_0.$$

The *square* of $\boldsymbol{x} \in \mathcal{E}^n$ is defined as $\boldsymbol{x}^2 \triangleq J_n \sum_{i=1}^{n-1} \boldsymbol{x}^{\langle i \rangle^2}$, where $\boldsymbol{x}^{\langle i \rangle^2} \triangleq (\lambda^{+}(\boldsymbol{x}^{\langle i \rangle}))^2 \boldsymbol{c}^{+}(x^{\langle i \rangle}) + (\lambda^{-}(\boldsymbol{x}^{\langle i \rangle}))^2 \boldsymbol{c}^{-}(\boldsymbol{x}^{\langle i \rangle})$. It can be seen that

$$\boldsymbol{x}^2 = J_n \begin{bmatrix} (n-1)x_0^2 + \sum_{i=1}^{n-1} x_i^2 \\ 2x_0x_1 \\ 2x_0x_2 \\ \vdots \\ 2x_0x_{n-1} \end{bmatrix} = \mathsf{Arw}(\boldsymbol{x})\boldsymbol{x},$$

where $\mathsf{Arw}(\boldsymbol{x})$ denotes the *arrow-shaped matrix of* $\boldsymbol{x}$ defined as

$$\mathsf{Arw}(\boldsymbol{x}) \triangleq J_n \begin{bmatrix} (n-1)x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_0 & 0 & \cdots & 0 \\ x_2 & 0 & x_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & 0 & 0 & 0 & x_0 \end{bmatrix}$$

for any $\boldsymbol{x} \in \mathcal{E}^n$. Note that $\mathsf{Arw}(\boldsymbol{x})$ can be also redefined as

$$\mathsf{Arw}(\boldsymbol{x}) = J_n \sum_{i=1}^{n-1} \mathsf{Arw}^{\langle i \rangle}(\boldsymbol{x}) = \sum_{i=1}^{n-1} J_n^{\langle i \rangle} \mathsf{Arw}\left(\boldsymbol{x}^{\langle i \rangle}\right),$$

where

$$\mathsf{Arw}^{\langle i \rangle}(\boldsymbol{x}) \triangleq \mathsf{Arw}\left(\boldsymbol{x}^{\langle i \rangle}\right) = \begin{bmatrix} x_0 & x_i \boldsymbol{u}_{n-1}^{\langle i \rangle^{\mathsf{T}}} \\ x_i \boldsymbol{u}_{n-1}^{\langle i \rangle} & x_0 I_{n-1}^{\langle i \rangle} \end{bmatrix},$$

for $i = 1, 2, \ldots, n-1$.
Let $\boldsymbol{x} \in \mathcal{E}^n$ be such that $\det(\boldsymbol{x}) \neq 0$. We define the *inverse* of $\boldsymbol{x} \in \mathcal{E}^n$ as

$$\boldsymbol{x}^{-1} \triangleq J_n \sum_{i=1}^{n-1} \boldsymbol{x}^{\langle i \rangle^{-1}}, \tag{2}$$

where

$$\boldsymbol{x}^{\langle i\rangle^{-1}} \triangleq \frac{1}{\lambda^{+}(\boldsymbol{x}^{\langle i\rangle})}\boldsymbol{c}^{+}\left(\boldsymbol{x}^{\langle i\rangle}\right)+\frac{1}{\lambda^{-}(\boldsymbol{x}^{\langle i\rangle})}\boldsymbol{c}^{-}\left(\boldsymbol{x}^{\langle i\rangle}\right)=\frac{1}{\det(\boldsymbol{x}^{\langle i\rangle})}\begin{bmatrix} x_0 \\ -x_i\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix}=\frac{1}{\det(\boldsymbol{x}^{\langle i\rangle})}R_n^{\langle i\rangle}\boldsymbol{x}^{\langle i\rangle}.$$

We define the *Jordan multiplication* $\Box : \mathcal{E}^n \times \mathcal{E}^n \longrightarrow \mathcal{E}^n$ as

$$\begin{aligned}\boldsymbol{x}\,\Box\,\boldsymbol{y} &\triangleq J_n\sum_{i=1}^{n-1}\mathsf{Arw}^{\langle i\rangle}(\boldsymbol{x})\boldsymbol{y}^{\langle i\rangle}\\ &=\sum_{i=1}^{n-1}J_n^{\langle i\rangle}\mathsf{Arw}\left(\boldsymbol{x}^{\langle i\rangle}\right)\boldsymbol{y}^{\langle i\rangle}\\ &=J_n\begin{bmatrix}(n-1)\,x_0y_0+\sum_{i=1}^{n-1}x_iy_i\\ x_0y_1+x_1y_0\\ x_0y_2+x_2y_0\\ \vdots\\ x_0y_{n-1}+x_{n-1}y_0\end{bmatrix}=\mathsf{Arw}(\boldsymbol{x})\boldsymbol{y},\end{aligned}$$

for $\boldsymbol{x},\boldsymbol{y}\in\mathcal{E}^n$. Therefore, $\boldsymbol{x}\,\Box\,\boldsymbol{e}_n=\boldsymbol{x}$. We also have

$$\boldsymbol{x}\,\Box\,\boldsymbol{x}^{-1}=J_n\sum_{i=1}^{n-1}\mathsf{Arw}^{\langle i\rangle}(\boldsymbol{x})\,\boldsymbol{x}^{\langle i\rangle^{-1}}=J_n\sum_{i=1}^{n-1}\boldsymbol{e}_n=\boldsymbol{e}_n=(1;\ \underbrace{0;0;\ldots;0}_{(n-1)-\text{times}}\ ),$$

and

$$\boldsymbol{x}\,\Box\,\boldsymbol{x}=J_n\sum_{i=1}^{n-1}\mathsf{Arw}^{\langle i\rangle}(\boldsymbol{x})\boldsymbol{x}^{\langle i\rangle}=J_n\sum_{i=1}^{n-1}\boldsymbol{x}^{\langle i\rangle^{2}}=\boldsymbol{x}^2.$$

The *quadratic representation* of $x\in\mathcal{E}^n$ is denoted by $\mathsf{Q}_{\boldsymbol{x}}$ and is defined as

$$\mathsf{Q}_{\boldsymbol{x}}\triangleq J_n\sum_{i=1}^{n-1}\mathsf{Q}_{\boldsymbol{x}}^{\langle i\rangle}=\sum_{i=1}^{n-1}J_n^{\langle i\rangle}\mathsf{Q}_{\boldsymbol{x}^{\langle i\rangle}}, \tag{3}$$

where

$$\mathsf{Q}_{\boldsymbol{x}}^{\langle i\rangle}\triangleq\mathsf{Q}_{\boldsymbol{x}^{\langle i\rangle}}=2\boldsymbol{x}^{\langle i\rangle}\boldsymbol{x}^{\langle i\rangle^{\mathsf{T}}}-\det\left(\boldsymbol{x}^{\langle i\rangle}\right)R_n^{\langle i\rangle}=\begin{bmatrix}x_0^2+x_i^2 & 2x_0x_i\boldsymbol{u}_{n-1}^{\langle i\rangle^{\mathsf{T}}}\\ 2x_0x_i\boldsymbol{u}_{n-1}^{\langle i\rangle} & \left(x_0^2+x_i^2\right)I_{n-1}^{\langle i\rangle}\end{bmatrix}, \tag{4}$$

for $i=1,2,\ldots,n-1$. One can easily find that $\mathsf{Q}_{\boldsymbol{x}^{\langle i\rangle}}\boldsymbol{e}_n=\boldsymbol{x}^{\langle i\rangle^{2}}$, $\mathsf{Q}_{\boldsymbol{x}^{\langle i\rangle^{-1}}}\boldsymbol{x}^{\langle i\rangle}=\boldsymbol{x}^{\langle i\rangle^{-1}}$ (hence $\mathsf{Q}^{-1}_{\boldsymbol{x}^{\langle i\rangle^{-1}}}\boldsymbol{x}^{\langle i\rangle^{-1}}=x^{\langle i\rangle}$), and $\mathsf{Arw}\left(\boldsymbol{x}^{\langle i\rangle^{-1}}\right)\mathsf{Q}_{\boldsymbol{x}^{\langle i\rangle}}x^{\langle i\rangle^{-1}}=\boldsymbol{e}_n$, for $i=1,2,\ldots,n-1$. Therefore $\mathsf{Q}_{\boldsymbol{x}}\boldsymbol{e}_n=x^2$, $\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{x}=\boldsymbol{x}^{-1}$ (hence $\mathsf{Q}^{-1}_{\boldsymbol{x}^{-1}}\boldsymbol{x}^{-1}=\boldsymbol{x}$), $\boldsymbol{x}^{-1}\Box\,\mathsf{Q}_{\boldsymbol{x}}\boldsymbol{x}^{-1}=\boldsymbol{e}_n$, and

$$\boldsymbol{x}^{-1}\Box\,\mathsf{Q}^{-1}_{\boldsymbol{x}^{-1}}x^{-1}=e_n. \tag{5}$$

We define the product $\blacksquare : \mathcal{E}^n \times \mathcal{E}^n \longrightarrow \mathbb{R}$ as

$$\boldsymbol{x} \blacksquare \boldsymbol{y} \triangleq \frac{1}{2}\text{trace}\,(\boldsymbol{x} \,\Box\, \boldsymbol{y}) = (n-1)x_0y_0 + \bar{x}^{\mathsf{T}}\bar{\boldsymbol{y}}, \text{ for } \boldsymbol{x},\boldsymbol{y} \in \mathcal{E}^n.$$

It is easy to find that

$$\boldsymbol{x} \blacksquare \boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}} J_n^{-1} \boldsymbol{y} = \sum_{i=1}^{n-1} \left({\boldsymbol{x}^{\langle i \rangle}}^{\mathsf{T}} \boldsymbol{y}^{\langle i \rangle} \right), \text{ for any } \boldsymbol{x},\boldsymbol{y} \in \mathcal{E}^n. \tag{6}$$

Let $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in \mathcal{E}^n$, and $\alpha,\beta \in \mathbb{R}$. It is not hard to check that $\boldsymbol{x} \,\Box\, (\alpha\boldsymbol{y} + \beta\boldsymbol{z}) = \alpha(\boldsymbol{x} \,\Box\, \boldsymbol{y}) + \beta(\boldsymbol{x} \,\Box\, \boldsymbol{z})$, and $(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \,\Box\, \boldsymbol{z} = \alpha(\boldsymbol{x} \,\Box\, \boldsymbol{z}) + \beta(\boldsymbol{y} \,\Box\, \boldsymbol{z})$, hence "$\Box$" is a bilinear map, and therefore the structure $(\mathcal{E}, \Box)$ is an algebra. One can also prove the following identities: $\boldsymbol{x} \,\Box\, \boldsymbol{y} = \boldsymbol{y} \,\Box\, \boldsymbol{x}$ (Commutativity), and $\boldsymbol{x} \,\Box\, \left(\boldsymbol{x}^2 \,\Box\, \boldsymbol{y}\right) = \boldsymbol{x}^2 \,\Box\, (\boldsymbol{x} \,\Box\, \boldsymbol{y})$ (Jordan identity), which in turn imply the result in the following proposition.

Proposition 1. *The algebra $(\mathcal{E}, \Box)$ is a Jordan algebra.*

Table 1 compares between the Jordan algebraic structures associated with the infinity norm cone and that with the second-order cone.
Finally, we define the *Frobenius norm* of $\boldsymbol{x} \in (\mathcal{E}^n, \Box)$ as

$$\|\boldsymbol{x}\|_{\mathrm{F}} \triangleq \sqrt{\sum_{i=1}^{n-1} \left(\left(\lambda_i^+(\boldsymbol{x})\right)^2 + \left(\lambda_i^-(\boldsymbol{x})\right)^2 \right)} = \sqrt{2\,\boldsymbol{x} \blacksquare \boldsymbol{x}}.$$

Let also $\boldsymbol{y} \in (\mathcal{E}^n, \Box)$, then

$$\left\|\boldsymbol{x}^2\right\|_{\mathrm{F}} \le \|\boldsymbol{x}\|_{\mathrm{F}}^2 \quad \text{and} \quad |\boldsymbol{x} \blacksquare \boldsymbol{y}| \le \frac{1}{2}\,\|\boldsymbol{x}\|_{\mathrm{F}}\,\|\boldsymbol{y}\|_{\mathrm{F}}\,.$$

This can be seen by noting that

$$\|\boldsymbol{x}^2\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{n-1} \left(\left(\lambda_i^+(\boldsymbol{x})\right)^4 + \left(\lambda_i^-(\boldsymbol{x})\right)^4 \right)} \le \sum_{i=1}^{n-1} \left(\left(\lambda_i^+(\boldsymbol{x})\right)^2 + \left(\lambda_i^-(\boldsymbol{x})\right)^2 \right) = \|\boldsymbol{x}\|_{\mathrm{F}}^2\,,$$

and that

$$\begin{array}{rcl} |\boldsymbol{x} \blacksquare \boldsymbol{y}| & = & \left|\boldsymbol{x}^{\mathsf{T}} J_n^{-1} \boldsymbol{y}\right| \\ & = & \left|\left(J_n^{-1/2}\boldsymbol{x}\right)^{\mathsf{T}} J_n^{-1/2}\boldsymbol{y}\right| \\ & \le & \left\|J_n^{-1/2}\boldsymbol{x}\right\|_2 \left\|J_n^{-1/2} y\right\|_2 \\ & = & \sqrt{x^{\mathsf{T}} J_n^{-1}\boldsymbol{x}}\sqrt{\boldsymbol{y}^{\mathsf{T}} J_n^{-1}\boldsymbol{y}} \\ & = & \sqrt{\boldsymbol{x} \blacksquare \boldsymbol{x}}\sqrt{\boldsymbol{y} \blacksquare \boldsymbol{y}} \;=\; \dfrac{1}{2}\,\|\boldsymbol{x}\|_{\mathrm{F}}\,\|\boldsymbol{y}\|_{\mathrm{F}}\,, \end{array}$$

where the last inequality was obtained using the Cauchy–Schwarz inequality.

Table 1: Comparing the Jordan algebraic notions and concepts associated with $\mathcal{I}^n$ and $\mathcal{C}_2^n$.

Cone	Infinity norm cone $\mathcal{I}^n$	Second-order cone $\mathcal{C}_2^n$ (see [7, Section 11.1])
Constraint	$x_0 \geq \Vert\bar{\boldsymbol{x}}\Vert_\infty$	$x_0 \geq \Vert\bar{\boldsymbol{x}}\Vert_2$
Inner product	$\boldsymbol{x} \blacksquare \boldsymbol{y} \triangleq (n-1)x_0y_0 + \bar{\boldsymbol{x}}^\mathsf{T}\bar{\boldsymbol{y}}$	$\boldsymbol{x} \bullet \boldsymbol{y} \triangleq x_0y_0 + \bar{\boldsymbol{x}}^\mathsf{T}\bar{\boldsymbol{y}}$
Jordan multiplication	$\boldsymbol{x} \square \boldsymbol{y} \triangleq J_n \begin{bmatrix} \boldsymbol{x} \blacksquare \boldsymbol{y} \\ x_0\bar{\boldsymbol{y}} + y_0\bar{\boldsymbol{x}} \end{bmatrix}$	$\boldsymbol{x} \circ \boldsymbol{y} \triangleq \begin{bmatrix} \boldsymbol{x} \bullet \boldsymbol{y} \\ x_0\bar{\boldsymbol{y}} + y_0\bar{\boldsymbol{x}} \end{bmatrix}$
Jordan algebra	$(\mathcal{E}^n, \square)$	$(\mathcal{E}^n, \circ)$
Identity	$e \triangleq \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}$	$e \triangleq \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}$
Eigenvalues	$\lambda_i^{\mp}(x) \triangleq x_0 \mp x_i$ for all $i = 1, \ldots, n-1$	$\lambda^{\mp}(x) \triangleq x_0 \mp \Vert\bar{\boldsymbol{x}}\Vert_2$
Eigenvectors	$\boldsymbol{c}_i^{\mp} \triangleq \frac{1}{2} \begin{bmatrix} 1 \\ \mp \boldsymbol{u}_{n-1}^{\langle i \rangle} \end{bmatrix}$ for all $i = 1, \ldots, n-1$	$\boldsymbol{c}^{\mp} \triangleq \frac{1}{2} \begin{bmatrix} 1 \\ \mp \frac{\bar{\boldsymbol{x}}}{\Vert\bar{\boldsymbol{x}}\Vert_2} \end{bmatrix}$
Rank	$\mathsf{rk}(\mathcal{I}^n) \triangleq 2(n-1) \in \mathcal{O}(n)$	$\mathsf{rk}(\mathcal{C}_2^n) \triangleq 2 \in \mathcal{O}(1)$
Trace	$\text{trace}(\boldsymbol{x}) \triangleq 2(n-1)x_0$	$\text{trace}(\boldsymbol{x}) \triangleq 2x_0$
Determinant	$\det(\boldsymbol{x}) \triangleq \prod_{i=1}^{n-1} (x_0^2 - x_i^2)$	$\det(\boldsymbol{x}) \triangleq x_0^2 - \Vert\bar{\boldsymbol{x}}\Vert_2^2$
Spectral decomposition	$\boldsymbol{x} \triangleq J_n \sum_{i=1}^{n-1} \left(\lambda_i^+(\boldsymbol{x})\boldsymbol{c}_i^+(\boldsymbol{x}) + \lambda_i^-(\boldsymbol{x})\boldsymbol{c}_i^-(\boldsymbol{x})\right)$	$\boldsymbol{x} \triangleq \lambda^+(\boldsymbol{x})\boldsymbol{c}^+(\boldsymbol{x}) + \lambda^-(\boldsymbol{x})\boldsymbol{c}^-(\boldsymbol{x})$
Square	$\boldsymbol{x}^2 \triangleq \boldsymbol{x} \square \boldsymbol{x} = \begin{bmatrix} \boldsymbol{x} \blacksquare \boldsymbol{x} \\ 2x_0\bar{\boldsymbol{x}} \end{bmatrix}$	$\boldsymbol{x}^2 \triangleq \boldsymbol{x} \circ \boldsymbol{x} = \begin{bmatrix} \boldsymbol{x} \bullet \boldsymbol{x} \\ 2x_0\bar{\boldsymbol{x}} \end{bmatrix}$
Inverse	$\boldsymbol{x}^{-1} \triangleq J_n \sum_{i=1}^{n-1} \left(\frac{1}{\lambda_i^+(\boldsymbol{x})}\boldsymbol{c}_i^+(\boldsymbol{x}) + \frac{1}{\lambda_i^-(\boldsymbol{x})}\boldsymbol{c}_i^-(\boldsymbol{x})\right)$	$\boldsymbol{x}^{-1} \triangleq \frac{1}{\lambda^+(\boldsymbol{x})}\boldsymbol{c}^+(\boldsymbol{x}) + \frac{1}{\lambda^-(\boldsymbol{x})}\boldsymbol{c}^-(\boldsymbol{x})$
Arrow-shaped matrix	$\mathsf{Arw}(\boldsymbol{x}) \triangleq J_n \begin{bmatrix} (n-1)x_0 & \bar{\boldsymbol{x}}^\mathsf{T} \\ \bar{\boldsymbol{x}} & x_0 I_{n-1} \end{bmatrix}$	$\mathsf{Arw}(\boldsymbol{x}) \triangleq \begin{bmatrix} x_0 & \bar{\boldsymbol{x}}^\mathsf{T} \\ \bar{\boldsymbol{x}} & x_0 I_{n-1} \end{bmatrix}$
Quadratic representation matrix	$\mathsf{Q}_{\boldsymbol{x}} \triangleq J_n \sum_{i=1}^{n-1} \left(2\boldsymbol{x}^{\langle i \rangle}{\boldsymbol{x}^{\langle i \rangle}}^\mathsf{T} - \det(\boldsymbol{x}^{\langle i \rangle})R_n^{\langle i \rangle}\right)$	$\mathsf{Q}_{\boldsymbol{x}} \triangleq 2\boldsymbol{x}\boldsymbol{x}^\mathsf{T} - \det(\boldsymbol{x})R_n$
Logarithmic barrier	$\ln\det(\boldsymbol{x}) \triangleq \sum_{i=1}^{n-1} \ln\left(x_0^2 - x_i^2\right)$	$\ln\det(\boldsymbol{x}) \triangleq \ln\left(x_0^2 - \Vert\bar{\boldsymbol{x}}\Vert_2^2\right)$
Gradient $\nabla_{\boldsymbol{x}} \ln\det(\boldsymbol{x})$	$2J_n^{-1}\boldsymbol{x}^{-1} = \sum_{i=1}^{n-1} \left(\frac{2}{\det(\boldsymbol{x}^{\langle i \rangle})} \begin{bmatrix} x_0 \\ -x_i\boldsymbol{u}_{n-1}^{\langle i \rangle} \end{bmatrix}\right)$	$2\boldsymbol{x}^{-1} = \frac{2}{\det(\boldsymbol{x})} \begin{bmatrix} x_0 \\ -\bar{\boldsymbol{x}} \end{bmatrix}$
Hessian $\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \ln\det(\boldsymbol{x}) = -2J_n^{-1}Q_{x^{-1}}$	$\sum_{i=1}^{n-1} \left(\frac{2}{(\det \boldsymbol{x}^{\langle i \rangle})^2} \begin{bmatrix} x_0^2 + x_i^2 & -2x_0x_i{\boldsymbol{u}_{n-1}^{\langle i \rangle}}^\mathsf{T} \\ -2x_0x_i\boldsymbol{u}_{n-1}^{\langle i \rangle} & (x_0^2 + x_i^2)\, I_{n-1}^{\langle i \rangle} \end{bmatrix}\right)$	$\frac{2}{(\det \boldsymbol{x}^{\langle i \rangle})^2} \begin{bmatrix} \Vert\boldsymbol{x}\Vert & -2x_0\bar{\boldsymbol{x}}^\mathsf{T} \\ -2x_0\bar{\boldsymbol{x}} & \det(\boldsymbol{x})I_{n-1} + 2\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\mathsf{T} \end{bmatrix}$

