A primal-dual interior point method for nonlinear optimization over second order cones *

Hiroshi Yamashita† and Hiroshi Yabe‡

Abstract

In this paper, we are concerned with nonlinear minimization problems with second order cone constraints. A primal-dual interior point method is proposed for solving the problems. We also propose a new primal-dual merit function by combining the barrier penalty function and the potential function within the framework of the line search strategy, and show the global convergence property of our method.

Key words. constrained optimization, second order cone, primal-dual interior point method, barrier penalty function, potential function, global convergence

AMS subject classifications. 90C30, 90C51, 90C53

1 Introduction

In this paper, we consider the following constrained optimization problem with the second order cone constraints:

$$\begin{array}{lll} \text{minimize} & f(x), & x \in \mathbf{R}^n, \\ \text{subject to} & g(x) = 0, & x \in \mathcal{K}, \end{array} \tag{1}$$

where we assume that the functions $f : \mathbf{R}^n \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}^m$ are sufficiently smooth, and $\mathcal{K}$ is the Cartesian product of socond order cones: $\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2 \times \cdots \times \mathcal{K}^s$, and $\mathcal{K}^i$ is an n_i dimensional second order cone which is define by

$$\mathcal{K}^i = \{(x_0^i, \bar{x}^i)^t \in \mathbf{R}^{n_i} \mid x_0^i \geq \|\bar{x}^i\|,\ x_0^i \in \mathbf{R},\ \bar{x}^i \in \mathbf{R}^{n_i - 1}\},$$

and $n_1 + n_2 + \cdots + n_s = n$, and $\|\cdot\|$ denotes the l_2 vector norm. Let $x = (x^1, x^2, \ldots, x^s)^t$ where $x^i = (x_0^i, \bar{x}^i)^t \in \mathbf{R}^{n_i}$. By $x \in \mathcal{K}$, we mean

$$x^i \in \mathcal{K}^i \subset \mathbf{R}^{n_i}, \quad i = 1, \ldots, s.$$

*The second author was supported in part by the Grant-in-Aid for Scientific Research (C) 16510123 of Japan Society for the Promotion of Sciences.

†Mathematical Systems, Inc., 2-4-3, Shinjuku, Shinjuku-ku, Tokyo, Japan. hy@msi.co.jp

‡Department of Mathematical Information Science, Faculty of Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo, Japan. yabe@rs.kagu.tus.ac.jp

We denote the conditions $x^i \in \mathcal{K}^i, x^i \in \text{int}\mathcal{K}^i, x \in \mathcal{K}, x \in \text{int}\mathcal{K}$ by $x^i \succeq 0, x^i \succ 0, x \succeq 0, x \succ 0$, respectively. If there exists a constraint like $h(x) \succeq 0, h : \mathbf{R}^n \to \mathbf{R}^{n'}$ in the problem to be solved, we transform the constraint to $h(x) - v = 0, v \succeq 0$ by introducing a slack variable $v \in \mathbf{R}^{n'}$ which results in the above form (1).

Various examples of SOCP (second order cone programming) problems are described in [10]. Examples in the paper are linear SOCPs, i.e., the functions $f(x)$ and $g(x)$ above are linear. However it is easy to extend these examples to nonlinear cases. For example, there is no reason that the robust optimization problem which is often referred to as a typical example of linear SOCP should not include a nonlinear objective function.

It is known that linear SOCP problems include linear and convex quadratic programming problems as special cases, and are special cases of SDP (semidefinite programming) problems. Interior point methods for solving these problems have been studied by many researchers in the past. On the other hand, some researchers have studied numerical methods for solving nonlinear SOCP or SDP problems. For example, Kocvara and Stingl [9] developed a computer program PENNON for solving nonlinear SDP, in which the augmented Lagrangian function method was used. Correa and Ramirez [4] proposed an algorithm for nonlinear SDP which modified the sequentially semidefinite programming method by using a nondifferentiable merit function. Kato and Fukushima [8] proposed an SQP-type algorithm for nonlinear SOCP problems. Related researches include Jarre [7], Freund and Jarre [6] and Bonnans and Ramirez [2]. However, there are not so much research has been done on interior point methods for solving nonlinear SOCP problems yet.

In this paper, we propose a primal-dual interior point method for solving nonlinear SOCP problems. The method is based on a line search algorithm in the primal-dual space. We show its global convergence. The present paper is organized as follows. In Section 2, the optimality condition for problem (1) and basic Jordan algebra are introduced. In Sections 3 and 4, our primal-dual interior point method is discussed. Specifically, in Section 4.1, we describe the Newton method for solving nonlinear equations that are obtained by modifying the optimality conditions given in Section 2. In Section 4.2, we propose a new primal-dual merit function that consists of the barrier penalty function and the potential function. Then Section 4.3 presents the algorithm called SOCPLS based on the line search strategy, and Section 4.4 shows its global convergence property. Finally, we give some concluding remarks in Section 5.

2 Optimality conditions and basic Jordan algebra

Let the Lagrangian function of problem (1) be defined by

$$L(w) = f(x) - y^t g(x) - z^t x,$$

where $w = (x, y, z)^t$, and $y \in \mathbf{R}^m$ and $z \in \mathbf{R}^n$ are the Lagrange multiplier vectors which correspond to the equality and second order cone constraints respectively. Then Karush-Kuhn-Tucker (KKT) conditions for optimality of problem (1) are given by the following (see [3]):

$$r_0(w) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ x \circ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{2}$$

and

$$x \succeq 0, \qquad z \succeq 0. \tag{3}$$

Here $\nabla_x L(w)$ is given by

$$\begin{aligned} \nabla_x L(w) &= \nabla f(x) - A(x)^t y - z, \\ A(x) &= \begin{pmatrix} \nabla g_1(x)^t \\ \vdots \\ \nabla g_m(x)^t \end{pmatrix}, \end{aligned}$$

and the multiplication $x \circ z$ is defined by

$$x \circ z = \begin{pmatrix} x^1 \circ z^1 \\ \vdots \\ x^s \circ z^s \end{pmatrix},$$

where

$$x^i \circ z^i = \begin{pmatrix} (x^i)^t z^i \\ x_0^i \bar{z}^i + z_0^i \bar{x}^i \end{pmatrix}.$$

The Jordan algebra used in this paper is surveyed in the paper by Alizadeh and Goldfarb [1] (see also [5]). We first define the following notations:

$$\begin{aligned} \mathrm{Arw}(x) &= \mathrm{Arw}(x^1) \oplus \mathrm{Arw}(x^2) \oplus \cdots \oplus \mathrm{Arw}(x^s), \\ \mathrm{Arw}(x^i) &= \begin{pmatrix} x_0^i & (\bar{x}^i)^t \\ \bar{x}^i & x_0^i I \end{pmatrix} \in \mathbf{R}^{n_i \times n_i}, \\ e &= (e^1, \ldots, e^s)^t, \\ e^i &= (1, 0)^t \in \mathbf{R}^{n_i} \quad \text{with} \quad 0 \in \mathbf{R}^{n_i - 1}, \\ \det(x^i) &= (x_0^i)^2 - \|\bar{x}^i\|^2, \\ R_i &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} \in \mathbf{R}^{n_i \times n_i}. \end{aligned}$$

Here $\det(x^i)$ is called the determinant of the vector x^i. We note that $\det(x^i) > 0$ for $x^i \succ 0$. We also note that $x^i \succ 0$ if and only if the matrix $\mathrm{Arw}(x^i)$ is positive definite. By using the notation above, the multiplication $x^i \circ z^i$ can be expressed as

$$x^i \circ z^i = \mathrm{Arw}(x^i) z^i = \mathrm{Arw}(x^i)\mathrm{Arw}(z^i)e. \tag{4}$$

The vector e^i is the unique identity in the sense that $v \circ e^i = v$ holds for any $v \in \mathbf{R}^{n_i}$. It is known that there exists a unique inverse $(x^i)^{-1}$ for any $x^i \succ 0$ in the sense that $x^i \circ (x^i)^{-1} = e^i$. Let

$$x^{-1} = ((x^1)^{-1}, (x^2)^{-1}, \ldots, (x^s)^{-1})^t.$$

In this case, x and x^i are said to be nonsingular. We note that the inverse of x^i is written as

$$(x^i)^{-1} = \frac{R_i x^i}{\det(x^i)}.$$

In the following, we also use the relation

$$x^{-1} = \mathrm{Arw}(x)^{-1} e,$$

which can be proved by confirming $\mathrm{Arw}(x^{-1})e = \mathrm{Arw}(x)^{-1}e$.

