ISE

Industrial and Systems Engineering

Quadratic convergence of Newton's method to the optimal solution of second-order conic optimization

ALI MOHAMMAD-NEZHAD AND TAMÁS TERLAKY

Department of Industrial and Systems Engineering, Lehigh University, USA

ISE Technical Report 17T-014

LEHIGH UNIVERSITY

Quadratic convergence of Newton's method to the optimal solution of second-order conic optimization

Ali Mohammad-Nezhad*[1] and Tamás Terlaky†[1]

[1]Department of Industrial and Systems Engineering, Lehigh University, USA

October 21, 2017

Abstract

Under strict complementarity and primal and dual nondegeneracy conditions we establish the quadratic convergence of Newton's method to the unique strictly complementary optimal solution of second-order conic optimization, when the initial point is sufficiently close to the optimal set. When strict complementarity fails but the primal and dual nondegeneracy conditions hold, we show that if the optimal partition of the problem is known, then the application of Newton's method to the optimality conditions of a reduced nonlinear optimization problem results in quadratic convergence to the unique maximally complementary optimal solution. For a special case of the optimal partition, we present a rounding procedure which gives an exact strictly complementary optimal solution in strongly polynomial time.

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

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

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

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

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, 15] for the applications of SOCO.

*Email: alm413@lehigh.edu
†Email: terlaky@lehigh.edu

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:

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

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

$$R_i := \operatorname{diag}\big(1,\ -1,\ -1, \ldots, -1\big). \tag{3}$$

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 [24]. The study on IPMs for SOCO problems was pioneered by the work of Nesterov and Nemirovskii [24], see also Nesterov and Todd [25, 26] 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, 3, 20, 21, 32, 33].

The identification of the optimal partition have been throughly studied for linear optimization (LO) and linear complementarity problems (LCP), see e.g., [12, 28], to generate either a *strictly* or a *maximally complementary* optimal solution, see e.g., [17, 18, 28, 34] for LO and [12, 13, 22] for LCP. The concept of optimal partition was generalized to SDO by Goldfarb and Scheinberg [10]. Yildirim [35] presented a facial description of the optimal partition for linear conic optimization. Bonnans and Ramírez [5] established the algebraic definition of the optimal partition for SOCO. Recently, Terlaky and Wang [31] investigated the identification of the optimal partition for SOCO, and Mohammad-Nezhad and Terlaky [19] considered the identification of the optimal partition for SDO.

The main contribution of this paper is to use the optimal partition of SOCO to formulate a reduced nonlinear optimization (NLO) problem. Then we apply Newton's method to the first-order optimality conditions of the reduced problem to generate the unique maximally complementary optimal solution. 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 maximally complementary optimal solution. Indeed, quadratic convergence does not depend on the strict complementarity condition. This result is notably stronger than what we could obtain by the direct application of Newton's method to the optimality conditions of SOCO. For a special case, we present a rounding procedure which gives an exact strictly complementary optimal solution in strongly polynomial time.

In the rest of Section 1, we provide the reader with the preliminaries regarding the optimal partition and primal and dual nondegeneracy conditions for SOCO. In Section 2, we specialize the second-order sufficient condition for the optimal solutions of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$), and highlight its connection to nondegeneracy conditions. In Section 3, we establish quadratic convergence of Newton's method to the unique strictly complementary optimal solution of SOCO. Section 4 presents the main result of this paper: we establish the quadratic convergence of Newton's method to the unique maximally complementary optimal solution of SOCO. For a special case, a strongly polynomial rounding procedure to an exact strictly complementary optimal solution is given in Section 5. Our conclusions and ideas for future research are presented in Section 6. Supplementary materials are given in the Appendix.

Throughout this paper, $\lambda_j(.)$ denotes an eigenvalue of a matrix, int(.), ri(.), and bd(.) stand for the interior, the relative interior, and the boundary of a set, and Ker(.) and $\mathcal{R}(.)$ denote the null space and the range space of a matrix, respectively. A maximally complementary optimal solution is indicated by $(x^*; y^*; s^*)$, and $B(x^*, r)$ denotes an open ball centered at x^* with radius r. For any $I \subseteq \{1, \ldots, p\}$ and an arbitrary matrix A, $|I|$ denotes the cardinality of I, and A_I represents the corresponding subset of the columns of A. Furthermore, $(N_i)_{i \in I}$ denotes the horizontal arrangement of N_i for $i \in I$. Finally, $\|.\|$ 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 ($\mathrm{P_{SOCO}}$) or ($\mathrm{D_{SOCO}}$), 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.

1.1 Optimality and complementarity

We assume that the coefficient matrix A has full row rank, and the interior point condition holds for both ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$), i.e., there exists a feasible solution $(x; y; s)$ such that for all $i = 1, \dots, p$ we have $x^i, s^i \in \mathrm{int}(\mathbb{L}_+^{n_i})$. 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 [4]. Since strong duality holds, the optimal set for ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$) can be represented as

$$\begin{aligned} Ax &= b, & x &\in \mathcal{L}_+^{\bar{n}}, \\ A^T y + s &= c, & s &\in \mathcal{L}_+^{\bar{n}}, \\ x \circ s &= 0, \end{aligned} \tag{4}$$

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

$$x^i \circ s^i := \begin{pmatrix} (x^i)^T s^i \\ x_1^i s_{2:n_i}^i + s_1^i x_{2:n_i}^i \end{pmatrix}, \qquad i = 1, \dots, p, \tag{5}$$

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 [5].

Definition 1 (Definition 23 in [2]). *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

$$x^* \in \mathrm{ri}(\mathcal{P}^*), \quad (y^*; s^*) \in \mathrm{ri}(\mathcal{D}^*).$$

Under the interior point condition, a maximally complementary optimal solution always exists for SOCO. In contrast with LO, which always has a strictly complementary optimal solution, a SDO or SOCO problem may not have a strictly complementary optimal solution.

1.2 Optimal partition in SOCO

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 $\{1, \dots, p\}$ of the second-order cones is partitioned into four sets $\mathcal{B}, \mathcal{N}, \mathcal{R}$, and $\mathcal{T} := (\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$ as defined in [5]:

$$\begin{aligned} \mathcal{B} &:= \big\{ i \mid x_1^i > \|x_{2:n_i}^i\|, \ \text{ for some } x \in \mathcal{P}^* \big\}, \\ \mathcal{N} &:= \big\{ i \mid s_1^i > \|s_{2:n_i}^i\|, \ \text{ for some } s \in \mathcal{D}^* \big\}, \\ \mathcal{R} &:= \big\{ i \mid x_1^i = \|x_{2:n_i}^i\| > 0, \ \ s_1^i = \|s_{2:n_i}^i\| > 0, \ \text{ for some } (x; y; s) \in \mathcal{P}^* \times \mathcal{D}^* \big\}, \\ \mathcal{T}_1 &:= \big\{ i \mid x^i = s^i = 0, \ \text{ for all } (x; y; s) \in \mathcal{P}^* \times \mathcal{D}^* \big\}, \\ \mathcal{T}_2 &:= \big\{ i \mid s^i = 0, \ \text{ for all } (y; s) \in \mathcal{D}^*, \ \ x_1^i = \|x_{2:n_i}^i\| > 0, \ \text{ for some } x \in \mathcal{P}^* \big\}, \\ \mathcal{T}_3 &:= \big\{ i \mid x^i = 0, \ \text{ for all } x \in \mathcal{P}^*, \ \ s_1^i = \|s_{2:n_i}^i\| > 0, \ \text{ for some } (y; s) \in \mathcal{D}^* \big\}. \end{aligned}$$

The convexity of the optimal set implies that the sets $\mathcal{B}, \mathcal{N}, \mathcal{R}$, and $\mathcal{T}$ are mutually disjoint and their union is the index set $\{1, \ldots, p\}$. For instance, for $i \in \mathcal{T}_2$ we have $s^i = 0$ for all $(y; s) \in \mathcal{D}^*$ since otherwise, we would have a complementary solution $(x^i; s^i)$ with $x^i, s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}$, implying that $i \in \mathcal{R}$. Therefore, it follows from the complementarity condition that for all $(x; y; s) \in \mathcal{P}^* \times \mathcal{D}^*$, $x^i = 0$ for all $i \in \mathcal{N}$, and $s^i = 0$ for all $i \in \mathcal{B}$, see e.g., [31]. An optimal solution $(x^*; y^*; s^*)$ is maximally complementary if

$$
\begin{array}{ll}
(x^*)^i \in \mathrm{int}(\mathbb{L}_+^{n_i}), & \forall i \in \mathcal{B}, \\
(s^*)^i \in \mathrm{int}(\mathbb{L}_+^{n_i}), & \forall i \in \mathcal{N}, \\
(x^*)^i, (s^*)^i \neq 0, & \forall i \in \mathcal{R},
\end{array}
\tag{6}
$$

see Definition 5 in [5]. Further, a strictly complementary optimal solution $(x^*; y^*; s^*)$ exists iff $\mathcal{T} = \emptyset$.