3. The barrier function associated with the cone

In this section, we introduce the logarithmic barrier function associated with the infinity norm cone, its derivatives, and its self-concordance properties.
Following the standard way in defining the logarithmic barriers in convex programming, we define the logarithmic barrier associated with the infinity norm cone as $\ell(\boldsymbol{x}) \triangleq -\ln\det(\boldsymbol{x})$ for $x = (x_0; \bar{\boldsymbol{x}}) \in \text{int}(\mathcal{I}^n)$. In our setting, we have

$$\ell(\boldsymbol{x}) \triangleq -\ln\det(x) = -\ln\left(\prod_{i=1}^{n-1} \left(x_0^2 - x_i^2\right)\right) = -\sum_{i=1}^{n-1} \ln\left(x_0^2 - x_i^2\right) = \sum_{i=1}^{n-1} \ell^{\langle i \rangle}(\boldsymbol{x}),$$

where

$$\ell^{\langle i \rangle}(\boldsymbol{x}) \triangleq -\ln\ \left(x_0^2 - x_i^2\right),\ i = 1, 2, \ldots, n-1.$$

Note that $\ell^{\langle i \rangle}(\cdot)$ is a strictly convex function on $\text{int}(\mathcal{I}^n)$ for $i = 1, 2, \ldots, n-1$. Since the sum of strictly convex functions is strictly convex, the logarithmic barrier function $\ell(\boldsymbol{x})$ is strictly convex.
The results in the following lemma are a handy tool for our subsequent development.

Lemma 1. *Let $\boldsymbol{x} \in$ int$(\mathcal{I}^n)$. We have that*

1. *The gradient $\nabla_{\boldsymbol{x}}\ell(\boldsymbol{x}) = -2J_n^{-1}\boldsymbol{x}^{-1}$, where $\boldsymbol{x}^{-1}$ is defined in (2).*

2. *The Hessian* $\nabla^2_{\boldsymbol{xx}}\ell(\boldsymbol{x}) = 2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}$*, and hence* $\mathsf{D}_{\boldsymbol{x}}\boldsymbol{x}^{-1} = -\mathsf{Q}_{\boldsymbol{x}^{-1}}$*, where* $\mathsf{Q}_{\cdot}$ *is defined in (3) and* $\mathsf{D}_{\boldsymbol{x}}$ *is the Jacobian matrix with respect to* $\boldsymbol{x}$*.*
3. *For any vector* $\boldsymbol{h} \in (\mathcal{E}^n, \square)$*, the third derivative* $\nabla^3_{\boldsymbol{xxx}}\ell(x)[\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}] = -4\,\boldsymbol{s}(\boldsymbol{x},\boldsymbol{h}) \blacksquare \boldsymbol{s}^2(\boldsymbol{x},\boldsymbol{h})$*, where the vector* $\boldsymbol{s} \triangleq \boldsymbol{s}(\boldsymbol{x},\boldsymbol{h}) = J_n \sum_{i=1}^{n-1} \boldsymbol{s}^{\langle i\rangle}(\boldsymbol{x},\boldsymbol{h}) \in (\mathcal{E}^n, \square)$*, and*

$$\boldsymbol{s}^{\langle i\rangle} \triangleq \boldsymbol{s}^{\langle i\rangle}(\boldsymbol{x},\boldsymbol{h}) \triangleq \mathsf{Arw}\left(\boldsymbol{x}^{\langle i\rangle^{-1}}\right)\boldsymbol{h} = \frac{1}{\det\left(\boldsymbol{x}^{\langle i\rangle}\right)} \begin{bmatrix} x_0h_0 - x_ih_i \\ (x_0h_i - x_ih_0)\,\boldsymbol{u}_{n-1}^{\langle i\rangle} \end{bmatrix},\ i = 1,2,\ldots,n-1.$$

Proof Note that

$$\begin{aligned}
\nabla_{\boldsymbol{x}} \ln\det(\boldsymbol{x}) &= \nabla_{\boldsymbol{x}}\left(\prod_{i=1}^{n-1} \ln\left((x_0+x_i)(x_0-x_i)\right)\right) \\
&= \nabla_{\boldsymbol{x}} \sum_{i=1}^{n-1} \left(\ln(x_0+x_i) + \ln(x_0-x_i)\right) \\
&= \sum_{i=1}^{n-1}\left(\frac{1}{x_0+x_i}\nabla_{\boldsymbol{x}}(x_0+x_i) + \frac{1}{x_0-x_i}\nabla_{\boldsymbol{x}}(x_0-x_i)\right) \\
&= \sum_{i=1}^{n-1}\left(\frac{1}{(x_0+x_i)}\begin{bmatrix}1\\ \boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix} + \frac{1}{(x_0-x_i)}\begin{bmatrix}1\\ -\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix}\right) \\
&= \sum_{i=1}^{n-1}\left(\frac{1}{(x_0+x_i)(x_0-x_i)}\begin{bmatrix}2x_0\\ -2x_i\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix}\right) \\
&= 2\sum_{i=1}^{n-1}\left(\frac{1}{\det(x^{\langle i\rangle})}\begin{bmatrix}x_0\\ -x_i\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix}\right) \;=\; 2J_n^{-1}\boldsymbol{x}^{-1}.
\end{aligned}$$

This proves item (1). To prove item (2), it suffices to show that $\mathsf{D}_{\boldsymbol{x}}\boldsymbol{x}^{\langle i\rangle^{-1}} = -\mathsf{Q}_{\boldsymbol{x}^{\langle i\rangle^{-1}}}$ for $i = 1,2,\ldots,n-1$. Note that

$$\begin{aligned}
\mathsf{D}_{\boldsymbol{x}}\boldsymbol{x}^{\langle i\rangle^{-1}} &= \mathsf{D}_{\boldsymbol{x}}\left(\frac{1}{2(x_0+x_i)}\begin{bmatrix}1\\ \boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix} + \frac{1}{2(x_0-x_i)}\begin{bmatrix}1\\ -\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix}\right) \\
&= \mathsf{D}_{x}\begin{bmatrix}\dfrac{x_0}{x_0^2-x_i^2}\\ \dfrac{-x_i}{x_0^2-x_i^2}\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix} \\
&= \frac{-1}{\left(\det\boldsymbol{x}^{\langle i\rangle}\right)^2}\begin{bmatrix} x_0^2+x_i^2 & -2x_0x_i\boldsymbol{u}_{n-1}^{\langle i\rangle\mathsf{T}} \\ -2x_0x_i\boldsymbol{u}_{n-1}^{\langle i\rangle} & (x_0^2+x_i^2)\,I_{n-1}^{\langle i\rangle}\end{bmatrix} \\
&= \frac{-1}{\left(\det x^{\langle i\rangle}\right)^2}\left(\begin{bmatrix} 2x_0^2 & -2x_0x_i\boldsymbol{u}_{n-1}^{\langle i\rangle\mathsf{T}} \\ -2x_0x_i\boldsymbol{u}_{n-1}^{\langle i\rangle} & 2x_i^2 I_{n-1}^{\langle i\rangle}\end{bmatrix} - (x_0^2-x_i^2)\begin{bmatrix}1 & 0^{\mathsf{T}}\\ 0 & -I_{n-1}^{\langle i\rangle}\end{bmatrix}\right) \\
&= -\left(2\boldsymbol{x}^{\langle i\rangle^{-1}}\boldsymbol{x}^{\langle i\rangle^{-1}\mathsf{T}} - \det\left(\boldsymbol{x}^{\langle i\rangle^{-1}}\right)R_n^{\langle i\rangle}\right) \;=\; -\mathsf{Q}^{\langle i\rangle}_{\boldsymbol{x}^{-1}},
\end{aligned}$$

where the last equality follows from (4).

Finally, the following sequence of equalities proves item (3).

$$
\begin{aligned}
\nabla^3_{\boldsymbol{xxx}}\ell(\boldsymbol{x})[\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}] &= \sum_{i=1}^{n-1} \nabla^3_{\boldsymbol{xxx}}\ell^{\langle i\rangle}(\boldsymbol{x})[\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}] \\
&= \sum_{i=1}^{n-1} \nabla_{\boldsymbol{x}}\left(\boldsymbol{h}^{\mathsf{T}}\nabla^2_{\boldsymbol{xx}}\ell^{\langle i\rangle}(\boldsymbol{x})\boldsymbol{h}\right)[\boldsymbol{h}] \\
&= \sum_{i=1}^{n-1}\left(\begin{bmatrix} h_0 & \bar{\boldsymbol{h}}^{\mathsf{T}}\end{bmatrix}\nabla_{\boldsymbol{x}}\left(\frac{2}{\left(\det\left(\boldsymbol{x}^{\langle i\rangle}\right)\right)^2}\left((x_0^2+x_i^2)(h_0^2+h_i^2)-4x_0x_ih_0h_i\right)\right)\right) \\
&= \sum_{i=1}^{n-1}\left(\frac{-4}{\left(\det\left(\boldsymbol{x}^{\langle i\rangle}\right)\right)^3}\begin{bmatrix} h_0 & \bar{\boldsymbol{h}}^{\mathsf{T}}\end{bmatrix}\begin{bmatrix}\left(h_0^2+h_i^2\right)\left(x_0^3+3x_0x_i^2\right)-2h_0h_i\left(x_i^3+3x_0^2x_i\right) \\ -\left(\left(h_0^2+h_i^2\right)\left(x_i^3+3x_0^2x_i\right)-2h_0h_i\left(x_0^3+3x_0x_i^2\right)\right)\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix}\right) \\
&= \sum_{i=1}^{n-1}\left(\frac{-4}{\left(\det\left(\boldsymbol{x}^{\langle i\rangle}\right)\right)^3}\left((x_0^3+3x_0x_i^2)(h_0^3+3h_0h_i^2)-(x_i^3+3x_0^2x_i)(h_i^3+3h_0^2h_i)\right)\right) \\
&= \sum_{i=1}^{n-1}\left(\frac{-4}{\left(\det\left(\boldsymbol{x}^{\langle i\rangle}\right)\right)^3}\begin{bmatrix} x_0h_0-x_ih_i \\ (x_0h_i-x_ih_0)\,\boldsymbol{u}_{n-1}^{\langle i\rangle}\end{bmatrix}^{\mathsf{T}}\begin{bmatrix}\left(x_0^2+x_i^2\right)\left(h_0^2+h_i^2\right)-4x_0x_ih_0h_i \\ \left(2h_0h_i\left(x_0^2+x_i^2\right)-2x_0x_i\left(h_0^2+h_i^2\right)\right)u_{n-1}^{\langle i\rangle}\end{bmatrix}\right) \\
&= -4\sum_{i=1}^{n-1}\left(\boldsymbol{s}^{\langle i\rangle^{\mathsf{T}}}\,\mathsf{Arw}\left(\boldsymbol{s}^{\langle i\rangle}\right)\boldsymbol{s}^{\langle i\rangle}\right) \\
&= -4\sum_{i=1}^{n-1}\left(\boldsymbol{s}^{\langle i\rangle^{\mathsf{T}}}\boldsymbol{s}^{\langle i\rangle^{2}}\right) \quad = \quad -4\,\boldsymbol{s} \blacksquare \boldsymbol{s}^2.
\end{aligned}
$$

The proof is complete. □

Now, we show that the function $\ell(\cdot)$ is a self-concordant barrier with complexity value 1.

Definition 1 (Definition 2.1.1 in [21]). Let V be a finite-dimensional real vector space, G be an open nonempty convex subset of V, and let f be a C^3, convex mapping from G to $\mathbb{R}$. Then f is called *α-self-concordant* on G with the parameter $\alpha > 0$ if for every $x \in G$ and $\boldsymbol{h} \in V$, the following inequality holds

$$
\left|\ \nabla^3_{\boldsymbol{xxx}}f(\boldsymbol{x})\ [\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}]\ \right| \ \le\ \frac{2}{\sqrt{\alpha}}\ \left(\nabla^2_{\boldsymbol{xx}}f(\boldsymbol{x})\ [\boldsymbol{h},\boldsymbol{h}]\right)^{3/2}. \tag{7}
$$

An α-self-concordant function f on G is called *strongly α-self-concordant* if f tends to infinity for any sequence approaching a boundary point of G.

Table 2: 1-self-concordant barriers for most well-known conic programs.

Linear program	Second-order cone program	Semidefinite program	Infinity norm program
$\boldsymbol{x} \in \mathbb{R}^n_+$	$\boldsymbol{x} \in \mathcal{C}^n_2$	$X \in \mathcal{S}^n_+$	$\boldsymbol{x} \in \mathcal{I}^n$
$\ell(\boldsymbol{x}) = -\sum_{i=1}^{n} \ln x_i$	$\ell(\boldsymbol{x}) = -\ln\left(x_0^2 - \lVert\bar{\boldsymbol{x}}\rVert_2^2\right)$	$\ell(X) = -\ln\det(X)$	$\ell(\boldsymbol{x}) = -\sum_{i=1}^{n}\ln\left(x_0^2 - x_i^2\right)$

The result in the following theorem is crucial to subsequent results in this paper. The result in this theorem is the counterpart of very well-known results in the interior-point theory of conic programming (see Table 2).

Theorem 1. *The logarithmic barrier function $\ell(\cdot)$ is 1-strongly self-concordant on $\mathcal{I}^n$.*

Proof Let $\boldsymbol{h} \in (\mathcal{E}^n, \square)$. From item (3) in Lemma 1, we have

$$\nabla^3_{\boldsymbol{xxx}} \ell(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}, \boldsymbol{h}] = -4\, \boldsymbol{s} \blacksquare \boldsymbol{s}^2,$$

where $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{x}, \boldsymbol{h}) \in (\mathcal{E}^n, \square)$ is given by

$$\boldsymbol{s}^{\langle i \rangle} \triangleq \mathsf{Arw}\left(\boldsymbol{x}^{\langle i \rangle^{-1}}\right) \boldsymbol{h} = \frac{1}{\det\left(\boldsymbol{x}^{\langle i \rangle}\right)} \begin{bmatrix} x_0 h_0 - x_i h_i \\ (x_0 h_i - x_i h_0)\, \boldsymbol{u}^{\langle i \rangle}_{n-1} \end{bmatrix}, \; i = 1, 2, \ldots, n-1.$$

Note that

$$\left(\lambda^+_i(\boldsymbol{s})\right)^2 + \left(\lambda^-_i(\boldsymbol{s})\right)^2 = \frac{2}{\left(\det\left(\boldsymbol{x}^{\langle i \rangle}\right)\right)^2} \left((x_0 h_0 - x_i h_i)^2 + (x_0 h_i - x_i h_0)^2\right), \; i = 1, 2, \ldots, n-1.$$

