



Research article

A comprehensive characterization of the robust isolated calmness of Ky Fan k -norm regularized convex matrix optimization problems

Ziran Yin¹, Chongyang Liu², Xiaoyu Chen¹, Jihong Zhang³ and Jinlong Yuan^{1,*}

¹ School of Science, Dalian Maritime University, Dalian, Liaoning, 116026, China

² School of Mathematics and Information Science, Shandong Technology and Business University, Yantai, Shandong, 264005, China

³ School of Science, Shenyang Ligong University, Shenyang, Liaoning, 110159, China

* Correspondence: Email: yuanjinlong@dlmu.edu.cn.

Abstract: This paper extends a result of isolated calmness for nuclear norm regularized convex optimization problems to Ky Fan k -norm regularized convex optimization problems. We find that there exists a certain equivalence relationship among the critical cones of the Ky Fan k -norm function and its conjugate as well as the “sigma term”, namely, the conjugate function of the parabolic second-order directional derivative of the Ky Fan k -norm. By establishing the equivalence between the primal (dual) strict Robinson constraint qualification (SRCQ) and the dual (primal) second-order sufficient condition (SOSC), we derive a series of complete characterizations of the robust isolated calmness of the Karush-Kuhn-Tucker (KKT) mapping for Ky Fan k -norm regularized convex matrix optimization problems. The obtained results enrich the stability theory of the Ky Fan k -norm regularized convex optimization problems and further enhance the usability of the related algorithms.

Keywords: isolated calmness; Ky Fan k -norm; critical cone; second-order sufficient condition; strict Robinson constraint qualification

Mathematics Subject Classification: 65K10, 90C25, 90C31

1. Introduction

Consider the Ky Fan k -norm regularized convex matrix optimization problem as follows:

$$\begin{aligned} \min_{X \in \mathbb{R}^{m \times n}} & h(AX) + \langle C, X \rangle + \|X\|_{(k)}, \\ \text{s.t.} & BX - b \in \mathcal{P}, \end{aligned} \tag{1.1}$$

where $h : \mathbb{R}^d \rightarrow \mathbb{R}$ is a twice continuously differentiable function on $\text{dom } h$, which is also strictly convex, $A : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^d$ and $B : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^l$ are two linear operators with $m \leq n$, $C \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^l$

are given data, $\mathcal{P} \subseteq \mathbb{R}^l$ is a nonempty convex polyhedral cone, $\|\cdot\|_{(k)}$ denotes the matrix Ky Fan k -norm for an integer $k \in [1, m]$, i.e., the sum of k largest singular values of a given matrix. In particular, when $k = 1$, $\|\cdot\|_{(1)}$ coincides with the spectral norm function $\|\cdot\|_2$, i.e., the largest singular values of matrices; when $k = m$, $\|\cdot\|_{(m)}$ is exactly the nuclear norm function $\|\cdot\|_*$, namely, the sum of singular value of matrices. It is well known that problem (1.1) has a wide range of applications in various fields, such as matrix Chebyshev polynomials [1], H_∞ synthesis problems [2, 3], control problems [4], deep learning and neural networks problems [5, 6], matrix completion and rank minimization problems [7, 8], matrix approximation problems [9], and fastest mixing Markov chain problems [10], etc.

Given the extensive application background of Problem (1.1), efficiently solving such problems has become a significant research focus. Stability analysis theory plays a crucial role in the convergence analysis of optimization algorithms [11–13]. Therefore, it is essential to study the stability properties of Ky Fan k -norm regularized optimization problems. Several studies in the existing literature have explored the stability analysis of Ky Fan k -norm regularized optimization problems. For example, Ding [14] studied the variational properties of the Ky Fan k -norm, laying the foundation for analyzing the optimality conditions and stability theory of Ky Fan k -norm regularized problems. Ding [15] provided a series of equivalent characterizations of the strong regularity of the KKT system for linear matrix cone programming involving the Ky Fan k -norm. Liu and Pan [16] gave sufficient conditions for the locally upper Lipschitz property of the KKT system for cone optimization problems induced by the Ky Fan k -norm epigraph. However, research on other stability properties, such as the characterization of isolated calmness, remains insufficiently comprehensive and mature. Therefore, this paper innovatively provides a comprehensive set of equivalent characterizations of the isolated calmness of the KKT mapping (ICKKTM) for Ky Fan k -norm regularized optimization problems.

The isolated calmness (see Definition 2.1), as one of the most important stability properties, can sometimes guarantee a linear convergence rate for some algorithms, rather than the generic sublinear rate, such as the alternating direction method of multipliers [17] and the proximal augmented Lagrangian method [18]. Numerous research works have been devoted to discovering the sufficient and even necessary conditions for the ICKKTM for optimization problems (cf. e.g., [16, 19–21]). It is highly worth mentioning that in [21], Ding, Sun, and Zhang studied the robust ICKKTM for a large class of conic programming problems (when the constraint sets are \mathcal{C}^2 -cone reducible), they proved that the KKT mapping is robustly isolated calm if and only if both the SRCQ and the SOSC hold. In addition, based on the special linear structure of problem (1.1) and its Lagrangian dual problem, more characterizations of the ICKKTM can be given from the dual perspective. For instance, Han, Sun and Zhang [22] found that for the convex composite quadratic semidefinite programming problem, the primal (dual) SRCQ and the dual (primal) SOSC are equivalent. Furthermore, they obtained five equivalent characterizations of the ICKKTM. Similarly, for the nuclear norm regularized convex optimization problem, Cui and Sun [23] also proved that the primal (dual) SRCQ is equivalent to the dual (primal) SOSC, thereby deriving a series of equivalent conditions of the ICKKTM. Motivated by [23], a natural question emerges: that is, whether the conclusions regarding the nuclear norm regularized problem can be extended to the Ky Fan k -norm regularized problem. We will give a positive answer in this paper.

Compared to existing works, the contributions of this paper are summarized as follows.

- 1) We originally discovered an ingenious relationship among the critical cone of the Ky Fan k -norm

function, the critical cone of its conjugate, and the “sigma term,” which has not been mentioned in the existing literature. This relationship is essential for deriving the main results of this paper. Although the mathematical derivation is not excessively complex, it represents a novel and important finding.

- 2) We derived a more comprehensive set of equivalent characterizations of the ICKKT for Ky Fan k -norm regularized convex optimization problems from the dual perspective, thereby enriching the stability analysis theory of Ky Fan k -norm regularized optimization problems.
- 3) The newly derived equivalent conditions are easier to verify in practical computations, which can enhance the feasibility of algorithm design for solving related problems.
- 4) The Ky Fan k -norm regularized optimization problem studied in this paper is a broader class than the nuclear norm regularized optimization problem. Therefore, the results we obtained have wider applications.

The content of this paper is organized as below. In Section 2, we provide some notations and preliminaries on the variational analysis of the Ky Fan k -norm function. We explore the connection among the critical cones of the Ky Fan k -norm and its conjugate function as well as the “sigma term” in Section 3, which forms the most essential part of this paper. In Section 4, we demonstrate that the primal (dual) SRCQ and the dual (primal) SOSC are equivalent. Furthermore, we obtain several equivalent characterizations for the ICKKT. We conclude this paper in Section 5.

Below, we present some basic notations and symbols for matrices.