1.3 Identification of the optimal partition

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

$$
\begin{array}{lll}
Ax & = b, & x \in \mathrm{int}(\mathcal{L}_+^{\bar{n}}), \\
A^T y + s & = c, & s \in \mathrm{int}(\mathcal{L}_+^{\bar{n}}), \\
x \circ s & = \mu e.
\end{array}
\tag{7}
$$

Under the full rank and the interior point conditions, for all $\mu > 0$ system (7) has a unique solution $(x(\mu); y(\mu); s(\mu))$, the so called central solution, where $x(\mu), s(\mu) \in \mathrm{int}(\mathcal{L}_+^{\bar{n}})$. For $\mu > 0$ the set of central solutions forms a smooth analytical curve which converges to a maximally complementary optimal solution, see [24] or Corollary 3.5 in [31].

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 [31] defined two condition numbers, σ_1 and σ_2, as

$$
\begin{aligned}
\sigma_{\mathcal{B}} &:= \min_{i \in \mathcal{B}} \max_{x \in \mathcal{P}^*} \left\{ x_1^i - \left\| x_{2:n_i}^i \right\| \right\}, \\
\sigma_{\mathcal{N}} &:= \min_{i \in \mathcal{N}} \max_{(y;s) \in \mathcal{D}^*} \left\{ s_1^i - \left\| s_{2:n_i}^i \right\| \right\}, \\
\sigma_1 &:= \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}, \\
\sigma_2 &:= \min_{i \in \mathcal{R}} \max_{(x;y;s) \in \mathcal{P}^* \times \mathcal{D}^*} \left\{ x_1^i + s_1^i - \left\| x_{2:n_i}^i + s_{2:n_i}^i \right\| \right\}.
\end{aligned}
\tag{8}
$$

We also define

$$
\sigma_3 := \max_{(x;y;s) \in \mathcal{P}^* \times \mathcal{D}^*} \left\{ \left\| (x; y; s) \right\| \right\}. \tag{9}
$$

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 [31].

Theorem 1 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 [31]. For $\mathcal{T}$ we reproduced the bounds using the error bound result for a linear mixed conic system, as stated in Theorem 7. See Appendix E for the sketch of the proof of the bounds for the $\mathcal{T}$ parts of the following theorem.

Theorem 1 (Theorem 3.4 in [31]). *Let $\big(x(\mu); y(\mu); s(\mu)\big)$ be a central solution such that $\mu \leq \hat{\mu}$, where $\hat{\mu}$ is defined in* (44). *Then there exist a positive scalar γ and condition number κ, as introduced in Theorem 7 in*

Appendix D, so that

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

The results of Theorem 1 can be extended to the case when IPMs generate approximate solutions in a neighborhood of the central path. See Section 4 in [31] for a detailed discussion.

From the bounds of Theorem 1 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

$$\mu < \tilde{\mu} := \min\left\{\frac{\sigma_1^2}{2p^2}, \frac{\sigma_1\sigma_2}{4p^2},\ \frac{1}{p}\left(\frac{1}{4\kappa}\min\left\{\frac{\sigma_1}{2p}, \frac{\sigma_2}{2p}\right\}\right)^{\frac{1}{\gamma}},\ \hat{\mu}\right\}. \tag{10}$$

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 1. Nevertheless, we assume in the convergence analysis of Section 4 that $\mu < \tilde{\mu}$ is small enough for a complete identification of $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$.

1.4 Primal and dual nondegeneracy conditions in SOCO

The primal and dual nondegeneracy conditions for a feasible solution of SOCO are presented in [2, 3] as follows:

Definition 2 (Definitions 17 and 18 in [2])**.** *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*

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

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 relationship between nondegeneracy and uniqueness of the optimal solution is stated in the following theorem.

Theorem 2 (Theorem 22 in [2])**.** *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.*

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 $(\mathrm{D_{SOCO}})$.

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

$$q_1^i := \frac{1}{2}\begin{pmatrix} 1 \\ \frac{-x^i_{2:n_i}}{\|x^i_{2:n_i}\|} \end{pmatrix}, \qquad q_2^i := \frac{1}{2}\begin{pmatrix} 1 \\ \frac{x^i_{2:n_i}}{\|x^i_{2:n_i}\|} \end{pmatrix},$$

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 3 (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*

$$\Big((A_i \bar{Q}_i^*)_{i \in \mathcal{R} \cup \mathcal{T}_2},\ A_{\mathcal{B}} \Big) \tag{11}$$

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

$$\Big((A_i R_i (s^*)^i)_{i \in \mathcal{R} \cup \mathcal{T}_3},\ A_{\mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2} \Big) \tag{12}$$

has full column rank, where R_i is defined in (3).

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 Second-order sufficient condition for SOCO

In this section, we highlight the connection between the second-order sufficient condition of Bonnans and Ramírez [5], see Appendix B, and the nondegeneracy conditions. We show that if the primal nondegeneracy condition holds and $\mathcal{R}$ is nonempty, then the second-order sufficient condition holds at the unique optimal solution of $(\mathrm{D_{SOCO}})$. In Section 4, we use the second-order sufficient condition to show the nonsingularity of the Jacobian of the equations in KKT conditions for a reduced NLO problem.

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

$$\sup\nolimits_{x \in \mathcal{P}^*} h^T H_{\mathrm{D}}(y, x) h > 0, \quad \forall h \in \mathcal{C}_{\mathrm{D}}(y) \setminus \{0\}, \tag{13}$$

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

$$h^T H_{\mathrm{D}}(y, x) h = \sum_{i \in \mathcal{R}} -\frac{x_1^i}{s_1^i} h^T A_i R_i A_i^T h,$$

and $\mathcal{C}_{\mathrm{D}}(y)$ is the cone of critical directions for $(\mathrm{D_{SOCO}})$ as defined in (35). Let

$$A_i^T h =: \big((A_i^T h)_1;\ (A_i^T h)_{2:n_i} \big), \quad \forall i.$$

Then for all $h \in \mathcal{C}_{\mathrm{D}}(y) \setminus \{0\}$ we have

$$\begin{aligned} 0 = (x^i)^T A_i^T h &= x_1^i (A_i^T h)_1 + (x^i_{2:n_i})^T (A_i^T h)_{2:n_i} \\ &\geq x_1^i (A_i^T h)_1 - \|x^i_{2:n_i}\| \|(A_i^T h)_{2:n_i}\| \\ &= x_1^i \Big((A_i^T h)_1 - \|(A_i^T h)_{2:n_i}\| \Big), \qquad \forall i \in \mathcal{R}, \end{aligned}$$

where the last equality follows from $x_1^i = \|x^i_{2:n_i}\|$. Since $x_1^i > 0$, we can conclude that

$$(A_i^T h)_1 - \|(A_i^T h)_{2:n_i}\| \leq 0, \qquad \forall i \in \mathcal{R}.$$

Analogously, we can derive

$$0 = (x^i)^T A_i^T h = x_1^i (A_i^T h)_1 + (x_{2:n_i}^i)^T (A_i^T h)_{2:n_i} \le x_1^i \Big((A_i^T h)_1 + \big\|(A_i^T h)_{2:n_i}\big\| \Big),$$

which implies

$$(A_i^T h)_1 + \big\|(A_i^T h)_{2:n_i}\big\| \ge 0, \qquad \forall i \in \mathcal{R}.$$

Consequently,

$$h^T A_i R_i A_i^T h = \Big((A_i^T h)_1 - \big\|(A_i^T h)_{2:n_i}\big\| \Big) \Big((A_i^T h)_1 + \big\|(A_i^T h)_{2:n_i}\big\| \Big) \le 0, \tag{14}$$

which implies

$$\sup_{x \in \mathcal{P}^*} h^T H_{\mathrm{D}}(y,x) h \ge 0, \quad \forall h \in \mathcal{C}_{\mathrm{D}}(y) \setminus \{0\}.$$

In a similar manner, we can show that under the dual nondegeneracy condition, the second-order sufficient condition holds at the unique optimal solution of $(\mathrm{P_{SOCO}})$. We redefine $h \in \mathbb{R}^{\bar{n}}$ to refer to a critical direction belonging to $\mathcal{C}_{\mathrm{P}}(x)$ as defined in (36) and let $(x;y;s) \in \mathcal{P}^* \times \mathcal{D}^*$. Assume that $\mathcal{R} \neq \emptyset$ with $x^i, s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}$ for some $i \in \mathcal{R}$. Then the second-order sufficient condition for $(\mathrm{P_{SOCO}})$ is given by

$$\sup_{(y;s) \in \mathcal{D}^*} h^T H_{\mathrm{P}}(x,s) h > 0, \quad \forall h \in \mathcal{C}_{\mathrm{P}}(x) \setminus \{0\}, \tag{15}$$

where, by $s^i = 0$ for $i \in \mathcal{T}_2$, we have

$$h^T H_{\mathrm{P}}(x,s) h = \sum_{i \in \mathcal{R}} -\frac{s_1^i}{x_1^i} \big((h_1^i)^2 - \|h_{2:n_i}^i\|^2 \big).$$

Then for all $h \in \mathcal{C}_{\mathrm{P}}(x) \setminus \{0\}$ we have

$$s_1^i \Big(h_1^i - \|h_{2:n_i}^i\| \Big) \le 0 = (s^i)^T h^i \le s_1^i \Big(h_1^i + \|h_{2:n_i}^i\| \Big), \qquad \forall i \in \mathcal{R},$$

which follows from $s_1^i = \|s_{2:n_i}^i\|$. Since $s_1^i > 0$, we have

$$(h_1^i)^2 - \big\|h_{2:n_i}^i\big\|^2 \le 0, \qquad \forall i \in \mathcal{R}. \tag{16}$$

The next lemma shows that under the primal nondegeneracy condition we have $h^T A_i R_i A_i^T h < 0$ for some $i \in \mathcal{R}$, and under the dual nondegeneracy condition (16) holds with strict inequality for some $i \in \mathcal{R}$. The connection between the dual nondegeneracy condition and a strong second-order sufficient condition was established for SDO, see Proposition 15 in [7]. The proof of the following lemma has been given in [14] for the more general case of symmetric cone optimization problems. For the sake of completeness, we provide an illustrative proof for SOCO in Appendix B. We refer to this proof in Lemma 6.