We next introduce the so-called spectral decomposition of a vector $x^i \in \mathbf{R}^{n_i}$, which is given by

$$x^i = \lambda_1^i c_1^i + \lambda_2^i c_2^i,$$

where λ_1^i, λ_2^i are called the eigenvalues and c_1^i, c_2^i are called the Jordan frame of the vector x^i, respectively. They are defined by

$$\lambda_1^i = x_0^i + \|\bar{x}^i\|, \qquad \lambda_2^i = x_0^i - \|\bar{x}^i\|$$

and

$$c_1^i = \frac{1}{2}\begin{pmatrix} 1 \\ \frac{\bar{x}^i}{\|\bar{x}^i\|} \end{pmatrix}, \qquad c_2^i = \frac{1}{2}\begin{pmatrix} 1 \\ -\frac{\bar{x}^i}{\|\bar{x}^i\|} \end{pmatrix}.$$

We note that the Jordan frame $\{c_1^i, c_2^i\}$ satisfies the relations

$$c_1^i \circ c_2^i = 0, \quad c_1^i \circ c_1^i = c_1^i, \quad c_2^i \circ c_2^i = c_2^i, \quad c_1^i + c_2^i = e^i, \quad c_1^i = R_i c_2^i \ \text{ and } \ c_2^i = R_i c_1^i.$$

Eigenvalues have the properties $\lambda_1^i \geq 0,\ \lambda_2^i \geq 0$ for $x^i \succeq 0$ and $\lambda_1^i > 0,\ \lambda_2^i > 0$ for $x^i \succ 0$. The inverse of a nonsingular vector x^i can be written as

$$(x^i)^{-1} = (\lambda_1^i)^{-1} c_1^i + (\lambda_2^i)^{-1} c_2^i.$$

Furthermore, for $x^i \succ 0$, we can define

$$(x^i)^{1/2} = (\lambda_1^i)^{1/2} c_1^i + (\lambda_2^i)^{1/2} c_2^i$$

and

$$(x^i)^{-1/2} = (\lambda_1^i)^{-1/2} c_1^i + (\lambda_2^i)^{-1/2} c_2^i,$$

which satisfy the properties $(x^i)^{1/2} \circ (x^i)^{1/2} = x^i$ and $(x^i)^{-1/2} \circ (x^i)^{-1/2} = (x^i)^{-1}$.

We call $w = (x, y, z)$ satisfying $x \succ 0$ and $z \succ 0$ an interior point. The algorithm of this paper will generate such interior points. To construct an interior point algorithm, we introduce a positive parameter μ, and try to find a point that satisfies the barrier KKT (BKKT) conditions:

$$r(w, \mu) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ x \circ z - \mu e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{5}$$

and

$$x \succ 0, \qquad z \succ 0. \tag{6}$$

In applying the Newton method to the system of equations (5), we usually consider an effective scaling of the primal-dual pair (x, z) (Tsuchiya [11]). For this purpose, we define the transformations

$$\begin{array}{rcl} T_p & = & T_{p^1} \oplus T_{p^2} \oplus \cdots \oplus T_{p^s}, \\ T_{p^i} & = & 2\mathrm{Arw}^2(p^i) - \mathrm{Arw}((p^i)^2) \end{array}$$

with respect to $p^i \succ 0,\ i = 1, \dots, s$. The matrix T_p is nonsingular if and only if the inverse of p exists. Using this transformation, we scale x and z by

$$\tilde{x} = T_p x \quad \text{and} \quad \tilde{z} = T_p^{-1} z.$$

Then we obtain (see Theorem 8 in [1])

$$\tilde{x}^{-1} = T_p^{-1} x^{-1} \quad \text{and} \quad \tilde{z}^{-1} = T_p z^{-1}. \tag{7}$$

Throughout this paper, we choose the transformation T_p such that the matrices $\mathrm{Arw}(\tilde{x})$ and $\mathrm{Arw}(\tilde{z})$ commute. In this case, the vectors $\tilde{x}^i$ and $\tilde{z}^i$ share a Jordan frame $\{c_1^i, c_2^i\}$, that is, they can be represented by

$$\tilde{x}^i = \lambda_1^i c_1^i + \lambda_2^i c_2^i \quad \text{and} \quad \tilde{z}^i = \tau_1^i c_1^i + \tau_2^i c_2^i,$$

where λ_1^i, λ_2^i and τ_1^i, τ_2^i are the eigenvalues of $\tilde{x}^i$ and $\tilde{z}^i$, respectively.

As examples of the transformation that makes $\mathrm{Arw}(\tilde{x})$ and $\mathrm{Arw}(\tilde{z})$ commute, the following choices of p are well known:

$$p = z^{1/2}, \qquad p = x^{-1/2} \tag{8}$$

and

$$p = \left[T_{x^{1/2}}(T_{x^{1/2}} z)^{-1/2}\right]^{-1/2} = \left[T_{z^{-1/2}}(T_{z^{1/2}} x)^{1/2}\right]^{-1/2}. \tag{9}$$

For the first two choices, we have

$$\tilde{z} = T_{z^{1/2}}^{-1} z = e \quad \text{and} \quad \tilde{x} = T_{x^{-1/2}} x = e,$$

respectively. The third choice (9) is the Nesterov-Todd direction and this yields $\tilde{x} = \tilde{z}$. See the paper by Alizadeh and Goldfarb [1] for more detailed exposition and references.

3 A procedure for satisfying KKT conditions

We first describe a procedure for finding a KKT point using the BKKT conditions. In this section, the subscript k denotes an iteration count of the outer iterations.

Algorithm SOCPIP

Step 0. (Initialize) Set $\varepsilon > 0$, $M_c > 0$ and $k = 0$. Let a positive sequence $\{\mu_k\}, \mu_k \downarrow 0$ be given.

Step 1. (Approximate BKKT point) Find an interior point w_{k+1} that satisfies

$$\|r(w_{k+1}, \mu_k)\| \le M_c \mu_k. \tag{10}$$

Step 2. (Termination) If $\|r_0(w_{k+1})\| \le \varepsilon$, then stop.

Step 3. (Update) Set $k := k+1$ and go to Step 1. □

We note that the barrier parameter sequence $\{\mu_k\}$ in Algorithm SOCPIP needs not be determined beforehand. The value of each μ_k may be set adaptively as the iteration proceeds. We call condition (10) the approximate BKKT condition, and call a point that satisfies this condition the approximate BKKT point.

The following theorem shows the convergence property of Algorithm SOCPIP.

Theorem 1 *Assume that the functions f and g are continuously differentiable. Let $\{w_k\}$ be an infinite sequence generated by Algorithm SOCPIP. Suppose that the sequences $\{x_k\}$ and $\{y_k\}$ are bounded. Then $\{z_k\}$ is bounded, and any accumulation point of $\{w_k\}$ satisfies KKT conditions* (2) *and* (3)*.*

Proof. Assume that $\{z_k\}$ is not bounded, i.e., that there exists an i such that $(z_k)_i \to \infty$. Equation (10) yields

$$\left| \frac{(\nabla f(x_k) - A(x_k)^t y_k)_i}{(z_k)_i} - 1 \right| \le M_c \frac{\mu_{k-1}}{(z_k)_i}.$$

The sequences $\{x_k\}$ and $\{y_k\}$ are bounded, and f and g are continuously differentiable, and $\mu_k \to +0$ as $k \to \infty$. This implies that $1 \le 0$, which is a contradiction. Thus the sequence $\{z_k\}$ is bounded.

