Journal of Global Optimization (2022) 83:671–697
https://doi.org/10.1007/s10898-021-01116-w

An approximate lower order penalty approach for solving second-order cone linear complementarity problems

Zijun Hao[1] · Chieu Thanh Nguyen[2] · Jein-Shan Chen[2]

Received: 24 March 2020 / Accepted: 15 November 2021 / Published online: 3 December 2021

Abstract
Based on a class of smoothing approximations to projection function onto second-order cone, an approximate lower order penalty approach for solving second-order cone linear complementarity problems (SOCLCPs) is proposed, and four kinds of specific smoothing approximations are considered. In light of this approach, the SOCLCP is approximated by asymptotic lower order penalty equations with penalty parameter and smoothing parameter. When the penalty parameter tends to positive infinity and the smoothing parameter monotonically decreases to zero, we show that the solution sequence of the asymptotic lower order penalty equations converges to the solution of the SOCLCP at an exponential rate under a mild assumption. A corresponding algorithm is constructed and numerical results are reported to illustrate the feasibility of this approach. The performance profile of four specific smoothing approximations is presented, and the generalization of two approximations are also investigated.

Keywords Second-order cone · Linear complementarity problem · Lower order penalty approach · Exponential convergence rate

Mathematics Subject Classification 90C25 · 90C30 · 90C33

The author's work is supported by the National Natural Science Foundation of China Nos. 11661002, 11871383), the Natural Science Fund of Ningxia (No. 2020AAC03236), the First-class Disciplines Foundation of Ningxia (No. NXYLXK2017B09). J.-S. Chen: The author's work is supported by Ministry of Science and Technology, Taiwan.

✉ Jein-Shan Chen
jschen@math.ntnu.edu.tw

Zijun Hao
zijunhao@126.com

Chieu Thanh Nguyen
thanhchieu90@gmail.com

1 School of Mathematics and Information Science, North Minzu University, Yinchuan 750021, China

2 Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan

1 Introduction

This paper targets the following second-order cone linear complementarity problem (SOCLCP), which is to find $x \in \mathbb{R}^n$, such that

$$x \in \mathcal{K},\ \ Ax - b \in \mathcal{K},\ \ x^T(Ax - b) = 0, \tag{1}$$

where A is an $n \times n$ matrix, b is a vector in $\mathbb{R}^n$, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOCs), also called Lorentz cones [7,18]. In other words,

$$\mathcal{K} := \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_r} \tag{2}$$

with $r, n_1, \ldots, n_r \geq 1$, $n_1 + \cdots + n_r = n$ and

$$\mathcal{K}^{n_i} := \left\{(x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n_i - 1} \mid \|x_2\| \leq x_1\right\}, i = 1, \ldots, r,$$

where $\|\cdot\|$ denotes the Euclidean norm and $(x_1, x_2) := (x_1, x_2^T)^T$. Note that $\mathcal{K}^1$ denotes the set of nonnegative real numbers $\mathbb{R}_+$. The SOCLCP, as an extension of the linear complementarity problem (LCP), has a wide range of applications in linear and quadratic programming problems, computer science, game theory, economics, finance, engineering, and network equilibrium problems [3,15,17,26,27,30].

During the past several years, there are many methods proposed for solving the SOCLCPs (1)–(2), including the interior-point method [1,28,32], the smoothing Newton method [14,19,24], the smoothing-regularization method [23], the semismooth Newton method [25,33], the merit function method [5,10,12], and the matrix splitting method [22,41] etc. Although the effectiveness of some methods has been improved substantially in recent years, the fact remains that there still have many complementarity problems require efficient and accurate numerical methods. The penalty methods are well-known for solving constrained optimization problems which possess many nice properties. More specifically, the l_1 exact penalty function method and lower order penalty function method are known as the approaches which hold many nice properties and attracts much attention [2,20,29,34,39,40]. The smoothing of the exact penalty methods are also proposed [35,37,38]. Besides, Wang and Yang [36] focus on the power of lower order penalty function, and propose a power penalty method for solving LCP based on approximating the LCP by nonlinear equations. It shows that under some mild assumptions, the solution sequence of the nonlinear equations converges to the solution of the LCP at an exponential rate when the penalty parameter tends to positive infinity. Based on the method in [36], Hao et al. [21] propose a power penalty method for solving the SOCLCP with a single $\mathcal{K} = \mathcal{K}^n$, i.e.,

$$x \in \mathcal{K}^n,\ \ Ax - b \in \mathcal{K}^n,\ \ x^T(Ax - b) = 0, \tag{3}$$

where $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$. In particular, they consider the power penalty equations:

$$Ax - \alpha[x]_-^{1/k} = b, \tag{4}$$

where $k \geq 1$ and $\alpha \geq 1$ are parameters,

$$[x]_-^{1/k} = [\lambda_1(x)]_-^{1/k} u_x^{(1)} + [\lambda_2(x)]_-^{1/k} u_x^{(2)}$$

with $[t]_- = \max\{0, -t\}$ and the spectral decomposition (will be introduced later in (5)). Under a mild assumption of matrix A, as $\alpha \to +\infty$, the solution sequence of (4) converges to the solution of the SOCLCP (3) at an exponential rate.

In this paper, we further enhance improvement and extension of the method and the problem studied in [21]. We first generalize $[x]_-^{1/k}$ in (4) to general lower order penalty function $[x]_-^\sigma$ with $\sigma \in (0, 1]$, then focus on a class of approximate function to $[x]_-^\sigma$ for solving the general SOCLCP (1) instead of the SOCLCP (3) with single SOC constraint. In addition, we construct a class of functions $\Phi^-(\mu, x)^\sigma$ to approximate $[x]_-^\sigma$ as $\mu \to 0^+$. Four kinds of specific smoothing approximations are studied. Theoretically, we prove that the solution sequence of the approximating lower order penalty equations converge to the solution of the SOCLCP (1) at an exponential rate $O(\alpha^{-1/\sigma})$ when $\alpha \to +\infty$ and $\mu \to 0^+$. This generalizes all its counterparts in the literature. Moreover, a corresponding algorithm is constructed and numerical results are also reported to examine the feasibility of the proposed method. The performance profile of those specific smoothing approximations is presented, and the generalization of two approximations are also investigated.

This paper is organized as follows: In Sect. 2, we review some properties related to the single SOC which is the basis for our subsequent analysis. In Sect. 3, a class of approximation functions for lower order penalty function is constructed, and four kinds of specific smoothing approximations are investigated. In Sect. 4, we study the approximating lower order penalty equations for solving the SOCLCP (1), and prove the convergence analysis. In Sect. 5, a corresponding algorithm is constructed and the preliminary numerical experiments are presented. The performance profiles of the considered four specific smoothing approximations and the generalization of two approximations are also considered. Finally, we draw the conclusion in Sect. 6.

For simplicity, we denote the interior of single SOC $\mathcal{K}^n$ by $\mathrm{int}(\mathcal{K}^n)$. For any x, y in $\mathbb{R}^n$, we write $x \succeq_{\mathcal{K}^n} y$ if $x - y \in \mathcal{K}^n$ and write $x \succ_{\mathcal{K}^n} y$ if $x - y \in \mathrm{int}(\mathcal{K}^n)$. In other words, we have $x \succeq_{\mathcal{K}^n} 0$ if and only if $x \in \mathcal{K}^n$, and $x \succ_{\mathcal{K}^n} 0$ if and only if $x \in \mathrm{int}(\mathcal{K}^n)$. We usually denote $(x, y) := (x^T, y^T)^T$ for the concatenation of two column vectors x, y for simplicity. The notation $\|\cdot\|_p$ denotes the usual l_p-norm on $\mathbb{R}^n$ for any $p \geq 1$. In particular, it is Euclidean norm $\|\cdot\|$ when $p = 2$.

2 Preliminary results

In this section, we first recall some basic concepts and preliminary results related to a single SOC $\mathcal{K} = \mathcal{K}^n$ that will be used in the subsequent analysis. All of the analysis are then carried over to the general structure $\mathcal{K}$ (2). For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, their *Jordan product* [7,18] is defined as

$$x \circ y := (\langle x, y \rangle, y_1 x_2 + x_1 y_2).$$

We write $x + y$ to mean the usual componentwise addition of vectors and x^2 to mean $x \circ x$. The identity element under this product is $e = (1, 0, \ldots, 0)^T \in \mathbb{R}^n$. It is known that $x^2 \in \mathcal{K}^n$ for all $x \in \mathbb{R}^n$. Moreover, if $x \in \mathcal{K}^n$, then there is a unique vector in $\mathcal{K}^n$, denoted by $x^{\frac{1}{2}}$, such that $(x^{\frac{1}{2}})^2 = x^{\frac{1}{2}} \circ x^{\frac{1}{2}} = x$. For any $x \in \mathbb{R}^n$, we define $x^0 = e$ if $x \neq 0$. For any integer $k \geq 1$, we recursively define the powers of element as $x^k = x \circ x^{k-1}$, and define $x^{-k} = (x^k)^{-1}$ if $x \in \mathrm{int}(\mathcal{K}^n)$. The Jordan product is *not associative* for $n > 2$, but it is power associated, i.e., $x \circ (x \circ x) = (x \circ x) \circ x$. Thus, for any positive integer p, the form x^p is definite, and $x^{m+n} = x^m \circ x^n$ for all positive integer m and n. Note that $\mathcal{K}^n$ is *not closed* under the Jordan product for $n > 2$.

In the following, we recall the expression of the spectral decomposition of x with respect to SOC, see [5–8,10–12,18,19,33]. For $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the spectral decomposition of x with respect to SOC is given by

$$x = \lambda_1(x) u_x^{(1)} + \lambda_2(x) u_x^{(2)}, \tag{5}$$

where for $i = 1, 2$,

$$\lambda_i(x) = x_1 + (-1)^i \|x_2\|, \quad u_x^{(i)} = \begin{cases} \frac{1}{2}(1, (-1)^i \frac{x_2}{\|x_2\|}) & \text{if } \|x_2\| \neq 0, \\ \frac{1}{2}(1, (-1)^i w) & \text{if } \|x_2\| = 0, \end{cases} \tag{6}$$

with $w \in \mathbb{R}^{n-1}$ being any unit vector. The two scalars $\lambda_1(x)$ and $\lambda_2(x)$ are called spectral values of x, while the two vectors $u_x^{(1)}$ and $u_x^{(2)}$ are called the spectral vectors of x. Moreover, it is obvious that the spectral decomposition of $x \in \mathbb{R}^n$ is unique if $x_2 \neq 0$.

Some basic properties of the spectral decomposition in the Jordan algebra associated with SOC are stated as below, whose proofs can be found in [6,7,18,19].

Proposition 2.1 *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with the spectral values $\lambda_1(x), \lambda_2(x)$ and spectral vectors $u_x^{(1)}, u_x^{(2)}$ given as (6), we have:*

(a) $u_x^{(1)} \circ u_x^{(2)} = 0$ *and* $u_x^{(i)} \circ u_x^{(i)} = u_x^{(i)}$, $\|u_x^{(i)}\|^2 = 1/2$ *for* $i = 1, 2$.
(b) $\lambda_1(x), \lambda_2(x)$ *are nonnegative* (*positive*) *if and only if* $x \in \mathcal{K}^n$ $(x \in \operatorname{int}(\mathcal{K}^n))$.
(c) *For any* $x \in \mathbb{R}^n$, $x \succeq_{\mathcal{K}^n} 0$ *if and only if* $\langle x, y\rangle \geq 0$ *for all* $y \succeq_{\mathcal{K}^n} 0$.

The spectral decomposition (5)–(6) and the Proposition 2.1 indicate that x^k can be described as $x^k = \lambda_1^k(x) u_x^{(1)} + \lambda_2^k(x) u_x^{(2)}$. For any $x \in \mathbb{R}^n$, let $[x]_+$ denote the projection of x onto $\mathcal{K}^n$, and $[x]_-$ be the projection of $-x$ onto the dual cone $(\mathcal{K}^n)^*$ of $\mathcal{K}^n$, where the dual cone $(\mathcal{K}^n)^*$ is defined by $(\mathcal{K}^n)^* := \{y \in \mathbb{R}^n \mid \langle x, y\rangle \geq 0, \forall x \in \mathcal{K}^n\}$. In fact, by Proposition 2.1, the dual cone of $\mathcal{K}^n$ being itself, i.e., $(\mathcal{K}^n)^* = \mathcal{K}^n$. Due to the special structure of $\mathcal{K}^n$, the explicit formula of projection of $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ onto $\mathcal{K}^n$ is obtained in [14,17,19] as below

$$[x]_+ = \begin{cases} x & \text{if } x \in \mathcal{K}^n, \\ 0 & \text{if } x \in -\mathcal{K}^n, \\ u & \text{otherwise}, \end{cases} \quad \text{where} \quad u = \begin{bmatrix} \frac{x_1 + \|x_2\|}{2} \\ \left(\frac{x_1 + \|x_2\|}{2}\right) \frac{x_2}{\|x_2\|} \end{bmatrix}.$$

Similarly, the expression of $[x]_-$ can be written out as

$$[x]_- = \begin{cases} 0 & \text{if } x \in \mathcal{K}^n, \\ -x & \text{if } x \in -\mathcal{K}^n, \\ v & \text{otherwise}, \end{cases} \quad \text{where} \quad v = \begin{bmatrix} -\frac{x_1 - \|x_2\|}{2} \\ \left(\frac{x_1 - \|x_2\|}{2}\right) \frac{x_2}{\|x_2\|} \end{bmatrix}.$$

It is easy to verify that $x = [x]_+ - [x]_-$ and

$$[x]_+ = [\lambda_1(x)]_+ u_x^{(1)} + [\lambda_2(x)]_+ u_x^{(2)}, \quad [x]_- = [\lambda_1(x)]_- u_x^{(1)} + [\lambda_2(x)]_- u_x^{(2)},$$

where $[\alpha]_+ = \max\{0, \alpha\}$ and $[\alpha]_- = \max\{0, -\alpha\}$ for $\alpha \in \mathbb{R}$. Thus, it can be seen that $[x]_+, [x]_- \in \mathcal{K}^n$ and $[x]_+ \circ [x]_- = 0$.

