

## Research Article

Ziran Yin\*, Xiaoyu Chen, and Jihong Zhang

# The robust isolated calmness of spectral norm regularized convex matrix optimization problems

<https://doi.org/10.1515/math-2025-0189>

received November 15, 2024; accepted July 6, 2025

**Abstract:** This article aims to provide a series of characterizations of the robust isolated calmness of the Karush-Kuhn-Tucker (KKT) mapping for spectral norm regularized convex optimization problems. By establishing the variational properties of the spectral norm function, we directly prove that the KKT mapping is isolated calm if and only if the strict Robinson constraint qualification (SRCQ) and the second-order sufficient condition (SOSC) hold. Furthermore, we obtain the crucial result that the SRCQ for the primal/dual problem and the SOSC for the dual/primal problem are equivalent. The obtained results can derive more equivalent conditions of the robust isolated calmness of the KKT mapping, thereby enriching the stability theory of spectral norm regularized optimization problems and enhancing the usability of isolated calmness in algorithm applications.

**Keywords:** isolated calmness, second-order sufficient condition, spectral norm, strict Robinson constraint qualification, critical cone

**MSC 2020:** 90C25, 90C31, 65K10

## 1 Introduction

Consider the spectral norm regularized matrix optimization problem

$$\begin{aligned} \min_{X \in \mathfrak{R}^{m \times n}} \quad & h(QX) + \langle C, X \rangle + \|X\|_2 \\ \text{s.t.} \quad & \mathcal{B}X - b \in \mathcal{P}, \end{aligned} \quad (1.1)$$

where  $h : \mathfrak{R}^d \rightarrow \mathfrak{R}$  is twice continuously differentiable convex function and is essentially strictly convex,  $Q : \mathfrak{R}^{m \times n} \rightarrow \mathfrak{R}^d$  and  $\mathcal{B} : \mathfrak{R}^{m \times n} \rightarrow \mathfrak{R}^l$  are two linear operators,  $C \in \mathfrak{R}^{m \times n}$  and  $b \in \mathfrak{R}^l$  are given data, and  $\mathcal{P} \subseteq \mathfrak{R}^l$  is a given convex polyhedra cone. The function  $\|X\|_2$  is the spectral norm function, namely, the largest singular value of  $X$ . Without loss of generality, we suppose that  $m \leq n$  in what follows. For convenience, in the following text, we always use  $\theta$  to represent the spectral norm function, i.e.,  $\theta(\cdot) = \|\cdot\|_2$ . Problem (1.1) has a wide range of applications in various fields, such as matrix approximation problems [1], matrix Chebyshev polynomial problems [2], and  $H_\infty$  synthesis problems [3–5], etc. Recently, spectral norm regularized problem has also been applied in deep learning and neural networks problems [6–8].

Stability analysis theory is very crucial in studying the convergence of algorithms. As one of the important concepts in stability properties, isolate calmness (Definition 1) may guarantee the linear convergence rate of

\* **Corresponding author: Ziran Yin**, School of Science, Dalian Maritime University, Dalian, Liaoning, 116026, P. R. China, e-mail: yinziran@dlnu.edu.cn

**Xiaoyu Chen:** School of Science, Dalian Maritime University, Dalian, Liaoning, 116026, P. R. China, e-mail: cxy11238556969@163.com

**Jihong Zhang:** School of Science, Shenyang Ligong University, Shenyang, Liaoning, 110159, P. R. China, e-mail: zjh7815040x@163.com

some algorithms, such as the alternating direction method of multipliers [9] and the proximal augmented Lagrangian method [10]. There are many publications on studying the isolated calmness of the Karush-Kuhn-Tucke (KKT) mapping (ICKKTM) for optimization problems. For the nonlinear semidefinite programming, Zhang and Zhang [11] show that the ICKKTM can be derived from the second-order sufficient condition (SOSC) and the strict Robinson constraint qualification (SRCQ). Zhang et al. [12] demonstrate that both the SOSC and the SRCQ are sufficient and necessary for the ICKKTM for the nonlinear second-order cone programming problem. It is important to note that Ding et al. [13] study the robust isolated calmness of a large class of cone programming problems. They prove that if and only if both the SOSC and the SRCQ hold, the KKT mapping is robustly isolated calm.

When the optimization problem has some kind of special linear structure, by establishing the equivalence between the constraint qualification of the primal (dual) problem and second-order optimality condition of the dual (primal) problem, one can obtain more characterizations of stability properties. For instance, for the standard semidefinite programming problem, Chan and Sun [14] show that the constraint nondegeneracy for the primal (dual) problem is equivalent to the strong SOSC for the dual (primal), then they obtain a series of equivalent characterizations of the strong regularity of the KKT point. Han et al. [15] discover the equivalence relationship between the dual (primal) SRCQ and the primal (dual) SOSC for convex composite quadratic semidefinite programming problem, thereby deriving a series of equivalent conditions of ICKKTM. For the case where  $\theta$  in problem (1.1) is the nuclear norm function, Cui and Sun [16] show that the primal (dual) SRCQ is equivalent to the dual (primal) SOSC. Therefore, they derive more equivalent conditions of the robust ICKKTM.

Inspired by the work in [16], given the widespread application of spectral norm regularized convex matrix optimization problems, a natural question is whether the results in [16] can be extended to spectral norm regularized convex optimization problems. We provide an positive answer in this article. Due to the special structure of the critical cone of spectral norm function, we provide several equivalent conditions of the robust ICKKTM. That is to say, the conclusions in [16] are still valid for problem (1.1). Compared to [16], the difference in this article is that we use the established variational properties of the proximal mapping of spectral norm to provide a direct proof of the ICKKTM for problem (1.1), while [16] provides an indirect proof of the isolated calmness based on [13, Theorem 24] regarding the optimization problems with a  $C^2$ -cone reducible constraint set. In addition, compared to the SOSC, the SRCQ for primal or dual problem is easier to verify, which can enhance the practicality of isolated calmness in algorithm research.

The organization of the subsequent content is as follows. We provide some notations and preliminaries on variational analysis, which will be used in the following text in Section 2. The variational properties of the spectral norm function are studied in Section 3, including the relationship between the critical cones of the spectral norm function and its conjugate function, and the explicit expression of the directional derivative of the proximal mapping of the spectral norm function. The results in Section 3 play a vital role in obtaining the important conclusions of this article. In Section 4, we prove that the primal (dual) SRCQ holds if and only if the dual (primal) SOSC holds and thus establish more equivalent conditions for the robust ICKKTM for problem (1.1). We conclude this article in Section 5.

Some common symbols and notations for matrices are as follows:

- For any positive integer  $t$ , we denote by  $[t]$  the index set  $\{1, \dots, t\}$ . For any  $Z \in \mathfrak{R}^{m \times n}$ , the  $(i, j)$ th entry of  $Z$  is denoted as  $Z_{ij}$ , where  $i \in [m]$ ,  $j \in [n]$ . Let  $\mu \subseteq [m]$  and  $\nu \subseteq [n]$  be two index sets. We write  $Z_\nu$  to be the submatrix of  $Z$  that only retains all columns in  $\nu$ , and  $Z_{\mu\nu}$  to be the  $|\mu| \times |\nu|$  submatrix of  $Z$  that only retains all rows in  $\mu$  and columns in  $\nu$ .
- For any  $d \in \mathfrak{R}^m$ ,  $\text{Diag}(d)$  represents the  $m \times m$  diagonal matrix, where the  $i$ th diagonal element is  $d_i$ ,  $i \in [m]$ .
- Using “trace” to represent the sum of diagonal elements in a given square matrix. For any two matrices  $P$  and  $Q$  in  $\mathfrak{R}^{m \times n}$ , the inner product of  $P$  and  $Q$  is written as  $\langle P, Q \rangle = \text{trace}(P^T Q)$ . The Hadamard product of matrices  $P$  and  $Q$  is represented by the symbol “ $\circ$ ”, i.e., the  $(i, j)$ th entry of  $P \circ Q \in \mathfrak{R}^{m \times n}$  is  $P_{ij}Q_{ij}$ .
- Let  $\mathcal{S}^w$  be the linear space of all  $w \times w$  real symmetric matrices, and  $\mathcal{S}_+^w$  and  $\mathcal{S}_-^w$  be the cones of all  $w \times w$  positive and negative semidefinite matrices, respectively.

## 2 Notation and preliminaries

Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two finite dimensional real Euclidean spaces. Let  $B_{\mathcal{Y}}$  be the unit ball in  $\mathcal{Y}$ . Let  $D \subseteq \mathcal{X}$  be a nonempty closed convex set. For any  $z \in D$ , the tangent cone [17, Definition 6.1] of  $D$  at  $z$  is defined by  $\mathcal{T}_D(z) := \{d \in \mathcal{X} | \exists z^k \rightarrow z \text{ with } z^k \in D \text{ and } \tau^k \downarrow 0 \text{ s.t. } (z^k - z)/\tau^k \rightarrow d\}$ . We use  $\mathcal{N}_D(x) := \{d \in \mathcal{X} : \langle d, z - x \rangle \leq 0, \forall z \in D\}$  to denote the normal cone of  $D$  at  $x \in D$ . Denote by  $\delta_D(x)$  the indicator function over  $D$ , i.e.,  $\delta_D(x) = 0$  if  $x \in D$ , and  $\delta_D(x) = \infty$  if  $x \notin D$ . Define the support function of the set  $D$  as  $\sigma(y, D) := \sup_{x \in D} \langle x, y \rangle, y \in \mathcal{X}$ . For a given  $x \in \mathcal{X}$ , define  $\Pi_D(x) := \arg \min\{\|d - x\| | d \in D\}$  as the projection mapping onto  $D$ . Suppose the function  $f: \mathcal{X} \rightarrow (-\infty, +\infty]$  is proper closed convex, denote by  $\text{dom} f := \{x | f(x) < \infty\}$  the effective domain of  $f$ , by  $f^*$  the conjugate of  $f$  and by  $\partial f$  the subdifferential of  $f$ . For more details, one can refer standard convex analysis [17]. For any convex cone  $K \subset \mathcal{Y}$ , denote by  $K^\circ := \{z' \in \mathcal{Y} | \langle z', z \rangle \leq 0, \forall z \in K\}$  the polar of  $K$ .

The definition of isolated calmness below, which is taken from [13, Definition 2], is the most important concept in this article.