Let $\hat{w}$ be any accumulation point of $\{w_k\}$. Since the sequences $\{w_k\}$ and $\{\mu_k\}$ satisfy (10) for each k and μ_k approaches zero, $r_0(\hat{w}) = 0$ follows from the definition of $r(w, \mu)$.

Therefore the proof is complete. □

4 An algorithm for finding a barrier KKT point

Using the transformation T_p described in Section 2, we replace the equation $x \circ z = \mu e$ by an equivalent form $\tilde{x} \circ \tilde{z} = \mu e$, and deal with the modified BKKT conditions

$$\tilde{r}(w, \mu) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ \tilde{x} \circ \tilde{z} - \mu e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{11}$$

instead of (5) to form Newton directions as described below.

4.1 The Newton method

In this subsection we consider a method for solving the BKKT conditions approximately for a given $\mu > 0$ (Step 1 of Algorithm SOCPIP). Throughout this section, the index k denotes the inner iteration count for a given $\mu > 0$. We note again that $x_k \succ 0$ and $z_k \succ 0$ for all k in the following.

For the above purpose, we apply a Newton-like method to the system of equations (11). Let the Newton directions for the primal and dual variables by Δx and Δz, respectively. Since $\tilde{x} \circ \tilde{z} = \mu e$ can be written as $(T_p x) \circ (T_p^{-1} z) = \mu e$, the equation $T_p(x + \Delta x) \circ T_p^{-1}(z + \Delta z) = \mu e$ yields

$$(T_p x) \circ (T_p^{-1} z) + (T_p x) \circ (T_p^{-1} \Delta z) + (T_p \Delta x) \circ (T_p^{-1} z) + (T_p \Delta x) \circ (T_p^{-1} \Delta z) = \mu e.$$

By neglecting the nonlinear part $(T_p \Delta x) \circ (T_p^{-1} \Delta z)$, we have the equation

$$(T_p x) \circ (T_p^{-1} z) + (T_p x) \circ (T_p^{-1} \Delta z) + (T_p \Delta x) \circ (T_p^{-1} z) = \mu e. \tag{12}$$

Then using (4), the Newton equations for solving (11) are defined by

$$\begin{aligned} G\Delta x - A(x)^t \Delta y - \Delta z &= -\nabla_x L(w), && (13) \\ A(x)\Delta x &= -g(x), && (14) \\ \mathrm{Arw}(\tilde{z}) T_p \Delta x + \mathrm{Arw}(\tilde{x}) T_p^{-1} \Delta z &= \mu e - \mathrm{Arw}(\tilde{x}) \mathrm{Arw}(\tilde{z}) e, && (15) \end{aligned}$$

or equivalently

$$J(w)\Delta w = -\tilde{r}(w, \mu), \tag{16}$$

where the matrix $J(w)$ is given by

$$J(w) = \begin{pmatrix} G & -A(x)^t & -I \\ A(x) & 0 & 0 \\ \mathrm{Arw}(\tilde{z}) T_p & 0 & \mathrm{Arw}(\tilde{x}) T_p^{-1} \end{pmatrix},$$

and the matrix G is $\nabla_x^2 L(w)$ or an approximation to $\nabla_x^2 L(w)$. We recommend to use a quasi-Newton approximation for G if $\nabla_x^2 L(w)$ is indefinite, because we will assume positive semidefiniteness of G in this paper. Since equation (15) was derived for a transformation T_p where p denpends on the current w at the k-th iteration, equations (16) are not the Newton equations, strictly speaking. However, in this paper, we call (16) the Newton equations for simplicity.

The following lemma gives a sufficient condition for equation (16) to be solvable.

Lemma 1 *If the matrix $G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p$ is positive definite and the matrix $A(x)$ is of full rank, then the matrix $J(w)$ is nonsingular.*

Proof. Consider the equation

$$J(w) \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = 0,$$

for $(v_x, v_y, v_z)^t \in \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n$. Since the equation above gives

$$v_z = -T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p v_x,$$

by eliminating v_z, we have

$$v_x = (G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p)^{-1} A(x)^t v_y.$$

The condition $A(x) v_x = 0$ yields

$$A(x)(G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p)^{-1} A(x)^t v_y = 0.$$

Since the matrix $G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p$ is positive definite and the matrix $A(x)$ is of full rank, we have $v_y = 0$. This implies that $v_x = v_z = 0$. Therefore the proof is complete. □

It is known that if p_k is chosen to make $\mathrm{Arw}(\tilde{x}_k)$ and $\mathrm{Arw}(\tilde{z}_k)$ commute, then the matrix $T_{p_k} \mathrm{Arw}(\tilde{x}_k)^{-1} \mathrm{Arw}(\tilde{z}_k) T_{p_k}$ becomes symmetric positive definite. In this case, if we choose a symmetric positive semidefinite matrix G_k, the matrix $G_k + T_{p_k} \mathrm{Arw}(\tilde{x}_k)^{-1} \mathrm{Arw}(\tilde{z}_k) T_{p_k}$ is symmetric positive definite. This is true for the choices of $p_k = x_k^{-1/2}$ and $p_k = z_k^{1/2}$, which are introduced in Section 2. Furthermore, if p_k is chosen to be the Nesterov-Todd direction (9), then we have $\mathrm{Arw}(\tilde{x}_k)^{-1} \mathrm{Arw}(\tilde{z}_k) = I$ and the matrix $T_{p_k} \mathrm{Arw}(\tilde{x}_k)^{-1} \mathrm{Arw}(\tilde{z}_k) T_{p_k}$ becomes the symmetric positive definite matrix $T_{p_k}^2$. These facts justify the assumption of the previous lemma.

The following lemma claims that a BKKT point is obtained if the Newton direction satisfies $\Delta x = 0$.

Lemma 2 *Assume that Δw solves* (16). *If $\Delta x = 0$, then $(x, y + \Delta y, z + \Delta z)$ is a BKKT point.*

Proof. It follows from the Newton equations that

$$\begin{aligned} \nabla f(x) - A(x)^t (y + \Delta y) - (z + \Delta z) &= 0, \\ g(x) &= 0, \\ (T_p x) \circ (T_p^{-1} \Delta z) &= \mu e - (T_p x) \circ (T_p^{-1} z). \end{aligned}$$

Since the last equation yields $T_p x \circ T_p^{-1}(z + \Delta z) = \mu e$, we have that $x \circ (z + \Delta z) = \mu e$, and then $z + \Delta z = \mu x^{-1} \succ 0$. Therefore the point $(x, y + \Delta y, z + \Delta z)$ satisfies the BKKT conditions. □

4.2 The primal-dual merit function

To force the global convergence of the algorithm described in this paper, we use a merit function in the primal-dual space. For this purpose, we propose the following merit function:

$$F(x, z) = F_{BP}(x) + \nu F_P(x, z), \tag{17}$$

where $F_{BP}(x)$ and $F_P(x,z)$ are the barrier penalty function and the potential function, respectively, and they are given by

$$
\begin{aligned}
(18) \qquad F_{BP}(x) &= f(x) - \frac{\mu}{2}\sum_{i=1}^{s}\log(\det(x^i)) + \rho\|g(x)\|_1, \\
(19) \qquad F_P(x,z) &= (s+\sigma)\log(\frac{x^tz}{s} + |\frac{x^tz}{s} - \mu|) - \frac{1}{2}\sum_{i=1}^{s}\log(\det(x^i)\det(z^i)),
\end{aligned}
$$

where ν, ρ and σ are positive parameters. The following lemma gives a lower bound on the value of the potential function, and the behavior of the function when $x^tz \downarrow 0$ and $x^tz \uparrow \infty$.