Putting these analyses into a single SOC $\mathcal{K}^{n_i}$, $i = 1, \ldots, r$ in (2), we can extend them to the general case $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_r}$. More specifically, for any $x = (x_1, \ldots, x_r) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r}$, $y = (y_1, \ldots, y_r) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r}$, their *Jordan product* is defined as

$$x \circ y := (x_1 \circ y_1, \ldots, x_r \circ y_r).$$

Let $[x]_+, [x]_-$ respectively denote the projection of x onto $\mathcal{K}$ and the projection of $-x$ onto the dual cone $\mathcal{K}^* = \mathcal{K}$, then

$$[x]_+ := ([x_1]_+, \ldots, [x_r]_+), \quad [x]_- := ([x_1]_-, \ldots, [x_r]_-), \tag{7}$$

where $[x_i]_+, [x_i]_-$ for $i = 1, \ldots, r$ respectively denote the projection of x_i onto the single SOC $\mathcal{K}^{n_i}$ and the projection of $-x_i$ onto $(\mathcal{K}^{n_i})^*$.

3 Approximation of projection function with power

This section is devoted to presenting a way to generate smoothing functions for the plus function $[t]_+ = \max\{0, t\}$ and minus function $[t]_- = \max\{0, -t\}$ via convolution which was proposed by Chen and Mangasarian [4]. First, we consider the piecewise continuous function $d(t)$ with finite number of pieces, which is a density (kernel) function. In other words, it satisfies

$$d(t) \geq 0 \quad \text{and} \quad \int_{-\infty}^{+\infty} d(t)dt = 1. \tag{8}$$

Next, we define $\hat{s}(\mu, t) := \frac{1}{\mu} d\left(\frac{t}{\mu}\right)$, where μ is a positive parameter. If $\int_{-\infty}^{+\infty} |t|\, d(t)dt < +\infty$, then a smoothing approximation for $[t]_+$ is formed. In particular,

$$\phi^+(\mu, t) = \int_{-\infty}^{+\infty} (t-s)_+ \hat{s}(\mu, s)ds = \int_{-\infty}^{t} (t-s)\hat{s}(\mu, s)ds \approx [t]_+. \tag{9}$$

The following proposition states the properties of $\phi^+(\mu, t)$, whose proofs can be found in [4, Proposition 2.2].

Proposition 3.1 *Let $d(t)$ be a density function satisfying (8) and $\hat{s}(\mu, t) = \frac{1}{\mu} d\left(\frac{t}{\mu}\right)$ with positive parameter μ. If $d(t)$ is piecewise continuous with finite number of pieces and $\int_{-\infty}^{+\infty} |t|\, d(t)dt < +\infty$. Then, the function $\phi^+(\mu, t)$ defined by (9) possesses the following properties.*

(a) *$\phi^+(\mu, t)$ is continuously differentiable.*
(b) *$-D_2\mu \leq \phi^+(\mu, t) - [t]_+ \leq D_1\mu$, where*

$$D_1 = \int_{-\infty}^{0} |t| d(t)dt \text{ and } D_2 = \max\left\{\int_{-\infty}^{+\infty} td(t)dt, 0\right\}.$$

(c) *$\frac{\partial}{\partial t}\phi^+(\mu, t)$ is bounded satisfying $0 \leq \frac{\partial}{\partial t}\phi^+(\mu, t) \leq 1$.*

From Proposition 3.1(b), we have

$$\lim_{\mu \to 0^+} \phi^+(\mu, t) = [t]_+$$

under the assumptions of this proposition. Applying the above way of generating smoothing function to approximate $[t]_- = \max\{0, -t\}$, which appears in equation (4), we also achieve a smoothing approximation as follows:

$$\phi^-(\mu, t) = \int_{-\infty}^{-t} (-t-s)\hat{s}(\mu, -s)ds = \int_{t}^{+\infty} (s-t)\hat{s}(\mu, s)ds \approx [t]_-. \tag{10}$$

Similar to Proposition 3.1, we have the below properties for $\phi^-(\mu, t)$.

Proposition 3.2 *Let $d(t)$ and $\hat{s}(\mu, t)$ be as in Proposition 3.1 with the same assumptions. Then, the function $\phi^-(\mu, t)$ defined by (10) possesses the following properties.*

(a) *$\phi^-(\mu, t)$ is continuously differentiable.*
(b) *$-D_2\mu \leq \phi^-(\mu, t) - [t]_- \leq D_1\mu$, where*

$$D_1 = \int_{0}^{+\infty} |t| d(t)dt \text{ and } D_2 = \max\left\{\int_{-\infty}^{+\infty} td(t)dt, 0\right\}.$$

(c) $\frac{\partial}{\partial t}\phi^-(\mu, t)$ *is bounded satisfying* $-1 \leq \frac{\partial}{\partial t}\phi^-(\mu, t) \leq 0$.

Similar to Proposition 3.1, we also obtain $\lim_{\mu \to 0^+} \phi^-(\mu, t) = [t]_-$. Therefore, in view of Proposition 3.1 and 3.2, we know that $\phi^+(\mu, t)$ defined by (9) and $\phi^-(\mu, t)$ defined by (10), are the smoothing functions of $[t]_+$ and $[t]_-$, respectively. Accordingly, using the continuity of compound function and $\phi^+(\mu, t) \geq 0$, $\phi^-(\mu, t) \geq 0$, we can generate approximate function (not necessarily smooth) for $[t]_+^\sigma$ and $[t]_-^\sigma$, see below lemma.

Lemma 3.1 *Under the assumptions of Proposition 3.1, let* $\phi^+(\mu, t), \phi^-(\mu, t)$ *be the smoothing functions of* $[t]_+, [t]_-$*, defined by* (9) *and* (10) *respectively. Then, for any* $\sigma > 0$*, we have*

(a) $\lim\limits_{\mu \to 0^+} \phi^+(\mu, t)^\sigma = [t]_+^\sigma$,
(b) $\lim\limits_{\mu \to 0^+} \phi^-(\mu, t)^\sigma = [t]_-^\sigma$.

By modifying the smoothing functions used in [4,9,31], we have four specific smoothing functions for $[t]_-$ as well:

$$\phi_1^-(\mu, t) = -t + \mu \ln\left(1 + e^{\frac{t}{\mu}}\right), \tag{11}$$

$$\phi_2^-(\mu, t) = \begin{cases} 0 & \text{if} \quad \mathrm{t} \geq \frac{-}{2}, \\ \frac{1}{2\mu}\left(-t + \frac{\mu}{2}\right)^2 & \text{if} \ -\frac{-}{2} < \mathrm{t} < \frac{-}{2}, \\ -t & \text{if} \quad \mathrm{t} \leq -\frac{-}{2}, \end{cases} \tag{12}$$

$$\phi_3^-(\mu, t) = \frac{\sqrt{4\mu^2 + t^2} - t}{2}, \tag{13}$$

$$\phi_4^-(\mu, t) = \begin{cases} 0 & \text{if} \quad \mathrm{t} > 0, \\ \frac{t^2}{2\mu} & \text{if} \ -^{-} \leq \mathrm{t} \leq 0, \\ -t - \frac{\mu}{2} & \text{if} \quad \mathrm{t} < -^{-}, \end{cases} \tag{14}$$

where the corresponding kernel functions are

$$d_1(t) = \frac{e^t}{(1 + e^t)^2},$$
$$d_2(t) = \begin{cases} 1 & \text{if} \ -\frac{1}{2} \leq \mathrm{t} \leq \frac{1}{2}, \\ 0 & \text{otherwise}, \end{cases}$$
$$d_3(t) = \frac{2}{(t^2 + 4)^{\frac{3}{2}}},$$
$$d_4(t) = \begin{cases} 1 & \text{if} \ -1 \leq \mathrm{t} \leq 0, \\ 0 & \text{otherwise}. \end{cases}$$

For those specific functions (11)–(14), they certainly obey Proposition 3.2 and Lemma 3.1. The graphs of $[t]_-$ and $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$ with $\mu = 0.1$ are depicted in Fig. 1.

From Fig. 1, we see that, for a fixed $\mu > 0$, the function $\phi_2^-(\mu, t)$ seems the one which best approximate the function $[t]_-$ among all $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$. Indeed, for a fixed $\mu > 0$ and all $t \in \mathbb{R}$, we have

$$\phi_3^-(\mu, t) \geq \phi_1^-(\mu, t) \geq \phi_2^-(\mu, t) \geq [t]_- \geq \phi_4^-(\mu, t). \tag{15}$$

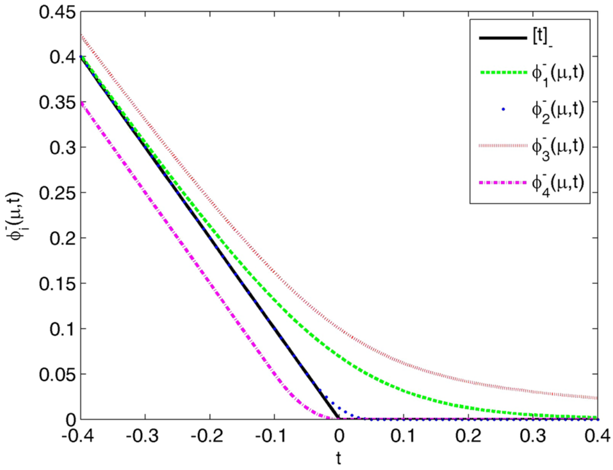

Fig. 1 Graphs of $[t]_-$ and $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$ with $\mu = 0.1$.

Furthermore, we shall show that $\phi_2^-(\mu, t)$ is the function closest to $[t]_-$ in the sense of the infinite norm. For any fixed $\mu > 0$, it is clear that

$$\lim_{|t|\to\infty} \left|\phi_i^-(\mu, t) - [t]_-\right| = 0,\ \ i = 1, 2, 3.$$

The functions $\phi_i^-(\mu, t) - [t]_-$, $i = 1, 3$ have no stable point but unique non-differentiable point $t = 0$, and $\phi_2^-(\mu, t) - [t]_-$ is non-zero only on the interval $(-\mu/2, \mu/2)$ with $\max_{t\in(-\mu/2,\mu/2)} \left|\phi_2^-(\mu, t) - [t]_-\right| = \phi_2^-(\mu, 0)$. These imply that

$$\max_{t\in\mathbb{R}} \left|\phi_i^-(\mu, t) - [t]_-\right| = \left|\phi_i^-(\mu, 0)\right|, i = 1, 2, 3.$$

Since $\phi_1^-(\mu, 0) = (\ln 2)\mu \approx 0.7\mu$, $\phi_2^-(\mu, 0) = \mu/8$, $\phi_3^-(\mu, 0) = \mu$, we obtain

$$\begin{aligned}
\|\phi_1^-(\mu, t) - [t]_-\|_\infty &= (\ln 2)\mu,\\
\|\phi_2^-(\mu, t) - [t]_-\|_\infty &= \mu/8,\\
\|\phi_3^-(\mu, t) - [t]_-\|_\infty &= \mu.
\end{aligned}$$

On the other hand, it is obvious that $\max_{t\in\mathbb{R}} \left|\phi_4^-(\mu, t) - [t]_-\right| = \mu/2$, which says

$$\|\phi_4^-(\mu, t) - [t]_-\|_\infty = \mu/2.$$

In summary, we have

$$\|\phi_3^-(\mu, t) - [t]_-\|_\infty > \|\phi_1^-(\mu, t) - [t]_-\|_\infty > \|\phi_4^-(\mu, t) - [t]_-\|_\infty > \|\phi_2^-(\mu, t) - [t]_-\|_\infty. \tag{16}$$

The orderings of (15) and (16) indicate the behavior of $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$ for fixed $\mu > 0$. When taking $\mu \to 0^+$, we know $\lim_{\mu\to 0^+} \phi_i^-(\mu, t) = [t]_-$, $i = 1, 2, 3, 4$ and $\phi_2^-(\mu, t)$ is the closest to $[t]_-$, which can be verified by geometric views depicted as in Fig. 2.

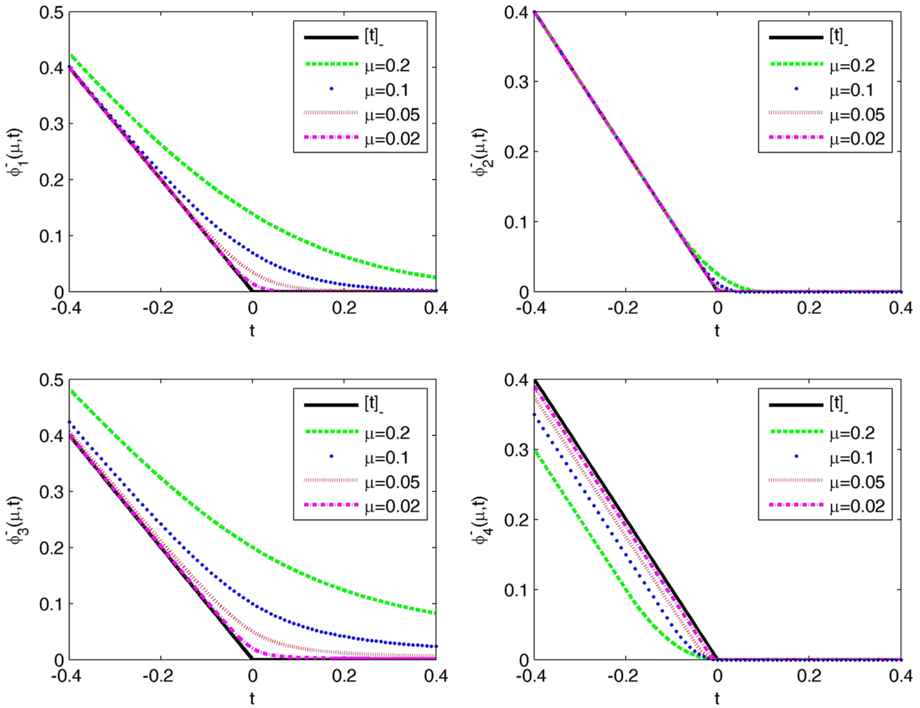

Fig. 2 Graphs of $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$ with different μ

Remark 3.1 For any $\mu > 0, \sigma > 0$ and continuously differentiable $\phi^-(\mu, t)$ defined by (10), it can be easily seen that, $\phi^-(\mu, t)^\sigma$ is continuous function about t, but may not be differentiable. For example, $\phi_1^-(\mu, t)^\sigma$, $\phi_3^-(\mu, t)^\sigma$ are continuously differentiable, but $\phi_2^-(\mu, t)^\sigma$, $\phi_4^-(\mu, t)^\sigma$ are not continuously differentiable for $\sigma = 1/2$ since the non-differentiable points are $t = \mu/2$ and $t = 0$ respectively. Their geometric views are depicted in Fig. 3.

