Industrial and
Systems Engineering

Quadratic convergence to the optimal solution of second-order conic optimization without strict complementarity

Ali Mohammad-Nezhad and Tamás Terlaky

Department of Industrial and Systems Engineering, Lehigh University, USA

ISE Technical Report 17T-014

LEHIGH UNIVERSITY

Quadratic convergence to the optimal solution of second-order conic optimization without strict complementarity

Ali Mohammad-Nezhad and Tamás Terlaky

[a]Department of Industrial and Systems Engineering, Lehigh University, Harold S. Mohler Laboratory, 200 West Packer Avenue, Bethlehem, PA 18015-1582

ARTICLE HISTORY
Compiled July 4, 2018

ABSTRACT
Under primal and dual nondegeneracy conditions, we establish the quadratic convergence of Newton's method to the unique optimal solution of second-order conic optimization. Only very few approaches have been proposed to remedy the failure of strict complementarity, mostly based on nonsmooth analysis of the optimality conditions. Our local convergence result depends on the optimal partition of the problem, which can be identified from a bounded sequence of interior solutions. We provide a theoretical complexity bound for identifying the quadratic convergence region of Newton's method from the trajectory of central solutions. By way of experimentation, we illustrate quadratic convergence of Newton's method on some SOCO problems which fail strict complementarity condition.

KEYWORDS
Second-order conic optimization; optimal partition; quadratic convergence; maximally complementary optimal solution; nondegeneracy conditions

AMS CLASSIFICATION
90C51, 90C22, 90C25

1. Introduction

Second-order conic optimization (SOCO) problems minimize a linear objective function over the intersection of an affine space and Cartesian product of p second-order (Lorentz) cones of dimension n_i, i.e.,

$$\mathcal{L}_+^{\bar{n}} := \mathbb{L}_+^{n_1} \times \ldots \times \mathbb{L}_+^{n_p}, \qquad \bar{n} := \sum_{i=1}^{p} n_i,$$

where

$$\mathbb{L}_+^{n_i} := \left\{ x^i := (x_1^i, \ldots, x_{n_i}^i)^T \in \mathbb{R}^{n_i} \mid x_1^i \geq \|x_{2:n_i}^i\| \right\}, \quad i = 1, \ldots, p. \tag{1}$$

CONTACT Ali Mohammad-Nezhad. Email: alm413@lehigh.edu; Tel.: +1(610)758-2903.

The primal and dual SOCO problems in standard form are represented as

$$
\begin{aligned}
&(\mathrm{P_{SOCO}}) \qquad \min\left\{c^T x \mid Ax = b,\ x \in \mathcal{L}_+^{\bar{n}}\right\},\\
&(\mathrm{D_{SOCO}}) \qquad \max\left\{b^T y \mid A^T y + s = c,\ s \in \mathcal{L}_+^{\bar{n}}\right\},
\end{aligned}
$$

where $b \in \mathbb{R}^m$, $A := (A_1, \ldots, A_p)$, $x := (x^1; \ldots; x^p)$, $s := (s^1; \ldots; s^p)$, and $c := (c^1; \ldots; c^p)$, in which $A_i \in \mathbb{R}^{m\times n_i}$, $s^i \in \mathbb{R}^{n_i}$, and $c^i \in \mathbb{R}^{n_i}$ for $i = 1, \ldots, p$. Notice that x, s, and c are concatenation of the column vectors x^i, s^i, and c^i, respectively. A wide range of applications in engineering, control, robust optimization, and combinatorial optimization can be modeled as SOCO problems, see e.g., [2, 21] for the applications of SOCO.

From an algebraic point of view, SOCO can be embedded in a semidefinite optimization (SDO) problem using the following equivalence between a second-order cone and a positive semidefinite cone:

$$
L(x^i) := \begin{pmatrix} x_1^i & (x_{2:n_i}^i)^T \\ x_{2:n_i}^i & x_1^i I_{n_i-1} \end{pmatrix} \succeq 0 \quad \Leftrightarrow \quad x^i \in \mathbb{L}_+^{n_i} \quad \Leftrightarrow \quad (x^i)^T R_i x^i \geq 0, \quad x_1^i \geq 0,
$$

where R_i is an $n_i \times n_i$ diagonal matrix given by

$$
R_i := \mathrm{diag}\big(1,\ -1, \ldots, -1\big), \tag{2}
$$

and $L(x^i) \succeq 0$ means that the matrix $L(x^i)$ is positive semidefinite. All this hints that SOCO problems are polynomially solvable using an interior point method (IPM) for SDO. However, a direct implementation of IPMs for SOCO is proven to be more efficient in terms of computational complexity than IPMs applied to the equivalent SDO formulation [27]. The study on IPMs for SOCO problems was pioneered by the work of Nesterov and Nemirovskii [27], see also [28, 29] for the theory of linear conic optimization problems over self-scaled cones. Various search directions of IPMs have been studied for SOCO, see e.g., [1, 5, 25, 26, 37, 38].

Quadratic convergence of a primal-dual IPM for SOCO follows from Theorem 28 in [2] under nondegeneracy and strict complementarity conditions. Under the same conditions, the application of Newton's method to the AHO system of optimality conditions enjoys a quadratic convergence, when the initial point is sufficiently close to the optimal set, see Corollary 3.2 in [4]. If strict complementarity condition fails, then this local convergence result is no longer maintained. To the best of our knowledge, only very few remedies are available to resolve the issue of strict complementarity for SDO and SOCO. Those mostly are based on nonsmooth analysis of the optimality conditions, see e.g., [9, 17–20]. In such cases, the complementarity condition is replaced by $x^i - \Pi_{\mathbb{L}_+^{n_i}}(x^i - s^i) = 0$ for $i = 1, \ldots, p$, where $\Pi_{\mathbb{L}_+^{n_i}}(.)$ denotes the Euclidean projection on $\mathbb{L}_+^{n_i}$. Due to non-differentiability of $\Pi_{\mathbb{L}_+^{n_i}}(.)$ at some points, smoothing functions have been proposed to reformulate the complementarity condition. Under primal and dual nondegeneracy conditions, Chan and Sun [9] and Kong [19] established the quadratic convergence of a smoothing Newton's method to the unique optimal solution of SDO and symmetric conic optimization, respectively. Extending a smoothing function from nonlinear complementarity problems, Chi and Liu [10] proposed a non-interior continuation method for SOCO with superlinear convergence rate in the absence of strict complementarity. See also [11] for another derivation of smoothing function and smoothing

Newton's method for SOCO with quadratic convergence rate.

In this paper, our goal is to establish quadratic convergence to the unique optimal solution of SOCO using the so called *optimal partition* of the problem, as introduced by Bonnans and Ramírez [7]. Analogous to e.g., [9, 19], we only assume primal and dual nondegeneracy conditions. However, in contrast to [9, 19], we do not consider the metric projection operator to reformulate the complementarity condition. In our case, quadratic convergence is established only when the solution is sufficiently centered and sufficiently close to the optimal set, which allows for the identification of the optimal partition. The optimal partition can be identified using a sequence of interior solutions which has accumulation points in the relative interior of the optimal set (e.g., interior solutions generated by a feasible primal-dual IPM). See Terlaky and Wang [36] for the identification of the optimal partition for SOCO. Given the optimal partition identified from a sequence of central solutions, we reformulate the SOCO problem as a reduced nonlinear optimization (NLO) problem and then apply Newton's method to the first-order optimality conditions of the reduced problem. We show that if the primal and dual nondegeneracy conditions hold, then the Jacobian of the equations in KKT system of the reduced NLO problem is nonsingular at the unique globally optimal solution of the NLO problem. As a result, starting from a solution sufficiently close to the optimal set, Newton's method converges quadratically to the unique optimal solution. We support the theory by some numerical results.

In Section 2, we review the concepts of strict and maximal complementarity, optimal partition, nondegeneracy conditions, and a second-order sufficient condition for SOCO. Furthermore, we reproduce the bounds for the identification of the optimal partition using an error bound result for linear conic systems. Section 3 presents the main result of this paper: we establish the quadratic convergence of Newton's method to the unique optimal solution of SOCO. Furthermore, we provide a theoretical complexity bound for the identification of the quadratic convergence region of Newton's method, which is dependent on some condition numbers of the problem. In Section 4, we derive our complexity bound for the case when the strict complementarity condition holds. In Section 5, we conduct numerical experiments to demonstrate quadratic convergence of Newton's method on some instances of SOCO problems. Finally, our conclusions and ideas for future research are presented in Section 6.

Throughout this paper, int(.), ri(.), and bd(.) stand for the interior, the relative interior, and the boundary of a convex set, respectively, and the orthogonal complement of a linear subspace $\mathcal{A}$ is denoted by $\mathcal{A}^\perp$. Moreover, Ker(.) and $\mathcal{R}(.)$ denote the null space and the range space of a matrix. An arbitrary optimal solution is denoted by $(\tilde{x}; \tilde{y}; \tilde{s})$, and a maximally complementary optimal solution is indicated by $(x^*; y^*; s^*)$. An open ball of radius r is denoted by $B(., r)$. For any $I \subseteq \{1, \dots, p\}$ and an arbitrary matrix A, $|I|$ denotes the cardinality of I, and A_I represents the corresponding subset of the columns of A. Furthermore, letting N_i be an $m \times n_i$ matrix for $i \in I$, $(N_i)_{i \in I}$ is an $m \times \sum_{i \in I} n_i$ matrix formed by N_i for $i \in I$ put side by side. Finally, $\sigma_{\min}(.)$ stands for the minimum singular value of a matrix, and $\|.\|$ denotes the l_2 norm and the induced 2-norm (spectral norm) for the vectors and matrices, respectively.

Note that a critical direction associated with $(\mathrm{D}_{\mathrm{SOCO}})$ and the reduced NLO problem is always denoted by h. The reader should understand from the context to which problem the critical direction h belongs to. Moreover, we define and use the auxiliary vectors ξ and η occasionally to prove the linear independence of rows and columns of a matrix.

2. Preliminaries

The following assumptions are made throughout this paper:

Assumption 1. The coefficient matrix A has full row rank.

Assumption 2. The interior point condition holds, i.e., there exists a feasible solution $(x; y; s)$ such that for all $i = 1, \ldots, p$ we have $x^i, s^i \in \mathrm{int}(\mathbb{L}_+^{n_i})$, where

$$\mathrm{int}(\mathbb{L}_+^{n_i}) := \left\{x^i \in \mathbb{R}^{n_i} \mid x_1^i > \|x_{2:n_i}^i\|\right\}, \quad i = 1, \ldots, p.$$

As a result, at optimality the duality gap is zero, and the optimal value of ($\mathrm{P_{SOCO}}$) as well as that of ($\mathrm{D_{SOCO}}$) is attained, see Theorem 2.4.1 in [6]. Since strong duality holds, the optimal set for ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) can be represented as

$$\begin{aligned} Ax &= b, \quad x \in \mathcal{L}_+^{\bar{n}}, \\ A^T y + s &= c, \quad s \in \mathcal{L}_+^{\bar{n}}, \\ x \circ s &= 0, \end{aligned} \tag{3}$$

in which $x \circ s = 0$ denotes the complementarity condition, where $x \circ s := (x^1 \circ s^1; \ldots; x^p \circ s^p)$, and the bilinear map

$$x^i \circ s^i := \begin{pmatrix} (x^i)^T s^i \\ x_1^i s_{2:n_i}^i + s_1^i x_{2:n_i}^i \end{pmatrix}, \qquad i = 1, \ldots, p, \tag{4}$$

is called the Jordan product. The Jordan product $x^i \circ s^i$ can also be given by

$$x^i \circ s^i = L(x^i) s^i.$$

Any solution $(x; y; s)$ satisfying $x \circ s = 0$ is called complementary. Let $\mathcal{P}^*$ and $\mathcal{D}^*$ denote the primal and dual optimal sets, respectively. The interior point condition implies that the optimal set $\mathcal{P}^* \times \mathcal{D}^*$ is nonempty and compact, see e.g., Lemma 2 in [7].

Let $(x^*; y^*; s^*) \in \mathcal{P}^* \times \mathcal{D}^*$. Then $(x^*; y^*; s^*)$ is called strictly complementary if $x^* + s^* \in \mathrm{int}(\mathcal{L}_+^{\bar{n}})$. An optimal solution $(x^*; y^*; s^*)$ is called maximally complementary if $x^* \in \mathrm{ri}(\mathcal{P}^*)$ and $(y^*; s^*) \in \mathrm{ri}(\mathcal{D}^*)$. Under the interior point condition, a maximally complementary optimal solution always exists for SOCO, but a strictly complementary optimal solution may not exist.

