McMaster University

Advanced Optimization Laboratory

Title:

The Q Method for Second-order Cone Programming

Authors:

Farid Alizadeh and Yu Xia

AdvOl-Report No. 2004/15

October 2004, Hamilton, Ontario, Canada

The Q Method for Second-Order Cone Programming

Farid Alizadeh * Yu Xia †

October 9, 2004

Abstract

Based on the Q method for SDP, we develop the Q method for SOCP. A modified Q method is also introduced. Properties of the algorithms are discussed. Convergence proofs are given. Finally, we present numerical results.

Key words. Second-order cone programming, infeasible interior point method, eigen space decomposition.

1 Introduction

The second-order cone programming (abbreviated as SOCP) is currently an active research area because it has many applications; see [2, 11] for a survey. It is somewhat between SDP and LP; so the computation time and approximation accuracy of SOCP are between LP and SDP. Most interior-point methods for LP and SDP have been extended to SOCP, but until now, there is no Q method([4]) for SOCP. The Q method for SDP is quite different from other methods. And it has many attractive properties: each iterate of the Q method is relatively cheap to obtain than other methods becaues no eigenvalue decomposition is needed, and the Schur complement can be calculated by Cholesky factorization; unlike some other interior point methods, this algorithm converges fast and is numerical stable near optimum since the Newton system is well defined and its Jacobian is nonsingular at the solution under certain conditions (see [4]). In this paper, we carry on the Q method to SOCP. We also give a modified Q method for SOCP. Convergence proofs are presented. These two methods for SOCP are also different from other methods, have the above properties. Preliminary numerical results show that they are promising. See also [6] for the Q method on symmetric programming and [13] for a Newton type algorithm on the Q method for SOCP.

This paper has eight parts. In § 2, we give the eigen space decomposition of any $\mathbf{x} \in \mathbb{R}^{n+1}$ and the update scheme of the orthogonal transformation. In § 3, we derive the Newton system and give the properties of the solution. In § 4, we give an algorithm that converges under certain conditions. In § 5, we further give some restrictions under which the algorithm finds an ϵ-optimal solution in finite iterations. Numerical results are given in § 6. A modified Q method which doesn't need to update the orthogonal matrix is presented in § 7. Finally, in § 8, we give conclusion and future work.

Notations

Throughout this paper, superscripts are used to represent iteration numbers while subscripts are for block numbers of the variables. We use capital letters for matrices, bold lower case letters for column vectors, lower case letters for entries of a vector. In this way, jth entry of vector $\mathbf{x}_i$ is written as

*RUTCOR and Business School, Ruters, the State University of New Jersey, U.S.A. alizadeh@rutcor.rutgers.edu
Rsearch supported in part by the U.S. National Science Foundation
†Computing and Software, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada yuxia@cas.mcmaster.ca

$(x_i)_j$. Primal and dual vectors $(\mathbf{x}, \mathbf{z})$ are indexed from 0. Superscript T represents matrix or vector transpose. Semicolon ";" is used to concatenate column vectors; so $(\mathbf{x}; \mathbf{y}; \mathbf{z}) = (\mathbf{x}^T, \mathbf{y}^T, \mathbf{z}^T)^T$.

We use $\bar{\mathbf{x}}$ to represent the sub-vector of $\mathbf{x}$ excluding x_0; thus $\mathbf{x} = (x_0, \bar{\mathbf{x}}^T)^T$.

As convension, we use $O(n)$ to represent the $n \times n$ real orthogonal groups.

For a vector $\boldsymbol{\lambda}$, we use $\operatorname{Diag} \boldsymbol{\lambda}$ to represent a diagonal matrix with $\boldsymbol{\lambda}$ on its diagonal. Sometimes we use corresponding upper case letters to represent the diagonal matrix. Thus, $\Lambda = \operatorname{Diag} \boldsymbol{\lambda}$.

$\|\cdot\|_2$ denotes the Euclidean or l_2 norm: $\|\mathbf{x}\|_2 \stackrel{\text{def}}{=} \left(\sum_{i=0}^n x_i^2\right)^{1/2}$.

$\|\cdot\|_1$ denotes the LAR(Least Absolute Residual) or l_1 norm: $\|\mathbf{x}\|_1 \stackrel{\text{def}}{=} \sum_{i=0}^n |x_i|$.

$\|\cdot\|_\infty$ denotes the Tchebycheff norm or l_∞ norm: $\|\mathbf{x}\|_\infty \stackrel{\text{def}}{=} \max_{0 \le i \le n} |x_i|$.

We denote an n-dimensional all zero vector as $\mathbf{0}_n$ and an n-dimensional vector of all ones as $\mathbf{1}_n$. We omit subscripts when the dimensions are undoubt. The identity matrix is denoted as I. The matrix R is defined as the following, whose dimension is clear from the context.

$$R \stackrel{\text{def}}{=} \begin{pmatrix} 1 & & \\ & -1 & \\ & & \ddots \end{pmatrix}.$$

A second-order cone in $\mathbb{R}^{n+1}$ is represented as

$$\mathcal{Q}_{n+1} \stackrel{\text{def}}{=} \left\{ \mathbf{x} \in \mathbb{R}^{n+1} \colon x_0 \ge \sqrt{\sum_{i=1}^n x_i^2} \right\}.$$

$\mathcal{Q}$ is also known as Loréntz cone, ice-cream cone, quadratic cone. We write $\mathbf{x} \ge_{\mathcal{Q}_{n+1}} \mathbf{0}$ interchangeably with $\mathbf{x} \in \mathcal{Q}_{n+1}$ since it is a partial order. We will also omit the subscript, just write $\mathcal{Q}$ when the dimension is clear from the context. Second-order cone is self dual. Therefore, the second-order cone program is generally written in the primal-dual pair as the following:

(1)

Primal		**Dual**	
$\min_{\mathbf{x}}$	$\mathbf{c}_1^T\mathbf{x}_1 + \cdots + \mathbf{c}_n^T\mathbf{x}_n$	$\max_{\mathbf{z},\,\mathbf{y}}$	$\mathbf{b}^T\mathbf{y}$
s.t.	$A_1\mathbf{x}_1 + \ldots A_n\mathbf{x}_n = \mathbf{b}$	s.t.	$A_i^T\mathbf{y} + \mathbf{z}_i = \mathbf{c}_i \ (i = 1, \ldots, n)$
	$\mathbf{x}_i \ge_Q \mathbf{0} \ (i = 1, \ldots, n)$		$\mathbf{z}_i \ge_Q \mathbf{0} \ (i = 1 \ldots, n)$.

Here, $\mathbf{x}_i \in \mathbb{R}^{n_i}$, $\mathbf{z}_i \in \mathbb{R}^{n_i}$, $\mathbf{y} \in \mathbb{R}^m$ are unknowns; $A_i \in \mathbb{R}^{m \times n_i}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c}_i \in \mathbb{R}^{n_i}$ are data. The dimensions n_i may not be the same.

2 Basic Properties

This section lays the basic tools for the Q method for SOCP. We first briefly sketch the Q method for semidefinite programming in §§ 2.1, then give the corresponding decompostion and update scheme for SOCP in §§ 2.2.

2.1 The Q Method for Semidefinite Programming

Basic idea of the Q method for SDP (see [4]) is the following. Let real symmetric matrices X, Z denote the primal and dual variables. When $X \bullet Z = \mu I$, it is not hard to see that X and Z commute; so they share a same complete system of eigenvectors, which can be described by an orthogonal matrix Q. Hence, the eigenvalue decompositions can be written as $X = [Q^T \Lambda Q]$ and $Z = [Q^T \Omega Q]$, where Λ and Ω are diagonal matrices with eigenvalues of X and Z as the diagonal elements respectively. The Q method employs Newton's method to the primal-dual system on the

central path by updating Q, Λ, Ω and $\mathbf{y}$ at each iteration seperately, instead of modifying X and Z as a whole.

At each iteration of the Q method, the orthogonal matrix Q is replaced by $Q(I+S)$, where S is skew-symmetric. Justification of it is the one to one correspondence between the group of real orthogonal matrices and the set of skew symmetric matrices via the exponential map – exp and the Cayley transforamtion – $\left(I+\frac{S}{2}\right)\left(I-\frac{S}{2}\right)^{-1}$. The linear approximation of each map at $S=0$ is $I+S$.

2.2 Foundations of the Q method for SOCP

To develop the Q method for SOCP, in this part, we

1. give the second-order cone related eigen space decomposition of any vector $\mathbf{x}\in\mathbb{R}^{n+1}$ (2) and corresponding approximation ($\mathfrak{l}$, $\mathfrak{l}_\pi$);
2. prove that primal and dual variables share a same orthogonal transformation (Proposition 2.1);
3. show how to update the orthogonal transformation (Proposition 2.2);
4. give linearization of the orthogonal transformation (Propositions 2.3, 2.4).

We first give the eigen space decomposition.

Given $\mathbf{x}\in\mathbb{R}^{n+1}$, denote the eigenvalues of $\mathbf{x}$ as

$$\boldsymbol{\lambda}_{\mathbf{x}}=\begin{pmatrix}x_0+\|\bar{\mathbf{x}}\|_2\\ x_0-\|\bar{\mathbf{x}}\|_2\end{pmatrix}.$$

Then, $\mathbf{x}\in\mathcal{Q}$ iff $\boldsymbol{\lambda}\geq\mathbf{0}$; $\mathbf{x}\in\operatorname{Int}\mathcal{Q}$ iff $\boldsymbol{\lambda}>\mathbf{0}$; $\mathbf{x}\in\operatorname{bd}\mathcal{Q}$ iff one of λ_i's is zero; and $\mathbf{x}=\mathbf{0}$ iff $\boldsymbol{\lambda}=\mathbf{0}$ (see [2, 7]).

We define a set of orthogonal matrices $K_{\mathbf{x}}$ related to $\mathbf{x}$ as the following:

$$K_{\mathbf{x}}\overset{\text{def}}{=}\left\{\begin{pmatrix}1 & \\ & \bar{Q}\end{pmatrix}:\bar{Q}\in O(n),\,\bar{Q}_1=\frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|_2}\text{ if }\|\bar{\mathbf{x}}\|_2\neq 0\right\}.$$

Note that each element in $K_{\mathbf{x}}$ maps $\begin{pmatrix}x_0\\ \|\bar{\mathbf{x}}\|_2\\ \mathbf{0}\end{pmatrix}$ to $\mathbf{x}$. Hence, $\mathbf{x}$ can be written as

$$\mathbf{x}=Q_{\mathbf{x}}\left[\lambda_1\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\\ \mathbf{0}\end{pmatrix}+\lambda_2\begin{pmatrix}\frac{1}{2}\\ -\frac{1}{2}\\ \mathbf{0}\end{pmatrix}\right]. \tag{2}$$

Remark 2.1 *In the above discussion, we assume the dimension of* $\mathbf{x}$ *is more than* 2*. When the dimension of* $\mathbf{x}$ *is 2, we can still write the decomposition of* $\mathbf{x}$ *in the form of* (2) *by letting* $Q_{\mathbf{x}}=I$*,* $\lambda_i=x_0\pm x_1$*. When the dimension of* x *is 1, we may let* $Q_x=1$*,* $\lambda_1=\lambda_2=x$*.*

Using the conventional notation $\Lambda=\operatorname{Diag}\boldsymbol{\lambda}$; then analogous to that of SDP, we have the following proposition showing that the primal and dual variables share a same orthogonal transformation on the central path.

Proposition 2.1 *The primal and dual pair* $\mathbf{x}$ *and* $\mathbf{z}$ *is on the analytic center for* (1) *iff for each block* $i=1,\ldots,n$*, there exists a real orthogonal matrix* Q_i *such that*

$$\mathbf{x}_i=Q_i\left[(\lambda_i)_1\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\\ \mathbf{0}\end{pmatrix}+(\lambda_i)_2\begin{pmatrix}\frac{1}{2}\\ -\frac{1}{2}\\ \mathbf{0}\end{pmatrix}\right],\quad \mathbf{z}_i=Q_i\left[(\omega_i)_1\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\\ \mathbf{0}\end{pmatrix}+(\omega_i)_2\begin{pmatrix}\frac{1}{2}\\ -\frac{1}{2}\\ \mathbf{0}\end{pmatrix}\right], \tag{3}$$

and

$$\Lambda_i\boldsymbol{\omega}_i=\mu\mathbf{1},\quad \boldsymbol{\lambda}_i\geq\mathbf{0},\quad \boldsymbol{\omega}_i\geq\mathbf{0}. \tag{4}$$

Proof: By [1], a pair $(\mathbf{x}, \mathbf{z})$ on the analytic center means for $(i = 1, \ldots, n,\ j = 1, \ldots, n_i)$ [recall n_i is the dimension of the *ith* block],

$$\mathbf{x}_i^T \mathbf{z}_i = \mu, \quad \mathbf{x} \geq_Q \mathbf{0},\ \mathbf{z} \geq_Q \mathbf{0}, \quad \text{and} \tag{5}$$

$$(x_i)_j (z_i)_0 + (x_i)_0 (z_i)_j = 0. \tag{6}$$

The sufficiency is easy to verify for (5) and (6).

Next, we prove the necessity.

For $i = 1, \ldots, n$, when $\mu = 0$, by Cauchy-Schwartz-Bomiakovsky inequality, also $(x_i)_0 \geq \|\bar{\mathbf{x}}_i\|_2$ and $(z_i)_0 \geq \|\bar{\mathbf{z}}\|_2$, we have

$$0 \leq (x_i)_0 (z_i)_0 - \|\bar{\mathbf{x}}_i\|_2 \|\bar{\mathbf{z}}_i\|_2 \leq \mathbf{x}_i^T \mathbf{z}_i = 0.$$

So one of $\mathbf{x}_i$ and $\mathbf{z}_i$ must be zero, or both of them must be in the boundary. If either $\mathbf{x}_i$ or $\mathbf{z}_i$ is zero, $\boldsymbol{\lambda}_i$ or $\boldsymbol{\omega}_i$ must also be zero correspondingly; hence, (3) and (4) are satisfied trivially.
When neither $(x_i)_0$ nor $(z_i)_0$ is zero, by (6),

$$\bar{\mathbf{z}}_i = -\frac{(z_i)_0}{(x_i)_0} \bar{\mathbf{x}}_i.$$

Setting

$$Q_i = Q_{\mathbf{x}_i} \in K_{\mathbf{x}_i},$$

we get (3) and (4).

When $\mu \neq 0$, it is proved in [1] that on the analytic center,

$$\mathbf{x}_i = \frac{\gamma(\mathbf{x}_i)}{\mu} R \mathbf{z}_i, \tag{7}$$

where $\gamma(\mathbf{x}_i) \stackrel{\text{def}}{=} \mathbf{x}_i^T R \mathbf{x}_i$.
By (2), $\exists\, Q_{\mathbf{z}_i} \in K_{\mathbf{z}_i}$, such that

$$\mathbf{z}_i = Q_{\mathbf{z}_i} \left[(\omega_i)_1 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \mathbf{0} \end{pmatrix} + (\omega_i)_2 \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \mathbf{0} \end{pmatrix} \right]. \tag{8}$$

Combining (7) and (8), we see that

$$\mathbf{x}_i = Q_{\mathbf{z}_i} \left[(\lambda_i)_1 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \mathbf{0} \end{pmatrix} + (\lambda_i)_2 \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \mathbf{0} \end{pmatrix} \right], \tag{9}$$

with

$$(\lambda_i)_1 = \frac{\gamma(\mathbf{x}_i)}{\mu} (\omega_i)_2, \qquad (\lambda_i)_2 = \frac{\gamma(\mathbf{x}_i)}{\mu} (\omega_i)_1. \tag{10}$$

That shows (3).

Substituting (8) and (9) to (5) with consideration of (10), we get (4).

Thus, we have proved the proposition. ■

Next, we will prove that the update of the orthogonal matrix can be obtained from some special orthogonal matrices.

Let L be a subset of

$$K \stackrel{\text{def}}{=} \bigcup_{\mathbf{x} \in \mathbb{R}^{n+1}} K_{\mathbf{x}},$$

defined as

$$L \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} 1 & 0 & \mathbf{0}^T \\ 0 & c_0 & -\bar{\mathbf{c}}^T \\ \mathbf{0} & \bar{\mathbf{c}} & I - \frac{\bar{\mathbf{c}}\bar{\mathbf{c}}^T}{1+c_0} \end{pmatrix} : \mathbf{c} \in \mathbb{R}^n,\ \|\mathbf{c}\|_2 = 1,\ c_0 \neq -1 \right\} \cup \left\{ \begin{pmatrix} 1 & 0 & 0 & \mathbf{0}^T \\ 0 & -1 & 0 & \mathbf{0}^T \\ 0 & 0 & -1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I \end{pmatrix} \right\}.$$

Apparently, L is a subgroup of $O(n)$. We have the following propositions regarding the group L.