Lemma 1. *Let $(x^*;y^*;s^*)$ be the unique optimal solution of $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$ so that $\mathcal{R} \neq \emptyset$. Then,*

1. *Under the primal nondegeneracy condition the second-order sufficient condition (13) holds at $(y^*;s^*)$.*
2. *Under the dual nondegeneracy condition the second-order sufficient condition (15) holds at x^*.*

3 Quadratic convergence to a strictly complementary optimal solution

Assume that the primal and dual nondegeneracy conditions hold, and the unique primal-dual optimal solution $(x^*;y^*;s^*)$ is strictly complementary. We show that from a solution, which is sufficiently close to the optimal set, Newton's method is quadratically convergent to the strictly complementary optimal solution $(x^*;y^*;s^*)$.

We can rewrite the optimality conditions (4) 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 defined as

$$F((x;y;s)) := \begin{pmatrix} Ax - b \\ A^T y + s - c \\ x \circ s \end{pmatrix}. \tag{17}$$

The Jacobian of F is given by

$$\nabla F((x;y;s)) := \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ L(s) & 0 & L(x) \end{pmatrix}, \tag{18}$$

where $L(x)$ and $L(s)$ are given by

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

The next lemma shows the Lipschitz continuity of ∇F.

Lemma 2. *The Jacobian ∇F is Lipschitz continuous with global Lipschitz constant $\tau_1 := 2$.*

Proof. Let $\xi := (\xi^1;\xi^2;\xi^3) \in \mathbb{R}^{\bar{n}} \times \mathbb{R}^m \times \mathbb{R}^{\bar{n}}$. Then from (2) and (18) we have

$$\begin{aligned}\|\nabla F((x;y;s)) - \nabla F((x';y';s'))\| &= \max_{\|\xi\|=1} \|\big(\nabla F((x;y;s)) - \nabla F((x';y';s'))\big)\xi\| \\ &= \max_{\|\xi\|=1} \|L(s-s')\xi^1 + L(x-x')\xi^3\| \\ &\le \max_{\|\xi\|=1} \|L(s-s')\xi^1\| + \max_{\|\xi\|=1} \|L(x-x')\xi^3\| \\ &\le \max_{\|\xi^1\|=1} \|L(s-s')\xi^1\| + \max_{\|\xi^3\|=1} \|L(x-x')\xi^3\| \\ &\le \|L(s-s')\| + \|L(x-x')\|,\end{aligned}$$

where $L(s-s')$ and $L(x-x')$ are block diagonal symmetric matrices. Then from Theorem 3 in [2] and the definition of the spectral norm we get

$$\begin{aligned}\|L(s-s')\| &= \max_{i=1,\ldots,p} \max_{j=1,\ldots,n} |\lambda_j(L(s^i - (s')^i))| \\ &\le \max_{i=1,\ldots,p} \left|s_1^i - (s')_1^i\right| + \|s_{2:n_i}^i - (s')_{2:n_i}^i\| \\ &\le \max_{i=1,\ldots,p} \sqrt{2}\|s^i - (s')^i\| \le \sqrt{2}\|s - s'\|.\end{aligned}$$

The case for $\|L(x-x')\|$ is similar. Consequently, we get

$$\begin{aligned}\|\nabla F((x;y;s)) - \nabla F((x';y';s'))\| &\le \sqrt{2}\|s-s'\| + \sqrt{2}\|x-x'\| \\ &\le 2\|(x-x';y-y';s-s')\|,\end{aligned}$$

which completes the proof. □

In order to apply Newton's method, ∇F must be nonsingular in a neighborhood of $(x^*;y^*;s^*)$. The following lemma is in order.

Lemma 3 (Theorem 28 in [2])**.** *If strict complementarity and the primal and dual nondegeneracy conditions hold, then $\nabla F((x^*;y^*;s^*))$ is nonsingular.*

Notice that without the strict complementarity condition, ∇F might be singular at an optimal solution, and thus Newton's method is not applicable. In Section 4, we release the dependence on the strict complementarity condition by considering the optimal partition of SOCO.

Since the Jacobian is Lipschitz continuous and nonsingular at the optimal solution $(x^*;y^*;s^*)$, it can be proven that $\nabla F((x;y;s))$ is nonsingular for all $(x;y;s)$ in a neighborhood of $(x^*;y^*;s^*)$, see the argument after Theorem 6 in Appendix C. Then, if the initial point is sufficiently close to the optimal set, Newton's method applied to F converges quadratically to $(x^*;y^*;s^*)$, see [23] for a similar result for SDO. The application of Newton's method to F is similar to its application in Section 4, and thus it is not presented here.

Theorem 4. *Assume that there exists $\beta_1 > 0$ so that*

$$\|\nabla F((x^*;y^*;s^*))^{-1}\| \le \beta_1.$$

Let $\hat{\mu}$ defined by (44) and a central solution $(x(\mu);y(\mu);s(\mu))$ with

$$\mu < \min\left\{p^{-1}\big(4\sqrt{3}\beta_1\kappa\big)^{-\frac{1}{\gamma}},\ \hat{\mu}\right\} \tag{19}$$

be given. Then starting from $(x(\mu); y(\mu); s(\mu))$ Newton's method is quadratically convergent to $(x^*; y^*; s^*)$.

Proof. Since F is continuously differentiable in $\mathbb{R}^n$ and Lipschitz continuous, the result of Theorem 6 is valid. Hence, Newton steps are well-defined in a neighborhood of $(x^*; y^*; s^*)$. Additionally, from the error bounds given in (45) we have

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

Then it is immediate from (40) that $(x(\mu); y(\mu); s(\mu))$ is in the convergence region of Newton's method if

$$\sqrt{3}\kappa(p\mu)^{\gamma} < \frac{1}{4\beta_1},$$

which yields the result. □

4 Quadratic convergence to the unique maximally complementary optimal solution

In this section, we release the assumption of strict complementarity, but we assume that the primal and dual nondegeneracy conditions hold. We aim to establish quadratic convergence of Newton's method to the unique maximally complementary optimal solution from an initial solution sufficiently close to the optimal set. 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 4. *Assume that the primal and dual nondegeneracy conditions hold, and let $(x^*; y^*; s^*)$ be the unique optimal solution. Then $\mathcal{R} = \emptyset$ implies $\mathcal{T} = \emptyset$.*

Proof. We refer to conditions (11) and (12) for the primal and dual nondegeneracy conditions, and consider all the possible cases of the optimal partition when $\mathcal{R} = \emptyset$. Note that both matrices in (11) and (12) have m rows, but they 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, i.e., the number of columns in (12) is strictly greater than the number of columns in (11) for every possible case in which $\mathcal{T}_1$, $\mathcal{T}_2$, or $\mathcal{T}_3$ is nonempty[1]. Thus, when $\mathcal{T} \neq \emptyset$, we cannot have simultaneously a full row rank matrix in (11), and have a full column rank matrix in (12). This completes the proof. □

As a result of Lemma 4, 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 $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$ 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 $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$ 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 $(\mathrm{D_{SOCO}})$ as

$$(\mathrm{D'_{SOCO}}) \qquad \max\Big\{ b^T y \;:\; A_i^T y + s^i = c^i,\; s^i \in \mathbb{L}_+^{n_i},\; i \in \{1,\dots,p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\} \Big\},$$

and its dual is written as

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

Since $(x^*)^i = 0$ for $i \in \mathcal{T}_1 \cup \mathcal{T}_3$, it follows from the optimality conditions, (11) and (12) that $((x^*)^i; y^*; (s^*)^i)$ for $i \in \{1,\dots,p\} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_3\}$ is a primal-dual optimal solution for $(\mathrm{P'_{SOCO}})$ and $(\mathrm{D'_{SOCO}})$, 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

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

[1] For each $i \in \mathcal{T}_2$, $A_i \bar{Q}_i^*$ is an $m \times n_i - 1$ matrix, where A_i is an $m \times n_i$ matrix. Hence, there is no way to have $A_i \bar{Q}_i^*$ and A_i full row rank and full column rank, respectively, at the same time.