2.1. Identification of the optimal partition

The notion of the optimal partition of LO can be extended to SOCO. Even though a SOCO problem can be embedded in SDO, the optimal partition in SOCO is more nuanced when it is defined and analyzed directly in the SOCO setting. In SOCO, the index set $\{1, \ldots, p\}$ of the second-order cones is partitioned into four subsets $\mathcal{B}, \mathcal{N}, \mathcal{R}$, and $\mathcal{T} := (\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$ as defined in [7]:

$$\begin{aligned}
\mathcal{B} &:= \{i \mid \tilde{x}_1^i > \|\tilde{x}_{2:n_i}^i\|, \quad \text{for some } \tilde{x} \in \mathcal{P}^*\},\\
\mathcal{N} &:= \{i \mid \tilde{s}_1^i > \|\tilde{s}_{2:n_i}^i\|, \quad \text{for some } \tilde{s} \in \mathcal{D}^*\},\\
\mathcal{R} &:= \{i \mid \tilde{x}_1^i = \|\tilde{x}_{2:n_i}^i\| > 0, \quad \tilde{s}_1^i = \|\tilde{s}_{2:n_i}^i\| > 0, \quad \text{for some } (\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^* \times \mathcal{D}^*\},\\
\mathcal{T}_1 &:= \{i \mid \tilde{x}^i = \tilde{s}^i = 0, \quad \text{for all } (\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^* \times \mathcal{D}^*\},\\
\mathcal{T}_2 &:= \{i \mid \tilde{s}^i = 0, \quad \text{for all } (\tilde{y};\tilde{s}) \in \mathcal{D}^*, \quad \tilde{x}_1^i = \|\tilde{x}_{2:n_i}^i\| > 0, \quad \text{for some } \tilde{x} \in \mathcal{P}^*\},\\
\mathcal{T}_3 &:= \{i \mid \tilde{x}^i = 0, \quad \text{for all } \tilde{x} \in \mathcal{P}^*, \quad \tilde{s}_1^i = \|\tilde{s}_{2:n_i}^i\| > 0, \quad \text{for some } (\tilde{y};\tilde{s}) \in \mathcal{D}^*\}.
\end{aligned}$$

The convexity of the optimal set implies that $\mathcal{B}, \mathcal{N}, \mathcal{R}$, and $\mathcal{T}$ are mutually disjoint and their union is the index set $\{1, \ldots, p\}$. Therefore, it follows from the complementarity condition that for all $(\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^* \times \mathcal{D}^*$, $\tilde{x}^i = 0$ for all $i \in \mathcal{N}$, and $\tilde{s}^i = 0$ for all $i \in \mathcal{B}$, see e.g., [36]. An optimal solution $(x^*; y^*; s^*)$ is maximally complementary if $(x^*)^i \in \operatorname{int}(\mathbb{L}_+^{n_i})$ for all $i \in \mathcal{B}$, $(s^*)^i \in \operatorname{int}(\mathbb{L}_+^{n_i})$ for all $i \in \mathcal{N}$, and $(x^*)^i, (s^*)^i \neq 0$ for all $i \in \mathcal{R}$, see Definition 5 in [7]. Further, a strictly complementary optimal solution $(x^*; y^*; s^*)$ exists iff $\mathcal{T} = \emptyset$.

Let $e_i := (1; \mathbf{0})$ denote the unit vector for the i^{th} second-order cone, and $e := (e_1; \ldots; e_p)$. Then, for $\mu > 0$, the central path is defined as the solution set of

$$\begin{aligned}
Ax = b, &\qquad x \in \operatorname{int}(\mathcal{L}_+^{\bar{n}}),\\
A^T y + s = c, &\qquad s \in \operatorname{int}(\mathcal{L}_+^{\bar{n}}),\\
x \circ s = \mu e. &
\end{aligned} \tag{5}$$

By Assumptions 1 and 2, for all $\mu > 0$ system (5) has a unique solution $(x(\mu); y(\mu); s(\mu))$, the so called central solution, where $x(\mu), s(\mu) \in \operatorname{int}(\mathcal{L}_+^{\bar{n}})$. For $\mu > 0$ the set of central solutions forms a smooth analytical curve which converges to a maximally complementary optimal solution, see [27] or Corollary 3.5 in [36].
To identify the optimal partition from a central solution, some proximity measures are needed to evaluate the magnitude of the solutions on the central path. Toward this end, Terlaky and Wang [36] defined two condition numbers, σ_1 and σ_2, as

$$\begin{aligned}
&\sigma_{\mathcal{B}} := \min_{i \in \mathcal{B}} \max_{\tilde{x} \in \mathcal{P}^*} \{\tilde{x}_1^i - \|\tilde{x}_{2:n_i}^i\|\}, \qquad \sigma_{\mathcal{N}} := \min_{i \in \mathcal{N}} \max_{(\tilde{y};\tilde{s}) \in \mathcal{D}^*} \{\tilde{s}_1^i - \|\tilde{s}_{2:n_i}^i\|\},\\
&\sigma_1 := \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\},\\
&\sigma_2 := \min_{i \in \mathcal{R}} \max_{(\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^* \times \mathcal{D}^*} \{\tilde{x}_1^i + \tilde{s}_1^i - \|\tilde{x}_{2:n_i}^i + \tilde{s}_{2:n_i}^i\|\}.
\end{aligned} \tag{6}$$

We also define

$$\sigma_3 := \max_{(\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^* \times \mathcal{D}^*} \{\|(\tilde{x};\tilde{y};\tilde{s})\|\}. \tag{7}$$

Note that the interior point condition implies that the condition numbers σ_1, σ_2, and σ_3 are finite positive values, see Lemma 3.3 in [36].

Theorem 2.3 presents the magnitude of the solutions on the central path for $\mathcal{B}, \mathcal{N}$, and $\mathcal{R}$, as given in Theorem 3.4 in [36]. For $\mathcal{T}$ we reproduce the bounds using the error bound result for a linear conic system defined as

$$\begin{cases} x \in \bar{b} + \mathcal{A}, \\ x \in \mathcal{L}_+^{\bar{n}}, \end{cases}$$

where $\mathcal{A}$ is a linear subspace of $\mathbb{R}^{\bar{n}}$ and $\bar{b} \in \mathbb{R}^{\bar{n}}$. The following Hölderian error bound is well-known from Theorem 7.4.2 in [22].

Lemma 2.1. *Let $x(\zeta)$ with $0 < \zeta \leq 1$ be a bounded set of solutions so that for all $0 < \zeta \leq 1$ we have*

$$\operatorname{dist}\left(x(\zeta),\ \bar{\mathcal{A}}\right) \leq \zeta, \quad x_1^i(\zeta) - \|x_{2:n_i}^i(\zeta)\| \geq -\zeta, \quad i = 1, \ldots, p, \tag{8}$$

where $\bar{\mathcal{A}}$ is the minimal linear subspace which contains $\bar{b} + \mathcal{A}$, and

$$\operatorname{dist}\left(x(\zeta),\ \bar{\mathcal{A}}\right) := \min\left\{\|\psi - x(\zeta)\| \mid \psi \in \bar{\mathcal{A}}\right\}$$

denotes the distance of $x(\zeta)$ from the subspace $\bar{\mathcal{A}}$. Then there exist a positive condition number κ' independent of ζ and a positive exponent γ' so that

$$\operatorname{dist}\left(x(\zeta),\ (\bar{b} + \mathcal{A}) \cap \mathcal{L}_+^{\bar{n}}\right) \leq \kappa' \zeta^{\gamma'},$$

where $\gamma' := 2^{-d(\bar{\mathcal{A}}, \mathcal{L}_+^{\bar{n}})}$ and $d(\bar{\mathcal{A}}, \mathcal{L}_+^{\bar{n}})$ denotes the degree of singularity[1] *of the subspace $\bar{\mathcal{A}}$.*

We use Lemma 2.1 to specify an upper bound on the distance of a central solution to the optimal set. To do so, let $(\hat{x}; \hat{y}; \hat{s})$ be a primal-dual optimal solution of $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$. Note that the primal and dual optimal sets can be equivalently written as the following linear conic systems

$$\begin{cases} x \in \hat{x} + \operatorname{Ker}(A), \\ \hat{s}^T x = 0, \\ x \in \mathcal{L}_+^{\bar{n}}, \end{cases} \qquad \begin{cases} s \in \hat{s} + \mathcal{R}(A^T), \\ \hat{x}^T s = 0, \\ s \in \mathcal{L}_+^{\bar{n}}, \end{cases} \tag{9}$$

see also Section 4 in [35]. In this case, the linear subspace $\bar{\mathcal{A}}(\mathcal{P}^*)$, which contains $\mathcal{P}^*$, is defined as

$$\bar{\mathcal{A}}(\mathcal{P}^*) := \left(\operatorname{Ker}(A) \cap (\mathbb{R}\hat{s})^\perp\right) + \mathbb{R}\hat{x},$$

where $\mathbb{R}\hat{x}$ and $\mathbb{R}\hat{s}$ denote the line generated by $\hat{x}$ and $\hat{s}$, respectively. Analogously, the linear subspace $\bar{\mathcal{A}}(\mathcal{D}^*)$, which is the orthogonal complement of $\bar{\mathcal{A}}(\mathcal{P}^*)$, is defined as

$$\bar{\mathcal{A}}(\mathcal{D}^*) := \left(\mathcal{R}(A^T) \cap (\mathbb{R}\hat{x})^\perp\right) + \mathbb{R}\hat{s}.$$

From the orthogonality of $(x(\mu) - \hat{x})$ and $(s(\mu) - \hat{s})$ we have

$$\hat{x}^T s(\mu) + \hat{s}^T x(\mu) = p\mu,$$

[1] The degree of singularity [35] is defined as the number of facial reduction steps to get the minimal face of the second-order cone $\mathcal{L}_+^{\bar{n}}$ which contains $\bar{b} + \mathcal{A}$.

which implies $0 \le \hat{s}^T x(\mu) \le p\mu$ and $0 \le \hat{x}^T s(\mu) \le p\mu$. Then the application of the Hoffman error bound [16] gives

$$\begin{aligned} \operatorname{dist}\big(x(\mu),\ \bar{\mathcal{A}}(\mathcal{P}^*)\big) &\le \operatorname{dist}\big(x(\mu),\ \{x \in \hat{x} + \operatorname{Ker}(A) \mid \hat{s}^T x = 0\}\big) \\ &= \operatorname{dist}\big(x(\mu),\ \{x \mid Ax = b,\ \hat{s}^T x = 0\}\big) \\ &\le \theta_1\big(\|Ax(\mu) - b\| + \hat{s}^T x(\mu)\big) = \theta_1 \hat{s}^T x(\mu) \le \theta_1 p\mu, \end{aligned}$$

where $\theta_1 > 0$ denotes the Hoffman condition number. Analogously, we derive

$$\operatorname{dist}\big(s(\mu),\ \bar{\mathcal{A}}(\mathcal{D}^*)\big) \le \operatorname{dist}\big(s(\mu),\ \{s \in \hat{s} + \mathcal{R}(A^T) \mid \hat{x}^T s = 0\}\big) \le \theta_2 p\mu,$$

where $\theta_2 > 0$ is the Hoffman condition number. Note that the condition numbers θ_1 and θ_2 depend on A and $\hat{s}$, and A and $\hat{x}$, respectively.

Lemma 2.2. *Let $(x(\mu); y(\mu); s(\mu))$ be a central solution with*

$$\mu \le \hat{\mu} := \frac{1}{p} \min\{\theta_1^{-1}, \theta_2^{-1}\}. \tag{10}$$

Then there exist $(x_\mu; y_\mu; s_\mu) \in \mathcal{P}^* \times \mathcal{D}^*$, a positive condition number κ independent of μ, and $\gamma > 0$ so that

$$\begin{aligned} \|x(\mu) - x_\mu\| &\le \kappa(p\mu)^\gamma, \\ \|y(\mu) - y_\mu\| &\le \kappa(p\mu)^\gamma, \\ \|s(\mu) - s_\mu\| &\le \kappa(p\mu)^\gamma. \end{aligned} \tag{11}$$

Proof. Conditions (8) hold if $\mu \le \hat{\mu}$. Moreover, Assumptions 1 and 2 imply that the set

$$\{(x(\mu); y(\mu); s(\mu)) \mid 0 < \mu \le \hat{\mu}\}$$

is contained in a compact set, see Lemma 3.2 in [12]. Therefore, the result of Lemma 2.1 is applicable to the linear conic systems in (9). Furthermore, the compactness of $\mathcal{P}^*$ and $\mathcal{D}^*$ implies the existence of $(x_\mu; y_\mu; s_\mu) \in \mathcal{P}^* \times \mathcal{D}^*$ so that

$$\|x(\mu) - x_\mu\| \le \kappa_1(p\mu)^{\gamma_1}, \quad \|s(\mu) - s_\mu\| \le \kappa_2(p\mu)^{\gamma_2},$$

where $\gamma_1 := 2^{-d(\bar{\mathcal{A}}(\mathcal{P}^*), \mathcal{L}_+^{\bar{n}})}$ and $\gamma_2 := 2^{-d(\bar{\mathcal{A}}(\mathcal{D}^*), \mathcal{L}_+^{\bar{n}})}$. Since the rows of A are assumed to be linearly independent, system $A^T(y(\mu) - y_\mu) = s(\mu) - s_\mu$ has a unique solution. Therefore, using Lemma 2.1 again, we get

$$\|y(\mu) - y_\mu\| \le \big\|(A^T)^\dagger\big\| \|s(\mu) - s_\mu\| \le \kappa_3(p\mu)^{\gamma_2},$$

where $(A^T)^\dagger := (AA^T)^{-1}A$ stands for the pseudo-inverse of A^T [33]. Then, taking $\kappa := \max\{\kappa_1, \kappa_2, \kappa_3\}$ and $\gamma := \min\{\gamma_1, \gamma_2\}$, we get the results as desired. □

Now, the bounds on the magnitude of the central solutions are summarized in Theorem 2.3.

Theorem 2.3. *Let $(x(\mu); y(\mu); s(\mu))$ be a central solution such that $\mu \le \hat{\mu}$, where $\hat{\mu}$ is defined in (10). Then we have*

$$
\begin{aligned}
& x_1^i(\mu) \ge x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\| > \frac{\sigma_1}{2p}, \quad \text{and} \quad s_1^i(\mu) \le \frac{p\mu}{\sigma_1}, && \forall i \in \mathcal{B}, \\
& s_1^i(\mu) \ge s_1^i(\mu) - \|s_{2:n_i}^i(\mu)\| > \frac{\sigma_1}{2p}, \quad \text{and} \quad x_1^i(\mu) \le \frac{p\mu}{\sigma_1}, && \forall i \in \mathcal{N}, \\
& x_1^i(\mu) > \frac{\sigma_2}{4p}, \quad \text{and} \quad s_1^i(\mu) > \frac{\sigma_2}{4p}, && \forall i \in \mathcal{R}, \\
& \big(x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\|\big) + \big(s_1^i(\mu) - \|s_{2:n_i}^i(\mu)\|\big) \le \frac{2p\mu}{\sigma_2}, && \forall i \in \mathcal{R}, \\
& x_1^i(\mu) + s_1^i(\mu) - \|x_{2:n_i}^i(\mu) + s_{2:n_i}^i(\mu)\| > \frac{\sigma_1}{2p}, && \forall i \in \mathcal{B} \cup \mathcal{N}, \\
& x_1^i(\mu) + s_1^i(\mu) - \|x_{2:n_i}^i(\mu) + s_{2:n_i}^i(\mu)\| > \frac{\sigma_2}{2p}, && \forall i \in \mathcal{R}, \\
& \frac{\mu}{2\kappa(p\mu)^\gamma} \le x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\| \le x_1^i(\mu) \le \kappa(p\mu)^\gamma, && \forall i \in \mathcal{T}_1, \\
& \frac{\mu}{2\kappa(p\mu)^\gamma} \le s_1^i(\mu) - \|s_{2:n_i}^i(\mu)\| \le s_1^i(\mu) \le \kappa(p\mu)^\gamma, && \forall i \in \mathcal{T}_1, \\
& \frac{\mu}{2\kappa(p\mu)^\gamma} \le x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\| \le \sqrt{2}\kappa(p\mu)^\gamma, && \forall i \in \mathcal{T}_2, \\
& \frac{\mu}{2\sqrt{2}\kappa(p\mu)^\gamma} \le s_1^i(\mu) \le \kappa(p\mu)^\gamma, && \forall i \in \mathcal{T}_2, \\
& \frac{\mu}{2\kappa(p\mu)^\gamma} \le s_1^i(\mu) - \|s_{2:n_i}^i(\mu)\| \le \sqrt{2}\kappa(p\mu)^\gamma, && \forall i \in \mathcal{T}_3, \\
& \frac{\mu}{2\sqrt{2}\kappa(p\mu)^\gamma} \le x_1^i(\mu) \le \kappa(p\mu)^\gamma, && \forall i \in \mathcal{T}_3, \\
& \frac{\mu}{2\kappa(p\mu)^\gamma} \le x_1^i(\mu) + s_1^i(\mu) - \|x_{2:n_i}^i(\mu) + s_{2:n_i}^i(\mu)\| \le 4\kappa(p\mu)^\gamma, && \forall i \in \mathcal{T},
\end{aligned}
$$

where κ and γ are defined in Lemma 2.2.

