A one-parametric class of merit functions for the symmetric cone complementarity problem

Shaohua Pan [a,1], Jein-Shan Chen [b,*,2]

[a] School of Mathematical Sciences, South China University of Technology, Guangzhou 510640, China
[b] Department of Mathematics, National Taiwan Normal University, Taipei, Taiwan 11677

ABSTRACT

In this paper, we extend the one-parametric class of merit functions proposed by Kanzow and Kleinmichel [C. Kanzow, H. Kleinmichel, A new class of semismooth Newton-type methods for nonlinear complementarity problems, Comput. Optim. Appl. 11 (1998) 227–251] for the nonnegative orthant complementarity problem to the general symmetric cone complementarity problem (SCCP). We show that the class of merit functions is continuously differentiable everywhere and has a globally Lipschitz continuous gradient mapping. From this, we particularly obtain the smoothness of the Fischer–Burmeister merit function associated with symmetric cones and the Lipschitz continuity of its gradient. In addition, we also consider a regularized formulation for the class of merit functions which is actually an extension of one of the NCP function classes studied by [C. Kanzow, Y. Yamashita, M. Fukushima, New NCP functions and their properties, J. Optim. Theory Appl. 97 (1997) 115–135] to the SCCP. By exploiting the Cartesian P-properties for a nonlinear transformation, we show that the class of regularized merit functions provides a global error bound for the solution of the SCCP, and moreover, has bounded level sets under a rather weak condition which can be satisfied by the monotone SCCP with a strictly feasible point or the SCCP with the joint Cartesian R_{02}-property. All of these results generalize some recent important works in [J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, Math. Program. 104 (2005) 293–327; C.-K. Sim, J. Sun, D. Ralph, A note on the Lipschitz continuity of the gradient of the squared norm of the matrix-valued Fischer–Burmeister function, Math. Program. 107 (2006) 547–553; P. Tseng, Merit function for semidefinite complementarity problems, Math. Program. 83 (1998) 159–185] under a unified framework.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Given a Euclidean Jordan algebra $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ where $\mathbb{V}$ is a finite-dimensional vector space over the real field $\mathbb{R}$ endowed with the inner product $\langle \cdot, \cdot \rangle$ and "∘" denotes the Jordan product. Let $\mathcal{K}$ be a symmetric cone in $\mathbb{V}$ and $G, F : \mathbb{V} \to \mathbb{V}$ be nonlinear transformations assumed to be continuously differentiable throughout this paper. Consider the symmetric cone complementarity problem (SCCP) of finding $\zeta \in \mathbb{V}$ such that

$$G(\zeta) \in \mathcal{K}, \quad F(\zeta) \in \mathcal{K}, \quad \langle G(\zeta), F(\zeta) \rangle = 0. \tag{1}$$

---

* Corresponding author.
  E-mail addresses: shhpan@scut.edu.cn (S. Pan), jschen@math.ntnu.edu.tw (J.-S. Chen).
[1] Partially supported by the Doctoral Starting-up Foundation (B13B6050640) of GuangDong Province.
[2] Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. Partially supported by National Science Council of Taiwan.

The model provides a simple, natural and unified framework for various existing complementarity problems such as the nonnegative orthant nonlinear complementarity problem (NCP), the second-order cone complementarity problem (SOCCP), and the semidefinite complementarity problem (SDCP). In addition, the model itself is closely related to the KKT optimality conditions for the convex symmetric cone program (CSCP):

$$\text{minimize } g(x),$$

$$\text{subject to } \langle a_i, x \rangle = b_i, \quad i = 1, 2, \ldots, m, \ x \in \mathcal{K}, \tag{2}$$

where $a_i \in \mathbb{V}$, $b_i \in \mathbb{R}$ for $i = 1, 2, \ldots, m$, and $g : \mathbb{V} \to \mathbb{R}$ is a convex twice continuously differentiable function. Therefore, the SCCP has wide applications in engineering, economics, management science and other fields; see [1,11,20,29] and references therein.

During the past several years, interior-point methods have been well used for solving the symmetric cone linear programming problem (SCLP), i.e., the CSCP with g being a linear function (see [7,8,23,24]). However, in view of the wide applications of the SCCP, it is worthwhile to explore other solution methods for the more general CSCP and SCCP. Recently, motivated by the successful applications of the merit function approach in the solution of NCPs, SOCCPs and SDCPs (see, e.g., [4,10,22,28]), some researchers started with the investigation of merit functions or complementarity functions associated with symmetric cones. For example, Liu, Zhang and Wang [21] extended a class of merit functions proposed in [18] to the following special SCCP:

$$\zeta \in \mathcal{K}, \ F(\zeta) \in \mathcal{K}, \quad \langle \zeta, F(\zeta) \rangle = 0; \tag{3}$$

Kong, Tuncel and Xiu [17] studied the extension of the implicit Lagrangian function proposed by Mangasarian and Solodov [22] to symmetric cones; and Kong, Sun and Xiu [16] proposed a regularized smoothing method for the SCCP (3) based on the natural residual complementarity function associated with symmetric cones. Following this line, in this paper we will consider the extension of the one-parametric class of merit functions proposed by Kanzow and Kleinmichel [14] and a class of regularized functions based on it.

We define the one-parametric class of vector-valued functions $\phi_\tau : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ by

$$\phi_\tau(x, y) := \left(x^2 + y^2 + (\tau - 2)(x \circ y)\right)^{1/2} - (x + y), \tag{4}$$

where $\tau \in (0, 4)$ is an arbitrary but fixed parameter, $x^2 = x \circ x$, $x^{1/2}$ is a vector such that $(x^{1/2})^2 = x$, and $x + y$ means the usual componentwise addition of vectors. When $\tau = 2$, ϕ_τ reduces to the vector-valued Fischer–Burmeister function given by

$$\phi_{\text{FB}}(x, y) := \left(x^2 + y^2\right)^{1/2} - (x + y); \tag{5}$$

whereas as $\tau \to 0$ it will become a multiple of the vector-valued residual function

$$\psi_{\text{NR}}(x, y) := x - (x - y)_+,$$

where $(\cdot)_+$ denotes the metric projection on $\mathcal{K}$. In this sense, the one-parametric class of vector-valued functions covers the two popular complementarity functions associated with the symmetric cone $\mathcal{K}$. In fact, from Lemma 3.1 later, it follows that the function ϕ_τ with any $\tau \in (0, 4)$ is a complementarity function associated with $\mathcal{K}$, that is,

$$\phi_\tau(x, y) = 0 \quad \Longleftrightarrow \quad x \in \mathcal{K}, \ y \in \mathcal{K}, \quad \langle x, y \rangle = 0.$$

Consequently, its squared norm yields a merit function associated with $\mathcal{K}$

$$\psi_\tau(x, y) := \frac{1}{2} \left\| \phi_\tau(x, y) \right\|^2, \tag{6}$$

where $\| \cdot \|$ is the norm induced by $\langle \cdot, \cdot \rangle$, and the SCCP can be reformulated as

$$\min_{\zeta \in \mathbb{V}} f_\tau(\zeta) := \psi_\tau\left(G(\zeta), F(\zeta)\right). \tag{7}$$

To apply the effective unconstrained optimization methods, such as the quasi-Newton method, the trust-region method and the conjugate gradient method, for solving the unconstrained minimization reformulation (7) of the SCCP, the smoothness of the merit function ψ_τ and the Lipschitz continuity of its gradient will play an important role. In Sections 3 and 4, we show that the function ψ_τ defined by (6) is continuously differentiable everywhere and has a globally Lipschitz continuous gradient with the Lipschitz constant being a positive multiple of $1 + \tau^{-1}$. These results generalize some recent important works in [4,25,28] under a unified framework, as well as improve the work [21] greatly in which only the differentiability of the merit function ψ_{FB} was given.

In addition, we also consider a class of regularized functions for f_τ defined as

$$\hat{f}_\tau(\zeta) := \psi_0\left(G(\zeta) \circ F(\zeta)\right) + \psi_\tau\left(G(\zeta), F(\zeta)\right), \tag{8}$$

where $\psi_0 : \mathbb{V} \to \mathbb{R}_+$ is continuously differentiable and satisfies

$$\psi_0(u) = 0 \quad \forall u \in -\mathcal{K} \quad \text{and} \quad \psi_0(u) \geqslant \beta \big\| (u)_+ \big\| \quad \forall u \in \mathbb{V} \tag{9}$$

for some constant $\beta > 0$. Using the properties of ψ_0 in (9), it is not hard to verify that $\hat{f}_\tau$ is a merit function for the SCCP. The class of functions will reduce to the one studied in [21] if $\tau = 2$ and G degenerates into an identity transformation. In Section 5, we show that the class of merit functions can provide a global error bound for the solution of the SCCP under the condition that G and F have the joint uniform Cartesian P-property. In Section 6, we establish the boundedness of the level sets of $\hat{f}_\tau$ under a weaker condition than the one used by [21], which can be satisfied by the monotone SCCP with a strictly feasible point or the SCCP with G and F having the joint Cartesian R_{02}-property.

Throughout this paper, $\mathcal{I}$ denotes an identity operator, $\|\cdot\|$ represents the norm induced by the inner product $\langle \cdot, \cdot \rangle$, and $\mathrm{int}(\mathcal{K})$ denotes the interior of the symmetric cone $\mathcal{K}$. All vectors are column ones and write the column vector $(x_1^T, \ldots, x_m^T)^T$ as $(x_1, \ldots, x_m)$, where x_i is a column vector from the subspace $\mathbb{V}_i$. For any $x \in \mathbb{V}$, we denote $(x)_+$ and $(x)_-$ by the metric projection of x onto $\mathcal{K}$ and $-\mathcal{K}$, respectively, i.e., $(x)_+ := \arg\min_{y \in \mathcal{K}} \{\|x - y\|\}$. For any symmetric matrix A, the notation $A \succeq O$ means that A is positive semidefinite. For a differentiable mapping $F : \mathbb{V} \to \mathbb{V}$, the notation $\nabla F(x)$ denotes the transposed Jacobian operator of F at a point x. We write $x = o(\alpha)$ (respectively, $x = O(\alpha)$) if $\|x\|/|\alpha| \to 0$ (respectively, uniformly bounded) as $\alpha \to 0$.

2. Preliminaries

In this section, we recall some concepts and materials of Euclidean Jordan algebras that will be used in the subsequent sections. More detailed expositions of Euclidean Jordan algebras can be found in the monograph by Faraut and Korányi [9]. Besides, one can find excellent summaries in the articles [2,12,24,26].

A Euclidean Jordan algebra is a triple $(\mathbb{V}, \circ, \langle \cdot, \cdot \rangle_\mathbb{V})$, where $\mathbb{V}$ is a finite-dimensional inner product space over the real field $\mathbb{R}$ and $(x, y) \mapsto x \circ y : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ is a bilinear mapping satisfying the following conditions:

(i) $x \circ y = y \circ x$ for all $x, y \in \mathbb{V}$,
(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in \mathbb{V}$, where $x^2 := x \circ x$, and
(iii) $\langle x \circ y, z \rangle_\mathbb{V} = \langle x, y \circ z \rangle_\mathbb{V}$ for all $x, y, z \in \mathbb{V}$.

We call $x \circ y$ the Jordan product of x and y. We also assume that there is an element $e \in \mathbb{V}$, called the unit element, such that $x \circ e = x$ for all $x \in \mathbb{V}$. For $x \in \mathbb{V}$, let $\zeta(x)$ be the degree of the minimal polynomial of x, which can be equivalently defined as

$$\zeta(x) := \min\big\{ k : \ \{e, x, x^2, \ldots, x^k\} \text{ are linearly dependent} \big\}.$$

Since $\zeta(x) \leqslant \dim(\mathbb{V})$ where $\dim(\mathbb{V})$ denotes the dimension of $\mathbb{V}$, the rank of $\mathbb{V}$ is well defined by $r := \max\{\zeta(x) : x \in \mathbb{V}\}$. In a Euclidean Jordan algebra $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle_\mathbb{V})$, we define the set of squares as $\mathcal{K} := \{x^2 : x \in \mathbb{V}\}$. Then, by Theorem III.2.1 of [9], $\mathcal{K}$ is a symmetric cone. This means that $\mathcal{K}$ is a self-dual closed convex cone with nonempty interior $\mathrm{int}(\mathcal{K})$ and for any two elements $x, y \in \mathrm{int}(\mathcal{K})$, there exists an invertible linear transformation $\mathcal{T} : \mathbb{V} \to \mathbb{V}$ such that $\mathcal{T}(\mathcal{K}) = \mathcal{K}$ and $\mathcal{T}(x) = y$.

A Euclidean Jordan algebra is said to be *simple* if it is not the direct sum of two Euclidean Jordan algebras. By Proposition III.4.4 of [9], each Euclidean Jordan algebra is, in a unique way, a direct sum of simple Euclidean Jordan algebras. A common simple Euclidean Jordan algebra is $(\mathbb{S}^n, \circ, \langle \cdot, \cdot \rangle_{\mathbb{S}^n})$, where $\mathbb{S}^n$ is the space of $n \times n$ real symmetric matrices with the inner product $\langle X, Y \rangle_{\mathbb{S}^n} := \mathrm{Tr}(XY)$, and the Jordan product is defined by $X \circ Y := (XY + YX)/2$. Here, XY is the usual matrix multiplication of X and Y and $\mathrm{Tr}(X)$ is the trace of X. The associate cone $\mathcal{K}$ is the set of all positive semidefinite matrices. Another one is the Lorentz algebra $(\mathbb{R}^n, \circ, \langle \cdot, \cdot \rangle_{\mathbb{R}^n})$, where $\mathbb{R}^n$ is the Euclidean space of dimension n with the standard inner product $\langle x, y \rangle_{\mathbb{R}^n} = x^T y$, and the Jordan product is defined by $x \circ y := (\langle x, y \rangle_{\mathbb{R}^n}, x_1 y_2 + y_1 x_2)$ for any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$. The associate cone, called the Lorentz cone or the second-order cone, is

$$\mathcal{K} := \big\{ x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} : \ \|x_2\| \leqslant x_1 \big\}.$$

Recall that an element $c \in \mathbb{V}$ is said to be *idempotent* if $c^2 = c$. Two idempotents c and d are said to be *orthogonal* if $c \circ d = 0$. One says that $\{c_1, c_2, \ldots, c_k\}$ is a complete system of orthogonal idempotents if

$$c_j^2 = c_j, \qquad c_j \circ c_i = 0 \ \text{ if } j \neq i, \ j, i = 1, 2, \ldots, k, \quad \text{and} \quad \sum_{j=1}^k c_j = e.$$

A nonzero idempotent is said to be *primitive* if it cannot be written as the sum of two other nonzero idempotents. We call a complete system of orthogonal primitive idempotents a *Jordan frame*. Then, we have the following spectral decomposition theorem.

