

A quadratically convergent semismooth Newton method for nonlinear semidefinite programming without generalized Jacobian regularity

Fuxiaoyue Feng · Chao Ding · Xudong Li

This version: November 24, 2024

Abstract We introduce a quadratically convergent semismooth Newton method for nonlinear semidefinite programming that eliminates the need for the generalized Jacobian regularity, a common yet stringent requirement in existing approaches. Our strategy involves identifying a single nonsingular element within the Bouligand generalized Jacobian, thus avoiding the standard requirement for nonsingularity across the entire generalized Jacobian set, which is often too restrictive for practical applications. The theoretical framework is supported by introducing the weak second order condition (W-SOC) and the weak strict Robinson constraint qualification (W-SRCQ). These conditions not only guarantee the existence of a nonsingular element in the generalized Jacobian but also forge a primal-dual connection in linearly constrained convex quadratic programming. The theoretical advancements further lay the foundation for the algorithmic design of a novel semismooth Newton method, which integrates a correction step to address degenerate issues. Particularly, this correction step ensures the local convergence as well as a superlinear/quadratic convergence rate of the proposed method. Preliminary numerical experiments corroborate our theoretical findings and underscore the practical effectiveness of our method.

Key words: Semismooth Newton method, singularity, generalized Jacobian, nonlinear semidefinite programming, second order conditions, constraint qualifications.

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

F.X.Y. Feng
Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, P.R. China, School of Mathematical Sciences, University of Chinese Academy of Science, Beijing, P.R. China.

E-mail: fengfuxiaoyue@amss.ac.cn

C. Ding
Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, P.R. China.

E-mail: dingchao@amss.ac.cn

X.D. Li
School of Data Science, Fudan University, Shanghai, P.R. China.
E-mail: lixudong@fudan.edu.cn

1 Introduction

Let \mathbb{X} and \mathbb{Y} be two Euclidean spaces each equipped with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Consider the following optimization problem (OP):

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & G(x) \in K, \end{aligned} \tag{1}$$

where $K \subseteq \mathbb{Y}$ is a nonempty closed convex set and $f : \mathbb{X} \rightarrow \mathbb{R}$, $G : \mathbb{X} \rightarrow \mathbb{Y}$ are twice continuously differentiable.

Let $L : \mathbb{X} \times \mathbb{Y} \rightarrow \mathbb{R}$ be the Lagrangian function of problem (1) defined by

$$L(x; y) := f(x) + \langle y, G(x) \rangle, \quad (x, y) \in \mathbb{X} \times \mathbb{Y}.$$

For any $y \in \mathbb{Y}$, denote the derivative of $L(\cdot; y)$ at $x \in \mathbb{X}$ by $L'_x(x; y)$ and denote the adjoint of $L'_x(x; y)$ by $\nabla_x L(x; y)$. The Karush-Kuhn-Tucker (KKT) optimality conditions for problem (1) take the following form:

$$\begin{cases} \nabla_x L(x, y) = 0, \\ y \in \mathcal{N}_K(G(x)), \end{cases} \tag{2}$$

where $\mathcal{N}_K(G(x))$ is the normal cone to K at $G(x)$ in the context of convex analysis [32]. We say that $(x, y) \in \mathbb{Z} := \mathbb{X} \times \mathbb{Y}$ is a KKT point of (1) if it solves (2). The set of Lagrange multipliers associated with a feasible point x is defined by $M(x) = \{y \mid (x, y) \text{ is a KKT point}\}$. For a given feasible point \bar{x} , if $M(\bar{x}) \neq \emptyset$, then we call it a stationary point of (1). It is well-known that (\bar{x}, \bar{y}) is a KKT point of (1) if and only if it is a solution of the nonsmooth system

$$F(x, y) := \begin{bmatrix} \nabla_x L(x, y) \\ -G(x) + \Pi_K(G(x) + y) \end{bmatrix} = 0, \tag{3}$$

where $\Pi_K : \mathbb{Y} \rightarrow \mathbb{Y}$ is the metric projection operator onto K , i.e., $\Pi_K(z) = \arg \min_{y' \in K} \frac{1}{2} \|z - y'\|^2$ for $z \in \mathbb{Y}$.

When solving $F(x, y) = 0$, one may encounter challenges related to the nonsmoothness of F , primarily arising from the nonsmooth nature of the metric projection operator $\Pi_K(\cdot)$, especially when K is not a subspace. Fortunately, for various problems, including nonlinear programming (NLP), nonlinear second-order cone programming, and nonlinear semidefinite programming (NLSDP), the corresponding nonsmooth map F exhibits the so-called (strong) semismoothness (see Definition 1). This property is conducive to the use of nonsmooth Newton methods. Qi and Sun [29] introduced a semismooth Newton method, which takes the following semismooth Newton step in each iteration:

$$(x^{k+1}; y^{k+1}) = (x^k; y^k) - (U^k)^{-1} F(x^k, y^k)$$

with U^k being selected from the Bouligand generalized Jacobian $\partial_B F(x^k, y^k)$ (see (4) for definition) or the Clarke generalized Jacobian $\partial_C F(x^k, y^k)$ (see (5) for definition). The semismooth Newton method, renowned for its high efficiency, has found widespread applications across various domains. It has been effectively utilized for tackling variational inequalities and constrained optimization problems in function spaces [38], infinite-dimensional linear complementarity problems [15] and computing the nearest correlation matrix [27]. Notably, its efficiency has been particularly pronounced in solving augmented Lagrangian subproblems, whose versatility is underscored by its implementation in several influential software packages, including SDPNAL [41],

SDPNAL+ [40], QSDPNAL [21], and SSNAL [20]. For an in-depth exploration, one may refer to [10, 15, 11].

It is well-known (see e.g., [29, 18, 28]) that the (quadratic) superlinear convergence of semismooth Newton methods hinges on the (strong) g -semismoothness (see Definition 1) of F and the uniform boundedness of $\{(U^k)^{-1}\}$. To guarantee such uniform boundedness, typically, one has to assume that the KKT mapping F satisfies a certain generalized Jacobian regularity condition at the solution point (\bar{x}, \bar{y}) , for example, all elements in some generalized Jacobian sets of F at (\bar{x}, \bar{y}) are nonsingular. Particularly, one may assume that every element in $\partial_B F(\bar{x}, \bar{y})$ or even $\partial_C F(\bar{x}, \bar{y})$ is nonsingular, known as the BD-regularity or CD-regularity, respectively. However, this condition is frequently demanding and may not always be fulfilled in practical applications. Thus, the following important question arises:

Can we design a locally superlinear or quadratic convergent semismooth Newton method for solving (3) without assuming the generalized Jacobian regularity of F ?

In this paper, we provide an affirmative answer to this question in the context of nonlinear semidefinite programming (NLSDP). Specifically, for the NLSDP, we design a semismooth Newton method with a novel correction step to avoid the requirement of nonsingularity across the entire generalized Jacobian of F , and prove its fast local convergence rates, thus advancing the traditional paradigms. As illustrated in Example 5, despite the presence of singular elements in the Bouligand generalized Jacobian $\partial_B F(\bar{x}, \bar{y})$ at the KKT point (\bar{x}, \bar{y}) and the lack of the calmness property for F , our method demonstrates local quadratic convergence, as shown in Figure 3.

For the optimization problem (1), the characterization of the generalized Jacobian regularity for the nonsmooth function F is closely intertwined with perturbation theory, a fundamental topic in optimization. Indeed, the perturbation theory is crucial for both theoretical insights and practical algorithm design. Over the past four decades, significant advancements in this field have been well-documented in the literature [1, 17, 33, 11, 24], especially concerning (1) with K being a polyhedral set (i.e., K is the intersection of a finite number of half-spaces). However, less exploration has been done regarding non-polyhedral cases, such as K being the set of positive semidefinite matrices.

In the realm of nonlinear semidefinite programming, Sun [34] established that under the Robinson constraint qualification (RCQ) (Definition 2), the nonsingularity of $\partial_C F(\bar{x}, \bar{y})$ (i.e., the CD-regularity of F) at a local optimal solution \bar{x} and an associated multiplier \bar{y} is equivalent to the strong second-order sufficient condition (S-SOSC) (Definition 6) and the constraint nondegeneracy (Definition 4), and is also equivalent with the strong regularity introduced by Robinson [31] of the solution to the KKT system (2). For linear SDP problems, Chan and Sun [2] provided deeper insights into the strong regularity, demonstrating that under the RCQ, primal and dual constraint nondegeneracies, the nonsingularity of $\partial_B F(\bar{x}, \bar{y})$ (i.e., the BD-regularity of F), CD-regularity and the strong regularity at (\bar{x}, \bar{y}) are all equivalent. This comprehensive understanding underscores that the conditions for the nonsingularity of all elements in the generalized Jacobians of F at (\bar{x}, \bar{y}) may be stringent and not consistently be satisfied in practice. For instance, the S-SOSC condition presupposes the isolatedness of a local optimal solution, a premise that is often violated in statistical optimization applications such as LASSO [37]. Consequently, it is crucial to develop a new framework of semismooth Newton methods that can bypass the generalized Jacobian regularity and still achieve a local superlinear or quadratic convergence rate.

In this paper, we propose to relax the stringent generalized Jacobian regularity conditions by identifying a single nonsingular element within the Bouligand generalized Jacobian. However, characterizing the nonsingularity of even a single element within the generalized Jacobians of

F at a KKT point (\bar{x}, \bar{y}) , especially in non-polyhedral cases, remains an unresolved challenge. When K is polyhedral, Proposition 6 in Izmailov and Solodov [16] indicated that the weak strict Robinson constraint qualification (W-SRCQ) and the weak second-order condition (W-SOC) (see Definitions 7 and 8, respectively) guarantee the existence of a nonsingular element in the parametric generalized Jacobian. Here the parametric generalized Jacobian, defined as differentials of all possible parametric systems, contains $\partial_B F(\bar{x}, \bar{y})$. Indeed, a straightforward deduction from their findings indicates that the W-SRCQ and W-SOC also guarantee the presence of a nonsingular element in $\partial_B F(\bar{x}, \bar{y})$. However, these conditions are inadequate when applied to NLSDP. Specifically, the W-SRCQ and W-SOC are insufficient to ensure the existence of a nonsingular element in $\partial_B F(\bar{x}, \bar{y})$, as demonstrated by our counterexample (Example 2). Meanwhile, in NLSDP, the presence of a nonsingular generalized Jacobian element, as in Example 5, does not necessarily imply calmness, an important property described in [33, Definition 9(30)]. The lack of calmness may lead to slow progress of algorithms for solving the KKT system (3), as the associated solution mapping exhibits heightened sensitivity to small perturbations.

In this paper, we show that the existence of a nonsingular element in $\partial_B F(\bar{x}, \bar{y})$ can be ensured under the W-SOC with constraint nondegeneracy or the S-SOSC with the W-SRCQ at a KKT point (\bar{x}, \bar{y}) . Additionally, our results indicate that the W-SRCQ and W-SOC are sufficient for the existence of a nonsingular element in $\partial_C F(\bar{x}, \bar{y})$, and are also necessary for convex cases. More interestingly, we uncover a profound primal-dual connection between the W-SOC and W-SRCQ for linearly constrained convex quadratic semidefinite programming. These connections mirror those found between the strict Robinson constraint qualification (SRCQ) (Definition 3) and the second-order sufficient condition (SOSC) (Definition 5) for linearly constrained quadratic semidefinite programming [14], as well as those between the constraint nondegeneracy and the S-SOSC in the setting of linear semidefinite programming [2]. We summarize our results and connections in Figure 1. Based on the theoretical findings, we are able to design an inexact semismooth Newton method with a correction step. Each iteration commences with a semismooth Newton step, followed by a correction step designed to ensure the nonsingularity. Under the W-SOC and constraint nondegeneracy (or the S-SOSC and W-SRCQ), we show the local convergence and superlinear convergence of the designed algorithm. Furthermore, if the function is twice Lipschitz continuously differentiable, the convergence rate can be further improved to quadratic.

The remaining parts of this paper are organized as follows. Section 2 provides some basic definitions and preliminary results concerning the variational properties of the metric projection projector over the positive semidefinite cone. Section 3 delves into the sufficient conditions and necessary conditions for the existence of one nonsingular element in the generalized Jacobians of F for NLSDP. The primal-dual connections of the W-SOC and the W-SRCQ are explored in Section 4 for convex quadratic semidefinite programming. Then, we present a semismooth Newton method with a correction step in Section 5, and some preliminary numerical results in Section 6. Finally, we conclude this paper and highlight potential directions for future research in Section 7.

2 Preliminaries

We first give the following notation that will be used throughout the paper. Let \mathbb{X} and \mathbb{Y} be two real Euclidean spaces. Given a set $S \subseteq \mathbb{X}$ and a point $x \in \mathbb{X}$, the distance from x to S is denoted by

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

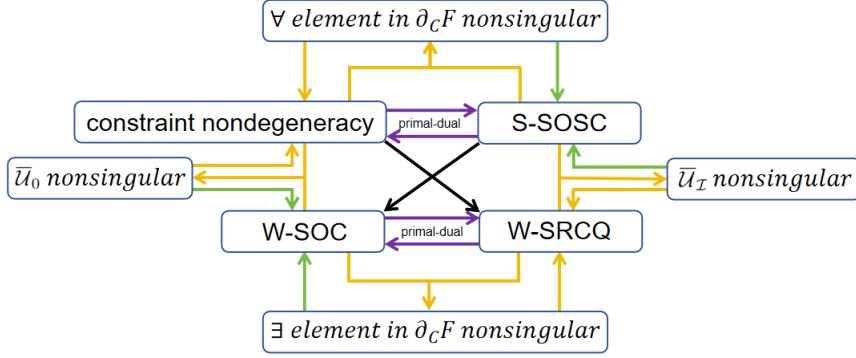


Fig. 1: Connections of second order conditions, constraint qualifications and nonsingularity of generalized Jacobians. Here the black arrow holds for (1); the orange arrow holds for NLSDP; the green arrow holds for CSDP; and the violet arrow holds for QSDP. Moreover, “primal-dual” means that the starting point of the arrow for the primal problem implies the arrival point of the arrow for the dual problem.