- For any positive integer t , we denote by $[t]$ the index set $\{1, \dots, t\}$. For any $Z \in \mathbb{R}^{m \times n}$, 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 use Z_ν to represent the sub-matrix of Z , which is obtained by retaining only all columns in ν , and use $Z_{\mu\nu}$ to represent the sub-matrix of Z , which is obtained by retaining only all rows in μ and columns in ν .
- For any $d \in \mathbb{R}^m$, $\text{Diag}(d)$ represents the $m \times m$ diagonal matrix, where the i -th diagonal element is d_i , $i \in [m]$.
- Let “trace” represent the sum of diagonal elements in a given square matrix. For any two matrices M and N in $\mathbb{R}^{m \times n}$, the inner product of M and N is written as $\langle M, N \rangle := \text{trace}(M^T N)$.
- Denote by S^w the linear space consisting of all $w \times w$ real symmetric matrices, and denote by S_+^w and S_-^w the cones of all $w \times w$ positive and negative semidefinite matrices, respectively.

2. Preliminaries

Let \mathcal{X} and \mathcal{Y} be two finite-dimensional real Euclidean spaces. We use the “ $\mathcal{X} \rightrightarrows \mathcal{Y}$ ” to denote a set-valued mapping from \mathcal{X} to \mathcal{Y} [24, Page 148]. Let $D \subseteq \mathcal{X}$ be a non-empty closed convex set. For any $s \in D$, the tangent cone (cf. [24, Definition 6.1]) of D at s is defined by $\mathcal{T}_D(s) := \{d \in \mathcal{X} \mid \exists s^k \rightarrow s \text{ with } s^k \in D \text{ and } \tau^k \downarrow 0 \text{ s.t. } (s^k - s)/\tau^k \rightarrow d\}$. For a given $x \in \mathcal{X}$, we define $\Pi_D(x) := \arg \min \{\|d - x\| \mid d \in D\}$ as the projection mapping onto D . For any closed convex cone $K \subset \mathcal{Y}$, denote by $K^\circ := \{s \in \mathcal{Y} \mid \langle s, z \rangle \leq 0, \forall z \in K\}$ the polar of K . Given $\rho > 0$ and $z \in \mathcal{X}$, define the ball $\mathbb{B}_\rho(z) := \{z' \in \mathcal{X} \mid \|z' - z\| \leq \rho\}$. Given that $f : \mathcal{X} \rightarrow (-\infty, \infty]$ is a proper closed convex function, we use f^* to represent the conjugate of f and ∂f to represent the subdifferential of f . For the sake of convenience, we denote by θ the Ky Fan k -norm function in the subsequent text, i.e., $\theta(\cdot) := \|\cdot\|_{(k)}$.

The following definition of isolated calmness, taken from [21, Definition 2], is the most important concept in this paper.

Definition 2.1. *The set-valued mapping $\Phi : \mathcal{X} \rightrightarrows \mathcal{Y}$ is said to be isolated calm at \bar{u} for \bar{v} if $\bar{v} \in \Phi(\bar{u})$ and there exist a positive constant κ and neighborhoods \mathcal{U} of \bar{u} and \mathcal{V} of \bar{v} such that*

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

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

For any closed convex function $l : \mathcal{X} \rightarrow (-\infty, +\infty]$, it is known from [25, Proposition 2.58] that l is directionally epidifferentiable. We use $l^\downarrow(x; \cdot)$ to represent the directional epiderivatives of l . Moreover, if $l^\downarrow(x; d)$ is finite for $x \in \text{dom } l$ and $d \in \mathcal{X}$, we define the lower second-order directional epiderivative of l for any $w \in \mathcal{X}$ by

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

Let $W \in \mathbb{R}^{m \times n}$ be any matrix and \mathcal{O}^p be the set of all $p \times p$ orthogonal matrices. We use $\sigma_1(W) \geq \sigma_2(W) \geq \dots \geq \sigma_m(W)$ to denote the singular values of W with multiplicity being taken into account. Define $\sigma(W) := (\sigma_1(W), \sigma_2(W), \dots, \sigma_m(W))^T$ and $\Sigma(W) := \text{Diag}(\sigma(W))$. Suppose that $W \in \mathbb{R}^{m \times n}$ has the singular value decomposition (SVD) as follows:

$$W = U[\Sigma(W) \ 0]V^T = U[\Sigma(W) \ 0][V_1 \ V_2]^T = U\Sigma(W)V_1^T, \quad (2.2)$$

where $U \in \mathcal{O}^m$ and $V = [V_1 \ V_2] \in \mathcal{O}^n$ with $V_1 \in \mathbb{R}^{n \times m}$ and $V_2 \in \mathbb{R}^{n \times (n-m)}$. Define three index sets as follows:

$$a := \{i \in [m] \mid \sigma_i(W) > 0\}, \quad b := \{i \in [m] \mid \sigma_i(W) = 0\}, \quad c := \{m+1, \dots, n\}. \quad (2.3)$$

We know from [26, Theorem 31.5] that (P, Q) is a solution of the generalized equation

$$0 \in -P + \partial\theta(Q) \iff 0 \in -Q + \partial\theta^*(P),$$

if and only if

$$P - \text{Prox}_\theta(P + Q) = 0 \iff Q - \text{Prox}_{\theta^*}(P + Q) = 0,$$

where $\text{Prox}_\theta : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ and $\text{Prox}_{\theta^*} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ are the Moreau-Yosida proximal mappings of θ (cf., e.g., [27, Definition 6.1]) and θ^* , respectively, i.e.,

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

Let $W = P + Q$, and assume that W admits the SVD as in (2.2). By [27, Theorem 7.29 and Example 7.31], we know that

$$P = U[\Sigma(P) \ 0]V^T, \quad Q = U[\Sigma(Q) \ 0]V^T.$$

Denote $\bar{\sigma} = \sigma(P)$ and $\bar{q} = \sigma(Q)$. According to [14, Lemma 3], we know that if $\bar{\sigma}_k > 0$, there exist two integers $0 \leq s_0 \leq k - 1$ and $k \leq s_1 \leq m$ such that

$$\bar{\sigma}_1 \geq \dots \geq \bar{\sigma}_{s_0} > \bar{\sigma}_{s_0+1} = \dots = \bar{\sigma}_k = \dots = \bar{\sigma}_{s_1} > \bar{\sigma}_{s_1+1} \geq \dots \geq \bar{\sigma}_m \geq 0.$$

Then,

$$\bar{q}_i = 1, \text{ if } i \in \alpha; \quad 0 \leq \bar{q}_i \leq 1, \text{ if } i \in \beta \text{ and } \sum_{i \in \beta} \bar{q}_i = k - s_0; \quad \bar{q}_i = 0, \text{ if } i \in \gamma,$$

where

$$\alpha = \{1, \dots, s_0\}, \quad \beta = \{s_0 + 1, \dots, s_1\}, \quad \gamma = \{s_1 + 1, \dots, m\}. \quad (2.4)$$

If $\bar{\sigma}_k = 0$, there exists an integer $0 \leq s_0 \leq k - 1$ such that

$$\bar{\sigma}_1 \geq \dots \geq \bar{\sigma}_{s_0} > \bar{\sigma}_{s_0+1} = \dots = \bar{\sigma}_k = \dots = \bar{\sigma}_m = 0.$$

Then,

$$\bar{q}_i = 1, \text{ if } i \in \alpha; \quad 0 \leq \bar{q}_i \leq 1, \text{ if } i \in \beta \text{ and } \sum_{i \in \beta} \bar{q}_i \leq k - s_0,$$

where

$$\alpha = \{1, \dots, s_0\}, \quad \beta = \{s_0 + 1, \dots, m\}. \quad (2.5)$$

For convenience, we subdivide the index set β into three index sets:

$$\beta_1 := \{i \in \beta \mid \bar{q}_i = 1\}, \quad \beta_2 := \{i \in \beta \mid 0 < \bar{q}_i < 1\}, \quad \beta_3 := \{i \in \beta \mid \bar{q}_i = 0\}. \quad (2.6)$$

