First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints

Chao Ding[*], Defeng Sun[†] and Jane J. Ye[‡]

Abstract

In this paper we consider a mathematical program with semidefinite cone complementarity constraints (SDCMPCC). Such a problem is a matrix analogue of the mathematical program with (vector) complementarity constraints (MPCC) and includes MPCC as a special case. We first derive explicit formulas for the proximal and limiting normal cone of the graph of the normal cone to the positive semidefinite cone. Using these formulas and classical nonsmooth first order necessary optimality conditions we derive explicit expressions for the strong-, Mordukhovich- and Clarke- (S-, M- and C-)stationary conditions. Moreover we give constraint qualifications under which a local solution of SDCMPCC is a S-, M- and C-stationary point. Moreover we show that applying these results to MPCC produces new and weaker necessary optimality conditions.

Key words: mathematical program with semidefinite cone complementarity constraints, necessary optimality conditions, constraint qualifications, S-stationary conditions, M-stationary conditions, C-stationary conditions.

AMS subject classification: 49K10, 49J52, 90C30, 90C22, 90C33

1 Introduction

Let $\mathcal{S}^n$ be the linear space of all $n \times n$ real symmetric matrices equipped with the usual Frobenius inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. For the given positive integer n, let $\mathcal{S}^n_+$ ($\mathcal{S}^n_-$) be the closed convex cone of all $n \times n$ positive (negative) semidefinite matrices in $\mathcal{S}^n$. Let n_i, $i = 1, \ldots, m$ be given positive integers. The *mathematical program with*

(semidefinite) cone complementarity constraints (MPSCCC or SDCMPCC) is defined as follows

$$
\begin{aligned}
\text{(SDCMPCC)} \quad \min \quad & f(z) \\
\text{s.t.} \quad & h(z) = 0\,, \\
& g(z) \preceq_{\mathcal{Q}} 0\,, \\
& \mathcal{S}_+^{n_i} \ni G_i(z) \perp H_i(z) \in \mathcal{S}_-^{n_i}, \quad i = 1, \ldots, m\,,
\end{aligned} \tag{1}
$$

where Z and $\mathcal{H}$ are two finite dimensional real Euclidean spaces; $f : Z \to \Re$, $h : Z \to \Re^p$, $g : Z \to \mathcal{H}$ and $G_i : Z \to \mathcal{S}^{n_i}$, $H_i : Z \to \mathcal{S}^{n_i}$, $i = 1, \ldots, m$ are continuously differentiable mappings; $\mathcal{Q} \subseteq \mathcal{H}$ is a closed convex symmetric cone with a nonempty interior (such as the nonnegative orthant, the second order cone, or the cone of symmetric and positive semidefinite real matrices); for each $i \in \{1, \ldots, m\}$, "$G_i(z) \perp H_i(z)$" means that the matrices $G_i(z)$ and $H_i(z)$ are perpendicular to each other, i.e., $\langle G_i(z), H_i(z) \rangle = 0$; "$g(z) \preceq_{\mathcal{Q}} 0$" means that $-g(z) \in \mathcal{Q}$. In particular, for a given symmetric matrix $Z \in \mathcal{S}^n$, we use $Z \preceq 0$ and $Z \succeq 0$ to denote $Z \in \mathcal{S}_-^n$ and $Z \in \mathcal{S}_+^n$, respectively.

Our research on SDCMPCC is motivated by a number of important applications in diverse areas. Below we describe some of them.

A rank constrained nearest correlation matrix problem. A matrix is said to be a correlation matrix if it is real symmetric positive semidefinite and its diagonal entries are all ones. Let C be a given matrix in $\mathcal{S}^n$. Let $1 \le r \le n$ be a given integer. The rank constrained nearest correlation matrix problem takes the following form

$$
\begin{aligned}
\min \quad & f_C(X) \\
\text{s.t.} \quad & X_{ii} = 1, \quad i = 1, \ldots, n\,, \\
& X \in \mathcal{S}_+^n\,, \\
& \operatorname{rank}(X) \le r\,,
\end{aligned} \tag{2}
$$

where $f_C : \mathcal{S}^n \to \Re$ is a given cost function that measures the closeness of X to a targeted matrix C. For instance, f_C can be simply chosen as $\frac{1}{2}\|X - C\|^2$ in some applications. Problem (2) has many important applications in quantitative finance and engineering, e.g., [58, 66, 7, 24, 8, 28, 49] and the references therein. We may easily cast (2) in a SDCMPCC form

$$
\begin{aligned}
\min_{X,U} \quad & f_C(X) \\
\text{s.t.} \quad & X_{ii} = 1, \quad i = 1, \ldots, n\,, \\
& \langle I, U \rangle = r, \quad U \in \mathcal{S}_+^n\,, \\
& \mathcal{S}_+^n \ni X \perp (U - I) \in \mathcal{S}_-^n\,.
\end{aligned} \tag{3}
$$

We refer to [23] for details on the equivalence of these two formulations. More SDCMPCC examples concerning the matrix rank minimization problems can be found in [67, 5].

A bilinear matrix inequality (BMI) problem. Bilinear matrix inequalities arise frequently from pooling and blending problems [55], system analysis and robust design [47, 18, 54]. In particular, many problems including robustness analysis [38, 11] and robust process design problems [46, 55, 56] can be stated as the following optimization problem with the BMI constraint

$$
\begin{aligned}
\min \quad & b^T u + d^T v \\
\text{s.t.} \quad & D + \sum_{i=1}^m u_i A^{(i)} + \sum_{j=1}^n v_j B^{(j)} + \sum_{i=1}^m \sum_{j=1}^n u_i v_j C^{(ij)} \preceq 0\,,
\end{aligned} \tag{4}
$$

where $u \in \Re^m$ and $v \in \Re^n$ are decision variables, $b \in \Re^m$ and $d \in \Re^n$ are given, and D, $A^{(i)}$, $B^{(j)}$, and $C^{(ij)}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ are given p by p symmetric matrices. Denote $x := (u, v) \in \Re^{m+n}$, $c := (b, d) \in \Re^{m+n}$. Then, the optimization problem (4) can be rewritten as the following optimization problem [15]

$$
\begin{aligned}
\min \quad & c^T x \\
\text{s.t.} \quad & D + \sum_{i=1}^{m+n} x_i \overline{A}^{(i)} + \sum_{i,j=1}^{m+n} W_{ij} \overline{C}^{(ij)} \preceq 0, \\
& W = xx^T,
\end{aligned} \tag{5}
$$

where $\overline{A}^{(i)} = (A^{(1)}, \ldots, A^{(m)}, B^{(1)}, \ldots, B^{(n)})$ and for each $i, j \in \{1, \ldots, m\} \times \{1, \ldots, n\}$, $\overline{C}^{(ij)} = C^{(ij)}$ if $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ and $\overline{C}^{(ij)} = 0$ otherwise. It is easy to see that the second constraint in the problem (5) can be replaced by the following constraints [15]

$$
Z = \begin{bmatrix} W & x \\ x^T & 1 \end{bmatrix} \succeq 0 \quad \text{and} \quad \operatorname{rank}(Z) \leq 1.
$$

Therefore, similarly as the previous example, we know that the problem (5) can be cast in the following SDCMPCC form

$$
\begin{aligned}
\min \quad & c^T x \\
\text{s.t.} \quad & D + \sum_{i=1}^{m+n} x_i \overline{A}^{(i)} + \sum_{i,j=1}^{m+n} W_{ij} \overline{C}^{(ij)} \preceq 0, \\
& \langle I, U \rangle = 1, \quad U \in \mathcal{S}_+^n, \\
& \mathcal{S}_+^n \ni \begin{bmatrix} W & x \\ x^T & 1 \end{bmatrix} \perp (U - I) \in \mathcal{S}_-^n.
\end{aligned}
$$

A single-firm model in electric power market with uncertain data. The electric power market is an oligopolistic market, which means that there are several dominant firms in this market. Each dominant firm has some number of generators, which submit the hourly bids to an independent system operator (ISO). The firm can be thought of as a leader of a Stackelberg game, which calculates its bids based on what it anticipates the followers would do, which is the ISO in this case. Without the uncertain data, it is well-known that this single-firm problem in the electric power market can be modeled as a bilevel programming problem [21]. In this bilevel programming model, the upper-level problem is the single firm's profit maximization problem and the lower-level problem is the ISO's single spatial price equilibrium problem. In practice it is more realistic to assume that the lower-level problem involves uncertainty. For instance, the coefficients of the marginal demand functions, which are decided by the information of consumers, usually contain uncertainty. Therefore, it makes sense to consider a robust bilevel programming problem where for a fixed upper-level decision variable x, the lower-level problem is replaced by its robust counterpart:

$$
P_x: \quad \min_y \{ f(x, y, \zeta) : g(x, y, \zeta) \leq 0 \quad \forall \zeta \in \mathcal{U} \},
$$

where $\mathcal{U}$ is some "uncertainty set" in the space of the data. It is well-known (see [3]) that if the uncertainty set $\mathcal{U}$ is given by a system of linear matrix inequalities, then the deterministic counterpart of the problem P_x is a semidefinite program. If this semidefinite programming problem can be equivalently replaced by its Karush-Kuhn-Tucker (KKT) condition, then it yields a SDCMPCC problem.

SDCMPCC is a broad framework, which includes the mathematical program with (vector) complementarity constraints (MPCC) as a special case. In fact, if $\mathcal{Q} \equiv \Re_+^q$, the nonnegative orthant in $\mathcal{H} \equiv \Re^q$ and $n_i \equiv 1$, $i = 1, \ldots, m$, the SDCMPCC becomes the following MPCC problem

$$
\begin{aligned}
\text{(MPCC)} \quad \min \quad & f(z) \\
\text{s.t.} \quad & h(z) = 0 \,, \\
& g(z) \leq 0 \,, \\
& \Re_+ \ni G_i(z) \perp H_i(z) \in \Re_-, \quad i = 1, \ldots, m \,. \quad (6)
\end{aligned}
$$

Denote $G(z) = (G_1(z), \ldots, G_m(z))^T : Z \to \Re^m$ and $H(z) = (H_1(z), \ldots, H_m(z))^T : Z \to \Re^m$. Then the constraints (6) can be replaced by the following standard vector complementarity constraint

$$
\Re_+^m \ni G(z) \perp H(z) \in \Re_-^m \,.
$$

MPCC is a class of very important problems since they arise frequently in applications where the constraints come from equilibrium systems and hence is also known as the mathematical program with equilibrium constraints (MPEC); see [26, 34] for references. One of the main sources of MPCCs comes from bilevel programming problems which have numerous applications; see [12].

In this paper, we study first order necessary optimality conditions for SDCMPCC. For simplicity, we consider the SDCMPCC problem which has only one semidefinite cone complementarity constraint. However all results can be generalized to the case of more than one semidefinite cone complementarity constraints in a straightforward manner.

MPCC is notoriously known as a difficult class of optimization problems since if one treats a MPCC as a standard nonlinear programming problem, then Mangasarian Fromovitz constraint qualification (MFCQ) fails to hold at each feasible point of the feasible region; see [65, Proposition 1.1]. One of the implications of the failure of MFCQ is that the classical KKT condition may not hold at a local optimizer. The classical KKT condition for MPCC is known to be equivalent to the strong stationary condition (S-stationary condition). Consequently weaker stationary conditions such as the Mordukhovich stationary condition (M-stationary condition) and the Clarke stationary condition (C-stationary condition) have been proposed and the constraint qualifications under which a local minimizer is a M-(C-)stationary point have been studied; see e.g., [48, 63] for a detailed discussion.

The same difficulties exist for SDCMPCC. The cone complementarity constraint (1) amounts to the following convex cone constraints:

$$
\langle G(z), H(z) \rangle = 0, \quad G(z) \in \mathcal{S}_+^n, \quad H(z) \in \mathcal{S}_-^n.
$$

For an optimization problem with convex cone constraints, the usual constraint qualification is Robinson's CQ. In this paper we show that if we consider SDCMPCC as an optimization problem with cone constraints, Robinson's CQ fails to hold at each feasible point of the SDCMPCC. Hence SDCMPCC is also a difficult class of optimization problems. One of the implications of the failure of Robinson's CQ is that the classical KKT condition may not hold at a local optimizer. It is obvious that the complementarity constraint (1) can be reformulated as a nonconvex cone constraint:

$$
(G(z), H(z)) \in \operatorname{gph} N_{\mathcal{S}_+^n},
$$

where gph $N_{\mathcal{S}_+^n}$ is the graph of the normal cone to the positive semidefinite cone. We first derive the exact expressions for the proximal and limiting normal cone of gph $N_{\mathcal{S}_+^n}$. As in the vector case, the first order necessary optimality condition based on the proximal and limiting normal cones are called S- and M-stationary condition respectively. To derive the C-stationary condition, we reformulate the complementarity constraint (1) as a nonsmooth equation constraint:

$$G(z) - \Pi_{\mathcal{S}_+^n}(G(z) + H(z)) = 0,$$

where $\Pi_{\mathcal{S}_+^n}$ denotes the metric projection to the positive semidefinite cone. As in the vector case, based on this reformulation and the classical nonsmooth necessary optimality condition we derive the necessary optimality condition in terms of the C-stationary condition. We also show that the classical KKT condition implies the S-stationary condition but not vice versa.

To the best of our knowledge, this is the first time explicit expressions for S-, M- and C-stationary conditions for SDCMPCC are given. In [60], a smoothing algorithm is given for mathematical program with symmetric cone complementarity constraints and the convergence to C-stationary points is shown. Although the problem studied in [60] may include our problem as a special case, there is no explicit expression for C-stationary condition given. It is also the first time precise formulas for the proximal and limiting normal cone of gph $N_{\mathcal{S}_+^n}$ are developed. In particular the precise expression for the limiting normal cone of gph $N_{\mathcal{S}_+^n}$ is not only important for deriving the M-stationary condition but also useful in the so-called Mordukhovich criterion for characterizing the Aubin continuity [45, Theorem 9.40] of a perturbed generalized equation such as:

$$S(x) := \{z : x \in H(z) + N_{\mathcal{S}_+^n}(z)\}.$$

We organize our paper as following. In §2 we introduce the preliminaries and preliminary results on the background in variational analysis, first order conditions for a general problem and background in variational analysis in matrix spaces. In §3, we give the precise expressions for the proximal and limiting normal cones of the graph of the normal cone $N_{\mathcal{S}_+^n}$. In §4, we show that if SDCMPCC is considered as an optimization problem with convex cone constraints then Robinson's CQ fails at every feasible solution of SDCMPCC and derive the classical KKT condition under the Clarke calmness condition. Explicit expressions for S-stationary conditions are given in §5 where it is also shown that the classical KKT condition implies the S-stationary condition. Explicit expressions for M- and C-stationary conditions are given in §6 and §7 respectively. In §8 we reformulate MPCC as a particular case of SDCMPCC by taking the vector complementarity functions as matrices with diagonal values. Comparisons between the S-, M- and C-stationary points are made. We show that the S-stationary condition for the two formulations are equivalent while the M- and C-stationary conditions for SDCMPCC may be weaker.

2 Preliminaries and preliminary Results

We first give the following notation that will be used throughout the paper. Let X and Y be finite dimensional spaces. We denote by $\|\cdot\|$ the Euclidean norm in X. We denote by $B(x,\delta) := \{y \in X \mid \|y - x\| < \delta\}$ the open ball centered at x with radius $\delta > 0$ and B the open unit ball centered at 0. Given a set $S \subseteq X$ and a point $x \in X$, the distance from x to S is denoted by

$$\text{dist}(x, S) := \inf\{\|y - x\| \mid y \in S\}.$$

Given a linear operator $\mathcal{A} : X \to Y$, $\mathcal{A}^*$ denotes the adjoint of the linear operator $\mathcal{A}$. Given a matrix A, we denote by A^T the transpose of the matrix A. For a mapping $F : X \to Y$ and $x \in X$, $F'(x)$ stands for the classical derivative or the Jacobian of F at x and $\nabla F(x)$ the adjoint of the Jacobian. We denote by $F'(x; d)$ the directional derivative of F at x in direction d. For a set-valued mapping $\Phi : X \rightrightarrows Y$, we denote by gph$\Phi$ the graph of Φ, i.e., gph $\Phi := \{(z, v) \in X \times Y \mid v \in \Phi(z)\}$. For a set $\mathcal{C}$, we denote by int$\mathcal{C}$, cl$\mathcal{C}$, co$\mathcal{C}$ its interior, closure and convex hull respectively. For a function $g : X \to \Re$, we denote $g^+(x) := \max\{0, g(x)\}$ and if it is vector-valued then the maximum is taken componentwise.

- Let $\mathcal{O}^n$ be the set of all $n \times n$ orthogonal matrices.

- For any $Z \in \Re^{m \times n}$, we denote by Z_{ij} the (i, j)-th entry of Z.

- For any $Z \in \Re^{m \times n}$ and a given index set $\mathcal{J} \subseteq \{1, \ldots, n\}$, we use $Z_{\mathcal{J}}$ to denote the sub-matrix of Z obtained by removing all the columns of Z not in $\mathcal{J}$. In particular, we use Z_j to represent the j-th column of Z, $j = 1, \ldots, n$.

- Let $\mathcal{I} \subseteq \{1, \ldots, m\}$ and $\mathcal{J} \subseteq \{1, \ldots, n\}$ be two index sets. For any $Z \in \Re^{m \times n}$, we use $Z_{\mathcal{I}\mathcal{J}}$ to denote the $|\mathcal{I}| \times |\mathcal{J}|$ sub-matrix of Z obtained by removing all the rows of Z not in $\mathcal{I}$ and all the columns of Z not in $\mathcal{J}$.

- We use " $\circ$ " to denote the Hardamard product between matrices, i.e., for any two matrices A and B in $\Re^{m \times n}$ the (i, j)-th entry of $Z := A \circ B \in \Re^{m \times n}$ is $Z_{ij} = A_{ij}B_{ij}$.

- Let $diag(\cdot) : \Re^m \to \mathcal{S}^m$ be a linear mapping defined by for any $x \in \Re^n$, $diag(x)$ denotes the diagonal matrix whose i-th diagonal entry is x_i, $i = 1, \ldots, n$.

2.1 Background in variational analysis

In this subsection we summarize some background materials on variational analysis which will be used throughout the paper. Detailed discussions on these subjects can be found in [9, 10, 31, 32, 45]. In this subsection X is a finite dimensional space.

Definition 2.1 *(see e.g., [10, Proposition 1.5(a)] or [45, page 213]) Let Ω be a nonempty subset of X. Given $\bar{x} \in \mathrm{cl}\,\Omega$, the following convex cone*

$$N_\Omega^\pi(\bar{x}) := \{\zeta \in X : \exists\, M > 0, \ such \ that \ \langle \zeta, x - \bar{x} \rangle \le M\|x - \bar{x}\|^2 \quad \forall\, x \in \Omega\} \qquad (7)$$

is called the proximal normal cone to set Ω at point $\bar{x}$.

Definition 2.2 *(see e.g., [10, page 62 and Theorem 6.1(b)]) Let Ω be a nonempty subset of X. Given $\bar{x} \in \mathrm{cl}\,\Omega$, the following closed cone*

$$N_\Omega(\bar{x}) := \{\lim_{i \to \infty} \zeta_i : \zeta_i \in N_\Omega^\pi(x_i), \quad x_i \to \bar{x}, \quad x_i \in \Omega\} \qquad (8)$$

is called the limiting normal cone (also known as Mordukhovich normal cone or basic normal cone) to set Ω at point $\bar{x}$ and the closed convex hull of the limiting normal cone

$$N_\Omega^c(\bar{x}) := \mathrm{clco}\, N_\Omega(\bar{x}).$$

is the Clarke normal cone ([9]) to set Ω at point $\bar{x}$.

