

## STRONG VARIATIONAL SUFFICIENCY FOR NONLINEAR SEMIDEFINITE PROGRAMMING AND ITS IMPLICATIONS\*

SHIWEI WANG<sup>†</sup>, CHAO DING<sup>‡</sup>, YANGJING ZHANG<sup>§</sup>, AND XINYUAN ZHAO<sup>¶</sup>

**Abstract.** Strong variational sufficiency is a newly proposed property, which turns out to be of great use in the convergence analysis of multiplier methods. However, what this property implies for nonpolyhedral problems remains a puzzle. In this paper, we prove the equivalence between the strong variational sufficiency and the strong second-order sufficient condition (SOSC) for nonlinear semidefinite programming (NLSDP) without requiring the uniqueness of the multiplier or any other constraint qualifications. Based on this characterization, the local convergence property of the augmented Lagrangian method (ALM) for NLSDP can be established under the strong SOSC in the absence of constraint qualifications. Moreover, under the strong SOSC, we can apply the semismooth Newton method to solve the ALM subproblems of NLSDP because the positive definiteness of the generalized Hessian of augmented Lagrangian function is satisfied.

**Key words.** strong variational sufficiency, nonlinear semidefinite programming, strong second-order sufficient condition, augmented Lagrangian method

**MSC codes.** 49J52, 90C22, 90C46

**DOI.** 10.1137/22M1530161

**1. Introduction.** The local optimality of the general optimization problem is a crucial topic for its theoretical importance and wide application. Traditionally, it is studied through the growth condition, e.g., the well-known first- or second-order optimality condition (see, e.g., [6]). Recently, Rockafellar [34] proposed a new property named strong variational sufficiency to deal with this topic geometrically. The key idea of this abstract definition originates from a so-called (strong) variational convexity, which indicates that the values and subgradients of a function are locally indistinguishable from those of a convex function. These two approaches seem to originate from different angles to understand the local optimality, but whether they possess deep connections is an essential issue, as it provides not only a better comprehension of optimization theory but also a solid theoretical foundation in algorithm design. Thus, an explicit characterization of strong variational sufficiency is demanding.

---

\*Received by the editors October 21, 2022; accepted for publication (in revised form) June 30, 2023; published electronically November 6, 2023.

<https://doi.org/10.1137/22M1530161>

**Funding:** The work of the second author was supported in part by the Beijing Natural Science Foundation (Z190002), the National Natural Science Foundation of China under project 12071464, and the CAS Project for Young Scientists in Basic Research (YSBR-034). The work of the third author was supported by the National Natural Science Foundation of China under project 12201617. The work of the fourth author was supported in part by the National Natural Science Foundation of China under projects 12271015 and 11871002.

<sup>†</sup>School of Mathematical Sciences, University of Chinese Academy of Science, Beijing, People's Republic of China and Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, People's Republic of China ([wangshiwei182@mails.ucas.ac.cn](mailto:wangshiwei182@mails.ucas.ac.cn)).

<sup>‡</sup>Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, People's Republic of China and School of Mathematical Sciences, University of Chinese Academy of Science, Beijing, People's Republic of China ([dingchao@amss.ac.cn](mailto:dingchao@amss.ac.cn)).

<sup>§</sup>Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, People's Republic of China ([yangjing.zhang@amss.ac.cn](mailto:yangjing.zhang@amss.ac.cn)).

<sup>¶</sup>Department of Mathematics, Beijing University of Technology, Beijing, People's Republic of China ([xyzhao@bjut.edu.cn](mailto:xyzhao@bjut.edu.cn)).

The definition of the (strong) variational sufficient condition (Definition 2.3) is officially given in [37] for the following general composite optimization problem:

$$(1.1) \quad \min_{x \in \mathbb{X}} f(x) + \theta(G(x)),$$

where  $\mathbb{X}$  and  $\mathbb{Y}$  are two given Euclidean spaces,  $f : \mathbb{X} \rightarrow \mathbb{R}$  and  $G : \mathbb{X} \rightarrow \mathbb{Y}$  are twice continuously differentiable, and  $\theta : \mathbb{Y} \rightarrow (-\infty, \infty]$  is a closed proper convex function. The strong variational sufficient condition is first introduced to deal with the local convexity of the primal augmented Lagrangian function and the augmented tilt stability. It has gained more and more attention for its mathematical elegance and wide applications in the convergence analysis of multiplier methods. Several characterizations of this abstract property are also given in [37]. For instance, it is equivalent to the positive definiteness of the Hessian bundle of the augmented Lagrangian function [37, Theorem 3] or the criterion [37, Theorem 5] involving the quadratic bundle (Definition 2.6).

In terms of the connection with traditional optimality conditions, Rockafellar [37, Theorem 4] shows that the strong variational sufficiency is equivalent to the well-known strong second-order sufficient condition (SOSC), which is expressed entirely via the program data, when the function  $\theta$  in (1.1) is polyhedral convex (i.e., the epigraph  $\text{epi } \theta$  of  $\theta$  is a polyhedral convex set). For nonpolyhedral problems, the equivalence is still valid if  $\theta$  in (1.1) is the indicator function of the second-order cone with  $G(\bar{x}) \neq 0$ , where  $\bar{x}$  is a local optimal solution (see [37, Example 3] for details). However, an explicit and verifiable characterization of the strong variational sufficient condition for general nonpolyhedral problems remains unknown. In this paper, without loss of generality, we mainly focus on the characterization of the strong variational sufficiency for the following nonlinear semidefinite programming (NLSDP):

$$(1.2) \quad \begin{array}{ll} \min_{x \in \mathbb{X}} & f(x) \\ \text{s.t.} & G(x) \in \mathbb{S}_+^n, \end{array}$$

where  $\mathbb{S}^n$  is the linear space of all  $n \times n$  real symmetric matrices equipped with the usual Frobenius inner product and its induced norm,  $\mathbb{S}_+^n$  ( $\mathbb{S}_-^n$ ), is the closed convex cone of all  $n \times n$  positive (negative) semidefinite matrices in  $\mathbb{S}^n$ . One of our contributions lies in uncovering the equivalence between the strong variational sufficient condition and the strong SOSC (see Definition 2.2 for details) for NLSDP problems without requiring any other constraint qualifications.

An important application of the strong variational sufficiency of NLSDP is the local convergence analysis of the augmented Lagrangian method (ALM). The ALM, which was first proposed by Hestenes and Powell in 1969, has attained fruitful achievements during the past fifty years. Especially, [33] uncovers the equivalence between the proximal point algorithm and the ALM for convex problems. This work is of great importance in establishing the Q-linear convergence rate of the dual sequence for the ALM. The conditions for the local linear convergence are further weakened (see, e.g., [24, 13]). For a nonconvex case, the properties built up under the convexity assumption have become inadequate. Many efforts have been made to explore the convergence properties of the ALM for nonconvex problems. See [4, 9, 10, 20, 17, 18] for more details. It is worth noting that, in [4], the author revealed a kind of local duality based on sufficient conditions for local optimality, which turns out to be the key to understanding the convergence of the ALM for nonconvex nonlinear programming. It follows that there may be a local reduction from nonconvex optimization to convex

optimization. This gives a hint on extending this local duality approach to broader nonconvex problems. Recently, the newly published work [38] opens new doors for the convergence rate analysis of the ALM for nonconvex problems by using the aforementioned characterization of the strong variational sufficient condition without any constraint qualifications.

For NLSDP (1.2), the Lagrangian function is defined by

$$(1.3) \quad L(x, Y) := f(x) + \langle Y, G(x) \rangle, \quad (x, Y) \in \mathbb{X} \times \mathbb{S}^n.$$

For any  $Y \in \mathbb{S}^n$ , denote the first-order and second-order derivatives of  $L(\cdot, Y)$  at  $x \in \mathbb{X}$  by  $L'_x(x, Y)$  and  $L''_{xx}(x, Y)$ , respectively. Given  $\rho > 0$ , the augmented Lagrangian function of (1.2) takes the following form (cf. [39, section 11.K] and [40]):

$$(1.4) \quad \mathcal{L}_\rho(x, Y) := f(x) + \frac{\rho}{2} \text{dist}^2 \left( G(x) + \frac{Y}{\rho}, \mathbb{S}_+^n \right) - \frac{\|Y\|^2}{2\rho},$$

where  $\text{dist}(z, \mathbb{S}_+^n)$  is the distance from point  $z$  to  $\mathbb{S}_+^n$ . For a given initial point  $(x^0, Y^0) \in \mathbb{X} \times \mathbb{S}^n$  and a constant  $\rho^0 > 0$ , the  $(k+1)$ th iteration of the (extended) ALM for NLSDP (1.2) proposed by [38] takes the following form:

$$(1.5) \quad \begin{cases} x^{k+1} \approx \arg \min \{ \mathcal{L}_{\rho^k}(x, Y^k) \}, \\ Y^{k+1} = Y^k + \tilde{\rho}^k [G(x^{k+1}) - \Pi_{\mathbb{S}_+^n} (G(x^{k+1}) + \frac{Y^k}{\tilde{\rho}^k})], \end{cases}$$

where  $\rho^k, \tilde{\rho}^k > 0$  and  $\Pi_{\mathbb{S}_+^n}(\cdot)$  is the metric projection onto  $\mathbb{S}_+^n$ . The (extended) ALM (1.5) reduces to the traditional ALM when  $\tilde{\rho}^k = \rho^k$ . To see how the ALM (1.5) works for nonconvex problems without constraint qualifications, consider the following simple example:

$$(1.6) \quad \begin{aligned} \min \quad & \frac{1}{2}x^3 \\ \text{s.t.} \quad & -x^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{S}_+^3. \end{aligned}$$

The optimal solution of (1.6) is  $\bar{x} = 0$  with the corresponding multiplier set  $\mathcal{M}(\bar{x}) := \{Y \mid Y \in \mathbb{S}_+^3\}$ . Due to the unboundedness of  $\mathcal{M}(\bar{x})$ , we know from [50, Theorem 4.1] that the Robinson constraint qualification [30] does not hold at  $\bar{x}$ . Pick a particular multiplier

$$\bar{Y} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \in \mathcal{M}(\bar{x}).$$

It is clear that  $L''_{xx}(\bar{x}, \bar{Y}) = 4 > 0$ , which implies that the strong SOSC (Definition 2.2) holds at  $(\bar{x}, \bar{Y})$ . We directly apply the ALM (Algorithm 4.1) to (1.6), and we find that the corresponding ALM subproblem in (1.5) can be solved exactly. Then, it can be observed from Figure 1 that, for fixed  $\rho^k$  and  $\tilde{\rho}^k$ ,  $\text{dist}(Y^k, \mathcal{M}(\bar{x}))$ , the distance between the  $k$ th iteration  $Y^k$  and  $\mathcal{M}(\bar{x})$  converges to zero linearly.

In this paper, combining [37] with our equivalence result, we obtain the Q-linear convergence of dual variables and the R-linear convergence of primal ones of the ALM for NLSDP under the merely strong SOSC, without requiring the uniqueness of the multiplier or any other constraint qualifications. It is worth noting that [44]

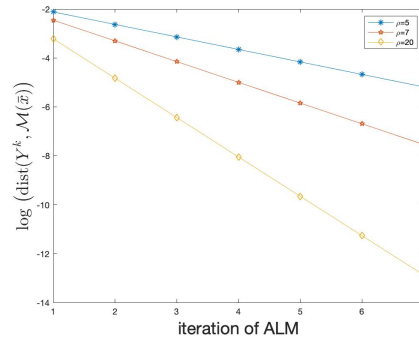


FIG. 1. The ALM for solving (1.6) with different penalty parameters  $\rho^k = \rho$ .

and [21] obtained the ALM local convergence result under the (strong) SOSC and nondegeneracy [41, Definition 3.2] or strict Robinson constraint qualification (SRCQ) (see, e.g., [15, equation (11)] for its definition) respectively, both of which imply the uniqueness of the multiplier. Although [48, Theorem 2] relaxed the uniqueness assumption of the multiplier, additional assumptions [48, assumption 1] with respect to multipliers are still needed.

Another important and practical issue in applying the ALM is how to solve the ALM subproblem (1.5). For convex problems, the successes of the Newton-CG augmented Lagrangian method for solving SDP (SDPNAL) [49], SDPNAL+ [45, 47], and a two-phase augmented Lagrangian method for solving convex quadratic SDP (QSDPNAL) [23] have proved the efficiency of the semismooth Newton-CG method. Its efficiency relies critically on the positive definiteness of a certain generalized Hessian of the augmented Lagrangian function. We are able to show the positive definiteness of the generalized Hessian under the strong SOSC, as a direct application of [37, Theorem 3] and our equivalence result.

The remaining parts of this paper are organized as follows. In the next section, we introduce some preliminary knowledge in semidefinite cone and strong variational sufficiency. In section 3, we propose our main equivalence result for both nonlinear second-order cone programming (NLSOC) and NLSDP. Section 4, as a direct application, copes with the convergence properties of the ALM for NLSDP. Moreover, how to solve the subproblems and the corresponding convergence analysis are discussed. In section 5, the numerical experiment of an example is given to show the validity of the aforementioned results. We conclude our paper and make some comments in the final section.