For a linear operator $\mathcal{A} : \mathbb{X} \rightarrow \mathbb{Y}$, \mathcal{A}^* denotes the adjoint of the linear operator \mathcal{A} . We denote by A^T the transpose of a given matrix A . For a mapping $\Phi : \mathbb{X} \rightarrow \mathbb{Y}$ and $x \in \mathbb{X}$, $\Phi'(x)$ stands for classical derivative or the Jacobian of Φ at x and $\nabla\Phi(x)$ the adjoint of the Jacobian. We use \mathbb{S}^n to denote the linear space of all $n \times n$ real symmetric matrices and use \mathbb{S}_+^n to denote the closed convex cone of all $n \times n$ positive semidefinite matrices in \mathbb{S}^n . For a function $g : \mathbb{X} \rightarrow \mathbb{R}$, we denote $g^+(x) := \max\{0, g(x)\}$ and if it is vector-valued then the maximum is taken componentwise. Some other notations are listed below:

- Let \mathcal{O}^n be the set of all $n \times n$ orthogonal matrices.
- For any $Z \in \mathbb{R}^{m \times n}$, we denote by Z_{ij} the (i, j) -th entry of Z .
- For any $Z \in \mathbb{R}^{m \times n}$ and a given index set $\mathcal{J} \subseteq \{1, \dots, n\}$, we use $Z_{\mathcal{J}}$ to denote the sub-matrix of Z obtained by removing all the columns of Z not in \mathcal{J} . In particular, we use Z_j to represent the j -th column of Z , $j = 1, \dots, n$.
- Let $\mathcal{I} \subseteq \{1, \dots, m\}$ and $\mathcal{J} \subseteq \{1, \dots, n\}$ be two index sets. For any $Z \in \mathbb{R}^{m \times n}$, we use $Z_{\mathcal{I}\mathcal{J}}$ to denote the $|\mathcal{I}| \times |\mathcal{J}|$ sub-matrix of Z obtained by removing all the rows of Z not in \mathcal{I} and all the columns of Z not in \mathcal{J} .
- We use “ \circ ” to denote the Hadamard product between matrices, i.e., for any two matrices A and B in $\mathbb{R}^{m \times n}$ the (i, j) -th entry of $Z := A \circ B \in \mathbb{R}^{m \times n}$ is $Z_{ij} = A_{ij}B_{ij}$.
- Let $\text{Diag}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{S}^n$ be the linear map such that for any $x \in \mathbb{R}^n$, $\text{Diag}(x)$ denotes the diagonal matrix whose i -th diagonal entry is x_i , $i = 1, \dots, n$.
- Let $\text{lin}(C)$ be the linearity space of a closed convex set $C \subseteq \mathbb{X}$, i.e., the largest subspace in C . Meanwhile, let $\text{aff}(C)$ be the affine hull of C , i.e., the smallest subspace containing C .
- For a map $\Phi : \mathbb{X} \rightarrow \mathbb{Y}$, define $\text{Im}(\Phi) := \{\Phi(x) \mid x \in \mathbb{X}\}$ and $\text{Null}(\Phi) := \{x \in \mathbb{X} \mid \Phi(x) = 0\}$.
- The support function of a set $S \subseteq \mathbb{X}$ is defined by $\sigma(x, S) = \sup_{s \in S} \langle x, s \rangle$ for $x \in \mathbb{X}$.
- For $r > 0$ and $\bar{X} \in \mathbb{X}$, the closed ball centered at \bar{X} with radius r is denoted by $\mathbb{B}(\bar{X}, r) = \{X \in \mathbb{X} \mid \|X - \bar{X}\| \leq r\}$.
- For $f, g : \mathbb{X} \rightarrow \mathbb{R}$, $f(x) = \Theta(g(x))$ for $x \rightarrow \bar{x}$ means that $f(x) = O(g(x))$ and $g(x) = O(f(x))$.

For a given map $\Phi : \mathbb{X} \rightarrow \mathbb{Y}$, the directional differential of Φ at a point $x \in \mathbb{X}$ along a direction $d \in \mathbb{X}$ is defined by

$$\Phi'(x; d) = \lim_{t \downarrow 0} \frac{F(x + td) - F(x)}{t} \quad \text{if this limit exists.}$$

Φ is directionally differentiable at $x \in \mathbb{X}$ if $\Phi'(x; d)$ exists for all $d \in \mathbb{X}$. Suppose that $\Phi : N \subseteq \mathbb{X} \rightarrow \mathbb{Y}$ is a locally Lipschitz continuous function on the open set N . Then, according to Rademacher's theorem [30], Φ is almost everywhere F-differentiable in N . Let D_Φ be the set of points in N where Φ is differentiable. Then the Bouligand generalized Jacobian of Φ at $x \in N$ is denoted by [28]:

$$\partial_B \Phi(x) := \left\{ \lim_{D_\Phi \ni x^k \rightarrow x} \Phi'(x^k) \right\}, \quad (4)$$

and the Clarke generalized Jacobian of Φ at $x \in N$ [5] takes the form:

$$\partial_C \Phi(x) = \text{conv} \{ \partial_B \Phi(x) \}, \quad (5)$$

where ‘‘conv’’ stands for the convex hull in the usual sense of convex analysis [32]. The following g-semismoothness of a locally Lipschitz continuous function is an extension of the semismoothness. The concept of semismoothness was initially introduced by Mifflin [23] for real-valued functions and later extended to vector-valued functions by Qi and Sun [29]. This concept was further extended to g-semismoothness by Gowda [13], which applies to functions without requiring directional differentiability.

Definition 1 Let $\Phi : N \subseteq \mathbb{X} \rightarrow \mathbb{Y}$ be a locally Lipschitz continuous function on the open set N . The function Φ is said to be g-semismooth at a point $x \in N$ if for any $y \rightarrow x$ and $V \in \partial_C \Phi(y)$,

$$\Phi(y) - \Phi(x) - V(y - x) = o(\|y - x\|).$$

The function Φ is said to be strongly g-semismooth at x if for any $y \rightarrow x$ and $V \in \partial_C \Phi(y)$,

$$\Phi(y) - \Phi(x) - V(y - x) = O(\|y - x\|^2).$$

Furthermore, the function Φ is said to be (strongly) semismooth at $x \in N$ if (i) Φ is directional differentiable at x ; and (ii) Φ is (strongly) g-semismooth.

2.1 Background in variational analysis

In this subsection, we introduce some variational properties related to problem (1), which will be useful in our subsequent discussions.

For the closed convex set $K \subseteq \mathbb{Y}$, the tangent cone to K at $y \in \mathbb{Y}$ is defined by

$$\mathcal{T}_K(y) = \{d \in \mathbb{Y} \mid \exists t^k \downarrow 0, \text{dist}(y + t^k d, K) = o(t^k)\}. \quad (6)$$

Let \bar{x} be a given feasible solution to (1). It is well-known [1, Theorem 3.9 and Proposition 3.17] that for the feasible solution \bar{x} , the following Robinson constraint qualification is equivalent to the nonemptiness and boundedness of the corresponding Lagrange multiplier set $M(\bar{x})$.

Definition 2 The Robinson constraint qualification (RCQ) is said to hold at a feasible point \bar{x} if

$$G'(\bar{x})\mathbb{X} + \mathcal{T}_K(G(\bar{x})) = \mathbb{Y}.$$

For a stationary solution \bar{x} of (1), the strict Robinson constraint qualification condition is defined as follows.

Definition 3 Let $\bar{x} \in \mathbb{X}$ be a stationary point of (1) with a Lagrange multiplier $\bar{y} \in M(\bar{x})$. The strict Robinson constraint qualification (SRCQ) is said to hold at \bar{x} with respect to \bar{y} if

$$G'(\bar{x})\mathbb{X} + \mathcal{T}_K(G(\bar{x})) \cap \bar{y}^\perp = \mathbb{Y},$$

where $y^\perp := \{s \in \mathbb{Y} \mid \langle s, y \rangle = 0\}$ for any vector $y \in \mathbb{Y}$.

It follows from [1, Proposition 4.50] that the set of Lagrange multipliers $M(\bar{x})$ is a singleton if the SRCQ hold.

For the feasible point x , the critical cone $\mathcal{C}(x)$ of (1) at x is defined by

$$\mathcal{C}(x) = \{\Delta x \in \mathbb{X} \mid G'(x)\Delta x \in \mathcal{T}_K(G(x)), f'(x)\Delta x = 0\}. \quad (7)$$

If \bar{x} is a stationary point with $\bar{y} \in M(\bar{x})$, then the critical cone $\mathcal{C}(\bar{x})$ takes the following form:

$$\mathcal{C}(\bar{x}) = \{\Delta x \in \mathbb{X} \mid G'(\bar{x})\Delta x \in \mathcal{T}_K(G(\bar{x})) \cap \bar{y}^\perp\}. \quad (8)$$

The following definition on the constraint nondegeneracy of (1), introduced by Robinson [31], is stronger than the SRCQ since $\text{lin}(\mathcal{T}_K(G(\bar{x}))) \subseteq \mathcal{T}_K(G(\bar{x})) \cap \bar{y}^\perp$ (cf. [1, Proposition 4.73]).

Definition 4 The constraint nondegeneracy of (1) is said to hold at \bar{x} if

$$G'(\bar{x})\mathbb{X} + \text{lin}(\mathcal{T}_K(G(\bar{x}))) = \mathbb{Y}.$$

Recall that the inner and outer second order tangent sets ([1, (3.49) and (3.50)]) to the given closed set $K \in \mathbb{Y}$ in the direction $d \in \mathbb{Y}$ are defined, respectively, by

$$\mathcal{T}_K^{i,2}(y, d) := \{w \in \mathbb{Y} \mid \text{dist}(y + td + \frac{1}{2}t^2w, C) = o(t^2), t \geq 0\}$$

and

$$\mathcal{T}_K^2(y, d) := \{w \in \mathbb{Y} \mid \exists t^k \downarrow 0, \text{dist}(y + t^k d + \frac{1}{2}(t^k)^2w, C) = o((t^k)^2)\}.$$

Note that in general, $\mathcal{T}_K^{i,2}(y, d) \neq \mathcal{T}_K^2(y, d)$ even for convex sets K (cf. [1, Section 3.3]). However, it follows from [1, Proposition 3.136] that if K is a C^2 -cone reducible [1, Definition 3.135] convex set (e.g., the polyhedron, second-order cone, positive semidefinite matrix cone and their Cartesian products), then equality always holds. In this case, $\mathcal{T}_K^2(y, d)$ will be simply called the second order tangent set to K at $y \in \mathbb{Y}$ in the direction $d \in \mathbb{Y}$.

Suppose that K in (1) is C^2 -cone reducible. The following second order sufficient condition is adopted from [1, (3.276)].

Definition 5 Let $\bar{x} \in \mathbb{X}$ be a stationary point of (1) with a Lagrange multiplier $\bar{y} \in M(\bar{x})$. The second order sufficient condition (SOSC) is said to hold at (\bar{x}, \bar{y}) if

$$\langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{y}) \Delta x \rangle - \sigma(\bar{y}, \mathcal{T}_K^2(G(\bar{x}), G'(\bar{x})\Delta x)) > 0 \quad \forall \Delta x \in \mathcal{C}(\bar{x}) \setminus \{0\}. \quad (9)$$

The following definition of the strong second order sufficient condition for (1) is an extension of [34, Definition 3.2] for the NLSDP. It is worth noting that when $M(\bar{x})$ is a singleton, both of them are the same.

Definition 6 Let $\bar{x} \in \mathbb{X}$ be a stationary point of (1) with a Lagrange multiplier $\bar{y} \in M(\bar{x})$. The strong second order sufficient condition (S-SOSC) is said to hold at (\bar{x}, \bar{y}) if

$$\langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{y}) \Delta x \rangle - \sigma(\bar{y}, \mathcal{T}_K^2(G(\bar{x}), G'(\bar{x})\Delta x)) > 0 \quad \forall \Delta x \in \text{app}(\bar{x}, \bar{y}) \setminus \{0\},$$

where $\text{app}(\bar{x}, \bar{y})$ is the outer approximation of the affine hull of the critical cone $\mathcal{C}(\bar{x})$ of (1) with respect to (\bar{x}, \bar{y}) , i.e.,

$$\text{app}(\bar{x}, \bar{y}) := \{\Delta x \in \mathbb{X} \mid G'(\bar{x})\Delta x \in \text{aff}(\mathcal{T}_K(G(\bar{x})) \cap \bar{y}^\perp)\}. \quad (10)$$

2.2 Eigenvalue decomposition of symmetric matrices

In this subsection, we introduce some useful preliminary results on eigenvalue decompositions of real symmetric matrices and the differentiability of the metric projectors over the positive semidefinite matrix cone \mathbb{S}_+^n .

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 $\Lambda(A) := \text{Diag}(\lambda(A))$. Consider the eigenvalue decomposition of A , i.e., $A = P\Lambda(A)P^T$, where $P \in \mathcal{O}^n$. By considering the index sets of positive, zero, and negative eigenvalues of A , we are able to write A in the following form

$$A = \begin{bmatrix} P_\alpha & P_\beta & P_\gamma \end{bmatrix} \begin{bmatrix} \Lambda(A)_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda(A)_{\gamma\gamma} \end{bmatrix} \begin{bmatrix} P_\alpha^T \\ P_\beta^T \\ P_\gamma^T \end{bmatrix}, \quad (11)$$

where α , β and γ are index sets defined by

$$\alpha := \{i \mid \lambda_i > 0\}, \quad \beta := \{i \mid \lambda_i = 0\} \quad \text{and} \quad \gamma := \{i \mid \lambda_i < 0\}. \quad (12)$$

We use $\mathcal{O}^n(A)$ to denote the set of all orthogonal matrices $P \in \mathcal{O}^n$ satisfying (11).

The next proposition, which was stated in [36, Proposition 4.4], illustrates that the eigenvectors of symmetric matrices, though not continuous, are upper Lipschitz continuous.

Proposition 1 *Given $A \in \mathbb{S}^n$, for any $H \in \mathbb{S}^n$, let P be an orthogonal matrix such that $X + H = P\Lambda(A + H)P^T$. Then, as $H \rightarrow 0$, we have*

$$\text{dist}(P, \mathcal{O}^n(A)) = O(\|H\|).$$

Let $\Pi_{\mathbb{S}_+^n} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ be the metric projection operator over the positive semidefinite matrix cone \mathbb{S}_+^n , i.e.,

$$\Pi_{\mathbb{S}_+^n}(A) := \arg \min_{X \in \mathbb{S}_+^n} \left\{ \frac{1}{2} \|X - A\|^2 \right\}, \quad A \in \mathbb{S}^n. \quad (13)$$

It is clear that $\Pi_{\mathbb{S}_+^n}(\cdot)$ is globally Lipschitz continuous. Furthermore, it follows from [35, Theorem 4.7] (see also [26, Proposition 9]) that $\Pi_{\mathbb{S}_+^n}(\cdot)$ is directionally differentiable at any $A \in \mathbb{S}^n$ and the directional derivative at A along the direction $H \in \mathbb{S}^n$ is given by

$$\Pi'_{\mathbb{S}_+^n}(A; H) = P \begin{bmatrix} \tilde{H}_{\alpha\alpha} & \tilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \tilde{H}_{\alpha\gamma} \\ \tilde{H}_{\alpha\beta}^T & \Pi_{\mathbb{S}_+^{|\beta|}}(\tilde{H}_{\beta\beta}) & 0 \\ \Xi_{\alpha\gamma}^T \circ \tilde{H}_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} P^T, \quad (14)$$

with $P \in \mathcal{O}^n(A)$, $\tilde{H} := P^T H P$ and $\Xi \in \mathbb{S}^n$ is defined by

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

where $0/0$ is defined to be 1. Meanwhile, it is well-known [35, Proposition 4.5 and Theorem 4.13] that the metric projection operator $\Pi_{\mathbb{S}_+^n}(\cdot)$ is strongly semismooth everywhere.