With the aforementioned discussions, for any $x = (x_1, \ldots, x_r) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r}$, we are ready to show how to construct a smoothing function for vectors $[x]_+$ and $[x]_-$ associated with $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_r}$. We start by constructing a smoothing function for vectors $[x_i]_+$, $[x_i]_-$ on a single SOC $\mathcal{K}^{n_i}$, $i = 1, \ldots, r$ since $[x]_+$ and $[x]_-$ are shown as (7). First, given smoothing functions ϕ^+, ϕ^- in (9),(10) and $x_i \in \mathbb{R}^{n_i}$, $i = 1, \ldots, r$, we define vector-valued function $\Phi_i^+, \Phi_i^- : \mathbb{R}_{++} \times \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$, $i = 1, \ldots, r$ as

$$\Phi_i^+(\mu, x_i) := \phi^+(\mu, \lambda_1(x_i))\, u_{x_i}^{(1)} + \phi^+(\mu, \lambda_2(x_i))\, u_{x_i}^{(2)}, \tag{17}$$

$$\Phi_i^-(\mu, x_i) := \phi^-(\mu, \lambda_1(x_i))\, u_{x_i}^{(1)} + \phi^-(\mu, \lambda_2(x_i))\, u_{x_i}^{(2)}, \tag{18}$$

where $\mu \in \mathbb{R}_{++}$ is a parameter, $\lambda_1(x_i)$, $\lambda_2(x_i)$ are the spectral values, and $u_{x_i}^{(1)}$, $u_{x_i}^{(2)}$ are the spectral vectors of x_i.

Consequently, $\Phi_i^+(\mu, x_i)$, $\Phi_i^-(\mu, x_i)$ are also smooth on $\mathbb{R}_{++} \times \mathbb{R}^{n_i}$ [8]. Moreover, it is easy to assert that

$$\lim_{\mu \to 0^+} \Phi_i^+(\mu, x_i) = [\lambda_1(x_i)]_+ u_{x_i}^{(1)} + [\lambda_2(x_i)]_+ u_{x_i}^{(2)} = [x_i]_+, \tag{19}$$

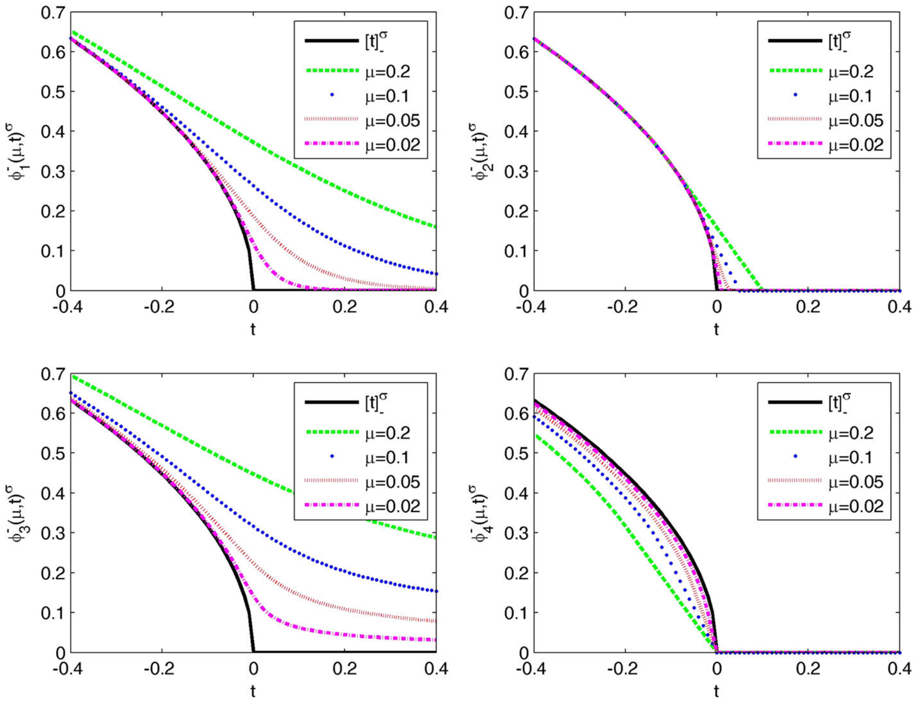

Fig. 3 Graphs of $\phi_i^-(\mu, t)^\sigma$, $i = 1, 2, 3, 4$ with different μ and $\sigma = 1/2$

$$\lim_{\mu\to 0^+} \Phi_i^-(\mu, x_i) = [\lambda_1(x_i)]_- u_{x_i}^{(1)} + [\lambda_2(x_i)]_- u_{x_i}^{(2)} = [x_i]_-, \tag{20}$$

which means each function $\Phi_i^+(\mu, x_i)$, $\Phi_i^-(\mu, x_i)$ serves as a smoothing function of $[x_i]_+$, $[x_i]_-$ associated with single SOC $\mathcal{K}^{n_i}$, $i = 1, \ldots, r$, respectively. Due to Lemma 3.1, Remark 3.1 and from definition of $\Phi_i^+(\mu, x_i)$, $\Phi_i^-(\mu, x_i)$ in (17), (18), it is not difficult to verify that for any $\sigma > 0$, the below two functions

$$\Phi_i^+(\mu, x_i)^\sigma := \phi^+(\mu, \lambda_1(x_i))^\sigma\, u_{x_i}^{(1)} + \phi^+(\mu, \lambda_2(x_i))^\sigma\, u_{x_i}^{(2)}, \tag{21}$$

$$\Phi_i^-(\mu, x_i)^\sigma := \phi^-(\mu, \lambda_1(x_i))^\sigma\, u_{x_i}^{(1)} + \phi^-(\mu, \lambda_2(x_i))^\sigma\, u_{x_i}^{(2)} \tag{22}$$

are continuous functions approximate to $[x_i]_+^\sigma$ and $[x_i]_-^\sigma$, respectively. In other words,

$$\lim_{\mu\to 0^+} \Phi_i^+(\mu, x_i)^\sigma = [\lambda_1(x_i)]_+^\sigma u_{x_i}^{(1)} + [\lambda_2(x_i)]_+^\sigma u_{x_i}^{(2)} = [x_i]_+^\sigma,$$

$$\lim_{\mu\to 0^+} \Phi_i^-(\mu, x_i)^\sigma = [\lambda_1(x_i)]_-^\sigma u_{x_i}^{(1)} + [\lambda_2(x_i)]_-^\sigma u_{x_i}^{(2)} = [x_i]_-^\sigma.$$

Now we construct smoothing function for vectors $[x]_+$ and $[x]_-$ associated with general cone (2). To this end, we define vector-valued function $\Phi^+, \Phi^- : \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}^n$ as

$$\Phi^+(\mu, x) := \left(\Phi_1^+(\mu, x_1), \ldots, \Phi_r^+(\mu, x_r)\right), \tag{23}$$

$$\Phi^-(\mu, x) := \left(\Phi_1^-(\mu, x_1), \ldots, \Phi_r^-(\mu, x_r)\right), \tag{24}$$

where $\Phi_i^+(\mu, x_i)$, $\Phi_i^-(\mu, x_i)$, $i = 1, \ldots, r$ are defined by (17), (18), respectively. Therefore, from (19), (20) and (7), $\Phi^+(\mu, x)$, $\Phi^-(\mu, x)$ serves as a smoothing function for $[x]_+$, $[x]_-$

associated with $\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_r}$, respectively. At the same time, from (21), (22),

$$\Phi^+(\mu, x)^\sigma := \left(\Phi_1^+(\mu, x_1)^\sigma, \ldots, \Phi_r^+(\mu, x_r)^\sigma\right), \tag{25}$$

$$\Phi^-(\mu, x)^\sigma := \left(\Phi_1^-(\mu, x_1)^\sigma, \ldots, \Phi_r^-(\mu, x_r)^\sigma\right) \tag{26}$$

are continuous functions approximate to $[x]_+^\sigma$ and $[x]_-^\sigma$, respectively.

In light of this idea, we establish an approximating lower order penalty equations for solving SOCLCP (1), which will be described in next section. To end this section, we present a technical lemma for subsequent needs.

Lemma 3.2 *Suppose that $\Phi^+(\mu, x)$ and $\Phi^-(\mu, x)$ are defined by (23), (24), respectively, and $\Phi^+(\mu, x)^\sigma$ and $\Phi^-(\mu, x)^\sigma$ are defined for any $\sigma > 0$ as in (25), (26), respectively. Then, the following results hold.*

(a) *Both $\Phi^+(\mu, x)$ and $\Phi^-(\mu, x)$ belong to $\mathcal{K}$,*
(b) *Both $\Phi^+(\mu, x)^\sigma$ and $\Phi^-(\mu, x)^\sigma$ belong to $\mathcal{K}$.*

Proof (a) For any $x_i \in \mathbb{R}^{n_i}$, $i = 1, \ldots, r$, since $\phi^+(\mu, \lambda_k(x_i)) \geq 0$, $\phi^-(\mu, \lambda_k(x_i)) \geq 0$ for $k = 1, 2$ from (9), (10), we have $\Phi_i^+(\mu, x_i)$, $\Phi_i^-(\mu, x_i) \in \mathcal{K}^{n_i}$ according to the definition (17), (18). Therefore, the conclusion holds due to the definitions (23), (24) and (2).
(b) From part (a) and knowing $\sigma > 0$, we have $\phi^+(\mu, \lambda_k(x_i))^\sigma \geq 0$, $\phi^-(\mu, \lambda_k(x_i))^\sigma \geq 0$, $k = 1, 2$. Applying (25) and (26), the desired result follows. □

4 Approximate lower order penalty approach and convergence analysis

In this section, we propose an approximate lower order penalty approach for solving SOCLCP (1). To this end, we consider the approximate lower order penalty equations (LOPEs):

$$Ax - \alpha\Phi^-(\mu, x)^\sigma = b, \tag{27}$$

where $\sigma \in (0, 1]$ is a given power parameter, $\alpha \geq 1$ is a penalty parameter and $\Phi^-(\mu, x)^\sigma$ is defined as (26). Throughout this section, $x_{\mu,\alpha}$ means the solution of (27), and corresponding to the structure of (2), we denote

$$x_{\mu,\alpha} = \left((x_{\mu,\alpha})_1, \ldots, (x_{\mu,\alpha})_r\right) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r}. \tag{28}$$

For simplicity and without causing confusion, we always denote the spectral values and spectral vectors of $(x_{\mu,\alpha})_i$, $i = 1, \ldots, r$ as $\lambda_k := \lambda_k((x_{\mu,\alpha})_i)$, $u^{(k)} := u^{(k)}_{(x_{\mu,\alpha})_i}$ for $k = 1, 2$. Accordingly, $[\lambda_k]_- := [\lambda_k((x_{\mu,\alpha})_i)]_-$ and $\phi^-(\mu, \lambda_k) := \phi^-(\mu, \lambda_k((x_{\mu,\alpha})_i))$, $k = 1, 2$ for instance. Note that for special case $\sigma = 1$, the nonlinear function in (27) is always smooth.

Note that the equations (27) are penalized equations corresponding to the SOCLCP (1) because the penalty term $\alpha\Phi^-(\mu, x)^\sigma$ penalizes the 'negative part' of x when $\mu \to 0^+$. By Lemma 3.2 and from equations (27), it is easy to see that $Ax_{\mu,\alpha} - b \in \mathcal{K}$ (noting $\alpha\Phi^-(\mu, x_{\mu,\alpha})^\sigma \in \mathcal{K}$). Our goal is to show that the solution sequence $\{x_{\mu,\alpha}\}$ converges to the solution of SOCLCP (1) when $\alpha \to +\infty$ and $\mu \to 0^+$. In order to achieve this, we need to make the assumption for matrix A as below.

Assumption 4.1 The matrix A is positive definite, but not necessarily symmetric, i.e., there exists a constant $a_0 > 0$, such that

$$y^T Ay \geq a_0\|y\|^2, \ \forall y \in \mathbb{R}^n. \tag{29}$$

This assumption just implies that $\bar{A} = (A + A^T)/2$ is symmetric positive definite with $a_0 = \lambda_{\min}(\bar{A})$ since $y^T Ay = y^T \bar{A} y$. Here is an example of A. Let

$$A = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix},$$

it is easy to see that matrix A is positive definite satisfying (29), but not symmetric. Under Assumption 4.1, the SOCLCP (1) has a unique solution and the LOPEs (27) also has a unique solution, see for more details in [17,21].