**Definition 1.** The set-valued mapping  $G: \mathcal{X} \rightrightarrows \mathcal{Y}$  is said to be isolated calm at  $\bar{u}$  for  $\bar{v}$  if  $\bar{v} \in G(\bar{u})$ , and there exist a positive constant  $\kappa$  and neighborhoods  $\mathcal{U}$  of  $\bar{u}$  and  $\mathcal{V}$  of  $\bar{v}$  such that

$$G(u) \cap \mathcal{V} \subset \{\bar{v}\} + \kappa \|u - \bar{u}\| B_{\mathcal{Y}} \quad \forall u \in \mathcal{U}. \quad (2.1)$$

Moreover, we say that  $G$  is robustly isolated calm at  $\bar{u}$  for  $\bar{v}$  if (2.1) holds, and for each  $u \in \mathcal{U}$ ,  $G(u) \cap \mathcal{V} \neq \emptyset$ .

For any closed convex function  $g: \mathcal{X} \rightarrow (-\infty, +\infty]$ , we know from [18, Proposition 2.58] that  $g$  is directionally epidifferentiable. We use  $g^l(x; \cdot)$  to denote the directional epiderivatives of  $g$ . Further, if  $g^l(x; d)$  is finite for  $x \in \text{dom } g$  and  $d \in \mathcal{X}$ , we define the lower second-order directional epiderivative of  $g$  for any  $w \in \mathcal{X}$  by

$$g_-^{ll}(x; d, w) := \liminf_{\substack{\tau \downarrow 0 \\ w' \rightarrow w}} \frac{g(x + \tau d + \frac{1}{2} \tau^2 w') - g(x) - \tau g^l(x; d)}{\frac{1}{2} \tau^2}.$$

## 3 Variational analysis of the spectral norm function

In this section, we will provide a direct expression for the directional derivative of the proximal mapping of the spectral norm function. Then discuss the relationship between directional derivatives of the proximal mapping, the so-called sigma term and the critical cones.

Given an arbitrary matrix  $Q \in \mathfrak{R}^{m \times n}$ , let  $\sigma_1(Q) \geq \sigma_2(Q) \geq \dots \geq \sigma_m(Q)$  be the singular values of  $Q$ . Denote  $\sigma(Q) := (\sigma_1(Q), \sigma_2(Q), \dots, \sigma_m(Q))^T$ . For any integer  $p > 0$ , let  $O^p$  be the set of all  $p \times p$  orthogonal matrices. We assume that  $Q \in \mathfrak{R}^{m \times n}$  admits the singular value decomposition (SVD) as follows:

$$Q = U[\text{Diag}(\sigma(Q)) \quad 0]V^T = U[\text{Diag}(\sigma(Q)) \quad 0][V_1 \quad V_2]^T = U\text{Diag}(\sigma(Q))V_1^T, \quad (3.1)$$

where  $U \in O^m$  and  $[V_1 \quad V_2] \in O^n$  with  $V_1 \in \mathfrak{R}^{n \times m}$  and  $V_2 \in \mathfrak{R}^{n \times (n-m)}$ . Define the following three index sets

$$a := \{1 \leq i \leq m | \sigma_i(Q) > 0\}, \quad b := \{1 \leq i \leq m | \sigma_i(Q) = 0\}, \quad c := \{m+1, \dots, n\}. \quad (3.2)$$

For any two matrices  $P, W \in S^n$ , the well known Fan's inequality [19] takes the form

$$\langle P, W \rangle \leq \lambda(P)^T \lambda(W), \quad (3.3)$$

where  $\lambda(P)$  represents the eigenvalue vector of  $P$  whose elements are arranged in nonincreasing order.

For any  $P, W \in \mathfrak{R}^{m \times n}$ , by [20, Theorem 31.5], we know that  $W \in \partial\theta(P)$  (or, equivalently,  $P \in \partial\theta^*(W)$ ) holds if and only if

$$\text{Prox}_{\theta}(P + W) = P \Leftrightarrow \text{Prox}_{\theta^*}(P + W) = W,$$

where  $\text{Prox}_\theta : \mathfrak{R}^{m \times n} \rightarrow \mathfrak{R}^{m \times n}$  denote the proximal mapping of  $\theta$  [21, Definition 6.1], namely,

$$\text{Prox}_\theta(Z) = \arg \min_{Z' \in \mathfrak{R}^{m \times n}} \left\{ \theta(Z') + \frac{1}{2} \|Z' - Z\|^2 \right\}, \quad Z \in \mathfrak{R}^{m \times n},$$

and  $\text{Prox}_{\theta^*} : \mathfrak{R}^{m \times n} \rightarrow \mathfrak{R}^{m \times n}$  denote the proximal mapping of  $\theta^*$ . Denote  $Q = P + W$ . Let  $Q$  have the SVD as (3.1). Then by [21, Theorem 7.29 and Example 7.31], we have that

$$P = U[\text{Diag}(\sigma(P)) \quad 0]V^T, \quad W = U[\text{Diag}(\sigma(W)) \quad 0]V^T$$

and

$$\sigma(P) = \kappa(\sigma(Q)) \quad \text{and} \quad \sigma(W) = \sigma(Q) - \kappa(\sigma(Q)),$$

where  $\kappa : \mathfrak{R}^m \rightarrow \mathfrak{R}^m$  is the proximal mapping of the  $l_\infty$  norm (i.e., the maximum absolute value of the elements in a vector in  $\mathfrak{R}^m$ ).

Define  $\phi : \mathfrak{R} \rightarrow \mathfrak{R}$  as the following scalar function:

$$\phi(x) = \min\{x, \lambda^*\}, \quad x \in \mathfrak{R},$$

where  $\lambda^* > 0$  and satisfies  $\sum_{i=1}^m [\sigma_i(Q) - \lambda^*]_+ = 1$  (for any  $x \in \mathfrak{R}$ ,  $[x]_+$  means the nonnegative part of  $x$ ). Then, we can conclude that  $\kappa(\sigma) = (\phi(\sigma_1), \phi(\sigma_2), \dots, \phi(\sigma_m))$  and

$$P = \text{Prox}_\theta(Q) = U[\text{Diag}(\phi(\sigma_1), \phi(\sigma_2), \dots, \phi(\sigma_m)) \quad 0]V^T.$$

Obviously,  $\text{Prox}_\theta(\cdot)$  can be regarded as Löwner's operator related to  $\phi$ .

Let  $v_1(Q), v_2(Q), \dots, v_r(Q)$  be the nonzero different singular values of  $Q$ , and there exists an integer  $1 \leq \tilde{r} \leq r$  such that

$$v_1(Q) > v_2(Q) > \dots > v_{\tilde{r}}(Q) \geq \lambda^* > v_{\tilde{r}+1}(Q) > \dots > v_r(Q) > 0.$$

To proceed, we further define several index sets. We denote  $a_{r+1} = b$  for convenience. Divide the set  $a$  into subsets as follows:

$$a_k = \{i \in a \mid \sigma_i(Q) = v_k(Q)\}, \quad k \in [r],$$

and define two index sets associated with  $\lambda^*$ :

$$\alpha = \bigcup_{k=1}^{\tilde{r}} a_k, \quad \beta = \bigcup_{k=\tilde{r}+1}^{r+1} a_k. \quad (3.4)$$

Specifically,  $\sigma_i(P) = \min\{\sigma_i(Q), \lambda^*\}$ ,  $i = 1, \dots, m$ , namely,

$$\sigma_i(P) = \begin{cases} \lambda^*, & \text{if } i \in \alpha, \\ \sigma_i(Q), & \text{if } i \in \beta. \end{cases}$$

In fact,  $\lambda^*$  coincides with the largest singular value of  $P$ , namely,  $\lambda^* = \|P\|_2$ . According to Watson [22], the subdifferential of  $\theta$  can be characterized as follows:

$$\partial\theta(P) = \{U_\alpha H V_\alpha^T, \forall H \in \mathcal{S}_+^{|\alpha|}, \text{trace}(H) = 1\}. \quad (3.5)$$

Therefore, we know that  $\sigma_i(W) = 0$  when  $i \in \beta$ , and the set  $\alpha$  can be divided into three subsets as follows:

$$\alpha_1 = \{i \in \alpha \mid \sigma_i(W) = 1\}, \quad \alpha_2 = \{i \in \alpha \mid 0 < \sigma_i(W) < 1\}, \quad \alpha_3 = \{i \in \alpha \mid \sigma_i(W) = 0\}. \quad (3.6)$$

Hence, there exist integers  $0 \leq r_0 \leq 1$  and  $\tilde{r} - 1 \leq r_1 \leq \tilde{r}$  such that

$$\alpha_1 = \bigcup_{k=1}^{r_0} a_k, \quad \alpha_2 = \bigcup_{k=r_0+1}^{r_1} a_k, \quad \alpha_3 = \bigcup_{k=r_1+1}^{\tilde{r}} a_k.$$

Since the spectral norm function is globally Lipschitz continuous with modulus 1 and convex,  $\theta$  is directional differentiable at any point in  $\mathfrak{R}^{m \times n}$ . From (3.5), for any  $D \in \mathfrak{R}^{m \times n}$ , the directional derivative of  $\theta$  at  $P$

in the direction  $D$  can be expressed as follows:

$$\theta'(P; D) = \sup_{S \in \partial\theta(P)} \langle D, S \rangle = \|U_a^T D V_a\|_2. \quad (3.7)$$

For  $W \in \partial\theta(P)$ , or, equivalently,  $P \in \partial\theta^*(W)$ , the critical cone of  $\theta$  at  $P$  for  $W$  and the critical cone of  $\theta^*$  at  $W$  for  $P$  are defined as follows:

$$C_\theta(P, W) := \{D \in \Re^{m \times n} | \theta'(P; D) = \langle D, W \rangle\} \quad (3.8)$$

and

$$C_{\theta^*}(W, P) := \{D \in \Re^{m \times n} | (\theta^*)'(W; D) = \langle D, P \rangle\}. \quad (3.9)$$

Next we will give the expression for the directional derivative of  $\text{Prox}_\theta(\cdot)$ . By [23], we know that  $\text{Prox}_\theta(\cdot)$  is directional differentiable and the directional derivative can be obtained via the directional derivative of  $\phi$ . Clearly the directional derivative of  $\phi$  is

$$\phi'(p; h) = \begin{cases} 0, & \text{if } p > \lambda^*, \\ \min\{h, 0\}, & \text{if } p = \lambda^*, \\ h, & \text{if } p < \lambda^*. \end{cases}$$

