Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 6, No. 3, pp. 551 – 571, September 2010

SMOOTHNESS OF A CLASS OF GENERALIZED MERIT FUNCTIONS FOR THE SECOND-ORDER CONE COMPLEMENTARITY PROBLEM*

Sheng-Long Hu, Zheng-Hai Huang† and Nan Lu

Abstract: In this paper, we consider the second-order cone complementarity problem (SOCCP). We propose a family of complementarity functions for the second-order cone complementarity problem (SOC C-functions), which contains several popular SOC C-functions as special cases. Based on the new SOC C-functions, a family of merit functions for the SOCCP is proposed. We show that the new merit functions are continuously differentiable and give their derivative formulae. These provide an important theoretical basis for designing some merit function methods to solve the SOCCP. Some preliminary numerical results indicate that the new SOC C-functions and the corresponding merit functions are worth investigating.

Key words: *second-order cone, complementarity function, merit function*

Mathematics Subject Classification: *90C26, 90C30, 90C33*

---

1 Introduction

In the last two decades, people have put a lot of their energy and attention on complementarity problems due to their various applications in operations research, economics, and engineering (see, for example, [8, 11, 16]). Many algorithms were proposed to solve the nonlinear complementarity problem (NCP) (see the excellent monograph [8]). Recently, there are great interests in designing various algorithms for solving some conic complementarity problems, such as the second-order cone complementarity problem (SOCCP) [4, 5, 7, 15], the semidefinite complementarity problem [6, 18], and the symmetric cone complementarity problem [10, 19, 20, 26]. In this paper, we are interested in the SOCCP which is to find a point $x \in \Re^n$ such that

$$x \succeq 0, \quad F(x) \succeq 0, \quad \langle x, F(x) \rangle = 0, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, $F : \Re^n \to \Re^n$ is a continuously differentiable mapping, and $\succeq$ is a partial order induced by a second-order cone $\mathcal{K}$ (i.e., $x \succeq 0$ means $x \in \mathcal{K}$; similarly, $x \succ 0$ means $x \in \text{int}\, \mathcal{K}$ (the interior of $\mathcal{K}$)) defined by

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

*This work was partially supported by the National Natural Science Foundation of China (Grant No. 10871144) and the Natural Science Foundation of Tianjin (Grant No. 07JCYBJC05200).

†Corresponding author.

here integers $m, n_1, \dots, n_m \geq 1, n_1 + \cdots + n_m = n$, and $\mathcal{K}^{n_i} := \{(x_1, x_2) \in \Re \times \Re^{n_i-1} : \|x_2\| \leq x_1\}$ with $\|\cdot\|$ denoting the Euclidean norm. For simplicity, we assume that $\mathcal{K} = \mathcal{K}^n$ without loss of generality.

Many solution methods have been developed to solve the SOCCP (1.1). One of the most popular methods is to reformulate the SOCCP (1.1) as an unconstrained optimization problem and then to solve the reformulated problem by using unconstrained optimization techniques. This kind of methods is called the merit function method, where the merit function is generally constructed by some SOC C-function.

Definition 1.1. A function $\phi : \Re^n \times \Re^n \to \Re^n$ is called an SOC C-function [5, 25], if it satisfies that $\phi(a, b) = 0$ if and only if $a \succeq 0, b \succeq 0, \langle a, b \rangle = 0$. In addition, if a function $\Psi : \Re^n \to \Re$ is nonnegative and $\Psi(x) = 0$ if and only if x solves the SOCCP (1.1), then Ψ is called a merit function for the SOCCP (1.1).

If ϕ is an SOC C-function, then it is easy to see that the function $\Psi : \Re^n \to \Re$ defined by $\Psi(x) := \frac{1}{2}\|\phi(x, F(x))\|^2$ is a merit function for the SOCCP (1.1). Thus, finding a solution of the SOCCP (1.1) is equivalent to finding a global minimum of the unconstrained minimization $\min_{x \in \Re^n} \Psi(x)$ with optimal value 0. It is well known that most effective unconstrained minimization methods require the smoothness of the objective function. Thus, in such a reformulation method for the SOCCP, a basic requirement is that the objective function (i.e., Ψ) is smooth. In this paper, we will propose a new class of generalized SOC C-functions and discuss the smoothness of their related merit functions.

When $\mathcal{K}^n = \Re^n_+ (:= \{(x_1, \dots, x_n) \in \Re^n \mid x_i \geq 0\})$, the SOCCP reduces to the NCP and the SOC C-function reduces to the NCP-function. Many NCP-functions have been proposed in the literature [1, 22, 21, 27, 29]. Among them, the FB function is one of the most popular NCP-functions, which is defined by

$$\phi(a, b) := \sqrt{a^2 + b^2} - a - b, \quad \forall (a, b) \in \Re^2.$$

One of the main generalizations of the FB function was given by Kanzow and Kleinmichel [21]:

$$\phi_\theta(a, b) := \sqrt{(a - b)^2 + \theta ab} - a - b, \quad \theta \in (0, 4),\ \forall (a, b) \in \Re^2. \tag{1.2}$$

Another main generalization was given by Luo and Tseng [22], and studied by Chen [2]:

$$\phi_p(a, b) := \sqrt[p]{|a|^p + |b|^p} - a - b, \quad p \in (1, \infty),\ \forall (a, b) \in \Re^2. \tag{1.3}$$

The NCP-functions given in (1.2) and (1.3) had been extended to the framework of second-order cones [24, 25], and it had been proved that they enjoy a lot of favorable properties. Both of the papers [24, 25] pointed out that many best numerical results of the functions ϕ_θ given in (1.2) and ϕ_p given in (1.3) do not appear in the case of $\theta = 2$ and $p = 2$, respectively. It is easy to see that both ϕ_θ and ϕ_p reduce to the well-known FB function when $\theta = 2$ and $p = 2$, respectively. Hence, it is of great interest and importance to investigate the theoretical properties and numerical behavior of these two classes of generalized SOC C-functions and their generalization. In this paper, we propose a new family of SOC C-functions which is a generalization of both the generalized SOC C-functions mentioned above. The proposed SOC C-functions are new even in the case of $\mathcal{K}^n = \Re^n_+$. Since the NCP is a special case of the SOCCP [4, 5, 24, 25], our results can be directly applied to the NCP. Now, we formally propose the following function, which will be proved to be an SOC C-function later.

$$\phi_{\theta p}(x, y) := (\theta(|x|^p + |y|^p) + (1 - \theta)|x + y|^p)^{1/p} - x - y, \tag{1.4}$$

where $x, y \in \Re^n$, and $\theta \in (0,2), p = 2$ or $\theta \in (0,1], p \in (1,2)\bigcup(2,+\infty)$, see (2.2) for the definition of $|\cdot|$.

Remark 1.2. (i) It is easy to see that the function defined by (1.4) reduces to ϕ_p under the framework of second-order cones when $\theta = 1$ and to ϕ_θ under the framework of second-order cones when $p = 2$ and $\theta \in (0,2)$. In particular, it reduces to the FB function under the framework of second-order cones when $\theta = 1$ and $p = 2$. Thus, the new family of the functions defined by (1.4) is a generalization of several SOC C-functions mentioned above.

(ii) It should be pointed out that it is possible that the function $\phi_{\theta p}$ defined by (1.4) is not an SOC C-function for $\theta \in (1,2)$ and $p \in (1,2)\bigcup(2,+\infty)$. For example, if we let $n = 1, \theta = 1.5, p = 4, x = 2, y = 2$, then $1.5 \times (2^4 + 2^4) - 0.5 \times (2+2)^4 = -80 < 0$. Hence, the function $\phi_{\theta p}$ in this case is even not well-defined.

(iii) When $\mathcal{K}^n = \Re^n_+$, a similar family of NCP-functions

$$\phi_{\theta p}(x,y) := (\theta(|x|^p + |y|^p) + (1-\theta)|x-y|^p)^{1/p} - x - y,\ \forall x, y \in \Re,\ \forall \theta \in (0,1], p > 1$$

was proposed in [17]. We don't know whether this family of functions can be extended to the case of second-order cones or not. In addition, it was mentioned in [3] that the best numerical results of the algorithms involving ϕ_θ happens when $\theta = 2.5$ or 3.0 for the SOCCP. Obviously, this case is not included in the function $\phi_{\theta p}$ proposed in [17], however, it is a special case of $\phi_{\theta p}$ defined by (1.4).

Since the family of generalized SOC C-functions (1.4) reduces to the SOC C-functions studied in [3] when $\theta \in (0,2]$ and $p = 2$, and it reduces to the generalized SOC C-function studied in [25] when $\theta = 1$, we only consider the case where $\theta \in (0,1)$ and $p \in (1,+\infty)$ in the rest of this paper.

The rest of this paper is organized as follows. In Section 2, we review some basic concepts and results of second-order cones. In Section 3, we show that the functions defined by (1.4) are SOC C-functions. Some other properties are also discussed. In Section 4, we propose a family of merit functions based on the new SOC C-functions; and show the smoothness of this family of merit functions. Some preliminary numerical results are reported in Section 5, where most of the very well performed numerical results do not appear in the case when $\phi_{\theta p}$ proposed in this paper reduces to the existing SOC C-functions. Some conclusions are given in Section 6.

2 Preliminaries

In the following, we review some basic concepts and results of second-order cones, see also the excellent summarizations [4, 12, 25]. It should be noted that the second-order cone is a special case of the square cone of some Euclidean Jordan algebra, see the monograph by Faraut and Korányi [9]. In the following, we will use (x_1, x_2) to denote $(x_1, x_2^T)^T$ for any $x_1 \in \Re$ and $x_2 \in \Re^{n-1}$ for convenience. For any $x = (x_1, x_2)$, $y = (y_1, y_2) \in \Re \times \Re^{n-1}$ with $n \geq 1$, we define

$$\langle x, y\rangle := x^T y, \quad \text{and} \quad x \circ y := (\langle x, y\rangle, x_1 y_2 + y_1 x_2),$$

where the former is just the usual inner product on the finite Euclidean space, and the latter is called the Jordan product [9]. It is well known that the second-order cone is a closed convex cone with nonempty interior [9], i.e., $\text{int}(\mathcal{K}^n) := \{x = (x_1, x_2) \in \Re \times \Re^{n-1} \mid x_1 > \|x_2\|\} \neq \emptyset$. The determinant and trace of x are defined by $\det(x) := x_1^2 - \|x_2\|^2$ and $\text{tr}(x) := 2x_1$, respectively. A vector $x \in \Re^n$ is invertible if and only if $\det(x) \neq 0$, and its inverse is denoted by x^{-1}. Given a vector $x = (x_1, x_2) \in \Re \times \Re^{n-1}$, we will use the following symmetric matrix to describe some key properties of second-order cones:

$$L_x := \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix},$$

where I is the identity matrix with appreciate size. We can view it as a linear operator from $\Re^n$ to $\Re^n$. It is easy to verify that $L_x y = x \circ y$ and $L_{x+y} = L_x + L_y$ for all $x, y \in \Re^n$. We note that $L_{\theta x} = \theta L_x$ holds for all $x \in \Re^n$ and $\theta \in \Re$. It should be noted that there are some connections between the second-order cone and the positive semidefinite cone (the set of all positive semidefinte symmetric matrices on $\Re^{n\times n}$), which are listed below:

$$x \in \mathcal{K}^n \Longleftrightarrow L_x \in \mathcal{S}_+^n;$$

$$x \in \text{int}\ \mathcal{K}^n \Longleftrightarrow L_x \in \mathcal{S}_{++}^n, \text{ and } L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \frac{\det(x)}{x_1} I + \frac{1}{x_1} x_2 x_2^T \end{bmatrix} \text{ in this case,}$$

where $\mathcal{S}_+^n$ denotes the positive semidefinite cone, while $\mathcal{S}_{++}^n$ is the interior of the positive semidefinite cone (i.e., the set of all positive definite symmetric matrices on $\Re^{n\times n}$).