Proof. The sketch of the proof for the subsets $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{R}$ is similar to Theorem 3.4 in [36]. Let $i \in \mathcal{T}$ denote a block $(x^i(\mu); y^i(\mu); s^i(\mu))$ of the central solution with $\mu \le \hat{\mu}$. From (4) and the central path equation $x^i(\mu) \circ s^i(\mu) = \mu e_i$ we get

$$
s^i(\mu) = \frac{\mu(x_1^i(\mu); -x_{2:n_i}^i(\mu))}{(x_1^i(\mu))^2 - \|x_{2:n_i}^i(\mu)\|^2}.
$$

Therefore, from $x_1^i(\mu) > \|x_{2:n_i}^i(\mu)\|$ we have

$$
\begin{aligned}
s_1^i(\mu) &= \frac{\mu x_1^i(\mu)}{(x_1^i(\mu))^2 - \|x_{2:n_i}^i(\mu)\|^2} = \frac{\mu}{x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\|} \frac{x_1^i(\mu)}{x_1^i(\mu) + \|x_{2:n_i}^i(\mu)\|} \\
&\ge \frac{\mu}{2(x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\|)}.
\end{aligned} \tag{12}
$$

Analogously, we can derive

$$x_1^i(\mu) \geq \frac{\mu}{2(s_1^i(\mu) - \|s_{2:n_i}^i(\mu)\|)}. \tag{13}$$

- $i \in \mathcal{T}_1$: In this case, we have $\tilde{x}^i = \tilde{s}^i = 0$ for all $(\tilde{x};\tilde{y};\tilde{s}) \in \mathcal{P}^* \times \mathcal{D}^*$, and thus the bounds in (11) reduce to

  $$\|x^i(\mu)\| \leq \kappa(p\mu)^\gamma, \quad \|s^i(\mu)\| \leq \kappa(p\mu)^\gamma. \tag{14}$$

  Consequently, it can be deducted from (12) and (14) that

  $$\begin{aligned} x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\| &\geq \frac{\mu}{2s_1^i(\mu)} \geq \frac{\mu}{2\|s^i(\mu)\|} \geq \frac{\mu}{2\kappa(p\mu)^\gamma}, \\ x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\| &\leq x_1^i(\mu) \leq \|x^i(\mu)\| \leq \kappa(p\mu)^\gamma. \end{aligned}$$

  In a similar manner, using (13) we can show that

  $$\frac{\mu}{2\kappa(p\mu)^\gamma} \leq s_1^i(\mu) - \|s_{2:n_i}^i(\mu)\| \leq s_1^i(\mu) \leq \kappa(p\mu)^\gamma,$$

  which completes the first part of the proof.

- $i \in \mathcal{T}_2$: In this case, the bound in (11) reduces to $\|s^i(\mu)\| \leq \kappa(p\mu)^\gamma$. Thus, we can conclude from (12) that

  $$x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\| \geq \frac{\mu}{2s_1^i(\mu)} \geq \frac{\mu}{2\|s^i(\mu)\|} \geq \frac{\mu}{2\kappa(p\mu)^\gamma}.$$

  Furthermore, it follows from $(x_\mu)_1^i = \|(x_\mu)_{2:n_i}^i\|$ that

  $$\begin{aligned} x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\| &= (x_1^i(\mu) - (x_\mu)_1^i) + \big(\|(x_\mu)_{2:n_i}^i\| - \|x_{2:n_i}^i(\mu)\|\big) \\ &\leq |x_1^i(\mu) - (x_\mu)_1^i| + \|(x_\mu)_{2:n_i}^i - x_{2:n_i}^i(\mu)\| \\ &\leq \sqrt{2}\|x^i(\mu) - x_\mu^i\| \leq \sqrt{2}\kappa(p\mu)^\gamma. \end{aligned} \tag{15}$$

  Therefore, using (12) and (15) we get

  $$\kappa(p\mu)^\gamma \geq s_1^i(\mu) \geq \frac{\mu}{2\big(x_1^i(\mu) - \|x_{2:n_i}^i(\mu)\|\big)} \geq \frac{\mu}{2\sqrt{2}\kappa(p\mu)^\gamma},$$

  which completes the proof for the second part.

- $i \in \mathcal{T}_3$: It immediately follows after reversing the roles of $x^i(\mu)$ and $s^i(\mu)$.

The rest of the theorem follows by applying the results from the previous parts as in Theorem 3.8 in [36]. □

From the bounds of Theorem 2.3 one can observe that a complete separation of the variables to the partition $\mathcal{B}$, $\mathcal{N}$, $\mathcal{R}$ and $\mathcal{T}$ can be made if

$$\frac{p\mu}{\sigma_1} < \min\left\{\frac{\sigma_1}{2p}, \frac{\sigma_2}{4p}\right\}, \quad \max\left\{\frac{p\mu}{\sigma_1}, \frac{2p\mu}{\sigma_2}\right\} < \frac{\sigma_1}{2p}, \quad 4\kappa(p\mu)^\gamma < \min\left\{\frac{\sigma_1}{2p}, \frac{\sigma_2}{2p}\right\},$$

which can be simplified to

$$\mu < \tilde{\mu} := \min\left\{\frac{\sigma_1^2}{2p^2}, \frac{\sigma_1\sigma_2}{4p^2}, \ \frac{1}{p}\left(\frac{1}{4\kappa}\min\left\{\frac{\sigma_1}{2p}, \frac{\sigma_2}{2p}\right\}\right)^{\frac{1}{\gamma}}, \ \hat{\mu}\right\}. \tag{16}$$

However, we do not have enough information for a further separation of $\mathcal{T}$ into $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$. To that end, we need positive lower bounds for $x_1^i(\mu)$ and $s_1^i(\mu)$ in $\mathcal{T}_2$ and $\mathcal{T}_3$, respectively, which cannot be directly obtained from Theorem 2.3. Nevertheless, we assume in the convergence analysis of Section 3 that $\mu < \tilde{\mu}$ is small enough for a complete identification of $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$.

2.2. *Primal and dual nondegeneracy conditions in SOCO*

The primal and dual nondegeneracy conditions for a feasible solution of SOCO are presented in [2, 5]. Let $\tan(x^i, \mathbb{L}_+^{n_i})$ be the tangent space to $\mathbb{L}_+^{n_i}$ at x^i, see [2] for the definition of tangent space. Then a primal feasible solution x is called nondegenerate if

$$\tan(x^1, \mathbb{L}_+^{n_1}) \times \ldots \times \tan(x^p, \mathbb{L}_+^{n_p}) + \operatorname{Ker}(A) = \mathbb{R}^{\bar{n}}.$$

A dual feasible solution $(y; s)$ is called nondegenerate if

$$\tan(s^1, \mathbb{L}_+^{n_1}) \times \ldots \times \tan(s^p, \mathbb{L}_+^{n_p}) + \mathcal{R}(A^T) = \mathbb{R}^{\bar{n}}.$$

The existence of a primal (dual) nondegenerate optimal solution implies the uniqueness of the dual (primal) optimal solution. If the strict complementarity condition holds, then the converse is true as well, see Theorem 22 in [2].

Since we assume both the primal and dual nondegeneracy almost everywhere in this paper, using the optimal partition we characterize the primal and dual nondegeneracy conditions for the unique optimal solutions of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$). To that end, the matrix of the eigenvectors of $L(x^i)$ is denoted by $Q_i := (\sqrt{2}q_1^i, \sqrt{2}q_2^i, \ \hat{Q}_i)$, where

$$q_1^i := \frac{1}{2}\left(1; \frac{-x_{2:n_i}^i}{\|x_{2:n_i}^i\|}\right), \qquad q_2^i := \frac{1}{2}\left(1; \frac{x_{2:n_i}^i}{\|x_{2:n_i}^i\|}\right),$$

and $\hat{Q}_i \in \mathbb{R}^{(n_i \times n_i - 2)}$ is a matrix with orthogonal columns. The eigenvectors of $L((x^*)^i)$ are indicated by superscript $*$.

Theorem 2.4 (Theorems 20 and 21 in [2]). *Let $(x^*; y^*; s^*)$ be the unique optimal solution of* ($\mathrm{P_{SOCO}}$) *and* ($\mathrm{D_{SOCO}}$). *Then x^* is primal nondegenerate iff the matrix*

$$\big((A_i\bar{Q}_i^*)_{i \in \mathcal{R}\cup\mathcal{T}_2}, \ A_{\mathcal{B}}\big) \tag{17}$$

has full row rank, where $\bar{Q}_i^* := (\sqrt{2}(q^*)_2^i,\ \hat{Q}_i^*)$. Furthermore, $(y^*;s^*)$ is dual nondegenerate iff the matrix

$$\big((A_i R_i (s^*)^i)_{i\in\mathcal{R}\cup\mathcal{T}_3},\ A_{\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2}\big) \tag{18}$$

has full column rank, where R_i is defined in (2).

For the sake of convenience, given the unique optimal solution $(x^*;y^*;s^*)$, the primal nondegeneracy of x^* and the dual nondegeneracy of $(y^*;s^*)$ are simply called the primal and dual nondegeneracy conditions, respectively.

2.3. Second-order sufficient condition for SOCO

We highlight the connection between the second-order sufficient condition of Bonnans and Ramírez [7] and the primal nondegeneracy condition. In Section 3, we use the second-order sufficient condition to show the nonsingularity of the Jacobian of the equations in the KKT conditions for a reduced NLO problem.

Let $h \in \mathbb{R}^m$, and assume that $\mathcal{R}$ is nonempty so that there exists $(\check{y};\check{s}) \in \mathcal{D}^*$ with $\check{s}^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})\setminus\{0\}$ for some $i \in \mathcal{R}$. Then the specialization of the second-order sufficient condition for $(\mathrm{D_{SOCO}})$, where the objective is to minimize $-b^Ty$, is given by

$$\sup_{\tilde{x}\in\mathcal{P}^*} h^T H(\check{y},\tilde{x})h > 0, \quad \forall h \in \mathcal{C}(\check{y})\setminus\{0\}, \tag{19}$$

where

$$H(\check{y},\tilde{x}) := \sum_{i=1}^p H^i(\check{y},\tilde{x}),$$

$$H^i(\check{y},\tilde{x}) := \begin{cases} -\frac{\tilde{x}_1^i}{\check{s}_1^i} A_i R_i A_i^T, & \check{s}^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})\setminus\{0\}, \\ \mathbf{0}_{m\times m}, & \text{otherwise}, \end{cases} \qquad i = 1,\dots,p,$$

and $\mathcal{C}(\check{y})$ is the cone of critical directions for $(\mathrm{D_{SOCO}})$ which is defined as follows

$$\begin{cases} h \in \mathbb{R}^m, & \\ -A_i^T h \in \mathbb{L}_+^{n_i}, & \tilde{x}^i, \check{s}^i = 0, \\ -A_i^T h \in \{d \mid d_{2:n_i}^T \check{s}_{2:n_i}^i - d_1\check{s}_1^i \le 0\}, & \tilde{x}^i = 0,\ \check{s}^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})\setminus\{0\}, \\ A_i^T h = 0, & \tilde{x}^i \in \mathrm{int}(\mathbb{L}_+^{n_i}), \\ (\tilde{x}^i)^T A_i^T h = 0, & \tilde{x}^i, \check{s}^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})\setminus\{0\}, \\ -A_i^T h \in \mathbb{R}_+(\tilde{x}_1^i; -\tilde{x}_{2:n_i}^i), & \tilde{x}^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})\setminus\{0\},\ \check{s}^i = 0. \end{cases} \tag{20}$$

Then, by $\tilde{x}^i = 0$ for $i \in \mathcal{T}_3$, we have

$$h^T H(\check{y},\tilde{x})h = \sum_{i\in\mathcal{R}} -\frac{\tilde{x}_1^i}{\check{s}_1^i} h^T A_i R_i A_i^T h.$$

The connection between the primal nondegeneracy condition and the second-order sufficient condition (19) is stated in the following lemma.