For any  $Q \in \Re^{m \times n}$ , denote  $\Theta_{aa}^2 : \Re^{m \times n} \rightarrow \Re^{|\alpha| \times |\alpha|}$ ,  $\Theta_{a\beta}^1 : \Re^{m \times n} \rightarrow \Re^{|\alpha| \times |\beta|}$ ,  $\Theta_{a\beta}^2 : \Re^{m \times n} \rightarrow \Re^{|\alpha| \times |\beta|}$ , and  $\Theta_{ac} : \Re^{m \times n} \rightarrow \Re^{|\alpha| \times |c|}$  as follows:

$$\begin{cases} (\Theta_{aa}^2(Q))_{ij} = \frac{2\lambda^*}{\sigma_i(Q) + \sigma_j(Q)}, & i \in [|\alpha|], j \in [|\alpha|], \\ (\Theta_{a\beta}^1(Q))_{ij} = \frac{\lambda^* - \sigma_{j+|\alpha|}(Q)}{\sigma_i(Q) - \sigma_{j+|\alpha|}(Q)}, & i \in [|\alpha|], j \in [|\beta|], \\ (\Theta_{a\beta}^2(Q))_{ij} = \frac{\lambda^* + \sigma_{j+|\alpha|}(Q)}{\sigma_i(Q) + \sigma_{j+|\alpha|}(Q)}, & i \in [|\alpha|], j \in [|\beta|], \\ (\Theta_{ac}(Q))_{ij} = \frac{\lambda^*}{\sigma_i(Q)}, & i \in [|\alpha|], j \in [n - m]. \end{cases}$$

Define two linear matrix operators  $S : \Re^{p \times p} \rightarrow \mathcal{S}^p$  and  $T : \Re^{p \times p} \rightarrow \Re^{p \times p}$  by

$$S(H) = \frac{1}{2}(H + H^T) \quad \text{and} \quad T(H) = \frac{1}{2}(H - H^T), \quad \forall H \in \Re^{p \times p}. \quad (3.10)$$

For all  $Q \in \Re^{m \times n}$  and  $D \in \Re^{m \times n}$ , let  $D = [D_1 \ D_2]$  with  $D_1 \in \Re^{m \times m}$ , and  $D_2 \in \Re^{m \times (n-m)}$ . We define four matrix mappings  $\Xi_1 : \Re^{m \times n} \times \Re^{m \times n} \rightarrow \Re^{|\alpha| \times |\alpha|}$ ,  $\Xi_2 : \Re^{m \times n} \times \Re^{m \times n} \rightarrow \Re^{|\alpha| \times |\beta|}$ ,  $\Xi_3 : \Re^{m \times n} \times \Re^{m \times n} \rightarrow \Re^{|\beta| \times |\alpha|}$ , and  $\Xi_4 : \Re^{m \times n} \times \Re^{m \times n} \rightarrow \Re^{|\alpha| \times |c|}$  by

$$\begin{cases} \Xi_1(Q, D) = \Theta_{aa}^2(Q) \circ T(D_1)_{aa}, \\ \Xi_2(Q, D) = \Theta_{a\beta}^1(Q) \circ S(D_1)_{a\beta} + \Theta_{a\beta}^2(Q) \circ T(D_1)_{a\beta}, \\ \Xi_3(Q, D) = (\Theta_{a\beta}^1(Q))^T \circ S(D_1)_{\beta a} + (\Theta_{a\beta}^2(Q))^T \circ T(D_1)_{\beta a}, \\ \Xi_4(Q, D) = \Theta_{ac}(Q) \circ D_{ac}. \end{cases} \quad (3.11)$$

Therefore, according to [23, Theorem 3], the directional derivative  $\text{Prox}'_\theta(Q; D)$  can be written as follows:

$$\text{Prox}'_\theta(Q; D) = U \begin{pmatrix} \Xi_1(Q, \tilde{D}) + \Xi_{aa}(\tilde{D}) & \Xi_2(Q, \tilde{D}) & \Xi_4(Q, \tilde{D}) \\ \Xi_3(Q, \tilde{D}) & \tilde{D}_{\beta\beta} & \tilde{D}_{\beta c} \end{pmatrix} V^T, \quad (3.12)$$

where  $\tilde{D} = U^T D V$  and

$$\Xi_{aa}(\tilde{D}) = \begin{pmatrix} 0_{a_1 \times a_1} & & \\ & 0_{a_2 \times a_2} & \\ & & \Pi_{\mathcal{S}^{|\alpha|}}(S(\tilde{D}_{a_3 a_3})) \end{pmatrix}. \quad (3.13)$$

Theorem 3.1 in [24] shows that the singular value function is second-order directionally differentiable in  $\Re^{m \times n}$ . This means that the spectral norm function is second-order directional differentiable, and  $\theta''(X; \cdot) = \theta^{(1)}(X; \cdot)$ . To analyze the SOSC of problem (1.1), which helps to characterize the ICKKT, we need to compute the “sigma term” of the problem (1.1), namely, the conjugate function of the second-order directional derivative of  $\theta$ . For any  $k = 1, \dots, r$ , define  $\Omega_{a_k} : \Re^{m \times n} \times \Re^{m \times n} \rightarrow \Re^{a_k \times a_k}$  by

$$\begin{aligned} \Omega_{a_k}(P, D) = & (S(\tilde{D}_1))_{a_k}^T (\Sigma(P) - v_k(P)I_m)^\dagger (S(\tilde{D}_1))_{a_k} - (2v_k(P))^{-1} \tilde{D}_{a_k c} \tilde{D}_{a_k c}^T \\ & + (T(\tilde{D}_1))_{a_k}^T (-\Sigma(P) - v_k(P)I_m)^\dagger (T(\tilde{D}_1))_{a_k}, \end{aligned}$$

where  $Z^\dagger$  is the Moore-Penrose pseudo-inverse of  $Z$ ,  $\Sigma(P) = \text{Diag}(\sigma(P))$ ,  $I_m$  is the  $m \times m$  identity matrix and  $\tilde{D} = [\tilde{D}_1 \quad \tilde{D}_2] = [U^T D V_1 \quad U^T D V_2]$ . By [25], we know that  $\forall D \in \Re^{m \times n}$ , the conjugate of  $\theta''(P; D, \cdot)$  is

$$\psi_{(P,D)}^*(W) = (\theta''(P; D, \cdot))^*(W) = \langle \Sigma(W) \rangle_{aa}, 2\Omega_a(P, D) \rangle = \sum_{k=1}^{\tilde{r}_1} \sigma_k(W) \text{trace}(2\Omega_{a_k}(P, D)), \quad (3.14)$$

where  $\sigma_k(W) = (\sigma_i(W))_{i \in a_k}$ . Clearly,  $v_k(P) = \lambda^*$  when  $k = 1, \dots, \tilde{r}$ . Continuing to calculate formula (3.14), we can obtain the explicit expression for  $\psi_{(P,D)}^*(W)$  as follows:

$$\begin{aligned} \psi_{(P,D)}^*(W) = & \sum_{\substack{1 \leq l \leq \tilde{r}_1 \\ \tilde{r}+1 \leq l \leq r+1}} \frac{2\sigma_l(W)}{v_l(P) - \lambda^*} \| (S(\tilde{D}_1))_{a_l a_l} \|^2 - \sum_{1 \leq l \leq \tilde{r}_1} \frac{\sigma_l(W)}{\lambda^*} \| \tilde{D}_{a_l c} \|^2 \\ & + \sum_{\substack{1 \leq l \leq \tilde{r}_1 \\ 1 \leq t \leq r+1}} \frac{2\sigma_l(W)}{-v_t(P) - \lambda^*} \| (T(\tilde{D}_1))_{a_l a_t} \|^2. \end{aligned}$$

We are now prepared to give the main conclusions of this section, that is, some properties of the critical cones and the conjugate function of second-order directional derivative  $\psi^*$ .

Propositions 10 and 12 in [25] characterized the critical cones of  $\theta$  and  $\theta^*$  defined in (3.8) and (3.9), respectively. We summarize the results as follows.

**Proposition 1.** Let  $W \in \partial\theta(P)$  and  $Q = P + W$  admit the SVD as in (3.1). Let the index sets  $a, \beta, \alpha_1, \alpha_2, \alpha_3$ , and  $c$  be defined as (3.2), (3.4), and (3.6). Given any  $D \in \Re^{m \times n}$ , denote  $\tilde{D} = U^T D V$ . Then the following results hold:

(i)  $D \in C_\theta(P, W)$  if and only if there exists some  $\tau \in \Re$  such that

$$\lambda_{|a_1|}(S(\tilde{D}_{a_1 a_1})) \geq \tau \geq \lambda_1(S(\tilde{D}_{a_3 a_3}))$$

and

$$S(\tilde{D}_{aa}) = \begin{pmatrix} S(\tilde{D}_{a_1 a_1}) & & \\ & \tau I_{|a_2|} & \\ & & S(\tilde{D}_{a_3 a_3}) \end{pmatrix},$$

where  $\lambda(Z)$  denote the real eigenvalue vector of symmetric matrix  $Z$  being arranged in nonincreasing order.

(ii)  $D \in C_{\theta^*}(W, P)$  if and only if

$$\text{trace}(\tilde{D}_{aa}) = 0, \quad S(\tilde{D}_{a_1 a_1}) \in S_-^{|\alpha_1|}, \quad \tilde{D}_{a_3 a_3} \in S_+^{|\alpha_3|}$$

and

$$\tilde{D} = \begin{pmatrix} \tilde{D}_{a_1 a_1} & \tilde{D}_{a_1 a_2} & \tilde{D}_{a_1 a_3} & \tilde{D}_{a_1 \beta} & \tilde{D}_{a_1 c} \\ \tilde{D}_{a_2 a_1} & \tilde{D}_{a_2 a_2} & \tilde{D}_{a_2 a_3} & \tilde{D}_{a_2 \beta} & \tilde{D}_{a_2 c} \\ \tilde{D}_{a_3 a_1} & \tilde{D}_{a_3 a_2} & \tilde{D}_{a_3 a_3} & 0 & 0 \\ \tilde{D}_{\beta a_1} & \tilde{D}_{\beta a_2} & 0 & 0 & 0 \end{pmatrix}.$$

**Remark 1.** For the spectral norm function  $\theta$ , the  $\tau$  in part (i) of proposition 1 is actually the largest eigenvalue of  $S(\tilde{D}_{aa})$ , i.e.,  $\tau = \lambda_1(S(\tilde{D}_{aa}))$ . This means that  $\tau$  is relevant to  $D$ . It is easy to observe that the index sets  $\alpha_1$  and  $\alpha_2$  cannot exist simultaneously. When  $\alpha_1 \neq \emptyset$ , there must be  $|\alpha_1| = 1$  and  $\alpha_2 = \emptyset$ , then  $D \in C_\theta(P, W)$  if and only if

$$S(\tilde{D}_{aa}) = \begin{pmatrix} \lambda_1(S(\tilde{D}_{aa})) & \\ & S(\tilde{D}_{\alpha_3\alpha_3}) \end{pmatrix}.$$

When  $\alpha_1 = \emptyset$ ,  $\alpha_2 \neq \emptyset$ ,  $D \in C_\theta(P, W)$  if and only if

$$S(\tilde{D}_{aa}) = \begin{pmatrix} \lambda_1(S(\tilde{D}_{aa}))I_{|\alpha_2|} & \\ & S(\tilde{D}_{\alpha_3\alpha_3}) \end{pmatrix}.$$

In brief, we can conclude that  $D \in C_\theta(P, W)$  if and only if

$$S(\tilde{D}_{aa}) = \begin{pmatrix} \lambda_1(S(\tilde{D}_{aa}))I_{|\alpha_1|+|\alpha_2|} & \\ & S(\tilde{D}_{\alpha_3\alpha_3}) \end{pmatrix}. \quad (3.15)$$

**Proposition 2.** Let  $W \in \partial\theta(P)$  and  $Q = P + W$  admit the SVD as in (3.1). Let the index sets  $\alpha$ ,  $\beta$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ , and  $c$  be defined as (3.2), (3.4), and (3.6). Given any  $H \in \mathfrak{R}^{m \times n}$ , denote  $\tilde{H} = U^T H V$ . Then the following results hold:

- (i) If  $H \in C_{\theta^*}(W, P)$  and  $\phi_{(W,H)}^*(P) = 0$ , then  $H \in (C_\theta(P, W))^{\circ}$ .
- (ii)  $H \in C_\theta(P, W)$  and  $\psi_{(P,H)}^*(W) = 0$  if and only if  $H \in (C_{\theta^*}(W, P))^{\circ}$ .