For any $x = (x_1, x_2) \in \Re \times \Re^{n-1}$, the spectral decomposition (or spectral factorization) of x is given by $x = \lambda_1(x)u^{(1)}(x) + \lambda_2(x)u^{(2)}(x)$, where $\lambda_i(x)$ $(i = 1, 2)$ and $u^{(i)}(x)$ $(i = 1, 2)$ are the spectral values and the associated spectral vectors of x, respectively, with

$$\lambda_i(x) := x_1 + (-1)^i \|x_2\|, \quad \text{and} \quad u^{(i)}(x) := \frac{1}{2}\left(1, (-1)^i \hat{x}_2\right), \quad i = 1, 2, \tag{2.1}$$

here, if $x_2 \neq 0$ then $\hat{x}_2 := \frac{x_2}{\|x_2\|}$, and otherwise, $\hat{x}_2 \in \Re^{n-1}$ is arbitrarily taken satisfying $\|\hat{x}_2\| = 1$. It is obvious that the spectral decomposition of x is unique if $x_2 \neq 0$.

Lemma 2.1. *Let $x, s \in \Re^n$. Then, $x \succeq 0$, $s \succeq 0$, $x \circ s = 0$ if and only if $x \succeq 0$, $s \succeq 0$, $\langle x, s\rangle = 0$. Moreover, x and s share a common Jordan system of spectral vectors in each case.*

Proof. The results are special cases of [13, Propositions 6 and 7]. □

We next review some basic results about Löwner operator [28]. For any scalar function $g : \Re \to \Re$, the spectral function $\bar{g}$ induced by g is defined by

$$\bar{g}(x) := g(\lambda_1(x))u^{(1)}(x) + g(\lambda_2(x))u^{(2)}(x), \quad \forall x \in \Re^n.$$

This function was first introduced and analyzed by Löwner [23], and hence, was called Löwner function (or Löwner operator) in honor of Löwner's contribution. From [4, Proposition 5] and [12, Proposition 5.2], $\bar{g}$ is (continuously) differentiable on $\Re^n$ if and only if g is (continuously) differentiable on $\Re$. When $g(\alpha) = |\alpha|^p$ $(p \geq 1)$ for any $\alpha \in \Re$,

$$|x|^p := \bar{g}(x) = |\lambda_1(x)|^p u^{(1)}(x) + |\lambda_2(x)|^p u^{(2)}(x), \quad \forall x \in \Re^n; \tag{2.2}$$

and when $g(\alpha) = \alpha^{1/p}$ $(p > 1)$ for any $\alpha \in \Re_+$,

$$x^{1/p} := \bar{g}(x) = (\lambda_1(x))^{1/p} u^{(1)}(x) + (\lambda_2(x))^{1/p} u^{(2)}(x), \quad \forall x \in \mathcal{K}^n.$$

Hence, the function $\phi_{\theta p}$ defined by (1.4) is well-defined. Similarly, we can define x^p for $p > 1$.

The following results are evident.

Proposition 2.2. *For any $x \in \Re^n$, let $\lambda_1(x)$, $\lambda_2(x)$ and $u^{(1)}(x)$, $u^{(2)}(x)$ be the spectral values and the associated spectral vectors of x, respectively. Then $x \succeq 0$ if and only if $0 \le \lambda_1(x) \le \lambda_2(x)$; and $x \succ 0$ if and only if $0 < \lambda_1(x) \le \lambda_2(x)$.*

3 Complementarity Function

In the following, we will show that the new functions defined by (1.4) are indeed SOC C-functions. Before starting our analysis, we list some symbols for convenience. For any $x = (x_1, x_2), y = (y_1, y_2) \in \Re \times \Re^{n-1}$ and $p > 1$, we have

$$|x|^p + |y|^p = \begin{pmatrix} \frac{|\lambda_1(x)|^p + |\lambda_2(x)|^p}{2} + \frac{|\lambda_1(y)|^p + |\lambda_2(y)|^p}{2} \\ \frac{|\lambda_2(x)|^p - |\lambda_1(x)|^p}{2}\hat{x}_2 + \frac{|\lambda_2(y)|^p - |\lambda_1(y)|^p}{2}\hat{y}_2 \end{pmatrix},$$

$$|x+y|^p = \begin{pmatrix} \frac{|\lambda_1(x+y)|^p + |\lambda_2(x+y)|^p}{2} \\ \frac{|\lambda_2(x+y)|^p - |\lambda_1(x+y)|^p}{2}\widehat{(x+y)}_2 \end{pmatrix},$$

where

$$\lambda_1 := \lambda_1(x) = x_1 - \|x_2\|, \quad \lambda_2 := \lambda_2(x) = x_1 + \|x_2\|; \tag{3.1}$$

$$\mu_1 := \lambda_1(y) = y_1 - \|y_2\|, \quad \mu_2 := \lambda_2(y) = y_1 + \|y_2\|; \tag{3.2}$$

$$\bar{\lambda}_1 := \lambda_1(x+y) = x_1 + y_1 - \|x_2 + y_2\|, \quad \bar{\lambda}_2 := \lambda_2(x+y) = x_1 + y_1 + \|x_2 + y_2\|, \tag{3.3}$$

and $\hat{x}_2, \hat{y}_2, \widehat{(x+y)}_2$ are defined similarly as those in (2.1).

Theorem 3.1. *The functions $\phi_{\theta p}$ defined by (1.4) are SOC C-functions.*

Proof. We divide the proof into the following two parts.

Part I For any $x, y \in \Re^n$ with $x \in \mathcal{K}^n, y \in \mathcal{K}^n$ and $x \circ y = 0$, it follows from Lemma 2.1 that x and y share a common system of spectral vectors:

$$x = \lambda_1(x)c_1 + \lambda_2(x)c_2; \quad y = \lambda_1(y)c_1 + \lambda_2(y)c_2.$$

Then, $x + y = (\lambda_1(x) + \lambda_1(y))c_1 + (\lambda_2(x) + \lambda_2(y))c_2$. So,

$$\begin{aligned} \phi_{\theta p}(x,y) &= (\theta(|\lambda_1(x)|^p + |\lambda_1(y)|^p) + (1-\theta)|\lambda_1(x) + \lambda_1(y)|^p)c_1 \\ &\quad + (\theta(|\lambda_2(x)|^p + |\lambda_2(y)|^p) + (1-\theta)|\lambda_2(x) + \lambda_2(y)|^p)c_2)^{1/p} \\ &\quad - ((\lambda_1(x) + \lambda_1(y))c_1 + (\lambda_2(x) + \lambda_2(y))c_2) \\ &= ((\theta(|\lambda_1(x)|^p + |\lambda_1(y)|^p) + (1-\theta)|\lambda_1(x) + \lambda_1(y)|^p)^{1/p} - (\lambda_1(x) + \lambda_1(y)))c_1 \\ &\quad + ((\theta(|\lambda_2(x)|^p + |\lambda_2(y)|^p) + (1-\theta)|\lambda_2(x) + \lambda_2(y)|^p)^{1/p} - (\lambda_2(x) + \lambda_2(y)))c_2 \\ &= 0c_1 + 0c_2 \\ &= 0, \end{aligned}$$

where the third equality follows from the fact that $\lambda_1(x)\lambda_1(y) = 0$ and $\lambda_2(x)\lambda_2(y) = 0$ because of $x \circ y = 0$, $x, y \in \mathcal{K}^n$, and Proposition 2.2.

Part II Suppose that $\phi_{\theta p}(x, y) = 0$. Then

$$(\theta(|x|^p + |y|^p) + (1-\theta)|x+y|^p)^{1/p} = x + y,$$

which implies $x+y\in\mathcal{K}^n$. Hence, $|x+y|^p=(x+y)^p$. So,

$$(\theta(|x|^p+|y|^p)+(1-\theta)|x+y|^p)^{1/p}=(\theta(|x|^p+|y|^p)+(1-\theta)(x+y)^p)^{1/p}=x+y,$$

which yields that $\theta(|x|^p+|y|^p)+(1-\theta)(x+y)^p$, i.e., $\theta(|x|^p+|y|^p)=\theta(x+y)^p$. Since $\theta\in(0,1)$, by using [25, Lemma 3.1] we can obtain the desired result.

Combining **Part I** with **Part II**, we complete the proof. $\square$

In the following, we establish a technique lemma which plays a crucial role in the sequel analysis. For convenience, we give some notations first. For all $\theta\in(0,1)$ and $p\in(1,+\infty)$, we define

$$s \;:=\; (s_1,s_2):=s(x,y):=\theta(|x|^p+|y|^p)+(1-\theta)|x+y|^p; \tag{3.4}$$

$$t \;:=\; (t_1,t_2):=t(x,y):=\left(\theta(|x|^p+|y|^p)+(1-\theta)|x+y|^p\right)^{1/p}, \tag{3.5}$$

where $s=(s_1,s_2)\in\Re\times\Re^{n-1}$ and $t=(t_1,t_2)\in\Re\times\Re^{n-1}$.

Lemma 3.2. *For any $\theta\in(0,1)$, $p>1$, $x=(x_1,x_2)$, and $y=(y_1,y_2)\in\Re^n$, let the function s be defined by (3.4) with $s\notin \mathrm{int}\,\mathcal{K}^n$. Then we have*

$$s_1=\|s_2\|=2^{p-1}(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p); \tag{3.6}$$

$$x_1^2=\|x_2\|^2;\quad y_1^2=\|y_2\|^2;\quad x_1y_1=x_2^Ty_2;\quad x_1y_2=y_1x_2. \tag{3.7}$$

Furthermore, when $s_2\neq 0$, we have

$$x_2^T\frac{s_2}{\|s_2\|}=x_1;\quad x_1\frac{s_2}{\|s_2\|}=x_2;\quad y_2^T\frac{s_2}{\|s_2\|}=y_1;\quad y_1\frac{s_2}{\|s_2\|}=y_2. \tag{3.8}$$

Proof. We will divide the proof into the following two parts.

Part I In this part, we prove (3.6) and (3.7).