**2. Preliminaries.** For a given Euclidean space  $\mathbb{X}$ , let  $C$  be any subset in  $\mathbb{X}$ , and let  $\text{aff} C$  be the affine hull of  $C$  [32, Page 6]. The Bouligand tangent/contingent cone of  $C$  at  $x$  is a closed cone defined by

$$\mathcal{T}_C(x) := \{d \in \mathbb{X} \mid \exists t^k \downarrow 0 \text{ and } d^k \rightarrow d \text{ with } x + t^k d^k \in C \text{ for all } k\}.$$

When  $C$  is convex, the normal cone in the sense of convex analysis [32] is defined as

$$\mathcal{N}_C(x) = \{d \in \mathbb{X} \mid \langle d, x' - x \rangle \leq 0 \text{ for all } x' \in C\}.$$

For any  $x \in C$ , the critical cone associated with  $y \in \mathcal{N}_C(x)$  of  $C$  is defined in [16, Page 98]; i.e.,  $\mathcal{C}_C(x, y) = \mathcal{T}_C(x) \cap (y)^\perp$ .

For a set-valued mapping  $\mathcal{F} : \mathbb{X} \rightrightarrows \mathbb{X}$ , its limsup means that

$$\limsup_{x \rightarrow \bar{x}} \mathcal{F}(x) := \{d \in \mathbb{X} \mid \exists x_k \rightarrow \bar{x}, y_k \rightarrow d \text{ such that } y_k \in \mathcal{F}(x_k) \text{ for all } k\}.$$

The following definitions of proximal and (general/Mordukhovich limiting) subdifferentials of functions are adopted from [39, Definitions 8.45 and 8.3].

**DEFINITION 2.1.** Consider a function  $f : \mathbb{X} \rightarrow (-\infty, +\infty]$  and a point  $\bar{x}$  with  $f(\bar{x})$  finite. The proximal subdifferential of  $f$  at  $\bar{x}$  is defined as

$$\partial^\pi f(\bar{x}) := \{d \mid \exists l > 0, r > 0 \text{ such that } f(x) \geq f(\bar{x}) + \langle d, x - \bar{x} \rangle - l\|x - \bar{x}\|^2 \text{ for all } x \in \mathcal{B}_r(\bar{x})\},$$

where  $\mathcal{B}_r(\bar{x})$  is the ball centered at  $\bar{x}$  with radius  $r$ . The (general/Mordukhovich limiting) subdifferential of  $f$  at  $\bar{x}$  is  $\partial f(\bar{x}) := \limsup_{x \xrightarrow{f} \bar{x}} \partial^\pi f(x)$ , where  $x \xrightarrow{f} \bar{x}$  signifies  $x \rightarrow \bar{x}$  with  $f(x) \rightarrow f(\bar{x})$ .

Note that in an Euclidean space, the (general) subdifferential also can be constructed via the regular/Fréchet subdifferentials of functions  $\hat{\partial}f$  [39, Definition 8.3]. It is well known that these two definitions are equivalent (see, e.g., [39, Page 345] or [26, Theorem 1.89] for details). In particular, when  $f$  is convex and finite at the corresponding point, the proximal and (general) subdifferentials coincide with the subdifferential in the sense of convex analysis [32].

**2.1. Variational analysis of NLSDP.** Let  $A \in \mathbb{S}^n$  be given. We use  $\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A)$  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  $A$ . Let  $\Lambda(A) := \text{Diag}(\lambda(A))$ . Also, we use  $v_1(A) > \dots > v_d(A)$  to denote the different eigenvalues and  $\zeta_l := \{i \mid \lambda_i(A) = v_l(A)\}$ . Consider the eigenvalue decomposition of  $A$ ; i.e.,  $A = P\Lambda(A)P^T$ , where  $P \in \mathcal{O}^n(A)$  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:

$$(2.1) \quad A = \begin{bmatrix} P_\alpha & P_\beta & P_\gamma \end{bmatrix} \begin{bmatrix} \Lambda(A)_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda(A)_{\gamma\gamma} \end{bmatrix} \begin{bmatrix} P_\alpha^T \\ P_\beta^T \\ P_\gamma^T \end{bmatrix},$$

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

The critical cone of  $\mathbb{S}_+^n$  at  $Y \in \mathcal{N}_{\mathbb{S}_+^n}(X)$  with  $A = X + Y$  is given by (see [8, equations (11) and (12)])

$$(2.2) \quad \mathcal{C}_{\mathbb{S}_+^n}(X, Y) = \{H \in \mathbb{S}^n \mid P_\beta^T H P_\beta \in \mathbb{S}_+^{|\beta|}, P_\beta^T H P_\gamma = 0, P_\gamma^T H P_\gamma = 0\},$$

and the affine hull of  $\mathcal{C}_{\mathbb{S}_+^n}(X, Y)$  is

$$(2.3) \quad \text{aff } \mathcal{C}_{\mathbb{S}_+^n}(X, Y) = \{H \in \mathbb{S}^n \mid P_\beta^T H P_\gamma = 0, P_\gamma^T H P_\gamma = 0\}.$$

The NLSDP (1.2) can be written in the following form:

$$(2.4) \quad \min f(x) + \delta_{\mathbb{S}_+^n}(G(x)),$$

where  $\delta_{\mathbb{S}_+^n} : \mathbb{S}^n \rightarrow (-\infty, \infty]$  is the indicator function of positive semidefinite cone  $\mathbb{S}_+^n$ . Denote  $S_{KKT}(a, b)$  as the solution set of the KKT optimality condition for problem

(2.4); i.e.,

$$(2.5) \quad S_{KKT}(a, b) = \left\{ (x, Y) \in \mathbb{X} \times \mathbb{S}^n \mid \begin{array}{l} L'_x(x, Y) - a = 0, \\ \mathbb{S}_+^n \ni (G(x) - b) \perp Y \in \mathbb{S}_-^n \end{array} \right\}.$$

For any KKT pair  $(\bar{x}, \bar{Y})$  that satisfies the KKT condition with  $(a, b) = (0, 0)$ , we call  $\bar{x}$  a stationary point. Suppose  $\bar{x}$  is a stationary point. Define  $\mathcal{M}(\bar{x})$  as the set of all multipliers  $Y \in \mathbb{S}^n$  satisfying the KKT condition (2.5); i.e.,

$$(2.6) \quad \mathcal{M}(\bar{x}) = \{Y \in \mathbb{S}^n \mid (\bar{x}, Y) \in S_{KKT}(0, 0)\}.$$

The definition of the strong SOSOC for NLSDP, which demands that the supremum of (2.7) over  $Y \in \mathcal{M}(\bar{x})$  holds, is originally given by [41] as an analogue for Robinson [31] for the nonlinear programming (NLP). However, the strong SOSOC mentioned here is slightly different in only requiring the validity of (2.7) at  $\bar{Y}$ . It is worth noting that, when  $\mathcal{M}(\bar{x})$  is a singleton, both of them are the same.

DEFINITION 2.2 (see [41, Definition 3.2]). *Let  $\bar{x}$  be a stationary point of NLSDP (2.4), and let  $\bar{Y} \in \mathcal{M}(\bar{x})$ . We say that the strong SOSOC holds at  $(\bar{x}, \bar{Y})$  if*

$$(2.7) \quad \langle L''_{xx}(\bar{x}, \bar{Y})d, d \rangle - \Upsilon_{G(\bar{x})}(\bar{Y}, G'(\bar{x})d) > 0 \quad \text{for all } 0 \neq G'(\bar{x})d \in \text{aff } \mathcal{C}_{\mathbb{S}_+^n}(G(\bar{x}), \bar{Y}),$$

where  $\Upsilon_{G(\bar{x})}(\bar{Y}, G'(\bar{x})d) = 2\langle \bar{Y}, (G'(\bar{x})d)G(\bar{x})^\dagger(G'(\bar{x})d) \rangle$  is the  $\sigma$ -term and  $G(\bar{x})^\dagger$  is the generalized inverse matrix of  $G(\bar{x})$ .

Suppose that  $A = G(\bar{x}) + \bar{Y}$  possesses the decomposition (2.1). From [8, Page 386], we know the  $\sigma$ -term takes the explicit form of

$$(2.8) \quad \Upsilon_{G(\bar{x})}(\bar{Y}, H) = 2 \sum_{i \in \alpha, j \in \gamma} \frac{\lambda_j(A)}{\lambda_i(A)} (\tilde{H}_{ij})^2,$$

where  $\tilde{H} = P^T H P$ .

Next, we introduce some essential notation for our main result. The (second-order) generalized differentiability of the augmented Lagrangian function  $\mathcal{L}_\rho$  defined by (1.4) for NLSDP (2.4) has been explicitly established in [44]. In fact, it is well known that  $\mathcal{L}_\rho$  is continuously differentiable with

$$(2.9) \quad (\mathcal{L}_\rho)'_x(x, Y) = f'(x) + \rho G'(x)^* \Pi_{\mathbb{S}_-^n}(G(x) + \rho^{-1}Y),$$

where  $\Pi_{\mathbb{S}_-^n}(\cdot)$  is the metric projection over  $\mathbb{S}_-^n$  and  $G'(x)^*$  denotes the adjoint of the corresponding linear mapping. Moreover, since  $(\mathcal{L}_\rho)'_x$  is Lipschitz continuous, by using Rademacher's theorem, the B(ouligand)-subdifferential of  $(\mathcal{L}_\rho)'_x$  at  $(x, Y)$  is given by

$$(2.10) \quad \partial_B((\mathcal{L}_\rho)'_x)(x, Y) := \left\{ \lim_{k \rightarrow \infty} ((\mathcal{L}_\rho)'_x)'(x^k, Y^k) \mid (x^k, Y^k) \in \mathcal{U}, (x^k, Y^k) \rightarrow (x, Y) \right\},$$

where  $\mathcal{U}$  is the set of Fréchet-differentiable points of  $(\mathcal{L}_\rho)'_x$ . It follows from [44, equation (18)] (using [44, Lemmas 2 and 3]) that

$$\begin{aligned} & \pi_x \partial_B((\mathcal{L}_\rho)'_x)(x, Y)(\Delta x) \\ &= L''_{xx}(x, \rho \Pi_{\mathbb{S}_-^n}(G(x) + \rho^{-1}Y))(\Delta x) + \rho G'(x)^* \partial_B \Pi_{\mathbb{S}_-^n}(G(x) + \rho^{-1}Y) G'(x)(\Delta x), \end{aligned}$$

where  $\pi_x \partial_B((\mathcal{L}_\rho)'_x)(x, Y)$  is the projection of  $\partial_B((\mathcal{L}_\rho)'_x)(x, Y)$  onto the space  $\mathbb{X}$  and  $\partial_B \Pi_{\mathbb{S}_-^n}(G(x) + \rho^{-1}Y)$  is the B-subdifferential of  $\Pi_{\mathbb{S}_-^n}(\cdot)$  at  $G(x) + \rho^{-1}Y$ . Actually, the

set  $\pi_x \partial_B((\mathcal{L}_\rho)'_x)(x, Y)$  was recently defined as the  $x$  part of the Hessian bundle [37, equation (3.1)] for the augmented Lagrangian function  $\mathcal{L}_\rho$  at  $(x, Y)$  (see [37, equation (3.6)] for detail).

Let  $\bar{x}$  be the stationary point of (2.4). For each  $\bar{Y} \in \mathcal{M}(\bar{x})$  and  $W \in \partial_B \Pi_{\mathbb{S}^n}(G(\bar{x}) + \rho^{-1}\bar{Y})$ , define the following mapping:

$$(2.11) \quad \mathcal{A}_\rho(\bar{Y}, W) := L''_{xx}(\bar{x}, \bar{Y}) + \rho G'(\bar{x})^* W G'(\bar{x}).$$

Let  $\bar{Z} = G(\bar{x}) + \bar{Y}$  possesses decomposition (2.1) with  $\bar{P} \in \mathcal{O}^n(\bar{Z})$ . It follows from [44, Lemma 5] that  $W \in \partial_B \Pi_{\mathbb{S}^n}(\bar{Z})$  if and only if there exists  $W_0 \in \partial_B \Pi_{\mathbb{S}^{|\beta|}}(0)$  such that, for all  $H \in \mathbb{S}^n$ ,

$$W(H) = \bar{P} \begin{bmatrix} 0 & 0 & \Sigma_{\alpha\gamma} \circ \bar{P}_\alpha^T H \bar{P}_\gamma \\ 0 & W_0(\bar{P}_\beta^T H \bar{P}_\beta) & \bar{P}_\beta^T H \bar{P}_\gamma \\ \Sigma_{\gamma\alpha} \circ \bar{P}_\gamma^T H \bar{P}_\alpha & \bar{P}_\gamma^T H \bar{P}_\beta & \bar{P}_\gamma^T H \bar{P}_\gamma \end{bmatrix} \bar{P}^T,$$

where

$$\begin{cases} \Sigma_{ij} = 1 - \frac{\max\{\lambda_i, 0\} + \max\{\lambda_j, 0\}}{|\lambda_i| + |\lambda_j|}, & (i, j) \notin \beta \times \beta, \\ \Sigma_{ij} \in [0, 1], & (i, j) \in \beta \times \beta, \end{cases}$$

with  $\lambda_i := \lambda_i(\bar{Z})$  for short. Let  $\mathcal{Q}^n$  denote the set of the  $n$ -dimensional orthogonal matrix. Also,  $W_0 \in \partial_B \Pi_{\mathbb{S}^{|\beta|}}(0)$  if and only if there exist  $Q \in \mathcal{Q}^{|\beta|}$  and  $\Omega \in \mathbb{S}^{|\beta|}$  with entries  $\Omega_{ij} \in [0, 1]$  such that, for all  $H \in \mathbb{S}^{|\beta|}$ ,

$$W_0(H) = Q(\Omega \circ (Q^T H Q))Q^T.$$

By [44, Lemma 9], we have

$$(2.12) \quad \begin{aligned} \langle d, \mathcal{A}_\rho(\bar{Y}, W)d \rangle &= \langle d, L''_{xx}(\bar{x}, \bar{Y})d \rangle + \rho \sum_{i,j \in \gamma} (\bar{P}^T (G'(\bar{x})d) \bar{P})_{ij}^2 + 2\rho \sum_{i \in \beta, j \in \gamma} (\bar{P}^T (G'(\bar{x})d) \bar{P})_{ij}^2 \\ &\quad + 2\rho \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j}{\rho\lambda_i - \lambda_j} (\bar{P}^T (G'(\bar{x})d) \bar{P})_{ij}^2 + \rho \sum_{i,j \in \beta} (\Omega_\rho)_{ij} (P^T (G'(\bar{x})d) P)_{ij}^2, \end{aligned}$$

where  $\Omega_\rho \in \mathbb{S}^{|\beta|}$  with entries  $(\Omega_\rho)_{ij} \in [0, 1]$  and  $P = [\bar{P}_\alpha \bar{P}_\beta Q \bar{P}_\gamma]$  with  $Q \in \mathcal{Q}^{|\beta|}$ .