**Proof.** We first prove part (i). By [25, Proposition 16],  $\forall H \in \mathfrak{R}^{m \times n}$ ,  $\psi_{(P,H)}^*(W) = 0$  if and only if  $\phi_{(W,H)}^*(P) = 0$ , which also is equivalent to

$$\begin{pmatrix} \tilde{H}_{\alpha_1\alpha_1} & \tilde{H}_{\alpha_1\alpha_2} \\ \tilde{H}_{\alpha_2\alpha_1} & \tilde{H}_{\alpha_2\alpha_2} \end{pmatrix} \in S^{|\alpha_1|+|\alpha_2|}, \quad \tilde{H}_{\alpha_1\alpha_3} = (\tilde{H}_{\alpha_3\alpha_1})^T, \quad \tilde{H}_{\alpha_2\alpha_3} = (\tilde{H}_{\alpha_3\alpha_2})^T, \\ \tilde{H}_{\alpha_1\beta} = (\tilde{H}_{\beta\alpha_1})^T = 0, \quad \tilde{H}_{\alpha_2\beta} = (\tilde{H}_{\beta\alpha_2})^T = 0, \quad \tilde{H}_{\alpha_1c} = 0, \quad \tilde{H}_{\alpha_2c} = 0.$$

Then, it follows from part (ii) of Proposition 1 that  $H \in C_{\theta^*}(W, P)$  and  $\phi_{(W,H)}^*(P) = 0$  are equivalent to

$$\tilde{H}_{aa} \in S^{|\alpha|}, \quad \text{trace}(\tilde{H}_{aa}) = 0, \quad H_{\alpha_1\alpha_1} \in S_-^{|\alpha_1|}, \quad H_{\alpha_3\alpha_3} \in S_+^{|\alpha_3|} \quad \text{and} \quad \tilde{H} = \begin{pmatrix} \tilde{H}_{aa} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.16)$$

By part (i) of Proposition 1, for any  $D \in C_\theta(P, W)$ , there exists  $\tau' \in \mathfrak{R}$  such that

$$\lambda_{|\alpha_1|}(S(\tilde{D}_{\alpha_1\alpha_1})) \geq \tau' \geq \lambda_1(S(\tilde{D}_{\alpha_3\alpha_3}))$$

and

$$S(\tilde{D}_{aa}) = \begin{pmatrix} S(\tilde{D}_{\alpha_1\alpha_1}) & & \\ & \tau' I_{|\alpha_2|} & \\ & & S(\tilde{D}_{\alpha_3\alpha_3}) \end{pmatrix}.$$

Hence,  $\forall D \in C_\theta(P, W)$  and  $H$  satisfies (3.16), based on the Fan's inequality (3.3), we have that

$$\begin{aligned} \langle H, D \rangle &= \langle \tilde{H}, \tilde{D} \rangle = \langle \tilde{H}_{aa}, \tilde{D}_{aa} \rangle = \langle S(\tilde{H}_{aa}), \tilde{D}_{aa} \rangle = \langle \tilde{H}_{aa}, S(\tilde{D}_{aa}) \rangle \\ &= \langle \tilde{H}_{\alpha_1\alpha_1}, S(\tilde{D}_{\alpha_1\alpha_1}) \rangle + \tau' \text{trace}(\tilde{H}_{\alpha_2\alpha_2}) + \langle \tilde{H}_{\alpha_3\alpha_3}, S(\tilde{D}_{\alpha_3\alpha_3}) \rangle \\ &\leq (\lambda(\tilde{H}_{\alpha_1\alpha_1}))^T \lambda(S(\tilde{D}_{\alpha_1\alpha_1})) + \tau' \text{trace}(\tilde{H}_{\alpha_2\alpha_2}) + (\lambda(\tilde{H}_{\alpha_3\alpha_3}))^T \lambda(S(\tilde{D}_{\alpha_3\alpha_3})) \\ &= (-\lambda(\tilde{H}_{\alpha_1\alpha_1}))^T (-\lambda(S(\tilde{D}_{\alpha_1\alpha_1}))) + \tau' \text{trace}(\tilde{H}_{\alpha_2\alpha_2}) + (\lambda(\tilde{H}_{\alpha_3\alpha_3}))^T \lambda(S(\tilde{D}_{\alpha_3\alpha_3})) \\ &\leq \tau' \text{trace}(\tilde{H}_{\alpha_1\alpha_1}) + \tau' \text{trace}(\tilde{H}_{\alpha_2\alpha_2}) + \tau' \text{trace}(\tilde{H}_{\alpha_3\alpha_3}) \\ &= \tau' \text{trace}(\tilde{H}_{aa}) \\ &= 0, \end{aligned}$$

which means that  $H \in (C_\theta(P, W))^\circ$ . Conversely, when  $\tau'$  in  $C_\theta(P, W)$  equals to 0, we cannot obtain  $\text{trace}(\widetilde{H}_{aa}) = 0$ , i.e., the reverse conclusion is not true. Then the proof of part (i) is completed.

To prove (ii), it is easy to see that  $\psi_{(P,H)}^*(W) = 0$  and  $H \in C_\theta(P, W)$  hold if and only if there exists  $\tau'' \in \mathfrak{A}$  such that

$$\lambda_{|\alpha_1|}(S(\widetilde{H}_{\alpha_1\alpha_1})) \geq \tau'' \geq \lambda_1(S(\widetilde{H}_{\alpha_3\alpha_3}))$$

and

$$\widetilde{H} = \begin{pmatrix} S(\widetilde{H}_{\alpha_1\alpha_1}) & 0 & 0 & 0 & 0 \\ 0 & \tau'' I_{|\alpha_2|} & 0 & 0 & 0 \\ 0 & 0 & \widetilde{H}_{\alpha_3\alpha_3} & \widetilde{H}_{\alpha_3\beta} & \widetilde{H}_{\alpha_3c} \\ 0 & 0 & \widetilde{H}_{\beta\alpha_3} & \widetilde{H}_{\beta\beta} & \widetilde{H}_{\beta_3c} \\ 0 & 0 & \widetilde{H}_{\beta_3\alpha_3} & \widetilde{H}_{\beta_3\beta} & \widetilde{H}_{\beta_3c} \end{pmatrix},$$

which can be proved to be equivalent to  $H \in (C_{\theta^*}(W, P))^\circ$  by using the same method as in part (i).  $\square$

The following propositions provide some properties of directional derivatives of the proximal mapping of  $\theta$ .

**Proposition 3.** Let  $W \in \partial\theta(P)$  and  $Q = P + W$  admit the SVD as in (3.1). Let the index sets  $\alpha, \beta, \alpha_1, \alpha_2, \alpha_3$ , and  $c$  be defined as (3.2), (3.4), and (3.6). Given any  $D \in \mathfrak{A}^{m \times n}$ , denote  $\widetilde{D} = U^T D V$ . The function  $\psi_{(P,D)}^*(W)$  is defined in (3.14). For any  $D \in \mathfrak{A}^{m \times n}$ , if  $H = \text{Prox}'_\theta(Q; H + D)$  holds, then  $H \in C_\theta(P, W)$  and  $\langle H, D \rangle = -\psi_{(P,D)}^*(W)$ .

**Proof.** Suppose  $H = \text{Prox}'_\theta(Q; H + D)$  holds, together with the expression for  $\text{Prox}'_\theta(Q; \cdot)$  in (3.12), we have that

$$\begin{aligned} \widetilde{H} &= U^T \text{Prox}'_\theta(Q; H + D) V \\ &= \begin{pmatrix} \Xi_1(Q, \widetilde{H} + \widetilde{D}) + \Xi_{aa}(\widetilde{H} + \widetilde{D}) & \Xi_2(Q, \widetilde{H} + \widetilde{D}) & \Xi_4(Q, \widetilde{H} + \widetilde{D}) \\ \Xi_3(Q, \widetilde{H} + \widetilde{D}) & \widetilde{H}_{\beta\beta} + \widetilde{D}_{\beta\beta} & \widetilde{H}_{\beta c} + \widetilde{D}_{\beta c} \end{pmatrix}, \end{aligned} \quad (3.17)$$

where  $\Xi_{aa}$  is defined in (3.13). We first verify  $H \in C_\theta(P, W)$ . The equality

$$\begin{aligned} \widetilde{H}_{aa} &= \Xi_1(Q, \widetilde{H} + \widetilde{D}) + \Xi_{aa}(\widetilde{H} + \widetilde{D}) \\ &= \Theta_{aa}^2(Q) \circ (T(\widetilde{H}_1)_{aa} + T(\widetilde{D}_1)_{aa}) + \begin{pmatrix} 0_{\alpha_1 \times \alpha_1} & & \\ & 0_{\alpha_2 \times \alpha_2} & \\ & & \Pi_{S^{|\alpha_3|}}(S(\widetilde{H}_{\alpha_3\alpha_3}) + S(\widetilde{D}_{\alpha_3\alpha_3})) \end{pmatrix} \end{aligned}$$

implies that

$$S(\widetilde{H}) = \begin{pmatrix} 0_{\alpha_1 \times \alpha_1} & & \\ & 0_{\alpha_2 \times \alpha_2} & \\ & & \Pi_{S^{|\alpha_3|}}(S(\widetilde{H}_{\alpha_3\alpha_3}) + S(\widetilde{D}_{\alpha_3\alpha_3})) \end{pmatrix}.$$

From part (i) of proposition 1, we obtain that  $H \in C_\theta(P, W)$ .



