MATHEMATICS OF OPERATIONS RESEARCH
Vol. 31, No. 4, November 2006, pp. 761–776
ISSN 0364-765X | EISSN 1526-5471 |06|3104|0761
DOI 10.1287/moor.xxxx.xxxx

The Strong Second Order Sufficient Condition and Constraint Nondegeneracy in Nonlinear Semidefinite Programming and Their Implications

Defeng Sun
Department of Mathematics, National University of Singapore, Singapore
email: matsundf@nus.edu.sg http://www.math.nus.edu.sg/~matsundf/

For a locally optimal solution to the nonlinear semidefinite programming problem, under Robinson's constraint qualification, the following conditions are proved to be equivalent: the strong second order sufficient condition and constraint nondegeneracy; the nonsingularity of Clarke's Jacobian of the Karush-Kuhn-Tucker system; the strong regularity of the Karush-Kuhn-Tucker point; and others.

Key words: nonlinear semidefinite programming; strong second order sufficient condition; constraint nondegeneracy; strong regularity.

MSC2000 Subject Classification: Primary: 90C31, 65K10; Secondary: 90C26, 90C22

OR/MS subject classification: Primary: Programming/nonlinear

History: Received: May 16, 2005; Revised: November 21, 2005.

---

1. Introduction

Consider the optimization problem

$$(OP) \qquad \begin{array}{ll} \min\limits_{x\in X} & f(x) \\ \text{s.t.} & G(x) \in K, \end{array} \tag{1}$$

where $f : X \to \Re$ and $G : X \to Y$ are twice continuously differentiable functions, X and Y are two finite dimensional real vector spaces each equipped with a scalar product denoted by $\langle\cdot,\cdot\rangle$ and its induced norm denoted by $\|\cdot\|$, and K is a closed convex set in Y. We reserve Z to represent an arbitrary real vector space with a scalar product $\langle\cdot,\cdot\rangle$. We denote by $\mathcal{J}_x f(x)$ and $\mathcal{J}^2_{xx} f(x)$ the derivative and the second order derivative of f with respect to $x \in X$, respectively. For any given linear operator A, we denote its adjoint by A^*. The first order optimality condition, namely the Karush-Kuhn-Tucker (KKT) condition, for (OP) takes the following form:

$$\mathcal{J}_x L(x,\mu) = 0 \quad \text{and} \quad \mu \in \mathcal{N}_K(G(x)), \tag{2}$$

where the Lagrangian function $L : X \times Y \to \Re$ is defined by

$$L(x,\mu) := f(x) + \langle \mu, G(x)\rangle\,, \quad (x,\mu) \in X \times Y, \tag{3}$$

$\mathcal{J}_x L(x,\mu)$ is the derivative of $L(x,\mu)$ at (x,μ) with respect to $x \in X$, and $\mathcal{N}_K(y)$ denotes the *normal cone* of K at y in the sense of convex analysis (Rockafellar [35]):

$$\mathcal{N}_K(y) = \begin{cases} \{d \in Y \,:\, \langle d, z-y\rangle \le 0 \ \ \forall\, z \in K\} & \text{if } y \in K, \\ \emptyset & \text{if } y \notin K. \end{cases}$$

For any (x,μ) satisfying (2), we call x a *stationary point* and (x,μ) a *KKT point* of (OP), respectively.

During the last three decades, tremendous progress has been achieved towards sensitivity and stability analysis of solutions to the optimization problem (OP) subject to data perturbation (Bonnans and Shapiro [6], Facchinei and Pang [12], Klatte and Kummer [18], Rockafellar and Wets [36]). When K is a polyhedral set, the corresponding theory is quite complete. This is especially the case for the conventional nonlinear programming

(NLP)

$$\begin{array}{ll} \min\limits_{x\in X} & f(x) \\ \text{s.t.} & h(x) = 0, \\ & g(x) \leq 0, \end{array} \tag{4}$$

where $f : X \to \Re$, $h : X \to \Re^m$, and $g : X \to \Re^p$ are twice continuously differentiable functions.

For the case that K is a general nonpolyhedral set, much less has been known. However, when K is $\mathcal{C}^2$-*cone reducible* in the sense of Bonnans and Shapiro [6, Definition 3.135], the full picture of sensitivity and stability of solutions for problem (OP) is emerging (Bonnans et al. [1, 2], Bonnans and Shapiro [5, 6]). The class of $\mathcal{C}^2$-cone reducible sets is rich. It includes notably the polyhedral set, the second order cone (ice-cream cone or Lorentz cone), the cone of symmetric positive semidefinite matrices, and their Cartesian product (Bonnans and Shapiro [6], Shapiro [39]).

Compared to the conventional nonlinear programming (NLP), the theory for (OP) with $\mathcal{C}^2$-cone reducible sets is evolving and yet to be completed. Let $\bar{x}$ be a feasible solution to (OP). Robinson's constraint qualification (CQ) (Robinson [29]) is said to hold at $\bar{x}$ if

$$0 \in \text{int}\{G(\bar{x}) + \mathcal{J}_x G(\bar{x})X - K\}, \tag{5}$$

where "int" denotes the topological interior part of a given set. If $\bar{x}$ is a locally optimal solution to (OP) and Robinson's CQ holds at $\bar{x}$, then there exists a Lagrangian multiplier $\bar{\mu} \in Y$, together with $\bar{x}$, satisfying the KKT condition:

$$\mathcal{J}_x L(\bar{x}, \bar{\mu}) = 0 \quad \text{and} \quad \bar{\mu} \in \mathcal{N}_K(G(\bar{x})). \tag{6}$$

For any closed (not necessary convex) set $D \subseteq Y$ and $y \in Y$, denote

$$\text{dist}(y, D) := \inf\{\|y - d\| \,:\, d \in D\}.$$

For any closed set $D \subseteq Y$, we write $\mathcal{T}_D^i(y)$ and $\mathcal{T}_D(y)$ for the *inner tangent cone* and the *contingent (Bouligand) cone* of D at y, respectively. That is,

$$\mathcal{T}_D^i(y) = \{d \in Y \,:\, \text{dist}(y + td, D) = o(t)\,,\, t \geq 0\}$$

and

$$\mathcal{T}_D(y) = \{d \in Y \,:\, \exists\, t_k \downarrow 0,\ \text{dist}(y + t_k d, D) = o(t_k)\}.$$

When D is a closed convex set, the inner tangent cone and the contingent cone are equal:

$$\mathcal{T}_D(y) = \mathcal{T}_D^i(y) = \{d \in Y \,:\, \text{dist}(y + td, D) = o(t)\,,\, t \geq 0\}, \quad y \in D$$

and will be simply called the *tangent cone* of D at y. Since Y is assumed to be a finite dimensional space and K is a closed convex set, Robinson's CQ (5) can be equivalently written as

$$\mathcal{J}_x G(\bar{x})X + \mathcal{T}_K(G(\bar{x})) = Y, \tag{7}$$

which reduces to the well known Mangasarian-Fromovitz constraint qualification (MFCQ) for the conventional nonlinear programming (NLP) (Mangasarian and Fromovitz [22]):

$$\left\{ \begin{array}{ll} \mathcal{J}_x h_i(\bar{x}), & i = 1, \ldots, m, \text{ are linearly independent,} \\ \exists\, d \in X : & \mathcal{J}_x h_i(\bar{x})d = 0\,,\ i = 1, \ldots, m\,,\ \mathcal{J}_x g_j(\bar{x})d < 0\,,\ j \in \mathcal{I}(\bar{x}), \end{array} \right. \tag{8}$$

where the active set $\mathcal{I}(\bar{x})$ of $g(\cdot)$ at $\bar{x}$ is defined by

$$\mathcal{I}(\bar{x}) := \{j \,:\, g_j(\bar{x}) = 0\,,\ j = 1, \ldots, p\}.$$

For a proof on this equivalence, see Robinson [28, Theorem 3]. A stronger notion than the MFCQ in (NLP) is the linear independence constraint qualification (LICQ):

$$\{\mathcal{J}_x h_i(\bar{x})\}_{i=1}^m \text{ and } \{\mathcal{J}_x g_j(\bar{x})\}_{j\in\mathcal{I}(\bar{x})} \text{ are linearly independent.} \tag{9}$$

Let $\mathcal{M}(\bar{x})$ denote the set of Lagrangian multipliers satisfying (6). Then $\mathcal{M}(\bar{x})$ is nonempty and bounded if and only if Robinson's CQ holds at $\bar{x}$ (Bonnans and Shapiro [6, Theorem 3.9 and Proposition 3.17]), which generalizes an analogous assertion for (NLP): $\mathcal{M}(\bar{x})$ is nonempty and bounded if and only if the MFCQ holds $\bar{x}$ (cf., Gauvin [13]). For (NLP), the LICQ implies that $\mathcal{M}(\bar{x})$ is a singleton.

In one of his seminal papers, Robinson [30] introduced the important concept of *strong regularity* for generalized equations, which include the KKT system (6) as a special case, and defined a *strong second* *order sufficient condition* for (NLP). He also showed that the strong second order sufficient condition and the LICQ imply the strong regularity of the solution to the KKT system (6). Interestingly, the converse is also true, see Jongen et al. [16], Bonnans and Sulem [7], Dontchev and Rockafellar [10], and Bonnans and Shapiro [6, Proposition 5.38].

The primary objective of this paper is to build up the connections between the strong second order sufficient condition and strong regularity for the nonlinear semidefinite programming

$$(NLSDP) \qquad \begin{array}{ll} \min\limits_{x\in X} & f(x) \\ \text{s.t.} & h(x) = 0, \\ & g(x) \in \mathcal{S}^p_+, \end{array} \tag{10}$$

where $f : X \to \Re$, $h : X \to \Re^m$, and $g : X \to \mathcal{S}^p$ are twice continuously differentiable, $\mathcal{S}^p$ is the linear space of all $p \times p$ real symmetric matrices, and $\mathcal{S}^p_+$ is the cone of all $p \times p$ positive semidefinite matrices. Problem $(NLSDP)$ is a special case of (OP) with

$$G(x) := (h(x), g(x)), \quad x \in X, \quad Y := \Re^m \times \mathcal{S}^p, \quad \text{and} \quad K := \{0\} \times \mathcal{S}^p_+. \tag{11}$$

We achieve this objective via the study of the nonsingularity of generalized Jacobian of the system of nonsmooth equations reformulated from (2). Consequently, we show that if $\bar{x}$ is a locally optimal solution to $(NLSDP)$ and Robinson's CQ holds at $\bar{x}$, then the nonsingularity of Clarke's Jacobian of the corresponding nonsmooth system is not only sufficient but also necessary for the strong regularity. Since the nonsingularity of Clarke's Jacobian is a stronger condition than many other conditions posed for general nonsmooth equations (Kummer [20], Pang et al. [26]), this actually establishes the equivalence of many conditions discussed in a wide range of literatures for $(NLSDP)$.

The organization of this paper is as follows. In Section 2, we study some useful properties of Clarke's Jacobian for Lipschitz functions, in particular for the metric projector over $\mathcal{S}^p_+$. We propose a strong second order sufficient condition for the nonlinear semidefinite programming $(NLSDP)$ in Section 3. It is shown that this strong second order sufficient condition and constraint nondegeneracy imply the nonsingularity of Clarke's Jacobian of the corresponding nonsmooth system. The promised equivalent conditions are discussed in Section 4. We conclude this paper by pointing out some possible research topics in Section 5.