Lemma 3 *The potential function satisfies*

$$(20) \qquad F_P(x,z) \geq \sigma \log \mu.$$

The equality holds in (20) *if and only if the vectors x and z satisfies the relation $x \circ z = \mu e$. Furthermore*

$$(21) \qquad \lim_{x^tz\downarrow 0} F_P(x,z) = \infty, \quad \lim_{x^tz\uparrow\infty} F_P(x,z) = \infty$$

Proof. Noting that $\tilde{x}^t\tilde{z} = x^tz$ and $\det(\tilde{x}^i)\det(\tilde{z}^i) = \det(p^i)^2\det(x^i)\cdot\det(p^i)^{-2}\det(z^i) = \det(x^i)\det(z^i)$ (see Theorem 8 in [1]), we have $F_P(\tilde{x},\tilde{z}) = F_P(x,z)$. Let the eigenvalues of $\tilde{x}^i$ and $\tilde{z}^i$ be λ_1^i, λ_2^i and τ_1^i, τ_2^i, respectively. Since $\tilde{x} \succ 0$ and $\tilde{z} \succ 0$ are satisfied and $\mathrm{Arw}(\tilde{x})$ and $\mathrm{Arw}(\tilde{z})$ commute, these eigenvalues are positive and the Jordan frame of $\tilde{x}^i$ and $\tilde{z}^i$, c_1^i and c_2^i say, is shared as stated in Section 2. Then $\tilde{x}^i$ and $\tilde{z}^i$ are written as

$$\tilde{x}^i = \lambda_1^ic_1^i + \lambda_2^ic_2^i \quad \text{and} \quad \tilde{z}^i = \tau_1^ic_1^i + \tau_2^ic_2^i,$$

and we have $\tilde{x}^t\tilde{z} = \sum_{i=1}^{s}(\tilde{x}^i)^t\tilde{z}^i = \frac{1}{2}\sum_{i=1}^{s}(\lambda_1^i\tau_1^i + \lambda_2^i\tau_2^i)$, $\det(\tilde{x}^i) = \lambda_1^i\lambda_2^i$ and $\det(\tilde{z}^i) = \tau_1^i\tau_2^i$. Thus it follows from the algebraic and geometric mean

$$\sum_{i=1}^{s}\frac{\lambda_1^i\tau_1^i + \lambda_2^i\tau_2^i}{2s} \geq \left(\prod_{i=1}^{s}\lambda_1^i\tau_1^i\lambda_2^i\tau_2^i\right)^{\frac{1}{2s}}$$

that

$$(22) \qquad \frac{x^tz}{s} \geq \left(\prod_{i=1}^{s}\det(x^i)\det(z^i)\right)^{\frac{1}{2s}}.$$

The equality holds in (22) if and only if the equality holds in the algebraic and geometric mean. This implies that

$$(23) \qquad \lambda_1^1\tau_1^1 = \lambda_2^1\tau_2^1 = \cdots = \lambda_1^s\tau_1^s = \lambda_2^s\tau_2^s.$$

From (19) and (22), we have

$$(24) \qquad F_P(x,z) \geq (s+\sigma)\log(\frac{x^tz}{s} + |\frac{x^tz}{s} - \mu|) - s\log(\frac{x^tz}{s}).$$

To see the behavior of the function in the right hand side, we introduce the variable $t = x^t z/s$, and define

$$\phi(t) = (s+\sigma)\log(t + |t-\mu|) - s\log t, \quad t > 0.$$

For $0 < t \le \mu$, we have

$$\phi(t) = (s+\sigma)\log\mu - s\log t. \tag{25}$$

In this region, $\phi(t)$ is convex and monotonically decreasing. We note $\phi(\mu) = \sigma\log\mu$. For $t > \mu$, we have

$$\phi(t) = (s+\sigma)\log(2t-\mu) - s\log t,$$

and

$$\phi'(t) = \frac{2(s+\sigma)}{2t-\mu} - \frac{s}{t} = \frac{2\sigma t + \mu s}{t(2t-\mu)} > 0.$$

Thus $\phi(t)$ is monotonically increasing in this region. Therefore $\phi(t)$ attains its unique minimum at $t = \mu$, and the minimum value is $\phi(\mu) = \sigma\log\mu$. This means that

$$F_P(x,z) \ge \sigma\log\mu.$$

The equality holds if and only if $x^t z/s = \mu$ and (23) hold. These two conditions are equivalent to

$$\lambda_1^1\tau_1^1 = \lambda_2^1\tau_2^1 = \cdots = \lambda_1^s\tau_1^s = \lambda_2^s\tau_2^s = \mu. \tag{26}$$

The condition (26) means

$$\tilde{x}^i \circ \tilde{z}^i = \lambda_1^i\tau_1^i c_1^i \circ c_1^i + \lambda_2^i\tau_2^i c_2^i \circ c_2^i = \mu(c_1^i + c_2^i) = \mu e^i$$

which implies that $x \circ z = \mu e$. Conversely, if we assume $x \circ z = \mu e$, then we have $\tilde{x}^i \circ \tilde{z}^i = \mu e^i$ for each i, i.e.,

$$\tilde{x}^i \circ \tilde{z}^i = \lambda_1^i\tau_1^i c_1^i + \lambda_2^i\tau_2^i c_2^i = \mu e^i = \mu(c_1^i + c_2^i),$$

and then (26).

The limits (21) are apparent, because of (25) for $0 < t \le \mu$, and

$$\phi(t) = \sigma\log(2t-\mu) + s\log(2 - \frac{\mu}{t})$$

for $t > \mu$.

This completes the proof. □

It is known that

$$\nabla_v(\log\det(v)) = 2v^{-1} \quad \text{for } v \succ 0.$$

Then we introduce the first order approximation F_l of the merit function by

$$F_l(x,z;\Delta x,\Delta z) = F(x,z) + \Delta F_l(x,z;\Delta x,\Delta z), \tag{27}$$

where

$$
\begin{array}{rrcl}
 & \Delta F_l(x,z;\Delta x,\Delta z) & = & \Delta F_{BPl}(x;\Delta x) + \nu \Delta F_{Pl}(x,z;\Delta x,\Delta z), \\
(28) & \Delta F_{BPl}(x;\Delta x) & = & \nabla f(x)^t \Delta x - \mu (x^{-1})^t \Delta x \\
 & & & + \rho \left(\|g(x) + A(x)\Delta x\|_1 - \|g(x)\|_1 \right), \\
(29) & \Delta F_{Pl}(x,z;\Delta x,\Delta z) & = & (s+\sigma) \left\{ \dfrac{(z^t \Delta x + x^t \Delta z)}{s} + \left| \dfrac{(x^t z + z^t \Delta x + x^t \Delta z)}{s} - \mu \right| \right. \\
 & & & \left. - \left| \dfrac{x^t z}{s} - \mu \right| \right\} \Big/ \left\{ \dfrac{x^t z}{s} + \left| \dfrac{x^t z}{s} - \mu \right| \right\} \\
 & & & - \left((x^{-1})^t \Delta x + (z^{-1})^t \Delta z \right).
\end{array}
$$

We now show that the search direction is a descent direction for both the barrier penalty function and the potential function. We first give an estimate of $\Delta F_{BPl}(x;\Delta x)$ for the barrier-penalty function.

Lemma 4 *Assume that Δw solves* (16)*. Then the following holds*

$$
\begin{array}{rrcl}
(30) & \Delta F_{BPl}(x;\Delta x) & \leq & -\Delta x^t \left(G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p \right) \Delta x \\
 & & & -(\rho - \|y + \Delta y\|_\infty) \|g(x)\|_1.
\end{array}
$$

Proof. It is clear from (14) and (28) that

$$
(31) \qquad \Delta F_{BPl}(x;\Delta x) = \nabla f(x)^t \Delta x - \mu (x^{-1})^t \Delta x - \rho \|g(x)\|_1.
$$

It follows from (13) that

$$
\nabla f(x)^t \Delta x = -\Delta x^t G \Delta x + \Delta x^t A(x)^t (y + \Delta y) + \Delta x^t (z + \Delta z).
$$

Since $T_p \mathrm{Arw}(\tilde{x})^{-1} e = x^{-1}$ holds from (7), equation (15) implies that

$$
\begin{array}{rcl}
z + \Delta z & = & T_p \mathrm{Arw}(\tilde{x})^{-1} (\mu e - \mathrm{Arw}(\tilde{z}) T_p \Delta x) \\
 & = & \mu x^{-1} - T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p \Delta x.
\end{array}
$$

Then we have

$$
\nabla f(x)^t \Delta x = -\Delta x^t \left(G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p \right) \Delta x - g(x)^t (y + \Delta y) + \mu \Delta x^t x^{-1}.
$$

Therefore equation (31) yields

$$
\begin{array}{rcl}
\Delta F_{BPl}(x;\Delta x) & = & -\Delta x^t \left(G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p \right) \Delta x - g(x)^t (y + \Delta y) \\
 & & + \mu \Delta x^t x^{-1} - \mu (x^{-1})^t \Delta x - \rho \|g(x)\|_1 \\
 & \leq & -\Delta x^t \left(G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p \right) \Delta x \\
 & & -(\rho - \|y + \Delta y\|_\infty) \|g(x)\|_1.
\end{array}
$$

The proof is complete. □

Next we estimate the difference $\Delta F_{Pl}(x,z;\Delta x,\Delta z)$ for the potential function.