For convenience, we denote  $\alpha' := \alpha_1 \cup \alpha_2$ . By directly calculating (3.17), we deduce that  $H = \text{Prox}'_{\theta}(Q; H + D)$  if and only if

$$\left\{ \begin{array}{ll} S(\tilde{H}_{\alpha'\alpha'}) = 0, & S(\tilde{H}_1)_{\alpha'\alpha_3} = 0, & S(\tilde{H}_1)_{\alpha_3\alpha'} = 0, \\ S_{-}^{|\alpha_3|} \ni S(\tilde{H}_{\alpha_3\alpha_3}) \perp S(\tilde{D}_{\alpha_3\alpha_3}) = \tilde{D}_{\alpha_3\alpha_3} \in S_{+}^{|\alpha_3|}, \\ T(\tilde{D}_{\alpha'\alpha'})_{ij} = \left( \frac{\sigma_i(Q) + \sigma_j(Q)}{2\lambda^*} - 1 \right) (\tilde{H}_{\alpha'\alpha'})_{ij}, & i \in [|\alpha'|], j \in [|\alpha'|], \\ (T(\tilde{D}_1)_{\alpha'\alpha_3})_{ij} = \frac{\sigma_i(Q) - \lambda^*}{2\lambda^*} (\tilde{H}_{\alpha'\alpha_3})_{ij}, & i \in [|\alpha'|], j \in [|\alpha_3|], \\ (T(\tilde{D}_1)_{\alpha_3\alpha'})_{ji} = \frac{\sigma_i(Q) - \lambda^*}{2\lambda^*} (\tilde{H}_{\alpha_3\alpha'})_{ji}, & i \in [|\alpha'|], j \in [|\alpha_3|], \\ (\tilde{D}_{\alpha\beta})_{ij} = \frac{\lambda^*(\sigma_i(Q) - \lambda^*)}{\lambda^{*2} - (\sigma_{j+|\alpha|}(Q))^2} (\tilde{H}_{\alpha\beta})_{ij} + \frac{\sigma_{j+|\alpha|}(Q)(\sigma_i(Q) - \lambda^*)}{\lambda^{*2} - (\sigma_{j+|\alpha|}(Q))^2} (\tilde{H}_{\beta\alpha})_{ji}, & i \in [|\alpha|], j \in [|\beta|], \\ (\tilde{D}_{\beta\alpha})_{ji} = \frac{\lambda^*(\sigma_i(Q) - \lambda^*)}{\lambda^{*2} - (\sigma_{j+|\alpha|}(Q))^2} (\tilde{H}_{\beta\alpha})_{ji} + \frac{\sigma_{j+|\alpha|}(Q)(\sigma_i(Q) - \lambda^*)}{\lambda^{*2} - (\sigma_{j+|\alpha|}(Q))^2} (\tilde{H}_{\alpha\beta})_{ij}, & i \in [|\alpha|], j \in [|\beta|], \\ (\tilde{D}_{ac})_{ij} = \left( \frac{\sigma_i(Q)}{\lambda^*} - 1 \right) (\tilde{H}_{ac})_{ij}, & i \in [|\alpha|], j \in [n - m], \\ \tilde{D}_{\beta\beta} = 0, \\ \tilde{D}_{\beta c} = 0. \end{array} \right. \quad (3.18)$$

Noting that  $v_k(Q) - \lambda^* = \sigma_k(W)$  for  $1 \leq k \leq r_1$ ,  $v_t(P) = \lambda^*$  for  $1 \leq t \leq \tilde{r}$  and  $v_t(P) = v_t(Q)$  for  $\tilde{r} + 1 \leq t \leq r + 1$ . Together with (3.18) and the fact that for any  $D \in \mathfrak{R}^{p \times p}$ ,

$$\begin{aligned} \langle D, D^T \rangle &= \|S(D)\|^2 - \|T(D)\|^2, \\ \langle D, D \rangle &= \|D\|^2 = \|S(D)\|^2 + \|T(D)\|^2, \end{aligned}$$

we obtain that

$$\begin{aligned} \langle H, D \rangle &= \langle \tilde{H}, \tilde{D} \rangle = \langle \tilde{H}_{aa}, \tilde{D}_{aa} \rangle + \langle \tilde{H}_{a\beta}, \tilde{D}_{a\beta} \rangle + \langle \tilde{H}_{\beta a}, \tilde{D}_{\beta a} \rangle + \langle \tilde{H}_{ac}, \tilde{D}_{ac} \rangle \\ &= \langle \tilde{H}_{aa}, T(\tilde{D}_{aa}) \rangle + \langle \tilde{H}_{a\beta}, \tilde{D}_{a\beta} \rangle + \langle \tilde{H}_{\beta a}, \tilde{D}_{\beta a} \rangle + \langle \tilde{H}_{ac}, \tilde{D}_{ac} \rangle \\ &= \sum_{\substack{1 \leq l \leq r_1 \\ 1 \leq t \leq \tilde{r}}} \frac{v_l(Q) - \lambda^*}{\lambda^*} \| (T(\tilde{H}_1))_{a_l a_t} \|^2 + \sum_{1 \leq l \leq r_1} \frac{v_l(Q) - \lambda^*}{\lambda^*} \| \tilde{H}_{a_l c} \|^2 \\ &\quad + \sum_{\substack{1 \leq l \leq r_1 \\ \tilde{r}+1 \leq t \leq r+1}} \left( - \frac{2(v_l(Q) - \lambda^*)}{v_t(Q) - \lambda^*} \| S(\tilde{H}_1)_{a_l a_t} \|^2 + \frac{2(v_l(Q) - \lambda^*)}{v_t(Q) + \lambda^*} \| T(\tilde{H}_1)_{a_l a_t} \|^2 \right) \\ &= - \sum_{\substack{1 \leq l \leq r_1 \\ \tilde{r}+1 \leq t \leq r+1}} \frac{2\sigma_l(W)}{v_t(P) - \lambda^*} \| S(\tilde{H}_1)_{a_l a_t} \|^2 + \sum_{1 \leq l \leq r_1} \frac{\sigma_l(W)}{\lambda^*} \| \tilde{H}_{a_l c} \|^2 \\ &\quad + \sum_{\substack{1 \leq l \leq r_1 \\ 1 \leq t \leq r+1}} \frac{2\sigma_l(W)}{v_t(P) + \lambda^*} \| T(\tilde{H}_1)_{a_l a_t} \|^2, \end{aligned}$$

so we have

$$\langle H, D \rangle = \langle \tilde{H}, \tilde{D} \rangle = -\psi_{(P,H)}^*(W).$$

This completes the proof.  $\square$

**Proposition 4.** Let  $W \in \partial\theta(P)$  and  $Q = P + W$  admit the SVD as in (3.1). Let the index sets  $\alpha, \beta, \alpha_1, \alpha_2, \alpha_3$ , and  $c$  be defined as (3.2), (3.4), and (3.6). For any  $D \in \mathfrak{R}^{m \times n}$ , then

$$\text{Prox}'_{\theta}(Q; D) = 0 \Leftrightarrow D \in (C_{\theta}^0(P, W))^{\circ},$$

where  $C_\theta^0(P, W)$  represents the critical cone  $C_\theta(P, W)$  in the case of  $\tau = 0$ , i.e.,

$$C_\theta^0(P, W) = \left\{ H \in \mathfrak{R}^{m \times n} \mid S(\tilde{H}_{aa}) = \begin{pmatrix} 0_{|a_1|+|a_2|} \\ S(\tilde{H}_{a_3a_3}) \end{pmatrix}, S(\tilde{H}_{a_3a_3}) \in S_-^{|a_3|} \right\}.$$

**Proof.** It is easy to verify either  $\text{Prox}'_\theta(Q; D) = 0$  or  $D \in (C_\theta^0(P, W))^\circ$  is equivalent to

$$\tilde{D} = \begin{pmatrix} \tilde{D}_{aa} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \tilde{D}_{aa} \in S^{|a|}, \quad D_{a_3a_3} \in S_+^{|a_3|}. \quad \square$$

## 4 The robust ICKKT

In this section, owing to the special linear structure of the problem (1.1), we first show the primal (dual) SRCQ is equivalent to the dual (primal) SOSC. In fact, when the spectral norm in problem (1.1) is substituted by the nuclear norm, the same results are already given in [16]. Therefore, we only briefly state here, and the main tools we used are Propositions 1 and 2. The Lagrangian dual of problem (1.1) is as follows:

$$\begin{aligned} \max_{y, w, S} \quad & -\langle b, y \rangle - \delta_{\mathcal{P}^\circ}(y) - h^*(w) \\ \text{s.t.} \quad & \mathcal{B}^*y + Q^*w + S + C = 0, \quad \|S\|_* \leq 1, \end{aligned} \quad (4.1)$$

where  $(y, w, S) \in \mathfrak{R}^l \times \mathfrak{R}^d \times \mathfrak{R}^{m \times n}$ ,  $\mathcal{B}^*$  and  $Q^*$  represent the adjoint of  $\mathcal{B}$  and  $Q$ , respectively. The first order optimality conditions, i.e., the KKT conditions for the primal problem (1.1) and the dual problem (4.1) are given by

$$\begin{cases} Q^*\nabla h(QX) + C + S + \mathcal{B}^*y = 0, \\ \mathcal{P}^\circ \ni y \perp \mathcal{B}X - b \in \mathcal{P}, \\ S \in \partial\theta(X) \end{cases} \quad (X, y, S) \in \mathfrak{R}^{m \times n} \times \mathfrak{R}^l \times \mathfrak{R}^{m \times n} \quad (4.2)$$

and

$$\begin{cases} \mathcal{B}^*y + Q^*w + S + C = 0, \\ \mathcal{P}^\circ \ni y \perp \mathcal{B}X - b \in \mathcal{P}, \\ QX \in \partial h^*(w), \quad X \in \partial\theta^*(S), \end{cases} \quad (y, w, S, X) \in \mathfrak{R}^l \times \mathfrak{R}^d \times \mathfrak{R}^{m \times n} \times \mathfrak{R}^{m \times n}, \quad (4.3)$$

respectively. We denote by  $\mathcal{M}_P(X)$  the set of Lagrangian multipliers respect to  $X$  for problem (1.1), namely,  $\mathcal{M}_P(X) = \{(y, S) \in \mathfrak{R}^l \times \mathfrak{R}^{m \times n} \mid (X, y, S) \text{ satisfies (4.2)}\}$ . We write  $\mathcal{M}_D(y, w, S)$  the set of Lagrangian multipliers respect to  $(y, w, S)$  for problem (4.1), i.e.,  $\mathcal{M}_D(y, w, S) = \{X \in \mathfrak{R}^{m \times n} \mid (y, w, S, X) \text{ satisfies (4.3)}\}$ .