Proposition 2.2 *Given* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n+1}$*, and* $Q_{\mathbf{x}} \in K_{\mathbf{x}}$*, there exists* $Q_{\mathbf{c}} \in L$*, such that* $Q_{\mathbf{x}} Q_{\mathbf{c}} \in K_{\mathbf{y}}$*. In addition,* $Q_{\mathbf{x}} L \subseteq K$*.*

Proof: 1) When $\|\bar{\mathbf{y}}\|_2 = 0$, any $Q_{\mathbf{c}} \in L$ satisfies $Q_{\mathbf{x}} Q_{\mathbf{c}} \in K_{\mathbf{y}}$.
When $\|\bar{\mathbf{y}}\|_2 \neq 0$, since $\bar{Q}_{\mathbf{x}}$ is nonsingular, there is a unique $\mathbf{c} \in \mathbb{R}^n$, such that $\bar{Q}_{\mathbf{x}}\mathbf{c} = \frac{\bar{\mathbf{y}}}{\|\bar{\mathbf{y}}\|_2}$. Observe $\|\mathbf{c}\|_2 = 1$. Not that each element in L is determined solely by a point on the unit sphere in $\mathbb{R}^n$. We form $Q_{\mathbf{c}} \in L$ as

$$Q_{\mathbf{c}} = \begin{cases} \begin{pmatrix} 1 & 0 & \mathbf{0}^T \\ 0 & c_0 & -\bar{\mathbf{c}}^T \\ \mathbf{0} & \bar{\mathbf{c}} & I - \frac{\bar{\mathbf{c}}\bar{\mathbf{c}}^T}{1+c_0} \end{pmatrix} & c_0 \neq -1, \\ \begin{pmatrix} 1 & 0 & 0 & \mathbf{0}^T \\ 0 & -1 & 0 & \mathbf{0}^T \\ 0 & 0 & -1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I \end{pmatrix} & c_0 = -1. \end{cases}$$

And it is easy to see that $Q_{\mathbf{x}} Q_{\mathbf{c}} \in K_{\mathbf{y}}$.

2) It is easy to verify that for any $Q_{\mathbf{x}} \in K$ and $Q_{\mathbf{c}} \in L$, we have $Q_{\mathbf{x}} Q_{\mathbf{c}} \in K$. Thus, $Q_{\mathbf{x}} L \subseteq K$. ■
The above proposition implies that to update the decomposition (9) of $\mathbf{x}$ to that of $\mathbf{x} + \Delta\mathbf{x}$, we only need to restrict orthogonal matrices in L.

To apply Newton's method to (14), next we will give the linear approximation of every element in L.

Define $\mathfrak{l}$ to be a set of skew-symmetric matrices in the following form:

$$\mathfrak{l} \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} 0 & 0 & \mathbf{0}^T \\ 0 & 0 & \mathbf{s}^T \\ \mathbf{0} & -\mathbf{s} & 0 \end{pmatrix} : \mathbf{s} \in \mathbb{R}^{n-1} \right\}.$$

Let $\mathfrak{l}_\pi$ be a subset of $\mathfrak{l}$:

$$\mathfrak{l}_\pi \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} 0 & 0 & \mathbf{0}^T \\ 0 & 0 & \mathbf{s}^T \\ \mathbf{0} & -\mathbf{s} & 0 \end{pmatrix} :\ \|\mathbf{s}\|_2 \leq \pi \right\}.$$

The following propositions relate L to $\mathfrak{l}$ or $\mathfrak{l}_\pi$.

Proposition 2.3 *The mapping* $\exp : \mathfrak{l}_\pi \mapsto L$ *is a bijection.*

Proof: For any $S \in \mathfrak{l}$,

$$S^2 = \begin{pmatrix} 0 & 0 & \mathbf{0}^T \\ 0 & -\mathbf{s}^T\mathbf{s} & \mathbf{0}^T \\ \mathbf{0} & \mathbf{0} & -\mathbf{s}\mathbf{s}^T \end{pmatrix}, \quad S^{(2k+1)} = (-\mathbf{s}^T\mathbf{s})^k S, \quad S^{(2k+2)} = (-\mathbf{s}^T\mathbf{s})^k S^2. \tag{11}$$

Hence, given $\|\mathbf{s}\|_2 \neq 0$,

$$\exp(S) = I + \frac{S^2}{\|\mathbf{s}\|_2^2}\left[\sum_{i=1}^{\infty}(-1)^{i+1}\frac{\|\mathbf{s}\|_2^{2i}}{(2i)!}\right] + \frac{S}{\|\mathbf{s}\|_2}\left[\sum_{i=0}^{\infty}(-1)^{i}\frac{\|\mathbf{s}\|_2^{2i+1}}{(2i+1)!}\right] \tag{12}$$
$$= I + \frac{1-\cos(\|\mathbf{s}\|_2)}{\|\mathbf{s}\|_2^2}S^2 + \frac{\sin(\|\mathbf{s}\|_2)}{\|\mathbf{s}\|_2}S.$$

We use $Q_{\mathbf{c}}$ to emphasize the dependence of an element in L on $\mathbf{c} \in \mathbb{R}^n$ (with $\|\mathbf{c}\|_2 = 1$) in this proof.

First we will prove that $\forall\, Q \in L$, $\exists\, S \in \mathfrak{l}_\pi$, such that $exp(S) = Q$.
Notice $\exp(0) = I = Q_{(1;\mathbf{0})}$; and for $\mathbf{s} = (\pi; \mathbf{0})$,

$$\exp(S) = \begin{pmatrix} 1 & 0 & 0 & \mathbf{0}^T \\ 0 & -1 & 0 & \mathbf{0}^T \\ 0 & 0 & -1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I \end{pmatrix} = Q_{(-1;\mathbf{0})}.$$

Now given $\mathbf{c} \in \mathbb{R}^n$, $\|\mathbf{c}\|_2 = 1$, $|c_0| \neq 1$, there is a unique $0 < \alpha < \pi$ such that $\cos\alpha = c_0$ and $\sin\alpha = \|\bar{\mathbf{c}}\|_2$. Notice $\|\bar{\mathbf{c}}\|_2 \neq 0$, we let $\mathbf{s} = -\frac{\alpha}{\|\bar{\mathbf{c}}\|_2}\bar{\mathbf{c}}$; then $\exp(S) = Q_{\mathbf{c}}$. Different $\mathbf{s}$ will give different $Q_{\mathbf{c}}$ since the $(2,2)$-entry of $\exp(S)$ is $\cos(\|\mathbf{s}\|_2)$, and the $(3:n,2)$-entry of $\exp(S)$ is $-\frac{\sin(\|\mathbf{s}\|_2)}{\|\mathbf{s}\|_2}\mathbf{s}$.

On the other hand, given $S \in \mathfrak{l}_\pi, S \neq 0$, let $\bar{\mathbf{c}} = -\frac{\sin(\|\mathbf{s}\|_2)}{\|\mathbf{s}\|_2}\mathbf{s}$, $c_0 = \cos\|\mathbf{s}\|_2$; then $\exp(S) = Q_{\mathbf{c}} \in L$. ■

Proposition 2.4 *The sets L and $\mathfrak{l}$ can be related by Cayley transformation $(I + \frac{S}{2})(I - \frac{S}{2})^{-1}$.*

Proof: When $\left\|\frac{1}{2}\mathbf{s}\right\|_2 < 1$, according to Neumann Lemma, $(I - \frac{1}{2}S)^{-1}$ can be expanded by power series. So the Cayley transformation is

$$(I + \frac{1}{2}S)(I - \frac{1}{2}S)^{-1} = I + \sum_{k=1}^{+\infty}\frac{1}{2^{k-1}}S^k.$$

By (11), the Cayley transformation is equivalent to

$$(I + \frac{1}{2}S)(I - \frac{1}{2}S)^{-1} = I + \frac{4}{4+\|\mathbf{s}\|_2^2}S + \frac{2}{4+\|\mathbf{s}\|_2^2}S^2. \tag{13}$$

Since the right hand side of (13) is well defined even for $\|\mathbf{s}\|_2 \geq 2$, we use the right hand side of (13) as the definition of Cayley transformation for any $S \in \mathfrak{l}$ (See Appendix for the justification of this definition).

It is not hard to see that given $S \in \mathfrak{l}$, the Cayley transformation of S is $Q_{\left(\frac{4-\|\mathbf{s}\|_2^2}{4+\|\mathbf{s}\|_2^2}; -\frac{4\mathbf{s}}{4+\|\mathbf{s}\|_2^2}\right)} \in L$.
Next we will show that given $Q_{\mathbf{c}} \in L$, there is an $S \in \mathfrak{l}$, such that the Cayley transformation of S is $Q_{\mathbf{c}}$.
Denote the first element of $\mathbf{s}$ as s_1. Then when $s_1 \to \infty$ and $\frac{|s_1|}{\|\mathbf{s}\|_2} \to 1$, the Cayley transformation of S converges to $Q_{(-1;\mathbf{0})}$.
Given $\mathbf{c} \in \mathbb{R}^n$, such that $\|\mathbf{c}\|_2 = 1$ and $c_0 \neq -1$, let $\mathbf{s} = -\frac{2\bar{\mathbf{c}}}{c_0+1}$. Then the Cayley transformation of S is $Q_{\mathbf{c}}$. The uniqueness of S can be proved similarly as that in Proposition 2.3. ■

Propositions 2.3, 2.4 show that the tangent space to L at the identity I is $\mathfrak{l}$.

3 The Newton System

In this section, we will first derive the Newton System, and then give some properties of its solution (16), including the nonsingularity.

By Proposition 2.1, on the central path, each iterate $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ satisfies

$$\begin{aligned} Q\tilde{P}\boldsymbol{\omega} + A^T\mathbf{y} &= \mathbf{c}, \\ AQ\tilde{P}\boldsymbol{\lambda} &= \mathbf{b}, \\ \Lambda\Omega &= \mu I, \end{aligned} \tag{14}$$

where $\tilde{P}$ is block diagonal, whose ith block, denoted as $\tilde{P}_i \in \mathbb{R}^{n_i \times 2}$, is in the form

$$\tilde{P}_i \stackrel{\text{def}}{=} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 \\ \vdots & \vdots \end{pmatrix}.$$

It is known that if both the primal and dual of (1) have an interior feasible solution and A has full row rank, then $\forall\, \mu > 0$, (14) has a unique solution $(\mathbf{x}_\mu, \mathbf{y}_\mu, \mathbf{z}_\mu)$, and as $\mu \to 0$, $(\mathbf{x}_\mu, \mathbf{y}_\mu, \mathbf{z}_\mu)$ tends to the optimum of (1)(see[1]). Assume $\mathbf{x} = Q_{\mathbf{x}}\boldsymbol{\lambda}$ by decomposition (2), then any perturbation of $\mathbf{x}$ can be written as $Q_{\mathbf{x}}Q_{\Delta\mathbf{x}}(\boldsymbol{\lambda} + \Delta\boldsymbol{\lambda})$, with $Q_{\Delta\mathbf{x}} \in L$ by Proposition 2.2.

By Proposition 2.3 and Proposition 2.4, we can replace each diagonal block of $Q_{\Delta\mathbf{x}}$ by $\exp(S_i)$ with $S_i \in \mathfrak{l}_\pi$, or by Cayley transformation of S_i with $S_i \in \mathfrak{l}$; and then discard the nonlinear terms. Notice when $\|\mathbf{s}\|_2 \to 0$, both the linear terms of its exponential and Cayley transformation converge to $I + S$. Define

$$\mathbf{r}_p \stackrel{\text{def}}{=} \mathbf{b} - A\mathbf{x}, \quad \mathbf{r}_d \stackrel{\text{def}}{=} \mathbf{c} - \mathbf{z} - A^T\mathbf{y}, \quad \mathbf{r}_c \stackrel{\text{def}}{=} \operatorname{vec}(\mu I - \Lambda\Omega).$$

Given the kth iterate $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{z}^k) = \left(Q^k\tilde{P}\boldsymbol{\lambda}^k, \mathbf{y}^k, Q^k\tilde{P}\boldsymbol{\omega}^k\right)$, we denote $B^k \stackrel{\text{def}}{=} AQ^k$.

Note that only the vector $\mathbf{s}$, not the matrix S is involved in calculation. Let P be a block diagonal matrix with each diagonal block in the form $\left(\begin{smallmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{smallmatrix}\right)$. It is obvious that $P^{-1} = 2P$. After collecting all the first two columns of B_i^k into $\bar{B}^k$, the remaining columns into $\hat{B}^k$, and splitting ${Q^k}^T\mathbf{r}_d^k$ accordingly as $\bar{\mathbf{r}}_d^k$ and $\hat{\mathbf{r}}_d^k$, we rewrite the Newton system as

$$\begin{aligned} P\,\Delta\boldsymbol{\omega} + (\bar{B}^k)^T\,\Delta\mathbf{y} &= \bar{\mathbf{r}}_d^k, \\ \frac{(\omega_i^k)_2 - (\omega_i^k)_1}{2}\mathbf{s}_i + (\hat{B}_i^k)^T\,\Delta\mathbf{y} &= (\hat{\mathbf{r}}_d)_i^k \quad (i = 1, \ldots, n), \\ \bar{B}^k P\,\Delta\boldsymbol{\lambda} + \sum_{i=1}^{n}\left[\frac{(\lambda_i^k)_2 - (\lambda_i^k)_1}{2}\hat{B}_i^k\mathbf{s}_i\right] &= \mathbf{r}_p^k, \\ \Lambda^k\,\Delta\boldsymbol{\omega} + \Omega^k\,\Delta\boldsymbol{\lambda} &= \mathbf{r}_c^k. \end{aligned} \tag{15}$$

For simplicity, in the next context we will omit k. Define $E_i \stackrel{\text{def}}{=} \frac{(\boldsymbol{\omega}_i)_2 - (\boldsymbol{\omega}_i)_1}{2} I$, $D_i \stackrel{\text{def}}{=} \frac{(\boldsymbol{\lambda}_i)_2 - (\boldsymbol{\lambda}_i)_1}{2} I$.

Correspondingly, define $E \stackrel{\text{def}}{=} \text{Diag}(E_i)$, $D \stackrel{\text{def}}{=} \text{Diag}(D_i)$. Hence, solution to (15) is

$$
\begin{aligned}
\Delta \mathbf{y} &= \left(\bar{B}P2\Omega^{-1}\Lambda P^T\bar{B}^T - \hat{B}DE^{-1}\hat{B}^T\right)^{-1}\left(\mathbf{r}_p - \bar{B}P\Omega^{-1}\mathbf{r}_c - \hat{B}DE^{-1}\hat{\mathbf{r}}_d\right. \\
&\quad \left. + \bar{B}P2\Omega^{-1}\Lambda P^T\bar{\mathbf{r}}_d\right) \\
\Delta \boldsymbol{\omega} &= P^{-1}\left(\bar{\mathbf{r}}_d - \bar{B}^T\Delta \mathbf{y}\right) \\
\Delta \boldsymbol{\lambda} &= \Omega^{-1}\left(\mathbf{r}_c - \Lambda\Delta\boldsymbol{\omega}\right) \\
\mathbf{s} &= 2E^{-1}\left(\hat{\mathbf{r}}_d - \hat{B}^T\Delta\mathbf{y}\right).
\end{aligned} \tag{16}
$$

Properties of the Solution

1. Though (14) is a primal-dual system, since we force primal and dual variables share a same orthogonal matrix in the decomposition, the number of variables and equations in (15) are about half of that required by other algorithms.

2. Each iterate is relative cheap to compute, because
   1) each block of the Schur complement is two dimensions less than that of other systems, which means less computaion for the search direction;
   2) to keep each iterate in $\mathcal{Q}$, i.e. $\mathbf{x}+\alpha\Delta\mathbf{x} \geq_Q \mathbf{0}$, instead of solving $x_0+\alpha\Delta x_0 \geq \|\bar{\mathbf{x}}+\alpha\Delta\bar{\mathbf{x}}\|_2$for α as that in other methods, one only needs to calculate $\alpha^{-1} \geq \max\{-\Delta\lambda_i/\lambda_i\colon \Delta\lambda_i < 0\}$.

3. The Schur complement of (16) is symmetric positive definite; so the Cholesky factorization is applicable for the computation of the search direction.

   This can be seen by writing the Schur complement as

   $$AQ\,\text{Diag}\left(\begin{smallmatrix} 2P\Omega_i^{-1}\Lambda_i P^T & \\ & -D_iE_i^{-1}\end{smallmatrix}\right)Q^TA^T,$$

   which is positive definite when A has full row rank, $(\lambda_i)_1 > (\lambda_i)_2 > 0$, and $(\omega_i)_2 > (\omega_i)_1 > 0$.

