# On the Coerciveness of Merit Functions for the Second-Order Cone Complementarity Problem

Guidance

| | |
|---|---|
| Professor | Masao Fukushima |
| Assistant Professor | Nobuo Yamashita |


Shunsuke Hayashi

2000 Graduate Course

in

Department of Applied Mathematics and Physics,

Graduate School of Informatics,

Kyoto University



February 2002

**Abstract**

The Second-Order Cone Complementarity Problem (SOCCP) is a wide class of problems, which includes the Nonlinear Complementarity Problem (NCP) and the Second-Order Cone Programming Problem (SOCP). Recently, Fukushima, Luo and Tseng extended some merit functions and their smoothing functions for NCP to SOCCP. Moreover, they derived computable formulas for the Jacobians of the smoothing functions and gave the conditions for the Jacobians to be invertible. In this paper, we focus on a merit function for SOCCP, and show that the merit function is coercive under the condition that the function involved in SOCCP is strongly monotone. Furthermore, we propose a globally convergent algorithm, which is based on smoothing and regularization methods, for solving merely monotone SOCCP, and examine its effectiveness by means of numerical experiments.

# Contents

# 1 Introduction

In this paper, we consider the *second-order cone complementarity problem* (SOCCP) [8]:

$$\begin{array}{rl} \text{Find} & (x, y, \zeta) \in \Re^n \times \Re^n \times \Re^l \\ \text{such that} & x \in \mathcal{K},\ y \in \mathcal{K},\ x^T y = 0,\ F(x, y, \zeta) = 0, \end{array} \tag{1.1}$$

where $F : \Re^n \times \Re^n \times \Re^l \to \Re^n \times \Re^l$ is a continuously differentiable mapping, and $\mathcal{K} \subset \Re^n$ is the direct product of second-order cones, that is,

$$\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m}$$

with $n = n_1 + \cdots + n_m$ and $n_i$-dimensional second-order cones $\mathcal{K}^{n_i} \subset \Re^{n_i}$ defined by

$$\mathcal{K}^{n_i} = \left\{ (z_1, z_2) \in \Re \times \Re^{n_i - 1} \;\middle|\; \|z_2\| \le z_1 \right\}. \tag{1.2}$$

Here and throughout $\| \cdot \|$ denotes the Euclidean norm.

The SOCCP includes a wide class of problems such as the Nonlinear Complementarity Problem (NCP) and the Second-Order Cone Programming Problem (SOCP) [10]. To see this, consider the SOCCP where $n_1 = n_2 = \cdots = n_m = 1$ and $F(x, y, \zeta) = f(x) - y$ with $f : \Re^n \to \Re^n$. Then this SOCCP is reduced to the NCP : Find $x \in \Re^n$ such that

$$x \ge 0,\ f(x) \ge 0,\ x^T f(x) = 0.$$

On the other hand, consider the SOCP :

$$\begin{array}{rl} \text{Minimize} & \theta(z) \\ \text{subject to} & \gamma(z) \in \mathcal{K}, \end{array} \tag{1.3}$$

where $\theta : \Re^s \to \Re$ and $\gamma : \Re^s \to \Re^t$. Then the Karush-Kuhn-Tucker (KKT) conditions of SOCP (1.3) can be written as

$$\begin{array}{l} \nabla\theta(z) - \nabla\gamma(z)\lambda = 0, \\ \lambda \in \mathcal{K},\ \ \gamma(z) \in \mathcal{K},\ \ \lambda^T \gamma(z) = 0, \end{array} \tag{1.4}$$

where $\lambda \in \Re^t$ is the Lagrange multiplier vector. Then, by setting $n = t$, $l = s$, $x = \lambda$, $y = \mu$, $\zeta = z$ and

$$F(x,\, y,\, \zeta) \;=\; \begin{pmatrix} \mu - \gamma(z) \\ \nabla\theta(z) - \nabla\gamma(z)\lambda \end{pmatrix},$$

the KKT conditions can be reduced to the form of SOCCP (1.1).

For solving SOCCP, Fukushima, Luo and Tseng [8] reformulated SOCCP as an equivalent system of equations, and consider the smoothing methods for them. In this paper, instead of an equivalent system of equations, we focus on the unconstrained optimization problem equivalent to SOCCP (1.1) :

$$\text{Minimize}\ \ \Theta(x, y, \zeta), \tag{1.5}$$

where $\Theta$ is a function from $\Re^n \times \Re^n \times \Re^l$ into $\Re$. The objective function $\Theta$ is called *a merit function* for SOCCP (1.1).

In order to construct a merit function for the SOCCP, it is convenient to introduce a function $\Phi : \Re^n \times \Re^n \to \Re^n$ satisfying

$$\Phi(x, y) = 0 \iff x \in \mathcal{K},\ y \in \mathcal{K},\ x^T y = 0. \tag{1.6}$$

By using such a function, SOCCP (1.1) can be rewritten as the following equivalent system of equations:

$$\Phi(x, y) = 0, \quad F(x, y, \zeta) = 0.$$

Now we define the function $\Psi : \Re^n \times \Re^n \times \Re^l \to \Re$ by

$$\Psi(x, y, \zeta) = \frac{1}{2}\|\Phi(x, y)\|^2 + \frac{1}{2}\|F(x, y, \zeta)\|^2. \tag{1.7}$$

Then, it is easy to see that $\Psi(x, y, \zeta) \geq 0$ for any $(x, y, \zeta) \in \Re^n \times \Re^n \times \Re^l$, and that $\Psi(x, y, \zeta) = 0$ if and only if $(x, y, \zeta)$ is a solution of SOCCP (1.1). Therefore, the function $\Psi$ defined by (1.7) can serve as a merit function for SOCCP (1.1).

Fukushima, Luo and Tseng [8] showed that the natural residual function satisfying (1.6) for the NCP can be extended to the SOCCP by means of Jordan algebra. However, they did not consider the merit function $\Psi$ explicitly, and there remains much to study on it. One of the important questions is under what conditions $\Psi$ is coercive. The function $\Psi : \Re^n \times \Re^n \times \Re^l \to \Re$ is said to be *coercive*, if

$$\lim_{\|(x,y,\zeta)\| \to \infty} \Psi(x, y, \zeta) = \infty.$$

The coerciveness of the merit function is an important property from the viewpoint of optimization. For example, if the merit function $\Psi$ is coercive, then the level sets $\mathcal{L}_\alpha = \{(x, y, \zeta) \mid \Psi(x, y, \zeta) \leq \alpha\}$ are bounded for all $\alpha \in \Re$, which guarantees that a sequence generated by an appropriate descent algorithm applied to the problem (1.5) has an accumulation point. One of the main purposes of this paper is to give a condition for the merit function $\Psi$ given by (1.7) with the natural residual to be coercive. More precisely, we will show that the merit function is coercive when $F$ is strongly monotone. It is well-known that sufficient conditions for coerciveness of merit functions are the strong monotonicity and Lipschitz continuity of $F$ [12]. The result of this paper weakens the conditions.

Another main purpose of this paper is to develop an algorithm for solving the SOCCP. In order to solve the equivalent minimization problem efficiently, we propose to combine a smoothing method with a regularization method. Smoothing methods, which aim to handle nondifferentiablity of functions, have been developed for solving various kinds of complementarity problems [2, 3, 4, 9, 13, 14]. On the other hand, regularization methods are used to deal with ill-posed problems [5, 6]. In this paper, we propose a hybrid method based on these two methods and show that it is globally convergent under the monotonicity assumption on the problem.

The paper is organized as follows. In Section 2, we introduce the spectral factorization for second-order cones, which plays a key role in analyzing the properties of merit functions for the SOCCP. Moreover, we construct a merit function $\Psi$ by using the natural residual for the SOCCP. In Section 3, we show that the merit function $\Psi$ defined in terms of the natural residual is coercive, provided that the function involved in the SOCCP is strongly monotone. In section 4, we propose an algorithm for solving the SOCCP and show that it has global convergence for monotone SOCCPs. In Section 5, we present some numerical results with the proposed algorithm. In Section 6, we conclude the paper with some remarks.

# 2 Preliminaries

## 2.1 Spectral Factorization

In this section, we briefly review some properties of the spectral factorization with respect to a second-order cone, which will be used in the subsequent analysis. Spectral factorization is one of the basic issues of Jordan algebra. For more detail, see [7, 8].

For any vector $x = (x_1, x_2) \in \Re \times \Re^{n-1}$ $(n \geq 2)$, its spectral factorization with respect to the second-order cone $\mathcal{K}^n$ is defined as

$$x = \kappa_1 v_1 + \kappa_2 v_2,$$

where $\kappa_1$ and $\kappa_2$ are *spectral values* given by

$$\kappa_i \;=\; x_1 + (-1)^i \|x_2\|, \quad i = 1, 2, \tag{2.1}$$

and $v_1$ and $v_2$ are *spectral vectors* given by

$$v_i \;=\; \begin{cases} \frac{1}{2}\left(1, (-1)^i \frac{x_2}{\|x_2\|}\right) & (x_2 \neq 0), \\ \frac{1}{2}(1, (-1)^i \omega) & (x_2 = 0), \end{cases} \qquad i = 1, 2, \tag{2.2}$$

with $\omega \in \Re^{n-1}$ such that $\|\omega\| = 1$.

The spectral values and vectors have the following properties. For any $x \in \Re^n$, the inequality $\kappa_1 \leq \kappa_2$ holds, and

$$\kappa_1 \geq 0 \iff x \in \mathcal{K}^n. \tag{2.3}$$

Moreover, for any $x \in \Re^n$, we have $\|v_i\| = 1/\sqrt{2}$ for $i = 1, 2$, and ${v_1}^T v_2 = 0$.

For any $x \in \Re^n$, let $[x]_+^{\mathcal{K}^n}$ denote the projection of $x$ onto the second-order cone $\mathcal{K}^n$, that is,

$$[x]_+^{\mathcal{K}^n} \;=\; \arg\min_{y \in \mathcal{K}^n} \|y - x\|.$$

We often denote $[x]_+^{\mathcal{K}^n}$ as $[x]_+$ when $\mathcal{K}^n$ is clear from the context. In particular, when $n = 1$,

$$[\alpha]_+ = \max(0, \alpha) \tag{2.4}$$

for any $\alpha \in \Re$. For $n \geq 2$, the projection $[\,\cdot\,]_+$ can be also calculated easily as shown in the following proposition.

**Proposition 2.1** [8, Proposition 3.3] *For any $x \in \Re^n$ $(n \geq 2)$,*

$$[x]_+ \;=\; [\kappa_1]_+ v_1 + [\kappa_2]_+ v_2,$$

*where $\kappa_1$ and $\kappa_2$ are the spectral values of $x$ defined by* (2.1)*, and $v_1$ and $v_2$ are the spectral vectors of $x$ defined by* (2.2)*.*

From Proposition 2.1, (2.1), (2.3) and (2.4), we can easily verify that

$$x \in \mathcal{K}^n \iff [x]_+ = x, \tag{2.5}$$

$$x \in -\mathcal{K}^n \iff [x]_+ = 0, \tag{2.6}$$

$$x \notin \mathcal{K}^n \cup -\mathcal{K}^n \implies [x]_+ = \kappa_2 v_2. \tag{2.7}$$

These relation will be useful in showing the coerciveness of a merit function based on the natural residual.

## 2.2 Merit Function

In this section, we aim to construct a merit function for SOCCP (1.1) by using the *natural residual* $\varphi_{\rm NR} : \Re^n \times \Re^n \to \Re^n$ defined in [8]. Firstly, we note that the complementarity conditions on $\mathcal{K}$ can be decomposed into those conditions on $\mathcal{K}^{n_i}$.

**Proposition 2.2** *Let* $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}$, $x = (x^1, \ldots, x^m) \in \Re^{n_1} \times \cdots \times \Re^{n_m}$ *and* $y = (y^1, \ldots, y^m) \in \Re^{n_1} \times \cdots \times \Re^{n_m}$. *Then the following relation holds:*

$$x \in \mathcal{K},\ y \in \mathcal{K},\ x^T y = 0 \quad \text{if and only if} \quad x^i \in \mathcal{K}^{n_i},\ y^i \in \mathcal{K}^{n_i},\ {x^i}^T y^i = 0 \ \ (i = 1, \ldots, m).$$

**Proof.** Since "if" part is evident, we only show "only if" part. Noticing that $x^i, y^i \in \mathcal{K}^{n_i}$ implies $x_1^i \geq \|x_2^i\|$ and $y_1^i \geq \|y_2^i\|$, we have ${x^i}^T y^i = x_1^i y_1^i + {x_2^i}^T y_2^i \geq \|x_2^i\|\|y_2^i\| + {x_2^i}^T y_2^i \geq 0$ for each $i$, where the last inequality follows from Cauchy-Schwarz inequality. On the other hand, $x^T y = 0$ yields $\sum_{i=1}^m {x^i}^T y^i = 0$. Thus ${x^i}^T y^i = 0$ holds for each $i$. ■

This proposition naturally leads us to construct a function $\Phi : \Re^n \times \Re^n \to \Re^n$ satisfying (1.6) as

$$\Phi(x, y) = \begin{pmatrix} \varphi^1(x^1, y^1) \\ \vdots \\ \varphi^m(x^m, y^m) \end{pmatrix},$$

where $\varphi^i : \Re^{n_i} \times \Re^{n_i} \to \Re^{n_i}$ is a function satisfying

$$\varphi^i(x^i, y^i) = 0 \iff x^i \in \mathcal{K}^{n_i},\ y^i \in \mathcal{K}^{n_i},\ {x^i}^T y^i = 0 \tag{2.8}$$

for each $i = 1, \ldots, m$. Fukushima, Luo and Tseng [8] showed that (2.8) holds for the natural residual function $\varphi_{\rm NR} : \Re^{n_i} \times \Re^{n_i} \to \Re^{n_i}$ defined by

$$\varphi_{\rm NR}(x^i, y^i) = x^i - [x^i - y^i]_+^{\mathcal{K}^{n_i}}.$$

This function is a natural extension of the corresponding function for the NCP. Using this function, we define the function $\Phi_{\rm NR} : \Re^n \times \Re^n \to \Re^n$ by

$$\Phi_{\rm NR}(x, y) = \begin{pmatrix} \varphi_{\rm NR}(x^1, y^1) \\ \vdots \\ \varphi_{\rm NR}(x^m, y^m) \end{pmatrix},$$

and then, we can construct a merit function for SOCCP (1.1) as

$$\begin{aligned} \Psi_{\rm NR}(x, y, \zeta) &= \frac{1}{2}\|\Phi_{\rm NR}(x, y)\|^2 + \frac{1}{2}\|F(x, y, \zeta)\|^2 \\ &= \frac{1}{2}\sum_{i=1}^m \|\varphi_{\rm NR}(x^i, y^i)\|^2 + \frac{1}{2}\|F(x, y, \zeta)\|^2. \end{aligned} \tag{2.9}$$