Let $P \in \partial\theta(Q)$ (or equivalently $Q \in \partial\theta^*(P)$). Define the critical cone of θ at P with respect to Q as

$$C_\theta(P, Q) = \{D \in \mathbb{R}^{m \times n} \mid \theta'(P; D) = \langle Q, D \rangle\}.$$

Likewise, define the critical cone of θ^* at Q with respect to P as

$$C_{\theta^*}(Q, P) = \{D \in \mathbb{R}^{m \times n} \mid (\theta^*)'(Q; D) = \langle D, P \rangle\}.$$

For any two matrices $Y, Z \in \mathbb{R}^{m \times n}$, according to the well-known von Neumann's trace inequality [28], we have the following result:

$$\langle Y, Z \rangle \leq \sigma(Y)^T \sigma(Z). \quad (2.7)$$

For any two matrices $Y, Z \in \mathcal{S}^p$, by Fan's inequality [29], we have that

$$\langle Y, Z \rangle \leq \lambda(Y)^T \lambda(Z), \quad (2.8)$$

where $\lambda(Y)$ represents the eigenvalue vector of Y with its elements arranged in non-increasing order.

For the sake of notational convenience, for any integer $p > 0$, we define a linear matrix operator $S : \mathbb{R}^{p \times p} \rightarrow \mathcal{S}^p$ by

$$S(N) := \frac{1}{2}(N + N^T).$$

As the principal tools for the subsequent analysis, we need to utilize the characterizations of the subdifferential of θ , the critical cones of θ and its conjugate function, as well as the characterization of the sigma term being zero. These contents have already been provided in [14]. We summarize them as follows.

Lemma 2.1. Given $P, Q \in \mathbb{R}^{m \times n}$. Suppose that $P \in \partial\theta(Q)$ and $W = P + Q$ admit the SVD as in (2.2). For any $D \in \mathbb{R}^{m \times n}$, denote $\tilde{D} = U^T D V$. Then $D \in C_\theta(P, Q)$ is equivalent to the subsequent conditions:

(i) If $\sigma_k(P) > 0$, then there exists some $\tau \in \mathbb{R}$ such that

$$\lambda_{|\beta_1|}(S(\tilde{D}_{\beta_1\beta_1})) \geq \tau \geq \lambda_1(S(\tilde{D}_{\beta_3\beta_3}))$$

and

$$S(\tilde{D}_{\beta\beta}) = \begin{bmatrix} S(\tilde{D}_{\beta_1\beta_1}) & 0 & 0 \\ 0 & \tau I_{|\beta_2|} & 0 \\ 0 & 0 & S(\tilde{D}_{\beta_3\beta_3}) \end{bmatrix},$$

where $I_{|\beta_2|}$ is the $|\beta_2| \times |\beta_2|$ identity matrix.

(ii) If $\sigma_k(P) = 0$ and $\|Q\|_* = k$, then there exists some $\tau \geq 0$ such that

$$\lambda_{|\beta_1|}(S(\tilde{D}_{\beta_1\beta_1})) \geq \tau \geq \sigma_1([\tilde{D}_{bb} \quad \tilde{D}_{bc}])$$

and

$$[\tilde{D}_{\beta\beta} \quad \tilde{D}_{\beta c}] = \begin{bmatrix} S(\tilde{D}_{\beta_1\beta_1}) & 0 & 0 & 0 \\ 0 & \tau I_{|\beta_2|} & 0 & 0 \\ 0 & 0 & \tilde{D}_{bb} & \tilde{D}_{bc} \end{bmatrix}.$$

(iii) If $\sigma_k(P) = 0$ and $\|Q\|_* < k$, then $S(\tilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_+^{|\beta_1|}$ and

$$[\tilde{D}_{\beta\beta} \quad \tilde{D}_{\beta c}] = \begin{bmatrix} S(\tilde{D}_{\beta_1\beta_1}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Lemma 2.2. Given $P, Q \in \mathbb{R}^{m \times n}$. Suppose that $Q \in \partial\theta^*(P)$ and $W = P + Q$ admit the SVD as in (2.2). Assume that $\|Q\|_{(k)}^* = 1$, where $\|\cdot\|_{(k)}^*$ denote the dual norm of $\|\cdot\|_{(k)}$. For any $D \in \mathbb{R}^{m \times n}$, denote $\tilde{D} = U^T D V$. Then $D \in C_{\theta^*}(Q, P)$ is equivalent to the subsequent conditions:

(i) If $\sigma_k(P) > 0$, then $\text{trace}(\tilde{D}_{\beta\beta}) = 0$,

$$S(\tilde{D}_{\alpha\cup\beta_1\alpha\cup\beta_1}) = \begin{bmatrix} 0 & 0 \\ 0 & S(\tilde{D}_{\beta_1\beta_1}) \end{bmatrix} \quad \text{with} \quad S(\tilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_-^{|\beta_1|}$$

and

$$\begin{bmatrix} \tilde{D}_{\beta_3\beta_3} & \tilde{D}_{\beta_3\gamma} & \tilde{D}_{\beta_3c} \\ \tilde{D}_{\gamma\beta_3} & \tilde{D}_{\gamma\gamma} & \tilde{D}_{\gamma c} \end{bmatrix} = \begin{bmatrix} S(\tilde{D}_{\beta_3\beta_3}) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{with} \quad S(\tilde{D}_{\beta_3\beta_3}) \in \mathcal{S}_+^{|\beta_3|};$$

(ii) If $\sigma_k(P) = 0$ and $\|Q\|_* = k$, then $\text{trace}(\tilde{D}_{\beta_1\cup\beta_2\beta_1\cup\beta_2}) + \|[\tilde{D}_{bb} \quad \tilde{D}_{bc}]\|_* \leq 0$,

$$S(\tilde{D}_{\alpha\cup\beta_1\alpha\cup\beta_1}) = \begin{bmatrix} 0 & 0 \\ 0 & S(\tilde{D}_{\beta_1\beta_1}) \end{bmatrix} \quad \text{with} \quad S(\tilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_-^{|\beta_1|};$$

(iii) If $\sigma_k(P) = 0$ and $\|Q\|_* < k$, then

$$S(\tilde{D}_{\alpha\cup\beta_1\alpha\cup\beta_1}) = \begin{bmatrix} 0 & 0 \\ 0 & S(\tilde{D}_{\beta_1\beta_1}) \end{bmatrix} \quad \text{with} \quad S(\tilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_-^{|\beta_1|}.$$

Since the Ky Fan k -norm function θ is convex and Lipschitz continuous, and the singular value function in $\mathbb{R}^{m \times m}$ is second-order directionally differentiable [30, Theorem 3.1], the lower second-order directional epiderivative of θ coincides with the parabolic second-order directional derivative of θ , i.e., for any $P \in \mathbb{R}^{m \times n}$, $\theta_{-}^{\downarrow\downarrow}(P; \cdot) = \theta''(P; \cdot)$. For any given $P \in \partial\theta(Q)$, $\forall D \in \mathbb{R}^{m \times n}$, denote

$$\psi_{(P,D)}^*(Q) := (\theta''(P; D, \cdot))^*(Q)$$

as the conjugate of $\theta''(P; D, \cdot)$ at Q . This can be regarded as the “sigma term” of problem (1.1). Similarly, we denote

$$\phi_{(Q,D)}^*(P) := ((\theta^*)''(Q; D, \cdot))^*(P)$$

as the conjugate of $(\theta^*)''(Q; D, \cdot)$ at P .