4. The Jacobian of the solution is nonsingular under mild conditions. See Theorem 3.1.

Therefore, we can expect low computation time, high convergence rate and numerical stable near optimum, and high accuracy of the algorithm under the assumption of the theorem. This property is not shared by some other search directions whose Jacobians become increasingly ill-conditioned near optimum.

Remark 3.1 *For each $1 \leq i \leq n$, we can always ensure that $(\lambda_i^{k+1})_1 > (\lambda_i^{k+1})_2 > 0$ and $(\omega_i^{k+1})_2 > (\omega_i^{k+1})_1 > 0$ by careful choice of step sizes. For example, if $(\omega_i^{k+1})_1 > (\omega_i^{k+1})_2$, we swap them. Assume $(\omega_i^k)_2 > (\omega_i^k)_1$, only when $\Delta(\omega_i)_1 \neq \Delta(\omega_i)_2$ and $\beta = \frac{(\omega_i^k)_2-(\omega_i^k)_1}{\Delta(\omega_i)_1-\Delta(\omega_i)_2}$, is it possible that $(\omega_i^k)_2 + \beta\Delta(\omega_i)_2 = (\omega_i^k)_1 + \beta\Delta(\omega_i)_1$. Under this case, we can use a smaller step size β_i'. It is obvious that β' can be at least as large as $\frac{\beta}{2}$. And β_i' are not necessarily the same for all i.*

Next, we will show the nonsingularity of Jacobian at optimum.
Assume $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ is a solution of (1). Suppose Q simultaneously diagonalize $\mathbf{x}$ and $\mathbf{z}$. And

$$(\boldsymbol{\lambda}_i)_1 > (\boldsymbol{\lambda}_i)_2 \geq 0\ ,\ 0 \leq (\boldsymbol{\omega}_i)_1 < (\boldsymbol{\omega}_i)_2 \tag{17}$$

for any nonzero block $\mathbf{x}_i$ or $\mathbf{z}_i$, $i \in \{1, \ldots, n\}$. We also assume $\mathbf{x} \neq \mathbf{0}$, since otherwise, $\mathbf{b} = \mathbf{0}$, the dual is trivial. Analogous to [4, Theorem 6.1], we have the following results.

Theorem 3.1 *Let* $(\mathbf{x},\mathbf{y},\mathbf{z}) = (Q\tilde{P}\boldsymbol{\lambda},\mathbf{y},Q\tilde{P}\boldsymbol{\omega})$ *be an optimal solution of* (1) *satisfying strict complementarity, primal and dual nondegeneracy conditions, and also condition* (17)*; then the Jacobian of* (15) *evaluated at* $(\mathbf{x},\mathbf{y},\mathbf{z})$ *is nonsingular.*

Proof: It is easy to verify that the strict complementarity (see [3]) equals to that one and only one of $(\lambda_i)_j$ and $(\omega_i)_j$ is zero for each $1 \le i \le n$, $j = 1, 2$.
As in [3], we partition the index set $\{1,\ldots,n\}$ into three subsets B, I, O, and write $\mathbf{x}$ as $(\mathbf{x}_B;\mathbf{x}_I;\mathbf{x}_O)$, where $\mathbf{x}_B$ includes all the boundary blocks, $\mathbf{x}_I$ includes all the interior blocks, and $\mathbf{x}_O$ includes all the zero blocks. Assume $\mathbf{x}_B = (\mathbf{x}_1,\ldots,\mathbf{x}_r)$. It is proved in [3] that primal nondegeneracy means matrix in the following form has linearly independent rows for all $\alpha_1,\ldots,\alpha_r$ and $\boldsymbol{\nu}$ that are not all zeros.

$$\begin{pmatrix} A_1 & \ldots & A_r & A_I & A_O \\ \alpha_1(R\mathbf{x}_1)^T & \ldots & \alpha_r(R\mathbf{x}_r)^T & \mathbf{0}^T & \boldsymbol{\nu}^T \end{pmatrix} \tag{18}$$

Let $\hat{P}$ be a block diagonal matrix with each diagonal block in the form $\left(\begin{smallmatrix} P & \\ & I \end{smallmatrix}\right)$, where $I \in \mathbb{R}^{(n_i-2)\times(n_i-2)}$ is the identity (a little abuse of notation, P here represents $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$). By [1, Lemma 2.1], at optimum, there exists a vector $\beta > \mathbf{0}$, such that $R\mathbf{x}_i = \beta_i\mathbf{z}_i$ for $(i = 1,\ldots,r)$. Substituting $\mathbf{z}_i$ by its eigen space decomposition, get $R\mathbf{x}_i = \beta_i Q_i\hat{P}_i \begin{pmatrix} 0 \\ (\boldsymbol{\omega}_i)_2 \\ \mathbf{0} \end{pmatrix}$ with $(\omega_i)_2 > 0$ for $(i = 1,\ldots,r)$ due to strict complementarity. Postmultiplying $Q\hat{P}$ to (18), we obtain the following matrix.

$$\begin{pmatrix} (\bar{B}_1P)_1 & (\bar{B}_1P)_2 & \hat{B}_1 & \ldots & (\bar{B}_rP)_1 & (\bar{B}_rP)_2 & \hat{B}_r & B_I\hat{P}_I & B_O\hat{P}_O \\ 0 & \frac{1}{2}\alpha_1\beta_1(\boldsymbol{\omega}_1)_2 & \mathbf{0}^T & \ldots & 0 & \frac{1}{2}\alpha_r\beta_r(\boldsymbol{\omega}_r)_2 & \mathbf{0}^T & \mathbf{0}^T & \boldsymbol{\nu}^TQ\hat{P}_O \end{pmatrix} \tag{19}$$

Notice (18) has full row rank, and right timing a nonsingular matrix doesn't change its rank; so (19) has full row rank for all $\alpha_1,\ldots,\alpha_r$ and $\boldsymbol{\nu}$ that are not all zeros. Hence the matrix

$$\begin{pmatrix} (\bar{B}_1P)_1 & \hat{B}_1 & \ldots & (\bar{B}_rP)_1 & \hat{B}_r & B_I\hat{P}_I \end{pmatrix} \tag{20}$$

has full row rank.

The solution satisfies dual nondegeneracy and strict complementarity iff the following matrix has linearly independent columns (see [3]).

$$\begin{pmatrix} A_1R\mathbf{z}_1 & \ldots & A_rR\mathbf{z}_r & A_I \end{pmatrix} \tag{21}$$

Because

$$A_iR\mathbf{z}_i = \frac{1}{\beta_i}A_i\mathbf{x}_i = \frac{1}{\beta_i}A_iQ_i\hat{P}\hat{P}^{-1}Q_i^T\mathbf{x}_i = \frac{1}{\beta_i}(\bar{B}_iP)_1(\boldsymbol{\lambda}_i)_1 \quad (i \in B),$$

(21) equals to the following matrix having full column rank:

$$\begin{pmatrix} (\bar{B}_1P)_1 & \ldots & (\bar{B}_rP)_1 & B_I\hat{P}_I \end{pmatrix}. \tag{22}$$

So (20) and (22) mean we can choose all columns of $(\bar{B}_iP)_1\,(i = 1,\ldots,r)$ and $B_I\hat{P}_I$, together with some columns from $\hat{B}_i\,(i = 1,\ldots,r)$ to form an $m \times m$ nonsingular matrix B_1.

Because of the above properties, we first premultiply P $\left(= \mathrm{Diag}\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}\right)\right)$ to the $1st$ block equations of (15); then form a nonsingular matrix B_1, collect all the remaining columns of $\bar{B}P$ to L_2, all the remaining columns of $\hat{B}$ to R_2; partition $D = \mathrm{Diag}(D_1, D_2)$ and $E = \mathrm{Diag}(E_1, E_2)$ accordingly. Since D_1 includes only $\boldsymbol{\lambda}_i$'s from boundary and interior blocks, E_2 including only $\boldsymbol{\omega}_i$'s from boundary and zero blocks, we see $-D_1 \succ 0$ and $E_2 \succ 0$. Define $\tilde{D} \stackrel{\text{def}}{=} \mathrm{Diag}(I, D_1)$, $\tilde{E} \stackrel{\text{def}}{=} \mathrm{Diag}(0, E_1)$, $\tilde{I} = \left(\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right)$.

After permuting the rows and columns of the Jacobian of (15) properly, we find the nonsingularity of Jacobian is the same as the nonsingularity of the following matrix:

$$\begin{pmatrix} \tilde{E} & & & B_1^T & \frac{1}{2}\tilde{I} & \\ & & & L_2^T & & \frac{1}{2}I \\ & & E_2 & R_2^T & & \\ B_1\tilde{D} & L_2 & R_2D_2 & & & \\ & & & & \Lambda_1 & \\ & \Omega_2 & & & & \end{pmatrix}.$$

We first interchange the $1st$ and the $4th$ block rows, the $2nd$ and the last block columns; then subtract $\tilde{E}\tilde{D}^{-1}B_1^{-1}$ timing the $1st$ block rows from the $4th$ block rows, add $\tilde{E}\tilde{D}^{-1}B_1^{-1}R_2D_2E_2^{-1}$ timing the $3rd$ block rows from the $4th$ block rows. Hence the nonsingularity of the above matrix equals to the nonsingularity of

$$B_1^T + \tilde{E}\tilde{D}^{-1}B_1^{-1}R_2D_2E_2^{-1}R_2^T. \tag{23}$$

Left multiplying (23) by B_1^{-T}, we get the matrix

$$I + (B_1^{-T}\tilde{E}\tilde{D}^{-1}B_1^{-1})(R_2D_2E_2^{-1}R_2^T),$$

which is nonsingular since $B_1^{-T}\tilde{E}\tilde{D}^{-1}B_1^{-1}$ and $R_2D_2E_2^{-1}R_2^T$ are symmetric negative semidefinite. ■

4 The Algorithm

In this section, we will give a convergent algorithm for the Q method for SOCP.

This algorithm is originally for infeasible LP with exact search directions [10], while the system for the Q method is nonlinear and the search direction is not exact. It can start from an arbitrary infeasible interior point. So it doesn't employ big M method; consequently, it doesn't have the drawback of the big M method – numerically instable and computationally inefficient, see [12]. Its accuracy measures for primal, dual infeasibility and complementarity can be chosen separately; primal and dual step sizes can be different.

The algorithm is described in §§ 4.1. Its convergent analysis is given in §§ 4.2.

4.1 Algorithm Description

Let ϵ_p, ϵ_d, and ϵ_c denote the accuracy requirement for the primal, dual feasibility and duality gap. The neighborhood we using is

$$\begin{aligned} \mathcal{N}(\gamma_c,\gamma_p,\gamma_d) \stackrel{\text{def}}{=} \Big\{ (\boldsymbol{\lambda},\boldsymbol{\omega},\mathbf{y},Q)\colon\ & \boldsymbol{\lambda}\in\mathbb{R}^{2n}, \boldsymbol{\omega}\in\mathbb{R}^{2n}, \mathbf{y}\in\mathbb{R}^m, Q\in K,\ \boldsymbol{\lambda}>0,\ \boldsymbol{\omega}>0, \\ & (\lambda_i)_j(\omega_i)_j \ge \gamma_c\frac{\boldsymbol{\lambda}^T\boldsymbol{\omega}}{2n}\ (j=1,2; i=1,\ldots,n), \\ & \boldsymbol{\lambda}^T\boldsymbol{\omega} \ge \gamma_p \left\|AQ\tilde{P}\boldsymbol{\lambda}-\mathbf{b}\right\|_2 \text{ or } \left\|AQ\tilde{P}\boldsymbol{\lambda}-\mathbf{b}\right\|_2 \le \epsilon_p, \\ & \boldsymbol{\lambda}^T\boldsymbol{\omega} \ge \gamma_d \left\|A^T\mathbf{y}+Q\tilde{P}\boldsymbol{\omega}-\mathbf{c}\right\|_2 \text{ or } \left\|A^T\mathbf{y}+Q\tilde{P}\boldsymbol{\omega}-\mathbf{c}\right\|_2 \le \epsilon_d. \Big\} \end{aligned}$$

The first inequality is the centrality condition. The second and third inequalities guarantee that the complementarity will not be achieved before the primal or the dual feasibility. Obviously, when $(\gamma_c',\gamma_p',\gamma_d') \le (\gamma_c,\gamma_p,\gamma_d)$,

$$\mathcal{N}(\gamma_c,\gamma_p,\gamma_d) \subseteq \mathcal{N}(\gamma_c',\gamma_p',\gamma_d').$$

And

$$\bigcup_{(\gamma_c,\gamma_p,\gamma_d)>0} \mathcal{N}(\gamma_c,\gamma_p,\gamma_d) = \{(\boldsymbol{\lambda},\boldsymbol{\omega},\mathbf{y},Q)\colon \boldsymbol{\lambda}>0,\boldsymbol{\omega}>0\}.$$

Clearly, when $\boldsymbol{\lambda}^T\boldsymbol{\omega}$ approaches 0, $\mathcal{N}$ tends to the optimal solution set of (1). The algorithm is the following.

Algorithm 1
Choose $0<\sigma_1<\sigma_2<\sigma_3<1$ and $\Upsilon>0$. To start from an arbitrary point $(\boldsymbol{\lambda}^0,\boldsymbol{\omega}^0,\mathbf{y}^0,Q^0)$, one may select $0<\gamma_c<1,\gamma_p>0,\gamma_d>0$, so that $(\boldsymbol{\lambda}^0,\boldsymbol{\omega}^0,\mathbf{y}^0,Q^0)\in\mathcal{N}(\gamma_c,\gamma_p,\gamma_d)$.
Do until (1) $\left\|\mathbf{r}_p^k\right\|_2<\epsilon_p$, $\left\|\mathbf{r}_d^k\right\|_2<\epsilon_d$, and ${\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k<\epsilon_c$; or (2) $\left\|(\boldsymbol{\lambda}^k,\boldsymbol{\omega}^k)\right\|_1>\Upsilon$.

1. Set $\mu=\sigma_1\frac{{\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k}{2n}$.
2. Compute the search direction $(\Delta\boldsymbol{\lambda},\Delta\boldsymbol{\omega},\Delta\mathbf{y},\mathbf{s})$ from (15).
3. Choose step sizes α, β, γ, set

$$\begin{aligned}
\Lambda^{k+1} &= \Lambda^k+\alpha\,\Delta\,\Lambda,\\
\mathbf{y}^{k+1} &= \mathbf{y}^k+\beta\,\Delta\,\mathbf{y},\\
\Omega^{k+1} &= \Omega^k+\beta\,\Delta\,\Omega,\\
Q^{k+1} &= Q^k\left(I+\frac{1}{2}\gamma S\right)\left(I-\frac{1}{2}\gamma S\right)^{-1}.
\end{aligned}$$

4. $k\leftarrow k+1$.

End

We use Cayley transformation here. Updating of orthogonal matrices through *exp* can be stated in a similar way; and the later analysis can also be carried over with slight modifications of constants.

Note that it doesn't require too much work to calculate the Cayley transformation or the exponential mapping from (13) or (12). When the dimension of $\mathbf{x}_i$ is two, $\forall\, k\geq 1$, we set $Q_i^k=I$ and $S_i^k=0$.

Let $\hat{\alpha}^k$ be the maximum of $\tilde{\alpha}\in[0,1]$, so that for any $\alpha\in[0,\tilde{\alpha}]$.

$$\begin{aligned}
\left(\boldsymbol{\lambda}^k+\alpha\,\Delta\,\boldsymbol{\lambda},\boldsymbol{\omega}^k+\alpha\,\Delta\,\boldsymbol{\omega},\mathbf{y}^k+\alpha\,\Delta\,\mathbf{y},Q^k(I+\frac{\alpha}{2}S)(I-\frac{\alpha}{2}S)^{-1}\right)\in\mathcal{N},\\
(\boldsymbol{\lambda}^k+\alpha\,\Delta\,\boldsymbol{\lambda})^T(\boldsymbol{\omega}^k+\alpha\,\Delta\,\boldsymbol{\omega})\leq[1-\alpha(1-\sigma_2)]\,{\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k,
\end{aligned}$$

The step sizes $\alpha\in(0,1]$, $\beta\in(0,1]$, $\gamma\in(0,1]$ are chosen so that

$$\begin{aligned}
(\boldsymbol{\lambda}^{k+1},\boldsymbol{\omega}^{k+1},\mathbf{y}^{k+1},Q^{k+1})\in\mathcal{N}(\gamma_c,\gamma_p,\gamma_d),\\
{\boldsymbol{\lambda}^{k+1}}^T\boldsymbol{\omega}^{k+1}\leq\left[1-\hat{\alpha}^k(1-\sigma_3)\right]{\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k.
\end{aligned}$$

Because $\sigma_1<\sigma_2<\sigma_3$, the primal and dual step sizes are not necessarily the same.