## 3 Coerciveness of Merit Function with Natural Residual

In this section, we focus on the merit function $\Psi_{\mathrm{NR}}$ defined by (2.9), and study conditions for $\Psi_{\mathrm{NR}}$ to be coercive. In the remainder of the paper, we restrict ourselves to the SOCCP in which (i) $F$ is given by $F(x,y,\zeta) = f(x) - y$ with a function $f : \Re^n \to \Re^n$ and (ii) $\mathcal{K} = \mathcal{K}^n$. Then, we can rewrite SOCCP (1.1) as follows: Find $(x,y) \in \Re^n \times \Re^n$ such that

$$x \in \mathcal{K}^n,\ y \in \mathcal{K}^n,\ x^T y = 0,\ y = f(x). \tag{3.1}$$

Note that the assumption (i) implies $F$ does not involve the additional variable $\zeta$. This assumption may seem rather restrictive. However, under the assumption (i), the KKT conditions for SOCP (1.3) can be written in the form of the SOCCP (see Section 6 of this paper). The assumption (ii) implies $\Phi_{\mathrm{NR}}(x,y) = \varphi_{\mathrm{NR}}(x,y)$. This assumption is only for simplicity of presentation, and the results obtained in this section can be extended to the general $\mathcal{K}$ in a straightforward manner.

The merit function $\Psi_{\mathrm{NR}}$ for SOCCP (3.1) can be constructed as

$$\Psi_{\mathrm{NR}}(x,y) = \frac{1}{2}\|\varphi_{\mathrm{NR}}(x,y)\|^2 + \frac{1}{2}\|f(x)-y\|^2. \tag{3.2}$$

Now we investigate the coerciveness of $\Psi_{\mathrm{NR}}$ defined by (3.2). The next proposition says that, to show the coerciveness of $\Psi_{\mathrm{NR}}(x,y)$, it is sufficient to show the coerciveness of $\tilde{\Psi}_{\mathrm{NR}} : \Re^n \to \Re$ defined by

$$\tilde{\Psi}_{\mathrm{NR}}(x) := \|\varphi_{\mathrm{NR}}(x, f(x))\|. \tag{3.3}$$

**Proposition 3.1** *$\Psi_{\mathrm{NR}}$ is coercive if $\tilde{\Psi}_{\mathrm{NR}}$ is coercive.*

**Proof.** Noticing that $\sqrt{2\|\xi\|^2 + 2\|\eta\|^2} \geq \|\xi\| + \|\eta\|$ for any $\xi$, $\eta \in \Re^n$, we have

$$\begin{aligned} 2\sqrt{\Psi_{\mathrm{NR}}(x,y)} &\geq \|x - [x-y]_+\| + \|f(x)-y\| \\ &\geq \|x - [x-f(x)]_+\| - \|[x-f(x)]_+ - [x-y]_+\| + \|f(x)-y\| \\ &\geq \|x - [x-f(x)]_+\| - \|f(x)-y\| + \|f(x)-y\| \\ &= \tilde{\Psi}_{\mathrm{NR}}(x), \end{aligned}$$

where the second inequality follows from the triangle inequality and the third inequality follows from the nonexpansiveness of the projection operator. Hence, $\Psi_{\mathrm{NR}}$ is coercive if $\tilde{\Psi}_{\mathrm{NR}}$ is coercive. ■

In order to investigate conditions for $\tilde{\Psi}_{\mathrm{NR}}$ to be coercive, we need the following five lemmas.

**Lemma 3.1** (a) *If $x - y \in \mathcal{K}^n$, then $\varphi_{\mathrm{NR}}(x,y) = y$.* (b) *If $x - y \in -\mathcal{K}^n$, then $\varphi_{\mathrm{NR}}(x,y) = x$.*

**Proof.** (a) From (2.5) we have $\varphi_{\mathrm{NR}}(x,y) = x - [x-y]_+ = x - (x-y) = y$. (b) From (2.6) we have $\varphi_{\mathrm{NR}}(x,y) = x - [x-y]_+ = x - 0 = x$. ■

**Lemma 3.2** *For any $x = (x_1, x_2) \in \Re \times \Re^{n-1}$ with $x_1 < 0$,*

$$\|\varphi_{\mathrm{NR}}(x,y)\| \geq \frac{1}{\sqrt{2}}\|x\|. \tag{3.4}$$

**Proof.** Let $\kappa_1$ and $\kappa_2$ be the spectral values of $x$ defined by (2.1), and let $v_1$ and $v_2$ be the spectral vectors of $x$ defined by (2.2). Since $\kappa_1 = x_1 - \|x_2\| < 0$ from $x_1 < 0$, we have $[x]_+ = [\kappa_2]_+ v_2$. Furthermore, since $[x]_+$ is the nearest point to $x$ in $\mathcal{K}^n$ and $[x-y]_+ \in \mathcal{K}^n$, we have $\|x - [x-y]_+\| \geq \|x - [x]_+\|$. Hence, we obtain

$$\|\varphi_{\text{NR}}(x,y)\| \;\; \geq \;\; \|x - [\kappa_2]_+ v_2\|\,. \tag{3.5}$$

When $\kappa_2 < 0$, we have $\|x - [\kappa_2]_+ v_2\| = \|x\| \geq (1/\sqrt{2})\|x\|$. When $\kappa_2 \geq 0$, we have $\|x - [\kappa_2]_+ v_2\|^2 = \|x - \kappa_2 v_2\|^2 = \|\kappa_1 v_1\|^2 = (1/2)(x_1^2 + \|x_2\|^2) - x_1\|x_2\| \geq (1/2)\|x\|^2$, where the last inequality holds from $x_1 < 0$. In either case, we have $\|x - [\kappa_2]_+ v_2\| \geq (1/\sqrt{2})\|x\|$. It then follows from (3.5) that (3.4) holds. ∎

The next lemma gives some properties of unbounded sequences $\{x^{(k)}\}$ and $\{y^{(k)}\}$ with $\{\varphi_{\text{NR}}(x^{(k)}, y^{(k)})\}$ being bounded.

**Lemma 3.3** *Assume that sequences $\{x^{(k)}\} \subseteq \Re^n$ and $\{y^{(k)}\} \subseteq \Re^n$ satisfy*

(i) $\lim_{k\to\infty} \|x^{(k)}\| = \infty$ *and* $\lim_{k\to\infty} \|y^{(k)}\| = \infty$;

(ii) $x^{(k)} - y^{(k)} \notin \mathcal{K}^n \cup -\mathcal{K}^n$ *for all* $k$;

(iii) $\{\varphi_{\text{NR}}(x^{(k)}, y^{(k)})\}$ *is bounded.*

*Then there exist a positive constant $M \in (2, \infty)$ and an infinite subsequence $K \subseteq \{0, 1, \ldots\}$ such that*

(a) $-M < |x_1^{(k)}| - \|x_2^{(k)}\| < M$ *and* $-M < |y_1^{(k)}| - \|y_2^{(k)}\| < M$ *for all* $k$;

(b-1) $\{|x_1^{(k)}|\}_{k\in K}$, $\{\|x_2^{(k)}\|\}_{k\in K}$, $\{|y_1^{(k)}|\}_{k\in K}$ *and* $\{\|y_2^{(k)}\|\}_{k\in K}$ *are unbounded*;

(b-2) $|x_1^{(k)}| > 10M$, $\|x_2^{(k)}\| > 10M$, $|y_1^{(k)}| > 10M$ *and* $\|y_2^{(k)}\| > 10M$ *for all* $k \in K$;

(c) $1 + \dfrac{x_1^{(k)} x_2^{(k)}}{|x_1^{(k)}||y_1^{(k)}|} \dfrac{{x_2^{(k)}}^T y_2^{(k)}}{\|x_2^{(k)}\|\|y_2^{(k)}\|} < \dfrac{10M}{9}\left(\dfrac{1}{|x_1^{(k)}|^2} + \dfrac{1}{|y_1^{(k)}|^2}\right) < \dfrac{1}{45M}$ *for all* $k \in K$;

(d) $\sqrt{2} - 0.1 < \dfrac{\|x^{(k)}\|}{|x_1^{(k)}|} < \sqrt{2} + 0.1$ *and* $\sqrt{2} - 0.1 < \dfrac{\|y^{(k)}\|}{|y_1^{(k)}|} < \sqrt{2} + 0.1$ *for all* $k \in K$.

**Proof.** Let $\lambda_1^{(k)}$ and $\lambda_2^{(k)}$ be the spectral values of $x^{(k)} - y^{(k)}$, and let $u_1^{(k)}$ and $u_2^{(k)}$ be the spectral vectors of $x^{(k)} - y^{(k)}$. Then, from assumption (ii) and (2.7) we have

$$\varphi_{\text{NR}}(x^{(k)}, y^{(k)}) \;=\; x^{(k)} - [x^{(k)} - y^{(k)}]_+ \;=\; x^{(k)} - \lambda_2^{(k)} u_2^{(k)}.$$

Moreover, from the definitions (2.1), (2.2) of spectral values and vectors, we have

$$x^{(k)} - \lambda_2^{(k)} u_2^{(k)} \;=\; (x_1^{(k)}, x_2^{(k)}) - \frac{1}{2}\left\{x_1^{(k)} - y_1^{(k)} + \left\|x_2^{(k)} - y_2^{(k)}\right\|\right\}\left(1, \frac{x_2^{(k)} - y_2^{(k)}}{\left\|x_2^{(k)} - y_2^{(k)}\right\|}\right).$$

From these equalities, we have

$$
\begin{aligned}
& \left\|\varphi_{\rm NR}(x^{(k)},y^{(k)})\right\|^2 \\
= \; & \left(|x_1^{(k)}|^2+\|x_2^{(k)}\|^2\right)-\left\{x_1^{(k)}-y_1^{(k)}+\left\|x_2^{(k)}-y_2^{(k)}\right\|\right\}\left\{x_1^{(k)}+(x_2^{(k)})^T\frac{x_2^{(k)}-y_2^{(k)}}{\left\|x_2^{(k)}-y_2^{(k)}\right\|}\right\} \\
& +\frac{1}{2}\left\{x_1^{(k)}-y_1^{(k)}+\left\|x_2^{(k)}-y_2^{(k)}\right\|\right\}^2 \\
= \; & (x_1^{(k)})^2+\|x_2^{(k)}\|^2-\left(x_1^{(k)}-y_1^{(k)}\right)x_1^{(k)}-\left(x_1^{(k)}-y_1^{(k)}\right)x_2^{(k)T}\frac{x_2^{(k)}-y_2^{(k)}}{\left\|x_2^{(k)}-y_2^{(k)}\right\|} \\
& -\left\|x_2^{(k)}-y_2^{(k)}\right\|x_1^{(k)}-x_2^{(k)T}\left(x_2^{(k)}-y_2^{(k)}\right) \\
& +\frac{1}{2}\left(x_1^{(k)}-y_1^{(k)}\right)^2+\left(x_1^{(k)}-y_1^{(k)}\right)\left\|x_2^{(k)}-y_2^{(k)}\right\|+\frac{1}{2}\left\|x_2^{(k)}-y_2^{(k)}\right\|^2 \\
= \; & \left\{(x_1^{(k)})^2+\|x_2^{(k)}\|^2-\left(x_1^{(k)}-y_1^{(k)}\right)x_1^{(k)}-x_2^{(k)T}\left(x_2^{(k)}-y_2^{(k)}\right)+\frac{1}{2}\left(x_1^{(k)}-y_1^{(k)}\right)^2+\frac{1}{2}\left\|x_2^{(k)}-y_2^{(k)}\right\|^2\right\} \\
& -\frac{1}{\left\|x_2^{(k)}-y_2^{(k)}\right\|}\left\{\left(x_1^{(k)}-y_1^{(k)}\right)x_2^{(k)T}\left(x_2^{(k)}-y_2^{(k)}\right)+\left\|x_2^{(k)}-y_2^{(k)}\right\|^2x_1^{(k)}-\left(x_1^{(k)}-y_1^{(k)}\right)\left\|x_2^{(k)}-y_2^{(k)}\right\|^2\right\} \\
= \; & \frac{1}{2}\left\{(x_1^{(k)})^2+\|x_2^{(k)}\|^2+(y_1^{(k)})^2+\|y_2^{(k)}\|^2\right\} \\
& -\frac{1}{\left\|x_2^{(k)}-y_2^{(k)}\right\|}\left(x_2^{(k)}-y_2^{(k)}\right)^T\left(x_1^{(k)}x_2^{(k)}-y_1^{(k)}y_2^{(k)}\right) \\
= \; & \frac{1}{2}\left\{\left(|x_1^{(k)}|-\|x_2^{(k)}\|\right)^2+\left(|y_1^{(k)}|-\|y_2^{(k)}\|\right)^2\right\} \\
& +\left(\left\|x_1^{(k)}x_2^{(k)}\right\|+\left\|y_1^{(k)}y_2^{(k)}\right\|-\left\|x_1^{(k)}x_2^{(k)}-y_1^{(k)}y_2^{(k)}\right\|\right) \\
& +\frac{1}{\left\|x_2^{(k)}-y_2^{(k)}\right\|}\left\{\left\|x_2^{(k)}-y_2^{(k)}\right\|\left\|x_1^{(k)}x_2^{(k)}-y_1^{(k)}y_2^{(k)}\right\|-\left(x_2^{(k)}-y_2^{(k)}\right)^T\left(x_1^{(k)}x_2^{(k)}-y_1^{(k)}y_2^{(k)}\right)\right\}. \qquad (3.6)
\end{aligned}
$$