Proposition 4.1 *For any $\alpha \geq 1$, $\sigma \in (0, 1]$ and sufficiently small μ, the solution of the LOPEs (27) is bounded, i.e., there exists a positive constant M, independent of $x_{\mu,\alpha}$, μ, α and σ, such that $\|x_{\mu,\alpha}\| \leq M$.*

Proof By multiplying $x_{\mu,\alpha}$ on both sides of (27), we observe that

$$x_{\mu,\alpha}^T Ax_{\mu,\alpha} = x_{\mu,\alpha}^T b + \alpha x_{\mu,\alpha}^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma = \sum_{i=1}^r \left((x_{\mu,\alpha})_i^T b_i + \alpha (x_{\mu,\alpha})_i^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma\right) \quad (30)$$

by (26),(28) and denoting $b = (b_1, \ldots, b_r) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r}$. For any $(x_{\mu,\alpha})_i$, $i = 1, \ldots, r$, to proceed, we consider three cases to evaluate the term

$$\Xi_i := (x_{\mu,\alpha})_i^T b_i + \alpha (x_{\mu,\alpha})_i^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma \leq \|x_{\mu,\alpha}\| \left(\|b\| + 1\right). \quad (31)$$

Case 1: $(x_{\mu,\alpha})_i \in \mathcal{K}^{n_i}$. From Cauchy-Schwarz inequality, spectral decomposition of $(x_{\mu,\alpha})_i$, and the fact that the norm of the piece component is less than that of the whole vector, we have

$$\begin{aligned} \Xi_i &\leq \|(x_{\mu,\alpha})_i\| \left(\|b_i\| + \alpha \|\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma\|\right) \\ &\leq \|x_{\mu,\alpha}\| \left(\|b\| + \alpha \|\phi^-(\mu, \lambda_1)^\sigma u^{(1)} + \phi^-(\mu, \lambda_2)^\sigma u^{(2)}\|\right) \\ &\leq \|x_{\mu,\alpha}\| \left(\|b\| + \sqrt{2}\alpha \phi^-(\mu, 0)^\sigma\right), \end{aligned} \quad (32)$$

where the second inequality holds by the definition of $\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma$ as in (22), and the last inequality holds by the triangle inequality, the nonnegativity of $\phi^-(\mu, 0)^\sigma$ from (10) and the monotone decreasing of $\phi^-(\mu, t)$ about t since $0 \leq \lambda_1 \leq \lambda_2$ in this case. Now, applying Lemma 3.1, we have $\lim_{\mu \to 0^+} \phi^-(\mu, 0)^\sigma = 0$. This means, for any penalty parameter α, there exists a positive real number ν, such that $\sqrt{2}\alpha\phi^-(\mu, 0)^\sigma \leq 1$ for all $\mu \in (0, \nu]$. Therefore, from (32), we obtain the conclusion (31).

Case 2: $(x_{\mu,\alpha})_i \in -\mathcal{K}^{n_i}$. In light of Lemma 3.2, we know $\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma \in \mathcal{K}^{n_i}$, and hence

$$(x_{\mu,\alpha})_i^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma \leq 0.$$

Thus, we have $\Xi_i \leq (x_{\mu,\alpha})_i^T b_i \leq \|(x_{\mu,\alpha})_i\| \|b_i\| \leq \|x_{\mu,\alpha}\| (\|b\| + 1)$, which says conclusion (31) holds.

Case 3: $(x_{\mu,\alpha})_i \notin \mathcal{K}^{n_i} \cup -\mathcal{K}^{n_i}$. In this case, we know that $\lambda_1 < 0 < \lambda_2$ and $[(x_{\mu,\alpha})_i]_+ = \lambda_2 u^{(2)}$. From the definition of $\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma$ as in (22), Proposition 2.1, we have

$$\begin{aligned} (x_{\mu,\alpha})_i^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma &= (\lambda_1 u^{(1)} + \lambda_2 u^{(2)})^T \left(\phi^-(\mu, \lambda_1)^\sigma u^{(1)} + \phi^-(\mu, \lambda_2)^\sigma u^{(2)}\right). \\ &= \tfrac{1}{2}\left(\lambda_1 \phi^-(\mu, \lambda_1)^\sigma + \lambda_2 \phi^-(\mu, \lambda_2)^\sigma\right) \\ &\leq \tfrac{\sqrt{2}}{2}(\tfrac{\sqrt{2}}{2}\lambda_2)\phi^-(\mu, \lambda_2)^\sigma \\ &\leq \tfrac{\sqrt{2}}{2}\|x_{\mu,\alpha}\|\phi^-(\mu, \lambda_2)^\sigma, \end{aligned} \quad (33)$$

where the first inequality holds due to $\lambda_1\phi^-(\mu,\lambda_1)^\sigma < 0 < \lambda_2\phi^-(\mu,\lambda_2)^\sigma$, and the second inequality holds due to $\frac{\sqrt{2}}{2}\lambda_2 = \|[(x_{\mu,\alpha})_i]_+\| \le \|(x_{\mu,\alpha})_i\| \le \|x_{\mu,\alpha}\|$. Substituting (33) from Ξ_i and using Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\Xi_i &\le \|(x_{\mu,\alpha})_i\|\|b_i\| + \tfrac{\sqrt{2}}{2}\alpha\|x_{\mu,\alpha}\|\phi^-(\mu,\lambda_2)^\sigma\\ &\le \|x_{\mu,\alpha}\|\|b\| + \tfrac{\sqrt{2}}{2}\alpha\|x_{\mu,\alpha}\|\phi^-(\mu,\lambda_2)^\sigma\\ &\le \|x_{\mu,\alpha}\|\left(\|b\| + \tfrac{\sqrt{2}}{2}\alpha\phi^-(\mu,0)^\sigma\right),\end{aligned} \tag{34}$$

where the third inequality holds by the monotone decreasing of $\phi^-(\mu,t)$ about t. Similar to case 1, for any penalty parameter α, there exists a positive real number ν, such that $\frac{\sqrt{2}}{2}\alpha\phi^-(\mu,0)^\sigma \le 1$ for all $\mu \in (0,\nu]$. Hence, we reach the conclusion (31) by (34).

From above three cases, the conclusion (31) holds, which shows an evaluation of Ξ_i. Thus, from (30) and Assumption 4.1, there exists a constants $a_0 > 0$ such that

$$a_0\|x_{\mu,\alpha}\|^2 \le x_{\mu,\alpha}^T A x_{\mu,\alpha} = \sum_{i=1}^r \Xi_i \le r\|x_{\mu,\alpha}\|\left(\|b\|+1\right).$$

This implies $\|x_{\mu,\alpha}\|\cdot\big(a_0\|x_{\mu,\alpha}\| - r(\|b\|+1)\big) \le 0$, and hence $\|x_{\mu,\alpha}\| \le \frac{r}{a_0}(\|b\|+1)$. By taking $M = \frac{r}{a_0}(\|b\|+1)$, the proof is completed. □

It is well-known that the affine function $g(x) := Ax - b$ is continuous function and by Proposition 4.1, $\|g(x_{\mu,\alpha})\|$ is bounded for any $\alpha \ge 1$, $\sigma \in (0,1]$ and sufficiently small μ. We are able to establish an upper bound for $\|\Phi^-(\mu, x_{\mu,\alpha})\|$ in next proposition. The upper bound is also applicable for $\|[x_{\mu,\alpha}]_-\|$ (see Remark 4.1), which plays an important role in the convergence analysis. The detailed proof is based on the definition of $\Phi_i^-(\mu,(x_{\mu,\alpha})_i)$ stated as in (18) and uses the same techniques as in [21, Proposition 3.2] by left multiplying $\Phi_i^-(\mu,(x_{\mu,\alpha})_i)$ on both sides of the ith block of (27):

$$(Ax_{\mu,\alpha})_i - \alpha\Phi_i^-(\mu,(x_{\mu,\alpha})_i)^\sigma = b_i.$$

Therefore, we omit it and only present the result, i.e., there exists a positive constant C_i, independent of $x_{\mu,\alpha}$, μ and α, such that

$$\|\Phi_i^-(\mu,(x_{\mu,\alpha})_i)\| \le \frac{C_i}{\alpha^{1/\sigma}} \tag{35}$$

holds for any $\alpha \ge 1$, $\sigma \in (0,1]$ and sufficiently small μ. By the definition of $\Phi^-(\mu, x_{\mu,\alpha})$ as shown in (24) and setting $C = C_1 + \cdots + C_r$, we obtain the following proposition.

Proposition 4.2 *For any $\alpha \ge 1$, $\sigma \in (0,1]$ and sufficiently small μ, there exists a positive constant C, independent of $x_{\mu,\alpha}$, μ and α, such that*

$$\|\Phi^-(\mu, x_{\mu,\alpha})\| \le \frac{C}{\alpha^{1/\sigma}}. \tag{36}$$

Remark 4.1 For any $\alpha \ge 1$, $\sigma \in (0,1]$ and sufficiently small μ, the ith ($i = 1,\ldots,r$) component vector $(x_{\mu,\alpha})_i$ is fixed since $x_{\mu,\alpha}$ with (28) means the solution of (27). For the fixed $(x_{\mu,\alpha})_i$ with spectral decomposition $(x_{\mu,\alpha})_i = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}$ and the expression $\Phi_i^-(\mu,(x_{\mu,\alpha})_i) = \phi^-(\mu,\lambda_1)u^{(1)} + \phi^-(\mu,\lambda_2)u^{(2)}$, by taking $\mu \to 0^+$ in $\phi^-(\mu,\lambda_1)$ and $\phi^-(\mu,\lambda_2)$, we obtain $\|[\lambda_1]_- u^{(1)} + [\lambda_2]_- u^{(2)}\| \le \frac{C_i}{\alpha^{1/\sigma}}$ from (35),
which yields

$$\left\|[(x_{\mu,\alpha})_i]_-\right\| \le \frac{C_i}{\alpha^{1/\sigma}}. \tag{37}$$

Also, by setting $C = C_1 + \cdots + C_r$, we obtain

$$\left\| [(x_{\mu,\alpha})]_- \right\| \le \frac{C}{\alpha^{1/\sigma}}. \tag{38}$$

By using Propositions 4.1, 4.2 and Remark 4.1, we are able to obtain the following desired convergence result of SOCLCP (1) is approximated by the LOPEs (27).

Theorem 4.1 *For any $\alpha \ge 1$, $\sigma \in (0, 1]$ and sufficiently small μ, let $x_{\mu,\alpha}$ be the solution of LOPEs (27), and x^* be the solution of SOCLCP (1). Then, there exists a positive constant C, independent of x^*, $x_{\mu,\alpha}$, μ and α, such that*

$$\|x^* - x_{\mu,\alpha}\| \le \frac{C}{\alpha^{1/\sigma}}. \tag{39}$$

Proof Follows from (28) and the definition (7), we get $x_{\mu,\alpha} = [x_{\mu,\alpha}]_+ - [x_{\mu,\alpha}]_-$, where

$$[x_{\mu,\alpha}]_+ = ([(x_{\mu,\alpha})_1]_+, \ldots, [(x_{\mu,\alpha})_r]_+), \quad [x_{\mu,\alpha}]_- = ([(x_{\mu,\alpha})_1]_-, \ldots, [(x_{\mu,\alpha})_r]_-)$$

respectively denotes the projection of $x_{\mu,\alpha}$ on $\mathcal{K}$ and $-x_{\mu,\alpha}$ on $\mathcal{K}^*$. Therefore, the vector $x^* - x_{\mu,\alpha}$ can be decomposed as

$$x^* - x_{\mu,\alpha} = x^* - [x_{\mu,\alpha}]_+ + [x_{\mu,\alpha}]_- = r_{\mu,\alpha} + [x_{\mu,\alpha}]_-, \tag{40}$$

where

$$r_{\mu,\alpha} = x^* - [x_{\mu,\alpha}]_+. \tag{41}$$

Let's consider the estimation of $r_{\mu,\alpha}$. If $r_{\mu,\alpha} = 0$, the inequality (39) is satisfied due to (38) and (40). Therefore, in the following, we only consider $r_{\mu,\alpha} \ne 0$. Noting that, the SOCLCP (1) is equivalent to the variational inequality problem: find $x^* \in \mathcal{K}$ (see [17, Proposition 1.1.3]), such that

$$(y - x^*)^T A x^* \ge (y - x^*)^T b, \ \forall y \in \mathcal{K}. \tag{42}$$

It is known that $[x_{\mu,\alpha}]_+ \in \mathcal{K}$, by (41) and substituting $[x_{\mu,\alpha}]_+$ for y in (42) yields

$$- r_{\mu,\alpha}^T A x^* \ge -r_{\mu,\alpha}^T b. \tag{43}$$

Then, multiplying both sides of (27) by $r_{\mu,\alpha}$ yields

$$r_{\mu,\alpha}^T A x_{\mu,\alpha} - \alpha r_{\mu,\alpha}^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma = r_{\mu,\alpha}^T b. \tag{44}$$

Adding up (43) and (44) leads to

$$r_{\mu,\alpha}^T A(x_{\mu,\alpha} - x^*) - \alpha r_{\mu,\alpha}^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma \ge 0. \tag{45}$$

Applying (41) again, we have

$$r_{\mu,\alpha}^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma = (x^* - [x_{\mu,\alpha}]_+)^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma. \tag{46}$$

Combining (45) and (46), we achieve $r_{\mu,\alpha}^T A(x_{\mu,\alpha} - x^*) \ge \alpha (x^* - [x_{\mu,\alpha}]_+)^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma$, which says

$$r_{\mu,\alpha}^T A(x^* - x_{\mu,\alpha}) \le \alpha([x_{\mu,\alpha}]_+ - x^*)^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma. \tag{47}$$