Theorem 2.1. *(See [9, Theorem III.1.2].) Suppose that $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle_{\mathbb{V}})$ is a Euclidean Jordan algebra and the rank of $\mathbb{A}$ is r. Then for any $x \in \mathbb{V}$, there exist a Jordan frame $\{c_1, c_2, \ldots, c_r\}$ and real numbers $\lambda_1(x), \lambda_2(x), \ldots, \lambda_r(x)$, arranged in the decreasing order $\lambda_1(x) \geqslant \lambda_2(x) \geqslant \cdots \geqslant \lambda_r(x)$, such that $x = \sum_{j=1}^{r} \lambda_j(x) c_j$.*

The numbers $\lambda_j(x)$ (counting multiplicities), which are uniquely determined by x, are called the eigenvalue, and we write the maximum eigenvalue and the minimum eigenvalue of x as $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$, respectively. The trace of x, denoted as $\mathrm{tr}(x)$, is defined by $\mathrm{tr}(x) := \sum_{j=1}^{r} \lambda_j(x)$; whereas the determinant of x is defined by $\det(x) := \prod_{j=1}^{r} \lambda_j(x)$.

By Proposition III.1.5 of [9], a Jordan algebra over $\mathbb{R}$ with a unit element $e \in \mathbb{V}$ is Euclidean if and only if the symmetric bilinear form $\mathrm{tr}(x \circ y)$ is positive definite. Hence, we may define an inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{V}$ by

$$\langle x, y \rangle := \mathrm{tr}(x \circ y), \quad \forall x, y \in \mathbb{V}. \tag{10}$$

Unless otherwise states, the inner product $\langle \cdot, \cdot \rangle$ appearing in this paper always means the one defined by (10). By the associativity of $\mathrm{tr}(\cdot)$ (see [9, Proposition II.4.3]), the inner product $\langle \cdot, \cdot \rangle$ is associative, i.e., $\langle x, y \circ z \rangle = \langle y, x \circ z \rangle$ for all $x, y, z \in \mathbb{V}$. Let

$$\mathcal{L}(x)y := x \circ y \quad \text{for every } y \in \mathbb{V}.$$

Then, the linear operator $\mathcal{L}(x)$ for each $x \in \mathbb{V}$ is symmetric with respect to the inner product $\langle \cdot, \cdot \rangle$ in the sense that $\langle \mathcal{L}(x)y, z \rangle = \langle y, \mathcal{L}(x)z \rangle$ for any $y, z \in \mathbb{V}$. Let $\| \cdot \|$ be the norm on $\mathbb{V}$ induced by the inner product $\langle \cdot, \cdot \rangle$, namely,

$$\|x\| := \sqrt{\langle x, x \rangle} = \left(\sum_{j=1}^{r} \lambda_j^2(x) \right)^{1/2}, \quad \forall x \in \mathbb{V}.$$

It is not difficult to verify that for any $x, y \in \mathbb{V}$, there always holds that

$$\langle x, y \rangle \leqslant \frac{1}{2} \left(\|x\|^2 + \|y\|^2 \right) \quad \text{and} \quad \|x + y\|^2 \leqslant 2 \left(\|x\|^2 + \|y\|^2 \right). \tag{11}$$

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a scalar valued function. Then, it is natural to define a vector-valued function associated with the Euclidean Jordan algebra $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ by

$$\varphi_{\mathbb{V}}(x) := \varphi(\lambda_1(x)) c_1 + \varphi(\lambda_2(x)) c_2 + \cdots + \varphi(\lambda_r(x)) c_r, \tag{12}$$

where $x \in \mathbb{V}$ has the spectral decomposition $x = \sum_{j=1}^{r} \lambda_j(x) c_j$. The function $\varphi_{\mathbb{V}}$ is also called the Löwner operator [26]. When $\varphi(t)$ is chosen as $\max\{0, t\}$ and $\min\{0, t\}$ for $t \in \mathbb{R}$, respectively, $\varphi_{\mathbb{V}}$ becomes the metric projection operator onto $\mathcal{K}$ and $-\mathcal{K}$:

$$(x)_+ := \sum_{j=1}^{r} \max\{0, \lambda_j(x)\} c_j \quad \text{and} \quad (x)_- := \sum_{j=1}^{r} \min\{0, \lambda_j(x)\} c_j. \tag{13}$$

Lemma 2.1. *(See [26, Theorem 13].) For any $x = \sum_{j=1}^{r} \lambda_j(x) c_j$, let $\varphi_{\mathbb{V}}$ be given as in (12). Then $\varphi_{\mathbb{V}}$ is (continuously) differentiable at x if and only if φ is (continuously) differentiable at each $\lambda_j(x)$, $j = 1, 2, \ldots, r$. The derivative of $\varphi_{\mathbb{V}}$ at x, for any $h \in \mathbb{V}$, is*

$$\varphi_{\mathbb{V}}'(x)h = \sum_{j=1}^{r} \left[\varphi^{[1]}(\lambda(x)) \right]_{jj} \langle c_j, h \rangle c_j + \sum_{1 \leqslant j < l \leqslant r} 4 \left[\varphi^{[1]}(\lambda(x)) \right]_{jl} c_j \circ (c_l \circ h),$$

where

$$\left[\varphi^{[1]}(\lambda(x)) \right]_{ij} := \begin{cases} \frac{\varphi(\lambda_i(x)) - \varphi(\lambda_j(x))}{\lambda_i(x) - \lambda_j(x)} & \text{if } \lambda_i(x) \neq \lambda_j(x), \\ \varphi'(\lambda_i(x)) & \text{if } \lambda_i(x) = \lambda_j(x), \end{cases} \quad i, j = 1, 2, \ldots, r.$$

In fact, the Jacobian $\varphi_{\mathbb{V}}'(\cdot)$ is a linear and symmetric operator, which can be written as

$$\varphi_{\mathbb{V}}'(x) = \sum_{j=1}^{r} \varphi'(\lambda_j(x)) \mathcal{Q}(c_j) + 2 \sum_{i,j=1, i \neq j}^{r} \left[\varphi^{[1]}(\lambda(x)) \right]_{ij} \mathcal{L}(c_j) \mathcal{L}(c_i), \tag{14}$$

where $\mathcal{Q}(x) := 2\mathcal{L}^2(x) - \mathcal{L}(x^2)$ for any $x \in \mathbb{V}$ is called the *quadratic representation* of $\mathbb{V}$.

In the sequel, unless otherwise stated, we assume that $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ is a simple Euclidean Jordan algebra of rank r and $\dim(\mathbb{V}) = n$.

An important part in the theory of Euclidean Jordan algebras is the *Peirce decomposition*. Let c be a nonzero idempotent in $\mathbb{A}$. Then, by [9, Proposition III.1.3], c satisfies $2\mathcal{L}^3(c) - 3\mathcal{L}^2(c) + \mathcal{L}(c) = 0$ and the distinct eigenvalues of the symmetric operator $\mathcal{L}(c)$ are 0, $\frac{1}{2}$ and 1. Let $\mathbb{V}(c, 1)$, $\mathbb{V}(c, \frac{1}{2})$ and $\mathbb{V}(c, 0)$ be the three corresponding eigenspaces, i.e.,

$$\mathbb{V}(c, \alpha) := \{x \in \mathbb{V} : \mathcal{L}(c)x = \alpha x\}, \quad \alpha = 1, \frac{1}{2}, 0.$$

Then $\mathbb{V}$ is the orthogonal direct sum of $\mathbb{V}(c, 1)$, $\mathbb{V}(c, \frac{1}{2})$ and $\mathbb{V}(c, 0)$. The decomposition

$$\mathbb{V} = \mathbb{V}(c, 1) \oplus \mathbb{V}\left(c, \frac{1}{2}\right) \oplus \mathbb{V}(c, 0)$$

is called the *Peirce decomposition* of $\mathbb{V}$ with respect to the nonzero idempotent c.

Let $\{c_1, c_2, \ldots, c_r\}$ be a Jordan frame of $\mathbb{A}$. For $i, j \in \{1, \ldots, r\}$, define the eigenspaces

$$\mathbb{V}_{ii} := \mathbb{V}(c_i, 1) = \mathbb{R}c_i,$$
$$\mathbb{V}_{ij} := \mathbb{V}\left(c_i, \frac{1}{2}\right) \cap \mathbb{V}\left(c_j, \frac{1}{2}\right), \quad i \neq j.$$

Then, from [9, Theorem IV.2.1], it follows that the following conclusion holds.

Theorem 2.2. *The space $\mathbb{V}$ is the orthogonal direct sum of subspaces $\mathbb{V}_{ij}$ ($1 \leqslant i \leqslant j \leqslant r$), i.e., $\mathbb{V} = \bigoplus_{i \leqslant j} V_{ij}$. Furthermore,*

$$\mathbb{V}_{ij} \circ \mathbb{V}_{ij} \subset \mathbb{V}_{ii} + \mathbb{V}_{jj},$$
$$\mathbb{V}_{ij} \circ \mathbb{V}_{jk} \subset \mathbb{V}_{ik}, \quad \text{if } i \neq k,$$
$$\mathbb{V}_{ij} \circ \mathbb{V}_{kl} = \{0\}, \quad \text{if } \{i, j\} \cap \{k, l\} = \emptyset.$$

Let $x \in \mathbb{V}$ have the spectral decomposition $x = \sum_{j=1}^{r} \lambda_j(x)c_j$. For $i, j \in \{1, 2, \ldots, r\}$, let $\mathcal{C}_{ij}(x)$ be the orthogonal projection operator onto $\mathbb{V}_{ij}$. Then,

$$\mathcal{C}_{ij}(x) = \mathcal{C}_{ij}^*(x), \qquad \mathcal{C}_{ij}^2(x) = \mathcal{C}_{ij}(x), \qquad \mathcal{C}_{ij}(x)\mathcal{C}_{kl}(x) = 0 \quad \text{if } \{i, j\} \neq \{k, l\}, \ i, j, k, l = 1, \ldots, r, \tag{15}$$

and

$$\sum_{1 \leqslant i \leqslant j \leqslant r} \mathcal{C}_{ij}(x) = \mathcal{I}, \tag{16}$$

where $\mathcal{C}_{ij}^*$ is the adjoint (operator) of $\mathcal{C}_{ij}$. In addition, by [9, Theorem IV.2.1],

$$\mathcal{C}_{jj}(x) = \mathcal{Q}(c_j) \quad \text{and} \quad \mathcal{C}_{ij}(x) = 4\mathcal{L}(c_i)\mathcal{L}(c_j) = 4\mathcal{L}(c_j)\mathcal{L}(c_i) = \mathcal{C}_{ji}(x), \quad i, j = 1, 2, \ldots, r.$$

Note that the original notation in [9] for orthogonal projection operator is P_{ij}. However, to avoid confusion with another orthogonal projector $P_i(c_j)$ onto $\mathbb{V}(c, \alpha)$ and orthogonal matrix P which will be used later (Sections 3 and 4), we adopt $\mathcal{C}_{ij}$ instead.

With the orthogonal projection operators $\{\mathcal{C}_{ij}(x): i, j = 1, 2, \ldots, r\}$, we have the following spectral decomposition theorem for $\mathcal{L}(x)$ and $\mathcal{L}(x^2)$; see [15, Chapters VI–V].

Lemma 2.2. *Let $x \in \mathbb{V}$ have the spectral decomposition $x = \sum_{j=1}^{r} \lambda_j(x)c_j$. Then the symmetric operator $\mathcal{L}(x)$ has the spectral decomposition*

$$\mathcal{L}(x) = \sum_{j=1}^{r} \lambda_j(x)\mathcal{C}_{jj}(x) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{2}\left(\lambda_j(x) + \lambda_l(x)\right)\mathcal{C}_{jl}(x)$$

with the spectrum $\sigma(\mathcal{L}(x))$ consisting of all distinct numbers in $\{\frac{1}{2}(\lambda_j(x) + \lambda_l(x)): j, l = 1, 2, \ldots, r\}$, and $\mathcal{L}(x^2)$ has the spectral decomposition

$$\mathcal{L}(x^2) = \sum_{j=1}^{r} \lambda_j^2(x)\mathcal{C}_{jj}(x) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{2}\left(\lambda_j^2(x) + \lambda_l^2(x)\right)\mathcal{C}_{jl}(x) \tag{17}$$

with the spectrum $\sigma(\mathcal{L}(x^2))$ consisting of all distinct numbers in $\{\frac{1}{2}(\lambda_j^2(x) + \lambda_l^2(x)): j, l = 1, 2, \ldots, r\}$.

Proposition 2.1. *For any $x \in \mathbb{V}$, the operator $\mathcal{L}(x^2) - \mathcal{L}^2(x)$ is positive semidefinite.*

Proof. By Lemma 2.2 and (15), we can verify that $\mathcal{L}^2(x)$ has the spectral decomposition

$$\mathcal{L}^2(x) = \sum_{j=1}^{r} \lambda_j^2(x)\mathcal{C}_{jj}(x) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{4}\big(\lambda_j(x) + \lambda_l(x)\big)^2 \mathcal{C}_{jl}(x). \tag{18}$$

This means that the operator $\mathcal{L}(x^2) - \mathcal{L}^2(x)$ has the spectral decomposition

$$\mathcal{L}(x^2) - \mathcal{L}^2(x) = \sum_{1 \leqslant j < l \leqslant r} \left[\frac{1}{2}\big(\lambda_j^2(x) + \lambda_l^2(x)\big) - \frac{1}{4}\big(\lambda_j(x) + \lambda_l(x)\big)^2 \right] \mathcal{C}_{jl}(x).$$

Noting that the orthogonal projection operator is positive semidefinite on $\mathbb{V}$ and

$$\frac{\lambda_j^2(x) + \lambda_l^2(x)}{2} \geqslant \frac{(\lambda_j(x) + \lambda_l(x))^2}{4} \quad \text{for all } j, l = 1, 2, \ldots, r,$$

we readily obtain the conclusion from the spectral decomposition of $\mathcal{L}(x^2) - \mathcal{L}^2(x)$. $\quad\square$

3. Differentiability of the function ψ_τ

In this section, we show that ψ_τ is a merit function associated with $\mathcal{K}$, and moreover, it is differentiable everywhere on $\mathbb{V} \times \mathbb{V}$. By the definition of Jordan product,

$$x^2 + y^2 + (\tau - 2)(x \circ y) = \left(x + \frac{\tau - 2}{2}y\right)^2 + \frac{\tau(4 - \tau)}{4}y^2 = \left(y + \frac{\tau - 2}{2}x\right)^2 + \frac{\tau(4 - \tau)}{4}x^2 \in \mathcal{K} \tag{19}$$

for any $x, y \in \mathbb{V}$, and consequently the function ϕ_τ in (4) is well defined. The following lemma states that ϕ_τ and ψ_τ is respectively a complementarity function and a merit function associated with $\mathcal{K}$.