Now, let $P_k$, $Q_k$ and $R_k$ be defined by

$$
\begin{aligned}
P_k &= \frac{1}{2}\left\{\left(|x_1^{(k)}|-\|x_2^{(k)}\|\right)^2+\left(|y_1^{(k)}|-\|y_2^{(k)}\|\right)^2\right\}, \\
Q_k &= \left\|x_1^{(k)}x_2^{(k)}\right\|+\left\|y_1^{(k)}y_2^{(k)}\right\|-\left\|x_1^{(k)}x_2^{(k)}-y_1^{(k)}y_2^{(k)}\right\|, \\
R_k &= \frac{1}{\left\|x_2^{(k)}-y_2^{(k)}\right\|}\left\{\left\|x_2^{(k)}-y_2^{(k)}\right\|\left\|x_1^{(k)}x_2^{(k)}-y_1^{(k)}y_2^{(k)}\right\|-\left(x_2^{(k)}-y_2^{(k)}\right)^T\left(x_1^{(k)}x_2^{(k)}-y_1^{(k)}y_2^{(k)}\right)\right\}.
\end{aligned}
$$

Then (3.6) can be written as

$$
\left\|\varphi_{\rm NR}(x^{(k)},y^{(k)})\right\|^2 = P_k+Q_k+R_k.
$$

Since $\{\|\varphi_{\rm NR}(x^{(k)},y^{(k)})\|^2\}$ is bounded, there exists a sufficiently large constant $M\in(2,\infty)$ such that

$$
P_k+Q_k+R_k < M
$$

for all $k>0$. Here we note that $P_k$ is nonnegative. By the triangle inequality, $Q_k$ is also nonnegative. Furthermore, $R_k$ is also nonnegative from Cauchy-Schwarz inequality. Thus, we have $P_k<M$, $Q_k<M$ and $R_k<M$ for all $k$.

We first show (a). $P_k < M$ implies that $\max\left\{\left||x_1^{(k)}| - \|x_2^{(k)}\|\right|,\ \left||y_1^{(k)}| - \|y_2^{(k)}\|\right|\right\} < \sqrt{2M} < M$, where the last inequality holds from $M > 2$. Therefore, we have

$$-M < |x_1^{(k)}| - \|x_2^{(k)}\| < M,\ -M < |y_1^{(k)}| - \|y_2^{(k)}\| < M,$$

that is (a).

By using (a), we show (b). From (a), we have $|x_1^{(k)}| - M < \|x_2^{(k)}\|$ and $\|x_2^{(k)}\| - M < |x_1^{(k)}|$. Similarly, we have $|y_1^{(k)}| - M < \|y_2^{(k)}\|$ and $\|y_2^{(k)}\| - M < |y_1^{(k)}|$. Moreover, since $\lim_{k\to\infty}\|x^{(k)}\| = \infty$ and $\lim_{k\to\infty}\|y^{(k)}\| = \infty$, there exists an infinite subsequence $K \subseteq \{0, 1, \ldots\}$ such that $\{|x_1^{(k)}|\}_{k\in K}$, $\{\|x_2^{(k)}\|\}_{k\in K}$, $\{|y_1^{(k)}|\}_{k\in K}$ and $\{\|y_2^{(k)}\|\}_{k\in K}$ are unbounded, and that $|x_1^{(k)}| > 10M$, $\|x_2^{(k)}\| > 10M$, $|y_1^{(k)}| > 10M$ and $\|y_2^{(k)}\| > 10M$ for all $k \in K$, that is (b).

Next we show that (c) holds for all $k \in K$. From $Q_k < M$ and the definition of $Q_k$, we have

$$\left\|x_1^{(k)}x_2^{(k)}\right\| + \left\|y_1^{(k)}y_2^{(k)}\right\| - M < \left\|x_1^{(k)}x_2^{(k)} - y_1^{(k)}y_2^{(k)}\right\|.$$

Noticing that the left-hand side of this inequality is positive for any $k \in K$ from (b-2), we obtain

$$\left(\left\|x_1^{(k)}x_2^{(k)}\right\| + \left\|y_1^{(k)}y_2^{(k)}\right\| - M\right)^2 \;<\; \left\|x_1^{(k)}x_2^{(k)} - y_1^{(k)}y_2^{(k)}\right\|^2.$$

By easy calculation, we have

$$\begin{aligned}
\left\|x_1^{(k)}x_2^{(k)}\right\|\left\|y_1^{(k)}y_2^{(k)}\right\| + x_1^{(k)}y_1^{(k)}{x_2^{(k)}}^T y_2^{(k)} &< M\left(\left\|x_1^{(k)}x_2^{(k)}\right\| + \left\|y_1^{(k)}y_2^{(k)}\right\|\right) - \frac{M^2}{2} \\
&< M\left(\left\|x_1^{(k)}x_2^{(k)}\right\| + \left\|y_1^{(k)}y_2^{(k)}\right\|\right).
\end{aligned}$$

Dividing both sides by $|x_1^{(k)}||y_1^{(k)}|\|x_2^{(k)}\|\|y_2^{(k)}\| > 0$ and using (a), we have

$$\begin{aligned}
1 + \frac{x_1^{(k)}y_1^{(k)}}{|x_1^{(k)}||y_1^{(k)}|}\frac{{x_2^{(k)}}^T y_2^{(k)}}{\|x_2^{(k)}\|\|y_2^{(k)}\|} &< M\left(\frac{1}{|x_1^{(k)}|\|x_2^{(k)}\|} + \frac{1}{|y_1^{(k)}|\|y_2^{(k)}\|}\right) \\
&< M\left\{\frac{1}{|x_1^{(k)}|\left(|x_1^{(k)}| - M\right)} + \frac{1}{|y_1^{(k)}|\left(|y_1^{(k)}| - M\right)}\right\} \\
&< \frac{10M}{9}\left(\frac{1}{|x_1^{(k)}|^2} + \frac{1}{|y_1^{(k)}|^2}\right) \\
&< \frac{1}{45M},
\end{aligned}$$

where the third inequality follows from the fact that $|x_1^{(k)}| - M > (9/10)|x_1^{(k)}|$ and $|y_1^{(k)}| - M > (9/10)|y_1^{(k)}|$ for any $k \in K$, and the last inequality follows from (b-2).

Finally, we show that (d) holds for all $k \in K$. (a) and (b-2) yield the following relations for all $k \in K$:

$$\begin{aligned}
\frac{\|x^{(k)}\|^2}{|x_1^{(k)}|^2} &= \frac{|x_1^{(k)}|^2 + \|x_2^{(k)}\|^2}{|x_1^{(k)}|^2} \\
&< \frac{|x_1^{(k)}|^2 + (|x_1^{(k)}| + M)^2}{|x_1^{(k)}|^2} \\
&= 1 + \left(1 + \frac{M}{|x_1^{(k)}|}\right)^2 \\
&< 1 + (1 + 0.1)^2 \\
&< (\sqrt{2} + 0.1)^2.
\end{aligned}$$

Furthermore, (a) and (b-2) yield the following relations for any $k \in K$ :

$$\begin{aligned} \frac{\|x^{(k)}\|^2}{|x_1^{(k)}|^2} &= \frac{|x_1^{(k)}|^2 + \|x_2^{(k)}\|^2}{|x_1^{(k)}|^2} \\ &> \frac{|x_1^{(k)}|^2 + (|x_1^{(k)}| - M)^2}{|x_1^{(k)}|^2} \\ &= 1 + \left(1 - \frac{M}{|x_1^{(k)}|}\right)^2 \\ &> 1 + (1 - 0.1)^2 \\ &> (\sqrt{2} - 0.1)^2. \end{aligned}$$

Hence, for any $k \in K$,

$$\sqrt{2} - 0.1 \; < \; \frac{\|x^{(k)}\|}{|x_1^{(k)}|} \; < \; \sqrt{2} + 0.1.$$

Similar inequalities hold for $\|y^{(k)}\|/|y_1^{(k)}|$. ∎

Before we show the fourth lemma, we recall the concepts of monotonicity and strong monotonicity of vector-valued functions.

**Definition 3.1** *The function $f : \Re^n \to \Re^n$ is called*

(a) *monotone if, for any $x$, $z \in \Re^n$,*

$$(x - z)^T (f(x) - f(z)) \geq 0;$$

(b) *strongly monotone with modulus $\varepsilon > 0$ if, for any $x$, $z \in \Re^n$,*

$$(x - z)^T (f(x) - f(z)) \geq \varepsilon \|x - z\|^2.$$

It is obvious that a strongly monotone function is monotone. Besides, we note that $f_\varepsilon$ is strongly monotone if $f$ is monotone, where $f_\varepsilon : \Re^n \to \Re^n$ is defined by $f_\varepsilon(x) = f(x) + \varepsilon x$ for $\varepsilon > 0$.

Lemma 3.3 assumed that $\|y^{(k)}\| \to \infty$. The next lemma shows that the assumption holds when $f$ is strongly monotone, $\|x^{(k)}\| \to \infty$ and $y^{(k)} = f(x^{(k)})$.

**Lemma 3.4** *Let $\{x^{(k)}\}$ be an arbitrary sequence satisfying $\lim_{k\to\infty} \|x^{(k)}\| = \infty$. If $f$ is strongly monotone with modulus $\varepsilon > 0$, then there exists $\overline{k} > 0$ such that*

$$\|f(x^{(k)})\| \;\; > \;\; \frac{\varepsilon}{2}\|x^{(k)}\|$$

*holds for all $k > \overline{k}$. Moreover we have $\lim_{k\to\infty} \|y^{(k)}\| = \infty$ when $y^{(k)} = f(x^{(k)})$ for all $k$.*

**Proof.** Setting $x = x^{(k)}$ and $z = 0$ in Definition 3.1 (b), we have

$$(x^{(k)})^T \left(f(x^{(k)}) - f(0)\right) \; \geq \; \varepsilon \|x^{(k)}\|^2.$$

Then, from Cauchy-Schwarz inequality we have

$$\begin{aligned} \varepsilon\|x^{(k)}\|^2 &\leq (x^{(k)})^T\left(f(x^{(k)}) - f(0)\right) \\ &\leq \|x^{(k)}\|\|f(x^{(k)}) - f(0)\| \\ &\leq \|x^{(k)}\|\left(\|f(x^{(k)})\| + \|f(0)\|\right). \end{aligned}$$

Moreover, dividing both sides by $\|x^{(k)}\|$, we have

$$\|f(x^{(k)})\| \geq \varepsilon\|x^{(k)}\| - \|f(0)\|.$$

Since $x^{(k)} \to \infty$, the last inequality implies that there exists $\overline{k}$ such that

$$\|f(x^{(k)})\| > \frac{\varepsilon}{2}\|x^{(k)}\|$$

for all $k > \overline{k}$. ■

We finally present the fifth lemma that gives two properties of infinite sequences. Since it can be easily shown, we omit the proof.

**Lemma 3.5** *Let $\mathcal{N}$ be the set of nonnegative integers and let $\mathcal{N}_1, \ldots, \mathcal{N}_m$ be subsets of $\mathcal{N}$ such that*

(i) $N_i \cap \mathcal{N}_j = \emptyset$ *for any* $i \neq j \in \{1, \ldots, m\}$,

(ii) $\mathcal{N} = \bigcup_{i=1}^{m} \mathcal{N}_i$.

*Then the following statements hold.*

(a) *Suppose that $\{\alpha_k\}$ is an arbitrary real sequence such that $\lim_{k\to\infty} \alpha_k = +\infty$. Then,*

$$|\mathcal{N}_i| = \infty \implies \lim_{k\to\infty,\, k\in\mathcal{N}_i} \alpha_k = +\infty.$$

(b) *Suppose $|\mathcal{N}_i| = \infty$ and $\{\beta_k\}$ is an arbitrary real sequence such that $\lim_{k\to\infty,\, k\in\mathcal{N}_i} \beta_k = +\infty$. Then,*

$$\lim_{k\to\infty} \beta_k = +\infty.$$

We now give the following main theorem of this section by using the five lemmas shown above.

**Theorem 3.1** *Let $\tilde{\Psi}_{\rm NR} : \Re^n \to \Re$ be defined by (3.3). If $f$ is strongly monotone, then $\tilde{\Psi}_{\rm NR}$ is coercive.*

**Proof.** Let $\{x^{(k)}\}$ be an arbitrary sequence such that $\lim_{k\to\infty} \|x^{(k)}\| = \infty$. Then our goal is to prove

$$\lim_{k\to\infty} \tilde{\Psi}_{\rm NR}(x^{(k)}) = \infty. \tag{3.7}$$

We note from Lemma 3.4 that

$$\lim_{k\to\infty} \|f(x^{(k)})\| \;=\; \infty. \tag{3.8}$$

We let $x^{(k)} = (x_1^{(k)}, x_2^{(k)}) \in \Re \times \Re^{n-1}$ and $f(x^{(k)}) = (f_1(x^{(k)}), f_2(x^{(k)})) \in \Re \times \Re^{n-1}$.