Lemma 5 *Assume that Δw solves* (16). *Then the following holds*

$$\Delta F_{Pl}(x,z;\Delta x,\Delta z) \leq 0. \tag{32}$$

The equality holds in (32) *if and only if the vectors x and z satisfies the relation $x \circ z = \mu e$.*

Proof. Equation (12) yields

$$(T_p^{-1}z)^t T_p \Delta x + (T_p x)^t T_p^{-1} \Delta z = \mu s - (T_p x)^t T_p^{-1} z$$

and

$$z^t \Delta x + x^t \Delta z = \mu s - x^t z.$$

Since matrices $\mathrm{Arw}(\tilde{x})$ and $\mathrm{Arw}(\tilde{z})$ commute, premultiplying (15) by $e^t \mathrm{Arw}(\tilde{x})^{-1}\mathrm{Arw}(\tilde{z})^{-1}$ implies

$$e^t \mathrm{Arw}(\tilde{x})^{-1} T_p \Delta x + e^t \mathrm{Arw}(\tilde{z})^{-1} T_p^{-1} \Delta z = \mu e^t \mathrm{Arw}(\tilde{x})^{-1}\mathrm{Arw}(\tilde{z})^{-1} e - e^t e.$$

Using (7) this yields

$$(T_p^{-1}x^{-1})^t T_p \Delta x + (T_p z^{-1})^t T_p^{-1} \Delta z = -s + \mu (T_p^{-1}x^{-1})^t T_p z^{-1},$$

and then

$$(x^{-1})^t \Delta x + (z^{-1})^t \Delta z = -s + \mu (x^{-1})^t z^{-1}.$$

Thus from (29) we obtain

$$\begin{aligned} \Delta F_{Pl}(x,z;\Delta x,\Delta z) &= (s+\sigma)\frac{(-x^t z/s + \mu) - |x^t z/s - \mu|}{x^t z/s + |x^t z/s - \mu|} - \left(-s + \mu (x^{-1})^t z^{-1}\right) \\ &= -\sigma + \frac{(s+\sigma)\mu}{x^t z/s + |x^t z/s - \mu|} - \mu (x^{-1})^t z^{-1}. \end{aligned} \tag{33}$$

We use the spectral decomposition of $\tilde{x}^i$ and $\tilde{z}^i$ as in the proof of Lemma 3. Then $(\tilde{x}^i)^{-1}$ and $(\tilde{z}^i)^{-1}$ are written as

$$(\tilde{x}^i)^{-1} = (\lambda_1^i)^{-1} c_1^i + (\lambda_2^i)^{-1} c_2^i \quad \text{and} \quad (\tilde{z}^i)^{-1} = (\tau_1^i)^{-1} c_1^i + (\tau_2^i)^{-1} c_2^i.$$

Therefore we obtain

$$\begin{aligned} \frac{(x^{-1})^t z^{-1}}{s} &= \frac{(\tilde{x}^{-1})^t \tilde{z}^{-1}}{s} = \sum_{i=1}^{s} \frac{((\tilde{x}^i)^{-1})^t (\tilde{z}^i)^{-1}}{s} \\ &= \sum_{i=1}^{s} \frac{(\lambda_1^i \tau_1^i)^{-1} + (\lambda_2^i \tau_2^i)^{-1}}{2s} \geq \left(\prod_{i=1}^{s} \lambda_1^i \tau_1^i \lambda_2^i \tau_2^i \right)^{-1/2s} \\ &\geq \frac{2s}{\sum_{i=1}^{s} \lambda_1^i \tau_1^i + \lambda_2^i \tau_2^i} = \frac{s}{\tilde{x}^t \tilde{z}} = \frac{s}{x^t z}. \end{aligned}$$

Thus from (33)

$$\begin{aligned}
\Delta F_{Pl}(x, z; \Delta x, \Delta z) &\leq -\sigma + \frac{(s+\sigma)\mu}{x^t z/s + |x^t z/s - \mu|} - \frac{\mu s^2}{x^t z} \\
&= -\sigma + \frac{(s+\sigma)\mu}{t + |t-\mu|} - \frac{\mu s}{t}, \quad \text{where } t = x^t z/s > 0 \\
&= -\sigma \frac{(t-\mu) + |t-\mu|}{t + |t-\mu|} - \mu s \left(\frac{1}{t} - \frac{1}{t + |t-\mu|} \right) \leq 0.
\end{aligned}$$

The equalities hold if and only if $\lambda_1^1 \tau_1^1 = \lambda_2^1 \tau_2^1 = \cdots = \lambda_1^s \tau_1^s = \lambda_2^s \tau_2^s$ and $x^t z/s = \mu$ as in the proof of Lemma 3.

Therefore the proof is complete. □

Now we obtain the following theorem by using the two lemmas given above. This theorem shows that the Newton direction Δw becomes a descent search direction for the proposed primal-dual merit function in (17).

Theorem 2 *Assume that Δw solves* (16) *and that the matrix $G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p$ is positive definite. Suppose that the penalty parameter ρ satisfies $\rho > \|y + \Delta y\|_\infty$. Then the following hold:*
(i) *The direction Δw becomes a descent search direction for the primal-dual merit function $F(x, z)$, i.e. $\Delta F_l(x, z; \Delta x, \Delta z) \leq 0$.*
(ii) *If $\Delta x \neq 0$, then $\Delta F_l(x, z; \Delta x, \Delta z) < 0$.*
(iii) *$\Delta F_l(x, z; \Delta x, \Delta z) = 0$ holds if and only if $(x, y + \Delta y, z)$ is a BKKT point.*

Proof. (i) and (ii) : It follows directly from Lemmas 4 and 5 that

$$\begin{aligned}
\Delta F_l(x, z; \Delta x, \Delta z) &\leq -\Delta x^t \left(G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p \right) \Delta x \\
&\qquad -(\rho - \|y + \Delta y\|_\infty) \|g(x)\|_1 \\
&\leq 0.
\end{aligned} \tag{34}$$

The last inequality becomes a strict inequality if $\Delta x \neq 0$. Therefore the results hold.
(iii) If $\Delta F_l(x, z; \Delta x, \Delta z) = 0$ holds, then $\Delta F_{BPl}(x; \Delta x) = 0$ and $\Delta F_{Pl}(x, z; \Delta x, \Delta z) = 0$ are satisfied, and equation (34) yields

$$\Delta x = 0, \quad g(x) = 0. \tag{35}$$

Since $\Delta F_{Pl}(x, z; \Delta x, \Delta z) = 0$, Lemma 5 gives $x \circ z = \mu e$. Since equation (15) yields $\mathrm{Arw}(\tilde{x}) T_p^{-1} \Delta z = 0$, we have $\Delta z = 0$. Then equation (13) implies that $\nabla f(x) - A(x)^t (y + \Delta y) - z = 0$. Hence $(x, y + \Delta y, z)$ is a BKKT point.