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 $(\mathrm{P}_{\mathrm{SOCO}})$ and $(\mathrm{D}_{\mathrm{SOCO}})$ by solving $(\mathrm{P}'_{\mathrm{SOCO}})$ and $(\mathrm{D}'_{\mathrm{SOCO}})$. At the risk of causing confusion, we refer to $(\bar{x}; \bar{y}; \bar{s})$ as the unique optimal solution of $(\mathrm{P}'_{\mathrm{SOCO}})$ and $(\mathrm{D}'_{\mathrm{SOCO}})$.

The algebraic definition (1) can be used to reformulate $(\mathrm{D}'_{\mathrm{SOCO}})$ as a nonconvex NLO problem. Then inspired by the optimal partition information and the characteristics of a maximally complementary optimal solution specified by (6), one can realize that the unique dual optimal solution $(\bar{y}; \bar{s})$ can be obtained by solving the NLO reformulation of $(\mathrm{D}'_{\mathrm{SOCO}})$ as

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

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

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

Let z denote the concatenation of the column vectors z^i for $i \in \mathcal{R} \cup \mathcal{N}$. It then follows that $(\mathrm{D}_{\mathrm{NLO}})$ 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

$$\bar{w} := \bar{y}, \qquad \bar{z}^i := \bar{s}^i, \quad i \in \mathcal{R} \cup \mathcal{N}. \tag{20}$$

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

$$\begin{array}{lll} (\mathrm{P}_{\mathrm{NLO}}) & \min & \displaystyle\sum_{i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2} (c^i)^T \nu^i \\ & \text{s.t.} & \displaystyle\sum_{i \in \mathcal{B} \cup \mathcal{R} \cup \mathcal{T}_2} A_i \nu^i = b, \\ & & (\nu^i)^T R_i \nu^i = 0, \qquad i \in \mathcal{R} \cup \mathcal{T}_2, \\ & & \nu \in \mathcal{V}, \end{array}$$

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

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

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

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 $(\mathrm{D}_{\mathrm{NLO}})$. The first-order optimality conditions[2] for $(\mathrm{D}_{\mathrm{NLO}})$, see Appendix A, are given by

$$\begin{cases} -\sum_{i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}} A_i u^i & = b, \\ -u^i - 2v_i R_i z^i & = 0, \quad i \in \mathcal{R}, \\ -u^i & = 0, \quad i \in \mathcal{N}, \\ A_i^T w & = c^i, \quad i \in \mathcal{B} \cup \mathcal{T}_2, \\ A_i^T w + z^i & = c^i, \quad i \in \mathcal{R} \cup \mathcal{N}, \\ (z^i)^T R_i z^i & = 0, \quad i \in \mathcal{R}, \\ z \in \mathcal{W}, \end{cases} \tag{21}$$

which bears a striking resemblance to the optimality conditions (4). 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 (21). Such a solution can be obtained by setting

$$\begin{aligned} \bar{u}^i &:= -\bar{x}^i, \qquad i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R}, \\ \bar{u}^i &:= 0, \qquad i \in \mathcal{N}, \\ \bar{v}_i &:= \frac{1}{2} \frac{\bar{x}_1^i}{\bar{s}_1^i}, \qquad i \in \mathcal{R}. \end{aligned} \tag{22}$$

[2]Here, we use a version of the first-order optimality conditions where the constraint set is defined on an open set.

We show in Lemma 5 that, under the dual nondegeneracy condition, the Lagrange multipliers $(\bar{u};\bar{v})$ are unique. Furthermore, we prove in Lemma 6 that, under the primal nondegeneracy condition, the second-order sufficient condition (24) holds at $(\bar{w};\bar{z})$.

Let $J((w;z))$ denote the Jacobian of the equality constraints in $(\mathrm{D_{NLO}})$ as follows

$$J((w;z)) := \begin{pmatrix} A_{\mathcal{B}}^T & 0 & 0 \\ A_{\mathcal{T}_2}^T & 0 & 0 \\ A_{\mathcal{R}}^T & I & 0 \\ A_{\mathcal{N}}^T & 0 & I \\ 0 & Z_{\mathcal{R}} & 0 \end{pmatrix}, \tag{23}$$

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 5. *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, under the dual nondegeneracy condition, $J((\bar{w};\bar{z}))^T\eta = 0$ 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

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

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 $(\mathrm{D'_{SOCO}})$, it follows from (12) 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 5, the linear independence constraint qualification (LICQ) [27] 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

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

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

$$V_{\mathcal{R}} := -2\mathrm{diag}(v_1R_1, v_2R_2, \ldots, v_iR_i, \ldots)$$

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 6, 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.,

$$h^T \nabla^2 L(\bar{\vartheta}) h > 0, \quad \forall h \in \mathrm{Ker}(J((\bar{w}; \bar{z}))) \setminus \{0\}, \tag{24}$$

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

Lemma 6. *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* (24) *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

$$\begin{cases} A_i^T h^1 = 0, & i \in \mathcal{B} \cup \mathcal{T}_2, \\ A_i^T h^1 + (h^2)^i = 0, & i \in \mathcal{R}, \\ A_i^T h^1 + (h^3)^i = 0, & i \in \mathcal{N}, \\ (\bar{z}^i)^T R_i (h^2)^i = 0, & i \in \mathcal{R}. \end{cases} \tag{25}$$

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 proof of Lemma 1, see Appendix B, for the unique primal optimal solution $\bar{x}$ system (38) has only a trivial solution. Analogous to the proof of Lemma 1, we should 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 (11). Hence, it follows from (22) and (25) 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 (25). □

In Example 1, we investigate the second-order sufficient condition (13) and the nondegeneracy conditions for a simple SOCO problem.

4.1 Quadratic convergence of Newton's method

We apply Newton's method to the first-order optimality conditions of ($\mathrm{D_{NLO}}$). The idea is to start from a central solution, for which μ satisfies (10), and take Newton steps to converge to $\bar{\vartheta}$. The first-order optimality conditions (21) 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

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

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 (21) 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

$$\vartheta^{(k+1)} := \vartheta^{(k)} + d\vartheta^{(k)}, \qquad d\vartheta_k := \left(dw^{(k)}; dz^{(k)}; du^{(k)}; dv^{(k)}\right), \tag{26}$$

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

$$\nabla G\left(\vartheta^{(k)}\right) d\vartheta^{(k)} = -G(\vartheta^{(k)}). \tag{27}$$

Lemma 5 shows that $J((\bar{w}; \bar{z}))$ is of full row rank, and by Lemma 6 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 7. *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. We consider the linear system of equations $\nabla G(\bar{\vartheta}) \eta =$

0 and show that $\eta = 0$ is the only solution. To that end, we have

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

From the first equation we have $(\eta^1)^T\nabla^2 L(\bar{\vartheta})\eta^1 = 0$, which implies $\eta^1 = 0$ by Lemma 6. Setting $\eta^1 = 0$, the first equation gives $J((\bar{w};\bar{z}))^T\eta^2$, which implies $\eta^2 = 0$ by Lemma 5. □

Now, we show the Lipschitz continuity of ∇G.

Lemma 8. *The Jacobian ∇G is Lipschitz continuous with global Lipschitz constant $\tau_2 := 2\sqrt{2}$.*

Proof. Consider two arbitrary solutions ϑ and ϑ'. Analogous to Lemma 2, let $\xi := (\xi^1;\ldots;\xi^8)$ be a vector of appropriate size. Then we have

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

Then from the properties of the spectral norm we get

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

Therefore, we get

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

which completes the proof. □

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

Lemma 9. *Let $\big(x(\mu);y(\mu);s(\mu)\big)$ be a central solution with $\mu \le \hat{\mu}$, where $\hat{\mu}$ is defined by (44), $(\bar{x};\bar{y};\bar{s})$ 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*

$$\sqrt{\sum_{i\in\mathcal{R}} \left(\frac{x^i_1(\mu)}{s^i_1(\mu)} - \frac{\bar{x}^i_1}{\bar{s}^i_1}\right)^2} \le \frac{4p\sqrt{|\mathcal{R}|}\kappa(p\mu)^\gamma}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right). \tag{28}$$

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

$$\frac{\bar{x}^i_{2:n_i}}{\|\bar{x}^i_{2:n_i}\|} = -\frac{\bar{s}^i_{2:n_i}}{\|\bar{s}^i_{2:n_i}\|}. \tag{29}$$

Since $\bar{x}^i = (x^*)^i$ and $\bar{s}^i = (s^*)^i$ for $i \in \mathcal{R}$, it follows from (8) and (29) that for every $i \in \mathcal{R}$

$$\begin{aligned}\sigma_2 \le \bar{x}^i_1 + \bar{s}^i_1 - \left\|\bar{x}^i_{2:n_i} + \bar{s}^i_{2:n_i}\right\| &= \bar{x}^i_1 + \bar{s}^i_1 - |\bar{x}^i_1 - \bar{s}^i_1| \\ &= 2\min\{\bar{x}^i_1, \bar{s}^i_1\}.\end{aligned} \tag{30}$$

Furthermore, it holds that

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

where the last inequality follows from (45). Now using (9), (30), and Theorem 1 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. ◻

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

$$\begin{aligned} w^{(0)} &:= y(\mu), && \\ (z^i)^{(0)} &:= s^i(\mu), && i \in \mathcal{R} \cup \mathcal{N}, \\ (u^i)^{(0)} &:= -x^i(\mu), && i \in \mathcal{B} \cup \mathcal{T}_2 \cup \mathcal{R} \cup \mathcal{N}, \\ v_i^{(0)} &:= \frac{1}{2}\frac{x_1^i(\mu)}{s_1^i(\mu)}, && i \in \mathcal{R}. \end{aligned} \tag{31}$$

Then a search direction is computed by using (27), and the new iterate is obtained by (26). The next theorem shows that if μ is sufficiently small, then Newton's method converges quadratically to the unique optimal solution $(\bar{x}; \bar{y}; \bar{s})$.