Now we define the index sets $\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3, \mathcal{N}_4$ and $\mathcal{N}_5$ by

$$\begin{aligned}
\mathcal{N}_1 &= \left\{ k \,\middle|\, x^{(k)} - f(x^{(k)}) \in \mathcal{K}^n \right\}, \\
\mathcal{N}_2 &= \left\{ k \,\middle|\, x^{(k)} - f(x^{(k)}) \in -\mathcal{K}^n \backslash \{0\} \right\}, \\
\mathcal{N}_3 &= \left\{ k \,\middle|\, x^{(k)} - f(x^{(k)}) \notin \mathcal{K}^n \cup -\mathcal{K}^n,\ x_1^{(k)} \le 0 \right\}, \\
\mathcal{N}_4 &= \left\{ k \,\middle|\, x^{(k)} - f(x^{(k)}) \notin \mathcal{K}^n \cup -\mathcal{K}^n,\ x_1^{(k)} > 0,\ f_1(x^{(k)}) < 0 \right\}, \\
\mathcal{N}_5 &= \left\{ k \,\middle|\, x^{(k)} - f(x^{(k)}) \notin \mathcal{K}^n \cup -\mathcal{K}^n,\ x_1^{(k)} > 0,\ f_1(x^{(k)}) \ge 0 \right\}.
\end{aligned}$$

Note that $\cup_{i=1}^5 \mathcal{N}_i = \{0, 1, \ldots\}$ and that $\mathcal{N}_i \cap \mathcal{N}_j = \emptyset$ for any $i \neq j \in \{1, 2, 3, 4, 5\}$. By setting $\alpha_k = \|x^{(k)}\|$ or $\alpha_k = \|f(x^{(k)})\|$ in Lemma 3.5 (a), we obtain the following Property A.

Property A. If $|\mathcal{N}_i| = \infty$, then

$$\lim_{k\to\infty,\, k\in\mathcal{N}_i} \|x^{(k)}\| = \infty, \quad \lim_{k\to\infty,\, k\in\mathcal{N}_i} \|f(x^{(k)})\| = \infty.$$

By setting $\beta_k = \tilde{\Psi}_{\text{NR}}(x^{(k)})$ in Lemma 3.5 (b), we obtain the following Property B.

Property B. If

$$\lim_{k\to\infty,\, k\in\mathcal{N}_i} \tilde{\Psi}_{\text{NR}}(x^{(k)}) = \infty$$

for $\mathcal{N}_i$ such that $|\mathcal{N}_i| = \infty$, then

$$\lim_{k\to\infty} \tilde{\Psi}_{\text{NR}}(x^{(k)}) = \infty.$$

Property B implies that, for proving (3.7), it is sufficient to show

$$|\mathcal{N}_i| = \infty \implies \lim_{k\to\infty,\, k\in\mathcal{N}_i} \tilde{\Psi}_{\text{NR}}(x^{(k)}) = \infty \tag{3.9}$$

for each $i \in \{1, 2, 3, 4, 5\}$. Let $y^{(k)} := f(x^{(k)})$ for all $k$. We note that $\lim_{k\to\infty} \|y^{(k)}\| = \infty$ holds from (3.8).

**Case 1.** $|\mathcal{N}_1| = \infty$:

Lemma 3.1 (a) implies that

$$\varphi_{\text{NR}}(x^{(k)}, y^{(k)}) \;=\; y^{(k)}$$

for all $k \in \mathcal{N}_1$. It then follows from Property A that

$$\lim_{k\to\infty,\, k\in\mathcal{N}_1} \tilde{\Psi}_{\text{NR}}(x^{(k)}) \;=\; \lim_{k\to\infty,\, k\in\mathcal{N}_1} \left\|\varphi_{\text{NR}}(x^{(k)}, y^{(k)})\right\| \;=\; \lim_{k\to\infty,\, k\in\mathcal{N}_1} \|y^{(k)}\| \;=\; \infty.$$

**Case 2.** $|\mathcal{N}_2| = \infty$:

Lemma 3.1 (b) implies that

$$\varphi_{\rm NR}(x^{(k)}, y^{(k)}) = x^{(k)}$$

for all $k \in \mathcal{N}_2$. It then follows from Property A that

$$\lim_{k\to\infty,\, k\in\mathcal{N}_2} \tilde{\Psi}_{\rm NR}(x^{(k)}) = \lim_{k\to\infty,\, k\in\mathcal{N}_2} \left\|\varphi_{\rm NR}(x^{(k)}, y^{(k)})\right\| = \lim_{k\to\infty,\, k\in\mathcal{N}_2} \|x^{(k)}\| = \infty.$$

**Case 3.** $|\mathcal{N}_3| = \infty$:

Lemma 3.2 implies that

$$\left\|\varphi_{\rm NR}(x^{(k)}, y^{(k)})\right\| \geq \frac{1}{\sqrt{2}}\|x^{(k)}\|$$

for all $k \in \mathcal{N}_3$. It then follows from Property A that

$$\lim_{k\to\infty,\, k\in\mathcal{N}_3} \tilde{\Psi}_{\rm NR}(x^{(k)}) = \lim_{k\to\infty,\, k\in\mathcal{N}_3} \left\|\varphi_{\rm NR}(x^{(k)}, y^{(k)})\right\| \geq \lim_{k\to\infty,\, k\in\mathcal{N}_3} \frac{1}{\sqrt{2}}\|x^{(k)}\| = \infty.$$

**Case 4.** $|\mathcal{N}_4| = \infty$:

Assuming $\liminf_{k\to\infty,\, k\in\mathcal{N}_4} \|\varphi_{\rm NR}(x^{(k)}, y^{(k)})\| < +\infty$, we will derive contradiction. Under this assumption, there exists an infinite subsequence $\overline{\mathcal{N}}_4 \subseteq \mathcal{N}_4$ such that $\{\varphi_{\rm NR}(x^{(k)}, y^{(k)})\}_{k\in\overline{\mathcal{N}}_4}$ is bounded. Note that $\lim_{k\to\infty,\, k\in\overline{\mathcal{N}}_4} \|x^{(k)}\| = \infty$ and $\lim_{k\to\infty,\, k\in\overline{\mathcal{N}}_4} \|y^{(k)}\| = \infty$ from Property A and Lemma 3.5 (a). Thus, since $\{x^{(k)}\}_{k\in\overline{\mathcal{N}}_4}$ and $\{y^{(k)}\}_{k\in\overline{\mathcal{N}}_4}$ satisfy assumptions (i)–(iii) of Lemma 3.3, there exist $K \subseteq \overline{\mathcal{N}}_4$ and $M \in (2, \infty)$ satisfying (a)–(d) in Lemma 3.3. Let us choose $k \in K$ arbitrarily.

From Lemma 3.3 (c) and $(x_1^{(k)} y_1^{(k)})/(|x_1^{(k)}||y_1^{(k)}|) = -1$, we have

$$\begin{aligned} {x_2^{(k)}}^T y_2^{(k)} &> \|x_2^{(k)}\|\|y_2^{(k)}\| \left(1 - \frac{1}{45M}\right) \\ &> 100M^2\left(1 - \frac{1}{45M}\right) \\ &> 90M^2, \end{aligned} \tag{3.10}$$

where the second inequality follows from Lemma 3.3 (b-2), and the third inequality follows from the fact that $M > 2 > 2/9$. Since $x^{(k)} - y^{(k)} \notin \mathcal{K}^n \cup -\mathcal{K}^n$, we have $-\|x_2^{(k)} - y_2^{(k)}\| < x_1^{(k)} - y_1^{(k)} < \|x_2^{(k)} - y_2^{(k)}\|$ from the definition (1.2) of second-order cone, that is,

$$\left(x_1^{(k)} - y_1^{(k)}\right)^2 < \left\|x_2^{(k)} - y_2^{(k)}\right\|^2.$$

We then have

$$\begin{aligned} {x_2^{(k)}}^T y_2^{(k)} &< x_1^{(k)} y_1^{(k)} + \frac{1}{2}\left\{\|x_2^{(k)}\|^2 - (x_1^{(k)})^2\right\} + \frac{1}{2}\left\{\|y_2^{(k)}\|^2 - (y_1^{(k)})^2\right\} \\ &< x_1^{(k)} y_1^{(k)} + \frac{1}{2}\left\{\left(|x_1^{(k)}| + M\right)^2 - (x_1^{(k)})^2\right\} + \frac{1}{2}\left\{\left(|y_1^{(k)}| + M\right)^2 - (y_1^{(k)})^2\right\} \\ &= -|x_1^{(k)}||y_1^{(k)}| + M|x_1^{(k)}| + M|y_1^{(k)}| + M^2 \\ &= -\left(|x_1^{(k)}| - M\right)\left(|y_1^{(k)}| - M\right) + 2M^2 \\ &< -79M^2, \end{aligned} \tag{3.11}$$

where the second inequality holds from Lemma 3.3 (a), the first equality holds from $x^{(k)} > 0$ and $y^{(k)} < 0$, and the third inequality holds from Lemma 3.3 (b-2). Since (3.10) and (3.11) contradict each other, we obtain

$$\lim_{k\to\infty,\, k\in\mathcal{N}_4} \tilde{\Psi}_{\rm NR}(x^{(k)}) = \lim_{k\to\infty,\, k\in\mathcal{N}_4} \left\|\varphi_{\rm NR}(x^{(k)}, y^{(k)})\right\|^2 = \infty.$$

**Case 5.** $|\mathcal{N}_5| = \infty$ :

Assuming $\liminf_{k\to\infty,\, k\in\mathcal{N}_5} \|\varphi_{\rm NR}(x^{(k)}, y^{(k)})\| < +\infty$, we will derive contradiction. Under this assumption, there exists an infinite subsequence $\overline{\mathcal{N}}_5 \subseteq \mathcal{N}_5$ such that $\{\varphi_{\rm NR}(x^{(k)}, y^{(k)})\}_{k\in\overline{\mathcal{N}}_5}$ is bounded. Note that $\lim_{k\to\infty,\, k\in\overline{\mathcal{N}}_5} \|x^{(k)}\| = \infty$ and $\lim_{k\to\infty,\, k\in\overline{\mathcal{N}}_5} \|y^{(k)}\| = \infty$ from Property A and Lemma 3.5 (a). Thus, since $\{x^{(k)}\}_{k\in\overline{\mathcal{N}}_5}$ and $\{y^{(k)}\}_{k\in\overline{\mathcal{N}}_5}$ satisfy assumptions (i)–(iii) of Lemma 3.3, there exist $K \subseteq \overline{\mathcal{N}}_5$ and $M \in (2, \infty)$ satisfying (a)–(d) in Lemma 3.3. Let us choose $k \in K$ arbitrarily.

From Lemma 3.3 and the fact that $x^{(k)}, y^{(k)} > 0$ for any $k \in K$, we have

$$-M < x_1^{(k)} - \|x_2^{(k)}\| < M, \;\; -M < y_1^{(k)} - \|y_2^{(k)}\| < M, \tag{3.12}$$

$$\{x_1^{(k)}\},\ \{\|x_2^{(k)}\|\},\ \{y_1^{(k)}\} \text{ and } \{\|y_2^{(k)}\|\} \text{ are unbounded,} \tag{3.13}$$

$$x_1^{(k)} > 10M,\ \|x_2^{(k)}\| > 10M,\ y_1^{(k)} > 10M \text{ and } \|y_2^{(k)}\| > 10M, \tag{3.14}$$

$$1 + \frac{{x_2^{(k)}}^T y_2^{(k)}}{\|x_2^{(k)}\|\|y_2^{(k)}\|} < \frac{10M}{9}\left\{\frac{1}{(x_1^{(k)})^2} + \frac{1}{(y_1^{(k)})^2}\right\}, \tag{3.15}$$

$$\sqrt{2} - 0.1 < \frac{\|x^{(k)}\|}{|x_1^{(k)}|} < \sqrt{2} + 0.1 \text{ and } \sqrt{2} - 0.1 < \frac{\|y^{(k)}\|}{|y_1^{(k)}|} < \sqrt{2} + 0.1. \tag{3.16}$$

Now let the spectral values of $x^{(k)}$ be $\kappa_1^{(k)}$ and $\kappa_2^{(k)}$, and the spectral vectors be $v_1^{(k)}$ and $v_2^{(k)}$. Furthermore, let the spectral values of $y^{(k)}$ be $\nu_1^{(k)}$ and $\nu_2^{(k)}$, and the spectral vectors be $w_1^{(k)}$ and $w_2^{(k)}$. Note that, from (3.12) and the definition (2.1) of spectral values, we have

$$-M < \kappa_1^{(k)} < M, \;\; -M < \nu_1^{(k)} < M. \tag{3.17}$$

Since $f$ is strongly monotone, there exists $\varepsilon > 0$ such that

$$(x - z)^T (f(x) - f(z)) \;\geq\; \varepsilon \|x - z\|^2 \tag{3.18}$$

for any $x,\, z \in \Re^n$. Let the sequences $\{\xi^{(k)}\}$ and $\{\eta^{(k)}\}$ be defined by

$$\xi^{(k)} := x^{(k)} - \kappa_2^{(k)} w_1^{(k)}, \tag{3.19}$$

$$\eta^{(k)} := f(\xi^{(k)}). \tag{3.20}$$

Substituting $x = x^{(k)}$ and $z = \xi^{(k)}$ in (3.18), we obtain

$$\begin{aligned}
\varepsilon\|\kappa_2^{(k)} w_1^{(k)}\|^2 &\leq \left(x^{(k)} - \xi^{(k)}\right)^T \left(y^{(k)} - \eta^{(k)}\right) \\
&= \kappa_2^{(k)} {w_1^{(k)}}^T \left(\nu_1^{(k)} w_1^{(k)} + \nu_2^{(k)} w_2^{(k)} - \eta^{(k)}\right) \\
&= \frac{1}{2}\kappa_2^{(k)} \nu_1^{(k)} - \kappa_2^{(k)} {w_1^{(k)}}^T \eta^{(k)} \\
&< \frac{1}{2} M \kappa_2^{(k)} + \frac{1}{\sqrt{2}} \|\eta^{(k)}\| \kappa_2^{(k)},
\end{aligned} \tag{3.21}$$

where the second equality follows from $\|w_1^{(k)}\|^2 = 1/2$ and $w_1^{(k)T} w_2^{(k)} = 0$, and the last inequality follows from $\kappa_2^{(k)} > 0$, $-M < \nu_1^{(k)} < M$, the Cauchy-Schwarz inequality and $\|w_1^{(k)}\| = 1/\sqrt{2}$. Since the left-hand side of (3.21) equals $(1/2)\varepsilon(\kappa_2^{(k)})^2$ and $\kappa_2^{(k)}$ is positive, dividing both sides of (3.21) yields

$$\|\eta^{(k)}\| > \frac{\varepsilon}{\sqrt{2}}\kappa_2^{(k)} - \frac{M}{\sqrt{2}}.$$

Since $\{\kappa_2^{(k)}\} = \{x_1^{(k)} + \|x_2^{(k)}\|\}$ is unbounded, $\{\eta^{(k)}\}$ is also unbounded.