Lemma 2.5 (Proposition 3.2 in [20]). *Let $\mathcal{R} \neq \emptyset$ and $(x^*; y^*; s^*)$ be the unique optimal solution of* ($\mathrm{P_{SOCO}}$) *and* ($\mathrm{D_{SOCO}}$). *Then, under the primal nondegeneracy condition, the second-order sufficient condition* (19) *holds at $(y^*; s^*)$.*

It follows from the definition, see e.g., (3.109) in [8], that $h^T H(\tilde{y}, \tilde{x}) h \geq 0$ for all $(\tilde{x}; \tilde{y}; \tilde{s}) \in \mathcal{P}^* \times \mathcal{D}^*$ and $h \in \mathcal{C}(\tilde{y})$. We can also observe from the proof of Lemma 2.5 that under the primal nondegeneracy condition we have $h^T A_i R_i A_i^T h < 0$ for some $i \in \mathcal{R}$. More precisely, notice from $((x^*)^i)^T A_i^T h = 0$ that $A_i^T h \notin \mathrm{int}(\mathbb{L}_+^{n_i})$ and $-A_i^T h \notin \mathrm{int}(\mathbb{L}_+^{n_i})$ for every $h \in \mathcal{C}(y^*) \setminus \{0\}$. Then from the characterization of the primal nondegeneracy condition in Theorem 2.4 we have that

$$
\begin{aligned}
A_i^T \eta &= 0, \quad i \in \mathcal{B}, \\
((x^*)^i)^T A_i^T \eta &= 0, \quad i \in \mathcal{R} \cup \mathcal{T}_2, \\
(\hat{Q}_i^*)^T A_i^T \eta &= 0, \quad i \in \mathcal{R} \cup \mathcal{T}_2
\end{aligned}
\tag{21}
$$

has only a trivial solution $\eta = 0$, where $\hat{Q}_i^*$ is defined in (17), and $\eta \in \mathbb{R}^m$. From (20) we can observe that a critical direction $h \in \mathcal{C}(y^*) \setminus \{0\}$ satisfies

$$
\begin{aligned}
A_i^T h &= 0, && i \in \mathcal{B}, \\
((x^*)^i)^T A_i^T h &= 0, && i \in \mathcal{R} \cup \mathcal{T}_2, \\
(\hat{Q}_i^*)^T A_i^T h &= 0, && i \in \mathcal{T}_2,
\end{aligned}
$$

where the last two equalities hold, because $-A_i^T h = \rho R_i (x^*)^i$ for some $\rho \geq 0$, and the columns of $\hat{Q}_i^*$ are orthogonal to both $(x^*)^i$ and $R_i (x^*)^i$ for $i \in \mathcal{T}_2$. Therefore, we have $(\hat{Q}_i^*)^T A_i^T h \neq 0$ for some $i \in \mathcal{R}$, since otherwise we would get a nontrivial solution η for (21). Consequently, from $(\hat{Q}_i^*)^T A_i^T h \neq 0$ and $((x^*)^i)^T A_i^T h = 0$ it can be deducted that $A_i^T h \notin \mathbb{L}_+^{n_i}$ and $-A_i^T h \notin \mathbb{L}_+^{n_i}$ for all $h \in \mathcal{C}(y^*) \setminus \{0\}$.

Example 2.6. Consider the following SOCO problem from [2]:

$$
\begin{aligned}
\min \quad & -x_2^1 \\
\text{s.t.} \quad & x_1^1 = 1, \\
& 2x_2^1 + x_3^1 - x_2^2 = 0, \\
& 2x_3^1 - x_3^2 = 0, \\
& x_1^2 = 2, \\
& x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2}, \\
& x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}.
\end{aligned}
\tag{22}
$$

The SOCO problem (22) satisfies the interior point condition, and it has the unique primal-dual optimal solution

$$
x^* = (1, 1, 0, 2, 2, 0)^T, \quad y^* = (-1, 0, 0, 0)^T, \quad s^* = (1, -1, 0, 0, 0, 0)^T,
$$

which fails strict complementarity. The optimal partition is given by

$$\mathcal{R} = \{1\}, \quad \mathcal{T}_2 = \{2\}, \quad \mathcal{B} = \mathcal{N} = \mathcal{T}_1 = \mathcal{T}_3 = \emptyset.$$

We can check that both the primal and dual nondegeneracy conditions hold. Note that

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \bar{Q}_1^* = \bar{Q}_2^* = \begin{pmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1 \end{pmatrix},$$

which gives

$$(A_1\bar{Q}_1^*, A_2\bar{Q}_2^*) = \begin{pmatrix} 1/\sqrt{2} & 0 & 0 & 0 \\ \sqrt{2} & 1 & -1/\sqrt{2} & 0 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 1/\sqrt{2} & 0 \end{pmatrix}, \ (A_1R_1(s^*)^1, A_2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Since both of these matrices are nonsingular, the unique optimal solution is primal and dual nondegenerate. For the dual problem the cone of critical directions is given by

$$\mathcal{C}(y^*) = \begin{cases} h \in \mathbb{R}^4, \\ ((x^*)^1)^T A_1^T h = 0, \\ -A_2^T h \in \mathbb{R}_+ \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} \end{cases},$$

which is equivalent to $\mathcal{C}(y^*) = \left\{h \in \mathbb{R}^4 \mid h_1 \geq 0,\ h_3 = 0,\ h_2 = h_4 = -\frac{1}{2}h_1 \leq 0\right\}$. Therefore, we get

$$-\frac{(x^*)_1^1}{(s^*)_1^1} A_1 R_1 A_1^T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 5 & 2 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

which implies that

$$h^T H(y^*, x^*) h = h_2^2 > 0, \qquad \forall h \in \mathcal{C}(y^*) \setminus \{0\}.$$

Thus, the second-order sufficient condition holds at $(y^*; s^*)$.

3. Quadratic convergence to the unique optimal solution

In this section, under the primal and dual nondegeneracy conditions, we establish quadratic convergence of Newton's method to the unique optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$). To that end, we need the optimal partition $(\mathcal{B}, \mathcal{N}, \mathcal{R}, \mathcal{T})$ to be known and $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$ to be correctly identified. Hence, it is assumed that $\mu < \tilde{\mu}$ allows for a complete identification of $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$.

Lemma 3.1. *Assume that the primal and dual nondegeneracy conditions hold. Then $\mathcal{R} = \emptyset$ implies $\mathcal{T} = \emptyset$.*

Proof. Suppose that $\mathcal{R} = \emptyset$ and $\mathcal{T} \neq \emptyset$, and $(x^*; y^*; s^*)$ is the unique[2] optimal solution of (P_{SOCO}) and (D_{SOCO}). Then, for every possible case in which $\mathcal{T}_1$, $\mathcal{T}_2$, or $\mathcal{T}_3$ is nonempty, the number of columns in (18) is strictly greater than the number of columns in (17), i.e., (17) and (18) have $\sum_{i \in \mathcal{B} \cup \mathcal{T}_2} n_i - |\mathcal{T}_2|$ and $\sum_{i \in \mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2} n_i + |\mathcal{T}_3|$ columns, respectively. Thus, (17) and (18) cannot be simultaneously of full row rank and full column rank. □

As a result of Lemma 3.1, if $\mathcal{R} = \emptyset$, then $A_{\mathcal{B}}$ is a nonsingular matrix by the primal and dual nondegeneracy conditions. Therefore, the unique optimal solutions of (P_{SOCO}) and (D_{SOCO}) can be obtained by solving two linear systems of equations. Hence, in the sequel we assume that $\mathcal{R} \neq \emptyset$.

Let $(x^*; y^*; s^*)$ be the unique optimal solution of (P_{SOCO}) and (D_{SOCO}) which satisfies the primal and dual nondegeneracy conditions. Further, let us assume that $\mathcal{T}_1, \mathcal{T}_3 \neq \emptyset$. If we drop the dual constraints $c^i - A_i^T y \in \mathbb{L}_+^{n_i}$ for $i \in \mathcal{T}_1 \cup \mathcal{T}_3$, then we obtain a relaxation of (D_{SOCO}) as

$$(\text{D}'_{\text{SOCO}}) \quad \max \left\{ b^T y \mid A_i^T y + s^i = c^i, \quad s^i \in \mathbb{L}_+^{n_i}, \quad i \in \{1, \ldots, p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\} \right\},$$

and its dual is written as

$$(\text{P}'_{\text{SOCO}}) \quad \min \Big\{ \sum_{i \in \{1,\ldots,p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\}} (c^i)^T x^i \mid \sum_{i \in \{1,\ldots,p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\}} A_i x^i = b, \quad x^i \in \mathbb{L}_+^{n_i}, \quad i \in \{1, \ldots, p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\} \Big\}.$$

Since $(x^*)^i = 0$ for $i \in \mathcal{T}_1 \cup \mathcal{T}_3$, it follows from the optimality conditions, (17), and (18) that $((x^*)^i; y^*; (s^*)^i)$ for $i \in \{1, \ldots, p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\}$ is a primal-dual optimal solution for (P'_{SOCO}) and (D'_{SOCO}), and it satisfies the primal and dual nondegeneracy conditions. To see this, the primal nondegeneracy condition is the same as the one for x^*, and the dual nondegeneracy condition needs

$$\big((A_i R_i (s^*)^i)_{i \in \mathcal{R}}, \; A_{\mathcal{B} \cup \mathcal{T}_2}\big)$$

to have linearly independent columns, which is true by the dual nondegeneracy of $(y^*; s^*)$. As a result, if we remove the columns of $\mathcal{T}_1$ and $\mathcal{T}_3$ from A and c, we can recover the unique optimal solutions of (P_{SOCO}) and (D_{SOCO}) by solving (P'_{SOCO}) and (D'_{SOCO}). At the risk of causing confusion, we refer to $(\bar{x}; \bar{y}; \bar{s})$ as the unique optimal solution of (P'_{SOCO}) and (D'_{SOCO}).

The algebraic definition (1) can be used to reformulate (D'_{SOCO}) as a nonconvex NLO problem. Then inspired by the optimal partition information and the characteristics of a maximally complementary optimal solution, one can realize that the unique dual optimal solution $(\bar{y}; \bar{s})$ can be obtained by solving the NLO reformulation of (D'_{SOCO})

[2] Otherwise, the nondegeneracy conditions fail.

as

$$
\begin{aligned}
(\mathrm{D_{NLO}}) \quad \min \quad & -b^T w \\
\text{s.t.} \quad & A_i^T w = c^i, && i \in \mathcal{B} \cup \mathcal{T}_2, \\
& A_i^T w + z^i = c^i, && i \in \mathcal{R} \cup \mathcal{N}, \\
& (z^i)^T R_i z^i = 0, && i \in \mathcal{R}, \\
& z \in \mathcal{W},
\end{aligned}
$$

where $w \in \mathbb{R}^m$, $z^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{R} \cup \mathcal{N}$, and $\mathcal{W}$ is a nonempty open convex cone defined as

$$
\mathcal{W} := \left\{ z \mid z_1^i > 0,\ i \in \mathcal{R}, \quad z^i \in \mathrm{int}(\mathbb{L}_+^{n_i}),\ i \in \mathcal{N} \right\}.
$$

Let z denote the concatenation of the column vectors z^i for $i \in \mathcal{R} \cup \mathcal{N}$. It then follows that $(\mathrm{D_{NLO}})$ has the unique globally optimal solution $(\bar{w}; \bar{z})$, since otherwise the optimality or the uniqueness of $(\bar{y}; \bar{s})$ is contradicted. The unique globally optimal solution is given by

$$
\bar{w} := \bar{y}, \qquad \bar{z}^i := \bar{s}^i, \quad i \in \mathcal{R} \cup \mathcal{N}. \tag{23}
$$

In a similar manner, the unique optimal solution $\bar{x}$ can be computed by solving

$$
\begin{aligned}
(\mathrm{P_{NLO}}) \quad \min \quad & \sum_{i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2} (c^i)^T \nu^i \\
\text{s.t.} \quad & \sum_{i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2} A_i \nu^i = b, \\
& (\nu^i)^T R_i \nu^i = 0, && i \in \mathcal{R} \cup \mathcal{T}_2, \\
& \nu \in \mathcal{V},
\end{aligned}
$$

where $\nu^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2$, and $\mathcal{V}$ is an open convex cone defined as

$$
\mathcal{V} := \left\{ \nu \mid \nu_1^i > 0,\ i \in \mathcal{R} \cup \mathcal{T}_2, \quad \nu^i \in \mathrm{int}(\mathbb{L}_+^{n_i}),\ i \in \mathcal{B} \right\}.
$$

For the sake of convenience, we only consider $(\mathrm{D_{NLO}})$. Analogous results can be derived for problem $(\mathrm{P_{NLO}})$.

Let $u^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}$ and $v \in \mathbb{R}^{|\mathcal{R}|}$ be the Lagrange multipliers associated with the constraints in $(\mathrm{D_{NLO}})$. The first-order optimality conditions for $(\mathrm{D_{NLO}})$ are given by

$$
\begin{aligned}
-\sum_{i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}} A_i u^i &= b, \\
-u^i - 2 v_i R_i z^i &= 0, && i \in \mathcal{R}, \\
-u^i &= 0, && i \in \mathcal{N}, \\
A_i^T w &= c^i, && i \in \mathcal{B} \cup \mathcal{T}_2, \\
A_i^T w + z^i &= c^i, && i \in \mathcal{R} \cup \mathcal{N}, \\
(z^i)^T R_i z^i &= 0, && i \in \mathcal{R}, \\
z &\in \mathcal{W},
\end{aligned} \tag{24}
$$

which bears a striking resemblance to the optimality conditions (3). Let u be the concatenation of the column vectors u^i for $i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}$. Then we can observe that for $\bar{z} \in \mathcal{W}$ there exist Lagrange multipliers $\bar{u}$ and $\bar{v}$ so that $(\bar{w}; \bar{z}; \bar{u}; \bar{v})$ satisfies the first-order optimality conditions (24). Such a solution can be obtained by setting

$$\begin{aligned} \bar{u}^i &:= -\bar{x}^i, & i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R}, \\ \bar{u}^i &:= 0, & i \in \mathcal{N}, \\ \bar{v}_i &:= \frac{1}{2}\frac{\bar{x}_1^i}{\bar{s}_1^i}, & i \in \mathcal{R}. \end{aligned} \tag{25}$$

We show in Lemma 3.2 that, under the dual nondegeneracy condition, the Lagrange multipliers $(\bar{u}; \bar{v})$ are unique. Let $J((w; z))$ denote the Jacobian of the equality constraints in $(\mathrm{D_{NLO}})$ as follows

$$J((w;z)) := \begin{pmatrix} A_{\mathcal{B}}^T & 0 & 0 \\ A_{\mathcal{T}_2}^T & 0 & 0 \\ A_{\mathcal{R}}^T & I & 0 \\ A_{\mathcal{N}}^T & 0 & I \\ 0 & Z_{\mathcal{R}} & 0 \end{pmatrix}, \tag{26}$$

where $Z_{\mathcal{R}}$ is given by

$$Z_{\mathcal{R}} := \begin{pmatrix} 2(z_1^1; -z_{2:n_1}^1)^T & 0 & 0 & 0 \\ & 2(z_1^2; -z_{2:n_2}^2)^T & & \\ 0 & 0 & \ddots & 0 \\ & & & 2(z_1^i; -z_{2:n_i}^i)^T \\ 0 & 0 & 0 & \ddots \end{pmatrix},$$

in which $i \in \mathcal{R}$. Note that $Z_{\mathcal{R}}$ has full row rank since $(z^i)^T R_i \neq 0$ for every $i \in \mathcal{R}$.