From item (2) in Lemma 1, we also have

$$\begin{aligned}
\boldsymbol{h}^\mathsf{T} \nabla^2_{\boldsymbol{xx}} \ell(x) \boldsymbol{h} &= \sum_{i=1}^{n-1} \boldsymbol{h}^\mathsf{T} \nabla^2_{\boldsymbol{xx}} \ell^{\langle i \rangle}(\boldsymbol{x}) \boldsymbol{h} \\
&= 2 \sum_{i=1}^{n-1} \boldsymbol{h}^\mathsf{T} \mathsf{Q}^{\langle i \rangle}_{\boldsymbol{x}^{-1}} \boldsymbol{h} \\
&= \sum_{i=1}^{n-1} \left(\frac{2}{\left(\det\left(\boldsymbol{x}^{\langle i \rangle}\right)\right)^2} \begin{bmatrix} h_0 & \bar{\boldsymbol{h}}^\mathsf{T} \end{bmatrix} \begin{bmatrix} x_0^2 + x_i^2 & -2x_0 x_i \boldsymbol{u}^{\langle i \rangle^\mathsf{T}}_{n-1} \\ -2x_0 x_i \boldsymbol{u}^{\langle i \rangle}_{n-1} & (x_0^2 + x_i^2)\, I^{\langle i \rangle}_{n-1} \end{bmatrix} \begin{bmatrix} h_0 \\ \bar{\boldsymbol{h}} \end{bmatrix} \right) \\
&= \sum_{i=1}^{n-1} \left(\frac{2}{\left(\det\left(\boldsymbol{x}^{\langle i \rangle}\right)\right)^2} \begin{bmatrix} h_0 & \bar{\boldsymbol{h}}^\mathsf{T} \end{bmatrix} \begin{bmatrix} h_0 \left(x_0^2 + x_i^2\right) - 2x_0 x_i h_i \\ \left(h_i \left(x_0^2 + x_i^2\right) - 2x_0 x_i h_0\right) \boldsymbol{u}^{\langle i \rangle}_{n-1} \end{bmatrix} \right) \\
&= \sum_{i=1}^{n-1} \left(\frac{2}{\left(\det\left(\boldsymbol{x}^{\langle i \rangle}\right)\right)^2} \left(h_0^2 \left(x_0^2 + x_i^2\right) - 2x_0 x_i h_0 h_i + h_i^2 \left(x_0^2 + x_i^2\right) - 2x_0 x_i h_0 h_i\right) \right) \\
&= \sum_{i=1}^{n-1} \left(\frac{2}{\left(\det\left(\boldsymbol{x}^{\langle i \rangle}\right)\right)^2} \left((x_0^2 + x_i^2)(h_0^2 + h_i^2) - 4x_0 x_i h_0 h_i\right) \right) \\
&= \sum_{i=1}^{n-1} \left(\frac{2}{\left(\det\left(\boldsymbol{x}^{\langle i \rangle}\right)\right)^2} \left((x_0 h_0 - x_i h_i)^2 + (x_0 h_i - x_i h_0)^2\right) \right) \\
&= \sum_{i=1}^{n-1} \left(\lambda_1^2(\boldsymbol{s}^{\langle i \rangle}) + \lambda_2^2(\boldsymbol{s}^{\langle i \rangle})\right) \\
&= 2\boldsymbol{s} \blacksquare \boldsymbol{s} \;=\; \|\boldsymbol{s}\|_\mathrm{F}^2.
\end{aligned}$$

The result immediately follows from the following:

$$\left|\nabla^3_{\boldsymbol{xxx}} \ell(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}, \boldsymbol{h}]\right| = 4 \left|\boldsymbol{s} \blacksquare \boldsymbol{s}^2\right| \leq 2 \left\|\boldsymbol{s}\right\|_\mathrm{F} \left\|\boldsymbol{s}^2\right\|_\mathrm{F} \leq 2 \left\|\boldsymbol{s}\right\|_\mathrm{F}^3 = 2 \left(\nabla^2_{\boldsymbol{xx}} \ell^{\langle i \rangle}[\boldsymbol{h}, \boldsymbol{h}]\right)^{3/2}.$$

Since $\ell(\boldsymbol{x})$ tends to ∞ for any sequence approaching a boundary point of $\mathcal{I}^n$, we deduce that $\ell(\boldsymbol{x})$ is 1-strongly self-concordant. This completes the proof. □

4. The composite recourse function

In this section, we compute the derivatives of the Composite recourse function and its self-concordance properties. The two-stage barrier SINP problem is defined as

$$\begin{array}{lll} \min \quad \eta(\boldsymbol{x},\mu) \triangleq \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} - \mu \ln\det(\boldsymbol{x}) + \sum_{k=1}^{K} \rho^{(k)}(\boldsymbol{x},\mu) & \text{s.t.} & A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \operatorname{int}\mathcal{I}^n, \\ \rho^{(k)}(\boldsymbol{x},\mu) \triangleq \min \ {\boldsymbol{d}^{(k)}}^{\mathsf{T}}\boldsymbol{y}^{(k)} - \mu \ln\det\left(\boldsymbol{y}^{(k)}\right) & \text{s.t.} & W^{(k)}\boldsymbol{y}^{(k)} = \boldsymbol{q}^{(k)} + T^{(k)}\boldsymbol{x},\ \boldsymbol{y}^{(k)} \in \operatorname{int}\mathcal{I}^m_+, \\ & & k = 1,\ldots,K, \end{array} \tag{8}$$

where $\operatorname{int}\mathcal{I}^n$ and $\operatorname{int}\mathcal{I}^m_+$ are the interiors of the infinity norm cones $\mathcal{I}^n$ and $\mathcal{I}^m_+$, for $k = 1,2,\ldots,K$, the function $\rho^{(k)}(\boldsymbol{x},\mu)$, for $k = 1,2,\ldots,K$, is called the *barrier recourse function*, $\eta(\boldsymbol{x},\mu)$ is called the *composite barrier function*, and μ is positive scalar.

In the next sections, we study common properties of all barrier recourse functions $\rho^{(k)}(\boldsymbol{x},\mu)$, for $k = 1,2,\ldots,K$. For this reason, we represent the barrier recourse function as

$$\begin{array}{l} \rho^{(k)}(\boldsymbol{x},\mu) \triangleq \min\left\{r\left(\boldsymbol{y}^{(k)},\mu\right) : W^{(k)}\boldsymbol{y}^{(k)} = \boldsymbol{q}^{(k)} + T^{(k)}\boldsymbol{x}\right\}, \quad \text{where} \\ r\left(\boldsymbol{y}^{(k)},\mu\right) \triangleq {\boldsymbol{d}^{(k)}}^{\mathsf{T}}\boldsymbol{y}^{(k)} - \mu \ln\det\left(\boldsymbol{y}^{(k)}\right). \end{array} \tag{9}$$

For the rest of this paper, we define the feasibility sets

$$\begin{array}{lll} \mathcal{L}^{(0)} & \triangleq & \left\{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}\right\}, \\ \mathcal{L}^{(k)}(\boldsymbol{x}) & \triangleq & \left\{\boldsymbol{y}^{(k)} \in \mathbb{R}^m : W^{(k)}\boldsymbol{y}^{(k)} = \boldsymbol{q}^{(k)} + T^{(k)}\boldsymbol{x}\right\}, \text{ for } k = 1,2,\ldots,K, \\ \mathcal{F}^{(k)} & \triangleq & \left\{\boldsymbol{x} \in \mathcal{I}^n \cap \mathcal{L}^{(0)} : \mathcal{I}^m_+ \cap \mathcal{L}^{(k)}(\boldsymbol{x}) \neq \varnothing\right\}, \text{ for } k = 1,2,\ldots,K, \\ \mathcal{F} & \triangleq & \bigcap_{k=1}^{K} \mathcal{F}^{(k)}. \end{array}$$

We also make the following assumptions.

Assumption 1. *The matrices A and $W^{(k)}$, $k = 1,2,\ldots,K$, have full row rank.*

Assumption 2. *The feasibility set $\mathcal{F}$ is nonempty.*

Assumption 1 is a standard assumption in linear and convex programming. Assumption 2 is the Slater condition, and based on this assumption strong duality holds. We have the following proposition.

Proposition 2. *Let $\boldsymbol{x} \in \operatorname{int}\mathcal{I}^n \cap \mathcal{L}^{(0)}$. The barrier recourse function $\rho^{(k)}(\boldsymbol{x},\mu)$, $k = 1,2,\ldots,K$, is convex in x.*

Proof Let $\alpha+\beta=1,\ \alpha,\beta\geq 0,$ and $\boldsymbol{x},\boldsymbol{z}\in \operatorname{int}\mathcal{I}^n\cap\mathcal{L}^{(0)}$. Then

$$\begin{aligned}\alpha\rho^{(k)}(\boldsymbol{x},\mu)+\beta\rho^{(k)}(\boldsymbol{z},\mu) &= \alpha\boldsymbol{d}^{(k)^{\mathsf{T}}}\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)-\alpha\mu\ln\det\left(\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)\right)\\ &\quad+\beta\,\boldsymbol{d}^{(k)^{\mathsf{T}}}\boldsymbol{y}^{(k)^\star}(\boldsymbol{z},\mu)-\beta\mu\ln\det\left(\boldsymbol{y}^{(k)^\star}(\boldsymbol{z},\mu)\right)\\ &\geq \boldsymbol{d}^{(k)^{\mathsf{T}}}\left(\alpha\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)+\beta\boldsymbol{y}^{(k)^\star}(\boldsymbol{z},\mu)\right)\\ &\quad-\mu\ln\det\left(\alpha\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)+\beta\boldsymbol{y}^{(k)^\star}(\boldsymbol{z},\mu)\right)\\ &\geq \rho^{(k)}\left(\alpha\boldsymbol{x}+\beta\boldsymbol{z},\mu\right),\end{aligned}$$

where the first inequality follows from the convexity of the barrier function $-\ln\det(\cdot)$, and the second inequality follows from the feasibility of $\alpha\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)+\beta\boldsymbol{y}^{(k)^\star}(\boldsymbol{z},\mu)$ for $W^{(k)}\boldsymbol{y}^{(k)}=\boldsymbol{q}^{(k)}+T^{(k)}(\alpha\boldsymbol{x}+\beta\boldsymbol{z})$.

Based on Proposition 2, we conclude that is strictly convex in $\boldsymbol{x}$.

Let $\boldsymbol{x}\in\operatorname{int}\mathcal{I}^n\cap\mathcal{L}^{(0)}$, $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)$ be the optimal primal solution for (9), $\boldsymbol{u}^{(k)^\star}(\boldsymbol{x},\mu)$ be the optimal Lagrange multiplier, and define

$$\vartheta\left(\boldsymbol{x},\boldsymbol{y}^{(k)}(\boldsymbol{x},\mu),\boldsymbol{u}^{(k)}(\boldsymbol{x},\mu),\mu\right)\triangleq\begin{bmatrix}\boldsymbol{d}^{(k)}-2\mu J_m^{-1}\left(\boldsymbol{y}^{(k)}(\boldsymbol{x},\mu)\right)^{-1}+W^{(k)^{\mathsf{T}}}\boldsymbol{u}^{(k)}(\boldsymbol{x},\mu)\\ W^{(k)}\boldsymbol{y}^{(k)}(\boldsymbol{x},\mu)-\boldsymbol{q}^{(k)}-T^{(k)}\boldsymbol{x}\end{bmatrix}.$$

Then the first-order KKT conditions $\vartheta(\boldsymbol{x},\boldsymbol{y}^{(k)}(\boldsymbol{x},\mu),\boldsymbol{u}^{(k)}(\boldsymbol{x},\mu),\mu)=0$ hold true for $\boldsymbol{y}^{(k)}=\boldsymbol{y}^{(k)^\star}$ and $\boldsymbol{u}^{(k)}=\boldsymbol{u}^{(k)^\star}$. That is, we have

$$\begin{array}{lcl}\boldsymbol{d}^{(k)}-2\mu J_m^{-1}\left(\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)\right)^{-1}+W^{(k)^{\mathsf{T}}}\boldsymbol{u}^{(k)^\star}(\boldsymbol{x},\mu) &=& \boldsymbol{0},\\ W^{(k)}\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)-\boldsymbol{q}^{(k)}-T^{(k)}\boldsymbol{x} &=& \boldsymbol{0}.\end{array}\tag{10}$$

The solution $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)$ is unique because the map $r(\boldsymbol{y}^{(k)},\mu)$ is strictly convex of $\boldsymbol{y}^{(k)}$ for a given $(\boldsymbol{x},\mu)$.
Now, we compute the gradient, Hessian and partial derivatives of the barrier recourse function $\rho^{(k)}(\boldsymbol{x},\mu)$ and the composite barrier function $\eta(\boldsymbol{x},\mu)$ associated with infinity norm cone $\mathcal{I}^n$. These derivatives will be used to prove fundamental properties of these functions.
Throughout this section and the rest of this paper, let $\mathsf{g}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)$ and $\mathsf{H}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)$ represent the gradient and Hessian of the barrier function $-\ln\det(\boldsymbol{y}^{(k)})$ with respect to $\boldsymbol{y}^{(k)}$ at $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)$. Then, using Lemma 1, we have

$$\mathsf{g}_{y}^{(k)}(\boldsymbol{x},\mu)\triangleq-2J_m^{-1}\left(y^{(k)^\star}(\boldsymbol{x},\mu)\right)^{-1},\ \text{ and }\ \mathsf{H}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)\triangleq 2J_m^{-1}\mathsf{Q}_{(\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu))^{-1}},\ k=1,2,\ldots,K.\tag{11}$$

Note that $\mathsf{H}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)$ is positive definite since logarithmic barrier is strictly convex. Furthermore, the matrix $W^{(k)}\,\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1}}(\boldsymbol{x},\mu)\,W^{(k)^{\mathsf{T}}}$ is invertible since $W^{(k)}$ has a full row rank (Assumption 1). In the rest of this paper, we also let

$$\begin{array}{rcl} \mathsf{S}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu) & \triangleq & W^{(k)}\,\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}}(\boldsymbol{x},\mu),\ k=1,2,\ldots,K,\\ \mathsf{P}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu) & \triangleq & \mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}}(\boldsymbol{x},\mu)\left(\mathsf{S}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)\,\mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}}(\boldsymbol{x},\mu)\right)^{-1}\mathsf{S}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu),\ k=1,2,\ldots,K. \end{array} \tag{12}$$

By applying the implicit function theoremto the KKT system (10), we conclude that the Lagrange multiplier $\boldsymbol{u}^{(k)^{\star}}(\boldsymbol{x},\mu)$ can be uniquely determined. Particularly, we have

$$\begin{array}{rcl} \boldsymbol{u}^{(k)^{\star}}(\boldsymbol{x},\mu) & \triangleq & -\left(W^{(k)}\,\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1}}(\boldsymbol{x},\mu)\,W^{(k)^{\mathsf{T}}}\right)^{-1}W^{(k)}\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1}}(\boldsymbol{x},\mu)\,\nabla_{\boldsymbol{y}^{(k)}}r\left(\boldsymbol{y}^{(k)}(\boldsymbol{x},\mu),\mu\right)\Big|_{\boldsymbol{y}^{(k)}=\boldsymbol{y}^{(k)^{\star}}(\boldsymbol{x},\mu)}\\ & = & -\left(\mathsf{S}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)\mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}}(\boldsymbol{x},\mu)\right)^{-1}\mathsf{S}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}}(\boldsymbol{x},\mu)\left(\boldsymbol{d}^{(k)}+\mu\mathsf{g}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)\right). \end{array} \tag{13}$$

For the second-stage problem (9), the Lagrangian function is given by

$$\Im\left(\boldsymbol{x},\boldsymbol{y}^{(k)},\boldsymbol{u}^{(k)},\mu\right)=\boldsymbol{d}^{(k)^{\mathsf{T}}}\boldsymbol{y}^{(k)}-\mu\ln\det\left(\boldsymbol{y}^{(k)}\right)+\boldsymbol{u}^{(k)^{\mathsf{T}}}W^{(k)}\boldsymbol{y}^{(k)}-\boldsymbol{u}^{(k)^{\mathsf{T}}}\boldsymbol{q}^{(k)}-\boldsymbol{u}^{(k)^{\mathsf{T}}}T^{(k)^{\mathsf{T}}}\boldsymbol{x}. \tag{14}$$

From (10), we have

$$\begin{array}{l} \nabla_{\boldsymbol{y}^{(k)}}\Im\left(\boldsymbol{x},\boldsymbol{y}^{(k)},\boldsymbol{u}^{(k)},\mu\right)\Big|_{\boldsymbol{y}^{(k)}=\boldsymbol{y}^{(k)^{\star}}(\boldsymbol{x},\mu),\ \boldsymbol{u}^{(k)}=\boldsymbol{u}^{(k)^{\star}}(\boldsymbol{x},\mu)}\\ =\left(\boldsymbol{d}^{(k)}-2\mu J_m^{-1}\boldsymbol{y}^{(k)^{-1}}+W^{(k)}\boldsymbol{u}^{(k)}\right)\Big|_{\boldsymbol{y}^{(k)}=\boldsymbol{y}^{(k)^{\star}}(\boldsymbol{x},\mu),\ \boldsymbol{u}^{(k)}=\boldsymbol{u}^{(k)^{\star}}(\boldsymbol{x},\mu)}\\ =\boldsymbol{d}^{(k)}-2\mu J_m^{-1}\left(\boldsymbol{y}^{(k)^{\star}}(\boldsymbol{x},\mu)\right)^{-1}+W^{(k)}\boldsymbol{u}^{(k)^{\star}}(\boldsymbol{x},\mu)\ =\ \boldsymbol{0}. \end{array} \tag{15}$$

and

$$\begin{array}{l} \nabla_{\boldsymbol{u}^{(k)}}\Im\left(\boldsymbol{x},\boldsymbol{y}^{(k)},\boldsymbol{u}^{(k)},\mu\right)\Big|_{\boldsymbol{y}^{(k)}=y^{(k)^{\star}}(\boldsymbol{x},\mu),\ \boldsymbol{u}^{(k)}=\boldsymbol{u}^{(k)^{\star}}(\boldsymbol{x},\mu)}\\ \qquad=\left(W^{(k)}\boldsymbol{y}^{(k)}-\boldsymbol{q}^{(k)}-T^{(k)}\boldsymbol{x}\right)\Big|_{\boldsymbol{y}^{(k)}=\boldsymbol{y}^{(k)^{\star}}(\boldsymbol{x},\mu)}\\ \qquad=W^{(k)}\boldsymbol{y}^{(k)^{\star}}(\boldsymbol{x},\mu)-\boldsymbol{q}^{(k)}-T^{(k)}\boldsymbol{x}\ =\ \boldsymbol{0}. \end{array} \tag{16}$$

Due to strong duality, we have

$$\rho^{(k)}(\boldsymbol{x},\mu)=\Im\left(\boldsymbol{x},\boldsymbol{y}^{(k)^{\star}}(\boldsymbol{x},\mu),\boldsymbol{u}^{(k)^{\star}}(\boldsymbol{x},\mu),\mu\right). \tag{17}$$

Throughout the rest of this paper, we let $\boldsymbol{y}^{(k)}\triangleq\boldsymbol{y}^{(k)}(\boldsymbol{x},\mu)$, $\boldsymbol{u}^{(k)}\triangleq\boldsymbol{u}^{(k)}(\boldsymbol{x},\mu)$, $\mathsf{H}_{\boldsymbol{y}}^{(k)}\triangleq\mathsf{H}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)$, and $\mathsf{S}_{\boldsymbol{y}}^{(k)}\triangleq\mathsf{S}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x},\mu)$. We need the following intermediate lemma.

Lemma 2. *Let $\boldsymbol{x} \in int\mathcal{I}^n \cap \mathcal{L}^{(0)}$, and $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x}, \mu)$ and $u^{(k)^\star}(\boldsymbol{x}, \mu)$ be the optimal solutions of (9) and (14), respectively. Then*

$$\mathsf{D}_{\boldsymbol{x}} \boldsymbol{y}^{(k)^\star}(\boldsymbol{x}, \mu) = \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \mathsf{S}_{\boldsymbol{y}}^{(k)} \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^\mathsf{T}} \right)^{-1} T^{(k)}, \tag{18}$$

$$\mathsf{D}_{\boldsymbol{x}} \boldsymbol{u}^{(k)^\star}(\boldsymbol{x}, \mu) = -\mu \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^\mathsf{T}} \right)^{-1} T^{(k)}, \tag{19}$$

$$\frac{\partial}{\partial \mu} \boldsymbol{y}^{(k)^\star}(\boldsymbol{x}, \mu) = -\frac{1}{\mu} \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \left(I - \mathsf{P}_{\boldsymbol{y}}^{(k)} \right) \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \mathsf{g}_{\boldsymbol{y}}^{(k)}, \tag{20}$$

$$\frac{\partial}{\partial \mu} \boldsymbol{u}^{(k)^\star}(\boldsymbol{x}, \mu) = -\left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^\mathsf{T}} \right)^{-1} \mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \mathsf{g}_{\boldsymbol{y}}^{(k)}, \tag{21}$$

where $\mathsf{g}_{\boldsymbol{y}}^{(k)}$ and $\mathsf{H}_{\boldsymbol{y}}^{(k)}$ are defined in (11), $\mathsf{S}_{\boldsymbol{y}}^{(k)}$ and $\mathsf{P}_{\boldsymbol{y}}^{(k)}$ are defined in (12).