Let  $(y, z) \in \mathfrak{R}^l \times \mathfrak{R}^l$  satisfies  $\mathcal{P}^\circ \ni y \perp z \in \mathcal{P}$ . Define the critical cone of  $\mathcal{P}$  at  $z$  for  $y$  as  $C_{\mathcal{P}}(z, y) = \mathcal{T}_{\mathcal{P}}(z) \cap y^\perp$  and the critical cone of  $\mathcal{P}^\circ$  at  $y$  for  $z$  as  $C_{\mathcal{P}^\circ}(y, z) = \mathcal{T}_{\mathcal{P}^\circ}(y) \cap z^\perp$ , respectively. Obviously,

$$(C_{\mathcal{P}}(z, y))^\circ = C_{\mathcal{P}^\circ}(y, z). \quad (4.4)$$

Now we are ready to demonstrate the crucial result of this article, namely, the relationship between the primal SRCQ and the dual SOSC. Actually, this result was given in [16, Theorem 5.2], but they omitted the proof details. For the sake of completeness, we supplement the proof below.

**Theorem 1.** Let  $(\bar{y}, \bar{w}, \bar{S}) \in \mathfrak{R}^l \times \mathfrak{R}^d \times \mathfrak{R}^{m \times n}$  be an optimal solution of problem (4.1) with  $\mathcal{M}_D(\bar{y}, \bar{w}, \bar{S}) \neq \emptyset$ . Let  $\bar{X} \in \mathcal{M}_D(\bar{y}, \bar{w}, \bar{S})$ . Then the following results are equivalent:

(i) The dual SOSC holds at  $(\bar{y}, \bar{w}, \bar{S})$  with respect to  $\bar{X}$ , i.e.,

$$\langle H_w, (\nabla h^*)(\bar{w}; H_w) \rangle - \phi_{(\bar{S}, H_S)}^*(\bar{X}) > 0, \quad \forall (H_y, H_w, H_S) \in C(\bar{y}, \bar{w}, \bar{S}) \setminus \{0\}, \quad (4.5)$$

where  $C(\bar{y}, \bar{w}, \bar{S})$  is the critical cone defined as follows:

$$C(\bar{y}, \bar{w}, \bar{S}) := \left\{ (D_y, D_w, D_S) \in \mathfrak{R}^l \times \mathfrak{R}^d \times \mathfrak{R}^{m \times n} \mid \begin{array}{l} \mathcal{B}^* D_y + Q^* D_w + D_S = 0, \\ D_y \in C_{\mathcal{P}^*}(\bar{y}, \mathcal{B}\bar{X} - b), \quad D_S \in C_{\theta^*}(\bar{S}, \bar{X}) \end{array} \right\}.$$

(ii) The primal SRCQ holds at  $\bar{X}$  with respect to  $(\bar{y}, \bar{S})$ , i.e.,

$$\begin{pmatrix} \mathcal{B} \\ \mathcal{I}_{\mathfrak{R}^{m \times n}} \end{pmatrix} \mathfrak{R}^{m \times n} + \begin{pmatrix} C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}) \\ C_{\theta}(\bar{X}, \bar{S}) \end{pmatrix} = \begin{pmatrix} \mathfrak{R}^l \\ \mathfrak{R}^{m \times n} \end{pmatrix}. \quad (4.6)$$

**Proof.** “(i)  $\Rightarrow$  (ii)”. Denote

$$\Psi := \begin{pmatrix} \mathcal{B} \\ \mathcal{I}_{\mathfrak{R}^{m \times n}} \end{pmatrix} \mathfrak{R}^{m \times n} + \begin{pmatrix} C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}) \\ C_{\theta}(\bar{X}, \bar{S}) \end{pmatrix}.$$

Suppose on the contrary that the SRCQ (4.6) does not hold at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ , i.e.,  $\Psi \neq \begin{pmatrix} \mathfrak{R}^l \\ \mathfrak{R}^{m \times n} \end{pmatrix}$ . Then  $\text{cl}\Psi \neq \begin{pmatrix} \mathfrak{R}^l \\ \mathfrak{R}^{m \times n} \end{pmatrix}$ .

Therefore, there exists  $\bar{H} \in \begin{pmatrix} \mathfrak{R}^l \\ \mathfrak{R}^{m \times n} \end{pmatrix}$  s.t.  $\bar{H} \notin \text{cl}\Psi$ . Since  $\text{cl}\Psi$  is a closed convex cone,  $\bar{H} - \Pi_{\text{cl}\Psi}(\bar{H}) = \Pi_{(\text{cl}\Psi)^\circ}(\bar{H}) \neq 0$ . Denote  $H := \Pi_{(\text{cl}\Psi)^\circ}(\bar{H}) = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} \in \begin{pmatrix} \mathfrak{R}^l \\ \mathfrak{R}^{m \times n} \end{pmatrix}$ . Obviously,  $\langle H, D \rangle \leq 0$  for any  $D \in \text{cl}\Psi$ , which implies that

$$\mathcal{B}^* H_1 = 0, \quad H_1 \in C_{\mathcal{P}^*}(\bar{y}, \mathcal{B}\bar{X} - b), \quad H_2 = 0.$$

Thus, we have that  $(H_1, 0, 0) \in C(\bar{y}, \bar{w}, \bar{S}) \setminus \{0\}$  and  $\langle 0, (\nabla h^*)(\bar{w}; 0) \rangle - \varphi_{(\bar{S}, 0)}^*(\bar{X}) = 0$ , which contradicts the SOSC (4.5) at  $(\bar{y}, \bar{w}, \bar{S})$  with respect to  $\bar{X}$ .

“(ii)  $\Rightarrow$  (i)” Suppose on the contrary that the SOSC (4.5) does not hold at  $(\bar{y}, \bar{w}, \bar{S})$  for  $\bar{X}$ . Then,  $\exists \bar{H} = (\bar{H}_y, \bar{H}_w, \bar{H}_S) \in C(\bar{y}, \bar{w}, \bar{S}) \setminus \{0\}$  such that

$$\langle \bar{H}_w, (\nabla h^*)(\bar{w}; \bar{H}_w) \rangle - \varphi_{(\bar{S}, \bar{H}_S)}^*(\bar{X}) = 0.$$

From [25, Proposition 16], we know that  $\varphi_{(\bar{S}, H)}^*(\bar{X}) \leq 0$  for any  $H \in \mathfrak{R}^{m \times n}$ . Moreover, because  $h$  is essentially strictly convex, by [20, Theorem 26.3],  $h^*$  is essentially smooth. Then,  $\nabla h^*$  is locally Lipschitz continuous and directionally differentiable on  $\text{int}(\text{dom } h^*)$ . Since  $(\bar{y}, \bar{w}, \bar{S})$  is the optimal solution, it is not hard to verify by the mean value theorem that  $\langle H_w, (\nabla h^*)(\bar{w}; H_w) \rangle > 0$  for any  $H_w \neq 0$  such that  $(H_y, H_w, H_S) \in C(\bar{y}, \bar{w}, \bar{S}) \setminus \{0\}$ . Therefore,

$$\bar{H}_w = 0, \quad \varphi_{(\bar{S}, \bar{H}_S)}^*(\bar{X}) = 0.$$

From  $\bar{H} \in C(\bar{y}, \bar{w}, \bar{S}) \setminus \{0\}$ , we know that

$$\mathcal{B}^* \bar{H}_y + \bar{H}_S = 0, \quad \bar{H}_y \in C_{\mathcal{P}^*}(\bar{y}, \mathcal{B}\bar{X} - b), \quad \bar{H}_S \in C_{\theta^*}(\bar{S}, \bar{X}).$$

Together with (4.4) and Proposition 2, we have

$$\bar{H}_y \in (C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}))^\circ, \quad \bar{H}_S \in (C_{\theta}(\bar{X}, \bar{S}))^\circ.$$

Because the SRCQ (4.6) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ , there exist  $\tilde{W} \in \mathfrak{R}^{m \times n}$  and  $\tilde{H} = \begin{pmatrix} \tilde{H}_1 \\ \tilde{H}_2 \end{pmatrix} \in \begin{pmatrix} C_{\mathcal{P}}(\mathcal{B}\bar{X} - b) \\ C_{\theta}(\bar{X}, \bar{S}) \end{pmatrix}$  such that

$$\begin{pmatrix} \mathcal{B} \\ \mathcal{I}_{\mathfrak{R}^{m \times n}} \end{pmatrix} \tilde{W} + \begin{pmatrix} \tilde{H}_1 \\ \tilde{H}_2 \end{pmatrix} = \begin{pmatrix} \bar{H}_y \\ \bar{H}_S \end{pmatrix},$$

i.e.,

$$\mathcal{B}\tilde{W} + \tilde{H}_1 = \bar{H}_y, \quad \tilde{W} + \tilde{H}_2 = \bar{H}_S.$$

Thus,

$$\begin{aligned}\langle \bar{H}, \bar{H} \rangle &= \langle \bar{H}_y, \mathcal{B}\bar{W} + \bar{H}_1 \rangle + \langle \bar{H}_s, \bar{W} + \bar{H}_2 \rangle \\ &= \langle \mathcal{B}^* \bar{H}_y + \bar{H}_s, \bar{W} \rangle + \langle \bar{H}_y, \bar{H}_1 \rangle + \langle \bar{H}_s, \bar{H}_2 \rangle \\ &= 0 + \langle \bar{H}_y, \bar{H}_1 \rangle + \langle \bar{H}_s, \bar{H}_2 \rangle \\ &\leq 0,\end{aligned}$$

which means that  $\bar{H} = 0$ . This contradicts the hypothesis that  $\bar{H} \neq 0$ .  $\square$

Switching the “primal” and “dual” in the aforementioned theorem, we can also obtain that the primal SOSC is equivalent to the dual SRCQ from Proposition 2. One may refer to [16, Theorem 5.1] for more details.