Now, using (40) and (47) further gives

$$(x^* - x_{\mu,\alpha} - [x_{\mu,\alpha}]_-)^T A(x^* - x_{\mu,\alpha}) \le \alpha([x_{\mu,\alpha}]_+ - x^*)^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma,$$

which implies

$$(x^*-x_{\mu,\alpha})^T A(x^*-x_{\mu,\alpha}) \le [x_{\mu,\alpha}]_-^T A(x^*-x_{\mu,\alpha}) + \alpha([x_{\mu,\alpha}]_+ - x^*)^T \Phi^-(\mu, x_{\mu,\alpha})^\sigma. \tag{48}$$

Follows from (26),(28) and the definition (7), by denoting

$$\Xi_i := [(x_{\mu,\alpha})_i]_-^T \left(A(x^*-x_{\mu,\alpha})\right)_i + \alpha\left([(x_{\mu,\alpha})_i]_+ - x_i^*\right)^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma, \tag{49}$$

the inequality (48) is reduced to

$$(x^*-x_{\mu,\alpha})^T A(x^*-x_{\mu,\alpha}) \le \sum_{i=1}^r \Xi_i. \tag{50}$$

To proceed, we discuss three cases of $(x_{\mu,\alpha})_i$ to proof the term (49) satisfying

$$\Xi_i \le \frac{1}{\alpha^{1/\sigma}}\left(C'\|x^*-x_{\mu,\alpha}\| + c\right), \tag{51}$$

where C' is a positive constant, independent of $x_{\mu,\alpha}$, μ, α and $c \in \mathbb{R}_{++}$ is undetermined.
Case 1: $(x_{\mu,\alpha})_i \in \mathcal{K}^{n_i}$. Under this case, we see that $[(x_{\mu,\alpha})_i]_- = 0$ and $[(x_{\mu,\alpha})_i]_+ = (x_{\mu,\alpha})_i$. Using (49) and Cauchy-Schwarz inequality, we have

$$\begin{aligned}\Xi_i &= \alpha\left((x_{\mu,\alpha})_i - x_i^*\right)^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma \\ &\le \alpha\|(x_{\mu,\alpha})_i - x_i^*\| \cdot \|\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma\| \\ &\le \alpha\|x_{\mu,\alpha} - x^*\| \cdot \|\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma\| \\ &= \alpha\|x^* - x_{\mu,\alpha}\|\|\phi^-(\mu,\lambda_1)^\sigma u^{(1)} + \phi^-(\mu,\lambda_2)^\sigma u^{(2)}\| \\ &\le \|x^*-x_{\mu,\alpha}\|\sqrt{2}\alpha\phi^-(\mu,0)^\sigma,\end{aligned} \tag{52}$$

where the second inequality holds by the fact that the norm of the piece component is less than that of the whole vector, the second equality holds by the definition as (18) and the last inequality holds by Proposition 2.1, the triangle inequality, the nonnegativity of $\phi^-(\mu,0)^\sigma$ from (10) and the monotone decreasing of $\phi^-(\mu,t)$ about t since $0 \le \lambda_1 \le \lambda_2$ in this case. By Lemma 3.1, we know $\lim_{\mu\to 0^+}\phi^-(\mu,0)^\sigma = 0$. Therefore, for any $\alpha \ge 1$ and $\sigma \in (0,1]$, there exists a positive real number ν, such that $\sqrt{2}\alpha\phi^-(\mu,0)^\sigma \le \frac{1}{\alpha^{1/\sigma}}$ for all $\mu \in (0,\nu]$. Thus, we achieve the conclusion (51) by setting $C' = 1$.
Case 2: $(x_{\mu,\alpha})_i \in -\mathcal{K}^{n_i}$. Under this case, it is clear that $[(x_{\mu,\alpha})_i]_+ = 0$, and we have $(x_i^*)^T\Phi_i^-(\mu,(x_{\mu,\alpha})_i)^\sigma \ge 0$ since $\Phi_i^-(\mu,(x_{\mu,\alpha})_i)^\sigma \in \mathcal{K}^{n_i}$ and $x_i^* \in \mathcal{K}^{n_i}$. Thus, it follows from (49) and Cauchy-Schwarz inequality that

$$\begin{aligned}\Xi_i &\le [(x_{\mu,\alpha})_i]_-^T\left(A(x^*-x_{\mu,\alpha})\right)_i \\ &\le \|[(x_{\mu,\alpha})_i]_-\| \cdot \|\left(A(x^*-x_{\mu,\alpha})\right)_i\| \\ &\le \|[(x_{\mu,\alpha})_i]_-\| \cdot \|A(x^*-x_{\mu,\alpha})\| \\ &\le \tfrac{C_i}{\alpha^{1/\sigma}}\|A\|\|x^*-x_{\mu,\alpha}\|,\end{aligned} \tag{53}$$

where the last inequality holds by (37) and norm compatibility. Thus, we also achieve the conclusion (51) by setting $C' = C_i\|A\|$.
Case 3: $(x_{\mu,\alpha})_i \notin \mathcal{K}^{n_i} \cup -\mathcal{K}^{n_i}$. Under this case, $\lambda_1 < 0 < \lambda_2$ and $[(x_{\mu,\alpha})_i]_+ = \lambda_2 u^{(2)}$. Because $\Phi_i^-(\mu,(x_{\mu,\alpha})_i)^\sigma \in \mathcal{K}^{n_i}$ and $x_i^* \in \mathcal{K}^{n_i}$, we have $(x_i^*)^T\Phi_i^-(\mu,(x_{\mu,\alpha})_i)^\sigma \ge 0$ and

$$\left([(x_{\mu,\alpha})_i]_+ - x_i^*\right)^T \Phi_i^-(\mu,(x_{\mu,\alpha})_i)^\sigma \le [(x_{\mu,\alpha})_i]_+^T \Phi_i^-(\mu,(x_{\mu,\alpha})_i)^\sigma.$$

Furthermore, from the definition of $\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma$ stated as in (18), Proposition 2.1 and Proposition 4.1, we also have

$$\begin{aligned}[(x_{\mu,\alpha})_i]_+^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma &= (\lambda_2 u^{(2)})^T \left(\phi^-(\mu, \lambda_1)^\sigma u^{(1)} + \phi^-(\mu, \lambda_2)^\sigma u^{(2)}\right) \\ &= \tfrac{1}{2}\lambda_2 \phi^-(\mu, \lambda_2)^\sigma \le M\phi^-(\mu, 0)^\sigma\end{aligned}$$

by the monotone decreasing of $\phi^-(\mu, t)$ about t since $\lambda_2 > 0$ in this case.

According to $\lim_{\mu\to 0^+} \phi^-(\mu, 0)^\sigma = 0$, there exists a positive real number ν, such that

$$\alpha[(x_{\mu,\alpha})_i]_+^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma \le \alpha M\phi^-(\mu, 0)^\sigma \le \frac{c}{\alpha^{1/\sigma}} \tag{54}$$

for all $\mu \in (0, \nu]$, where $c \in \mathbb{R}_{++}$ is undetermined. It follows from (49),(53) and (54) that

$$\begin{aligned}\Xi_i &\le [(x_{\mu,\alpha})_i]_-^T \left(A(x^* - x_{\mu,\alpha})\right)_i + \alpha[(x_{\mu,\alpha})_i]_+^T \Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma \\ &\le \tfrac{C_i}{\alpha^{1/\sigma}} \|A\|\|x^* - x_{\mu,\alpha}\| + \tfrac{c}{\alpha^{1/\sigma}} = \tfrac{1}{\alpha^{1/\sigma}}(C_i\|A\|\|x^* - x_{\mu,\alpha}\| + c).\end{aligned}$$

By setting $C' = C_i\|A\|$, we also achieve the conclusion (51).

From above three cases, we all conclude the conclusion (51), together with (50) and Assumption 4.1, there exists a constants $a_0 > 0$, such that

$$a_0\|x^* - x_{\mu,\alpha}\|^2 \le (x^* - x_{\mu,\alpha})^T A(x^* - x_{\mu,\alpha}) \le \sum_{i=1}^r \Xi_i \le \frac{r}{\alpha^{1/\sigma}} \left(C'\|x^* - x_{\mu,\alpha}\| + c\right).$$

Thus, we obtain a quadratic inequality with respect to $\|x^* - x_{\mu,\alpha}\|$, i.e.,

$$a_0\|x^* - x_{\mu,\alpha}\|^2 - \frac{rC'}{\alpha^{1/\sigma}}\|x^* - x_{\mu,\alpha}\| - \frac{rc}{\alpha^{1/\sigma}} \le 0.$$

According to the solution formula of quadratic inequality, we know

$$\|x^* - x_{\mu,\alpha}\| \le \frac{1}{2a_0}\left(\frac{rC'}{\alpha^{1/\sigma}} + \sqrt{\left(\frac{rC'}{\alpha^{1/\sigma}}\right)^2 + 4a_0\frac{rc}{\alpha^{1/\sigma}}}\right).$$

Especially, by taking $c = \frac{r(C')^2}{4a_0\alpha^{1/\sigma}}$, yields $\|x^* - x_{\mu,\alpha}\| \le \frac{1+\sqrt{2}}{2a_0}\frac{rC'}{\alpha^{1/\sigma}}$, which is the conclusion (39) by setting $C = \frac{1+\sqrt{2}}{2a_0}rC'$. The proof is complete. □

In light of Theorem 4.1, to deal with the SOCLCP (1) with any $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$ satisfying Assumption 4.1, we only solve the LOPEs (27) as $\alpha \to +\infty$ and $\mu \to 0^+$. However, in order to implementing the algorithm, it is necessary to balance $\alpha \to +\infty$ and $\mu \to 0^+$. The key part is on how to evaluate the sufficiently small μ in Theorem 4.1, in which there exists a positive real number ν such that the conclusion holds for all $\mu \in (0, \nu]$. The exact value of ν is unknown, on the other hand, it is not necessary to know the exact value of ν because it is only a bound of μ. Therefore, in the implementations, it would be better to provide a simple criterion to evaluate whether μ is sufficiently small. To this end, we discuss from the system of linear equations $Ax - b = 0$, which has a unique solution $A^{-1}b$ according to Assumption 4.1. If $A^{-1}b \in \mathcal{K}$, then of course this solution is the solution to SOCLCP (1). Hence, we only need to focus on the case where the system $Ax - b = 0$ has no solution in $\mathcal{K}$.

Proposition 4.3 *Consider the SOCLCP (1) with $Ax - b = 0$ having no solution in $\mathcal{K}$, and the LOPEs (27) with $\alpha \ge 1$, $\sigma \in (0, 1]$, $\mu \in (0, 1)$. If parameter μ is sufficiently small, then the numerical solution $x_{\mu,\alpha} \notin \mathcal{K}$.*

Proof For the SOCLCP (1) with $Ax - b = 0$ having no solution in $\mathcal{K}$, there is at least one $i \in \{1, \ldots, r\}$, such that $(Ax - b)_i \neq 0$

for any $x \in \mathcal{K}$. Therefore, we obtain $\|Ax - b\| > 0$ for all $x \in \mathcal{K}$ and there exists a positive number δ, such that

$$\min_{x \in \mathcal{K}} \|Ax - b\| = \delta > 0 \tag{55}$$

since $\mathcal{K}$ is a closed convex set and $g(x) = Ax - b$ is a continuous affine function. For LOPEs (27) with $\alpha \geq 1$, $\sigma \in (0, 1]$ and sufficiently small parameter μ, we will claim that the numerical solution must satisfy $x_{\mu,\alpha} \notin \mathcal{K}$. Suppose not, i.e., $x_{\mu,\alpha} \in \mathcal{K}$, we have

$$Ax_{\mu,\alpha} - b = \alpha \Phi^-(\mu, x_{\mu,\alpha})^\sigma = \alpha \left(\Phi_1^-(\mu, (x_{\mu,\alpha})_1)^\sigma, \ldots, \Phi_r^-(\mu, (x_{\mu,\alpha})_r)^\sigma \right)$$

from LOPEs (27). For any $\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma$, using Proposition 2.1, triangle inequality, the nonnegativity of $\phi^-(\mu, 0)^\sigma$ and the monotone descending of $\phi^-(\mu, t)$, we have

$$\|\Phi_i^-(\mu, (x_{\mu,\alpha})_i)^\sigma\| = \|\phi^-(\mu, \lambda_1)^\sigma u^{(1)} + \phi^-(\mu, \lambda_2)^\sigma u^{(2)}\| \leq \sqrt{2}\phi^-(\mu, 0)^\sigma$$

since $0 \leq \lambda_1 \leq \lambda_2$ from $(x_{\mu,\alpha})_i \in \mathcal{K}^{n_i}$. Accordingly, this gives

$$\|Ax_{\mu,\alpha} - b\| = \alpha \|\Phi^-(\mu, x_{\mu,\alpha})^\sigma\| \leq \alpha\sqrt{2r}\phi^-(\mu, 0)^\sigma.$$

For any $\alpha \geq 1$, $\sigma \in (0, 1]$ and positive number δ in (55), due to $\lim_{\mu \to 0^+} \phi^-(\mu, 0)^\sigma = 0$ from Lemma 3.1, there exists a positive real number ν, such that $\alpha\sqrt{2r}\phi^-(\mu, 0)^\sigma < \delta/2$ for all $\mu \in (0, \nu]$. In other words, $x_{\mu,\alpha} \in \mathcal{K}$, but $\|Ax_{\mu,\alpha} - b\| \leq \delta/2$, which contradicts the formula (55). This completes the proof. □

As a result, from Proposition 4.3, in practical implementation, an approximate simple criterion is developed to estimate the parameter μ as following, which will be applied to Algorithm 5.1 in next section.

Remark 4.2 For the SOCLCP (1) with $Ax - b = 0$ having no solution in $\mathcal{K}$, and the LOPEs (27) with $\alpha \geq 1, \sigma \in (0, 1], \mu \in (0, 1)$. If $x_{\mu,\alpha} \notin \mathcal{K}$, we regard parameter μ as sufficiently small; otherwise, μ is not quite small yet, and we should take μ much smaller.