Conversely, suppose that $(x, y + \Delta y, z)$ is a BKKT point. The Newton equations imply that

$$G\Delta x - \Delta z = 0, \quad \text{and} \quad \mathrm{Arw}(\tilde{z}) T_p \Delta x + \mathrm{Arw}(\tilde{x}) T_p^{-1} \Delta z = 0.$$

It follows that $(G + T_p \mathrm{Arw}(\tilde{x})^{-1} \mathrm{Arw}(\tilde{z}) T_p) \Delta x = 0$ holds, which yields $\Delta x = 0$. Using equation (31) and Lemma 5, we have

$$\Delta F_{BPl}(x; \Delta x) = 0 \quad \text{and} \quad \Delta F_{Pl}(x, z; \Delta x, \Delta z) = 0,$$

which implies $\Delta F_l(x, z; \Delta x, \Delta z) = 0$. Therefore, the theorem is proved. $\square$

We close this subsection by giving a lemma that gives a basis for Armijo's line search rule and its convergence described in the next section. This lemma corresponds to Lemma 2 and Lemma 3 of the paper by Yamashita [12], so we omit the proof.

Lemma 6 *Let $d_x \in \mathbf{R}^n$ and $d_z \in \mathbf{R}^n$ be given. Define $F'(x, z; d_x, d_z)$ by*

$$F'(x, z; d_x, d_z) = \lim_{t \downarrow 0} \frac{F(x + td_x, z + td_z) - F(x, z)}{t}.$$

Then the following hold:
(i) *The function $F_l(x, z; \alpha d_x, \alpha d_z)$ is convex with respect to the variable α.*
(ii) *The relation*

$$F(x, z) + F'(x, z; d_x, d_z) \leq F_l(x, z; d_x, d_z)$$

holds.
(iii) *There exists a $\theta \in (0, 1)$ such that*

$$F(x + d_x, z + d_z) \leq F(x, z) + F'(x + \theta d_x, z + \theta d_z; d_x, d_z),$$

whenever $x + d_x \succ 0$ and $z + d_z \succ 0$.
(iv) *Let $\varepsilon_0 \in (0, 1)$ be given. If $\Delta F_l(x, z; d_x, d_z) < 0$, then*

$$F(x + \alpha d_x, z + \alpha d_z) - F(x, z) \leq \varepsilon_0 \alpha \Delta F_l(x, z; d_x, d_z),$$

for sufficiently small $\alpha > 0$. $\square$

4.3 The line search algorithm

To obtain a globally convergent algorithm to a BKKT point for a fixed $\mu > 0$, we modify the basic Newton iteration. Our iterations take the form

$$x_{k+1} = x_k + \alpha_k \Delta x_k, \quad z_{k+1} = z_k + \alpha_k \Delta z_k \quad \text{and} \quad y_{k+1} = y_k + \Delta y_k$$

where α_k is a step size determined by the line search procedure described below.

The main iteration is to decrease the value of the merit function $F(x, z)$ for fixed μ. Thus the step size is determined by the sufficient decrease rule of the merit function. We adopt Armijo's rule. At the point (x_k, z_k), we calculate the maximum allowed step to the boundary of the feasible region by

$$\alpha_{xk\max} = \operatorname{argmin} \left\{ \det(x_k^i + \alpha \Delta x_k^i) = 0,\ (x_k^i)_0 + \alpha (\Delta x_k^i)_0 \geq 0,\ i = 1, \ldots, s,\ \alpha > 0 \right\}$$

and

$$\alpha_{zk\max} = \operatorname{argmin} \left\{ \det(z_k^i + \alpha \Delta z_k^i) = 0,\ (z_k^i)_0 + \alpha (\Delta z_k^i)_0 \geq 0,\ i = 1, \ldots, s,\ \alpha > 0 \right\}.$$

Specifically, the equation $\det(x_k^i + \alpha \Delta x_k^i) = 0$ implies the quadratic equation of α

$$\det(\Delta x_k^i)\alpha^2 + 2(x_k^i)^t R_i \Delta x_k^i \alpha + \det(x_k^i) = 0.$$

Thus we can easily get $\alpha_{xk\max}$, and we obtain $\alpha_{zk\max}$ in a similar way. The step sizes are set to be infinity if there is no step size that satisfies these conditions.

A step to the next iterate is given by

$$\alpha_k = \bar{\alpha}_k \beta^{l_k}, \qquad \bar{\alpha}_k = \min\left\{\gamma\alpha_{xk\max}, \gamma\alpha_{zk\max},\ 1\right\},$$

where $\gamma \in (0,1)$ and $\beta \in (0,1)$ are fixed constants and l_k is the smallest nonnegative integer such that

$$F(x_k + \bar{\alpha}_k\beta^{l_k}\Delta x_k, z_k + \bar{\alpha}_k\beta^{l_k}\Delta z_k) \le F(x_k, z_k) + \varepsilon_0\bar{\alpha}_k\beta^{l_k}\Delta F_l(x_k, z_k; \Delta x_k, \Delta z_k), \tag{36}$$

where $\varepsilon_0 \in (0,1)$.

Now we give a line search algorithm called Algorithm SOCPLS below. This algorithm should be regarded as the inner iteration of Algorithm SOCPIP (see Step 1 of Algorithm SOCPIP). We also note that ε' given below corresponds to $M_c\mu$ in Algorithm SOCPIP.

Algorithm SOCPLS

Step 0. (Initialize) Let $w_0 \in \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n$ ($x_0 \succ 0, z_0 \succ 0$), and $\mu > 0$, $\rho > 0$, $\rho' > 0$, $\nu > 0$. Set $\varepsilon' > 0$, $\gamma \in (0,1)$, $\beta \in (0,1)$ and $\varepsilon_0 \in (0,1)$. Let $k = 0$.

Step 1. (Termination) If $\|r(w_k, \mu)\| \le \varepsilon'$, then stop.

Step 2. (Compute direction) Calculate the matrix G_k and the vector p_k. Determine the direction Δw_k by solving (16).

Step 3. (Step size) Find the smallest nonnegative integer l_k that satisfies the criterion (36), and calculate

$$\alpha_k = \bar{\alpha}_k\beta^{l_k}.$$

Step 4. (Update variables) Set

$$\begin{aligned} \begin{pmatrix} x_{k+1} \\ z_{k+1} \end{pmatrix} &= \begin{pmatrix} x_k \\ z_k \end{pmatrix} + \alpha_k \begin{pmatrix} \Delta x_k \\ \Delta z_k \end{pmatrix}, \\ y_{k+1} &= y_k + \Delta y_k. \end{aligned}$$

Step 5. Set $k := k+1$ and go to Step 1. □

4.4 Global convergence

Now we prove global convergence of Algorithm SOCPLS. For this purpose, we make the following assumptions.

Assumptions

(A1) The functions f and $g_i, i = 1, ..., m$, are twice continuously differentiable.

(A2) The sequence $\{x_k\}$ generated by Algorithm SOCPLS remains in a compact set Ω of $\mathbf{R}^n$.

(A3) The matrix $A(x_k)$ is of full rank for all x_k in Ω.

(A4) The matrix G_k is positive semidefinite and uniformly bounded.

(A5) The vector p_k is so chosen that $\mathrm{Arw}(\tilde{x}_k)$ and $\mathrm{Arw}(\tilde{z}_k)$ commute. The sequence $\{p_k\}$ is bounded, and $\liminf_{k\to\infty} \det(p_k) > 0$.

(A6) The penalty parameter ρ is sufficiently large so that $\rho > \|y_k + \Delta y_k\|_\infty$ holds for all k. □

Remarks (i) The compactness of the generated sequence $\{x_k\}$ in (A2) is derived if we assume the compactness of the level set of the function $F_{BP}(x)$ at the initial point, for example, because the iterates give decreasing merit function values, and $F_P(x, z) \geq \sigma \log \mu$. Another case which automatically assures the compactness of $\{x_k\}$ is when all primal variables have finite upper bounds. This is not so uncommon in practical applications.

(ii) We should note that if a quasi-Newton approximation is used for computing the matrix G_k, then we only need the continuity of the first order derivatives of functions in Assumption (A1).