Lemma 3.1. *For any $x, y \in \mathbb{V}$, let ϕ_τ and ψ_τ be given by (4) and (6), respectively. Then,*

$$\psi_\tau(x, y) = 0 \quad \Longleftrightarrow \quad \phi_\tau(x, y) = 0 \quad \Longleftrightarrow \quad x \in \mathcal{K}, \ y \in \mathcal{K}, \quad \langle x, y \rangle = 0.$$

Proof. The first equivalence is clear by the definition of ψ_τ, and we only need to prove the second equivalence. Suppose that $\phi_\tau(x, y) = 0$. Then,

$$\big[x^2 + y^2 + (\tau - 2)(x \circ y)\big]^{1/2} = (x + y). \tag{20}$$

Squaring the two sides of (20) yields that

$$x^2 + y^2 + (\tau - 2)(x \circ y) = x^2 + y^2 + 2(x \circ y),$$

which implies $x \circ y = 0$ since $\tau \in (0, 4)$. Substituting $x \circ y = 0$ into (20), we have that

$$x = \big(x^2 + y^2\big)^{1/2} - y \quad \text{and} \quad y = \big(x^2 + y^2\big)^{1/2} - x.$$

Since $x^2 + y^2 \in \mathcal{K}$, $x^2 \in \mathcal{K}$ and $y^2 \in \mathcal{K}$, from [12, Proposition 8] or [19, Corollary 9] it follows that $x, y \in \mathcal{K}$. Consequently, the necessity holds. For the other direction, suppose $x, y \in \mathcal{K}$ and $x \circ y = 0$. Then, $(x + y)^2 = x^2 + y^2$. This, together with $x \circ y = 0$, implies that

$$\big[x^2 + y^2 + (\tau - 2)(x \circ y)\big]^{1/2} - (x + y) = 0.$$

Consequently, the sufficiency follows. The proof is thus completed. $\quad\square$

In what follows, we concentrate on the differentiability of the merit function ψ_τ. For this purpose, we need the following two crucial technical lemmas.

Lemma 3.2. *For any $x, y \in \mathbb{V}$, let $u(x, y) := (x^2 + y^2)^{1/2}$. Then, the function $u(x, y)$ is continuously differentiable at any point (x, y) such that $x^2 + y^2 \in \text{int}(\mathcal{K})$. Furthermore,*

$$\nabla_x u(x, y) = \mathcal{L}(x)\mathcal{L}^{-1}\big(u(x, y)\big) \quad \text{and} \quad \nabla_y u(x, y) = \mathcal{L}(y)\mathcal{L}^{-1}\big(u(x, y)\big). \tag{21}$$

Proof. The first part is due to Lemma 2.1. It remains to derive the formulas in (21). From the definition of $u(x, y)$, it follows that

$$u^2(x, y) = x^2 + y^2, \quad \forall x, y \in \mathbb{V}. \tag{22}$$

By the formula (14), it is easy to verify that $\nabla_x(x^2) = 2\mathcal{L}(x)$. Differentiating on both sides of (22) with respect to x then yields that

$$2\nabla_x u(x, y)\mathcal{L}(u(x, y)) = 2\mathcal{L}(x).$$

This implies that $\nabla_x u(x, y) = \mathcal{L}(x)\mathcal{L}^{-1}(u(x, y))$ since, by $u(x, y) \in \text{int}(\mathcal{K})$, $\mathcal{L}(u(x, y))$ is positive definite on $\mathbb{V}$. Similarly, we have that $\nabla_y u(x, y) = \mathcal{L}(y)\mathcal{L}^{-1}(u(x, y))$. $\square$

To present another lemma, we first introduce some related notations. For any $0 \neq z \in \mathcal{K}$ and $z \notin \text{int}(\mathcal{K})$, suppose that z has the spectral decomposition $z = \sum_{j=1}^{r} \lambda_j(z)c_j$, where $\{c_1, c_2, \ldots, c_r\}$ is a Jordan frame and $\lambda_1(z), \ldots, \lambda_r(z)$ are the eigenvalues arranged in the decreasing order $\lambda_1(z) \geqslant \lambda_2(z) \geqslant \cdots \geqslant \lambda_r(z) = 0$. Define the index

$$j^* := \min\{j \mid \lambda_j(z) = 0, \ j = 1, 2, \ldots, r\} \tag{23}$$

and let

$$c_J := \sum_{l=1}^{j^*-1} c_l. \tag{24}$$

Clearly, j^* and c_J are well defined since $0 \neq z \in \mathcal{K}$ and $z \notin \text{int}(\mathcal{K})$. Since c_J is an idempotent and $c_J \neq 0$ (otherwise $z = 0$), $\mathbb{V}$ can be decomposed as the orthogonal direct sum of the subspaces $\mathbb{V}(c_J, 1)$, $\mathbb{V}(c_J, \frac{1}{2})$ and $\mathbb{V}(c_J, 0)$. In the sequel, we write $P_1(c_J)$, $P_{\frac{1}{2}}(c_J)$ and $P_0(c_J)$ as the orthogonal projection onto $\mathbb{V}(c_J, 1)$, $\mathbb{V}(c_J, \frac{1}{2})$ and $\mathbb{V}(c_J, 0)$, respectively. From [21], we know that $\mathcal{L}(z)$ is positive definite on $\mathbb{V}(c_J, 1)$ and is a one-to-one mapping from $\mathbb{V}(c_J, 1)$ to $\mathbb{V}(c_J, 1)$. This means that $\mathcal{L}(z)$ an inverse $\mathcal{L}^{-1}(z)$ on $\mathbb{V}(c_J, 1)$, i.e., for any $u \in \mathbb{V}(c_J, 1)$, $\mathcal{L}^{-1}(z)u$ is the unique $v \in \mathbb{V}(c_J, 1)$ such that $z \circ v = u$.

Lemma 3.3. *For any $x, y \in \mathbb{V}$, let $z : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ be the mapping defined as*

$$z = z(x, y) := \left[x^2 + y^2 + (\tau - 2)(x \circ y)\right]^{1/2}. \tag{25}$$

If $(x, y) \neq (0, 0)$ such that $z(x, y) \notin \text{int}(\mathcal{K})$, then the following results hold:

(a) *The vectors $x, y, x + y, x + \frac{\tau-2}{2}y$ and $y + \frac{\tau-2}{2}x$ belong to the subspace $\mathbb{V}(c_J, 1)$.*
(b) *For any $h \in \mathbb{V}$ such that $z^2(x, y) + h \in \mathcal{K}$, let $w = w(x, y) := [z^2(x, y) + h]^{1/2} - z(x, y)$. Then, $P_1(c_J)w = \frac{1}{2}\mathcal{L}^{-1}(z(x, y)) \times [P_1(c_J)h] + o(\|h\|)$.*

Proof. From (19) and the definition of z, it is clear that $z(x, y) \in \mathcal{K}$ for all $x, y \in \mathbb{V}$. Hence, using the similar arguments as Lemma 11 of [21] yields the desired result. $\square$

Now by Lemmas 3.2 and 3.3, we prove the differentiability of the merit function ψ_τ.

Proposition 3.1. *The function ψ_τ defined by (6) is differentiable everywhere on $\mathbb{V} \times \mathbb{V}$. Furthermore, $\nabla_x \psi_\tau(0, 0) = \nabla_y \psi_\tau(0, 0) = 0$, and if $(x, y) \neq (0, 0)$, then*

$$\nabla_x \psi_\tau(x, y) = \left[\mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}(z(x, y)) - \mathcal{I}\right]\phi_\tau(x, y),$$

$$\nabla_y \psi_\tau(x, y) = \left[\mathcal{L}\left(y + \frac{\tau - 2}{2}x\right)\mathcal{L}^{-1}(z(x, y)) - \mathcal{I}\right]\phi_\tau(x, y), \tag{26}$$

where $z(x, y)$ is given by (25).

Proof. We prove the conclusion by the following three cases.

Case (1): $(x, y) = (0, 0)$. For any $u, v \in \mathbb{V}$, suppose that $u^2 + v^2 + (\tau - 2)(u \circ v)$ has the spectrum decomposition $u^2 + v^2 + (\tau - 2)(u \circ v) = \sum_{j=1}^{r} \mu_j d_j$, where $\{d_1, d_2, \ldots, d_r\}$ is the corresponding Jordan frame. Then, for $j = 1, 2, \ldots, r$, we have

$$\mu_j = \frac{1}{\|d_j\|^2}\left\langle\sum_{j=1}^{r}\mu_j d_j, d_j\right\rangle = \langle u^2 + v^2 + (\tau-2)(u \circ v), d_j\rangle = \left\langle\left(u + \frac{\tau-2}{2}v\right)^2 + \frac{\tau(4-\tau)}{4}v^2, d_j\right\rangle$$

$$\leqslant \left\langle\left(u + \frac{\tau-2}{2}v\right)^2 + \frac{\tau(4-\tau)}{4}v^2, e\right\rangle = \|u\|^2 + (\tau-2)\langle u, v\rangle + \|v\|^2$$

$$\leqslant (\tau/2)\big(\|u\|^2 + \|v\|^2\big), \tag{27}$$

where the second equality is by $\|d_j\| = 1$, the first inequality is due to $e = \sum_{j=1}^{r}d_j$ and $d_j \in \mathcal{K}$ for $j = 1, 2, \ldots, r$, and the last inequality is due to (11). Therefore,

$$\psi_\tau(u, v) - \psi_\tau(0, 0) = \frac{1}{2}\left\|\big[u^2 + v^2 + (\tau-2)(u \circ v)\big]^{1/2} - (u+v)\right\|^2 = \frac{1}{2}\left\|\sum_{j=1}^{r}\sqrt{\mu_j}d_j - (u+v)\right\|^2$$

$$\leqslant \left\|\sum_{j=1}^{r}\sqrt{\mu_j}d_j\right\|^2 + \|u+v\|^2 \leqslant \sum_{j=1}^{r}\mu_j\|d_j\|^2 + 2\big(\|u\|^2 + \|v\|^2\big)$$

$$\leqslant \left(\frac{1}{2}\tau r + 2\right)\big(\|u\|^2 + \|v\|^2\big),$$

where the first two inequalities are due to (11), and the last one is from (27). This shows that ψ_τ is differentiable at $(0,0)$ with $\nabla_x\psi_\tau(0,0) = \nabla_y\psi_\tau(0,0) = 0$.

Case (2): $z(x, y) \in \text{int}(\mathcal{K})$. Since $\phi_\tau(x, y) = z(x, y) - (x+y)$, we have from Lemma 2.1 that ϕ_τ is continuously differentiable under this case. Notice that

$$\psi_\tau(x, y) = \frac{1}{2}\langle e, \phi_\tau^2(x, y)\rangle,$$

and hence the function ψ_τ is continuously differentiable. Applying the chain rule yields

$$\nabla_x\psi_\tau(x, y) = \nabla_x\phi_\tau(x, y)\mathcal{L}\big(\phi_\tau(x, y)\big)e = \nabla_x\phi_\tau(x, y)\phi_\tau(x, y). \tag{28}$$

On the other hand, from (19) it follows that

$$\phi_\tau(x, y) = \left[\left(x + \frac{\tau-2}{2}y\right)^2 + \frac{\tau(4-\tau)}{4}y^2\right]^{1/2} - (x+y),$$

and therefore using the formulas in (21) gives that

$$\nabla_x\phi_\tau(x, y) = \mathcal{L}\left(x + \frac{\tau-2}{2}y\right)\mathcal{L}^{-1}\big(z(x, y)\big) - \mathcal{I}.$$

This, together with (28), immediately yields that

$$\nabla_x\psi_\tau(x, y) = \left[\mathcal{L}\left(x + \frac{\tau-2}{2}y\right)\mathcal{L}^{-1}\big(z(x, y)\big) - \mathcal{I}\right]\phi_\tau(x, y).$$

For symmetry of x and y in $\psi_\tau(x, y)$, we also have that

$$\nabla_y\psi_\tau(x, y) = \left[\mathcal{L}\left(y + \frac{\tau-2}{2}x\right)\mathcal{L}^{-1}\big(z(x, y)\big) - \mathcal{I}\right]\phi_\tau(x, y).$$

Case (3): $(x, y) \neq (0, 0)$ and $z(x, y) \notin \text{int}(\mathcal{K})$. For any $u, v \in \mathbb{V}$, define

$$\hat{z} := 2\hat{x} \circ u + 2\hat{y} \circ v + u^2 + v^2 + (\tau-2)u \circ v$$

with $\hat{x} = x + \frac{\tau-2}{2}y$ and $\hat{y} = y + \frac{\tau-2}{2}x$. It is not difficult to verify that

$$z^2(x, y) + \hat{z} = \left((x+u) + \frac{\tau-2}{2}(y+v)\right)^2 + \frac{\tau(4-\tau)}{4}(y+v)^2 = z^2(x+u, y+v) \in \mathcal{K}.$$

Let

$$w(x, y) := \big(z^2(x, y) + \hat{z}\big)^{1/2} - z(x, y).$$

From the definitions of ψ_τ and $z(x, y)$, it then follows that

$$\psi_\tau(x+u, y+v) - \psi_\tau(x, y) = \frac{1}{2}\Big[\big\|[z^2(x, y) + \hat{z}]^{1/2} - (x+u+y+v)\big\|^2 - \|z(x, y) - (x+y)\|^2\Big]$$

$$= \frac{1}{2}\big[\langle\hat{z}, e\rangle + \|u+v\|^2\big] - \big\langle w(x, y), x+u+y+v\big\rangle + \big\langle x+y-z(x, y), u+v\big\rangle$$

$$= -\big\langle w(x, y), x+y\big\rangle + \big\langle x+y-z(x, y), u+v\big\rangle + \langle\hat{x}, u\rangle + \langle\hat{y}, v\rangle + o\big(\|(u, v)\|\big). \qquad (29)$$