Alternatively in a finite dimensional space, the limiting normal cone can be also defined by the Fréchet (also called regular) normal cone instead of the proximal normal cone, see [31, Definition 1.1 (ii)]. In the case when Ω is convex, the proximal normal cone, the limiting normal cone and the Clarke normal cone coincide with the normal cone in the sense of the convex analysis [44], i.e., $N_\Omega(\bar{x}) := \{\zeta \in X : \langle \zeta, x - \bar{x} \rangle \leq 0 \quad \forall\, x \in \Omega\}$.

Definition 2.3 *Let $f : X \to \Re \cup \{+\infty\}$ be a lower semicontinuous function and finite at $\bar{x} \in X$. The proximal subdifferential ([45, Definition 8.45]) of f at $\bar{x}$ is defined as*

$$\partial^\pi f(\bar{x}) \quad := \quad \{\zeta \in X : \exists\, \sigma > 0, \delta > 0 \text{ such that } f(x) \geq f(\bar{x}) + \langle \zeta, x - \bar{x} \rangle - \sigma \|x - \bar{x}\|^2$$
$$\forall\, x \in B(\bar{x}, \delta)\}$$

and the limiting (Mordukhovich or basic [31]) subdifferential of f at $\bar{x}$ is defined as

$$\partial f(\bar{x}) := \{\lim_{k \to \infty} \zeta_k : \zeta_k \in \partial^\pi f(x_k),\ x_k \to \bar{x},\ f(x_k) \to f(\bar{x})\}.$$

When f is Lipschitz continuous near $\bar{x}$,

$$\partial^c f(\bar{x}) := \mathrm{co}\, \partial f(\bar{x})$$

is the Clarke subdifferential [9] of f at $\bar{x}$.

Note that in a finite dimensional space, alternatively the limiting subgradient can be also constructed via Fréchet subgradients (also known as regular subgradients), see [31, Theorem 1.89]. The equivalence of the two definitions is well-known, see the commentary by Rockafellar and Wets [45, page 345]. In the case when f is convex and locally Lipschitz, the proximal subdifferential, the limiting subdifferential and the Clarke subdifferential coincide with the subdifferential in the sense of convex analysis [44]. In the case when f is strictly differentiable, the limiting subdifferenial and the Clarke subdifferential reduce to the classical gradient of f at $\bar{x}$, i.e., $\partial^c f(\bar{x}) = \partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

2.2 First order optimality conditions for a general problem

In this subsection we discuss constraint qualifications and first order necessary optimality conditions for the following general optimization problem:

$$\begin{aligned}
(GP) \qquad \min \quad & f(z) \\
\text{s.t.} \quad & h(z) = 0\,, \\
& g(z) \leq 0\,, \\
& G(z) \in K\,,
\end{aligned}$$

where Y, Z are finite dimensional spaces, K is a closed subset of Y, $f : Z \to \Re$, $h : Z \to \Re^p$, $g : Z \to \Re^q$ and $G : Z \to Y$ are locally Lipschitz mappings.

We denote the set of feasible solutions for (GP) by $\mathcal{F}$ and the perturbed feasible region by

$$\mathcal{F}(r, s, P) := \{z \in Z : h(z) + r = 0, \quad g(z) + s \leq 0, \quad G(z) + P \in K\}. \tag{9}$$

Then $\mathcal{F}(0, 0, 0) = \mathcal{F}$. The following definition is the Clarke calmness [9] adapted to our setting.

Definition 2.4 (Clarke calmness) *We say that problem (GP) is (Clarke) calm at a local optimal solution $\bar{z}$ if there exist positive ε and μ such that, for all (r, s, P) in εB, for all $z \in (\bar{z} + \varepsilon B) \cap \mathcal{F}(r, s, P)$, one has*

$$f(z) - f(\bar{z}) + \mu\|(r, s, P)\| \geq 0\,.$$

The following equivalence is obvious.

Proposition 2.1 *Problem (GP) is Clarke calm at a local optimal solution $\bar{z}$ if and only if $(\bar{z}, G(\bar{z}))$ is a local optimal solution to the penalized problem for some $\mu > 0$:*

$$(GP)_\mu \quad \min_{z,X} \quad f(z) + \mu(\|h(z)\| + \|\max\{g(z), 0\}\| + \|G(z) - X\|)$$

$$\text{s.t.} \quad X \in K\,.$$

Theorem 2.1 *Let $\bar{z}$ be a local optimal solution of (GP). Suppose that (GP) is Clarke calm at $\bar{z}$. Then there exist $\lambda^h \in \Re^p$, $\lambda^g \in \Re^q$ and $\Omega^G \in \mathcal{S}^n$ such that*

$$0 \in \partial f(\bar{z}) + \partial \langle h, \lambda^h \rangle(\bar{z}) + \partial \langle g, \lambda^g \rangle(\bar{z}) + \partial \langle G, \Omega^G \rangle(\bar{z})\,,$$

$$\lambda^g \geq 0, \quad \langle g(\bar{z}), \lambda^g \rangle = 0 \quad \Omega^G \in N_K(G(\bar{z}))\,.$$

Proof. The results follow from applying the limiting subdifferential version of the generalized Lagarange multiplier rule (see e.g., Mordukhovich [32, Proposition 5.3]), calculus rules for limiting subdifferentials in particular the chain rule in Mordukhovich and Shao [33, Proposition 2.5 and Corollary 6.3]). ∎

The calmness condition involves both the constraint functions and the objective function. It is therefore not a constraint qualification in classical sense. Indeed it is a sufficient condition under which KKT type necessary optimality conditions hold. The calmness condition may hold even when the weakest constraint qualification does not hold. In practice one often uses some verifiable constraint qualifications sufficient to the calmness condition.

Definition 2.5 *(Calmness of a set-valued map) A set-valued map $\Phi : Z \rightrightarrows Y$ is said to be calm at a point $(\bar{z}, \bar{v}) \in \text{gph}\,\Phi$ if there exist a constant $M > 0$ and a neighborhood U of $\bar{z}$, a neighborhood V of $\bar{v}$ such that*

$$\Phi(z) \cap V \subseteq \Phi(\bar{z}) + M\|z - \bar{z}\|clB \quad \forall z \in U.$$

Although the term "calmness" was coined in Rockafellar and Wets [45], the concept of calmness of a set-valued map was first introduced by Ye and Ye in [64] under the term "pseudo upper-Lipschitz continuity" which comes from the fact that it is a combination of Aubin's pseudo Lipschitz continuity [1] and Robinson's upper-Lipschitz continuity [40, 41]. For recent discussion on the properties and the criterion of calmness of a set-valued mapping, see Henrion and Outrata ([19, 20]). In what follows, we consider the calmness of the perturbed feasible region $\mathcal{F}(r, s, P)$ at $(r, s, P) = (0, 0, 0)$ to establish the Clarke calmness of the problem.

The proposition below is an easy consequence of Clarke's exact penalty principle [9, Proposition 2.4.3] and the calmness of the perturbed feasible region of the problem. See [62, Proposition 4.2] for a proof.

Proposition 2.2 *If the objective function of (GP) is Lipschitz near $\bar{z} \in \mathcal{F}$ and the perturbed feasible region of the constraint system $\mathcal{F}(r, s, P)$ defined as in (9) is calm at $(0, 0, 0, \bar{z})$, then the problem (GP) is Clarke calm at $\bar{z}$.*

From the definition it is easy to verify that the set-valued mapping $\mathcal{F}(r, s, P)$ is calm at $(0, 0, 0, \bar{z})$ if and only if there exist a constant $M > 0$ and U, a neighborhood of $\bar{z}$, such that

$$\text{dist}\,(z, \mathcal{F}) \le M\|(r, s, P)\| \quad \forall\, z \in U \cap \mathcal{F}(r, s, P)\,.$$

The above property is also referred to the existence of a local error bound for the feasible region $\mathcal{F}$. Hence any results on the existence of a local error bound of the constraint system may be used as a sufficient condition for calmness of the perturbed feasible region (see e.g., Wu and Ye [59] for such sufficient conditions).

By virtue of Proposition 2.2, the following four constraint qualifications are stronger than the Clarke calmness of (GP) at a local minimizer when the objective function of the problem (GP) is Lipschitz continuous.

Proposition 2.3 *Let $\mathcal{F}(r, s, P)$ be defined as in (9) and $\bar{z} \in \mathcal{F}$. Then the set-valued map $\mathcal{F}(r, s, P)$ is calm at $(0, 0, 0, \bar{z})$ under one of the following constraint qualifications:*

(i) *There is no singular Lagrange multiplier for problem (GP) at $\bar{z}$:*

$$\begin{cases} 0 \in \partial\langle h, \lambda^h\rangle(\bar{z}) + \partial\langle g, \lambda^g\rangle(\bar{z}) + \partial\langle G, \Omega^G\rangle(\bar{z}), \\ \Omega^G \in N_K(G(\bar{z})), \quad \lambda^g \ge 0, \ \langle g(\bar{z}), \lambda^g\rangle = 0 \end{cases} \implies (\lambda^h, \lambda^g, \Omega^G) = 0.$$

(ii) *Robinson's CQ ([42]) holds at $\bar{z}$: h, g and G are continuously differentiable at $\bar{z}$. K is a closed convex cone with a nonempty interior. The gradients $h_i'(\bar{z})^* (i = 1, \ldots, p)$ are linearly independent and there exists a vector $d \in Z$ such that*

$$h_i(\bar{z})'d = 0, \quad i = 1, \ldots, p\,,$$
$$g_i(\bar{z})'d < 0, \quad i \in I_g(\bar{z})\,,$$
$$G(\bar{z}) + G'(\bar{z})d \in \text{int}\, K\,,$$

where $I_g(\bar{z}) := \{i : g_i(\bar{z}) = 0\}$ is the index of active inequality constraints.

(iii) *Linear Independence Constraint Qualification (LICQ) holds at $\bar{z}$:*

$$0 \in \partial\langle h, \lambda^h\rangle(\bar{z}) + \partial\langle g, \lambda^g\rangle(\bar{z}) + \partial\langle G, \Omega^G\rangle(\bar{z}), \Omega^G \in N_K(G(\bar{z})) \implies (\lambda^h, \lambda^g, \Omega^G) = 0.$$

(iv) *h, g and G are affine mappings and the set K is a union of finitely many polyhedral convex sets.*

Proof. It is obvious that (iii) implies (i). By [6, Propositions 3.16 (ii) and 3.19 (iii)], Robinson's CQ (ii) is equivalent to (i) when all functions h, g, G are continuously differentiable and K is a closed convex cone with a nonempty interior. By Mordukhovich's criteria for pseudo-Lipschitz continuity, (i) implies that the set-valued map $\mathcal{F}(r, s, P)$ is pseudo-Lipschitz continuous around $(r, s, P) = (0, 0, 0)$ (see e.g., [33, Theorem 6.1]) and hence calm. By Robinson [43], (iv) implies the upper-Lipschitz continuity and hence the calmness of the set-valued map $\mathcal{F}(r, s, P)$ at $(0, 0, 0, \bar{z})$. ∎

Combining Theorem 2.1 and Propositions 2.2 and 2.3, we have the following.

Theorem 2.2 *Let $\bar{z}$ be a local optimal solution of (GP). Suppose the problem is Clarke calm at $\bar{z}$; in particular one of the constraint qualifications in Proposition 2.3 holds. Then the KKT condition in Theorem 2.1 holds at $\bar{z}$.*

2.3 Background in variational analysis in matrix spaces

Let $A \in \mathcal{S}^n$ be given. We use $\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A)$ to denote the eigenvalues of A (all real and counting multiplicity) arranging in nonincreasing order and use $\lambda(A)$ to denote the vector of the ordered eigenvalues of A. Denote $\Lambda(A) := \mathrm{diag}(\lambda(A))$. Consider the eigenvalue decomposition of A, i.e., $A = \overline{P}\Lambda(A)\overline{P}^T$, where $\overline{P} \in \mathcal{O}^n$ is a corresponding orthogonal matrix of the orthonormal eigenvectors. By considering the index sets of positive, zero, and negative eigenvalues of A, we are able to write A in the following form

$$A = \begin{bmatrix} \overline{P}_\alpha & \overline{P}_\beta & \overline{P}_\gamma \end{bmatrix} \begin{bmatrix} \Lambda(A)_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda(A)_{\gamma\gamma} \end{bmatrix} \begin{bmatrix} \overline{P}_\alpha^T \\ \overline{P}_\beta^T \\ \overline{P}_\gamma^T \end{bmatrix}. \tag{10}$$

where $\alpha := \{i : \lambda_i(A) > 0\}$, $\beta := \{i : \lambda_i(A) = 0\}$ and $\gamma := \{i : \lambda_i(A) < 0\}$.

Proposition 2.4 *(see e.g., [16, Theorem 2.1]) For any $X \in \mathcal{S}^n_+$ and $Y \in \mathcal{S}^n_-$,*

$$
\begin{aligned}
N_{\mathcal{S}^n_+}(X) &= \{X^* \in \mathcal{S}^n_- : \langle X, X^* \rangle = 0\} = \{X^* \in \mathcal{S}^n_- : XX^* = 0\}, \\
N_{\mathcal{S}^n_-}(Y) &= \{Y^* \in \mathcal{S}^n_+ : \langle Y, Y^* \rangle = 0\} = \{Y^* \in \mathcal{S}^n_+ : YY^* = 0\}.
\end{aligned}
$$

We say that $X, Y \in \mathcal{S}^n$ have a simultaneous ordered eigenvalue decomposition provided that there exists $P \in \mathcal{O}^n$ such that $X = P\Lambda(X)P^T$ and $Y = P\Lambda(Y)P^T$. The following theorem is well-known and can be found in e.g., [22].

Theorem 2.3 *[von Neumann-Theobald] Any matrices X and Y in $\mathcal{S}^n$ satisfy the inequality*

$$\langle X, Y \rangle \leq \lambda(X)^\top \lambda(Y);$$

the equality holds if and only if X and Y admit a simultaneous ordered eigenvalue decomposition.

Proposition 2.5 *The graph of the set-valued map $N_{\mathcal{S}^n_+}$ can be written as*

$$
\begin{aligned}
\mathrm{gph}\, N_{\mathcal{S}^n_+} &= \{(X,Y) \in \mathcal{S}^n_+ \times \mathcal{S}^n_- : \Pi_{\mathcal{S}^n_+}(X+Y) = X\} \tag{11} \\
&= \{(X,Y) \in \mathcal{S}^n_+ \times \mathcal{S}^n_- : \Pi_{\mathcal{S}^n_-}(X+Y) = Y\} \tag{12} \\
&= \{(X,Y) \in \mathcal{S}^n_+ \times \mathcal{S}^n_- : XY = YX = 0, \langle X, Y \rangle = 0\}. \tag{13}
\end{aligned}
$$

Proof. Equations (11) and (12) are well-known (see [13]). Let $X \in \mathcal{S}^n_+$. Since $N_{\mathcal{S}^n_+}(X) = \partial\delta_{\mathcal{S}^n_+}(X)$, where δ_C is the indicate function of a set C, by [22, Theorem 3], since the function $\delta_{\mathcal{S}^n_+}(X)$ is an eigenvalue function, for any $Y \in N_{\mathcal{S}^n_+}(X)$, X and Y commute. Equation (13) then follows from the expression for the normal cone in Proposition 2.4. ∎

From [51, Theorem 4.7] we know that the metric projection operator $\Pi_{\mathcal{S}^n_+}(\cdot)$ is directionally differentiable at any $A \in \mathcal{S}^n$ and the directional derivative of $\Pi_{\mathcal{S}^n_+}(\cdot)$ at A along direction $H \in \mathcal{S}^n$ is given by

$$\Pi'_{\mathcal{S}^n_+}(A; H) = \overline{P} \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & \Sigma_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}_{\alpha\beta}^T & \Pi_{\mathcal{S}^{|\beta|}_+}(\widetilde{H}_{\beta\beta}) & 0 \\ \Sigma_{\alpha\gamma}^T \circ \widetilde{H}_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} \overline{P}^T, \tag{14}$$

10

where $\widetilde{H} := \overline{P}^T H \overline{P}$, $\circ$ is the Hadamard product and

$$\Sigma_{ij} := \frac{\max\{\lambda_i(A), 0\} - \max\{\lambda_j(A), 0\}}{\lambda_i(A) - \lambda_j(A)}, \quad i, j = 1, \ldots, n, \tag{15}$$

where $0/0$ is defined to be 1. Since $\Pi_{\mathcal{S}_+^n}(\cdot)$ is global Lipschitz continuous on $\mathcal{S}^n$, it is well-known that $\Pi_{\mathcal{S}_+^n}(\cdot)$ is *B(ouligand)-differentiable* (c.f. [14, Definition 3.1.2]) on $\mathcal{S}^n$. In the following proposition, we will show that $\Pi_{\mathcal{S}_+^n}(\cdot)$ is also *calmly B(ouligand)-differentiable* on $\mathcal{S}^n$. This result is not only of its own interest, but also is crucial for the study of the proximal and limiting normal cone of the normal cone mapping $N_{\mathcal{S}_+^n}$ in the next section.

Proposition 2.6 *The metric projection operator $\Pi_{\mathcal{S}_+^n}(\cdot)$ is calmly B-differentiable for any given $A \in \mathcal{S}^n$, i.e., for $\mathcal{S}^n \ni H \to 0$,*

$$\Pi_{\mathcal{S}_+^n}(A + H) - \Pi_{\mathcal{S}_+^n}(A) - \Pi'_{\mathcal{S}_+^n}(A; H) = O(\|H\|^2). \tag{16}$$

Proof. See the Appendix. ∎

3 Expression of the proximal and limiting normal cones

In order to characterize the S-stationary and M-stationary conditions, we need to give the precise expressions for the proximal and limiting normal cones of the graph of the normal cone mapping $N_{S_+^n}$ at any given point $(X, Y) \in \text{gph } N_{\mathcal{S}_+^n}$. The purpose of this section is to provide such formulas. The result is also of independent interest.

3.1 Expression of the proximal normal cone

By using the directional derivative formula (14), Qi and Fusek [39] characterized the Fréchet normal cone of $\text{gph } N_{S_+^n}$. In this subsection, we will establish the representation of the desired proximal normal cone by using the same formula and the just proved calmly B-differentiability of the metric projection operator. The proximal normal cone is in general smaller than the Fréchet normal cone. For the set $\text{gph } N_{\Re_+^n}$, however, it is well-known that the Fréchet normal cone coincides with the proximal normal cone. The natural question to ask is that whether this statement remains true for the set $\text{gph } N_{S_+^n}$. Our computations in this section give an affirmative answer, that is, the expression for the proximal normal cone coincides with the one for the Fréchet normal cone derived by Qi and Fusek in [39].