(iii) It will be shown after Lemma 7 that Assumption (A5) is valid for well known examples of p_k.

(iv) In practice, the value of ρ should be updated in the course of computation to satisfy the condition in Assumption (A6). One such procedure is described in 5.1.7 of [14] in which a primal-dual interior point method for general nonlinear optimization problems is proposed and tested numerically. In this paper we assume the above for simplicity of exposition. □

Lemma 7 *Let an infinite sequence* $\{w_k\}$ *be generated by Algorithm SOCPLS. Suppose that Assumptions* (A1), (A2) *and* (A6) *hold. Then the following hold.*
(i) $\liminf_{k\to\infty} \det(x_k) > 0$ *and* $\liminf_{k\to\infty} \det(z_k) > 0$.
(ii) *The sequence* $\{w_k\}$ *is bounded.*
Suppose further that Assumptions (A3), (A4) *and* (A5) *hold. Then the following hold.*
(iii) *There exists a positive constant* M *such that*

$$\frac{1}{M}\|v\|^2 \leq v^t(G_k + T_{p_k}\mathrm{Arw}(\tilde{x}_k)^{-1}\mathrm{Arw}(\tilde{z}_k)T_{p_k})v \leq M\|v\|^2 \tag{37}$$

for any $v \in \mathbf{R}^n$.
(iv) *The sequence* $\{\Delta w_k\}$ *is bounded.*

Proof. (i) Since $\{F_P(x_k, z_k)\}$ is bounded below from Lemma 3, $\{F_{BP}(x_k)\}$ is bounded above because of descent property of $\{F(x_k, z_k)\}$. Therefore $\det(x_k)$ is bounded away from zero because of the log barrier term in $F_{BP}(x)$, and $\liminf_{k\to\infty} \det(x_k) > 0$. Then we also have $\liminf_{k\to\infty} \det(z_k) > 0$, because $\{F_P(x_k, z_k)\}$ is bounded above and below.

(ii) Let the spectral decomposition of z_k^i be

$$z_k^i = \kappa_{k1}^i c_{k1}^i + \kappa_{k2}^i c_{k2}^i,$$

where $\kappa_{1k}^i, \kappa_{2k}^i$ are the eigenvalues and c_{k1}^i, c_{k2}^i are the Jordan frame of the vector z_k^i. We have

$$\begin{aligned} x_k^t z_k &= \sum_{i=1}^{s} (\kappa_{k1}^i x_k^t c_{k1}^i + \kappa_{k2}^i x_k^t c_{k2}^i) \\ &= \frac{1}{2} \sum_{i=1}^{s} \kappa_{k1}^i \left(x_{k0} + \frac{\bar{x}_k^t \bar{z}_k^i}{\|\bar{z}_k^i\|} \right) + \frac{1}{2} \sum_{i=1}^{s} \kappa_{k2}^i \left(x_{k0} - \frac{\bar{x}_k^t \bar{z}_k^i}{\|\bar{z}_k^i\|} \right). \end{aligned} \tag{38}$$

From Cauchy-Schwarz inequality, we have

$$x_{k0} \pm \frac{\bar{x}_k^t \bar{z}_k^i}{\|\bar{z}_k^i\|} \geq x_{k0} - \|\bar{x}_k^i\| = \frac{\det(x_k)}{x_{k0} + \|\bar{x}_k^i\|}.$$

As shown above, the right hand side of the above inequality is strictly bounded away from zero. If $\limsup_{k\to\infty} \|z_k\| = \infty$, then $\limsup_{k\to\infty} \kappa_{k1}^i = \limsup_{k\to\infty} (z_{k0}^i + \|\bar{z}_k^i\|) = \infty$ for some i. Then from (38), we have $\limsup_{k\to\infty} x_k^t z_k \to \infty$. This is impossible because the merit function is decreasing and $\limsup_{k\to\infty} F_p(x_k, z_k) = \infty$ from Lemma 3. Therefore $\{\|z_k\|\}$ is bounded. From Assumption (A6), the sequence $\{y_k\}$ is bounded, and therefore the sequence $\{w_k\}$ is bounded.

(iii) From Assumption (A5), the bounded sequence $\{x_k\}$ implies the bounded sequence $\{\tilde{x}_k\}$. Therefore the assertions similar to (i) and (ii) also hold for $\{\tilde{x}_k\}$ and $\{\tilde{z}_k\}$. Thus there exists a positive constant M' such that $\|v\|^2/M' \leq v^t(T_{p_k} \mathrm{Arw}(\tilde{x}_k)^{-1} \mathrm{Arw}(\tilde{z}_k) T_{p_k})v \leq M'\|v\|^2$ for any $v \in \mathbf{R}^n$, because we have $1/M_0 \leq \|T_{p_k}\| \leq M_0$ for a positive constant M_0 from Assumption (A5). This implies (37) from Assumption (A4).

(iv) From Lemma 1, (ii) and (iii), the matrix $J(w_k)$ is nonsingular and $\|\Delta w_k\|$ is uniformly bounded. $\square$

We now show that Assumption (A5) holds for the examples (8) and (9). Commutability of $\mathrm{Arw}(\tilde{x}_k)$ and $\mathrm{Arw}(\tilde{z}_k)$ for these examples is well known and does not need further proof. The sequence $\{p_k\}$ is bounded and $\liminf_{k\to\infty} \det(p_k) > 0$, because $\{x_k\}$ and $\{z_k\}$ are bounded, and $\liminf_{k\to\infty} \det(x_k) > 0$ and $\liminf_{k\to\infty} \det(z_k) > 0$ from Assumption (A2) and Lemma 7.

The following theorem gives the global convergence of an infinite sequence generated by Algorithm SOCPLS.

Theorem 3 *Suppose that Assumptions* (A1) – (A6) *hold. Let an infinite sequence* $\{w_k\}$ *be generated by Algorithm SOCPLS. Then there exists at least one accumulation point of* $\{w_k\}$*, and any accumulation point of the sequence* $\{w_k\}$ *is an BKKT point.*

Proof. In the proof, we define the following notations

$$u_k = \begin{pmatrix} x_k \\ z_k \end{pmatrix} \quad \text{and} \quad \Delta u_k = \begin{pmatrix} \Delta x_k \\ \Delta z_k \end{pmatrix}$$

for simplicity. In view of Lemma 2, we can assume $\Delta x_k \neq 0$ for all k. By Lemma 7, the sequence $\{w_k\}$ has at least one accumulation point. From Lemma 7, each component of x_k and z_k is bounded away from the boundary of the second order cone. Hence we have $\liminf_{k\to\infty} \bar{\alpha}_k > 0$.

From (37) and (34), we have

$$\Delta F_l(u_k; \Delta u_k) \leq -\frac{\|\Delta x_k\|^2}{M} < 0, \tag{39}$$

and from (36),

$$\begin{aligned} F(u_{k+1}) - F(u_k) &\leq \varepsilon_0 \bar{\alpha}_k \beta^{l_k} \Delta F_l(u_k; \Delta u_k) \\ &\leq -\varepsilon_0 \bar{\alpha}_k \beta^{l_k} \frac{\|\Delta x_k\|^2}{M} \\ &< 0. \end{aligned} \tag{40}$$

Because the sequence $\{F(u_k)\}$ is decreasing and bounded below, the left-hand side of (40) converges to 0.