By Lemma 3.3(a), $x+y \in \mathbb{V}(c_J, 1)$. Thus, using Lemma 3.3(b), we have that

$$\big\langle w(x, y), x+y\big\rangle = \big\langle P_1(c_J)w(x, y), x+y\big\rangle = \Big\langle \frac{1}{2}\mathcal{L}^{-1}(z(x, y))[P_1(c_J)\hat{z}] + o(\|\hat{z}\|), x+y\Big\rangle$$

$$= \frac{1}{2}\big\langle P_1(c_J)\hat{z}, \mathcal{L}^{-1}(z(x, y))[x+y]\big\rangle + o(\|\hat{z}\|)$$

$$= \big\langle P_1(c_J)[\hat{x}\circ u + \hat{y}\circ v], \mathcal{L}^{-1}(z(x, y))[x+y]\big\rangle + o\big(\|(u, v)\|\big)$$

$$= \big\langle \hat{x}\circ u + \hat{y}\circ v, P_1(c_J)[\mathcal{L}^{-1}(z(x, y))(x+y)]\big\rangle + o\big(\|(u, v)\|\big)$$

$$= \big\langle \hat{x}\circ u + \hat{y}\circ v, \mathcal{L}^{-1}(z(x, y))(x+y)\big\rangle + o\big(\|(u, v)\|\big)$$

$$= \big\langle[\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\hat{x}, u\big\rangle + \big\langle[\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\hat{y}, v\big\rangle + o\big(\|(u, v)\|\big), \qquad (30)$$

where the first equality is since $\mathbb{V} = \mathbb{V}(c_J, 1) \oplus \mathbb{V}(c_J, \frac{1}{2}) \oplus \mathbb{V}(c_J, 0)$, the fifth one is due to $P_1(c_J) = P_1^*(c_J)$, and the sixth is from the fact that $\mathcal{L}^{-1}(z(x, y))(x+y) \in \mathbb{V}(c_J, 1)$. Combining (29) with (30), we obtain that

$$\psi_\tau(x+u, y+v) - \psi_\tau(x, y) = \big\langle\hat{x}+x+y-z(x, y) - [\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\hat{x}, u\big\rangle$$

$$+ \big\langle\hat{y}+x+y-z(x, y) - [\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\hat{y}, v\big\rangle + o\big(\|(u, v)\|\big).$$

This implies that the function ψ_τ is differentiable at (x, y), and furthermore,

$$\nabla_x\psi_\tau(x, y) = \hat{x}+x+y-z(x, y) - [\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\hat{x},$$

$$\nabla_y\psi_\tau(x, y) = \hat{y}+x+y-z(x, y) - [\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\hat{y}.$$

Notice that

$$\hat{x}+x+y-z(x, y) - [\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\hat{x} = \hat{x} - \phi_\tau(x, y) - [\mathcal{L}^{-1}(z(x, y))(x+y)]\circ\Big(x+\frac{\tau-2}{2}y\Big)$$

$$= x + \frac{\tau-2}{2}y - \phi_\tau(x, y) - \mathcal{L}\Big(x+\frac{\tau-2}{2}y\Big)[\mathcal{L}^{-1}(z(x, y))(x+y)]$$

$$= \mathcal{L}\Big(x+\frac{\tau-2}{2}y\Big)\mathcal{L}^{-1}(z(x, y))[z(x, y) - x - y] - \phi_\tau(x, y)$$

$$= \Big[\mathcal{L}\Big(x+\frac{\tau-2}{2}y\Big)\mathcal{L}^{-1}(z(x, y)) - \mathcal{I}\Big]\phi_\tau(x, y),$$

where the third equality is due to $\mathcal{L}^{-1}(z(x, y))z(x, y) = e$ and the fact that

$$x + \frac{\tau-2}{2}y = \mathcal{L}\Big(x+\frac{\tau-2}{2}y\Big)e = \mathcal{L}\Big(x+\frac{\tau-2}{2}y\Big)\mathcal{L}^{-1}(z(x, y))z(x, y).$$

Therefore,

$$\nabla_x\psi_\tau(x, y) = \Big[\mathcal{L}\Big(x+\frac{\tau-2}{2}y\Big)\mathcal{L}^{-1}(z(x, y)) - \mathcal{I}\Big]\phi_\tau(x, y).$$

Similarly, we also have that

$$\nabla_y\psi_\tau(x, y) = \Big[\mathcal{L}\Big(y+\frac{\tau-2}{2}x\Big)\mathcal{L}^{-1}(z(x, y)) - \mathcal{I}\Big]\phi_\tau(x, y).$$

This shows that the conclusion holds under this case. The proof is thus completed. □

It should be pointed out that the formula (26) is well defined even if $z(x, y) \notin \text{int}(\mathcal{K})$ since in this case $\phi_\tau(x, y) \in \mathbb{V}(c_J, 1)$ by Lemma 3.3(a). When $\mathbb{A}$ is specified as the Lorentz algebra $(\mathbb{R}^n, \circ, \langle\cdot,\cdot\rangle_{\mathbb{R}^n})$, the formula reduces to the one of [3, Proposition 3.2]; whereas when $\mathbb{A}$ is specified as $(\mathbb{S}^n, \circ, \langle\cdot,\cdot\rangle_{\mathbb{S}^n})$ and $\tau = 2$, the formula is same as the one in [28, Lemma 6.3(b)] by noting that $z(x, y) = (x^2 + y^2)^{1/2}$ and

$$\nabla_x \psi_\tau(x, y) = \mathcal{L}(x)\mathcal{L}^{-1}\big(z(x, y)\big)\phi_{FB}(x, y) - \phi_{FB}(x, y)$$

$$= x \circ \big[\mathcal{L}^{-1}\big(z(x, y)\big)\phi_{FB}(x, y)\big] - \mathcal{L}\big(z(x, y)\big)\mathcal{L}^{-1}\big(z(x, y)\big)\phi_{FB}(x, y)$$

$$= x \circ \big[\mathcal{L}^{-1}\big(z(x, y)\big)\phi_{FB}(x, y)\big] - z(x, y) \circ \big[\mathcal{L}^{-1}\big(z(x, y)\big)\phi_{FB}(x, y)\big]$$

$$= \big[\mathcal{L}^{-1}\big(z(x, y)\big)\phi_{FB}(x, y)\big] \circ \big(x - z(x, y)\big).$$

Thus, the formula (26) provides a unified framework for the SOCCP and the SDCP cases.

From Proposition 3.1, we readily obtain the following properties of $\nabla \psi_\tau$, which have been given in the setting of NCP [14] and the SOCCP [3], respectively.

Proposition 3.2. *Let ψ_τ be given as in (6). Then, for any $(x, y) \in \mathbb{V} \times \mathbb{V}$, we have*

(a) $\langle x, \nabla_x \psi_\tau(x, y) \rangle + \langle y, \nabla_y \psi_\tau(x, y) \rangle = \|\phi_\tau(x, y)\|^2$.
(b) $\nabla \psi_\tau(x, y) = 0$ *if and only if $x \in \mathcal{K}$, $y \in \mathcal{K}$, $\langle x, y \rangle = 0$.*

Proof. (a) If $(x, y) = (0, 0)$, the result is clear. Otherwise, from (26) it follows that

$$\langle x, \nabla_x \psi_\tau(x, y) \rangle + \langle y, \nabla_y \psi_\tau(x, y) \rangle = \left\langle x, \left(x + \frac{\tau - 2}{2}y\right) \circ \big[\mathcal{L}^{-1}\big(z(x, y)\big)\phi_\tau(x, y)\big]\right\rangle - \langle x, \phi_\tau(x, y) \rangle$$

$$+ \left\langle y, \left(y + \frac{\tau - 2}{2}x\right) \circ \big[\mathcal{L}^{-1}\big(z(x, y)\big)\phi_\tau(x, y)\big]\right\rangle - \langle y, \phi_\tau(x, y) \rangle$$

$$= \left\langle x \circ \left(x + \frac{\tau - 2}{2}y\right), \mathcal{L}^{-1}\big(z(x, y)\big)\phi_\tau(x, y)\right\rangle - \langle x, \phi_\tau(x, y) \rangle$$

$$+ \left\langle y \circ \left(y + \frac{\tau - 2}{2}x\right), \mathcal{L}^{-1}\big(z(x, y)\big)\phi_\tau(x, y)\right\rangle - \langle y, \phi_\tau(x, y) \rangle$$

$$= \left\langle z^2(x, y), \mathcal{L}^{-1}\big(z(x, y)\big)\phi_\tau(x, y)\right\rangle - \langle x + y, \phi_\tau(x, y) \rangle$$

$$= \langle z(x, y), \phi_\tau(x, y) \rangle - \langle x + y, \phi_\tau(x, y) \rangle = \|\phi_\tau(x, y)\|^2,$$

where the next to last equality is by $z^2 = \mathcal{L}(z)z$ and the symmetry of $\mathcal{L}(z)$.

(b) The proof is direct by part (a), Lemma 3.1 and Proposition 3.1. $\square$

4. Lipschitz continuity of $\nabla \psi_\tau$

In this section, we investigate the continuity of the gradients $\nabla_x \psi_\tau(x, y)$ and $\nabla_y \psi_\tau(x, y)$. To this end, for any $\epsilon > 0$, we define the mapping $z_\epsilon : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ by

$$z_\epsilon = z_\epsilon(x, y) := \big(x^2 + y^2 + (\tau - 2)(x \circ y) + \epsilon e\big)^{1/2}. \tag{31}$$

From (19), clearly, $z_\epsilon(x, y) \in \text{int}(\mathcal{K})$ for any $x, y \in \mathbb{V}$, and hence the operator $\mathcal{L}(z_\epsilon(x, y))$ is positive definite on $\mathbb{V}$. Since the spectral function induced by $\varphi(t) = \sqrt{t}$ ($t \geqslant 0$) is continuous by Lemma 2.1, it follows that $z_\epsilon(x, y) \to z(x, y)$ as $\epsilon \to 0^+$ for any $(x, y) \in \mathbb{V} \times \mathbb{V}$, where $z(x, y)$ is given by (25). This means that $\mathcal{L}(z_\epsilon(x, y)) \to \mathcal{L}(z(x, y))$ as $\epsilon \to 0^+$.

In what follows, we prove that the gradients $\nabla_x \psi_\tau(x, y)$ and $\nabla_y \psi_\tau(x, y)$ are Lipschitz continuous by arguing the Lipschitz continuity of $z_\epsilon(x, y)$ and the mapping

$$H_\epsilon(x, y) := \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y). \tag{32}$$

To establish the Lipschitz continuity of $z_\epsilon(x, y)$, we need the following crucial lemma.

Lemma 4.1. *For any $(x, y) \in \mathbb{V} \times \mathbb{V}$ and $\epsilon > 0$, let $z_\epsilon(x, y)$ be defined as in (31). Then the function $z_\epsilon(x, y)$ is continuously differentiable everywhere with*

$$\nabla_x z_\epsilon(x, y) = \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big),$$

$$\nabla_y z_\epsilon(x, y) = \mathcal{L}\left(y + \frac{\tau - 2}{2}x\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big). \tag{33}$$

Furthermore, there exists a constant $C > 0$, independent of x, y and ϵ, τ, such that

$$\big\|\nabla_x z_\epsilon(x, y)\big\| \leqslant C \quad and \quad \big\|\nabla_y z_\epsilon(x, y)\big\| \leqslant C.$$

Proof. The first part follows from Lemma 3.2 and the following fact that

$$z_\epsilon(x, y) = \left[\left(x + \frac{\tau - 2}{2} y\right)^2 + \frac{\tau(4 - \tau)}{4} y^2 + \epsilon e\right]^{1/2} = \left[\left(y + \frac{\tau - 2}{2} x\right)^2 + \frac{\tau(4 - \tau)}{4} x^2 + \epsilon e\right]^{1/2}. \tag{34}$$

We next prove that the operator $\nabla_x z_\epsilon(x, y)$ is bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$. Let $\{u_1, u_2, \ldots, u_n\}$ be an orthonormal basis of $\mathbb{V}$. For any $x, y \in \mathbb{V}$, let $L(z^2)$, $L(x + \frac{\tau-2}{2} y)$, $L(z_\epsilon)$ and $L((x + \frac{\tau-2}{2} y)^2)$ be the corresponding matrix representation of the operators $\mathcal{L}(z^2)$, $\mathcal{L}(x + \frac{\tau-2}{2} y)$, $\mathcal{L}(z_\epsilon)$ and $\mathcal{L}((x + \frac{\tau-2}{2} y)^2)$ with respect to the basis $\{u_1, u_2, \ldots, u_n\}$. Then, by the formula (33), it suffices to prove that the matrix $L(x + \frac{\tau-2}{2} y)L^{-1}(z_\epsilon)$ is bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$. The verifications are given as below.

Suppose that $z = z(x, y)$ has the spectral decomposition $z = \sum_{j=1}^r \lambda_j(z) c_j$, where $\lambda_1(z) \geqslant \lambda_2(z) \geqslant \cdots \geqslant \lambda_r(z) \geqslant 0$ are the eigenvalue of z and $\{c_1, c_2, \ldots, c_r\}$ is the corresponding Jordan frame. From Lemma 2.2, $\mathcal{L}(z)$ has the spectral decomposition

$$\mathcal{L}(z) = \sum_{j=1}^r \lambda_j(z) \mathcal{C}_{jj}(z) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{2}(\lambda_j(z) + \lambda_l(z)) \mathcal{C}_{jl}(z) \tag{35}$$

with the spectrum $\sigma(\mathcal{L}(z))$ consisting of all distinct numbers in $\{\frac{1}{2}(\lambda_j(z) + \lambda_l(z)): j, l = 1, 2, \ldots, r\}$, and $\mathcal{L}(z^2)$ has the spectral decomposition

$$\mathcal{L}(z^2) = \sum_{j=1}^r \lambda_j^2(z) \mathcal{C}_{jj}(z) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{2}(\lambda_j^2(z) + \lambda_l^2(z)) \mathcal{C}_{jl}(z) \tag{36}$$

with $\sigma(\mathcal{L}(z^2))$ consisting of all distinct numbers in $\{\frac{1}{2}(\lambda_j^2(z) + \lambda_l^2(z)): j, l = 1, 2, \ldots, r\}$. By the definition of $z_\epsilon(x, y)$, it is easy to verify that $z_\epsilon = \sum_{j=1}^r \sqrt{\lambda_j^2(z) + \epsilon} c_j$, and consequently the symmetric operator $\mathcal{L}(z_\epsilon)$ has the spectral decomposition

$$\mathcal{L}(z_\epsilon) = \sum_{j=1}^r \sqrt{\lambda_j^2(z) + \epsilon} \mathcal{C}_{jj}(z) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{2}\left(\sqrt{\lambda_j^2(z) + \epsilon} + \sqrt{\lambda_l^2(z) + \epsilon}\right) \mathcal{C}_{jl}(z) \tag{37}$$

with the spectrum $\sigma(\mathcal{L}(z_\epsilon))$ consisting of all distinct numbers in

$$\left\{\frac{1}{2}\left(\sqrt{\lambda_j^2(z) + \epsilon} + \sqrt{\lambda_l^2(z) + \epsilon}\right): j, l = 1, 2, \ldots, r\right\}.$$