2. Jacobian Properties Let X, Y, and Z be three arbitrary finite dimensional real vector spaces each equipped with a scalar product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\|$. Let $\mathcal{O}$ be an open set in Y and $\Xi : \mathcal{O} \subseteq Y \to Z$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. The well known Rademacher's theorem (Rockafellar and Wets [36, Section 9.J]) says that Ξ is almost everywhere F(réchet)-differentiable in $\mathcal{O}$. We denote by $\mathcal{O}_\Xi$ the set of points in $\mathcal{O}$ where Ξ is F-differentiable. Then Clarke's generalized Jacobian of Ξ at y is well defined (Clarke [9]):

$$\partial\Xi(y) := \text{conv}\{\partial_B\Xi(y)\},$$

where "conv" denotes the convex hull and

$$\partial_B\Xi(y) := \{V \,:\, V = \lim_{k\to\infty} \mathcal{J}_y\Xi(y^k)\,,\, y^k \to y\,,\, y^k \in \mathcal{O}_\Xi\}.$$

The next lemma is about the generalized Jacobian for composite functions.

Lemma 2.1 *Let $\Psi : X \to Y$ be a continuously differentiable function on an open neighborhood $\widehat{N}$ of $\bar{x}$ and $\Xi : \mathcal{O} \subseteq Y \to Z$ be a locally Lipschitz continuous function on an open set $\mathcal{O}$ containing $\bar{y} := \Psi(\bar{x})$. Suppose that Ξ is directionally differentiable at every point in $\mathcal{O}$ and that $J_x\Psi(\bar{x}) : X \to Y$ is onto. Then it holds that*

$$\partial_B\Phi(\bar{x}) = \partial_B\Xi(\bar{y})\mathcal{J}_x\Psi(\bar{x}), \tag{12}$$

where $\Phi : \widehat{N} \to Z$ is defined by $\Phi(x) := \Xi(\Psi(x))$, $x \in \widehat{N}$.

Proof. By shrinking $\widehat{N}$ if necessary, we may assume that $\Psi(\widehat{N}) \subseteq \mathcal{O}$. Then Φ is Lipschitz continuous and directionally differentiable on $\widehat{N}$. By further shrinking $\widehat{N}$ if necessary, we may also assume that for each $x \in \widehat{N}$, $\mathcal{J}_x\Psi(x)$ is onto.

We shall first show that Φ is F-differentiable at $x \in \widehat{N}$ if and only if Ξ is F-differentiable at $\Psi(x)$, which ensures that

$$\partial_B \Phi(\bar{x}) \subseteq \partial_B \Xi(\bar{y}) \mathcal{J}_x \Psi(\bar{x}).$$

Certainly, Φ is F-differentiable at $x \in \widehat{N}$ if Ξ is F-differentiable at $\Psi(x)$. Now, suppose that Φ is F-differentiable at $x \in \widehat{N}$. Then, since Ξ is directionally differentiable at $\Psi(x)$, for any $d \in X$ we have

$$\mathcal{J}_x \Phi(x) d = \Xi'(\Psi(x); \mathcal{J}_x \Psi(x) d),$$

which implies that for any $s, t \in \Re$ and $u, v \in X$,

$$\begin{aligned} \Xi'(\Psi(x); s\mathcal{J}_x\Psi(x)u + t\mathcal{J}_x\Psi(x)v) &= \Xi'(\Psi(x); \mathcal{J}_x\Psi(x)(su + tv)) \\ &= \mathcal{J}_x\Phi(x)(su + tv) \\ &= s\mathcal{J}_x\Phi(x)u + t\mathcal{J}_x\Phi(x)v \\ &= s\Xi'(\Psi(x); \mathcal{J}_x\Psi(x)u) + t\Xi'(\Psi(x); \mathcal{J}_x\Psi(x)v). \end{aligned}$$

By the surjectivity of $\mathcal{J}_x \Psi(x)$, we can conclude that $\Xi'(\Psi(x); \cdot)$ is a linear operator and so Ξ is Gâteau differentiable at $\Psi(x)$. Since Ξ is assumed to be locally Lipschitz continuous on $\mathcal{O}$, Ξ is F-differentiable at $\Psi(x)$.

Next, we show that the second half inclusion holds:

$$\partial_B \Phi(\bar{x}) \supseteq \partial_B \Xi(\bar{y}) \mathcal{J}_x \Psi(\bar{x}).$$

Let $W \in \partial_B \Xi(\bar{y})$. Then there exists a sequence $\{y^k\}$ in $\mathcal{O}$ converging to $\bar{y}$ such that Ξ is F-differentiable at y^k and $W = \lim_{k\to\infty} \mathcal{J}_y \Xi(y^k)$. By applying the classical Inverse Function Theorem to

$$\Psi\left(\bar{x} + \mathcal{J}_x\Psi(\bar{x})^*(y - \bar{y})\right) - \Psi(\bar{x}) = 0,$$

we obtain that there exists a sequence $\{\tilde{y}^k\}$ in $\mathcal{O}$ converging to $\bar{y}$ such that

$$\Psi\left(\bar{x} + \mathcal{J}_x\Psi(\bar{x})^*(\tilde{y}^k - \bar{y})\right) - \Psi(\bar{x}) = y^k - \Psi(\bar{x})$$

for all k sufficiently large. Let $\tilde{x}^k := \bar{x} + \mathcal{J}_x\Psi(\bar{x})^*(\tilde{y}^k - \bar{y})$. Then $y^k = \Psi(\tilde{x}^k)$ and Φ is F-differentiable at $\tilde{x}^k$ with

$$\mathcal{J}_x \Phi(\tilde{x}^k) = \mathcal{J}_y \Xi(y^k) \mathcal{J}_x \Psi(\tilde{x}^k).$$

By using the fact that $\tilde{y}^k \to \bar{y}$ implies $\tilde{x}^k \to \bar{x}$, we know that there exists a $V \in \partial_B \Phi(\bar{x})$ such that

$$W\mathcal{J}_x\Psi(\bar{x}) = \lim_{k\to\infty} \mathcal{J}_y\Xi(y^k) \lim_{k\to\infty} \mathcal{J}_x\Psi(\tilde{x}^k) = \lim_{k\to\infty} \mathcal{J}_x\Phi(\tilde{x}^k) = V \in \partial_B\Phi(\bar{x}).$$

This completes the proof. □

Let D be a closed convex set in Z. Let $\Pi_D : Z \to Z$ denote the metric projector over D. That is, for any $y \in Z$, $\Pi_D(Z)$ is the unique optimal solution to the convex programming problem:

$$\begin{array}{ll} \min & \frac{1}{2}\langle z - y, z - y\rangle \\ \text{s.t.} & z \in D. \end{array} \tag{13}$$

It is well known (Zarantonello [42]) that the metric projector $\Pi_D(\cdot)$ is contractive, i.e., for any two vectors $y, z \in Z$,

$$\|\Pi_D(y) - \Pi_D(z)\| \le \|y - z\|.$$

Hence, $\Pi_D(\cdot)$ is F-differentiable almost everywhere in Z and for any $y \in Z$, $\partial\Pi_D(y)$ is well defined.

LEMMA 2.2 (Meng et al. [23, Proposition 1]) *Let $D \subseteq Z$ be a closed convex set. Then, for any $y \in Z$ and $V \in \partial\Pi_D(y)$, it holds that*

(a) V *is self-adjoint.*

(b) $\langle d, Vd\rangle \ge 0 \quad \forall\, d \in Z$.

(c) $\langle Vd, d - Vd\rangle \ge 0 \quad \forall\, d \in Z$.

Lemma 2.2 provides general properties about $\partial\Pi_D(\cdot)$. In our analysis, we need a finer characterization about Clarke's Jacobian of $\Pi_{\mathcal{S}^p_+}(\cdot)$. We write $A \succeq 0$ and $A \succ 0$ to mean that A is a symmetric positive semidefinite matrix and a symmetric positive definite matrix, respectively. For any two matrices A and B in $\mathcal{S}^p$, we write

$$\langle A, B\rangle := \mathrm{Tr}\left(A^T B\right)$$

for the *Frobenius inner product* between A and B, where "Tr" denotes the trace of a matrix. Under the Frobenius inner product, the projection $A_+ := \Pi_{\mathcal{S}^p_+}(A)$ of a matrix $A \in \mathcal{S}^p$ onto the cone $\mathcal{S}^p_+$ satisfies the following complementarity condition:

$$\mathcal{S}^p_+ \ni A_+ \perp (A_+ - A) \in \mathcal{S}^p_+, \tag{14}$$

where for any two matrices B and S in $\mathcal{S}^p$, $B \perp S \Longleftrightarrow \langle B, S\rangle = 0$. Let A have the following spectral decomposition

$$A = P\Lambda P^T, \tag{15}$$

where Λ is the diagonal matrix of eigenvalues of A and P is a corresponding orthogonal matrix of orthonormal eigenvectors. Then

$$A_+ = P\Lambda_+ P^T,$$

where Λ_+ is the diagonal matrix whose diagonal entries are the nonnegative parts of the respective diagonal entries of Λ (Higham [15], Tseng [41]). Define three index sets of positive, zero, and negative eigenvalues of A, respectively, as

$$\alpha := \{i \,:\, \lambda_i > 0\}, \quad \beta := \{i \,:\, \lambda_i = 0\}, \quad \gamma := \{i \,:\, \lambda_i < 0\}.$$

Write

$$\Lambda = \begin{bmatrix} \Lambda_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda_\gamma \end{bmatrix} \quad \text{and} \quad P = [\; P_\alpha \quad P_\beta \quad P_\gamma \;]$$

with $P_\alpha \in \Re^{p\times|\alpha|}$, $P_\beta \in \Re^{p\times|\beta|}$, and $P_\gamma \in \Re^{p\times|\gamma|}$. Define the matrix $U \in \mathcal{S}^p$ with entries

$$U_{ij} := \frac{\max\{\lambda_i, 0\} + \max\{\lambda_j, 0\}}{|\,\lambda_i\,| + |\,\lambda_j\,|}, \quad i, j = 1, \ldots, p,$$

where $0/0$ is defined to be 1. Bonnans et al. [1, 2], showed, among many other important things, that $\Pi_{\mathcal{S}^p_+}$ is directionally differentiable everywhere in $\mathcal{S}^p$. Sun and Sun [40] showed that $\Pi_{\mathcal{S}^p_+}$ is a strongly semismooth matrix-valued function and for any $H \in \mathcal{S}^p$, gave an explicit formula for the directional derivative of $\Pi'_{\mathcal{S}^p_+}(A; H)$:

$$\Pi'_{\mathcal{S}^p_+}(A; H) = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & U_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}^T_{\alpha\beta} & \Pi_{\mathcal{S}^{|\beta|}_+}(\widetilde{H}_{\beta\beta}) & 0 \\ \widetilde{H}^T_{\alpha\gamma} \circ U^T_{\alpha\gamma} & 0 & 0 \end{bmatrix} P^T, \tag{16}$$

where $\widetilde{H} := P^T H P$ and $\circ$ denotes the Hadamard product. Hence, we have

$$\mathcal{T}_{\mathcal{S}^p_+}(A_+) = \{B \in \mathcal{S}^p \,:\, B = \Pi'_{\mathcal{S}^p_+}(A_+; B)\} = \{B \in \mathcal{S}^p \,:\, P^T_{\bar\alpha} B P_{\bar\alpha} \succeq 0\}, \tag{17}$$

where $\bar\alpha := \{1, \ldots, p\}\backslash\alpha$ and $P_{\bar\alpha} := [P_\beta \; P_\gamma]$. The lineality space of $\mathcal{T}_{\mathcal{S}^p_+}(A_+)$, i.e., the largest linear space in $\mathcal{T}_{\mathcal{S}^p_+}(A_+)$, denoted by $\mathrm{lin}\left(\mathcal{T}_{\mathcal{S}^p_+}(A_+)\right)$, takes the following form:

$$\mathrm{lin}\left(\mathcal{T}_{\mathcal{S}^p_+}(A_+)\right) = \{B \in \mathcal{S}^p \,:\, P^T_{\bar\alpha} B P_{\bar\alpha} = 0\}. \tag{18}$$

The *critical cone* of $\mathcal{S}^p_+$ at $A \in \mathcal{S}^p$, associated with the complementarity problem (14), is defined as

$$C(A; \mathcal{S}^p_+) := \mathcal{T}_{\mathcal{S}^p_+}(A_+) \cap (A_+ - A)^\perp,$$

which can be completely described:

$$C(A; \mathcal{S}^p_+) = \left\{B \in \mathcal{S}^p \,:\, P^T_\beta B P_\beta \succeq 0,\; P^T_\beta B P_\gamma = 0,\; P^T_\gamma B P_\gamma = 0\right\}. \tag{19}$$

The affine hull of $C(A;\mathcal{S}^p_+)$, which we denote $\mathrm{aff}(C(A;\mathcal{S}^p_+))$, can thus be written as

$$\mathrm{aff}\left(C(A;\mathcal{S}^p_+)\right) = \left\{B \in \mathcal{S}^p \;:\; P_\beta^T B P_\gamma = 0,\; P_\gamma^T B P_\gamma = 0\right\}. \tag{20}$$

Since $0 \in C(A;\mathcal{S}^p_+)$, $\mathrm{aff}\left(C(A;\mathcal{S}^p_+)\right)$ is the linear space generated by $C(A;\mathcal{S}^p_+)$.