Next, we derive a contradiction by showing the boundedness of $\{\eta^{(k)}\}$. To do this, since $\eta^{(k)} = f(\xi^{(k)})$, it is sufficient to show the boundedness of $\{\xi^{(k)}\}$. Substituting $x^{(k)} = \kappa_1^{(k)} v_1^{(k)} + \kappa_2^{(k)} v_2^{(k)}$ in (3.19), we have

$$\xi^{(k)} = \kappa_2^{(k)} \left(v_2^{(k)} - w_1^{(k)}\right) + \kappa_1^{(k)} v_1^{(k)}.$$

Since $\|v_1^{(k)}\| = 1/\sqrt{2}$ and (3.17) holds, $\{\kappa_1^{(k)} v_1^{(k)}\}$ is bounded. Then, we can show the boundedness of $\left\{\kappa_2^{(k)}\left(v_2^{(k)} - w_1^{(k)}\right)\right\}$ as follows:

$$\begin{aligned}
\left\|\kappa_2^{(k)}\left(v_2^{(k)} - w_1^{(k)}\right)\right\|^2 &= \left(x_1^{(k)} + \|x_2^{(k)}\|\right)^2 \left\|\left\{\frac{1}{2}\left(1, \frac{x_2^{(k)}}{\|x_2^{(k)}\|}\right) - \frac{1}{2}\left(1, -\frac{y_2^{(k)}}{\|y_2^{(k)}\|}\right)\right\}\right\|^2 \\
&= \left(x_1^{(k)} + \|x_2^{(k)}\|\right)^2 \left\|\frac{1}{2}\left(0, \frac{x_2^{(k)}}{\|x_2^{(k)}\|} + \frac{y_2^{(k)}}{\|y_2^{(k)}\|}\right)\right\|^2 \\
&= \frac{1}{2}\left(x_1^{(k)} + \|x_2^{(k)}\|\right)^2 \left(1 + \frac{x_2^{(k)T} y_2^{(k)}}{\|x_2^{(k)}\|\|y_2^{(k)}\|}\right) \\
&< \frac{5}{9} M \left(2x_1^{(k)} + M\right)^2 \left\{\frac{1}{(x_1^{(k)})^2} + \frac{1}{(y_1^{(k)})^2}\right\} \\
&< \frac{5 \cdot 2.1^2}{9} M \left\{1 + \left(\frac{x_1^{(k)}}{y_1^{(k)}}\right)^2\right\} \\
&= \frac{5 \cdot 2.1^2}{9} M \left\{1 + \left(\frac{\|x^{(k)}\|}{\|y^{(k)}\|} \frac{\|y^{(k)}\|/y_1^{(k)}}{\|x^{(k)}\|/x_1^{(k)}}\right)^2\right\} \\
&< 3M \left\{1 + \left(\frac{2(\sqrt{2} + 0.1)}{\varepsilon(\sqrt{2} - 0.1)}\right)^2\right\},
\end{aligned}$$

where the first inequality follows from $x^{(k)} > 0$, $y^{(k)} > 0$, (3.12) and (3.15), the second inequality follows from the fact that (3.14) implies $M < 0.1x_1^{(k)}$, and the third inequality follows from Lemma 3.4 and (3.16). Since $\{\xi^{(k)}\}$ is bounded, $\{\eta^{(k)}\} = \{f(\xi^{(k)})\}$ is also bounded. However, this contradicts the unboundedness of $\{\eta^{(k)}\}$. Hence, we obtain

$$\lim_{k\to\infty,\, k\in\mathcal{N}_5} \tilde{\Psi}_{\mathrm{NR}}(x^{(k)}) = \lim_{k\to\infty,\, k\in\mathcal{N}_5} \left\|\varphi_{\mathrm{NR}}(x^{(k)}, y^{(k)})\right\| = \infty.$$

Consequently, we have shown that (3.9) holds for all $i \in \{1, 2, 3, 4, 5\}$. This completes the proof. ■

# 4 Algorithm for Solving SOCCP

In Section 3, we have shown that, if $f$ is strongly monotone, then the merit function $\Psi_{\rm NR}$ given by (3.2) is coercive. Therefore, we can solve the strongly monotone SOCCP (3.1) by minimizing $\Psi_{\rm NR}$ with an appropriate descent algorithm. However, the merit function $\Psi_{\rm NR}$ is not differentiable in general. Therefore, methods that use the gradient of the objective function, such as the steepest descent method and Newton's method, are not applicable. Furthermore, the assumption for $f$ to be strongly monotone is quite restrictive from a practical standpoint. In order to overcome such shortcomings, we consider a smoothing method and a regularization method. In this section, we propose a globally convergent algorithm for SOCCP, based on these methods.

## 4.1 Smoothing and Regularization Methods

In this section, we introduce the smoothing and regularization methods. Firstly, we describe the smoothing method. For a nondifferentiable function $h : \Re^n \to \Re^m$, we consider a function $h_\mu : \Re^n \to \Re^m$ with a parameter $\mu > 0$ which has the following properties:

(a) $h_\mu$ is differentiable for any $\mu > 0$,

(b) $\lim_{\mu \to +0} h_\mu(x) = h(x)$ for any $x \in \Re^n$.

Such a function $h_\mu$ is called *a smoothing function* of $h$. Instead of solving the original problem $h(x) = 0$, the smoothing method solves subproblems $h_\mu(x) = 0$ for $\mu > 0$, and obtain a solution of the original problem by letting $\mu \to +0$. Fukushima, Luo and Tseng [8] extended Chen and Mangasarian's class [2] of smoothing functions for NCP to SOCCP, which may be regarded as a smoothing function of natural residual $\varphi_{\rm NR}$.

Now, in order to define $\varphi_\mu$, we consider the following continuously differentiable convex function $\hat{g} : \Re \to \Re$ satisfying

$$\lim_{\alpha \to -\infty} \hat{g}(\alpha) = 0, \quad \lim_{\alpha \to \infty} (\hat{g}(\alpha) - \alpha) = 0, \quad 0 < \hat{g}'(\alpha) < 1. \tag{4.1}$$

For example, $\hat{g}_1(\alpha) = (\sqrt{\alpha^2 + 4} + \alpha)/2$ and $\hat{g}_2(\alpha) = \ln(e^\alpha + 1)$ satisfy the conditions. Furthermore, by using $\hat{g}$, we define $g : \Re^n \to \Re^n$ by

$$g(z) \;=\; \hat{g}(\lambda_1)u_1 + \hat{g}(\lambda_2)u_2, \tag{4.2}$$

where $\lambda_1$ and $\lambda_2$ are the spectral values of $z$, and $u_1$ and $u_2$ are the spectral vectors of $z$. Then, for $\mu > 0$, the function $\varphi_\mu$ is given by

$$\varphi_\mu(x, y) = x - \mu g((x - y)/\mu). \tag{4.3}$$

Fukushima, Luo and Tseng [8] showed that $\varphi_\mu$ with $\mu > 0$ is a smoothing function of $\varphi_{\rm NR}$. For $\mu = 0$, we denote $\varphi_0(x, y) := \varphi_{\rm NR}(x, y)$ for convenience.

By using $\varphi_\mu$, we define $\Psi_\mu : \Re^n \times \Re^n \to \Re$ by

$$\Psi_\mu(x,\, y) \;=\; \frac{1}{2}\,\|\varphi_\mu(x,\, y)\|^2 + \frac{1}{2}\|f(x) - y\|^2. \tag{4.4}$$

Since $\varphi_\mu$ is differentiable for $\mu > 0$, $\Psi_\mu$ is also differentiable for $\mu > 0$. Next, by applying Theorem 3.1, we give a sufficient condition for the coerciveness of $\Psi_\mu$.

**Theorem 4.1** *If $f$ is strongly monotone, then, for any $\mu \geq 0$, the function $\Psi_\mu(x, y)$ defined by (4.4) is coercive.*

**Proof.** From [8, Proposition 5.1], for any $\mu \geq 0$, there exists a constant $\rho > 0$ such that

$$\|\varphi_\mu(x, y) - \varphi_{\text{NR}}(x, y)\| \leq \rho\mu, \quad \forall x, y \in \Re^n.$$

It then follows from Proposition 3.1 and Theorem 3.1 that $\Psi_\mu$ is coercive if $f$ is strongly monotone. ■

When $f$ is strongly monotone, this theorem enables us to obtain a solution of SOCCP (3.1) by using the smoothing method combined with a suitable descent method. However, strong monotonicity of $f$ is quite a severe condition.

As a remedy for this inconvenience, we employ the regularization method. In the regularization method, we consider the function $f_\varepsilon : \Re^n \to \Re^n$ defined by $f_\varepsilon(x) := f(x) + \varepsilon x$. This method solves SOCCP with $f_\varepsilon$ for each $\varepsilon > 0$ as a subproblem, and obtains a solution of SOCCP by letting $\varepsilon \to +0$. If $f$ is monotone, then $f_\varepsilon$ is strongly monotone for any $\varepsilon > 0$. Therefore, the coerciveness of the function

$$\Psi_{\mu,\varepsilon}(x, y) = \frac{1}{2}\|\varphi_\mu(x, y)\|^2 + \frac{1}{2}\|f_\varepsilon(x) - y\|^2 \tag{4.5}$$

is guaranteed from Theorem 4.1 and Proposition 3.1.

## 4.2 Stationary Point of Smoothing Function

In this section, we show that any stationary point of $\Psi_{\mu,\varepsilon}$ is a global minimum of $\Psi_{\mu,\varepsilon}$. To this end, we first give an explicit representation of $\nabla\Psi_{\mu,\varepsilon}$ under the assumption that $f$ is differentiable.

Now, let us define $H_{\mu,\varepsilon} : \Re^n \times \Re^n \to \Re^n \times \Re^n$ by

$$H_{\mu,\varepsilon}(x, y) := \begin{pmatrix} \varphi_\mu(x, y) \\ f_\varepsilon(x) - y \end{pmatrix}. \tag{4.6}$$

Then, noticing

$$\Psi_{\mu,\varepsilon}(x, y) = \frac{1}{2}\|H_{\mu,\varepsilon}(x, y)\|^2,$$

we have

$$\nabla\Psi_{\mu,\varepsilon}(x, y) = \nabla H_{\mu,\varepsilon}(x, y)\ H_{\mu,\varepsilon}(x, y) \tag{4.7}$$

for $\mu > 0$.

From the definition of $H_{\mu,\varepsilon}$, $\varphi_\mu$ and $f_\varepsilon$, $\nabla H_{\mu,\varepsilon}(x, y)$ can be written as

$$\begin{aligned} \nabla H_{\mu,\varepsilon}(x, y) &= \begin{pmatrix} \nabla_x \varphi_\mu(x, y) & \nabla f(x) + \varepsilon I \\ \nabla_y \varphi_\mu(x, y) & -I \end{pmatrix} \\ &= \begin{pmatrix} I - \nabla g(z) & \nabla f(x) + \varepsilon I \\ \nabla g(z) & -I \end{pmatrix}, \end{aligned} \tag{4.8}$$

where $z = (x-y)/\mu$ and $g$ is the function defined by (4.2). Moreover, from [8, Proposition 5.2], $\nabla g(z)$ is written as follows:

$$\nabla g(z) \; = \; \begin{cases} \hat{g}'(z_1) I & \text{if } z_2 = 0, \\ \begin{pmatrix} b & \dfrac{c\, z_2^T}{\|z_2\|} \\ \dfrac{c\, z_2}{\|z_2\|} & aI + (b-a)\dfrac{z_2 z_2^T}{\|z_2\|^2} \end{pmatrix} & \text{if } z_2 \neq 0, \end{cases} \tag{4.9}$$

where

$$a = \frac{\hat{g}(\lambda_2) - \hat{g}(\lambda_1)}{\lambda_2 - \lambda_1}, \;\; b = \frac{1}{2}(\hat{g}'(\lambda_2) + \hat{g}'(\lambda_1)), \;\; c = \frac{1}{2}(\hat{g}'(\lambda_2) - \hat{g}'(\lambda_1)) \tag{4.10}$$

and $\lambda_i$, $i = 1, 2$ are the spectral values of $z = (x-y)/\mu$.

It is important to see when $\nabla H_{\mu,\varepsilon}(x,y)$ is nonsingular since (4.7) implies that, if $\nabla H_{\mu,\varepsilon}(x,y)$ is nonsingular everywhere, then every stationary point of $\Psi_{\mu,\varepsilon}$ is a global minimum of $\Psi_{\mu,\varepsilon}$. The following proposition gives a sufficient condition for $\nabla H_{\mu,\varepsilon}(x,y)$ to be nonsingular.