Lemma 3.2. *Let $(\bar{w}; \bar{z})$ be the unique globally optimal solution of* $(\mathrm{D_{NLO}})$. *Then, under the dual nondegeneracy condition, $J((\bar{w}; \bar{z}))$ has full row rank.*

Proof. We show that $J((\bar{w}; \bar{z}))^T \eta = 0$ has only the trivial solution $\eta = 0$, where $\eta := (\eta^1; \ldots; \eta^5)$ is a vector of appropriate size. Then from $J((\bar{w}; \bar{z}))^T \eta = 0$ we have

$$\begin{aligned} A_{\mathcal{B}}\eta^1 + A_{\mathcal{T}_2}\eta^2 + A_{\mathcal{R}}\eta^3 + A_{\mathcal{N}}\eta^4 &= 0, \\ \eta^3 + \bar{Z}_{\mathcal{R}}^T \eta^5 &= 0, \\ \eta^4 &= 0, \end{aligned}$$

which implies

$$A_{\mathcal{B}}\eta^1 + A_{\mathcal{T}_2}\eta^2 - A_{\mathcal{R}}\bar{Z}_{\mathcal{R}}^T\eta^5 = 0,$$

where $A_{\mathcal{R}}\bar{Z}_{\mathcal{R}}^T = (2A_1R_1\bar{z}^1, \ldots, 2A_iR_i\bar{z}^i, \ldots)$ for $i \in \mathcal{R}$. Since $(\bar{y}; \bar{s})$ is the unique dual nondegenerate optimal solution of $(\mathrm{D}'_{\mathrm{SOCO}})$, it follows from (18) that $(A_{\mathcal{R}}\bar{Z}_{\mathcal{R}}^T,\ A_{\mathcal{B}\cup\mathcal{T}_2})$ has full column rank, and thus $\eta = 0$ is the unique solution of $J((\bar{w}; \bar{z}))^T \eta = 0$. □

Under the full rank result of Lemma 3.2, the linear independence constraint qualification (LICQ) [30] holds at $(\bar{w}; \bar{z})$. This regularity condition guarantees that the set of Lagrange multipliers associated with $(\bar{w}; \bar{z})$ is a singleton.

For the sake of simplicity let $\vartheta := (w; z; u; v)$. The Lagrange function of $(\mathrm{D_{NLO}})$ is defined as

$$L(\vartheta) := \\ -b^T w - \sum_{i \in \mathcal{B}\cup\mathcal{T}_2} (u^i)^T (A_i^T w - c^i) - \sum_{i \in \mathcal{R}\cup\mathcal{N}} (u^i)^T (A_i^T w + z^i - c^i) - \sum_{i \in \mathcal{R}} v_i (z^i)^T R_i z^i,$$

and the Hessian of $L(\vartheta)$ is given by

$$\nabla^2 L(\vartheta) := \begin{pmatrix} 0 & 0 & 0 \\ 0 & V_{\mathcal{R}} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where

$$V_{\mathcal{R}} := -2\mathrm{diag}(v_1 R_1, v_2 R_2, \ldots, v_i R_i, \ldots)$$

is a block diagonal matrix, in which $i \in \mathcal{R}$. Let $h = (h^1; h^2; h^3) \in \mathrm{Ker}(J((\bar{w}; \bar{z})))$, where $h^1 \in \mathbb{R}^m$ and h^2 as well as h^3 is the concatenation of the vectors $(h^2)^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{R}$ and $(h^3)^i \in \mathbb{R}^{n_i}$ for $i \in \mathcal{N}$, respectively. In Lemma 3.3, we show that under the primal nondegeneracy condition, the second-order sufficient condition for ($\mathrm{D_{NLO}}$) holds at $(\bar{w}; \bar{z})$, i.e.,

$$h^T \nabla^2 L(\bar{\vartheta}) h > 0, \quad \forall h \in \mathrm{Ker}(J((\bar{w}; \bar{z}))) \setminus \{0\}, \tag{27}$$

in which $\bar{\vartheta} := (\bar{w}; \bar{z}; \bar{u}; \bar{v})$.

Lemma 3.3. *Let $(\bar{w}; \bar{z})$ be the unique globally optimal solution of ($\mathrm{D_{NLO}}$). Then, under the primal nondegeneracy condition, the second-order sufficient condition (27) holds at $(\bar{w}; \bar{z})$.*

Proof. Note that $\mathrm{Ker}(J((\bar{w}; \bar{z})))$ can be equivalently written as the solution set of

$$\begin{aligned} A_i^T h^1 &= 0, \quad i \in \mathcal{B} \cup \mathcal{T}_2, \\ A_i^T h^1 + (h^2)^i &= 0, \quad i \in \mathcal{R}, \\ A_i^T h^1 + (h^3)^i &= 0, \quad i \in \mathcal{N}, \\ (\bar{z}^i)^T R_i (h^2)^i &= 0, \quad i \in \mathcal{R}. \end{aligned} \tag{28}$$

Then we get

$$h^T \nabla^2 L(\bar{\vartheta}) h = -2 \sum_{i \in \mathcal{R}} \bar{v}_i ((h^2)^i)^T R_i (h^2)^i = -2 \sum_{i \in \mathcal{R}} \bar{v}_i (h^1)^T A_i R_i A_i^T h^1.$$

By the primal nondegeneracy condition and the argument after Lemma 2.5, for the unique primal optimal solution $\bar{x}$ system (21) has only a trivial solution. Thus, we have $(\hat{Q}_i^*)^T A_i^T h^1 \neq 0$ for some $i \in \mathcal{R}$, where $\hat{Q}_i^*$ is defined as in (17). Hence, it follows from (25) and (28) that $\bar{v}_i (h^1)^T A_i R_i A_i^T h^1 < 0$ for Lagrange multipliers $(\bar{u}; \bar{v})$ for all $h^1 \neq 0$ satisfying (28). □

Example 3.4. Problem (22) in the nonlinear format ($\mathrm{P_{NLO}}$) has four isolated locally optimal solutions

$$\begin{aligned} \nu_{(1)} &= (1,\ -0.2425,\ 0.9701,\ 2,\ 0.4851,\ 1.9403)^T, \\ \nu_{(2)} &= (1,\ -1,\ 0,\ 2,\ -2,\ 0)^T, \\ \nu_{(3)} &= (1,\ 0.2425,\ -0.9701,\ 2,\ -0.4851,\ -1.9403)^T, \\ \nu_{(4)} &= (1,\ 1,\ 0,\ 2,\ 2,\ 0)^T, \end{aligned}$$

where the objective values are 0.2425, 1, -0.2425, and -1, respectively. The second-order constraint $x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}$ is weakly inactive at $\nu_{(4)}$, i.e., its removal does not affect the optimality of $\nu_{(4)}$. Removing the weakly inactive constraint reduces the set of locally optimal solutions to $\{\nu_{(2)}, \nu_{(4)}\}$ but leaves the set of globally optimal solutions $\{\nu_{(4)}\}$ unchanged. Note that the Jacobian matrix (26) is nonsingular at the unique globally optimal solution $(\bar{w}; \bar{z})$. Therefore, the second-order sufficient condition (27) trivially holds at $(\bar{w}; \bar{z})$.

3.1. Quadratic convergence of Newton's method

We apply Newton's method to the first-order optimality conditions of (D_{NLO}). The idea is to start from a central solution, for which μ satisfies (16), and take Newton steps to converge to $\bar{\vartheta}$. The first-order optimality conditions (24) can be written as $G(\vartheta) = 0$ and $z \in \mathcal{W}$, where the mapping $G : \mathbb{R}^{\bar{n}_c} \to \mathbb{R}^{\bar{n}_c}$ is defined as

$$G(\vartheta) := \begin{pmatrix} -\sum_{i\in\mathcal{B}\cup\mathcal{T}_2\cup\mathcal{R}\cup\mathcal{N}} A_i u^i - b & \\ -u^i - 2v_i R_i z^i & i \in \mathcal{R} \\ -u^i & i \in \mathcal{N} \\ A_i^T w - c^i & i \in \mathcal{B}\cup\mathcal{T}_2 \\ A_i^T w + z^i - c^i & i \in \mathcal{R}\cup\mathcal{N} \\ (z^i)^T R_i z^i & i \in \mathcal{R} \end{pmatrix}, \tag{29}$$

in which

$$\bar{n}_c := \sum_{i\in\mathcal{B}\cup\mathcal{T}_2\cup\mathcal{R}\cup\mathcal{N}} n_i + \sum_{i\in\mathcal{R}\cup\mathcal{N}} n_i + |\mathcal{R}| + m.$$

For ease of exposition, the equations of (24) are indexed in mapping G. The Jacobian of G is given by

$$\nabla G(\vartheta) := \begin{pmatrix} \nabla^2 L(\vartheta) & -J((w;z))^T \\ J((w;z)) & 0 \end{pmatrix}.$$

Letting $\vartheta^{(k)}$ be the k^{th} iterate, a Newton step is taken by computing

$$\vartheta^{(k+1)} := \vartheta^{(k)} + d\vartheta^{(k)}, \qquad d\vartheta^{(k)} := \big(dw^{(k)}; dz^{(k)}; du^{(k)}; dv^{(k)}\big), \tag{30}$$

where the search direction $d\vartheta^{(k)}$ is obtained by solving

$$\nabla G(\vartheta^{(k)}) d\vartheta^{(k)} = -G(\vartheta^{(k)}). \tag{31}$$

Lemma 3.2 shows that $J((\bar{w};\bar{z}))$ is of full row rank, and by Lemma 3.3 it holds that $L(\bar{\vartheta})$ has a positive curvature in the null space of $J((\bar{w};\bar{z}))$. Now, we show that $\nabla G(\bar{\vartheta})$ is nonsingular.

Lemma 3.5. *Assume that the primal and dual nondegeneracy conditions hold. Then $\nabla G(\bar{\vartheta})$ is nonsingular.*

Proof. Let $\eta := (\eta^1; \eta^2)$ be a vector of appropriate size and consider the linear system $\nabla G(\bar{\vartheta})\eta = 0$. Then we have

$$\begin{aligned} \nabla^2 L(\bar{\vartheta})\eta^1 - J((\bar{w};\bar{z}))^T \eta^2 &= 0, \\ J((\bar{w};\bar{z}))\eta^1 &= 0. \end{aligned}$$

From the first equation we have $(\eta^1)^T \nabla^2 L(\bar{\vartheta})\eta^1 = 0$, which implies $\eta^1 = 0$ by Lemma 3.3. Setting $\eta^1 = 0$, the first equation gives $J((\bar{w};\bar{z}))^T \eta^2$, which implies $\eta^2 = 0$ by Lemma 3.2. □

The next lemma shows that ∇G is Lipschitz continuous, regardless of any regularity condition.

Lemma 3.6. *The Jacobian ∇G is Lipschitz continuous with global Lipschitz constant $\tau_1 := 2\sqrt{2}$.*

Proof. Let $\xi := (\xi^1; \ldots; \xi^8)$ be a vector of appropriate size. Then we have

$$\begin{aligned}\|\nabla G(\vartheta) - \nabla G(\vartheta')\| &\le \max_{\|\xi\|=1} \|(V_{\mathcal{R}} - V'_{\mathcal{R}})\xi^2\| \\ &\quad + \max_{\|\xi\|=1} \|((Z'_{\mathcal{R}})^T - Z^T_{\mathcal{R}})\xi^8\| + \max_{\|\xi\|=1} \|(Z_{\mathcal{R}} - Z'_{\mathcal{R}})\xi^2\| \\ &\le \max_{\|\xi^2\|=1} \|(V_{\mathcal{R}} - V'_{\mathcal{R}})\xi^2\| + 2 \max_{\|\xi^2\|=1} \|(Z_{\mathcal{R}} - Z'_{\mathcal{R}})\xi^2\| \\ &= \|V_{\mathcal{R}} - V'_{\mathcal{R}}\| + 2\|Z_{\mathcal{R}} - Z'_{\mathcal{R}}\|.\end{aligned}$$

Then from the properties of the spectral norm we get

$$\begin{aligned}\|V_{\mathcal{R}} - V'_{\mathcal{R}}\| &\le 2 \max_{i\in\mathcal{R}} |v_i - v'_i| \le 2\|v - v'\|, \\ \|Z_{\mathcal{R}} - Z'_{\mathcal{R}}\| &\le \sqrt{\max_{i\in\mathcal{R}} \|z^i - (z')^i\|^2} \le \|z - z'\|.\end{aligned}$$

All this gives

$$\|\nabla G(\vartheta) - \nabla G(\vartheta')\| \le 2(\|v - v'\| + \|z - z'\|) \le 2\sqrt{2}\|\vartheta - \vartheta'\|,$$

for all ϑ and ϑ'. □

The following lemma will be useful for establishing the quadratic convergence of Newton's method.