We will prove that

$$\lim_{k\to\infty} \Delta F_l(u_k; \Delta u_k) = 0, \tag{41}$$

by contradiction. Suppose that there exists an infinite subsequence $K \subset \{0, 1, \cdots\}$ and a δ such that

$$|\Delta F_l(u_k; \Delta u_k)| \geq \delta > 0, \quad \text{for all } k \in K. \tag{42}$$

Since the fact that the left most expression in (40) tends to zero yields $\beta^{l_k} \to 0$, we have $l_k \to \infty, k \in K$, and therefore we can assume $l_k > 0$ for sufficiently large $k \in K$ without loss of generality. In particular, the point $u_k + \alpha_k \Delta u_k/\beta$ does not satisfy condition (36). Thus, we get

$$F(u_k + \alpha_k \Delta u_k/\beta) - F(u_k) > \varepsilon_0 \alpha_k \Delta F_l(u_k; \Delta u_k)/\beta. \tag{43}$$

By Lemma 6, there exists a $\theta_k \in (0, 1)$ such that for $k \in K$,

$$\begin{aligned} &F(u_k + \alpha_k \Delta u_k/\beta) - F(u_k) \\ &\quad \leq \alpha_k F'(u_k + \theta_k \alpha_k \Delta u_k/\beta; \Delta u_k)/\beta \\ &\quad \leq \alpha_k \Delta F_l(u_k + \theta_k \alpha_k \Delta u_k/\beta; \Delta u_k)/\beta. \end{aligned} \tag{44}$$

Then, from (43) and (44), we see that

$$\varepsilon_0 \Delta F_l(u_k; \Delta u_k) < \Delta F_l(u_k + \theta_k \alpha_k \Delta u_k/\beta; \Delta u_k).$$

This inequality yields

$$\Delta F_l(u_k + \theta_k \alpha_k \Delta u_k/\beta; \Delta u_k) - \Delta F_l(u_k; \Delta u_k) \\ > (\varepsilon_0 - 1)\Delta F_l(u_k; \Delta u_k) > 0. \tag{45}$$

Thus by the property $l_k \to \infty$, we have $\alpha_k \to 0$ and thus $\|\theta_k \alpha_k \Delta u_k/\beta\| \to 0, k \in K$, because $\|\Delta u_k\|$ is uniformly bounded above. This implies that the left-hand side of (45) and therefore $\Delta F_l(u_k; \Delta u_k)$ converges to zero when $k \to \infty, k \in K$. This contradicts assumption (42). Therefore we have proved (41).

Since equation (41) implies that

$$\Delta F_{BPl}(x_k; \Delta x_k) \to 0 \quad \text{and} \quad \Delta F_{Pl}(x_k, z_k; \Delta x_k, \Delta z_k) \to 0,$$

it follows from (34), (14) and Lemma 5 that

$$\Delta x_k \to 0, \quad g(x_k) \to 0, \quad x_k \circ z_k \to \mu e \quad (\tilde{x}_k \circ \tilde{z}_k \to \mu e). \tag{46}$$

Therefore, the third equation (15) of the Newton equations yields

$$\lim_{k\to\infty} \|\mathrm{Arw}(\tilde{x}_k) T_{p_k}^{-1} \Delta z_k\| = \lim_{k\to\infty} \|(\mu e - \tilde{x}_k \circ \tilde{z}_k) - \mathrm{Arw}(\tilde{z}_k) T_{p_k} \Delta x_k\| = 0.$$

Since $\{\mathrm{Arw}(\tilde{x}_k)\}$ is uniformly positive definite and $\{T_{p_k}^{-1}\}$ is uniformly bounded, we get

$$\Delta z_k \to 0.$$

By equation (13), we have

$$\nabla_x L(x_k, y_k + \Delta y_k, z_k) \to 0,$$

which implies that

$$r(x_k, y_k + \Delta y_k, z_k, \mu) \to 0.$$

Since $x_{k+1} = x_k + \alpha_k \Delta x_k, z_{k+1} = z_k + \alpha_k \Delta z_k, \Delta x_k \to 0, \Delta z_k \to 0$ and $y_{k+1} = y_k + \Delta y_k$, the result follows. Therefore, the theorem is proved. □

The preceding theorem guarantees that any accumulation point of the sequence $\{(x_k, y_k, z_k)\}$ satisfies the BKKT conditions. If we adopt a common step size α_k as $w_{k+1} = w_k + \alpha_k \Delta w_k$ in Step 4 of Algorithm SOCPLS, where α_k is determined in Step 3, then the result of the theorem is replaced by the statement that any accumulation point of the sequence $\{(x_k, y_k + \Delta y_k, z_k)\}$ satisfies the BKKT conditions.

5 Concluding Remarks

In this paper, we have proposed a primal-dual interior point method for solving nonconvex programming problems over second order cones. Within the line search strategy, we have proposed the primal-dual merit function that consists of the barrier penalty function and the potential function, and we have proved the global convergence property of our method.

If we set $s = n$ and $n_i = 1$, i.e. $\mathcal{K}^i = \{x_i \geq 0\}$, for $i = 1, \ldots, s$, then problem (1) reduces to the usual constrained optimization problem:

$$\begin{array}{lll} \text{minimize} & f(x), & x \in \mathbf{R}^n, \\ \text{subject to} & g(x) = 0, & x \geq 0. \end{array} \tag{47}$$

In this case, the merit function reduces to

$$\begin{aligned} F(x,z) &= F_{BP}(x) + \nu F_P(x,z), \\ F_{BP}(x) &= f(x) - \mu \sum_{i=1}^n \log x_i + \rho \|g(x)\|_1, \\ F_P(x,z) &= (n+\sigma)\log(x^t z/n + \left|x^t z/n - \mu\right|) - \sum_{i=1}^n \log(x_i z_i). \end{aligned} \tag{48}$$

Therefore, as a special case, the results of the present paper include the global convergence property of the usual primal-dual interior point method for solving problem (47) by using the primal-dual merit function (48) within the framework of the line search strategy. This relates to the convergence result by Yamashita and Yabe [13] in which the primal-dual quadratic barrier penalty function was used in the whole space of (x, y, z). In this case, the merit function (48) may be modified as

$$F(x,y,z) = f(x) - \mu \sum_{i=1}^n \log x_i + \rho \|g(x) + \mu y\|_1 + \nu \log \frac{(x^t z/n + |x^t z/n - \mu|)^{n+\sigma}}{\Pi_{i=1}^n x_i z_i},$$

and gives a slightly different form from the one given in [13].

Analysis of the rate of convergence and numerical experiments of our method are under further research. In addition, we plan to construct a method within the framework of the trust region globalization strategy.

References

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

[2] J. F. Bonnans and H. Ramirez, *Perturbation analysis of second-order cone programming problems*, Technical report, INRIA, August 2004.

[3] S. Boyd and L. Vanderberghe, *Convex Optimization*, Cambridge University Press, 2004.

[4] R. Correa and C.H. Ramirez, A global algorithm for nonlinear semidefinite programming, *SIAM Journal on Optimization*, 15 (2004), pp.303-318.

[5] L. Faybusovich, Linear systems in Jordan algebras and primal-dual interior-point algorithms, *Journal of Computational and Applied Mathematics*, 86 (1997), pp.149-175.

[6] R. W. Freund and F. Jarre, *A sensitivity analysis and a convergence result for a sequential semidefinite programming method*, Numerical Analysis Manuscript No. 03-4-08, Bell Laboratories, April 2003.

[7] F. Jarre, *A QQP-minimization method for semidefinite and smooth nonconvex programs*, Technical report, Institute für Angewandte Mathematik und Statistik, Universität Würzburg, August 1998.

[8] H. Kato and M. Fukushima, *An SQP-type algorithm for nonlinear second-order cone programs*, Technical report, Kyoto University, March 2006.

[9] M. Kocvara and M. Stingl, *PENNON : A code for convex nonlinear and semidefinite programming*, Technical Report.

[10] M. S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret, *Applications of second-order cone programing*, Technical report, Stanford University, January 1998.

[11] T. Tsuchiya, A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming, *Optimization Methods and Software*, 11/12 (1999), pp.141-182.

[12] H. Yamashita, A globally convergent primal-dual interior point method for constrained optimization, *Optimization Methods and Software*, 10 (1998), pp.443-469.

[13] H. Yamashita and H. Yabe, An interior point method with a primal-dual quadratic barrier penalty function for nonlinear optimization, *SIAM Journal on Optimization*, 14 (2003), pp.479-499.

[14] H. Yamashita, H. Yabe and T. Tanabe, A globally and superlinearly convergent primal-dual interior point trust region method for large scale constrained optimization, Mathematical Programming, 102 (2005), pp111-151.