The following result on the characterizations of the Bouligand generalized Jacobian and Clarke generalized Jacobian of $\Pi_{\mathbb{S}_+^n}(\cdot)$ is taken from [26, Lemma 11].

Proposition 2 *Given $A \in \mathbb{S}^n$, it holds that $\mathcal{V} \in \partial_B \Pi_{\mathbb{S}_+^n}(A)$ ($\partial_C \Pi_{\mathbb{S}_+^n}(A)$) if and only if there is $\mathcal{W} \in \partial_B \Pi_{\mathbb{S}^{|\beta|}}(0)$ ($\partial_C \Pi_{\mathbb{S}^{|\beta|}}(0)$) such that for any $H \in \mathbb{S}^n$,*

$$\mathcal{V}(H) = P \begin{bmatrix} \tilde{H}_{\alpha\alpha} & \tilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \tilde{H}_{\alpha\gamma} \\ \tilde{H}_{\beta\alpha} & \mathcal{W}(\tilde{H}_{\beta\beta}) & 0 \\ \Xi_{\gamma\alpha} \circ \tilde{H}_{\gamma\alpha} & 0 & 0 \end{bmatrix} P^T,$$

where $P \in \mathcal{O}^n(A)$ and $\tilde{H} = P^T H P$. In particular, since both the zero mapping $\mathcal{W} \equiv 0$ and the identity mapping $\mathcal{W} \equiv \mathcal{I}$ from $\mathbb{S}^{|\beta|} \rightarrow \mathbb{S}^{|\beta|}$ are elements of $\partial_B \Pi_{\mathbb{S}^{|\beta|}}(0)$, the following two mappings \mathcal{V}_0 and $\mathcal{V}_{\mathcal{I}}$ defined by

$$\mathcal{V}_0(H) = P \begin{bmatrix} \tilde{H}_{\alpha\alpha} & \tilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \tilde{H}_{\alpha\gamma} \\ \tilde{H}_{\beta\alpha} & 0 & 0 \\ \Xi_{\gamma\alpha} \circ \tilde{H}_{\gamma\alpha} & 0 & 0 \end{bmatrix} P^T, \quad H \in \mathbb{S}^n \quad (16)$$

and

$$\mathcal{V}_{\mathcal{I}}(H) = P \begin{bmatrix} \tilde{H}_{\alpha\alpha} & \tilde{H}_{\alpha\beta} & \Xi_{\alpha\gamma} \circ \tilde{H}_{\alpha\gamma} \\ \tilde{H}_{\beta\alpha} & \tilde{H}_{\beta\beta} & 0 \\ \Xi_{\gamma\alpha} \circ \tilde{H}_{\gamma\alpha} & 0 & 0 \end{bmatrix} P^T, \quad H \in \mathbb{S}^n \quad (17)$$

are elements of $\partial_B \Pi_{\mathbb{S}_+^n}(A)$. Moreover, \mathcal{V}_0 and $\mathcal{V}_{\mathcal{I}}$ are independent of the choice of $P \in \mathcal{O}^n(A)$.

3 Existence of nonsingular elements of the generalized Jacobians

As mentioned in the introduction, our preliminary objective is to explore the existence of nonsingular elements within the generalized Jacobians of F defined in (3). This investigation sets the stage for the subsequent development of a semismooth Newton method that achieves quadratic convergence without the BD-regularity. To commence, we introduce the concept of the weak strict Robinson constraint qualification (W-SRCQ), which serves as an extension to the SRCQ condition (Definition 3).

Definition 7 Let $\bar{x} \in \mathbb{X}$ be a stationary point of (1) with a Lagrange multiplier $\bar{y} \in M(\bar{x})$. The weak strict Robinson constraint qualification (W-SRCQ) is said to hold at \bar{x} with respect to \bar{y} if

$$G'(\bar{x})\mathbb{X} + \text{aff}(\mathcal{T}_K(G(\bar{x})) \cap \bar{y}^\perp) = \mathbb{Y}.$$

Remark 1 The W-SRCQ condition is referred to as the *weak linear independence constraint qualification* (W-LICQ) in the context of nonlinear variational inequalities, as explored in [16]. Nevertheless, we choose not to adopt the W-LICQ terminology in this context. This is because the W-SRCQ is not simply a weaker form of classical LICQ; rather, it is a distinct concept that arises by incorporating the affine hull into the SRCQ.

As can be inferred directly from the definitions (Definitions 3 and 7), the W-SRCQ is weaker than the SRCQ. This implies that the W-SRCQ is also milder than the constraint non-degeneracy condition as outlined in Definition 4. Moreover, the example provided below demonstrates that W-SRCQ is, in fact, even weaker than the RCQ defined by Definition 2.

Example 1 Consider the following quadratic SDP problem:

$$\begin{aligned} \min \quad & \frac{1}{2}\|X_{11}\|^2 - \frac{1}{2}\|X_{22}\|^2 \\ \text{s.t.} \quad & X_{12} = 0, \quad X_{22} = 0, \\ & X \in \mathbb{S}_+^n, \end{aligned} \tag{18}$$

where $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \in \mathbb{S}^n$ with $X_{11} \in \mathbb{R}^{l_1 \times l_1}$ and $X_{22} \in \mathbb{R}^{l_2 \times l_2}$. It is easy to verify that $\bar{X} = 0$ is a local optimal solution with the following Lagrange multiplier set

$$M(\bar{X}) = \left\{ (\xi_{12}, \xi_{22}, \Gamma) \in \mathbb{R}^{l_1 \times l_2} \times \mathbb{R}^{l_2 \times l_2} \times \mathbb{S}_+^n \mid \begin{bmatrix} 0 & \xi_{12} \\ \xi_{12}^T & \xi_{22} \end{bmatrix} + \Gamma = 0 \right\}.$$

Thus, $(\bar{\xi}_{12}, \bar{\xi}_{22}, \bar{\Gamma}) = (0, 0, 0) \in M(\bar{X})$. Consider the KKT point $(\bar{X}, \bar{\xi}_{12}, \bar{\xi}_{22}, \bar{\Gamma}) = (0, 0, 0, 0)$. We know that the index sets defined by (12) are $\alpha = \gamma = \emptyset$ and $\beta = \{1, \dots, n\}$. It then can be checked directly by Definition 7 that the W-SRCQ holds at \bar{X} with respect to $(\bar{\xi}_{12}, \bar{\xi}_{22}, \bar{\Gamma})$. Meanwhile, since the multiplier set $M(\bar{X})$ is unbounded, we know that the RCQ does not hold at \bar{X} .

Next, we are to introduce the concept of the weak second order condition for a KKT solution (\bar{x}, \bar{y}) of (1), which is a generalization of the SOSC (Definition 5).

Definition 8 Let $\bar{x} \in \mathbb{X}$ be a stationary point of (1) with a Lagrange multiplier $\bar{y} \in M(\bar{x})$. The weak second order condition (W-SOC) is said to hold at (\bar{x}, \bar{y}) if

$$\langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{y}) \Delta x \rangle - \sigma(\bar{y}, \mathcal{T}_K^2(G(\bar{x}), G'(\bar{x})\Delta x)) > 0 \quad \forall \Delta x \in \text{appl}(\bar{x}, \bar{y}) \setminus \{0\}, \tag{19}$$

where $\text{appl}(\bar{x}, \bar{y})$ is defined by

$$\text{appl}(\bar{x}, \bar{y}) := \{ \Delta x \in \mathbb{X} \mid G'(\bar{x})\Delta x \in \text{lin}(\mathcal{T}_K(G(\bar{x})) \cap \bar{y}^\perp) \}. \tag{20}$$

Remark 2 The W-SOC is referred to as the *second order condition* for nonlinear variational inequalities in [16]. However, to emphasize Definition 8 is a weaker variant of the SOSC, we prefer the term of W-SOC here. Moreover, it can be seen from the definitions that at a KKT point, the following implications hold:

$$\text{S-SOSC} \implies \text{SOSC} \implies \text{W-SOC}.$$

Furthermore, it is important to note that the W-SOC does not suffice to ensure local optimality for general optimization problems. Additionally, within the domain of convex problems, the W-SOC alone does not necessarily lead to the isolation of the optimal solution.

Henceforth, our primary focus will be on the following nonlinear semidefinite programming (NLSDP) problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0, \\ & g(x) \in \mathbb{S}_+^n. \end{aligned} \tag{21}$$

It should be noted that the results derived herein can be similarly extended through straightforward generalizations to cases where K in (1) represents the Cartesian product of a finite number of positive semidefinite cones and zero vectors. Furthermore, considering that \mathbb{R}^n can be represented as the Cartesian product of n one-dimensional positive semidefinite cones \mathbb{S}^1 , our

analysis encompasses classical nonlinear programming (NLP) as well. In addition, we call the NLSDP (21) a convex semidefinite programming (CSDP) problem if f is convex, h is affine and g is \mathbb{S}_+^n -convex, i.e., for any $x, y \in \mathbb{X}$ and $t \in (0, 1)$,

$$g(tx + (1-t)y) - tg(x) - (1-t)g(y) \in \mathbb{S}_+^n.$$

For the NLSDP (21), the nonsmooth system (3) which is equivalent to the KKT optimality condition (2), takes the following form:

$$F(x, \xi, \Gamma) = \begin{bmatrix} \nabla_x L(x, \xi, \Gamma) \\ h(x) \\ -g(x) + \Pi_{\mathbb{S}_+^n}(g(x) + \Gamma) \end{bmatrix} = 0, \quad (22)$$

where the associated Lagrange function is given by $L(x, \xi, \Gamma) = f(x) + \langle \xi, h(x) \rangle + \langle \Gamma, g(x) \rangle$ and $\Pi_{\mathbb{S}_+^n}(\cdot)$ is the metric projection operator over \mathbb{S}_+^n defined by (13).

Let \bar{x} be a stationary solution of (22) with the Lagrange multipliers $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$. Denote $\bar{A} = g(\bar{x}) + \bar{\Gamma}$. Consider the corresponding eigenvalue decomposition (11) of \bar{A} with the index sets α , β and γ defined by (12), i.e.,

$$A = g(\bar{x}) + \bar{\Gamma} = \bar{P} \begin{bmatrix} \bar{\Lambda}_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{\Lambda}_{\gamma\gamma} \end{bmatrix} \bar{P}^T, \quad (23)$$

where $\bar{\Lambda} = \Lambda(A) = \text{Diag}(\bar{\lambda}_i)$ with $\bar{\lambda} := \lambda(A)$ and $\bar{P} \in \mathcal{O}^n(A)$. Thus, the tangent cone $\mathcal{T}_{\mathbb{S}_+^n}(g(\bar{x}))$ of \mathbb{S}_+^n defined by (6) has the following explicit expression (cf. e.g., [34, (17)]):

$$\mathcal{T}_{\mathbb{S}_+^n}(g(\bar{x})) = \left\{ \bar{P}B\bar{P}^T \in \mathbb{S}^n \mid \begin{bmatrix} B_{\beta\beta} & B_{\beta\gamma} \\ B_{\gamma\beta} & B_{\gamma\gamma} \end{bmatrix} \succeq 0 \right\}. \quad (24)$$

Furthermore, the detailed characterizations of the following sets can be found in [34].

$$\mathcal{T}_{\mathbb{S}_+^n}(g(\bar{x})) \cap \bar{\Gamma}^\perp = \left\{ \bar{P}B\bar{P}^T \in \mathbb{S}^n \mid B_{\beta\beta} \succeq 0, B_{\beta\gamma} = 0, B_{\gamma\gamma} = 0 \right\}, \quad (25)$$

$$\text{lin}(\mathcal{T}_{\mathbb{S}_+^n}(g(\bar{x}))) = \left\{ \bar{P}B\bar{P}^T \in \mathbb{S}^n \mid B_{\beta\beta} = 0, B_{\beta\gamma} = 0, B_{\gamma\gamma} = 0 \right\}, \quad (26)$$

$$\text{aff}(\mathcal{T}_{\mathbb{S}_+^n}(g(\bar{x})) \cap \bar{\Gamma}^\perp) = \left\{ \bar{P}B\bar{P}^T \in \mathbb{S}^n \mid B_{\beta\gamma} = 0, B_{\gamma\gamma} = 0 \right\}. \quad (27)$$

For the notational simplicity, we further define the linear operator $\bar{\mathcal{B}}: \mathbb{X} \rightarrow \mathbb{S}^n$ by

$$\bar{\mathcal{B}}(\Delta x) = \bar{P}^T (g'(\bar{x})\Delta x) \bar{P}, \quad \Delta x \in \mathbb{X}. \quad (28)$$

Thus, together with (8), we know from (25) that the critical cone $\mathcal{C}(\bar{x})$ of the NLSDP (21) at the stationary point \bar{x} takes the following form:

$$\mathcal{C}(\bar{x}) = \left\{ \Delta x \in \mathbb{X} \mid h'(\bar{x})\Delta x = 0, \bar{\mathcal{B}}(\Delta x)_{\beta\beta} \succeq 0, \bar{\mathcal{B}}(\Delta x)_{\beta\gamma} = 0, \bar{\mathcal{B}}(\Delta x)_{\gamma\gamma} = 0 \right\}. \quad (29)$$

It follows from (27) and (26) that for the stationary point $\bar{x} \in \mathbb{X}$ of (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$, we have

$$\text{app}(\bar{x}, \bar{\xi}, \bar{\Gamma}) = \left\{ \Delta x \in \mathbb{X} \mid h'(\bar{x})\Delta x = 0, \bar{\mathcal{B}}(\Delta x)_{\beta\gamma} = 0, \bar{\mathcal{B}}(\Delta x)_{\gamma\gamma} = 0 \right\} \quad (30)$$

and

$$\text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma}) = \{ \Delta x \in \mathbb{X} \mid h'(\bar{x})\Delta x = 0, \bar{\mathcal{B}}(\Delta x)_{\beta\beta} = 0, \bar{\mathcal{B}}(\Delta x)_{\beta\gamma} = 0, \bar{\mathcal{B}}(\Delta x)_{\gamma\gamma} = 0 \}, \quad (31)$$

where the sets $\text{app}(\bar{x}, \bar{\xi}, \bar{\Gamma})$ and $\text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma})$ are defined by (10) and (20), respectively. Moreover, by employing the explicit formulas (27) and (26), we obtain the following useful characterizations of the W-SRCQ (Definition 7) and the constraint nondegeneracy (Definition 4) for the NLSDP (21). We omit the detailed proof here.