From Proposition 2.6, we know that for any given $X^* \in \mathcal{S}^n$ and any fixed $X \in \mathcal{S}^n$ there exist $M_1, M_2 > 0$ (depending on X and X^* only) such that for any $X' \in \mathcal{S}^n$ sufficiently close to X,

$$\langle X^*, \Pi_{\mathcal{S}_+^n}(X') - \Pi_{\mathcal{S}_+^n}(X) \rangle \leq \langle X^*, \Pi'_{\mathcal{S}_+^n}(X; X' - X) \rangle + M_1 \|X' - X\|^2, \tag{17}$$

$$\langle X^*, \Pi_{\mathcal{S}_-^n}(X') - \Pi_{\mathcal{S}_-^n}(X) \rangle \leq \langle X^*, \Pi'_{\mathcal{S}_-^n}(X; X' - X) \rangle + M_2 \|X' - X\|^2. \tag{18}$$

Proposition 3.1 *For any given $(X, Y) \in \text{gph } N_{\mathcal{S}_+^n}$, $(X^*, Y^*) \in N^\pi_{\text{gph } N_{\mathcal{S}_+^n}}(X, Y)$ if and only if $(X^*, Y^*) \in \mathcal{S}^n \times \mathcal{S}^n$ satisfies*

$$\langle X^*, \Pi'_{\mathcal{S}_+^n}(X + Y; H) \rangle + \langle Y^*, \Pi'_{\mathcal{S}_-^n}(X + Y; H) \rangle \leq 0 \quad \forall H \in \mathcal{S}^n. \tag{19}$$

Proof. " $\Longleftarrow$ " Suppose that $(X^*, Y^*) \in \mathcal{S}^n \times \mathcal{S}^n$ is given and satisfies the condition (19). By Proposition 2.5, (17) and (18), we know that there exist a constant $\delta > 0$ and a constant $\widetilde{M} > 0$ such that for any $(X', Y') \in \text{gph } N_{\mathcal{S}^n_+}$ and $\|(X', Y') - (X, Y)\| \leq \delta$,

$$
\begin{aligned}
& \langle (X^*, Y^*), (X', Y') - (X, Y) \rangle \\
= \ & \langle (X^*, Y^*), (\Pi_{\mathcal{S}^n_+}(X' + Y'), \Pi_{\mathcal{S}^n_-}(X' + Y')) - (\Pi_{\mathcal{S}^n_+}(X + Y), \Pi_{\mathcal{S}^n_-}(X + Y)) \rangle \\
\leq \ & \widetilde{M} \|(X', Y') - (X, Y)\|^2 .
\end{aligned}
$$

By taking $M = \max \left\{ \widetilde{M}, \|(X^*, Y^*)\| / \delta \right\}$, we know that for any $(X', Y') \in \text{gph } N_{\mathcal{S}^n_+}$,

$$
\langle (X^*, Y^*), (X', Y') - (X, Y) \rangle \leq M \|(X', Y') - (X, Y)\|^2 ,
$$

which implies, by the definition of the proximal normal cone, that $(X^*, Y^*) \in N^\pi_{\text{gph } N_{\mathcal{S}^n_+}}(X, Y)$.

" $\Longrightarrow$ " Let $(X^*, Y^*) \in N^\pi_{\text{gph } N_{\mathcal{S}^n_+}}(X, Y)$ be given. Then there exists $M > 0$ such that for any $(X', Y') \in \text{gph } N_{\mathcal{S}^n_+}$,

$$
\langle (X^*, Y^*), (X', Y') - (X, Y) \rangle \leq M \|(X', Y') - (X, Y)\|^2 . \tag{20}
$$

Let $H \in \mathcal{S}^n$ be arbitrary but fixed. For any $t \downarrow 0$, let

$$
X'_t = \Pi_{\mathcal{S}^n_+}(X + Y + tH) \quad \text{and} \quad Y'_t = \Pi_{\mathcal{S}^n_-}(X + Y + tH) .
$$

By noting that $(X'_t, Y'_t) \in \text{gph } N_{\mathcal{S}^n_+}$ (c.f., (11)-(12) in Proposition 2.5) and $\Pi_{\mathcal{S}^n_+}(\cdot)$ and $\Pi_{\mathcal{S}^n_-}(\cdot)$ are globally Lipschitz continuous with modulus 1, we obtain from (20) that

$$
\begin{aligned}
& \langle X^*, \Pi'_{\mathcal{S}^n_+}(X + Y; H) \rangle + \langle Y^*, \Pi'_{\mathcal{S}^n_-}(X + Y; H) \rangle \\
\leq \ & M \lim_{t \downarrow 0} \frac{1}{t} (\|X'_t - X\|^2 + \|Y'_t - Y\|^2) \leq M \lim_{t \downarrow 0} \frac{1}{t} (2t^2 \|H\|^2) = 0 .
\end{aligned}
$$

Therefore, we know that $(X^*, Y^*) \in \mathcal{S}^n \times \mathcal{S}^n$ satisfies the condition (19). The proof is completed. $\blacksquare$

For any given $(X, Y) \in \text{gph } N_{\mathcal{S}^n_+}$, let $A = X + Y$ have the eigenvalue decomposition (10). From (11)-(12), we know that $X = \Pi_{\mathcal{S}^n_+}(A)$ and $Y = \Pi_{\mathcal{S}^n_-}(A)$. It follows from the directional derivative formula (14) that for any $H \in \mathcal{S}^n$,

$$
\Pi'_{\mathcal{S}^n_-}(A; H) = \overline{P} \begin{bmatrix} 0 & 0 & (E_{\alpha\gamma} - \Sigma_{\alpha\gamma}) \circ \widetilde{H}_{\alpha\gamma} \\ 0 & \Pi_{\mathcal{S}^{|\beta|}_-}(\widetilde{H}_{\beta\beta}) & \widetilde{H}_{\beta\gamma} \\ (E_{\alpha\gamma} - \Sigma_{\alpha\gamma})^T \circ \widetilde{H}^T_{\alpha\gamma} & \widetilde{H}_{\beta\gamma} & \widetilde{H}_{\gamma\gamma} \end{bmatrix} \overline{P}^T , \tag{21}
$$

where E is a $n \times n$ matrix whose entries are all ones. Denote

$$
\Theta_1 := \begin{bmatrix} E_{\alpha\alpha} & E_{\alpha\beta} & \Sigma_{\alpha\gamma} \\ E^T_{\alpha\beta} & 0 & 0 \\ \Sigma^T_{\alpha\gamma} & 0 & 0 \end{bmatrix} \quad \text{and} \quad \Theta_2 := \begin{bmatrix} 0 & 0 & E_{\alpha\gamma} - \Sigma_{\alpha\gamma} \\ 0 & 0 & E_{\beta\gamma} \\ (E_{\alpha\gamma} - \Sigma_{\alpha\gamma})^T & E^T_{\beta\gamma} & E_{\gamma\gamma} \end{bmatrix} . \tag{22}
$$

We are now in a position to derive the precise expression of the proximal normal cone to $\text{gph } N_{\mathcal{S}^n_+}$.

Proposition 3.2 *For any* $(X, Y) \in \operatorname{gph} N_{\mathcal{S}^n_+}$, *let* $A = X + Y$ *have the eigenvalue decomposition (10). Then*

$$N^\pi_{\operatorname{gph} N_{\mathcal{S}^n_+}}(X, Y) = \left\{ (X^*, Y^*) \in \mathcal{S}^n \times \mathcal{S}^n : \Theta_1 \circ \widetilde{X}^* + \Theta_2 \circ \widetilde{Y}^* = 0, \ \widetilde{X}^*_{\beta\beta} \preceq 0 \text{ and } \widetilde{Y}^*_{\beta\beta} \succeq 0 \right\},$$

where $\widetilde{X}^* := \overline{P}^T X^* \overline{P}$ *and* $\widetilde{Y}^* := \overline{P}^T Y^* \overline{P}$.

Proof. By Proposition 3.1, $(X^*, Y^*) \in N^\pi_{\operatorname{gph} N_{\mathcal{S}^n_+}}(X, Y)$ if and only if

$$\langle X^*, \Pi'_{\mathcal{S}^n_+}(A; H) \rangle + \langle Y^*, \Pi'_{\mathcal{S}^n_-}(A; H) \rangle \leq 0 \quad \forall H \in \mathcal{S}^n,$$

which, together with the directional derivative formulas (14) and (21) implies that $(X^*, Y^*) \in N^\pi_{\operatorname{gph} N_{\mathcal{S}^n_+}}(X, Y)$ if and only if

$$\langle \Theta_1 \circ \widetilde{X}^*, \widetilde{H} \rangle + \langle \Theta_2 \circ \widetilde{Y}^*, \widetilde{H} \rangle + \langle \widetilde{X}^*_{\beta\beta}, \Pi_{\mathcal{S}^{|\beta|}_+}(\widetilde{H}_{\beta\beta}) \rangle + \langle \widetilde{Y}^*_{\beta\beta}, \Pi_{\mathcal{S}^{|\beta|}_-}(\widetilde{H}_{\beta\beta}) \rangle \leq 0 \quad \forall H \in \mathcal{S}^n.$$

The conclusion of the proposition holds. ∎

3.2 Expression of the limiting normal cone

In this subsection, we will use the formula of the proximal normal cone $N^\pi_{\operatorname{gph} N_{\mathcal{S}^n_+}}(X, Y)$ obtained in Proposition 3.2 to characterize the limiting normal cone $N_{\operatorname{gph} N_{\mathcal{S}^n_+}}(X, Y)$.

For any given $(X, Y) \in \operatorname{gph} N_{\mathcal{S}^n_+}$, let $A = X + Y$ have the eigenvalue decomposition (10) and β be the index set of zero eigenvalues of A. Denote the set of all partitions of the index set β by $\mathscr{P}(\beta)$. Let $\Re^{|\beta|}_{\gtrsim}$ be the set of all vectors in $\Re^{|\beta|}$ whose components being arranged in non-increasing order, i.e.,

$$\Re^{|\beta|}_{\gtrsim} := \left\{ z \in \Re^{|\beta|} : z_1 \geq \cdots \geq z_{|\beta|} \right\}.$$

For any $z \in \Re^{|\beta|}_{\gtrsim}$, let $D(z)$ represent the generalized first divided difference matrix for $f(t) = \max\{t, 0\}$ at z, i.e.,

$$(D(z))_{ij} = \begin{cases} \dfrac{\max\{z_i, 0\} - \max\{z_j, 0\}}{z_i - z_j} \in [0, 1] & \text{if } z_i \neq z_j, \\ 1 & \text{if } z_i = z_j > 0, \\ 0 & \text{if } z_i = z_j \leq 0, \end{cases} \quad i, j = 1, \ldots, |\beta|. \quad (23)$$

Denote

$$\mathcal{U}_{|\beta|} := \{ \overline{\Omega} \in \mathcal{S}^{|\beta|} : \overline{\Omega} = \lim_{k \to \infty} D(z^k), \ z^k \to 0, \ z^k \in \Re^{|\beta|}_{\gtrsim} \}. \quad (24)$$

Let $\Xi_1 \in \mathcal{U}_{|\beta|}$. Then, from (23), it is easy to see that there exists a partition $\pi(\beta) := (\beta_+, \beta_0, \beta_-) \in \mathscr{P}(\beta)$ such that

$$\Xi_1 = \begin{bmatrix} E_{\beta_+\beta_+} & E_{\beta_+\beta_0} & (\Xi_1)_{\beta_+\beta_-} \\ E^T_{\beta_+\beta_0} & 0 & 0 \\ (\Xi_1)^T_{\beta_+\beta_-} & 0 & 0 \end{bmatrix}, \quad (25)$$

13

where each element of $(\Xi_1)_{\beta_+\beta_-}$ belongs to $[0,1]$. Let

$$\Xi_2 := \begin{bmatrix} 0 & 0 & E_{\beta_+\beta_-} - (\Xi_1)_{\beta_+\beta_-} \\ 0 & 0 & E_{\beta_0\beta_-} \\ (E_{\beta_+\beta_-} - (\Xi_1)_{\beta_+\beta_-})^T & E_{\beta_0\beta_-}^T & E_{\beta_-\beta_-} \end{bmatrix}. \tag{26}$$

We first characterize the limiting normal cone $N_{\mathrm{gph}\,N_{\mathcal{S}_+^n}}(X,Y)$ for the special case when $(X,Y)=(0,0)$ and $\beta=\{1,2,\dots,n\}$.

Proposition 3.3 *The limiting normal cone to the graph of the limiting normal cone mapping $N_{\mathcal{S}_+^n}$ at $(0,0)$ is given by*

$$N_{\mathrm{gph}\,N_{\mathcal{S}_+^n}}(0,0) = \bigcup_{\substack{Q\,\in\,\mathcal{O}^n \\ \Xi_1\,\in\,\mathcal{U}_n}} \left\{ \begin{array}{cc} (U^*,V^*): & \Xi_1\circ Q^T U^* Q + \Xi_2\circ Q^T V^* Q = 0, \\ & Q_{\beta_0}^T U^* Q_{\beta_0} \preceq 0, \quad Q_{\beta_0}^T V^* Q_{\beta_0} \succeq 0 \end{array} \right\}. \tag{27}$$

Proof. See the Appendix. ∎

We now characterize the limiting normal cone $N_{\mathrm{gph}\,N_{\mathcal{S}_+^n}}(X,Y)$ for any $(X,Y)\in\mathrm{gph}\,N_{\mathcal{S}_+^n}$ for the general case in the following theorem.

Theorem 3.1 *For any $(X,Y)\in\mathrm{gph}\,N_{\mathcal{S}_+^n}$, let $A=X+Y$ have the eigenvalue decomposition (10). Then, $(X^*,Y^*)\in N_{\mathrm{gph}\,N_{\mathcal{S}_+^n}}(X,Y)$ if and only if*

$$X^* = \overline{P} \begin{bmatrix} 0 & 0 & \widetilde{X}_{\alpha\gamma}^* \\ 0 & \widetilde{X}_{\beta\beta}^* & \widetilde{X}_{\beta\gamma}^* \\ \widetilde{X}_{\gamma\alpha}^* & \widetilde{X}_{\gamma\beta}^* & \widetilde{X}_{\gamma\gamma}^* \end{bmatrix} \overline{P}^T \quad\text{and}\quad Y^* = \overline{P} \begin{bmatrix} \widetilde{Y}_{\alpha\alpha}^* & \widetilde{Y}_{\alpha\beta}^* & \widetilde{Y}_{\alpha\gamma}^* \\ \widetilde{Y}_{\beta\alpha}^* & \widetilde{Y}_{\beta\beta}^* & 0 \\ \widetilde{Y}_{\gamma\alpha}^* & 0 & 0 \end{bmatrix} \overline{P}^T \tag{28}$$

with

$$(\widetilde{X}_{\beta\beta}^*, \widetilde{Y}_{\beta\beta}^*) \in N_{\mathrm{gph}\,N_{\mathcal{S}_+^{|\beta|}}}(0,0) \quad\text{and}\quad \Sigma_{\alpha\gamma}\circ\widetilde{X}_{\alpha\gamma}^* + (E_{\alpha\gamma}-\Sigma_{\alpha\gamma})\circ\widetilde{Y}_{\alpha\gamma}^* = 0, \tag{29}$$

where Σ is given by (15), $\widetilde{X}^* = \overline{P}^T X^* \overline{P}$, $\widetilde{Y}^* = \overline{P}^T Y^* \overline{P}$ and

$$N_{\mathrm{gph}\,N_{\mathcal{S}_+^{|\beta|}}}(0,0) = \bigcup_{\substack{Q\,\in\,\mathcal{O}^{|\beta|} \\ \Xi_1\,\in\,\mathcal{U}_{|\beta|}}} \left\{ \begin{array}{cc} (U^*,V^*): & \Xi_1\circ Q^T U^* Q + \Xi_2\circ Q^T V^* Q = 0, \\ & Q_{\beta_0}^T U^* Q_{\beta_0} \preceq 0, \quad Q_{\beta_0}^T V^* Q_{\beta_0} \succeq 0 \end{array} \right\}.$$

Proof. See the Appendix. ∎

Remark 3.1 *For any given $(X,Y)\in\mathrm{gph}\,N_{\mathcal{S}_+^n}$, the (Mordukhovich) coderivative $D^*N_{\mathcal{S}_+^n}(X,Y)$ of the normal cone to the set $\mathcal{S}_+^n$ can be calculated by using Theorem 3.1 and the definition of coderivative, i.e., for given $Y^*\in\mathcal{S}^n$,*

$$X^*\in D^*N_{\mathcal{S}_+^n}(X,Y)(Y^*) \quad\Longleftrightarrow\quad (X^*,-Y^*)\in N_{\mathrm{gph}\,N_{\mathcal{S}_+^n}}(X,Y).$$

Furthermore, by (11) in Proposition 2.5, we know that

$$\mathrm{gph}\,N_{\mathcal{S}_+^n} = \{(X,Y)\in\mathcal{S}^n\times\mathcal{S}^n : L(X,Y)\in\mathrm{gph}\,\Pi_{S_+^n}\},$$

where $L: \mathcal{S}^n \times \mathcal{S}^n \to \mathcal{S}^n \times \mathcal{S}^n$ is a linear function defined by

$$L(X, Y) := (X + Y, X), \quad (X, Y) \in \mathcal{S}^n \times \mathcal{S}^n.$$

By noting that the derivative of L is nonsingular and self-adjoint, we know from [30, Theorem 6.10] that for any given $(X, Y) \in \mathrm{gph}\, N_{\mathcal{S}^n_+}$ and $Y^* \in \mathcal{S}^n$,

$$D^* N_{\mathcal{S}^n_+}(X, Y)(-Y^*) = \{X^* \in \mathcal{S}^n : (X^*, Y^*) \in L'(X, Y) N_{\mathrm{gph}\, \Pi_{\mathcal{S}^n_+}}(X + Y, X)\}.$$

Thus, for any given $U^* \in \mathcal{S}^n$, $V^* \in D^* \Pi_{\mathcal{S}^n_+}(X + Y)(U^*)$ if and only if there exists $(X^*, Y^*) \in N_{\mathrm{gph}\, N_{\mathcal{S}^n_+}}(X, Y)$ such that $(X^*, Y^*) = L(V^*, -U^*)$, that is,

$$X^* = V^* - U^* \quad \text{and} \quad Y^* = V^*.$$

Note that for any given $Z \in \mathcal{S}^n$, there exists a unique element $(X, Y) \in \mathrm{gph}\, N_{\mathcal{S}^n_+}$ such that $Z = X + Y$. Hence, the coderivative of the metric projector operator $\Pi_{\mathcal{S}^n_+}(\cdot)$ at any $Z \in \mathcal{S}^n$ can also be computed by Theorem 3.1.

4 Failure of Robinson's CQ

Since for any $(G(z), H(z)) \in \mathcal{S}^n_+ \times \mathcal{S}^n_-$, by the von Neumann-Theobald theorem (Theorem 2.3), one always has

$$\langle G(z), H(z) \rangle \le \lambda(G(z))^T \lambda(H(z)) \le 0.$$

Consequently one can rewrite the SDCMPCC problem in the following form:

$$
\begin{aligned}
(CP - SDCMPCC) \qquad \min \quad & f(z) \\
s.t. \quad & h(z) = 0, \\
& g(z) \preceq_{\mathcal{Q}} 0, \\
& \langle G(z), H(z) \rangle \ge 0, \\
& (G(z), H(z)) \in \mathcal{S}^n_+ \times \mathcal{S}^n_-.
\end{aligned}
$$

Rewriting the constraints $g(z) \preceq_{\mathcal{Q}} 0$ and $(G(z), H(z)) \in \mathcal{S}^n_+ \times \mathcal{S}^n_-$ as the cone constraint

$$(g(z), G(z), H(z)) \in -\mathcal{Q} \times \mathcal{S}^n_+ \times \mathcal{S}^n_-,$$

we know that the above problem belongs to the class of general optimization problems with a cone constraint (GP) as discussed in §2.2. Hence, the necessary optimality condition stated in §2.2 can be applied to obtain the following classical KKT condition.