When $\theta\in(0,1)$ and $p>1$, we have $\theta|x|^p\in\mathcal{K}^n,\theta|y|^p\in\mathcal{K}^n$ and $(1-\theta)|x+y|^p\in\mathcal{K}^n$ for all $x=(x_1,x_2),y=(y_1,y_2)\in\Re^n$. Since $s\notin\mathrm{int}\ \mathcal{K}^n$ and $\mathcal{K}^n$ is a closed convex cone with nonempty interior, it follows that $\theta|x|^p\notin\mathrm{int}\ \mathcal{K}^n,\theta|y|^p\notin\mathrm{int}\ \mathcal{K}^n$ and $(1-\theta)|x+y|^p\notin\mathrm{int}\ \mathcal{K}^n$, and hence, $|x|\notin\mathrm{int}\ \mathcal{K}^n,|y|\notin\mathrm{int}\ \mathcal{K}^n$ and $|x+y|\notin\mathrm{int}\ \mathcal{K}^n$. Furthermore,

$$x_1^2=\|x_2\|^2;\quad y_1^2=\|y_2\|^2;\quad (x_1+y_1)^2=\|x_2+y_2\|^2$$

by a simple analysis of spectral values. Then $x_1y_1=x_2^Ty_2$ follows directly from the above equalities. So, if one of x_1 and y_1 is zero, then the last equality in (3.7) holds trivially; otherwise, we have $\|x_2\|\|y_2\|=|x_1y_1|=|x_2^Ty_2|$. That is, $x_2=\alpha y_2$ for some $\alpha\neq 0$ since both $x_1\neq 0$ and $y_1\neq 0$ (so are x_2 and y_2 by the above analysis). It follows from $x_1y_1=x_2^Ty_2=\alpha y_2^Ty_2=\alpha y_1^2$ that $\alpha=\frac{x_1}{y_1}$. Hence, the last equality of (3.7) holds in this case. Thus, (3.7) is proved to be true. While $s_1=\frac{1}{2}(\theta(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)+(1-\theta)(|\bar\lambda_1|^p+|\bar\lambda_2|^p))$, (3.6) follows from (3.7) and $s\notin\mathrm{int}\ \mathcal{K}^n$ immediately.

Part II We prove (3.8) by discussing several cases in the following way.

(i) $x_2=y_2=0$. This case can not happen since $s_2\neq 0$.

(ii) $x_2=0,y_2\neq 0$. In this case, we have $x=0$ by (3.7). Hence, $s=\theta|y|^p+(1-\theta)|y|^p=|y|^p$. It is evident that (3.8) holds.

(iii) $x_2\neq 0,y_2=0$. In this case, the results are evidently satisfied by the symmetry of x and y in s and the item (ii) above.

(iv) $x_2 \neq 0, y_2 \neq 0$ while $x_2 + y_2 = 0$. In this case, $|x+y|^p \notin \text{int } \mathcal{K}^n$ since $\theta \in (0,1)$. We have $x + y = 0$, and hence, it reduces to the one considered in [25, Lemma 3.3] by multiplying a constant. So, the result holds from a similar proof as the one in [25, Lemma 3.3].

(v) $x_2 \neq 0, y_2 \neq 0, x_2 + y_2 \neq 0$. Using (3.7), we could get

$$\|x_2 + y_2\| = \|\frac{x_1}{y_1} y_2 + y_2\| = \frac{|x_1 + y_1|}{|y_1|}\|y_2\| = \frac{|x_1 + y_1|}{|x_1|}\|x_2\|.$$

So, if $x_1 = \|x_2\|$ and $y_1 = -\|y_2\|$, we have

$$\begin{array}{l} \lambda_1 = 0; \quad \lambda_2 = 2x_1; \quad \mu_1 = 2y_1; \quad \mu_2 = 0; \\ \bar{\lambda}_1 = 0; \quad \bar{\lambda}_2 = 2(x_1 + y_1), \quad \text{if } x_1 + y_1 > 0; \\ \bar{\lambda}_1 = 2(x_1 + y_1); \quad \bar{\lambda}_2 = 0, \quad \text{if } x_1 + y_1 < 0. \end{array}$$

Furthermore,

$$\begin{aligned} 2x_1 s_2 &= \theta \left((|\lambda_2|^p - |\lambda_1|^p) \frac{x_1}{\|x_2\|} x_2 + (|\mu_2|^p - |\mu_1|^p) \frac{x_1}{\|y_2\|} y_2 \right) \\ &\quad + (1-\theta)(|\bar{\lambda}_2|^p - |\bar{\lambda}_1|^p) \frac{x_1 x_2 + x_1 y_2}{\|x_2 + y_2\|} \\ &= \theta \left((|\lambda_2|^p - |\lambda_1|^p) \frac{x_1}{\|x_2\|} x_2 + (|\mu_2|^p - |\mu_1|^p) \frac{y_1}{\|y_2\|} x_2 \right) \\ &\quad + (1-\theta)(|\bar{\lambda}_2|^p - |\bar{\lambda}_1|^p) \frac{x_1 x_2 + y_1 x_2}{\|x_2 + y_2\|}. \end{aligned} \tag{3.9}$$

If $x_1 + y_1 > 0$, from (3.9) we have

$$\begin{aligned} 2x_1 s_2 &= \theta \left((|\lambda_2|^p - |\lambda_1|^p) \frac{x_1}{\|x_2\|} + (|\mu_2|^p - |\mu_1|^p) \frac{y_1}{\|y_2\|} \right) x_2 \\ &\quad + (1-\theta)(|\bar{\lambda}_2|^p - |\bar{\lambda}_1|^p) \frac{x_1 + y_1}{\|x_2 + y_2\|} x_2 \\ &= 2^p \theta \left(|x_1|^p + |y_1|^p\right) x_2 + 2^p (1-\theta)|x_1 + y_1|^p x_2 \\ &= 2\|s_2\| x_2; \end{aligned}$$

and if $x_1 + y_1 < 0$, similar analysis will yield $x_1 s_2 = \|s_2\| x_2$. Similarly, $x_1 s_2 = \|s_2\| x_2$ can be obtained under each of the following cases: (a) $x_1 = \|x_2\|$ and $y_1 = \|y_2\|$; (b) $x_1 = -\|x_2\|$ and $y_1 = \|y_2\|$; (c) $x_1 = -\|x_2\|$ and $y_1 = -\|y_2\|$. So, we obtain that the second equality of (3.8) holds. Furthermore, $x_2^T s_2 = x_2^T x_2 \|s_2\|/x_1 = x_1 \|s_2\|$, i.e., the first equality of (3.8) holds. The rest equalities of (3.8) can be obtained directly from the symmetry of x and y in s.

Thus, **Part II** holds from (i)-(v).

By combining **Part I** with **Part II**, we complete the proof. □

Unless stated otherwise, we always assume in the following that $g(\alpha) := |\alpha|^p$ $(\alpha \in \Re)$ for all $p > 1$. Let $\text{sgn}(\alpha)$ denote the sign function. By [12, Proposition 5.2], we have

$$\nabla \bar{g}(x) = p \cdot \text{sgn}(x_1)|x_1|^{p-1} I \quad \text{if } x_2 = 0, \tag{3.10}$$

and

$$\nabla \bar{g}(x) = \begin{bmatrix} b(x) & c(x)\hat{x}_2^T \\ c(x)\hat{x}_2 & a(x)I + (b(x) - a(x))\hat{x}_2 \hat{x}_2^T \end{bmatrix} \quad \text{if } x_2 \neq 0, \tag{3.11}$$

where

$$\begin{array}{l}\hat{x}_2=\frac{x_2}{\|x_2\|},\quad a(x)=\frac{|\lambda_2(x)|^p-|\lambda_1(x)|^p}{\lambda_2(x)-\lambda_1(x)},\\ b(x)=\frac{p[\mathrm{sgn}(\lambda_2(x))|\lambda_2(x)|^{p-1}+\mathrm{sgn}(\lambda_1(x))|\lambda_1(x)|^{p-1}]}{2},\\ c(x)=\frac{p[\mathrm{sgn}(\lambda_2(x))|\lambda_2(x)|^{p-1}-\mathrm{sgn}(\lambda_1(x))|\lambda_1(x)|^{p-1}]}{2}.\end{array} \tag{3.12}$$

Through a similar discussion as the one given in [25, Page 10], we have that $\nabla\bar{g}(t)\in\mathcal{S}^n_{++}$ for any (x,y) satisfying $s\in\mathrm{int}\,\mathcal{K}^n$, and furthermore, for $q>1$ satisfying $\frac{1}{p}+\frac{1}{q}=1$,

$$\nabla\bar{g}(t)^{-1}=\frac{1}{2p}\begin{bmatrix}\frac{1}{(\lambda_2(s))^{1/q}}+\frac{1}{(\lambda_1(s))^{1/q}} & \frac{\hat{s}_2^T}{(\lambda_2(s))^{1/q}}-\frac{\hat{s}_2^T}{(\lambda_1(s))^{1/q}}\\ \frac{\hat{s}_2}{(\lambda_2(s))^{1/q}}-\frac{\hat{s}_2}{(\lambda_1(s))^{1/q}} & \frac{\hat{s}_2\hat{s}_2^T}{(\lambda_2(s))^{1/q}}+\frac{\hat{s}_2\hat{s}_2^T}{(\lambda_1(s))^{1/q}}+\frac{2p(I-\hat{s}_2\hat{s}_2^T)}{a(t)}\end{bmatrix}. \tag{3.13}$$

From the above analysis, the following lemma can be easily obtained.

Lemma 3.3. *The function t defined by (3.5) is continuously differentiable at any $(x,y)\in\Re^n\times\Re^n$ satisfying $s\in$ int $\mathcal{K}^n$, and furthermore,*

$$\begin{array}{r}\nabla_x t(x,y)=(\theta\nabla\bar{g}(x)+(1-\theta)\nabla\bar{g}(x+y))\nabla\bar{g}(t)^{-1};\\ \nabla_y t(x,y)=(\theta\nabla\bar{g}(y)+(1-\theta)\nabla\bar{g}(x+y))\nabla\bar{g}(t)^{-1},\end{array}$$

where $\bar{g}$ is the Löwner function induced by the scalar function g.

4 Smoothness of Merit Function $\Psi_{\theta p}$

For $x,y\in\Re^n$, and $\theta\in(0,1), p\in(1,+\infty)$, we define

$$\psi_{\theta p}(x,y):=\frac{1}{2}\|\phi_{\theta p}(x,y)\|^2\quad\text{and}\quad\Psi_{\theta p}(x):=\psi_{\theta p}(x,F(x)), \tag{4.1}$$

where F is given in (1.1). Then, $\Psi_{\theta p}$ is a merit function of (1.1). In this section, we investigate the smoothness of the functions $\Psi_{\theta p}$ defined by (4.1).

Theorem 4.1. *The function $\psi_{\theta p}$ defined by (4.1) is differentiable everywhere when $\theta\in(0,1)$ and $p\in(1,2)$. Furthermore, the following results hold.*

(i) If $(x,y)=(0,0)$, then $\nabla_x\psi_{\theta p}(x,y)=\nabla_y\psi_{\theta p}(x,y)=0$.

(ii) If $s\in$ int $\mathcal{K}^n$, then

$$\begin{array}{l}\nabla_x\psi_{\theta p}(x,y)=\left((\theta\nabla\bar{g}(x)+(1-\theta)\nabla\bar{g}(x+y))\nabla\bar{g}(t)^{-1}-I\right)\phi_{\theta p}(x,y);\\ \nabla_y\psi_{\theta p}(x,y)=\left((\theta\nabla\bar{g}(y)+(1-\theta)\nabla\bar{g}(x+y))\nabla\bar{g}(t)^{-1}-I\right)\phi_{\theta p}(x,y),\end{array} \tag{4.2}$$

with $\nabla\bar{g}$ given by (3.10)-(3.12) and $\nabla\bar{g}^{-1}$ given by (3.13).