Lemma 1 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$.*

(i) *The W-SRCQ holds at \bar{x} with respect to $(\bar{\xi}, \bar{\Gamma})$ if and only if*

$$\begin{cases} h'(\bar{x})^* \Delta \xi + \bar{\mathcal{B}}^* \Delta \Gamma = 0, \\ \Delta \Gamma_{\alpha\alpha} = 0, \Delta \Gamma_{\alpha\beta} = 0, \Delta \Gamma_{\alpha\gamma} = 0, \Delta \Gamma_{\beta\beta} = 0 \end{cases} \implies (\Delta \xi, \Delta \Gamma) = 0. \quad (32)$$

(ii) *The constraint nondegeneracy holds at \bar{x} if and only if*

$$\begin{cases} h'(\bar{x})^* \Delta \xi + \bar{\mathcal{B}}^* \Delta \Gamma = 0, \\ \Delta \Gamma_{\alpha\alpha} = 0, \Delta \Gamma_{\alpha\beta} = 0, \Delta \Gamma_{\alpha\gamma} = 0 \end{cases} \implies (\Delta \xi, \Delta \Gamma) = 0. \quad (33)$$

For the stationary point $\bar{x} \in \mathbb{X}$ of the NLSDP (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$, the support function of the second order tangent set in the second order conditions (Definitions 5, 6 and 8) has the following explicit characterization [34, Lemma 3.1]:

$$\sigma(\bar{y}, \mathcal{T}_K^2(G(\bar{x}), G'(\bar{x})\Delta x)) = \sigma(\bar{\Gamma}, \mathcal{T}_{\mathbb{S}_+^n}^2(g(\bar{x}), g'(\bar{x})\Delta x)) = 2 \sum_{i \in \alpha} \sum_{j \in \gamma} \frac{\bar{\lambda}_j}{\bar{\lambda}_i} \bar{\mathcal{B}}(\Delta x)_{ij}^2, \quad \Delta x \in \mathbb{X}, \quad (34)$$

where $K = \{0\} \times \mathbb{S}_+^n$, $G(x) = (h(x), g(x))$ and $\bar{y} = (\bar{\xi}, \bar{\Gamma})$.

Define $\mathbb{Z} := \mathbb{X} \times \mathbb{R}^m \times \mathbb{S}^n$. Let $(\bar{x}, \bar{\xi}, \bar{\Gamma})$ be a KKT solution of the NLSDP (21). Consider the Bouligand generalized Jacobian $\partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ and Clarke generalized Jacobian $\partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ of the KKT nonsmooth mapping F given by (22). It follows from [34, Lemma 2.1] (see also [2, Lemma 1]) that $\bar{U} \in \partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ ($\partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma})$) if and only if there exists $\bar{V} \in \partial_B \Pi_{\mathbb{S}_+^m}(g(\bar{x}) + \bar{\Gamma})$ ($\partial_C \Pi_{\mathbb{S}_+^m}(g(\bar{x}) + \bar{\Gamma})$) such that for any $(\Delta x, \Delta \xi, \Delta \Gamma) \in \mathbb{Z}$,

$$\bar{U}(\Delta x, \Delta \xi, \Delta \Gamma) = \begin{bmatrix} \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x + h'(\bar{x})^* \Delta \xi + g'(\bar{x})^* \Delta \Gamma \\ h'(\bar{x}) \Delta x \\ -g'(\bar{x}) \Delta x + \bar{V}(g'(\bar{x}) \Delta x + \Delta \Gamma) \end{bmatrix}. \quad (35)$$

In particular, let \bar{V}_0 and $\bar{V}_{\mathcal{I}}$ be two mappings defined by (16) and (17) for $A = g(\bar{x}) + \bar{\Gamma}$, respectively. Define the following two elements \bar{U}_0 and $\bar{U}_{\mathcal{I}}$ in $\partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ with \bar{V}_0 and $\bar{V}_{\mathcal{I}}$ by (35), respectively, i.e., for any $(\Delta x, \Delta \xi, \Delta \Gamma) \in \mathbb{Z}$,

$$\bar{U}_0(\Delta x, \Delta \xi, \Delta \Gamma) = \begin{bmatrix} \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x + h'(\bar{x})^* \Delta \xi + g'(\bar{x})^* \Delta \Gamma \\ h'(\bar{x}) \Delta x \\ -g'(\bar{x}) \Delta x + \bar{V}_0(g'(\bar{x}) \Delta x + \Delta \Gamma) \end{bmatrix} \quad (36)$$

and

$$\bar{U}_{\mathcal{I}}(\Delta x, \Delta \xi, \Delta \Gamma) = \begin{bmatrix} \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x + h'(\bar{x})^* \Delta \xi + g'(\bar{x})^* \Delta \Gamma \\ h'(\bar{x}) \Delta x \\ -g'(\bar{x}) \Delta x + \bar{V}_{\mathcal{I}}(g'(\bar{x}) \Delta x + \Delta \Gamma) \end{bmatrix}. \quad (37)$$

In the following two subsections, we will study the existence of nonsingular elements in generalized Jacobians at a KKT point $(\bar{x}, \bar{\xi}, \bar{\Gamma})$.

3.1 Sufficient conditions for the existence of nonsingular elements in generalized Jacobians

In this subsection, we study the sufficient conditions to ensure the existence of nonsingular elements in generalized Jacobians of the KKT nonsmooth mapping F (22). The following results illustrate that under the W-SOC and constraint nondegeneracy (or the S-SOSC and W-SRCQ), the mapping \bar{U}_0 (or \bar{U}_I) defined in (36) (or (37)) is indeed a nonsingular element of the Bouligand generalized Jacobian of F .

Theorem 1 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$.*

- (i) *If the W-SOC and constraint nondegeneracy hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, then $\bar{U}_0 \in \partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ in (36) is nonsingular.*
- (ii) *If the S-SOSC and W-SRCQ hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, then $\bar{U}_I \in \partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ in (37) is nonsingular.*

Proof Let $A = g(\bar{x}) + \bar{\Gamma}$ satisfy the eigenvalue decomposition (23) and $\bar{P} \in \mathcal{O}^n(A)$. We only show (i) holds, as (ii) can be established through a similar argument.

Suppose that $\bar{U}_0 \in \partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ is singular, i.e., there is $0 \neq (\Delta x, \Delta \xi, \Delta \Gamma) \in \mathbb{Z}$ such that $\bar{U}_0(\Delta x, \Delta \xi, \Delta \Gamma) = 0$. It then follows from (36) and Proposition 2 that

$$\nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x + h'(\bar{x})^* \Delta \xi + g'(\bar{x})^* \Delta \Gamma = 0, \quad (38)$$

$$h'(\bar{x}) \Delta x = 0, \quad (39)$$

$$\widetilde{\Delta \Gamma}_{\alpha\alpha} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta} = 0, \quad (40)$$

$$\bar{B}(\Delta x)_{\beta\beta} = 0, \quad \bar{B}(\Delta x)_{\beta\gamma} = 0, \quad \bar{B}(\Delta x)_{\gamma\gamma} = 0, \quad (41)$$

$$(E_{\alpha\gamma} - \bar{\Xi}_{\alpha\gamma}) \circ \bar{B}(\Delta x)_{\alpha\gamma} - \bar{\Xi}_{\alpha\gamma} \circ \widetilde{\Delta \Gamma}_{\alpha\gamma} = 0, \quad (42)$$

where $\widetilde{\Delta \Gamma} := \bar{P}^T \Delta \Gamma \bar{P}$, the linear operator $\bar{B} : \mathbb{X} \rightarrow \mathbb{S}^n$ is defined by (28), the matrix $\bar{\Xi} \in \mathbb{S}^n$ is given by (15) for A and $E \in \mathbb{S}^n$ denotes the matrix whose elements are all ones. Then, by (38)-(42) and (34), we obtain that

$$\begin{aligned} 0 &= \langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x + h'(\bar{x})^* \Delta \xi + g'(\bar{x})^* \Delta \Gamma \rangle \\ &= \langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x \rangle + \langle h'(\bar{x}) \Delta x, \Delta \xi \rangle + \langle \bar{B}(\Delta x), \widetilde{\Delta \Gamma} \rangle \\ &= \langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x \rangle + 2 \sum_{i \in \alpha} \sum_{j \in \gamma} \frac{-\bar{\lambda}_j}{\bar{\lambda}_i} (\bar{B}(\Delta x)_{ij})^2 \\ &= \langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x \rangle - \sigma(\bar{\Gamma}, \mathcal{T}_{\mathbb{S}_+^n}^2(g(\bar{x}), g'(\bar{x}) \Delta x)). \end{aligned} \quad (43)$$

Meanwhile, by (41) and (39), we know from (31) that $\Delta x \in \text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma})$. Since the W-SOC holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, we know from (43) that $\Delta x = 0$. It then follows from (42) that $\widetilde{\Delta \Gamma}_{\alpha\gamma} = 0$. This, together with (40) and (38), yields that

$$\begin{cases} h'(\bar{x})^* \Delta \xi + \bar{B}^* \widetilde{\Delta \Gamma} = 0, \\ \widetilde{\Delta \Gamma}_{\alpha\alpha} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\gamma} = 0. \end{cases}$$

Since the constraint nondegeneracy holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, we know from (33) in Lemma 1 (ii) that $(\Delta \xi, \Delta \Gamma) = 0$, which contradicts with $(\Delta x, \Delta \xi, \Delta \Gamma) \neq 0$. The proof is then completed. \square

Remark 3 It should be noted that Theorem 1 can be easily extended to cases where K in (1) is a Cartesian product of a finite number of positive semidefinite cones and zero vectors. By regarding a polyhedron as a Cartesian product of a finite number of one-dimensional positive semidefinite cones and zero vectors, the results derived in Theorem 1 are consistent with the findings presented in [16, P644].

It is natural to question whether the W-SRCQ and W-SOC are sufficient for the existence of nonsingular elements in the Bouligand generalized Jacobian $\partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$. We address this question with the following proposition, which confirms the sufficiency under the condition that $|\beta| \leq 1$, i.e., the number of zero eigenvalues of $g(\bar{x}) + \bar{\Gamma}$ is less than or equal to one.

Proposition 3 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$. Suppose that $|\beta| \leq 1$. If the W-SOC and W-SRCQ hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, then at least one of $\{\bar{\mathcal{U}}_0, \bar{\mathcal{U}}_{\mathcal{I}}\}$ is nonsingular.*

Proof Consider the following two cases.

Case 1: $|\beta| = 0$. In this case, from Lemma 1, we know that the W-SRCQ is equivalent to the constraint nondegeneracy. Moreover, (30) and (31) imply that $\text{app}(\bar{x}, \bar{\xi}, \bar{\Gamma}) = \text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma})$, and thus further imply that the W-SOC is equivalent to the S-SOSC. We also have from Proposition 2 that $\partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma}) = \{F'(\bar{x}, \bar{\xi}, \bar{\Gamma})\} = \{\bar{\mathcal{U}}_0\} = \{\bar{\mathcal{U}}_{\mathcal{I}}\}$. The desired result follows directly from Theorem 1.

Case 2: $|\beta| = 1$. Since the W-SOC holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, we know from (i) of Theorem 1 that if the constraint nondegeneracy holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, then $\bar{\mathcal{U}}_0$ is nonsingular. Suppose that the constraint nondegeneracy does not hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$. Since the W-SRCQ holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, we know from Lemma 1 that there exists $(\Delta\xi, \Delta\Gamma) \neq 0$ such that

$$\begin{cases} h'(\bar{x})^* \Delta\xi + \bar{\mathcal{B}}^* \Delta\Gamma = 0, \\ \Delta\Gamma_{\alpha\alpha} = 0, \Delta\Gamma_{\alpha\beta} = 0, \Delta\Gamma_{\alpha\gamma} = 0, \\ \Delta\Gamma_{\beta\beta} \neq 0. \end{cases}$$

Next, we show that $\text{app}(\bar{x}, \bar{\xi}, \bar{\Gamma}) = \text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma})$ with their explicit expressions given in (30) and (31), respectively. Clearly, $\text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma}) \subseteq \text{app}(\bar{x}, \bar{\xi}, \bar{\Gamma})$. For any $\Delta x \in \text{app}(\bar{x}, \bar{\xi}, \bar{\Gamma})$, we have

$$0 = \langle \Delta x, h'(\bar{x})^* \Delta\xi + \bar{\mathcal{B}}^* \Delta\Gamma \rangle = \langle h'(\bar{x}) \Delta x, \Delta\xi \rangle + \langle \bar{\mathcal{B}}(\Delta x), \Delta\Gamma \rangle = \langle \bar{\mathcal{B}}(\Delta x)_{\beta\beta}, \Delta\Gamma_{\beta\beta} \rangle, \quad (44)$$

where the last equation follows from (30). Since $|\beta| = 1$, by noting that $\Delta\Gamma_{\beta\beta} \neq 0$, we know that $\bar{\mathcal{B}}(\Delta x)_{\beta\beta} = 0$. Therefore, we obtain from (31) that $\text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma}) \subseteq \text{app}(\bar{x}, \bar{\xi}, \bar{\Gamma})$. Then, by Definitions 8 and 6, the W-SOC and S-SOSC are equivalent. Therefore, we know from (ii) of Theorem 1 that $\bar{\mathcal{U}}_{\mathcal{I}}$ is nonsingular. The proof is then completed. \square

Remark 4 The condition $|\beta| = 0$ is satisfied if and only if the well-known strict complementarity condition holds for (21). In this case, it can be verified that the sets in (25), (26), and (27) are identical, and the sets in (30) and (30) are likewise identical. Therefore, the W-SRCQ is equivalent to the constraint nondegeneracy, and the W-SOC is equivalent to the S-SOSC. Moreover, in this scenario, the mapping F is smooth near $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, reducing our conditions to those applicable to the classical Newton method.