Proof Note that the Jacobian of $\vartheta(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu)$ with respect to $(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})$ is

$$\begin{aligned}
&\mathsf{D}_{(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})}\, \vartheta\left(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu \right) \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)^\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)^\star}} \\
&= \left[\mathsf{D}_{(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})}\, \vartheta\left(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu \right) \ \vdots\ \mathsf{D}_{\boldsymbol{x}}\, \vartheta\left(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu \right) \right] \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)^\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)^\star}} \\
&= \begin{bmatrix} \mu \mathsf{H}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x}, \mu) & W^{(k)^\mathsf{T}} & \vdots & O \\ W^{(k)} & O & \vdots & -T^{(k)} \end{bmatrix}.
\end{aligned}$$

The matrix $\mathsf{D}_{(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})}\, \vartheta\left(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)} \right)$ is invertible since $\mathsf{H}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x}, \mu)$ positive definite and $W^{(k)}$ has a full rank. In particular, one can verify that

$$\begin{aligned}
&\left[\mathsf{D}_{(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})}\, \vartheta\left(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu \right) \right]^{-1} \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)^\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)^\star}} \\
&= \begin{bmatrix} \frac{1}{\mu} \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \left(I - \mathsf{P}_{\boldsymbol{y}}^{(k)} \right) \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} & \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \mathsf{S}_{\boldsymbol{y}}^{(k)^\mathsf{T}} \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^\mathsf{T}} \right)^{-1} \\ \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^\mathsf{T}} \right)^{-1} \mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} & -\mu \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^\mathsf{T}} \right)^{-1} \end{bmatrix}.
\end{aligned}$$

Since the hypotheses of the implicit function theorem are satisfied at $(\boldsymbol{x}, \boldsymbol{y}^{(k)^\star}, \boldsymbol{u}^{(k)^\star})$ in (10), we have

$$\begin{aligned}
\mathsf{D}_{\boldsymbol{x}} \begin{bmatrix} \boldsymbol{y}^{(k)^\star} \\ \boldsymbol{u}^{(k)^\star} \end{bmatrix} &= -\left[\mathsf{D}_{(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})}\, \vartheta\left(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu \right) \right]^{-1} \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)^\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)^\star}} \\
&\quad \times\ \mathsf{D}_{\boldsymbol{x}}\, \vartheta\left(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu \right) \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)^\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)^\star}},
\end{aligned}$$

giving us the desired results in (18) and (19).

To obtain the results in (20) and (21), note that the Jacobian of $\vartheta(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu)$ with respect to $(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu)$ is

$$\begin{aligned}
&\mathsf{D}_{(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu)}\, \vartheta\Big(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu\Big) \\
&= \left[\mathsf{D}_{(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})}\, \vartheta\Big(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu\Big) \ \vdots\ \mathsf{D}_{\mu}\, \vartheta\Big(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu\Big)\right] \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)\star}} \\
&= \begin{bmatrix} \mu \mathsf{H}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x}, \mu) & W^{(k)^{\mathsf{T}}} & \vdots & \mathsf{g}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x}, \mu) \\ W^{(k)} & O & \vdots & O \end{bmatrix}.
\end{aligned}$$

Again, by the implicit function theorem, the mapping from $(\boldsymbol{x}, \mu)$ to $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x}, \mu)$ and that from $(\boldsymbol{x}, \mu)$ to $\boldsymbol{u}^{(k)^\star}(\boldsymbol{x}, \mu)$ are differentiable in μ with

$$\begin{aligned}
\frac{\partial}{\partial \mu} \begin{bmatrix} \boldsymbol{y}^{(k)^\star}(\boldsymbol{x}, \mu) \\ \boldsymbol{u}^{(k)^\star}(\boldsymbol{x}, \mu) \end{bmatrix} &= -\left[\mathsf{D}_{(\boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)})}\, \vartheta\Big(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu\Big)\right]^{-1} \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)\star}} \\
&\quad \times\ \mathsf{D}_{\mu}\, \vartheta\Big(\boldsymbol{x}, \boldsymbol{y}^{(k)}, \boldsymbol{u}^{(k)}, \mu\Big) \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)\star}, \boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k)\star}},
\end{aligned}$$

giving us the desired results in (20) and (21). □

Lemma 3. *Let $\boldsymbol{x} \in int\mathcal{I}^n \cap \mathcal{L}^{(0)}$, and $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x}, \mu)$ and $\boldsymbol{u}^{(k)^\star}(\boldsymbol{x}, \mu)$ be the optimal solutions of (9) and (14), respectively. Then*

$$\nabla_{\boldsymbol{x}} \eta(\boldsymbol{x}, \mu) = \boldsymbol{c} - 2\mu J_n^{-1} \boldsymbol{x}^{-1} + \sum_{k=1}^{K} \left(T^{(k)^{\mathsf{T}}} \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}} \right)^{-1} \mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \left(\boldsymbol{d}^{(k)} + \mu \mathsf{g}_{\boldsymbol{y}}^{(k)} \right) \right), \tag{22}$$

$$\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \eta(\boldsymbol{x}, \mu) = 2\mu\, J_n^{-1} \mathsf{Q}_{\boldsymbol{x}^{-1}} + \mu \sum_{k=1}^{K} \left(T^{(k)^{\mathsf{T}}} \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}} \right)^{-1} T^{(k)} \right), \tag{23}$$

$$\frac{\partial}{\partial \mu} (\nabla_{\boldsymbol{x}} \eta(\boldsymbol{x}, \mu)) = -2 J_n^{-1} \boldsymbol{x}^{-1} + \sum_{k=1}^{K} \left(T^{(k)^{\mathsf{T}}} \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}} \right)^{-1} \mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \mathsf{g}_{\boldsymbol{y}}^{(k)} \right), \tag{24}$$

$$\begin{aligned}
\frac{\partial}{\partial \mu} \left(\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \eta(\boldsymbol{x}, \mu) \right) &= 2 J_n^{-1} \mathsf{Q}_{\boldsymbol{x}^{-1}} + \sum_{k=1}^{K} \left(\mathsf{R}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}} \left(\mathsf{H}_{y}^{(k)} + \left(\nabla_{y^{(k)}} \mathsf{H}_{\boldsymbol{y}}^{(k)} \right) \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \right.\right. \\
&\times \left.\left. \left(I - \mathsf{P}_{\boldsymbol{y}}^{(k)} \right) \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \mathsf{g}_{\boldsymbol{y}}^{(k)} \right) \mathsf{R}_{\boldsymbol{y}}^{(k)} \right),
\end{aligned} \tag{25}$$

where $\mathsf{g}_{\boldsymbol{y}}^{(k)}$ and $\mathsf{H}_{\boldsymbol{y}}^{(k)}$ are defined in (11), $\mathsf{S}_{\boldsymbol{y}}^{(k)}$ and $\mathsf{P}_{\boldsymbol{y}}^{(k)}$ are defined in (12), and

$$\mathsf{R}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x}, \mu) \triangleq \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1}}(\boldsymbol{x}, \mu)\, W^{(k)^{\mathsf{T}}} \left(\mathsf{S}_{\boldsymbol{y}}^{(k)}(\boldsymbol{x}, \mu)\, \mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}}(\boldsymbol{x}, \mu) \right)^{-1} T^{(k)},\ k = 1, 2, \ldots, K. \tag{26}$$

Proof Using (8) and (18), and applying the chain rule, we have

$$\begin{aligned}
\nabla_{\boldsymbol{x}} \rho^{(k)}(\boldsymbol{x}, \mu) &= \left(\mathsf{D}_{\boldsymbol{x}} \boldsymbol{y}^{(k)^\star}(\boldsymbol{x}, \mu) \right)^{\mathsf{T}} \nabla_{\boldsymbol{y}^{(k)}} \rho^{(k)}(\boldsymbol{x}, \mu) \Big|_{\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)\star}} \\
&= T^{(k)^{\mathsf{T}}} \left(\mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{S}_{\boldsymbol{y}}^{(k)^{\mathsf{T}}} \right)^{-1} \mathsf{S}_{\boldsymbol{y}}^{(k)} \mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}} \left(\boldsymbol{d}^{(k)} + \mu \mathsf{g}_{\boldsymbol{y}}^{(k)} \right).
\end{aligned} \tag{27}$$

Using (14), (15), (16), (17) and (19), and by applying the chain rule, we have

$$\begin{aligned} \nabla^2_{\boldsymbol{xx}}\rho^{(k)}(\boldsymbol{x},\mu) &= \nabla^2_{\boldsymbol{xx}}\Im\Big(\boldsymbol{x},\boldsymbol{y}^{(k)^\star},\boldsymbol{u}^{(k)^\star},\mu\Big) \\ &= \nabla_{\boldsymbol{x}}\Big(\nabla_{\boldsymbol{x}}\Im + \nabla_{\boldsymbol{y}^{(k)}}\Im \mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)} + \nabla_{\boldsymbol{u}^{(k)}}\Im \mathsf{D}_{\boldsymbol{x}}\boldsymbol{u}^{(k)}\Big)\Big|_{\boldsymbol{y}^{(k)}=\boldsymbol{y}^{(k)^\star},\boldsymbol{u}^{(k)}=\boldsymbol{u}^{(k)^\star}} \\ &= \nabla_{\boldsymbol{x}}\Big(-T^{(k)^\mathsf{T}}\boldsymbol{u}^{(k)^\star}(\boldsymbol{x},\mu)\Big) \\ &= -T^{(k)^\mathsf{T}} \times \mathsf{D}_{\boldsymbol{x}}\boldsymbol{u}^{(k)}(\boldsymbol{x},\mu)\Big|_{\boldsymbol{u}^{(k)}=\boldsymbol{u}^{(k)^\star}(\boldsymbol{x},\mu)} = \mu T^{(k)^\mathsf{T}}\Big(\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big)^{-1}T^{(k)}. \end{aligned} \tag{28}$$

The gradient and Hessian in (22) and (23) are immediately obtained by plugging the results in (27), (28) and Lemma 1 into the gradient and Hessian of the function $\eta(\boldsymbol{x},\mu)=\boldsymbol{c}^\mathsf{T}\boldsymbol{x}-\mu\ln\det(\boldsymbol{x})+\sum_{k=1}^K\rho^{(k)}(\boldsymbol{x},\mu)$.
Next, we obtain the partial derivative in (24). Differentiating (22) with respect to μ, and applying the chain rule, we get

$$\frac{\partial}{\partial\mu}(\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x},\mu)) = -2J_n^{-1}x^{-1}+\sum_{k=1}^K\frac{\partial}{\partial\mu}\Big(\nabla_{\boldsymbol{x}}\rho^{(k)}(\boldsymbol{x},\mu)\Big) = -2J_n^{-1}\boldsymbol{x}^{-1}-\sum_{k=1}^K T^{(k)^\mathsf{T}}\frac{\partial}{\partial\mu}\boldsymbol{u}^{(k)^\star}(\boldsymbol{x},\mu). \tag{29}$$

The partial derivative in (24) is now immediately obtained by plugging (21) into (29). Finally, we obtain the partial derivative in (25). Differentiating (23) with respect to μ, using the first equation in (12), and applying the chain rule, we get

$$\begin{aligned} \frac{\partial}{\partial\mu}\left(\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu)\right) &= \frac{\partial}{\partial\mu}\left(2\mu J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}+\sum_{k=1}^K\mu\, T^{(k)^\mathsf{T}}\Big(\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big)^{-1}T^{(k)}\right) \\ &= 2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}} \\ &\quad -\sum_{k=1}^K\left(T^{(k)^\mathsf{T}}\Big(\frac{1}{\mu}\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big)^{-1}\frac{\partial}{\partial\mu}\Big(\frac{1}{\mu}\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big)\Big(\frac{1}{\mu}\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big)^{-1}T^{(k)}\right) \\ &= 2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}} \\ &\quad -\sum_{k=1}^K\left(T^{(k)^\mathsf{T}}\Big(\frac{1}{\mu}\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big)^{-1}W^{(k)}\frac{\partial}{\partial\mu}\Big(\mu\mathsf{H}^{(k)}_{\boldsymbol{y}}\Big)^{-1}W^{(k)^\mathsf{T}}\Big(\frac{1}{\mu}\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big)^{-1}T^{(k)}\right) \\ &= 2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}+\sum_{k=1}^K\left(\mathsf{R}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\frac{\partial}{\partial\mu}\Big(\mu\mathsf{H}^{(k)}_{\boldsymbol{y}}\Big)\mathsf{R}^{(k)}_{\boldsymbol{y}}\right) \\ &= 2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}+\sum_{k=1}^K\left(\mathsf{R}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big(\mathsf{H}^{(k)}_{\boldsymbol{y}}+\mu\frac{\partial}{\partial\mu}\mathsf{H}^{(k)}_{\boldsymbol{y}}\Big)\mathsf{R}^{(k)}_{\boldsymbol{y}}\right) \\ &= 2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}+\sum_{k=1}^K\left(\mathsf{R}^{(k)^\mathsf{T}}_{\boldsymbol{y}}\Big(\mathsf{H}^{(k)}_{\boldsymbol{y}}+\mu\nabla_{\boldsymbol{y}^{(k)}}\mathsf{H}^{(k)}_{\boldsymbol{y}}\frac{\partial}{\partial\mu}\boldsymbol{y}^{(k)^\star}\Big)\mathsf{R}^{(k)}_{\boldsymbol{y}}\right). \end{aligned} \tag{30}$$

The partial derivative in (25) is then immediately obtained by plugging (20) into (30). □

The following corollary is a direct consequence of Lemma 3.

Corollary 1. *The derivatives $\nabla^2_{\boldsymbol{x}\boldsymbol{x}}\eta(\boldsymbol{x},\mu)$ and $\frac{\partial}{\partial\mu}(\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x},\mu))$ can be written as*

$$\nabla^2_{\boldsymbol{x}\boldsymbol{x}}\eta(\boldsymbol{x},\mu) = \hat{I}\;\mathsf{B}_1(\boldsymbol{x},\mu)\mathsf{B}_1{}^{\mathsf{T}}(\boldsymbol{x},\mu)\;\hat{I}^{\mathsf{T}} \quad \text{and} \quad \frac{\partial}{\partial\mu}(\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x},\mu)) = \hat{I}\;\mathsf{B}_1(\boldsymbol{x},\mu)\mathsf{B}_2(\boldsymbol{x},\mu)\mathsf{g}_{\boldsymbol{y}}(\boldsymbol{x},\mu),$$

where $\mathsf{g}_{\boldsymbol{y}}(\boldsymbol{x},\mu) \triangleq (-2J_n^{-1}x^{-1}; \mathsf{g}_{\boldsymbol{y}}^{(1)}(\boldsymbol{x},\mu); \mathsf{g}_{\boldsymbol{y}}^{(2)}(\boldsymbol{x},\mu); \ldots; \mathsf{g}_{\boldsymbol{y}}^{(K)}(\boldsymbol{x},\mu))$, $\hat{I} \triangleq \begin{bmatrix} I_n & I_n & \cdots & I_n \end{bmatrix} \in \mathbb{R}^{n\times(Kn+n)}$*, and* $\mathsf{B}_1(\boldsymbol{x},\mu)$ *and* $\mathsf{B}_2(\boldsymbol{x},\mu)$ *are the block diagonal matrices given by*

$$\mathsf{B}_1(\boldsymbol{x},\mu) \triangleq \sqrt{\mu}\begin{bmatrix} \left(2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\right)^{1/2} & O & \cdots & O \\ O & T^{(1)^{\mathsf{T}}}\left(\mathsf{S}_{\boldsymbol{y}}^{(1)}\mathsf{S}_{\boldsymbol{y}}^{(1)^{\mathsf{T}}}\right)^{-1/2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & T^{(K)^{\mathsf{T}}}\left(\mathsf{S}_{\boldsymbol{y}}^{(K)}\mathsf{S}_{\boldsymbol{y}}^{(K)^{\mathsf{T}}}\right)^{-1/2} \end{bmatrix},$$

$$\mathsf{B}_2(\boldsymbol{x},\mu) \triangleq \frac{1}{\sqrt{\mu}}\begin{bmatrix} \left(2J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\right)^{-1/2} & O & \cdots & O \\ O & \left(\mathsf{S}_{\boldsymbol{y}}^{(1)}\mathsf{S}_{\boldsymbol{y}}^{(1)^{\mathsf{T}}}\right)^{-1/2}\mathsf{S}_{\boldsymbol{y}}^{(1)}\mathsf{H}_{\boldsymbol{y}}^{(1)^{-1/2}} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \left(\mathsf{S}_{\boldsymbol{y}}^{(K)}\mathsf{S}_{\boldsymbol{y}}^{(K)^{\mathsf{T}}}\right)^{-1/2}\mathsf{S}_{\boldsymbol{y}}^{(K)}\mathsf{H}_{\boldsymbol{y}}^{(K)^{-1/2}} \end{bmatrix}.$$

Now, we are ready to state and prove the self-concordant properties of the family of composite barrier functions $\{\eta^{(k)}(\boldsymbol{x},\mu) : \mu > 0\}$. We have the following definition. This definition uses $\mathbb{R}_{++}$ to denote the set of all positive real numbers.

Definition 2 (Definition 3.1.1 in [21]). Let G be an open nonempty convex subset of $\mathbb{R}^n$. Let also $\mu \in \mathbb{R}_{++}$ and $f_\mu : \mathbb{R}_{++} \times G \to \mathbb{R}$ be a family of functions indexed by μ. Let $\alpha_1(\mu), \alpha_2(\mu), \alpha_3(\mu), \alpha_4(\mu), \alpha_5(\mu) : \mathbb{R}_{++} \longrightarrow \mathbb{R}_{++}$ be continuously differentiable functions on μ. Then the family of functions $f_{\mu\in\mathbb{R}_{++}}$ is called *strongly self-concordant with the parameters* $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$, if the following conditions hold:

(i) The function f_μ is continuous on $\mathbb{R}_{++} \times G$, and for fixed $\mu \in \mathbb{R}_{++}$, f_μ is convex on G and has three partial derivatives on G, which are continuous on $\mathbb{R}_{++} \times G$ and continuously differentiable with respect to μ on $\mathbb{R}_{++}$.

(ii) For any $\mu \in \mathbb{R}_{++}$, the function f_μ is strongly $\alpha_1(\mu)$-self-concordant.