(iii) If $s\notin$ int $\mathcal{K}^n$, then

$$\nabla_x\psi_{\theta p}(x,y)=D_1(x,y)\phi_{\theta p}(x,y);\quad\nabla_y\psi_{\theta p}(x,y)=D_2(x,y)\phi_{\theta p}(x,y), \tag{4.3}$$

with

$$\begin{array}{rcl}D_1(x,y)&:=&\dfrac{\theta\mathrm{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{1/q}}-1;\\ D_2(x,y)&:=&\dfrac{\theta\mathrm{sgn}(y_1)|y_1|^{p-1}+(1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{1/q}}-1.\end{array}$$

Proof. We divide the proof into three parts, i.e., **Part I**, **Part II**, and **Part III**, to show the results in (i), (ii), and (iii) listed in this theorem, respectively. In **Part III**, we first show the differentiability of $\psi_{\theta p}$ in **Step 1** and **Step 2**, and then derive the formulae given in case (iii) in **Step 3**.

Part I In this part, we prove the results given in the case (i). Obviously, $\phi_{\theta p}(x,y)=0$. For any $(h,k)\in\Re^n\times\Re^n$, let l_1 and l_2 ($l_1\le l_2$) denote the spectral values of $s(h,k)$, and c_1 and c_2 denote the corresponding spectral vectors. Then,

$$\begin{aligned}\|\phi_{\theta p}(h,k)\| &= \|l_1^{\frac{1}{p}}c_1+l_2^{\frac{1}{p}}c_2-h-k\|\le\|l_1^{\frac{1}{p}}c_1\|+\|l_2^{\frac{1}{p}}c_2\|+\|h\|+\|k\|\\ &\le 2l_2^{\frac{1}{p}}\frac{1}{\sqrt{2}}+\|h\|+\|k\|=\sqrt{2}l_2^{\frac{1}{p}}+\|h\|+\|k\|,\end{aligned}$$

while

$$\begin{aligned}l_2 &= \theta\left(\frac{|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p}{2}\right)+(1-\theta)\frac{|\bar\lambda_1|^p+|\bar\lambda_2|^p}{2}\\ &\quad+\|\theta\left(\frac{|\lambda_2|^p-|\lambda_1|^p}{2}\hat h_2+\frac{|\mu_2|^p-|\mu_1|^p}{2}\hat k_2\right)+(1-\theta)\frac{|\bar\lambda_2|^p-|\bar\lambda_1|^p}{2}(\widehat{h+k})_2\|\\ &\le \frac{\theta}{2}(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)+\frac{(1-\theta)}{2}(|\bar\lambda_1|^p+|\bar\lambda_2|^p)\\ &\quad+\frac{\theta}{2}(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)+\frac{(1-\theta)}{2}(|\bar\lambda_1|^p+|\bar\lambda_2|^p)\\ &= \theta(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)\\ &\quad+(1-\theta)(|h_1+k_1-\|h_2+k_2\||^p+|h_1+k_1+\|h_2+k_2\||^p)\\ &\le \theta(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)+2^{p+1}(1-\theta)(|h_1+k_1|^p+\|h_2+k_2\|^p)\\ &\le \theta(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)+2^{2p+1}(1-\theta)(|h_1|^p+|k_1|^p+\|h_2\|^p+\|k_2\|^p)\\ &\le \theta(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)+2^{2p+3}(1-\theta)(|\lambda_1|^p+|\mu_1|^p+|\lambda_2|^p+|\mu_2|^p)\\ &= \kappa_0(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p),\end{aligned}$$

where the notations are similarly defined as (2.1), (3.1)-(3.3) with h,k instead of x,y; the second inequality follows from the fact that $|a+b|^p\le||a|+|b||^p\le(2\max\{|a|,|b|\})^p\le 2^p(|a|^p+|b|^p)$ for any $a,b\in\Re$ and $p\in(1,2)$; the third inequality from the above fact and that $\|h_2+k_2\|^p\le(\|h_2\|+\|k_2\|)^p$; the last inequality from the above fact and $|h_1|^p\le(\max\{|\lambda_1|,|\lambda_2|\})^p$; and κ_0 used in the last equality is given by $\kappa_0=2^{2p+3}+(1-2^{2p+3})\theta$. Furthermore,

$$\begin{aligned}l_2^{\frac{1}{p}} &\le (\kappa_0(|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p))^{\frac{1}{p}}=\kappa_0^{\frac{1}{p}}\left((|\lambda_1|^p+|\lambda_2|^p+|\mu_1|^p+|\mu_2|^p)\right)^{\frac{1}{p}}\\ &\le \kappa_0^{\frac{1}{p}}(|\lambda_1|+|\lambda_2|+|\mu_1|+|\mu_2|)\le\kappa_0^{\frac{1}{p}}\sqrt{2}(\|h\|+\|k\|).\end{aligned}$$

Hence, $\psi_{\theta p}(h,k)=\frac{1}{2}\|\phi_{\theta p}(h,k)\|^2=O(\|h\|^2+\|k\|^2)$, which demonstrates that $\psi_{\theta p}$ is differentiable at $(0,0)$ with the gradient $(0,0)$.

Part II In this part, we prove the results given in the case (ii). From Lemma 3.3 and the definition of $\psi_{\theta p}$, it is easy to obtain that $\psi_{\theta p}$ is continuously differentiable in this case with the gradient formulae given by (4.2).

Part III In this part, we prove the results given in the case (iii). It is sufficient to consider the case of $(x,y)\neq(0,0)\in\Re^n\times\Re^n$ and $s(x,y)\notin\text{int }\mathcal{K}^n$ (hence $s_2\neq 0$ by using Lemma 3.2 when $(x,y)\neq(0,0)$). Denote $s=s(x,y)=(s_1,s_2)=\eta_1c_1+\eta_2c_2$ with η_1 and

η_2 being the spectral values of s and c_1 and c_2 being the corresponding spectral vectors. It follows from the definition of $\psi_{\theta p}(x,y)$ that

$$2\psi_{\theta p}(x,y) = \|\phi_{\theta p}(x,y)\|^2 = \|s^{\frac{1}{p}} - x - y\|^2 = \|s^{\frac{1}{p}}\|^2 + \|x+y\|^2 - 2\langle s^{\frac{1}{p}}, x+y\rangle,$$

while

$$\|s^{\frac{1}{p}}\|^2 = \frac{1}{2}\eta_1^{\frac{2}{p}} + \frac{1}{2}\eta_2^{\frac{2}{p}}; \tag{4.4}$$

$$\begin{aligned} 2\langle s^{\frac{1}{p}}, x+y\rangle &= 2\left\langle \begin{pmatrix} \frac{\eta_1^{\frac{1}{p}}+\eta_2^{\frac{1}{p}}}{2} \\ \frac{\eta_2^{\frac{1}{p}}-\eta_1^{\frac{1}{p}}}{2}\hat{s}_2 \end{pmatrix}, \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix} \right\rangle \\ &= \eta_1^{\frac{1}{p}}(x_1+y_1 - \frac{s_2^T(x_2+y_2)}{\|s_2\|}) + \eta_2^{\frac{1}{p}}(x_1+y_1+\frac{s_2^T(x_2+y_2)}{\|s_2\|}). \end{aligned} \tag{4.5}$$

Hence, in order to prove that $\psi_{\theta p}(\cdot,\cdot)$ is differentiable at (x,y), it is sufficient to prove that both functions given in (4.4) and (4.5) are differentiable at (x,y). We complete the proof by the following three steps.

Step 1 This step is dedicated to the differentiability of s_1 and s_2, which are viewed as functions of x and y. For any $(\tilde{x},\tilde{y}) \in \Re^n \times \Re^n$, with $\tilde{x}_2 \neq 0, \tilde{y}_2 \neq 0$ and $\tilde{x}_2+\tilde{y}_2 \neq 0$, sufficiently close to (x,y), we define

$$\begin{aligned} \tilde{s}_1 &= \theta\left(\tfrac{|\tilde{\lambda}_1|^p+|\tilde{\lambda}_2|^p+|\tilde{\mu}_1|^p+|\tilde{\mu}_2|^p}{2}\right) + (1-\theta)\tfrac{|\tilde{\tilde{\lambda}}_1|^p+|\tilde{\tilde{\lambda}}_2|^p}{2}; \\ \tilde{s}_2 &= \theta\left(\tfrac{|\tilde{\lambda}_2|^p-|\tilde{\lambda}_1|^p}{2}\tfrac{\tilde{x}_2}{\|\tilde{x}_2\|} + \tfrac{|\tilde{\mu}_2|^p-|\tilde{\mu}_1|^p}{2}\tfrac{\tilde{y}_2}{\|\tilde{y}_2\|}\right) + (1-\theta)\tfrac{|\tilde{\tilde{\lambda}}_2|^p-|\tilde{\tilde{\lambda}}_1|^p}{2}\tfrac{\tilde{x}_2+\tilde{y}_2}{\|\tilde{x}_2+\tilde{y}_2\|}, \end{aligned}$$

where the related notations are similar to those in the above analysis. Then, $\tilde{s}_1$ and $\tilde{s}_2$, viewed as functions of $(\tilde{x},\tilde{y})$, are differentiable at $(\tilde{x},\tilde{y}) = (x,y)$. Indeed, when $x_2 \neq 0, y_2 \neq 0$ and $x_2+y_2 \neq 0$, $\tilde{s}_1$ and $\tilde{s}_2$ are clearly differentiable at $(\tilde{x},\tilde{y}) = (x,y)$ since $\tilde{\lambda}_1$, $\tilde{\lambda}_2$, $\tilde{\mu}_1$, $\tilde{\mu}_2$, $\tilde{\tilde{\lambda}}_1$, $\tilde{\tilde{\lambda}}_2$, $\|\tilde{x}_2\|$, $\|\tilde{y}_2\|$, $\|\tilde{x}_2+\tilde{y}_2\|$ are differentiable at $(\tilde{x},\tilde{y}) = (x,y)$ and $|\alpha|^p$ is continuously differentiable on $\Re$ for $p>1$. When $x_2 = 0$ and $y_2 \neq 0$ (hence $x_2+y_2 \neq 0$),

$$\frac{|\tilde{\mu}_1|^p+|\tilde{\mu}_2|^p}{2},\quad \frac{|\tilde{\tilde{\lambda}}_1|^p+|\tilde{\tilde{\lambda}}_2|^p}{2},\quad \frac{|\tilde{\mu}_2|^p-|\tilde{\mu}_1|^p}{2}\frac{\tilde{y}_2}{\|\tilde{y}_2\|},\quad \text{and}\quad \frac{|\tilde{\tilde{\lambda}}_2|^p-|\tilde{\tilde{\lambda}}_1|^p}{2}\frac{\tilde{x}_2+\tilde{y}_2}{\|\tilde{x}_2+\tilde{y}_2\|}$$

are differentiable at $(\tilde{x},\tilde{y}) = (x,y)$. By using the continuous differentiability of $|\alpha|^p$ and the Mean-Valued Theorem, we obtain that

$$\begin{aligned} |\tilde{\lambda}_1|^p &= |\lambda_1|^p + p\cdot\mathrm{sgn}[(1-\alpha_1)\tilde{\lambda}_1+\alpha_1\lambda_1]|\alpha_1\lambda_1+(1-\alpha_1)\tilde{\lambda}_1|^{p-1}(\tilde{\lambda}_1-\lambda_1) \\ &= p\cdot\mathrm{sgn}((1-\alpha_1)\tilde{\lambda}_1)|(1-\alpha_1)\tilde{\lambda}_1|^{p-1}(\tilde{x}_1-\|\tilde{x}_2\|) \quad (0<\alpha_1<1), \\ |\tilde{\lambda}_2|^p &= p\cdot\mathrm{sgn}((1-\alpha_2)\tilde{\lambda}_2)|(1-\alpha_2)\tilde{\lambda}_2|^{p-1}(\tilde{x}_1+\|\tilde{x}_2\|) \quad (0<\alpha_2<1), \\ |\tilde{\lambda}_2|^p &= |\tilde{\lambda}_1|^p + 2p\cdot\mathrm{sgn}[\alpha_3\tilde{\lambda}_1+(1-\alpha_3)\tilde{\lambda}_2]|\alpha_3\tilde{\lambda}_1+(1-\alpha_3)\tilde{\lambda}_2|^{p-1}\|\tilde{x}_2\| \quad (0<\alpha_3<1). \end{aligned}$$