Definition 4.1 *Let $\bar{z}$ be a feasible solution of SDCMPCC. We call $\bar{z}$ a classical KKT point if there exists $(\lambda^h, \lambda^g, \lambda^e, \Omega^G, \Omega^H) \in \Re^p \times \mathcal{H} \times \Re \times \mathcal{S}^n \times \mathcal{S}^n$ with $\lambda^g \in \mathcal{Q}$, $\lambda^e \le 0$, $\Omega^G \preceq 0$ and $\Omega^H \succeq 0$ such that*

$$0 = \nabla f(\bar{z}) + h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + \lambda^e [H'(\bar{z})^* G(\bar{z}) + G'(\bar{z})^* H(\bar{z})] + G'(\bar{z})^* \Omega^G + H'(\bar{z})^* \Omega^H,$$

$$\langle g(\bar{z}), \lambda^g \rangle = 0, \quad G(\bar{z})\Omega^G = 0, \quad H(\bar{z})\Omega^H = 0.$$

Theorem 4.1 *Let $\bar{z}$ be a local optimal solution of SDCMPCC. Suppose that the problem CP-SDCMPCC is Clarke calm at $\bar{z}$; in particular the set-valued map*

$$\mathcal{F}(r, s, t, P) :=$$
$$\{z : h(z) + r = 0, g(z) + s \preceq_{\mathcal{Q}} 0, -\langle G(z), H(z)\rangle + t \le 0, (G(z), H(z)) + P \in \mathcal{S}_+^n \times \mathcal{S}_-^n\}$$

(30)

is calm at $(0, 0, 0, 0, \bar{z})$. Then $\bar{z}$ is a classical KKT point.

Proof. By Theorem 2.2, there exists a Lagrange multiplier $(\lambda^h, \lambda^e, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \Re^q \times \Re \times \mathcal{H} \times \mathcal{S}^n \times \mathcal{S}^n$ with $\lambda^e \le 0$ such that

$$0 = \nabla f(\bar{z}) + h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + \lambda^e [H'(\bar{z})^* G(\bar{z}) + G'(\bar{z})^* H(\bar{z})] + G'(\bar{z})^* \Gamma^G + H'(\bar{z})^* \Gamma^H,$$
$$(\lambda^g, \Gamma^G, \Gamma^H) \in N_{-\mathcal{Q} \times \mathcal{S}_+^n \times \mathcal{S}_-^n}(g(\bar{z}), G(\bar{z}), H(\bar{z})).$$

Since $\mathcal{Q}$ is a symmetric cone it follows that $\lambda^g \in \mathcal{Q}$ and $\langle g(\bar{z}), \lambda^g \rangle = 0$. The desired result follows from the normal cone expressions in Proposition 2.4. ∎

Definition 4.2 *We say that $(\lambda^h, \lambda^g, \lambda^e, \Omega^G, \Omega^H) \in \Re^p \times \mathcal{H} \times \Re \times \mathcal{S}^n \times \mathcal{S}^n$ with $\lambda^g \in \mathcal{Q}, \lambda^e \le 0, \Omega^G \preceq 0, \Omega^H \succeq 0$ is a singular Lagrange multiplier for CP-SDCMPCC if it is not equal to zero and*

$$0 = h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + \lambda^e [H'(\bar{z})^* G(\bar{z}) + G'(\bar{z})^* H(\bar{z})] + G'(\bar{z})^* \Omega^G + H'(\bar{z})^* \Omega^H,$$
$$\langle g(\bar{z}), \lambda^g \rangle = 0, \quad G(\bar{z})\Omega^G = 0, \quad H(\bar{z})\Omega^H = 0.$$

For a general optimization problem with a cone constraint such as CP-SDCMPCC, the following Robinson's CQ is considered to be a usual constraint qualification:

$$h'(\bar{z}) \text{ is onto} \quad (\text{ equivalently } h_i'(\bar{z})(i = 1, \ldots, p) \text{ are linearly independent}),$$

$$\exists \, d \text{ such that } \begin{cases} h_i'(\bar{z})d = 0, \quad i = 1, \ldots, p, \\ -g(\bar{z}) - g'(\bar{z})d \in \text{int } \mathcal{Q} \\ (H'(\bar{z})^* G(\bar{z}) + G'(\bar{z})^* H(\bar{z}))d > 0, \\ G(\bar{z}) + G'(\bar{z})d \in \text{int } \mathcal{S}_+^n, \\ H(\bar{z}) + H'(\bar{z})d \in \text{int } \mathcal{S}_-^n. \end{cases}$$

It is well-known that the MFCQ never holds for MPCCs. We now show that Robinson's CQ will never hold for CP-SDCMPCC.

Proposition 4.1 *For CP-SDCMPCC, Robinson's constraint qualification fails to hold at every feasible solution of SDCMPCC.*

Proof. By the von Neumann-Theobald theorem, $G(z) \succeq 0$ and $H(z) \preceq 0$ implies that $\langle G(z), H(z)\rangle \le 0$. Hence any feasible solution $\bar{z}$ of SDCMPCC must be a solution to the following nonlinear semidefinite program:

$$\min \quad -\langle G(z), H(z)\rangle$$
$$\text{s.t.} \quad G(z) \succeq 0, \quad H(z) \preceq 0.$$

16

Since for this problem, $f(z) = -\langle G(z), H(z) \rangle$, we have $\nabla f(z) = -H'(z)^* G(z) - G'(z)^* H(z)$. By the first order necessary optimality condition, there exist $\lambda^e = 1, \Omega^G \preceq 0, \Omega^H \succeq 0$ such that

$$0 = -\lambda^e[H'(\bar{z})^* G(\bar{z}) + G'(\bar{z})^* H(\bar{z})] + G'(\bar{z})^* \Omega^G + H'(\bar{z})^* \Omega^H \,,$$
$$G(\bar{z})\Omega^G = 0, \quad H(\bar{z})\Omega^H = 0 \,.$$

Since $(-\lambda^e, \Omega^G, \Omega^H) \neq 0$, it is clear that $(0, 0, 0, -\lambda^e, \Omega^G, \Omega^H)$ is a singular Lagrange multiplier of CP-SDCMPCC. By [6, Propositions 3.16 (ii) and 3.19(iii)]), a singular Lagrange multiplier exists if and only if Robinson's CQ does not hold. Therefore we conclude that the Robinson's CQ does not hold at $\bar{z}$ for CP-SDCMPCC. ∎

5 S-stationary conditions

In the MPCC literature [26, 61], using the so-called "piecewise programming approach" to rewrite the feasible region as a union of branches which consist of only ordinary equality and inequality constraints, one derives the S-stationary condition as a necessary optimality condition for a local optimal solution under the condition that each branch has a common multiplier. Moreover it is well-known that the S-stationary condition is equivalent to the classical KKT condition; see e.g., [17]. In this section we introduce the concept of S-stationary condition and show that the classical KKT condition implies the S-stationary condition. Unfortunately for SDCMPCC, "piecewise programming approach" is not applicable any more. Hence the converse implication may not be true in general.

For MPCC, the S-stationary condition is shown to be equivalent to the necessary optimality condition of a reformulated problem involving the proximal normal cone to the graph of the normal cone operator (see [61, Theorem 3.2]). Motivated by this fact and the precise expression for the proximal normal cone formula in Proposition 3.2, we introduce the concept of a S-stationary point for SDCMPCC.

Definition 5.1 *Let $\bar{z}$ be a feasible solution of SDCMPCC. Let $A := G(\bar{z}) + H(\bar{z})$ have the eigenvalue decomposition (10). We say that $\bar{z}$ is a S-stationary point of SDCMPCC if there exists $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \mathcal{H} \times \mathcal{S}^n \times \mathcal{S}^n$ such that*

$$0 = \nabla f(\bar{z}) + h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + G'(\bar{z})^* \Gamma^G + H'(\bar{z})^* \Gamma^H \,, \tag{31}$$

$$\lambda^g \in \mathcal{Q}, \quad \langle \lambda^g, g(\bar{z}) \rangle = 0 \,, \tag{32}$$

$$\widetilde{\Gamma}^G_{\alpha\alpha} = 0, \quad \widetilde{\Gamma}^G_{\alpha\beta} = 0, \quad \widetilde{\Gamma}^G_{\beta\alpha} = 0 \,, \tag{33}$$

$$\widetilde{\Gamma}^H_{\gamma\gamma} = 0, \quad \widetilde{\Gamma}^H_{\beta\gamma} = 0, \quad \widetilde{\Gamma}^H_{\gamma\beta} = 0 \,, \tag{34}$$

$$\Sigma_{\alpha\gamma} \circ \widetilde{\Gamma}^G_{\alpha\gamma} + (E_{\alpha\gamma} - \Sigma_{\alpha\gamma}) \circ \widetilde{\Gamma}^H_{\alpha\gamma} = 0 \,, \tag{35}$$

$$\widetilde{\Gamma}^G_{\beta\beta} \preceq 0, \quad \widetilde{\Gamma}^H_{\beta\beta} \succeq 0 \,, \tag{36}$$

where E is a $n \times n$ matrix whose entries are all ones and $\Sigma \in \mathcal{S}^n$ is defined by (15), and $\widetilde{\Gamma}^G = \overline{P}^T \Gamma^G \overline{P}$ and $\widetilde{\Gamma}^H = \overline{P}^T \Gamma^H \overline{P}$.

We now show that for SDCMPCC, the classical KKT condition implies the S-stationary condition.

Proposition 5.1 *Let $\bar{z}$ be a feasible solution of SDCMPCC. If $\bar{z}$ is a classical KKT point, i.e., there exists a classical Lagrange multiplier $(\lambda^h, \lambda^g, \lambda^e, \Omega^G, \Omega^H) \in \Re^p \times \mathcal{H} \times \Re \times \mathcal{S}^n \times \mathcal{S}^n$ with $\lambda^g \in \mathcal{Q}$, $\lambda^e \leq 0$, $\Omega^G \preceq 0$ and $\Omega^H \succeq 0$ such that*

$$0 = \nabla f(\bar{z}) + h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + \lambda^e [H'(\bar{z})^* G(\bar{z}) + G'(\bar{z})^* H(\bar{z})] + G'(\bar{z})^* \Omega^G + H'(\bar{z})^* \Omega^H,$$

$$\langle \lambda^g, g(\bar{z}) \rangle = 0, \quad G(\bar{z}) \Omega^G = 0, \quad H(\bar{z}) \Omega^H = 0,$$

then it is also a S-stationary point.

Proof. Denote $\overline{\Lambda} := \Lambda(A)$. Define $\Gamma^G := \Omega^G + \lambda^e H(\bar{z})$ and $\Gamma^H := \Omega^H + \lambda^e G(\bar{z})$. Then (31) and (32) hold. It remains to show (33)-(36). By the assumption we have

$$\mathcal{S}^n_+ \ni G(\bar{z}) \perp \Omega^G \in \mathcal{S}^n_- \quad \text{and} \quad \mathcal{S}^n_- \ni H(\bar{z}) \perp \Omega^H \in \mathcal{S}^n_+.$$

By Theorem 2.3, we know that $G(\bar{z})$ and Ω^G ($H(\bar{z})$ and Ω^H) admit a simultaneous ordered eigenvalue decomposition, i.e., there exist two orthogonal matrices $\widetilde{P}, \widehat{P} \in \mathcal{O}^n$ such that

$$\Omega^G = \widetilde{P} \begin{bmatrix} 0 & 0 \\ 0 & \Lambda(\Omega^G)_{\gamma'\gamma'} \end{bmatrix} \widetilde{P}^T, \quad G(\bar{z}) = \widetilde{P} \begin{bmatrix} \overline{\Lambda}_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \widetilde{P}^T$$

and

$$\Omega^H = \widehat{P} \begin{bmatrix} \Lambda(\Omega^H)_{\alpha'\alpha'} & 0 \\ 0 & 0 \end{bmatrix} \widehat{P}^T, \quad H(\bar{z}) = \widehat{P} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \overline{\Lambda}_{\gamma\gamma} \end{bmatrix} \widehat{P}^T,$$

where $\alpha' := \{i \,|\, \lambda_i(\Omega^H) > 0\}$ and $\gamma' := \{i \,|\, \lambda_i(\Omega^G) < 0\}$. Moreover, we have

$$\gamma' \subseteq \bar{\alpha} \quad \text{and} \quad \alpha' \subseteq \bar{\gamma}, \tag{37}$$

where $\bar{\alpha} := \beta \cup \gamma$, $\bar{\gamma} := \alpha \cup \beta$. On the other hand, we know that

$$G(\bar{z}) = \Pi_{\mathcal{S}^n_+}(A) = \overline{P} \begin{bmatrix} \overline{\Lambda}_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \overline{P}^T \quad \text{and} \quad H(\bar{z}) = \Pi_{\mathcal{S}^n_-}(A) = \overline{P} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \overline{\Lambda}_{\gamma\gamma} \end{bmatrix} \overline{P}^T.$$

Therefore, it is easy to check that there exist two orthogonal matrices $S, T \in \mathcal{O}^n$ such that

$$\overline{P} = \widetilde{P} S \quad \text{and} \quad \overline{P} = \widehat{P} T,$$

with

$$S = \begin{bmatrix} S_{\alpha\alpha} & 0 \\ 0 & S_{\bar{\alpha}\bar{\alpha}} \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} T_{\bar{\gamma}\bar{\gamma}} & 0 \\ 0 & T_{\gamma\gamma} \end{bmatrix},$$

where $S_{\alpha\alpha} \in \mathcal{O}^{|\alpha|}$, $S_{\bar{\alpha}\bar{\alpha}} \in \mathcal{O}^{|\bar{\alpha}|}$ and $T_{\bar{\gamma}\bar{\gamma}} \in \mathcal{O}^{|\bar{\gamma}|}$, $T_{\gamma\gamma} \in \mathcal{O}^{|\gamma|}$. Denote

$$S_{\bar{\alpha}\bar{\alpha}} = [S_1 \quad S_2] \quad \text{and} \quad T_{\bar{\gamma}\bar{\gamma}} = [T_1 \quad T_2]$$

with $S_1 \in \Re^{|\bar{\alpha}| \times |\beta|}$, $S_2 \in \Re^{|\bar{\alpha}| \times |\gamma|}$ and $T_1 \in \Re^{|\bar{\gamma}| \times |\alpha|}$ and $T_2 \in \Re^{|\bar{\gamma}| \times |\beta|}$. Then, we have

$$\widetilde{\Gamma}^G = \overline{P}^T(\Omega^G + \lambda^e H(\bar{z}))\overline{P} = S^T \widetilde{P}^T \Omega^G \widetilde{P} S + \lambda^e \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \overline{\Lambda}_{\gamma\gamma} \end{bmatrix}$$

$$= \begin{bmatrix} S_{\alpha\alpha}^T & 0 \\ 0 & S_{\bar{\alpha}\bar{\alpha}}^T \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} \end{bmatrix} \begin{bmatrix} S_{\alpha\alpha} & 0 \\ 0 & S_{\bar{\alpha}\bar{\alpha}} \end{bmatrix} + \lambda^e \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \overline{\Lambda}_{\gamma\gamma} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & S_1^T \Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} S_1 & S_1^T \Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} S_2 \\ 0 & S_2^T \Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} S_1 & S_2^T \Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} S_2 + \lambda^e \overline{\Lambda}_{\gamma\gamma} \end{bmatrix}$$

and

$$\widetilde{\Gamma}^H = \overline{P}^T(\Omega^H + \lambda^e G(\bar{z}))\overline{P} = T^T \widetilde{P}^T \Omega^H \widetilde{P} T + \lambda^e \begin{bmatrix} \overline{\Lambda}_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} T_{\bar{\gamma}\bar{\gamma}}^T & 0 \\ 0 & T_{\gamma\gamma}^T \end{bmatrix} \begin{bmatrix} \Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} T_{\bar{\gamma}\bar{\gamma}} & 0 \\ 0 & T_{\gamma\gamma} \end{bmatrix} + \lambda^e \begin{bmatrix} \overline{\Lambda}_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} T_1^T \Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} T_1 + \lambda^e \overline{\Lambda}_{\alpha\alpha} & T_1^T \Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} T_2 & 0 \\ T_2^T \Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} T_1 & T_2^T \Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} T_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore it is easy to see that (33)-(35) hold. Since $\Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} \preceq 0$, $\Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} \succeq 0$ and $S_{\bar{\alpha}\bar{\alpha}}$, $T_{\bar{\gamma}\bar{\gamma}}$ are orthogonal, we know that

$$S_{\bar{\alpha}\bar{\alpha}}^T \Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} S_{\bar{\alpha}\bar{\alpha}} \preceq 0 \quad \text{and} \quad T_{\bar{\gamma}\bar{\gamma}}^T \Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} T_{\bar{\gamma}\bar{\gamma}} \succeq 0.$$

Hence, we have

$$\widetilde{\Gamma}_{\beta\beta}^G = S_1^T \Lambda(\Omega^G)_{\bar{\alpha}\bar{\alpha}} S_1 \preceq 0 \quad \text{and} \quad \widetilde{\Gamma}_{\beta\beta}^H = T_2^T \Lambda(\Omega^H)_{\bar{\gamma}\bar{\gamma}} T_2 \succeq 0,$$

which implies (36). Therefore $\bar{z}$ is also a S-stationary point. ■

Combining Theorem 4.1 and Proposition 5.1 we have the following necessary optimality condition in terms of S-stationary conditions.

Corollary 5.1 *Let $\bar{z}$ be an optimal solution of SDCMPCC. Suppose the problem CP-SDCMPCC is Clarke calm at $\bar{z}$; in particular the set-valued map defined by (30) is calm at $(0, 0, 0, 0, \bar{z})$. Then $\bar{z}$ is a S-stationary point.*

6   M-stationary conditions

In this section we study the M-stationary conditon for SDCMPCC. For this purpose rewrite the SDCMPCC as an optimization problem with a cone constraint:

$$\begin{aligned} \text{(GP-SDCMPCC)} \quad \min \quad & f(z) \\ \text{s.t.} \quad & h(z) = 0, \\ & g(z) \preceq_Q 0, \\ & (G(z), H(z)) \in \text{gph}\, N_{\mathcal{S}_+^n}. \end{aligned}$$

19

Definition 6.1 *Let $\bar{z}$ be a feasible solution of SDCMPCC. Let $A = G(\bar{z}) + H(\bar{z})$ have the eigenvalue decomposition (10). We say that $\bar{z}$ is a M-stationary point of SDCMPCC if there exists $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \mathcal{H} \times \mathcal{S}^n \times \mathcal{S}^n$ such that (31)-(35) hold and there exist $Q \in \mathcal{O}^{|\beta|}$ and $\Xi_1 \in \mathcal{U}_{|\beta|}$ (with a partition $\pi(\beta) = (\beta_+, \beta_0, \beta_-)$ of β and the form (25)) such that*

$$\Xi_1 \circ Q^T \widetilde{\Gamma}^G_{\beta\beta} Q + \Xi_2 \circ Q^T \widetilde{\Gamma}^H_{\beta\beta} Q = 0\,, \tag{38}$$

$$Q^T_{\beta_0} \widetilde{\Gamma}^G_{\beta\beta} Q_{\beta_0} \preceq 0\,, \quad Q^T_{\beta_0} \widetilde{\Gamma}^H_{\beta\beta} Q_{\beta_0} \succeq 0\,, \tag{39}$$

where $\widetilde{\Gamma}^G = \overline{P}^T \Gamma^G \overline{P}$, $\widetilde{\Gamma}^H = \overline{P}^T \Gamma^H \overline{P}$ and Ξ_2 is defined by (26). We say that $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \mathcal{H} \times \mathcal{S}^n \times \mathcal{S}^n$ is a singular M-multiplier for SDCMPCC if it is not equal to zero and all conditions above hold except the term $\nabla f(\bar{z})$ vanishes in (31).