Ding [14, Proposition 16] provides the properties of the “sigma term” as well as the equivalent characterizations in the case where the “sigma term” equals 0, which are presented as follows:

Lemma 2.3. *Given that $P = \partial\theta(Q)$. For any $D \in \mathbb{R}^{m \times n}$, denote $\tilde{D} = U^T D V = [\tilde{D}_1 \quad \tilde{D}_2]$ with $\tilde{D}_1 \in \mathbb{R}^{m \times m}$ and $\tilde{D}_2 \in \mathbb{R}^{m \times (n-m)}$. Then, $\psi_{(P,D)}^*(Q) \leq 0$, $\phi_{(Q,D)}^*(P) \leq 0$. Moreover,*

$$\psi_{(P,D)}^*(Q) = 0 \iff \phi_{(Q,D)}^*(P) = 0,$$

which is also equivalent to the subsequent conditions:

(i) If $\sigma_k(P) > 0$,

$$\tilde{D} = \begin{pmatrix} S(\tilde{D}_1)_{\alpha\alpha} & S(\tilde{D}_1)_{\alpha\beta_1} & 0 & 0 & 0 & 0 \\ S(\tilde{D}_1)_{\beta_1\alpha} & S(\tilde{D}_1)_{\beta_1\beta_1} & S(\tilde{D}_1)_{\beta_1\beta_2} & S(\tilde{D}_1)_{\beta_1\beta_3} & 0 & 0 \\ 0 & S(\tilde{D}_1)_{\beta_2\beta_1} & S(\tilde{D}_1)_{\beta_2\beta_2} & S(\tilde{D}_1)_{\beta_2\beta_3} & 0 & 0 \\ 0 & S(\tilde{D}_1)_{\beta_3\beta_1} & S(\tilde{D}_1)_{\beta_3\beta_2} & \tilde{D}_{\beta_3\beta_3} & \tilde{D}_{\beta_3\gamma} & \tilde{D}_{\beta_3c} \\ 0 & 0 & 0 & \tilde{D}_{\gamma\beta_3} & \tilde{D}_{\gamma\gamma} & \tilde{D}_{\gamma c} \end{pmatrix}.$$

(ii) If $\sigma_k(P) = 0$,

$$\tilde{D} = \begin{pmatrix} S(\tilde{D}_1)_{\alpha\alpha} & S(\tilde{D}_1)_{\alpha\beta_1} & 0 & 0 & 0 \\ S(\tilde{D}_1)_{\beta_1\alpha} & \tilde{D}_{\beta_1\beta_1} & \tilde{D}_{\beta_1\beta_2} & \tilde{D}_{\beta_1b} & \tilde{D}_{\beta_1c} \\ 0 & \tilde{D}_{\beta_2\beta_1} & \tilde{D}_{\beta_2\beta_2} & \tilde{D}_{\beta_2b} & \tilde{D}_{\beta_2c} \\ 0 & \tilde{D}_{b\beta_1} & \tilde{D}_{b\beta_2} & \tilde{D}_{bb} & \tilde{D}_{bc} \end{pmatrix}.$$

3. The properties of critical cones

In this section, we present the relationships among the critical cone of the Ky Fan k -norm function, the critical cone of its conjugate, and the “sigma term”, which are of vital significance in this paper.

Proposition 3.1. *Assume that $W \in \mathbb{R}^{m \times n}$ has the SVD as in (2.2). Denote $P = \text{Prox}_\theta(W)$ and $Q = \text{Prox}_{\theta^*}(W)$. Suppose the index sets $\alpha, \beta, \gamma, \beta_1, \beta_2, \beta_3, b, c$ are defined by (2.3), (2.4) or (2.5) and (2.6). For any $D \in \mathbb{R}^{m \times n}$, denote $\tilde{D} = U^T D V$. Then the following conclusions hold:*

- (i) If $\sigma_k(P) > 0$, then $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P)$ imply that $D \in (C_\theta(P, Q))^\circ$.
- (ii) If $\sigma_k(P) = 0$, $\|Q\|_* = k$, then $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P)$ imply that $D \in (C_\theta(P, Q))^\circ$.

(iii) If $\sigma_k(P) = 0$, $\|Q\|_* < k$, then $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P) \iff D \in (C_\theta(P, Q))^\circ$.

Proof. Consider the following three cases.

Case 1. $\sigma_k(P) > 0$. If $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P)$, by part (i) of Lemma 2.2 and part (i) of Lemma 2.3, we can obtain that

$$\tilde{D} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & S(\tilde{D}_{\beta_1\beta_1}) & S(\tilde{D}_{\beta_1\beta_2}) & S(\tilde{D}_{\beta_1\beta_3}) & 0 & 0 \\ 0 & S(\tilde{D}_{\beta_2\beta_1}) & S(\tilde{D}_{\beta_2\beta_2}) & S(\tilde{D}_{\beta_2\beta_3}) & 0 & 0 \\ 0 & S(\tilde{D}_{\beta_3\beta_1}) & S(\tilde{D}_{\beta_3\beta_2}) & S(\tilde{D}_{\beta_3\beta_3}) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (3.1)$$

and

$$S(\tilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_-^{|\beta_1|}, S(\tilde{D}_{\beta_3\beta_3}) \in \mathcal{S}_+^{|\beta_3|} \quad \text{and} \quad \text{trace}(\tilde{D}_{\beta\beta}) = 0. \quad (3.2)$$

For any $H \in C_\theta(P, Q)$, from part (i) of Lemma 2.1, $\exists \tau \in \mathfrak{R}$ s.t.

$$\lambda_{|\beta_1|}(S(\tilde{H}_{\beta_1\beta_1})) \geq \tau \geq \lambda_1(S(\tilde{H}_{\beta_3\beta_3}))$$

and

$$S(\tilde{H}_{\beta\beta}) = \begin{bmatrix} S(\tilde{H}_{\beta_1\beta_1}) & 0 & 0 \\ 0 & \tau I_{|\beta_2|} & 0 \\ 0 & 0 & S(\tilde{H}_{\beta_3\beta_3}) \end{bmatrix}.$$

Thus, $\forall H \in C_\theta(P, Q)$ and D satisfies (3.1) and (3.2), according to Fan's inequality (2.8), we have that

$$\begin{aligned} \langle H, D \rangle &= \langle \tilde{H}, \tilde{D} \rangle = \langle \tilde{H}_{\beta\beta}, \tilde{D}_{\beta\beta} \rangle = \langle \tilde{H}_{\beta\beta}, S(\tilde{D}_{\beta\beta}) \rangle = \langle S(\tilde{H}_{\beta\beta}), \tilde{D}_{\beta\beta} \rangle \\ &= \langle S(\tilde{H}_{\beta_1\beta_1}), \tilde{D}_{\beta_1\beta_1} \rangle + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + \langle S(\tilde{H}_{\beta_3\beta_3}), \tilde{D}_{\beta_3\beta_3} \rangle \\ &\leq (\lambda(S(\tilde{H}_{\beta_1\beta_1})))^T \lambda(\tilde{D}_{\beta_1\beta_1}) + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + (\lambda(S(\tilde{H}_{\beta_3\beta_3})))^T \lambda(\tilde{D}_{\beta_3\beta_3}) \\ &= (-\lambda(S(\tilde{H}_{\beta_1\beta_1})))^T (-\lambda(\tilde{D}_{\beta_1\beta_1})) + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + (\lambda(S(\tilde{H}_{\beta_3\beta_3})))^T \lambda(\tilde{D}_{\beta_3\beta_3}) \\ &\leq \tau \text{trace}(\tilde{D}_{\beta_1\beta_1}) + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + \tau \text{trace}(\tilde{D}_{\beta_3\beta_3}) \\ &= \tau \text{trace}(\tilde{D}_{\beta\beta}) \\ &= 0. \end{aligned}$$

This implies that $D \in (C_\theta(P, Q))^\circ$. The proof of the part (i) is complete.