Since $x=0$ by Lemma 3.2 in this case (i.e., $x_2 = 0$), it follows that $\tilde{\lambda}_2, \tilde{\lambda}_1 \to 0$ when $(\tilde{x},\tilde{y})$ tends to (x,y), and hence, the above equalities imply that

$$|\tilde{\lambda}_1|^p = o(\|\tilde{x}\|),\quad |\tilde{\lambda}_2|^p = o(\|\tilde{x}\|),\quad \text{and}\quad \frac{|\tilde{\lambda}_2|^p-|\tilde{\lambda}_1|^p}{2}\frac{\tilde{x}_2}{\|\tilde{x}_2\|} = o(\|\tilde{x}\|).$$

Therefore, $|\tilde{\lambda}_1|^p + |\tilde{\lambda}_2|^p$ and $\frac{|\tilde{\lambda}_2|^p - |\tilde{\lambda}_1|^p}{2}\frac{\tilde{x}_2}{\|\tilde{x}_2\|}$ are differentiable at $(\tilde{x}, \tilde{y}) = (x, y)$. Similarly, $\tilde{s}_1$ and $\tilde{s}_2$ are differentiable at $(\tilde{x}, \tilde{y}) = (x, y)$ in each of the following cases: (a) $y_2 = 0$ and $x_2 \neq 0$ (hence $x_2 + y_2 \neq 0$); (b) $x_2 \neq 0$, $y_2 \neq 0$ and $x_2 + y_2 = 0$.

Step 2 In this step, we prove the differentiability of $\psi_{\theta p}$ given in the case (iii). It follows from **Step 1** that the second items in the right-hand side of both (4.4) and (4.5) are differentiable at $(\tilde{x}, \tilde{y}) = (x, y)$. Similarly, $\tilde{\eta}_1$ (defined similarly), viewed as a function of $(\tilde{x}, \tilde{y})$, is also differentiable at $(\tilde{x}, \tilde{y}) = (x, y)$ since $\tilde{s}_2 \neq 0$. Since $\tilde{s} \notin \text{int } \mathcal{K}^n$, we have $\eta_1 = 0$. Using the first Taylor's expansion of $\tilde{\eta}_1$ at (x, y) we obtain that $\tilde{\eta}_1 = O(\|\tilde{x} - x\| + \|\tilde{y} - y\|)$, so, $\tilde{\eta}_1^{\frac{2}{p}} = O[(\|\tilde{x} - x\| + \|\tilde{y} - y\|)^{\frac{2}{p}}]$. Thus, the first item in the right-hand side of (4.4) is differentiable at $(\tilde{x}, \tilde{y}) = (x, y)$ since $2 > p > 1$. Furthermore, $\tilde{x}_1 + \tilde{y}_1 - \frac{\tilde{s}_2^T(\tilde{x}_2 + \tilde{y}_2)}{\|\tilde{s}_2\|}$ is differentiable at $(\tilde{x}, \tilde{y}) = (x, y)$ by **Step 1**. From the fact that $s_2 \neq 0$, $s \notin \text{int } \mathcal{K}^n$ and (3.8), we get $x_1 + y_1 - \frac{s_2^T(x_2 + y_2)}{\|s_2\|} = 0$. So, $\tilde{x}_1 + \tilde{y}_1 - \frac{\tilde{s}_2^T(\tilde{x}_2 + \tilde{y}_2)}{\|\tilde{s}_2\|} = O(\|\tilde{x} - x\| + \|\tilde{y} - y\|)$, and hence,

$$\begin{aligned} \tilde{\eta}_1^{\frac{1}{p}} \left(\tilde{x}_1 + \tilde{y}_1 - \frac{\tilde{s}_2^T(\tilde{x}_2 + \tilde{y}_2)}{\|\tilde{s}_2\|} \right) &= O\left((\|\tilde{x} - x\| + \|\tilde{y} - y\|)^{1 + \frac{1}{p}} \right) \\ &= o(\|\tilde{x} - x\| + \|\tilde{y} - y\|). \end{aligned}$$

This implies that the first item in the right-hand side of (4.5) is differentiable at $(\tilde{x}, \tilde{y}) = (x, y)$ with the gradient being zero. Thus, the differentiability of $\psi_{\theta p}$ is proved in this case.

Step 3 In this step, we derive the derivative formulae given in the case (iii). Without loss of generality, we assume that $x_2 \neq 0, y_2 \neq 0, x_2 + y_2 \neq 0$. It follows from the definition of $\psi_{\theta p}$ and **Step 2** that the gradient of $2\psi_{\theta p}$ is the sum of the gradients of $\|x + y\|^2$, the right-hand side of (4.4), and the second item in the right-hand side of (4.5). The gradient of $\|\tilde{x} + \tilde{y}\|^2$ with respect to $\tilde{x}$ evaluated at $(\tilde{x}, \tilde{y}) = (x, y)$ is $2(x + y)$; and the gradients of $\tilde{s}_1$ and $\|\tilde{s}_2\|$ with respect to $\tilde{x}$ evaluated at $(\tilde{x}, \tilde{y}) = (x, y)$ are

$$\frac{p}{2} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \quad \text{and} \quad \frac{p}{2\|s_2\|} \begin{pmatrix} m_3 \\ m_4 \end{pmatrix}$$

with

$$\begin{aligned} m_1 &= \theta(\text{sgn}(\lambda_2)|\lambda_2|^{p-1} + \text{sgn}(\lambda_1)|\lambda_1|^{p-1}) + (1-\theta)(\text{sgn}(\bar{\lambda}_2)|\bar{\lambda}_2|^{p-1} + \text{sgn}(\bar{\lambda}_1)|\bar{\lambda}_1|^{p-1}); \\ m_2 &= \theta(\text{sgn}(\lambda_2)|\lambda_2|^{p-1} - \text{sgn}(\lambda_1)|\lambda_1|^{p-1})\frac{x_2}{\|x_2\|} \\ &\quad + (1-\theta)(\text{sgn}(\bar{\lambda}_2)|\bar{\lambda}_2|^{p-1} - \text{sgn}(\bar{\lambda}_1)|\bar{\lambda}_1|^{p-1})\frac{x_2 + y_2}{\|x_2 + y_2\|}; \\ m_3 &= \theta(\text{sgn}(\lambda_2)|\lambda_2|^{p-1} - \text{sgn}(\lambda_1)|\lambda_1|^{p-1})\frac{x_2^T s_2}{\|x_2\|} \\ &\quad + (1-\theta)(\text{sgn}(\bar{\lambda}_2)|\bar{\lambda}_2|^{p-1} - \text{sgn}(\bar{\lambda}_1)|\bar{\lambda}_1|^{p-1})\frac{(x_2 + y_2)^T s_2}{\|x_2 + y_2\|}; \\ m_4 &= \theta((\text{sgn}(\lambda_2)|\lambda_2|^{p-1} + \text{sgn}(\lambda_1)|\lambda_1|^{p-1})\frac{x_2 x_2^T}{\|x_2\|^2} + (|\lambda_2|^p - |\lambda_1|^p)\frac{\|x_2\|^2 - x_2^T x_2}{\|x_2\|^3})s_2 \\ &\quad + (1-\theta)(\text{sgn}(\bar{\lambda}_2)|\bar{\lambda}_2|^{p-1} + \text{sgn}(\bar{\lambda}_1)|\bar{\lambda}_1|^{p-1})\frac{(x_2 + y_2)(x_2 + y_2)^T}{\|x_2 + y_2\|^2}s_2 \\ &\quad + (1-\theta)(|\bar{\lambda}_2|^p - |\bar{\lambda}_1|^p)\frac{\|x_2 + y_2\|^2 - (x_2 + y_2)^T(x_2 + y_2)}{\|(x_2 + y_2)\|^3}s_2. \end{aligned}$$

By using the equalities $x_1^2 = \|x_2\|^2$, $x_2^T \frac{s_2}{\|s_2\|} = x_1$, $x_1 \frac{s_2}{\|s_2\|} = x_2$, $y_1^2 = \|y_2\|^2$, $y_2^T \frac{s_2}{\|s_2\|} = y_1$, and $y_1 \frac{s_2}{\|s_2\|} = y_2$ in Lemma 3.2, both the above gradient formulae could be simplified as

$$\begin{aligned}&\begin{pmatrix} 2^{p-2}p(\theta\mathrm{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}) \\ 2^{p-2}p(\theta\mathrm{sgn}(x_1)|x_1|^{p-1}\frac{x_2}{\|x_2\|} + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1})\frac{x_2+y_2}{\|x_2+y_2\|} \end{pmatrix} \\ &\quad = 2^{p-2}p(\theta\mathrm{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}) \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix}.\end{aligned}$$

Hence, the gradient of $\tilde{\eta}_2 = \tilde{s}_1 + \|\tilde{s_2}\|$ with respect to $\tilde{x}$ evaluated at $(\tilde{x}, \tilde{y}) = (x, y)$ is

$$2^{p-1}p(\theta\mathrm{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}) \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix}.$$

Furthermore, by using the product and quotient rules of differentiation, the gradient of $x_1 + y_1 + \frac{s_2^T(x_2+y_2)}{\|s_2\|}$ with respect to $\tilde{x}$ evaluated at $(\tilde{x}, \tilde{y}) = (x, y)$ is

$$\begin{aligned}&\begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} + \frac{\nabla_{x_2} s_2 (x_2+y_2)\|s_2\| - \nabla_{x_2} s_2 \frac{s_2}{\|s_2\|} s_2^T(x_2+y_2)}{\|s_2\|^2} \end{pmatrix} \\ &\quad = \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} + \nabla_{x_2} s_2 \frac{(x_2+y_2)\|s_2\| - \frac{s_2}{\|s_2\|} s_2^T(x_2+y_2)}{\|s_2\|^2} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix},\end{aligned}$$

where the last equality follows from Lemma 3.2 (i.e., we use the fact that $\frac{s_2^T(x_2+y_2)}{\|s_2\|} = x_1 + y_1$ and $(x_1+y_1)\frac{s_2}{\|s_2\|} = x_2 + y_2$). Hence, the gradients of the second items in the right-hand sides of (4.4) and (4.5) with respect to $\tilde{x}$ evaluated at $(\tilde{x}, \tilde{y}) = (x, y)$ are respectively given by

$$(\eta_2)^{\frac{2}{p}-1} 2^{p-1}(\theta\mathrm{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}) \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix}$$

and

$$\begin{aligned}&(x_1+y_1)(\eta_2)^{\frac{1}{p}-1} 2^{p}(\theta\mathrm{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}) \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix} \\ &\quad + (\eta_2)^{\frac{1}{p}} \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix}.\end{aligned}$$