The following result is on the first order necessary optimality condition of SDCMPCC in terms of M-stationary conditions.

Theorem 6.1 *Let $\bar{z}$ be a local optimal solution of SDCMPCC. Suppose that the problem GP-SDCMPCC is Clarke calm at $\bar{z}$; in particular one of the following constraint qualifications holds.*

(i) *There is no singular M-multiplier for problem SDCMPCC at $\bar{z}$.*

(ii) *SDCMPCC LICQ holds at $\bar{z}$: there is no nonzero $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \mathcal{H} \times \mathcal{S}^n \times \mathcal{S}^n$ such that*

$$h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + G'(\bar{z})^* \Gamma^G + H'(\bar{z})^* \Gamma^H = 0\,, \tag{40}$$

$$\widetilde{\Gamma}^G_{\alpha\alpha} = 0\,, \quad \widetilde{\Gamma}^G_{\alpha\beta} = 0\,, \quad \widetilde{\Gamma}^G_{\beta\alpha} = 0\,,$$

$$\widetilde{\Gamma}^H_{\gamma\gamma} = 0\,, \quad \widetilde{\Gamma}^H_{\beta\gamma} = 0\,, \quad \widetilde{\Gamma}^H_{\gamma\beta} = 0\,,$$

$$\Sigma_{\alpha\gamma} \circ \widetilde{\Gamma}^G_{\alpha\gamma} + (E_{\alpha\gamma} - \Sigma_{\alpha\gamma}) \circ \widetilde{\Gamma}^H_{\alpha\gamma} = 0\,.$$

(iii) *Assume that there is no inequality constraint $g(z) \preceq_{\mathcal{Q}} 0$. Assume also that $Z = X \times \mathcal{S}^n$ where X is a finite dimensional space and $G(x, u) = u$. The following generalized equation is strongly regular in the sense of Robinson:*

$$0 \in -F(x, u) + N_{\Re^q \times \mathcal{S}^n_+}(x, u)\,,$$

where $F(x, u) = (h(x, u), H(x, u))$.

(iv) *Assume that there is no inequality constraint $g(z) \preceq_{\mathcal{Q}} 0$. Assume also that $Z = X \times \mathcal{S}^n$, $G(z) = u$ and $F(x, u) = (h(x, u), H(x, u))$. $-F$ is locally strongly monotone in u uniformly in x with modulus $\delta > 0$, i.e., there exist neighborhood U_1 of $\bar{x}$ and U_2 of $\bar{u}$ such that*

$$\langle -F(x, u) + F(x, v), u - v \rangle \geq \delta \|u - v\|^2 \quad \forall u \in U_2 \cap \mathcal{S}^n_+, \; v \in \mathcal{S}^n_+, \; x \in U_1\,.$$

Then $\bar{z}$ is a M-stationary point of SDCMPCC.

Proof. Condition (ii) is obviously stronger than (i). Condition (i) is a necessary and sufficient condition for the perturbed feasible region of the constraint system to be pseudo Lipschitz continuous; see e.g., [33, Theorem 6.1]. See [62, Theorem 4.7] for the proof of the implication of (iii) to (i). (iv) is a sufficient condition for (iii) and the direct proof can be found in [64, Theorem 3.2(b)]. The desired result follows from Theorem 2.2 and the expression of the limiting normal cone in Theorem 3.1. ∎

Next, we give two SDCMPCC examples to illustrate the M-stationary conditions. Note that in the first example the local solution is a M-stationary point, but not a S-stationary point.

Example 6.1 *Consider the following SDCMPCC problem*

$$\begin{aligned} \min \quad & -\langle I, X \rangle + \langle I, Y \rangle \\ \text{s.t.} \quad & X + Y = 0, \\ & \mathcal{S}_+^n \ni X \perp Y \in \mathcal{S}_-^n. \end{aligned} \tag{41}$$

Since the unique feasible point of (41) *is* $(0,0)$*, we know that* $(X^*, Y^*) = (0,0)$ *is the optimal solution of* (41)*. Note that* $A = X^* + Y^* = 0$*, which implies that*

$$\alpha = \emptyset, \quad \beta = \{1, \ldots, n\} \quad \text{and} \quad \gamma = \emptyset.$$

Without loss of generality, we may choose $\overline{P} = I$*. Therefore, by considering the equation* (31)*, we know that*

$$\begin{bmatrix} -I \\ I \end{bmatrix} + \begin{bmatrix} \Gamma^e \\ \Gamma^e \end{bmatrix} + \begin{bmatrix} \Gamma^G \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \Gamma^H \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which implies that

$$\Gamma^G = I - \Gamma^e \quad \text{and} \quad \Gamma^H = -I - \Gamma^e, \tag{42}$$

where $\Gamma^e \in \mathcal{S}^n$*. Let* $\Gamma^e = I$*. Then, it is clear that the equation* (38) *holds for* $\Gamma^G = 0$ *and* $\Gamma^H = -2I$ *with* $\beta_+ = \beta = \{1, \ldots, n\}$*,* $\beta_0 = \beta_- = \emptyset$*, and* $Q = I \in \mathcal{O}^n$*. By noting that* $\beta_0 = \emptyset$*, we know that the optimal solution* $(X^*, Y^*) = (0,0)$ *is a M-stationary point with the multiplier* $(I, 0, -2I) \in \mathcal{S}^n \times \mathcal{S}^n \times \mathcal{S}^n$*. However, the optimal solution* $(X^*, Y^*) = (0,0)$ *is not a S-stationary point. In fact, we know from* (42) *that if there exists some* $\Gamma^e \in \mathcal{S}^n$ *such that* (36) *holds, then*

$$\Gamma^e \succeq I \quad \text{and} \quad \Gamma^e \preceq -I,$$

which is a contradiction.

Example 6.2 *As a direct application, we characterize the M-stationary condition of the rank constrained nearest correlation matrix problem* (2)*. For any given feasible point* $\overline{X} \in \mathcal{S}^n$ *of* (2)*, suppose that* $\overline{X}$ *has the eigenvalue decomposition* $\overline{X} = \overline{P} \Lambda(\overline{X}) \overline{P}^T$*. It is easy to check that* $(\overline{X}, U) \in \mathcal{S}^n \times \mathcal{S}^n$ *is a feasible solution of* (3) *if and only if* $U = \sum_{i=1}^r \overline{P}_i \overline{P}_i^T$ *(see e.g., [35, 36, 27, 23] for details). Assume* $\operatorname{rank}(\overline{X}) = \bar{r} \leq r$*. Then, the index sets of positive, zero and negative eigenvalues of* $A = \overline{X} + (U - I)$ *are given by* $\alpha = \{1, \ldots, \bar{r}\}$*,* $\beta = \{\bar{r}+1, \ldots, r\}$ *and* $\gamma = \{r+1, \ldots, n\}$*, respectively. Therefore, we say that the feasible* $\overline{X} \in \mathcal{S}^n$ *is a M-stationary point of* (2)*, if there exist* $(\lambda_1^h, \lambda_2^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^n \times \Re \times \mathcal{S}^n \times \mathcal{S}^n \times \mathcal{S}^n$*,* $Q \in \mathcal{O}^{|\beta|}$

and $\Xi_1 \in \mathcal{U}_{|\beta|}$ such that

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \nabla f_C(\overline{X}) \\ 0 \end{bmatrix} + \begin{bmatrix} \text{diag}(\lambda_1^h) \\ \lambda_2^h I + \lambda^g \end{bmatrix} + \begin{bmatrix} \Gamma^G \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \Gamma^H \end{bmatrix}, \tag{43}$$

$$0 \preceq U \perp \lambda^g \preceq 0, \tag{44}$$

$$\widetilde{\Gamma}_{\alpha\alpha}^G = 0, \quad \widetilde{\Gamma}_{\alpha\beta}^G = 0, \quad \widetilde{\Gamma}_{\beta\alpha}^G = 0, \quad \widetilde{\Gamma}_{\gamma\gamma}^H = 0, \quad \widetilde{\Gamma}_{\beta\gamma}^H = 0, \quad \widetilde{\Gamma}_{\gamma\beta}^H = 0, \tag{45}$$

$$\frac{\lambda_i(\overline{X})}{\lambda_i(\overline{X}) + 1} \widetilde{\Gamma}_{ij}^G + (1 - \frac{\lambda_i(\overline{X})}{\lambda_i(\overline{X}) + 1}) \widetilde{\Gamma}_{ij}^H = 0, \quad i \in \alpha, \ j \in \gamma, \tag{46}$$

$$\Xi_1 \circ Q^T \widetilde{\Gamma}_{\beta\beta}^G Q + \Xi_2 \circ Q^T \widetilde{\Gamma}_{\beta\beta}^H Q = 0, \tag{47}$$

$$Q_{\beta_0}^T \widetilde{\Gamma}_{\beta\beta}^G Q_{\beta_0} \preceq 0, \quad Q_{\beta_0}^T \widetilde{\Gamma}_{\beta\beta}^H Q_{\beta_0} \succeq 0, \tag{48}$$

where $\widetilde{\Gamma}^G = \overline{P}^T \Gamma^G \overline{P}$, $\widetilde{\Gamma}^H = \overline{P}^T \Gamma^H \overline{P}$ and Ξ_2 is defined by (26).

Remark 6.1 *SDCMPCC LICQ is the analogue of the well-known MPCC LICQ (also called MPEC LICQ). However, we would like to remark that unlike in MPCC case, we can only show that SDCMPCC LICQ is a constraint qualification for a M-stationary condition instead of a S-stationary condition.*

7 C-stationary conditions

In this section, we consider the C-stationary condition by reformulating SDCMPCC as a nonsmooth problem:

$$\begin{aligned} (\text{NS} - \text{SDCMPCC}) \qquad \min \quad & f(z) \\ \text{s.t.} \quad & h(z) = 0, \\ & g(z) \preceq_{\mathcal{Q}} 0, \\ & G(z) - \Pi_{\mathcal{S}_+^n}(G(z) + H(z)) = 0. \end{aligned}$$

From (11), we know that the reformulation NS-SDCMPCC is equivalent to SDCMPCC. As in the MPCC case, the C-stationary condition introduced below is the nonsmooth KKT condition of NS-SDCMPCC by using the Clarke subdifferential.

Definition 7.1 *Let $\bar{z}$ be a feasible solution of SDCMPCC. Let $A = G(\bar{z}) + H(\bar{z})$ have the eigenvalue decomposition (10). We say that $\bar{z}$ is a C-stationary point of SDCMPCC if there exists $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \Re^q \times \mathcal{S}^n \times \mathcal{S}^n$ such that (31)-(35) hold and*

$$\langle \widetilde{\Gamma}_{\beta\beta}^G, \widetilde{\Gamma}_{\beta\beta}^H \rangle \leq 0, \tag{49}$$

where $\widetilde{\Gamma}^G = \overline{P}^T \Gamma^G \overline{P}$ and $\widetilde{\Gamma}^H = \overline{P}^T \Gamma^H \overline{P}$. We say that $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \mathcal{H} \times \mathcal{S}^n \times \mathcal{S}^n$ is a singular C-multiplier for SDCMPCC if it is not equal to zero and all conditions above hold except the term $\nabla f(\bar{z})$ vanishes in (31).

Remark 7.1 *It is easy to see that as in MPCC case,*

S-stationary condition $\implies$ M-stationary condition $\implies$ C-stationary condition.

Indeed, since the proximal normal cone is included in the limiting normal cone, it is obvious that the S-stationary condition implies the M-stationary condition. We now show that the M-stationary condition implies the C-stationary condition. In fact, suppose that $\bar{z}$ is a M-stationary point of SDCMPCC. Then, there exists $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \Re^q \times \mathcal{S}^n \times \mathcal{S}^n$ such that (31)-(35) hold and there exist $Q \in \mathcal{O}^{|\beta|}$ and $\Xi_1 \in \mathcal{U}_{|\beta|}$ (with a partition $\pi(\beta) = (\beta_+, \beta_0, \beta_-)$ of β and the form (25)) such that (38) and (39) hold. Let $A = G(\bar{z}) + H(\bar{z})$ have the eigenvalue decomposition (10). Therefore, we know that

$$Q_{\beta_+}^T \widetilde{\Gamma}_{\beta\beta}^G Q_{\beta_+} = 0, \quad Q_{\beta_+}^T \widetilde{\Gamma}_{\beta\beta}^G Q_{\beta_-} = 0, \quad Q_{\beta_-}^T \widetilde{\Gamma}_{\beta\beta}^G Q_{\beta_+} = 0,$$

$$Q_{\beta_-}^T \widetilde{\Gamma}_{\beta\beta}^H Q_{\beta_-} = 0, \quad Q_{\beta_0}^T \widetilde{\Gamma}_{\beta\beta}^H Q_{\beta_-} = 0, \quad Q_{\beta_-}^T \widetilde{\Gamma}_{\beta\beta}^H Q_{\beta_0} = 0,$$

which implies that

$$\begin{aligned}
\langle \widetilde{\Gamma}_{\beta\beta}^G, \widetilde{\Gamma}_{\beta\beta}^H \rangle &= \langle Q^T \widetilde{\Gamma}_{\beta\beta}^G Q, Q^T \widetilde{\Gamma}_{\beta\beta}^H Q \rangle \\
&= 2\langle Q_{\beta_+}^T \widetilde{\Gamma}_{\beta\beta}^G Q_{\beta_-}, Q_{\beta_+}^T \widetilde{\Gamma}_{\beta\beta}^H Q_{\beta_-} \rangle + 2\langle Q_{\beta_0}^T \widetilde{\Gamma}_{\beta\beta}^G Q_{\beta_0}, Q_{\beta_0}^T \widetilde{\Gamma}_{\beta\beta}^H Q_{\beta_0} \rangle.
\end{aligned}$$

Note that for each $(i,j) \in \beta_+ \times \beta_-$, $(\Xi_1)_{ij} \in [0,1]$ and $(\Xi_2)_{ij} = 1 - (\Xi_1)_{ij}$. Therefore, we know from (38) that

$$\langle Q_{\beta_+}^T \widetilde{\Gamma}_{\beta\beta}^G Q_{\beta_-}, Q_{\beta_+}^T \widetilde{\Gamma}_{\beta\beta}^H Q_{\beta_-} \rangle \leq 0.$$

Finally, together with (39), we know that

$$\langle \widetilde{\Gamma}_{\beta\beta}^G, \widetilde{\Gamma}_{\beta\beta}^H \rangle \leq 0,$$

which implies that $\bar{z}$ is also a C-stationary point of SDCMPCC.

We present the first order optimality condition of SDCMPCC in terms of C-stationary conditions in the following result.

Theorem 7.1 *Let $\bar{z}$ be a local optimal solution of SDCMPCC. Suppose that the problem NS-SDCMPCC is Clarke calm at $\bar{z}$; in particular suppose that there is no singular C-multiplier for problem SDCMPCC at $\bar{z}$. Then $\bar{z}$ is a C-stationary point of SDCMPCC.*

Proof. By Theorem 2.2 with $K = \{0\}$, we know that there exist $\lambda^h \in \Re^p$, $\lambda^g \in \Re^q$ and $\Gamma \in S^n$ such that

$$0 \in \partial_z^c L(\bar{z}, \lambda^h, \lambda^g, \Gamma), \quad \lambda^g \geq 0 \quad \text{and} \quad \langle \lambda^g, g(\bar{z}) \rangle = 0, \tag{50}$$

where $L(z, \lambda^h, \lambda^g, \Gamma) := f(z) + \langle \lambda^h, h(z) \rangle + \langle \lambda^g, g(z) \rangle + \langle \Gamma, G(z) - \Pi_{\mathcal{S}_+^n}(G(z) + H(z)) \rangle$.

Consider the Clarke subdifferential of the nonsmooth part $S(z) := \langle \Gamma, \Pi_{\mathcal{S}_+^n}(G(z) + H(z)) \rangle$ of L. By the chain rule [9, Corollary pp.75], for any $v \in Z$, we have

$$\partial^c S(\bar{z})v \subseteq \langle \Gamma, \partial^c \Pi_{\mathcal{S}_+^n}(A)(G'(\bar{z})v + H'(\bar{z})v) \rangle.$$

Therefore, since any element of the Clarke subdifferential of the metric projection operator to a close convex set is self-adjoint (see e.g., [29, Proposition 1(a)]), we know from (50) that there exists $V \in \partial^c \Pi_{\mathcal{S}_+^n}(A)$ such that

$$\nabla f(\bar{z}) + h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + G'(\bar{z})^* \Gamma - (G'(\bar{z})^* + H'(\bar{z})^*)V(\Gamma) = 0. \tag{51}$$

Define $\Gamma^G := \Gamma - V(\Gamma)$ and $\Gamma^H := -V(\Gamma)$. Then (31)-(32) follow from (50) and (51) immediately. By [50, Proposition 2.2], we know that there exists $W \in \partial^c \Pi_{\mathcal{S}_+^{|\beta|}}(0)$ such that

$$V(\Gamma) = \overline{P} \begin{bmatrix} \widetilde{\Gamma}_{\alpha\alpha} & \widetilde{\Gamma}_{\alpha\beta} & \Sigma_{\alpha\gamma} \circ \widetilde{\Gamma}_{\alpha\gamma} \\ \widetilde{\Gamma}_{\alpha\beta}^T & W(\widetilde{\Gamma}_{\beta\beta}) & 0 \\ \widetilde{\Gamma}_{\alpha\gamma}^T \circ \Sigma_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} \overline{P}^T \,,$$

where $\Sigma \in \mathcal{S}^n$ is defined by (15). Therefore, it is easy to see that (33)-(35) hold. Moreover, from [29, Proposition 1(c)], we know that

$$\langle W(\widetilde{\Gamma}_{\beta\beta}), \widetilde{\Gamma}_{\beta\beta} - W(\widetilde{\Gamma}_{\beta\beta}) \rangle \geq 0 \,,$$

which implies $\langle \widetilde{\Gamma}_{\beta\beta}^G, \widetilde{\Gamma}_{\beta\beta}^H \rangle \leq 0$. Hence, we know $\bar{z}$ is a C-stationary point of SDCMPCC. $\blacksquare$

Next, we give an example whose optimal solution is a C-stationary point but not a M-stationary point.