Theorem 5. *Assume that the primal and dual nondegeneracy conditions hold. Let*

$$\mu < \min\left\{p^{-1}\left(4\sqrt{2}\beta_2\kappa\left(\sqrt{3} + \frac{2p\sqrt{|\mathcal{R}|}}{\sigma_2}\left(1 + \frac{2\sigma_3}{\sigma_2}\right)\right)\right)^{-\frac{1}{\gamma}},\ \tilde{\mu}\right\}, \tag{32}$$

in which β_2 denotes an upper bound for $\left\|\nabla G(\bar{\vartheta})^{-1}\right\|$, and $\tilde{\mu}$ is defined in (10)*. Then, initiated as given in* (31)*, 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 7 and 8, the conditions of Theorem 6 hold, and we get

$$\epsilon := \frac{1}{4\sqrt{2}\beta_2}.$$

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 (20) and (22). Using the bounds in Theorem 1 and (28) we get

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

Then, considering the error bounds given in (45), we obtain

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

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

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

or equivalently,

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

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 problem $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$ by appending the $\mathcal{T}_1$ and $\mathcal{T}_3$ parts 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. *Bound* (32)*, relying on the condition numbers σ_2, σ_3, and the exponent γ, is significantly more complicated than* (19)*. In fact, the intricacy of bound* (32) *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.*

5 Special case: a strongly polynomial rounding procedure

In a special case when the sets $\mathcal{R}$ and $\mathcal{T}$ are empty, a strictly complementary optimal solution can be obtained as easily as in LO [18, 34], regardless of the nondegeneracy conditions. More precisely, a central solution $(x(\mu); y(\mu); s(\mu))$, with sufficiently small μ, can be rounded to an exact strictly complementary optimal solution in strongly polynomial time through solving two least squares problems.

Let $(x^*; y^*; s^*) \in \mathrm{ri}(\mathcal{P}^* \times \mathcal{D}^*)$ be a maximally complementary optimal solution of $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$. Then the primal-dual feasibility constraints imply

$$\sum_{i\in\mathcal{B}} A_i (x^*)^i = b, \qquad \sum_{i\in\mathcal{B}\cup\mathcal{N}} A_i x^i(\mu) = b,$$

and

$$\begin{array}{llll} A_i^T y^* &= c^i, & A_i^T y(\mu) + s^i(\mu) &= c^i, \quad i \in \mathcal{B}, \\ A_i^T y^* + (s^*)^i &= c^i, & A_i^T y(\mu) + s^i(\mu) &= c^i, \quad i \in \mathcal{N}. \end{array}$$

Subtracting the right hand side equations from the left hand side ones we get

$$\begin{aligned} \sum_{i\in\mathcal{B}} A_i \Delta x^i(\mu) &= \sum_{i\in\mathcal{N}} A_i x^i(\mu), \\ A_i^T \Delta y(\mu) &= s^i(\mu), \quad i \in \mathcal{B}, \\ A_i^T \Delta y(\mu) + \Delta s^i(\mu) &= 0, \qquad i \in \mathcal{N}, \end{aligned}$$

where $\Delta y(\mu) := y^* - y(\mu)$, $\Delta x^i(\mu) := (x^*)^i - x^i(\mu)$ for $i \in \mathcal{B}$, and $\Delta s^i(\mu) := (s^*)^i - s^i(\mu)$ for $i \in \mathcal{N}$. Thus, we get a primal-dual solution with zero complementarity gap by solving the least squares problem

$$\begin{aligned} \min \ & \frac{1}{2}\sum_{i\in\mathcal{B}} \|\Delta x^i\|^2 \\ \text{s.t.} \ & \sum_{i\in\mathcal{B}} A_i \Delta x^i = \sum_{i\in\mathcal{N}} A_i x^i(\mu), \end{aligned} \tag{33}$$

for the primal solution, and the least squares problem

$$\begin{aligned} \min \ & \frac{1}{2}\sum_{i\in\mathcal{N}} \|\Delta s^i\|^2 + \frac{1}{2}\|\Delta y\|^2 \\ \text{s.t.} \ & A_i^T \Delta y = s^i(\mu), \qquad i \in \mathcal{B}, \\ & A_i^T \Delta y + \Delta s^i = 0, \quad i \in \mathcal{N}, \end{aligned}$$

for the dual solution. The last one is equivalent to

$$\begin{aligned} \min \ & \frac{1}{2}\sum_{i\in\mathcal{N}} \|A_i^T \Delta y\|^2 + \frac{1}{2}\|\Delta y\|^2 \\ \text{s.t.} \ & A_i^T \Delta y = s^i(\mu), \quad i \in \mathcal{B}. \end{aligned} \tag{34}$$

The solutions of least squares problems (33) and (34) give the complementary primal-dual pair $(\tilde{x}, \tilde{y}, \tilde{s})$, where

$$\begin{aligned} \tilde{x}^i &:= x^i(\mu) + (\Delta x^*)^i, \quad i \in \mathcal{B}, \\ \tilde{y} &:= y(\mu) + \Delta y^*, \\ \tilde{s}^i &:= s^i(\mu) + (\Delta s^*)^i, \quad i \in \mathcal{N}. \end{aligned}$$

It can be easily shown that $(\tilde{x}, \tilde{y}, \tilde{s})$ is feasible with respect to the primal and dual affine constraints. Further, the feasibility of $(\tilde{x}, \tilde{y}, \tilde{s})$ with respect to the second-order cones can be established when μ is sufficiently small, see e.g., [18].

6 Conclusions and discussions

Using the optimal partition of a SOCO problem, under the primal and dual nondegeneracy conditions, we established quadratic convergence of Newton's method to the unique maximally complementary optimal solution of ($\mathrm{P_{SOCO}}$) and ($\mathrm{D_{SOCO}}$). We showed that if the primal and dual nondegeneracy conditions hold, then $\nabla G(\bar{\vartheta})$ is nonsingular. In contrast to the application of Newton's method to (17), quadratic convergence is not dependent on the strict complementarity condition. For a special case where $\mathcal{R}, \mathcal{T} = \emptyset$, we presented a rounding procedure which yields an exact strictly complementary optimal solution in a strongly polynomial time.

The optimal partition approach in Section 4 can be directly applied to the optimality conditions for ($\mathrm{P'_{SOCO}}$) and ($\mathrm{D'_{SOCO}}$). 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 (17), 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 for quadratic convergence of Newton's method.

To establish quadratic convergence of Newton's method for SOCO, we assumed that the optimal partition is known, and that $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$ can be identified from $\mathcal{T}$. For SDO, it still remains an open question if we can establish quadratic convergence to the unique optimal solution using the optimal partition of SDO.

7 Acknowledgements

This work is supported by 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. Technical Report RRR-111-95, RUTCOR, Rutgers University, USA, 1995.

[2] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming*, 95(1):3–51, 2003.

[3] F. Alizadeh and S. Schmieta. Optimization with semidefinite, quadratic and linear constraints. Technical Report RRR 23-97, RUTCOR, Rutgers University, USA, 1997.

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

[5] J. F. Bonnans and H. Ramírez C. Perturbation analysis of second-order cone programming problems. *Mathematical Programming*, 104(2):205–227, 2005.

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

[7] Z. X. Chan and D. Sun. Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming. *SIAM Journal on Optimization*, 19(1):370–396, 2008.

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

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

[10] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programming. *SIAM Journal on Optimization*, 8(4):871–886, 1998.

[11] A. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards*, 49(4):263–265, 1952.

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

[13] M. Kojima, S. Mizuno, and A. Yoshise. A polynomial-time algorithm for a class of linear complementarity problems. *Mathematical Programming*, 44(1):1–26, 1989.

[14] L. Kong, L. Tunçel, and N. Xiu. Equivalent conditions for Jacobian nonsingularity inlinear symmetric cone programming. *Journal of Optimization Theory and Applications*, 148(2):364–389, Feb 2011.

[15] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284(1):193 – 228, 1998.

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

[17] S. Mehrotra. Finite termination and superlinear convergence in primal-dual methods. Technical Report TR91-13, Department of Industrial Engineering and Management Sciences, Northwestern University, 1991.