**2.2. The strong variational sufficiency.** The definition of the (strong) variational sufficient condition is given in [37] for the general composite optimization problem (1.1). We can recast (1.1) in the form

$$(2.13) \quad \min \phi(x, u) \text{ subject to } u = 0, \quad \text{where } \phi(x, u) = f(x) + \theta(G(x) + u).$$

The first-order local optimality condition for (1.1) of  $\bar{x}$  is the existence of  $\bar{Y}$  such that

$$L'_x(\bar{x}, \bar{Y}) = 0 \quad \text{with } \bar{Y} \in \partial \theta(G(\bar{x})).$$

Define

$$(2.14) \quad \phi_r := \phi(x, u) + \frac{r}{2}|u|^2.$$

The variational (strong) convexity, which is firstly proposed in [34], refers to the existence of open convex neighborhoods  $\mathcal{W}$  of  $(\bar{x}, 0)$  and  $\mathcal{Z}$  of  $(0, \bar{Y})$  such that there exists a proper closed (strongly) convex function  $\psi \leq \phi_r/2$  on  $\mathcal{W}$  such that

$$(\mathcal{W} \times \mathcal{Z}) \cap \text{gph } \partial \psi = (\mathcal{W} \times \mathcal{Z}) \cap \text{gph } \partial \phi_r,$$

and for  $(x, u; v, y)$  belonging to this common set,  $\psi(x, u) = \phi_r(x, u)$ .

DEFINITION 2.3 (see [37]). *The (strong) variational sufficient condition for local optimality in (2.13) holds with respect to  $\bar{x}$  and  $\bar{Y}$  satisfying the first-order condition if there exists  $r > 0$  such that  $\phi_r(x, u)$  is variationally (strongly) convex with respect to the pair  $((\bar{x}, 0), (0, \bar{Y}))$  in  $\text{gph } \partial\phi_r$ .*

To characterize the above property, the tools of the second subderivative and generalized quadratic form are needed.

DEFINITION 2.4 (see [39, Definition 13.6]). *Let  $\bar{x}$  be a point where the function  $f: \mathbb{X} \rightarrow [-\infty, +\infty]$  is finite.  $f$  is twice epidifferentiable at  $\bar{x}$  for  $v$  if the functions*

$$\Delta_t^2 f(\bar{x} | v)(u) = \frac{f(\bar{x} + tu) - f(\bar{x}) - t\langle v, u \rangle}{\frac{1}{2}t^2}$$

*epiconverge to  $d^2 f(\bar{x} | v)$  as  $t \downarrow 0$ , where  $d^2 f(\bar{x} | v)$  is the second subderivative of  $f$  at  $\bar{x}$  for  $v$  defined as*

$$d^2 f(\bar{x} | v)(w) = \liminf_{t \downarrow 0, u \rightarrow w} \Delta_t^2 f(\bar{x} | v)(u).$$

We know from [25, Theorem 3.6] that  $\delta_{\mathbb{S}_+^n}$  is twice epidifferentiable at  $X$  for  $Y$  with  $Y \in \mathcal{N}_{\mathbb{S}_+^n}(X)$ .

DEFINITION 2.5 (see [37]). *A generalized linear mapping  $\mathcal{R}$  from  $\mathbb{X}$  to  $\mathbb{Y}$  is a set-valued mapping for which  $\text{gph } \mathcal{R}$  is a subspace of  $\mathbb{X} \times \mathbb{Y}$ . This means that  $\text{dom } \mathcal{R}$  is a subspace  $\mathbb{Z}$  of  $\mathbb{X}$ ,  $\mathcal{R}(0)$  is a subspace  $\mathbb{Z}'$  of  $\mathbb{Y}$ , and there is an ordinary linear mapping  $\mathcal{R}_0: \mathbb{Z} \rightarrow \mathbb{Y}$  such that  $\mathcal{R}(x) = \mathcal{R}_0(x) + \mathbb{Z}'$  for  $x \in \mathbb{Z}$ . We call that a function  $q: \mathbb{X} \rightarrow (-\infty, +\infty]$  is a generalized quadratic form on  $\mathbb{X}$  if  $q(0) = 0$  and if the subgradient mapping  $\partial q: \mathbb{X} \rightrightarrows \mathbb{X}$  is generalized linear. A function  $g$  on  $\mathbb{X}$  will be called generalized twice differentiable at  $x$  for a subgradient  $y$  if it is twice epidifferentiable at  $x$  for  $y$ , with the second-order subderivative  $d^2 g(x | y)$  being a generalized quadratic form.*

The following definition of the quadratic bundle is taken from [37], which is essential in characterizing the strong variational sufficiency.

DEFINITION 2.6. *For the general optimization problem (1.1), suppose that  $(\bar{x}, \bar{Y})$  is a KKT pair. The quadratic bundle of  $\theta$  is defined as*

$$\text{quad } \theta(G(\bar{x}) | \bar{Y}) = \begin{cases} \text{the collection of generalized quadratic forms } q \text{ for which} \\ \exists (X^k, Y^k) \rightarrow (G(\bar{x}), \bar{Y}) \text{ with } \theta \text{ generalized twice differentiable} \\ \text{at } X^k \text{ for } Y^k \text{ and such that the generalized quadratic} \\ \text{forms } q_k = \frac{1}{2} d^2 \theta(X^k | Y^k) \text{ converge epigraphically to } q. \end{cases}$$

As mentioned in [37], the variational sufficient condition guarantees the local optimality for (2.13). The following result is taken from [37, Theorem 5], which is useful for the subsequent analysis.

PROPOSITION 2.7. *For general optimization problem (1.1), strong variational sufficient condition for local optimality with respect to  $(\bar{x}, \bar{Y})$  is equivalent to every  $q \in \text{quad } \theta(G(\bar{x}) | \bar{Y})$  having*

$$(2.15) \quad \frac{1}{2} \langle L''_{xx}(\bar{x}, \bar{Y})d, d \rangle + q(G'(\bar{x})d) > 0 \quad \text{when } d \neq 0,$$

*where  $\text{quad } \theta(G(\bar{x}) | \bar{Y})$  is the quadratic bundle of  $\theta$ .*



**3. The characterization of the strong variational sufficient condition for NLSDP.** In this section, we will study the relationship between the strong variational sufficient condition and the well-known strong SOSOC (2.7) for NLSDP by combining (2.15) and [25, Theorem 3.3] together. First, we use NLSOC as an example to illustrate our approach. The general NLSOC can be written as

$$(3.1) \quad \begin{aligned} & \min_{x \in \mathbb{X}} f(x) \\ & \text{s.t. } g(x) \in \mathcal{K}, \end{aligned}$$

where  $f, g$  are both twice continuously differentiable and  $\mathcal{K}$  is the second-order cone defined as  $\mathcal{K} = \{(u_1, \dots, u_m) \in \mathbb{R}^m \mid u_1 \geq \|(u_2, \dots, u_m)\|\} = \{u \mid h(u) \leq 0\}$  with  $h(u) = -u_1 + \|(u_2, \dots, u_m)\|$ . The ordinary Lagrangian function of (3.1) is defined by

$$(3.2) \quad L(x, y) := f(x) + \langle y, g(x) \rangle, \quad (x, y) \in \mathbb{X} \times \mathbb{R}^m.$$

Given a stationary point  $\bar{x}$ , let

$$\mathcal{M}_{\text{soc}}(\bar{x}) := \left\{ y \in \mathbb{R}^m \mid \begin{array}{l} L'_x(\bar{x}, y) = 0, \\ \mathcal{K} \ni g(\bar{x}) \perp y \in \mathcal{K}^\circ \end{array} \right\}$$

be the set of all multipliers  $y \in \mathbb{R}^m$  satisfying the KKT condition for (3.1), where  $\mathcal{K}^\circ$  is the polar cone of  $\mathcal{K}$  defined in [32, section 14]. The strong SOSOC at  $(\bar{x}, \bar{y})$  is defined as

$$(3.3) \quad \langle L''_{xx}(\bar{x}, \bar{y})d, d \rangle - \Upsilon_{g(\bar{x})}(\bar{y}, g'(\bar{x})d) > 0 \quad \text{for all } 0 \neq g'(\bar{x})d \in \text{aff } \mathcal{C}_{\mathcal{K}}(g(\bar{x}), \bar{y}),$$

where the explicit form of  $\Upsilon_{g(\bar{x})}(\bar{y}, g'(\bar{x})d)$  is given in [5, Theorem 29].

As illustrated in [37, Example 3], the explicit form of the quadratic bundle quad  $\delta_{\mathcal{K}}(g(\bar{x}) \mid \bar{y})$  for certain reference KKT point  $(\bar{x}, \bar{y})$  can be obtained by using corresponding results for polyhedral problems. For example, if  $g(\bar{x}) \in \text{int } \mathcal{K}$ , we have  $\bar{y} = 0$ . It can be checked directly by using [25, Theorem 3.3] and [5, equations (36) and (43)] that the quadratic bundle quad  $\delta_{\mathcal{K}}(g(\bar{x}) \mid \bar{y})$  consists of solely  $q \equiv 0$ . If  $g(\bar{x}) \in \text{bd } \mathcal{K} \setminus \{0\}$ , for any nonzero  $\bar{y} \in \mathcal{N}_{\mathcal{K}}(g(\bar{x}))$ , the quadratic bundle quad  $\delta_{\mathcal{K}}(g(\bar{x}) \mid \bar{y})$  consists of the generalized quadratic form

$$(3.4) \quad q = \frac{1}{2} d^2 \delta_{\mathcal{K}}(g(\bar{x}) \mid \bar{y}) \quad \text{with} \quad q(w) = \begin{cases} \frac{1}{2} w \cdot h''(g(\bar{x}))w & \text{if } h'(g(\bar{x}))w = 0, \\ \infty & \text{otherwise.} \end{cases}$$

If  $g(\bar{x}) = 0$  and  $\bar{y} \in \text{int}(-\mathcal{K})$ , the quadratic bundle consists of  $q = \delta_{\{0\}}$ . If  $g(\bar{x}) = 0$  and  $\bar{y} \in \text{bd}(-\mathcal{K}) \setminus \{0\}$ , both  $\delta_{\{0\}}$  and (3.4) constitute the quadratic bundle. It can be checked directly that for NLSOC, the strong variational sufficient condition is equivalent to the strong SOSOC when the reference point  $(\bar{x}, \bar{y})$  lies in one of the above circumstances.

Thus, only two circumstances for NLSOC remain to be discussed. The first one is  $g(\bar{x}) = 0$  and  $\bar{y} = 0$ . Since  $\mathcal{K}$  is  $\mathcal{C}^2$ -cone reducible, we know from [25, Theorem 3.3] and [5, Theorem 29] that, for any  $(g^k, y^k) \rightarrow (g(\bar{x}), \bar{y})$  and  $w \in \mathbb{R}^m$ ,

$$(3.5) \quad d^2 \delta_{\mathcal{K}}(g^k \mid y^k)(w) = w^T \mathcal{H}(g^k, y^k)w + \delta_{\mathcal{C}_{\mathcal{K}}(g^k, y^k)}(w) \geq 0,$$

where  $\mathcal{H}(g^k, y^k) = \frac{y^k}{g_1^k} (g^k)'^T (1 \ 0^T; 0 \ -I_{m-1})(g^k)'$  if  $g^k \in \text{bd } \mathcal{K} \setminus \{0\}$ , and  $\mathcal{H}(g^k, y^k) = 0$  otherwise. If we pick  $g^k \in \text{int } \mathcal{K} \rightarrow g(\bar{x})$  and  $y^k = 0$  for all  $k$ , it is easy to see from [5, Theorem 29 and equations (35) and (36)] that

$$d^2 \delta_{\mathcal{K}}(g^k \mid y^k)(w) = 0 + \delta_{\mathcal{C}_{\mathcal{K}}(g^k, y^k)}(w),$$

where  $\mathcal{C}_{\mathcal{K}}(g^k, y^k) = \mathbb{R}^m$ . Thus, we have proved  $0 \in \text{quad } \delta_{\mathcal{K}}(g(\bar{x}) \mid \bar{y})$ . Combining this with (3.5), it follows that, in this case, (2.15) is equivalent to the strong SOSOC.