We summarize some differential properties of $\Pi_{\mathcal{S}^p_+}(\cdot)$ in the following proposition. For details, see Pang et al. [26, Corollary 10 & Lemma 11].

Proposition 2.1 *The following three statements are true.*

(a) $\Pi_{\mathcal{S}^p_+}(\cdot)$ *is* F*-differentiable at* $A \in \mathcal{S}^p$ *if and only if* A *is nonsingular.*

(b) *For any* $A \in \mathcal{S}^p$, *the directional derivative* $\Pi'_{\mathcal{S}^p_+}(A;\cdot)$ *is* F*-differentiable at* $H \in \mathcal{S}^p$ *if and only if* $\widetilde{H}_{\beta\beta}$ *is nonsingular, where* $\widetilde{H} := P^T H P$.

(c) *Let* $A \in \mathcal{S}^p$ *be arbitrary and* $\Theta(\cdot) := \Pi'_{\mathcal{S}^p_+}(A;\cdot)$. *It holds that*

$$\partial_B \Pi_{\mathcal{S}^p_+}(A) = \partial_B \Theta(0).$$

Proposition 2.1 and Lemma 2.1 allow us to prove the following useful result on $\partial \Pi_{\mathcal{S}^p_+}(\cdot)$.

Proposition 2.2 *Suppose that* $A \in \mathcal{S}^p$ *has the spectral decomposition as in* (15). *Then for any* $V \in \partial_B \Pi_{\mathcal{S}^p_+}(A)$ *(respectively,* $\partial \Pi_{\mathcal{S}^p_+}(A)$*), there exists a* $W \in \partial_B \Pi_{\mathcal{S}^{|\beta|}_+}(0)$ *(respectively,* $\partial \Pi_{\mathcal{S}^{|\beta|}_+}(0)$*) such that*

$$V(H) = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & U_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}_{\alpha\beta}^T & W(\widetilde{H}_{\beta\beta}) & 0 \\ \widetilde{H}_{\alpha\gamma}^T \circ U_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} P^T \quad \forall\, H \in \mathcal{S}^p, \tag{21}$$

where $\widetilde{H} := P^T H P$. *Conversely, for any* $W \in \partial_B \Pi_{\mathcal{S}^{|\beta|}_+}(0)$ *(respectively,* $\partial \Pi_{\mathcal{S}^{|\beta|}_+}(0)$*) , there exists a* $V \in \partial_B \Pi_{\mathcal{S}^p_+}(A)$ *(respectively,* $\partial \Pi_{\mathcal{S}^p_+}(A)$*) such that* (21) *holds.*

Proof. We only need to prove that (21) holds for $V \in \partial_B \Pi_{\mathcal{S}^p_+}(A)$ and $W \in \partial_B \Pi_{\mathcal{S}^{|\beta|}_+}(0)$.

Let $\Theta(\cdot) := \Pi'_{\mathcal{S}^p_+}(A;\cdot)$. Define $\Psi : \mathcal{S}^p \to \mathcal{S}^p$ by $\Psi(H) := P^T H P$, $H \in \mathcal{S}^p$ and $\Xi : \mathcal{S}^p \to \mathcal{S}^p$ by

$$\Xi(B) := P \begin{bmatrix} B_{\alpha\alpha} & B_{\alpha\beta} & U_{\alpha\gamma} \circ B_{\alpha\gamma} \\ B_{\alpha\beta}^T & \Pi_{\mathcal{S}^{|\beta|}_+}(B_{\beta\beta}) & 0 \\ B_{\alpha\gamma}^T \circ U_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} P^T, \quad B \in \mathcal{S}^p.$$

Then, by (16), we have

$$\Theta(H) = \Xi(\Psi(H)), \quad H \in \mathcal{S}^p.$$

Since $\Pi_{\mathcal{S}^{|\beta|}_+}$ is directionally differentiable everywhere and $\mathcal{J}_H \Psi(H) : \mathcal{S}^p \to \mathcal{S}^p$ is onto, we know from Lemma 2.1 that

$$\partial_B \Theta(0) = \partial_B \Xi(0) \mathcal{J}_H \Psi(0).$$

This, together with (c) of Proposition 2.1, completes the proof. □

Remark 2.1 *Relation* (21) *in* Proposition 2.2 *holds for any orthogonal matrix* P *such that the spectral decomposition* (15) *holds. One may further characterize* $\partial_B \Pi_{\mathcal{S}^{|\beta|}_+}(0)$ *as in* Pang et al. [26, Lemma 11]. *In this paper, we do not need the structure of* $\partial_B \Pi_{\mathcal{S}^{|\beta|}_+}(0)$.

Motivated by Shapiro [38, p. 313] and Bonnans and Shapiro [6, p. 487], for any given $B \in \mathcal{S}^p$ we introduce a linear-quadratic function $\Upsilon_B : \mathcal{S}^p \times \mathcal{S}^p \to \Re$ in the next definition.

Definition 2.1 *For any given $B \in \mathcal{S}^p$, define the linear-quadratic function $\Upsilon_B : S^p \times \mathcal{S}^p \to \Re$, which is linear in the first argument and quadratic in the second argument, by*

$$\Upsilon_B(\Gamma, A) := 2\left\langle \Gamma, AB^\dagger A\right\rangle, \quad (\Gamma, A) \in \mathcal{S}^p \times \mathcal{S}^p,$$

where $B^\dagger$ is the Moore-Penrose pseudo-inverse of B.

The following result plays an important role in our subsequent analysis.

Proposition 2.3 *Suppose that $B \in \mathcal{S}^p_+$ and $\Gamma \in \mathcal{N}_{\mathcal{S}^p_+}(B)$. Then for any $V \in \partial\Pi_{\mathcal{S}^p_+}(B+\Gamma)$ and $\Delta B, \Delta\Gamma \in \mathcal{S}^p$ such that $\Delta B = V(\Delta B + \Delta\Gamma)$, it holds that*

$$\langle \Delta B, \Delta\Gamma\rangle \geq -\Upsilon_B(\Gamma, \Delta B). \tag{22}$$

Proof. Let $A := B + \Gamma$. Then we know from Eaves [11] that

$$B = \Pi_{\mathcal{S}^p_+}(B+\Gamma) = \Pi_{\mathcal{S}^p_+}(A) \quad \text{and} \quad B\Gamma = \Gamma B = 0.$$

Thus, we can assume that A has the spectral decomposition as in (15),

$$B = P\begin{bmatrix} \Lambda_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}P^T, \quad \text{and} \quad \Gamma = P\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda_\gamma \end{bmatrix}P^T.$$

Let $\widetilde{B} := P^T BP$, $\widetilde{\Gamma} := P^T\Gamma P$, $\Delta\widetilde{B} := P^T\Delta BP$, and $\Delta\widetilde{\Gamma} := P^T\Delta\Gamma P$. Then, by Proposition 2.2, there exists a $W \in \partial\Pi_{\mathcal{S}^{|\beta|}_+}(0)$ such that

$$V(\Delta B + \Delta\Gamma) = P\begin{bmatrix} \Delta\widetilde{B}_{\alpha\alpha} + \Delta\widetilde{\Gamma}_{\alpha\alpha} & \Delta\widetilde{B}_{\alpha\beta} + \Delta\widetilde{\Gamma}_{\alpha\beta} & U_{\alpha\gamma} \circ (\Delta\widetilde{B}_{\alpha\gamma} + \Delta\widetilde{\Gamma}_{\alpha\gamma}) \\ (\Delta\widetilde{B}_{\alpha\beta} + \Delta\widetilde{\Gamma}_{\alpha\beta})^T & W(\Delta\widetilde{B}_{\beta\beta} + \Delta\widetilde{\Gamma}_{\beta\beta}) & 0 \\ (\Delta\widetilde{B}_{\alpha\gamma} + \Delta\widetilde{\Gamma}_{\alpha\gamma})^T \circ U^T_{\alpha\gamma} & 0 & 0 \end{bmatrix}P^T,$$

which, together with the assumption that $\Delta B = V(\Delta B + \Delta\Gamma)$, implies that

$$\Delta\widetilde{\Gamma}_{\alpha\alpha} = 0, \quad \Delta\widetilde{\Gamma}_{\alpha\beta} = 0, \quad \Delta\widetilde{B}_{\beta\gamma} = 0, \quad \Delta\widetilde{B}_{\gamma\gamma} = 0, \tag{23}$$

$$\Delta\widetilde{B}_{\beta\beta} = W(\Delta\widetilde{B}_{\beta\beta} + \Delta\widetilde{\Gamma}_{\beta\beta}), \tag{24}$$

and

$$\Delta\widetilde{B}_{\alpha\gamma} - U_{\alpha\gamma} \circ \Delta\widetilde{B}_{\alpha\gamma} = U_{\alpha\gamma} \circ \Delta\widetilde{\Gamma}_{\alpha\gamma}. \tag{25}$$

By (c) of Lemma 2.2 and equation (24), we obtain that

$$\left\langle \Delta\widetilde{B}_{\beta\beta}, \Delta\widetilde{\Gamma}_{\beta\beta}\right\rangle = \left\langle W(\Delta\widetilde{B}_{\beta\beta} + \Delta\widetilde{\Gamma}_{\beta\beta}), (\Delta\widetilde{B}_{\beta\beta} + \Delta\widetilde{\Gamma}_{\beta\beta}) - W(\Delta\widetilde{B}_{\beta\beta} + \Delta\widetilde{\Gamma}_{\beta\beta})\right\rangle \geq 0. \tag{26}$$

Hence, by equations (23), (25), and (26),

$$\begin{aligned} \langle \Delta B, \Delta\Gamma\rangle &= \left\langle \Delta\widetilde{B}, \Delta\widetilde{\Gamma}\right\rangle = 2\operatorname{Tr}\left((\Delta\widetilde{B}_{\alpha\gamma})^T\Delta\widetilde{\Gamma}_{\alpha\gamma}\right) + \operatorname{Tr}\left(\Delta\widetilde{B}_{\beta\beta}\Delta\widetilde{\Gamma}_{\beta\beta}\right) \\ &\geq 2\operatorname{Tr}\left((\Delta\widetilde{B}_{\alpha\gamma})^T\Delta\widetilde{\Gamma}_{\alpha\gamma}\right) = 2\sum_{i\in\alpha, j\in\gamma}(\Delta\widetilde{B})_{ij}(\Delta\widetilde{\Gamma})_{ij} \\ &= 2\sum_{i\in\alpha,\, j\in\gamma}\frac{|\lambda_j|}{\lambda_i}\left((\Delta\widetilde{B})_{ij}\right)^2 = -2\sum_{i\in\alpha,\, j\in\gamma}\frac{\lambda_j}{\lambda_i}\left((\Delta\widetilde{B})_{ij}\right)^2. \end{aligned} \tag{27}$$