From LOPEs (27) and Lemma 3.2, we always have $Ax_{\mu,\alpha} - b = \alpha\Phi^-(\mu, x_{\mu,\alpha})^\sigma \in \mathcal{K}$. Therefore, for the SOCLCP (3) with single SOC $\mathcal{K}^n$ and $A^{-1}b \notin \mathcal{K}^n$, if the exact solution x^* and $Ax^* - b$ are in different boundary of $\mathcal{K}^n$ without origin, a more concise criterion will be obtained, which is to treat the parameters as sufficiently small if $x_{\mu,\alpha}^T(Ax_{\mu,\alpha} - b) \leq 0$ since $x_{\mu,\alpha}$ and $Ax_{\mu,\alpha} - b$ approximate x^* and $Ax^* - b$, respectively.

5 Algorithm and numerical experiments

In this section, an algorithm is constructed first for solving the SOCLCP (1), and some numerical experiments are implemented. Then, we present the performance profiles of four kinds specific smoothing approximations in (11)–(14). Finally, we consider the generalization of two approximations (12) and (14), the performance profiles of the generalization are also reported.

5.1 Algorithmic construction and numerical experiments

According to Theorem 4.1 and the criterion in Proposition 4.3, Remark 4.2, we give the algorithm as following.

Algorithm 5.1

Step 0 Given a vector $b \in \mathbb{R}^n$, and a matrix $A \in \mathbb{R}^{n \times n}$ satisfying Assumption 4.1

Step 1 If $A^{-1}b \in \mathcal{K}$, set $\tilde{x} = A^{-1}b$, go to step 5; else, go to step 2.

Step 2 Given the power parameter $\sigma \in (0, 1]$, the penalty parameter $\alpha \geq 1$, the smoothing parameter $0 < \mu < 1$, the error bound `eps` and the multiple parameter $c_1 > 1$ and $c_2 < 1$, select an initial point $x^{(0)} = (x_1^{(0)}, \ldots, x_r^{(0)}) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_r}$ with $x_i^{(0)} = (x_{i1}^{(0)}, x_{i2}^{(0)}) \in \mathbb{R} \times \mathbb{R}^{n_i - 1}$ by taking $x_{i2}^{(0)} \neq 0$ while $x_{i1}^{(0)} \leq 0, i = 1, \ldots, r$.

Step 3 For the parameters α, μ and the initial point $x^{(0)}$, solve the nonlinear equations

$$Ax - \alpha \Phi^{-}(\mu, x)^{\sigma} = b. \tag{56}$$

Suppose that $x_{\mu,\alpha}$ is the solution of (56) and let $\mathtt{Tol} = x_{\mu,\alpha}^T (Ax_{\mu,\alpha} - b)$.

Step 4 If $|\mathtt{Tol}| \leq \mathtt{eps}$, set $\tilde{x} = x_{\mu,\alpha}$, go to step 5; else, let $x^{(0)} = x_{\mu,\alpha}$, $\alpha = c_1 \alpha$ if $x_{\mu,\alpha} \notin \mathcal{K}$, and $x^{(0)} = x_{\mu,\alpha}$, $\mu = c_2 \mu$ if $x_{\mu,\alpha} \in \mathcal{K}$, go to step 3.

Step 5 The vector $\tilde{x}$ is the approximate optimal solution of SOCLCP (1), stop.

We test some examples to show the efficiency of Algorithm 5.1. Especially, we start with a simple SOCLCP (1) whose solution can be obtained by algebraic way.

Example 5.1 Consider the SOCLCP (1) on $\mathcal{K}^2$, where

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 4 \end{bmatrix}.$$

We consider the influence of parameter variation on numerical results. To see this, noting that the exact solution of this problem is $x^* = (1, 1)^T$. We take an initial point $x^{(0)} = (0, 2)^T$ and consider the tendency of numerical results while parameters α, μ, σ change. In the following tables, “Err" denotes $\|\tilde{x} - x^*\|$, “Val" denotes $\tilde{x}^T (A\tilde{x} - b)$, where $\tilde{x}$ is the numerical solution.

We do the numerical experiment according to the following three steps:

- First, we take $\sigma = 1/2$, $\mu = 1e - 5$ and some $\alpha = 50 \times 2^i$, $i = 1, \cdots, 6$, the numerical results are listed in Table 1;
- Second, we take $\sigma = 1/2$, $\alpha = 1000$ and some $\mu = 1e - i$, $i = 2, \cdots, 7$, the numerical results are listed in Table 2;
- Finally, we take $\mu = 1e - 5$, $\alpha = 1000$ and some $\sigma = 1, 1/2, \cdots, 1/6$, the numerical results are listed in Table 3, where “-" denotes the Jacobian matrix close to singular or badly scaled.

From these tables, we can draw the following conclusions:

- From Table 1, we see that, the overall situation is that as α increases, the Err decreases, but it is not true for ϕ_3^- and ϕ_1^-. For ϕ_3^- for example, by looking at the value of Val, we can conclude that the penalty paremeter $\alpha = 200$ is large enough under the existing conditions. At this point, we won't get better results by simply increasing the value of α. The bottleneck here is the need for a smaller μ.

Table 1 The numerical results for α change ($\sigma = 1/2$, $\mu = 1\text{e}{-5}$)

$\alpha \rightarrow$		100	200	400	800	1600	3200
ϕ_1^-	Err	0.001264	3.1682e−4	7.9820e−5	1.9566e−5	5.8716e−7	1.3108e−5
	Val	−0.003198	−8.0142e−4	−2.0193e−4	−4.9497e−5	1.4854e−6	3.3160e−5
ϕ_2^-	Err	0.001264	3.1682e−4	7.9820e−5	2.0655e−5	5.4553e−6	9.9050e−7
	Val	−0.003198	−8.0142e−4	−2.0193e−4	−5.2252e−5	−1.3801e−5	−2.5058e−6
ϕ_3^-	Err	0.001174	4.5673e−4	9.6690e−4	0.001991	0.003997	0.007999
	Val	−0.002490	6.1347e−4	0.002629	0.005618	0.011347	0.022814
ϕ_4^-	Err	0.001268	3.2077e−4	8.3773e−5	2.4608e−5	9.4082e−6	4.9433e−6
	Val	−0.003208	−8.1142e−4	−2.1193e−4	−6.2252e−5	−2.3801e−5	−1.2506e−5

Table 2 The numerical results for μ change ($\sigma = 1/2$, $\alpha = 1000$)

$\mu \rightarrow$		1e−2	1e−3	1e−4	1e−5	1e−6	1e−7
ϕ_1^-	Err	0.050637	0.003261	1.3748e−4	1.1602e−5	1.3242e−5	1.3242e−5
	Val	0.130153	0.008257	3.4782e−4	−2.9352e−5	−3.3500e−5	−3.3500e−5
ϕ_2^-	Err	0.003504	2.5302e−4	5.7373e−6	1.3242e−5	1.3242e−5	1.3242e−5
	Val	0.008875	6.4015e−4	−1.4514e−5	−3.3500e−5	−3.3500e−5	−3.3500e−5
ϕ_3^-	Err	2.934882	0.255396	0.025001	0.002494	2.4442e−4	2.2246e−5
	Val	19.36984	0.890571	0.072549	0.007056	6.7396e−4	3.6514e−5
ϕ_4^-	Err	4.4813e−4	1.4225e−4	4.5265e−5	1.7195e−5	1.3637e−5	1.3282e−5
	Val	−0.001134	−3.5984e−4	−1.1451e−4	−4.3500e−5	−3.4500e−5	−3.3600e−5

Table 3 The numerical results for σ change ($\mu = 1\text{e}{-5}$, $\alpha = 1000$)

$\sigma \rightarrow$		1	1/2	1/3	1/4	1/5	1/6
ϕ_1^-	Err	–	1.1602e−5	3.8325e−5	8.1231e−5	1.3490e−4	–
	Val	–	−2.9352e−5	9.6957e−5	2.0577e−4	3.7235e−4	–
ϕ_2^-	Err	0.003159	1.3242e−5	2.4262e−6	3.5321e−6	3.6108e−6	3.8182e−6
	Val	−0.007984	−3.3500e−5	6.1378e−6	8.9355e−6	9.1346e−6	9.6594e−6
ϕ_3^-	Err	0.003159	0.002494	–	–	–	–
	Val	−0.007984	0.007056	–	–	–	–
ϕ_4^-	Err	0.003163	1.7195e−5	1.5267e−6	4.2077e−7	3.4207e−7	1.3464e−7
	Val	−0.007994	−4.3500e−5	−3.8622e−6	−1.0645e−6	−8.6539e−7	−3.4062e−7

- We see from Table 2 that, the overall situation is that as μ decreases, the Err gets decrease, but it is not true for ϕ_2^- and ϕ_1^-. For ϕ_2^-, by observing the value of Val, we can conclude that the smoothing parameter $\mu = 1e - 4$ is small enough under the existing conditions. At this point, we will not obtain better results by simply decreasing the value of μ. What we really need to improve is to increase the penalty parameter α.

Table 4 The numerical results for different initial points

IP($x^{(0)}$)	Val(ϕ_1^-)	Val(ϕ_2^-)	Val(ϕ_3^-)	Val(ϕ_4^-)	Val($[t]_-$)
$(1, 1, 1, 1, 1)^T$	5.9114e−7	1.8896e−7	4.7370e−7	5.1100e−7	7.3415e−7
$(-1, \cdots, -1)^T$	5.9119e−7	1.8902e−7	4.5308e−7	5.1100e−7	7.3419e−7
$(10, \cdots, 10)^T$	5.9124e−7	1.8908e−7	3.3751e−7	5.1101e−7	7.3423e−7
$(-10, \cdots, -10)^T$	5.9128e−7	1.8913e−7	4.3411e−7	5.1101e−7	7.3426e−7
$(10^3, \cdots, 10^3)^T$	5.9100e−7	1.8878e−7	3.6168e−7	5.1099e−7	7.3403e−7
$(10^6, \cdots, 10^6)^T$	5.9100e−7	1.8880e−7	4.4352e−7	5.1099e−7	7.3404e−7

- From Table 3, we also see that, the roughly situation is that as σ decreases, the Err gets decreases, but ϕ_3^-, ϕ_1^- are less stable than ϕ_2^-, ϕ_4^-. In addition, we see that, with the decreasing of σ, the minimum α that can satisfy the penalty is correspondingly smaller.

All these numerical results listed above coincide with Theorem 4.1. Note that if different initial points are chosen, even far away from x^*, the similar numerical results can be obtained. Yet, the points on $\mathbb{R}_- \times \{0\}$, cannot be taken as the initial point, since the denominator of LOPEs (27) in calculation equals to zero in this case. Therefore, we choose $x_{i2}^{(0)} \neq 0$ while $x_{i1}^{(0)} \leq 0$ for initial point $x^{(0)}$ in *step 2* of Algorithm 5.1.

Now, we test more examples to evaluate the efficiency of Algorithm 5.1. We use discrete Newton method to solve nonlinear equations for all examples. All numerical experiments are performed under the MATLAB 2012a running on a PC with Inter(R) Core(TM) i5-2410M CPU of 2.3GHz and RAM of 1GB.

The following two test examples are employed from [13], which will be solved by Algorithm 5.1. In our tests, we employ `eps`$=1e-6$ as the termination criterion. In the following tables, IP($x^{(0)}$) denotes the initial points, Val denotes $|\tilde{x}^T(A\tilde{x}-b)|$, where $\tilde{x}$ is the numerical solution.

Example 5.2 Consider the SOCLCP (1) on $\mathcal{K}^5$, where

$$A = \begin{bmatrix} 15 & -5 & -1 & 4 & -5 \\ 0 & 5 & 0 & 0 & 1 \\ -1 & -3 & 8 & 2 & -3 \\ 2 & -4 & 2 & 9 & -4 \\ 0 & -5 & 0 & 0 & 10 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

In this example, the matrix A is positive definite, but not symmetric, i.e., Assumption 4.1 holds. The exact solution $x^* \approx (0.049185, -0.0030997, 0.0096024, 0.0031883, 0.048033)^T$ [13]. For different initial points, by taking $\sigma = 1/2$, $c_1 = 10$, $c_2 = 0.1$, initial $\alpha = 100$ and proper initial μ($1e-6$ for ϕ_1^- and $1e-5$ for ϕ_i^-, $i = 2, 3, 4$), the test results are listed in Table 4, in which, the numerical results based on (4) are also listed.

From Table 4, we see that Algorithm 5.1 is insensitive for initial point in this example.