Remark 5 It is worth noting that the condition $|\beta| \leq 1$ is always satisfied for nonlinear programming (NLP). Actually, as we mentioned before, the positive orthant \mathbb{R}_+^n can be represented as the Cartesian product of n one-dimensional positive semidefinite cones. Consequently, based on Proposition 3, the existence of a nonsingular element in the Bouligand generalized Jacobian $\partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ is ensured under the W-SOC and the W-SRCQ for the conventional NLP, which recovers the corresponding result obtained by Izmailov and Solodov [16, Proposition 6].

However, the following simple example illustrates that for the NLSDP (21), if the number of zero eigenvalues of $g(\bar{x}) + \bar{\Gamma}$ is greater than one, i.e., $|\beta| > 1$, the W-SRCQ and W-SOC are not sufficient to guarantee the existence of nonsingular elements in $\partial_B F(\bar{x}, \bar{\xi}, \bar{\Gamma})$. This demonstrates the fundamental differences between the NLSDP and the NLP.

Example 2 Consider the following SDP problem:

$$\begin{aligned} \min \quad & 0 \\ \text{s.t.} \quad & x_{12} = 0, \\ & X \in \mathbb{S}_+^2, \end{aligned}$$

where $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \in \mathbb{S}^2$. It is clear that $(\bar{X}, \bar{\xi}, \bar{\Gamma}) = (0, 0, 0)$ is a KKT point. We have $\beta = \{1, 2\}$ and $\alpha = \gamma = \emptyset$. It is easy to verify that the W-SRCQ and W-SOC hold at $(\bar{X}, \bar{\xi}, \bar{\Gamma})$. Moreover, it follows from [26, Lemma 11] that $\mathcal{U} \in \partial_B F(\bar{X}, \bar{\xi}, \bar{\Gamma})$ if and only if there exists $\Omega \in R$ such that for any $(\Delta X, \Delta \xi, \Delta \Gamma) \in \mathbb{S}^2 \times \mathbb{R} \times \mathbb{S}^2$,

$$\mathcal{U}(\Delta X, \Delta \xi, \Delta \Gamma) = \begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \Delta \xi \\ \Delta \xi & 0 \end{bmatrix} + \Delta \Gamma \\ \Delta X_{12} \\ -\Delta X + \Omega \circ (\Delta X + \Delta \Gamma) \end{bmatrix}, \quad (45)$$

where $R \in \mathbb{S}^n$ is the set of matrices defined by

$$R := \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\} \cup \left\{ \begin{bmatrix} 0 & t \\ t & 1 \end{bmatrix} \mid t \in [0, 1] \right\} \cup \left\{ \begin{bmatrix} 1 & t \\ t & 0 \end{bmatrix} \mid t \in [0, 1] \right\}.$$

Then, one can check easily that all elements in $\partial_B F(\bar{X}, \bar{\xi}, \bar{\Gamma})$ are singular.

Next, let us consider the Clarke generalized Jacobian defined by (5) of the KKT nonsmooth mapping (22). Interestingly, one may observe that for Example 2, the convex combination $t\mathcal{U}_0 + (1-t)\mathcal{U}_{\mathcal{I}} \in \partial_C F(\bar{X}, \bar{\xi}, \bar{\Gamma})$ for any $t \in (0, 1)$ is actually nonsingular, where \mathcal{U}_0 and $\mathcal{U}_{\mathcal{I}}$ are elements in $\partial_B F(\bar{X}, \bar{\xi}, \bar{\Gamma})$ defined by (45) with $\Omega = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, respectively. In fact, we have the following general results, which demonstrate that the W-SOC and W-SRCQ guarantee the existence of a nonsingular element in the Clarke generalized Jacobian.

Proposition 4 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$. Suppose that the W-SOC and W-SRCQ hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$. Let $\bar{\mathcal{U}}_0$ and $\bar{\mathcal{U}}_{\mathcal{I}}$ be two elements defined in (36) and (37), respectively.*

(i) *There exists $t \in (0, 1)$ such that*

$$t\bar{\mathcal{U}}_0 + (1-t)\bar{\mathcal{U}}_{\mathcal{I}} \in \partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma}) \text{ is nonsingular.}$$

(ii) *If, additionally, the NLSDP (21) is convex, that is, it becomes a CSDP, then for any $t \in (0, 1)$,*

$$t\bar{\mathcal{U}}_0 + (1-t)\bar{\mathcal{U}}_{\mathcal{I}} \in \partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma}) \text{ is nonsingular.}$$

Proof Let $A = g(\bar{x}) + \bar{\Gamma}$ satisfy the eigenvalue decomposition (23) and $\bar{P} \in \mathcal{O}^n(A)$ with the index sets α , β and γ defined by (12). Recall the linear operator $\bar{\mathcal{B}} : \mathbb{X} \rightarrow \mathbb{S}^n$ defined in (28).

(i) Assume that for any $t \in (0, 1)$, $t\bar{\mathcal{U}}_0 + (1-t)\bar{\mathcal{U}}_{\mathcal{I}} \in \partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ is singular. Define a sequence of elements $\{\bar{\mathcal{U}}^k\}$ as $\bar{\mathcal{U}}^k = \frac{k}{k+1}\bar{\mathcal{U}}_0 + \frac{1}{k+1}\bar{\mathcal{U}}_{\mathcal{I}} \in \partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma})$ for each k . Since $\bar{\mathcal{U}}^k$ is singular, we

know that for each k , there are $0 \neq (\Delta x^k, \Delta \xi^k, \Delta \Gamma^k) \in \mathbb{Z}$ such that $\bar{U}^k(\Delta x^k, \Delta \xi^k, \Delta \Gamma^k) = 0$. Thus, for each k , we know from (35) and Proposition 2 that

$$\nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x^k + h'(\bar{x})^* \Delta \xi^k + g'(\bar{x})^* \Delta \Gamma^k = 0, \quad (46)$$

$$h'(\bar{x}) \Delta x^k = 0, \quad (47)$$

$$\widetilde{\Delta \Gamma}_{\alpha\alpha}^k = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta}^k = 0, \quad (48)$$

$$\bar{\mathcal{B}}(\Delta x^k)_{\beta\gamma} = 0, \quad \bar{\mathcal{B}}(\Delta x^k)_{\gamma\gamma} = 0, \quad (49)$$

$$(E_{\alpha\gamma} - \bar{\Xi}_{\alpha\gamma}) \circ \bar{\mathcal{B}}(\Delta x^k)_{\alpha\gamma} - \bar{\Xi}_{\alpha\gamma} \circ \widetilde{\Delta \Gamma}_{\alpha\gamma}^k = 0, \quad (50)$$

$$k\bar{\mathcal{B}}(\Delta x^k)_{\beta\beta} - \widetilde{\Delta \Gamma}_{\beta\beta}^k = 0, \quad (51)$$

where $\widetilde{\Delta \Gamma}^k = \bar{P}^T \Delta \Gamma^k \bar{P}$ with $\bar{P} \in \mathcal{O}^n(A)$ and $A = g(\bar{x}) + \bar{T}$, the matrix $\bar{\Xi} \in \mathbb{S}^n$ is given by (15) and $E \in \mathbb{S}^n$ denotes the matrix whose elements are all ones. We then conclude that for each k , $\Delta x^k \neq 0$. Indeed, if there exists some \bar{k} such that $\Delta x^{\bar{k}} = 0$, then by (48), (50) and (51), we obtain that

$$\widetilde{\Delta \Gamma}_{\alpha\alpha}^{\bar{k}} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta}^{\bar{k}} = 0, \quad \widetilde{\Delta \Gamma}_{\beta\beta}^{\bar{k}} = 0 \quad \text{and} \quad \widetilde{\Delta \Gamma}_{\alpha\gamma}^{\bar{k}} = 0. \quad (52)$$

Thus, combining with (46), we know that

$$\begin{cases} h'(\bar{x})^* \Delta \xi^{\bar{k}} + \bar{\mathcal{B}}^* \widetilde{\Delta \Gamma}^{\bar{k}} = 0, \\ \widetilde{\Delta \Gamma}_{\alpha\alpha}^{\bar{k}} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta}^{\bar{k}} = 0, \quad \widetilde{\Delta \Gamma}_{\beta\beta}^{\bar{k}} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\gamma}^{\bar{k}} = 0. \end{cases} \quad (53)$$

It then follows from (32) in Lemma 1 (i) that $(\Delta \xi^{\bar{k}}, \widetilde{\Delta \Gamma}^{\bar{k}}) = 0$, which, together with $\Delta x^{\bar{k}} = 0$, contradicts with the assumption. Now, since $\Delta x^k \neq 0$ for all k , without loss of generality, we may assume that $\|\Delta x^k\| = 1$ and $\Delta x^k \rightarrow \Delta x^\infty \neq 0$ as $k \rightarrow \infty$. By taking the inner product with Δx^k in (46), we obtain from (47)-(51) that for each k ,

$$\langle \Delta x^k, \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x^k \rangle + 2 \sum_{i \in \alpha} \sum_{j \in \gamma} \frac{-\bar{\lambda}_j}{\bar{\lambda}_i} \bar{\mathcal{B}}(\Delta x^k)_{ij}^2 + \sum_{i \in \beta} \sum_{j \in \beta} k \bar{\mathcal{B}}(\Delta x^k)_{ij}^2 = 0. \quad (54)$$

Let $\mathcal{Q} : \mathbb{X} \rightarrow \mathbb{X}$ be the linear map such that

$$\langle d, \mathcal{Q}(d) \rangle = \langle d, \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) d \rangle + 2 \sum_{i \in \alpha} \sum_{j \in \gamma} \frac{-\bar{\lambda}_j}{\bar{\lambda}_i} \bar{\mathcal{B}}(d)_{ij}^2, \quad \forall d \in \mathbb{X}.$$

Then there exists $q > 0$ that $|\mathcal{Q}(d)| \leq q \|d\|^2$ for any $d \in \mathbb{X}$. Therefore, we have from (54) that for each k ,

$$|\bar{\mathcal{B}}(\Delta x^k)_{ij}|^2 \leq \frac{1}{k} q, \quad (i, j) \in \beta \times \beta.$$

As a result, $\bar{\mathcal{B}}(\Delta x^\infty)_{\beta\beta} = 0$, which, together with (49), implies $\Delta x^\infty \in \text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma})$. Further note that (54) implies that

$$\langle \Delta x^k, \mathcal{Q} \Delta x^k \rangle \leq 0.$$

Taking the limit as $k \rightarrow +\infty$, we obtain $\langle \Delta x^\infty, \mathcal{Q} \Delta x^\infty \rangle \leq 0$, which contradicts the W-SOC given that $0 \neq \Delta x^\infty \in \text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma})$. Therefore, there exists a nonsingular \bar{U}^k .

(ii) For the CSDP, if there is $t \in (0, 1)$ that $\bar{U}_t = t\bar{U}_0 + (1-t)\bar{U}_T$ is singular. Then, there is $(\Delta x, \Delta\xi, \Delta\Gamma) \neq 0$ such that $\bar{U}_t(\Delta x, \Delta\xi, \Delta\Gamma) = 0$. Similarly to the proof of (i), we obtain from (52)-(54) that $\Delta x \neq 0$ and

$$\langle \Delta x, \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x \rangle + 2 \sum_{i \in \alpha} \sum_{j \in \gamma} \frac{-\bar{\lambda}_j}{\bar{\lambda}_i} \bar{\mathcal{B}}(\Delta x)_{ij}^2 + \sum_{i \in \beta} \sum_{j \in \beta} \frac{t}{1-t} \bar{\mathcal{B}}(\Delta x)_{ij}^2 = 0. \quad (55)$$

By convexity, we know that $\nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma})$ is positive semidefinite. Thus, $\bar{\mathcal{B}}(\Delta x)_{\beta\beta} = 0$. This, together with (47) and (49), yields $0 \neq \Delta x \in \text{appl}(\bar{x}, \bar{\xi}, \bar{\Gamma})$, and further implies the contradiction between (55) the W-SOC. Therefore, we conclude that $\bar{U}_t = t\bar{U}_0 + (1-t)\bar{U}_T$ is nonsingular for any $t \in (0, 1)$. \square

Remark 6 According to the proof of part (i) of Proposition 4, there exists only a finite number of singular elements on the line segment connecting \bar{U}_0 and \bar{U}_T . In fact, by considering the matrix representations of \bar{U}_0 and \bar{U}_T and invoking the fundamental theorem of algebra, we are able to deduce that the non-Zero polynomial $\det(t\bar{U}_0 + (1-t)\bar{U}_T)$ has a number of roots that does not exceed its degree.

3.2 Necessary conditions for the existence of nonsingular elements in generalized Jacobians

We first show that when the NLSDP (21) is convex, i.e., for the CSDP, the conditions of Theorem 1 are actually necessary. The proof relies on the following lemma regarding the characterization of W-SOC violations.

Lemma 2 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of the NLSDP with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$. Suppose that the W-SOC does not hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$ and $\nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma})$ is positive semidefinite, there exists $\Delta x \neq 0$ that*

$$\begin{cases} \nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x = 0, \\ h'(\bar{x}) \Delta x = 0, \\ \bar{\mathcal{B}}(\Delta x)_{\alpha\gamma} = 0, \quad \bar{\mathcal{B}}(\Delta x)_{\beta\beta} = 0, \quad \bar{\mathcal{B}}(\Delta x)_{\beta\gamma} = 0, \quad \bar{\mathcal{B}}(\Delta x)_{\gamma\gamma} = 0. \end{cases} \quad (56)$$

The proof is omitted here for the sake of simplicity. Based on this lemma, we have the following characterizations of the nonsingularity of \bar{U}_0 and \bar{U}_T defined by (36) and (37), respectively, for the CSDP.

Proposition 5 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of the CSDP with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$.*

- (i) *The W-SOC and constraint nondegeneracy hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$ if and only if \bar{U}_0 defined in (36) is nonsingular;*
- (ii) *the S-SOSC and W-SRCQ hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$ if and only if \bar{U}_T defined in (37) is nonsingular.*

Proof Let $A = g(\bar{x}) + \bar{\Gamma}$ satisfy the eigenvalue decomposition (23) and $\bar{P} \in \mathcal{O}^n(A)$ with the index sets α , β and γ defined by (12). Recall the linear operator $\bar{\mathcal{B}} : \mathbb{X} \rightarrow \mathbb{S}^n$ defined by (28). We only show (i), as (ii) can be established through a similar argument.