On the other hand, since

$$B^\dagger = P\begin{bmatrix} (\Lambda_\alpha)^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}P^T,$$

we obtain from (23) and the spectral decomposition of Γ that

$$\begin{aligned}
\Upsilon_B(\Gamma,\Delta B) &= 2\left\langle \Gamma, (\Delta B)B^\dagger(\Delta B)\right\rangle \;=\; 2\langle(\Delta B)\Gamma, B^\dagger(\Delta B)\rangle \\
&= 2\left\langle (\Delta\widetilde{B})\widetilde{\Gamma}, (P^TB^\dagger P)(\Delta\widetilde{B})\right\rangle \\
&= 2\,\mathrm{Tr}\left((\Delta\widetilde{B}_{\alpha\gamma}\Lambda_\gamma)^T(\Lambda_\alpha)^{-1}\Delta\widetilde{B}_{\alpha\gamma}\right) \\
&= 2\sum_{i\in\alpha,\, j\in\gamma}\frac{\lambda_j}{\lambda_i}\left((\Delta\widetilde{B})_{ij}\right)^2. \qquad (28)
\end{aligned}$$

By combining (27) and (28), we get (22). □

3. Strong Second Order Sufficient Condition and Constraint Nondegeneracy Let D be an arbitrary closed set in a given finite dimensional real vector space Z. The *inner* and *outer second order tangent sets* (Bonnans and Shapiro [6, Section 3.2.1]) to the set D at the point $y\in D$ and in the direction $d\in Z$ can be defined, respectively, by

$$\mathcal{T}_D^{i,2}(y,d) := \{w\in Z\,:\,\mathrm{dist}(y+td+\frac{1}{2}t^2w, D) = o(t^2)\,,\; t\geq 0\} \qquad (29)$$

and

$$\mathcal{T}_D^{2}(y,d) := \{w\in Z\,:\,\exists\, t_k\downarrow 0 \;\text{ such that }\; \mathrm{dist}(y+t_kd+\frac{1}{2}t_k^2w, D) = o(t_k^2)\}. \qquad (30)$$

From the definitions of inner and outer second order tangent sets, we can see directly that $\mathcal{T}_D^{i,2}(z,d)\subseteq \mathcal{T}_D^2(y,d)$ and $\mathcal{T}_D^{i,2}(z,d)=\emptyset$ (respectively, $\mathcal{T}_D^2(z,d)=\emptyset$) if $d\notin\mathcal{T}_D^i(y)$ (respectively, $d\notin\mathcal{T}_D(y)$). In general, $\mathcal{T}_D^{i,2}(z,d)\neq\mathcal{T}_D^2(z,d)$ even if D is convex (Bonnans and Shapiro [6, Section 3.3]). However, if D is $\mathcal{C}^2$-cone reducible, the equality always holds (Bonnans and Shapiro [6, Proposition 3.136]). In particular, when $K:=\{0\}\times\mathcal{S}_+^p\subset Y:=\Re^m\times\mathcal{S}^p$,

$$\mathcal{T}_K^{i,2}(y,d) = \mathcal{T}_K^2(y,d) \quad \forall\, y,d\in Y.$$

Let $\bar{x}$ be a feasible solution to the nonlinear semidefinite programming $(NLSDP)$. The *critical cone* $C(\bar{x})$ of $(NLSDP)$ at $\bar{x}$ is defined by

$$C(\bar{x}) := \{d\,:\,\mathcal{J}_xG(\bar{x})d\in\mathcal{T}_K(G(\bar{x})),\; \mathcal{J}_xf(\bar{x})d\leq 0\}\,, \qquad (31)$$

i.e.,

$$C(\bar{x}) = \left\{d\,:\,\mathcal{J}_xh(\bar{x})d=0,\; \mathcal{J}_xg(\bar{x})d\in\mathcal{T}_{\mathcal{S}_+^p}(g(\bar{x})),\; \mathcal{J}_xf(\bar{x})d\leq 0\right\}. \qquad (32)$$

If $\bar{x}$ is a stationary point of $(NLSDP)$, i.e., if $\mathcal{M}(\bar{x})$ is nonempty, then

$$C(\bar{x}) = \left\{d\,:\,\mathcal{J}_xh(\bar{x})d=0,\; \mathcal{J}_xg(\bar{x})d\in\mathcal{T}_{\mathcal{S}_+^p}(g(\bar{x})),\; \mathcal{J}_xf(\bar{x})d= 0\right\}. \qquad (33)$$

Let $\bar{x}$ be a stationary point of $(NLSDP)$. Then there exists $(\bar{\zeta},\overline{\Gamma})\in\mathcal{M}(\bar{x})$ such that

$$\mathcal{J}_xL(\bar{x},\bar{\zeta},\overline{\Gamma})=0, \quad -h(\bar{x})=0, \quad \text{and} \quad \overline{\Gamma}\in\mathcal{N}_{\mathcal{S}_+^p}(g(\bar{x})).$$

By using the fact that

$$\overline{\Gamma}\in\mathcal{N}_{S_+^p}(g(\bar{x}))\Longleftrightarrow \mathcal{S}_+^p\ni(-\overline{\Gamma})\perp g(\bar{x})\in\mathcal{S}_+^p,$$

we may assume that $A:=g(\bar{x})+\overline{\Gamma}$ has the spectral decomposition as in (15),

$$g(\bar{x}) = P\begin{bmatrix}\Lambda_\alpha & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}P^T, \quad \text{and} \quad \overline{\Gamma} = P\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & \Lambda_\gamma\end{bmatrix}P^T. \qquad (34)$$

Then, by (17) and (18), we have

$$\mathcal{T}_{\mathcal{S}_+^p}(g(\bar{x})) = \left\{B\in\mathcal{S}^p\,:\,[P_\beta\;P_\gamma]^TB[P_\beta\;P_\gamma]\succeq 0\right\},$$

$$\mathcal{T}_{\mathcal{S}_+^p}(g(\bar{x}))\cap\overline{\Gamma}^\perp = \left\{B\in\mathcal{S}^p\,:\,P_\beta^TBP_\beta\succeq 0, \quad P_\beta^TBP_\gamma=0, \quad P_\gamma^TBP_\gamma=0\right\}, \qquad (35)$$

and

$$\text{lin}\left(\mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x}))\right) = \left\{B \in \mathcal{S}^p \,:\, [P_\beta \; P_\gamma]^T B [P_\beta \; P_\gamma] = 0\right\}. \tag{36}$$

Furthermore, since $\mathcal{M}(\bar{x})$ is nonempty,

$$\begin{aligned} C(\bar{x}) &= \left\{d \,:\, \mathcal{J}_x h(\bar{x})d = 0, \;\; [P_\beta \; P_\gamma]^T(\mathcal{J}_x g(\bar{x})d)[P_\beta \; P_\gamma] \succeq 0, \; P_\gamma^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0\right\} \\ &= \Big\{d \,:\, \mathcal{J}_x h(\bar{x})d = 0, \;\; P_\beta^T(\mathcal{J}_x g(\bar{x})d)P_\beta \succeq 0, \; P_\beta^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0, \\ &\qquad P_\gamma^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0\Big\} \\ &= \left\{d \,:\, \mathcal{J}_x h(\bar{x})d = 0, \; \mathcal{J}_x g(\bar{x})d \in C(A; \mathcal{S}^p_+)\right\}, \end{aligned} \tag{37}$$

where $C(A; \mathcal{S}^p_+)$ is the critical cone of $\mathcal{S}^p_+$ at $A = g(\bar{x}) + \overline{\Gamma}$. However, it is not easy to give an explicit formula to the affine hull of $C(\bar{x})$, which we denote $\text{aff}(C(\bar{x}))$.[1] Instead, we define the following outer approximation set to $\text{aff}(C(\bar{x}))$ with respect to $(\bar{\zeta}, \overline{\Gamma})$ by

$$\text{app}(\bar{\zeta}, \overline{\Gamma}) := \left\{d \,:\, \mathcal{J}_x h(\bar{x})d = 0, \quad \mathcal{J}_x g(\bar{x})d \in \text{aff}\left(C(A; \mathcal{S}^p_+)\right)\right\}. \tag{38}$$

By (20), it holds that

$$\text{app}(\bar{\zeta}, \overline{\Gamma}) = \left\{d \,:\, \mathcal{J}_x h(\bar{x})d = 0, \; P_\beta^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0, \; P_\gamma^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0\right\}. \tag{39}$$

Then by the definition of $\text{aff}(C(\bar{x}))$, we have for any $(\bar{\zeta}, \overline{\Gamma}) \in \mathcal{M}(\bar{x})$ that

$$\text{aff}(C(\bar{x})) \subseteq \text{app}(\bar{\zeta}, \overline{\Gamma}). \tag{40}$$

Obviously, the two sets in (40) coincide if the *strict complementary* condition holds at $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$:

$$\text{rank}(g(\bar{x})) + \text{rank}(\overline{\Gamma}) = p,$$

where "rank" denotes the rank of a square matrix. In general, these two sets may be different even if $\mathcal{M}(\bar{x})$ is a singleton as in the case for the conventional nonlinear programming (NLP).

The next proposition shows that the equality in (40) holds if $(\bar{\zeta}, \overline{\Gamma}) \in \mathcal{M}(\bar{x})$ satisfies a constraint qualification stronger than Robinson's CQ (7) at $\bar{x}$, which, in the context of $(NLSDP)$, is equivalent to

$$\begin{pmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{pmatrix} X + \begin{pmatrix} \{0\} \\ \mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x})) \end{pmatrix} = \begin{pmatrix} \Re^m \\ \mathcal{S}^p \end{pmatrix}. \tag{41}$$

Proposition 3.1 [2] *Let $\bar{x}$ be a feasible solution to the nonlinear semidefinite programming* $(NLSDP)$ *and $(\bar{\zeta}, \overline{\Gamma}) \in \mathcal{M}(\bar{x})$. Suppose that $(\bar{\zeta}, \overline{\Gamma})$ satisfies the following strict constraint qualification:*

$$\begin{pmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{pmatrix} X + \begin{pmatrix} \{0\} \\ \mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x})) \cap \overline{\Gamma}^\perp \end{pmatrix} = \begin{pmatrix} \Re^m \\ \mathcal{S}^p \end{pmatrix}. \tag{42}$$

Then $\mathcal{M}(\bar{x})$ is a singleton, i.e., $\mathcal{M}(\bar{x}) = \{(\bar{\zeta}, \overline{\Gamma})\}$, and $\text{aff}(C(\bar{x})) = \text{app}(\bar{\zeta}, \overline{\Gamma})$.