4.2 Convergence Analysis

The global convergence of the preceding algorithm can also be proved by contradiction as [10].

Theorem 4.1 *If Algorithm 1 doesn't stop after finite steps, the smallest singular value of the Jacobian of* (15) *must converge to zero.*

Proof: The key to the proof is to show that the step sizes are bounded below. Assume the algorithm doesn't stop after finite steps. Let $\epsilon^* \stackrel{\text{def}}{=} \min(\epsilon_c, \gamma_p\epsilon_p, \gamma_d\epsilon_d)$. Then for each iteration k,

$$\boldsymbol{\lambda}^{k^T}\boldsymbol{\omega}^k \geq \epsilon^*, \quad \text{and } \left\|(\boldsymbol{\lambda}^k, \boldsymbol{\omega}^k)\right\|_1 \leq \Upsilon,$$

because otherwise, the iteration will terminate due to the stopping criteria. Boundedness of $\mathbf{y}^k$ is due to the dual feasible constraint. Also observe that Q^k is orthogonal, and the set of orthogonal matrices is compact. Assume the smallest singular value of (15) doesn't converge to zero. Then there must exist a positive scalar d, and a subsequence $\{(\boldsymbol{\lambda}^{m_i}, \boldsymbol{\omega}^{m_i}, \mathbf{y}^{m_i}, Q^{m_i})\}_{i=1}^{\infty}$ such that for all m_i, the largest singular value of the inverse of (15) to zero is at most $\frac{1}{d}$.

Both the right hand side and the left hand side of (15) depend continuously on the iterate $(\boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{y}, Q)$, which is in a compact set; so, the Newton's direction of (15) is a continuous function of $(\boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{y}, Q)$. Therefore, the solution of (15) is uniformly bounded for the subsequence $\{m_i\}$. Hence, there exists a positive constant η, such that the search direction computed by (15) satisfies

$$\left|\Delta(\lambda_i)_j\,\Delta(\omega_i)_j - \frac{\gamma_c}{2n}\,\Delta\,\boldsymbol{\lambda}^T\,\Delta\,\boldsymbol{\omega}\right| \leq \eta, \quad \left|\Delta\,\boldsymbol{\lambda}^T\,\Delta\,\boldsymbol{\omega}\right| \leq \eta, \quad \|\Delta\,\boldsymbol{\lambda}\|_2 \leq \eta,$$
$$\|\Delta\,\boldsymbol{\omega}\|_2 \leq \eta, \quad \|\mathbf{s}_i\|_2^2 \leq \eta, \quad \text{for } i = 1, \ldots, n;\ j = 1, 2.$$

Note that $\|S_i\|_2 = \|\mathbf{s}_i\|_2$ (for $i = 1, \ldots, n$), $\|S\|_2 = \max_i \|\mathbf{s}_i\|_2$.

For $k \in \{m_i\}_{i=1}^{\infty}$, following the notations of [10], we define

$$\begin{aligned}
f_{ij}(\alpha) &\stackrel{\text{def}}{=} \left[(\lambda_i^k)_j + \alpha(\Delta\,\lambda_i)_j\right]\left[(\omega_i^k)_j + \alpha(\Delta\,\omega_i)_j\right] - \frac{\gamma_c}{2n}(\boldsymbol{\lambda}^k + \alpha\,\Delta\,\boldsymbol{\lambda})^T(\boldsymbol{\omega}^k + \alpha\,\Delta\,\boldsymbol{\omega}),\\
g_p(\alpha) &\stackrel{\text{def}}{=} (\boldsymbol{\lambda}^k + \alpha\,\Delta\,\boldsymbol{\lambda})^T(\boldsymbol{\omega}^k + \alpha\,\Delta\,\boldsymbol{\omega})\\
&\quad - \gamma_p\left\|AQ^k(I + \frac{\alpha}{2}S)(I - \frac{\alpha}{2}S)^{-1}\tilde{P}(\boldsymbol{\lambda}^k + \alpha\,\Delta\,\boldsymbol{\lambda}) - \mathbf{b}\right\|_2,\\
g_d(\alpha) &\stackrel{\text{def}}{=} (\boldsymbol{\lambda}^k + \alpha\,\Delta\,\boldsymbol{\lambda})^T(\boldsymbol{\omega}^k + \alpha\,\Delta\,\boldsymbol{\omega})\\
&\quad - \gamma_d\left\|A^T(\mathbf{y}^k + \Delta\,\mathbf{y}) + Q^k(I + \frac{\alpha}{2}S)(I - \frac{\alpha}{2}S)^{-1}\tilde{P}(\boldsymbol{\omega}^k + \alpha\,\Delta\,\boldsymbol{\omega}) - \mathbf{c}\right\|_2,\\
h(\alpha) &\stackrel{\text{def}}{=} [1 - \alpha(1 - \sigma_2)]\,\boldsymbol{\lambda}^{k^T}\boldsymbol{\omega}^k - (\boldsymbol{\lambda}^k + \alpha\,\Delta\,\boldsymbol{\lambda})^T(\boldsymbol{\omega}^k + \alpha\,\Delta\,\boldsymbol{\omega}).
\end{aligned}$$

Therefore, $\hat{\alpha}^k$ is determined by the following inequalities:

$$\begin{aligned}
f_{ij}(\alpha) &\geq 0 \quad (i = 1, \ldots, n;\ j = 1, 2),\\
g_p(\alpha) &\geq 0 \text{ or } \left\|AQ^k\tilde{P}\boldsymbol{\lambda}^k - \mathbf{b}\right\|_2 \leq \epsilon_p,\\
g_d(\alpha) &\geq 0 \text{ or } \left\|A^T\mathbf{y}^k + Q^k\tilde{P}\boldsymbol{\omega}^k - \mathbf{c}\right\|_2 \leq \epsilon_d,\\
h(\alpha) &\geq 0.
\end{aligned}$$

Next, we will show that there is a lower bound for each $\hat{\alpha}^k$.

Each block of the Cayley transformation is equivalent to

$$(I + \frac{\alpha}{2}S_i)(I - \frac{\alpha}{2}S_i)^{-1} = I + \alpha S_i - \frac{\alpha^3\,\|\mathbf{s}_i\|_2^2}{4 + \alpha^2\,\|\mathbf{s}_i\|_2^2}S_i + \frac{2\alpha^2}{4 + \alpha^2\,\|\mathbf{s}_i\|_2^2}S_i^2. \tag{24}$$

The inequalities for f_{ij} and h are obtained by the similar arguments as those in [10].

$$\begin{aligned}
f_{ij}(\alpha) &\geq \sigma_1\frac{\epsilon^*}{2n}(1 - \gamma_c)\alpha - \eta\alpha^2,\\
h(\alpha) &\geq (\sigma_2 - \sigma_1)\epsilon^*\alpha - \eta\alpha^2.
\end{aligned}$$

Next, we will estimate $g_p(\alpha)$ and $g_d(\alpha)$. Note that the first column of S_i^2 is zero; and the only nonzero entry of its second column is $-\mathbf{s}_i^T\mathbf{s}_i$. Let Q_2^k denote the matrix consisting of only the $2nd$ column of each block of Q^k, $\boldsymbol{\lambda}_1$ be the vector of all the first eigenvalues of $\mathbf{x}_i$, $\boldsymbol{\lambda}_2$ be the vector of all the second eigenvalues of $\mathbf{x}_i$ $(i = 1, \ldots, n)$.

When ${\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k \geq \gamma_p \left\| AQ^k\tilde{P}\boldsymbol{\lambda}^k - \mathbf{b}\right\|_2$,

$$
\begin{aligned}
(25)\quad g_p(\alpha) &\geq (1-\alpha){\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k + \alpha\sigma_1{\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k + \alpha^2\,\Delta\boldsymbol{\lambda}^T\,\Delta\boldsymbol{\omega} - \gamma_p(1-\alpha)\left\|AQ^k\tilde{P}\boldsymbol{\lambda}^k - \mathbf{b}\right\|_2 \\
&\quad - \gamma_p\alpha^2\left[\left\|AQ^kS\tilde{P}\,\Delta\boldsymbol{\lambda}\right\|_2 + \frac{\max_i\|\mathbf{s}_i\|_2^2}{4}\left\|AQ_2^k(\boldsymbol{\lambda}_1^k - \boldsymbol{\lambda}_2^k + \alpha\,\Delta\boldsymbol{\lambda}_1 - \alpha\,\Delta\boldsymbol{\lambda}_2)\right\|_2\right. \\
&\qquad \left. + \frac{\alpha\max_i\|\mathbf{s}_i\|_2^2}{4}\left\|AQ^kS\tilde{P}(\boldsymbol{\lambda}^k + \alpha\,\Delta\boldsymbol{\lambda})\right\|_2\right] \\
&\geq \alpha\sigma_1\epsilon^* - \alpha^2\eta - \gamma_p\alpha^2\|A\|_2\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right).
\end{aligned}
$$

The first inequality is due to the Newton system of search directions from (15), ${\boldsymbol{\lambda}^k}^T\boldsymbol{\omega} \geq \gamma_p\left\|AQ^k\tilde{P}\boldsymbol{\lambda} - \mathbf{b}\right\|_2$, and the expansion of Cayley transformation of (24); the second inequality is because of the bound on the variables and search directions, and $\alpha \leq 1$, also the fact

$$\left\|\tilde{P}\boldsymbol{\lambda}\right\|_2 = \frac{1}{\sqrt{2}}\|\boldsymbol{\lambda}\|_2 \leq \frac{1}{\sqrt{2}}\|\boldsymbol{\lambda}\|_1.$$

If $\left\|AQ^k\tilde{P}\boldsymbol{\lambda}^k - \mathbf{b}\right\|_2 \leq \epsilon_p$, then

$$
\begin{aligned}
(26)\quad &\left\|AQ^k(I + \frac{\alpha}{2}S)(I - \frac{\alpha}{2}S)^{-1}\tilde{P}(\boldsymbol{\lambda}^k + \alpha\,\Delta\boldsymbol{\lambda}) - \mathbf{b}\right\|_2 \leq (1-\alpha)\left\|AQ^k\tilde{P}\boldsymbol{\lambda}^k - \mathbf{b}\right\|_2 \\
&\quad + \alpha^2\left[\left\|AQ^kS\tilde{P}\,\Delta\boldsymbol{\lambda}\right\|_2 + \frac{\max_i\|\mathbf{s}_i\|_2^2}{4}\left\|AQ_2^k(\boldsymbol{\lambda}_1^k - \boldsymbol{\lambda}_2^k + \alpha\,\Delta\boldsymbol{\lambda}_1 - \alpha\,\Delta\boldsymbol{\lambda}_2)\right\|_2\right. \\
&\qquad \left. + \frac{\alpha\max_i\|\mathbf{s}_i\|_2^2}{4}\left\|AQ^kS\tilde{P}(\boldsymbol{\lambda}^k + \alpha\,\Delta\boldsymbol{\lambda}^k)\right\|_2\right] \\
&\quad \leq (1-\alpha)\epsilon_p + \alpha^2\|A\|_2\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right).
\end{aligned}
$$

So when

$$\alpha \leq \frac{\epsilon_p}{\|A\|_2\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right)},$$

$\left\|AQ^{k+1}\tilde{P}\boldsymbol{\lambda}^{k+1} - \mathbf{b}\right\|_2 \leq \epsilon_p$.

Next, we will consider the dual constraints.
When ${\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k \geq \gamma_d \left\| A^T\mathbf{y}^k + Q^k\tilde{P}\boldsymbol{\omega}^k - \mathbf{c}\right\|_2$,

$$
\begin{aligned}
(27)\quad g_d(\alpha) &\geq (1-\alpha){\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k + \alpha\sigma_1{\boldsymbol{\lambda}^k}^T\boldsymbol{\omega}^k + \alpha^2\,\Delta\boldsymbol{\lambda}^T\,\Delta\boldsymbol{\omega} \\
&\quad - \gamma_d(1-\alpha)\left\| A^T\mathbf{y}^k + Q^k\tilde{P}\boldsymbol{\omega}^k - \mathbf{c}\right\|_2 - \gamma_d\alpha^2\Bigg[\left\|Q^kS\tilde{P}\,\Delta\boldsymbol{\omega}\right\|_2 \\
&\quad + \frac{2}{4}\left\|Q^kS^2\tilde{P}(\boldsymbol{\omega}^k+\alpha\,\Delta\boldsymbol{\omega})\right\|_2 + \frac{\alpha\max_i\|\mathbf{s}_i\|_2^2}{4}\left\|Q^kS\tilde{P}(\boldsymbol{\omega}^k+\alpha\,\Delta\boldsymbol{\omega})\right\|_2\Bigg] \\
&\geq \alpha\sigma_1\epsilon^* - \alpha^2\eta - \gamma_d\alpha^2\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right).
\end{aligned}
$$

When $\left\|A^T\mathbf{y}^k + Q^k\tilde{P}\boldsymbol{\omega}^k - \mathbf{c}\right\|_2 \leq \epsilon_d$,

$$
\begin{aligned}
(28)\quad &\left\|A^T(\mathbf{y}^k+\alpha\,\Delta\mathbf{y}) + Q^k(I+\frac{\alpha}{2}S)(I-\frac{\alpha}{2}S)^{-1}\tilde{P}(\boldsymbol{\omega}^k+\alpha\,\Delta\boldsymbol{\omega}) - \mathbf{c}\right\|_2 \\
&\leq (1-\alpha)\left\|A^T\mathbf{y}^k + Q^k\tilde{P}\boldsymbol{\omega}^k - \mathbf{c}\right\|_2 + \alpha^2\Bigg[\left\|Q^kS\tilde{P}\,\Delta\boldsymbol{\omega}\right\|_2 \\
&\quad + \frac{2}{4}\left\|Q^kS^2\tilde{P}(\boldsymbol{\omega}^k+\alpha\,\Delta\boldsymbol{\omega})\right\|_2 + \alpha\frac{\max_i\|\mathbf{s}_i\|_2^2}{4}\left\|Q^kS\tilde{P}(\boldsymbol{\omega}^k+\alpha\,\Delta\boldsymbol{\omega})\right\|_2\Bigg] \\
&\leq (1-\alpha)\epsilon_d + \alpha^2\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right).
\end{aligned}
$$

Thus, the lower bound on $\hat{\alpha}$'s is

$$
\begin{aligned}
\alpha^* = \min\Bigg\{1,\ &\frac{(\sigma_2-\sigma_1)\epsilon^*}{\eta}, \frac{\sigma_1(1-\gamma_c)\epsilon^*}{2n\eta}, \\
&\frac{\sigma_1\epsilon^*}{\eta + \gamma_p\|A\|_2\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right)}, \\
&\frac{\epsilon_p}{\|A\|_2\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right)}, \\
&\frac{\sigma_1\epsilon^*}{\eta + \gamma_d\left(\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}\right)}, \\
&\frac{\epsilon_d}{\frac{1}{\sqrt{2}}\eta^{3/2} + \frac{1}{4}\eta\Upsilon + \frac{1}{2\sqrt{2}}\eta^2 + \frac{1}{4\sqrt{2}}\eta^{3/2}\Upsilon + \frac{1}{4\sqrt{2}}\eta^{5/2}}\Bigg\}.
\end{aligned}
$$

After the perturbations of step sizes to ensure $\lambda_1 > \lambda_2$ and $\omega_2 > \omega_1$, the lower bound on $\hat{\alpha}^k$ is at least $\frac{\alpha^*}{2}$. The algorithm imposes the decrease of the sequence $\left\{{\boldsymbol{\lambda}^j}^T\boldsymbol{\omega}^j\right\}_{j=1}^{\infty}$. So for each m_i in the subsequence, by $h(\alpha) \geq 0$, we see

$$
\begin{aligned}
&{\boldsymbol{\lambda}^{m_i+1}}^T\boldsymbol{\omega}^{m_i+1} \\
&\leq \left[1 - \frac{\alpha^*}{2}(1-\sigma_3)\right]{\boldsymbol{\lambda}^{m_i}}^T\boldsymbol{\omega}^{m_i} \leq \left[1 - \frac{\alpha^*}{2}(1-\sigma_3)\right]{\boldsymbol{\lambda}^{m_{i-1}+1}}^T\boldsymbol{\omega}^{m_{i-1}+1} \\
&\leq \left[1 - \frac{\alpha^*}{2}(1-\sigma_3)\right]^2{\boldsymbol{\lambda}^{m_{i-1}}}^T\boldsymbol{\omega}^{m_{i-1}} \leq \cdots \leq \left[1 - \frac{\alpha^*}{2}(1-\sigma_3)\right]^i{\boldsymbol{\lambda}^{m_1}}^T\boldsymbol{\omega}^{m_1}.
\end{aligned}
$$

That means the whole sequence $\{{\boldsymbol{\lambda}^j}^T\boldsymbol{\omega}^j\}_{j=1}^{\infty}$ converges to 0, which contradicts to the assumption.