**Proposition 4.1** *If $f : \Re^n \to \Re^n$ is monotone, then the Jacobian $\nabla H_{\mu,\varepsilon}(x,y)$ is nonsingular for any $\mu > 0$, $\varepsilon \geq 0$ and $(x,y) \in \Re^n \times \Re^n$.*

**Proof.** Firstly, we show that $O \prec \nabla g(z) \prec I$ for any $z \in \Re^n$. When $z_2 = 0$, it is clear that $O \prec \nabla g(z) \prec I$ since $\nabla g(z) = \hat{g}'(z_1) I$ from (4.9) and $0 < \hat{g}'(z_1) < 1$ from (4.1). So we only consider the case where $z_2 \neq 0$. Noticing that $\nabla g(z)$ is given by (4.9), in order to show $\nabla g(z) \succ O$, it is sufficient to show that $b$ is positive and the Schur complement of $\nabla g(z)$ with respect to $b$ is positive definite. Since $b = (1/2)(\hat{g}'(\lambda_1)) + (1/2)(\hat{g}'(\lambda_2))$, we have $b > 0$ from (4.1). On the other hand, the Schur complement of $\nabla g(z)$ with respect to $b$ is given by

$$\left\{ aI + (b-a)\frac{z_2 z_2^T}{\|z_2\|^2} \right\} - \frac{c^2}{b}\frac{z_2 z_2^T}{\|z_2\|^2} \; = \; a\left( I - \frac{z_2 z_2^T}{\|z_2\|^2} \right) + \frac{b^2 - c^2}{b}\frac{z_2 z_2^T}{\|z_2\|^2}.$$

If $a = (\hat{g}(\lambda_2) - \hat{g}(\lambda_1))/(\lambda_2 - \lambda_1) \leq 0$, then the continuous differentiability of $\hat{g}$ and the mean value theorem guarantee the existence of $\tau \in [\lambda_1, \lambda_2]$ such that $\hat{g}'(\tau) \leq 0$. Since this fact contradicts (4.1), we have $a > 0$. Moreover, we have $(b^2 - c^2)/b^2 = 2/\hat{g}(\lambda_1) + 2/\hat{g}(\lambda_2) > 0$ from (4.1). Furthermore, both $z_2 z_2^T/\|z_2\|^2$ and $I - z_2 z_2^T/\|z_2\|^2$ are positive semidefinite and their sum is the identity matrix. Hence, any positive combination of $z_2 z_2^T/\|z_2\|^2$ and $I - z_2 z_2^T/\|z_2\|^2$ is positive definite. Therefore, the Schur complement of $\nabla g(z)$ with respect to $b$ is positive definite, and hence we obtain $\nabla g(z) \succ O$. In a similar way, we can show

$$I - \nabla g(z) \; = \; \begin{pmatrix} 1 - b & -\dfrac{c\, z_2^T}{\|z_2\|} \\ -\dfrac{c\, z_2}{\|z_2\|} & (1-a)I - (b-a)\dfrac{z_2 z_2^T}{\|z_2\|^2} \end{pmatrix} \succ O,$$

by showing $1 - b > 0$ and the positive definiteness of the Schur complement of $I - \nabla g(z)$ with respect to $1 - b$.

Secondly, we show the nonsingularity of $\nabla H_{\mu,\varepsilon}(x,y)^T$ instead of $\nabla H_{\mu,\varepsilon}(x,y)$. Let us denote $\mathcal{G} := \nabla g((x-y)/\mu)$ and $\mathcal{F} := \nabla f(x)$ for convenience. Then $\nabla H_{\mu,\varepsilon}(x,y)^T$ can be written as

$$\nabla H_{\mu,\varepsilon}(x,y)^T \; = \; \begin{pmatrix} I - \mathcal{G} & \mathcal{G} \\ \mathcal{F} + \varepsilon I & -I \end{pmatrix}.$$

Let $\xi, \eta \in \Re^n$ satisfy

$$\nabla H_{\mu,\varepsilon}(x,y)^T \begin{pmatrix} \xi \\ \eta \end{pmatrix} = 0,$$

that is,

$$(I - \mathcal{G})\xi + \mathcal{G}\eta = 0, \tag{4.11}$$

$$(\mathcal{F} + \varepsilon I)\xi - \eta = 0. \tag{4.12}$$

Multiplying the left-hand side of (4.11) by $\mathcal{G}^{-1}$ and combining with (4.12), we have

$$(\mathcal{G}^{-1} - I + \mathcal{F} + \varepsilon I)\xi = 0.$$

Since $O \prec \mathcal{G} \prec I$ implies $\mathcal{G}^{-1} \succ I$ and monotonicity of $f$ implies $\mathcal{F} \succeq O$, $\mathcal{G}^{-1} - I + \mathcal{F} + \varepsilon I$ is positive definite. So we have $\xi = 0$, and then $\eta = 0$ from (4.12). Hence, $\nabla H_{\mu,\varepsilon}(x,y)^T$ is nonsingular, that is, $\nabla H_{\mu,\varepsilon}(x,y)$ is nonsingular. ■

Finally, by using the result of Proposition 4.1, we give the main result of this section.

**Proposition 4.2** *If $f : \Re^n \to \Re^n$ is monotone, then, for any $\mu > 0$ and $\varepsilon \geq 0$, every stationary point $(\overline{x}, \overline{y})$ of the function $\Psi_{\mu,\varepsilon}$ defined by (4.5) satisfies $\Psi_{\mu,\varepsilon}(\overline{x}, \overline{y}) = 0$.*

**Proof.** Note that $\nabla \Psi_{\mu,\varepsilon}(\overline{x}, \overline{y}) = \nabla H_{\mu,\varepsilon}(\overline{x}, \overline{y})\, H_{\mu,\varepsilon}(\overline{x}, \overline{y}) = 0$. By Proposition 4.1, $\nabla H_{\mu,\varepsilon}(\overline{x}, \overline{y})$ is nonsingular. Hence, we have $H_{\mu,\varepsilon}(\overline{x}, \overline{y}) = 0$, that is, $\Psi_{\mu,\varepsilon}(\overline{x}, \overline{y}) = (1/2)\|H_{\mu,\varepsilon}(\overline{x}, \overline{y})\|^2 = 0$. ■

## 4.3 Globally Convergent Algorithm

In Section 4.1, we showed that, for $\mu > 0$ and $\varepsilon > 0$, the function $\Psi_{\mu,\varepsilon}$ defined by (4.5) is coercive if $f$ is monotone. Hence, by applying an appropriate descent method, we can obtain a minimum $(x(\mu,\varepsilon), y(\mu,\varepsilon))$ of the function $\Psi_{\mu,\varepsilon}$. Moreover, letting $(\mu,\varepsilon)$ converge to $(0,0)$, we may expect that $(x(\mu,\varepsilon), y(\mu,\varepsilon))$ converges to a solution of SOCCP. However, in practice, it is usually impossible to find an exact minimum of $\Psi_{\mu,\varepsilon}$. So, we consider the following algorithm in which the function $\Psi_{\mu,\varepsilon}$ is minimized only approximately at each iteration.

**Algorithm 4.1** *Let $\{\varepsilon_k\}$, $\{\mu_k\}$ and $\{\alpha_k\}$ be sequences of positive numbers converging to 0.*
(Step 0) *Choose $(x^{(0)}, y^{(0)}) \in \Re^n \times \Re^n$ and set $k := 0$.*
(Step 1) *Terminate the iteration if an appropriate stopping criterion is satisfied.*
(Step 2) *Find a pair $(x^{(k+1)}, y^{(k+1)})$ such that*

$$\Psi_{\mu_{k+1},\,\varepsilon_{k+1}}(x^{(k+1)}, y^{(k+1)}) \leq \alpha_{k+1}.$$

*Set $k := k+1$, and go back to Step 1.*

To obtain $(x^{(k+1)}, y^{(k+1)})$ in Step 2, we can use any unconstrained minimization technique such as the steepest descent method and Newton's method. We note that this algorithm is well-defined by the following reasons. Since $\Psi_{\mu,\varepsilon}$ is differentiable and coercive from Theorem 4.1, it has a stationary point. Moreover, Proposition 4.2 implies that the value of $\Psi_{\mu,\varepsilon}$ at the stationary point is 0. Hence, there exists $(x^{(k+1)}, y^{(k+1)})$ satisfying Step 2 for any $\alpha_{k+1} > 0$.

We next show the global convergent property of Algorithm 4.1, by extending the result of [6] for NCP to SOCCP. To this end, we give two lemmas. The following lemma implies that $\Psi_{\mu,\varepsilon}$ is uniformly continuous on a compact set not only in $x$ and $y$ but also in $\mu$ and $\varepsilon$

**Lemma 4.1** *Let $C \subset \Re^n \times \Re^n$ be a compact set. Then, for any $\delta > 0$, there exists $\overline{\varepsilon} > 0$ and $\overline{\mu} > 0$ such that*

$$|\Psi_{\mu,\varepsilon}(x, y) - \Psi_{\rm NR}(x, y)| \leq \delta$$

*for any $(x, y) \in C$, $\varepsilon \in [0, \overline{\varepsilon}]$ and $\mu \in [0, \overline{\mu}]$.*

**Proof.** Define the function $\Omega : \Re^n \times \Re^n \times \Re \times \Re \to \Re$ by $\Omega(x, y, \mu, \varepsilon) := \Psi_{\mu,\varepsilon}(x, y)$. Then, $\Omega$ is continuous and satisfies $\Omega(x, y, 0, 0) = \Psi_{\rm NR}(x, y)$. Since any continuous function is uniformly continuous on a compact set, $\Omega$ is uniformly continuous on $C \times [0, \overline{\varepsilon}] \times [0, \overline{\mu}]$. ■

The next lemma is known as the mountain pass theorem, which is useful for our analysis. For more detail, see Theorem 9.2.7 in [11].

**Lemma 4.2 (Mountain Pass Theorem)** *Let $\theta : \Re^n \to \Re$ be a continuously differentiable and coercive function. Let $C \subset \Re^n$ be a nonempty and compact set and let $m$ be the minimum value of $\theta$ on the boundary of $C$, that is,*

$$m := \min_{x \in \partial C} \theta(x).$$

*Assume moreover that there exist points $a \in C$ and $b \notin C$ such that $\theta(a) < m$ and $\theta(b) < m$. Then, there exists a point $c \in \Re^n$ such that $\nabla\theta(c) = 0$ and $\theta(c) \geq m$.*

Finally, by using the above lemmas. we establish the global convergence property of Algorithm 4.1.

**Theorem 4.2** *Let $f : \Re^n \to \Re^n$ be a monotone function, and assume that the solution set $\mathcal{S}$ of SOCCP* (3.1) *is nonempty and bounded. Let $\{(x^{(k)}, y^{(k)})\}$ be a sequence generated by Algorithm 4.1. Then, $\{(x^{(k)}, y^{(k)})\}$ is bounded, and every accumulation point is a solution of* (3.1).

**Proof.** From a simple continuity argument, we can easily show that every accumulation point of $\{(x^{(k)}, y^{(k)})\}$ is a solution of SOCCP (3.1). So we only show the boundedness of $\{(x^{(k)}, y^{(k)})\}$. For a contradiction purpose, we assume that $\{(x^{(k)}, y^{(k)})\}$ is not bounded. Then, there exists a subsequence $\{(x^{(k)}, y^{(k)})\}_{k\in K}$ such that $\lim_{k\to\infty,\, k\in K} \|(x^{(k)}, y^{(k)})\| = \infty$. From the boundedness of $\mathcal{S}$, there exists a compact set $C \subset \Re^n \times \Re^n$ such that $\mathcal{S} \subset \operatorname{int} C$ and

$$(x^{(k)}, y^{(k)}) \notin C \tag{4.13}$$

for all $k \in K$ sufficiently large. Moreover, we have

$$\overline{m} := \min_{(x,y)\in\partial C} \Psi_{\rm NR}(x, y) > 0, \tag{4.14}$$

since any $(x, y) \in \partial C$ does not belong to $\mathcal{S}$ and $C$ is compact. Now, applying Lemma 4.1 with $\delta := \overline{m}/4 > 0$, we have

$$\Psi_{\mu_k,\,\varepsilon_k}(x, y) - \Psi_{\rm NR}(x, y) \leq \frac{1}{4}\overline{m} \tag{4.15}$$

and

$$\Psi_{\mu_k,\varepsilon_k}(x,y) - \Psi_{\rm NR}(x,y) \ \geq\ -\frac{1}{4}\overline{m} \tag{4.16}$$

for any $(x,y) \in C$ and $k \in K$ sufficiently large. Let $(\overline{x},\overline{y}) \in \mathcal{S} \subset C$ be a solution of SOCCP. Then, from (4.15), we have

$$\Psi_{\mu_k,\varepsilon_k}(\overline{x},\overline{y}) - \Psi_{\rm NR}(\overline{x},\overline{y}) \ =\ \Psi_{\mu_k,\varepsilon_k}(\overline{x},\overline{y}) \ \leq\ \frac{1}{4}\overline{m} \tag{4.17}$$

for all $k \in K$ sufficiently large. On the other hand, letting $(\tilde{x}^{(k)}, \tilde{y}^{(k)})$ be a solution of $\min_{(x,y)\in\partial C}$ $\Psi_{\mu_k,\varepsilon_k}(x,y)$, we have, for all $k \in K$ sufficiently large,

$$\begin{aligned}
\min_{(x,y)\in\partial C} \Psi_{\mu_k,\varepsilon_k}(x,y) &= \Psi_{\mu_k,\varepsilon_k}(\tilde{x}^{(k)},\tilde{y}^{(k)}) \\
&\geq -\frac{1}{4}\overline{m} + \Psi_{\rm NR}(\tilde{x}^{(k)},\tilde{y}^{(k)}) \\
&\geq -\frac{1}{4}\overline{m} + \overline{m} \\
&= \frac{3}{4}\overline{m},
\end{aligned} \tag{4.18}$$

where the first inequality follows from (4.16) and the second inequality follows from (4.14) and $(\tilde{x}^{(k)},\tilde{y}^{(k)}) \in \partial C$. Furthermore, since $\Psi_{\mu_k,\varepsilon_k}(x^{(k)},y^{(k)}) \leq \alpha_k$ from Step 2 of Algorithm 4.1, we have

$$\Psi_{\mu_k,\varepsilon_k}(x^{(k)},y^{(k)}) \ \leq\ \frac{1}{4}\overline{m} \tag{4.19}$$

for all $k \in K$ sufficiently large. Now, let $\overline{k} \in K$ be a sufficiently large integer satisfying (4.13), (4.17), (4.18) and (4.19). Then, applying Lemma 4.2 to $\Psi_{\mu_{\bar{k}},\varepsilon_{\bar{k}}}$ with $a := (\overline{x},\overline{y})$, $b := (x^{(\bar{k})},y^{(\bar{k})})$, $m := (3/4)\overline{m}$, we obtain the existence of $(\hat{x}^{(k)},\hat{y}^{(k)}) \in \Re^n \times \Re^n$ such that

$$\nabla\Psi_{\mu_{\bar{k}},\varepsilon_{\bar{k}}}(\hat{x}^{(k)},\hat{y}^{(k)}) = 0 \quad \text{and} \quad \Psi_{\mu_{\bar{k}},\varepsilon_{\bar{k}}}(\hat{x}^{(k)},\hat{y}^{(k)}) \geq \frac{3}{4}\overline{m} > 0.$$

However, this result contradicts Proposition 4.2. Hence, $\{(x^{(k)},y^{(k)})\}$ is bounded. ■

# 5 Numerical Experiments

In this section we present some numerical results for Algorithm 4.1.