Example 7.1 *Consider the following SDCMPCC problem*

$$\begin{aligned} \min \quad & \tfrac{1}{2}z_1 - \tfrac{1}{2}z_2 - z_3 - \tfrac{1}{2}z_4 \\ \text{s.t.} \quad & -2z_1 + z_3 + z_4 \leq 0 \,, \\ & 2z_2 + z_3 \leq 0 \,, \\ & z_4^2 \leq 0 \,, \\ & \mathcal{S}_+^3 \ni G(z) \perp H(z) \in \mathcal{S}_-^3 \,, \end{aligned} \qquad (52)$$

where $G : \Re^4 \to \mathcal{S}^3$ and $H : \Re^4 \to \mathcal{S}^3$ are the linear operators defined as follows for any $z = (z_1, z_2, z_3, z_4)^T \in \Re^4$,

$$G(z) := \begin{bmatrix} 1 + \frac{z_1}{6} & -1 + \frac{z_1}{6} & -\frac{z_1}{3} \\ -1 + \frac{z_1}{6} & 1 + \frac{z_1}{6} & -\frac{z_1}{3} \\ -\frac{z_1}{3} & -\frac{z_1}{3} & \frac{2z_1}{3} \end{bmatrix} \quad \text{and} \quad H(z) := \begin{bmatrix} \frac{z_2}{6} - 1 & \frac{z_2}{6} - 1 & -\frac{z_2}{3} - 1 \\ \frac{z_2}{6} - 1 & \frac{z_2}{6} - 1 & -\frac{z_2}{3} - 1 \\ -\frac{z_2}{3} - 1 & -\frac{z_2}{3} - 1 & \frac{2z_2}{3} - 1 \end{bmatrix} .$$

Since $\langle G(z), H(z) \rangle = z_1 z_2$, one can verify that $\bar{z} = (0,0,0,0)$ is the unique optimal solution of the problem (52). Thus, we have

$$A = G(\bar{z}) + H(\bar{z}) = \overline{P} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -3 \end{bmatrix} \overline{P}^T \,,$$

where $\overline{P}$ is the 3 by 3 orthogonal matrix given by

$$\overline{P} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} \,,$$

and the index sets of positive, zero and negative eigenvalues are $\alpha = \{1\}$, $\beta = \{2\}$ and $\gamma = \{3\}$. In the following we denote by $\frac{\partial G}{\partial z_1}$ the derivative of the mapping G with respect

to variable z_1. Since $G(z)$ only depends on z_1 and $H(z)$ only depends on z_2, (31) can be written as

$$
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -1 \\ -\frac{1}{2} \end{bmatrix} + \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} \lambda_1^g + \begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} \lambda_2^g + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \lambda_3^g + \begin{bmatrix} \langle \frac{\partial G}{\partial z_1}, \Gamma^G \rangle \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \langle \frac{\partial H}{\partial z_2}, \Gamma^H \rangle \\ 0 \\ 0 \end{bmatrix},
$$

for some $(\lambda^g, \Gamma^G, \Gamma^H) \in \Re^3 \times \mathcal{S}^3 \times \mathcal{S}^3$. From the above equation and (32), we obtain that $\lambda_1^g = \lambda_2^g = \frac{1}{2} > 0$, $\lambda_3^g \geq 0$. Let $\widetilde{\Gamma}^G = \overline{P}^T \Gamma^G \overline{P}$ and $\widetilde{\Gamma}^H = \overline{P}^T \Gamma^H \overline{P}$. Let (Γ^G, Γ^H) be such that all entries are zero except the entries $(\widetilde{\Gamma}_{22}^G, \widetilde{\Gamma}_{22}^H)$ left to be determined. Then (33)-(35) hold and

$$
\langle \frac{\partial G}{\partial z_1}, \Gamma^G \rangle = \langle \frac{\partial G}{\partial z_1}, \overline{P} \widetilde{\Gamma}^G \overline{P}^T \rangle = \langle \overline{P}^T \frac{\partial G}{\partial z_1} \overline{P}, \widetilde{\Gamma}^G \rangle = \widetilde{\Gamma}_{22}^G.
$$

Similarly we have $\langle \frac{\partial H}{\partial z_2}, \Gamma^H \rangle = \widetilde{\Gamma}_{22}^H$. Therefore we obtain $\widetilde{\Gamma}_{22}^G = \frac{1}{2} > 0$, $\widetilde{\Gamma}_{22}^H = -\frac{1}{2} < 0$. Since $\widetilde{\Gamma}_{22}^G \widetilde{\Gamma}_{22}^H < 0$, we know that there exists a multiplier $(\lambda^g, \Gamma^G, \Gamma^H) \in \Re^3 \times \mathcal{S}^3 \times \mathcal{S}^3$ such that (31)-(35) and (49) hold. Thus, the optimal solution $\bar{z} = (0, 0, 0, 0)$ is a C-stationary point. We now verify that the conditions (38) and (39) do not hold. Since $|\beta| = 1$, $\mathcal{O}^{|\beta|} = \{1, -1\}$. Let $\Xi_1 \in \mathcal{U}_1$ and $Q \in \{1, -1\}$. If $\beta_0 \neq \emptyset$, then it is obvious that (39) does not hold. On the other hand if $\beta_0 = \emptyset$ then $\beta = \beta_+$ or $\beta = \beta_-$. If $\beta = \beta_+$, then $\Xi_1 = [1]$ and $\Xi_2 = [0]$ and hence it is clear that the condition (38) does not hold. Alternatively if $\beta = \beta_-$, then $\Xi_1 = [0]$ and $\Xi_2 = [1]$ and hence the condition (38) does not hold. Therefore, we know that the optimal solution $\bar{z} = (0, 0, 0, 0)$ is not a M-stationary point.

8 New optimality conditions for MPCC via SDCMPCC

As we mentioned in the introduction, the vector MPCC problem (6) can be considered as a SDCMPCC problem with m one dimensional SDP complementarity constraints. Consequently, in this way, all the stationary conditions developed for SDCMPCC coincide with those for MPCC. On the other hand, the vector MPCC problem (6) can also be considered as the following SDCMPCC with one m dimensional SDP complementarity constraint:

$$
\begin{aligned}
\min \quad & f(z) \\
\text{s.t.} \quad & h(z) = 0, \\
& g(z) \leq 0, \\
& \mathcal{S}_+^m \ni \mathcal{D}(G(z)) \perp \mathcal{D}(H(z)) \in \mathcal{S}_-^m,
\end{aligned} \tag{53}
$$

where $G(z) = (G_1(z), \dots, G_m(z))^T : Z \to \Re^m$ and $H(z) = (H_1(z), \dots, H_m(z))^T : Z \to \Re^m$ and $\mathcal{D} : \Re^m \to \mathcal{S}^m$ is the linear operator defined by $\mathcal{D}(y) = \text{diag}(y)$ for any $y \in \Re^m$. We now compare the resulting S-, M- and C-stationary conditions for the two formulations. Since in this SDCMPCC reformulation the multipliers for the matrix complementarity constraints are matrices, it may provide more flexibilities and hence the resulting necessary optimality conditions may be weaker and more likely to hold at an optimal solution. We now demonstrate this point.

First, consider the S-stationary condition. It is easy to see that if a feasible point $\bar{z}$ is a S-stationary point (see e.g., [48, 63] for the definitions) of the original vector MPCC problem, then $\bar{z}$ is a S-stationary point of the special SDCMPCC problem (53). We now show that

the converse is also true. In fact, by the Definition 5.1, we know that if the feasible point $\bar{z}$ of (53) is a S-stationary point, then there exists $(\lambda^h, \lambda^g, \Gamma^G, \Gamma^H) \in \Re^p \times \Re^q \times \mathcal{S}^m \times \mathcal{S}^m$ such that (31)-(36) hold. In particular, we have

$$0 = \nabla f(\bar{z}) + h'(\bar{z})^* \lambda^h + g'(\bar{z})^* \lambda^g + G'(\bar{z})^* \mathcal{D}^*(\Gamma^G) + H'(\bar{z})^* \mathcal{D}^*(\Gamma^H),$$

where $\mathcal{D}^* : \mathcal{S}^m \to \Re^m$ is the adjoint of the linear operator $\mathcal{D}$ given by

$$\mathcal{D}^*(A) = (a_{11}, \ldots, a_{mm})^T, \quad A \in \mathcal{S}^m.$$

Denote by $\eta^G := \mathcal{D}^*(\Gamma^G) \in \Re^m$ and $\eta^H := \mathcal{D}^*(\Gamma^H) \in \Re^m$. Also, since $A = \mathcal{D}(G(\bar{z})) + \mathcal{D}(H(\bar{z}))$ is a diagonal matrix, we can just choose $\overline{P} \equiv I$ in the eigenvalue decomposition (10) of A. Therefore, by (33) and (34), we have that

$$\eta_i^G = 0 \quad \text{if } G_i(\bar{z}) > 0 \text{ and } H_i(\bar{z}) = 0,$$

$$\eta_i^H = 0 \quad \text{if } G_i(\bar{z}) = 0 \text{ and } H_i(\bar{z}) < 0.$$

Moreover, since $\Gamma_{\beta\beta}^G = \widetilde{\Gamma}_{\beta\beta}^G \preceq 0$ and $\Gamma_{\beta\beta}^H = \widetilde{\Gamma}_{\beta\beta}^H \succeq 0$, we know that the diagonal elements η^G and η^H satisfy

$$\eta_i^G \leq 0 \quad \text{and} \quad \eta_i^H \geq 0 \quad \text{if } G_i(\bar{z}) = 0 \text{ and } H_i(\bar{z}) = 0. \tag{54}$$

Therefore, we conclude that the feasible point $\bar{z}$ is also a S-stationary point of the original vector MPCC problem with the Lagrange multiplier $(\lambda^h, \lambda^g, \eta^G, \eta^H) \in \Re^p \times \Re^q \times \Re^m \times \Re^m$.

For the M- and C-stationary conditions, it is easy to check that if a feasible point $\bar{z}$ is a M- (or C-)stationary point (see e.g., [48, 63] for the definitions) of the original MPCC problem, then $\bar{z}$ is also a M- (or C-)stationary point of the SDCMPCC problem (53). However, the converse may not hold. For example, consider the following vector MPCC problem

$$\begin{aligned} \min \quad & z_1 - \tfrac{25}{8}z_2 - z_3 - \tfrac{1}{2}z_4 \\ s.t. \quad & z_4^2 \leq 0, \\ & 0 \leq G(z) \perp H(z) \leq 0, \end{aligned} \tag{55}$$

where $G : \Re^4 \to \Re^2$ and $H : \Re^4 \to \Re^2$ are defined as

$$G(z) := \begin{bmatrix} 6z_1 - z_3 - z_4 \\ z_1 \end{bmatrix} \quad \text{and} \quad H(z) := \begin{bmatrix} 6z_2 + z_3 \\ z_2 \end{bmatrix}, \quad z \in \Re^4.$$

It is easy to see that $z^* = (0,0,0,0)$ is the unqiue optimal solution of (55). By considering the weakly stationary condition (see e.g., [63] for the definition) of (55), we know that the corresponding Lagrange multiplier $(\lambda^g, \eta^G, \eta^H) \in \Re \times \Re^2 \times \Re^2$ satisfies

$$\lambda^g \geq 0, \quad \eta^G = \begin{bmatrix} -1/2 \\ 2 \end{bmatrix} \quad \text{and} \quad \eta^H = \begin{bmatrix} 1/2 \\ 1/8 \end{bmatrix}.$$

Therefore, the optimal solution $z^* = (0,0,0,0)$ is a weakly stationary point. However, by noting that $z_1^* = z_2^* = 0$, but $\eta_2^G > 0$ and $\eta_2^H > 0$, we know that z^* is neither the M-stationary point nor the C-stationary point.

Next, consider the corresponding SDCMPCC problem (53), i.e.,

$$\begin{aligned} \min \quad & z_1 - \tfrac{25}{8}z_2 - z_3 - \tfrac{1}{2}z_4 \\ s.t. \quad & z_4^2 \leq 0, \\ & \mathcal{S}_+^2 \ni \mathcal{D}(G(z)) \perp \mathcal{D}(H(z)) \in \mathcal{S}_-^2. \end{aligned} \tag{56}$$

26

We know that the Lagrange multiplier $(\lambda^g, \Gamma^G, \Gamma^H) \in \Re \times \mathcal{S}^2 \times \mathcal{S}^2$ with respect to the optimal solution z^* satisfies

$$\lambda^g \geq 0, \quad \Gamma_{11}^G = -1/2, \quad \Gamma_{22}^G = 2, \quad \Gamma_{11}^H = 1/2 \quad \text{and} \quad \Gamma_{22}^H = 1/8\,.$$

Choose

$$\Gamma^G = \begin{bmatrix} -1/2 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \Gamma^H = \begin{bmatrix} 1/2 & 1/4 \\ 1/4 & 1/8 \end{bmatrix}.$$

Let

$$Q = \begin{bmatrix} -2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \in \mathcal{O}^2\,.$$

Then, we have

$$Q^T \Gamma^G Q = \begin{bmatrix} 0 & 1 \\ 1 & 3/2 \end{bmatrix} \quad \text{and} \quad Q^T \Gamma^H Q = \begin{bmatrix} 5/8 & 0 \\ 0 & 0 \end{bmatrix}.$$

Conisder the partition $\beta_+ = \emptyset$, $\beta_0 = \{1\}$, $\beta_- = \{2\}$. Since $Q_{\beta_0}^T \Gamma^G Q_{\beta_0} = 0$ and $Q_{\beta_0}^T \Gamma^H Q_{\beta_0} = 5/8$, we know that there exist $Q \in \mathcal{O}^{|\beta|}$ and a partition $\pi(\beta) = (\beta_+, \beta_0, \beta_-)$ of β such that the Lagrange multiplier $(\lambda^g, \Gamma^G, \Gamma^H) \in \Re \times \mathcal{S}^2 \times \mathcal{S}^2$ satisfies (31)-(35) and (38)-(39). Therefore, although the optimal solution $z^* = (0,0,0,0)$ is not even a C-stationary point of the original MPCC (55), it is a M-stationary point (also a C-stationary point) of the corresponding SDCMPCC (56).

9   Acknowledgements

References

[1] J.-P. AUBIN, *Lipschitz behavior of solutions to convex minimization problems*, Mathematics of Operations Research, **9** (1984), pp. 87-111.

[2] A. BEN-TAL AND A. NEMIROVSKI. *Robust Convex Optimization*, Mathematics of Operations Research, **23** (1998), pp. 769–805.

[3] A. BEN-TAL AND A. NEMIROVSKI, *Robust convex optimization-methodology and applications*, Mathematical Programming, **92** (2002), pp. 453–480.

[4] R. BHATIA. *Matrix Analysis*, Springer-Verlag, New York, 1997.

[5] S. BI, L. HAN AND S. PAN, *Approximation of rank function and its application to the nearest low-rank correlation matrix*, Journal of Global Optimization, **57** (2013), pp. 1113–1137.

[6] J. F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[7] D. Brigo and F. Mercurio, *Calibrating LIBOR*, Risk Magazine, **15** (2002), pp. 117–122.

[8] J. P. Burge, D. G. Luenberger and D. L. Wenger, *Estimation of structured covariance matrices*, Proceedings of the IEEE, **70** (1982), pp. 963–974.

[9] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.

[10] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.

[11] R. R. E. de Gaston and M. G. Safonov, *Exact calculation of the multiloop stability margin*, IEEE Transactions on Automatic Control, **33** (1988), pp. 156–171.

[12] S. Dempe, *Foundations of Bilevel Programming*, Kluwer Academic Publishers, 2002.

[13] B. C. Eaves, *On the basic theorem for complementarity*, Mathematical Programming, **1** (1971), pp. 68–75.

[14] F. Faccchinei and J. S. Pang, *Finite-dimensional Variational Inequalities and Complementarity Problem*, Springer-Verlag, New York, 2003.

[15] M. Fazel, *Matrix Rank Minimization with Applications*, PhD thesis, Stanford University, 2002.

[16] R. Fletcher, *Semi-definite matrix constraints in optimization*, SIAM Journal on Control and Optimization, **23** (1985), pp. 493–513.

[17] M. L. Flegel and C. Kanzow, *On the Guignard constraint qualification for mathematical programs with equilibrium constraints*, Optimization, **54** (2005), pp. 517–534.

[18] K. C. Goh, J. C. Ly, M. G. Safonov, G. Papavassilopoulos and L. Turan *Biaffine matrix inequality properties and computational methods*, In: Proceeding of the American Control Conference, Baltimore, Maryland (1994), pp. 850–855.

[19] R. Henrion and J. Outrata, *On the calmness of a class of multifunctions*, SIAM Journal on Optimization, **13** (2002), pp. 603–618.

[20] R. Henrion and J. Outrata, *Calmness of constraint systems with applications*, Mathematical Programming, Seies. B, **104** (2005), pp. 437–464.

[21] B. F. Hobbs, C. B. Metzler and J. S. Pang, Strategic gaming analysis for electric power systems: an MPEC approach, *IEEE Transactions on Power Systems*, **15** (2000) 638-645.

[22] A. S. Lewis, *Nonsmooth analysis of eigenvalues*, Mathematical Programming, **84** (1999), pp. 1–24.

[23] Q. N. Li and H. D. Qi, *A sequential semismooth Newton method for the nearest low-rank correlation matrix problem*, SIAM Journal on Optimization, **21** (2011), pp. 1641–1666.

[24] F. LILLO AND R. N. MANTEGNA, *Spectral density of the correlation matrix of factor models: A random matrix theory approach*, Physical Review E, **72** (2005), pp. 016219-1–016219-10.

[25] K. LÖWNER, *Über monotone matrixfunktionen*, Mathematische Zeitschrift, **38** (1934), pp. 177–216.

[26] Z. Q. LUO, J. S. PANG AND D. RALPH, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[27] J.-B. HIRIART-URRUTY AND D. YE, *Sensitivity analysis of all eigenvalues of a symmetric matrix*, Numerische Mathematik, **70** (1995), pp. 4572.

[28] W. HOGE, *A subspace identification extension to the phase correlation method*, IEEE Transactions on Medical Imaging, **22** (2003), pp. 277–280.

[29] F. MENG, D. F. SUN AND G. Y. ZHAO, *Semismoothness of solutions to generalized equations and Moreau-Yosida regularization*, Mathematical Programming, **104** (2005), pp. 561–581.

[30] B. S. MORDUKHOVICH , *Generalized differential calculus for nonsmooth and set-valued mappings*, Journal of Mathematical Analysis and Applications, **183** (1994), pp.250–288.

[31] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 330, Springer, 2006.

[32] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, II: Applications*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 331, Springer, 2006.

[33] B. S. MORDUKHOVICH AND Y. SHAO, *Nonsmooth sequential analysis in Asplund space*, Transactions of the American Mathematical Society, **348** (1996), pp. 215–220.

[34] J. V. OUTRATA, M. KOČVARA AND J. ZOWE, *Nonsmooth Approach to Optimization Problem with Equilibrium Constraints: Theory, Application and Numerical Results*, Kluwer, Dordrecht, The Netherlands, 1998.

[35] M. OVERTON AND R. S. WOMERSLEY, *On the sum of the largest eigenvalues of a symmetric matrix*, SIAM Journal on Matrix Analysis and Applications, **13** (1992), pp. 41–45.

[36] M. OVERTON AND R. S. WOMERSLEY, *Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices*, Mathematical Programming, **62** (1993), pp. 321–357.

[37] J. S. PANG, D. F. SUN AND J. SUN, *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz cone complementarity problems*, Mathematics of Operations Research, **28** (2003), pp. 39–63.

[38] P. PSARRIS, C. A. FLOUDAS, *Robust stability analysis of linear and nonlinear systems with real parameter uncertainty, In: AIChE Annual Meeting, Miami Beach, Florida*, 1992, pp. 127e.

[39] H. D. Qi and P. Fusek, *Metric regularity and strong regularity in linear and non-linear semidefinite programming*, Technical Report, School of Mathematics, University of Southampton (2007).