We have proved that if the smallest singular value of (15) doesn't converge to zero, either the algorithm finds an $(\epsilon_p, \epsilon_d, \epsilon_c)$-optimal solution in finite iterations, or the iterate is unbounded. ■

5 Finite Convergence

Algorithm may abort due to unboundedness of eigenvalues or singularity of Jacobians. In this section, we will give some conditions under which Algorithm 1 converges to an $(\epsilon_p, \epsilon_d, \epsilon_c)$-optimum in finite iterations.

Conditions ensure boundedness is given in §§ 5.1, while that for nonsingularity is given in §§ 5.2.

5.1 Boundedness of Iterates

To make sure that each iterate is bounded, we use some ideas in [8], which is also for LP, further impose some restrictions on the problem.

Let ρ represent a positive scalar no larger than the smallest singular value of A. Suppose (1) has an interior feasible solution $(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{y}})$. Denote the eigenvalues of $\hat{\mathbf{x}}$ as $\hat{\boldsymbol{\lambda}}$, the eigenvalues of $\hat{\mathbf{z}}$ as $\hat{\boldsymbol{\omega}}$. Assume $\nu_p\mathbf{1} \leq \hat{\boldsymbol{\lambda}} \leq \chi_p\mathbf{1}$, $\nu_d\mathbf{1} \leq \hat{\boldsymbol{\omega}} \leq \chi_d\mathbf{1}$. We require the feasibility constraints to be calculated to a certain accuracy. That is, each iterate satisfies

$$\begin{aligned} Q\tilde{P}\boldsymbol{\omega} + A^T\mathbf{y} &= \mathbf{c} + \tilde{\mathbf{c}} \\ AQ\tilde{P}\boldsymbol{\lambda} &= \mathbf{b} + \tilde{\mathbf{b}} \end{aligned} \tag{29}$$

with $\left\|\tilde{\mathbf{b}}\right\|_2 \leq \frac{1}{2\sqrt{2}}\nu_p\rho$, $\|\tilde{\mathbf{c}}\|_2 \leq \frac{1}{2\sqrt{2}}\nu_d$.
It is shown in [8], that by some transformation, the smallest singular value of a matrix can be larger than 1, which means $\left\|\tilde{\mathbf{b}}\right\|_2$ and $\|\tilde{\mathbf{c}}\|_2$ are not too small.

If $\epsilon_p > \frac{1}{2\sqrt{2}}\rho\nu_p$, we replace ϵ_p with $\frac{1}{2\sqrt{2}}\rho\nu_p$; If $\epsilon_d > \frac{1}{2\sqrt{2}}\nu_d$, we replace ϵ_d with $\frac{1}{2\sqrt{2}}\nu_d$.
We modify the algorithm in §4 so that each iterate is in the neighborhood $\tilde{\mathcal{N}}$.

$$\begin{aligned} \tilde{\mathcal{N}} \stackrel{\text{def}}{=} \Big\{(\boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{y}, Q)\colon\ & \boldsymbol{\lambda} \in \mathbb{R}^{2n}, \boldsymbol{\omega} \in \mathbb{R}^{2n}, \mathbf{y} \in \mathbb{R}^m, Q \in K,\ \boldsymbol{\lambda} > 0,\ \boldsymbol{\omega} > 0; \\ & (\boldsymbol{\lambda}_i)_j(\boldsymbol{\omega}_i)_j \geq \gamma_c\frac{\boldsymbol{\lambda}^T\boldsymbol{\omega}}{2n}\ (j = 1, 2; i = 1, \ldots, n); \\ & \boldsymbol{\lambda}^T\boldsymbol{\omega} \geq \gamma_p\left\|AQ\tilde{P}\boldsymbol{\lambda} - \mathbf{b}\right\|_2 \text{ and } \left\|AQ\tilde{P}\boldsymbol{\lambda} - \mathbf{b}\right\|_2 \leq \frac{1}{2\sqrt{2}}\rho\nu_p, \\ & \text{or } \left\|AQ\tilde{P}\boldsymbol{\lambda} - \mathbf{b}\right\|_2 \leq \epsilon_p; \\ & \boldsymbol{\lambda}^T\boldsymbol{\omega} \geq \gamma_d\left\|A^T\mathbf{y} + Q\tilde{P}\boldsymbol{\omega} - \mathbf{c}\right\|_2 \text{ and } \left\|A^T\mathbf{y} + Q\tilde{P}\boldsymbol{\omega} - \mathbf{c}\right\|_2 \leq \frac{1}{2\sqrt{2}}\nu_d, \\ & \text{or } \left\|A^T\mathbf{y} + Q\tilde{P}\boldsymbol{\omega} - \mathbf{c}\right\|_2 \leq \epsilon_d.\Big\} \end{aligned}$$

Other parts of the algorithm is the same as that in § 4. For further reference, we name the algorithm in this section Algorithm 2.

As the proofs of (26) and (28), we find that when the step size $\hat{\alpha}^k \leq \alpha^{**}$, where α^{**} is defined

as the following,

$$\alpha^{**} \stackrel{\text{def}}{=} \min\left\{ \frac{\nu_p \rho}{\|A\|_2 \left(2\eta^{3/2} + \frac{\sqrt{2}}{2}\eta\Upsilon + \eta^2 + \frac{1}{2}\eta^{3/2}\Upsilon + \frac{1}{2}\eta^{5/2}\right)}, \right.$$
$$\left. \frac{\nu_d}{\left(2\eta^{3/2} + \frac{\sqrt{2}}{2}\eta\Upsilon + \eta^2 + \frac{1}{2}\eta^{3/2}\Upsilon + \frac{1}{2}\eta^{5/2}\right)} \right\},$$

condition (29) is satisfied. Hence, $\hat{\alpha}$ in the algorithm of this section has a lower bound: $\min\{\alpha^*, \alpha^{**}\}$. Thus, by the results in § 4, if the initial point is in $\tilde{\mathcal{N}}$, assume the smallest singular value of each element in $\tilde{\mathcal{N}}$ is distance at least d from 0; then the iterates of Algorithm 2 will converge to a solution of (1) in finite iterations, if each iterate is bounded.

Next, we will use two lemmas to show the boundedness of each iterate. Lemma 5.1 gives the existence of an interior feasible solution, under which Lemma 5.2 guarantees the boundedness.

We consider the perturbed system:

$$\begin{aligned} \mathbf{z} + A^T\mathbf{y} &= \mathbf{c} + \tilde{\mathbf{c}} \\ A\mathbf{x} &= \mathbf{b} + \tilde{\mathbf{b}} \\ \mathbf{x} &\geq_Q \mathbf{0} \\ \mathbf{z} &\geq_Q \mathbf{0}. \end{aligned} \tag{30}$$

Lemma 5.1 *Suppose* (1) *has an interior feasible solution* $(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{y}})$ *with* $\nu_p\mathbf{1} \leq \hat{\boldsymbol{\lambda}} \leq \chi_p\mathbf{1}$, $\nu_d\mathbf{1} \leq \hat{\boldsymbol{\omega}} \leq \chi_d\mathbf{1}$*; then for all* $\left\|\tilde{\mathbf{b}}\right\|_2 \leq \frac{1}{2\sqrt{2}}\nu_p\rho$ *and* $\|\tilde{\mathbf{c}}\|_2 \leq \frac{1}{2\sqrt{2}}\nu_d$, (30) *has a feasible solution* $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\omega}}, \tilde{\mathbf{y}}, \tilde{Q})$ *with* $\frac{1}{2}\nu_p\mathbf{1} \leq \tilde{\boldsymbol{\lambda}} \leq \frac{3}{2}\chi_p\mathbf{1}$, $\frac{1}{2}\nu_d\mathbf{1} \leq \tilde{\boldsymbol{\omega}} \leq \frac{3}{2}\chi_d\mathbf{1}$.

Proof: Let $\mathbf{h} \stackrel{\text{def}}{=} A^+\tilde{\mathbf{b}}$, where A^+ is the Moore-Penrose generalized inverse of A. Denote the decomposition (2) of $\mathbf{h}$ as $\mathbf{h} = Q_h\tilde{P}\boldsymbol{\lambda}_h$; then

$$\|\boldsymbol{\lambda}_h\|_\infty \leq \|\boldsymbol{\lambda}_h\|_2 = \sqrt{2}\,\|\mathbf{h}\|_2 \leq \sqrt{2}\left\|A^+\right\|_2\left\|\tilde{\mathbf{b}}\right\|_2 \leq \frac{1}{2}\nu_p.$$

Let

$$\tilde{\mathbf{x}} \stackrel{\text{def}}{=} \hat{\mathbf{x}} + \mathbf{h}, \quad \tilde{\mathbf{y}} \stackrel{\text{def}}{=} \hat{\mathbf{y}}, \quad \tilde{\mathbf{z}} \stackrel{\text{def}}{=} \hat{\mathbf{z}} + \tilde{\mathbf{c}}.$$

Write the smaller eigenvalue of $\tilde{\mathbf{x}}_i$ as $(\tilde{\lambda}_i)_{\text{small}}$; then

$$(\tilde{\lambda}_i)_{\text{small}} = (\hat{\mathbf{x}}_i)_0 + (\mathbf{h}_i)_0 - \sqrt{\sum_{j=1}^{n_i-1}\left((\hat{\mathbf{x}}_i)_j + (\mathbf{h}_i)_j\right)^2} = (\hat{\mathbf{x}}_i)_0 + (\mathbf{h}_i)_0$$
$$- \sqrt{\left\|\bar{\hat{\mathbf{x}}}_i\right\|_2^2 + \left\|\bar{\mathbf{h}}_i\right\|_2^2 + 2\sum_{j=1}^{n_i-1}\left((\hat{\mathbf{x}}_i)_j(\mathbf{h}_i)_j\right)} \geq (\hat{\mathbf{x}}_i)_0 + (\mathbf{h}_i)_0 - \left\|\bar{\hat{\mathbf{x}}}_i\right\|_2 - \left\|\bar{\mathbf{h}}_i\right\|_2$$
$$\geq (\hat{\lambda}_i)_{\text{small}} - \|(\lambda_h)_i\|_\infty \geq \frac{1}{2}\nu_p.$$

The first inequality above is due to the Cauchy-Schwartz-Bomiakovsky inequality.

Similarly, denote the bigger eigenvalue of $\tilde{\mathbf{x}}_i$ as $(\tilde{\lambda}_i)_{\text{big}}$; then,

$$(\tilde{\lambda}_i)_{\text{big}} = (\hat{\mathbf{x}}_i)_0 + (\mathbf{h}_i)_0 + \sqrt{\sum_{j=1}^{n_i-1}\left((\hat{\mathbf{x}}_i)_j + (\mathbf{h}_i)_j\right)^2} \leq (\hat{\mathbf{x}}_i)_0 + (\mathbf{h}_i)_0 + \left\|\bar{\hat{\mathbf{x}}}_i\right\|_2 + \left\|\bar{\mathbf{h}}_i\right\|_2$$
$$\leq (\hat{\lambda}_i)_{\text{big}} + \|(\lambda_h)_i\|_\infty \leq \frac{3}{2}\chi_p.$$

Thus, $\frac{1}{2}\nu_p\mathbf{1} \leq \tilde{\boldsymbol{\lambda}} \leq \frac{3}{2}\chi_p\mathbf{1}$.
The inequalities $\frac{1}{2}\nu_d\mathbf{1} \leq \tilde{\boldsymbol{\omega}} \leq \frac{3}{2}\chi_d\mathbf{1}$ can be proved in a same way. ■

Lemma 5.2 *If* (1) *has an interior feasible solution* $(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{y}})$ *with* $\nu_p\mathbf{1} \leq \hat{\boldsymbol{\lambda}} \leq \chi_p\mathbf{1}$, $\nu_d\mathbf{1} \leq \hat{\boldsymbol{\omega}} \leq \chi_d\mathbf{1}$*; then there exists a positive scalar* Γ*, such that for any iterate* $(\boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{y}, Q) \in \tilde{\mathcal{N}}$, $\nu_d \|\boldsymbol{\lambda}\|_1 + \nu_p \|\boldsymbol{\omega}\|_1 \leq \Gamma$.

Proof: Given an iterate $(\boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{y}, Q) \in \tilde{\mathcal{N}}$, there exists $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{b}}$, so that $(\boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{y}, Q)$ is a solution to the following system of equations.

$$Q\tilde{P}\boldsymbol{\omega} + A^T\mathbf{y} = \mathbf{c} + \tilde{\mathbf{c}}$$
$$AQ\tilde{P}\boldsymbol{\lambda} = \mathbf{b} + \tilde{\mathbf{b}}$$

Then, according to Lemma 5.1, there exists $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\omega}}, \tilde{\mathbf{y}}, \tilde{Q})$ satisfying the above perturbed constraints with $\frac{1}{2}\nu_p\mathbf{1} \leq \tilde{\boldsymbol{\lambda}} \leq \frac{3}{2}\chi_p\mathbf{1}$, $\frac{1}{2}\nu_d\mathbf{1} \leq \tilde{\boldsymbol{\omega}} \leq \frac{3}{2}\chi_d\mathbf{1}$; so

$$A(\mathbf{x} - \tilde{\mathbf{x}}) = \mathbf{0}, \quad A^T(\mathbf{y} - \tilde{\mathbf{y}}) + \mathbf{z} - \tilde{\mathbf{z}} = \mathbf{0}.$$

Hence,

$$(\mathbf{x} - \tilde{\mathbf{x}})^T(\mathbf{z} - \tilde{\mathbf{z}}) = -(\mathbf{x} - \tilde{\mathbf{x}})^T A^T(\mathbf{y} - \tilde{\mathbf{y}}) = 0.$$

Therefore,

$$\begin{aligned}(31)\quad \mathbf{x}^T\mathbf{z} + \tilde{\mathbf{x}}^T\tilde{\mathbf{z}} = \mathbf{x}^T\tilde{\mathbf{z}} + \tilde{\mathbf{x}}^T\mathbf{z} &\geq \sum_{i=1}^n \left[(\mathbf{x}_i)_0(\tilde{\mathbf{z}}_i)_0 - \|\bar{\mathbf{x}}_i\|_2 \left\|\bar{\tilde{\mathbf{z}}}_i\right\|_2\right] \\ &\quad + \sum_{i=1}^n \left[(\tilde{\mathbf{x}}_i)_0(\mathbf{z}_i)_0 - \left\|\bar{\tilde{\mathbf{x}}}_i\right\|_2 \|\bar{\mathbf{z}}_i\|_2\right] \\ &\geq \sum_{i=1}^n \frac{(\mathbf{x}_i)_0 - \|\bar{\mathbf{x}}_i\|_2 + (\mathbf{x}_i)_0 + \|\bar{\mathbf{x}}_i\|_2}{2}\left((\tilde{\mathbf{z}}_i)_0 - \left\|\bar{\tilde{\mathbf{z}}}_i\right\|_2\right) \\ + \sum_{i=1}^n \frac{(\mathbf{z}_i)_0 - \|\bar{\mathbf{z}}_i\|_2 + (\mathbf{z}_i)_0 + \|\bar{\mathbf{z}}_i\|_2}{2}\left((\tilde{\mathbf{x}}_i)_0 - \left\|\bar{\tilde{\mathbf{x}}}_i\right\|_2\right) &\geq \frac{1}{4}\|\boldsymbol{\lambda}\|_1\nu_d + \frac{1}{4}\|\boldsymbol{\omega}\|_1\nu_p.\end{aligned}$$

The first inequality is due to $(x_i)_0 \geq 0$, $(\tilde{z}_i)_0 \geq 0$, $(\tilde{x}_i)_0 \geq 0$, $(z_i)_0 \geq 0$, and Cauchy-Schwartz-Bomiakovsky inequality. The second one is because of $(x_i)_0 \geq \|\bar{\mathbf{x}}_i\|_2$, $(z_i)_0 \geq \|\bar{\mathbf{z}}_i\|_2$. The last one is obtained by the eigenvalue representations of the second-order cone (2), and the lower bounds on $\tilde{\boldsymbol{\lambda}}$ and $\tilde{\boldsymbol{\omega}}$.
We also have

$$\begin{aligned}(32)\quad \mathbf{x}^T\mathbf{z} + \tilde{\mathbf{x}}^T\tilde{\mathbf{z}} = \boldsymbol{\lambda}^T\tilde{P}^T\tilde{P}\boldsymbol{\omega} + \tilde{\mathbf{x}}^T\tilde{\mathbf{z}} &\leq \frac{1}{2}\boldsymbol{\lambda}^T\boldsymbol{\omega} + \sum_{i=1}^n \left((\tilde{\mathbf{x}}_i)_0(\tilde{\mathbf{z}}_i)_0 + \left\|\bar{\tilde{\mathbf{x}}}_i\right\|_2 \left\|\bar{\tilde{\mathbf{z}}}_i\right\|_2\right) \\ &\leq \frac{1}{2}\boldsymbol{\lambda}^{0T}\boldsymbol{\omega}^0 + \sum_{i=1}^n \left((\tilde{\mathbf{x}}_i)_0 + \left\|\bar{\tilde{\mathbf{x}}}_i\right\|_2\right)\left(\tilde{\mathbf{z}}_0 + \left\|\bar{\tilde{\mathbf{z}}}_i\right\|_2\right) \leq \frac{1}{2}\boldsymbol{\lambda}^{0T}\boldsymbol{\omega}^0 + \frac{9}{4}n\chi_p\chi_d\end{aligned}$$

We use Cauchy-Schwartz-Bomiakovsky inequality to get the first and second inequalities in the above. The second inequality is also from $\boldsymbol{\lambda}^{0T}\boldsymbol{\omega}^0 \geq \boldsymbol{\lambda}^{kT}\boldsymbol{\omega}^k$, which is forced by the algorithm. Combining (31) and (32), we obtain

$$\|\boldsymbol{\lambda}\|_1\nu_d + \|\boldsymbol{\omega}\|_1\nu_p \leq 2\boldsymbol{\lambda}^{0T}\boldsymbol{\omega}^0 + 9n\chi_p\chi_d.$$

■

We have proved that Algorithm 2 will terminate at an $(\epsilon_c, \epsilon_p, \epsilon_d)$ solution to (1) in finite iterations, provided that the smallest singular value of the Jacobian of the Newton's system doesn't converge to zero.