Case 2. $\sigma_k(P) = 0$, $\|Q\|_* = k$. If $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P)$, according to part (ii) of Lemma 2.2 and part (ii) of Lemma 2.3, we have that

$$\tilde{D} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \tilde{D}_{\beta_1\beta_1} & \tilde{D}_{\beta_1\beta_2} & \tilde{D}_{\beta_1b} & \tilde{D}_{\beta_1c} \\ 0 & \tilde{D}_{\beta_2\beta_1} & \tilde{D}_{\beta_2\beta_2} & \tilde{D}_{\beta_2b} & \tilde{D}_{\beta_2c} \\ 0 & \tilde{D}_{b\beta_1} & \tilde{D}_{b\beta_2} & \tilde{D}_{bb} & \tilde{D}_{bc} \end{pmatrix} \quad (3.3)$$

and

$$S(\tilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_-^{|\beta_1|} \quad \text{and} \quad \text{trace}(\tilde{D}_{\beta_1 \cup \beta_2 \beta_1 \cup \beta_2}) + \|[\tilde{D}_{bb} \quad \tilde{D}_{bc}] \|_* \leq 0. \quad (3.4)$$

Therefore, $\forall H \in C_\theta(P, Q)$, based on part (ii) of Lemma 2.1, we know that there exists $\tau \geq 0$ such that

$$\lambda_{|\beta_1|}(S(\tilde{H}_{\beta_1\beta_1})) \geq \tau \geq \sigma_1([\tilde{H}_{bb} \quad \tilde{H}_{bc}])$$

and

$$[\tilde{H}_{\beta\beta} \quad \tilde{H}_{\beta c}] = \begin{bmatrix} S(\tilde{H}_{\beta_1\beta_1}) & 0 & 0 & 0 \\ 0 & \tau I_{|\beta_2|} & 0 & 0 \\ 0 & 0 & \tilde{H}_{bb} & \tilde{H}_{bc} \end{bmatrix}.$$

Hence, $\forall H \in C_\theta(P, Q)$ and D satisfies (3.3) and (3.4), by von Neumann’s trace inequality (2.7), we have that

$$\begin{aligned} \langle H, D \rangle &= \langle \tilde{H}, \tilde{D} \rangle = \langle \tilde{H}_{\beta\beta}, \tilde{D}_{\beta\beta} \rangle + \langle \tilde{H}_{\beta c}, \tilde{D}_{\beta c} \rangle \\ &= \langle S(\tilde{H}_{\beta_1\beta_1}), S(\tilde{D}_{\beta_1\beta_1}) \rangle + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + \langle [\tilde{H}_{bb} \quad \tilde{H}_{bc}], [\tilde{D}_{bb} \quad \tilde{D}_{bc}] \rangle \\ &\leq (\lambda(S(\tilde{H}_{\beta_1\beta_1})))^T \lambda(S(\tilde{D}_{\beta_1\beta_1})) + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + (\sigma([\tilde{H}_{bb} \quad \tilde{H}_{bc}]))^T \sigma([\tilde{D}_{bb} \quad \tilde{D}_{bc}]) \\ &= (-\lambda(S(\tilde{H}_{\beta_1\beta_1})))^T (-\lambda(S(\tilde{D}_{\beta_1\beta_1}))) + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + (\sigma([\tilde{H}_{bb} \quad \tilde{H}_{bc}]))^T \sigma([\tilde{D}_{bb} \quad \tilde{D}_{bc}]) \\ &\leq \tau \text{trace}(\tilde{D}_{\beta_1\beta_1}) + \tau \text{trace}(\tilde{D}_{\beta_2\beta_2}) + \tau \|[\tilde{D}_{bb} \quad \tilde{D}_{bc}]\|_* \\ &\leq 0. \end{aligned}$$

This means that $D \in (C_\theta(P, Q))^\circ$. The proof of the part (ii) is complete.

Case 3. $\sigma_k(P) = 0, \|Q\|_* < k$. According to part (iii) of Lemma 2.2 and part (ii) of Lemma 2.3, we know that $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P)$ hold if and only if $S(\tilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_-^{|\beta_1|}$ and

$$\tilde{D} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \tilde{D}_{\beta_1\beta_1} & \tilde{D}_{\beta_1\beta_2} & \tilde{D}_{\beta_1b} & \tilde{D}_{\beta_1c} \\ 0 & \tilde{D}_{\beta_2\beta_1} & \tilde{D}_{\beta_2\beta_2} & \tilde{D}_{\beta_2b} & \tilde{D}_{\beta_2c} \\ 0 & \tilde{D}_{b\beta_1} & \tilde{D}_{b\beta_2} & \tilde{D}_{bb} & \tilde{D}_{bc} \end{pmatrix}.$$

In view of part (iii) of Lemma 2.1, this is equivalent to $D \in (C_\theta(P, Q))^\circ$. This completes the proof. \square

Remark 3.1. *The conclusions in the converse directions of (i) and (ii) in Proposition 3.1 are not always valid. That is because when the parameter τ within $C_\theta(P, Q)$ equals 0, we are unable to obtain the conditions such as $\text{trace}(\tilde{D}_{\beta\beta}) = 0$ or $\text{trace}(\tilde{D}_{\beta_1 \cup \beta_2 \beta_1 \cup \beta_2}) + \|[\tilde{D}_{bb} \quad \tilde{D}_{bc}]\|_* \leq 0$. However, there is a situation where the conditions $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P)$ are equivalent to $D \in (C_\theta(P, Q))^\circ$ in Proposition 3.1 when the index set β_2 is empty. In particular, when $k = m$, that is, when θ is the nuclear norm function, Cui and Sun [23, Proposition 4.2] have already presented this equivalent property.*

By exchanging the positions of $C_\theta(P, Q)$ and $C_{\theta^*}(Q, P)$ in the above proposition, we can obtain the following similar yet enhanced results.

Proposition 3.2. *Assume that $W \in \mathbb{R}^{m \times n}$ has the SVD as in (2.2). Denote $P = \text{Prox}_\theta(W)$ and $Q = \text{Prox}_{\theta^*}(W)$. Suppose the index sets $\alpha, \beta, \gamma, \beta_1, \beta_2, \beta_3, b, c$ are defined by (2.3), (2.4) or (2.5), and (2.6). For any $D \in \mathbb{R}^{m \times n}$, denote $\tilde{D} = U^T D V$. Then the subsequent two conditions are equivalent:*

- (i) $\psi_{(P,D)}^*(Q) = 0$ and $D \in C_\theta(P, Q)$.

(ii) $D \in (C_{\theta^*}(Q, P))^{\circ}$.