Example 5.3 Consider the SOCLCP (1) on $\mathcal{K}^3$, where

$$A = \begin{bmatrix} 21 & -9 & 18 \\ -9 & 4 & -7 \\ 18 & -7 & 19 \end{bmatrix}, \quad b = \begin{bmatrix} -3 \\ -7 \\ -1 \end{bmatrix}.$$

Table 5 The numerical results for different initial points

IP($x^{(0)}$)	Val(ϕ_1^-)	Val(ϕ_2^-)	Val(ϕ_3^-)	Val(ϕ_4^-)	Val($[t]_-$)
$(1, 1, 1)^T$	2.1794e−7	3.6749e−7	8.4605e−7	3.9977e−7	3.6749e−7
$(-1, -1, -1)^T$	2.1834e−7	3.6597e−7	8.4941e−7	4.0048e−7	3.6597e−7
$(10, 10, 10)^T$	2.2027e−7	3.6591e−7	8.4419e−7	4.0214e−7	3.6591e−7
$(-10, -10, -10)^T$	2.1966e−7	3.6655e−7	8.4676e−7	4.0082e−7	3.6654e−7
$(10^3, 10^3, 10^3)^T$	2.1843e−7	3.6598e−7	8.5017e−7	3.9908e−7	3.6598e−7
$(10^6, 10^6, 10^6)^T$	2.1922e−7	3.6355e−7	8.4671e−7	3.9940e−7	3.6355e−7

Table 6 The numerical results for different initial points

IP($x^{(0)}$)	Val(ϕ_1^-)	Val(ϕ_2^-)	Val(ϕ_3^-)	Val(ϕ_4^-)	Val($[t]_-$)
$(1, 1, 1, 1, 1)^T$	2.1450e−7	9.3391e−8	3.3642e−7	4.8759e−7	1.7123e−7
$(-1, \cdots, -1)^T$	–	9.3326e−8	3.6425e−7	4.8759e−7	1.6881e−7
$(10, \cdots, 10)^T$	2.1475e−7	9.3362e−8	4.1324e−7	4.8759e−7	1.7106e−7
$(-10, \cdots, -10)^T$	–	9.3362e−8	4.4575e−7	4.8759e−7	1.6955e−7
$(10^3, \cdots, 10^3)^T$	2.1475e−7	9.3361e−8	3.3902e−7	4.8759e−7	1.7101e−7
$(10^6, \cdots, 10^6)^T$	–	9.3358e−8	4.5595e−7	4.8759e−7	1.7156e−7

In this example, the symmetric matrix A is positive semidefinite, but not positive definite. As indicated in [13], it has one solution $x^* \approx (-0.183606, 0.154346, 0.099440)^T$. For different initial points, we test this problem by taking $\sigma = 1/2, c_1 = 10, c_2 = 0.1$, initial $\alpha = 1000$ and proper initial μ ($1e - 7$ for ϕ_1^- and $1e - 6$ for $\phi_i^-, i = 2, 3, 4$),
the results are listed in Table 5. The numerical results based on (4) are also listed in this table. This example indicates that, the Algorithm 5.1 is also applicable to those SOCLCPs, in which the matrix A is only positive semidefinite.

Examples 5.2 and 5.3 are two examples of SOCLCP with a single SOC $\mathcal{K} = \mathcal{K}^n$. Next, we construct two examples of the SOCLCP (1) with multiple SOCs.

Example 5.4 Consider the SOCLCP (1) on $\mathcal{K}^3 \times \mathcal{K}^2$, where A is shown as in Example 5.2, and $b = (3, 0, 2, 2, 5)^T$.

The Assumption 4.1 also holds in this example. By computation, the exact solution is about $x^* \approx (0.255103, -0.053464, 0.249438, 0.367316, 0.367316)^T$. For different initial points, the test results are listed in Table 6 by taking $\sigma = 1/2, c_1 = 10, c_2 = 0.1$, initial $\alpha = 100$ and proper initial μ ($1e - 7$ for ϕ_1^-, $1e - 5$ for $\phi_i^-, i = 2, 3, 4$).

Example 5.5 Consider the SOCLCP (1) on $\mathcal{K}^3 \times \mathcal{K}^4$, where

$$A = \begin{bmatrix} 3.9475 & 1.1370 & -0.3462 & -0.1258 & -1.2034 & -0.4979 & -1.0337 \\ 1.1370 & 3.5593 & -1.2955 & -0.4391 & -0.3009 & -0.6016 & -0.0404 \\ -0.3462 & -1.2955 & 5.0908 & -1.1187 & -0.6652 & -1.5541 & -1.0419 \\ -0.1258 & -0.4391 & -1.1187 & 3.5778 & -0.4033 & -0.1402 & -0.1991 \\ -1.2034 & -0.3009 & -0.6652 & -0.4033 & 2.9766 & 0.3725 & 0.0995 \\ -0.4979 & -0.6016 & -1.5541 & -0.1402 & 0.3725 & 4.8431 & -0.5048 \\ -1.0337 & -0.0404 & -1.0419 & -0.1991 & 0.0995 & -0.5048 & 4.0049 \end{bmatrix}$$

and $b = (2, -1, 3, -2, 4, -1, 3)^T$.

The Assumption 4.1 also holds since A is symmetry positive definite. The exact solution is about $x^* \approx (1.309502, 0.079978, 1.307057, 1.373952, 1.081021, 0.235470, 0.814673)^T$. For different initial points, we can obtain similar results to Example 5.4 by taking proper parameters, which we omit here.

In summary, according to the numerical experiments, we observe the following facts regarding parameter selection.

- The selection of initial μ. In general, for the Algorithm 5.1 to be successful and effective, the initial μ for ϕ_3^- and ϕ_1^- must be taken much smaller than the initial μ for ϕ_2^-, whereas any value of μ for ϕ_4^- is applicable.
- The selection of initial α. In practical situation, the penalty parameter α increases gradually by iteration. If we choose the initial α too large, the Algorithm 5.1 may fail because its corresponding matrix may be singular.
- The selection of power σ. Theoretically, the smaller the parameter σ is, the faster the convergence of Algorithm 5.1 will be. Nonetheless, in the implementations, the parameter σ cannot be taken too small, otherwise, the matrix singularity will likely occur. In most cases, it is appropriate to take $\sigma \geq 0.2$, and we usually use $\sigma = 1/2$ in our implementations.
- The selection of multiple parameter c_1 and c_2. The parameter c_1 represents the magnitude of increase of α, whereas the parameter c_2 represents the magnitude of decrease of μ. The closer the distance between c_1, c_2 and 1, the higher probability of success for the Algorithm 5.1 will be. For example, by keeping other parameters the same, and we have tried to implement two cases $c_1 = 10, c_2 = 0.1$ or $c_1 = 2, c_2 = 0.5$ which show that the numerical experiment success rate of the second case is higher than that of the first case. If both cases are successful, then the second case will take more time than the first case.

5.2 Performance profile of different $\boldsymbol{\phi^-(\mu, t)}$

In order to compare the performance of functions $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$ and $[t]_-$, we consider the performance profile which is introduced in [16] as a means. Assume that there are n_s solvers and n_p test problems from the test set $\mathcal{P}$. We are interested in using computing time or iteration number as a performance measure. In the following, we take computing time as a performance measure, the idea also applicable to iteration number. For each problem p and solver s, we define

$$f_{p,s} = \text{computing time required to solve problem } p \text{ by solver } s.$$

We employ the performance ratio

$$r_{p,s} = \frac{f_{p,s}}{\min\{f_{p,s} | s \in \mathcal{S}\}},$$

where $\mathcal{S}$ is the solver set. We assume that a parameter r_M, such that $r_M \geq r_{p,s}$ for all p, s is chosen, and $r_{p,s} = r_M$ if and only if solver s does not solve problem p. The choice of r_M does not affect the performance evaluation. In order to obtain an overall assessment for each solver, we define

$$\rho_s(\tau) = \frac{1}{n_p} \text{size}\{p \in \mathcal{P} | r_{p,s} \leq \tau\}.$$

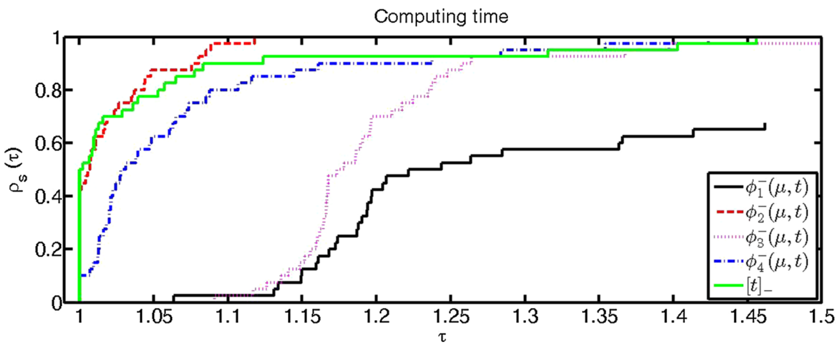

Fig. 4 Performance profile of $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$ and $[t]_-$

The function $\rho_s(\tau)$ is the cumulative performance ratio, which is called the performance profile. In the performance profile, we use functions $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$ in Algorithm 5.1 and $[t]_-$ in [21, Algorithm 4.1] as five solvers, and take randomly generated 40 SOCLCPs with single SOC, in which the matrices are symmetric positive definite. All these problems have an average distribution of dimensions from 3 to 10. The performance plot based on computing time is depicted in Fig. 4. By overall looking, from Fig. 4, we see that the function $\phi_2^-(\mu, t)$ has the best performance, then followed by $[t]_-$ and $\phi_4^-(\mu, t)$. Note that the time efficiency of $\phi_1^-(\mu, t)$ is the worst. In other words, in view of computing time, there has

$$\phi_2^-(\mu, t) > [t]_- > \phi_4^-(\mu, t) > \phi_3^-(\mu, t) > \phi_1^-(\mu, t), \tag{57}$$

where "$>$" means "better performance". If we concern the iteration number, although iteration numbers are pretty much the same, there are about 20% problems failed for $\phi_1^-(\mu, t)$. As a result, the similar results (57) are obtained. In summary, for the SOCLCPs (3), when we use Algorithm 5.1 by applying functions $\phi_i^-(\mu, t)$, $i = 1, 2, 3, 4$, no matter the iteration number or the computing time is taken into account, the function $\phi_2^-(\mu, t)$ is the best choice. Meanwhile, $\phi_2^-(\mu, t)$ in Algorithm 5.1 has better performance than $[t]_-$ in [21, Algorithm 4.1].

5.3 Generalization of $\phi_2^-(\mu, t)$ and $\phi_4^-(\mu, t)$

Among the given four typical smoothing functions (11)–(14) for $[t]_-$, from last subsection, the function $\phi_2^-(\mu, t)$ has the best performance for solving SOCLCPs (1) by using (27), then followed by $\phi_4^-(\mu, t)$. The functions $\phi_2^-(\mu, t)$ and $\phi_4^-(\mu, t)$ both smoothize the non-differentiable point of $[t]_-$ by a quadratic function, which can be further generalized by a p-power function [9,31]. Indeed, for $\phi_2^-(\mu, t)$ in (12) and $\phi_4^-(\mu, t)$ in (14), we can respectively consider a family of new smoothing functions, which include the function $\phi_2^-(\mu, t)$ and $\phi_4^-(\mu, t)$ as a special case, for solving the SOCLCP (1). More specifically, we consider the family of smoothing functions as below:

$$\psi_2^p(\mu, t) = \begin{cases} -t & \text{if } t \le \frac{-\mu}{p-1}, \\ \frac{\mu}{p-1}\left[\frac{(p-1)(-t+\mu)}{p\mu}\right]^p & \text{if } \frac{-\mu}{p-1} < t < \mu, \\ 0 & \text{if } t \ge \mu, \end{cases} \tag{58}$$

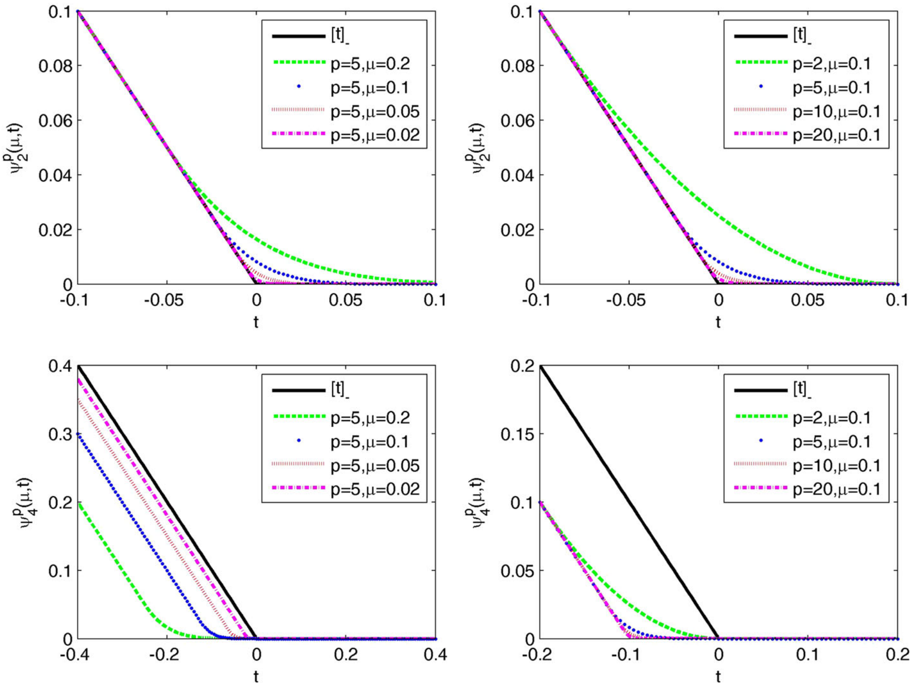

Fig. 5 Graphs of $\psi_2^p(\mu, t)$ and $\psi_4^p(\mu, t)$ with different μ and p

and

$$\psi_4^p(\mu, t) = \begin{cases} -t - u & \text{if } t \le \frac{-p\mu}{p-1}, \\ \frac{\mu}{p-1}\left[\frac{(p-1)(-t)}{p\mu}\right]^p & \text{if } \frac{-p\mu}{p-1} < t < 0, \\ 0 & \text{if } t \ge 0, \end{cases} \tag{59}$$

where $\mu > 0$ and $p \ge 2$. Note that when $p = 2$, $\psi_2^p(\frac{\mu}{2}, t)$ reduces to the smoothing function $\phi_2^-(\mu, t)$ in (12) and $\psi_4^p(\frac{\mu}{2}, t)$ reduces to the smoothing function $\phi_4^-(\mu, t)$ in (14). The graphs of $\psi_2^p(\mu, t)$ and $\psi_4^p(\mu, t)$ with different values of p and various μ are depicted as in Fig. 5. Moreover, the graphs of lower order power $(\psi_2^p(\mu, t))^\sigma$ and $(\psi_4^p(\mu, t))^\sigma$ corresponding to (27) are depicted in Fig. 6.