The second case is  $g(\bar{x}) \in \text{bd}\mathcal{K} \setminus \{0\}$  and  $\bar{y} = 0$ . Let  $g^k \in \text{int } \mathcal{K} \rightarrow g(\bar{x})$  and  $y^k = 0$  for each  $k$ . We know that

$$d^2 \delta_{\mathcal{K}}(g^k \mid y^k)(w) = 0 + \delta_{\mathcal{C}_{\mathcal{K}}(g^k, y^k)}(w),$$

where  $\mathcal{C}_{\mathcal{K}}(g^k, y^k) = \mathbb{R}^m$ . As  $k \rightarrow \infty$ , we know that  $\lim_{k \rightarrow \infty} d^2 \delta_{\mathcal{K}}(g^k \mid y^k)(w) = 0$ . Using [5, equations (35), (36), and (43)] again, we also have that (2.15) is equivalent to the strong SOSOC.

Thus, we immediately obtain the following characterization of strong variational sufficiency for NLSOC, which is a supplement for [37, Example 3].

**PROPOSITION 3.1.** *Let  $\bar{x} \in \mathbb{X}$  be a stationary point to the NLSOC (3.1), and let  $\bar{y} \in \mathcal{M}(\bar{x})$ . The strong variational sufficient condition (2.15) with respect to  $(\bar{x}, \bar{y})$  holds if and only if the strong SOSOC (3.3) holds at  $(\bar{x}, \bar{y})$ .*

It is worth noting that the above proposition is not surprising. If  $g(\bar{x}) \neq 0$ , we can obtain this result by regarding the NLSOC problem as a polyhedral problem as shown in [37, Example 3]. If  $g(\bar{x}) = 0$ , the  $\sigma$ -term happens to be 0, which makes the calculation of the quadratic bundle much simpler. When it comes to the NLSDP problem (2.4), things are not so easy as we can neither regard SDP as a polyhedral problem nor have a simplification for its  $\sigma$ -term (2.8). However, the success of this approach to NLSOC gives us a hint that this approach may also work for NLSDP. Before we put forward our main result, we need the following proposition on the quadratic bundle defined by Definition 2.6 for  $\delta_{\mathbb{S}_+^n}$ .

**PROPOSITION 3.2.** *Let  $\bar{x} \in \mathbb{X}$  be a stationary point to NLSDP (2.4), and let  $\bar{Y} \in \mathcal{M}(\bar{x})$ , where  $\mathcal{M}(\bar{x})$  is given in (2.6). Let  $A = G(\bar{x}) + \bar{Y}$ , which possesses the decomposition (2.1). Then, there exists  $q \in \text{quad } \delta_{\mathbb{S}_+^n}(G(\bar{x}) \mid \bar{Y})$  such that, for all  $H \in \mathbb{S}^n$ ,*

$$\begin{aligned} q(H) &= -\frac{1}{2} \mathcal{Y}_{G(\bar{x})}(\bar{Y}, H) + \delta_{\text{aff } \mathcal{C}_{\mathbb{S}_+^n}(G(\bar{x}), \bar{Y})}(H) \\ (3.6) \quad &= \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j(A)}{\lambda_i(A)} (\tilde{H}_{ij})^2 + \delta_{\text{aff } \mathcal{C}_{\mathbb{S}_+^n}(G(\bar{x}), \bar{Y})}(H), \end{aligned}$$

where  $\tilde{H} = P^T H P$ .

*Proof.* For each  $k$ , choose

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

where  $z^k \downarrow 0$  (this notation means that, for each  $k$ ,  $z^k > 0$  and  $z^k \rightarrow 0$  as  $k \rightarrow \infty$ ), with each  $(z^k)_i$  nonincreasing on  $k$ . Let  $A^k = X^k + Y^k$  for each  $k$ . Since  $\mathbb{S}_+^n$  is  $\mathcal{C}^2$ -cone reducible [6, Example 3.140], we know from [25, Theorem 3.3], (2.2), and (2.8) that, for any  $H \in \mathbb{S}^n$ ,

$$\begin{aligned} \frac{1}{2}d^2\delta_{\mathbb{S}_+^n}(X^k | Y^k)(H) &= \sum_{i \in \alpha \cup \beta, j \in \gamma} \frac{-\lambda_j(A^k)}{\lambda_i(A^k)} (\tilde{H}_{ij})^2 + \delta_{\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k)}(H) \\ &= \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j(A)}{\lambda_i(A)} (\tilde{H}_{ij})^2 + \sum_{i \in \beta, j \in \gamma} \frac{-\lambda_j(A)}{\lambda_i(X^k)} (\tilde{H}_{ij})^2 \\ &\quad + \delta_{\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k)}(H), \end{aligned}$$

where  $\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k) = \{H \in \mathbb{S}^n \mid \tilde{H}_{\gamma\gamma} = 0\}$ . It is worth noting that  $d^2\delta_{\mathbb{S}_+^n}(X^k | Y^k)(H)$  is a generalized quadratic form by using Definition 2.5, as  $d^2\delta_{\mathbb{S}_+^n}(X^k | Y^k)(0) = 0$  and  $\partial d^2\delta_{\mathbb{S}_+^n}(X^k | Y^k)(H) = \mathcal{R}(H) + \mathcal{N}_{\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k)}(H)$ , with  $\mathcal{N}_{\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k)}(H)$  being a subspace and

$$\mathcal{R}(H)(\Delta H) = -4 \sum_{i \in \alpha, j \in \gamma} \frac{\lambda_j(A)}{\lambda_i(A)} (\tilde{H})_{ij} \cdot (P^T \Delta H P)_{ij} - 4 \sum_{i \in \beta, j \in \gamma} \frac{\lambda_j(A)}{\lambda_i(A)} (\tilde{H})_{ij} \cdot (P^T \Delta H P)_{ij}$$

being linear on  $H$ . This also implies that  $\delta_{\mathbb{S}_+^n}$  is generalized twice differentiable at  $X^k$  for  $Y^k$ .

Let  $f_1^k(H) = \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j(A)}{\lambda_i(A)} (\tilde{H}_{ij})^2$ ,  $f_2^k(H) = \delta_{\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k)}(H)$ , and  $f_3^k(H) = \sum_{i \in \beta, j \in \gamma} \frac{-\lambda_j(A)}{\lambda_i(X^k)} (\tilde{H}_{ij})^2$ . By the definition of continuous convergence mentioned in [39, Page 250], we know that  $f_1^k$  converges continuously to  $f_1$  with

$$f_1(H) = \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j(A)}{\lambda_i(A)} (\tilde{H}_{ij})^2.$$

It follows from [39, Theorem 7.11] that  $f_1^k$  epiconverges to  $f_1$ . It can be checked easily that  $f_1^k$  also pointwise converges [39, Page 239] to  $f_1$ . Since  $\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k)$  is a constant closed set, we know that  $f_2^k$  pointwise converges and epiconverges [39, Proposition 4.4] to  $f_2$  with  $f_2(H) = \delta_{\mathcal{C}_{\mathbb{S}_+^n}(X^k, Y^k)}(H)$ . It follows from [39, Theorem 7.46(a)] that  $f_1^k + f_2^k$  epiconverges to  $f_1 + f_2$ . Also, it can be checked directly that  $f_1^k + f_2^k$  pointwise converges to  $f_1 + f_2$ . By the construction of  $z^k$ , the sequence of  $\{f_3^k\}$  is nondecreasing ( $f_3^k \leq f_3^{k+1}$ ). We know from [39, Proposition 7.4] that  $f_3^k$  epiconverges to  $\sup_k \{cl f_3^k\}$ , where  $cl f_3^k$  is the closure of  $f_3^k$ . It is easy to see that  $\sup_k \{cl f_3^k\} = \sup_k \{f_3^k\} = \delta_{\mathcal{V}}$  with  $\mathcal{V} = \{H \in \mathbb{S}^n \mid \tilde{H}_{\beta\gamma} = 0\}$  and  $f_3^k$  pointwise converges to  $\delta_{\mathcal{V}}$ . Since  $f_1^k + f_2^k$  pointwise and epiconverges to  $f_1 + f_2$ , by using [39, Theorem 7.46(a)] again, we obtain that  $f_1^k + f_2^k + f_3^k$  epiconverges to  $f_1 + f_2 + f_3$ .

Combining the above discussion with (2.3), the generalized quadratic form  $\frac{1}{2}d^2\delta_{\mathbb{S}_+^n}(X^k | Y^k)(H)$  converges epigraphically to the generalized quadratic form

$$\sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j(A)}{\lambda_i(A)} (\tilde{H}_{ij})^2 + \delta_{\text{aff } \mathcal{C}_{\mathbb{S}_+^n}(G(\bar{x}), \bar{Y})}(H).$$

Thus, we have verified this proposition.  $\square$

The following result is on the explicit characterization of the strong variational sufficiency of local optimality for NLSDP, which is the main result of this paper.

**THEOREM 3.3.** *Let  $\bar{x} \in \mathbb{X}$  be a stationary point to the NLSDP (2.4), and let  $\bar{Y} \in \mathcal{M}(\bar{x})$ . Then, the following three conditions are equivalent:*

- (i) *the strong variational sufficient condition with respect to  $(\bar{x}, \bar{Y})$  holds;*

- (ii) the strong SOSC (2.7) holds at  $(\bar{x}, \bar{Y})$ ;
- (iii) there exist  $\rho_0 > 0$  and  $\underline{\eta} > 0$  such that, for any  $\rho \geq \rho_0$  and any  $W \in \partial_B \Pi_{\mathbb{S}^n}(G(\bar{x}) + \rho^{-1}\bar{Y})$ ,

$$\langle d, \mathcal{A}_\rho(\bar{Y}, W)d \rangle \geq \underline{\eta} \|d\|^2 \quad \text{for all } d \in \mathbb{X},$$

where  $\mathcal{A}_\rho(\bar{Y}, W)$  is defined by (2.11).

*Proof.* “(i)  $\implies$  (ii)”: This direction can be obtained directly from Proposition 3.2 and Proposition 2.7 as we only need to substitute (3.6) into (2.15).

“(ii)  $\implies$  (iii)”: This proof sketch is similar to [44, Proposition 4], although they require the validity of nondegeneracy, which is superfluous here. It follows from Definition 2.7 that there exists  $\eta_0 > 0$  such that

$$\langle L''_{xx}(\bar{x}, \bar{Y})d, d \rangle - \Upsilon_{G(\bar{x})}(\bar{Y}, G'(\bar{x})d) \geq \eta_0 \|d\|^2 \quad \text{for all } G'(\bar{x})d \in \text{aff} \mathcal{C}_{\mathbb{S}^n_+}(G(\bar{x}), \bar{Y}).$$

By using [44, Lemma 7], there exist two positive numbers  $\rho_1$  and  $\underline{\eta} \in (0, \eta_0/2]$  such that, for any  $\rho' \geq \rho_1$ ,

$$\begin{aligned} & \langle L''_{xx}(\bar{x}, \bar{Y})d, d \rangle - \Upsilon_{G(\bar{x})}(\bar{Y}, G'(\bar{x})d) \\ & + \rho' \|\bar{P}_\gamma^T(G'(\bar{x})d)\bar{P}_\gamma\|^2 + \rho' \|\bar{P}_\beta^T(G'(\bar{x})d)\bar{P}_\gamma\|^2 \geq 2\underline{\eta} \|d\|^2 \quad \text{for all } d \in \mathbb{X}. \end{aligned}$$

Suppose a sufficiently large  $\rho_0 \geq \rho_1$ . For any  $\rho \geq \rho_0$  and  $d \in \mathcal{X}$ , we have

$$\begin{aligned} & -\Upsilon_{G(\bar{x})}(\bar{Y}, G'(\bar{x})d) - 2\rho \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j}{\rho\lambda_i - \lambda_j} (\bar{P}^T(G'(\bar{x})d)\bar{P})_{ij}^2 \\ & = 2 \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j}{\lambda_i} (\bar{P}^T(G'(\bar{x})d)\bar{P})_{ij}^2 - 2\rho \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j}{\rho\lambda_i - \lambda_j} (\bar{P}^T(G'(\bar{x})d)\bar{P})_{ij}^2 \\ & = 2 \sum_{i \in \alpha, j \in \gamma} \frac{\lambda_j^2}{\lambda_i(\rho\lambda_i - \lambda_j)} (\bar{P}^T(G'(\bar{x})d)\bar{P})_{ij}^2 \leq \underline{\eta} \|d\|^2, \end{aligned}$$

where the last inequality can be obtained by sufficiently large  $\rho$ . Combining the above two relations together, we have, for any  $\rho \geq \rho_0, \rho' \geq \rho_0 \geq \rho_1$ ,

$$\begin{aligned} & \langle L''_{xx}(\bar{x}, \bar{Y})d, d \rangle + 2\rho \sum_{i \in \alpha, j \in \gamma} \frac{-\lambda_j}{\rho\lambda_i - \lambda_j} (\bar{P}^T(G'(\bar{x})d)\bar{P})_{ij}^2 \\ & + \rho' \|\bar{P}_\gamma^T(G'(\bar{x})d)\bar{P}_\gamma\|^2 + \rho' \|\bar{P}_\beta^T(G'(\bar{x})d)\bar{P}_\gamma\|^2 \geq \underline{\eta} \|d\|^2. \end{aligned}$$

Then, it can be checked directly that, for all  $d \in \mathbb{X}$  and  $\rho \geq \rho_0, \langle d, \mathcal{A}_\rho(\bar{Y}, W)d \rangle \geq \underline{\eta} \|d\|^2$ .