Proof. Case 1. $\sigma_k(P) > 0$. “(i) \implies (ii)”. It is easy to verify $\psi_{(P,D)}^*(Q) = 0$ and $D \in C_{\theta}(P, Q)$ are equivalent to

$$\tilde{D} = \begin{pmatrix} S(\tilde{D}_1)_{\alpha\alpha} & S(\tilde{D}_1)_{\alpha\beta_1} & 0 & 0 & 0 & 0 \\ S(\tilde{D}_1)_{\beta_1\alpha} & S(\tilde{D}_1)_{\beta_1\beta_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau I_{|\beta_2|} & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{D}_{\beta_3\beta_3} & \tilde{D}_{\beta_3\gamma} & \tilde{D}_{\beta_3c} \\ 0 & 0 & 0 & \tilde{D}_{\gamma\beta_3} & \tilde{D}_{\gamma\gamma} & \tilde{D}_{\gamma c} \end{pmatrix}, \quad (3.5)$$

where τ is some real number that satisfies

$$\lambda_{|\beta_1|}(S(\tilde{D}_{\beta_1\beta_1})) \geq \tau \geq \lambda_1(S(\tilde{D}_{\beta_3\beta_3})). \quad (3.6)$$

Hence, for any D that satisfies (3.5) and (3.6), $\forall H \in C_{\theta^*}(Q, P)$, i.e., $\text{trace}(\tilde{H}_{\beta\beta}) = 0$,

$$S(\tilde{H}_{\alpha\cup\beta_1\alpha\cup\beta_1}) = \begin{bmatrix} 0 & 0 \\ 0 & S(\tilde{H}_{\beta_1\beta_1}) \end{bmatrix} \quad \text{with} \quad S(\tilde{H}_{\beta_1\beta_1}) \in \mathcal{S}_-^{|\beta_1|}$$

and

$$\begin{bmatrix} \tilde{H}_{\beta_3\beta_3} & \tilde{H}_{\beta_3\gamma} & \tilde{H}_{\beta_3c} \\ \tilde{H}_{\gamma\beta_3} & \tilde{H}_{\gamma\gamma} & \tilde{H}_{\gamma c} \end{bmatrix} = \begin{bmatrix} S(\tilde{H}_{\beta_3\beta_3}) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{with} \quad S(\tilde{H}_{\beta_3\beta_3}) \in \mathcal{S}_+^{|\beta_3|},$$

we have that

$$\begin{aligned} \langle H, D \rangle &= \langle \tilde{H}, \tilde{D} \rangle = \langle \tilde{H}_{\beta_1\beta_1}, S(\tilde{D}_{\beta_1\beta_1}) \rangle + \langle \tilde{H}_{\beta_2\beta_2}, \tau I_{|\beta_2|} \rangle + \langle S(\tilde{H}_{\beta_3\beta_3}), \tilde{D}_{\beta_3\beta_3} \rangle \\ &= \langle \tilde{H}, \tilde{D} \rangle = \langle \tilde{H}_{\beta_1\beta_1}, S(\tilde{D}_{\beta_1\beta_1}) \rangle + \langle \tilde{H}_{\beta_2\beta_2}, \tau I_{|\beta_2|} \rangle + \langle \tilde{H}_{\beta_3\beta_3}, S(\tilde{D}_{\beta_3\beta_3}) \rangle \\ &\leq (\lambda(S(\tilde{D}_{\beta_1\beta_1})))^T \lambda(\tilde{H}_{\beta_1\beta_1}) + \tau \text{trace}(\tilde{H}_{\beta_2\beta_2}) + (\lambda(S(\tilde{D}_{\beta_3\beta_3})))^T \lambda(\tilde{H}_{\beta_3\beta_3}) \\ &= (-\lambda(S(\tilde{D}_{\beta_1\beta_1})))^T (-\lambda(\tilde{H}_{\beta_1\beta_1})) + \tau \text{trace}(\tilde{H}_{\beta_2\beta_2}) + (\lambda(S(\tilde{D}_{\beta_3\beta_3})))^T \lambda(\tilde{H}_{\beta_3\beta_3}) \\ &\leq \tau \text{trace}(\tilde{H}_{\beta_1\beta_1}) + \tau \text{trace}(\tilde{H}_{\beta_2\beta_2}) + \tau \text{trace}(\tilde{H}_{\beta_3\beta_3}) \\ &= \tau \text{trace}(\tilde{H}_{\beta\beta}) \\ &= 0. \end{aligned}$$

This means that $D \in (C_{\theta^*}(Q, P))^{\circ}$. Through reversing the prior arguments, it is possible to prove the converse of this statement.

Case 2. $\sigma_k(P) = 0$, $|Q|_* = k$. Obviously, both (i) and (ii) are equivalent to

$$\tilde{D} = \begin{pmatrix} S(\tilde{D}_1)_{\alpha\alpha} & S(\tilde{D}_1)_{\alpha\beta_1} & 0 & 0 & 0 \\ S(\tilde{D}_1)_{\beta_1\alpha} & S(\tilde{D}_1)_{\beta_1\beta_1} & 0 & 0 & 0 \\ 0 & 0 & \tau I_{|\beta_2|} & 0 & 0 \\ 0 & 0 & 0 & \tilde{D}_{bb} & \tilde{D}_{bc} \end{pmatrix},$$

where $\tau \geq 0$ and satisfies

$$\lambda_{|\beta_1|}(S(\tilde{D}_{\beta_1\beta_1})) \geq \tau \geq \sigma_1([\tilde{D}_{bb} \quad \tilde{D}_{bc}]).$$

This result can be obtained in a similar way as in part (ii) of proposition 3.1. For simplicity, the detailed proof is omitted here.

Case 3. $\sigma_k(P) = 0$, $\|Q\|_* < k$. According to part (iii) of Lemma 2.1, part (iii) of Lemma 2.2, and part (ii) of Lemma 2.3, we obtain that either $\psi_{(P,D)}^*(Q) = 0$ with $D \in C_\theta(P, Q)$ or $D \in (C_{\theta^*}(Q, P))^\circ$ is equivalent to $S(\widetilde{D}_{\beta_1\beta_1}) \in \mathcal{S}_+^{|\beta_1|}$ and

$$\widetilde{D} = \begin{pmatrix} S(\widetilde{D}_1)_{\alpha\alpha} & S(\widetilde{D}_1)_{\alpha\beta_1} & 0 & 0 & 0 \\ S(\widetilde{D}_1)_{\beta_1\alpha} & S(\widetilde{D}_{\beta_1\beta_1}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The proof is complete. \square

4. Equivalent conditions

In this section, from the dual perspective, we will establish several equivalent characterizations of the robust ICKKT for problem (1.1). The key tools we use are Propositions 3.1 and 3.2. Indeed, when $k = m$, the similar results have already been provided in [23]. As a result, we will only make a brief statement here.