[40] S. M. Robinson, *Stability theory for systems of inequalities, part I: linear systems*, SIAM Journal on Numerical Analysis, **12** (1975), pp. 754–769.

[41] S. M. Robinson, *Stability theory for systems of inequalities, part II: nonlinear systems*, SIAM Journal on Numerical Analysis, **13** (1976), pp. 473–513.

[42] S. M. Robinson, *First order conditions for general nonlinear optimization*, SIAM Journal on Applied Mathematics, **30** (1976), pp. 597–607.

[43] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Mathematical Programming Studies, **14** (1981), pp. 206–214.

[44] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[45] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[46] H. S. Ryoo, N. V. Sahinidis, *Global optimization of nonconvex NLPs and MINLPs with applications in process design*, Computers and Chemical Engineering, **19** (1995), pp. 551–566.

[47] M. G. Safonov, K. C. Goh, and J. H. Ly *Control system synthesis via bilinear matrix inequalities*, In: Proceeding of the American Control Conference, Baltimore, Maryland (1994), pp. 45–49.

[48] H. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity*, Mathematics of Operations Research, **25** (2000), pp. 1–22.

[49] D. Simon, *Reduced order kalman filtering without model reduction*, Control and Intelligent Systems, **35** (2007), pp. 169–174.

[50] D. F. Sun, *The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*, Mathematics of Operations Research, **31** (2006), pp. 761–776.

[51] D. F. Sun and J. Sun, *Semismooth matrix valued functions.* Mathematics of Operations Research, **27** (2002), pp. 150–169.

[52] D. F. Sun and J. Sun, *Strong semismoothness of eigenvalues of symmetric matrices and its applications in inverse eigenvalue problems*, SIAM Journal on Numerical Analysis, **40** (2003), pp. 2352–2367.

[53] N. K. Tsing, M. K. H. Fan, and E.I. Verriest, *On analyticity of functions involving eigenvalues*, Linear Algebra and its Applications, **207** (1994), pp. 159–180.

[54] J. G. VanAntwerp, R. D. Braatz, N. V. Sahinidis, *Globally optimal robust control for systems with nonlinear time-varying perturbations*, Computers and Chemical Engineering, **21** (1997), S125–S130.

[55] V. VISWESWARAN, C. A. FLOUDAS, *A global optimization algorithm (GOP) for certain classes of nonconvex NLPs – I. Theory*, Computers and Chemical Engineering, **14** (1990), pp. 1397–1417.

[56] V. VISWESWARAN, C. A. FLOUDAS, *A global optimization algorithm (GOP) for certain classes of nonconvex NLPs – II. Application of theory and test problems*, Computers and Chemical Engineering, **14** (1990), pp. 1419–1434.

[57] A. J. WOOD AND B. F. WOLLENBERG, *Power Generation, Operation and Control*, 2nd ed, N.Y.: J. Wiley, 1995.

[58] L. X. WU, *Fast at-the-money calibration of the LIBOR market model using Lagrange multipliers*, Journal of Computational Finance, **6** (2003), pp. 39–77.

[59] Z. WU AND J. J. YE, *First and second order condition for error bounds*, SIAM Journal on Optimization, **14** (2003), pp. 621–645.

[60] T. YAN AND M. FUKUSHIMA, *Smoothing method for mathematical programs with symmetric cone complementarity constraints*, Optimization, **60** (2011), pp. 113–128.

[61] J. J. YE, *Optimality conditions for optimization problems with complementarity constraints*, SIAM Journal on Optimization, **9** (1999), pp. 374–387.

[62] J. J. YE, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM Journal on Optimization, **10** (2000), pp. 943–962.

[63] J. J. YE, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, Journal of Mathematical Analysis and Applications, **307** (2005), pp. 305–369.

[64] J. J. YE AND X. Y. YE, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Mathematics of Operations Research, **22** (1997), pp. 977–977.

[65] J. J. YE, D. L. ZHU AND Q. J. ZHU, *Exact penalization and necessary optimality conditions for generalized bilevel programming problems*, SIAM Journal on Optimization, **7** (1997), pp. 481–507.

[66] Z. Y. ZHANG AND L. X. WU, *Optimal low-rank approximation to a correlation matrix*, Linear Algebra and Its Applications, **364** (2003), pp. 161–187.

[67] Y. B. ZHAO, *An approximation theory of matrix rank minimization and its application to quadratic equations*, Linear Algebra and Its Applications, **437** (2012), pp. 77–93.

Appendix

Proof of Proposition 2.6: Firstly, we will show that (16) holds for the case that $A = \Lambda(A)$. For any $H \in \mathcal{S}^n$, denote $Y := A + H$. Let $P \in \mathcal{O}^n$ (depending on H) be such that

$$\Lambda(A) + H = P\Lambda(Y)P^T. \tag{57}$$

Let $\delta > 0$ be any fixed number such that $0 < \delta < \frac{\lambda_{|\alpha|}}{2}$ if $\alpha \neq \emptyset$ and be any fixed positive number otherwise. Then, define the following continuous scalar function

$$f(t) := \begin{cases} t & \text{if } t > \delta, \\ 2t - \delta & \text{if } \frac{\delta}{2} < t < \delta, \\ 0 & \text{if } t < \frac{\delta}{2}. \end{cases}$$

Therefore, we have

$$\{\lambda_1(A), \ldots, \lambda_{|\alpha|}(A)\} \in (\delta, +\infty) \quad \text{and} \quad \{\lambda_{|\alpha|+1}(A), \ldots, \lambda_n(A)\} \in \left(-\infty, \frac{\delta}{2}\right).$$

For the scalar function f, let $F : \mathcal{S}^n \to \mathcal{S}^n$ be the corresponding Löwner's operator [25], i.e., for any $Z \in \mathcal{S}^n$,

$$F(Z) := \sum_{i=1}^n f(\lambda_i(Z)) u_i u_i^T, \tag{58}$$

where $U \in \mathcal{O}^n$ satisfies that $Z = U\Lambda(Z)U^T$. Since f is real analytic on the open set $(-\infty, \frac{\delta}{2}) \cup (\delta, +\infty)$, we know from [53, Theorem 3.1] that F is analytic at A. Therefore, since $A = \Lambda(A)$, it is well-known (see e.g., [4, Theorem V.3.3]) that for H sufficiently close to zero,

$$F(A + H) - F(A) - F'(A)H = O(\|H\|^2) \tag{59}$$

and

$$F'(A)H = \begin{bmatrix} H_{\alpha\alpha} & H_{\alpha\beta} & \Sigma_{\alpha\gamma} \circ H_{\alpha\gamma} \\ H_{\alpha\beta}^T & 0 & 0 \\ \Sigma_{\alpha\gamma}^T \circ H_{\alpha\gamma}^T & 0 & 0 \end{bmatrix},$$

where $\Sigma \in \mathcal{S}^n$ is given by (15) . Let $R(\cdot) := \Pi_{\mathcal{S}_+^n}(\cdot) - F(\cdot)$. By the definition of f, we know that $F(A) = A_+ := \Pi_{\mathcal{S}_+^n}(A)$, which implies that $R(A) = 0$. Meanwhile, it is clear that the matrix valued function R is directionally differentiable at A, and from (14), the directional derivative of R for any given direction $H \in \mathcal{S}^n$, is given by

$$R'(A; H) = \Pi'_{\mathcal{S}_+^n}(A; H) - F'(A)H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Pi_{\mathcal{S}_+^{|\beta|}}(H_{\beta\beta}) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{60}$$

By the Lipschitz continuity of $\lambda(\cdot)$, we know that for H sufficiently close to zero,

$$\{\lambda_1(Y), \ldots, \lambda_{|\alpha|}(Y)\} \in (\delta, +\infty), \quad \{\lambda_{|\alpha|+1}(Y), \ldots, \lambda_{|\beta|}(Y)\} \in \left(-\infty, \frac{\delta}{2}\right)$$

and

$$\{\lambda_{|\beta|+1}(Y), \ldots, \lambda_n(Y)\} \in (-\infty, 0).$$

Therefore, by the definition of F, we know that for H sufficiently close to zero,

$$R(A + H) = \Pi_{\mathcal{S}_+^n}(A + H) - F(A + H) = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & (\Lambda(Y)_{\beta\beta})_+ & 0 \\ 0 & 0 & 0 \end{bmatrix} P^T. \tag{61}$$

Since P satisfies (57), we know that for any $\mathcal{S}^n \ni H \to 0$, there exists an orthogonal matrix $Q \in \mathcal{O}^{|\beta|}$ such that

$$P_\beta = \begin{bmatrix} O(\|H\|) \\ P_{\beta\beta} \\ O(\|H\|) \end{bmatrix} \quad \text{and} \quad P_{\beta\beta} = Q + O(\|H\|^2)\,, \tag{62}$$

which was stated in [52] and was essentially proved in the derivation of Lemma 4.12 in [51]. Therefore, by noting that $(\Lambda(Y)_{\beta\beta})_+ = O(\|H\|)$, we obtain from (60), (61) and (62) that

$$\begin{aligned}
R(A+H) - R(A) - R'(A;H) &= \begin{bmatrix} O(\|H\|^3) & O(\|H\|^2) & O(\|H\|^3) \\ O(\|H\|^2) & P_{\beta\beta}(\Lambda(Y)_{\beta\beta})_+ P_{\beta\beta}^T - \Pi_{\mathcal{S}_+^{|\beta|}}(H_{\beta\beta}) & O(\|H\|^2) \\ O(\|H\|^3) & O(\|H\|^2) & O(\|H\|^3) \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & Q(\Lambda(Y)_{\beta\beta})_+ Q^T - \Pi_{\mathcal{S}_+^{|\beta|}}(H_{\beta\beta}) & 0 \\ 0 & 0 & 0 \end{bmatrix} + O(\|H\|^2)\,.
\end{aligned}$$

By (57) and (62), we know that

$$\Lambda(Y)_{\beta\beta} = P_\beta^T \Lambda(A) P_\beta + P_\beta^T H P_\beta = P_{\beta\beta}^T H_{\beta\beta} P_{\beta\beta} + O(\|H\|^2) = Q^T H_{\beta\beta} Q + O(\|H\|^2)\,.$$

Since $Q \in \mathcal{O}^{|\beta|}$, we have
$$H_{\beta\beta} = Q\Lambda(Y)_{\beta\beta} Q^T + O(\|H\|^2)\,.$$

By noting that $\Pi_{\mathcal{S}_+^{|\beta|}}(\cdot)$ is globally Lipschitz continuous and $\Pi_{\mathcal{S}_+^{|\beta|}}(Q\Lambda(Y)_{\beta\beta}Q^T) = Q(\Lambda(Y)_{\beta\beta})_+ Q^T$, we obtain that

$$Q(\Lambda(Y)_{\beta\beta})_+ Q^T - \Pi_{\mathcal{S}_+^{|\beta|}}(H_{\beta\beta}) = Q(\Lambda(Y)_{\beta\beta})_+ Q^T - \Pi_{\mathcal{S}_+^{|\beta|}}(Q\Lambda(Y)_{\beta\beta}Q^T) + O(\|H\|^2) = O(\|H\|^2)\,.$$

Therefore,
$$R(A+H) - R(A) - R'(A;H) = O(\|H\|^2)\,. \tag{63}$$

By combining (59) and (63), we know that for any $\mathcal{S}^n \ni H \to 0$,

$$\Pi_{\mathcal{S}_+^n}(\Lambda(A) + H) - \Pi_{\mathcal{S}_+^n}(\Lambda(A)) - \Pi'_{\mathcal{S}_+^n}(\Lambda(A);H) = O(\|H\|^2)\,. \tag{64}$$

Next, consider the case that $A = \overline{P}^T \Lambda(A) \overline{P}$. Re-write (57) as

$$\Lambda(A) + \overline{P}^T H \overline{P} = \overline{P}^T P \Lambda(Y) P^T \overline{P}\,.$$

Let $\widetilde{H} := \overline{P}^T H \overline{P}$. Then, we have

$$\Pi_{\mathcal{S}_+^n}(A + H) = \overline{P}\,\Pi_{\mathcal{S}_+^n}(\Lambda(A) + \widetilde{H})\overline{P}^T\,.$$

Therefore, since $\overline{P} \in \mathcal{O}^n$, we know from (64) and (14) that for any $\mathcal{S}^n \ni H \to 0$, (16) holds. ∎

Proof of Proposition 3.3: Denote the set in the righthand side of (27) by $\mathcal{N}$. We first show that $N_{\text{gph}\,N_{\mathcal{S}_+^n}}(0,0) \subseteq \mathcal{N}$. By the definition of the limiting normal cone in (8), we know that

$(U^*, V^*) \in N_{\text{gph } N_{\mathcal{S}_+^{|\beta|}}}(0,0)$ if and only if there exist two sequences $\{(U^{k^*}, V^{k^*})\}$ converging to (U^*, V^*) and $\{(U^k, V^k)\}$ converging to $(0,0)$ with $(U^{k^*}, V^{k^*}) \in N^\pi_{\text{gph } N_{\mathcal{S}_+^n}}(U^k, V^k)$ and $(U^k, V^k) \in \text{gph } N_{\mathcal{S}_+^n}$ for each k.

For each k, denote $A^k := U^k + V^k \in \mathcal{S}^n$ and let $A^k = P^k \Lambda(A^k)(P^k)^T$ with $P^k \in \mathcal{O}^n$ be the eigenvalue decomposition of A^k. Then for any $i \in \{1, \cdots, n\}$, we have

$$\lim_{k \to \infty} \lambda_i(A^k) = 0.$$

Since $\{P^k\}_{k=1}^\infty$ is uniformly bounded, by taking a subsequence if necessary, we may assume that $\{P^k\}_{k=1}^\infty$ converges to an orthogonal matrix $Q := \lim_{k \to \infty} P^k \in \mathcal{O}^n$. For each k, we know that the vector $\lambda(A^k)$ is an element of $\Re_{\gtrless}^n$. By taking a subsequence if necessary, we may assume that for each k, $\Lambda(A^k)$ has the same form, i.e.,

$$\Lambda(A^k) = \begin{bmatrix} \Lambda(A^k)_{\beta_+ \beta_+} & 0 & 0 \\ 0 & \Lambda(A^k)_{\beta_0 \beta_0} & 0 \\ 0 & 0 & \Lambda(A^k)_{\beta_- \beta_-} \end{bmatrix},$$

where β_+, β_0 and β_- are the three index sets defined by

$$\beta_+ := \{i : \lambda_i(A^k) > 0\}, \quad \beta_0 := \{i : \lambda_i(A^k) = 0\} \quad \text{and} \quad \beta_- := \{i : \lambda_i(A^k) < 0\}.$$

Since $(U^{k^*}, V^{k^*}) \in N^\pi_{\text{gph } N_{\mathcal{S}_+^n}}(U^k, V^k)$, we know from Proposition 3.2 that for each k, there exist

$$\Theta_1^k = \begin{bmatrix} E_{\beta_+ \beta_+} & E_{\beta_+ \beta_0} & \Sigma_{\beta_+ \beta_-}^k \\ E_{\beta_+ \beta_0}^T & 0 & 0 \\ (\Sigma_{\beta_+ \beta_-}^k)^T & 0 & 0 \end{bmatrix}$$

and

$$\Theta_2^k = \begin{bmatrix} 0 & 0 & E_{\beta_+ \beta_-} - \Sigma_{\beta_+ \beta_-}^k \\ 0 & 0 & E_{\beta_0 \beta_-} \\ (E_{\beta_+ \beta_-} - \Sigma_{\beta_+ \beta_-}^k)^T & (E_{\beta_0 \beta_-})^T & E_{\beta_- \beta_-} \end{bmatrix}$$

such that

$$\Theta_1^k \circ \widetilde{U^*}^k + \Theta_2^k \circ \widetilde{V^*}^k = 0, \quad \widetilde{U^*}^*_{\beta_0 \beta_0} \preceq 0 \quad \text{and} \quad \widetilde{V^*}^*_{\beta_0 \beta_0} \succeq 0, \tag{65}$$

where $\widetilde{U^*}^k = (P^k)^T U^{k^*} P^k$, $\widetilde{V^*}^k = (P^k)^T V^{k^*} P^k$ and

$$(\Sigma^k)_{i,j} = \frac{\max\{\lambda_i(A^k), 0\} - \max\{\lambda_j(A^k), 0\}}{\lambda_i(A^k) - \lambda_j(A^k)} \quad \forall\, (i,j) \in \beta_+ \times \beta_-. \tag{66}$$

Since for each k, each element of $\Sigma_{\beta_+ \beta_-}^k$ belongs to the interval $[0,1]$, by further taking a subsequence if necessary, we may assume that the limit of $\{\Sigma_{\beta_+ \beta_-}^k\}_{k=1}^\infty$ exists. Therefore, by the definition of $\mathcal{U}_n$ in (24), we know that

$$\lim_{k \to \infty} \Theta_1^k = \Xi_1 \in \mathcal{U}_n \quad \text{and} \quad \lim_{k \to \infty} \Theta_2^k = \Xi_2,$$

where Ξ_1 and Ξ_2 are given by (26). Therefore, we obtain from (65) that $(U^*, V^*) \in \mathcal{N}$.