“(iii)  $\implies$  (i)”: It follows directly from [37, Theorem 3] because  $\mathcal{A}_\rho(\bar{Y}, W)$  coincides with [37, equation (3.6)].  $\square$

*Remark 3.4.* Next, we will discuss the relationship between tilt stability and variationally strong convexity/the strong variational sufficient condition. As shown in [35], for the general optimization problem, the variationally strong convexity implies the tilt stability [35, Definition 3] of the local minimizer. However, it is worth to note that tilt stability usually does not imply variationally strong convexity, as explained in [22, Remark 2.8] and [35]. For NLSDP, the general objective function of (1.2) is amenable under the Robinson constraint qualification (RCQ) [30]. It follows from [22, Proposition 2.9] that variationally strong convexity is equivalent to tilt stability under

the RCQ for NLSDP. For more information on the characterization of variationally strong convexity, the readers may refer to [37, 38, 22].

As an application of the tilt stability to the minimization of  $\phi_\rho$  (2.14), Rockafellar defines the so-called augmented tilt stability [37, equation (2.18)]. In [37, Theorem 2], Rockafellar also shows that augmented tilt stability is equivalent to the strong variational sufficient condition, i.e., the variationally strong convexity of  $\phi_\rho$ , without constraint qualifications. By using Theorem 3.3, the strong SOSC is also equivalent to the augmented tilt stability.

Recently, Khanh et al. [22, Theorem 6.5] also provide a sufficient condition for (strong) variational sufficiency of a strongly amenable problem in the form of (1.1). However, the strong variational sufficiency mentioned in [22] is different from the one used here. In [22], the strong variational sufficiency is defined as the variationally strong convexity of the function  $x \rightarrow f(x) + \theta(G(x))$ , while this paper focuses on that of the perturbed function  $(x, u) \rightarrow f(x) + \theta(G(x) + u) + \frac{\tau}{2}\|u\|^2$ . In their result, they require a condition named second-order qualification condition [27, equation (3.15)] at  $\bar{x}$ ; i.e.,

$$(3.7) \quad \ker G'(\bar{x})^* \cap \partial^2\theta(G(\bar{x}), \bar{Y})(0) = \{0\},$$

where  $\partial^2\theta(G(\bar{x}), \bar{Y})$  is defined in [27, Definition 2.1]. It follows from [14, Theorem 3.1] that (3.7) is equivalent to the nondegeneracy [41, Definition 3.3] for NLSDP. By combining [22, Theorem 6.2], Theorem 3.3, and [7, Theorem 5.6 and Lemma 6.3] together, we know that under the nondegenerate condition, the two strong variational sufficiencies are the same for NLSDP. However, their relationship without the nondegeneracy condition remains unclear.

Moreover, by using [46, Example 2.2], it can be proved in a similar manner to [37, Theorem 3] that the variational sufficiency of local optimality for NLSDP (2.4) is equivalent to the existence of  $\rho_0 > 0$  and a convex open neighborhood  $\mathcal{V}$  of  $(\bar{x}, \bar{Y})$  such that, for any  $\rho \geq \rho_0$ ,  $(x, Y) \in \mathcal{V}$  and any  $W \in \partial_B \Pi_{\mathbb{S}^n}(G(x) + \rho^{-1}Y)$ ,

$$\langle d, \mathcal{A}_\rho(Y, W)d \rangle \geq 0 \quad \text{for all } d \in \mathbb{X}.$$

However, whether the variational sufficiency of the local optimality is equivalent to a certain second-order optimality condition without any constraint qualification for NLSDP (2.4) is still a future consideration that we are working on.

**4. The semismooth Newton-CG based ALM for nonconvex NLSDP.** In this section, we apply the main result of Theorem 3.3 to study the local convergence rate of the (extended) ALM (1.5) for solving NLSDP. The detailed algorithm is stated in Algorithm 4.1. By [38, Theorems 1.1 and 1.2], the strong variational sufficient condition with respect to  $(\bar{x}, \bar{Y})$  for local optimality holds if and only if there exist  $\bar{\rho} > 0$  and a closed convex neighborhood  $\mathcal{X} \times \mathcal{Y}$  of  $(\bar{x}, \bar{Y})$  such that, for all  $\rho \geq \bar{\rho}$ ,  $\mathcal{L}_\rho(x, Y)$  is strongly convex in  $x \in \mathcal{X}$  with modulus  $s > 0^1$  for all  $Y \in \mathcal{Y}$  and concave in  $Y \in \mathcal{Y}$  for all  $x \in \mathcal{X}$ .

In the above algorithm, subproblem (4.1) is solved inexactly. Three increasing tightness stopping criteria for the updating of  $x^{k+1}$  are illustrated in [38, equation (1.15)]:

<sup>1</sup>A function  $\psi: \mathbb{X} \rightarrow \mathbb{R}$  is said to be strongly convex with modulus  $s > 0$  if  $\psi - \frac{s}{2}\|\cdot\|^2$  is convex.

---

**Algorithm 4.1** the ALM for solving (2.4).

---

**Require:** Let  $(x^0, Y^0) \in \mathbb{X} \times \mathbb{S}^n$  and  $\rho^0 > \bar{\rho}$ . Set  $k := 0$ .

- 1: If  $(x^k, Y^k)$  satisfies a suitable termination criterion: STOP.
- 2: Compute  $x^{k+1}$  such that

$$(4.1) \quad x^{k+1} \approx \bar{x}^{k+1} = \arg \min_{x \in \mathcal{X}} \mathcal{L}_{\rho^k}(x, Y^k).$$

- 3: Update the vector of multipliers to

$$Y^{k+1} := Y^k + \tilde{\rho}^k [G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + Y^k / \rho^k)],$$

where  $\tilde{\rho}^k = \rho^k - \bar{\rho}$ .

- 4: Update nondecreasing positive sequence  $\rho^{k+1}$  according to certain rules.
  - 5: Set  $k \leftarrow k + 1$  and go to 1.
- 

(4.2)

$$(2\tilde{\rho}^k [\mathcal{L}_{\rho^k}(x^{k+1}, Y^k) - \inf_{\mathcal{X}} \mathcal{L}_{\rho^k}(\cdot, Y^k)])^{1/2} \leq \begin{cases} \text{(a)} \ \epsilon_k, \\ \text{(b)} \ \epsilon_k \min\{1, \|\tilde{\rho}^k (\mathcal{L}_{\rho^k})'_Y(x^{k+1}, Y^k)\|\}, \\ \text{(c)} \ \epsilon_k \min\{1, \|\tilde{\rho}^k (\mathcal{L}_{\rho^k})'_Y(x^{k+1}, Y^k)\|^2\}, \end{cases}$$

$$\text{with } \epsilon_k \in (0, 1) \sum_{k=0}^{\infty} \epsilon_k = \sigma < \infty \text{ and } \rho^k \rightarrow \rho^\infty \leq \infty.$$

It is worth to note that (4.2)(c) is first introduced in [36] to support linear convergence in partnership with strong variational sufficiency. It follows from [38, Theorem 3.1] that (4.2) can be replaced by

$$(4.3) \quad \sqrt{\tilde{\rho}^k} \|(\mathcal{L}_{\rho^k})'_x(x^{k+1}, Y^k)\| \leq \begin{cases} \text{(a)} \ \epsilon'_k, \\ \text{(b)} \ \epsilon'_k \min\{1, \|\tilde{\rho}^k (\mathcal{L}_{\rho^k})'_Y(x^{k+1}, Y^k)\|\}, \\ \text{(c)} \ \epsilon'_k \min\{1, \|\tilde{\rho}^k (\mathcal{L}_{\rho^k})'_Y(x^{k+1}, Y^k)\|^2\}, \end{cases}$$

where  $\epsilon'_k = \epsilon_k \sqrt{s}$ , as the strong convexity of  $\mathcal{L}_{\rho^k}(\cdot, Y^k)$  with modulus  $s$ , guarantees  $\mathcal{L}_{\rho^k}(x^{k+1}, Y^k) - \inf_{\mathcal{X}} \mathcal{L}_{\rho^k}(\cdot, Y^k) \leq \frac{1}{2s} \|(\mathcal{L}_{\rho^k})'_x(x^{k+1}, Y^k)\|^2$ .

By applying the local duality, which comes from the strong variational sufficient condition through [37, Theorem 1], we suppose that  $S_{KKT}(0, 0) \cap \mathcal{X} \times \mathcal{Y} \neq \emptyset$ , where  $\mathcal{X} \times \mathcal{Y}$  is the neighborhood mentioned at the beginning of this section. As mentioned in [38, Pages 9–10], we can define the associated local primal and dual problems to  $\mathcal{X} \times \mathcal{Y}$  in the following sense. The associated local primal problem is

$$(4.4) \quad \min \hat{f}(x) := \sup_{Y \in \mathcal{Y}} \mathcal{L}_{\bar{\rho}}(x, Y) \quad \text{over } x \in \mathcal{X}.$$

The associated local dual problem is

$$(4.5) \quad \max \hat{h}(Y) := \inf_{x \in \mathcal{X}} \mathcal{L}_{\bar{\rho}}(x, Y) \quad \text{over } Y \in \mathcal{Y}.$$

In [38, Theorem 2.1], the author reveals the connection between (2.4), (4.4), and (4.5), which is of great use in the following discussion.

**4.1. Convergence analysis of the ALM.** By taking advantage of Theorem 3.3 and the genius work [38], we can view the convergence analysis of the ALM for NLSDP as a direct extension. To explore this topic, we need the definition of bounded linear regularity of a collection of closed convex sets, which can be found in, e.g., [2, Definition 5.6].

**DEFINITION 4.1.** *Let  $D_1, D_2, \dots, D_m \subseteq \mathbb{X}$  be closed convex sets for some positive integer  $m$ . Suppose that  $D := D_1 \cap D_2 \cap \dots \cap D_m$  is nonempty. The collection  $\{D_1, D_2, \dots, D_m\}$  is said to be boundedly linearly regular if, for every bounded set  $\mathcal{B} \subseteq \mathbb{X}$ , there exists a constant  $\kappa > 0$  such that*

$$\text{dist}(x, D) \leq \kappa \max \{ \text{dist}(x, D_1), \dots, \text{dist}(x, D_m) \} \quad \text{for all } x \in \mathcal{B}.$$

A sufficient condition to guarantee the property of bounded linear regularity is established in [3, Corollary 3]. Denote

$$(4.6) \quad \mathcal{G}_1(\bar{x}) = \{Y \in \mathbb{S}^n \mid f'(\bar{x}) + G'(\bar{x})^* Y = 0\} \text{ and } \mathcal{G}_2(\bar{x}) = \{Y \in \mathbb{S}^n \mid Y \in \mathcal{N}_{\mathbb{S}_+^n}(G(\bar{x}))\}.$$

It is easy to see that  $\mathcal{G}_1(\bar{x})$  is a polyhedron and  $\mathcal{G}_2(\bar{x})$  is convex.

For a stationary point  $\bar{x}$  of NLSDP (2.4), define

$$(4.7) \quad \xi(G'(\bar{x})) = \min \{ \|G'(\bar{x})^* \eta\| : \eta \in \mathcal{G}_1(\bar{x})^\perp, \|\eta\| = 1 \}.$$

The following condition is adopted from [38] as it is essential in verifying the convergence rate.

**CONDITION 4.2.** *There exist  $b > 0$  and  $\varepsilon > 0$  such that the local dual problem (4.5) satisfies  $\hat{h}(Y) \leq \max_{\mathcal{Y}} \hat{h} - b \text{dist}^2(Y, \mathcal{H})$  when  $\|Y - \bar{Y}\| < \varepsilon$ , where  $\mathcal{Y}$  is the closed convex neighborhood of  $\bar{Y}$  given at the beginning of section 4,  $\mathcal{H} := \arg \max_{\mathcal{Y}} \hat{h}$ .*

The following result, which is originally proposed in [38, Theorem 4.2] for the polyhedral case, provides the sufficiency of condition 4.2. By applying the bounded linear regularity [3], one can obtain the results in a similar approach as that of [38, Theorem 4.2] directly. We omit it here for simplicity. Moreover, it follows from [38, Page 34] that condition 4.2 trivially holds if  $G'(\bar{x}) = 0$ .