Proof. The uniqueness of $(\bar{\zeta}, \overline{\Gamma})$ follows from Bonnans and Shapiro [6, Proposition 4.50]. We only need to show

$$\text{app}(\bar{\zeta}, \overline{\Gamma}) \subseteq \text{aff}(C(\bar{x})).$$

Let d be an arbitrary vector in $\text{app}(\bar{\zeta}, \overline{\Gamma})$. Let $A := g(\bar{x}) + \overline{\Gamma}$. We may assume that A has the spectral decomposition as in (15) and the two matrices $g(\bar{x})$ and $\overline{\Gamma}$ satisfy (34). Let S be any matrix in $\mathcal{S}^p$ such that

$$P^T S P = \begin{bmatrix} P_\alpha^T S P_\alpha & P_\alpha^T S P_\beta & P_\alpha^T S P_\gamma \\ P_\beta^T S P_\alpha & P_\beta^T S P_\beta & 0 \\ P_\gamma^T S P_\alpha & 0 & 0 \end{bmatrix} \quad \text{with} \;\; P_\beta^T S P_\beta \succ 0.$$

[1] One referee pointed out that a characterization on the set $\text{aff}(C(\bar{x}))$ was given in a recent report (Bonnans and Ramírez C. [4, Lemma 2.2]) by using a direction $d \in \text{ri}(C(\bar{x}))$, the relative interior of $C(\bar{x})$.

[2] It was observed by one referee that in order to obtain $\text{aff}(C(\bar{x})) = \text{app}(\bar{\zeta}, \overline{\Gamma})$, it suffices to assume the existence of a direction $\bar{d} \in C(\bar{x})$ such that $P_\beta^T(\mathcal{J}_x g(\bar{x})\bar{d})P_\beta \succ 0$. See Bonnans and Ramírez C. [4, Corollary 2.3].

By the strict constraint qualification (42), we know that there exist a vector $\bar{d} \in X$ and a matrix $U \in \mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x})) \cap \overline{\Gamma}^{\perp}$ such that

$$\begin{cases} \mathcal{J}_x h(\bar{x})(-\bar{d}) = 0, \\ \mathcal{J}_x g(\bar{x})(-\bar{d}) + U = -S, \end{cases} \tag{43}$$

which, together with (35), implies that

$$\mathcal{J}_x g(\bar{x})\bar{d} = U + S \in \mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x})) \cap \overline{\Gamma}^{\perp}, \quad P_\beta^T(\mathcal{J}_x g(\bar{x})\bar{d})P_\beta \succ 0, \quad \text{and} \quad \bar{d} \in C(\bar{x}).$$

Let $\bar{\tau} > 0$ be sufficiently large such that

$$P_\beta^T[\mathcal{J}_x g(\bar{x})(\bar{\tau}\bar{d} - d)]P_\beta = \bar{\tau}P_\beta^T(\mathcal{J}_x g(\bar{x})\bar{d})P_\beta - P_\beta^T(\mathcal{J}_x g(\bar{x})d)P_\beta \succeq 0.$$

Furthermore, since

$$P_\beta^T(\mathcal{J}_x g(\bar{x})\bar{d})P_\gamma = P_\beta^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0 \quad \text{and} \quad P_\gamma^T(\mathcal{J}_x g(\bar{x})\bar{d})P_\gamma = P_\gamma^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0,$$

it holds that

$$\bar{\tau}\bar{d} - d \in C(\bar{x}).$$

Therefore, by using the facts that $d = \bar{\tau}\bar{d} - (\bar{\tau}\bar{d} - d)$ and both $\bar{\tau}\bar{d}$ and $\bar{\tau}\bar{d} - d$ are in the critical cone $C(\bar{x})$, we complete the proof. □

Before we state the second order conditions for $(NLSDP)$, we need below the concept of $\mathcal{C}^2$-cone reducibility, which is adapted from Bonnans and Shapiro [6, Definition 3.135].

DEFINITION 3.1 *A closed (not necessarily convex) set $D \subseteq Y$ is called $\mathcal{C}^2$-cone reducible at a point $\bar{y} \in D$ if there exist a neighborhood $\mathcal{V} \subseteq Y$ of $\bar{y}$, a pointed closed convex cone Q (a cone is said to be pointed if and only its lineality space is the origin) in a finite dimensional space Z and a twice continuously differentiable mapping $\Xi : \mathcal{V} \to Z$ such that: (i) $\Xi(\bar{y}) = 0 \in Z$, (ii) the derivative mapping $J_y\Xi(\bar{y}) : Y \to Z$ is onto, and (iii) $D \cap \mathcal{V} = \{y \in \mathcal{V} \,|\, \Xi(y) \in Q\}$. We say that D is $\mathcal{C}^2$-cone reducible if D is $\mathcal{C}^2$-cone reducible at every point $\bar{y} \in Y$ (possibly to a different pointed cone Q).*

Many interesting sets such as the polyhedral convex set, the second-order cone, and the cone S^p_+ are all $\mathcal{C}^2$-cone reducible, and the Cartesian product of $\mathcal{C}^2$-cone reducible sets is again $\mathcal{C}^2$-cone reducible (Bonnans and Shapiro [6], Shapiro [39]). In particular, $K = \{0\} \times \mathcal{S}^p_+$ is $\mathcal{C}^2$-cone reducible.

Recall that for any set $D \subseteq Z$, the support function of the set D is defined as

$$\sigma(y, D) := \sup_{z \in D}\langle z, y\rangle, \quad y \in Z.$$

Combining Theorem 3.45 and Proposition 3.136 with Theorem 3.137 in Bonnans and Shapiro [6], we can state in the following theorem the second order necessary condition and the second order sufficient condition for the nonlinear semidefinite programming $(NLSDP)$. See also Bonnans et al. [2].

THEOREM 3.1 (*Second order conditions.*) *Let $K = \{0\} \times \mathcal{S}^p_+ \subset \Re^m \times \mathcal{S}^p$. Suppose that $\bar{x}$ is a locally optimal solution to the nonlinear semidefinite programming $(NLSDP)$ and Robinson's CQ holds at $\bar{x}$. Then the following inequality holds:*

$$\sup_{\mu \in \mathcal{M}(\bar{x})} \left\{ \left\langle d, \mathcal{J}^2_{xx}L(\bar{x}, \mu)d \right\rangle - \sigma\left(\mu, \mathcal{T}^2_K(G(\bar{x}), \mathcal{J}_x G(\bar{x})d)\right) \right\} \geq 0 \quad \forall\, d \in C(\bar{x}). \tag{44}$$

Conversely, let $\bar{x}$ be a feasible solution to $(NLSDP)$ such that $\mathcal{M}(\bar{x})$ is nonempty. Suppose that Robinson's CQ holds at $\bar{x}$. Then the following condition

$$\sup_{\mu \in \mathcal{M}(\bar{x})} \left\{ \left\langle d, \mathcal{J}^2_{xx}L(\bar{x}, \mu)d \right\rangle - \sigma\left(\mu, \mathcal{T}^2_K(G(\bar{x}), \mathcal{J}_x G(\bar{x})d)\right) \right\} > 0 \quad \forall\, d \in C(\bar{x})\backslash\{0\} \tag{45}$$

is necessary and sufficient for the quadratic growth condition at the point $\bar{x}$:

$$f(x) \geq f(\bar{x}) + c\|x - \bar{x}\|^2 \quad \forall\, x \in \widehat{N} \text{ such that } G(x) \in K \tag{46}$$

for some constant $c > 0$ and a neighborhood $\widehat{N}$ of $\bar{x}$ in X.

Obviously, when the second order growth condition (46) holds, $\bar{x}$ is a strictly local solution of $(NLSDP)$. So there exists no gap between the above second order sufficient condition (45) and the second order necessary condition (44).

We write $\mu = (\zeta, \Gamma) \in \Re^m \times \mathcal{S}^p$ for any $\mu \in \mathcal{M}(\bar{x})$. Then for $\mu \in \mathcal{M}(\bar{x})$ and $d \in C(\bar{x})$ the "sigma term" in (44) and (45) can be written as

$$\begin{aligned} \sigma\left(\mu, \mathcal{T}_K^2(G(\bar{x}), \mathcal{J}_x G(\bar{x})d)\right) &= \sigma\left(\zeta, \mathcal{T}^2_{\{0\}}(h(\bar{x}), \mathcal{J}_x h(\bar{x})d)\right) + \sigma\left(\Gamma, \mathcal{T}^2_{\mathcal{S}^p_+}(g(\bar{x}), \mathcal{J}_x g(\bar{x})d)\right) \\ &= 0 + \sigma\left(\Gamma, \mathcal{T}^2_{\mathcal{S}^p_+}(g(\bar{x}), \mathcal{J}_x g(\bar{x})d)\right) \\ &= \sigma\left(\Gamma, \mathcal{T}^2_{\mathcal{S}^p_+}(g(\bar{x}), \mathcal{J}_x g(\bar{x})d)\right), \end{aligned}$$

which becomes $-\infty$ for any $d \in X$ such that $\mathcal{J}_x g(\bar{x})d \notin \mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x}))$. This means that in order to define a stronger second order sufficient condition over a set larger than $C(\bar{x})$ one needs to find a substitute for this sigma term. The following lemma, due to Shapiro [38, p. 313] and Bonnans and Shapiro [6, p. 487], makes it possible.

Lemma 3.1 *Let $\bar{x}$ be a feasible solution to $(NLSDP)$ such that $\mathcal{M}(\bar{x})$ is nonempty. Then for any $(\zeta, \Gamma) \in \mathcal{M}(\bar{x})$ with $\zeta \in \Re^m$ and $\Gamma \in \mathcal{S}^p$, one has*

$$\Upsilon_{g(\bar{x})}(\Gamma, \mathcal{J}_x g(\bar{x})d) = \sigma\left(\Gamma, \mathcal{T}^2_{\mathcal{S}^p_+}(g(\bar{x}), \mathcal{J}_x g(\bar{x})d)\right) \quad \forall\, d \in C(\bar{x}).$$

Now, we are ready to define a strong second order sufficient condition, which extends an analogue for the conventional nonlinear programming (NLP) introduced by Robinson [30] to the nonlinear semidefinite programming $(NLSDP)$.

Definition 3.2 *Let $\bar{x}$ be a stationary point of the nonlinear semidefinite programming $(NLSDP)$. We say that the strong second order sufficient condition holds at $\bar{x}$ if*

$$\sup_{(\zeta,\Gamma)\in\mathcal{M}(\bar{x})} \left\{ \left\langle d, \mathcal{J}^2_{xx} L(\bar{x}, \zeta, \Gamma) d \right\rangle - \Upsilon_{g(\bar{x})}(\Gamma, \mathcal{J}_x g(\bar{x})d) \right\} > 0 \quad \forall\, d \in \widehat{C}(\bar{x}) \backslash \{0\}, \tag{47}$$

where for any $(\zeta, \Gamma) \in \mathcal{M}(\bar{x})$, $(\zeta, \Gamma) \in \Re^m \times \mathcal{S}^p$ and

$$\widehat{C}(\bar{x}) := \bigcap_{(\zeta,\Gamma)\in\mathcal{M}(\bar{x})} \operatorname{app}(\zeta, \Gamma).$$

Next, we define a nondegeneracy condition for $(NLSDP)$, which is an analogue of the LICQ for the conventional nonlinear programming (NLP). The concept of nondegeneracy originally appeared in Robinson [31] for the general optimization problem (OP). Here we adopt a somewhat slightly different version from Robinson's original definition.

Definition 3.3 *We say that a feasible point $\bar{x}$ to the optimization problem (OP) is constraint nondegenerate if*

$$\mathcal{J}_x G(\bar{x}) X + \operatorname{lin}(\mathcal{T}_K(\bar{y})) = Y, \tag{48}$$

where $\bar{y} := G(\bar{x})$.