It is easy to see that the gradient of the first item in the right-hand side of (4.4) with respect to $\tilde{x}$ evaluated at $(\tilde{x}, \tilde{y}) = (x, y)$ is zero when $1 < p < 2$. Hence,

$$\begin{aligned} 2\nabla_x \psi_{\theta p} &= (\eta_2)^{\frac{2}{p}-1} 2^{p-1}(\theta\mathrm{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}) \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix} \\ &\quad -(x_1+y_1)(\eta_2)^{\frac{1}{p}-1} 2^{p}(\theta\mathrm{sgn}(x_1)|x_1|^{p-1} - (\eta_2)^{\frac{1}{p}} \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix} + 2(x+y) \\ &\quad +(1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}) \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix}, \quad 1 < p < 2, \quad \forall \theta \in (0,1). \end{aligned}$$

From $s \notin \text{int}\,\mathcal{K}^n$, we have $\eta_1 = 0$, and hence,

$$\phi_{\theta p}(x,y) = t(x,y) - x - y = \frac{1}{2}(\eta_2)^{\frac{1}{p}} \begin{pmatrix} 1 \\ \frac{s_2}{\|s_2\|} \end{pmatrix} - (x+y).$$

The last two equalities, together with $x_1 \frac{s_2}{\|s_2\|} = x_2$ and $y_1 \frac{s_2}{\|s_2\|} = y_2$ by Lemma 3.2, yield that

$$\begin{aligned} 2\nabla_x \psi_{\theta p} &= (\eta_2)^{\frac{1}{p}-1} 2^p (\theta \text{sgn}(x_1)|x_1|^{p-1} \\ &\quad +(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1})(\phi_{\theta p}(x,y)+x+y) - 2\phi_{\theta p}(x,y) \\ &\quad -(\eta_2)^{\frac{1}{p}-1} 2^p (\theta \text{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1})(x+y) \\ &= (\eta_2)^{\frac{1}{p}-1} 2^p (\theta \text{sgn}(x_1)|x_1|^{p-1} \\ &\quad +(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1})(\phi_{\theta p}(x,y)) - 2\phi_{\theta p}(x,y) \\ &= 2\left(\frac{\theta \text{sgn}(x_1)|x_1|^{p-1} + (1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p + |y_1|^p) + (1-\theta)|x_1+y_1|^p)^{1/q}} - 1 \right) \phi_{\theta p}(x,y), \\ &\quad (1 < p < 2, \quad \forall \theta \in (0,1)) \end{aligned}$$

where the last equality follows from $\eta_2 = 2s_1 = 2^p(\theta(|x_1|^p + |y_1|^p) + (1-\theta)|x_1+y_1|^p)$. Hence, the first equality of (4.3) follows. In addition, the second equality of (4.3) follows immediately from the symmetry of x and y in s.

Combining **Step 1**, **Step 2** and **Step 3**, the proof of **Part III** is complete.

Combining **Part I**, **Part II** and **Part III**, the theorem is proved. □

Remark 4.2. It is hard for us to prove that the function $\psi_{\theta p}$ defined by (4.1) is differentiable for $\theta \in (0,1)$ and $p > 2$ just following the proof of Theorem 4.1, since $\alpha^{\frac{2}{p}}$ $(p > 2)$ is not differentiable at $\alpha = 0$ (hence, it is hard to get the differentiability of the function given in Eq. (4.4)).

In the following, we will show the continuity of the gradient of $\psi_{\theta p}$ defined by (4.1), i.e., the smoothness of $\psi_{\theta p}$, when $\theta \in (0,1)$ and $p \in (1,2)$.

Lemma 4.3. *Suppose that $\theta \in (0,1)$ and $p \in (1,2)$. Let s and t be defined by (3.4) and (3.5), respectively. Then, there exists a constant $\hat{\kappa} > 0$ such that for all $x, y \in \Re^n$ satisfying $s \in$ int $\mathcal{K}^n$,*

$$\|L_{\theta|x|^{p-1}} L^{-1}_{t^{p-1}}\|_F \le \hat{\kappa}, \quad \|L_{\theta|y|^{p-1}} L^{-1}_{t^{p-1}}\|_F \le \hat{\kappa}, \quad \|L_{(1-\theta)|x+y|^{p-1}} L^{-1}_{t^{p-1}}\|_F \le \hat{\kappa},$$

where $\|\cdot\|_F$ denotes the matrix Frobenius norm on $\Re^{n\times n}$.

Proof. The proof is similar to that in [25, Lemma 4.1], we omit it here. □

Under the assumptions of Lemma 4.3, it follows from Lemma 4.3 and $L_{\theta|x|^{p-1}} = \theta L_{|x|^{p-1}}$ that $\|L_{|x|^{p-1}} L^{-1}_{t^{p-1}}\|_F \le \frac{\hat{\kappa}}{\theta}$. Thus, we have

$$\lambda_2(s)^{\frac{1}{q}} \ge \left(\frac{\theta(|\lambda_1|^p + |\lambda_2|^p)}{2} \right)^{\frac{1}{q}} \ge \left(\theta \frac{(|\lambda_1|^2 + |\lambda_2|^2)^{\frac{p}{2}}}{2} \right)^{\frac{1}{q}} = \theta^{\frac{1}{q}} 2^{\frac{p-2}{2q}} \|x\|^{\frac{p}{q}}.$$

Furthermore, by a similar analysis as the one in [25, Remark 4.1], it follows that for all $x, y \in \Re^n$ satisfying $s \in \text{int}\,\mathcal{K}^n$,

$$\frac{|\lambda_2|^{p-1}(1 - \hat{x}_2^T \hat{s}_2) + |\lambda_1|^{p-1}(1 + \hat{x}_2^T \hat{s}_2)}{\lambda_1(s)^{\frac{1}{q}}} = O(1),$$

and

$$\frac{|\lambda_2|^{2p-2}(1-\hat{x}_2^T\hat{s}_2)+|\lambda_1|^{2p-2}(1+\hat{x}_2^T\hat{s}_2)}{\lambda_1(s)^{\frac{2}{q}}}=O(1).$$

Lemma 4.4. *Suppose that $\theta\in(0,1)$ and $p\in(1,2)$. Let s and t be defined by (3.4) and (3.5), respectively, and $\nabla\bar{g}(x)$ be given by (3.10)-(3.12). Then, there exists a constant $\kappa>0$ such that for all $x,y\in\Re^n$ satisfying $s\in$ int $\mathcal{K}^n$,*

$$\|(\theta\nabla\bar{g}(x)+(1-\theta)\nabla\bar{g}(x+y))\nabla\bar{g}(t)^{-1}\|_F\leq\kappa,$$
$$\|(\theta\nabla\bar{g}(y)+(1-\theta)\nabla\bar{g}(x+y))\nabla\bar{g}(t)^{-1}\|_F\leq\kappa.$$

Proof. The results could be similarly proved as those in [25, Lemma 4.2]. We omit it here. □

In the following, we establish one main result of this section, i.e., the continuity of the gradient function of $\psi_{\theta p}$ defined by (4.1).

Theorem 4.5. *The function $\psi_{\theta p}$ defined by (4.1) is smooth everywhere on $\Re^n\times\Re^n$ for all $\theta\in(0,1)$ and $p\in(1,2)$.*

Proof. It is sufficient to prove $\nabla_x\psi_{\theta p}$ is continuous on $\Re^n\times\Re^n$ for all $\theta\in(0,1)$ and $p\in(1,2)$ by Theorem 4.1 and the symmetry of x and y in $\nabla\psi_{\theta p}$. We divide the proof into the following three parts.

Part I For any $(x,y)\in\Re^n\times\Re^n$ satisfying $t\in$ int $\mathcal{K}^n$, the result follows form **Part II** in the proof of Theorem 4.1 immediately.

Part II If $(x,y)=(0,0)$, then $\nabla_x\psi_{\theta p}(0,0)=0$ by Theorem 4.1. Let $(h,k)\in\Re^n\times\Re^n$. On one hand, if $t(h,k)\in$ int $\mathcal{K}^n$, then by Theorem 4.1,

$$\nabla_x\psi_{\theta p}(h,k)=\left((\theta\nabla\bar{g}(h)+(1-\theta)\nabla\bar{g}(h+k))\nabla\bar{g}(t(h,k))^{-1}-I\right)\phi_{\theta p}(h,k).$$

Thus, by using Lemma 4.4, the continuity of $\phi_{\theta p}$, and $\phi_{\theta p}(0,0)=0$ by Theorem 3.1, we have that $\nabla_x\psi_{\theta p}(h,k)\to 0$ as $(h,k)\to 0$. On the other hand, if $(h,k)\neq(0,0)$ and $t(h,k)\notin$ int $\mathcal{K}^n$, then by Theorem 4.1,

$$\nabla_x\psi_{\theta p}(h,k)=\left(\frac{\theta\mathrm{sgn}(h_1)|h_1|^{p-1}+(1-\theta)\mathrm{sgn}(h_1+k_1)|h_1+k_1|^{p-1}}{(\theta(|h_1|^p+|k_1|^p)+(1-\theta)|h_1+k_1|^p)^{1/q}}-1\right)\phi_{\theta p}(h,k).$$

Thus, by combining the continuity of $\phi_{\theta p}$ with $\phi_{\theta p}(0,0)=0$ and the uniform boundedness of the function

$$\frac{\theta\mathrm{sgn}(h_1)|h_1|^{p-1}+(1-\theta)\mathrm{sgn}(h_1+k_1)|h_1+k_1|^{p-1}}{(\theta(|h_1|^p+|k_1|^p)+(1-\theta)|h_1+k_1|^p)^{1/q}}-1,$$

we obtain that $\nabla_x\psi_{\theta p}(h,k)\to 0$ as $(h,k)\to 0$.

Part III For any $(x,y)\neq(0,0)\in\Re^n\times\Re^n$ satisfying $t\notin$ int $\mathcal{K}^n$, we consider the case of $(h,k)\to(x,y)$ with $(h,k)\in\Re^n\times\Re^n$. Since $(x,y)\neq(0,0)$, it follows that $(h,k)\neq(0,0)$ for any (h,k) sufficiently close to (x,y).

On one hand, if $(h,k)\neq(0,0)$ and $t(h,k)\notin$ int $\mathcal{K}^n$, then by Theorem 4.1,

$$\nabla_x\psi_{\theta p}(h,k)=\left(\frac{\theta\mathrm{sgn}(h_1)|h_1|^{p-1}+(1-\theta)\mathrm{sgn}(h_1+k_1)|h_1+k_1|^{p-1}}{(\theta(|h_1|^p+|k_1|^p)+(1-\theta)|h_1+k_1|^p)^{1/q}}-1\right)\phi_{\theta p}(h,k).$$

Thus, from the continuality of both $\phi_{\theta p}$ and

$$\frac{\theta \mathrm{sgn}(h_1)|h_1|^{p-1} + (1-\theta)\mathrm{sgn}(h_1+k_1)|h_1+k_1|^{p-1}}{(\theta(|h_1|^p + |k_1|^p) + (1-\theta)|h_1+k_1|^p)^{1/q}} - 1,$$

it follows that $\nabla_x \psi_{\theta p}(h,k) \to \nabla_x \psi_{\theta p}(x,y)$.