We first prove that the matrix $L(x + \frac{\tau-2}{2} y)(L(z^2) + \epsilon I)^{-1/2}$ is bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$. For this purpose, let P be an $n \times n$ orthogonal matrix such that

$$PL(z^2)P^T = \text{diag}(\lambda_1(L(z^2)), \lambda_2(L(z^2)), \ldots, \lambda_n(L(z^2))), \tag{38}$$

where $\lambda_1(L(z^2)) \geqslant \lambda_2(L(z^2)) \cdots \geqslant \lambda_n(L(z^2)) \geqslant 0$ are the eigenvalues of $L(z^2)$. Then, it is not hard to verify that for any $\epsilon > 0$,

$$P(L(z^2) + \epsilon I)^{-1/2} P^T = \text{diag}\left(\frac{1}{\sqrt{\lambda_1(L(z^2)) + \epsilon}}, \ldots, \frac{1}{\sqrt{\lambda_n(L(z^2)) + \epsilon}}\right).$$

Write $\widetilde{U} := PL(x + \frac{\tau-2}{2} y)P^T$. We can compute that

$$L\left(x + \frac{\tau - 2}{2} y\right)(L(z^2) + \epsilon I)^{-1/2} = P^T \widetilde{U} \, \text{diag}\left(\frac{1}{\sqrt{\lambda_1(L(z^2)) + \epsilon}}, \ldots, \frac{1}{\sqrt{\lambda_n(L(z^2)) + \epsilon}}\right) P$$

$$= P^T \left[\frac{\widetilde{U}_{ik}}{\sqrt{\lambda_k(L(z^2)) + \epsilon}}\right]_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant k \leqslant n}} P. \tag{39}$$

Since $\mathcal{L}(z^2) = \mathcal{L}((x + \frac{\tau-2}{2} y)^2) + \mathcal{L}(\frac{\tau(4-\tau)}{4} y^2)$ and $\mathcal{L}(y^2)$ is positive semidefinite, we get

$$L(z^2) - L\left(\left(x + \frac{\tau - 2}{2} y\right)^2\right) \succcurlyeq 0.$$

In addition, by Proposition 2.1 $\mathcal{L}[(x + \frac{\tau-2}{2} y)^2] - \mathcal{L}(x + \frac{\tau-2}{2} y)\mathcal{L}(x + \frac{\tau-2}{2} y)$ is positive semidefinite, and hence we have that

$$L\left(\left(x + \frac{\tau - 2}{2} y\right)^2\right) - L\left(x + \frac{\tau - 2}{2} y\right) L\left(x + \frac{\tau - 2}{2} y\right) \succcurlyeq 0.$$

The last two equations thus imply that

$$L(z^2) - L\left(x + \frac{\tau - 2}{2}y\right)L\left(x + \frac{\tau - 2}{2}y\right) \succeq 0. \tag{40}$$

Now, for any given $k \in \{1, 2, \ldots, n\}$, from (38) and (40) it follows that

$$\lambda_k(L(z^2)) = [PL(z^2)P^T]_{kk} \geqslant \left[PL\left(x + \frac{\tau - 2}{2}y\right)L\left(x + \frac{\tau - 2}{2}y\right)P^T\right]_{kk}$$

$$= [\widetilde{U}\widetilde{U}]_{kk} = \sum_{i=1}^{n} \widetilde{U}_{ik}^2,$$

where the inequality is by the fact that the diagonal entries of a positive semidefinite matrix are nonnegative. This immediately implies that

$$\sqrt{\lambda_k(L(z^2)) + \epsilon} \geqslant \sqrt{\sum_{i=1}^{n} \widetilde{U}_{ik}^2} \geqslant \widetilde{U}_{ik} \quad \forall i = 1, 2, \ldots, n.$$

Combining with Eq. (39), there exists a constant $C_1 > 0$ such that

$$\left\| L\left(x + \frac{\tau - 2}{2}y\right)(L(z^2) + \epsilon I)^{-1/2} \right\| \leqslant C_1 \quad \forall x, y \in \mathbb{V} \text{ and } \epsilon > 0. \tag{41}$$

We next prove that the matrix $(L(z^2) + \epsilon I)^{1/2}L^{-1}(z_\epsilon)$ is bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$. Let $C_{jl}(z)$ for $1 \leqslant j, l \leqslant r$ be the matrix representation of $\mathcal{C}_{jl}(z)$ with respect to the basis $\{u_1, u_2, \ldots, u_n\}$. From Eqs. (36)–(37), it then follows that

$$(L(z^2) + \epsilon I)^{1/2} = \sum_{j=1}^{r} \sqrt{\lambda_j^2(z) + \epsilon}\, C_{jj}(z) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{2}\sqrt{2(\lambda_j^2(z) + \lambda_l^2(z) + 2\epsilon)}\, C_{jl}(z),$$

$$L^{-1}(z_\epsilon) = \sum_{j=1}^{r} \frac{1}{\sqrt{\lambda_j^2(z) + \epsilon}}\, C_{jj}(z) + \sum_{1 \leqslant j < l \leqslant r} \frac{1}{(\sqrt{\lambda_j^2(z) + \epsilon} + \sqrt{\lambda_l^2(z) + \epsilon})/2}\, C_{jl}(z).$$

Using the last two equalities and (15), it is easy to compute that

$$(L(z^2) + \epsilon I)^{1/2}L^{-1}(z_\epsilon) = \sum_{j=1}^{r} C_{jj}(z) + \sum_{1 \leqslant j < l \leqslant r} \frac{\sqrt{2(\lambda_j^2(z) + \lambda_l^2(z) + 2\epsilon)}}{\sqrt{\lambda_j^2(z) + \epsilon} + \sqrt{\lambda_l^2(z) + \epsilon}}\, C_{jl}(z). \tag{42}$$

Notice that the projection matrix $C_{jl}(z)$ with $1 \leqslant j, l \leqslant r$ is bounded for any $x, y \in \mathbb{V}$, and for any $x, y \in \mathbb{V}$ and $\epsilon > 0$

$$\sqrt{\lambda_j^2(z) + \epsilon} + \sqrt{\lambda_l^2(z) + \epsilon} \geqslant \sqrt{\lambda_j^2(z) + \lambda_l^2(z) + 2\epsilon} \quad \forall 1 \leqslant j, l \leqslant r.$$

Hence, from (42) we can deduce that $(L(z^2) + \epsilon I)^{1/2}L^{-1}(z_\epsilon)$ is bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$, i.e., there exists a positive constant C_2 such that

$$\left\| (L(z^2) + \epsilon I)^{1/2}L^{-1}(z_\epsilon) \right\| \leqslant C_2 \quad \forall x, y \in \mathbb{V} \text{ and } \epsilon > 0. \tag{43}$$

Combining (43) and (41), we have that the matrix $L(x + \frac{\tau - 2}{2}y)L^{-1}(z_\epsilon)$ is bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$, because

$$\left\| L\left(x + \frac{\tau - 2}{2}y\right)L^{-1}(z_\epsilon) \right\| = \left\| L\left(x + \frac{\tau - 2}{2}y\right)[(L(z^2) + \epsilon I)^{-1/2}(L(z^2) + \epsilon I)^{1/2}]L^{-1}(z_\epsilon) \right\|$$

$$= \left\| \left[L\left(x + \frac{\tau - 2}{2}y\right)(L(z^2) + \epsilon I)^{-1/2}\right][(L(z^2) + \epsilon I)^{1/2}L^{-1}(z_\epsilon)] \right\|$$

$$\leqslant \left\| L\left(x + \frac{\tau - 2}{2}y\right)(L(z^2) + \epsilon I)^{-1/2} \right\| \cdot \left\| (L(z^2) + \epsilon I)^{1/2}L^{-1}(z_\epsilon) \right\|$$

$$\leqslant C_1 C_2 \quad \forall x, y \in \mathbb{V} \text{ and } \epsilon > 0.$$

Consequently, there exists a constant $C > 0$ such that $\|\nabla_x z_\epsilon(x, y)\| \leqslant C$ for any $x, y \in \mathbb{V}$ and $\epsilon > 0$. For the symmetry, $\|\nabla_y z_\epsilon(x, y)\| \leqslant C$ also holds for any $x, y \in \mathbb{V}$ and $\epsilon > 0$. From the discussions above, we see that the constant C is also independent of τ. $\square$

By Lemma 4.1 and the Mean-Value Theorem, we can establish the global Lipschitz continuity of the mapping $z_\epsilon(x, y)$, which is summarized in the following proposition.

Proposition 4.1. *For any $x, y \in \mathbb{V}$ and $\epsilon > 0$, let $z_\epsilon(x, y)$ be defined as in (31). Then the function $z_\epsilon(x, y)$ is globally Lipschitz continuous.*

Proof. For any $(x, y), (a, b) \in \mathbb{V} \times \mathbb{V}$, by the Mean-Value Theorem we have that

$$
\begin{aligned}
\left\| z_\epsilon(x, y) - z_\epsilon(a, b) \right\| &= \left\| \left[z_\epsilon(x, y) - z_\epsilon(a, y) \right] + \left[z_\epsilon(a, y) - z_\epsilon(a, b) \right] \right\| \\
&= \left\| \int_0^1 \nabla_x z_\epsilon \left(a + t(x - a), y \right)(x - a)\, dt + \int_0^1 \nabla_y z_\epsilon \left(a, b + t(y - b) \right)(y - b)\, dt \right\| \\
&\leqslant \sqrt{2} \int_0^1 \left\| \nabla_x z_\epsilon \left(a + t(x - a), y \right) \right\| \cdot \| x - a \|\, dt + \sqrt{2} \int_0^1 \left\| \nabla_y z_\epsilon \left(a, b + t(y - b) \right) \right\| \cdot \| y - b \|\, dt \\
&\leqslant \sqrt{2} C \left(\| x - a \| + \| y - b \| \right) \leqslant 2C \left\| (x, y) - (a, b) \right\|,
\end{aligned}
$$

where the last two inequalities are respectively by Lemma 4.1 and (11). This shows that the function $z_\epsilon(x, y)$ is globally Lipschitz continuous. □

Next, we focus on the Lipschitz continuity of the mapping H_ϵ given by (32). We achieve the goal by proving that $\nabla_x H_\epsilon(x, y)$ and $\nabla_y H_\epsilon(x, y)$ are bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$. To compute $\nabla_x H_\epsilon(x, y)$ and $\nabla_y H_\epsilon(x, y)$, the following lemma is needed.

Lemma 4.2. *For any $x, y \in \mathbb{V}$ and $\epsilon > 0$, let $h \in \mathbb{V}$ be such that $z_\epsilon^2(x, y) + h \in \mathcal{K}$ and write $w := [z_\epsilon^2(x, y) + h]^{1/2} - z_\epsilon(x, y)$. Then, $w = \frac{1}{2} \mathcal{L}^{-1}(z_\epsilon(x, y))h + o(\|h\|)$.*

Proof. From the definition of w, it immediately follows that

$$
\left[w + z_\epsilon(x, y) \right]^2 = z_\epsilon^2(x, y) + h,
$$

which is equivalent to saying that

$$
w^2 + 2w \circ z_\epsilon(x, y) = h \tag{44}
$$

or

$$
h = 2\mathcal{L}\left(z_\epsilon(x, y) \right)w + w^2. \tag{45}
$$

We claim that, as $\|h\| \to 0$, we must have $\|w\| \to 0$. Indeed, let $\|h\| \to 0$, then we obtain from (44) that $w^2 + 2w \circ z_\epsilon(x, y) = 0$. Adding $z_\epsilon^2(x, y)$ to both sides gives

$$
\left(w + z_\epsilon(x, y) \right)^2 = z_\epsilon^2(x, y).
$$

This, by the fact that $w + z_\epsilon(x, y) \in \mathcal{K}$ and $z_\epsilon(x, y) \in \text{int}(\mathcal{K})$, implies that

$$
w + z_\epsilon(x, y) = z_\epsilon(x, y),
$$

and hence $w = 0$. Since $\mathcal{L}(z_\epsilon(x, y))$ is invertible on $\mathbb{V}$ and $\|w\| \to 0$ as $\|h\| \to 0$, using the implicit function theorem and Eq. (45) yields that

$$
w_\epsilon = \frac{1}{2} \mathcal{L}^{-1}\left(z_\epsilon(x, y) \right)h + o\left(\|h\| \right).
$$

Consequently, the proof is completed. □

Lemma 4.3. *For any $x, y \in \mathbb{V}$ and $\epsilon > 0$, let $H_\epsilon(x, y)$ be given as in (32). Then, $H_\epsilon(x, y)$ is differentiable everywhere. Moreover, for any given $u, v \in \mathbb{V}$,*

$$
\begin{aligned}
\nabla_x H_\epsilon(x, y)u &= \left[\mathcal{L}^{-1}\left(z_\epsilon(x, y) \right)(x + y) \right] \circ u + \mathcal{L}\left(x + \frac{\tau - 2}{2} y \right) \mathcal{L}^{-1}\left(z_\epsilon(x, y) \right) \\
&\quad \times \left[u - \mathcal{L}^{-1}\left(z_\epsilon(x, y) \right)(x + y) \circ \mathcal{L}^{-1}\left(z_\epsilon(x, y) \right) \mathcal{L}\left(x + \frac{\tau - 2}{2} y \right)u \right],
\end{aligned}
$$

$$\nabla_y H_\epsilon(x, y)v = \frac{\tau - 2}{2}\left[\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y)\right] \circ v + \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)$$
$$\times\left[v - \mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y) \circ \mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\mathcal{L}\left(y + \frac{\tau - 2}{2}x\right)v\right]. \tag{46}$$

Proof. For any $x, y \in \mathbb{V}$ and any given $u, v \in \mathbb{V}$, let $x' = x + \frac{\tau-2}{2}y$, $y' = y + \frac{\tau-2}{2}x$ and

$$h := 2x' \circ u + 2y' \circ v + u^2 + v^2 + (\tau - 2)u \circ v.$$

It is easy to compute that

$$z_\epsilon^2(x, y) + h = (x + u)^2 + (y + v)^2 + (\tau - 2)\big[(x + u) \circ (y + v)\big] + \epsilon e$$
$$= z_\epsilon^2(x + u, y + v) \in \text{int}(\mathcal{K}).$$