Lemma 3.7. *Let $(x(\mu); y(\mu); s(\mu))$ be a central solution with $\mu \le \hat{\mu}$, where $\hat{\mu}$ is defined by (10), $(\bar{x}; \bar{y}; \bar{s})$ be the unique optimal solution of $(\mathrm{P}'_{\mathrm{SOCO}})$ and $(\mathrm{D}'_{\mathrm{SOCO}})$, and $(x^*; y^*; s^*)$ be the unique optimal solution of $(\mathrm{P}_{\mathrm{SOCO}})$ and $(\mathrm{D}_{\mathrm{SOCO}})$. Then, under the primal and dual nondegeneracy conditions, we have*

$$\sqrt{\sum_{i\in\mathcal{R}} \left(\frac{x^i_1(\mu)}{s^i_1(\mu)} - \frac{\bar{x}^i_1}{\bar{s}^i_1}\right)^2} \le \frac{4p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right). \tag{32}$$

Proof. Note that for every $i \in \mathcal{R}$ we have

$$\frac{\bar{x}^i_{2:n_i}}{\|\bar{x}^i_{2:n_i}\|} = -\frac{\bar{s}^i_{2:n_i}}{\|\bar{s}^i_{2:n_i}\|}. \tag{33}$$

Since $\bar{x}^i = (x^*)^i$ and $\bar{s}^i = (s^*)^i$ for $i \in \mathcal{R}$, it follows from (6) and (33) that for every $i \in \mathcal{R}$

$$\begin{aligned}\sigma_2 \le \bar{x}^i_1 + \bar{s}^i_1 - \|\bar{x}^i_{2:n_i} + \bar{s}^i_{2:n_i}\| &= \bar{x}^i_1 + \bar{s}^i_1 - |\bar{x}^i_1 - \bar{s}^i_1| \\ &= 2\min\{\bar{x}^i_1, \bar{s}^i_1\}.\end{aligned} \tag{34}$$

Furthermore, it holds that

$$\begin{aligned}\left|\frac{x^i_1(\mu)}{s^i_1(\mu)} - \frac{\bar{x}^i_1}{\bar{s}^i_1}\right| &= \left|\left(\frac{x^i_1(\mu)}{s^i_1(\mu)} - \frac{\bar{x}^i_1}{s^i_1(\mu)}\right) + \left(\frac{\bar{x}^i_1}{s^i_1(\mu)} - \frac{\bar{x}^i_1}{\bar{s}^i_1}\right)\right| \\ &\le \frac{1}{s^i_1(\mu)}|x^i_1(\mu) - \bar{x}^i_1| + \bar{x}^i_1\left|\frac{\bar{s}^i_1 - s^i_1(\mu)}{s^i_1(\mu)\bar{s}^i_1}\right| \\ &\le \frac{1}{s^i_1(\mu)}\|x^i(\mu) - \bar{x}^i\| + \frac{\bar{x}^i_1}{s^i_1(\mu)\bar{s}^i_1}\|s^i(\mu) - \bar{s}^i\| \le \frac{\kappa(p\mu)^\gamma}{s^i_1(\mu)}\left(1 + \frac{\bar{x}^i_1}{\bar{s}^i_1}\right),\end{aligned}$$

where the last inequality follows from (11). Now using (7), (34), and Theorem 2.3 we get

$$\left|\frac{x^i_1(\mu)}{s^i_1(\mu)} - \frac{\bar{x}^i_1}{\bar{s}^i_1}\right| \le \frac{4p\kappa(p\mu)^\gamma}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right),$$

which completes the proof. $\square$

Let Newton's method be initiated with a given interior solution

$$\begin{aligned} w^{(0)} &:= y(\mu), && \\ (z^i)^{(0)} &:= s^i(\mu), && i \in \mathcal{R} \cup \mathcal{N}, \\ (u^i)^{(0)} &:= -x^i(\mu), && i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}, \\ v_i^{(0)} &:= \frac{1}{2}\frac{x_1^i(\mu)}{s_1^i(\mu)}, && i \in \mathcal{R}. \end{aligned} \tag{35}$$

Then a search direction is computed by using (31), and the new iterate is obtained by (30). Theorem 3.9 shows that if μ is sufficiently small, then Newton's method converges quadratically to the unique optimal solution $(\bar{x};\bar{y};\bar{s})$. To that end, we adopt the quadratic convergence result of Newton's method from Theorem 5.2.1 in [13].

Lemma 3.8. *Consider a continuously differentiable mapping $G : \mathbb{R}^n \to \mathbb{R}^n$ on an open convex set $\mathcal{C} \subseteq \mathbb{R}^n$. Let $x^* \in \mathbb{R}^n$ be a root of $G(x) = 0$ so that $B(x^*, r) \subseteq \mathcal{C}$ for some $r > 0$. If ∇G is Lipschitz continuous with constant τ on $B(x^*, r)$ and $\|\nabla G(x^*)^{-1}\| \leq \beta$ for some $\beta > 0$, then for a given $x^{(0)} \in B(x^*, \epsilon)$, where*

$$\epsilon := \min\left\{r, \frac{1}{2\beta\tau}\right\}, \tag{36}$$

the Newton iterates $x^{(k)}$ are well-defined and converge to x^* so that

$$\|x^{(k+1)} - x^*\| \leq \beta\tau \|x^{(k)} - x^*\|^2, \qquad k \geq 0.$$

Now, we present the following result, as planned.

Theorem 3.9. *Assume that the primal and dual nondegeneracy conditions hold. Let*

$$\mu < \min\left\{p^{-1}\left(4\sqrt{2}\beta_1\kappa\Big(\sqrt{3} + \frac{2p\sqrt{|\mathcal{R}|}}{\sigma_2}(1 + \frac{2\sigma_3}{\sigma_2})\Big)\right)^{-\frac{1}{\gamma}},\ \tilde{\mu}\right\}, \tag{37}$$

in which β_1 denotes an upper bound on $\|\nabla G(\bar{\vartheta})^{-1}\|$, and $\tilde{\mu}$ is defined in (16). Then, initiated as given in (35), Newton's method converges to $\bar{\vartheta}$ with quadratic rate. In particular, the convergence to the unique optimal solution $(\bar{x};\bar{y};\bar{s})$ is quadratic.

Proof. By Lemmas 3.5 and 3.6, the conditions of Lemma 3.8 hold, and we get

$$\epsilon := \frac{1}{4\sqrt{2}\beta_1}.$$

Therefore, the Newton steps are well-defined in the neighborhood $B(\bar{\vartheta}, \epsilon)$, and the convergence of Newton's method to $\bar{\vartheta}$ is quadratic if $\vartheta^{(0)} \in B(\bar{\vartheta}, \epsilon)$. The quadratic convergence to $(\bar{x};\bar{y};\bar{s})$ follows from (23) and (25). Using the bounds in Theorem 2.3 and (32) we get

$$\|v^{(0)} - \bar{v}\| = \sqrt{\frac{1}{4}\sum_{i\in\mathcal{R}}\left(\frac{x_1^i(\mu)}{s_1^i(\mu)} - \frac{\bar{x}_1^i}{\bar{s}_1^i}\right)^2} \leq \frac{2p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\Big(1 + \frac{2\sigma_3}{\sigma_2}\Big).$$

Then, considering the error bounds given in (11), we obtain

$$\begin{aligned}\|\vartheta^{(0)} - \bar{\vartheta}\| &\leq \|(w^{(0)} - \bar{w}; z^{(0)} - \bar{z}; u^{(0)} - \bar{u})\| + \|v^{(0)} - \bar{v}\| \\ &\leq \|(x(\mu) - x^*; y(\mu) - y^*; s(\mu) - s^*)\| + \|v^{(0)} - \bar{v}\| \\ &\leq \sqrt{3}\kappa(p\mu)^\gamma + \frac{2p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\Big(1 + \frac{2\sigma_3}{\sigma_2}\Big),\end{aligned}$$

where $(x^*; y^*; s^*)$ is the unique optimal solution of $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$. The result of the theorem follows if we satisfy

$$\sqrt{3}\kappa(p\mu)^\gamma + \frac{2p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\Big(1 + \frac{2\sigma_3}{\sigma_2}\Big) < \epsilon,$$

or equivalently,

$$(p\mu)^\gamma < \frac{\epsilon}{\kappa\Big(\sqrt{3} + \frac{2p\sqrt{|\mathcal{R}|}}{\sigma_2}\big(1 + \frac{2\sigma_3}{\sigma_2}\big)\Big)}.$$

This completes the proof. □

Recall that $(\bar{x}; \bar{y}; \bar{s})$ is the unique optimal solution for $(\mathrm{P'_{SOCO}})$ and $(\mathrm{D'_{SOCO}})$. If $\mathcal{T}_1, \mathcal{T}_3 \neq \emptyset$, then we can recover the unique optimal solutions of the original problems $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$ by appending $\mathcal{T}_1$ and $\mathcal{T}_3$ so that

$$\begin{aligned} (s^*)^i &:= c^i - A_i^T\bar{y}, && i \in \mathcal{T}_3,\\ (s^*)^i &:= 0, && i \in \mathcal{T}_1,\\ (x^*)^i &:= 0, && i \in \mathcal{T}_1 \cup \mathcal{T}_3.\end{aligned}$$

Remark 1. We can obtain a nontrivial lower bound $\gamma \geq 2^{-p}$ from Theorem 7.4.1 in [22]. Furthermore, a positive lower bound can be computed for the condition numbers σ_1 and σ_2 using the method given in [24]. Under the uniqueness condition, an upper bound can be computed for the condition number σ_3, by representing the set of primal-dual optimal solutions of SOCO as a semi-algebraic set, see Proposition 1.3 in [32]. However, we are not aware of any upper bound on β_1, θ_1, θ_2, and the condition number κ.

Remark 2. Theorem 2.3 can be extended to the case when IPMs generate approximate solutions in a neighborhood of the central path, see Section 4 in [36] for a detailed discussion. Hence, an upper bound analogous to (37) can be derived for approximate solutions in a neighborhood of the central path.

Remark 3. In a special case when the subsets $\mathcal{R}$ and $\mathcal{T}$ are empty, a strictly complementary optimal solution can be obtained as easily as in LO [23, 40], regardless of the nondegeneracy conditions. In fact, a central solution, with sufficiently small μ, can be rounded to an exact strictly complementary optimal solution in strongly polynomial time through solving two least squares problems.

4. Discussion

4.1. Stronger results with strict complementarity condition

When the strict complementarity condition holds in addition to the primal and dual nondegeneracy conditions, the quadratic convergence of Newton's method follows from Theorem 28 in [2]. In this case, a stronger complexity bound can be obtained in order to identify the quadratic convergence region. Note that the optimality conditions (3) can be written as $F((x;y;s)) = 0$ and $x, s \in \mathcal{L}^{\bar{n}}_+$, where the mapping

$$F : \mathbb{R}^{\bar{n}} \times \mathbb{R}^m \times \mathbb{R}^{\bar{n}} \to \mathbb{R}^m \times \mathbb{R}^{\bar{n}} \times \mathbb{R}^{\bar{n}}$$

is given by

$$F((x;y;s)) := \begin{pmatrix} Ax - b \\ A^T y + s - c \\ x \circ s \end{pmatrix}. \tag{38}$$

The Jacobian of F is given by

$$\nabla F((x;y;s)) := \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ L(s) & 0 & L(x) \end{pmatrix},$$

where

$$L(x) := \operatorname{diag}(L(x^1), \ldots, L(x^p)),$$
$$L(s) := \operatorname{diag}(L(s^1), \ldots, L(s^p)).$$

By Theorem 28 in [2], $\nabla F((x^*;y^*;s^*))$ is nonsingular, where $(x^*;y^*;s^*)$ is the unique optimal solution. Furthermore, analogous to Lemma 3.6, we can show that ∇F is Lipschitz continuous with global Lipschitz constant $\tau_2 := 2$. Then the following result is immediate from Lemma 3.8.

Theorem 4.1. *Assume that there exists $\beta_2 > 0$ so that*

$$\|\nabla F((x^*;y^*;s^*))^{-1}\| \leq \beta_2.$$

Let $\hat{\mu}$ be defined by (10) and a central solution $(x(\mu);y(\mu);s(\mu))$ with

$$\mu < \min\left\{p^{-1}\left(4\sqrt{3}\beta_2\kappa\right)^{-\frac{1}{\gamma}},\ \hat{\mu}\right\} \tag{39}$$

be given, where κ and γ are defined as in (11). Then starting from a central solution $(x(\mu);y(\mu);s(\mu))$, Newton's method is quadratically convergent to $(x^*;y^*;s^*)$.

Proof. Since F is continuously differentiable, the result of Lemma 3.8 is valid. Hence, Newton steps are well-defined in a neighborhood of $(x^*;y^*;s^*)$. Additionally, from Lemma 2.2 there exist positive κ and γ so that

$$\|(x(\mu) - x^*; y(\mu) - y^*; s(\mu) - s^*)\| \leq \sqrt{3}\kappa(p\mu)^\gamma.$$

Then it is immediate from (36) that $(x(\mu);y(\mu);s(\mu))$ is in the quadratic convergence region of Newton's method if

$$\sqrt{3}\kappa(p\mu)^\gamma < \frac{1}{4\beta_2},$$

which yields the result. □

Remark 4. Bound (37), relying on the condition numbers σ_1, σ_2, σ_3, and κ, is significantly more complicated than (39). In fact, the intricacy of bound (37) indicates that quadratic convergence is harder to achieve in the absence of strict complementarity. To that end, μ has to be small enough so that the optimal partition can be identified.

4.2. Magnitude of the condition numbers

The condition number σ_1 may have a magnitude in the order of 2^{-L} for LO, where L is the binary length of the parameters, while this condition number could be doubly exponentially small for some instances of SOCO, see Example 22 in [31]. It can be inferred from this comparison that an exact solution of SOCO is generally harder to compute than for an LO problem. Nevertheless, even for the case of LO, very high accuracy, far beyond the double precision arithmetic, might be needed for the computation of an exact solution. For instance, LO solvers, regardless of the algorithm used, fail to produce an accurate solution for an LO problem with a Hilbert matrix of size larger than 20.