[18] S. Mehrotra and Y. Ye. Finding an interior point in the optimal face of linear programs. *Mathematical Programming*, 62(1):497–515, 1993.

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

[20] A. Mohammad-Nezhad and T. Terlaky. A polynomial primal-dual affine scaling algorithm for symmetric conic optimization. *Computational Optimization and Applications*, 66(3):577–600, Apr 2017.

[21] 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(1):61–83, Jun 2000.

[22] R. D. Monteiro and S. J. Wright. Local convergence of interior-point algorithms for degenerate monotone LCP. *Computational Optimization and Applications*, 3(2):131–155, 1994.

[23] M. V. Nayakkankuppam and M. L. Overton. Conditioning of semidefinite programs. *Mathematical Programming*, 85(3):525–540, Aug 1999.

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

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

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

[27] J. Nocedal and S. Wright. *Numerical Optimization.* Springer, second edition, 2006.

[28] C. Roos, T. Terlaky, and J.-P. Vial. *Interior Point Methods for Linear Optimization.* Springer, 2005.

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

[30] J. F. Sturm. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*, 10(4):1228–1248, 2000.

[31] T. Terlaky and Z. Wang. On the identification of the optimal partition of second order cone optimization problems. *SIAM Journal on Optimization*, 24(1):385–414, 2014.

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

[33] T. Tsuchiya. A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. *Optimization Methods and Software*, 11(1-4):141–182, 1999.

[34] Y. Ye. On the finite convergence of interior-point algorithms for linear programming. *Mathematical Programming*, 57(1):325–335, 1992.

[35] E. Yildirim. Unifying optimal partition approach to sensitivity analysis in conic optimization. *Journal of Optimization Theory and Applications*, 122(2):405–423, 2004.

Appendices

A First-order optimality conditions for classical NLO problems

An NLO problem is defined as

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \geq 0, \quad i = 1, \dots, m_1, \\ & g_i(x) = 0, \quad i = m_1 + 1, \dots, m_1 + m_2, \\ & x \in X, \end{array}$$

where $X \in \mathbb{R}^n$ is a nonempty open set, and $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \dots, m_1 + m_2$ are continuously differentiable functions in a neighborhood of $x^* \in X$. Then the first-order optimality conditions at x^* are given by

$$\begin{array}{ll} \multicolumn{2}{l}{\nabla f(x^*) - \sum_{i=1}^{m_1} u_i^* \nabla g_i(x^*) - \sum_{i=m_1+1}^{m_1+m_2} u_i^* \nabla g_i(x^*) = 0,} \\ g_i(x^*) \geq 0, & i = 1, \dots, m_1, \\ g_i(x^*) = 0, & i = m_1 + 1, \dots, m_1 + m_2, \\ u_i^* g_i(x^*) = 0, & i = 1, \dots, m_1, \\ u_i^* \geq 0, & i = 1, \dots, m_1, \\ x^* \in X, & \end{array}$$

where u_i^* for $i = 1, \dots, m_1 + m_2$ denote the Lagrange multipliers.

B Second-order sufficient condition for SOCO

The second-order sufficient condition for a nonlinear SOCO has been studied by Bonnans and Ramírez [5] which relies on the concepts of the tangent cone and the cone of critical directions. We specialize the second-order sufficient condition for ($\mathrm{D}_{\mathrm{SOCO}}$), where the objective is the minimization of $-b^T y$. Let $(x; y; s) \in \mathcal{P}^* \times \mathcal{D}^*$. Then the cone of critical directions $\mathcal{C}_{\mathrm{D}}(y)$ is defined as

$$\begin{cases} h \in \mathbb{R}^m, & \\ -A_i^T h \in \mathbb{L}_+^{n_i}, & x^i, s^i = 0, \\ -A_i^T h \in \left\{d \mid d_{2:n_i}^T s_{2:n_i}^i - d_1 s_1^i \leq 0\right\}, & x^i = 0,\ s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ A_i^T h = 0, & x^i \in \mathrm{int}(\mathbb{L}_+^{n_i}), \\ (x^i)^T A_i^T h = 0, & x^i, s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ -A_i^T h \in \mathbb{R}_+(x_1^i; -x_{2:n_i}^i), & x^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\},\ s^i = 0. \end{cases} \tag{35}$$

Since the objective function is linear in ($\mathrm{D}_{\mathrm{SOCO}}$), the second-order sufficient condition is shortened to

$$\sup_{x \in \mathcal{P}^*} h^T H_{\mathrm{D}}(y, x) h > 0, \qquad \forall h \in \mathcal{C}_{\mathrm{D}}(y) \setminus \{0\},$$

where

$$H_{\mathrm{D}}(y, x) := \sum_{i=1}^{p} H_{\mathrm{D}}^i(y, x),$$

$$H_{\mathrm{D}}^i(y, x) = \begin{cases} -\frac{x_1^i}{s_1^i} A_i R_i A_i^T, & s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ \mathbf{0}_{m \times m}, & \text{otherwise}, \end{cases} \qquad i = 1, \dots, p.$$

It is straightforward to derive the cone of critical directions and the second-order sufficient condition for ($\mathrm{P}_{\mathrm{SOCO}}$). To that end, note that ($\mathrm{P}_{\mathrm{SOCO}}$) can be equivalently written as

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax - b \in \{0\}, \\ & x^i \in \mathbb{L}_+^{n_i}, \qquad i = 1, \dots, p. \end{aligned}$$

Lemma 10. *Let $h := (h^1; \dots; h^p)$, where $h^i \in \mathbb{R}^{n_i}$, and $(x; y; s)$ be a primal-dual optimal solution of* ($\mathrm{P}_{\mathrm{SOCO}}$) *and* ($\mathrm{D}_{\mathrm{SOCO}}$)*. Then the cone of critical directions $\mathcal{C}_{\mathrm{P}}(x)$ is given by*

$$\begin{cases} Ah = 0, & \\ h^i = 0, & s^i \in \mathrm{int}(\mathbb{L}_+^{n_i}), \\ h^i \in \mathbb{L}_+^{n_i}, & x^i, s^i = 0, \\ h^i \in \left\{d \mid d_{2:n_i}^T x_{2:n_i}^i - d_1 x_1^i \leq 0\right\}, & x^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\},\ s^i = 0, \\ h^i \in \mathbb{R}^{n_i}, & x^i \in \mathrm{int}(\mathbb{L}_+^{n_i}), \\ (s^i)^T h^i = 0, & x^i, s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ h^i \in \mathbb{R}_+(s_1^i; -s_{2:n_i}^i), & x^i = 0,\ s^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}. \end{cases} \tag{36}$$

Proof. The proof is straightforward, and it easily follows from the complementarity of x and s and Proposition 3.10 in [6]. □

The second-order sufficient condition at x is given by

$$\sup_{(y;s) \in \mathcal{D}^*} h^T H_{\mathrm{P}}(x, s) h > 0, \qquad \forall h \in \mathcal{C}_{\mathrm{P}}(x) \setminus \{0\},$$

where

$$H_{\mathrm{P}}(x, s) := \sum_{i=1}^{p} H_{\mathrm{P}}^i(x, s),$$

$$H_{\mathrm{P}}^i(x, s) := \begin{cases} -\frac{s_1^i}{x_1^i} \mathrm{diag}(\mathbf{0}, R_i, \mathbf{0}), & x^i \in \mathrm{bd}(\mathbb{L}_+^{n_i}) \setminus \{0\}, \\ \mathbf{0}_{\bar{n} \times \bar{n}}, & \text{otherwise}, \end{cases} \qquad i = 1, \dots, p,$$

in which $\mathrm{diag}(\mathbf{0}, R_i, \mathbf{0})$ is a block diagonal matrix whose i^{th} block is R_i and 0 elsewhere.

Proof of Lemma 1. Recall that $(x^*; y^*; s^*)$ is the unique primal-dual optimal solution of $(\mathrm{P_{SOCO}})$ and $(\mathrm{D_{SOCO}})$. Let $h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}$. Notice from (14) and (35) that $h^T A_i R_i A_i^T h = 0$ holds if and only if either $A_i^T h \in \mathrm{bd}(\mathbb{L}_+^{n_i})$ or $-A_i^T h \in \mathrm{bd}(\mathbb{L}_+^{n_i})$ holds for $i \in \mathcal{R}$. Since $A_i^T h \notin \mathrm{int}(\mathbb{L}_+^{n_i})$ and $-A_i^T h \notin \mathrm{int}(\mathbb{L}_+^{n_i})$, then $h^T A_i R_i A_i^T h < 0$ is equivalent to

$$A_i^T h \notin \mathbb{L}_+^{n_i}, \quad -A_i^T h \notin \mathbb{L}_+^{n_i}, \quad \forall h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}. \tag{37}$$

Hence, we only need to show that (37) is satisfied under the primal nondegeneracy condition. Then from the characterization of the primal nondegeneracy condition as in Theorem 3 we have that

$$\begin{aligned} A_i^T \eta &= 0, \;\; i \in \mathcal{B}, \\ ((x^*)^i)^T A_i^T \eta &= 0, \;\; i \in \mathcal{R} \cup \mathcal{T}_2, \\ (\hat{Q}_i^*)^T A_i^T \eta &= 0, \;\; i \in \mathcal{R} \cup \mathcal{T}_2, \end{aligned} \tag{38}$$

has only a trivial solution $\eta = 0$, where $\hat{Q}_i^*$ is defined in (11), and $\eta \in \mathbb{R}^m$. From (35) we can observe that a critical direction $h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}$ satisfies

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

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 (38), which is in contradiction to the primal nondegeneracy condition. 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 $A_i^T h$ is orthogonal to the linear subspace spanned by $(x^*)^i$ and $R_i(x^*)^i$, and thus $A_i^T h \notin \mathbb{L}_+^{n_i}$ and $-A_i^T h \notin \mathbb{L}_+^{n_i}$ for all $h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\}$.