By Theorem 1, we only need to show that the nonsingularity of \bar{U}_0 implies the W-SOC and constraint nondegeneracy. Suppose $\bar{U}_0(\Delta x, \Delta\xi, \Delta\Gamma) = 0$. We know from (36) and Proposition 2

that

$$\nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma}) \Delta x + h'(\bar{x})^* \Delta \xi + g'(\bar{x})^* \Delta \Gamma = 0, \quad (57)$$

$$h'(\bar{x}) \Delta x = 0, \quad (58)$$

$$\widetilde{\Delta \Gamma}_{\alpha\alpha} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta} = 0, \quad (59)$$

$$\bar{\mathcal{B}}(\Delta x)_{\beta\beta} = 0, \quad \bar{\mathcal{B}}(\Delta x)_{\beta\gamma} = 0, \quad \bar{\mathcal{B}}(\Delta x)_{\gamma\gamma} = 0, \quad (60)$$

$$(E_{\alpha\gamma} - \bar{\Xi}_{\alpha\gamma}) \circ \bar{\mathcal{B}}(\Delta x)_{\alpha\gamma} - \bar{\Xi}_{\alpha\gamma} \circ \widetilde{\Delta \Gamma}_{\alpha\gamma} = 0, \quad (61)$$

where $\widetilde{\Delta \Gamma} = \bar{P}^T \Delta \Gamma \bar{P}$, $\bar{\Xi}$ is defined in (15) for A and E represents the matrix consisting of all ones in \mathbb{S}^n . Suppose that the constraint nondegeneracy does not hold. Then, by (ii) of Lemma 1, we know that there exist $0 \neq (\Delta \xi, \Delta \Gamma)$ such that

$$\begin{cases} h'(\bar{x})^* \Delta \xi + \bar{\mathcal{B}}^* \widetilde{\Delta \Gamma} = 0, \\ \widetilde{\Delta \Gamma}_{\alpha\alpha} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\gamma} = 0. \end{cases}$$

Thus, it follows from (57)-(61) that $(\Delta x, \Delta \xi, \Delta \Gamma) = (0, \Delta \xi, \Delta \Gamma) \neq 0$ satisfy $\bar{\mathcal{U}}_0(\Delta x, \Delta \xi, \Delta \Gamma) = 0$, which contradicts with the nonsingularity of $\bar{\mathcal{U}}_0$. Thus, the constraint nondegeneracy holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$.

On the other hand, suppose that the W-SOC does not hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$. Since $\nabla_{xx}^2 L(\bar{x}, \bar{\xi}, \bar{\Gamma})$ is positive semidefinite, by Lemma 2, there exists $\Delta x \neq 0$ such that (56) holds. It then yields that $\bar{\mathcal{U}}_0(\Delta x, 0, 0) = 0$, which again contradicts with the nonsingularity of $\bar{\mathcal{U}}_0$. The proof of part (i) has been established. \square

Remark 7 In the context of linear semidefinite programming, a particular instance of the CSDP, Chan and Sun [2, Proposition 17] demonstrated that the nonsingularity of $\bar{\mathcal{U}}_0$ ($\bar{\mathcal{U}}_{\mathcal{I}}$) implies primal (dual) constraint nondegeneracy. In contrast, Proposition 5 in the current work provides a complete characterization of the nonsingularity of $\bar{\mathcal{U}}_0$ and $\bar{\mathcal{U}}_{\mathcal{I}}$ for more general CSDP.

Remark 8 For the general NLSDP, one can still use the first part of the proof to show that the nonsingularity of $\bar{\mathcal{U}}_0$ (or $\bar{\mathcal{U}}_{\mathcal{I}}$) implies the constraint nondegeneracy (or the W-SRCQ).

Finally, we can establish in the forthcoming proposition that, for the general NLSDP, the W-SRCQ is necessary to ensure the existence of nonsingular elements in the Clarke generalized Jacobian. Additionally, for the CSDP, both the W-SRCQ and the W-SOC are requisites.

Proposition 6 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of the NLSDP (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$.*

- (i) *If there is a nonsingular element $\bar{\mathcal{U}} \in \partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma})$, then the W-SRCQ holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$.*
- (ii) *For the CSDP, if there is a nonsingular element $\bar{\mathcal{U}} \in \partial_C F(\bar{x}, \bar{\xi}, \bar{\Gamma})$, then the W-SRCQ and W-SOC hold at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$.*

Proof Let $A = g(\bar{x}) + \bar{\Gamma}$ satisfy the eigenvalue decomposition (23) and $\bar{P} \in \mathcal{O}^n(A)$ with the index sets α, β and γ defined by (12). Recall the linear operator $\bar{\mathcal{B}} : \mathbb{X} \rightarrow \mathbb{S}^n$ defined by (28).

(i) Let $(\Delta \xi, \Delta \Gamma) \in \mathbb{R}^m \times \mathbb{S}^n$ be a solution to the following system:

$$\begin{cases} h'(\bar{x})^* \Delta \xi + \bar{\mathcal{B}}^* \widetilde{\Delta \Gamma} = 0, \\ \widetilde{\Delta \Gamma}_{\beta\beta} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\alpha} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\beta} = 0, \quad \widetilde{\Delta \Gamma}_{\alpha\gamma} = 0, \end{cases}$$

where $\widetilde{\Delta\Gamma} = \overline{P}^T \Delta\Gamma \overline{P}$. Then, we know from (35) and Proposition 2 that

$$\overline{U}(0, \Delta\xi, \Delta\Gamma) = \left[\begin{array}{c} h'(\overline{x})^* \Delta\xi + \overline{B}^* \widetilde{\Delta\Gamma} \\ 0 \\ \overline{P} \left[\begin{array}{ccc} \widetilde{\Delta\Gamma}_{\alpha\alpha} & \widetilde{\Delta\Gamma}_{\alpha\beta} & \overline{\Xi}_{\alpha\gamma} \circ \widetilde{\Delta\Gamma}_{\alpha\gamma} \\ \widetilde{\Delta\Gamma}_{\beta\alpha} & \mathcal{W}(\widetilde{\Delta\Gamma}_{\beta\beta}) & 0 \\ \overline{\Xi}_{\gamma\alpha} \circ \widetilde{\Delta\Gamma}_{\gamma\alpha} & 0 & 0 \end{array} \right] \overline{P}^T \end{array} \right] = 0, \quad (62)$$

where $\mathcal{W} \in \partial_C \Pi_{\mathbb{S}_+^{|\beta|}}(0)$. Since \overline{U} is nonsingular, we obtain that $(\Delta\xi, \Delta\Gamma) = 0$. Therefore, we know from (32) in Lemma 1 (i) that the W-SRCQ holds at $(\overline{x}, \overline{\xi}, \overline{\Gamma})$.

(ii) We only need to show that the W-SOC (Definition (8)) holds at $(\overline{x}, \overline{\xi}, \overline{\Gamma})$. Suppose on the contrary that the W-SOC does not hold. Since $\nabla_{xx}^2 L(\overline{x}, \overline{\xi}, \overline{\Gamma})$ is positive semidefinite, by Lemma 2, there exists $\Delta x \neq 0$ such that (56) holds. It, together with (62), implies that

$$\overline{U}(\Delta x, 0, 0) = \left[\begin{array}{c} \nabla_{xx}^2 L(\overline{x}, \overline{\xi}, \overline{\Gamma}) \Delta x \\ h'(\overline{x}) \Delta x \\ -\overline{P}(\overline{B}\Delta x)\overline{P}^T + \overline{P} \left[\begin{array}{ccc} (\overline{B}\Delta x)_{\alpha\alpha} & (\overline{B}\Delta x)_{\alpha\beta} & \overline{\Xi}_{\alpha\gamma} \circ (\overline{B}\Delta x)_{\alpha\gamma} \\ (\overline{B}\Delta x)_{\beta\alpha} & \mathcal{W}((\overline{B}\Delta x)_{\beta\beta}) & 0 \\ \overline{\Xi}_{\gamma\alpha} \circ (\overline{B}\Delta x)_{\gamma\alpha} & 0 & 0 \end{array} \right] \overline{P}^T \end{array} \right] = 0,$$

which contradicts with the nonsingularity of \overline{U} . The proof is then completed. \square

4 The primal-dual characterizations for convex quadratic semidefinite programming

In this section, our objective is to investigate the primal-dual characterizations of the W-SRCQ (Definition 7) and W-SOC (Definition 8) for the following convex quadratic semidefinite programming (QSDP):

$$\begin{aligned} \min \quad & \frac{1}{2} \langle x, \mathcal{Q}x \rangle + \langle c, x \rangle \\ \text{s.t.} \quad & \mathcal{H}x - p = 0, \\ & \mathcal{G}x - q \in \mathbb{S}_+^n, \end{aligned} \quad (63)$$

where $c \in \mathbb{X}$, $p \in \mathbb{R}^m$, $q \in \mathbb{S}^n$, the self-adjoint linear operator $\mathcal{Q} : \mathbb{X} \rightarrow \mathbb{X}$ is positive semi-definite, and $\mathcal{H} : \mathbb{X} \rightarrow \mathbb{R}^m$ and $\mathcal{G} : \mathbb{X} \rightarrow \mathbb{S}^n$ are two given linear operators. The dual problem (in the restricted Wolfe sense [39,21]) of the QSDP (63) takes the form:

$$\begin{aligned} \min_{\omega \in \mathbb{W}, \xi \in \mathbb{R}^m, \Gamma \in \mathbb{S}^n} \quad & \frac{1}{2} \langle \omega, \mathcal{Q}\omega \rangle + \langle p, \xi \rangle + \langle q, \Gamma \rangle \\ \text{s.t.} \quad & -\mathcal{Q}\omega - c - \mathcal{H}^* \xi - \mathcal{G}^* \Gamma = 0, \\ & \Gamma \in \mathbb{S}_-^n, \end{aligned} \quad (64)$$

where $\mathbb{W} \subseteq \mathbb{X}$ is a linear subspace containing $\text{Im}(\mathcal{Q})$. Thus, the KKT system of (64) is given by

$$\left\{ \begin{array}{l} \left[\begin{array}{c} \mathcal{Q}\omega \\ p \\ q \end{array} \right] + \left[\begin{array}{c} -\mathcal{Q} \\ -\mathcal{H} \\ -\mathcal{G} \end{array} \right] \tau + \left[\begin{array}{c} 0 \\ 0 \\ \mathcal{I} \end{array} \right] \Upsilon = 0, \\ \mathcal{Q}\omega + c + \mathcal{H}^* \xi + \mathcal{G}^* \Gamma = 0, \\ \Upsilon \in \mathcal{N}_{\mathbb{S}_-^n}(\Gamma), \\ \omega \in \mathbb{W}, \quad \xi \in \mathbb{R}^m, \quad \Gamma \in \mathbb{S}^n, \quad \tau \in \mathbb{X}, \quad \Upsilon \in \mathbb{S}^n. \end{array} \right. \quad (65)$$

Proposition 7 Consider the convex quadratic semidefinite programming (63) and its dual problem (64).

- (i) Suppose that $\mathbb{W} = \mathbb{X}$. Let $\bar{x} \in \mathbb{X}$ be a stationary point of the convex quadratic semidefinite programming (63) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$. The primal W-SOC holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$ if and only if the dual W-SRCQ holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma}, \bar{x}, \mathcal{G}\bar{x} - q)$.
- (ii) Suppose that $\mathbb{W} = \text{Im}(\mathcal{Q})$. Let $(\bar{\omega}, \bar{\xi}, \bar{\Gamma}) \in \mathbb{W} \times \mathbb{R}^m \times \mathbb{S}^n$ be a stationary point of the dual problem (64) with a Lagrange multiplier $(\bar{x}, \bar{\Upsilon}) \in M(\bar{\omega}, \bar{\xi}, \bar{\Gamma})$. The dual W-SOC holds at $(\bar{\omega}, \bar{\xi}, \bar{\Gamma}, \bar{x}, \bar{\Upsilon})$ if and only if the primal W-SRCQ holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$.

Proof We demonstrate (i) here, as (ii) follows from it with only minor modifications.

Let $A = g(\bar{x}) + \bar{\Gamma} = \mathcal{G}\bar{x} - q + \bar{\Gamma}$ satisfy the eigenvalue decomposition (23) with $\bar{P} \in \mathcal{O}^n(A)$, the index sets α, β and γ defined by (12), and $\bar{\mathcal{B}} : \mathbb{X} \rightarrow \mathbb{S}^n$ defined by (28). By the convexity and Lemma 2, we know that the primal W-SOC holds at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$ if and only if

$$\begin{cases} \mathcal{Q}\Delta x = 0, & \mathcal{H}\Delta x = 0, \\ \bar{\mathcal{B}}(\Delta x)_{\alpha\gamma} = 0, & \bar{\mathcal{B}}(\Delta x)_{\beta\beta} = 0, & \bar{\mathcal{B}}(\Delta x)_{\beta\gamma} = 0, & \bar{\mathcal{B}}(\Delta x)_{\gamma\gamma} = 0 \end{cases} \implies \Delta x = 0. \quad (66)$$

On the other hand, it is obvious that $(\bar{x}, \bar{\xi}, \bar{\Gamma}, \bar{x}, \mathcal{G}\bar{x} - q)$ is a KKT point of (64) by (65). By Definition 7, we know that the W-SRCQ of the dual problem (64) at $(\bar{x}, \bar{\xi}, \bar{\Gamma}, \bar{x}, \mathcal{G}\bar{x} - q)$ takes the form:

$$\begin{bmatrix} \mathcal{Q} & \mathcal{H}^* & \mathcal{G}^* \\ 0 & 0 & \mathcal{I} \end{bmatrix} \begin{bmatrix} \mathbb{X} \\ \mathbb{R}^m \\ \mathbb{S}^n \end{bmatrix} + \begin{bmatrix} \{0\} \\ \text{aff}(\mathcal{T}_{\mathbb{S}^n}(\bar{\Gamma}) \cap [\mathcal{G}\bar{x} - q]^\perp) \end{bmatrix} = \begin{bmatrix} \mathbb{X} \\ \mathbb{S}^n \end{bmatrix}. \quad (67)$$

Since $\text{aff}(\mathcal{T}_{\mathbb{S}^n}(\bar{\Gamma}) \cap [\mathcal{G}\bar{x} - q]^\perp) = \{\bar{P}B\bar{P}^T \in \mathbb{S}^n \mid B_{\alpha\beta} = 0, B_{\alpha\alpha} = 0\}$, we know that (67) is equivalent to

$$\text{Null} \left(\begin{bmatrix} \mathcal{Q} & 0 \\ \mathcal{H} & 0 \\ \bar{\mathcal{B}} & \mathcal{I} \end{bmatrix} \right) \cap \left[\{B \in \mathbb{S}^n \mid B_{\alpha\beta} = 0, B_{\alpha\alpha} = 0\}^\perp \right] = \{0\}. \quad (68)$$

Meanwhile, it is not difficult to observe that (68) and (66) are equivalent. Consequently, the primal W-SOC is satisfied at $(\bar{x}, \bar{\xi}, \bar{\Gamma})$ if and only if the dual W-SRCQ is satisfied at $(\bar{x}, \bar{\xi}, \bar{\Gamma}, \bar{x}, \mathcal{G}\bar{x} - q)$. This completes the proof. \square

Remark 9 By employing a similar argument, we are able to obtain the primal-dual characterizations of the S-SOSC (Definition 6) and the constraint nondegeneracy (Definition 4) for the QSDP (63). This extends the corresponding results obtained in Chan and Sun [2] for the linear SDP. Also, for the QSDP (63), a primal-dual characterizations of the SRCQ (Definition 3) and SOSC (Definition 5) is established by [14].