The Lagrangian dual of problem (1.1) is presented below:

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

where $(y, w, S) \in \mathbb{R}^l \times \mathbb{R}^d \times \mathbb{R}^{m \times n}$, B^* and A^* represent the adjoint of B and A , respectively. The KKT conditions for problem (1.1) and problem (4.1) are given by

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

and

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

respectively. We use $\mathcal{M}_P(X) \subseteq \mathbb{R}^l \times \mathbb{R}^{m \times n}$ to represent the set of Lagrangian multipliers with respect to X for problem (1.1), namely, $\mathcal{M}_P(X) := \{(y, S) \in \mathbb{R}^l \times \mathbb{R}^{m \times n} \mid (X, y, S) \text{ satisfies (4.2)}\}$. We denote by $\mathcal{M}_D(y, w, S)$ the set of Lagrangian multipliers with respect to (y, w, S) for problem (4.1), i.e., $\mathcal{M}_D(y, w, S) := \{X \in \mathbb{R}^{m \times n} \mid (y, w, S, X) \text{ satisfies (4.3)}\}$. Let $(y, z) \in \mathbb{R}^l \times \mathbb{R}^l$ satisfy $\mathcal{P}^\circ \ni y \perp z \in \mathcal{P}$. Define the critical cone of \mathcal{P} at z for y as $\mathcal{C}_{\mathcal{P}}(z, y) := \mathcal{T}_{\mathcal{P}}(z) \cap y^\perp$ and the critical cone of \mathcal{P}° at y for z as $\mathcal{C}_{\mathcal{P}^\circ}(y, z) := \mathcal{T}_{\mathcal{P}^\circ}(y) \cap z^\perp$, respectively.

Let $\bar{X} \in \mathbb{R}^{m \times n}$ be an optimal solution of problem (1.1) and $(\bar{y}, \bar{S}) \in \mathcal{M}_P(\bar{X})$. The primal SOS is said to hold at \bar{X} with respect to (\bar{y}, \bar{S}) if

$$\langle AD, \nabla^2 h(A\bar{X})AD \rangle - \psi_{(\bar{X}, D)}^*(\bar{S}) > 0, \quad \forall D \in C(\bar{X}) \setminus \{0\}, \quad (4.4)$$

where $C(\bar{X}) := C_{\mathcal{P}}(B\bar{X} - b, \bar{y}) \cap C_{\theta}(\bar{X}, \bar{S})$. We say that the primal SRCQ holds at \bar{X} with respect to (\bar{y}, \bar{S}) , if

$$\begin{pmatrix} B \\ I_{\mathbb{R}^{m \times n}} \end{pmatrix} \mathbb{R}^{m \times n} + \begin{pmatrix} C_{\mathcal{P}}(B\bar{X} - b, \bar{y}) \\ C_{\theta}(\bar{X}, \bar{S}) \end{pmatrix} = \begin{pmatrix} \mathbb{R}^l \\ \mathbb{R}^{m \times n} \end{pmatrix}. \quad (4.5)$$

Similarly, let $(\bar{y}, \bar{w}, \bar{S}) \in \mathbb{R}^l \times \mathbb{R}^d \times \mathbb{R}^{m \times n}$ be an optimal solution of problem (4.1) and $\bar{X} \in \mathcal{M}_D(\bar{y}, \bar{w}, \bar{S})$. We say that the dual SOSOC holds at $(\bar{y}, \bar{w}, \bar{S})$ with respect to \bar{X} , if

$$\langle H_w, (\nabla h^*)'(\bar{w}; H_w) \rangle - \varphi_{(\bar{S}, H_S)}^*(\bar{X}) > 0, \quad \forall (H_y, H_w, H_S) \in C(\bar{y}, \bar{w}, \bar{S}) \setminus \{0\}, \quad (4.6)$$

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

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

The dual SRCQ is said to hold at $(\bar{y}, \bar{w}, \bar{S})$ with respect to \bar{X} if

$$A^* \mathbb{R}^d + B^* C_{\mathcal{P}^0}(\bar{y}, B\bar{X} - b) + C_{\theta^*}(\bar{S}, \bar{X}) = \mathbb{R}^{m \times n}. \quad (4.7)$$

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

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

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

$$G(X, y, S) := \begin{bmatrix} A^* \nabla h(AX) + C + S + B^* y \\ BX - b - \Pi_{\mathcal{P}}(BX - b + y) \\ X - \text{Prox}_{\theta}(X + S) \end{bmatrix}, \quad (X, y, S) \in \mathbb{R}^{m \times n} \times \mathbb{R}^l \times \mathbb{R}^{m \times n}.$$

Denote the KKT mapping for problem (1.1) by

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

The Robinson constraint qualification (RCQ) is said to hold at $\bar{X} \in \mathbb{R}^{m \times n}$ for problem (1.1) if

$$B\mathbb{R}^{m \times n} + \mathcal{T}_{\mathcal{P}}(B\bar{X} - b) = \mathbb{R}^l. \quad (4.8)$$

At this point, we are fully prepared to present the most crucial conclusion within this paper, namely, a series of equivalent characterizations of the ICKKTM for problem (1.1).

Theorem 4.1. *Let $\bar{X} \in \mathbb{R}^{m \times n}$ and $(\bar{y}, \bar{w}, \bar{S}) \in \mathbb{R}^l \times \mathbb{R}^d \times \mathbb{R}^{m \times n}$ be optimal solutions of problem (1.1) and (4.1), respectively. Assume that $(\bar{y}, \bar{S}) \in \mathcal{M}_{\mathcal{P}}(\bar{X})$, $\bar{X} \in \mathcal{M}_D(\bar{y}, \bar{w}, \bar{S})$ and the RCQ (4.8) holds at \bar{X} . Then the subsequent conditions are equivalent:*

- The KKT mapping S_{KKT} is robustly isolated calm at the origin for $(\bar{X}, \bar{y}, \bar{S})$.
- The primal SOSOC (4.4) holds at \bar{X} for (\bar{y}, \bar{S}) and the primal SRCQ (4.5) holds at \bar{X} for (\bar{y}, \bar{S}) .
- The primal SOSOC (4.4) holds at \bar{X} for (\bar{y}, \bar{S}) and the dual SOSOC (4.6) holds at $(\bar{y}, \bar{w}, \bar{S})$ for \bar{X} .
- The dual SRCQ (4.7) holds at $(\bar{y}, \bar{w}, \bar{S})$ for \bar{X} and the dual SOSOC (4.6) 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 primal SRCQ (4.5) holds at \bar{X} for (\bar{y}, \bar{S}) .

Proof. Since the Ky Fan k -norm function $\|\cdot\|_{(k)}$ is C^2 -cone reducible [15, Proposition 4.3], “(a) \iff (b)” can be directly derived from [23, Proposition 3.3]. Besides, by Proposition 3.1 and [31, Theorem 4.1], we can obtain that the primal SRCQ (4.5) and the dual SOSC (4.6) are equivalent. Then we have “(b) \iff (c)”. Similarly, the equivalence of the primal SOSC (4.4) and the dual SRCQ (4.7) can also be derived from Proposition 3.2 and [23, Theorem 5.1]. It follows that “(b) \iff (e)” and “(c) \iff (d)”. In summary, we can conclude that “(a) \iff (b) \iff (c) \iff (d) \iff (e)”. \square

Remark 4.1. Although the conditions $\phi_{(Q,D)}^*(P) = 0$ and $D \in C_{\theta^*}(Q, P)$ in Proposition 3.1 are not always equivalent to $D \in (C_{\theta}(P, Q))^{\circ}$, this does not affect the equivalence between the primal SRCQ (4.5) and the dual SOSC (4.6) in the proof of the above theorem. For further details, one may refer to [31, Theorem 4.1].

5. Conclusions

This paper focuses on establishing a more comprehensive set of equivalent conditions of the robust ICKKT for Ky Fan k -norm regularized convex matrix optimization problems. In other words, we extend the results for nuclear norm regularized convex optimization problems in [23] to Ky Fan k -norm regularized convex optimization problems. Among the newly obtained equivalent conditions, the SRCQ is more easily verified than the SOSC. Therefore, the results presented in this paper not only enrich the stability analysis theory of the Ky Fan k -norm regularized convex optimization problems but also enhance the feasibility of designing related algorithms. The most crucial point is the establishment of Propositions 3.1 and 3.2. Therefore, exploring what kind of θ , that is, seeking the axiomatic properties of θ satisfying the results in Propositions 3.1 and 3.2, is a topic of our future interest. Additionally, integrating these theoretical results into algorithm design to efficiently solve optimization problems remains one of our future research directions.