**PROPOSITION 4.3.** *Let  $\bar{x} \in \mathbb{X}$  be a stationary point to the NLSDP (2.4), and let  $\bar{Y} \in \mathcal{M}(\bar{x})$ , where  $\mathcal{M}(\bar{x})$  is given in (2.6). Suppose that  $G'(\bar{x}) \neq 0$ . If the strong SOSC with respect to  $(\bar{x}, \bar{Y})$  holds and the collection  $\{\mathcal{G}_1(\bar{x}), \mathcal{G}_2(\bar{x})\}$  is boundedly linearly regular, where  $\mathcal{G}_1(\bar{x})$  and  $\mathcal{G}_2(\bar{x})$  are given in (4.6), then we have  $\xi(G'(\bar{x})) > 0$ , where  $\xi(G'(\bar{x}))$  is defined by (4.7) and condition 4.2 holds for*

$$(4.8) \quad b = \frac{\kappa}{a_2 + a_1} \text{ with } a_2 = b_0^{-1} + 2\bar{\rho} \text{ and } a_1 = \frac{2\|L''_{xx}(\bar{x}, \bar{Y}) + \hat{\rho}I\|}{\xi(G'(\bar{x}))^2},$$

where  $\hat{\rho} = \lambda_{\max}(\bar{\rho}G'(\bar{x})^*G'(\bar{x}) + \varepsilon')$ ,  $\varepsilon'$  is some positive constant, and  $b_0$  is the quadratic parameter  $\kappa$  given in [12, Proposition 2.1].

**Remark 4.4.** It follows from [12, Proposition 3.2] (see also [11, Proposition 17]) that the bounded linear regularity of the collection  $\{\mathcal{G}_1(\bar{x}), \mathcal{G}_2(\bar{x})\}$  holds under one of the following two conditions:

- (i)  $\mathcal{G}_2(\bar{x})$  is a polyhedron; i.e.,  $|\gamma| \geq n - 1$ , where  $\gamma$  is the negative eigenvalue index set of matrix  $G(\bar{x}) + \bar{Y}$ ;

- (ii) there exists a strict complementarity KKT pair  $(\bar{x}, \tilde{Y})$  ( $\tilde{Y}$  does not have to be  $\bar{Y}$ ); i.e.,  $\text{rank}(G(\bar{x})) + \text{rank}(\tilde{Y}) = n$ .

Next, we shall present the local convergence result of the ALM for NLSDP. We say that a sequence  $z^k > 0$  converges Q-linearly to 0 at a rate  $c$  if  $\limsup_{k \rightarrow \infty} \frac{z^{k+1}}{z^k} \leq c < \infty$ . When  $c = 0$ , we say  $z^k$  converges Q-superlinearly to 0. Moreover, a sequence  $y^k > 0$  converges R-linearly to 0 at a rate  $c$  if  $y^k \leq z^k$  with  $z^k > 0$  converges Q-linearly to 0 at that rate. The following closedness condition relative to the closed convex set  $\mathcal{M}(\bar{x})$  is taken from [38, Theorem 2.2]. Recall that  $\mathcal{X}, \mathcal{Y}$  are the closed convex neighborhood of  $\bar{x}, \bar{Y}$  mentioned in the beginning of section 4.

CONDITION 4.5. *We say the initial point  $Y^0$  and  $\sigma > 0$  in (4.2) satisfies the following closedness condition relative to the closed convex set  $\mathcal{M}(\bar{x})$  (2.6) if there exists  $\eta > \text{dist}(Y^0, \mathcal{M}(\bar{x})) + \sigma$  such that*

$$\{Y \mid \|Y - Y^0\| \leq 3\eta\} \subset \mathcal{Y}.$$

It is easy to see that this condition indicates  $Y^0$  to be sufficiently close to  $\mathcal{M}(\bar{x})$  and  $\bar{Y}$ . Also, the computations of subproblems at each iteration need to be sufficiently accurate. If subproblems are solved exactly, a sufficient condition for the above one is that there exists  $\eta > 0$  such that  $\mathcal{B}_\eta(Y^0) \subseteq \text{int } \mathcal{Y}$  and  $\text{dist}(Y^0, \mathcal{M}(\bar{x})) \leq \eta/3$ .

By using Theorem 3.3 and [38, Theorems 2.2, 2.3, 3.1, and 3.2], we immediately get the following local convergence result of the ALM for solving NLSDP.

THEOREM 4.6. *Let  $\bar{x} \in \mathcal{X}$  be a stationary point to the NLSDP (2.4), and let  $\bar{Y} \in \mathcal{M}(\bar{x})$ . Suppose that the strong SOSC (2.7) holds at  $(\bar{x}, \bar{Y})$ . Let the initial point  $Y^0$  and  $\sigma$  in (4.2) satisfy condition 4.5. Suppose that the set  $\{x \mid -(\mathcal{L}_\rho)'_x(x, Y) \in \mathcal{N}_\mathcal{X}(x)\}$  is nonempty and bounded when  $Y \in \text{int } \mathcal{Y}$ .*

- (i) *Under stopping criterion (4.3)(a), we have the sequence  $\{Y^k\}$  converge within  $\text{int } \mathcal{Y}$  to a particular  $\hat{Y} \in \text{int } \mathcal{Y}$ . Moreover, both  $x^k$  and  $\bar{x}^k$  converge to  $\bar{x}$ .*
- (ii) *The stopping criterion in (i) is strengthened into (4.3)(b), and suppose that  $\{\mathcal{G}_1(\bar{x}), \mathcal{G}_2(\bar{x})\}$  is also boundedly linearly regular, where  $\mathcal{G}_1(\bar{x})$  and  $\mathcal{G}_2(\bar{x})$  are given in (4.6). Then, we have  $\text{dist}(Y^k, \mathcal{M}(\bar{x})) \rightarrow 0$  with*

$$\text{dist}(Y^{k+1}, \mathcal{M}(\bar{x})) \leq \frac{1}{\sqrt{1 + b^2(\rho^\infty)^2}} \text{dist}(Y^k, \mathcal{M}(\bar{x}))$$

and  $\bar{x}^k \rightarrow \bar{x}$  with

$$\|\bar{x}^k - \bar{x}\| \leq \frac{1}{s} \text{dist}(Y^k, \mathcal{M}(\bar{x})),$$

where  $\bar{x}^k$  is the exact solution of subproblems in (4.1),  $b$  is given in condition 4.2, and  $s$  is the strong convexity modulus of  $\mathcal{L}_\rho(x, Y)$  on  $x$  with  $\rho \geq \bar{\rho}$ .

- (iii) *If the stopping criterion in (ii) is strengthened into (4.3)(c), we have  $Y^k \rightarrow \hat{Y}$  with*

$$\|Y^{k+1} - \hat{Y}\| \leq \frac{1}{\sqrt{1 + b^2(\rho^\infty)^2}} \|Y^k - \hat{Y}\|.$$

Moreover, if the stopping criterion is further supplemented by

$$\|(\mathcal{L}_{\rho^k})'_x(x^{k+1}, Y^k)\| \leq c \|Y^{k+1} - Y^k\| \text{ for some fixed } c,$$

we have  $x^k \rightarrow \bar{x}$  with

$$\|x^k - \bar{x}\| \leq p \|Y^k - \hat{Y}\|$$

for some  $p > 0$ .



As illustrated in [38, Theorem 2.3], the condition “set  $\{x \mid -(\mathcal{L}_{\bar{\rho}})'_x(x, Y) \in \mathcal{N}_{\mathcal{X}}(x)\}$  is nonempty and bounded when  $Y \in \text{int } \mathcal{Y}$ ” can be reduced to “the existence of  $Y \in \text{int } \mathcal{Y}$  such that  $\{x \mid -(\mathcal{L}_{\bar{\rho}})'_x(x, Y) \in \mathcal{N}_{\mathcal{X}}(x)\}$  is nonempty and bounded”. This condition is trivially satisfied as  $\mathcal{X}$  is a neighborhood of  $\bar{x}$  and  $(\bar{x}, \bar{Y})$  belongs to the set. It is worth noting that, as illustrated in [38, Theorem 2.3], the result in Theorem 4.6(i) only requires variational sufficiency. Under variational sufficiency, the convergence of  $x^k$  to  $\bar{x}$  can not be obtained. Meanwhile, under the stopping criterion (4.3)(b), we may not be able to obtain from Theorem 4.6(ii) the convergence of the primal iteration sequence  $\{x^k\}$ , since the exact solution  $\bar{x}^k$  of the subproblems in (4.1)  $\bar{x}^k$  is unknown in practice. Next, we shall show that the KKT residual of NLSDP (2.4) also converges R-linearly, which means that the KKT residual can be used as a verifiable stopping criterion for ALM. Its proof sketch is inspired by [13, Theorem 2].

PROPOSITION 4.7. *Suppose the conditions in Theorem 4.6(ii) hold. Define the following residual function*

$$(4.9) \quad R(x, Y) := \|L'_x(x, Y)\| + \|G(x) - \Pi_{\mathbb{S}_+^n}(G(x) + Y)\|.$$

Then, for  $k$  sufficiently large, if  $\sqrt{\tilde{\rho}^k} \epsilon_k < 1$ , we have that there exists  $a > 0$  such that

$$R(x^{k+1}, Y^{k+1}) \leq c^k \text{dist}(Y^k, \mathcal{M}(\bar{x})),$$

with  $c^k = (\epsilon'_k \sqrt{\tilde{\rho}^k} + (1 + \bar{\rho} + a\bar{\rho})(\tilde{\rho}^k)^{-1})(1 - \sqrt{\tilde{\rho}^k} \epsilon_k)^{-1}$ .

*Proof.* Let  $\hat{Y}^{k+1} = Y^k + \rho^k(G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + Y^k) + (\rho^k)^{-1}Y^k)$  and  $z^{k+1} = \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + (\rho^k)^{-1}Y^k)$ . We have  $G(x^{k+1}) - z^{k+1} = (\tilde{\rho}^k)^{-1}(Y^{k+1} - Y^k)$ . It follows directly from (4.3)(b) that, for each  $k$ ,

$$\|L'_x(x^{k+1}, \hat{Y}^{k+1})\| = \|(\mathcal{L}_{\rho^k})'_x(x^{k+1}, Y^k)\| \leq \epsilon'_k \sqrt{\tilde{\rho}^k} \|Y^{k+1} - Y^k\|.$$

Then, there exists  $a > 0$  such that

$$(4.10) \quad \begin{aligned} \|L'_x(x^{k+1}, Y^{k+1})\| &= \|L'_x(x^{k+1}, \hat{Y}^{k+1}) + G'(x^{k+1})^*(Y^{k+1} - \hat{Y}^{k+1})\| \\ &\leq \|L'_x(x^{k+1}, \hat{Y}^{k+1})\| + \|G'(x^{k+1})^*(Y^{k+1} - \hat{Y}^{k+1})\| \\ &\leq \|L'_x(x^{k+1}, \hat{Y}^{k+1})\| + a \|Y^{k+1} - \hat{Y}^{k+1}\| \\ &= \|L'_x(x^{k+1}, \hat{Y}^{k+1})\| + a\bar{\rho} \|G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + (\rho^k)^{-1}Y^k)\| \\ &= \|L'_x(x^{k+1}, \hat{Y}^{k+1})\| + a\bar{\rho}(\tilde{\rho}^k)^{-1} \|Y^{k+1} - Y^k\| \leq (\epsilon'_k \sqrt{\tilde{\rho}^k} + a\bar{\rho}(\tilde{\rho}^k)^{-1}) \|Y^{k+1} - Y^k\|, \end{aligned}$$

where the second inequality follows from the twice differentiable continuity and the boundedness of  $x^k$  obtained from Theorem 4.6(a). It can be verified directly from [39, Theorem 2.26] that  $\hat{Y}^{k+1} \in \partial\delta_{\mathbb{S}_+^n}(z^{k+1})$  and

$$\begin{aligned} &\|G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + \hat{Y}^{k+1})\| \\ &= \|G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + \hat{Y}^{k+1})\| - \|z^{k+1} - \Pi_{\mathbb{S}_+^n}(\hat{Y}^{k+1} + z^{k+1})\| \\ &\leq \|G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + \hat{Y}^{k+1}) - (z^{k+1} - \Pi_{\mathbb{S}_+^n}(\hat{Y}^{k+1} + z^{k+1}))\| \\ &\leq \|G(x^{k+1}) - z^{k+1}\| = (\tilde{\rho}^k)^{-1} \|Y^{k+1} - Y^k\|. \end{aligned}$$

It then follows that

$$\begin{aligned}
 (4.11) \quad & \|G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + Y^{k+1})\| \\
 & \leq \|G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + \widehat{Y}^{k+1})\| \\
 & \quad + \|\Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + \widehat{Y}^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + Y^{k+1})\| \\
 & \leq (\widehat{\rho}^k)^{-1} \|Y^{k+1} - Y^k\| + \|\widehat{Y}^{k+1} - Y^{k+1}\| \\
 & = (\widehat{\rho}^k)^{-1} \|Y^{k+1} - Y^k\| + \bar{\rho} \|G(x^{k+1}) - \Pi_{\mathbb{S}_+^n}(G(x^{k+1}) + (\rho^k)^{-1}Y^k)\| \\
 & = (\widehat{\rho}^k)^{-1} \|Y^{k+1} - Y^k\| + \bar{\rho}(\widehat{\rho}^k)^{-1} \|Y^{k+1} - Y^k\| = (1 + \bar{\rho})(\widehat{\rho}^k)^{-1} \|Y^{k+1} - Y^k\|.
 \end{aligned}$$

Combining (4.10) and (4.11) together, we obtain

$$(4.12) \quad R(x^{k+1}, Y^{k+1}) \leq (\epsilon'_k \sqrt{\widehat{\rho}^k} + (1 + \bar{\rho} + a\bar{\rho})(\widehat{\rho}^k)^{-1}) \|Y^{k+1} - Y^k\|.$$