**Theorem 2.** Let  $\bar{X} \in \mathfrak{R}^{m \times n}$  be an optimal solution of problem (1.1) with  $\mathcal{M}_P(\bar{X}) \neq \emptyset$ . Let  $(\bar{y}, \bar{S}) \in \mathcal{M}_P(\bar{X})$ . Then the following results are equivalent:

(i) The dual SRCQ holds at  $\bar{y}$  for  $\bar{X}$ , i.e.,

$$Q^* \mathfrak{R}^d + \mathcal{B}^* C_{\mathcal{P}}(\bar{y}, \mathcal{B}\bar{X} - b) + C_{\theta^*}(\bar{S}, \bar{X}) = \mathfrak{R}^{m \times n}. \quad (4.7)$$

(ii) The primal SOSC holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ , i.e.,

$$\langle QD, \nabla^2 h(Q\bar{X})QD \rangle - \psi_{(\bar{X}, D)}^*(\bar{S}) > 0, \quad \forall D \in C(\bar{X}) \setminus \{0\}, \quad (4.8)$$

where  $C(\bar{X}) := \{D \in \mathfrak{R}^{m \times n} | \mathcal{B}D \in C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}), D \in C_{\theta}(\bar{X}, \bar{S})\}$ .

In the following, we first give the definition of the KKT mapping for problem (1.1). Then, we demonstrate that both the SOSC and the SRCQ are sufficient and necessary for the ICKKTM. Finally, based on the aforementioned result, we give a series of characterizations of the robust ICKKTM for problem (1.1).

The KKT conditions (4.2) for problem (1.1) can be equivalently expressed as the nonsmooth equation:

$$F(X, y, S) = 0,$$

where  $F : \mathfrak{R}^{m \times n} \times \mathfrak{R}^l \times \mathfrak{R}^{m \times n} \rightarrow \mathfrak{R}^{m \times n} \times \mathfrak{R}^l \times \mathfrak{R}^{m \times n}$  is defined by

$$F(X, y, S) := \begin{bmatrix} C + \mathcal{B}^* y + Q^* \nabla h(QX) + S \\ \mathcal{B}X - b - \Pi_Q(\mathcal{B}X - b + y) \\ X - \text{Prox}_{\theta}(X + S) \end{bmatrix}, \quad (X, y, S) \in \mathfrak{R}^{m \times n} \times \mathfrak{R}^l \times \mathfrak{R}^{m \times n}.$$

Denote the KKT mapping for problem (1.1) by

$$S_{\text{KKT}}(\delta) := \{(X, y, S) \in \mathfrak{R}^{m \times n} \times \mathfrak{R}^l \times \mathfrak{R}^{m \times n} | F(X, y, S) = \delta\}.$$

Thus, we know from [13] that the isolated calmness of  $S_{\text{KKT}}$  at the origin for  $(\bar{X}, \bar{y}, \bar{S})$  is equivalent to  $F'((\bar{X}, \bar{y}, \bar{S}); (\Delta Z, \Delta y, \Delta S)) = 0 \Rightarrow (\Delta Z, \Delta y, \Delta S) = 0$ .

Define the strong SRCQ (SSRCQ) at  $\bar{X}$  for  $(\bar{y}, \bar{S})$  as follows:

$$\begin{bmatrix} \mathcal{B} \\ I_{\mathfrak{R}^{m \times n}} \end{bmatrix} \mathfrak{R}^{m \times n} + \begin{bmatrix} C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}) \\ C_{\theta}^0(\bar{X}, \bar{S}) \end{bmatrix} = \begin{bmatrix} \mathfrak{R}^l \\ \mathfrak{R}^{m \times n} \end{bmatrix}. \quad (4.9)$$

Similarly, define the weak SOSC (WSOSC) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ , i.e.,

$$\langle QD, \nabla^2 h(Q\bar{X})QD \rangle - \psi_{(\bar{X}, D)}^*(\bar{S}) > 0, \quad \forall D \in C^0(\bar{X}) \setminus \{0\}, \quad (4.10)$$

where  $C^0(\bar{X}) := \{D \in \mathfrak{R}^{m \times n} | \mathcal{B}D \in C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}), D \in C_{\theta}^0(\bar{X}, \bar{S})\}$ .

**Theorem 3.** Let  $\bar{X} \in \mathfrak{R}^{m \times n}$  be an optimal solution of problem (1.1) with  $(\bar{y}, \bar{S}) \in \mathcal{M}_P(\bar{X}) \neq \emptyset$ . Then the following results are equivalent:

- (i) The SSRCQ (4.9) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$  and the WSOSC (4.10) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ .
- (ii) The KKT mapping  $S_{\text{KKT}}$  is isolated calm at the origin for  $(\bar{X}, \bar{y}, \bar{S})$ .

**Proof.** “(i)  $\Rightarrow$  (ii)”. Let  $(\Delta Z, \Delta y, \Delta S) \in \mathfrak{R}^{m \times n} \times \mathfrak{R}^l \times \mathfrak{R}^{m \times n}$  satisfies

$$F'((\bar{X}, \bar{y}, \bar{S}); (\Delta Z, \Delta y, \Delta S)) = \begin{bmatrix} Q^* \nabla^2 h(Q\bar{X}) Q \Delta Z + \mathcal{B}^* \Delta y + \Delta S \\ \mathcal{B} \Delta Z - \Pi'_Q(\mathcal{B}\bar{X} - b + \bar{y}; \mathcal{B} \Delta Z + \Delta y) \\ \Delta Z - \text{Prox}'_{\theta}(\bar{X} + \bar{S}; \Delta Z + \Delta S) \end{bmatrix} = 0. \quad (4.11)$$

From the second equation of (4.11) and [9, Lemma 4.2], we know that

$$\mathcal{B} \Delta Z \in C_Q(\mathcal{B}\bar{X} - b, \bar{y}), \quad \langle \Delta y, \mathcal{B} \Delta Z \rangle = 0. \quad (4.12)$$

Proposition 3 and the third equation of (4.11) means that

$$\Delta Z \in C_{\theta}^0(\bar{X}, \bar{S}), \quad \langle \Delta Z, \Delta S \rangle = -\psi_{(\bar{X}, \Delta Z)}^*(\bar{S}). \quad (4.13)$$

Thus, we can conclude from (4.12) and (4.13) that

$$\Delta Z \in C^0(\bar{X}).$$

By taking the inner product of  $\Delta Z$  and the first equation of (4.11), together with (4.12) and (4.13), we can obtain

$$\begin{aligned} & \langle \Delta Z, Q^* \nabla^2 h(Q\bar{X}) Q \Delta Z + \mathcal{B}^* \Delta y + \Delta S \rangle \\ &= \langle \Delta Z, Q^* \nabla^2 h(Q\bar{X}) Q \Delta Z \rangle + \langle \Delta Z, \mathcal{B}^* \Delta y \rangle + \langle \Delta Z, \Delta S \rangle \\ &= \langle Q \Delta Z, \nabla^2 h(Q\bar{X}) Q \Delta Z \rangle + \langle \mathcal{B} \Delta Z, \Delta y \rangle + \langle \Delta Z, \Delta S \rangle \\ &= \langle Q \Delta Z, \nabla^2 h(Q\bar{X}) Q \Delta Z \rangle - \psi_{(\bar{X}, \Delta Z)}^*(\bar{S}) \\ &= 0. \end{aligned}$$

Then, from the WSOSC (4.10), we know that

$$\Delta Z = 0.$$

Therefore, (4.11) reduces to

$$\begin{bmatrix} \mathcal{B}^* \Delta y + \Delta S \\ \Pi'_Q(\mathcal{B}\bar{X} - b + \bar{y}; \Delta y) \\ \text{Prox}'_{\theta}(\bar{X} + \bar{S}; \Delta S) \end{bmatrix} = 0. \quad (4.14)$$

From Lemma 4.3 in [9] and Proposition 4, we obtain that (4.14) is equivalent to the following formula:

$$\begin{pmatrix} \Delta y \\ \Delta S \end{pmatrix} \in \left[ \begin{pmatrix} \mathcal{B} \\ \mathcal{I}_{\mathfrak{R}^{m \times n}} \end{pmatrix} \mathfrak{R}^{m \times n} + \begin{pmatrix} C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}) \\ C_{\theta}^0(\bar{X}, \bar{S}) \end{pmatrix} \right]^{\circ}. \quad (4.15)$$

According to the SSRCQ (4.9),

$$\langle \Delta y, \Delta S \rangle = 0.$$

This, together with  $\Delta Z = 0$ , imply that the KKT mapping is isolated calm at the origin for  $(\bar{X}, \bar{y}, \bar{S})$ .

We prove “(ii)  $\Rightarrow$  (i)” by contradiction. Assume that the SSRCQ (4.9) does not hold at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ . Then  $\exists (\Delta y, \Delta S) \neq 0$  s.t. (4.15) holds, which also means that (4.14) holds. Hence,  $F'((\bar{X}, \bar{y}, \bar{S}); (0, \Delta y, \Delta S)) = 0$ . This contradicts with the ICKKTM at the origin for  $(\bar{X}, \bar{y}, \bar{S})$ .

Because  $\bar{X}$  is an optimal solution of problem (1.1) and the SSRCQ (4.9) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ , the Lagrangian multipliers  $\mathcal{M}_P(\bar{X})$  is a singleton and the second-order necessary condition holds at  $\bar{X}$ , i.e.,

$$\langle QD, \nabla^2 h(Q\bar{X}) QD \rangle - \psi_{(\bar{X}, D)}^*(\bar{S}) \geq 0, \quad \forall D \in C^0(\bar{X}).$$

Suppose that the WSOSC (4.10) does not hold at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ , which means that  $\exists \Delta Z' \in C^0(\bar{X}) \setminus \{0\}$  such that

$$\langle Q \Delta Z', \nabla^2 h(Q\bar{X}) Q \Delta Z' \rangle - \psi_{(\bar{X}, \Delta Z')}^*(\bar{S}) = 0.$$

By [25, Proposition 16],  $\forall H \in \mathfrak{R}^{m \times n}$ ,  $\psi_{(\bar{X}, H)}^*(\bar{S}) \leq 0$ . Besides, since  $h$  is essentially strictly convex,  $\langle QH, \nabla^2 h(Q\bar{X})QH \rangle > 0$  for any  $H \in \mathfrak{R}^{m \times n}$  satisfies  $QH \neq 0$ . Consequently,

$$Q\Delta Z' = 0, \quad \psi_{(\bar{X}, \Delta Z')}^*(\bar{S}) = 0.$$

From  $\Delta Z' \in C^0(\bar{X}) \setminus \{0\}$ , we know that

$$\mathcal{B}\Delta Z' \in C_{\mathcal{P}}(\mathcal{B}\bar{X} - b, \bar{y}), \quad \Delta Z' \in C_{\theta}^0(\bar{X}, \bar{S}). \quad (4.16)$$

Thus, from [9, Lemma 4.2], the first formula in (4.16) is equivalent to

$$\mathcal{B}\Delta Z' = \Pi'_{\mathcal{Q}}(\mathcal{B}\bar{X} - b + \bar{y}; \mathcal{B}\Delta Z'). \quad (4.17)$$

We can conclude from the second formula in (4.16) and (3.12) that

$$\Delta Z' = \text{Prox}'_{\theta}(\bar{X} + \bar{S}; \Delta Z'). \quad (4.18)$$

Therefore, formulas (4.17), (4.18), and  $Q\Delta Z' = 0$  imply that

$$F'((\bar{X}, \bar{y}, \bar{S}); (\Delta Z', 0, 0)) = 0.$$

Notice that  $\Delta Z' \neq 0$ , which contradicts with the isolated calmness of  $S_{\text{KKT}}$  at the origin for  $(\bar{X}, \bar{y}, \bar{S})$ . That is, the WSOSC (4.10) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ . This completes the proof.  $\square$

It is said that the Robinson constraint qualification (RCQ) holds at  $\bar{X} \in \mathfrak{R}^{m \times n}$  for problem (1.1) if

$$\mathcal{B}\mathfrak{R}^{m \times n} + \mathcal{T}_{\mathcal{P}}(\mathcal{B}\bar{X} - b) = \mathfrak{R}^l. \quad (4.19)$$

Finally, we can establish a series of equivalent conditions of the robust ICKKT for problem (1.1).