Let

$$w := \big[z_\epsilon^2(x, y) + h\big]^{1/2} - z_\epsilon(x, y).$$

Then,

$$w + z_\epsilon(x, y) = \big[z_\epsilon^2(x, y) + h\big]^{1/2} = z_\epsilon(x + u, y + v) \in \text{int}(\mathcal{K}).$$

Applying Lemma 4.2 then yields that

$$w = \frac{1}{2}\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)h + o\big(\|(u, v)\|\big), \tag{47}$$

which implies that $w \to 0$ as $u \to 0$, $v \to 0$ and $w = O(\|(u, v)\|)$. Write

$$g := \mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y) \quad \text{and} \quad g + s := \mathcal{L}^{-1}\big(z_\epsilon(x, y) + w\big)(x + u + y + v). \tag{48}$$

We next express s in terms of g, w, u, v and $z_\epsilon(x, y)$. By (48), clearly,

$$\mathcal{L}\big(z_\epsilon(x, y)\big)g = x + y \quad \text{and} \quad \mathcal{L}\big(z_\epsilon(x, y) + w\big)(g + s) = x + u + y + v,$$

which in turn implies that

$$\mathcal{L}\big(z_\epsilon(x, y)\big)s = u + v - w \circ g - w \circ s,$$

and

$$s = \mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(u + v - w \circ g - w \circ s). \tag{49}$$

Using (47) and (49), we have that $\|s\| \to 0$ as $\|(u, v)\| \to 0$. This, together with (47), means that $w \circ s = o(\|w\|) = o(\|(u, v)\|)$. Therefore,

$$\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(w \circ s) = o\big(\|(u, v)\|\big)$$

and

$$s = \mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(u + v - w \circ g) + o\big(\|(u, v)\|\big).$$

Now, from the above discussions and the definition of H_ϵ, it follows that

$$H_\epsilon(x + u, y + v) - H_\epsilon(x, y) = \mathcal{L}\left(x + u + \frac{\tau - 2}{2}(y + v)\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y) + w\big)(x + u + y + v)$$
$$- \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y)$$
$$= \mathcal{L}\left(x + u + \frac{\tau - 2}{2}(y + v)\right)(g + s) - \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)g$$
$$= \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)s + \mathcal{L}\left(u + \frac{\tau - 2}{2}v\right)(g + s)$$
$$= \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\big[\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(u + v - g \circ w)\big] + \mathcal{L}\left(u + \frac{\tau - 2}{2}v\right)g + o\big(\|(u, v)\|\big)$$

$$= \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(u + v) - \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)$$
$$\times \left[\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y) \circ \big(\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\mathcal{L}(x')u + \mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\mathcal{L}(y')v\big)\right]$$
$$+ \mathcal{L}\left(u + \frac{\tau - 2}{2}v\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y) + o\big(\|(u, v)\|\big).$$

This means that H_ϵ is differentiable at the point (x, y). Also, the formulas of $\nabla_x H_\epsilon(x, y)u$ and $\nabla_x H_\epsilon(x, y)v$ are exactly given by (46). The proof is then completed. □

Lemma 4.4. *For any $x, y \in \mathbb{V}$ and $\epsilon > 0$, let $H_\epsilon(x, y)$ be defined as in (32). Then, for any given $u, v \in \mathbb{V}$, there exists a constant $C > 0$ independent of x, y and ϵ, τ such that*

$$\big\|\nabla_x H_\epsilon(x, y)u\big\| \leqslant C\tau^{-1}\|u\| \quad \text{and} \quad \big\|\nabla_y H_\epsilon(x, y)v\big\| \leqslant C\tau^{-1}\|v\|.$$

Proof. By Lemma 4.1, there exists a constant $\bar{C} > 0$ independent of x, y, ϵ, τ such that

$$\left\|\mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\right\| \leqslant \bar{C} \quad \text{and} \quad \left\|\mathcal{L}\left(y + \frac{\tau - 2}{2}x\right)\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\right\| \leqslant \bar{C}.$$

Hence, their adjoint operators $\mathcal{L}^{-1}(z_\epsilon(x, y))\mathcal{L}(x + \frac{\tau-2}{2}y)$ and $\mathcal{L}^{-1}(z_\epsilon(x, y))\mathcal{L}(y + \frac{\tau-2}{2}x)$ are also bounded for any $x, y \in \mathbb{V}$ and $\epsilon > 0$, i.e.,

$$\left\|\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\right\| \leqslant \bar{C} \quad \text{and} \quad \left\|\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\mathcal{L}\left(y + \frac{\tau - 2}{2}x\right)\right\| \leqslant \bar{C}.$$

Noting that

$$\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y) = \frac{2}{\tau}\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)\left[\mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)e + \mathcal{L}\left(y + \frac{\tau - 2}{2}x\right)e\right],$$

we also have

$$\big\|\mathcal{L}^{-1}\big(z_\epsilon(x, y)\big)(x + y)\big\| \leqslant 4\bar{C}\tau^{-1}. \tag{50}$$

Thus, by the formulas of $\nabla_x H_\epsilon(x, y)u$ and $\nabla_y H_\epsilon(x, y)v$, we get the desired result. □

Using Lemmas 4.3 and 4.4 and the same arguments as for Proposition 4.1, we obtain the global Lipschitz continuity of $H_\epsilon(x, y)$, which is summarized in the following proposition.

Proposition 4.2. *For any $x, y \in \mathbb{V}$ and $\epsilon > 0$, let $H_\epsilon(x, y)$ be defined as in (32). Then the function $H_\epsilon(x, y)$ is globally Lipschitz continuous with the Lipschitz constant being $C\tau^{-1}$, where $C > 0$ is independent of x, y and ϵ, τ.*

Now we are in a position to establish the Lipschitz continuity of $\nabla_x \psi_\tau$ and $\nabla_y \psi_\tau$.

Proposition 4.3. *The function ψ_τ has a Lipschitz continuous gradient with the Lipschitz constant being positive multiple of $1 + \tau^{-1}$, i.e., there exists a constant $C > 0$ such that*

$$\big\|\nabla_x \psi_\tau(x, y) - \nabla_x \psi_\tau(a, b)\big\| \leqslant C\big(1 + \tau^{-1}\big)\big\|(x, y) - (a, b)\big\|,$$
$$\big\|\nabla_y \psi_\tau(x, y) - \nabla_y \psi_\tau(a, b)\big\| \leqslant C\big(1 + \tau^{-1}\big)\big\|(x, y) - (a, b)\big\|$$

for any $(x, y), (a, b) \in \mathbb{V} \times \mathbb{V}$, where C is independent of $(x, y), (a, b)$ and ϵ, τ.

Proof. For the symmetry, we only need to prove the first inequality. By (26),

$$\nabla_x \psi_\tau(x, y) = 2x + \frac{\tau}{2}y - z(x, y) - \mathcal{L}\left(x + \frac{\tau - 2}{2}y\right)\mathcal{L}^{-1}\big(z(x, y)\big)(x + y).$$

For any $\epsilon > 0$, let $G_\epsilon : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ be the mapping defined by

$$G_\epsilon(x, y) := 2x + \frac{\tau}{2}y - z_\epsilon(x, y) - H_\epsilon(x, y).$$

Then, from Propositions 4.1 and 4.2, it follows that $G_\epsilon(x, y)$ is globally Lipschitz continuous with the Lipschitz constant being $C(1 + \tau^{-1})$, i.e., for all $(x, y), (a, b) \in \mathbb{V} \times \mathbb{V}$,

$$\left\| G_\epsilon(x, y) - G_\epsilon(a, b) \right\| \leqslant C\left(1 + \tau^{-1}\right)\left\| (x, y) - (a, b) \right\|. \tag{51}$$

We next show that for any $(x, y) \in \mathbb{V} \times \mathbb{V}$, $G_\epsilon(x, y) \to \nabla_x \psi_\tau(x, y)$ as $\epsilon \to 0^+$. Indeed, if $(x, y) = (0, 0)$, then $H_\epsilon(0, 0) \to 0$ by (32) and (50), and so $G_\epsilon(0, 0) \to 0 = \nabla_x \psi_\tau(0, 0)$. If $(x, y) \neq (0, 0)$, noting that $z_\epsilon(x, y) \to z(x, y)$ as $\epsilon \to 0^+$, it suffices to prove that

$$\mathcal{L}^{-1}\left(z_\epsilon(x, y)\right)(x + y) \to \mathcal{L}^{-1}\left(z(x, y)\right)(x + y). \tag{52}$$

If $(x, y) \neq (0, 0)$ such that $z \in \text{int}(\mathcal{K})$, then $\mathcal{L}(z)$ is positive definite on $\mathbb{V}$. If $(x, y) \neq (0, 0)$ such that $z \notin \text{int}(\mathcal{K})$, then from Section 3 it follows that $\mathcal{L}(z)$ is positive definite on the subspace $\mathbb{V}(c_J, 1)$. By the proof of [13, Lemma 4.1(ii)], $\mathcal{L}^{-1}(z)$ is then continuous on $\mathbb{V}$ or $\mathbb{V}(c_J, 1)$, which implies the result of (52). Thus, we show that $G_\epsilon(x, y) \to \nabla_x \psi_\tau(x, y)$ as $\epsilon \to 0^+$ for any $(x, y) \in \mathbb{V} \times \mathbb{V}$. Now taking $\epsilon \to 0^+$ in (51) and applying the relation between $G_\epsilon(x, y)$ and $\nabla_x \psi_\tau(x, y)$ shown as above, we get the desired result. $\quad\square$

Propositions 3.1 and 4.3 show that the one-parametric class of merit functions ψ_τ is smooth everywhere on $\mathbb{V} \times \mathbb{V}$ and has a Lipschitz continuous gradient with the Lipschitz constant related to τ^{-1}. From this, we particularly obtain that the squared norm of the FB function ϕ_{FB} in (5) is also a smooth merit function with a global Lipschitz continuous gradient, generalizing the main results of [4,5].

5. Global error bound

From this section, we start with the case that $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ is a general Euclidean Jordan algebra. By Propositions III.4.4 and III.4.5 and Theorem V.3.7 of [9], we know that for a given Euclidean Jordan algebra $\mathbb{A}$ and the corresponding symmetric cone $\mathcal{K}$,

$$\mathbb{V} = \mathbb{V}_1 \times \mathbb{V}_2 \times \cdots \times \mathbb{V}_m \quad \text{and} \quad \mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2 \times \cdots \times \mathcal{K}^m,$$

with each $\mathbb{A}_i = (\mathbb{V}_i, \circ, \langle \cdot, \cdot \rangle)$ being a simple Euclidean Jordan algebra where $\dim(\mathbb{V}_i) = n_i$ and the corresponding symmetric cone is $\mathcal{K}^i$. Moreover, for any $x = (x_1, \ldots, x_m)$, $y = (y_1, \ldots, y_m) \in \mathbb{V}$ with $x_i, y_i \in \mathbb{V}_i$, one also has

$$x \circ y = (x_1 \circ y_1, \ldots, x_m \circ y_m) \quad \text{and} \quad \langle x, y \rangle = \langle x_1, y_1 \rangle + \cdots + \langle x_m, y_m \rangle.$$

Therefore, the complementarity condition involved in the SCCP is equivalent to

$$x_i \in \mathcal{K}^i, \quad y_i \in \mathcal{K}^i, \quad \langle x_i, y_i \rangle = 0 \quad \text{for all } i = 1, 2, \ldots, m,$$

and the functions $\phi_\tau(x, y)$ and $\psi_\tau(x, y)$ can be respectively rewritten as

$$\phi_\tau(x, y) = \left(\phi_\tau(x_1, y_1), \ldots, \phi_\tau(x_m, y_m)\right) \quad \text{and} \quad \psi_\tau(x, y) = \sum_{i=1}^{m} \psi_\tau(x_i, y_i).$$

The gradients of ψ_τ with respect to x and y then become respectively

$$\nabla_x \psi_\tau(x, y) = \left(\nabla_{x_1} \psi_\tau(x_1, y_1), \ldots, \nabla_{x_m} \psi_\tau(x_m, y_m)\right),$$
$$\nabla_y \psi_\tau(x, y) = \left(\nabla_{y_1} \psi_\tau(x_1, y_1), \ldots, \nabla_{y_m} \psi_\tau(x_m, y_m)\right).$$

By this, it is not hard to verify that all the results in Sections 3 and 4 are valid in the general Euclidean Jordan algebra $\mathbb{A}$ since they hold in each simple Euclidean Jordan algebra $\mathbb{A}_i$. In the sequel, corresponding to the structure of $\mathbb{V}$, we write the mappings G and F as $G = (G_1, \ldots, G_m)$, $F = (F_1, \ldots, F_m)$ with $G_i, F_i : \mathbb{V} \to \mathbb{V}_i$.

From the last two sections, we see that the SCCP is equivalent to the global minimization of the smooth functions f_τ or $\hat{f}_\tau$. In this section, we show that the regularized merit function $\hat{f}_\tau$ can provide a global error bound for the solution of the SCCP under a different condition from [21]. To the end, we need the concepts of Cartesian P-properties introduced in [17] for a nonlinear transformation, which are natural extensions of the P-properties on Cartesian products in $\mathbb{R}^n$ established by Facchinei and Pang [11] and the Cartesian P-properties in the setting of $\mathbb{S}^n$ developed by Chen and Qi [6].

Definition 5.1. The mappings $G = (G_1, \ldots, G_m)$ and $F = (F_1, \ldots, F_m)$ are said to have

(a) the joint uniform Cartesian P-property if there exists a constant $\rho > 0$ such that, for every $\zeta, \xi \in \mathbb{V}$, there is an index $i \in \{1, 2, \ldots, m\}$ such that

$$\left\langle G_i(\zeta) - G_i(\xi), F_i(\zeta) - F_i(\xi) \right\rangle \geqslant \rho \|\zeta - \xi\|^2;$$

(b) the joint Cartesian P-property if for every $\zeta, \xi \in \mathbb{V}$ with $\zeta \neq \xi$, there is an index $i \in \{1, 2, \ldots, m\}$ such that

$$\zeta_i \neq \xi_i \quad \text{and} \quad \left\langle G_i(\zeta) - G_i(\xi), F_i(\zeta) - F_i(\xi) \right\rangle > 0.$$

To establish the main result of this section, we also need the following two lemmas. Since the proof of the first lemma is similar to that of Lemma 7 in [4], we here omit it.