5 A semismooth Newton method with correction

In this section, based on the theoretical results obtained in previous sections, we introduce a new semismooth Newton method with a correction step, which enjoys the locally quadratically convergence rate even without the generalized Jacobian regularity.

Denote $\mathbb{Z} := \mathbb{X} \times \mathbb{R}^m \times \mathbb{S}^n$. Let $\delta > 0$ be some given constant. Define the correction mapping $\mathcal{P}_\delta : \mathbb{Z} \rightarrow \mathbb{Z}$ by

$$\mathcal{P}_\delta(Z) := \left(x, \xi, \Gamma - \sum_{i \in \theta} \lambda_i(A) P_i P_i^T \right), \quad Z = (x, \xi, \Gamma) \in \mathbb{Z}, \quad (69)$$

where $A := g(x) + \Gamma$ has the eigenvalue decomposition (23) with $P \in \mathcal{O}^n(A)$ and $\theta := \{i \mid |\lambda_i(A)| \leq \delta\}$.

Given $Z = (x, \xi, \Gamma) \in \mathbb{Z}$, define two elements \mathcal{U}_0 and $\mathcal{U}_{\mathcal{I}} \in \partial_B F(Z)$ at Z similarly as those in (36) and (37), i.e., for any $(\Delta x, \Delta \xi, \Delta \Gamma) \in \mathbb{Z}$,

$$\mathcal{U}_l(\Delta x, \Delta \xi, \Delta \Gamma) = \begin{bmatrix} \nabla_{xx}^2 L(x, \xi, \Gamma) \Delta x + h'(x)^* \Delta \xi + g'(x)^* \Delta \Gamma \\ h'(x) \Delta x \\ -g'(x) \Delta x + \mathcal{V}_l(g'(x) \Delta x + \Delta \Gamma) \end{bmatrix}, \quad l \in \{0, \mathcal{I}\} \quad (70)$$

where \mathcal{V}_0 and $\mathcal{V}_{\mathcal{I}} \in \partial_B \Pi_{\mathbb{S}_+^m}(A)$ is given by (16) and (17) in Proposition 2, respectively. The following proposition addresses the legitimacy of the correction mapping. It essentially shows that under weaker regularity conditions, when \widehat{Z} is sufficiently close to \overline{Z} , the corresponding \mathcal{U}_0 (or $\mathcal{U}_{\mathcal{I}}$) in the Bouligand generalized Jacobian of F at the corrected point $Z = \mathcal{P}_\delta(\widehat{Z})$ with properly chosen δ is nonsingular and its inverse is uniformly bounded. As one can see later, this will be crucial for the local convergence analysis of the proposed semismooth Newton algorithm.

Proposition 8 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of the NLSDP (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$. Let δ be any fixed constant in $(0, \min(\bar{\lambda}_\alpha, |\bar{\lambda}_\gamma|))$.*

(i) *If the W-SOC (Definition 8) and constraint nondegeneracy (Definition 4) hold at $\overline{Z} = (\bar{x}, \bar{\xi}, \bar{\Gamma}) \in \mathbb{Z}$, then there exists $\kappa_0 > 0$ such that for any $\widehat{Z} = (\widehat{x}, \widehat{\xi}, \widehat{\Gamma})$ sufficient close to \overline{Z} and $Z = \mathcal{P}_\delta(\widehat{Z})$, the corresponding $\mathcal{U}_0 \in \partial_B F(Z)$ defined by (70) with respect to Z is nonsingular and*

$$\kappa_0^{-1} \|\Delta Z\| \leq \|\mathcal{U}_0(\Delta Z)\| \quad \forall \Delta Z \in \mathbb{Z}. \quad (71)$$

(ii) *If the S-SOSC (Definition 6) and W-SRCQ (Definition 7) hold at $\overline{Z} = (\bar{x}, \bar{\xi}, \bar{\Gamma}) \in \mathbb{Z}$, then there exists $\kappa_1 > 0$ such that for any $\widehat{Z} = (\widehat{x}, \widehat{\xi}, \widehat{\Gamma})$ sufficient close to \overline{Z} and $Z = \mathcal{P}_\delta(\widehat{Z})$, the corresponding $\mathcal{U}_{\mathcal{I}} \in \partial_B F(Z)$ defined by (70) with respect to Z is nonsingular and*

$$\kappa_1^{-1} \|\Delta Z\| \leq \|\mathcal{U}_{\mathcal{I}}(\Delta Z)\| \quad \forall \Delta Z \in \mathbb{Z}. \quad (72)$$

Proof We focus on the result (i), since (ii) can be obtained similarly.

It suffices to show that the inequality (71) holds, as the nonsingularity of \mathcal{U}_0 is a direct implication of (71). Assume, on the contrary, there exist a sequence $\{\widehat{Z}^\nu\}$ converging to \overline{Z} and a nonzero sequence $\{\Delta Z^\nu = (\Delta x^\nu, \Delta \xi^\nu, \Delta \Gamma^\nu)\} \subseteq \mathbb{Z}$ such that $\mathcal{U}_0^\nu \in \partial_B F(Z^\nu)$ defined by (70) satisfy

$$\|\mathcal{U}_0^\nu(\Delta Z^\nu)\| < 1/\nu \|\Delta Z^\nu\|.$$

Without loss of generality, we may assume that $\|\Delta Z^\nu\| = 1$ for each ν and $\Delta Z^\nu \rightarrow \Delta Z^\infty \neq 0$ as $\nu \rightarrow \infty$. For each ν , let $\widehat{A}^\nu = g(\widehat{x}^\nu) + \widehat{\Gamma}^\nu$ enjoy the eigenvalue decomposition (23) with $\widehat{P}^\nu \in \mathcal{O}^n(\widehat{A}^\nu)$. For each ν , denote

$$\zeta^\nu := \{i \mid \widehat{\lambda}_i^\nu \geq \delta\}, \quad \theta^\nu := \{i \mid |\widehat{\lambda}_i^\nu| \leq \delta\} \quad \text{and} \quad \kappa^\nu := \{i \mid \widehat{\lambda}_i^\nu \leq -\delta\},$$

where $\widehat{\lambda}^\nu := \lambda(\widehat{A}^\nu)$ are the eigenvalues of \widehat{A}^ν . Since $\widehat{A}^\nu \rightarrow \overline{A} = g(\bar{x}) + \bar{\Gamma}$ as $\nu \rightarrow \infty$, we know from the continuity of eigenvalues [19] that $\widehat{\lambda}^\nu \rightarrow \lambda(A)$ as $\nu \rightarrow \infty$. Notice that $\delta \in (0, \min(\bar{\lambda}_\alpha, |\bar{\lambda}_\gamma|))$, for ν large enough,

$$\zeta^\nu \equiv \alpha, \quad \theta^\nu \equiv \beta \quad \text{and} \quad \kappa^\nu \equiv \gamma, \quad (73)$$

where α , β and γ are the index sets defined by (11) for \overline{A} . Since $Z^\nu = \mathcal{P}_\delta(\widehat{Z}^\nu)$, it holds that

$$\|Z^\nu - \widehat{Z}^\nu\| = \|\mathcal{P}_\delta(\widehat{Z}^\nu) - \widehat{Z}^\nu\| = \left\| \sum_{i \in \beta} \lambda_i(\widehat{A}^\nu) \widehat{P}_i^\nu (\widehat{P}_i^\nu)^T \right\| \leq \|\widehat{\lambda}_\beta^\nu\| \rightarrow 0, \quad \text{as } \nu \rightarrow \infty.$$

Thus, we know that $Z^\nu \rightarrow \bar{Z}$ as $\nu \rightarrow \infty$. It also holds that for each ν ,

$$A^\nu = g(x^\nu) + \Gamma^\nu = \hat{P}^\nu \begin{bmatrix} \hat{\Lambda}_{\alpha\alpha}^\nu & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{\Lambda}_{\gamma\gamma}^\nu \end{bmatrix} (\hat{P}^\nu)^T \quad (74)$$

with $\hat{\Lambda}^\nu = \text{Diag}(\hat{\lambda}^\nu)$, and $A^\nu \rightarrow \bar{A}$ as $\nu \rightarrow \infty$.

Let \bar{U}_0 be given by (36) with \bar{V}_0 defined by (16). By noting that $\|\mathcal{U}_0^\nu(\Delta Z^\nu)\| < 1/\nu$, $\|\Delta Z^\nu\| = 1$, f , h and g are twice continuously differentiable and $Z^\nu \rightarrow \bar{Z}$, we have

$$\begin{aligned} \|\bar{U}_0(\Delta Z^\infty)\| &\leq \|\bar{U}_0(\Delta Z^\infty) - \mathcal{U}_0^\nu(\Delta Z^\nu)\| + \|\mathcal{U}_0^\nu(\Delta Z^\nu)\| \\ &\leq \|\bar{U}_0(\Delta Z^\infty) - \bar{U}_0(\Delta Z^\nu)\| + \|\bar{U}_0(\Delta Z^\nu) - \mathcal{U}_0^\nu(\Delta Z^\nu)\| + 1/\nu \\ &\leq O(\|\Delta Z^\infty - \Delta Z^\nu\|) + \|\bar{V}_0(g'(\bar{x})\Delta x^\nu + \Delta\Gamma^\nu) - \mathcal{V}_0^\nu(g'(x^\nu)\Delta x^\nu + \Delta\Gamma^\nu)\| + 1/\nu, \end{aligned} \quad (75)$$

where for each ν , \mathcal{V}_0^ν is defined by (16) with respect to A^ν . By Proposition 1 and taking subsequence if necessary, we may further assume that $\hat{P}^\nu \rightarrow P^\infty \in \mathcal{O}^n(\bar{A})$. Then, (74) and Proposition 2 imply that $\lim_{\nu \rightarrow \infty} \mathcal{V}_0^\nu = \bar{V}_0$. Since \mathcal{V}_0^ν is uniformly bounded [22, Proposition 1], we thus have as $\nu \rightarrow \infty$,

$$\begin{aligned} &\|\mathcal{V}_0^\nu(g'(x^\nu)\Delta x^\nu + \Delta\Gamma^\nu) - \bar{V}_0(g'(\bar{x})\Delta x^\nu + \Delta\Gamma^\nu)\| \\ &\leq \|\mathcal{V}_0^\nu(g'(x^\nu)\Delta x^\nu + \Delta\Gamma^\nu) - \mathcal{V}_0^\nu(g'(\bar{x})\Delta x^\nu + \Delta\Gamma^\nu)\| \\ &\quad + \|\mathcal{V}_0^\nu(g'(\bar{x})\Delta x^\nu + \Delta\Gamma^\nu) - \bar{V}_0(g'(\bar{x})\Delta x^\nu + \Delta\Gamma^\nu)\| \\ &\leq O(\|(g'(x^\nu) - g'(\bar{x}))\Delta x^\nu\|) + o(\|g'(\bar{x})\Delta x^\nu + \Delta\Gamma^\nu\|) = o(1). \end{aligned} \quad (76)$$

Combining (75) and (76), we obtain $\|\bar{U}_0(\Delta Z^\infty)\| = 0$, which implies that \bar{U}_0 is singular. This contradicts Theorem 1 (i), since the W-SOC and constraint nondegeneracy are assumed. The proof is then completed. \square

Using Proposition 8, we can further establish a local error bound for the nonsmooth map F at the corrected points around \bar{Z} in the following proposition.

Proposition 9 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of the NLSDP (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$ and the constant $\delta \in (0, \min(\bar{\lambda}_\alpha, |\bar{\lambda}_\gamma|))$. Then, for any given $\varrho > 0$, there exists a neighborhood \mathcal{N} of \bar{Z} such that for any $\hat{Z} \in \mathcal{N}$, $Z = \mathcal{P}_\delta(\hat{Z})$, and $U \in \partial_C F(Z)$*

$$\|F(Z) - F(\bar{Z}) - \mathcal{U}(Z - \bar{Z})\| \leq \varrho \|Z - \bar{Z}\|. \quad (77)$$

Recall the definitions of constants κ_0 and κ_1 in Proposition 8. It further holds that

(i) *if the W-SOC and the constraint nondegeneracy hold at $\bar{Z} = (\bar{x}, \bar{\xi}, \bar{\Gamma}) \in \mathbb{Z}$, then there exists a neighborhood \mathcal{N} of \bar{Z} such that for any $\hat{Z} \in \mathcal{N}$ and $Z = \mathcal{P}_\delta(\hat{Z})$,*

$$\|F(Z)\| \geq \frac{1}{2\kappa_0} \|Z - \bar{Z}\|;$$

(ii) *if the S-SOSC and the W-SRCQ hold at $\bar{Z} = (\bar{x}, \bar{\xi}, \bar{\Gamma}) \in \mathbb{Z}$, then there exists a neighborhood \mathcal{N} of \bar{Z} such that for any $\hat{Z} \in \mathcal{N}$ and $Z = \mathcal{P}_\delta(\hat{Z})$,*

$$\|F(Z)\| \geq \frac{1}{2\kappa_1} \|Z - \bar{Z}\|.$$

Proof By the semismoothness of $II_{\mathbb{S}^n_+}(\cdot)$ on \mathbb{S}^n and the twice continuous differentiability of f , g and h , we know F is semismooth at \bar{Z} . This implies that for any given $\varrho > 0$, we can find a neighborhood \mathcal{N}_1 of \bar{Z} such that for any $Z \in \mathcal{N}_1$ and $U \in \partial_C F(Z)$

$$\|F(Z) - F(\bar{Z}) - \mathcal{U}(Z - \bar{Z})\| \leq \varrho \|Z - \bar{Z}\|. \quad (78)$$

Recall the eigenvalue decomposition of $\bar{A} = g(\bar{x}) + \bar{\Gamma}$ in (23) with index sets α, β and γ defined in (12). For any \hat{Z} with $Z = \mathcal{P}_\delta(\hat{Z})$, let $\hat{A} := g(\hat{x}) + \hat{\Gamma}$ satisfy the eigenvalue decomposition (11) with $\hat{P} \in \mathcal{O}^n(\hat{A})$ and $\hat{\Lambda} = \Lambda(\hat{A})$. Then, similar to (73) and (74), there exists a neighborhood \mathcal{N} of \bar{Z} such that for any $\hat{Z} \in \mathcal{N}$

$$A := g(x) + \Gamma = \hat{P} \begin{bmatrix} \hat{\Lambda}_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{\Lambda}_{\gamma\gamma} \end{bmatrix} (\hat{P})^T.$$

Thus,

$$\begin{aligned} \|Z - \bar{Z}\| &\leq \|Z - \hat{Z}\| + \|\hat{Z} - \bar{Z}\| \leq \|\hat{\Lambda}_{\beta\beta}\| + \|\hat{Z} - \bar{Z}\| \leq \|\lambda(\hat{A}) - \lambda(\bar{A})\| + \|\hat{Z} - \bar{Z}\| \\ &\leq \|g(\hat{x}) + \hat{\Gamma} - g(\bar{x}) - \bar{\Gamma}\| + \|\hat{Z} - \bar{Z}\| \leq (L_g + 2)\|\hat{Z} - \bar{Z}\|, \end{aligned} \quad (79)$$

where the last inequality is due to the Lipschitz continuity of g over \mathcal{N} . From (79), by shrinking \mathcal{N} if necessary, we can assert that $Z = \mathcal{P}_\delta(\hat{Z}) \in \mathcal{N}_1$ for any $\hat{Z} \in \mathcal{N}$. This, together with (78), implies the desired inequality (77).