(iii) For any $(\mu, \boldsymbol{x}) \in \mathbb{R}_{++} \times G$ and any $\boldsymbol{h} \in \mathbb{R}^n$,

$$\begin{aligned} \left|\frac{\partial}{\partial\mu}\left(\boldsymbol{h}^{\mathsf{T}}\nabla_{\boldsymbol{x}}f_\mu(\boldsymbol{x},\mu)\right) - \frac{\partial}{\partial\mu}(\ln\alpha_3(\mu))\,\boldsymbol{h}^{\mathsf{T}}\nabla_{\boldsymbol{x}}f_\mu(\boldsymbol{x},\mu)\right| &\le \alpha_4(\mu)(\alpha_1(\mu))^{1/2}\left(\boldsymbol{h}^{\mathsf{T}}\nabla^2_{\boldsymbol{x}\boldsymbol{x}}f_\mu(\boldsymbol{x},\mu)\boldsymbol{h}\right)^{1/2}, \\ \left|\frac{\partial}{\partial\mu}\left(\boldsymbol{h}^{\mathsf{T}}\nabla^2_{\boldsymbol{x}\boldsymbol{x}}f_\mu(\boldsymbol{x},\mu)\boldsymbol{h}\right) - \frac{\partial}{\partial\mu}(\ln\alpha_2(\mu))\,\boldsymbol{h}^{\mathsf{T}}\nabla^2_{\boldsymbol{x}\boldsymbol{x}}f_\mu(\boldsymbol{x},\mu)\;\boldsymbol{h}\right| &\le 2\alpha_5(\mu)\;\boldsymbol{h}^{\mathsf{T}}\nabla^2_{\boldsymbol{x}\boldsymbol{x}}f_\mu(x,\mu)\;\boldsymbol{h}. \end{aligned}$$

The following theorem contains a fundamental result because it specifies appropriate barrier parameters to comprise a self-concordant family from the set of composite barrier functions. This enables us to prove the polynomiality of the proposed algorithms.

Theorem 2. *The family $\{\eta(.,\mu) : \mu > 0\}$ is a strongly self-concordant family with the following parameters*

$$\alpha_1(\mu) = \mu,\;\; \alpha_2(\mu) = \alpha_3(\mu) = 1,\;\; \alpha_4(\mu) = \frac{\sqrt{2((n-1)+K(m-1))}}{\mu},\;\; \alpha_5(\mu) = \frac{1+2^{3/2}\sqrt{m-1}}{2\mu}.$$

The proof of Theorem 2 depends on a sequence of some intermediate lemmas that we state and prove below.

Lemma 4. *For any fixed $\mu > 0$, the barrier recourse function $\rho^{(k)}(\boldsymbol{x}, \mu)$ is strongly μ-self-concordant on $\mathcal{F}^{(k)}$ for $k = 1, 2, \ldots, K$.*

Proof Let $\mu > 0$ be fixed, $\boldsymbol{y}^{(k)} = \boldsymbol{y}^{(k)}(\boldsymbol{x}, \mu) \in \mathcal{E}^m$, and $\boldsymbol{h} \in \mathbb{R}^m$. Then

$$
\begin{aligned}
\left|\nabla^3_{\boldsymbol{xxx}} \rho^{(k)}(x,\mu)[\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}]\right| &= \left|\nabla^3_{\boldsymbol{xxx}} \left(\boldsymbol{d}^{(k)\mathsf{T}} \boldsymbol{y}^{(k)}(\boldsymbol{x},\mu) - \mu \ln\det\left(\boldsymbol{y}^{(k)}(\boldsymbol{x},\mu)\right)\right)[\boldsymbol{h},\boldsymbol{h},\boldsymbol{h}]\right| \\
&= \left|\nabla^3_{\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)}} \left(\boldsymbol{d}^{(k)\mathsf{T}} \boldsymbol{y}^{(k)} - \mu \ln\det\left(\boldsymbol{y}^{(k)}\right)\right)\left[\mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)}\boldsymbol{h}, \mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)}\boldsymbol{h}, \mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)}\boldsymbol{h}\right]\right| \\
&= \mu \left|\nabla^3_{\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)}} \ell\left(\boldsymbol{y}^{(k)}\right)\left[\mathsf{D}_{\boldsymbol{x}}y^{(k)}\boldsymbol{h}, \mathsf{D}_{\boldsymbol{x}}y^{(k)}\boldsymbol{h}, \mathsf{D}_{\boldsymbol{x}}y^{(k)}\boldsymbol{h}\right]\right| \\
&\le 2\mu \left(\nabla^2_{\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)}} \ell\left(\boldsymbol{y}^{(k)}\right)\left[\mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)}\boldsymbol{h}, \mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)}\boldsymbol{h}\right]\right)^{3/2} \\
&= \frac{2}{\sqrt{\mu}} \left(\nabla^2_{\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)}} \left(\boldsymbol{d}^{(k)\mathsf{T}} \boldsymbol{y}^{(k)} - \mu \ln\det\left(\boldsymbol{y}^{(k)}\right)\right)\left[\mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)}\boldsymbol{h}, \mathsf{D}_{\boldsymbol{x}}\boldsymbol{y}^{(k)}h\right]\right)^{3/2} \\
&= \frac{2}{\sqrt{\mu}} \left(\nabla^2_{\boldsymbol{xx}} \left(\boldsymbol{d}^{(k)\mathsf{T}} \boldsymbol{y}^{(k)} - \mu \ln\det\left(\boldsymbol{y}^{(k)}\right)\right)[\boldsymbol{h},\boldsymbol{h}]\right)^{3/2} \\
&= \frac{2}{\sqrt{\mu}} \left(\nabla^2_{\boldsymbol{xx}} \rho^{(k)}(\boldsymbol{x},\mu)[\boldsymbol{h},\boldsymbol{h}]\right)^{3/2},
\end{aligned}
$$

where the inequality holds since $\ell(\cdot)$ is a self-concordant barrier of complexity value 1 (see Theorem 1). Therefore, the inequality in (7) holds for $\rho^{(k)}(\boldsymbol{x}, \mu)$. Finally, for any sequence $\{\boldsymbol{x}_i\}_{i=1}^{\infty}$ in $\mathcal{F}^{(k)}$ approaching a point from boundary of $\mathcal{F}^{(k)}$, the map $\rho^{(k)}(\boldsymbol{x}_i, \mu)$ approaches infinity. Thus, $\rho^{(k)}(\boldsymbol{x}, \mu)$ is strongly μ-self-concordant on $\mathcal{F}^{(k)}$ for $k = 1, 2, \ldots, K$. □

Lemma 5. *For any fixed $\mu > 0$, the composite barrier function $\eta(\cdot, \mu)$ is strongly μ-self-concordant on $\mathcal{F}$.*

Proof Recall that $\eta(\boldsymbol{x}, \mu) = \boldsymbol{c}^\mathsf{T}\boldsymbol{x} - \mu \ln\det(\boldsymbol{x}) + \sum_{k=1}^{K} \rho^{(k)}(\boldsymbol{x}, \mu)$. It is trivial to show that the linear map $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}$ is strongly μ-self-concordant on $\mathcal{I}^n \cap \mathcal{L}^{(0)}$ (both sides of the inequality in (7) are simply zeros). Theorem 1 shows that the barrier $-\mu \ln\det(\boldsymbol{x})$ is strongly μ-self-concordant on $\mathcal{I}^n \cap \mathcal{L}^{(0)}$, and Lemma 4 shows that the map $\rho^{(k)}(\boldsymbol{x}, \mu)$ is strongly μ-self-concordant on $\mathcal{F}^{(k)}$. The result then immediately follows from [21, Proposition 2.1.1(ii)]. □

Lemma 6. *For any $\mu > 0, \boldsymbol{x} \in \mathcal{F}$ and $\boldsymbol{h} \in \mathbb{R}^n$, we have*

$$\left| \frac{\partial}{\partial \mu} \left(\boldsymbol{h}^{\mathsf{T}} \nabla_{\boldsymbol{x}} \eta(\boldsymbol{x}, \mu) \right) \right| \leq \sqrt{\frac{2((n-1)+K(m-1))}{\mu}} \sqrt{\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu)[\boldsymbol{h}, \boldsymbol{h}]}, \tag{31}$$

$$\left| \frac{\partial}{\partial \mu} \left(\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu) \, [\boldsymbol{h}, \boldsymbol{h}] \right) \right| \leq \frac{1 + 2^{3/2}\sqrt{m-1}}{\mu} \nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu)[\boldsymbol{h}, \boldsymbol{h}]. \tag{32}$$

Proof This proof uses the block vector $\mathsf{g}_{\boldsymbol{y}} \triangleq \mathsf{g}_{\boldsymbol{y}}(\boldsymbol{x}, \mu)$, the block-diagonal matrices $\hat{I}$, $\mathsf{B}_1 \triangleq \mathsf{B}_1(\boldsymbol{x}, \mu)$ and $\mathsf{B}_2 \triangleq \mathsf{B}_2(\boldsymbol{x}, \mu)$ defined in Corollary 1, and the orthogonal projection matrix $\mathsf{P}^{(k)}_{\boldsymbol{y}}$ defined in (12). To prove (31), by using Corollary 1, we have

$$\begin{aligned}
\left| \frac{\partial}{\partial \mu} \left(\boldsymbol{h}^{\mathsf{T}} \nabla_{\boldsymbol{x}} \eta(\boldsymbol{x}, \mu) \right) \right| &= \left| \boldsymbol{h}^{\mathsf{T}} \hat{I} \, \mathsf{B}_1 \mathsf{B}_2 \mathsf{g}_{\boldsymbol{y}} \right| \\
&\leq \sqrt{\boldsymbol{h}^{\mathsf{T}} \hat{I} \, \mathsf{B}_1(\boldsymbol{x}, \mu) {\mathsf{B}_1}^{\mathsf{T}}(\boldsymbol{x}, \mu) \, \hat{I}^{\mathsf{T}} \boldsymbol{h}} \sqrt{\mathsf{g}^{\mathsf{T}}_{\boldsymbol{y}}(\boldsymbol{x}, \mu) {\mathsf{B}_2}^{\mathsf{T}}(\boldsymbol{x}, \mu) \mathsf{B}_2(\boldsymbol{x}, \mu) \mathsf{g}_{\boldsymbol{y}}(\boldsymbol{x}, \mu)} \\
&= \sqrt{\boldsymbol{h}^{\mathsf{T}} \nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu) \boldsymbol{h}} \sqrt{\mathsf{g}^{\mathsf{T}}_{\boldsymbol{y}}(\boldsymbol{x}, \mu) {\mathsf{B}_2}^{\mathsf{T}}(\boldsymbol{x}, \mu) \mathsf{B}_2(\boldsymbol{x}, \mu) \mathsf{g}_{\boldsymbol{y}}(\boldsymbol{x}, \mu)} \\
&\leq \sqrt{\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu)[\boldsymbol{h}, \boldsymbol{h}]} \\
&\quad \times \sqrt{\frac{1}{\mu} \left(2 \left(J_n^{-1} \boldsymbol{x}^{-1} \right)^{\mathsf{T}} \left((\mathsf{Q}_{\boldsymbol{x}^{-1}})^{-1} \boldsymbol{x}^{-1} \right) + \sum_{k=1}^{K} \left({\mathsf{g}^{(k)}_{\boldsymbol{y}}}^{\mathsf{T}} \left({\mathsf{H}^{(k)}_{\boldsymbol{y}}}^{-1} \mathsf{g}^{(k)}_{\boldsymbol{y}} \right) \right) \right)} \\
&= \sqrt{\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu)[\boldsymbol{h}, \boldsymbol{h}]} \\
&\quad \times \sqrt{\frac{1}{\mu} \left(2\boldsymbol{x}^{-1} \blacksquare \left((\mathsf{Q}_{\boldsymbol{x}^{-1}})^{-1} \boldsymbol{x}^{-1} \right) + \sum_{k=1}^{K} \left(\left(J_m \mathsf{g}^{(k)}_{\boldsymbol{y}} \right) \blacksquare \left({\mathsf{H}^{(k)}_{\boldsymbol{y}}}^{-1} \mathsf{g}^{(k)}_{\boldsymbol{y}} \right) \right) \right)} \\
&= \sqrt{\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu)[\boldsymbol{h}, \boldsymbol{h}]} \sqrt{\frac{1}{\mu} \left(\mathrm{trace}(\boldsymbol{e}_n) + \sum_{k=1}^{K} \mathrm{trace}(\boldsymbol{e}_m) \right)} \\
&= \sqrt{\frac{2((n-1)+K(m-1))}{\mu}} \sqrt{\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}, \mu)[\boldsymbol{h}, \boldsymbol{h}]},
\end{aligned} \tag{33}$$

where the second inequality follows from the fact that $\mathsf{P}^{(k)}_{\boldsymbol{y}}$ is an orthogonal projection matrix, and the third and forth equalities follow from (6) and (5), respectively. This proves (31).

To prove (32), let $\boldsymbol{h} \in \mathbb{R}^n$ and $\mathsf{R}^{(k)}_{\boldsymbol{yy}} \triangleq \mathsf{R}^{(k)}_{\boldsymbol{y}} \boldsymbol{h}$ where $\mathsf{R}^{(k)}_{\boldsymbol{y}}$ is the matrix defined in (26). Then, using the last equality in (30) and (20), one can show that

$$
\begin{aligned}
\left|\frac{\partial}{\partial\mu}\left(\boldsymbol{h}^{\mathsf T}\nabla^2_{\boldsymbol{x}\boldsymbol{x}}\eta(\boldsymbol{x},\mu)\boldsymbol{h}\right)\right| &:= \Bigg|2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}\\
&+\sum_{k=1}^{K}\left({\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}+{\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\left(\mu\nabla_{\boldsymbol{y}^{(k)}}\mathsf{H}^{(k)}_{\boldsymbol{y}}\frac{\partial}{\partial\mu}\boldsymbol{y}^{(k)\star}\right)\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right)\Bigg|\\
&\le\left|2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}\right|\\
&+\sum_{k=1}^{K}\left(\left|{\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right|+\left|\nabla^3_{\boldsymbol{y}^{(k)}}\ell^{(k)}_{\boldsymbol{y}}\left[\mu\frac{\partial}{\partial\mu}\boldsymbol{y}^{(k)\star}(\boldsymbol{x},\mu),\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}},\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right]\right|\right)\\
&\le 2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}+\sum_{k=1}^{K}{\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\\
&+\sum_{k=1}^{K}2\sqrt{\nabla^2_{\boldsymbol{y}^{(k)}}\ell^{(k)}_{\boldsymbol{y}}\left[\mu\frac{\partial}{\partial\mu}\boldsymbol{y}^{(k)\star}(\boldsymbol{x},\mu),\mu\frac{\partial}{\partial\mu}\boldsymbol{y}^{(k)\star}(\boldsymbol{x},\mu)\right]}\Bigg){\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\\
&\le 2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}\boldsymbol{x}^{-1}}\boldsymbol{h}\\
&+\sum_{k=1}^{K}\left({\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}+2\sqrt{{\mathsf{g}^{(k)}_{\boldsymbol{y}}}^{\mathsf T}{\mathsf{H}^{(k)}_{\boldsymbol{y}}}^{-1}\mathsf{g}^{(k)}_{\boldsymbol{y}}}\,{\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right).
\end{aligned}
$$

where we used [21, Proposition 9.1.1] to obtain the second inequality.
Now, using (6) and (5), respectively, it follows that

$$
\begin{aligned}
\left|\frac{\partial}{\partial\mu}\left(\boldsymbol{h}^{\mathsf T}\nabla^2_{\boldsymbol{x}\boldsymbol{x}}\eta(\boldsymbol{x},\mu)\boldsymbol{h}\right)\right| &= 2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}\\
&+\sum_{k=1}^{K}\left(\left(1+2\sqrt{\left(J_m\mathsf{g}^{(k)}_{\boldsymbol{y}}\right)\blacksquare\left({\mathsf{H}^{(k)}_{\boldsymbol{y}}}^{-1}\mathsf{g}^{(k)}_{\boldsymbol{y}}\right)}\right){\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right)\\
&= 2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}\\
&+\sum_{k=1}^{K}\left(\left(1+2\sqrt{2\boldsymbol{y}^{(k)-1}\blacksquare\mathsf{Q}^{-1}_{\boldsymbol{y}^{(k)-1}}\boldsymbol{y}^{(k)-1}}\right){\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right)\\
&= 2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}\\
&+\sum_{k=1}^{K}\left(\left(1+2\sqrt{\operatorname{trace}(\boldsymbol{y}^{(k)-1}\,\square\,\mathsf{Q}^{-1}_{\boldsymbol{y}^{(k)-1}}\boldsymbol{y}^{(k)-1})}\right){\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right)\\
&= 2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}\\
&+\sum_{k=1}^{K}\left(\left(1+2\sqrt{\operatorname{trace}(\boldsymbol{e}_m)}\right){\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right)\\
&= 2\boldsymbol{h}^{\mathsf T}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h}+\sum_{k=1}^{K}\left(\left(1+2\sqrt{2(m-1)}\right){\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}\right).
\end{aligned}
$$

Note that by (26) and the fact that $W^{(k)}{\mathsf{H}^{(k)}_y}^{-1}{W^{(k)}}^{\mathsf T}=\mathsf{S}^{(k)}_y{\mathsf{S}^{(k)}_y}^{\mathsf T}$, we have

$$
\begin{aligned}
{\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}}}^{\mathsf T}\mathsf{H}^{(k)}_{\boldsymbol{y}}\mathsf{R}^{(k)}_{\boldsymbol{y}\boldsymbol{y}} &= \boldsymbol{h}^{\mathsf T}{T^{(k)}}^{\mathsf T}\left(\mathsf{S}^{(k)}_{\boldsymbol{y}}{\mathsf{S}^{(k)}_{\boldsymbol{y}}}^{\mathsf T}\right)^{-1}W^{(k)}{\mathsf{H}^{(k)}_{\boldsymbol{y}}}^{-1}{W^{(k)}}^{\mathsf T}\left(\mathsf{S}^{(k)}_{\boldsymbol{y}}{\mathsf{S}^{(k)}_{\boldsymbol{y}}}^{\mathsf T}\right)^{-1}T^{(k)}\boldsymbol{h}\\
&= \boldsymbol{h}^{\mathsf T}{T^{(k)}}^{\mathsf T}\left(\mathsf{S}^{(k)}_{\boldsymbol{y}}{\mathsf{S}^{(k)}_{\boldsymbol{y}}}^{\mathsf T}\right)^{-1}T^{(k)}\boldsymbol{h}. \qquad (34)
\end{aligned}
$$

Using (23), it follows that

$$\begin{aligned}\left|\frac{\partial}{\partial\mu}\left(h^{\mathsf{T}}\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu)\,\boldsymbol{h}\right)\right| &= \boldsymbol{h}^{\mathsf{T}}J_n^{-1}\mathsf{Q}_{\boldsymbol{x}^{-1}}\boldsymbol{h} \\ &\quad+\sum_{k=1}^{K}\left(\left(1+2\sqrt{\operatorname{trace}(2(m-1))}\right)\boldsymbol{h}^{\mathsf{T}}T^{(k)^{\mathsf{T}}}\left(\mathsf{S}^{(k)}_{\boldsymbol{y}}\mathsf{S}^{(k)^{\mathsf{T}}}_{\boldsymbol{y}}\right)^{-1}T^{(k)}\boldsymbol{h}\right) \\ &\le \frac{1+2\sqrt{2(m-1)}}{\mu}\,\boldsymbol{h}^{\mathsf{T}}\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu)\boldsymbol{h}. \qquad \square\end{aligned}$$

Proof of Theorem 2. Condition (i) in Definition 2 is satisfied. Lemma 5 satisfies Condition (ii), and Lemma 6 shows that Condition (iii) holds. □

5. The algorithm and its complexity

In this section, we present a path-following primal interior-point algorithm for the two-stage SINP problem and see that the short- and long-step versions of the proposed algorithm obtain an ε-optimal solution in polynomial number of first-stage Newton iterations. This analysis assumes that the second stage barrier problems are solved exactly, and hence $\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x},\mu)$ and $\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu)$ are computed exactly as shown earlier in Lemma 3.