5.2 Nonsingularity of Iterates

In this subsection, we will give some conditions under which the smallest singular value of Jacobian doesn't converge to zero.

Each iterate satisfies a system of equations in the following form.

$$\begin{aligned} Q\tilde{P}\boldsymbol{\omega} + A^T\mathbf{y} &= \mathbf{c} + \mathbf{r}_d, \\ AQ\tilde{P}\boldsymbol{\lambda} &= \mathbf{b} + \mathbf{r}_p, \\ \Lambda\boldsymbol{\omega} &= \mu\mathbf{1} + \mathbf{r}_c. \end{aligned} \tag{33}$$

The algorithm ensures $\lambda_1 \neq \lambda_2$; hence $\|\bar{\mathbf{x}}\|_2 \neq 0$. Given $(\mathbf{x}, \mathbf{y}, \mathbf{z})$, because only the first two columns of Q contribute to (33), the left hand side of (33) is the same for any decomposition of $\mathbf{x}$. Keeping only the first two columns of Q_i, we see each iterate is a solution of the following system.

$$\begin{aligned} \left(\begin{smallmatrix} 1 & 0 \\ \mathbf{0} & \mathbf{q} \end{smallmatrix}\right) \tilde{P}\boldsymbol{\omega} + A^T\mathbf{y} &= \mathbf{c} + \mathbf{r}_d, \\ A\left(\begin{smallmatrix} 1 & 0 \\ \mathbf{0} & \mathbf{q} \end{smallmatrix}\right) \tilde{P}\boldsymbol{\lambda} &= \mathbf{b} + \mathbf{r}_p, \\ \Lambda\boldsymbol{\omega} &= \mu\mathbf{1} + \mathbf{r}_c, \\ \|\mathbf{q}\|_2 &= 1, \\ (\lambda_i)_1 > (\lambda_i)_2,\ (\omega_i)_2 > (\omega_i)_1 &\quad (i = 1, \ldots, n). \end{aligned} \tag{34}$$

Lemma 5.3 *For each triple* $(\mathbf{r}_p, \mathbf{r}_d, \mu\mathbf{1} + \mathbf{r}_c)$ *with* $\mu\mathbf{1} + \mathbf{r}_c > \mathbf{0}$*, if* (34) *has a finite solution, it is unique.*

Proof: Consider the constrained minimization problem:

$$\begin{array}{ll} \min_{\mathbf{x}} & (\mathbf{c} + \mathbf{r}_d)^T\mathbf{x} - \sum_{i=1}^n (\mu + ((r_c)_i)_1) \ln\left((x_i)_0 + \|\bar{\mathbf{x}}_i\|_2\right) \\ & - \sum_{i=1}^n (\mu + ((r_c)_i)_2) \ln\left((x_i)_0 - \|\bar{\mathbf{x}}_i\|_2\right) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} + \mathbf{r}_p. \end{array} \tag{35}$$

Since the Hessian of the objective function is positive definite, the objective is strictly convex; so for each $(\mathbf{r}_p, \mathbf{r}_d, \mathbf{r}_c)$, if (35) has a finite solution, it is unique. The Lagrangian of (35) is

$$\begin{aligned} L = (\mathbf{c} + \mathbf{r}_d)^T\mathbf{x} &- \sum_{i=1}^n \frac{\mu + ((r_c)_i)_1}{2} \ln\left((x_i)_0 + \|\bar{\mathbf{x}}_i\|_2\right) \\ &- \sum_{i=1}^n \frac{\mu + ((r_c)_i)_2}{2} \ln\left((x_i)_0 - \|\bar{\mathbf{x}}_i\|_2\right) - \mathbf{y}^T\left(A\mathbf{x} - \mathbf{b} - \mathbf{r}_p\right). \end{aligned}$$

Notice (35) has only linear constraints, and A has full row rank. So the solution to $\nabla L = \mathbf{0}$ is the same thing as the solution to (35). The logarithmic terms force $\mathbf{x}_i$ to be in the interior of the second-order cone. So we can set

$$\mathbf{z}_i = \frac{\mu + ((r_c)_i)_1}{2\left((x_i)_0 + \|\bar{\mathbf{x}}\|_2\right)} \begin{pmatrix} 1 \\ \frac{\bar{\mathbf{x}}_i}{\|\bar{\mathbf{x}}\|_2} \end{pmatrix} + \frac{\mu + ((r_c)_i)_2}{2\left((x_i)_0 - \|\bar{\mathbf{x}}\|_2\right)} \begin{pmatrix} 1 \\ -\frac{\bar{\mathbf{x}}_i}{\|\bar{\mathbf{x}}\|_2} \end{pmatrix},$$

and get the system

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} + \mathbf{r}_p \\ A^T\mathbf{y} + \mathbf{z} &= \mathbf{c} + \mathbf{r}_d \end{aligned}$$

has a unique solution, because it is just $\nabla L = \mathbf{0}$.

Given $\mathbf{x} \in \mathbb{R}^{n+1}$ with $\bar{\mathbf{x}} \neq \mathbf{0}$, the decomposition

$$\mathbf{x} = \frac{\lambda_1}{2} \begin{pmatrix} 1 \\ \mathbf{q} \end{pmatrix} + \frac{\lambda_2}{2} \begin{pmatrix} 1 \\ -\mathbf{q} \end{pmatrix}$$

with $\lambda_1 \geq \lambda_2$, $\|\mathbf{q}\|_2 = 1$ is unique if we assume $\mathbf{q} = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}$ for $\bar{\mathbf{x}} = \mathbf{0}$. This can be seen by directly solving the above equation for λ_1, λ_2 and $\mathbf{q}$: $\lambda_1 = x_0 + \|\bar{\mathbf{x}}\|_2$, $\lambda_2 = x_0 - \|\bar{\mathbf{x}}\|_2$.

The lemma is proved by letting

$$(\omega_i)_1 = \frac{\mu + ((r_c)_i)_1}{(x_i)_0 + \|\bar{\mathbf{x}}\|_2}, \quad (\omega_i)_2 = \frac{\mu + ((r_c)_i)_2}{(x_i)_0 - \|\bar{\mathbf{x}}\|_2}.$$

■

For briefness, we denote $\mathbf{w} \stackrel{\text{def}}{=} (\boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{y}, Q)$, and use G to represent the left hand side of (15).

Lemma 5.4 *Let $\mathbf{w}^*$ be a solution to* (1) *satisfying the conditions of Theorem 3.1. Then there are positive numbers δ, and ζ, such that if ${\boldsymbol{\lambda}^0}^T \boldsymbol{\omega}^0 \leq \zeta$, then $\partial G(\mathbf{w}^k)$ is distance at least δ from* 0 *for $k = 0, 1, \ldots$, where $\mathbf{w}^k$ is generated by the Algorithm.*

Proof: By Theorem 3.1, $\partial G(\mathbf{w}^*)$ is nonsingular. Let B denote the open unit ball. Since G is Lipschitz continuous, by implicit function theorem, there exist positive numbers δ and r, such that for any $\mathbf{w} \in \mathbf{w}^* + rB$, the smallest singular value of $\partial G(\mathbf{w})$ is at least distance δ from 0, and $G(\mathbf{w}^* + rB)$ contains $G(\mathbf{w}^*) + r\delta B$. Suppose $\mathbf{r}_p^k > \epsilon_p$, $\mathbf{r}_d^k > \epsilon_d$. By the definition of the algorithm, ${\boldsymbol{\lambda}^k}^T \boldsymbol{\omega}^k$ is decreasing with k, $\left\|\mathbf{r}_p^k\right\|_2 \leq \frac{1}{\gamma_p} {\boldsymbol{\lambda}^k}^T \boldsymbol{\omega}^k$, $\left\|\mathbf{r}_d^k\right\|_2 \leq \frac{1}{\gamma_d} {\boldsymbol{\lambda}^k}^T \boldsymbol{\omega}^k$. Hence, if we assume $\max\left(1, \frac{1}{\gamma_p}, \frac{1}{\gamma_d}\right) {\boldsymbol{\lambda}^0}^T \boldsymbol{\omega}^0 \leq r\delta$, then $G(\mathbf{w}^k) \in G(\mathbf{w}^*) + r\delta B$ for $k = 0, 1, \ldots$. By Lemma 5.3 and the relationship between (33) and (34), we get $\mathbf{w}^k$ must be in $\mathbf{w}^* + rB$; therefore, the smallest singular value of $\partial G(\mathbf{w}^k)$ is distance at least δ from 0 for $k = 0, 1, \ldots$. Observe the assumption, $\mathbf{r}_p^k > \epsilon_p$ or $\mathbf{r}_d^k > \epsilon_d$ is not necessary in the above proof. ■

Combining Lemma 5.2 and Lemma 5.4, we have the following theorem.

Theorem 5.1 *Under the conditions of Theorem 3.1 and Lemma 5.2, there is a positive number ζ, such that if ${\boldsymbol{\lambda}^0}^T \boldsymbol{\omega}^0 \leq \zeta$, Algorithm 2 converges to an $(\epsilon_p, \epsilon_d, \epsilon_c)$-solution of* (1) *in finite steps.*

6 Numerical Results

To test the Q method, we have implemented the basic algorithm in MATLAB. Below are the results of our test on randomly generated 1,000 problems with known solutions. For the step sizes, simply, we choose $\alpha = \min(1, \tau\alpha')$, $\beta = \min(1, \tau\beta')$, $\gamma = \sqrt{\alpha\beta}$, where α' and β' are the maximum stepsizes to the boundary of the second-order cone.

We used $\mathbf{x}_i = (2; 1; \mathbf{0})$, $\mathbf{s}_i = (2; -1; \mathbf{0})$, $\mathbf{y} = \mathbf{0}$ as starting point. We picked $\sigma = 0.25$, $\tau = 0.99$, which may not be the best choice of parameters. Our code reduced the l_2 norm of primal infeasibility, l_2 norm of dual infeasibility, and l_1 norm of duality gap to less than $5.0e-12$ for all the problems. The range of every element in our randomly generated problem is $(-0.5, 0.5)$; therefore, we didn't use relative measurement for accuracy, as done by other algorithms. Note that our accuracy requirement is much more stringent than most other algorithms. Below is the results.

bk	dimension of each block	type of each block	m	r_p0	r_d0	it
10	[2,2,2,2,2,2,2,2,2,2]	[b,i,o,b,i,b,o,i,i,b]	12	342.20	45.59	27.07
10	[10,10,10,10,10,10,10,10, 10,10]	[b,o,i,b,b,i,o,b,b, o]	30	299.69	142.30	34.16
10	[3,10,8,9,12,4,6,3,14,8]	[b,i,o,b,i,o,i,i,b,o]	45	539.07	146.97	31.46
10	[20,10,8,9,12,15,6,3,14,8]	[b,i,b,i,i,o,b,i,b,o]	55	861.28	190.32	33.31
10	[20,15,15,15,15,15,15,15, 15,15]	[b,i,b,i,i,o,b,i,b,o]	75	1331.71	269.07	32.16
12	[10,10,10,10,10,10,10,10, 10,10,10,10]	[b,o,i,b,b,i,o,b,b,o, b,i]	50	420.43	197.80	31.96
15	[10,10,10,10,10,10,10,10 10, 10,10,10,10,10,10]	[b,o,i,b,b,i,o,b,b,o, b,o,i,i,o]	70	558.19	262.10	32.46
15	[15,15,15,15,15,15,15,15, 15,15,15,15,15,15,]	[i,o,b,i,i,b,o,i,b,b, i,o,b,b,o]	100	1748.47	375.81	33.46
20	[10,20,13,20,24,20,3,8,26, 30,9,12,21,3,11,23,5,2,20, 18]	[b,o,i,b,b,i,o,b,b,o, b,b,i,o,i,b,b,b,i,b]	130	1478.57	496.35	31.97
20	[20,20,20,20,20,20,20,20 20,20,20,20,20,20,20,20, 20,20,20,20]	[b,o,i,b,b,i,o,b,b,o, b,b,i,o,i,b,b,b]	130	1348.60	572.31	33.94

In the above table, each row is the summary of 100 instances of problem with the same number of blocks, dimension of each block, optimum variable type, and number of constraints.
"bk" represents the number of blocks;
"type of each block" shows at optimum, whether each block is in the boundary(b), zero(o), or in the interior(i);
"m" is the number of constraints;
"r_p0" is the average l_2 norm of initial primal infeasibility for the 100 instances;
"r_d0" is the average l_2 norm of initial dual infeasibility for the 100 instances;
"it" is the average number of iterations for the 100 instances.

All the instances were terminated at ϵ solutions within 50 iterations, which shows our algorithm is indeed stable and can get high accuracy. The first row shows our algorithm can solve LP problems, since 2-dimensional SOCP is just LP [5].
Notice that the problem type and size have little effect on the total number of iterations, which is a property of interior point methods for SOCP.

Following is a typical instance of the $2nd$ type of problem. We use "gap" to represents the duality gap.

it	r_p	r_d	gap
0	2.652511e+002	1.417364e+002	6.000000e+001
1	2.621261e+002	9.163063e+001	4.860382e+001
2	2.662837e+002	9.141678e+001	4.184019e+001
3	2.003016e+002	4.515775e+001	2.398375e+001
4	1.442203e+002	3.350677e+001	1.337688e+001
5	1.005877e+002	1.827108e+001	5.221998e+000
6	4.741106e+001	1.212349e+001	3.047301e+000
7	9.310901e+001	1.194968e+001	3.763861e+000
8	7.729366e+001	7.665960e+000	1.255710e+000
9	5.298811e+001	5.902304e+000	8.522297e-001
10	3.468176e+001	5.231683e+000	6.016774e-001
11	3.071388e+001	4.333329e+000	4.714262e-001
12	5.204887e+001	8.923042e-001	2.086165e-001
13	9.172353e+000	3.863479e-001	8.796919e-002
14	5.336919e+000	7.467896e-003	2.347799e-002
15	2.565761e+000	3.172762e-005	5.879534e-003
16	4.317529e-001	2.780582e-006	1.470310e-003
17	4.520871e-003	1.163669e-008	3.676338e-004
18	1.355981e-006	2.237868e-010	9.190920e-005
19	2.704528e-009	1.096969e-011	2.297735e-005
20	1.681707e-010	6.830279e-013	5.744339e-006
21	1.049679e-011	6.096007e-014	1.436085e-006
22	6.527763e-013	4.299406e-014	3.590212e-007
23	1.052646e-013	3.547093e-014	8.975531e-008
24	1.767013e-013	3.275357e-014	2.243883e-008
25	2.032021e-013	4.052649e-014	5.609707e-009
26	1.709005e-013	3.516538e-014	1.402427e-009
27	1.450471e-013	4.051040e-014	3.506067e-010
28	1.471216e-013	3.918672e-014	8.765167e-011
29	1.570758e-013	3.542638e-014	2.191292e-011
30	1.426681e-013	4.027960e-014	5.478230e-012
31	1.446933e-013	3.891518e-014	1.369557e-012

Note that the closer the iterates to optimum, the faster the duality gap, primal and dual infeasibility gap reduce respectively, a property not shared by some other algorithms. Observe that the duality gap reduces much slower than the primal or dual infeasibility as iterations goes on. Hence we have also used l_2 norm as the measure of duality gap, and have found that the number of total iterations reduced about 5 on average.