The other direction, i.e., $N_{\text{gph } N_{\mathcal{S}_+^n}}(0,0) \supseteq \mathcal{N}$ can be proved in a similar but simpler way to that of the second part of Theorem 3.1. We omit it here. $\blacksquare$

Proof of Theorem 3.1: "⟹" Suppose that $(X^*, Y^*) \in N_{\mathrm{gph}\, N_{\mathcal{S}_+^n}}(X, Y)$. By the definition of the limiting normal cone in (8), we know that $(X^*, Y^*) = \lim\limits_{k \to \infty}(X^{k^*}, Y^{k^*})$ with

$$(X^{k^*}, Y^{k^*}) \in N^\pi_{\mathrm{gph}\, N_{\mathcal{S}_+^n}}(X^k, Y^k) \quad k = 1, 2, \dots,$$

where $(X^k, Y^k) \to (X, Y)$ and $(X^k, Y^k) \in \mathrm{gph}\, N_{\mathcal{S}_+^n}$. For each k, denote $A^k := X^k + Y^k$ and let $A^k = P^k \Lambda(A^k)(P^k)^T$ be the eigenvalue decomposition of A^k. Since $\Lambda(A) = \lim\limits_{k\to\infty}\Lambda(A^k)$, we know that $\Lambda(A^k)_{\alpha\alpha} \succ 0$, $\Lambda(A^k)_{\gamma\gamma} \prec 0$ for k sufficiently large and $\lim\limits_{k\to\infty}\Lambda(A^k)_{\beta\beta} = 0$. Since $\{P^k\}_{k=1}^\infty$ is uniformly bounded, by taking a subsequence if necessary, we may assume that $\{P^k\}_{k=1}^\infty$ converges to an orthogonal matrix $\widehat{P} \in \mathcal{O}^n(A)$. We can write $\widehat{P} = [\overline{P}_\alpha \;\; \overline{P}_\beta Q \;\; \overline{P}_\gamma]$, where $Q \in \mathcal{O}^{|\beta|}$ can be any $|\beta| \times |\beta|$ orthogonal matrix. By further taking a subsequence if necessary, we may also assume that there exists a partition $\pi(\beta) = (\beta_+, \beta_0, \beta_-)$ of β such that for each k,

$$\lambda_i(A^k) > 0 \quad \forall i \in \beta_+, \quad \lambda_i(A^k) = 0 \quad \forall i \in \beta_0 \quad \text{and} \quad \lambda_i(A^k) < 0 \quad \forall i \in \beta_-.$$

This implies that for each k,

$$\{i : \lambda_i(A^k) > 0\} = \alpha \cup \beta_+, \quad \{i : \lambda_i(A^k) = 0\} = \beta_0 \quad \text{and} \quad \{i : \lambda_i(A^k) < 0\} = \beta_- \cup \gamma.$$

Then, for each k, since $(X^{k^*}, Y^{k^*}) \in N^\pi_{\mathrm{gph}\, N_{\mathcal{S}_+^n}}(X^k, Y^k)$, we know from Proposition 3.2 that there exist

$$\Theta_1^k = \begin{bmatrix} E_{\alpha\alpha} & E_{\alpha\beta_+} & E_{\alpha\beta_0} & \Sigma_{\alpha\beta_-}^k & \Sigma_{\alpha\gamma}^k \\ E_{\alpha\beta_+}^T & E_{\beta_+\beta_+} & E_{\beta_+\beta_0} & \Sigma_{\beta_+\beta_-}^k & \Sigma_{\beta_+\gamma}^k \\ E_{\alpha\beta_0}^T & E_{\beta_+\beta_0}^T & 0 & 0 & 0 \\ {\Sigma_{\alpha\beta_-}^k}^T & {\Sigma_{\beta_+\beta_-}^k}^T & 0 & 0 & 0 \\ {\Sigma_{\alpha\gamma}^k}^T & {\Sigma_{\beta_+\gamma}^k}^T & 0 & 0 & 0 \end{bmatrix}$$

and

$$\Theta_2^k = \begin{bmatrix} 0 & 0 & 0 & E_{\alpha\beta_-} - \Sigma_{\alpha\beta_-}^k & E_{\alpha\gamma} - \Sigma_{\alpha\gamma}^k \\ 0 & 0 & 0 & E_{\beta_+\beta_-} - \Sigma_{\beta_+\beta_-}^k & E_{\beta_+\gamma} - \Sigma_{\beta_+\gamma}^k \\ 0 & 0 & 0 & E_{\beta_0\beta_-} & E_{\beta_0\gamma} \\ (E_{\alpha\beta_-} - \Sigma_{\alpha\beta_-}^k)^T & (E_{\beta_+\beta_-} - \Sigma_{\beta_+\beta_-}^k)^T & E_{\beta_0\beta_-}^T & E_{\beta_-\beta_-} & E_{\beta_-\gamma} \\ (E_{\alpha\gamma} - \Sigma_{\alpha\gamma}^k)^T & (E_{\beta_+\gamma} - \Sigma_{\beta_+\gamma}^k)^T & E_{\beta_0\gamma}^T & E_{\beta_-\gamma}^T & E_{\gamma\gamma} \end{bmatrix}$$

such that
$$\Theta_1^k \circ \widetilde{X^k}^* + \Theta_2^k \circ \widetilde{Y^k}^* = 0, \quad \widetilde{X^k}^*_{\beta_0\beta_0} \preceq 0 \quad \text{and} \quad \widetilde{Y^k}^*_{\beta_0\beta_0} \succeq 0, \tag{67}$$

where $\widetilde{X^k}^* = (P^k)^T X^{k^*} P^k$, $\widetilde{Y^k}^* = (P^k)^T Y^{k^*} P^k$ and

$$(\Sigma^k)_{i,j} = \frac{\max\{\lambda_i(A^k), 0\} - \max\{\lambda_j(A^k), 0\}}{\lambda_i(A^k) - \lambda_j(A^k)} \quad \forall (i,j) \in (\alpha \cup \beta_+) \times (\beta_- \cup \gamma). \tag{68}$$

By taking limits as $k \to \infty$, we obtain that

$$\widetilde{X^k}^* \to \widehat{P}^T X^* \widehat{P} = \begin{bmatrix} \widetilde{X}_{\alpha\alpha}^* & \widetilde{X}_{\alpha\beta}^* Q & \widetilde{X}_{\alpha\gamma}^* \\ (\widetilde{X}_{\alpha\beta}^* Q)^T & Q^T \widetilde{X}_{\beta\beta}^* Q & Q^T \widetilde{X}_{\beta\gamma}^* \\ (\widetilde{X}_{\alpha\gamma}^*)^T & (Q^T \widetilde{X}_{\beta\gamma}^*)^T & \widetilde{X}_{\gamma\gamma}^* \end{bmatrix}$$

35

and

$$\widetilde{Y^k}^* \to \widehat{P}^T Y^* \widehat{P} = \begin{bmatrix} \widetilde{Y}^*_{\alpha\alpha} & \widetilde{Y}^*_{\alpha\beta}Q & \widetilde{Y}^*_{\alpha\gamma} \\ (\widetilde{Y}^*_{\alpha\beta}Q)^T & Q^T\widetilde{Y}^*_{\beta\beta}Q & Q^T\widetilde{Y}^*_{\beta\gamma} \\ (\widetilde{Y}^*_{\alpha\gamma})^T & (Q^T\widetilde{Y}^*_{\beta\gamma})^T & \widetilde{Y}_{\gamma\gamma} \end{bmatrix}.$$

By simple calculations, we obtain from (68) that

$$\lim_{k\to\infty} \Sigma^k_{\alpha\beta_-} = E_{\alpha\beta_-}, \quad \lim_{k\to\infty} \Sigma^k_{\beta_+\gamma} = 0 \quad \text{and} \quad \lim_{k\to\infty} \Sigma^k_{\alpha\gamma} = \Sigma_{\alpha\gamma}.$$

This, together with the definition of $\mathcal{U}_{|\beta|}$, shows that there exist $\Xi_1 \in \mathcal{U}_{|\beta|}$ and the corresponding Ξ_2 such that

$$\lim_{k\to\infty} \Theta^k_1 = \begin{bmatrix} E_{\alpha\alpha} & E_{\alpha\beta} & \Sigma_{\alpha\gamma} \\ E_{\beta\alpha} & \Xi_1 & 0 \\ \Sigma^T_{\alpha\gamma} & 0 & 0 \end{bmatrix} = \Theta_1 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Xi_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$\lim_{k\to\infty} \Theta^k_2 = \begin{bmatrix} 0 & 0 & E_{\alpha\gamma} - \Sigma_{\alpha\gamma} \\ 0 & \Xi_2 & E_{\beta\gamma} \\ (E_{\alpha\gamma} - \Sigma_{\alpha\gamma})^T & E_{\gamma\beta} & E_{\gamma\gamma} \end{bmatrix} = \Theta_2 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Xi_2 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where Θ_1 and Θ_2 are given by (22). Meanwhile, since $Q \in \mathcal{O}^{|\beta|}$, by taking limits in (67) as $k \to \infty$, we obtain that

$$\Theta_1 \circ \widetilde{X}^* + \Theta_2 \circ \widetilde{Y}^* = 0, \quad \Xi_1 \circ Q^T \widetilde{X}^*_{\beta\beta}Q + \Xi_2 \circ Q^T \widetilde{Y}^*_{\beta\beta}Q = 0 \tag{69}$$

and

$$Q^T_{\beta_0} \widetilde{X}^*_{\beta\beta} Q_{\beta_0} \preceq 0 \quad \text{and} \quad Q^T_{\beta_0} \widetilde{Y}^*_{\beta\beta} Q_{\beta_0} \succeq 0.$$

Hence, by Proposition 3.3, we conclude that $(\widetilde{X}^*_{\beta\beta}, \widetilde{Y}^*_{\beta\beta}) \in N_{\mathrm{gph}\, N_{\mathcal{S}^{|\beta|}_+}}(0,0)$. From (69), it is easy to check that (X^*, Y^*) satisfies the conditions (28) and (29).

"$\Longleftarrow$" Let (X^*, Y^*) satisfies (28) and (29). We shall show that there exist two sequences $\{(X^k, Y^k)\}$ converging to (X, Y) and $\{(X^{k^*}, Y^{k^*})\}$ converging to (X^*, Y^*) with $(X^k, Y^k) \in \mathrm{gph}\, N_{\mathcal{S}^n_+}$ and $(X^{k^*}, Y^{k^*}) \in N^\pi_{\mathrm{gph}\, N_{\mathcal{S}^n_+}}(X^k, Y^k)$ for each k.

Since $(\widetilde{X}^*_{\beta\beta}, \widetilde{Y}^*_{\beta\beta}) \in N_{\mathrm{gph}\, N_{\mathcal{S}^{|\beta|}_+}}(0,0)$, by Proposition 3.3, we know that there exist an orthogonal matrix $Q \in \mathcal{O}^{|\beta|}$ and $\Xi_1 \in \mathcal{U}_{|\beta|}$ such that

$$\Xi_1 \circ Q^T \widetilde{X}^*_{\beta\beta}Q + \Xi_2 \circ Q^T \widetilde{Y}^*_{\beta\beta}Q = 0, \quad Q^T_{\beta_0} \widetilde{X}^*_{\beta\beta} Q_{\beta_0} \preceq 0 \quad \text{and} \quad Q^T_{\beta_0} \widetilde{Y}^*_{\beta\beta} Q_{\beta_0} \succeq 0. \tag{70}$$

Since $\Xi_1 \in \mathcal{U}_{|\beta|}$, we know that there exists a sequence $\{z^k\} \in \Re^{|\beta|}_{\gtrsim}$ converging to 0 such that $\Xi_1 = \lim_{k\to\infty} D(z^k)$. Without loss of generality, we can assume that there exists a partition $\pi(\beta) = (\beta_+, \beta_0, \beta_-) \in \mathscr{P}(\beta)$ such that for all k,

$$z^k_i > 0 \quad \forall i \in \beta_+, \quad z^k_i = 0 \quad \forall i \in \beta_0 \quad \text{and} \quad z^k_i < 0 \quad \forall i \in \beta_-.$$

For each k, let

$$X^k = \widehat{P} \begin{bmatrix} \Lambda(A)_{\alpha\alpha} & 0 & 0 & 0 & 0 \\ 0 & (z^k)_+ & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \widehat{P}^T \quad \text{and} \quad Y^k = \widehat{P} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (z^k)_- & 0 \\ 0 & 0 & 0 & 0 & \Lambda(A)_{\gamma\gamma} \end{bmatrix} \widehat{P}^T,$$

36

where $\widehat{P} = \begin{bmatrix} \overline{P}_\alpha & \overline{P}_\beta Q & \overline{P}_\gamma \end{bmatrix} \in \mathcal{O}^n(A)$. Then, it is clear that $\{(X^k, Y^k)\} \in \mathrm{gph}\, N_{\mathcal{S}^n_+}$ converging to (X, Y). For each k, denote

$$A^k = X^k + Y^k, \quad \Theta_1^k = \begin{bmatrix} E_{\alpha\alpha} & E_{\alpha\beta_+} & E_{\alpha\beta_0} & \Sigma^k_{\alpha\beta_-} & \Sigma_{\alpha\gamma} \\ E^T_{\alpha\beta_+} & E_{\beta_+\beta_+} & E_{\beta_+\beta_0} & \Sigma^k_{\beta_+\beta_-} & \Sigma^k_{\beta_+\gamma} \\ E^T_{\alpha\beta_0} & E^T_{\beta_+\beta_0} & 0 & 0 & 0 \\ (\Sigma^k_{\alpha\beta_-})^T & (\Sigma^k_{\beta_+\beta_-})^T & 0 & 0 & 0 \\ (\Sigma_{\alpha\gamma})^T & (\Sigma^k_{\beta_+\gamma})^T & 0 & 0 & 0 \end{bmatrix}$$

and

$$\Theta_2^k = \begin{bmatrix} 0 & 0 & 0 & E_{\alpha\beta_-} - \Sigma^k_{\alpha\beta_-} & E_{\alpha\gamma} - \Sigma_{\alpha\gamma} \\ 0 & 0 & 0 & E_{\beta_+\beta_-} - \Sigma^k_{\beta_+\beta_-} & E_{\beta_+\gamma} - \Sigma^k_{\beta_+\gamma} \\ 0 & 0 & 0 & E_{\beta_0\beta_-} & E_{\beta_0\gamma} \\ (E_{\alpha\beta_-} - \Sigma^k_{\alpha\beta_-})^T & (E_{\beta_+\beta_-} - \Sigma^k_{\beta_+\beta_-})^T & E^T_{\beta_0\beta_-} & E_{\beta_-\beta_-} & E_{\beta_-\gamma} \\ (E_{\alpha\gamma} - \Sigma_{\alpha\gamma})^T & (E_{\beta_+\gamma} - \Sigma^k_{\beta_+\gamma})^T & E^T_{\beta_0\gamma} & E^T_{\beta_-\gamma} & E_{\gamma\gamma} \end{bmatrix},$$

where

$$(\Sigma^k)_{i,j} = \frac{\max\{\lambda_i(A^k)), 0\} - \max\{\lambda_j(A^k)), 0\}}{\lambda_i(A^k) - \lambda_j(A^k)} \quad \forall\, (i,j) \in (\alpha \cup \beta_+) \times (\beta_- \cup \gamma).$$

Next, for each k, we define two matrices $\widehat{X^k}^*, \widehat{Y^k}^* \in \mathcal{S}^n$. Let $i, j \in \{1, \ldots, n\}$. If (i,j) and $(j,i) \notin (\alpha \times \beta_-) \cup (\beta_+ \times \gamma) \cup (\beta \times \beta)$. We define

$$\widehat{X^k}^*_{i,j} \equiv \widetilde{X}^*_{i,j}, \quad \widehat{Y^k}^*_{i,j} \equiv \widetilde{Y}^*_{i,j}, \quad k = 1, 2, \ldots. \tag{71}$$

Otherwise, denote $c^k := (\Sigma^k)_{i,j}$, $k = 1, 2, \ldots$. We consider the following four cases.

Case 1: (i,j) or $(j,i) \in \alpha \times \beta_-$. In this case, we know from (28) that $\widetilde{X}^*_{i,j} = 0$. Since $c_k \neq 0$ for all k and $c^k \to 1$ as $k \to \infty$, we define

$$\widehat{Y^k}^*_{i,j} \equiv \widetilde{Y}^*_{i,j} \quad \text{and} \quad \widehat{X^k}^*_{i,j} = \frac{c^k - 1}{c^k} \widehat{Y^k}^*_{i,j}, \quad k = 1, 2, \ldots. \tag{72}$$

Then, we have

$$c_k \widehat{X^k}^*_{i,j} + (1 - c_k) \widehat{Y^k}^*_{i,j} = 0 \quad \forall\, k \quad \text{and} \quad (\widehat{X^k}^*_{i,j}, \widehat{Y^k}^*_{i,j}) \to (\widetilde{X}^*_{i,j}, \widetilde{Y}^*_{i,j}) \quad \text{as } k \to \infty.$$

Case 2: (i,j) or $(j,i) \in \beta_+ \times \gamma$. In this case, we know from (28) that $\widetilde{Y}^*_{i,j} = 0$. Since $c_k \neq 1$ for all k and $c^k \to 0$ as $k \to \infty$, we define

$$\widehat{X^k}^*_{i,j} \equiv \widetilde{X}^*_{i,j} \quad \text{and} \quad \widehat{Y^k}^*_{i,j} = \frac{c^k}{c^k - 1} \widehat{X^k}^*_{i,j}, \quad k = 1, 2, \ldots. \tag{73}$$

Then, we know that

$$c_k \widehat{X^k}^*_{i,j} + (1 - c_k) \widehat{Y^k}^*_{i,j} = 0 \quad \forall\, k \quad \text{and} \quad (\widehat{X^k}^*_{i,j}, \widehat{Y^k}^*_{i,j}) \to (\widetilde{X}^*_{i,j}, \widetilde{Y}^*_{i,j}) \quad \text{as } k \to \infty.$$

Case 3: (i,j) or $(j,i) \in (\beta \times \beta) \setminus (\beta_+ \times \beta_-)$. In this case, we define

$$\widehat{X^k}^*_{i,j} \equiv Q_i^T \widetilde{X}^*_{\beta\beta} Q_j, \quad \widehat{Y^k}^*_{i,j} \equiv Q_i^T \widetilde{Y}^*_{\beta\beta} Q_j, \quad k = 1, 2, \ldots. \tag{74}$$

Case 4: (i,j) or $(j,i) \in \beta_+ \times \beta_-$. Since $c \in [0,1]$, we consider the following two sub-cases:

Case 4.1: $c \neq 1$. Since $c_k \neq 1$ for all k large enough, we define

$$\widehat{X^k}^*_{i,j} \equiv Q_i^T \widetilde{X}^*_{\beta\beta} Q_j \quad \text{and} \quad \widehat{Y^k}^*_{i,j} = \frac{c^k}{c^k - 1} \widehat{X^k}^*_{i,j}, \quad k = 1,2,\ldots. \tag{75}$$

Then, from (70), we know that

$$\widehat{Y^k}^*_{i,j} \to \frac{c}{c-1} Q_i^T \widetilde{X}^*_{\beta\beta} Q_j = Q_i^T \widetilde{Y}^*_{\beta\beta} Q_j \quad \text{as } k \to \infty.$$

Case 4.2: $c = 1$. Since $c_k \neq 0$ for all k large enough, we define

$$\widehat{Y^k}^*_{i,j} \equiv Q_i^T \widetilde{Y}^*_{\beta\beta} Q_j \quad \text{and} \quad \widehat{X^k}^*_{i,j} = \frac{c^k - 1}{c^k} \widehat{Y^k}^*_{i,j}, \quad k = 1,2,\ldots. \tag{76}$$

Then, again from (70), we know that

$$\widehat{X^k}^*_{i,j} \to \frac{c-1}{c} Q_i^T \widetilde{Y}^*_{\beta\beta} Q_j = Q_i^T \widetilde{X}^*_{\beta\beta} Q_j \quad \text{as } k \to \infty.$$

For each k, define $X^{k*} = \widehat{P} \widehat{X^k}^* \widehat{P}^T$ and $Y^{k*} = \widehat{P} \widehat{Y^k}^* \widehat{P}^T$. Then, from (71)-(76) we obtain that

$$\Theta_1^k \circ \widehat{P}^T X^{k*} \widehat{P} + \Theta_2^k \circ \widehat{P}^T Y^{k*} \widehat{P} = 0, \quad k = 1,2,\ldots.$$

and

$$(\widehat{P}^T X^{k*} \widehat{P}, \widehat{P}^T Y^{k*} \widehat{P}) \to (\widehat{P}^T X^* \widehat{P}, \widehat{P}^T Y^* \widehat{P}) \quad \text{as } k \to \infty. \tag{77}$$

Moreover, from (74) and (70), we have

$$Q_{\beta_0}^T \widetilde{X^k}^*_{\beta\beta} Q_{\beta_0} \equiv Q_{\beta_0}^T \widetilde{X}^*_{\beta\beta} Q_{\beta_0} \preceq 0 \quad \text{and} \quad Q_{\beta_0}^T \widetilde{Y^k}^*_{\beta\beta} Q_{\beta_0} \equiv Q_{\beta_0}^T \widetilde{Y}^*_{\beta\beta} Q_{\beta_0} \succeq 0, \quad k = 1,2,\ldots.$$

From Proposition 3.2 and (77), we know that

$$(X^{k*}, Y^{k*}) \in N^\pi_{\text{gph } N_{\mathcal{S}^n_+}}(X^k, Y^k) \quad \text{and} \quad (X^*, Y^*) = \lim_{k \to \infty} (X^{k*}, Y^{k*}).$$

Hence, the assertion of the theorem follows. ∎