The first-stage Newton step $\triangle\boldsymbol{x}$ is defined at a feasible solution $\boldsymbol{x}$ of the problem $\{min\ \ \eta(\boldsymbol{x},\mu)\mid A\boldsymbol{x}=\boldsymbol{b}\}$ as

$$\begin{aligned}\triangle\boldsymbol{x} \triangleq &-(\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu))^{-1}\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x},\mu)\\ &+(\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu))^{-1}A^{\mathsf{T}}\left(A\,(\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu))^{-1}A^{\mathsf{T}}\right)^{-1}A\,(\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu))^{-1}\,\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x},\mu),\end{aligned} \tag{35}$$

where (35) is a closed solution of the system:

$$\begin{aligned}\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu)\triangle\boldsymbol{x}+A^{\mathsf{T}}\triangle v &= -\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x},\mu),\\ A\triangle\boldsymbol{x} &= 0.\end{aligned} \tag{36}$$

We also define

$$\delta(\boldsymbol{x},\mu)\triangleq\sqrt{\frac{1}{\mu}\triangle\boldsymbol{x}^{\mathsf{T}}\nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu)\triangle\boldsymbol{x}}. \tag{37}$$

The algorithm is formally stated in Algorithm 1 and is graphically visualized in Figure 2.

Algorithm 1 starts with $(\boldsymbol{x}^0,\mu^0)$, where $\boldsymbol{x}^0$ satisfies $\delta(\boldsymbol{x}^0,\mu^0)<\kappa\triangleq(2-\sqrt{3})/2$. It generates a sequence of $(\boldsymbol{x}^k,\mu^k)$ with $\mu^{k+1}=\varpi\mu^k$ until $\mu^k<\varepsilon$. The algorithm needs to ensure that the proximity condition of $\boldsymbol{x}^k$ to $\boldsymbol{x}^\star(\mu^k)$ is maintained by using the criteria $\delta(\boldsymbol{x}^k,\mu^k)<\kappa$. The process of updating $(\boldsymbol{x}^k,\mu^k)$ to $(\boldsymbol{x}^{k+1},\mu^{k+1})$ is called an *outer iteration*. The value $\mu^N=\varpi\mu^0<\epsilon$ is achieved after N outer iterations, where

$$N\le\ln\left(\frac{\mu^0}{\epsilon}\right)\Big/\left(\ln\varpi^{-1}\right)+1.$$

Algorithm 1: The primal interior-point decomposition algorithm for two-stage SINP problem.

Initialize $i = 0, \boldsymbol{x}^0, \mu^0, \varpi, \epsilon$; **Ensure** $\mu^0, \varpi \in (0,1), \epsilon \in (0,1), \boldsymbol{x}^0$ **is feasible,** $\delta(\boldsymbol{x}^0, \mu^0) \leq \kappa$;
while $\mu^i > \epsilon$ **do**
  $\mu^{i+1} \triangleq \varpi \mu^i$;
  $j \triangleq 0, x^{i0} \triangleq x^i$;
  for $k = 1, 2, \ldots, K$ **do**
    solve subproblems $\rho^{(k)}(\boldsymbol{x}^{ij}, \mu^{i+1})$ **to obtain** $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x}^{ij}, \mu^{i+1})$ **and** $\boldsymbol{u}^{(k)^\star}(\boldsymbol{x}^{ij}, \mu^{i+1})$;
  compute $\nabla_{\boldsymbol{x}} \eta(\boldsymbol{x}^{ij}, \mu^{i+1})$ **using (22)** ;
  compute $\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}^{ij}, \mu^{i+1})$ **using (23)** ;
  compute the Newton direction $\triangle^{ij} x$ **using (35)** ;
  compute $\delta(\boldsymbol{x}^{ij}, \mu^{i+1})$ **using (37)**;
  while $\delta(\boldsymbol{x}^{ij}, \mu^{i+1}) > \kappa$ **do**
    perform line search $\theta(\geq 0)$ **to minimize** $\eta(\boldsymbol{x}^{ij} + \theta \triangle^{ij} \boldsymbol{x})$;
    $\boldsymbol{x}^{i(j+1)} \triangleq \boldsymbol{x}^{ij} + \theta \triangle^{ij} \boldsymbol{x}$;
    for $k = 1, 2, \ldots, K$ **do**
      solve subproblems $\rho^{(k)}(\boldsymbol{x}^{ij}, \mu^{i+1})$ **to obtain** $\boldsymbol{y}^{(k)^\star}(\boldsymbol{x}^{ij}, \mu^{i+1})$ **and** $\boldsymbol{u}^{(k)^\star}(\boldsymbol{x}^{ij}, \mu^{i+1})$;
    compute $\nabla_{\boldsymbol{x}} \eta(\boldsymbol{x}^{ij}, \mu^{i+1})$ **using (22)**;
    compute $\nabla^2_{\boldsymbol{xx}} \eta(\boldsymbol{x}^{ij}, \mu^{i+1})$ **using (23)** ;
    compute the Newton direction $\triangle^{ij} \boldsymbol{x}$ **using (35)** ;
    compute $\delta(\boldsymbol{x}^{ij}, \mu^{i+1})$ **using (37)** ;
    $j \triangleq j + 1$;
  $\boldsymbol{x}^{i+1} \triangleq \boldsymbol{x}^{ij}$;
  $i \triangleq i + 1$;

Table 3: Comparison of some features between the short- and long-step algorithms for SINP.

Feature	Short-step algorithm	Long-step algorithm
Factor ϖ	$\varpi = 1 - \iota/\sqrt{n + Km}$, $\iota \leq 0.0755$	Constant rate $\varpi \in (0,1)$
Inner iterations	Single inner iteration	Several inner iterations
Outer iterations	$\mathcal{O}(\sqrt{n + Km} \ln(\mu^0/\varepsilon))$	$\mathcal{O}((n + Km) \ln(\mu^0/\varepsilon))$

Let $\boldsymbol{x}^{k0} = \boldsymbol{x}^k$. The long-step algorithm generates a sequence $\boldsymbol{x}^{kj}$, $j = 1, 2, \ldots, M$, till $\boldsymbol{x}^{kM}$ satisfies the desired condition $\delta(\boldsymbol{x}^{kM}, \mu^{k+1}) < \kappa$. The process of updating $(\boldsymbol{x}^{kj}, \mu^{k+1})$ to $(\boldsymbol{x}^{kj+1}, \mu^{k+1})$ is called an inner iteration. After updating $k\mu^k$, the short-step algorithm restores the proximity condition in only one step, but the long-step algorithm may perform many steps to restore this proximity. The following theorem states the complexity result for the short-step algorithm.

Theorem 3. *Let μ^0 be the initial barrier parameter, $\mu^{k+1} = \varpi \mu^k$, and ε be the target precision. If $\delta(\boldsymbol{x}^0, \mu^0) \leq \kappa = (2 - \sqrt{3})/2$, $\varpi = 1 - \iota/\sqrt{n + Km}$, where $0 < \iota \leq 0.0755$, then the short-step algorithm terminates with $(\boldsymbol{x}^N, \mu^N)$ satisfying $\delta(\boldsymbol{x}^N, \mu^N) \leq \kappa$, and $\mu^N \leq \varepsilon$ in $\mathcal{O}(\sqrt{n + Km} \ln(\mu^0/\varepsilon))$ outer iterations. Each inner iteration requires calculation of a single Newton direction by solving (36).*

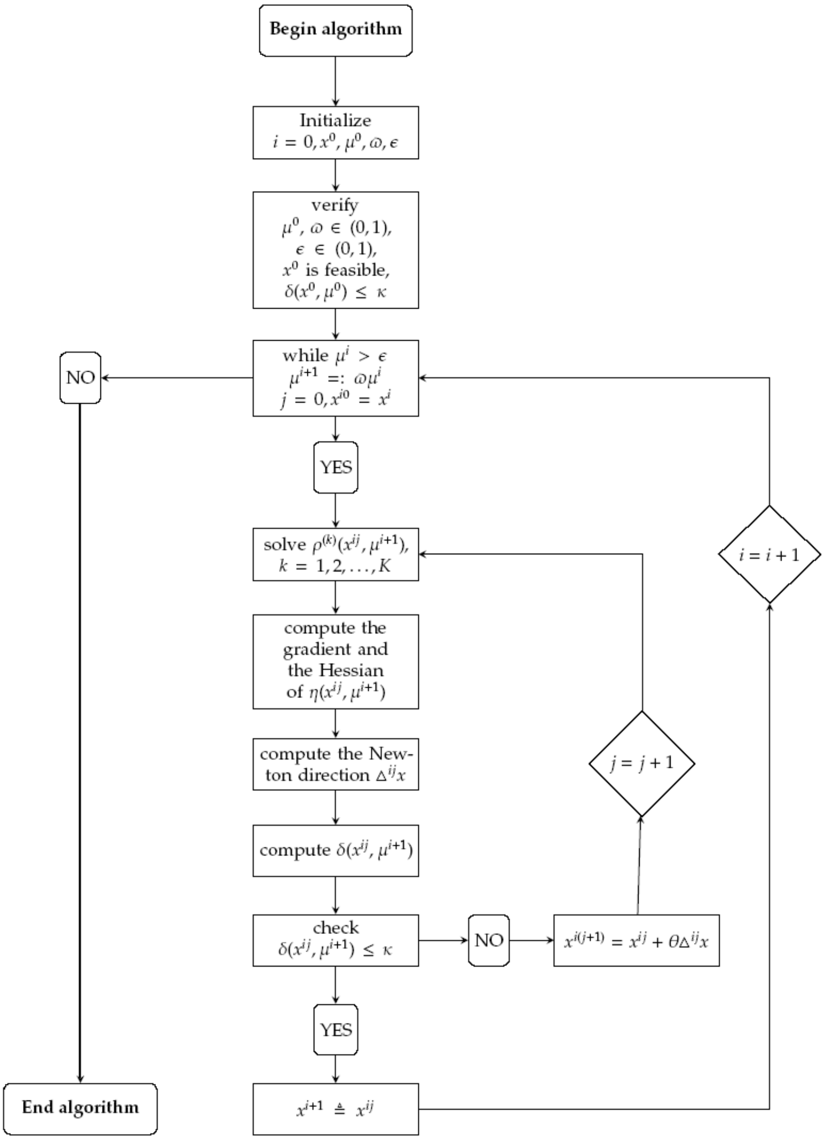


Figure 2: A flowchart of Algorithm 1.

Proof: See Sub-appendix A.1. □

The long-step algorithm takes a constant value for ϖ, say $\varpi = 0.1$, and may perform several inner iterations to restore the condition $\delta(\boldsymbol{x}^{iM}, \mu^{i+1}) < \kappa$. The following theorem states the complexity result for the long-step algorithm.

Theorem 4. *Let μ^0 be the initial barrier parameter, $\mu^{k+1} = \varpi\mu^k$, and ε be the target precision. If $\delta(\boldsymbol{x}^0, \mu^0) \leq \kappa = (2-\sqrt{3})/2$, and the long-step algorithm reduces μ^k at a constant rate ϖ, where $0 < \varpi < 1$, then the long-step algorithm terminates with $(\boldsymbol{x}^N, \mu^N)$ satisfying $\delta(\boldsymbol{x}^N, u^N) \leq \kappa$, and $\mu^N \leq \varepsilon$ in $\mathcal{O}((n + Km)\ln(\mu^0/\varepsilon))$ inner iterations. Each inner iteration requires calculation of a single Newton direction by solving (36).*

Proof: See Sub-appendix A.2. □

Table 3 compares some features between the two variants of the algorithm. From Theorems 3 and 4, it is clear that the dominant terms in the complexity expressions are given in terms of the number of realizations, and, most notably the ranks of the underlying infinity norm cones. The complexity results in Theorems 3 and 4 are the counterparts of those in Theorems 1 and 2 in [24] for two-stage stochastic linear programs with recourse those in Theorems 4.1 and 4.2 in [1] for two-stage stochastic second-order cone programs with recourse, and those in Theorems 4.1 and 4.2 in [20] for two-stage stochastic semidefinite programs with recourse. It is interesting to note there is a complete matching in terms of rank between our complexity results and their counterpart's complexity results found in [1, 20, 24] for other stochastic conic programming (see Table 4). This matching is despite that the asymmetry of the infinity norm cone when it is compared with the symmetry of the three other cones shown in Table 4.

Table 4: Comparing long-step algorithm complexities of some two-stage stochastic conic programs with K number of realizations.

Two-stage stochastic conic program	Cone rank	Complexity of the long-step algorithm
Nonnegative orthant cone: $\boldsymbol{x} \in \mathbb{R}^n_+, \boldsymbol{y}^{(k)} \in \mathbb{R}^m_+$	$\mathsf{rk}(\mathbb{R}^n_+) = \mathcal{O}(n)$	$\mathcal{O}\left((\mathsf{rk}(\mathbb{R}^n_+) + K\,\mathsf{rk}(\mathbb{R}^m_+))\ln(\mu^0/\varepsilon)\right)$
Second-order cone: $\boldsymbol{x} \in \mathcal{C}^n_2, \boldsymbol{y}^{(k)} \in \mathcal{C}^m_2$	$\mathsf{rk}(\mathcal{C}^n_2) = \mathcal{O}(1)$	$\mathcal{O}\left((\mathsf{rk}(\mathcal{C}^n_2) + K\,\mathsf{rk}(\mathcal{C}^m_2))\ln(\mu^0/\varepsilon)\right)$
Semidefinite cone: $X \in \mathcal{S}^n_+, Y^{(k)} \in \mathcal{S}^m_+$	$\mathsf{rk}(\mathcal{S}^n_+) = \mathcal{O}(n)$	$\mathcal{O}\left((\mathsf{rk}(\mathcal{S}^n_+) + K\,\mathsf{rk}(\mathcal{S}^m_+))\ln(\mu^0/\varepsilon)\right)$
Infinity norm cone: $\boldsymbol{x} \in \mathcal{I}^n, \boldsymbol{y}^{(k)} \in \mathcal{I}^m$	$\mathsf{rk}(\mathcal{I}^n) = \mathcal{O}(n)$	$\mathcal{O}\left((\mathsf{rk}(\mathcal{I}^n) + K\,\mathsf{rk}(\mathcal{I}^m))\ln(\mu^0/\varepsilon)\right)$

6. Numerical results

In order to see how the algorithm proposed in this paper works, it has been implemented to solve numerical examples. In this section, we present two numerical examples to show the computational performance of the long-step algorithm. In Example 1, we test the proposed algorithm on the two-stage stochastic facility location problem. In Example 2, we test the proposed algorithm on randomly-generated problems. Numerical results were obtained using MATLAB R2018a (Version: 9.4.0.813654) and on Windows XP Enterprise 64-bit operating system.

Example 1 (Stochastic uniform facility location problems). We consider instances of the stochastic uniform facility location problem formulated as an SINP problem. One way of classifying facility location problems is based on the distance measures. There are three distance measures: Manhattan distance (measured by l_1-norm), Euclidean distance (measured by l_2-norm), and Chebyshev distance (measured by l_∞-norm). In this paper, we

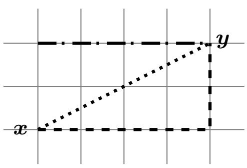
are interested in the so-called uniform facility location problem which uses the Chebyshev distance (see Figure 3).
Assume that we are given f existing fixed facilities with coordinates represented by fixed points, say $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_f \in \mathbb{R}^n$, and r random fixed facilities with coordinates represented by random points, say $\boldsymbol{b}_1(\omega), \boldsymbol{b}_2(\omega), \ldots, \boldsymbol{b}_r(\omega) \in \mathbb{R}^n$, whose realizations depend on underlying outcomes ω in an event space Ω with a known probability measure P. In the two-stage stochastic uniform facility location problem, we plan to add a new facility in $\mathbb{R}^n$ among the existing (fixed and random) facilities so that the sum of its weighted Chebyshev distances to the fixed facilities and the sum of its weighted expected Chebyshev distances to the realizations of the random facilities are both minimized.
Assume that we do not know the realizations of r random facilities at present time, and that these realizations become known at some point in the future. Assume also that the location of the new facility is to be determined so that the total sum is minimized. This decision must be made before the random facility realizations become available. Consequently, when the random facility realizations do become available, the new facility location that has already been determined, say by the point $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$, may or may not minimize the sum of its weighted expected distances to the realized random facilities. In order to make the location of the new facility minimizing the total sum of all weighted distances described above, we are allowed to change its location, say to the point $\boldsymbol{x}^{(0)} + \boldsymbol{x}(\omega) \in \mathbb{R}^n$, depending on the realized outcome $\omega \in \Omega$, if necessary. Given this, we are interested in a two-stage stochastic model of the form

$$\min_{\boldsymbol{x}^{(0)}} \sum_{i=1}^{f} \xi_i \, \|\boldsymbol{x}^{(0)} - \boldsymbol{a}_i\|_\infty \; + \; \mathbb{E}[Q(\boldsymbol{x}^{(0)}, \omega)], \tag{38}$$

where $\mathbb{E}[Q(\boldsymbol{x}^{(0)}, \omega)] \triangleq \int_{\omega \in \Omega} Q(\boldsymbol{x}^{(0)}, \omega) \; P(d\omega)$, and $Q(\boldsymbol{x}^{(0)}, \omega)$ is the minimum value of the unconstrained minimization problem

$$\min_{\boldsymbol{x}(\omega)} \sum_{j=1}^{r} \zeta_j(\omega) \; \|\boldsymbol{x}^{(0)} + \boldsymbol{x}(\omega) - \boldsymbol{b}_j(\omega)\|_\infty, \tag{39}$$

where $\xi_i \geq 0$ is the weight associated with the distance between the new facility and the i^{th} existing facility for $i = 1, 2, \ldots, f$, and $\zeta_j(\omega) \geq 0$ is the weight associated with the expected distance between the new facility and the realization of the j^{th} random facility for $j = 1, 2, \ldots, r$.



Chebyshev distance: $d_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$.

Euclidean distance: $d_2(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.

Manhattan distance: $d_1(\boldsymbol{x}, \boldsymbol{y}) = |x_1 - y_1| + |x_2 - y_2|$.

Figure 3: A geometric representation of Chebyshev, Manhattan and Euclidean distances.