**Theorem 4.** Let  $\bar{X} \in \mathfrak{R}^{m \times n}$  and  $(\bar{y}, \bar{w}, \bar{S}) \in \mathfrak{R}^l \times \mathfrak{R}^d \times \mathfrak{R}^{m \times n}$  be optimal solutions of problems (1.1) and (4.1), respectively. Assume that  $(\bar{y}, \bar{S}) \in \mathcal{M}_P(\bar{X})$  and  $\bar{X} \in \mathcal{M}_D(\bar{y}, \bar{w}, \bar{S})$ , and the RCQ (4.19) holds at  $\bar{X}$ . Then the following statements are equivalent:

- (a) The KKT mapping  $S_{\text{KKT}}$  is (robustly) isolated calm at the origin for  $(\bar{X}, \bar{y}, \bar{S})$ .
- (b) The primal WSOSC (4.10) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$  and the primal SSRQC (4.9) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ .
- (c) The primal SOS (4.8) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$  and the primal SRCQ (4.6) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ .
- (d) The primal SOS (4.8) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$  and the dual SOS (4.5) holds at  $(\bar{y}, \bar{w}, \bar{S})$  for  $\bar{X}$ .
- (e) The dual SRCQ (4.7) holds at  $(\bar{y}, \bar{w}, \bar{S})$  for  $\bar{X}$  and the dual SOS (4.5) holds at  $(\bar{y}, \bar{w}, \bar{S})$  for  $\bar{X}$ .
- (f) The dual SRCQ (4.7) holds at  $(\bar{y}, \bar{w}, \bar{S})$  for  $\bar{X}$  and the primal SRCQ (4.6) holds at  $\bar{X}$  for  $(\bar{y}, \bar{S})$ .

**Proof.** Since the spectral norm function  $\|\cdot\|_2$  is  $C^2$ -cone reducible, the “robustly” in (a) of the Theorem 4, which means that  $S_{\text{KKT}}$  is also lower semicontinuous at the origin for  $(\bar{X}, \bar{y}, \bar{S})$ , can be obtained automatically from [13, Theorem 17]. Besides, since the spectral norm function is Lipschitz continuous at  $\bar{X}$ , by combining Theorems 1, 2, 3, [13, Theorem 17], and [16, Proposition 3.3], we can obtain that (a)  $\Leftrightarrow$  (b)  $\Leftrightarrow$  (c)  $\Leftrightarrow$  (d)  $\Leftrightarrow$  (e)  $\Leftrightarrow$  (f).  $\square$

**Remark 2.** Proposition 3.3 in [16] is obtained indirectly by transforming problem (1.1) into a matrix conic optimization induced by nuclear norm and then applying [13, Theorem 24]. In the aforementioned theorem, we give a direct proof. The difference is that we need to use the variational properties of the directional derivatives of the proximal mapping of spectral norm, while [13] uses the projection mapping.

## 5 Conclusion

This article is devoted to establishing a series of equivalent conditions of the robust ICKKT for spectral norm regularized convex matrix optimization problems. We extend the results for nuclear norm regularized convex

optimization problems in [16] to spectral norm regularized convex optimization problems. There is still much work to be done, such as how to apply the variational properties of proximal mapping of spectral norm function to directly prove the equivalence between (a) and (c) in Theorem 4, whether the conclusions in this article or in [16] can be extended to the Ky Fan  $k$ -norm regularized problems, or whether the variational properties of the spectral norm function obtained in this article can be applied to provide stability results for the generalized variational inequality problem with the spectral norm function. These are all the topics we will study in the future.

**Acknowledgments:** The authors are very thankful to the referees for their valuable comments, which have improved the presentation of this manuscript.

**Funding information:** This work was supported by the National Natural Science Foundation of China (Grant numbers 12101101); the China Postdoctoral Science Foundation (Grant number 2022M720631); and the Fundamental Research Funds for the Central Universities (Grant number 3132025206).

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and consented to its submission to the journal, reviewed all the results, and approved the final version of the manuscript. ZY prepared the manuscript with contributions from all co-authors. All authors contributed to the review and editing of this article.

**Conflict of interest:** The authors state no conflict of interests.

## References

- [1] G. A. Watson, *On matrix approximation problems with Ky Fan  $k$  norms*, Numer. Algorithms **5** (1993), 263–272, DOI: <https://doi.org/10.1007/BF02210386>.
- [2] K. C. Toh and L. N. Trefethen, *The Chebyshev polynomials of a matrix*, SIAM J. Matrix Anal. Appl. **20** (1998), 400–419, DOI: <https://doi.org/10.1137/S0895479896303739>.
- [3] P. Apkarian and D. Noll, *Nonsmooth  $H^\infty$  synthesis*, IEEE Trans. Autom. Control **51** (2006), 71–86.
- [4] V. Bompert, D. Noll, and P. Apkarian, *Second-order nonsmooth optimization for  $H^\infty$  synthesis*, Numer. Math. **107** (2007), 433–454, DOI: <https://doi.org/10.1007/s00211-007-0095-9>.
- [5] J. Yuan, D. Yang, D. Xun, K. Teo, C. Wu, A. Li, et al., *Sparse optimal control of cyber-physical systems via PQA approach*, Pac. J. Optim. **21** (2025), no. 3, 553–569, DOI: <https://doi.org/10.61208/pjo-2025-027>.
- [6] A. Johansson, N. Engsner, C. Strannegård, and P. Mostad, *Improved spectral norm regularization for neural networks*, in: Modeling Decisions for Artificial Intelligence: 20th International Conference, MDAI 2023, Umeå, Sweden, June 19–22, 2023, Proceedings, Springer-Verlag, Berlin, Heidelberg, pp. 181–201, DOI: [https://doi.org/10.1007/978-3-031-33498-6\\_13](https://doi.org/10.1007/978-3-031-33498-6_13).
- [7] Y. Yoshida and M. Takeru, *Spectral norm regularization for improving the generalizability of deep learning*, arXiv:1705.10941 (2017).
- [8] R. Gao, W. Yang, and X. Sun, *Defying forgetting in continual relation extraction via batch spectral norm regularization*, 2024 International Joint Conference on Neural Networks (IJCNN), 2024, pp. 1–8, DOI: <https://doi.org/10.1109/IJCNN60899.2024.10651110>.
- [9] D. R. Han, D. F. Sun, and L. W. Zhang, *Linear rate convergence of the alternating direction method of multipliers for convex composite programming*, Math. Oper. Res. **43** (2018), 622–637, DOI: <https://doi.org/10.1287/moor.2017.0875>.
- [10] R. T. Rockafellar, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Oper. Res. **1** (1976), 97–116, DOI: <https://doi.org/10.1287/moor.1.2.97>.
- [11] Y. L. Zhang and L. W. Zhang, *On the upper Lipschitz property of the KKT mapping for nonlinear semidefinite optimization*, Oper. Res. Lett. **44** (2016), 474–478, DOI: <https://doi.org/10.1016/j.orl.2016.04.012>.
- [12] Y. Zhang, L. W. Zhang, J. Wu, and K. Wang, *Characterizations of local upper Lipschitz property of perturbed solutions to nonlinear second-order cone programs*, Optimization **66** (2017), 1079–1103, DOI: <https://doi.org/10.1080/02331934.2017.1325887>.
- [13] C. Ding, D. F. Sun, and L. W. Zhang, *Characterization of the robust isolated calmness for a class of conic programming problems*, SIAM J. Optim. **5** (2017), 67–90, DOI: <https://doi.org/10.1137/16M1058753>.
- [14] Z. X. Chan and D. F. Sun, *Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming*, SIAM J. Optim. **19** (2008), 370–396, DOI: <https://doi.org/10.1137/070681235>.
- [15] D. R. Han, D. F. Sun, and L. W. Zhang, *Linear rate convergence of the alternating direction method of multipliers for convex composite quadratic and semi-definite programming*, arXiv:1508.02134 (2015).

- [16] Y. Cui and D. F. Sun, *A complete characterization of the robust isolated calmness of nuclear norm regularized convex optimization problems*, J. Comput. Math. **36** (2018), 441–458, DOI: <https://doi.org/10.4208/jcm.1709-m2017-0034>.
- [17] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, Berlin, Heidelberg, 1998, DOI: <https://doi.org/10.1007/978-3-642-02431-3>.
- [18] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
- [19] K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations*, Proc. Natl. Acad. Sci. USA **35** (1949), 652–655, DOI: <https://doi.org/10.1073/pnas.36.1.31>.
- [20] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
- [21] A. Beck, *First-Order Methods in Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, 2017, DOI: <https://doi.org/10.1137/1.9781611974997>.
- [22] G. A. Watson, *Characterization of the subdifferential of some matrix norms*, Linear Algebra Appl. **170** (1992), 33–45, DOI: [https://doi.org/10.1016/0024-3795\(92\)90407-2](https://doi.org/10.1016/0024-3795(92)90407-2).
- [23] C. Ding, D. F. Sun, J. Sun, and K.-C. Toh, *Spectral operators of matrices*, Math. Program. **168** (2018), 509–531, DOI: <https://doi.org/10.1007/s10107-017-1162-3>.
- [24] L. W. Zhang, N. Zhang, and X. T. Xiao, *On the second-order directional derivatives of singular values of matrices and symmetric matrix-valued functions*, Set-Valued Var. Anal. **21** (2013), 557–586, DOI: <https://doi.org/10.1007/s11228-013-0237-4>.
- [25] C. Ding, *Variational analysis of the Ky Fan  $k$ -norm*, Set-Valued Var. Anal. **25** (2017), 265–296, DOI: <https://doi.org/10.1007/s11228-016-0378-3>.