The name "constraint nondegeneracy" was coined by Robinson in [34]. A related concept called rank-reducibility was introduced by Shapiro [37]. The nondegeneracy condition (48) given here is consistent with the version given in Robinson [32] and has been extensively used in Bonnans and Shapiro [6] and Shapiro [39] for sensitivity and stability analysis in optimization and variational inequalities. See Bonnans and Shapiro [6] and Shapiro [39] for various equivalent forms. Certainly, the constraint nondegenerate condition (48) implies Robinson's CQ (7). For the conventional (NLP), as observed in Robinson [31] and Shapiro [39], the LICQ is equivalent to the constraint nondegeneracy. For the nonlinear semidefinite programming $(NLSDP)$, the constraint nondegeneracy takes the following form:

$$\begin{pmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{pmatrix} X + \begin{pmatrix} \{0\} \\ \operatorname{lin}\left(\mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x}))\right) \end{pmatrix} = \begin{pmatrix} \Re^m \\ \mathcal{S}^p \end{pmatrix}. \tag{49}$$

Let x be a feasible point to $(NLSDP)$ such that $\mathcal{M}(x)$ is nonempty. Then there exists $(\zeta, \Gamma) \in \Re^m \times \mathcal{S}^p$, together with x, satisfying the following KKT condition:

$$\mathcal{J}_x L(x, \zeta, \Gamma) = 0, \quad -h(x) = 0, \quad \text{and} \quad \Gamma \in \mathcal{N}_{\mathcal{S}^p_+}(g(x)), \tag{50}$$

where

$$L(x, \zeta, \Gamma) = f(x) + \langle \zeta, h(x) \rangle + \langle \Gamma, g(x) \rangle.$$

Since, from Eaves [11],

$$\Gamma \in \mathcal{N}_{\mathcal{S}^p_+}(g(x)) \iff g(x) = \Pi_{\mathcal{S}^p_+}(g(x) + \Gamma),$$

we can write the KKT condition (50) equivalently as

$$F(x, \zeta, \Gamma) := \begin{bmatrix} \mathcal{J}_x L(x, \zeta, \Gamma) \\ -h(x) \\ -g(x) + \Pi_{\mathcal{S}^p_+}(g(x) + \Gamma) \end{bmatrix} = \begin{bmatrix} \mathcal{J}_x L(x, \zeta, \Gamma) \\ -h(x) \\ \Gamma - \Pi_{\mathcal{S}^p_-}(\Gamma + g(x)) \end{bmatrix} = 0, \tag{51}$$

where S^p_- is the cone of negative semidefinite symmetric matrices in $\mathcal{S}^p$, i.e., $S^p_- = -S^p_+$. Both (50) and (51) are equivalent to

$$0 \in \begin{bmatrix} \mathcal{J}_x L(x, \zeta, \Gamma) \\ -h(x) \\ -g(x) \end{bmatrix} + \begin{bmatrix} \mathcal{N}_X(x) \\ \mathcal{N}_{\Re^m}(\zeta) \\ \mathcal{N}_{S^p_-}(\Gamma) \end{bmatrix}. \tag{52}$$

Problem (52) is in the form of the following generalized equation:

$$0 \in \phi(z) + \mathcal{N}_D(z), \tag{53}$$

where ϕ is a continuously differentiable mapping from a given finite dimensional real vector space Z to itself and D is a closed convex set in Z.

Robinson [30] introduced the far reaching concept of *strong regularity* for a solution of the generalized equation (53).

Definition 3.4 *Let $\bar{z}$ be a solution of the generalized equation* (53). *We say that $\bar{z}$ is a strongly regular solution of the generalized equation* (53) *if there exist neighborhoods $\mathcal{B}$ of the origin $0 \in Z$ and $\mathcal{V}$ of $\bar{z}$ such that for every $\delta \in \mathcal{B}$, the following linearized generalized equation*

$$\delta \in \phi(\bar{z}) + \mathcal{J}_z \phi(\bar{z})(z - \bar{z}) + \mathcal{N}_D(z)$$

has a unique solution in $\mathcal{V}$, denoted by $z_{\mathcal{V}}(\delta)$, and the mapping $z_{\mathcal{V}} : \mathcal{B} \to \mathcal{V}$ is Lipschitz continuous.

Remark 3.1 *Recall that a function $\Xi : \mathcal{O} \subseteq Z \to Z$ is said to be a locally Lipschitz homeomorphism near $\bar{z} \in \mathcal{O}$ if there exists an open neighborhood $\mathcal{V} \subseteq \mathcal{O}$ of $\bar{z}$ such that the restricted mapping $\Xi\,|_{\mathcal{V}}: \mathcal{V} \to \Xi(\mathcal{V})$ is Lipschitz continuous and bijective, and its inverse is also Lipschitz continuous. Define two mappings $\widehat{\Xi},\ \Xi : Z \to Z$ by*

$$\widehat{\Xi}(z) := z - \Pi_D(z - \widehat{\phi}(z)) \quad \text{and} \quad \Xi(z) := z - \Pi_D(z - \phi(z)),$$

where $\widehat{\phi}(z) := \phi(\bar{z}) + \mathcal{J}_z \phi(\bar{z})(z - \bar{z})$, $z \in Z$. From Lemma 3.1 *in* Robinson [33] *or* Theorem 3.1 *in* Kummer [20] *we know that $\widehat{\Xi}$ is a locally Lipschitz homeomorphism near $\bar{z}$ if and only if Ξ is so. See also* Theorem 5.2.8 *in* Facchinei and Pang [12]*. Thus, $\bar{z}$ is a strongly regular solution of the generalized equation* (53) *is equivalent to say that $\widehat{\Xi}$ or Ξ is a locally Lipschitz homeomorphism near $\bar{z}$.*

The next proposition relates the strong second order sufficient condition and constraint nondegeneracy to the nonsingularity of Clarke's Jacobian of the mapping F and the strong regularity of a solution to the generalized equation (52).

Proposition 3.2 *Let $\bar{x}$ be a feasible solution to the nonlinear semidefinite programming $(NLSDP)$. Let $\bar{\zeta} \in \Re^m$ and $\overline{\Gamma} \in \mathcal{S}^p$ be such that $(\bar{\zeta}, \overline{\Gamma}) \in \mathcal{M}(\bar{x})$, i.e., let $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ be a KKT point of $(NLSDP)$. Consider the following three statements:*

(a) *The strong second order sufficient condition* (47) *holds at $\bar{x}$ and $\bar{x}$ is constraint nondegenerate.*

(b) *Any element in $\partial F(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ is nonsingular.*

(c) *The KKT point* $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ *is a strongly regular solution of the generalized equation* (52).

It holds that (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c).

Proof. "(a) $\Longrightarrow$ (b)" Since the constraint nondegeneracy condition (49) is assumed to hold at $\bar{x}$, $(\bar{\zeta}, \overline{\Gamma})$ satisfies the strict constraint qualification (42). Thus, by Proposition 3.1, $\mathcal{M}(\bar{x}) = \{(\bar{\zeta}, \overline{\Gamma})\}$ and $\mathrm{aff}(C(\bar{x})) = \mathrm{app}(\bar{\zeta}, \overline{\Gamma})$. The strong second order sufficient condition (47) then takes the following form:

$$\left\langle d, \mathcal{J}^2_{xx} L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) d \right\rangle - \Upsilon_{g(\bar{x})} \left(\overline{\Gamma}, \mathcal{J}_x g(\bar{x}) d \right) > 0 \quad \forall\, d \in \mathrm{aff}(C(\bar{x})) \backslash \{0\}. \tag{54}$$

Let W be an arbitrary element in $\partial F(\bar{x}, \bar{\zeta}, \overline{\Gamma})$. We shall show that W is nonsingular. Let $(\Delta x, \Delta\zeta, \Delta\Gamma) \in X \times \Re^m \times \mathcal{S}^p$ be such that

$$W(\Delta x, \Delta\zeta, \Delta\Gamma) = 0.$$

Let $A := g(\bar{x}) + \overline{\Gamma}$. Without loss of generality, we assume that A has the spectral decomposition as in (15) and $g(\bar{x})$ and $\overline{\Gamma}$ satisfy (34). Then, by Lemma 2.1, we know that there exists a $V \in \partial \Pi_{\mathcal{S}^p_+}(A)$ such that

$$W(\Delta x, \Delta\zeta, \Delta\Gamma) = \begin{bmatrix} \mathcal{J}^2_{xx} L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \Delta x + \mathcal{J}_x h(\bar{x})^* \Delta\zeta + \mathcal{J}_x g(\bar{x})^* \Delta\Gamma \\ -\mathcal{J}_x h(\bar{x}) \Delta x \\ -\mathcal{J}_x g(\bar{x}) \Delta x + V\left(\mathcal{J}_x g(\bar{x}) \Delta x + \Delta\Gamma\right) \end{bmatrix} = 0. \tag{55}$$

From Proposition 2.2, (39), and the second and the third equations of (55) we know that

$$\Delta x \in \mathrm{app}(\bar{\zeta}, \overline{\Gamma}) = \mathrm{aff}(C(\bar{x})). \tag{56}$$

By the first and second equations of (55), we obtain that

$$\begin{aligned} 0 &= \left\langle \Delta x, \mathcal{J}^2_{xx} L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \Delta x + \mathcal{J}_x h(\bar{x})^* \Delta\zeta + \mathcal{J}_x g(\bar{x})^* \Delta\Gamma \right\rangle \\ &= \left\langle \Delta x, \mathcal{J}^2_{xx} L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \Delta x \right\rangle + \langle \Delta x, \mathcal{J}_x h(\bar{x})^* \Delta\zeta \rangle + \langle \Delta x, \mathcal{J}_x g(\bar{x})^* \Delta\Gamma \rangle \\ &= \left\langle \Delta x, \mathcal{J}^2_{xx} L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \Delta x \right\rangle + \langle \Delta\zeta, \mathcal{J}_x h(\bar{x}) \Delta x \rangle + \langle \Delta\Gamma, \mathcal{J}_x g(\bar{x}) \Delta x \rangle \\ &= \left\langle \Delta x, \mathcal{J}^2_{xx} L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \Delta x \right\rangle + \langle \mathcal{J}_x g(\bar{x}) \Delta x, \Delta\Gamma \rangle, \end{aligned}$$

which, together with the third equation of (55) and Proposition 2.3, implies that

$$0 \geq \left\langle \Delta x, \mathcal{J}^2_{xx} L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \Delta x \right\rangle - \Upsilon_{g(\bar{x})} \left(\overline{\Gamma}, \mathcal{J}_x g(\bar{x}) \Delta x \right). \tag{57}$$

Hence, we can conclude from (56), (57), and the strong second order sufficient condition (54) that

$$\Delta x = 0.$$

Thus, (55) reduces to

$$\begin{bmatrix} \mathcal{J}_x h(\bar{x})^* \Delta\zeta + \mathcal{J}_x g(\bar{x})^* \Delta\Gamma \\ V(\Delta\Gamma) \end{bmatrix} = 0. \tag{58}$$

From Proposition 2.2 and $V(\Delta\Gamma) = 0$, we obtain that

$$P_\alpha{}^T \Delta\Gamma P_\alpha = 0, \quad P_\alpha{}^T \Delta\Gamma P_\beta = 0, \quad \text{and} \quad P_\alpha{}^T \Delta\Gamma P_\gamma = 0. \tag{59}$$

By the constraint nondegeneracy condition (49), there exist a vector $d \in X$ and a matrix $S \in \mathrm{lin}\left(\mathcal{T}_{\mathcal{S}^p_+}(g(\bar{x}))\right)$ such that

$$\mathcal{J}_x h(\bar{x}) d = \Delta\zeta \quad \text{and} \quad \mathcal{J}_x g(\bar{x}) d + S = \Delta\Gamma. \tag{60}$$

Hence, by (60) and the first equation of (58), we obtain

$$\begin{aligned} \langle \Delta\zeta, \Delta\zeta \rangle + \langle \Delta\Gamma, \Delta\Gamma \rangle &= \langle \mathcal{J}_x h(\bar{x}) d, \Delta\zeta \rangle + \langle \mathcal{J}_x g(\bar{x}) d + S, \Delta\Gamma \rangle \\ &= \langle d, \mathcal{J}_x h(\bar{x})^* \Delta\zeta \rangle + \langle d, \mathcal{J}_x g(\bar{x})^* \Delta\Gamma \rangle + \langle S, \Delta\Gamma \rangle \\ &= \langle d, \mathcal{J}_x h(\bar{x})^* \Delta\zeta + \mathcal{J}_x g(\bar{x})^* \Delta\Gamma \rangle + \langle S, \Delta\Gamma \rangle \\ &= \langle S, \Delta\Gamma \rangle \\ &= \langle P^T S P, P^T \Delta\Gamma P \rangle, \end{aligned}$$

which, together with (59) and (36), implies that

$$\langle \Delta\zeta, \Delta\zeta\rangle + \langle \Delta\Gamma, \Delta\Gamma\rangle = \langle P^T SP, P^T \Delta\Gamma P\rangle = 0.$$

Thus,

$$\Delta\zeta = 0 \quad \text{and} \quad \Delta\Gamma = 0.$$

This, together with $\Delta x = 0$, shows that W is nonsingular.