Then, we will prove  $\|Y^{k+1} - Y^k\| \leq \text{dist}(Y^k, \mathcal{M}(\bar{x}))$ . Let  $P_k(Y^k) = \arg \max\{\widehat{h}(Y) - \frac{1}{2\bar{\rho}^k} \|Y - Y^k\|^2\}$ . Then, we have

$$\begin{aligned}
 (4.13) \quad & \|Y^{k+1} - Y^k\| \leq \|Y^{k+1} - P_k(Y^k)\| + \|P_k(Y^k) - Y^k\| \\
 & \leq (2\widehat{\rho}^k (\mathcal{L}_{\rho^k}(x^{k+1}, Y^k) - \inf_{\mathcal{X}} \mathcal{L}_{\rho^k}(\cdot, Y^k)))^{1/2} + \|P_k(Y^k) - Y^k\| \\
 & \leq \sqrt{\widehat{\rho}^k} \epsilon_k \|Y^{k+1} - Y^k\| + \|P_k(Y^k) - Y^k\|,
 \end{aligned}$$

where the second inequality follows from [38, equation (2.19)] and the last one follows from (4.3)(b) and the paragraph under (4.3). By using [13, Proposition 1(b)] and the proof of Lemma 3] and [38, Theorem 2.1], we have

$$(4.14) \quad \|P_k(Y^k) - Y^k\| \leq \text{dist}(Y^k, \mathcal{M}_D),$$

where  $\mathcal{M}_D$  denotes the solution set of (4.5). It follows from [38, Theorem 2.1] that  $\text{dist}(Y^k, \mathcal{M}_D) = \text{dist}(Y^k, \mathcal{M}(\bar{x}) \cap \mathcal{Y}) = \text{dist}(Y^k, \mathcal{M}(\bar{x}))$  since, for  $k$  sufficiently large,  $Y^k \in \text{int } \mathcal{Y}$ . Based on (4.12), (4.13), and (4.14), we obtain that, for sufficiently large  $k$ ,

$$R(x^{k+1}, Y^{k+1}) \leq (\epsilon'_k \sqrt{\widehat{\rho}^k} + (1 + \bar{\rho} + a\bar{\rho})(\widehat{\rho}^k)^{-1})(1 - \sqrt{\widehat{\rho}^k} \epsilon_k)^{-1} \text{dist}(Y^k, \mathcal{M}(\bar{x})).$$

This completes the proof.  $\square$

*Remark 4.8.* We compare our results with existing ALM convergence results for nonconvex nonpolyhedral problems. [21] justified the primal-dual linear convergence of the ALM under the SOSC and SRCQ for  $\mathcal{C}^2$ -cone reducible constrained problems, which include NLSDP and NLSOC. [44] proved the convergence rate of NLSDP under the strong SOSC together with nondegeneracy. [48] does provide the linear rate under the SOSC and semi-isolated calmness of the KKT pair without requiring the multiplier to be unique. But some other assumptions are also needed. In this paper, to obtain the Q-linear convergence for the multiplier and R-linear convergence for the primal variable, we assume the strong SOSC (Definition 2.2), which is much stronger than the aforementioned SOSC. However, we do not assume any restriction on the dual variable.

In [38, Example 5.3], the author also studied the ALM convergence for second-order cone programming when  $G(\bar{x}) \neq 0$ . Moreover, [19] gives the primal-dual linear

---

**Algorithm 4.2** the semismooth Newton-CG method for solving (4.1).

---

**Require:** Set the initial point  $x_0 = x^k$ , where  $x^k$  is obtained from the  $k$ th iteration of Algorithm 4.1. Let  $\mu \in (0, 1/2)$ ,  $\tau \in (0, 1]$ ,  $\tau_1, \tau_2, \bar{\nu} \in (0, 1)$ , and  $\theta \in (0, 1)$ . Set  $j := 0$ .

1: Given a maximum number of CG iterations  $t_j > 0$ , compute

$$\nu_j = \min\{\bar{\nu}, \|(\mathcal{L}_{\rho^k})'_x(x_j, Y^k)\|^{1+\tau}\}.$$

2: Choose  $W_j \in \partial \Pi_{\mathbb{S}^n}(G(x_j) + (\rho^k)^{-1}Y^k)$ . Let  $V_j = L''_{xx}(x_j, \Pi_{\mathbb{S}^n}(G(x_j) + (\rho^k)^{-1}Y^k) + \rho^k G'(x_j)^* W_j G'(x_j))$  and  $\varepsilon_j = \tau_1 \min\{\tau_2, \|\nabla_x \mathcal{L}_{\rho^k}(x_j, Y^k)\|\}$ . Apply the CG algorithm (CG( $\nu_j, t_j$ )) mentioned in [49, Algorithm 1] to find an approximate solution  $d_j \in \mathbb{X}$  to

$$(V_j + \varepsilon_j I)d_j = -(\mathcal{L}_{\rho^k})'_x(x_j, Y^k)$$

such that

$$\|(V_j + \varepsilon_j I)d_j + (\mathcal{L}_{\rho^k})'_x(x_j, Y^k)\| \leq \nu_j.$$

3: Set  $\zeta_j = \theta^{m_j}$ , where  $m_j$  is the first nonnegative number such that

$$\mathcal{L}_{\rho^k}(x_j + \theta^{m_j} d_j, Y^k) \leq \mathcal{L}_{\rho^k}(x_j, Y^k) + \mu \theta^{m_j} \langle (\mathcal{L}_{\rho^k})'_x(x_j, Y^k), d^j \rangle.$$

4: Set  $x_{j+1} = x_j + \zeta_j d_j$  and  $j = j + 1$ .

---

convergence for NLSOC when the multiplier is unique, while they only require the SOSC instead of the strong SOSC. By using Proposition 3.1 without any constraint qualification, the local convergence results of ALM for NLSOC can be obtained immediately.

#### 4.2. Solving the ALM subproblem: Semismooth Newton-CG method.

Although the convergence properties of the ALM have been established, it is equally important to study how to solve the subproblem (4.1). To guarantee the convergence rate of ALM, we need to employ the stopping criteria (4.3)(b) and (c). To meet the requirement of these stopping criteria, we consider using the semismooth Newton-CG method to solve the subproblem because it possesses the quadratic convergence rate under suitable conditions.

It follows from the proposed characterization of the strong variational sufficiency (Theorem 3.3) that, for NLSQP (2.4), if the strong SOSC (2.7) holds, then there exists a neighborhood  $\mathcal{B}_r(\bar{x}, \bar{Y})$  of  $(\bar{x}, \bar{Y})$  such that, for all  $(x, Y) \in \mathcal{B}_r(\bar{x}, \bar{Y})$ , every element in  $\pi_x \partial_B((\mathcal{L}_{\rho})'_x)(x, Y)$  is positive definite. Thus, the validity of the semismooth Newton-CG method to solve the subproblems in (4.1) is ensured. The algorithm to solve the  $(k+1)$ th subproblem is stated below (see Algorithm 4.2). The convergence analysis framework for Algorithm 4.2 is well known [29, Theorem 3.2] (see also [49, Theorems 3.4 and 3.5]). We omit the detailed proof here for simplicity.

**PROPOSITION 4.9.** *Suppose that the strong SOSC (2.7) holds at  $(\bar{x}, \bar{Y}) \in S_{KKT}(0, 0)$ . Then, Algorithm 4.2 is well defined, and any accumulation point  $\hat{x}$  of  $\{x_j\}$  generated*

by Algorithm 4.2 is the optimal solution to the subproblem (4.1). Furthermore, suppose that, at each step  $j$  when CG terminates, i.e.,

$$\|(V_j + \varepsilon_j I)d_j + (\mathcal{L}_{\rho^k})'_x(x_j, Y^k)\| \leq \nu_j,$$

then the whole sequence  $\{x_j\}$  converges to  $\hat{x}$ , and

$$\|x_{j+1} - \hat{x}\| = O(\|x_j - \hat{x}\|^{1+\tau}).$$

*Remark 4.10.* It is worth noting that, for convex NLSDP, the convergence result of the ALM can be established under weaker conditions such that the optimal  $\bar{x}$  can be nonunique (see [11, Theorem 20]). This implies that when the problem is reduced to a convex case, Theorem 4.6 is much weaker than [11, Theorem 20] as it requires the strong SOSC, which implies the local uniqueness of  $\bar{x}$ . However, the semismooth Newton algorithm may fail to solve the subproblem in the absence of the strong SOSC since it is equivalent to the positive definiteness of the generalized Hessian of Newton's equation of the subproblem (4.1) (Theorem 3.3). Thus, the strong SOSC seems to be not only sufficient to the local fast linear convergence rate of the ALM but also necessary for the invertibility of the generalized Hessian of the augmented Lagrangian function for NLSDP, which is crucial for the semismooth Newton CG method for solving the ALM subproblem (4.1).

**5. Numerical experiments.** Consider the optimization problem

$$(5.1) \quad \begin{aligned} \min_{X \in \mathbb{S}^n} \quad & \frac{1}{2} \langle X, Q \circ X \rangle \\ \text{s.t.} \quad & X \in \mathbb{S}_+^n \\ & B \circ X = 0, \end{aligned}$$

where “ $\circ$ ” denotes the Hadamard product of matrices; i.e., for any  $U$  and  $V \in \mathbb{R}^{p \times q}$ ,  $(U \circ V)_{ij} = U_{ij}V_{ij}$ ,

$$B = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} q & 1 & \cdots & 1 & 0 \\ 1 & q & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & q & 0 \\ 0 & 0 & \cdots & 0 & -1 \end{bmatrix}$$

with a given  $q \geq n - 1$ . The Lagrangian function of (5.1) is given by

$$L(X, Y, Z) = \frac{1}{2} \langle X, Q \circ X \rangle + \langle X, Y \rangle + \langle B \circ X, Z \rangle.$$

It can be checked directly that  $\bar{X} = 0$  is a local optimal solution with the multiplier  $\bar{Y} = \text{Diag}(0, \dots, 0, -1)$  and  $\bar{Z} = \text{Diag}(0, \dots, 0, 1)$ . In fact, the corresponding multiplier set of  $\bar{X}$  is given by

$$\mathcal{M}(\bar{X}) = \{(Y, Z) \in \mathbb{S}^n \times \mathbb{S}^n \mid Y + B \circ Z = 0, Y \in \mathbb{S}_-^n\}.$$

Like the problem (1.6) mentioned in the introduction, it follows from [50, Theorem 4.1] that the RCQ [30] does not hold at  $\bar{X}$  due to the unboundedness of  $\mathcal{M}(\bar{x})$ . It is clear that  $L''_{XX}(\bar{X}, \bar{Y}, \bar{Z}) = Q$ , which is positive definite over  $d \in \{d \in \mathbb{S}^n \mid B \circ d = 0, d \in \text{aff}\mathcal{C}_{\mathbb{S}_+^n}(\bar{X}, \bar{Y})\}$ , implies the validity of the strong SOSC (2.7). Also,

TABLE 1  
*Numerical results of the semismooth Newton-CG based ALM for (5.1).*

$n$	$q$	Iteration	KKT residual	CPU(s)
3	2	8	8.27e-06	0.22
100	200	11	8.98e-06	3.11
1000	1500	21	9.57e-06	1083.34

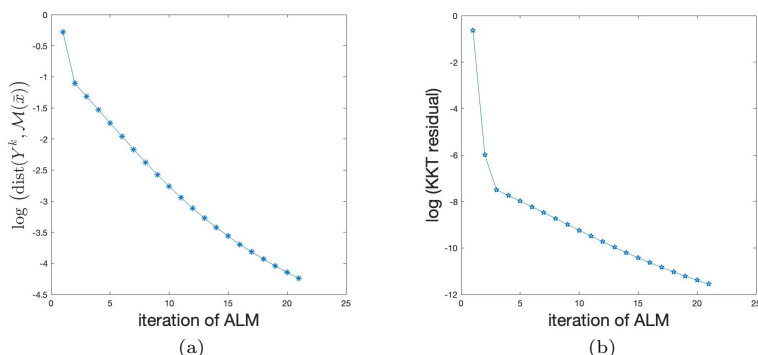


FIG. 2. *The semismooth Newton-CG based ALM for (5.1) with  $n = 1000$  and  $q = 1500$ .*

boundedly linear regularity is satisfied at  $(\bar{X}, \bar{Y}, \bar{Z})$  since  $(Y, Z) \in \mathcal{G}_1(\bar{X}) \cap \mathcal{G}_2(\bar{X})$  with  $Y = \text{Diag}(-1, \dots, -1)$ ,  $Z = \text{Diag}(0, \dots, 0, 1)$ .