The two-stage stochastic facility location model (38, 39) with K scenarios is written as

$$\begin{array}{lll} \min & \displaystyle\sum_{i=1}^{f} \xi_i \; u_i \; + \sum_{k=1}^{K} \hat{\rho}^{(k)}(\boldsymbol{x}^{(0)}) & \\ \text{s.t.} & u_i \geq \left\| \boldsymbol{x}^{(0)} - \boldsymbol{a}_i \right\|_\infty, & i = 1, 2, \ldots, f, \end{array} \tag{40}$$

where $\hat{\rho}^{(k)}(\boldsymbol{x}^{(0)}), k = 1, 2, \ldots, K$, is the minimum value of the constrained minimization problem

$$\begin{array}{ll} \min & \sum_{j=1}^{r} \zeta_j^{(k)}\ v_j^{(k)} \\ \text{s.t.} & v_j^{(k)} \geq \left\| \boldsymbol{x}^{(0)} + \boldsymbol{x}^{(k)} - \boldsymbol{b}_j^{(k)} \right\|_{\infty}, \quad j = 1, 2, \ldots, r. \end{array} \tag{41}$$

Algorithm 1 is performed with an accuracy $\epsilon = 10^{-5}$ for a number of two stage SINP problems. By "Iter" we denote the required inner iteration numbers, and by "CPU(s)" we denote the CPU time (in s) required to obtain an ϵ-approximate optimal solution of the underlying problem.

We run Algorithm 1 to solve the SINP problem (40) and (41), where the dimensions of this problem take the values $n = 4; 12; 20$, the numbers of scenarios take the values $K = 5; 10; 15; 20$, the number of fixed facilities take the values $f = 3; 10; 20$, and the number of random facilities take the values $r = 2; 10; 20$. For each quadruple (n, f, r, K), we generate 36 instances each with $\boldsymbol{a}_i$ and $\boldsymbol{b}_j^{(k)}$ chosen at random from the standard normal distribution. Finally, we choose the distance weights ξ_i and $\zeta_j^{(k)}$ randomly from a uniform distribution on $[0, 1]$. The numerical results of Algorithm 1 are displayed in Table 5 and are graphically visualized in Figure 4.

Table 5: The numerical results of Algorithm 1 for the stochastic uniform facility location problem.

n	f	r	$f+r$	K	Iter.	CPU(s)	n	f	r	$f+r$	K	Iter.	CPU(s)
4	3	2	5	5	5	1.1719	12	10	10	20	15	53	102.0140
4	3	2	5	10	5	1.2125	12	10	10	20	20	55	152.500
4	3	2	5	15	6	1.3344	12	20	20	40	5	29	120.6250
4	3	2	5	20	17	2.4219	12	20	20	40	10	49	128.8281
4	10	10	20	5	13	7.1719	12	20	20	40	15	56	93.2031
4	10	10	20	10	23	7.6250	12	20	20	40	20	60	99.0091
4	10	10	20	15	22	17.7656	20	3	2	5	5	57	91.1871
4	10	10	20	20	23	17.8938	20	3	2	5	10	65	23.1406
4	20	20	40	5	23	26.0781	20	3	2	5	15	44	93.2031
4	20	20	40	10	24	26.7355	20	3	2	5	20	84	98.7813
4	20	20	40	15	23	45.2188	20	10	10	20	5	45	107.2656
4	20	20	40	20	33	52.6406	20	10	10	20	10	10	112.6520
12	3	2	5	5	27	7.0625	20	10	10	20	15	61	120.5010
12	3	2	5	10	31	6.8281	20	10	10	20	20	77	140.4008
12	3	2	5	15	35	13.1563	20	20	10	30	5	75	100.0101
12	3	2	5	20	41	96.4036	20	20	10	30	10	56	119.9630
12	10	10	20	5	28	58.3906	20	20	20	40	15	56	130.2250
12	10	10	20	10	48	78.3594	20	20	20	40	20	88	135.3012

Example 2 (Randomly-generated problems). We run Algorithm 1 on random instances of K scenarios of the SINP problem (1) with values $K = 5; 15; 25; 35$. We assume that $\boldsymbol{c} \in \mathbb{R}^n, \boldsymbol{d}^{(k)} \in \mathbb{R}^m, \boldsymbol{b} \in \mathbb{R}^s, \boldsymbol{q}^{(k)} \in \mathbb{R}^l$ and $\boldsymbol{q}^{(k)} \in \mathbb{R}^s$, where the dimension of the problem takes the values $n = 10; 20; \ldots; 120$, $m = 5; 10; \ldots; 60$, $s = 5; 10; \ldots; 60$, and $l = 6; 12; \ldots; 72$. The parameters of Algorithm 1 are given as $\epsilon = 10^{-5}$, $\mu^0, \varpi \in (0, 1)$, with $\mu^0 > \epsilon$. For each

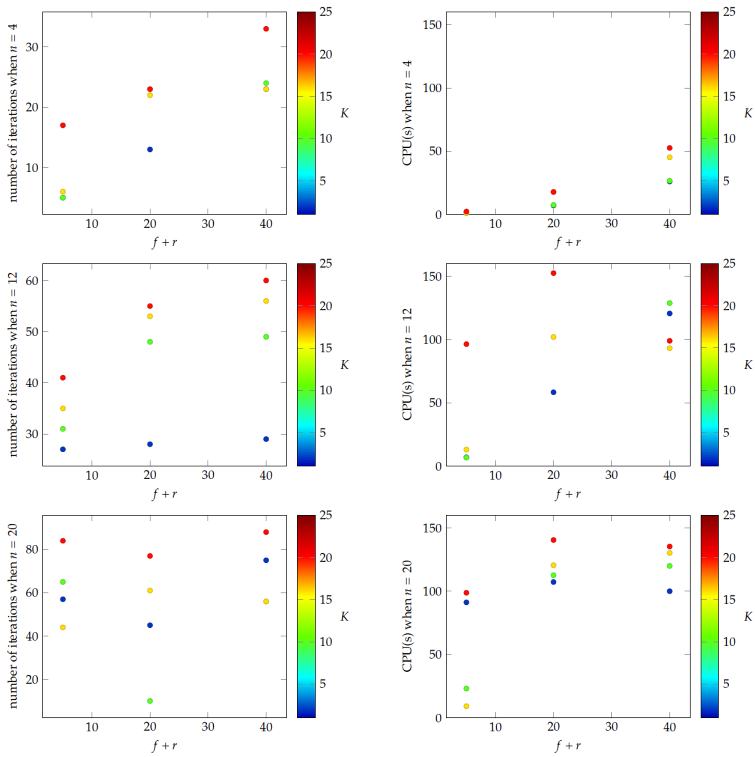


Figure 4: Dot plots of the numerical results obtained for the stochastic uniform facility location problem.

quintuple (m, n, s, l, K), we generate 48 instances each with $A, \boldsymbol{c}, \boldsymbol{b}, W^{(k)}, T^{(k)}, \boldsymbol{d}^{(k)}$ and $\boldsymbol{q}^{(k)}$ chosen at randomly generated values for $k = 1, 2, \ldots, K$.
Algorithm 1 is performed with an accuracy $\epsilon = 10^{-5}$ for a number of two stage SINP problems. By "Inn. Iter" we denote the required inner iteration numbers, by "Out. Iter" we denote the required outer iteration numbers, and by "CPU(s)" we denote the CPU time (in s) required to obtain an ϵ-approximate optimal solution of the underlying problem. The numerical results of Algorithm 1 are displayed in Table 6 and are graphically visualized in Figure 5. The initials of $\boldsymbol{x}^0$ and ${\boldsymbol{y}^{(k)}}^0$, for $k = 1, 2, \ldots, K$, were taken to be the unit vectors.

The implementation in Examples 1 and 2 shows Algorithm 1 is simple and efficient. From the numerical results displayed in Tables 5 and 6 and visualized in Figures 4 and 5, we notice that the number of iterations and the CPU(s) using Algorithm 1 are increasing when the number of scenarios K increases. We also find that the increase in the number of iterations is not only influenced by the number of scenarios, but also by dimensions of the underlying infinity norm cones. This totally agrees with the theoretical findings stated in Theorems 3 and 4.
Finally, we point out that we have also used the solver CVX to solve Examples 1

Table 6: Numerical results of Algorithm 1 for randomly-generated problems.

(m,n)	(s,l)	K	Out. Iter.	Inn. Iter.	CPU(s)
(5,10)	(5,6)	5	4	2	0.0312
(5,10)	(5,6)	15	6	4	0.3125
(5,10)	(5,6)	25	7	6	0.5000
(5,10)	(5,6)	35	9	7	0.6468
(10,20)	(10,12)	5	13	10	0.2806
(10,20)	(10,12)	15	16	12	0.7343
(10,20)	(10,12)	25	18	13	1.6406
(10,20)	(10,12)	35	24	15	4.6718
(15,30)	(15,18)	5	25	11	0.3906
(15,30)	(15,18)	15	31	17	2.9062
(15,30)	(15,18)	25	38	17	5.3750
(15,30)	(15,18)	35	39	21	11.2812
(20,40)	(20,24)	5	21	15	2.8280
(20,40)	(20,24)	15	30	19	5.8106
(20,40)	(20,24)	25	35	13	12.4610
(20,40)	(20,24)	35	38	21	14.7969
(25,50)	(25,30)	5	32	30	1.3562
(25,50)	(25,30)	15	40	36	6.7975
(25,50)	(25,30)	25	48	39	15.2818
(25,50)	(25,30)	35	50	41	32.1023
(30,60)	(30,36)	5	46	41	2.2904
(30,60)	(30,36)	15	49	43	7.7344
(30,60)	(30,36)	25	53	46	16.0625
(30,60)	(30,36)	35	56	47	33.0156
(35,70)	(35,42)	5	50	43	3.5625
(35,70)	(35,42)	15	58	48	10.1406
(35,70)	(35,42)	25	60	53	25.2969
(35,70)	(35,42)	35	67	56	36.8125
(40,80)	(40,48)	5	58	52	3.4219
(40,80)	(40,48)	15	57	58	13.6531
(40,80)	(40,48)	25	58	63	27.9250
(40,80)	(40,48)	35	63	67	40.0190
(45,90)	(45,52)	5	65	59	7.5938
(45,90)	(45,52)	15	75	61	15.9956
(45,90)	(45,52)	25	80	71	40.8125
(45,90)	(45,52)	35	83	80	62.0750
(50,100)	(50,58)	5	78	63	11.1875
(50,100)	(50,58)	15	89	69	29.7969
(50,100)	(50,58)	25	93	75	53.7813
(50,100)	(50,58)	35	110	84	114.4688
(55,110)	(55,64)	5	80	79	17.8125
(55,110)	(55,64)	15	93	84	33.9375
(55,110)	(55,64)	25	98	84	52.4063
(55,110)	(55,64)	35	121	89	117.6875
(60,120)	(60,72)	5	104	82	29.8594
(60,120)	(60,72)	15	116	95	66.7813
(60,120)	(60,72)	25	129	96	109.2031
(60,120)	(60,72)	35	136	87	126.6250

and 2. We have found that Algorithm 1 has no remarkable superiority to CVX in

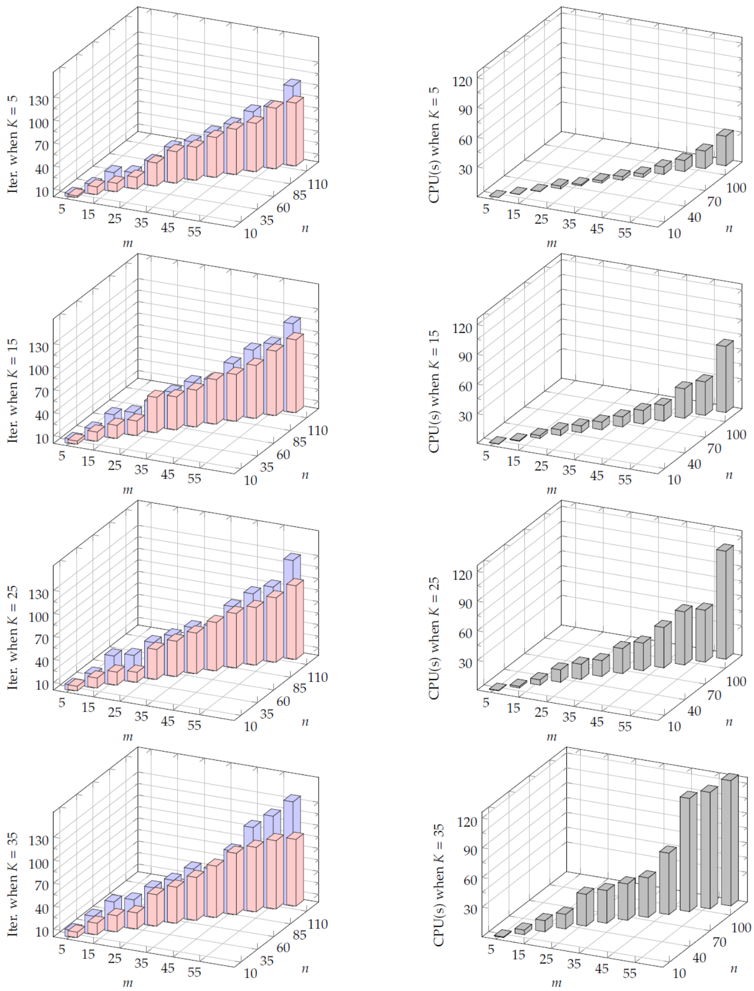


Figure 5: Three-dimensional plots of the numerical results obtained for randomly-generated problems. The number of iterations results are shown to the left (inner iter. is in light blue, and outer iter. is in light red). The CPU(s) results are shown to the right.

terms of number of iterations or running time. It is our belief that this does not lower the academic and practical value of this paper. Furthermore, when the number of realizations and the ranks of the underlying cones are large (typically, more than 50 each), we have found that the results show a small tendency toward Algorithm 1. Since the observed tendency is small and because we are keen to be cautious, we have decided not to include these results in this section.

7. Conclusions

In this paper, we have studied the two-stage stochastic infinity norm optimization problem. We have developed a novel Jordan algebra specialized and associated with the infinity norm cone. This allowed us not only to give explicit expressions for the derivatives of the logarithmic barrier functions associated with the cone, but also to specify the explicit barrier parameters for the self-concordant family of the composite barrier functions. These properties opened the door to develop primal decomposition-based interior-point algorithms for solving our optimization problem. We have shown that the worst case iteration complexity of the developed algorithms is the same as that for the short- and long-step interior-point algorithms applied to the two-stage stochastic linear programming. More specifically, we have found from our complexity results that the dominant complexity terms are linear in the number of realizations and linear in the ranks of the underlying infinity norm cones. We have seen that this exactly matches in terms of rank the best known results in the literature for stochastic second-order cone programming and stochastic semidefinite programming. Numerical experiments on stochastic uniform facility location problems as well as randomly-generated problems have demonstrated that the proposed algorithm not only has efficient worst case theoretical complexity, but also gives a good performance in practice. Future work may be devoted to study the two-stage stochastic first-order cone programming problem in which the underlying cone is $\mathcal{C}_1^n$.

Acknowledgment. The authors thank Yifan Dou from Ohio State University for reading the article and leaving comments.

Conflict of interests. The authors declare no potential conflict of interests.

Data Availability. Data sharing not applicable to this article as no datasets were generated or analysed during this study.

Appendix A: Complexity proofs

In this appendix, we present proofs for the complexity results stated in Section 6 that bound the number of iterations. The general scheme of our proofs follows the lines of the proofs from [16] and [5]. The proof of Theorem 3 for the short-step algorithm is given in Sub-appendix A.1, and the proof of Theorem 4 for the long-step algorithm is given in Sub-appendix A.2.

1. Complexity proof of the short-step algorithm

In this part, we show that the short-step algorithm takes only one inner iteration when $\varpi = 1 - \iota/\sqrt{n + Km}$ where $0 < \iota \leq 0.0755$. This value for ϖ is derived by

using the parameter functions $\alpha_1(\mu), \alpha_2(\mu), \ldots, \alpha_5(\mu)$ of the self-concordant family $\{\eta(\boldsymbol{x}, \mu) : \mu > 0\}$ established in Theorem 2.
Let v be a positive parameter. Let also $\{f_\mu\}_{\mu\in\mathbb{R}_{++}}$ be a family of strongly self-concordant functions with the parameters $\alpha_1(\mu), \alpha_2(\mu), \ldots, \alpha_5(\mu)$. The *v-metric function* associated with the family $\{f_\mu\}_{\mu\in\mathbb{R}_{++}}$ is denoted by $\varphi_v(f; t, \tau)$ and is defined as [21]

$$\varphi_v(f;t,\tau) \triangleq \max_{u,v\in[t,\tau]} \left| \ln \frac{\sqrt{\alpha_1(u)\alpha_2(u)}\alpha_3(v)}{\sqrt{\alpha_1(v)\alpha_2(v)}\alpha_3(u)} \right| + \frac{1}{v} \left| \int_t^\tau \alpha_4(w)\, dw \right| + \left| \int_t^\tau \alpha_5(w)\, dw \right|.$$

Proof of Theorem 3. In the short-step algorithm, we update μ^i using $\mu^{i+1} = \varpi\mu^i$, where $\varpi = 1 - \iota/\sqrt{n+Km}$, and ι is a small positive constant. Firstly, we show that if $0 < \iota \leq 0.0755$, then the proximity condition can be restored with only one inner iteration. By Theorem 2, the 2κ-metric function associated with the family $\{\rho(\boldsymbol{x}, \mu) : \mu > 0\}$ is

$$\begin{aligned}
\varphi_{2\kappa}\left(\eta;\mu^i,\mu^{i+1}\right) &= \frac{1}{2\kappa}\left|\int_{\mu^i}^{\mu^{i+1}} \frac{\sqrt{2((n-1)+K(m-1))}}{w} dw\right| + \left|\int_{\mu^i}^{\mu^{i+1}} \frac{1+2^{3/2}\sqrt{m-1}}{2w} dw\right| \\
&= \left(\frac{\sqrt{2((n-1)+K(m-1))}}{2\kappa} + \frac{1+2^{3/2}\sqrt{m-1}}{2}\right)\left|\int_{\mu^i}^{\mu^{i+1}} \frac{1}{w} dw\right| \\
&= \left(\frac{\sqrt{2((n-1)+K(m-1))}}{2\kappa} + \frac{1+2^{3/2}\sqrt{m-1}}{2}\right)\ln\left(\frac{\mu^i}{\mu^{i+1}}\right) \\
&= \left(\frac{\sqrt{2((n-1)+K(m-1))}}{2\kappa} + \frac{1+2^{3/2}\sqrt{m-1}}{2}\right)\ln\left(\frac{1}{\varpi}\right) \\
&= \left(\frac{\sqrt{2((n-1)+K(m-1))}}{2\kappa} + \frac{1+2^{3/2}\sqrt{m-1}}{2}\right)\ln\left(\frac{\sqrt{n+Km}}{\sqrt{n+Km}-\iota}\right).
\end{aligned}$$

Assume we have that

$$\delta\left(\boldsymbol{x}^i,\mu^i\right) \leq \kappa, \text{ or equivalently } \frac{1}{2} \leq 1 - \frac{1}{2\kappa}\delta\left(\boldsymbol{x}^i,\mu^i\right). \tag{A1}$$