We see from (58) and (59) that, for a fixed $p \ge 2$, as μ goes down to zero, $\psi_2^p(\mu, t)$ and $\psi_4^p(\mu, t)$ all tend to $[t]_-$. At the same time, for a fixed μ, $\psi_2^p(\mu, t)$ also tends to $[t]_-$ as p goes to positive infinite, which shows that the increasing of p and the decreasing of μ both play a positive role for $\psi_2^p(\mu, t)$. However, the increasing of p does not play a positive role for $\psi_4^p(\mu, t)$, since for a fixed μ and a negative small t, the distance of $\psi_4^p(\mu, t)$ and $[t]_-$ equals the constant μ. All of these are reflected in Figs. 5 and 6.

Next, we present the performance profile for various p. We take randomly generated 40 SOCLCPs as in Sect. 5.2, and replace $\phi^-(\mu, t)$ in (27) by $\psi_2^p(\mu, t)$, $p = 2, 3, 5, 10$ and $\psi_4^p(\mu, t)$, $p = 2, 3, 5, 10$, respectively, in Algorithm 5.1 as four solvers. The performance profiles based on computing time are presented in Figs. 7 and 8, and the performance based on iteration number are similar.

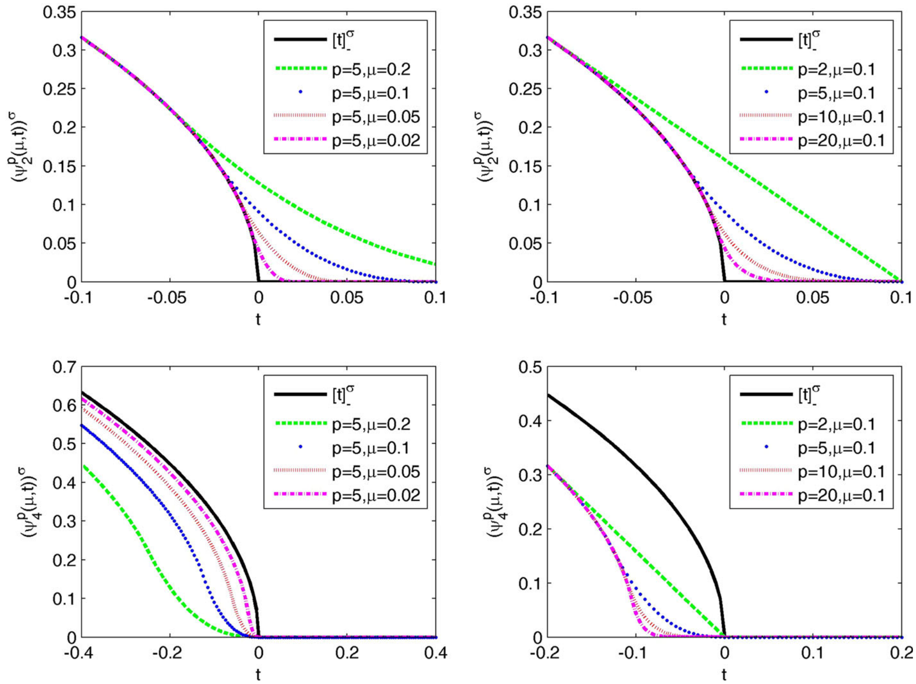

Fig. 6 Graphs of $(\psi_2^p(\mu,t))^\sigma$ and $(\psi_4^p(\mu,t))^\sigma$ with different p, μ and $\sigma = 1/2$

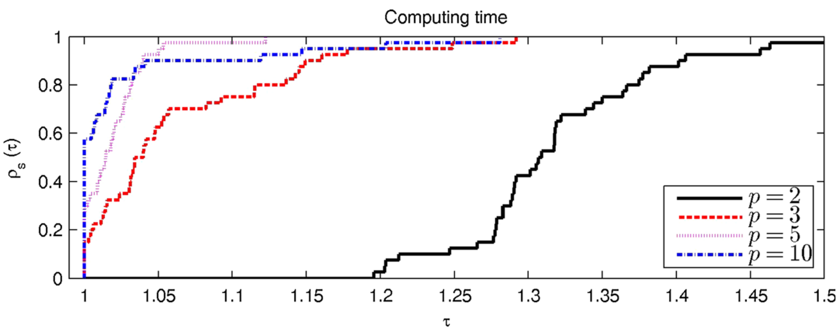

Fig. 7 Performance profile of $\psi_2^p(\mu,t)$, $p = 2, 3, 5, 10$

We conclude from Fig. 7 that, in general, the greater value of p is, the better performance of $\psi_2^p(\mu,t)$ will be, the visibility decreases with increasing of p. But, this phenomenon is totally different for $\psi_4^p(\mu,t)$. As we see from Fig. 8, the greater value of p is, the worse performance of $\psi_4^p(\mu,t)$ will be. In summary, for the SOCLCPs (1), when we replace $\phi^-(\mu,t)$ in (27) by $\psi_i^p(\mu,t)$, $i = 2, 4$ in (58) and (59), the appropriate greater value of p could be better choice for $\psi_2^p(\mu,t)$, whereas $p = 2$ is the best choice for $\psi_4^p(\mu,t)$.

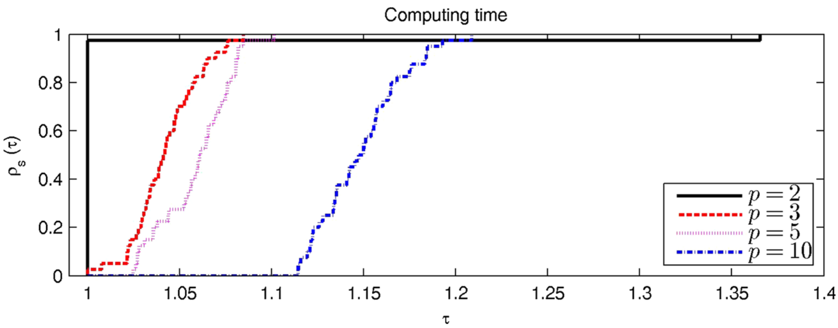

Fig. 8 Performance profile of $\psi_4^p(\mu, t)$, $p = 2, 3, 5, 10$

6 Conclusions

Based on the asymptotic approximate LOPEs (27), we propose an approximate lower order penalty approach for solving SOCLCPs (1). The main result is Theorem 4.1, which shows that, under Assumption 4.1, the solution sequence of asymptotic approximate LOPEs (27) converges to the solution of SOCLCP (1) at an exponential rate. Four specific approximate LOPEs (27) are considered, corresponding to $\phi_i^-(\mu, t), i = 1, 2, 3, 4$. Numerical experiments indicate that, the Algorithm 5.1 with $\phi_2^-(\mu, t)$ has the best performance, then followed by $\phi_4^-(\mu, t)$, and $\phi_1^-(\mu, t)$ has the worst performance. Meanwhile, $\phi_2^-(\mu, t)$ in Algorithm 5.1 has better performance than $[t]_-$ in [21, Algorithm 4.1] for some certain problems.

By generalizing quadratic function to a p-power ($p \geq 2$) function, we also investigate more general functions, $\psi_2^p(\mu, t)$ in (58) and $\psi_4^p(\mu, t)$ in (59), which are extensions of the functions $\phi_2^-(\mu, t)$ and $\phi_4^-(\mu, t)$ respectively. For $\psi_2^p(\mu, t)$, the appropriate greater p is the better choice, meanwhile, for $\psi_4^p(\mu, t)$, the best choice is $p = 2$.

To sum up, the study of this paper not only build up theoretical bricks for the proposed approach, but also suggest possible good choices of smoothing functions that work along with the algorithm. Further work includes the consideration of more extensive problems, such as second order cone nonlinear complementarity problem, and weaker assumptions etc.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. **95**(1), 3–51 (2003)
2. Bertsekas, D.P., Nedić, A., Ozdaglar, A.E.: Convex Analysis and Optimization. Athena Scientific, Belmont (2003)
3. Billupsa, S.C., Murtyb, K.G.: Complementarity problems. J. Comput. Appl. Math. **124**, 303–318 (2000)
4. Chen, C., Mangasarian, O.L.: A class of smoothing functions for nonlinear and mixed complementarity problems. Comput. Optim. Appl. **5**, 97–138 (1996)
5. Chen, J.-S.: Two classes of merit functions for the second-order cone complementarity problem. Math. Methods Oper. Res. **64**, 495–519 (2006)
6. Chen, J.-S.: The convex and monotone functions associated with second-order cone. Optimization **55**, 363–385 (2006)
7. Chen, J.-S.: SOC Functions and Their Applications, Springer Optimization and Its Applications 143. Springer, Singapore (2019)
8. Chen, J.-S., Chen, X., Tseng, P.: Analysis of nonsmooth vector-valued functions associated with second-order cones. Math. Program. **101**, 95–117 (2004)
9. Chen, J.-S., Ko, C.-H., Liu, Y.-D., Wang, S.-P.: New smoothing functions for solving a system of equalities and inequalities. Pac. J. Optim. **12**(1), 185–206 (2016)

10. Chen, J.-S., Pan, S.-H.: A descent method for a reformulation of the second-order cone complementarity problem. J. Comput. Appl. Math. **213**, 547–558 (2008)
11. Chen, J.-S., Pan, S.-H.: A survey on SOC complementarity functions and solution methods for SOCPs and SOCCPs. Pac. J. Optim. **8**, 33–74 (2012)
12. Chen, J.-S., Tseng, P.: An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. Math. Program. **104**, 297–327 (2005)
13. Chen, L.-J., Ma, C.-F.: A modified smoothing and regularized Newton method for monotone second-order cone complementarity problems. Comput. Math. Appl. **61**, 1407–1418 (2011)
14. Chen, X.-D., Sun, D., Sun, J.: Complementarity functions and numerical experiments for second-order cone complementarity problems. Comput. Optim. Appl. **25**, 39–56 (2003)
15. Cottle, R.W., Pang, J.S., Stone, R.E.: The Linear Complementarity Problem. Academic Press, San Diego (1992)
16. Dolan, E.D., Moré, J.J.: Benchmarking optimization software with performance profiles. Math. Program. **91**(2), 201–213 (2002)
17. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I and II. Springer-Verlag, New York (2003)
18. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford University Press, New York, Oxford Mathematical Monographs (1994)
19. Fukushima, M., Luo, Z.-Q., Tseng, P.: Smoothing functions for second-order-cone complementarity problems. SIAM J. Optim. **12**, 436–460 (2002)
20. Han, S.-P., Mangasrian, O.L.: Exact penalty function in nonlinear programming. Math. Program. **17**, 251–269 (1979)
21. Hao, Z., Wan, Z., Chi, X.: A power penalty method for second-order cone linear complementarity problems. Oper. Res. Lett. **43**, 137–142 (2015)
22. Hayashi, S., Yamaguchi, T., Yamashita, N., Fukushima, M.: A matrix-splitting method for symmetric affine second-order cone complementarity problems. J. Comput. Appl. Math. **175**, 335–353 (2005)
23. Hayashi, S., Yamashita, N., Fukushima, M.: A combined smoothing and regularization method for monotone second-order cone complementarity problems. SIAM J. Optim. **15**, 593–615 (2005)
24. Huang, Z.-H., Ni, T.: Smoothing algorithms for complementarity problems over symmetric cones. Comput. Optim. Appl. **45**, 557–579 (2010)
25. Kanzow, C., Ferenczi, I., Fukushima, M.: On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity. SIAM J. Optim. **20**, 297–320 (2009)
26. Lemke, C.E.: Bimatrix Equilibrium Points and Mathematical Programming. Manage. Sci. **11**(7), 681–689 (1965)
27. Lemke, C.E., Howson, J.T., Jr.: Equilibrium points of bimatrix games. J. Soc. Ind. Appl. Math. **12**(2), 413–423 (1964)
28. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Applications of second order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
29. Luo, Z.-Q., Pang, J.S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
30. Mangasarian, O.L., Pang, J.S.: The extended linear complementarity problem. SIAM J. Matrix Anal. Appl. **16**, 359–368 (1995)
31. Miao, X.-H., Qi, N., Saheya, B., Chen, J.-S.: Applying a type of SOC-functions to solve a system of equalities and inequalities under the order induced by second-order cone. Pac. J. Optim. **15**(1), 1–22 (2019)
32. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of primal algorithms for the second-order cone programs based on the MZ-family of directions. Math. Program. **88**, 61–83 (2000)
33. Pan, S.-H., Chen, J.-S.: A damped Gauss-Newton method for the second-order cone complementarity problem. Appl. Math. Optim. **59**, 293–318 (2009)
34. Di Pillo, G., Grippo, L.: An exact penalty function method with global convergence properties for nonlinear programming problems. Math. Program. **36**, 1–18 (1986)
35. Pinar, M.C., Zenios, S.A.: On smoothing exact penalty functions for convex constrained optimization. SIAM J. Optim. **4**, 486–511 (1994)
36. Wang, S., Yang, X.-Q.: A power penalty method for linear complementarity problems. Oper. Res. Lett. **36**, 211–214 (2008)
37. Wu, Z.-Y., Bai, F.-S., Yang, X.-Q., Zhang, L.-S.: An exact lower order penalty function and its smoothing in nonlinear programming. Optimization **53**, 51–68 (2004)
38. Xu, X.-S., Meng, Z.-Q., Sun, J.-W., Shen, R.: A penalty function method based on smoothing lower order penalty function. J. Comput. Appl. Math. **235**, 4047–4058 (2011)

39. Yang, X.-Q., Huang, X.-X.: Lower-order penalty methods for mathematical programs with complementarity constraints. Optim. Methods Softw. **19**(6), 693–720 (2004)
40. Zangwill, W.I.: Nonlinear programming via penalty function. Manag. Sci. **13**, 334–358 (1967)
41. Zhang, L.-H., Yang, W.-H.: An efficient matrix splitting method for the second-order cone complementarity problem. SIAM J. Optim. **24**, 1178–1205 (2014)