## 5.1 Newton's Method

To obtain the next iterate in Step 2 of Algorithm 4.1, we use Newton's method with Armijo's step size rule [1, Sec. 1.2.1]. The specific algorithm is described as follows. In the algorithm, we denote $w^{(k)} := \binom{x^{(k)}}{y^{(k)}}$ for convenience.

**Algorithm 5.1** *Choose* $\rho$, $\beta$, $\eta \in (0,1)$, $\sigma \in (0, 1-\beta)$ *and* $\nu \in (0,\infty)$.

(Step 0) *Choose* $w^{(0)} \in \Re^{2n}$, $\mu_0 \in (0,\infty)$ *and* $\varepsilon_0 \in (0,\infty)$. *Set* $\alpha_0 := \eta\, \Psi_{\mu_0,\varepsilon_0}(w^{(0)})$ *and* $k := 0$.

(Step 1) *If* $\Psi_{\rm NR}(w^{(k)}) = 0$, *terminate the iteration.*

(Step 2)

(Step 2.0) *Set* $v^{(0)} := w^{(k)}$ *and* $j := 0$.

(Step 2.1) *Find the solution* $\hat{d}^{(j)}$ *of the system of equations*

$$H_{\mu_k,\varepsilon_k}(v^{(j)}) + \nabla H_{\mu_k,\varepsilon_k}(v^{(j)})^T \hat{d}^{(j)} = 0.$$

(Step 2.2) *If* $\Psi_{\mu_k,\varepsilon_k}(v^{(j)} + \hat{d}^{(j)}) \le \alpha_k$, *let* $w^{(k+1)} := v^{(j)} + \hat{d}^{(j)}$ *and go to Step 3. Otherwise go to Step 2.3.*

(Step 2.3) *Let* $m_j$ *be the smallest nonnegative integer such that*

$$\Psi_{\mu_k,\varepsilon_k}(v^{(j)}) - \Psi_{\mu_k,\varepsilon_k}(v^{(j)} + \rho^{m_j}\hat{d}^{(j)}) \ \ge\ -\sigma\rho^{m_j}\nabla\Psi_{\mu_k,\varepsilon_k}(v^{(j)})^T\hat{d}^{(j)}.$$

*Let* $v^{(j+1)} := v^{(j)} + \rho^{m_j}\hat{d}^{(j)}$.

(Step 2.4) *If*

$$\Psi_{\mu_k,\varepsilon_k}(v^{(j+1)}) \ \le\ \alpha_k,$$

*set* $w^{(k+1)} := v^{(j+1)}$, *and go to Step 3. Otherwise, set* $j := j+1$, *and go back to Step 2.1.*

(Step 3) *Let*

$$\alpha_{k+1} := \eta\, \Psi_{\mu_k,\varepsilon_k}(w^{(k+1)}),\ \ \mu_{k+1} := \min\left(\nu\sqrt{2\alpha_{k+1}}, \frac{\mu_k}{2}\right) \ \ \text{and}\ \ \varepsilon_{k+1} := \mu_{k+1}.$$

*Set* $k := k+1$, *and go back to Step 1.*

Note that $\{\alpha_k\}$ converges to 0 since $\alpha_{k+1} = \eta\, \Psi_{\mu_k,\varepsilon_k}(w^{(k)}) \le \eta\alpha_k$ and $\eta \in (0,1)$. The sequences $\{\mu_k\}$ and $\{\varepsilon_k\}$ also converge to 0 since $\mu_{k+1} \le \mu_k/2$ and $\varepsilon_{k+1} \le \varepsilon_k/2$. Hence, this algorithm has global convergence from Theorem 4.2.

## 5.2 Results of Experiments

In order to evaluate the efficiency of Algorithm 5.1, we have conducted some numerical experiments. In our experiments, we chose $\hat{g}(\alpha) = (\sqrt{\alpha^2+4}+\alpha)/2$, $\beta = 0.5$, $\rho = 0.5$, $\eta = 0.99$, $\sigma = 0.5$ and $\nu = 0.45$. Moreover, we let $\mu_0 := \|w^{(0)}\|$ and $\varepsilon_0 := \|w^{(0)}\|$. We employed $\Psi_{\rm NR}(w^{(k)})$ <1e-20 as the termination criterion. Our algorithm was coded in MATLAB 5.3.1 and run on a SUN ULTRA 60.

We solved three SOCCPs by Algorithm 5.1. The functions $f$ and the second-order cone constraints $\mathcal{K}$ of the problems are given by:

$$\text{Problem 1 } (n=3): \quad f(x) := \begin{pmatrix} 21 & -9 & 18 \\ -9 & 4 & -7 \\ 18 & -7 & 19 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 3 \\ 7 \\ 1 \end{pmatrix}, \quad \mathcal{K} = \mathcal{K}^3,$$

$$\text{Problem 2 } (n=3): \quad f(x) := \begin{pmatrix} 0.07x_1^3 - 4 \\ 0.04x_2^3 - 3.93 \\ 0.03x_3^3 - 5.72 \end{pmatrix}, \quad \mathcal{K} = \mathcal{K}^3,$$

$$\text{Problem 3 } (n=5): \quad f(x) := \begin{pmatrix} 24(2x_1 - x_2)^3 + \exp(x_1 - x_3) - 4x_4 + x_5 \\ -12(2x_1 - x_2)^3 + 3(3x_2 + 5x_3)/\sqrt{1 + (3x_2 + 5x_3)^2} - 6x_4 - 7x_5 \\ -\exp(x_1 - x_3) + 5(3x_2 + 5x_3)/\sqrt{1 + (3x_2 + 5x_3)^2} - 3x_4 + 5x_5 \\ 4x_1 + 6x_2 + 3x_3 - 1 \\ -x_1 + 7x_2 - 5x_3 + 2 \end{pmatrix},$$
$$\mathcal{K} = \mathcal{K}^3 \times \mathcal{K}^2.$$

We note that these three functions are monotone but not strongly monotone. In Problem 1, $f : \Re^3 \to \Re^3$ is an affine function whose coefficient matrix is symmetric and positive semidefinite but not positive definite. In Problem 2, $f : \Re^3 \to \Re^3$ is a strictly monotone function comprised of cubic and constant terms only. Noticing that its Jacobian is given by $\nabla f(x) = \mathrm{diag}(0.21x_1^2,\, 0.12x_2^2,\, 0.09x_3^2)$, it can be easily seen that $f$ is monotone but not strongly monotone. In Problem 3, $f : \Re^5 \to \Re^5$ is the function which appears in the KKT conditions for the SOCP:

$$\begin{array}{ll} \text{Minimize} & \exp(z_1 - z_3) + 3(2z_1 - z_2)^4 + \sqrt{1 + (3z_2 + 5z_3)^2} \\ \text{subject to} & \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \in \mathcal{K}^3, \quad \begin{pmatrix} 4 & 6 & 3 \\ -1 & 7 & -5 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix} \in \mathcal{K}^2. \end{array}$$

From the convexity of the objective function, it can be easily shown that $f$ is monotone.

Table 1 shows the behavior of Algorithm 5.1 starting from the initial point $x^{(0)} := (100, \ldots, 100)^T$ and $y^{(0)} := (-100, \ldots, -100)^T$. The table contains the following items for each $k$: The number of inner iterations $j$, the smoothing parameter $\mu_k$, the regularization parameter $\varepsilon_k$, the inexact minimization parameter $\alpha_k$, the functional value $\Psi_{\mu_k,\varepsilon_k}(w^{(k)})$, and the value of the merit function $\Psi_{\mathrm{NR}}(w^{(k)})$. We can see from Table 1 that the generated sequences converge very quickly. Especially near the solution, the behavior exhibits superlinear convergence.

Next, in Table 2, we give the results for Algorithm 5.1 from several different starting points randomly chosen from a sphere centered at the origin. For Problems 1, 2 and 3, their radius of the sphere is set to be 1e+10, 1e+3, and 50, respectively. Table 2 contains the following items: The norm of the starting point $\|(x^{(0)}, y^{(0)})\|$, the total number of the iterations (iter.), and the final value of the merit function $\Psi_{\mathrm{NR}}$. From Table 2, we confirm that our algorithm has global convergence. Moreover, we see that the choice of an initial point does not influence the number of iterations substantially.

Table 1: Convergent Behavior

| Problem | $k$ | $j$ | $\mu_k(=\varepsilon_k)$ | $\alpha_k$ | $\Psi_{\mu_k,\,\varepsilon_k}(w^{(k)})$ | $\Psi_{\mathrm{NR}}(w^{(k)})$ |
|---|---|---|---|---|---|---|
| Problem 1 | 0 | 0 | 2.449e+02 | 1.035e+09 | 1.035e+09 | 1.023e+07 |
| | 1 | 1 | 5.883e+01 | 8.546e+03 | 4.935e+04 | 8.538e+04 |
| | 2 | 1 | 1.516e+00 | 5.676e+00 | 5.652e+04 | 5.955e+04 |
| | 3 | 1 | 7.581e-01 | 4.637e+00 | 3.833e+00 | 3.551e+00 |
| | 4 | 1 | 1.564e-01 | 6.036e-02 | 8.186e-02 | 1.144e-01 |
| | 5 | 1 | 3.614e-02 | 3.225e-03 | 1.629e-03 | 1.882e-03 |
| | 6 | 1 | 7.075e-04 | 1.236e-06 | 4.333e-05 | 4.503e-05 |
| | 7 | 1 | 3.893e-06 | 3.741e-11 | 1.671e-08 | 1.690e-08 |
| | 8 | 1 | 8.642e-10 | 1.844e-18 | 5.106e-13 | 5.108e-13 |
| | 9 | 1 | 2.619e-14 | 1.694e-27 | 2.517e-20 | 2.517e-20 |
| | 10 | 1 | 6.843e-16 | 1.156e-30 | 1.930e-29 | 1.890e-29 |
| Problem 2 | 0 | 0 | 2.449e+02 | 8.050e+09 | 8.050e+09 | 3.713e+09 |
| | 1 | 1 | 1.225e+02 | 7.863e+08 | 3.932e+08 | 1.406e+08 |
| | 2 | 1 | 6.124e+01 | 3.763e+07 | 1.601e+07 | 3.583e+06 |
| | 3 | 1 | 3.062e+01 | 1.624e+06 | 6.231e+05 | 1.211e+05 |
| | 4 | 1 | 1.531e+01 | 4.850e+04 | 1.315e+04 | 7.141e+02 |
| | 5 | 1 | 7.655e+00 | 7.507e+02 | 2.136e+01 | 5.670e+02 |
| | 6 | 2 | 1.445e+00 | 5.158e+00 | 2.211e+01 | 4.093e+01 |
| | 7 | 1 | 7.226e-01 | 1.343e+00 | 1.978e+00 | 1.400e+01 |
| | 8 | 2 | 1.209e-01 | 3.607e-02 | 9.624e-01 | 2.313e+00 |
| | 9 | 1 | 5.264e-02 | 6.842e-03 | 6.777e-02 | 2.637e-01 |
| | 10 | 1 | 3.196e-03 | 2.522e-05 | 5.712e-02 | 6.489e-02 |
| | 11 | 1 | 1.598e-03 | 1.830e-05 | 2.354e-05 | 1.563e-04 |
| | 12 | 2 | 8.485e-07 | 1.778e-12 | 1.592e-05 | 1.595e-05 |
| | 13 | 1 | 4.243e-07 | 1.292e-12 | 1.667e-12 | 1.104e-11 |
| | 14 | 2 | 5.896e-14 | 8.583e-27 | 1.125e-12 | 1.125e-12 |
| | 15 | 1 | 2.948e-14 | 6.526e-27 | 7.896e-27 | 5.276e-26 |
| Problem 3 | 0 | 0 | 3.162e+02 | 3.604e+14 | 3.604e+14 | 3.600e+14 |
| | 1 | 1 | 1.581e+02 | 1.322e+14 | 1.321e+14 | 1.320e+14 |
| | 2 | 1 | 7.906e+01 | 1.185e+13 | 1.184e+13 | 1.184e+13 |
| | 3 | 1 | 3.953e+01 | 1.069e+12 | 1.068e+12 | 1.067e+12 |
| | 4 | 1 | 1.976e+01 | 9.710e+10 | 9.700e+10 | 9.691e+10 |
| | 5 | 1 | 9.882e+00 | 8.892e+09 | 8.882e+09 | 8.872e+09 |
| | 6 | 1 | 4.941e+00 | 8.222e+08 | 8.211e+08 | 8.200e+08 |
| | 7 | 1 | 2.471e+00 | 7.698e+07 | 7.686e+07 | 7.674e+07 |
| | 8 | 1 | 1.235e+00 | 7.291e+06 | 7.277e+06 | 7.263e+06 |
| | 9 | 1 | 6.176e-01 | 6.536e+05 | 6.518e+05 | 6.500e+05 |
| | 10 | 1 | 3.088e-01 | 6.366e+04 | 6.338e+04 | 6.311e+04 |
| | 11 | 1 | 1.544e-01 | 1.670e+04 | 1.667e+04 | 1.664e+04 |
| | 12 | 1 | 7.720e-02 | 6.276e+03 | 6.274e+03 | 6.271e+03 |
| | 13 | 1 | 3.860e-02 | 7.459e+01 | 7.431e+01 | 7.403e+01 |
| | 14 | 1 | 1.930e-02 | 6.428e+01 | 6.426e+01 | 6.423e+01 |
| | 15 | 1 | 9.651e-03 | 1.161e+00 | 1.163e+00 | 1.166e+00 |
| | 16 | 1 | 4.825e-03 | 2.720e-03 | 2.663e-03 | 2.620e-03 |
| | 17 | 1 | 2.413e-03 | 2.232e-03 | 2.210e-03 | 2.192e-03 |
| | 18 | 1 | 1.206e-03 | 1.833e-03 | 1.824e-03 | 1.816e-03 |
| | 19 | 1 | 6.032e-04 | 1.498e-03 | 1.494e-03 | 1.491e-03 |
| | 20 | 1 | 3.016e-04 | 1.217e-03 | 1.215e-03 | 1.213e-03 |
| | 21 | 1 | 1.508e-04 | 1.031e-03 | 1.030e-03 | 1.029e-03 |
| | 22 | 1 | 7.539e-05 | 7.309e-06 | 7.322e-06 | 7.340e-06 |
| | 23 | 1 | 1.377e-06 | 4.682e-12 | 1.852e-09 | 1.922e-09 |
| | 24 | 1 | 2.440e-09 | 1.470e-17 | 6.399e-13 | 6.422e-13 |
| | 25 | 1 | 7.731e-13 | 1.476e-24 | 2.017e-18 | 2.018e-18 |
| | 26 | 1 | 3.728e-16 | 3.432e-31 | 2.025e-25 | 2.027e-25 |