5. Numerical experiments

We demonstrate quadratic convergence of Newton's method, applied to the first-order optimality conditions of ($\mathrm{D_{NLO}}$), on some instances of SOCO problems. For the first part of numerical experiments we solve (22) and the following SOCO problem

$$\begin{aligned}
\min \quad & -\frac{1}{2}x_2^1 - \frac{1}{2}x_3^1 \\
\text{s.t.} \quad & x_1^1 = 1, \\
& x_1^2 - x_4^1 = 1, \\
& x_2^2 - x_2^1 = 0, \\
& x_3^2 - x_3^1 = 0, \\
& x_3^1 - x_4^1 - x_1^3 = -1, \\
& x_1^1 \geq \sqrt{(x_2^1)^2 + (x_3^1)^2 + (x_4^1)^2}, \\
& x_1^2 \geq \sqrt{(x_2^2)^2 + (x_3^2)^2}, \\
& x_1^3 \geq 0.
\end{aligned} \tag{40}$$

The SOCO problem (40) has the unique primal-dual optimal solution

$$\begin{aligned}
x^* &= (1,\ 1/\sqrt{2},\ 1/\sqrt{2},\ 0,\ 1,\ 1/\sqrt{2},\ 1/\sqrt{2},\ 1 + 1/\sqrt{2})^T, \\
y^* &= (-1/\sqrt{2},\ 0,\ 0,\ 0,\ 0)^T, \\
s^* &= (1/\sqrt{2},\ -1/2,\ -1/2,\ 0,\ 0,\ 0,\ 0,\ 0)^T,
\end{aligned}$$

and its optimal partition is given by

$$\mathcal{B} = \{3\}, \quad \mathcal{R} = \{1\}, \quad \mathcal{T}_2 = \{2\}, \quad \mathcal{N} = \mathcal{T}_1 = \mathcal{T}_3 = \emptyset.$$

For the second part of this section we generate a set of 10 random SOCO problems which fail the strict complementarity condition, but satisfy both the primal and dual nondegeneracy conditions, see [39] for degenerate random SDO problems. Specifically, we generate random problems with a unique primal-dual optimal solution. We choose m and the optimal partition $(\mathcal{B}, \mathcal{N}, \mathcal{R}, \mathcal{T})$, see Table 1, in such a way that the necessary conditions for primal and dual nondegeneracy are satisfied, i.e.,

$$|\mathcal{R} \cup \mathcal{T}_3| + \sum_{i \in \mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2} n_i \leq m \leq \sum_{i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2} n_i - |\mathcal{R} \cup \mathcal{T}_2|.$$

Then the interior point condition automatically holds by the uniqueness of the optimal solution, see Theorem 5.81 in [8].

Table 1.: The optimal partition and dimension of random problems.

Problem	m	p	$(n_1,\dots,n_p)$	$\mathcal{B}$	$\mathcal{N}$	$\mathcal{R}$	$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_3$
1	7	4	$(3,6,2,2)$	$\{1\}$	$\emptyset$	$\{2\}$	$\{3\}$	$\emptyset$	$\{4\}$
2	6	2	$(3,5)$	$\emptyset$	$\emptyset$	$\{2\}$	$\emptyset$	$\{1\}$	$\emptyset$
3	6	3	$(3,6,2)$	$\{1\}$	$\emptyset$	$\{2\}$	$\emptyset$	$\{3\}$	$\emptyset$
4	4	2	$(5,3)$	$\emptyset$	$\emptyset$	$\{1\}$	$\{2\}$	$\emptyset$	$\emptyset$
5	9	5	$(5,6,4,2,3)$	$\{3\}$	$\{2\}$	$\{1,5\}$	$\emptyset$	$\{4\}$	$\emptyset$
6	11	6	$(5,6,5,2,3,2)$	$\{1\}$	$\{3\}$	$\{2,6\}$	$\{5\}$	$\emptyset$	$\{4\}$
7	7	5	$(3,9,3,3,4)$	$\emptyset$	$\{1\}$	$\{2,5\}$	$\emptyset$	$\{4\}$	$\{3\}$
8	18	6	$(10,5,7,8,2,8)$	$\{6\}$	$\emptyset$	$\{1,3,4\}$	$\emptyset$	$\{2,5\}$	$\emptyset$
9	22	7	$(10,5,7,8,2,8,5)$	$\emptyset$	$\emptyset$	$\{1,3,4\}$	$\{6\}$	$\{2,5\}$	$\{7\}$
10	35	7	$(8,8,8,8,8,8,8)$	$\{1,4\}$	$\emptyset$	$\{2,5\}$	$\{7\}$	$\{6\}$	$\{3\}$

For the random problems we generate the unique optimal solution $(x^*;y^*;s^*)$ as follows

$$\begin{aligned}
(x^*)^i_1 &\sim U\big(\|(x^*)^i_{2:n_i}\|+0.1,\ \|(x^*)^i_{2:n_i}\|+100.1\big), && i\in\mathcal{B},\\
(s^*)^i_1 &\sim U\big(\|(s^*)^i_{2:n_i}\|+0.1,\ \|(s^*)^i_{2:n_i}\|+100.1\big), && i\in\mathcal{N},\\
(x^*)^i &\sim U\big(0.1,100.1\big)\times\big(1;(\varrho^*)^i/\|(\varrho^*)^i\|\big), && i\in\mathcal{R}\cup\mathcal{T}_2,\\
(s^*)^i &\sim U\big(0.1,100.1\big)\times\big(1;-(\varrho^*)^i/\|(\varrho^*)^i\|\big), && i\in\mathcal{R}\cup\mathcal{T}_3,\\
y^*_i &\sim U(-100,100) && i=1,\dots,m,
\end{aligned}$$

where

$$\begin{aligned}
(x^*)^i_j &\sim U(-100,100), && j=2,\dots,n_i, && i\in\mathcal{B},\\
(s^*)^i_j &\sim U(-100,100), && j=2,\dots,n_i, && i\in\mathcal{N},\\
(\varrho^*)^i_j &\sim U(-100,100), && j=1,\dots,n_i-1, && i\in\mathcal{R}\cup\mathcal{T}_2\cup\mathcal{T}_3,
\end{aligned}$$

in which $U(.,.)$ denotes the uniform distribution. For the rest of the variables we have $(x^*)^i=0$ for $i\in\mathcal{N}\cup\mathcal{T}_1\cup\mathcal{T}_3$ and $(s^*)^i=0$ for $i\in\mathcal{B}\cup\mathcal{T}_1\cup\mathcal{T}_2$ by definition. We then normalize the vectors by

$$x^* := \frac{x^*}{\|x^*\|},\quad y^* := \frac{y^*}{\|y^*\|},\quad s^* := \frac{s^*}{\|s^*\|}.$$

Analogous to the optimal solution, the entries of A are uniformly distributed in $(-100,100)$. However, we keep generating random A until Assumption 1 and the nondegeneracy conditions in Theorem 2.4 hold[3]. We then compute $A := A/\|A\|$ and generate the right hand side and objective vectors by $b := Ax^*$ and $c := A^Ty^*+s^*$.

We solve all the SOCO problems by using SeDuMi 1.3 [34] included in the CVX optimization package [14, 15] and applying Newton's method to (29) and (38) throughout this section. The codes are run in MATLAB 9.2 environment on a MacBook Pro with Intel Core i5 CPU @ 2.3 GHz and 8GB of RAM. The Newton based approaches are referred to as NLO-Newton and SOCO-Newton, respectively. To solve (22) and (40) by Newton based approaches, we choose a central solution with $\mu=10^{-2}$ as the initial point. However, we choose smaller values of μ for the random problems, as specified in Table 8. For the Newton based approaches $\|F((x^{(k)};y^{(k)};s^{(k)}))\|\le 10^{-14}$ is set as the terminating condition, and the optimality tolerance for SeDuMi is fixed at 10^{-15}. For NLO-Newton we remove the rows and columns associated with $\mathcal{T}_1$ and $\mathcal{T}_3$ from both the primal and dual problems. Additionally, we assign $(x(\mu);y(\mu);s(\mu))$ to $\vartheta^{(0)}$

[3]Since Assumption 1 and the nondegeneracy conditions hold generically, see e.g., [3], the expected number of iterations to get the desired coefficient matrix is one.

according to (35), and we form the solution $(x^{(k)}; y^{(k)}; s^{(k)})$ by setting

$$\begin{aligned}
(x^i)^{(k)} &= -(u^i)^{(k)}, && i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N},\\
y^{(k)} &= w^{(k)},\\
(s^i)^{(k)} &= (z^i)^{(k)}, && i \in \mathcal{R} \cup \mathcal{N},\\
(s^i)^{(k)} &= 0, && i \in \mathcal{B} \cup \mathcal{T}_2,
\end{aligned}$$

since there is no s^i corresponding to $\mathcal{B}$ and $\mathcal{T}_2$ in (29).

Tables 2 to 4 illustrate the numerical results of the Newton based approaches on SOCO problem (22). For NLO-Newton we report both the Newton residuals $\|G(\vartheta^{(k)})\|$ and $\|F((x^{(k)}; y^{(k)}; s^{(k)}))\|$. NLO-Newton meets the stopping condition in only 4 iterations while this number is 21 for SOCO-Newton. As can be observed from Tables 2 and 3, the convergence of NLO-Newton to the unique optimal solution of (22) is quadratic while the convergence for SOCO-Newton is no better than linear. Additionally, SeDuMi arrives at the Newton residual 1.445657×10^{-12}, and in that sense it is less accurate than NLO-Newton and SOCO-Newton.

Table 2.: The numerical results of NLO-Newton on SOCO problem (22).

k	$\|Ax^{(k)} - b\|$	$c^T x^{(k)}$	$\|x^{(k)} \circ s^{(k)}\|$	$\|G(.)\|$	$\|F(.)\|$
0	5.551115E-17	-9.922138E-01	1.000000E-02	9.420176E-02	9.432417E-02
1	5.551115E-17	-1.004244E+00	1.132375E-02	2.952235E-02	1.132375E-02
2	0.000000E+00	-1.000041E+00	5.817148E-05	4.551179E-05	5.817148E-05
3	0.000000E+00	-1.000000E+00	2.149505E-10	1.600968E-10	2.149505E-10
4	0.000000E+00	-1.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00

Table 3.: The numerical results of SOCO-Newton on SOCO problem (22).

k	$\|Ax^{(k)} - b\|$	$c^T x^{(k)}$	$\|x^{(k)} \circ s^{(k)}\|$	$\sigma_{\min}(\nabla F(.))$	$\|F(.)\|$
15	0.000000E+00	-1.000000E+00	1.540356E-11	1.270995E-06	1.540356E-11
16	0.000000E+00	-1.000000E+00	3.850891E-12	6.354972E-07	3.850891E-12
17	2.220446E-16	-1.000000E+00	9.627227E-13	3.177485E-07	9.627227E-13
18	2.220446E-16	-1.000000E+00	2.406807E-13	1.588743E-07	2.406808E-13
19	0.000000E+00	-1.000000E+00	6.017017E-14	7.943713E-08	6.017017E-14
20	0.000000E+00	-1.000000E+00	1.504254E-14	3.971856E-08	1.504254E-14
21	2.220446E-16	-1.000000E+00	3.760636E-15	1.985928E-08	3.768821E-15

Table 4.: The k^{th} iterate of NLO-Newton and SOCO-Newton on SOCO problem (22).

NLO-Newton			SOCO-Newton		
k	$(x_3^1)^{(k)}$	$(x_3^2)^{(k)}$	k	$(x_3^1)^{(k)}$	$(x_3^2)^{(k)}$
0	-6.410764E-02	-1.282153E-01	17	-5.387688E-07	-1.077538E-06
1	8.783536E-03	1.756707E-02	18	-2.693844E-07	-5.387688E-07
2	0.000000E+00	0.000000E+00	19	-1.346922E-07	-2.693844E-07
3	0.000000E+00	0.000000E+00	20	-6.734610E-08	-1.346922E-07
4	0.000000E+00	0.000000E+00	21	-3.367305E-08	-6.734610E-08

The numerical results of Newton based approaches on SOCO problem (40) are summarized in Tables 5 to 7. From Tables 5 and 6 we can observe the quadratic convergence of NLO-Newton, versus linear convergence of SOCO-Newton. Furthermore, Table 7 confirms that NLO-Newton evolves faster toward the unique optimal solution of (40) than SOCO-Newton. NLO-Newton arrives at the Newton residual 1.110223×10^{-16} in only 4 iterations, while SeDuMi ends up with the Newton residual 1.777953×10^{-9}.

Table 5.: The numerical results of NLO-Newton on SOCO problem (40).

k	$\|Ax^{(k)} - b\|$	$c^T x^{(k)}$	$\|x^{(k)} \circ s^{(k)}\|$	$\|G(.)\|$	$\|F(.)\|$
0	1.110223E-16	-6.995520E-01	1.000000E-02	7.979631E-02	8.014507E-02
1	2.220446E-16	-7.102731E-01	6.634877E-03	8.185201E-03	6.634877E-03
2	3.330669E-16	-7.071020E-01	9.464771E-06	6.556691E-06	9.464771E-06
3	0.000000E+00	-7.071068E-01	6.959588E-11	4.962837E-11	6.959588E-11
4	0.000000E+00	-7.071068E-01	1.110223E-16	1.110223E-16	1.110223E-16

Table 6.: The numerical results of SOCO-Newton on SOCO problem (40).

k	$\|Ax^{(k)} - b\|$	$c^T x^{(k)}$	$\|x^{(k)} \circ s^{(k)}\|$	$\sigma_{\min}(\nabla F(.))$	$\|F(.)\|$
15	1.110223E-16	-7.071068E-01	9.289517E-12	1.937402E-06	9.289517E-12
16	1.110223E-16	-7.071068E-01	2.322376E-12	9.687002E-07	2.322376E-12
17	1.110223E-16	-7.071068E-01	5.805962E-13	4.843499E-07	5.805962E-13
18	0.000000E+00	-7.071068E-01	1.452016E-13	2.421749E-07	1.452016E-13
19	2.482534E-16	-7.071068E-01	3.629592E-14	1.210874E-07	3.629677E-14
20	0.000000E+00	-7.071068E-01	9.118149E-15	6.054371E-08	9.118149E-15

Table 7.: The k^{th} iterate of NLO-Newton and SOCO-Newton on SOCO problem (40).