"(b) $\Longrightarrow$ (c)" By Clarke's inverse function theorem (Clarke [8, 9]), F is a locally Lipschitz homeomorphism near $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$. Thus, by Remark 3.1, $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ is a strongly regular solution of the generalized equation (52). □

In Proposition 3.2, it is shown that (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c). In the next section, we shall show that if $\bar{x}$ is a locally optimal solution to the nonlinear semidefinite programming $(NLSDP)$ and Robinson's CQ holds at $\bar{x}$, then these three statements are actually equivalent to each other.

4. Equivalent Conditions We first introduce a uniform version of the second order growth condition defined in Bonnans and Shapiro [6, Definition 5.16]. Let U be a Banach space and $f : X \times U \to \Re$ and $G : X \times U \to Y$. We say that $(f(x, u), G(x, u))$, with $u \in U$, is a $\mathcal{C}^2$*-smooth parameterization* of the optimization problem (OP) if $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$ are twice continuously differentiable and there exists a $\bar{u} \in U$ such that $f(\cdot, \bar{u}) = f(\cdot)$ and $G(\cdot, \bar{u}) = G(\cdot)$. The corresponding parameterized problem takes the form:

$$(OP_u) \qquad \begin{array}{ll} \min\limits_{x \in X} & f(x, u) \\ \text{s.t.} & G(x, u) \in K\,. \end{array} \tag{61}$$

We say that a parameterization is *canonical* if $U := X \times Y$, $\bar{u} = (0, 0) \in X \times Y$, and

$$(f(x, u), G(x, u)) := (f(x) - \langle u_1, x\rangle, G(x) + u_2), \quad x \in X,\ u := (u_1, u_2) \in X \times Y.$$

Definition 4.1 *Let $\bar{x}$ be a stationary point of the optimization problem (OP). We say that the uniform second order (quadratic) growth condition holds at $\bar{x}$ with respect to a $\mathcal{C}^2$-smooth parameterization $(f(x, u), G(x, u))$ if there exist $c > 0$ and neighborhoods $\mathcal{V}_X$ of $\bar{x}$ and $\mathcal{V}_U$ of $\bar{u}$ such that for any $u \in \mathcal{V}_U$ and any stationary point $x(u) \in \mathcal{V}_X$ of the corresponding parameterized problem (OP_u), the following holds:*

$$f(x, u) \geq f(x(u), u) + c\|x - x(u)\|^2 \quad \forall\, x \in \mathcal{V}_X \text{ such that } G(x, u) \in K. \tag{62}$$

We say that the uniform second order growth condition holds at $\bar{x}$ if (62) holds for every $\mathcal{C}^2$-smooth parameterization of (OP).

The next lemma shows that for the nonlinear semidefinite programming $(NLSDP)$ the uniform second order growth condition implies the strong second order sufficient condition.

Lemma 4.1 *Let $\bar{x}$ be a stationary point of the nonlinear semidefinite programming $(NLSDP)$. Suppose that Robinson's CQ holds at $\bar{x}$. If the uniform second order growth condition holds at $\bar{x}$ with respect to the canonical parameterization, then the strong second order sufficient condition (47) holds at $\bar{x}$.*

Proof. Let $(\bar{\zeta}, \overline{\Gamma}) \in \mathcal{M}(\bar{x})$. We may assume that $A := g(\bar{x}) + \bar{\Gamma}$ has the spectral decomposition as in (15) and $g(\bar{x})$ and $\overline{\Gamma}$ satisfy (34). Consider the following parameterized nonlinear semidefinite programming problem:

$$\begin{array}{ll} \min\limits_{x \in X} & f(x) \\ \text{s.t.} & h(x) = 0, \\ & g(x) + \tau P_\beta P_\beta^T \in \mathcal{S}^p_+, \end{array} \tag{63}$$

where $\tau \in \Re$ is a parameter. Then, for any $\tau > 0$, $(\bar{\zeta}, \overline{\Gamma})$, together with $\bar{x}$, satisfies the following KKT condition of the parameterized problem (63):

$$\mathcal{J}_x L_\tau(\bar{x}, \zeta, \Gamma) = \mathcal{J}_x L(\bar{x}, \zeta, \Gamma) = 0, \quad -h(\bar{x}) = 0, \quad \text{and} \quad \Gamma \in \mathcal{N}_{\mathcal{S}^p_+}\left(g(\bar{x}) + \tau P_\beta P_\beta^T\right), \tag{64}$$

where for each $\tau \in \Re$,

$$L_\tau(x, \zeta, \Gamma) := L(x, \zeta, \Gamma) + \tau\langle \Gamma, P_\beta P_\beta^T\rangle, \quad (x, \zeta, \Gamma) \in X \times \Re^m \times \mathcal{S}^p.$$

Let $\mathcal{M}_\tau(\bar{x})$ be the set consisting of all $(\zeta, \Gamma) \in \Re^m \times \mathcal{S}^p$ that satisfy (64). Thus, since $\mathrm{rank}(g(\bar{x}) + \tau P_\beta P_\beta^T) + \mathrm{rank}(\overline{\Gamma}) = p$ for any $\tau > 0$, the critical cone $C_\tau(\bar{x})$ of the parameterized problem (63) at $\bar{x}$ for $\tau > 0$ takes the form:

$$C_\tau(\bar{x}) = \{d \,:\, \mathcal{J}_x h(\bar{x})d = 0, \quad P_\gamma^T(\mathcal{J}_x g(\bar{x})d)P_\gamma = 0\} \supseteq \mathrm{app}(\bar{\zeta}, \overline{\Gamma}), \tag{65}$$

where $\mathrm{app}(\bar{\zeta}, \overline{\Gamma})$ satisfies (39). Therefore, by Lemma 3.1 and the second part of Theorem 3.1, we have for all $\tau > 0$ that

$$\sup_{(\zeta,\Gamma)\in\mathcal{M}_\tau(\bar{x})} \left\{ \left\langle d, \mathcal{J}_{xx}^2 L_\tau(\bar{x}, \zeta, \Gamma)d \right\rangle - \Upsilon_{\left(g(\bar{x})+\tau P_\beta P_\beta^T\right)}\left(\Gamma, \mathcal{J}_x g(\bar{x})d\right) \right\} > 0 \quad \forall\, d \in C_\tau(\bar{x})\backslash\{0\},$$

which, together with the fact that for any $(\zeta, \Gamma) \in \mathcal{M}_\tau(\bar{x})$,

$$\Upsilon_{\left(g(\bar{x})+\tau P_\beta P_\beta^T\right)}\left(\Gamma, \mathcal{J}_x g(\bar{x})d\right) = \Upsilon_{g(\bar{x})}\left(\Gamma, \mathcal{J}_x g(\bar{x})d\right) \quad \forall\, d \in \mathrm{app}(\bar{\zeta}, \overline{\Gamma}),$$

$\mathcal{J}_{xx}^2 L_\tau(\bar{x}, \zeta, \Gamma) = \mathcal{J}_{xx}^2 L(\bar{x}, \zeta, \Gamma)$, and (65), implies

$$\sup_{(\zeta,\Gamma)\in\mathcal{M}_\tau(\bar{x})} \left\{ \left\langle d, \mathcal{J}_{xx}^2 L(\bar{x}, \zeta, \Gamma)d \right\rangle - \Upsilon_{g(\bar{x})}\left(\Gamma, \mathcal{J}_x g(\bar{x})d\right) \right\} > 0 \quad \forall\, d \in \mathrm{app}(\bar{\zeta}, \overline{\Gamma})\backslash\{0\}. \tag{66}$$

Since for any $\tau > 0$, $\mathcal{M}_\tau(\bar{x}) \subseteq \mathcal{M}(\bar{x})$, we derive from (66) that

$$\sup_{(\zeta,\Gamma)\in\mathcal{M}(\bar{x})} \left\{ \left\langle d, \mathcal{J}_{xx}^2 L(\bar{x}, \zeta, \Gamma)d \right\rangle - \Upsilon_{g(\bar{x})}\left(\Gamma, \mathcal{J}_x g(\bar{x})d\right) \right\} > 0 \quad \forall\, d \in \mathrm{app}(\bar{\zeta}, \overline{\Gamma})\backslash\{0\}.$$

This shows that the strong second order sufficient condition (47) holds. □

Another important concept in sensitivity and stability analysis of the optimization problem (OP) is the *strong stability* of a stationary point, introduced by Kojima in [19]. Below the definition of strong stability is from Bonnans and Shapiro [6, Definition 5.33].

Definition 4.2 *Let $\bar{x}$ be a stationary point of the optimization problem (OP). We say that $\bar{x}$ is strongly stable with respect to a $\mathcal{C}^2$-smooth parameterization $(f(x,u), G(x,u))$ if there exist neighborhoods $\mathcal{V}_X$ of $\bar{x}$ and $\mathcal{V}_U$ of $\bar{u}$ such that for any $u \in \mathcal{V}_U$, the corresponding perturbed problem (OP_u) has a unique stationary point $x(u) \in \mathcal{V}_X$ and $x(\cdot)$ is continuous on $\mathcal{V}_U$. If this holds for any $\mathcal{C}^2$-smooth parameterization, we say that $\bar{x}$ is strongly stable.*

Let $F : X \times \Re^m \times \mathcal{S}^p \to X \times \Re^m \times \mathcal{S}^p$ be defined as in (51) and $(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \in X \times \Re^m \times \mathcal{S}^p$ be a KKT point of $(NLSDP)$. Then

$$F(\bar{x}, \bar{\zeta}, \overline{\Gamma}) = \begin{bmatrix} \mathcal{J}_x L(\bar{x}, \bar{\zeta}, \overline{\Gamma}) \\ -h(\bar{x}) \\ -g(\bar{x}) + \Pi_{\mathcal{S}_+^p}(g(\bar{x}) + \overline{\Gamma}) \end{bmatrix} = 0.$$

Let $A := g(\bar{x}) + \overline{\Gamma}$. Assume that A has the spectral decomposition as in (15) and $g(\bar{x})$ and $\overline{\Gamma}$ satisfy (34). Then, F is directionally differentiable at $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ and for any $\delta := (\delta_1, \delta_2, \delta_3) \in X \times \Re^m \times \mathcal{S}^p$,

$$\Phi(\delta) := F'(\bar{x}, \bar{\zeta}, \overline{\Gamma}; \delta) = \begin{bmatrix} \mathcal{J}_{xx}^2 L(\bar{x}, \bar{\zeta}, \overline{\Gamma})\delta_1 + \mathcal{J}_x h(\bar{x})^*\delta_2 + \mathcal{J}_x g(\bar{x})^*\delta_3 \\ -\mathcal{J}_x h(\bar{x})\delta_1 \\ -\mathcal{J}_x g(\bar{x})\delta_1 + \Pi'_{\mathcal{S}_+^p}(A; \mathcal{J}_x g(\bar{x})\delta_1 + \delta_3) \end{bmatrix}, \tag{67}$$

where $\Pi'_{\mathcal{S}_+^p}(A; \cdot)$ is given by (16). Since $\Phi(\cdot)$ is Lipschitz continuous, $\partial_B \Phi(0)$ is well defined.