On the other hand, if $t(h,k) \in \mathrm{int}\, \mathcal{K}^n$, then

$$\begin{aligned} \nabla_x \psi_{\theta p}(h,k) &= ((\theta \nabla \bar{g}(h) + (1-\theta)\nabla \bar{g}(h+k))\nabla \bar{g}(t(h,k))^{-1} - I)\phi_{\theta p}(h,k) \\ &= (\theta \nabla \bar{g}(h) + (1-\theta)\nabla \bar{g}(h+k))\nabla \bar{g}(t(h,k))^{-1} t(h,k) - \phi_{\theta p}(h,k) \\ &\quad -(\theta \nabla \bar{g}(h) + (1-\theta)\nabla \bar{g}(h+k))\nabla \bar{g}(t(h,k))^{-1}(h+k). \end{aligned} \tag{4.6}$$

Since $s_2(x,y) \neq 0$ by Lemma 3.2, we have $s_2(h,k) \neq 0$ for any (h,k) sufficiently close to (x,y). Thus, from (3.13) and the spectral decomposition, we have

$$t(h,k) = \lambda_1(s(h,k))^{\frac{1}{p}} \begin{pmatrix} 1 \\ -\frac{s_2(h,k)}{\|s_2(h,k)\|} \end{pmatrix} + \lambda_2(s(h,k))^{\frac{1}{p}} \begin{pmatrix} 1 \\ \frac{s_2(h,k)}{\|s_2(h,k)\|} \end{pmatrix}.$$

Furthermore, we have

$$\nabla \bar{g}(t(h,k))^{-1} t(h,k) = \frac{1}{2p}\left(\lambda_1(s(h,k))^{\frac{1}{p}-\frac{1}{q}} + \lambda_2(s(h,k))^{\frac{1}{p}-\frac{1}{q}}\right) \begin{pmatrix} 1 \\ \frac{s_2(h,k)}{\|s_2(h,k)\|} \end{pmatrix}.$$

Since $s_2(x,y) \neq 0$, $\frac{1}{p} - \frac{1}{q} \geq 0$ and $t \geq 0$, it follows that s_2 and $t^{\frac{1}{p}-\frac{1}{q}}$ are continuous at (x,y). From Lemma 3.2, when $(h,k) \to (x,y)$, we have

$$\nabla \bar{g}(t(h,k))^{-1} t(h,k) \to \frac{1}{p} 2^{-\frac{p}{q}} (\theta(|x_1|^p + |y_1|^p) + (1-\theta)|x_1+y_1|^p)^{\frac{1}{p}-\frac{1}{q}} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix}.$$

Since $\nabla \bar{g}(\cdot)$ is continuous, it follows that $\nabla \bar{g}(h) \to \nabla \bar{g}(x)$ when $(h,k) \to (x,y)$. From Lemma 3.2 we have that $\min\{|\lambda_1|, |\lambda_2|\} = 0$ and $\min\{|\bar{\lambda}_1|, |\bar{\lambda}_2|\} = 0$, and hence,

$$\begin{aligned} &a(x) = 2^{p-1}\mathrm{sgn}(x_1)|x_1|^{p-1}, \quad b(x) = 2^{p-2}p \cdot \mathrm{sgn}(x_1)|x_1|^{p-1}, \quad c(x) = 2^{p-2}p|x_1|^{p-1}; \\ &a(x+y) = 2^{p-1}\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}, \\ &b(x+y) = 2^{p-2}p \cdot \mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1}, \quad c(x+y) = 2^{p-2}p|x_1+y_1|^{p-1}, \end{aligned}$$

where $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$ are defined by (3.12). So,

$$\begin{aligned} \nabla \bar{g}(x) &= 2^{p-2}p \cdot \mathrm{sgn}(x_1)|x_1|^{p-1} \begin{bmatrix} 1 & \frac{x_2^T}{x_1} \\ \frac{x_2}{x\ 1} & \frac{2}{p}I + (1-\frac{2}{p})\frac{x_2 x_2^T}{x_1^2} \end{bmatrix}; \\ \nabla \bar{g}(x+y) &= 2^{p-2}p \cdot \mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1} \begin{bmatrix} 1 & \frac{(x_2+y_2)^T}{x_1+y_1} \\ \frac{x_2+y_2}{x_1+y_1} & \frac{2}{p}I + (1-\frac{2}{p})\frac{(x_2+y_2)(x_2+y_2)^T}{(x_1+y_1)^2} \end{bmatrix}. \end{aligned}$$

Hence, when $(h,k) \to (x,y)$, we have

$$\begin{aligned} &(\theta \nabla \bar{g}(h) + (1-\theta)\nabla \bar{g}(h+k))\nabla \bar{g}(t(h,k))^{-1} t(h,k) \\ &\quad \to \frac{\theta \mathrm{sgn}(x_1)|x_1|^{p-1} m_5 + (1-\theta)\mathrm{sgn}(x_1+y_1)|x_1+y_1|^{p-1} m_6}{2(\theta(|x_1|^p + |y_1|^p) + (1-\theta)|x_1+y_1|^p)^{\frac{1}{q}-\frac{1}{p}}}, \end{aligned} \tag{4.7}$$

where

$$m_5 = \begin{bmatrix} 1 & \frac{x_2^T}{x_1} \\ \frac{x_2}{x_1} & \frac{2}{p}I + (1-\frac{2}{p})\frac{x_2x_2^T}{x_1^2} \end{bmatrix} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix};$$

$$m_6 = \begin{bmatrix} 1 & \frac{(x_2+y_2)^T}{x_1+y_1} \\ \frac{x_2+y_2}{x_1+y_1} & \frac{2}{p}I + (1-\frac{2}{p})\frac{(x_2+y_2)(x_2+y_2)^T}{(x_1+y_1)^2} \end{bmatrix} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix}.$$

Thus, the right-hand side of (4.7) can be simplified by Lemma 3.2 as

$$\begin{aligned} &\frac{\theta \text{sgn}(x_1)|x_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix} \\ &+\frac{(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix} \\ &= \frac{\theta \text{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix}. \end{aligned} \tag{4.8}$$

In addition,

$$\begin{aligned} \nabla_x \psi_{\theta p}(x,y) &= \left(\frac{\theta \text{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{1/q}} - 1\right)\phi_{\theta p}(x,y) \\ &= \frac{\theta \text{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix} \\ &\quad -\frac{\theta \text{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}}(x+y) \\ &\quad -\phi_{\theta p}(x,y), \end{aligned} \tag{4.9}$$

where the first equality follows from Theorem 4.1; and the second follows from

$$\begin{aligned} (\theta(|x|^p+|y|^p)+(1-\theta)|x+y|^p) &= \frac{(2s_1(x,y))^{\frac{1}{p}}}{2} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix} \\ &= (\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{\frac{1}{p}} \begin{pmatrix} 1 \\ \frac{s_2(x,y)}{\|s_2(x,y)\|} \end{pmatrix}. \end{aligned}$$

To complete the proof, by (4.6), (4.8), and (4.9), we only need to prove that

$$\begin{aligned} &(\theta\nabla\bar{g}(h)+(1-\theta)\nabla\bar{g}(h+k))\nabla\bar{g}(t(h,k))^{-1}(h+k) \\ &\quad \to \frac{\theta \text{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}}(x+y) \end{aligned} \tag{4.10}$$

when $(h,k) \to (x,y)$. Let

$$(l_1, l_2) := (\theta\nabla\bar{g}(h)+(1-\theta)\nabla\bar{g}(h+k))\nabla\bar{g}(t(h,k))^{-1}(h+k).$$

Then (4.10) reduces to

$$l_1 \to \frac{\theta \text{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}}(x_1+y_1); \tag{4.11}$$

$$l_2 \to \frac{\theta \text{sgn}(x_1)|x_1|^{p-1}+(1-\theta)\text{sgn}(x_1+y_1)|x_1+y_1|^{p-1}}{(\theta(|x_1|^p+|y_1|^p)+(1-\theta)|x_1+y_1|^p)^{-\frac{1}{p}+\frac{1}{q}}}(x_2+y_2). \tag{4.12}$$

The proofs of (4.11) and (4.12) are omitted here. Hence, the proof of **Part III** is complete.

Combining **Part I**, **Part II** with **Part III**, we complete the proof. □

From Theorem 4.5, we obtain the main result of this paper as follows.

Theorem 4.6. *If F is smooth everywhere on $\Re^n$, then the merit function $\Psi_{\theta p}$ defined by (4.1) for the SOCCP (1.1) is smooth everywhere on $\Re^n$ for all $\theta \in (0,1)$ and $p \in (1,2)$.*

5 Numerical Results

In this section, we report some numerical results for solving

$$\min_{x\in\Re^n} \Psi_{\theta p}(x), \tag{5.1}$$

where $\Psi_{\theta p}(x)$ is the merit function defined by (4.1) for the SOCCP (1.1). The purpose of the numerical testings is to show some intuitive usefulness of the merit functions proposed in this paper, which is also one of the motivations of this work. All experiments are done on a PC with CPU of 2.4GHz and RAM of 2.0GB, and all codes are written in MATLAB.

We use an iterative algorithm to solve problem (5.1), where the iterative direction is chosen as the steepest descent direction, and the iterative step-length is obtained by a non-monotone Armijo line search [14], i.e., we compute the smallest nonnegative integer h such that

$$\Psi_{\theta p}(x_k + \rho^h d_k) \le C_k - \sigma\rho^h \Psi_{\theta p}(x_k),$$

where

$$d_k := -\nabla\Psi_{\theta p}(x_k),\ \ C_k := \max_{i=k-m_k,\dots,k} \Psi_{\theta p}(x_i),\ \text{ and }\ m_k := \begin{cases} 0 & \text{if } k \le s, \\ \min\{m_{k-1}+1, \hat{m}\} & \text{otherwise.}\end{cases}$$

Throughout the experiments, the parameters we used are: $\hat{m} = 5$, $s = 5$, $\rho = 0.25$ and $\sigma = 10^{-6}$. We adopt the following stopping rule: gap $\le 10^{-3}$ and $\Psi_{\theta p}(x_k) \le 10^{-6}$, where gap $:= \sum_{i=1}^{n_m} |x_i^T F_i(x)|$ with $x_i \in \mathcal{K}^{n_i}$ and $F_i(x)$ being the corresponding part to x_i of $F(x)$. We test problem (1.1) using (5.1) with different $\theta \in (0,1]$, $p \in (1,2)$ or $\theta \in (0,2)$, $p = 2$. The test problems are the following Examples 1 and 2. We will map the number of iterations for the two examples into Figures 1 and 2 for numerical analysis. We divide the cases we tested into four groups according to different values of p (1.25, 1.5, 1.75 and 2). For every p, a marked point on the corresponding curve in Figures 1 and 2 represents a tested θ.

Example 5.1. We test the SOCCP (2.2) with $F(x) := Mx + q$, where $M \in \Re^{100\times 100}$ and $q \in \Re^{100}$; and we let $\mathcal{K} := \mathcal{K}^{10} \times \mathcal{K}^{10} \times \cdots \mathcal{K}^{10}$. We generate a random matrix $A \in \Re^{100\times 100}$ and a vector $q \in \Re^{100}$ uniformly on $[-1,1]$ for its every element, respectively, then we set $M := A + 10I$. The starting point x_0 is chosen randomly uniformly on $[-1,1]$ for its every element. We test every case ten times, and record the average number of iteration for every case for numerical analysis. The numerical results are mapped in Figure 1.