By [21, Theorem 3.1.1], we deduce that

$$\varphi_{2\kappa}(\eta;\mu^i,\mu^{i+1}) \leq 1 - \frac{\delta\left(\boldsymbol{x}^i,\mu^i\right)}{2\kappa}, \text{ which implies that } \delta(\boldsymbol{x}^i,\mu^{i+1}) \leq 2\kappa. \tag{A2}$$

From the right-hand side inequality in (A2) and [21, Theorem 2.2.2(ii)], we have

$$\delta\left(\boldsymbol{x}^{i+1},\mu^{i+1}\right) \leq \frac{1}{2}\delta\left(\boldsymbol{x}^i,\mu^{i+1}\right). \tag{A3}$$

Note that one Newton step can give $\delta(\boldsymbol{x}^{i+1},\mu^{i+1}) \leq \kappa$ again if (A3) is combined with the right-hand side inequality in (A2). Note also that the right-hand side inequality in (A1) and

$$\varphi_{2\kappa}\left(\eta;\mu^i,\mu^{i+1}\right) = \left(\frac{\sqrt{2((n-1)+K(m-1))}}{2\kappa} + \frac{1+2^{3/2}\sqrt{m-1}}{2}\right)\ln\left(\frac{\sqrt{n+Km}}{\sqrt{n+Km}-\iota}\right) \leq \frac{1}{2} \tag{A4}$$

ensure that the left-hand side inequality in (A2) is satisfied. As a result, the inequality in (A4) can give us an upper bound of ι. It is clear from (A4) that

$$\iota \leq \sqrt{n+Km}\left(1-\exp\left(\frac{-1}{\frac{\sqrt{2((n-1)+K(m-1))}}{\kappa}+1+2^{3/2}\sqrt{m-1}}\right)\right) \geq \psi(n+Km), \quad \text{(A5)}$$

where $\psi(\cdot)$ is defined as

$$\psi(t) \triangleq \sqrt{t}\left(1-\exp\left(\frac{-1}{2\left(1+2\sqrt{t-1}+\frac{\sqrt{t}}{2\kappa}\right)}\right)\right), \text{ for } t \in [2,\infty).$$

After a lengthy but not complex computation, one can find that $\psi''(t) < 0$ on $[2,\infty)$, $\psi'(2) > 0$, and $\lim_{t\to\infty}\psi(t) = 0$. This means that $\psi(t)$ is an increasing function on $[2,\infty)$, and hence $\psi(t) \geq \psi(2) > 0.0755$, for any $t \in [2,\infty)$. In particular, $\psi(n+Km) \geq \psi(2)$. Thus, for $\iota \leq 0.0755$, the left-hand side inequality in (A5) is met ensuring that the number of inner iterations equals one. Because $\ln \varpi^{-1} = -\ln(1-\iota/\sqrt{n+Km}) \approx \iota/\sqrt{n+Km}$ and due to the fact that the number of outer iterations is given by $N \leq \ln(\mu^0/\varepsilon)/\ln \varpi^{-1}$, we conclude that with $\mathcal{O}(\sqrt{n+Km}\ln(\mu^0/\varepsilon))$ outer iterations we can reduce μ^0 to ε or less. The proof is complete. □

2. Complexity proof of the long-step algorithm

The complexity proof of the long-step algorithm makes use of Theorems 2.1.1(i) and 2.2.3 in [21] and Lemma A.3 in [16]. We also define

$$\phi(\boldsymbol{x},\mu) \triangleq \eta(\boldsymbol{x},\mu) - \eta(\boldsymbol{x}^\star(\mu),\mu), \quad \text{(A6)}$$

$$\tilde{\delta}(\boldsymbol{x},\mu) \triangleq \sqrt{\frac{1}{\mu}\tilde{\triangle}\boldsymbol{x}^\intercal \nabla^2_{\boldsymbol{xx}}\eta(\boldsymbol{x},\mu)\tilde{\triangle}\boldsymbol{x}}, \text{ where } \tilde{\triangle}\boldsymbol{x} \triangleq \boldsymbol{x} - \boldsymbol{x}^\star(\mu). \quad \text{(A7)}$$

The self-concordance family property of $\{\eta(\boldsymbol{x},\mu) : \mu > 0\}$, an upper bound on $\phi(\boldsymbol{x},\mu)$, and lower bound on the decrement of $\eta(\boldsymbol{x},\mu)$ per inner iteration are all employed in the complexity analysis of long-step algorithm. We give some technical lemmas that shall be used to derive the complexity result for the long-step algorithm.

Lemma 7. *Let $\tilde{\delta} \triangleq \tilde{\delta}(\boldsymbol{x},\mu) < 1$ and $\phi(\boldsymbol{x},\mu)$ be defined in (A6) and (A7). Then*

$$\left|\frac{\partial}{\partial\mu}\phi(\boldsymbol{x},\mu)\right| \leq -\sqrt{2((n-1)+K(m-1))}\ln\left(1-\tilde{\delta}\right).$$

Proof By using the KKT condition for the first-stage problem, and applying the fundamental theorem of calculus to $\frac{\partial}{\partial\mu}\phi(\boldsymbol{x},\mu)$, we have

$$\begin{aligned}
\left|\frac{\partial}{\partial\mu}\phi(\boldsymbol{x},\mu)\right| &= \left|\frac{\partial}{\partial\mu}\eta(\boldsymbol{x},\mu) - \frac{\partial}{\partial\mu}\eta(\boldsymbol{x}(\mu),\mu) - \Big(\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x}(\mu),\mu)\Big)^{\mathsf{T}}\frac{\partial}{\partial\mu}\boldsymbol{x}(\mu)\right| \\
&= \left|\frac{\partial}{\partial\mu}\eta(\boldsymbol{x},\mu) - \frac{\partial}{\partial\mu}\eta(\boldsymbol{x}(\mu),\mu)\right| \\
&= \left|\int_0^1 \left(\frac{\partial}{\partial\mu}\nabla_{\boldsymbol{x}}\eta\Big(\boldsymbol{x}(\mu)+\alpha\tilde{\triangle}\boldsymbol{x},\mu\Big)\right)^{\mathsf{T}}\tilde{\triangle}\boldsymbol{x}\, d_\alpha\right| \\
&\leq \int_0^1 \left|\left(\frac{\partial}{\partial\mu}\nabla_{\boldsymbol{x}}\eta\Big(\boldsymbol{x}(\mu)+\alpha\tilde{\triangle}\boldsymbol{x},\mu\Big)\right)^{\mathsf{T}}\tilde{\triangle}\boldsymbol{x}\right| d_\alpha \\
&\leq \int_0^1 \sqrt{\frac{2((n-1)+K(m-1))}{\mu}}\sqrt{\tilde{\triangle}\boldsymbol{x}^{\mathsf{T}}\nabla^2_{\boldsymbol{xx}}\eta\Big(\boldsymbol{x}(\mu)+\alpha\tilde{\triangle}\boldsymbol{x},\mu\Big)\tilde{\triangle}\boldsymbol{x}}\, d_\alpha \\
&\leq \int_0^1 \sqrt{\frac{2((n-1)+K(m-1))}{\mu}}\,\frac{\sqrt{\mu}\tilde{\delta}}{1+(\alpha-1)\tilde{\delta}}\, d_\alpha \\
&= -\sqrt{2((n-1)+K(m-1))}\ln\Big(1-\tilde{\delta}\Big),
\end{aligned}$$

The second equality is founded on the fact that $\boldsymbol{x}(\mu)$ is the optimal solution for the first stage, it fulfills the optimality condition $\eta(\boldsymbol{x}(\mu),\mu)=\mathbf{0}$. The second inequality used Lemma 6, and we used Theorem 2.1.1(i) in [21] to obtain the third inequality. The proof is complete. □

Lemma 8. *If* $\tilde{\delta} \triangleq \tilde{\delta}(\boldsymbol{x},\mu) \leq \zeta$, *for* $\zeta \in (0,1)$, *then*

$$\phi\left(\boldsymbol{x},\mu^{i+1}\right) = \eta\left(\boldsymbol{x},\mu^{i+1}\right) - \eta\left(\boldsymbol{x}\left(\mu^{i+1}\right),\mu^{i+1}\right) \leq \mathcal{O}(n+Km)\mu^{i+1}.$$

Proof By differentiating (A6), we have

$$\begin{aligned}
\frac{\partial^2}{\partial\mu^2}\phi(\boldsymbol{x},\mu) &= \frac{\partial^2}{\partial\mu^2}\eta(\boldsymbol{x},\mu) - \frac{\partial^2}{\partial\mu^2}\eta(\boldsymbol{x}(\mu),\mu) - (\nabla_{\boldsymbol{x}}\eta(\boldsymbol{x}(\mu),\mu))^{\mathsf{T}}\frac{\partial}{\partial\mu}\boldsymbol{x}(\mu) \\
&\leq -\frac{\partial^2}{\partial\mu^2}\eta(\boldsymbol{x}(\mu),\mu) \\
&= -\sum_{k=1}^{K}\frac{\partial^2}{\partial\mu^2}\rho^{(k)}(\boldsymbol{x}(\mu),\mu) \\
&= \sum_{k=1}^{K}\left(\mathsf{g}_{\boldsymbol{y}}^{(k)\mathsf{T}}\frac{\partial}{\partial\mu}\boldsymbol{y}^{(k)^\star}(\boldsymbol{x},\mu)\right) \\
&= \frac{1}{\mu}\sum_{k=1}^{K}\left(\mathsf{g}_{\boldsymbol{y}}^{(k)\mathsf{T}}\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}}\left(I-\mathsf{P}_{\boldsymbol{y}}^{(k)}\right)\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1/2}}\mathsf{g}_{\boldsymbol{y}}^{(k)}\right) \\
&\leq \frac{1}{\mu}\sum_{k=1}^{K}\left(\mathsf{g}_{\boldsymbol{y}}^{(k)\mathsf{T}}\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1}}\mathsf{g}_{\boldsymbol{y}}^{(k)}\right) \\
&= \frac{1}{\mu}\sum_{k=1}^{K}\left(\left(J_m\mathsf{g}_{\boldsymbol{y}}^{(k)}\right)\blacksquare\left(\mathsf{H}_{\boldsymbol{y}}^{(k)^{-1}}\mathsf{g}_{\boldsymbol{y}}^{(k)}\right)\right) \\
&= \frac{1}{\mu}\sum_{k=1}^{K}\operatorname{trace}(\boldsymbol{e}_m) = \frac{2K(m-1)}{\mu}.
\end{aligned}$$

From Lemma A.3 in [16], we have $\phi(\boldsymbol{x},\mu^i) \leq (\tilde{\delta}/(1-\tilde{\delta}) + \ln(1-\tilde{\delta}))\mu^i$. It follows that

$$\begin{aligned}
\phi(\boldsymbol{x},\mu^{i+1}) &= \phi\left(\boldsymbol{x},\mu^i\right) + \left(\mu^{i+1}-\mu^i\right) + \frac{\partial}{\partial\mu}\phi(\boldsymbol{x},\mu)\bigg|_{\mu=\mu^i} + \int_{\mu^{i+1}}^{\mu^i}\int_{\alpha}^{\mu^i}\frac{\partial^2}{\partial t^2}\phi(\boldsymbol{x},t)\,dt\,d_\alpha \\
&\leq \left(\frac{\tilde{\delta}}{1-\tilde{\delta}} + \ln\left(1-\tilde{\delta}\right)\right)\mu^i - \left(\mu^i-\mu^{i+1}\right)\sqrt{2((n-1)+K(m-1))}\ln\left(1-\tilde{\delta}\right) \\
&\quad + 2K(m-1)\int_{\mu^{i+1}}^{\mu^i}\int_{\alpha}^{\mu^i}\frac{1}{t}\,dt\,d_\alpha \\
&\leq \left(\frac{\tilde{\delta}}{1-\tilde{\delta}} + \ln\left(1-\tilde{\delta}\right)\right)\mu^i - \left(\mu^i-\mu^{i+1}\right)\sqrt{2((n-1)+K(m-1))}\ln\left(1-\tilde{\delta}\right) \\
&\quad + 2K(m-1)\left(\mu^i-\mu^{i+1}\right)\ln\varpi.
\end{aligned}$$

The desired result is obtained since $\tilde{\delta} \leq \zeta$, $\mu^{i+1} = \varpi\mu^i$, and ζ and ϖ are constants.□

Now, we are ready to prove Theorem 4. The focus of this complexity proof is on bounding the number of inner iterations.

Proof of Theorem 4. We use an arbitrary constant factor $\varpi \in (0,1)$ to decrease the barrier parameter μ in the long-step variant of the algorithm. Each outer iterate $(\boldsymbol{x}^k,\mu^i)$ satisfies $(\boldsymbol{x}^k,\mu^k) \leq \kappa$. Since ϖ is a constant, the number of outer iterations required to reduce μ^0 to ϵ is equal to $\ln(\mu^0/\epsilon)/\ln\varpi^{-1}$. We bound the number of inner Newton iterations. We assume that after updating μ^k to μ^{k+1}, $\delta(\boldsymbol{x}^k,\mu^k) > \kappa$, and that we start the inner loop by letting $\boldsymbol{x}^{k0} \triangleq \boldsymbol{x}^k$. From [21, Theorem 2.2.3], at any inner iterate $\boldsymbol{x}^{kj}$, if $\delta(\boldsymbol{x}^{kj},\mu^{k+1}) > 2\kappa$, then

$$\eta\left(\boldsymbol{x}^{kj},\mu^{k+1}\right) - \eta\left(\boldsymbol{x}^{k(j+1)},\mu^{k+1}\right) \geq \mu^{k+1}\left(\delta\left(\boldsymbol{x}^{kj},\mu^{k+1}\right) - \ln\left(1+\delta\left(\boldsymbol{x}^{kj},\mu^{k+1}\right)\right)\right) > 0.03\mu^{i+1}.$$

That is, the difference is decreased by at least $0.03\mu^{k+1}$ at each inner Newton iteration. On the other hand, Lemma 8 shows that $\eta(\boldsymbol{x}^k,\mu^{k+1}) - \eta(\boldsymbol{x}^\star(\mu^{k+1}),\mu^{k+1}) \leq \mathcal{O}(n+Km)\mu^{k+1}$. When the difference is equal to or less than 2κ, by [21, Theorem 2.2.3], one Newton iteration will produce $\boldsymbol{x}^{kM}$ with $\delta(\boldsymbol{x}^{kM},\mu^{k+1}) \leq \kappa$, then $\boldsymbol{x}^{k+1}$ is taken to be $\boldsymbol{x}^{kM}$ and the inner loop is thus terminated. Thus, each inner loop takes only $\mathcal{O}(n+Km)$ Newton iterations. This therefore bounds the total number of Newton iterations at the order $\mathcal{O}((n+Km)\ln(\mu^0/\epsilon))$. □

References

[1] B. Alzalg, *Decomposition-based interior point methods for stochastic quadratic second-order cone programming*, Appl. Math. Comput. **249** (2014), 1–18.

[2] ———, *Homogeneous self-dual algorithms for stochastic second-order cone programming*, J. Optim. Theory Appl. **163** (2014), 148–164.

[3] ———, *Volumetric barrier decomposition algorithms for stochastic quadratic second-order cone programming*, Appl. Math. Comput. **265** (2015), 494–508.

[4] ———, *The Jordan algebraic structure of the circular cone*, Oper Matrices **11** (2017), no. 1, 1–21.

[5] ———, *Primal interior-point decomposition algorithms for two-stage stochastic extended second-order cone programming*, Optimization **67** (2018), no. 12, 2291–2323.

[6] ———, *Logarithmic-barrier decomposition interior-point methods for stochastic linear optimization in a hilbert space*, Numer. Func. Anal. Optim. **41** (2020), no. 8, 901–928.

[7] ———, *Combinatorial and Algorithmic Mathematics: From Foundation to Optimization*, Kindle Direct Publishing, Seatle WA, 2022.

[8] B. Alzalg and K.A. Ariyawansa, *Logarithmic barrier decomposition-based interior point methods for stochastic symmetric programming*, J. Math. Anal. Appl. **409** (2014), no. 2, 973–995.

[9] B. Alzalg, K. Badarneh, and A. Ababneh, *An infeasible interior-point algorithm for stochastic second-order cone optimization*, J. Optim. Theory Appl. **181** (2019), 324–346.

[10] B. Alzalg, A. Gafour, and L. Alzaleq, *Volumetric barrier cutting plane algorithms for stochastic linear semi-infinite optimization*, IEEE Access **8** (2019), 4995–5008.

[11] B. Alzalg and M. Pirhaji, *Elliptic cone optimization and primal–dual path-following algorithms*, Optimization **66** (2017), no. 12, 2245–2274.

[12] R. Andreani, E.H. Fukuda, G. Haeser, D.O. Santos, and L.D. Secchin, *On the use of Jordan Algebras for improving global convergence of an Augmented Lagrangian method in nonlinear semidefinite programming*, Comput. Optim. Appl. **79** (2021), no. 3, 633–648.

[13] K. Ariyawansa and Y. Zhu, *A class of polynomial volumetric barrier decomposition algorithms for stochastic semidefinite programming*, Math. Comput. **80** (2011), no. 275, 1639–1661.

[14] P. Benner and T. Mitchell, *Faster and more accurate computation of the $\mathcal{H}_\infty$ norm via optimization.*, SIAM J Sci Comput. **40** (2018), A3609–A3635.

[15] T. Cai and W.X. Zhou, *A max-norm constrained minimization approach to 1-bit matrix completion.*, J. Mach. Learn. Res. **14** (2013), no. 1, 3619–3647.

[16] M. Chen and S. Mehrotra, *Self-concordance and decomposition-based interior point methods for the two-stage stochastic convex optimization problem*, SIAM J. Optim. **21** (2011), no. 4, 1667–1687.

[17] I. Gravagne and I.D. Walker, *Properties of minimum infinity-norm optimization applied to kinematically redundant robots*, Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst., IEEE, 1998, pp. 152–160.

[18] S. Kakihara, A. Ohara, and T. Tsuchiya, *Curvature integrals and iteration complexities in SDP and symmetric cone programs*, Comput. Optim. Appl. **57** (2014), no. 3, 623–665.

[19] T. Liu, M.T. Hoang, Y. Yang, and M. Pesavento, *A parallel optimization approach on the infinity norm minimization problem*, 27th Proc Eur Signal Process Conf., IEEE, 2019, pp. 1–5.

[20] S. Mehrotra and M.G. Özevin, *Decomposition-based interior point methods for two-stage stochastic semidefinite programming*, SIAM J. Optim. **18** (2007), no. 1, 206–222.

[21] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, SIAM, 1994.

[22] M.S. Sayadi Shahraki, H. Mansouri, and M. Zangiabadi, *Two wide neighborhood interior-point methods for symmetric cone optimization*, Comput Optim Appl. **68** (2017), 29–55.

[23] G. Yu, *Min-max optimization of several classical discrete optimization problems*, J. Optim. Theory Appl. **98** (1998), 221–242.

[24] G. Zhao, *A log-barrier method with benders decomposition for solving two-stage stochastic linear programs*, Math. Program. **90** (2001), 507–536.

[25] ———, *A lagrangian dual method with self-concordant barriers for multi-stage stochastic convex programming*, Math. Program. **102** (2005), 1–24.