Lemma 5.1. *For any $x \in \mathbb{V}$, let $(x)_+$ and $(x)_-$ be the metric projection of x onto $\mathcal{K}$ and $-\mathcal{K}$, respectively. Then, the following results hold*:

(a) $x = (x)_+ + (x)_-$, $\langle (x)_+, (x)_- \rangle = 0$, and $\|x\|^2 = \|(x)_+\|^2 + \|(x)_-\|^2$.
(b) *For any $x \in \mathcal{K}$ and $y \in \mathbb{V}$ with $x^2 - y^2 \in \mathcal{K}$, we have $x - y \in \mathcal{K}$.*
(c) *For any $x \in \mathbb{V}$ and $y \in \mathcal{K}$, we have that $\langle x, y \rangle \leqslant \langle (x)_+, y \rangle$ and $\|(x + y)_+\| \geqslant \|(x)_+\|$.*

Lemma 5.2. *For any $x, y \in \mathbb{V}$, let ψ_τ be defined as in (6). Then,*

$$4\psi_\tau(x, y) \geqslant 2\left\|\left[\phi_\tau(x, y)\right]_+\right\|^2 \geqslant \frac{4 - \tau}{2}\left[\left\|(-x)_+\right\|^2 + \left\|(-y)_+\right\|^2\right].$$

Proof. The first inequality is due to Lemma 5.1(a) and the definition of ψ_τ. We next prove the second inequality. From (19) and Lemma 5.1(b), it follows that

$$\left[x^2 + y^2 + (\tau - 2)(x \circ y)\right]^{1/2} - \left(x + \frac{\tau - 2}{2}y\right) \in \mathcal{K},$$

$$\left[x^2 + y^2 + (\tau - 2)(x \circ y)\right]^{1/2} - \left(y + \frac{\tau - 2}{2}x\right) \in \mathcal{K}.$$

Combining with Lemma 5.1(c), we then obtain that

$$2\left\|\left[(x^2 + y^2 + (\tau - 2)x \circ y)^{1/2} - x - y\right]_+\right\|^2 = \left\|\left[(x^2 + y^2 + (\tau - 2)(x \circ y))^{1/2} - \left(x + \frac{\tau - 2}{2}y\right) - \frac{4 - \tau}{2}y\right]_+\right\|^2$$

$$+ \left\|\left[(x^2 + y^2 + (\tau - 2)(x \circ y))^{1/2} - \left(y + \frac{\tau - 2}{2}x\right) - \frac{4 - \tau}{2}x\right]_+\right\|^2$$

$$\geqslant \frac{4 - \tau}{2}\left\|(-y)_+\right\|^2 + \frac{4 - \tau}{2}\left\|(-x)_+\right\|^2.$$

Thus, we complete the proof of the second inequality. □

Now we prove that the regularized merit function $\hat{f}_\tau$ provides a global error bound for the solution of the SCCP if G and F have the joint uniform Cartesian P-property.

Proposition 5.1. *Let $\hat{f}_\tau$ be defined as in (8)–(9). Suppose that G and F have the joint uniform Cartesian P-property and the SCCP has a solution, denoted by ζ^*. Then, there exists a constant $\kappa > 0$ such that for any $\zeta \in \mathbb{V}$,*

$$\kappa\|\zeta - \zeta^*\|^2 \leqslant \beta^{-1}\hat{f}_\tau(\zeta) + \frac{4}{\sqrt{4 - \tau}}\left(\hat{f}_\tau(\zeta)\right)^{1/2}. \tag{53}$$

Proof. Since G and F have the joint uniform Cartesian P-property, there exists a constant $\rho > 0$ such that, for any $\zeta \in \mathbb{V}$, there is an index $i \in \{1, 2, \ldots, m\}$ such that

$$\rho\|\zeta - \zeta^*\|^2 \leqslant \langle G_i(\zeta) - G_i(\zeta^*), F_i(\zeta) - F_i(\zeta^*) \rangle = \langle G_i(\zeta), F_i(\zeta) \rangle + \langle F_i(\zeta^*), -G_i(\zeta) \rangle + \langle -F_i(\zeta), G_i(\zeta^*) \rangle$$

$$\leqslant \langle G_i(\zeta), F_i(\zeta) \rangle + \langle [-G_i(\zeta)]_+, F_i(\zeta^*) \rangle + \langle [-F_i(\zeta)]_+, G_i(\zeta^*) \rangle$$

$$\leqslant \lambda_{\max}\left[G_i(\zeta) \circ F_i(\zeta)\right] + \|F_i(\zeta^*)\|\left\|[-G_i(\zeta)]_+\right\| + \|G_i(\zeta^*)\|\left\|[-F_i(\zeta)]_+\right\|$$

$$\leqslant \max\{1, \|G_i(\zeta^*)\|, \|F_i(\zeta^*)\|\}\left[\lambda_{\max}\left[(G_i(\zeta) \circ F_i(\zeta))_+\right] + \left\|[-G_i(\zeta)]_+\right\| + \left\|[-F_i(\zeta)]_+\right\|\right]$$

$$\leqslant \max\{1, \|G(\zeta^*)\|, \|F(\zeta^*)\|\}\left[\left\|[G_i(\zeta) \circ F_i(\zeta)]_+\right\| + \left\|[-G_i(\zeta)]_+\right\| + \left\|[-F_i(\zeta)]_+\right\|\right],$$

where the equality is since $\langle G_i(\zeta^*), F_i(\zeta^*) \rangle = 0$, the second inequality is due to Lemma 5.1(b), and the third one follows from Proposition 2.1 of [27] and the Cauchy–Schwartz inequality. Setting $\kappa := \frac{\rho}{\max\{1, \|G(\zeta^*)\|, \|F(\zeta^*)\|\}}$, we immediately obtain

$$\kappa\|\zeta - \zeta^*\|^2 \leqslant \left\|[G_i(\zeta) \circ F_i(\zeta)]_+\right\| + \left\|[-G_i(\zeta)]_+\right\| + \left\|[-F_i(\zeta)]_+\right\|.$$

From the conditions given by (9), clearly, for any $i \in \{1, 2, \ldots, m\}$ we have

$$\left\|[G_i(\zeta) \circ F_i(\zeta)]_+\right\| \leqslant \beta^{-1}\psi_0\left(G_i(\zeta) \circ F_i(\zeta)\right) \leqslant \beta^{-1}\hat{f}_\tau(\zeta).$$

In addition, applying Lemma 5.2, we obtain that

$$\big\|[-G_i(\zeta)]_+\big\| + \big\|[-F_i(\zeta)]_+\big\| \leqslant \sqrt{2}\big(\big\|[-G(\zeta)]_+\big\|^2 + \big\|[-F(\zeta)]_+\big\|^2\big)^{1/2} \leqslant \frac{4}{\sqrt{4-\tau}}\psi_\tau\big(G(\zeta), F(\zeta)\big)^{1/2}$$
$$\leqslant \frac{4}{\sqrt{4-\tau}}\big(\hat{f}_\tau(\zeta)\big)^{1/2}.$$

Combining the last three inequalities immediately yields the desired result. □

For the NCP, Kanzow and Kleinmichel [14] showed that the merit function f_τ can provide a global error bound if F is a Lipschitz continuous uniform P-function. But, for the regularized merit function $\hat{f}_\tau$, Proposition 5.1 indicates that the Lipschitz continuity of F is not needed. The main reason is that the former is based on the global error bound result of ψ_{NR} and the same growth behavior of ψ_τ and ψ_{NR}, whereas the regularized merit function $\hat{f}_\tau$ attains this goal via the regularizing term ψ_0 involved in it.

For the SCCP (3), by Proposition 3.2 and Corollary 3.1 of [27], the assumption for the existence of the solution x^* can be removed from Proposition 5.1, since the uniform Cartesian P-property implies the uniform Jordan P-property. In addition, we note that for the SCCP (3), a similar result was given in [21] under the assumption that F having the uniform P^*-property. This assumption is stronger under some cases than F having the uniform Cartesian P-property; for example, when F is a affine function $M\zeta + q$ and $\mathbb{A}$ is the Lorentz algebra with $\dim(\mathbb{V}) \geqslant 5$. In this case, Professor Defeng Sun showed that F has the uniform P^*-property if and only if M is positive definite. Clearly, the positive definiteness of M implies F having Cartesian P-property, but F having Cartesian P-property does not imply M being positive definite. For example, for $\mathbb{V} = \mathbb{V}_1 \times \mathbb{V}_2 \times \mathbb{V}_3$ with $\dim(\mathbb{V}_1) = \dim(\mathbb{V}_2) = 2$ and $\dim(\mathbb{V}_3) = 1$, let M be a block diagonal matrix composed of $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right)$ and 1. It is easy to verify that $F(\zeta) = M\zeta + q$ for any $q \in \mathbb{V}$ has the Cartesian P-property, but M is not positive definite.

6. Bounded level sets

In this section, we provide a condition to guarantee the boundedness of the level sets

$$L_{\hat{f}_\tau}(\gamma) := \big\{\zeta \in \mathbb{V} \,\big|\, \hat{f}_\tau(\zeta) \leqslant \gamma\big\},$$

for any $\gamma \geqslant 0$. Specifically, we will prove that the following condition is sufficient.

Condition 6.1. For any sequence $\{\zeta^k\} \subseteq \mathbb{V}$ such that

$$\big\|\zeta^k\big\| \to +\infty, \qquad \big\|[-G(\zeta^k)]_+\big\| < +\infty, \qquad \big\|[-F(\zeta^k)]_+\big\| < +\infty, \tag{54}$$

there holds that

$$\max_{1\leqslant i\leqslant m} \lambda_{\max}\big[G_i(\zeta^k) \circ F_i(\zeta^k)\big] \to +\infty. \tag{55}$$

Proposition 6.1. *If the mappings G and F satisfy Condition 6.1, then the level sets $L_{\hat{f}_\tau}(\gamma)$ of $\hat{f}_\tau$ for all $\gamma \geqslant 0$ are bounded.*

Proof. Assume on the contrary that there is an unbounded sequence $\{\zeta^k\} \subseteq L_{\hat{f}_\tau}(\gamma)$ for some $\gamma \geqslant 0$. Then, $f_\tau(\zeta^k) \leqslant \hat{f}_\tau(\zeta^k) \leqslant \gamma$ for all k. By Lemma 5.2, G and F satisfy (54). Hence, there is $\nu \in \{1, \dots, m\}$ such that $\lambda_{\max}[G_\nu(\zeta^k) \circ F_\nu(\zeta^k)] \to +\infty$. Noting that

$$\lambda_{\max}\big[G_\nu(\zeta^k) \circ F_\nu(\zeta^k)\big] \leqslant \lambda_{\max}\big[\big(G_\nu(\zeta^k) \circ F_\nu(\zeta^k)\big)_+\big] \leqslant \big\|[G_\nu(\zeta^k) \circ F_\nu(\zeta^k)]_+\big\| \leqslant \beta^{-1}\hat{f}_\tau(\zeta^k),$$

we have $\hat{f}_\tau(\zeta^k) \to +\infty$. This contradicts the fact that $\{\zeta^k\} \subseteq L_{\hat{f}_\tau}(\gamma)$. □

Condition 6.1 is rather weak to guarantee the boundedness of the level sets of $\hat{f}_\tau$. Indeed, the SCCP with the jointly monotone mappings and a strictly feasible point, or the SCCP with joint Cartesian R_{02}-property, all satisfy this condition. To demonstrate the fact, we need the following technical lemma which can be regarded as an extension of Lemma 9(b) of [4] to the setting of symmetric cones.

Lemma 6.1. *Let $\{x^k\} \subseteq \mathbb{V}$ be any sequence satisfying $\|x^k\| \to +\infty$. If the sequence $\{\lambda_{\min}(x^k)\}$ is bounded below, then $\langle (x^k)_+, \hat{x}\rangle \to +\infty$ for any $\hat{x} \in \mathrm{int}(\mathcal{K})$.*

Proof. For every k, let x^k have the spectral decomposition $x^k = \sum_{j=1}^{r} \lambda_j(x^k) q_j^k$ with $\{q_1^k, \ldots, q_r^k\}$ being the corresponding Jordan frame. Let $\hat{x}$ have the spectral decomposition $x = \sum_{j=1}^{r} \lambda_j(\hat{x}) c_j$ with $\{c_1, \ldots, c_r\}$ being the corresponding Jordan frame. Without loss of generality, suppose that $\lambda_{l_k}(x^k) = \lambda_{\max}(x^k)$, where $1 \leqslant l_k \leqslant r$. Then, for every k,

$$
\begin{aligned}
\langle (x^k)_+, \hat{x} \rangle &= \left\langle \sum_{j=1}^{r} (\lambda_j(x^k))_+ q_j^k, \sum_{j=1}^{r} \lambda_j(\hat{x}) c_j \right\rangle \geqslant \lambda_{\max}((x^k)_+) \lambda_{\min}(\hat{x}) \left\langle q_{l_k}^k, \sum_{j=1}^{r} c_j \right\rangle \\
&= \lambda_{\max}((x^k)_+) \lambda_{\min}(\hat{x}) \langle q_{l_k}^k, e \rangle,
\end{aligned}
\tag{56}
$$

where the inequality holds since $q_j^k, c_j \in \mathcal{K}$ and $\lambda_j((x^k)_+), \lambda_j(\hat{x}) \geqslant 0$ for all $j = 1, 2, \ldots, r$. Notice that $\|(x^k)_-\| < +\infty$ as $k \to \infty$ since $\{\lambda_{\min}(x^k)\}$ is bounded below. Using the fact that $\|x^k\|^2 = \|(x^k)_+\|^2 + \|(x^k)_-\|^2$ and $\|x^k\| \to +\infty$, we then have that $\|(x^k)_+\| \to +\infty$. This, together with $(x^k)_+ \in \mathcal{K}$, immediately implies that

$$
\lambda_{\max}((x^k)_+) \to +\infty.
\tag{57}
$$

Since $\{q_{l_k}^k\}$ is bounded, we assume (subsequencing if necessary) that $\lim_{k \to +\infty} q_{l_k}^k = q^*$. By the closedness of $\mathcal{K}$ and $\|q_{l_k}^k\| = 1$ for each k, we have $q^* \in \mathcal{K} \setminus \{0\}$. From Proposition I.1.4 of [9], it then follows that $\langle q^*, e \rangle > 0$ since $e \in \mathrm{int}(\mathcal{K})$. Thus, taking the limit on the both sides of (56) and using Eq. (57), we readily obtain that $\langle (x^k)_+, \hat{x} \rangle \to +\infty$. $\quad\square$

Definition 6.1. The mappings $G = (G_1, \ldots, G_m)$ and $F = (F_1, \ldots, F_m)$ are said to have the joint Cartesian R_{02}-property if for any sequence $\{\zeta^k\}$ satisfying the condition:

$$
\|\zeta^k\| \to +\infty, \qquad \frac{[-G(\zeta^k)]_+}{\|\zeta^k\|} \to 0, \qquad \frac{[-F(\zeta^k)]_+}{\|\zeta^k\|} \to 0,
\tag{58}
$$

there exists an index $\nu \in \{1, 2, \ldots, m\}$ such that

$$
\liminf_{k \to +\infty} \frac{\lambda_{\max}[G_\nu(\zeta^k) \circ F_\nu(\zeta^k)]}{\|\zeta^k\|} > 0.
$$