The above results are generated by algorithm without Mehrotra's predictor-corrector procedure. We have tried pc method also. Numerical results show that although in most cases algorithm with pc procedure requires less number of iterations, even up to one third of that without pc procedure; in some cases, algorithm with pc procedure needs much more number of iterations, occasionally two times of that without pc procedure. Average these cases, algorithm with pc procedure can save only a few number of iterations, but each iteration requires more work.

We have tested the Algorithm with the orthogonal matrix updated by *exp*, too. The solutions can achieve 12-digit accuracy as well, but it needs a couple of more iterations on average.

Second-order cone programming has many applications (see the references in § 1). We have also tested the algorithm on one application – SMT problem from [14]. The SMT problem is to find a shortest network spanning a set of given points, called regular points, on the Euclidean plane. The solution is always a tree, called the Steiner minimal stree (SMT), including some additional vertices, called Steiner points. Assume the number of regular points are N; then there are at most $N-2$ Steiner points and the degree of each Steiner point is at most 3. A tree whose vertices including just the N given regular points and $N-2$ Steiner points with the degree of each Steiner point being 3 is called a full Steiner topology of the N regular points. In [14], the problem of finding the coordinates of the $N-2$ Steiner points to form the shortest network under a known full Steiner topology is transformed into an SOCP and solved by interior point method. Their numerical examples gave better computational results than that of existing algorithms did. Their formulation

is the following.
Denote $p \stackrel{\text{def}}{=} 2N-3$, which is the number of edges; $q \stackrel{\text{def}}{=} 2N-4$, which is the total number of coordinates of the Steiner points. Let

$$\mathbf{b} = \begin{pmatrix} -\mathbf{1}_p \\ \mathbf{0}_q \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} (0;\mathbf{c}_1) \\ (0;\mathbf{c}_2) \\ \vdots \\ (0;\mathbf{c}_p) \end{pmatrix}, \quad A^T = \begin{pmatrix} -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & A_1^T \\ 0 & -1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & A_2^T \\ & & \ddots & & \\ 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & \cdots & 0 & A_p^T \end{pmatrix} \in \mathbb{R}^{3p\times(p+q)},$$

where $A_i^T \in \mathbb{R}^{2\times q}$ is a row of $(N-2)$ 2×2 block matrices. The edges are ordered so that each of the first N edges connects a regular point to a Steiner point. For $i = 1,\ldots,N$, $\mathbf{c}_i$ is the coordinates of regular point i_1, where i_1 is the index of the regular point on the ith edge; the only non-zero block of A_i^T is the i_2nd, which is I_2, where i_2 is the index of the Steiner point on the ith edge. For $i = N+1,\ldots p$, $\mathbf{c}_i = \mathbf{0}$; assume the indices of the two Steiner points on the ith edge are i_1 and i_2; then the i_1st block of A_i^T is $-I_2$, the i_2nd block of A_i^T is I_2, the rest blocks of A_i^T are zero. For $i = 1,\ldots,p$, let y_i represents the length of the ith edge. Let $\mathbf{y}_{p+1:p+q}$ be the coordinates of the Steiner points. Therefore, the SMT problem is to find $\mathbf{y}$ satisfying the dual SOCP:

$$\begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ s.t. & A^T\mathbf{y} + \mathbf{s} = \mathbf{c} \\ & \mathbf{s} \geq_Q \mathbf{0}. \end{array} \tag{36}$$

We tested our code on example 1 in [14]. The two tables below are coordinates of the 10 regular points and the tree topology taken from [14]. The Steiner points are indexed before the regular points.

The coordinates of the 10 regular points in example 1.0

index	x-xoordinate	y-coordinate	index	x-cocordinate	y-coordinate
9	2.30946900	9.20821100	14	7.59815200	0.61583600
10	0.57736700	6.48093800	15	8.56812900	3.07917900
11	0.80831400	3.51906200	16	4.75750600	3.75366600
12	1.68591200	1.23167200	17	3.92609700	7.00879800
13	4.11085500	0.82111400	18	7.43649000	7.68328400

The indices of the two vertices of each edge are listed next to the index of the edge.

The tree topology

edge-index	ea-index	eb-index	edge-index	ea-index	eb-index
1	9	7	10	18	8
2	10	1	11	5	6
3	11	2	12	6	4
4	12	3	13	4	3
5	13	4	14	3	2
6	14	5	15	2	1
7	15	5	16	1	7
8	16	6	17	7	8
9	17	8			

Our starting points and accuracy requirements are the same as those for the randomly generated problems. Following is the result.

it	network-cost	r_p	r_d	gap
0	67.4046273974	5.744563e+000	2.631707e+001	1.020000e+002
1	46.4651882048	1.612280e+000	1.531418e+001	4.650618e+001
2	54.7067952019	2.814354e+000	1.253095e+001	2.718293e+001
3	46.3246649054	2.594265e+000	1.055171e+001	2.233545e+001
4	51.9713738708	2.618248e+000	9.177738e+000	1.840966e+001
5	41.7525188087	3.162782e+000	8.840996e+000	1.282556e+001
6	43.0043202564	3.149857e+000	8.887436e+000	1.223603e+001
7	40.3247180996	2.693144e+000	7.602036e+000	1.048399e+001
8	36.9653337403	2.371615e+000	5.992973e+000	8.623604e+000
9	36.9960141482	2.330662e+000	5.951867e+000	8.716129e+000
10	30.2772082577	1.664164e+000	3.598923e+000	5.686488e+000
11	26.5944387822	1.323526e+000	9.137568e-001	3.894575e+000
12	25.9013032614	1.428789e+000	6.976833e-001	1.968024e+000
13	25.7432878386	9.999986e-001	5.091734e-001	1.336884e+000
14	25.6618153793	1.058082e+000	1.420753e-001	4.118998e-001
15	25.3652128634	8.955233e-001	1.264806e-002	1.177352e-001
16	25.3582980135	8.882891e-001	3.105220e-003	3.270312e-002
17	25.3568833418	5.713701e-001	2.243341e-003	8.855324e-003
18	25.3562545601	2.877725e-001	6.910573e-004	2.587132e-003
19	25.3561029084	4.306751e-002	3.716272e-005	7.113014e-004
20	25.3560767365	2.733321e-004	4.986347e-006	1.920240e-004
21	25.3560701290	1.065202e-007	9.328195e-007	5.100677e-005
22	25.3560683817	8.318181e-009	7.883323e-008	1.300715e-005
23	25.3560679321	1.902448e-010	1.798652e-009	3.257782e-006
24	25.3560678175	4.292413e-012	3.994189e-011	8.145884e-007
25	25.3560677888	2.365748e-013	2.196946e-012	2.036550e-007
26	25.3560677817	1.481275e-014	1.362715e-013	5.091425e-008
27	25.3560677799	9.087567e-016	8.584688e-015	1.272859e-008
28	25.3560677794	5.861620e-016	2.057660e-015	3.182151e-009
29	25.3560677793	5.112227e-016	1.182791e-015	7.955377e-010
30	25.3560677793	3.578748e-016	1.463433e-015	1.988844e-010
31	25.3560677793	5.613218e-016	1.643918e-015	4.972111e-011
32	25.3560677793	4.406061e-016	1.495714e-015	1.243028e-011
33	25.3560677793	5.530259e-016	1.646727e-015	3.107569e-012

Our initial network-cost is the same as that of [14]. The network-cost at our $27th$ iteration is better than their final cost, which shows that our accuracy requirements are higher than theirs. Our method starts from an infeasible point, while their inintial point must be feasible.

7 Modified Q Method

In this section, we will give a variant of the Q method for SOCP, which has similar properties and convergence results as that of (16).

7.1 The System

Formulation (14) shows that only the first two columns of the orthogonal matrix Q are involved in calculation. The first column of Q is $(1;\mathbf{0})$; the second column is a unit vector and its first element is zero. Partition A_i as $A_i = [(A_i)_0\ \bar{A}_i]$. Denote $\bar{A} \stackrel{\text{def}}{=} [\bar{A}_1\ \cdots\ \bar{A}_n]$. Decompose $\mathbf{x}_i$ and $\mathbf{z}_i$ as $\mathbf{x}_i = \left(\frac{(\lambda_i)_1+(\lambda_i)_2}{2}; \frac{(\lambda_i)_1-(\lambda_i)_2}{2}\mathbf{q}_i\right)$, $\mathbf{z}_i = \left(\frac{(\omega_i)_1+(\omega_i)_2}{2}; \frac{(\omega_i)_1-(\omega_i)_2}{2}\mathbf{q}_i\right)$. We let $\mathbf{q} = \binom{1}{\mathbf{0}}$ when $\bar{\mathbf{x}} = \mathbf{0}$. Then the decomposition is unique under the assumptions $(\lambda_i)_1 \geq (\lambda_i)_2$, $(\omega_i)_1 \leq (\omega_i)_2$. Substitute the decompositions into (14), and add a constraint

$$\mathbf{q}_i^T\mathbf{q}_i = 1,\ (i = 1,\ldots,n).$$

Let $\mathbf{r}_p^k \stackrel{\text{def}}{=} \mathbf{b} - A\mathbf{x}^k$, $\mathbf{r}_d^k \stackrel{\text{def}}{=} \mathbf{c} - A^T\mathbf{y}^k - \mathbf{z}^k$, $\mathbf{r}_c^k \stackrel{\text{def}}{=} \mu\mathbf{1} - \Lambda^k\boldsymbol{\omega}^k$. We use $(\tilde{r}_d)_i^k$ to represent the first element of $(\mathbf{r}_d)_i^k$, and $(\check{\mathbf{r}}_d)_i^k$ to represent the remaining subvector. Then the resulting Newton system

is

$$
\begin{aligned}
\frac{\Delta(\omega_i)_1 + \Delta(\omega_i)_2}{2} + (A_i)_0^T \, \Delta \mathbf{y} &= (\tilde{r}_d)_i^k \\
\frac{\Delta(\omega_i)_1 - \Delta(\omega_i)_2}{2} \mathbf{q}_i^k + \frac{(\omega_i)_1^k - (\omega_i)_2^k}{2} \, \Delta \mathbf{q}_i + (\bar{A}_i)^T \, \Delta \mathbf{y} &= (\check{\mathbf{r}}_d)_i^k \\
\sum_{i=1}^{n} \left(\frac{\Delta(\lambda_i)_1 + \Delta(\lambda_i)_2}{2} (A_i)_0 + \bar{A}_i \frac{\Delta(\lambda_i)_1 - (\Delta \lambda_i)_2}{2} \mathbf{q}_i^k + \bar{A}_i \frac{(\lambda_i^k)_1 - (\lambda_i^k)_2}{2} \, \Delta \mathbf{q}_i^k \right) &= \mathbf{r}_p^k \\
{\mathbf{q}_i^k}^T \, \Delta \mathbf{q}_i = 0 \quad (i = 1, \ldots, n) & \\
\Lambda^k \, \Delta \boldsymbol{\omega} + \Omega^k \, \Delta \boldsymbol{\lambda} &= \mathbf{r}_c^k.
\end{aligned} \tag{37}
$$

The algorithm is the same as that in the previous sections, except that the orthogonalization is substituted by normalization:

$$
\mathbf{q}_i^{k+1} = \frac{\mathbf{q}_i^k + \gamma \, \Delta \mathbf{q}_i}{\left\| \mathbf{q}_i^k + \gamma \, \Delta \mathbf{q}_i \right\|_2}.
$$

7.2 Properties of the Solution

Let

$$
u_i \stackrel{\text{def}}{=} \frac{(\lambda_i)_1}{(\omega_i)_1} + \frac{(\lambda_i)_2}{(\omega_i)_2}, \quad v_i \stackrel{\text{def}}{=} \frac{(\lambda_i)_1}{(\omega_i)_1} - \frac{(\lambda_i)_2}{(\omega_i)_2};
$$

E_i and D_i are defined as that in previous sections, but with proper dimensions. Omitting k, the solution to (37) is

$$
\begin{aligned}
\Delta \mathbf{y} = M^{-1} \Bigg(& \mathbf{r}_p + \frac{1}{2} \sum_{i=1}^{n} \Big(- (A_i)_0 (\mathbf{r}_c)_i^T \boldsymbol{\omega}_i^{-1} - \bar{A}_i \mathbf{q}_i (\mathbf{r}_c)_i^T \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \boldsymbol{\omega}_i^{-1} \\
& + u_i (A_i)_0 (\tilde{r}_d)_i + v_i \bar{A}_i \mathbf{q}_i (\tilde{r}_d)_i + v_i (A_i)_0 \mathbf{q}_i^T (\check{\mathbf{r}}_d)_i \\
& + (u_i + 2 D_i E_i^{-1}) \bar{A}_i \mathbf{q}_i \mathbf{q}_i^T (\check{\mathbf{r}}_d)_i - 2 D_i E_i^{-1} \bar{A}_i (\check{\mathbf{r}}_d)_i \Big) \Bigg) \\
\Delta(\omega_i)_2 = & -\mathbf{q}_i^T \left((\check{\mathbf{r}}_d)_i - \mathbf{q}_i (\tilde{r}_d)_i \right) + \left(\mathbf{q}_i^T \bar{A}_i^T - (A_i)_0^T \right) \Delta \mathbf{y} \\
\Delta \mathbf{q}_i = & E_i^{-1} \left((\check{\mathbf{r}}_d)_i - \mathbf{q}_i (\tilde{r}_d)_i + \mathbf{q}_i \, \Delta(\omega_i)_2 - (\bar{A}_i^T - \mathbf{q}_i (A_i)_0^T) \, \Delta \mathbf{y} \right) \\
\Delta(\omega_i)_1 = & \, 2(\tilde{r}_d)_i - 2 (A_i)_0^T \, \Delta \mathbf{y} - \Delta(\omega_i)_2 \\
\Delta(\boldsymbol{\lambda}) = & \, \Omega^{-1} \left(\mathbf{r}_c - \Lambda \, \Delta \boldsymbol{\omega} \right).
\end{aligned}
$$

The Schur complement M in the above formula is:

$$
M = \sum_{i=1}^{n} \frac{v_i}{2} \left((A_i)_0 \mathbf{q}_i^T \bar{A}_i^T + \bar{A}_i \mathbf{q}_i (A_i)_0^T \right) + \bar{A} \, \mathrm{Diag} \begin{pmatrix} \frac{u_i}{2} & \\ & (D_i E_i^{-1} + \frac{u_i}{2} I) \mathbf{q}_i \mathbf{q}_i^T - D_i E_i^{-1} \end{pmatrix} \bar{A}^T.
$$

Each block of the Schur complement is one dimension less than those of other systems. When $(\lambda_i)_1 > (\lambda_i)_2 > 0$ and $(\omega_i)_2 > (\omega_i)_1 > 0$, we have $u_i > v_i > 0$, and $-D_i E_i^{-1}$ is a positive scalar matrix. Because one is the only nonzero eigenvalue of $\mathbf{q}_i \mathbf{q}_i^T$, the second part of M is symmetric positive definite. Observe the first part of M is symmetric positive semidefinite. Therefore M is symmetric positive definite; so we can use Cholesky factorization to calculate the Schur complement. The number of variables and equations used by the modified Q method are also about half of that required by the other method. So it is also efficient in storage and calculation per iteration.

The dimension of the Schur complement M is one dimensional less than other methods for each block; to keep each iterate in $\mathcal{Q}$, one only needs to compute $\alpha^{-1} \geq \max\{-\Delta \lambda_i / \lambda_i : \, \Delta \lambda_i < 0\}$, not a solution of the second-order equation.

To use (37), we don't need to update the orthogonal matrix, but the price we pay is n more variables and equations.

Similar to Theorem 3.1, we have

Theorem 7.1 *Let* $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ *be an optimal solution of* (1) *satisfying strict complementarity, primal and dual nondegeneracy conditions, and also condition* (17). *Assume* $\mathbf{x} \neq \mathbf{0}$ *at optimum. Decompose* $\mathbf{x}_i = \left(\frac{(\lambda_i)_1+(\lambda_i)_2}{2}; \frac{(\lambda_i)_1-(\lambda_i)_2}{2}\mathbf{q}_i\right)$, $\mathbf{z}_i = \left(\frac{(\omega_i)_1+(\omega_i)_2}{2}; \frac{(\omega_i)_1-(\omega_i)_2}{2}\mathbf{q}_i\right)$. *Then the Jacobian of* (37) *evaluated at* $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ *is nonsingular.*

Proof: For any unit vector $\mathbf{q} \in \mathbb{R}^n$, define an orthogonal matrix $\bar{Q}_{\mathbf{q}}$ as

$$\bar{Q}_{\mathbf{q}} = \begin{cases} \begin{pmatrix} q_0 & -\bar{\mathbf{q}}^T \\ \bar{\mathbf{q}} & I - \frac{\bar{\mathbf{q}}\bar{\mathbf{q}}^T}{1+q_0} \end{pmatrix} & q_0 \neq -1 \\ \begin{pmatrix} -1 & & \\ & -1 & \\ & & I \end{pmatrix} & q_0 = -1 \end{cases}.$$

After dropping the iteration number k, we write each block of the Jacobian (37) as the following.