Author contributions

Ziran Yin: Conceptualization, investigation, writing-review and editing, funding acquisition; Chongyang Liu: Methodology and supervision; Xiaoyu Chen: Original draft preparation; Jihong Zhang: Writing-review and editing; Jinlong Yuan: Validation, writing-review and editing.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant numbers 12101101, 12271307]; the China Postdoctoral Science Foundation [grant number 2022M720631]; the Fundamental Research Funds for the Central Universities; the Shandong Natural Science Foundation, China [grant number ZR2023MA054]; the Nature Science Foundation of Liaoning Province of China [grant number 2024-MS-015].

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. K. C. Toh, L. N. Trefethen, The Chebyshev polynomials of a matrix, *SIAM J. Matrix Anal. A.*, **20** (1998), 400–419. <https://doi.org/10.1137/S0895479896303739>
2. P. Apkarian, D. Noll, Nonsmooth H_∞ synthesis, *IEEE T. Automat. Contr.*, **51** (2006), 71–86. <https://doi.org/10.1109/TAC.2005.860290>
3. V. Bompard, D. Noll, P. Apkarian, Second-order nonsmooth optimization for H_∞ synthesis, *Numer. Math.*, **107** (2007), 433–454. <https://doi.org/10.1007/s00211-007-0095-9>
4. J. L. Yuan, D. Y. Yang, D. Y. Xun, K. L. Teo, C. Z. Wu, A. Li, et al., Sparse optimal control of cyber-physical systems via PQA approach, *Pac. J. Optim.*, 2024.
5. A. Johansson, N. Engsner, C. Strannegård, P. Mostad, Improved spectral norm regularization for neural networks, In: *Modeling Decisions for Artificial Intelligence*, Cham: Springer, 2023, 181–201. https://doi.org/10.1007/978-3-031-33498-6_13
6. R. D. Gao, W. K. Yang, X. Sun, Defying forgetting in continual relation extraction via batch spectral norm regularization, In: *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024, 1–8. <https://doi.org/10.1109/IJCNN60899.2024.10651110>
7. E. J. Candès, B. Recht, Exact matrix completion via convex optimization, *Found. Comput. Math.*, **9** (2009), 717–772. <https://doi.org/10.1007/s10208-009-9045-5>
8. B. Recht, M. Fazel, P. A. Parrilo, Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization, *SIAM Rev.*, **52** (2010), 471–501. <https://doi.org/10.1137/070697835>
9. G. A. Watson, On matrix approximation problems with Ky Fan k norms, *Numer. Algor.*, **5** (1993), 263–272. <https://doi.org/10.1007/BF02210386>
10. S. Boyd, P. Diaconis, P. A. Parrilo, L. Xiao, Fastest mixing Markov chain on graphs with symmetries, *SIAM J. Optim.*, **20** (2009), 792–819. <https://doi.org/10.1137/070689413>
11. Z. K. Yao, F. Y. Xu, G. P. Jiang, J. Y. Yao, Data-driven control of hydraulic manipulators by reinforcement learning, *IEEE-ASME T. Mech.*, **29** (2024), 2673–2684. <https://doi.org/10.1109/TMECH.2023.3336070>
12. X. T. Yang, Q. X. Zhu, H. Wang, Exponential stabilization of stochastic systems via novel event-triggered switching controls, *IEEE T. Automat. Contr.*, **69** (2024), 7948–7955. <https://doi.org/10.1109/TAC.2024.3406668>
13. Q. X. Zhu, Event-triggered sampling problem for exponential stability of stochastic nonlinear delay systems driven by Lévy processes, *IEEE T. Automat. Contr.*, **70** (2025), 1176–1183. <https://doi.org/10.1109/TAC.2024.3448128>
14. C. Ding, Variational analysis of the Ky Fan k -norm, *Set-Valued Var. Anal.*, **25** (2017), 265–296. <https://doi.org/10.1007/s11228-016-0378-3>
15. C. Ding, *An introduction to a class of matrix optimization problems*, Singapore: National University of Singapore, 2012.

16. Y. L. Liu, S. H. Pan, Locally upper Lipschitz of the perturbed KKT system of Ky Fan k-norm matrix conic optimization problems, 2015. <https://doi.org/10.48550/arXiv.1509.00681>
17. D. R. Han, D. F. Sun, L. W. Zhang, Linear rate convergence of the alternating direction method of multipliers for convex composite programming, *Math. Oper. Res.*, **43** (2018), 622–637. <https://doi.org/10.1287/moor.2017.0875>
18. R. T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.*, **1** (1976), 97–116. <https://doi.org/10.1287/moor.1.2.97>
19. Y. L. Zhang, L. W. Zhang, On the upper Lipschitz property of the KKT mapping for nonlinear semidefinite optimization, *Oper. Res. Lett.*, **44** (2016), 474–478. <https://doi.org/10.1016/j.orl.2016.04.012>
20. Y. Zhang, L. W. Zhang, J. Wu, K. Wang, Characterizations of local upper Lipschitz property of perturbed solutions to nonlinear second-order cone programs, *Optimization*, **66** (2017), 1079–1103. <https://doi.org/10.1080/02331934.2017.1325887>
21. C. Ding, D. F. Sun, L. W. Zhang, Characterization of the robust isolated calmness for a class of conic programming problems, *SIAM J. Optim.*, **5** (2017), 67–90. <https://doi.org/10.1137/16M1058753>
22. D. R. Han, D. F. Sun, L. W. Zhang, Linear rate convergence of the alternating direction method of multipliers for convex composite quadratic and semi-definite programming, 2015. <https://doi.org/10.48550/arXiv.1508.02134>
23. Y. Cui, 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. <https://doi.org/10.4208/jcm.1709-m2017-0034>
24. R. T. Rockafellar, R. J. B. Wets, *Variational analysis*, Berlin: Heidelberg, 1998. <https://doi.org/10.1007/978-3-642-02431-3>
25. J. F. Bonnans, A. Shapiro, *Perturbation analysis of optimization problems*, New York: Springer, 2000. <https://doi.org/10.1007/978-1-4612-1394-9>
26. R. T. Rockafellar, *Convex analysis*, Princeton university press, 1997.
27. A. Beck, *First-order methods in optimization*, Society for Industrial and Applied Mathematics, 2017.
28. J. von Neumann, *Some matrix-inequalities and metrization of matrix-space*, 1937.
29. K. Fan, On a theorem of Weyl concerning eigenvalues of linear transformations I, *P. Nat. Acad. Sci. USA*, **35** (1949), 652–655. <https://doi.org/10.1073/pnas.35.11.652>
30. L. W. Zhang, N. Zhang, X. T. Xiao, On the second-order directional derivatives of singular values of matrices and symmetric matrix-valued functions, *Set-Valued Var.*, **21** (2013), 557–586. <https://doi.org/557-586.10.1007/s11228-013-0237-4>
31. Z. R. Yin, X. Y. Chen, J. H. Zhang, The robust isolated calmness of spectral norm regularized convex matrix optimization problems, 2024. <https://doi.org/10.48550/arXiv.2410.16697>