Table 2: Global Convergence

| Problem | $\|(x^{(0)}, y^{(0)})\|$ | iter. | $\Psi_{NR}$ |
|---|---|---|---|
| Problem 1 | 1.245e-03 | 8 | 7.511e-25 |
| | 4.005e-02 | 8 | 2.096e-23 |
| | 1.186e-01 | 8 | 4.133e-25 |
| | 4.223e+00 | 7 | 2.462e-24 |
| | 7.663e+00 | 8 | 5.279e-25 |
| | 1.309e+01 | 9 | 3.437e-25 |
| | 2.153e+01 | 11 | 4.223e-29 |
| | 4.510e+01 | 12 | 2.098e-26 |
| | 1.719e+03 | 11 | 6.993e-27 |
| | 1.209e+04 | 15 | 9.754e-21 |
| | 8.439e+04 | 12 | 1.043e-22 |
| | 1.003e+05 | 12 | 1.821e-21 |
| | 2.661e+06 | 13 | 6.054e-30 |
| | 4.685e+06 | 12 | 4.386e-28 |
| | 4.041e+07 | 12 | 7.409e-21 |
| | 4.660e+07 | 13 | 9.241e-27 |
| | 3.178e+08 | 21 | 2.258e-26 |
| | 4.038e+09 | 16 | 8.342e-29 |
| | 5.114e+09 | 23 | 3.869e-29 |
| | 5.686e+09 | 17 | 1.184e-28 |
| Problem 2 | 1.111e-03 | 12 | 2.465e-29 |
| | 5.470e-02 | 10 | 1.610e-23 |
| | 9.352e-02 | 10 | 1.172e-25 |
| | 6.271e-01 | 11 | 1.824e-21 |
| | 7.094e+00 | 17 | 9.044e-21 |
| | 1.445e+01 | 13 | 6.013e-22 |
| | 2.231e+01 | 16 | 2.193e-27 |
| | 4.078e+01 | 15 | 6.027e-21 |
| | 4.639e+01 | 18 | 2.948e-21 |
| | 7.285e+01 | 17 | 1.715e-28 |
| | 1.154e+02 | 19 | 1.144e-29 |
| | 1.902e+02 | 18 | 7.987e-21 |
| | 4.161e+02 | 18 | 7.676e-24 |
| | 5.172e+02 | 21 | 5.424e-24 |
| | 5.551e+02 | 22 | 3.843e-21 |
| | 5.812e+02 | 31 | 2.636e-21 |
| | 6.223e+02 | 21 | 7.691e-30 |
| | 6.492e+02 | 23 | 7.396e-30 |
| | 6.868e+02 | 17 | 1.043e-24 |
| | 8.000e+02 | 20 | 1.589e-25 |
| Problem 3 | 0.029 | 8 | 2.111e-21 |
| | 0.078 | 11 | 2.245e-27 |
| | 0.121 | 10 | 6.932e-23 |
| | 0.311 | 10 | 2.692e-31 |
| | 1.621 | 11 | 4.127e-30 |
| | 5.172 | 15 | 6.268e-31 |
| | 9.615 | 14 | 8.243e-26 |
| | 11.42 | 14 | 1.086e-24 |
| | 16.13 | 16 | 2.041e-31 |
| | 19.03 | 19 | 7.510e-21 |
| | 26.21 | 15 | 1.120e-26 |
| | 27.08 | 22 | 3.289e-22 |
| | 32.02 | 21 | 5.445e-22 |
| | 33.15 | 23 | 1.958e-31 |
| | 33.32 | 17 | 6.546e-26 |
| | 33.45 | 21 | 5.161e-30 |
| | 34.88 | 15 | 4.933e-23 |
| | 41.40 | 19 | 1.984e-24 |
| | 44.03 | 16 | 8.486e-31 |
| | 48.33 | 22 | 1.293e-30 |

# 6 Final Remarks

In this paper, we have shown that the merit function $\Psi_{\text{NR}}$ defined by (3.2) for SOCCP (3.1) is coercive if $f$ is strongly monotone, and that the smoothing function $\Psi_{\mu}$ defined by (4.4) is also coercive under the same condition. Moreover, based on the idea of the smoothing method and the regularization method, we have proposed a globally convergent algorithm for solving monotone SOCCP (3.1).

In Section 3 and 4, the function $F$ in SOCCP (1.1) is assumed to be of the form by $F(x, y, \zeta) = f(x) - y$. This assumption may seem rather restrictive. However, the KKT conditions for SOCP (1.3) can be written as the SOCCP with $F(x, y, \zeta) = f(x) - y$ as follows: In SOCP (1.3), let $z = z' - z''$ with $z' \in \Re^s_+$ and $z'' \in \Re^s_+$ , and denote $\hat{z} := \binom{z'}{z''} \in \Re^{2s}_+$, where $\Re^n_+$ is the $n$-dimensional nonnegative orthant. Moreover, define $\hat{\theta} : \Re^{2s} \to \Re$ by $\hat{\theta}(\hat{z}) = \theta(z' - z'')$, and $\hat{\gamma} : \Re^{2s} \to \Re^t$ by $\hat{\gamma}(\hat{z}) = \gamma(z' - z'')$. Then the SOCP (1.3) can be reformulated as

$$\begin{array}{ll} \text{Minimize} & \hat{\theta}(\hat{z}) \\ \text{subject to} & \begin{pmatrix} \hat{\gamma}(\hat{z}) \\ \hat{z} \end{pmatrix} \in \mathcal{K} \times \Re^{2s}_+, \end{array} \tag{6.1}$$

and the KKT conditions for (6.1) are written as

$$\begin{aligned} \nabla\hat{\theta}(\hat{z}) - \begin{pmatrix} \nabla\hat{\gamma}(\hat{z}) & I \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix} &= 0, \\ \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix} \in \mathcal{K} \times \Re^{2s}_+, \quad \begin{pmatrix} \hat{\gamma}(\hat{z}) \\ \hat{z} \end{pmatrix} \in \mathcal{K} \times \Re^{2s}_+, \quad \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix}^T \begin{pmatrix} \hat{\gamma}(\hat{z}) \\ \hat{z} \end{pmatrix} &= 0. \end{aligned} \tag{6.2}$$

Now, let $\hat{\mu}_1 = \hat{\gamma}(\hat{z})$, and notice that Proposition 2.2 holds. Then, (6.2) can be rewritten as

$$\begin{aligned} \begin{pmatrix} \hat{\gamma}(\hat{z}) \\ \nabla\hat{\theta}(\hat{z}) - \nabla\hat{\gamma}(\hat{z})\hat{\lambda}_1 \end{pmatrix} &= \begin{pmatrix} \hat{\mu}_1 \\ \hat{\lambda}_2 \end{pmatrix}, \\ \begin{pmatrix} \hat{\lambda}_1 \\ \hat{z} \end{pmatrix} \in \mathcal{K} \times \Re^{2s}_+, \quad \begin{pmatrix} \hat{\mu}_1 \\ \hat{\lambda}_2 \end{pmatrix} \in \mathcal{K} \times \Re^{2s}_+, \quad \begin{pmatrix} \hat{\lambda}_1 \\ \hat{z} \end{pmatrix}^T \begin{pmatrix} \hat{\mu}_1 \\ \hat{\lambda}_2 \end{pmatrix} &= 0. \end{aligned} \tag{6.3}$$

Setting

$$x = \begin{pmatrix} \hat{\lambda}_1 \\ \hat{z} \end{pmatrix}, \quad y = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\lambda}_2 \end{pmatrix}, \quad f(x) = \begin{pmatrix} \hat{\gamma}(\hat{z}) \\ \nabla\hat{\theta}(\hat{z}) - \nabla\hat{\gamma}(\hat{z})\hat{\lambda}_1 \end{pmatrix}, \tag{6.4}$$

the KKT conditions (6.3) for SOCP (1.3) can be reduced to the SOCCP with $F(x, y, \zeta) = f(x) - y$.

Note that the KKT conditions (6.3) for SOCP (1.3) contain more variables than the original KKT conditions (1.4). Furthermore, some desirable properties of the functions involved in SOCP (1.3) may be lost. For example, even if $\gamma$ in SOCP (1.3) is strictly convex, $\hat{\gamma}$ in (6.3) is convex but not strictly convex in general. Hence, it will be useful to develop a method that can directly deal with the KKT conditions (1.4), or more generally, SOCCP involving the function $F(x, y, \zeta)$ which is not restricted to be of the form $F(x, y, \zeta) = f(x) - y$.

In Sections 3 and 4, it is also assumed that $\mathcal{K} = \mathcal{K}^n$. For the general case where $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}$, it can be shown that the merit function $\Psi_{\text{NR}}$ defined by (2.9) with $F(x, y, \zeta) = f(x) - y$ is coercive when $f$ has the following property.

**Property 1** *For* $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}$, *let* $x = (x^1, \ldots, x^m) \in \Re^{n_1} \times \cdots \times \Re^{n_m}$, $z = (z^1, \ldots, z^m) \in \Re^{n_1} \times \cdots \times \Re^{n_m}$. *Moreover, let* $f : \Re^n \to \Re^{n_1} \times \cdots \times \Re^{n_m}$ *be represented as* $f(x) = (f^1(x), \ldots, f^m(x))$ *with* $f^i : \Re^n \to \Re^{n_i}$, $i = 1, \ldots, m$. *Then, there exists* $\varepsilon > 0$ *such that*

$$\max_i (x^i - z^i)^T \left\{ f^i(x) - f^i(z) \right\} \quad \geq \quad \varepsilon \|x - z\|^2$$

*holds for any* $x$, $z \in \Re^n$.

Note that strongly monotone functions have Property 1. When $n_1 = \cdots = n_m = 1$, a function satisfying Property 1 is reduced to a uniform $P$ function. As a future research issue, it is interesting to see whether the condition for coerciveness of the merit function $\Psi_{\rm NR}$ can be weakened. In the case of NCP, it has been shown that some merit functions are coercive under a condition weaker than the uniform $P$ property [6].


**Acknowledgments**

First of all, I would like to express my sincere thanks and appreciation to Professor Masao Fukushima and Assistant Professor Nobuo Yamashita. Although I often troubled them due to my greenness, they kindly looked after me and gave me plenty of precise advice. Without their help and advice, I could never write up this paper. I would also like to tender my acknowledgments to Associate Professor Tetsuya Takine. He gave me a lot of comments at research conference beyond the difference of the research field. Finally, I would like to thank all the members in Fukushima's Laboratory.


# References


[1] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientifics, Belmont, Massachusetts, 1995.

[2] C. Chen and O. L. Mangasarian, *A class of smoothing functions for nonlinear and mixed complementarity problems,* Computational Optimization and Applications, Vol. 5, 1996, pp. 97–138.

[3] X. Chen, *Smoothing methods for complementarity problems and their applications: A survey*, Journal of the Operations Research Society of Japan, Vol. 43, 2000, pp. 32–47.

[4] X. Chen and L. Qi, *Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities*, Mathematics of Computation, Vol. 67, 1998, pp. 519–540.

[5] A. L. Dontchev and T. Zolezzi, *Well-posed optimization problems*, Lecture Notes in Mathematics 1543, Springer-Verlag, New York, 1993.

[6] F. Facchinei and C. Kanzow, *Beyond monotonicity in regularization methods for nonlinear complementarity problems*, SIAM Journal on Control and Optimization, Vol. 37, 1999, pp. 1150–1161.

[7] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[8] M. Fukushima, Z.-Q. Luo and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM Journal on Optimization, Vol. 12, 2001, pp. 436–460.

[9] C. Kanzow and H. Pieper, *Jacobian smoothing methods for general complementarity problems*, SIAM Journal on Optimization, Vol. 9, 1999, pp. 342–373.

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

[11] R. S. Palais and C.-L. Terng, *Critical Point Theory and Submanifold Geometry*, Lecture Notes in Mathematics 1353, Springer-Verlag, Berlin, 1988.

[12] M. Patriksson, *Merit functions and descent algorithms for a class of variational inequality problems*, Optimization, Vol. 41, 1997, pp. 37–55.

[13] L. Qi and X. Chen, *A globally convergent successive approximation method for severely nonsmooth equations*, SIAM Journal on Control and Optimization, Vol. 33, 1995, pp. 402–418.

[14] L. Qi and D. Sun, *Smoothing functions and a smoothing Newton method for complementarity problems and variational inequality problems*, School of Mathematics, University of New South Wales, 1998.