	$(\omega_i)_1$	$(\omega_i)_2$	$(\lambda_i)_1$	$(\lambda_i)_2$	$\mathbf{q}_i$	$\mathbf{y}_i$
$(\mathbf{r}_p)_i$			$\frac{1}{2}A_i \begin{pmatrix} 1 \\ \mathbf{q}_i \end{pmatrix}$	$\frac{1}{2}A_i \begin{pmatrix} 1 \\ -\mathbf{q}_i \end{pmatrix}$	$D_i\bar{A}_i$	
$(\mathbf{r}_d)_i$	$\frac{1}{2}\begin{pmatrix} 1 \\ \mathbf{q}_i \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 \\ -\mathbf{q}_i \end{pmatrix}$			$\begin{pmatrix} 0 \\ E_i \end{pmatrix}$	A_i^T
$((\mathbf{r}_c)_i)_1$	$(\lambda_i)_1$		$(\omega_i)_1$			
$((\mathbf{r}_c)_i)_2$		$(\lambda_i)_2$		$(\omega_i)_2$		
0					$\mathbf{q}_i^T$	

We first left multiply $\mathrm{Diag}\left(\begin{bmatrix} 1 & \\ & \bar{Q}_{\mathbf{q}_i}^T \end{bmatrix}\right)$ to the block of dual feasibility equations; then right time $\mathrm{Diag}\left(\bar{Q}_{\mathbf{q}_i}^T\right)$ to the columns corresponding to $\mathbf{q}$. Notice $\bar{A}_i\mathbf{q}_i = \bar{A}_i\bar{Q}_{\mathbf{q}_i}\bar{Q}_{\mathbf{q}_i}^T\mathbf{q}_i = \bar{A}_i\bar{Q}_{\mathbf{q}_i}\begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}$. After crossing out columns $(\mathbf{q}_i)_1$ and rows $\mathbf{q}_i^T\mathbf{q}_i$ (for $i = 1, \ldots, n$), we find the Jacobian of (37) is the same as that of (15) with $B = A\,\mathrm{Diag}\left(\begin{bmatrix} 1 & \\ & \bar{Q}_{\mathbf{q}_i}^T \end{bmatrix}\right)$. Hence all the proof of Theorem 3.1 are applicable here. ■

So, as for the Q method, we can expect that when the iterates of modified Q method are close to the optimum, they converge fast, and the solutions are accurate and numerically stable.

7.3 Convergence Analysis

All the convergence proofs in the previous sections can be adapted to the modified Q method. For example, replace $\|\mathbf{s}_i\|_2^2 \le \eta$ by $\|\Delta\,\mathbf{q}_i\|_2^2 \le \eta$ in the proof of Theorem 4.1. Then

$$
\begin{aligned}
(38)\quad g_p(\alpha) &= (\boldsymbol{\lambda}^k + \alpha\,\Delta\,\boldsymbol{\lambda})^T(\boldsymbol{\omega}^k + \alpha\,\Delta\,\boldsymbol{\omega}) - \gamma_p\,\|A\mathbf{x} - \mathbf{b}\|_2 \\
&= (1-\alpha)\boldsymbol{\lambda}^{k^T}\boldsymbol{\omega}^k + \alpha\sigma_1\boldsymbol{\lambda}^{k^T}\boldsymbol{\omega}^k + \alpha^2\,\Delta\,\boldsymbol{\lambda}^T\,\Delta\,\boldsymbol{\omega} - \gamma_p\left\|\sum_{i=1}^{n}\left[(A_i)_0\frac{(\lambda_i)_1 + (\lambda_i)_2}{2}\right.\right. \\
&+ \bar{A}\frac{(\lambda_i)_1 - (\lambda_i)_2}{2}\mathbf{q}_i - \mathbf{b} + \alpha(A_i)_0\frac{\Delta(\lambda_i)_1 + \Delta(\lambda_i)_2}{2} + \alpha\bar{A}_i\frac{\Delta(\lambda_i)_1 - \Delta(\lambda_i)_2}{2}\mathbf{q}_i \\
&+ \alpha\bar{A}_i\frac{(\lambda_i)_1 - (\lambda_i)_2}{2}\,\Delta\,\mathbf{q}_i + \alpha^2\bar{A}_i\frac{\Delta(\lambda_i)_1 - \Delta(\lambda_i)_2}{2}\frac{\Delta\,\mathbf{q}_i}{\|\mathbf{q}_i + \alpha\,\Delta\,\mathbf{q}_i\|_2} \\
&+ \bar{A}_i\left(\frac{(\lambda_i)_1 - (\lambda_i)_2}{2} + \alpha\frac{\Delta(\lambda_i)_1 + \Delta(\lambda_i)_2}{2}\right)\mathbf{q}_i\left(\frac{1}{\|\mathbf{q}_i + \alpha\,\Delta\,\mathbf{q}_i\|_2} - 1\right) \\
&\left.\left. + \bar{A}\frac{(\lambda_i)_1 - (\lambda_i)_2}{2}\alpha\,\Delta\,\mathbf{q}_i\left(\frac{1}{\|\mathbf{q}_i + \alpha\,\Delta\,\mathbf{q}_i\|_2} - 1\right)\right]\right\|_2.
\end{aligned}
$$

Since $\frac{1}{\sqrt{1+a}} - 1 + \frac{a}{2}$ is increasing for $a \ge 0$, one can easily see that

$$-\frac{a}{2} \le \frac{1}{\sqrt{1+a}} - 1 \le 0.$$

Also notice $\|q_i\|_2 = 1$ and $\mathbf{q}_i^T\,\Delta\,\mathbf{q}_i = 0$. Therefore,

$$1 - \frac{1}{\|\mathbf{q}_i + \alpha\,\Delta\,\mathbf{q}_i\|_2} = \left|\frac{1}{\sqrt{1 + \alpha^2\,\Delta\,\mathbf{q}_i^T\,\Delta\,\mathbf{q}_i}} - 1\right| \le \frac{\alpha^2}{2}\,\Delta\,\mathbf{q}_i^T\,\Delta\,\mathbf{q}_i \le \frac{\alpha^2\eta}{2}.$$

Hence,

$$(38) \ge \alpha\sigma_1\epsilon^* - \alpha^2\eta - \gamma_p\left\|\bar{A}\right\|_2\alpha^2\left(\eta^{\frac{3}{2}} + \left(\frac{\Upsilon}{2} + \eta\right)\frac{\eta}{2} + \frac{1}{2}\Upsilon\eta^{\frac{3}{2}}\right).$$

In other words, a lower bound on α for $g_p(\alpha) \ge 0$ is

$$\frac{\sigma_1\epsilon^*}{\eta + \gamma_p\left\|\bar{A}\right\|_2\left(\eta^{\frac{3}{2}} + \left(\frac{\Upsilon}{2} + \eta\right)\frac{\eta}{2} + \frac{1}{2}\Upsilon\eta^{\frac{3}{2}}\right)}.$$

Therefore, all the lemmas and theorems in the previous sections are fitting here.

7.4 Numerical Examples

We have implemented the basic algorithm for the modified Q method in MATLAB and have tested on 1,000 randomly generated problems. Step sizes α, β, γ are chosen as that for the Q method. The problem types, accuracy requirement, starting points, and parameters are the same as that in § 6. Below is the results.

problem	r_p0	r_d0	it
1	344.73	45.12	27.04
2	300.98	143.38	36.71
3	537.54	144.78	38.14
4	855.71	188.39	31.91
5	1343.14	268.49	32.75
6	422.70	199.52	32.84
7	549.52	260.49	38.08
8	1735.21	376.44	33.13
9	1496.93	493.70	37.54
10	1362.12	576.44	34.56

Although the algorithm finds an ϵ-optimal solution for all the 1,000 problems, a small portion of them need more than 100 iterations to reach the required accuracy, which brings up the number of average iterations.

Following is the results on SMT problem.

it	network-cost	r_p	r_d	gap
0	67.404627397	5.744563e+000	2.631707e+001	1.020000e+002
1	46.4651882048	1.570830e+000	1.536167e+001	4.650618e+001
2	57.9523360177	2.854242e+000	1.322494e+001	2.962662e+001
3	43.2491319152	2.266570e+000	8.886548e+000	1.977556e+001
4	36.5711471818	1.564727e+000	6.352045e+000	1.321500e+001
5	31.7493416355	1.773279e+000	4.411577e+000	8.430118e+000
6	27.6120403168	1.706357e+000	2.941044e+000	5.720180e+000
7	26.0413165945	7.411994e-001	6.780735e-001	2.293613e+000
8	25.5095803443	2.379125e-001	2.814603e-001	9.947367e-001
9	25.3737639737	2.115945e-002	1.835768e-002	2.627086e-001
10	25.3607982437	9.699106e-004	1.368075e-003	7.079254e-002
11	25.3573940424	6.257227e-005	3.953843e-004	1.934408e-002
12	25.3564281678	1.305314e-005	1.163975e-004	5.320336e-003
13	25.3561678287	4.022397e-006	3.361298e-005	1.473695e-003
14	25.3560924356	1.020580e-006	8.861424e-006	4.043510e-004
15	25.3560732242	2.106188e-007	1.953306e-006	1.080564e-004
16	25.3560690742	2.801981e-008	2.647413e-007	2.788748e-005
17	25.3560681075	1.282968e-009	1.213864e-008	7.011850e-006
18	25.3560678616	2.592161e-011	2.430832e-010	1.753810e-006
19	25.3560677999	1.120376e-012	1.041288e-011	4.384902e-007
20	25.3560677844	6.817490e-014	6.326124e-013	1.096249e-007
21	25.3560677806	4.091383e-015	3.977645e-014	2.740636e-008
22	25.3560677796	8.569464e-016	2.755012e-015	6.851598e-009
23	25.3560677794	7.306626e-016	1.534069e-015	1.712900e-009
24	25.3560677793	7.636567e-016	1.625551e-015	4.282250e-010
25	25.3560677793	8.355535e-016	2.122419e-015	1.070563e-010
26	25.3560677793	6.747517e-016	1.693835e-015	2.676407e-011
27	25.3560677793	8.039540e-016	1.946316e-015	6.691016e-012
28	25.3560677793	1.009651e-015	1.802288e-015	1.672754e-012

Note the total number of iteration required to reach the final network-cost of [14] is 21, one less than that of [14].

8 Conclusion and Future Research

We have developed and analyzed the Q method and its variant for SOCP. Preliminary numerical results show that the algorithm is promising. In the future, we intend to investigate sparse matrix issues and large-scale application.

Appendix

In this section, we will show that (13) is valid for any $S \in \mathfrak{l}$. We use the notion of primary matrix function (see [9, 6.2.4, p. 410]) to define a matrix valued function. The definition is the following.

Definition 1 *Let A be a given square matrix with Jordan canonical form $A = UJU^{-1}$. Assume*

$$J = \begin{pmatrix} J_{n_1}(\lambda_{\nu_1}) & & \\ & \ddots & \\ & & J_{n_r}(\lambda_{\nu_r}) \end{pmatrix},$$

where each $J_k(\lambda)$ is a k-by-k Jordan block with eigenvalue λ. Let c_i be the dimension of the largest Jordan block corresponding to λ_i. Let $f(t)$ be a scalar valued function of t such that each λ_i with

$c_i > 1$ is in the interior of the domain of $f(t)$ and $f(t)$ is $(c_i - 1)$-times differentiable at λ_i. Then the primary matrix function $f(A)$ associated with the scalar-valued stem function $f(t)$ is defined as

$$f(A) \stackrel{\text{def}}{=} Uf(J)U^{-1} \stackrel{\text{def}}{=} U \begin{pmatrix} f\left(J_{n_1}(\lambda_{\nu_1})\right) & & \\ & \ddots & \\ & & f\left(J_{n_r}(\lambda_{\nu_r})\right) \end{pmatrix} U^{-1},$$

where

$$f\left(J_k(\lambda)\right) \stackrel{\text{def}}{=} \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) & \cdots & \frac{1}{(k-1)!}f^{k-1}(\lambda) \\ & f(\lambda) & f'(\lambda) & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2}f''(\lambda) \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}.$$

Notice for any element $S \in \mathfrak{l}$, its eigenvalues are $\|\mathbf{s}\|_2 i$, $-\|\mathbf{s}\|_2 i$, and 0 with multiplicity $n-1$. Its minimal polynomial is $t(t - \|\mathbf{s}\|_2 i)(t + \|\mathbf{s}\|_2 i)$. Define some scalar-valued functions:

$$h_1(t) \stackrel{\text{def}}{=} \left(1 + \frac{t}{2}\right), \quad h_2(t) \stackrel{\text{def}}{=} \left(1 - \frac{t}{2}\right), \quad f(t) \stackrel{\text{def}}{=} h_1(t)/h_2(t).$$

It is obvious that $h_2(t) \neq 0$ when evaluated at any eigenvalue of $S \in \mathfrak{l}$. Hence by [9, 6.2.10 (e), p. 416], $h_2(S)$ is nonsingular for any $S \in \mathfrak{l}$; and $f(S) = h_1(S)\left[h_2(S)\right]^{-1}$. So $f(S) = \left(I + \frac{S}{2}\right)\left(I - \frac{S}{2}\right)$ is well defined on $\mathfrak{l}$ as the primary matrix function associated with the stem function $f(t)$.

Next, we define a family of scalar-valued function $g_{\mathbf{s}}(t)$ with parameter $\mathbf{s}$ corresponding to $S \in \mathfrak{l}$:

$$g_{\mathbf{s}}(t) \stackrel{\text{def}}{=} 1 + \frac{4t}{4 + \|\mathbf{s}\|_2^2} + \frac{2t^2}{4 + \|\mathbf{s}\|_2^2}.$$

For any $S \in \mathfrak{l}$, it is easy to verify that $g_{\mathbf{s}}(t) = f(t)$ for $t = 0, \|\mathbf{s}\|_2 i$, or $-\|\mathbf{s}\|_2 i$. Therefore, $g_{\mathbf{s}}(S) = f(S)$ by [9, 6.2.9(e), p. 413].

Hence, we have proved that the equality (13) is valid for any $S \in \mathfrak{l}$.

References

[1] I. Adler and F. Alizadeh. Primal-dual interior point algorithms for convex quadratically con strained and semidefinite optimization problems. Technical Report RRR 46-95, RUTCOR, Rutgers University, 1995.

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

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

[4] Farid Alizadeh, Jean-Pierre A. Haeberly, and Michael L. Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM J. Optim.*, 8(3):746–768 (electronic), 1998.

[5] Farid Alizadeh and Yu Xia. A note on 2-dimensional second-order cone programming and its simplex method. Technical Report AdvOl-Report No. 2004/16, McMaster University, 2004.

[6] Farid Alizadeh and Yu Xia. The Q method for symmetric cone programming. Technical Report AdvOl-Report No. 2004/18, McMaster University, 2004.

[7] Jacques Faraut and Adam Korányi. *Analysis on symmetric cones*. The Clarendon Press Oxford University Press, New York, 1994. Oxford Science Publications.

[8] Roland W. Freund, Florian Jarre, and Shinji Mizuno. Convergence of a class of inexact interior-point algorithms for linear programs. *Math. Oper. Res.*, 24(1):50–71, 1999.

[9] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[10] Masakazu Kojima, Nimrod Megiddo, and Shinji Mizuno. A primal-dual infeasible-interior-point algorithm for linear programming. *Math. Programming*, 61(3, Ser. A):263–280, 1993.

[11] Miguel Sousa Lobo, Lieven Vandenberghe, Stephen Boyd, and Hervé Lebret. Applications of second-order cone programming. *Linear Algebra Appl.*, 284(1-3):193–228, 1998. ILAS Symposium on Fast Algorithms for Control, Signals and Image Processing (Winnipeg, MB, 1997).

[12] Irvin J. Lustig. Feasibility issues in a primal-dual interior-point method for linear programming. *Math. Programming*, 49(2, (Ser. A)):145–162, 1990/91.

[13] Yu Xia. An algorithm for perturbed second-order cone programs. Technical Report AdvOl-Report No. 2004/17, McMaster University, 2004.

[14] Guoliang Xue and Yinyu Ye. An efficient algorithm for minimizing a sum of Euclidean norms with applications. *SIAM J. Optim.*, 7(4):1017–1036, 1997.