Lemma 4.2 *Let $A = g(\bar{x}) + \overline{\Gamma}$ and Φ be defined by (67). It holds that*

$$\partial_B \Phi(0) = \partial_B F(\bar{x}, \bar{\zeta}, \overline{\Gamma}).$$

Proof. Define $\Xi : X \times \Re^m \times \mathcal{S}^p \to \mathcal{S}^p$ by

$$\Xi(\delta) := \Pi'_{\mathcal{S}_+^p}(A; \Psi(\delta)),$$

where $\Psi(\delta) := \mathcal{J}_x g(\bar{x})\delta_1 + \delta_3$, $\delta := (\delta_1, \delta_2, \delta_3) \in X \times \Re^m \times \mathcal{S}^p$. Then, by Lemma 2.1 and (c) of Proposition 2.1, we obtain that

$$\partial_B \Xi(0) = \partial_B \Pi_{\mathcal{S}_+^p}(A)\mathcal{J}_\delta \Psi(0),$$

which, together with Lemma 2.1, proves the conclusion of this lemma. □

Let $\mathrm{ind}(\phi, \bar{z})$ denote the index of a continuous function $\phi : Z \to Z$ at an isolated zero $\bar{z} \in Z$ used in degree theory (Lloyd [21], Ortega and Rheinboldt [25]). Now, we are ready to state the main result of this paper.

Theorem 4.1 *Let $\bar{x}$ be a locally optimal solution to the nonlinear semidefinite programming $(NLSDP)$. Suppose that Robinson's CQ* (41) *holds at $\bar{x}$ so that $\bar{x}$ is necessarily a stationary point of $(NLSDP)$. Let $(\bar{\zeta}, \overline{\Gamma}) \in \Re^m \times \mathcal{S}^p$ be such that $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ is a KKT point of $(NLSDP)$. Then the following statements are equivalent:*

(a) *The strong second order sufficient condition* (47) *holds at $\bar{x}$ and $\bar{x}$ is constraint nondegenerate.*

(b) *Any element in $\partial F(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ is nonsingular.*

(c) *The KKT point $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$ is a strongly regular solution of the generalized equation* (52).

(d) *The uniform second order growth condition holds at $\bar{x}$ and $\bar{x}$ is constraint nondegenerate.*

(e) *The point $\bar{x}$ is strongly stable and $\bar{x}$ is constraint nondegenerate.*

(f) *F is a locally Lipschitz homeomorphism near the KKT point $(\bar{x}, \bar{\zeta}, \overline{\Gamma})$.*

(g) *For every $V \in \partial_B F(\bar{x}, \bar{\zeta}, \overline{\Gamma})$, $\operatorname{sgn}\det V = \operatorname{ind}(F, (\bar{x}, \bar{\zeta}, \overline{\Gamma})) = \pm 1$.*

(h) *Φ is a globally Lipschitz homeomorphism.*

(i) *For every $V \in \partial_B \Phi(0)$, $\operatorname{sgn}\det V = \operatorname{ind}(\Phi, 0) = \pm 1$.*

(j) *Any element in $\partial \Phi(0)$ is nonsingular.*

Proof. We have already known from Proposition 3.2 that (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c) and from Remark 3.1 that (c) $\Longleftrightarrow$ (f). The relations (c) $\Longleftrightarrow$ (d) $\Longleftrightarrow$ (e) follow from Bonnans and Shapiro [6, Theorems 5.24 & 5.35]. Since $\Pi_{\mathcal{S}^p_+}(\cdot)$ is strongly semismooth everywhere (Sun and Sun [40]), F is a semismooth function (see Mifflin [24], Qi and Sun [27] for discussions on semismooth functions). Then, by Gowda [14, Theorem 3 & Corollary 4], we know that (f) $\Longleftrightarrow$ (g). Furthermore, by the semismoothness of $\Pi_{\mathcal{S}^p_+}(\cdot)$, Lemma 4.2, and Theorem 6 in Pang et al. [26], it holds that (g) $\Longleftrightarrow$ (h) $\Longleftrightarrow$ (i). By Lemma 4.2, the relation (b) $\Longleftrightarrow$ (j) holds. The proof of this theorem will be completed if one can show that (d) $\Longrightarrow$ (a). The latter, however, is implied by Lemma 4.1. □

Remark 4.1 *As mentioned in* Section 1*, the equivalence between* (a) *and* (c) *has already been known for the conventional nonlinear programming* (NLP)*. This equivalence for* (NLP) *follows from* Jongen et al. [16, Theorem 3.1]*,* Robinson [30, Theorem 4.1]*, and* Kojima [19, Corollary 6.6]*. For different proofs, see* Bonnans and Shapiro [6, Proposition 5.38]*,* Bonnans and Sulem [7, Theorem 4.10]*, and* Dontechev and Rockafellar [10, Theorem 6]*. By assuming $\bar{x}$ to be a stationary point* (*not necessarily a local optimal solution*)*, for* (NLP)*,* Jongen et al. [17] *proved* (b) $\Longleftrightarrow$ (e) *for a different but equivalent KKT system. By focusing on the local optimal solution case only, we extend these results in* Theorem 4.1 *from* (NLP) *to* $(NLSDP)$*.*

5. Conclusions In this paper, we discussed a strong second order sufficient condition for the nonlinear semidefinite programming $(NLSDP)$. This strong second order sufficient condition, together with constraint nondegeneracy, is shown to be equivalent to many conditions, notably the strong regularity of the KKT point and the nonsingularity of Clarke's Jacobian of the mapping F at the KKT point. There are many important questions not addressed in this paper. For example, it would be interesting to know whether these equivalent results given in Theorem 4.1 can be generalized to other $\mathcal{C}^2$-cone reducible sets.[3] Another possibility is to see which of these conditions are still equivalent to (b) or (j) if $\bar{x}$ is assumed to be a stationary point only.

Acknowledgments. We would like to thank the two anonymous referees for their constructive suggestions which helped improve the presentation of this paper. Thanks also go to Alexander Shapiro at Georgia Institute of Technology for bringing Shapiro [37] to our attention and Houduo Qi at University of Southampton and Liwei Zhang at Dalian University of Technology for their useful comments on an early version of this paper. The author's research was partly supported by Grant R146-000-061-112 from the National University of Singapore.

[3]In their newly published paper, Bonnans and Ramírez C. [3] gave a positive answer to nonlinear second order cone programming.

References

[1] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Sensitivity analysis of optimization problems under second order regularity constraints*, Mathematics of Operations Research **23** (1998), 803–832.

[2] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Second order optimality conditions based on parabolic second order tangent sets*, SIAM Journal on Optimization **9** (1999), 466–493.

[3] J. F. Bonnans and H. Ramírez C., *Perturbation analysis of second order cone programming problems*, Mathematical Programming, Series B **104** (2005), 205–227.

[4] J. F. Bonnans and H. Ramírez C., *Strong regularity of semidefinite programming problems*, Techincal Report CMM-DIM B-05/06-137, Department of Mathematical Engineering, Universida de Chile, June 2005.

[5] J. F. Bonnans and A. Shapiro, *Optimization problems with perturbations: A guided tour*, SIAM Review **40** (1998), 202–227.

[6] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[7] J. F. Bonnans and A. Sulem, *Pseudopower expansion of solutions of generalized equations and constrained optimization problems*, Mathematical Programming **70** (1995), 123–148.

[8] F. H. Clarke, *On the inverse function theorem*, Pacific Journal of Mathematics **64** (1976), 97–102.

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

[10] A. L. Dontchev and R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM Journal on Optimization **6** (1996), 1087–1105.

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

[12] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems: Volume I*, Springer, New York, 2003.

[13] J. Gauvin, *A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming*, Mathematical Programming **12** (1977), 136–138.

[14] M. S. Gowda, *Inverse and implicit function theorems for H-differentiable and semismooth functions*, Optimization Methods and Software **19** (2004), 443-461.

[15] N. J. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra and Its Applications **103** (1988), 103–118.

[16] H. Th. Jongen, T. Mobert, J. Rückmann, and K. Tammer, *On inertia and Schur complement in optimization*, Linear Algebra and Its Applications **95** (1987), 97–109.

[17] H. Th. Jongen, J. Rückmann, and K. Tammer, *Implicit functions and sensitivity of stationary points*, Mathematical Programming **49** (1990), 123–138.

[18] D. Klatte and B. Kummer, *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, Kluwer Academic Publishers, Boston, 2002.

[19] M. Kojima, *Strongly stable stationary solutions in nonlinear programs*, Analysis and Computation of Fixed Points (S. M. Robinson, ed.), Academic Press, New York, 1980, pp. 93-138.

[20] B. Kummer, *Lipschitzian inverse functions, directional derivatives, and applications in $C^{1,1}$-optimization*, Journal of Optimization Theory and Applications **70** (1991), 559–580.

[21] N. G. Lloyd, *Degree Theory*, Cambridge University Press, Cambridge, 1978.

[22] O. L. Mangasarian and S. Fromovitz, *The Fritz-John necessary optimality conditions in the presence of equality and inequality constraints*, Journal of Mathematical Analysis and Applications **7** (1967), 37–47.

[23] F. Meng, D. Sun, and G. Zhao, *Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization*, Mathematical Programming, Series B **104** (2005), 561–581.

[24] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM Journal on Control and Optimization **15** (1977), 957–972.

[25] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[26] J. S. Pang, D. Sun, and J. Sun, *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems*, Mathematics of Operations Research **28** (2003), 39–63.

[27] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical Programming **58** (1993), 353–367.

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

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

[30] S. M. Robinson, *Strongly regular generalized equations*, Mathematics of Operations Research **5** (1980), 43–62.

[31] S. M. Robinson, *Local structure of feasible sets in nonlinear programming, Part II: Nondegeneracy*, Mathematical Programming Study **22** (1984), 217–230.

[32] S. M. Robinson, *Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity*, Mathematical Programming Study **30** (1987), 45–66.

[33] S. M. Robinson, *An implicit-function theorem for a class of nonsmooth functions*, Mathematics of Operations Research **16** (1991), 292–309.

[34] S. M. Robinson, *Constraint nondegeneracy in variational analysis*, Mathematics of Operations Research **28** (2003), 201–232.

[35] R. T. Rockafellar, *Convex Analysis*, Princeton, New Jersey, 1970.

[36] R. T. Rockafellar and R. J. -B. Wets, *Variational Analysis*, Springer, New York, 1998.

[37] A. Shapiro, *Rank-Reducibility of a symmetric matrix and sampling theory of minimum trace factor analysis*, Psychometrika **47** (1982), 187–199.

[38] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Mathematical Programming, Series B **77** (1997), 301–320.

[39] A. Shapiro, *Sensitivity analysis of generalized equations*, Journal of Mathematical Sciences **115** (2003), 2554–2565.

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

[41] P. Tseng, *Merit functions for semi-definite complementarity problems*, Mathematical Programming **83** (1998), 159–185.

[42] E. H. Zarantonello, *Projections on convex sets in Hilbert space and spectral theory I and II.*, Contributions to Nonlinear Functional Analysis (E. H. Zarantonello, ed.), Academic Press, New York, 1971, pp. 237–424.