Example 5.2. Consider a nonlinear SOCCP (1.1), which is taken from [15] with

$$F(x) := \begin{pmatrix} 24(2x_1 - x_2)^3 + \exp(x_1 - x_3) - 4x_4 + x_5 \\ -12(2x_1 - x_2)^3 + \frac{3(3x_2+5x_3)}{\sqrt{1+(3x_2+5x_3)^2}} - 6x_4 - 7x_5 \\ -\exp(x_1 - x_3) + \frac{5(3x_2+5x_3)}{\sqrt{1+(3x_2+5x_3)^2}} - 3x_4 + 5x_5 \\ 4x_1 + 6x_2 + 3x_3 - 1 \\ -x_1 + 7x_2 - 5x_3 + 2 \end{pmatrix}$$

and $\mathcal{K} := \mathcal{K}^3 \times \mathcal{K}^2$. The starting point x_0 is chosen as $(1, 1, 1, 1, 1)^T$. We record the number of iterations for every case and map the results in Figure 2.

In our testings, all cases are solved by the above method without failure. From Figures 1 and 2, we see that the best numerical results do not appear when $\theta = 1$ and $p = 2$, i.e., the case of the FB function. We see also that $\Psi_{\theta p}$ with parameters $p = 1.5$ and $p = 1.75$ while $\theta = 0.5$ work well for Example 5.1, and $p = 1.5$ and $p = 1.75$ while $\theta = 0.9$ for Example 5.2. However, these cases are not included in the cases given in [2, 22, 21, 24, 25]. Therefore, from the view of the numerical results obtained above, the SOC C-functions $\phi_{\theta p}$ and the corresponding merit functions $\Psi_{\theta p}$ proposed in this paper are valuable for the tested problems that we considered. We have also tested some other problems, and the computation effect is similar.

6 Conclusions

Based on the two generalized SOC C-functions proposed in [24, 25], we gave a generalization of the two generalized SOC C-functions in this paper. In particular, we proved the smoothness of the merit functions generated by the new SOC C-functions and their related merit functions. Such a property is important in designing some algorithms for solving the SOCCP, such as the merit function method for the SOCCP [24, 25]. The preliminary numerical results indicate that the SOC C-functions proposed in this paper are valuable for investigating the SOCCP. It is worthy to investigate the numerical behaviors of various algorithms for solving the SOCCP when the new SOC C-functions are used. The second further issue is to investigate other properties of the new SOC C-functions, such as, the coercivity, the (strong) semismoothness, etc.. The third issue is to extend various results related the new SOC C-functions to the framework of symmetric cones [19, 20].

Acknowledgment

The authors are very grateful to the two referees for their constructive comments and valuable suggestions, which have considerably improved the paper.

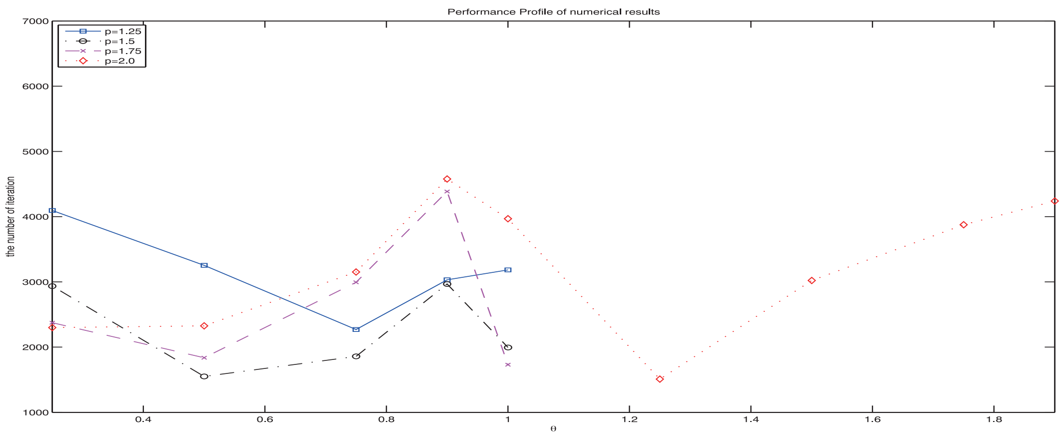

Figure 1: Performance profile for Example 5.1

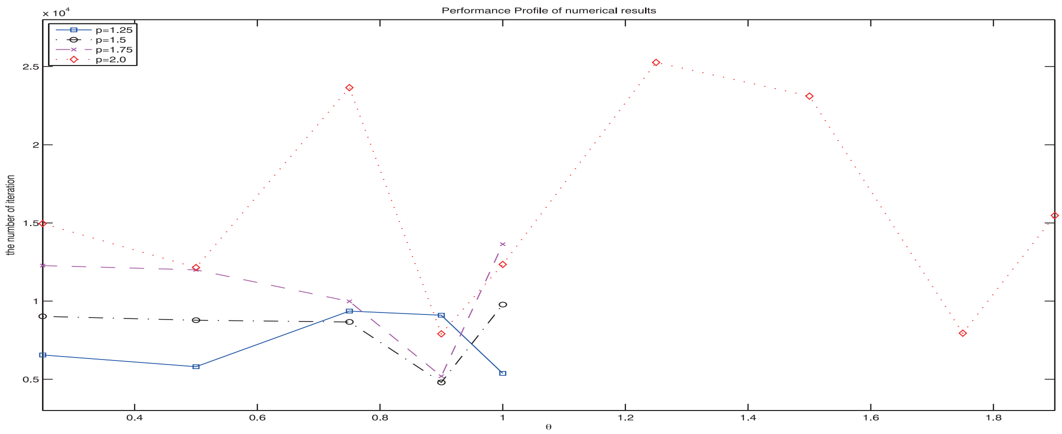

Figure 2: Performance profile for Example 5.2

References

[1] B. Chen, X. Chen and C. Kanzow, A penalized Fischer-Burmeister NCP-function, *Math. Program.* 88 (2000) 211–216.

[2] J.S. Chen, A family of NCP functions and a desent method for the nonlinear complementarity problem, *Comput. Optim. Appl.* 40 (2008) 389-404.

[3] J.S. Chen and S.H. Pan, A one-parametric class of merit function for the second-order cone complementarity problem, *Comput. Optim. Appl.* 45 (2010) 581–606.

[4] J.S. Chen, X. Chen and P. Tseng, Analysis of nonsmooth vector-valued functions associated with second order cone, *Math. Program.* 101 (2004) 95–117.

[5] J.S. Chen and P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. *Math. Program.* 104 (2005) 293–372.

[6] X. Chen and P. Tseng, Non-interior continuous methods for solving semidefinite complementarity problems, *Math. Program.* 95 (2003)431–474.

[7] X.D. Chen, D. Sun and J. Sun, Complementarity functions and numerical experiments for second-order cone complementarity problrms, *Comput. Optim. Appl.* 25 (2003) 39–56.

[8] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Verlag, New York, 2003.

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

[10] L. Faybusovich, Euclidean Jordan algebras and interior-point algorithms, *Positivity* 1 (1997) 331–357.

[11] M.C. Ferris and J.S. Pang, Engineering and economic applications of complementarity problems. *SIAM J. Rev.* 39 (1997) 669–713.

[12] M. Fukushima, Z.Q. Luo and P. Tseng, Smoothing function for second order cone complementarity problems, *SIAM J. Optim.* 12 (2002) 436–460.

[13] M.S. Gowda, R. Sznajder and J. Tao, Some P-properties for linear transformations on Euclidean Jordan algebras, *Linear Algebra Appl.* 393 (2004) 203–232.

[14] L. Grippo, F. Lampariello and S. Lucidi, A nonmonotone line search technique for Newton's method, *SIAM J. Numer. Anal.* 23 (1986) 707–716.

[15] S. Hayashi, N. Yamashita and M. Fukushima, A combined smoothing and regularization method for monotone second-order cone complementarity problems, *SIAM J. Optim.* 15 (2005) 593–615.

[16] P.T. Harker and J.S. Pang, Finite dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications, *Math. Program.* 48 (1990) 161–220.

[17] S.L. Hu, Z.H. Huang and J.S. Chen, Properties of a family of generalized NCP-functions and a derivative free algorithm for complementarity problems, *J Comput. Appl. Math.* 230 (2009) 69–82.

[18] Z.H. Huang and J. Han, Non-interior continuation method for solving the monotone semidefinite complementarity problem, *Appl. Math. Optim.* 47 (2003) 195–211.

[19] Z.H. Huang, S.L. Hu and J. Han, Global convergence of a smoothing algorithm for symmetric cone complementarity problems with a nonmonotone line search, *Sci China Ser A-Math* 52 (2009) 833–848.

[20] Z.H. Huang and T. Ni, Smoothing algorithms for complementarity problems over symmetric cones, *Comput. Optim. Appl.* 45 (2010) 557–579.

[21] C. Kanzow and H. Kleinmichel, A new class of semismooth Newton method for nonlinear complementarity problems, *Comput. Optim. Appl.* 11 (1998) 227–251.

[22] Z.Q. Luo and P. Tseng, A new class of merit functions for the nonlinear com- plementarity problem, *Complementarity and Variational Problems: State of the Art.* Edited by M. C. Ferris and J.-S. Pang, SIAM, Philadelphia, 204–225, (1997).

[23] K. Löwner, Über monotone matrixfunktionen, *Mathe. Zeits.* 38 (1934) 177–216.

[24] S.H. Pan and J.S. Chen, A semismooth Newton method for SOCCPs based on a one-parametric class of complementarity functions, *Comput. Optim. Appl.* 45 (2010) 59–88.

[25] S.H. Pan, S. Kum, Y. Lim and J.S. Chen, A generalized Fisher-Burmeister merit funvtion for the second-order cone complementarity problem, Manuscript, (2008)

[26] S.H. Schmieta and F. Alizadeh, Extension of primal-dual interior-point algorithms to symmetric cones, *Math. Program.* 96 (2003) 409–438.

[27] D. Sun and L. Qi, On NCP-functions, *Comput. Optim. Appl.* 13 (1999) 201–220.

[28] D. Sun and J. Sun, Löwner's operator and spectral functions in Euclidean Jordan algebras, *Math. Oper. Res.* 33 (2008) 421–445.

[29] P. Tseng, Growth behaviour of a class of merit functions for the nonlinear complementarity problem, *J. Optim. Theory Appl.* 89 (1996) 17–37.

Manuscript received 13 November 2008
revised 22 June 2009, 23 November 2009, 13 January 2010
accepted for publication 18 January 2010

Sheng-Long Hu
Department of Mathematics, School of Science
Tianjin University, Tianjin 300072, P.R. China
Current address: Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
E-mail address: `senoghoo@tju.edu.cn`

Zheng-Hai Huang
Department of Mathematics, School of Science
Tianjin University, Tianjin 300072, P.R. China
E-mail address: `huangzhenghai@tju.edu.cn`

Nan Lu
Department of Mathematics, School of Science
Tianjin University, Tianjin 300072, P.R. China
E-mail address: `xiaonanddup@yahoo.com.cn`