Analogous to the above case, let $h \in \mathcal{C}_{\mathrm{P}}(x^*) \setminus \{0\}$, and assume that $(h_1^i)^2 - \|h_{2:n_i}^i\|^2 = 0$ for every $i \in \mathcal{R}$, which implies either $h^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})$ or $-h^i \in \mathrm{bd}(\mathbb{L}_+^{n_i})$. Then, from $((s^*)^i)^T h^i = 0$, it follows that

$$h^i \in \mathbb{R}\big((s^*)_1^i; -(s^*)_{2:n_i}^i\big), \qquad i \in \mathcal{R}.$$

Noting that $h^i = 0$ for $i \in \mathcal{N}$ and $Ah = 0$, we get

$$\sum_{i \in \mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2} A_i h^i + \sum_{i \in \mathcal{R} \cup \mathcal{T}_3} A_i h^i = \sum_{i \in \mathcal{B} \cup \mathcal{T}_1 \cup \mathcal{T}_2} A_i h^i + \sum_{i \in \mathcal{R} \cup \mathcal{T}_3} \alpha_i A_i R_i (s^*)^i = 0,$$

where $\alpha_i \geq 0$ for $i \in \mathcal{T}_3$. All this implies that

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

has linearly dependent columns, which contradicts the dual nondegeneracy condition. This completes the proof.

Example 1 ([2])**.** *Consider the following SOCO problem:*

$$\begin{array}{lll} \min & -x_2 & \\ \text{s.t.} & x_1 & = 1, \\ & 2x_2 + x_3 - x_5 & = 0, \\ & 2x_3 - x_6 & = 0, \\ & x_4 & = 2, \\ & x_1 & \geq \sqrt{x_2^2 + x_3^2}, \\ & x_4 & \geq \sqrt{x_5^2 + x_6^2}. \end{array} \tag{39}$$

Both primal problem (39) *and its dual satisfy the interior point condition, and they have the unique primal-dual optimal solution*

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

with optimal objective value -1. The optimal partition is $\mathcal{R} = \{1\}$, $\mathcal{T}_2 = \{2\}$, $\mathcal{B} = \mathcal{N} = \mathcal{T}_1 = \mathcal{T}_3 = \emptyset$. Then *we have*

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

Further, we can derive that

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

Since both matrices are nonsingular, the unique optimal solution is primal and dual nondegenerate. Nevertheless, the optimal solution is not strictly complementary since $\mathcal{T}_2 \neq \emptyset$. In fact, the second-order constraint $x_4 \geq \sqrt{x_5^2 + x_6^2}$ is weakly inactive, i.e., its removal does not affect the optimality of the current solution. For the dual problem, the cone of critical directions is given by

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

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

$$-\frac{(x^*)_1^1}{(s^*)_1^1} A_1 R_1 A_1^T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 5 & 2 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad h^T H_{\mathrm{D}}(y^*, x^*) h = h_2^2 > 0,\ \forall h \in \mathcal{C}_{\mathrm{D}}(y^*) \setminus \{0\},$$

which confirms that the second-order sufficient condition holds.

This problem in the nonlinear format $(\mathrm{P_{NLO}})$ *has four locally optimal solutions given by*

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

where the objective values are 0.2425, 1, -0.2425, and -1, respectively. 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 unchanged, i.e., $\{\nu_{(4)}\}$. Note that the Jacobian matrix (23) *is nonsingular at the unique globally optimal solution $(\bar{w}; \bar{z})$. Therefore, the second-order sufficient condition* (24) *trivially holds at $(\bar{w}; \bar{z})$.*

C Local convergence of Newton's method

Consider a mapping $G : \mathbb{R}^n \to \mathbb{R}^n$ which is continuously differentiable in 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 a neighborhood $B(x^*, r) \subseteq \mathcal{C}$ for some $r > 0$ exists. The following theorem provides sufficient conditions for the quadratic convergence of Newton's method to x^* provided that the initial solution is sufficiently close to x^*.

Theorem 6 (Theorem 5.2.1 in [9])**.** *Assume that ∇G is Lipschitz continuous with constant τ on $B(x^*, r)$ and that $\|\nabla G(x^*)^{-1}\| \leq \beta$ for some $\beta > 0$. Then for a given $x^{(0)} \in B(x^*, \epsilon)$, where*

$$\epsilon := \min\left\{r, \frac{1}{2\beta\tau}\right\}, \tag{40}$$

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

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

Note that ϵ determines the region around x^* where the Newton directions are well-defined. In fact, from the Lipschitz continuity of ∇G we have

$$\begin{aligned}\big\|\nabla G(x^*)^{-1}(\nabla G(x^{(0)}) - \nabla G(x^*))\big\| \le& \big\|\nabla G(x^*)^{-1}\big\|\big\|\nabla G(x^{(0)}) - \nabla G(x^*)\big\| \\ \le& \beta\tau\|x^{(0)} - x^*\| \le \frac{1}{2}.\end{aligned}$$

Then it turns out, see e.g., Theorem 3.1.4 in [9], that

$$\big\|\nabla G(x^{(0)})^{-1}\big\| \le \frac{\big\|\nabla G(x^*)^{-1}\big\|}{1 - \big\|\nabla G(x^*)^{-1}(\nabla G(x^{(0)}) - \nabla G(x^*))\big\|} \le 2\beta.$$

The nonsingularity of $\nabla G(x^{(k)})$ for $k \ge 1$ follows by noting that $\|x^{(1)} - x^*\| \le \frac{1}{2}\|x^{(0)} - x^*\|$.

D Error bound for a linear mixed conic system

Let $\mathcal{A}$ be a linear subspace of $\mathbb{R}^{\bar{n}}$, $\bar{b} \in \mathbb{R}^{\bar{n}}$, and $\mathcal{K} \subset \mathbb{R}^{\bar{n}}$ be a Cartesian product of p second-order and q positive semidefinite cones as

$$\mathcal{K} := \mathbb{L}_+^{n_1} \times \ldots \times \mathbb{L}_+^{n_p} \times \mathbb{S}_+^{n_{p+1}} \times \ldots \times \mathbb{S}_+^{n_{p+q}},$$

where $\mathbb{S}_+^{n_i}$ denotes the cone of vectorized positive semidefinite matrices of size n_i and $\bar{n} = \sum_{i=1}^{p} n_i + \sum_{i=p+1}^{p+q} \frac{n_i(n_i+1)}{2}$. Then a linear mixed conic system is defined as

$$\begin{cases} x \in \bar{b} + \mathcal{A}, \\ x \in \mathcal{K}. \end{cases} \tag{41}$$

In a given solution x the amount of constraint violation is defined as

$$\operatorname{dist}\big(x,\ \bar{b} + \mathcal{A}\big) + [-\lambda_{\min}(x)]_+,$$

where $[-\lambda_{\min}(x)]_+ := \max\{-\lambda_{\min}(x), 0\}$ and $\operatorname{dist}(.,.)$ denotes a distance function with respect to a norm. If $x^i \in \mathbb{L}_+^{n_i}$, then $\lambda_{\min}(x^i) := x_1^i - \|x_{2:n_i}^i\|$. The following theorem provides an upper bound for the distance between x and the set of solutions of (41).

Theorem 7 (Theorem 7.4.2 in [16])**.** *Let $\bar{\mathcal{A}}$ be the minimal linear subspace which contains $\bar{b} + \mathcal{A}$, i.e.,*

$$\bar{\mathcal{A}} := \{x \mid x + t\bar{b} \in \mathcal{A} \quad \text{for some } t \in \mathbb{R}\}.$$

Consider a bounded set of solutions $x(\zeta)$ with $0 < \zeta \le 1$ so that

$$\operatorname{dist}\big(x(\zeta),\ \bar{\mathcal{A}}\big) \le \zeta, \quad \lambda_{\min}\big(x(\zeta)\big) \ge -\zeta, \quad \forall\ 0 < \zeta \le 1. \tag{42}$$

Then we have

$$\operatorname{dist}\big(x(\zeta),\ (\bar{b} + \mathcal{A}) \cap \mathcal{K}\big) \le \kappa\zeta^{\gamma},$$

where $\gamma = 2^{-d(\bar{\mathcal{A}},\mathcal{K})}$, $d(\bar{\mathcal{A}},\mathcal{K})$ denotes the degree of singularity of the subspace $\bar{\mathcal{A}}$, and κ is a positive condition number.