Next, we shall apply Algorithm 4.1 to solve (5.1) with different dimensions. The subproblem (4.1) is solved by Algorithm 4.2, where the exactness in (4.3) is chosen as  $\epsilon'_k = 0.01 \times (1/1.05)^{k-1}$  and the stopping criterion (4.3)(b) is employed. The algorithm is stopped when the KKT residual  $R(x^k, Y^k)$  defined in (4.9) is less than  $1e-5$ . The codes are implemented in MATLAB (R2018a), and the numerical experiments are run under a 64-bit MacOS on an Intel Cores i5 2.3 GHz CPU with 8 GB memory. The following table (Table 1) shows the numerical results of different dimensions of (5.1). Noting that the distance from  $(Y^k, Z^k)$  to  $\mathcal{M}(\bar{X})$  is difficult to compute, we use the following alternative  $\text{dist}((Y^k, Z^k), \mathcal{M}(\bar{X})) = O(\|Y^k + B \circ Z^k\| + \|\Pi_{\mathcal{S}_+^n}(Y^k)\|)$  because bounded linear regularity is satisfied. The detail iterative performance of the ALM for solving (5.1) with  $n = 1000$  and  $q = 1500$  is also reported in Figure 2. It can be seen from Figure 2(a) that, although the problem is nonconvex and the multiplier set is not a singleton, the ALM also possesses a linear rate of convergence of  $\text{dist}((Y, Z), \mathcal{M}(\bar{X}))$  for the validity of the strong variational sufficiency. Meanwhile, the right one shows that the KKT residual also converges to 0 as the algorithm proceeds. Moreover, by strong variational sufficiency, we know from Theorem 3.3 that the Hessian matrix  $V_j$  in Algorithm 4.2, step 2 is always positive definite when  $k$  is sufficiently large. Indeed, the minimum eigenvalue of the Hessian matrix  $V_j$  is positive (e.g.,  $\lambda_{\min}(V_j) \approx 1.01$  for the case  $n = 3$  and  $q = 2$ ;  $\lambda_{\min}(V_j) \approx 1.26$  for the case  $n = 100$  and  $q = 200$ ). It is also worth noting that the distance from  $(Y^k, Z^k)$  to  $(\bar{Y}, \bar{Z})$  does not converge to 0, which meets our theory as the limit point  $(\hat{Y}, \hat{Z})$  in Theorem 4.6 may be different from the reference point  $(\bar{Y}, \bar{Z})$ .

It is well known that the condition number of  $V_j$  in Algorithm 4.2, step 2 is proportional to  $\rho^k$ . In fact, the estimation of the condition number of  $V_j$  can be obtained in a similar manner as [44, Lemma 10] because the smallest eigenvalue lies

in a constant interval, while the largest one lies in an interval that is propositional to  $\rho^k$ . Thus, in practice, we usually set an upper bound for  $\rho^k$ , although  $\rho^k$  can be infinity in theory.

It can be seen from Theorem 4.6 that the convergence analysis of Algorithm 4.1 requires a good starting point of the multiplier, while no restriction to the initial primal variable  $x^0$  is needed. To verify the convergence theory (Theorem 4.6), in the numerical experiments of (5.1) presented in Table 1, we choose  $X^0$  randomly.  $(Y^0, Z^0)$  is chosen to be  $(\bar{Y}, \bar{Z}) + \eta(P_1, P_2)$ , where  $P_1, P_2$  are symmetry matrices that are uniform randomly generated. Typically,  $\eta$  is chosen to be small, e.g., 0.1.

Next, we will discuss how to generate the initial point more practically. Condition 4.5 indicates that  $(Y^0, Z^0)$  should be sufficiently close to  $(\bar{Y}, \bar{Z})$ , at which the strong SOSC is satisfied. How to find  $(Y^0, Z^0)$  satisfying condition 4.5 for the nonconvex optimization problem is very challenging. A natural idea is to apply first-order methods as a warm start to find a satisfactory initial point. Moreover, we know from Theorem 3.3 that a good starting point may be the one at which the positive definiteness of the generalized Hessian is satisfied. In practice, when the generalized Hessian is not positive definite, we may apply accelerate proximal gradient (APG) instead of the semismooth Newton to solve the ALM subproblem.

**6. Conclusion.** In this paper, we derive the equivalence between the strong variational sufficiency and the strong SOSC for NLSDP and NLSOC without requiring the uniqueness of the multiplier or any other constraint qualification. By using the equivalence result, the local convergence property of the ALM for NLSDP can be established under the merely strong SOSC instead of it plus any constraint qualification. As a direct application, we are able to show that the positive definiteness of the generalized Hessian of the augmented Lagrangian function, which is critical in the use of the semismooth Newton method for NLSDP, is satisfied under the strong SOSC. However, there are still many issues that deserve to be explored further. For example, we still do not have a satisfactory characterization for variational sufficiency. Moreover, for the implementation of the ALM for solving nonconvex cases, it is still very challenging for finding a good starting point, e.g., the one close enough to a KKT pair of a local optimal solution that satisfies the strong SOSC condition.

**Acknowledgments.** The authors would like to thank the two anonymous referees and the associate editor for their many valuable comments and constructive suggestions, which have substantially helped improve the quality and the presentation of this paper.

#### REFERENCES

- [1] F. J. ARAGON AND M. F. GEOFFROY, *Characterization of metric regularity of subdifferential mappings*, *J. Convex Anal.*, 15 (2008), pp. 365–380.
- [2] H. H. BAUSCHKE AND J. M. BORWEIN, *On projection algorithms for solving convex feasibility problems*, *SIAM Rev.*, 38 (1996), pp. 367–426.
- [3] H. H. BAUSCHKE, J. M. BORWEIN, AND W. LI, *Strong conical hull intersection property, bounded linear regularity, Jameson’s property (G), and error bounds in convex optimization*, *Math. Program.*, 86 (1999), pp. 135–160.
- [4] D. BERTSEKAS, *Constrained Optimization and Lagrange Multipliers Methods*, Academic Press, New York, 1982.
- [5] J. F. BONNANS AND H. RAMÍREZ CABRERA, *Perturbation analysis of second-order cone programming problems*, *Math. Program.*, 104 (2005), pp. 205–227.
- [6] J. F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

- [7] B. S. MORDUKHOVICH, T. T. A. NGHIA, AND R. T. ROCKAFELLAR, *Full stability in finite-dimensional optimization*, Math. Oper. Res., 40 (2015), pp. 226–252.
- [8] Z. X. CHAN AND D. F. SUN, *Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming*, SIAM J. Optim., 19 (2008), pp. 370–396.
- [9] A. R. CONN, N. I. M. GOULD, AND P. TOINT, *A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds*, SIAM J. Numer. Anal., 28 (1991), pp. 545–572.
- [10] L. CONTESSE-BECKER, *Extended convergence results for the method of multipliers for non-strictly binding inequality constraints*, J. Optim. Theory Appl., 79 (1993), pp. 273–310.
- [11] Y. CUI, C. DING, AND X. Y. ZHAO, *Quadratic growth conditions for convex matrix optimization problems associated with spectral functions*, SIAM J. Optim., 27 (2017), pp. 2332–2355.
- [12] Y. CUI, D. F. SUN, AND K. C. TOH, *On the Asymptotic Superlinear Convergence of the Augmented Lagrangian Method for Semidefinite Programming with Multiple Solutions*, preprint, arXiv:1610.00875, 2016.
- [13] Y. CUI, D. F. SUN, AND K. C. TOH, *On the R-superlinear convergence of the KKT residuals generated by the augmented Lagrangian method for convex composite conic programming*, Math. Program., 178 (2019), pp. 381–415.
- [14] C. DING, D. F. SUN, AND J. J. YE, *First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints*, Math. Program., 147 (2014), pp. 539–579.
- [15] C. DING, D. F. SUN, AND L. W. ZHANG, *Characterization of the robust isolated calmness for a class of conic programming problems*, SIAM J. Optim., 27 (2017), pp. 67–90.
- [16] A. L. DONTCHEV AND R. T. ROCKAFELLAR, *Implicit Functions and Solution Mappings*, Springer, New York, 2009.
- [17] D. FERNANDEZ AND M. V. SOLODOV, *Local convergence of exact and inexact augmented Lagrangian methods under the second-order sufficient optimality condition*, SIAM J. Optim., 22 (2012), pp. 384–407.
- [18] N. T. V. HANG AND M. E. SARABI, *Local convergence analysis of augmented Lagrangian methods for piecewise linear-quadratic composite optimization problems*, SIAM J. Optim., 31 (2021), pp. 2665–2694.
- [19] N. T. V. HANG, B. MORDUKHOVICH, AND E. SARABI, *Augmented Lagrangian method for second-order cone programs under second-order sufficiency*, J. Global Optim., (2021), pp. 1–31.
- [20] K. ITO AND K. KUNISCH, *The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces*, Math. Program., 46 (1990), pp. 341–360.
- [21] C. KANZOW AND D. STECK, *Improved local convergence results for augmented Lagrangian methods in  $C^2$ -cone reducible constrained optimization*, Math. Program., 177 (2019), pp. 425–438.
- [22] P. D. KHANH, B. S. MORDUKHOVICH, AND V. T. PHAT, *Variational Convexity of Functions and Variational Sufficiency in Optimization*, preprint, arXiv:2208.14399, 2022.
- [23] X. D. LI, D. F. SUN, AND K. C. TOH, *QSDPNAL: A two-phase augmented Lagrangian method for convex quadratic semidefinite programming*, Math. Program. Comput., 10 (2018), pp. 703–743.
- [24] F. J. LUQUE, *Asymptotic convergence analysis of the proximal point algorithm*, SIAM J. Optim., 22 (1984), pp. 277–293.
- [25] A. MOHAMMADI, B. S. MORDUKHOVICH, AND M. E. SARABI, *Parabolic regularity in geometric variational analysis*, Trans. Amer. Math. Soc., 374 (2021), pp. 1711–1763.
- [26] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Grundlehren Math. Wiss., 330, Springer, Cham, 2006.
- [27] B. S. MORDUKHOVICH AND R. T. ROCKAFELLAR, *Second-order subdifferential calculus with applications to tilt stability in optimization*, SIAM J. Optim., 22 (2012), pp. 953–986.
- [28] T. PENNANEN, *Local convergence of the proximal point algorithm and multiplier methods without monotonicity*, Math. Oper. Res., 27 (2002), pp. 170–191.
- [29] L. QI AND J. SUN, *A nonsmooth version of Newton’s method*, Math. Program., 58 (1993), pp. 353–367.
- [30] S. M. ROBINSON, *First order conditions for general nonlinear optimization*, SIAM J. Appl. Math., 30 (1976), pp. 597–607.
- [31] S. M. ROBINSON, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
- [32] R. T. ROCKAFELLAR, *Convex Analysis*, Vol. 36, Princeton University Press, Princeton, NJ, 1970.
- [33] R. T. ROCKAFELLAR, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Oper. Res., 1 (1976), pp. 97–116.

- [34] R. T. ROCKAFELLAR, *Progressive decoupling of linkages in monotone variational inequalities and convex optimization*, in Proceedings of the 10th International Conference on Nonlinear Analysis and Convex Analysis, Chitose, Japan, Yokohama Publishers, 2019, pp. 271–291.
- [35] R. T. ROCKAFELLAR, *Variational convexity and the local monotonicity of subgradient mappings*, Vietnam J. Math., 47 (2019), pp. 547–561.
- [36] R. T. ROCKAFELLAR, *Advances in convergence and scope of the proximal point algorithm*, J. Nonlinear Convex Anal., 22 (2021), pp. 2347–2375.
- [37] R. T. ROCKAFELLAR, *Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality*, Math. Program., 198 (2022), pp. 159–194.
- [38] R. T. ROCKAFELLAR, *Convergence of augmented Lagrangian methods in extensions beyond nonlinear programming*, Math. Program., 199 (2022), pp. 375–420.
- [39] R. T. ROCKAFELLAR AND R. J-B. WETS, *Variational Analysis*, Springer, New York, 1998.
- [40] A. SHAPIRO AND J. SUN, *Some properties of the augmented Lagrangian in cone constrained optimization*, Math. Oper. Res., 29 (2004), pp. 479–491.
- [41] D. F. SUN, *The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*, Math. Oper. Res., 31 (2006), pp. 761–776.
- [42] D. F. SUN AND J. SUN, *Semismooth matrix valued functions*, Math. Oper. Res., 27 (2002), pp. 150–169.
- [43] D. F. SUN AND J. SUN, *Strong semismoothness of eigenvalues of symmetric matrices and its applications in inverse eigenvalue problems*, SIAM J. Numer. Anal., 40 (2003), pp. 2352–2367.
- [44] D. F. SUN, J. SUN, AND L. W. ZHANG, *The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming*, Math. Program., 114 (2008), pp. 349–391.
- [45] D. F. SUN, K. C. TOH, Y. C. YUAN, AND X. Y. ZHAO, *SDPNAL+: A Matlab software for semidefinite programming with bound constraints (version 1.0)*, Optim. Methods Softw., 35 (2020), pp. 87–115.
- [46] J.-B. H. URRUTY, J.-J. STRODIOT, AND V. H. NGUYEN, *Generalized Hessian matrix and second-order optimality conditions for problems with  $C^{1,1}$  data*, Appl. Math. Optim., 11 (1984), pp. 43–56.
- [47] L. Q. YANG, D. F. SUN, AND K. C. TOH, *SDPNAL+: A majorized semismooth Newton-CG augmented Lagrangian method for semidefinite programming with nonnegative constraints*, Math. Program. Comput., 7 (2015), pp. 331–366.
- [48] S. W. WANG AND C. DING, *Local Convergence Analysis of Augmented Lagrangian Method for Nonlinear Semidefinite Programming*, preprint, arXiv:2110.10594, 2021.
- [49] X. Y. ZHAO, D. F. SUN, AND K. C. TOH, *A Newton-CG augmented Lagrangian method for semidefinite programming*, SIAM J. Optim., 20 (2010), pp. 1737–1765.
- [50] J. ZOWE AND S. KURCZYUSZ, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.