NLO-Newton			SOCO-Newton		
k	$(x_4^1)^{(k)}$	$(x_1^3)^{(k)}$	k	$(x_4^1)^{(k)}$	$(x_1^3)^{(k)}$
0	7.86081643E-02	1.62522185E+00	16	1.28146875E-06	1.70710550E+00
1	-4.67865604E-03	1.71469713E+00	17	6.40734373E-07	1.70710614E+00
2	0.00000000E+00	1.70710195E+00	18	3.20367187E-07	1.70710646E+00
3	0.00000000E+00	1.70710678E+00	19	1.60183593E-07	1.70710662E+00
4	0.00000000E+00	1.70710678E+00	20	8.00917970E-08	1.70710670E+00

We draw a sample of 100 instances for each random SOCO problem and report the average results in Tables 8 to 10. To ensure convergence of NLO-Newton to the unique optimal solution, we choose initial solutions with sufficiently small μ when solving the random problems. Table 8 reports the values of μ as well as the distance of the initial solution and the SeDuMi's solution from the unique optimal solution $\omega^* := (x^*; y^*; s^*)$, where ω^*_{se}, ω^*_{nn}, and ω^*_{sn} stand for the solution output by SeDuMi, NLO-Newton, and SOCO-Newton. For NLO-Newton the solution ω^*_{nn} is formed by $(x^*_{nn}; y^*_{nn}; s^*_{nn})$, where

$$
\begin{aligned}
(x^*_{nn})^i &= -(u^i)^{(\bar{k})}, && i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}, \\
(x^*_{nn})^i &= 0, && i \in \mathcal{T}_1 \cup \mathcal{T}_3, \\
y^*_{nn} &= w^{(\bar{k})}, && \\
(s^*_{nn})^i &= (z^i)^{(\bar{k})}, && i \in \mathcal{R} \cup \mathcal{N}, \\
(s^*_{nn})^i &= c^i - A_i^T w^{(\bar{k})}, && i \in \mathcal{T}_3, \\
(s^*_{nn})^i &= 0, && i \in \mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2,
\end{aligned}
$$

where $\bar{k}$ denotes the index of final iterate.

Tables 9 and 10 demonstrate the numerical results of Newton based approaches on the random SOCO problems. The values of $\|\omega^*_{nn} - \omega^*\|$, $\|\omega^*_{sn} - \omega^*\|$, and #Iter indicate the accuracy and fast convergence of NLO-Newton in comparison to SOCO-Newton.

Table 8.: The initial and optimal solutions for random SOCO problems.

Problem	μ	$\|(x^{(0)}; y^{(0)}; s^{(0)}) - \omega^*\|$	$\|\omega^*_{se} - \omega^*\|$
1	1.00E-05	1.912776E-02	1.592721E-06
2	1.00E-05	9.125473E-02	6.738418E-06
3	1.00E-05	1.643906E-02	2.662254E-06
4	1.00E-05	1.795054E-02	1.607865E-06
5	1.00E-05	2.629707E-02	2.755846E-07
6	1.00E-05	4.050471E-02	3.003945E-05
7	1.00E-05	8.766428E-03	4.772793E-07
8	1.00E-06	1.555120E-02	3.128175E-06
9	1.00E-06	1.049691E-02	6.437721E-07
10	1.00E-06	2.267326E-02	1.741151E-06

Table 9.: The numerical results of NLO-Newton on random SOCO problems.

Problem	#Iter	$\|F((x^{(0)}; y^{(0)}; s^{(0)}))\|$	$\|F(\omega^*_{nn})\|$	$\|\omega^*_{nn} - \omega^*_{se}\|$	$\|\omega^*_{nn} - \omega^*\|$	$\|\omega^*_{nn} - \omega^*_{sn}\|$
1	3	1.958945E-03	2.059106E-16	1.592721E-06	2.021790E-13	7.452722E-07
2	3	1.888355E-02	2.589828E-16	6.738418E-06	1.390557E-14	1.056285E-06
3	3	8.310096E-04	1.908819E-16	2.662245E-06	8.634383E-12	5.002596E-07
4	3	7.542060E-04	2.555563E-16	6.078650E-07	1.555339E-13	4.794085E-07
5	4	2.494958E-03	5.557528E-16	2.755846E-07	1.975277E-14	6.050262E-07
6	4	1.962349E-03	6.526095E-16	3.003945E-05	2.171982E-12	2.831458E-04
7	4	1.149507E-03	6.627321E-16	4.772793E-07	1.107431E-14	4.519775E-07
8	3	1.088135E-04	6.271617E-16	3.128175E-06	7.187394E-13	9.655715E-07
9	4	1.720883E-03	3.192262E-16	6.437721E-07	7.164252E-14	4.071836E-07
10	4	1.960026E-03	5.365570E-16	1.741151E-06	8.266871E-13	8.583896E-07

Table 10.: The numerical results of SOCO-Newton on random SOCO problems.

Problem	#Iter	$\|F((x^{(0)}; y^{(0)}; s^{(0)}))\|$	$\|F(\omega^*_{sn})\|$	$\|\nabla F(\omega^*_{sn})\|$	$\|\omega^*_{sn} - \omega_{se}\|$	$\|\omega^*_{sn} - \omega^*\|$
1	15	1.996282E-06	6.229178E-15	9.947566E-09	1.226415E-06	7.452723E-07
2	16	1.414214E-05	5.999936E-15	1.413926E-08	5.774589E-06	1.056285E-06
3	14	1.732051E-06	4.582760E-15	1.458812E-08	2.463535E-06	5.002683E-07
4	14	1.414214E-06	6.101894E-15	1.575818E-08	1.226030E-06	4.794086E-07
5	15	2.236068E-05	6.644257E-15	1.941088E-08	3.546504E-07	6.050262E-07
6	14	2.449490E-06	6.379693E-15	8.065773E-09	3.108220E-04	2.831458E-04
7	14	2.236068E-06	6.201575E-15	1.642138E-08	5.217249E-07	4.519775E-07
8	13	2.449491E-07	5.319422E-15	4.816730E-09	2.633089E-06	9.655719E-07
9	15	2.645751E-06	4.950853E-15	9.779408E-09	3.668787E-07	4.071836E-07
10	15	2.645751E-06	5.222010E-15	4.778974E-09	1.179792E-06	8.583897E-07

6. Conclusion and future studies

Using the optimal partition of a SOCO problem, we established quadratic convergence of Newton's method to the unique optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) without strict complementarity condition. We showed that if the primal and dual nondegeneracy conditions hold, then $\nabla G(\bar{\vartheta})$ is nonsingular. Furthermore, we derived a complexity bound for identifying the quadratic convergence region of Newton's method from a sequence of central solutions. The numerical results confirmed the quadratic convergence of Newton's method to the unique optimal solution of SOCO in the absence of strict complementarity.

The optimal partition approach in Section 3 can be directly applied to the optimality conditions for ($\mathrm{P}'_{\mathrm{SOCO}}$) and ($\mathrm{D}'_{\mathrm{SOCO}}$). If we assume the primal and dual nondegeneracy conditions, then the optimality conditions are written as

$$\begin{aligned}
\sum_{i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R}} A_i x^i &= b, & & \\
A_i^T y &= c^i, & & i \in \mathcal{B} \cup \mathcal{T}_2, \\
A_i^T y + s^i &= c^i, & & i \in \mathcal{R} \cup \mathcal{N}, \\
x^i \circ s^i &= 0, & & i \in \mathcal{R}, \\
x^i &\in \mathbb{L}_+^{n_i}, & & i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R}, \\
s^i &\in \mathbb{L}_+^{n_i}, & & i \in \mathcal{R} \cup \mathcal{N},
\end{aligned}$$

where the zero variables are set aside. Analogous to (38), the equality constraints in the reduced system can be represented by the mapping $F' : \mathbb{R}^{n'} \to \mathbb{R}^{n'}$, where

$$n' := m + \sum_{i \in \mathcal{B} \cup \mathcal{N} \cup \mathcal{T}_2} n_i + 2 \sum_{i \in \mathcal{R}} n_i.$$

Hence, Newton's method is applicable to the mapping F', since the domain and the range of F' are equal. Nevertheless, the nonsingularity of $\nabla F'$ at the optimal solution should be investigated.

To establish quadratic convergence of Newton's method, we assumed that the optimal partition is known, and that $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$ can be identified from $\mathcal{T}$. The derivation of upper bounds on θ_1, θ_2, κ, and β_1 deserves further research. Further, it is worth investigating the application of numerical algebraic geometry to establish quadratic convergence to an isolated solution without primal or dual nondegeneracy condition.

At this point, we are investigating the parametric analysis of SOCO problems using the notion of the optimal partition.

Funding

This work is supported by the Air force Office of Scientific Research (AFOSR) Grant # FA9550-15-1-0222.

References

[1] I. Adler and F. Alizadeh, *Primal-dual interior point algorithms for convex quadratically constrained and semidefinite optimization problems*, Tech. Rep. RRR-111-95, RUTCOR, Rutgers University, USA, 1995.

[2] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming, 95 (2003), pp. 3–51.

[3] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, *Complementarity and nondegeneracy in semidefinite programming*, Mathematical Programming, 77 (1997), pp. 111–128.

[4] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, *Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results*, SIAM Journal on Optimization, 8 (1998), pp. 746–768.

[5] F. Alizadeh and S. Schmieta, *Optimization with semidefinite, quadratic and linear constraints*, Tech. Rep. RRR 23-97, RUTCOR, Rutgers University, USA, 1997.

[6] A. Ben-Tal and A. Nemirovskii, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, Society for Industrial and Applied Mathematics, 2001.

[7] J. F. Bonnans and H. Ramírez, *Perturbation analysis of second-order cone programming problems*, Mathematical Programming, 104 (2005), pp. 205–227.

[8] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, 2000.

[9] Z. X. Chan and D. Sun, *Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming*, SIAM Journal on Optimization, 19 (2008), pp. 370–396.

[10] X. Chi and S. Liu, *A non-interior continuation method for second-order cone programming*, Optimization, 58 (2009), pp. 965–979.

[11] ——, *A one-step smoothing newton method for second-order cone programming*, Journal of Computational and Applied Mathematics, 223 (2009), pp. 114 – 123.

[12] E. de Klerk, *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*, vol. 65 of the series Applied Optimization, Springer, 2006.

[13] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, 1983.

[14] M. Grant and S. Boyd, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, V. Blondel, S. Boyd, and H. Kimura, eds., Lecture Notes in Control and Information Sciences, Springer-Verlag Limited, 2008, pp. 95–110. http://stanford.edu/~boyd/graph_dcp.html.

[15] ——, *CVX: Matlab software for disciplined convex programming, version 2.1.* http://cvxr.com/cvx, Mar. 2014.

[16] A. Hoffman, *On approximate solutions of systems of linear inequalities*, Journal of Research of the National Bureau of Standards, 49 (1952), pp. 263–265.

[17] C. Kanzow, I. Ferenczi, and M. Fukushima, *On the local convergence of semismooth newton methods for linear and nonlinear second-order cone programs without strict complementarity*, SIAM Journal on Optimization, 20 (2009), pp. 297–320.

[18] C. Kanzow and C. Nagel, *Quadratic convergence of a nonsmooth newton-type method for semidefinite programs without strict complementarity*, SIAM Journal on Optimization, 15 (2005), pp. 654–672.

[19] L. Kong, *Quadratic convergence of a smoothing newton method for symmetric cone programming without strict complementarity*, Positivity, 16 (2012), pp. 297–319.

[20] L. Kong, L. Tunçel, and N. Xiu, *Equivalent conditions for Jacobian nonsingularity in linear symmetric cone programming*, Journal of Optimization Theory and Applications, 148 (2011), pp. 364–389.

[21] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, *Applications of second-order cone programming*, Linear Algebra and its Applications, 284 (1998), pp. 193 – 228.

[22] Z.-Q. Luo and J. Sturm, *Error bounds for mixed semidefinite and second order cone programming*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., vol. 33 of the series Applied Optimization, Springer, 2000, pp. 163–190.

[23] S. Mehrotra and Y. Ye, *Finding an interior point in the optimal face of linear programs*, Mathematical Programming, 62 (1993), pp. 497–515.

[24] A. Mohammad-Nezhad and T. Terlaky, *On the identification of optimal partition for semidefinite optimization*, Tech. Rep. 17T-008, Department of Industrial and Systems Engineering, Lehigh University, 2017.

[25] A. Mohammad-Nezhad and T. Terlaky, *A polynomial primal-dual affine scaling algorithm for symmetric conic optimization*, Computational Optimization and Applications, 66 (2017).

[26] R. D. Monteiro and T. Tsuchiya, *Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions*, Mathematical

Programming, 88 (2000), pp. 61–83.

[27] Y. E. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, vol. 13, Society for Industrial and Applied Mathematics, 1994.

[28] Y. E. Nesterov and M. J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research, 22 (1997), pp. 1–42.

[29] Y. E. Nesterov and M. J. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM Journal on Optimization, 8 (1998), pp. 324–364.

[30] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, second ed., 2006.

[31] M. V. Ramana, *An exact duality theory for semidefinite programming and its complexity implications*, Mathematical Programming, 77 (1997), pp. 129–162.

[32] J. Renegar, *On the computational complexity of approximating solutions for real algebraic formulae*, SIAM Journal on Computing, 21 (1992), pp. 1008–1025.

[33] G. W. Stewart and J.-g. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.

[34] J. F. Sturm, *Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones*, Optimization Methods and Software, 11 (1999), pp. 625–653. Available at http://sedumi.ie.lehigh.edu/.

[35] J. F. Sturm, *Error bounds for linear matrix inequalities*, SIAM Journal on Optimization, 10 (2000), pp. 1228–1248.

[36] T. Terlaky and Z. Wang, *On the identification of the optimal partition of second order cone optimization problems*, SIAM Journal on Optimization, 24 (2014), pp. 385–414.

[37] T. Tsuchiya, *A polynomial primal-dual path-following algorithm for second-order cone programming*, Tech. Rep. 649, The Institute of Statistical Mathematics, Tokyo, Japan, 1997.

[38] T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optimization Methods and Software, 11 (1999), pp. 141–182.

[39] H. Wei and H. Wolkowicz, *Generating and measuring instances of hard semidefinite programs*, Mathematical Programming, 125 (2010), pp. 31–45.

[40] Y. Ye, *On the finite convergence of interior-point algorithms for linear programming*, Mathematical Programming, 57 (1992), pp. 325–335.