The reader is referred to [30] for the definition of the degree of singularity. The degree of singularity of a linear mixed conic system is zero if the system satisfies the interior point condition. An upper bound for the degree of singularity is given in Theorem 8.

Theorem 8. *For the linear mixed conic system (41) we have*

$$d(\bar{\mathcal{A}},\mathcal{K}) \le \min\Big\{p + \sum_{i=p+1}^{p+q}(n_i - 1), \dim(\bar{\mathcal{A}}), \dim(\bar{\mathcal{A}}^{\perp})\Big\},$$

where $\bar{\mathcal{A}}^{\perp}$ denotes the orthogonal complement of $\bar{\mathcal{A}}$.

Theorem 7 can be applied to derive an upper bound for the distance between a central solution and the optimal set. 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

$$\begin{cases} x \in \hat{x} + \mathrm{Ker}(A), \\ \hat{s}^T x = 0, \\ x \in \mathcal{L}^{\bar{n}}_+, \end{cases} \qquad \begin{cases} s \in \hat{s} + \mathcal{R}(A^T), \\ \hat{x}^T s = 0, \\ s \in \mathcal{L}^{\bar{n}}_+, \end{cases} \tag{43}$$

see Section 4 in [30]. In this case, the linear subspace $\bar{\mathcal{A}}(\mathcal{P}^*)$, which contains $\mathcal{P}^*$, is defined as

$$\bar{\mathcal{A}}(\mathcal{P}^*) := (\mathrm{Ker}(A) \cap \{\hat{s}\}^\perp) + \mathbb{R}\hat{x}.$$

The linear subspace $\bar{\mathcal{A}}(\mathcal{D}^*)$, which is the orthogonal complement of $\bar{\mathcal{A}}(\mathcal{P}^*)$, is defined as

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

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

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

which implies $0 \leq \hat{s}^T x(\mu) \leq p\mu$ and $0 \leq \hat{x}^T s(\mu) \leq p\mu$. Then the application of Hoffman error bound [11] gives

$$\begin{aligned} \mathrm{dist}\Big(x(\mu),\ \bar{\mathcal{A}}(\mathcal{P}^*)\Big) &\leq \mathrm{dist}\Big(x(\mu),\ \big\{x \in \hat{x} + \mathrm{Ker}(A) \mid \hat{s}^T x = 0\big\}\Big) \\ &\leq \theta_1 \Big(\mathrm{dist}\Big(x(\mu),\ \big\{x \in \hat{x} + \mathrm{Ker}(A)\big\}\Big) + \hat{s}^T x(\mu)\Big) \leq \theta_1 p\mu, \end{aligned}$$

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

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

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

Lemma 11. *Let $\big(x(\mu); y(\mu); s(\mu)\big)$ be a central solution with*

$$\mu \leq \hat{\mu} := \min\Big\{\frac{1}{\theta_1 p}, \frac{1}{\theta_2 p}\Big\}. \tag{44}$$

Then there exists $(x; y; s) \in \mathcal{P}^* \times \mathcal{D}^*$ and $\kappa > 0$ so that

$$\|x(\mu) - x\| \leq \kappa(p\mu)^\gamma, \quad \|y(\mu) - y\| \leq \kappa(p\mu)^\gamma, \quad \|s(\mu) - s\| \leq \kappa(p\mu)^\gamma, \tag{45}$$

where $\gamma \geq \frac{1}{2^p}$.

Proof. Conditions (42) hold if $\mu \leq \hat{\mu}$. Moreover, the full row rank and the interior point conditions imply that the set

$$\Big\{\big(x(\mu); y(\mu); s(\mu)\big) \;\big|\; 0 < \mu \leq \hat{\mu}\Big\}$$

is contained in a compact set, see Lemma 3.2 in [8]. Therefore, the result of Theorem 7 is applicable to the linear conic systems in (43). Furthermore, the compactness of $\mathcal{P}^*$ and $\mathcal{D}^*$ implies the existence of $(x; y; s) \in \mathcal{P}^* \times \mathcal{D}^*$ so that

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

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 $A^T(y(\mu) - y) = s(\mu) - s$ has a unique solution. Therefore, using Theorem 7 again, we get

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

where $(A^T)^\dagger := (AA^T)^{-1}A$ stands for the pseudo-inverse of A^T [29]. Then, taking $\kappa := \max\{\kappa_1, \kappa_2, \kappa_3\}$ and $\gamma := \min\{\gamma_1, \gamma_2\}$, we get the results as desired. The lower bound for γ follows from Theorem 8. □

E Bounds for the $\mathcal{T}$ parts in Theorem 1

Let $i \in \mathcal{T}$ denote a block $\big(x^i(\mu); y^i(\mu); s^i(\mu)\big)$ of the central solution with $\mu \leq \hat{\mu}$. The sketch of the proof is similar to the one for Theorem 3.8 in [31]. From (5) and the central path equation $x^i(\mu) \circ s^i(\mu) = \mu e_i$ we get

$$s^i(\mu) = \frac{\mu(x^i_1(\mu); -x^i_{2:n_i}(\mu))}{(x^i_1(\mu))^2 - \|x^i_{2:n_i}(\mu)\|^2}.$$

Therefore, from $x^i_1(\mu) > \|x^i_{2:n_i}(\mu)\|$ we have

$$\begin{aligned} s^i_1(\mu) &= \frac{\mu x^i_1(\mu)}{(x^i_1(\mu))^2 - \|x^i_{2:n_i}(\mu)\|^2} = \frac{\mu}{x^i_1(\mu) - \|x^i_{2:n_i}(\mu)\|} \frac{x^i_1(\mu)}{x^i_1(\mu) + \|x^i_{2:n_i}(\mu)\|} \\ &\geq \frac{\mu}{2(x^i_1(\mu) - \|x^i_{2:n_i}(\mu)\|)}. \end{aligned} \tag{46}$$

Analogously, we can derive

$$x^i_1(\mu) \geq \frac{\mu}{2(s^i_1(\mu) - \|s^i_{2:n_i}(\mu)\|)}. \tag{47}$$

- $i \in \mathcal{T}_1$: In this case, we have $x^i = s^i = 0$ for all $(x; y; s) \in \mathcal{P}^* \times \mathcal{D}^*$, and thus the bounds in (45) reduce to

  $$\|x^i(\mu)\| \leq \kappa(p\mu)^\gamma, \quad \|s^i(\mu)\| \leq \kappa(p\mu)^\gamma. \tag{48}$$

  Consequently, it can be deducted from (46) and (48) that

  $$x^i_1(\mu) - \|x^i_{2:n_i}(\mu)\| \geq \frac{\mu}{2s^i_1(\mu)} \geq \frac{\mu}{2\|s^i(\mu)\|} \geq \frac{\mu}{2\kappa(p\mu)^\gamma},$$
  $$x^i_1(\mu) - \|x^i_{2:n_i}(\mu)\| \leq x^i_1(\mu) \leq \|x^i(\mu)\| \leq \kappa(p\mu)^\gamma.$$

  In a similar manner, using (47) we can show that

  $$\frac{\mu}{2\kappa(p\mu)^\gamma} \leq s^i_1(\mu) - \|s^i_{2:n_i}(\mu)\| \leq s^i_1(\mu) \leq \kappa(p\mu)^\gamma,$$

  which completes the first part of the proof.

- $i \in \mathcal{T}_2$: In this case, the bound in (45) reduces to $\|s^i(\mu)\| \leq \kappa(p\mu)^\gamma$. Thus, we can conclude from (46) that

  $$x^i_1(\mu) - \left\|x^i_{2:n_i}(\mu)\right\| \geq \frac{\mu}{2s^i_1(\mu)} \geq \frac{\mu}{2\|s^i(\mu)\|} \geq \frac{\mu}{2\kappa(p\mu)^\gamma}.$$

  Furthermore, it follows from $x^i_1 = \|x^i_{2:n_i}\|$ that

  $$\begin{aligned} x^i_1(\mu) - \left\|x^i_{2:n_i}(\mu)\right\| &= (x^i_1(\mu) - x^i_1) + \left(\left\|x^i_{2:n_i}\right\| - \left\|x^i_{2:n_i}(\mu)\right\|\right) \\ &\leq \left|x^i_1(\mu) - x^i_1\right| + \left\|x^i_{2:n_i} - x^i_{2:n_i}(\mu)\right\| \\ &\leq \sqrt{2}\left\|x^i(\mu) - x^i\right\| \leq \sqrt{2}\kappa(p\mu)^\gamma. \end{aligned} \tag{49}$$

  Therefore, using (46) and (49) we get

  $$\kappa(p\mu)^\gamma \geq s^i_1(\mu) \geq \frac{\mu}{2\left(x^i_1(\mu) - \left\|x^i_{2:n_i}\right\|\right)} \geq \frac{\mu}{2\sqrt{2}\kappa(p\mu)^\gamma},$$

  which completes the proof for the second part.

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

The rest of the theorem follows by applying the results from the previous parts as analogously as in Theorem 3.8 in [31].