Proposition 6.2. *Condition* 6.1 *is satisfied if one of the following assumptions holds*:

(a) *G and F are jointly monotone mappings with $\|G(\zeta)\| + \|F(\zeta)\| \to +\infty$ as $\|\zeta\| \to +\infty$ and there exists a point $\hat{\zeta} \in \mathbb{V}$ such that $G(\hat{\zeta}), F(\hat{\zeta}) \in \mathrm{int}(\mathcal{K})$;*
(b) *G and F have the joint Cartesian R_{02}-property.*

Proof. (a) Let $\{\zeta^k\}$ be a sequence satisfying (54). Since G and F are jointly monotone,

$$
\langle G(\zeta^k) - G(\hat{\zeta}), F(\zeta^k) - F(\hat{\zeta}) \rangle \geqslant 0,
$$

which by Lemma 5.1(a) is equivalent to

$$
\langle G(\zeta^k), F(\zeta^k) \rangle + \langle G(\hat{\zeta}), F(\hat{\zeta}) \rangle \geqslant \langle [G(\zeta^k)]_+, F(\hat{\zeta}) \rangle + \langle [G(\zeta^k)]_-, F(\hat{\zeta}) \rangle + \langle [F(\zeta^k)]_+, G(\hat{\zeta}) \rangle + \langle [F(\zeta^k)]_-, G(\hat{\zeta}) \rangle.
$$

Notice that the sequences $\{\lambda_{\min}(G(\zeta^k))\}$ and $\{\lambda_{\min}(F(\zeta^k))\}$ are bounded below by (54), $\|G(\zeta^k)\| + \|F(\zeta^k)\| \to +\infty$ and $G(\hat{\zeta}), F(\hat{\zeta}) \in \mathrm{int}(\mathcal{K})$. Using Lemma 6.1 then yields that

$$
\langle [G(\zeta^k)]_+, F(\hat{\zeta}) \rangle + \langle [F(\zeta^k)]_+, G(\hat{\zeta}) \rangle \to +\infty.
$$

In addition, by (54) it is easy to verify that

$$
\langle [G(\zeta^k)]_-, F(\hat{\zeta}) \rangle > -\infty \quad \text{and} \quad \langle [F(\zeta^k)]_-, G(\hat{\zeta}) \rangle > -\infty.
$$

Therefore, from the last three equations it follows that

$$
\sum_{i=1}^{m} \langle G_i(\zeta^k), F_i(\zeta^k) \rangle = \langle G(\zeta^k), F(\zeta^k) \rangle \to +\infty,
$$

which in turn implies that there exists an index ν such that $\langle G_\nu(\zeta^k), F_\nu(\zeta^k) \rangle \to +\infty$. By Proposition 2.1(ii) of [27], we have $\lambda_{\max}[G_\nu(\zeta^k) \circ F_\nu(\zeta^k)] \to +\infty$, which implies (55).
   (b) The proof is direct by Definition 6.1. $\quad\square$

When $G(\zeta) = \zeta$ for any $\zeta \in \mathbb{V}$, Liu, Zhang and Wang [21] established the boundedness of the level sets of $L_{\hat{f}_\tau}(\gamma)$ for $\tau = 2$ under the condition that F is an R_{02}-function. The condition, as shown by Lemma 6.2 below, is stronger than the one of Proposition 6.2(b). Thus, Proposition 6.2(b) generalizes the result of [21, Theorem 7].

Lemma 6.2. *Assume that $G(\zeta) \equiv \zeta$ for any $\zeta \in \mathbb{V}$ and F is an R_{02}-function. Then, G and F have the joint Cartesian R_{02}-property.*

Proof. Suppose that F is an R_{02}-function. From Definition 3 of [21], for any sequence $\{\zeta^k\}$ satisfying the condition (58), there holds that

$$\liminf_{k \to +\infty} \frac{\lambda_{\max}[\zeta^k \circ F(\zeta^k)]}{\|\zeta^k\|^2} > 0. \tag{59}$$

For each k, let $z^k = \zeta^k \circ F(\zeta^k)$ and suppose that it has the spectral decomposition $z^k = \sum_{j=1}^r \lambda_j(z^k) c_j^k$, where $\{c_1^k, \ldots, c_r^k\} \subseteq \mathbb{V}$ is a Jordan frame. For convenience, we also denote $z^k = (z_1^k, \ldots, z_m^k)$ with $z_i^k \in \mathbb{V}_i$. By the spectral decomposition of z^k, clearly,

$$z_i^k = \sum_{j=1}^r \lambda_j(z^k)(c_j^k)_i, \quad i = 1, 2, \ldots, m, \tag{60}$$

with $c_j^k = ((c_j^k)_1, \ldots, (c_j^k)_m)$ for every $j \in \{1, 2, \ldots, r\}$. Now, without loss of generality we assume that $\lambda_{l_k}(z^k) = \lambda_{\max}(z^k)$ with $1 \leqslant l_k \leqslant r$. Then,

$$\lambda_{\max}(z^k) = \sum_{j=1}^r \lambda_j(z^k) \langle c_j^k, c_{l_k}^k \rangle = \sum_{i=1}^m \sum_{j=1}^r \lambda_j(z^k) \langle (c_j^k)_i, (c_{l_k}^k)_i \rangle.$$

Combining with (59) and (60), there exists an index $\nu \in \{1, 2, \ldots, m\}$ such that

$$0 < \liminf_{k \to +\infty} \frac{\sum_{j=1}^r \lambda_j(z^k) \langle (c_j^k)_\nu, (c_{l_k}^k)_\nu \rangle}{\|\zeta^k\|^2} = \liminf_{k \to +\infty} \frac{\langle z_\nu^k, (c_{l_k}^k)_\nu \rangle}{\|\zeta^k\|^2}. \tag{61}$$

Suppose that z_ν^k as an element in the simple Euclidean Jordan algebra $\mathbb{A}_\nu = (\mathbb{V}_\nu, \circ, \langle \cdot, \cdot \rangle)$ has the following spectral decomposition

$$z_\nu^k = \sum_{j=1}^{\bar{r}} \lambda_j(z_\nu^k) q_{\nu j}^k,$$

where $\{q_{\nu 1}^k, \ldots, q_{\nu \bar{r}}^k\} \subseteq \mathbb{V}_\nu$ be the corresponding Jordan frame. Then,

$$\langle z_\nu^k, (c_{l_k}^k)_\nu \rangle \leqslant \lambda_{\max}(z_\nu^k) \left\langle \sum_{j=1}^{\bar{r}} q_{\nu j}^k, (c_{l_k}^k)_\nu \right\rangle = \lambda_{\max}(z_\nu^k) \langle e_\nu, (c_{l_k}^k)_\nu \rangle, \tag{62}$$

where e_ν is the identity element in $\mathbb{V}_\nu$ and the inequality is since $(c_{l_k}^k)_\nu \in \mathcal{K}^\nu$ and $q_{\nu j}^k \in \mathcal{K}^\nu$ for every $j = 1, 2, \ldots, \bar{r}$. From (61) and (62), it then follows that

$$0 < \liminf_{k \to +\infty} \frac{\langle z_\nu^k, (c_{l_k}^k)_\nu \rangle}{\|\zeta^k\|^2} \leqslant \liminf_{k \to +\infty} \frac{\lambda_{\max}(z_\nu^k)}{\|\zeta^k\|} \cdot \frac{\langle e_\nu, (c_{l_k}^k)_\nu \rangle}{\|\zeta^k\|}.$$

Noting that $\langle e_\nu, (c_{l_k}^k)_\nu \rangle$ is bounded for each k, we have that

$$\liminf_{k \to +\infty} \frac{\langle e_\nu, (c_{l_k}^k)_\nu \rangle}{\|\zeta^k\|} = 0.$$

From the last two inequalities, it readily follows that

$$\liminf_{k \to +\infty} \frac{\lambda_{\max}(z_\nu^k)}{\|\zeta^k\|} > 0.$$

By Definition 6.1, the mappings G and F have the joint Cartesian R_{02}-property. $\quad \square$

7. Conclusions and final remarks

We have extended the one-parametric class of merit functions proposed in [14] for the NCP to the SCCP setting, which provides a unified framework for several existing merit functions for the SDCP and the SOCCP, for example, the squared norm of the matrix-valued FB function [28] and the vector-valued FB function [4]. We have established the smoothness and the global Lipschitz continuity of the class of merit functions, thereby generalizing some recent important results of [4,21,25,28]. In addition, we also considered a class of regularized merit function $\hat{f}_\tau$ based on ψ_τ. The class of regularized functions was

shown to provide a global error bound for the solution of SCCP and have bounded level sets under weaker conditions than those used by [21].

There is a natural question which is not yet answered in this paper: under what conditions every stationary point of f_τ is a solution of the SCCP. This is very important in the merit function approach and the semismooth Newton method for solving the SCCP. The main hurdle which prevents us from achieving this goal is that we cannot extend the first result of [10, Proposition 3.4] to the setting of symmetric cones. In particular, the following important property of $\nabla \psi_\tau$

$$\langle \nabla_x \psi_\tau(x, y), \nabla_y \psi_\tau(x, y) \rangle \geqslant 0 \quad \forall x, y \in \mathbb{V}, \tag{63}$$

a suitable characterization like [3, Proposition 4.1] for the SOCCP, cannot be proved yet. We suspect that this particular inequality will be useful in other context if it is proved. It deserves further investigation. To the contrast, inequality (63) can be easily verified for the Implicit Lagrangian merit function in SCCP case (see [17]), and conditions for each stationary point being a solution of the SCCP is also established therein. However, the analysis cannot be carried over to ψ_τ due to the more complicated structure of $\nabla \psi_\tau$.

References

[1] F. Alizadeh, D. Goldfarb, Second-order cone programming, Math. Program. 95 (2003) 3–51.
[2] M. Baes, Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras, Linear Algebra Appl. 422 (2007) 664–700.
[3] J.-S. Chen, S.-H. Pan, A one-parametric class of merit functions for the second-order cone complementarity problem, Comput. Optim. Appl. (2008), doi:10.1007/s10589-008-9182-9, in press.
[4] J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, Math. Program. 104 (2005) 293–327.
[5] J.-S. Chen, D. Sun, J. Sun, The SC^1 property of the squared norm of SOC Fischer–Burmeister function, Oper. Res. Lett. 36 (2008) 385–392.
[6] X. Chen, H.-D. Qi, Cartesian P-property and its applications to the semidefinite linear complementarity problem, Math. Program. 106 (2006) 177–201.
[7] L. Faybusovich, Euclidean Jordan algebras and interior point algorithm, Positivity 1 (1997) 331–357.
[8] L. Faybusovich, Linear systems in Jordan algebras and primal-dual interior point algorithm, J. Comput. Appl. Math. 86 (1997) 149–175.
[9] J. Faraut, A. Korányi, Analysis on Symmetric Cones, Oxford Math. Monogr., Oxford University Press, New York, 1994.
[10] M. Fukushima, Z.-Q. Luo, P. Tseng, Smoothing functions for second-order cone complementarity problems, SIAM J. Optim. 12 (2002) 436–460.
[11] F. Facchinei, J.-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, vols. I and II, Springer-Verlag, New York, 2003.
[12] M.S. Gowda, R. Sznajder, J. Tao, Some P-properties for linear transformations on Euclidean Jordan algebras, Linear Algebra Appl. 393 (2004) 203–232.
[13] Z.H. Huang, X.H. Liu, Extension of smoothing Newton algorithms to solve linear programming over symmetric cones, Technique Report, Department of Mathematics, School of Science, Tianjin University, PR China, 2007.
[14] C. Kanzow, H. Kleinmichel, A new class of semismooth Newton-type methods for nonlinear complementarity problems, Comput. Optim. Appl. 11 (1998) 227–251.
[15] M. Koecher, The Minnesota Notes on Jordan Algebras and Their Applications, edited and annotated by A. Brieg and S. Walcher, Springer, Berlin, 1999.
[16] L.C. Kong, J. Sun, N.H. Xiu, A regularized smoothing Newton method for symmetric cone complementarity problems, SIAM J. Optim. 19 (2008) 1028–1047.
[17] L.C. Kong, L. Tuncel, N.H. Xiu, Vector-valued implicit Lagrangian for symmetric cone complementarity problems, Asia-Pac. J. Oper. Res. (2008), in press.
[18] C. Kanzow, Y. Yamashita, M. Fukushima, New NCP functions and their properties, J. Optim. Theory Appl. 97 (1997) 115–135.
[19] Y.-D. Lim, Applications of geometric means on symmetric cones, Math. Ann. 319 (2001) 457–468.
[20] M.S. Lobo, L. Vandenberghe, S. Boyd, H. Lebret, Application of second-order cone programming, Linear Algebra Appl. 284 (1998) 193–228.
[21] Y. Liu, L. Zhang, Y. Wang, Some properties of a class of merit functions for symmetric cone complementarity problems, Asia-Pac. J. Oper. Res. 23 (2006) 473–496.
[22] O.L. Mangasarian, M.V. Solodov, Nonlinear complementarity as unconstrained and constrained minimization, Math. Program. 62 (1993) 277–297.
[23] S.H. Schmieta, F. Alizadeh, Associative and Jordan algebras, and polynomial time interior point algorithms for symmetric cones, Math. Oper. Res. 26 (2001) 543–564.
[24] S.H. Schmieta, F. Alizadeh, Extension of primal-dual interior point algorithms to symmetric cones, Math. Program. 96 (2003) 409–438.
[25] C.-K. Sim, J. Sun, D. Ralph, A note on the Lipschitz continuity of the gradient of the squared norm of the matrix-valued Fischer–Burmeister function, Math. Program. 107 (2006) 547–553.
[26] D. Sun, J. Sun, Löwner's operator and spectral functions on Euclidean Jordan algebras, Math. Oper. Res. 33 (2008) 421–445.
[27] J. Tao, M.S. Gowda, Some P-properties for nonlinear transformations on Euclidean Jordan algebras, Math. Oper. Res. 30 (2005) 985–1004.
[28] P. Tseng, Merit function for semidefinite complementarity problems, Math. Program. 83 (1998) 159–185.
[29] L. Vandenberghe, S. Boyd, A primal-dual potential reduction method for problems involving matrix inequalities, Math. Program. 69 (1995) 205–236.