Next, we focus on the rest results (i) and (ii). Here, only the proof of (i) is provided, as (ii) can be proved via similar arguments.

From (77) and Proposition 8 (i), we know that there exists a neighborhood \mathcal{N} of \bar{Z} such that for any $\hat{Z} \in \mathcal{N}$ and $Z = \mathcal{P}_\delta(\hat{Z})$,

$$\|F(Z) - \mathcal{U}_0(Z - \bar{Z})\| \leq \frac{1}{2\kappa_0} \|Z - \bar{Z}\| \quad \text{and} \quad \|\mathcal{U}_0\| \geq \frac{1}{\kappa_0},$$

where $\mathcal{U}_0 \in \partial_B F(Z)$ is given in (70). Thus,

$$\|F(Z)\| \geq \|\mathcal{U}_0(Z - \bar{Z})\| - \|F(Z) - \mathcal{U}_0(Z - \bar{Z})\| \geq \frac{1}{\kappa_0} \|Z - \bar{Z}\| - \frac{1}{2\kappa_0} \|Z - \bar{Z}\| = \frac{1}{2\kappa_0} \|Z - \bar{Z}\|.$$

The proof is completed. \square

Now we are ready to present our inexact semismooth Newton method with correction in Algorithm 1 for solving the nonsmooth system $F(Z) = 0$ with F given in (22). As one can observe, the algorithm is quite similar to classical inexact semismooth Newton method [11]. The main difference is that in each iteration, an extra correction step, i.e., step 6 in Algorithm 1, involving the correction mapping \mathcal{P}_δ is taken. This correction step, based on Proposition 8, ensures locally the nonsingularity of \mathcal{U}_0^k (or \mathcal{U}_T^k), and hence the well-definedness of the proposed algorithm. Meanwhile, it is worth noting that the additional computational costs brought by the correction step are negligible, as the eigendecomposition of $\hat{A}^{k+1} = g(\hat{x}^{k+1}) + \hat{\Gamma}^{k+1}$ in step 6 can be reused to obtain $\mathcal{U}^{k+1} \in \partial_B F(Z^{k+1})$ in step 3. Moreover, the inexact computation in step 4 indicates that our proposed algorithm has the potential for solving large-scale problems.

As promised, in the following theorem, we establish the well-definedness and local convergence results of Algorithm 1 without assuming the generalized Jacobian regularity.

Theorem 2 *Let $\bar{x} \in \mathbb{X}$ be a stationary point of the NLSDP (21) with a Lagrange multiplier $(\bar{\xi}, \bar{\Gamma}) \in M(\bar{x})$. Let $\delta \in (0, \min(\bar{\lambda}_\alpha, |\bar{\lambda}_\gamma|))$. Then,*

Algorithm 1 The semismooth Newton method with correction using \mathcal{U}_0 (or \mathcal{U}_T)

- 1: Initialize $\widehat{Z}^0 = (\widehat{x}^0, \widehat{\xi}^0, \widehat{T}^0)$, $\delta > 0$, $\eta \geq 0$, $\tau \in (0, 1]$ and $Z^0 = \mathcal{P}_\delta(\widehat{Z}^0)$.
- 2: **for** $k = 0, 1, \dots$ **do**
- 3: Get $\mathcal{U}^k = \mathcal{U}_0^k \in \partial_B F(Z^k)$ by (70) ($\mathcal{U}^k = \mathcal{U}_T^k \in \partial_B F(Z^k)$ by (70)).
- 4: Find an approximate solution d^k to the linear system $\mathcal{U}^k d + F(Z^k) = 0$ such that

$$\|\mathcal{U}^k d^k + F(Z^k)\| \leq \min(\eta, \|F(Z^k)\|^\tau) \|F(Z^k)\|.$$

- 5: $\widehat{Z}^{k+1} = Z^k + d^k$.
 - 6: $Z^{k+1} = \mathcal{P}_\delta(\widehat{Z}^{k+1})$.
 - 7: **end for**
-

- (i) if the W-SOC and the constraint nondegeneracy hold at $\bar{Z} = (\bar{x}, \bar{\xi}, \bar{T}) \in \mathbb{Z}$, then there exists $\bar{\eta} > 0$ such that for any \widehat{Z}^0 sufficiently close to \bar{Z} and $\eta \leq \bar{\eta}$, Algorithm 1 using \mathcal{U}_0 is well defined and the sequence $\{Z^k\}$ so generated converges to \bar{Z} superlinearly;
- (ii) if the S-SOSC and W-SRCQ hold at $\bar{Z} = (\bar{x}, \bar{\xi}, \bar{T}) \in \mathbb{Z}$, then there exists $\bar{\eta} > 0$ such that for any \widehat{Z}^0 sufficiently close to \bar{Z} and $\eta \leq \bar{\eta}$, Algorithm 1 using \mathcal{U}_T is well-defined and the sequence $\{Z^k\}$ so generated converges to \bar{Z} superlinearly.

Moreover, if f , g and h are twice Lipschitz continuously differentiable, i.e., they have Lipschitz continuous second order differentials, and $\tau = 1$, then the corresponding convergence rates obtained in (i) and (ii) are quadratic.

Proof Here, we focus on the case (i) where the W-SOC and the constraint nondegeneracy hold at \bar{Z} and \mathcal{U}_0 is employed. The proof for case (ii) where the S-SOSC and W-SRCQ hold at \bar{Z} can be obtained similarly.

By Proposition 8, Proposition 9 and (79), we can find a positive constant ς , a neighborhood $\mathbb{B}(\bar{Z}, \varsigma)$ of \bar{Z} and a positive constant κ_0 such that for any $\widehat{Z} \in \mathbb{B}(\bar{Z}, \varsigma)$ and $Z = \mathcal{P}_\delta(\widehat{Z})$,

$$\|F(Z) - F(\bar{Z}) - \mathcal{U}_0(Z - \bar{Z})\| \leq \frac{1}{4(L_g + 2)\kappa_0} \|Z - \bar{Z}\| \quad \text{and} \quad \|(\mathcal{U}_0)^{-1}\| \leq \kappa_0, \quad (80)$$

and

$$\|Z - \bar{Z}\| \leq (L_g + 2)\|\widehat{Z} - \bar{Z}\|, \quad (81)$$

where $\mathcal{U}_0 \in \partial_B F(Z)$ is given by (70) and $L_g > 0$ is the Lipschitz constant of g over $\mathbb{B}(\bar{Z}, \varsigma)$.

Now, let \widehat{Z}^0 be any point in $\mathbb{B}(\bar{Z}, \varsigma)$. Recall that $Z^0 = \mathcal{P}_\delta(\widehat{Z}^0)$. Then, (80) implies that $\mathcal{U}_0^0 \in \partial_B F(Z^0)$ is nonsingular, and therefore d^0 and \widehat{Z}^1 are well-defined. Since $F(\bar{Z}) = 0$, we further have

$$\begin{aligned} \|\widehat{Z}^1 - \bar{Z}\| &= \|Z^0 - \bar{Z} - (\mathcal{U}_0^0)^{-1}F(Z^0) + d^0 + (\mathcal{U}_0^0)^{-1}F(Z^0)\| \\ &\leq \| -(\mathcal{U}_0^0)^{-1}[F(Z^0) - F(\bar{Z}) - \mathcal{U}_0^0(Z^0 - \bar{Z}) - \mathcal{U}_0^0 d^0 - F(Z^0)] \| \\ &\leq \kappa_0 \|F(Z^0) - F(\bar{Z}) - \mathcal{U}_0^0(Z^0 - \bar{Z})\| + \eta \kappa_0 \|F(Z^0) - F(\bar{Z})\| \\ &\leq \left(\frac{1}{4(L_g + 2)} + \eta \kappa_0 L_F \right) \|Z^0 - \bar{Z}\|, \end{aligned} \quad (82)$$

where the last inequality follows from (80) and the Lipschitz continuity of F over $\mathbb{B}(\bar{Z}, \varsigma)$ with L_F being the corresponding Lipschitz constant. Meanwhile, we know from (81) that $\|Z^0 - \bar{Z}\| \leq$

$(L_g + 2)\|\widehat{Z}^0 - \bar{Z}\| \leq (L_g + 2)\varsigma$. Then, for any $0 \leq \eta \leq \bar{\eta} = 1/(4\kappa_0 L_F(L_g + 2))$, we see from (82) that $\|\widehat{Z}^1 - \bar{Z}\| \leq \|Z^0 - \bar{Z}\|/(2(L_g + 2)) \leq \varsigma$ and

$$\|Z^1 - \bar{Z}\| \leq (L_g + 2)\|\widehat{Z}^1 - \bar{Z}\| \leq \frac{1}{2}\|Z^0 - \bar{Z}\|. \quad (83)$$

In summary, the derivation above shows that if \widehat{Z}^0 belongs to $\mathbb{B}(\bar{Z}, \varsigma)$, then every \widehat{Z}^k and Z^k produced by Algorithm 1 satisfy

$$\widehat{Z}^k \in \mathbb{B}(\bar{Z}, \varsigma) \quad \text{and} \quad Z^k \in \mathbb{B}(\bar{Z}, (L_g + 2)\varsigma),$$

$\{Z^k\}$ converges to \bar{Z} due to the contraction inequality (83). Moreover, the semismoothness of F , (80) and (82) further imply that

$$\begin{aligned} \|\widehat{Z}^{k+1} - \bar{Z}\| &= \left\| -(\mathcal{U}_0^k)^{-1}[F(Z^k) - F(\bar{Z}) - \mathcal{U}_0^k(Z^k - \bar{Z}) - \mathcal{U}_0^k d^k - F(Z^k)] \right\| \\ &= o(\|Z^k - \bar{Z}\|) + O(\|Z^k - \bar{Z}\|^{1+\tau}). \end{aligned} \quad (84)$$

This, together with the fact that $\|Z^{k+1} - \bar{Z}\| \leq (L_g + 2)\|\widehat{Z}^{k+1} - \bar{Z}\|$, implies the superlinear convergence of $\{Z^k\}$.

If f, g, h are twice Lipschitz continuously differentiable, by the strong semismoothness of $\Pi_{\mathbb{S}_+^n}(\cdot)$ on \mathbb{S}^n and the Lipschitz continuous differentiability of $\nabla_x L$, g and h , we know F is strongly semismooth at \bar{Z} . The strong semismoothness of F and $\tau = 1$ allow us to replace (84) by the following:

$$\begin{aligned} \|\widehat{Z}^{k+1} - \bar{Z}\| &= \left\| -(\mathcal{U}_0^k)^{-1}[F(Z^k) - F(\bar{Z}) - \mathcal{U}_0^k(Z^k - \bar{Z}) - \mathcal{U}_0^k d^k - F(Z^k)] \right\| \\ &= O(\|Z^k - \bar{Z}\|^2). \end{aligned}$$

Then, the quadratic convergence of $\{Z^k\}$ follows directly. \square

Remark 10 Let $\{Z^k\}$ be the sequence generated by Algorithm 1. For a given tolerance $\varepsilon > 0$, Proposition 9 implies that $F(Z^k) < \varepsilon$ is a reasonable stopping criterion for Algorithm 1. In fact, the local Lipschitz continuity of F around \bar{Z} and Proposition 9 imply that $F(Z^k) = \Theta(\|Z^k - \bar{Z}\|)$.

Remark 11 (Choice of \widehat{Z}^0 and δ) It is not difficult to observe from Theorem 2 and its proof that the choice of \widehat{Z}^0 and δ is crucial to the performance of Algorithm 1. A larger δ may lead to a more restrictive choice of \widehat{Z}^0 in order to ensure the nonsingularity of \mathcal{U}^0 . Meanwhile, accurately determining the value of $\min(\bar{\lambda}_\alpha, |\lambda_\gamma|)$ for the general NLSDP is often challenging in practical settings, though obtaining a rough estimate may be feasible. In this study, our primary emphasis lies in the local convergence analysis of Algorithm 1. We defer the considerations regarding the efficient and robust implementation of the algorithm, including the selection of \widehat{Z}^0 and δ , to be explored in future research.

To conclude this section, we will illustrate with a simple example why the correction step is essential for semismooth Newton methods. Specifically, without the correction step, the classical semismooth Newton method may fail to converge or even fail to execute due to the singularity of the Jacobian. Moreover, the example illustrates that our algorithm can handle some differentiable but degenerate points, which is beyond the reach of the classical semismooth Newton method.

Example 3 Consider the following polyhedral problem, which can be viewed as a special instance of semidefinite programming:

$$\begin{aligned} \min & \frac{1}{2} \|x - (0, 1, 0)^T\|^2 \\ \text{s.t.} & \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ & x \in \mathbb{R}_+^3, \end{aligned}$$

where it is evident that $(0, 1, 0)^T$ is the unique solution with the multiplier set $M = \{((0, t)^T, (t, 0, t)^T) \mid t \leq 0\}$.

Verification reveals that the W-SRCQ and S-SOSC hold at the point

$$(\bar{x}, \bar{\xi}, \bar{\Gamma}) = ((0, 1, 0)^T, (0, 0)^T, (0, 0, 0)^T),$$

and the associated $\bar{\mathcal{U}}_{\mathcal{I}}$ is nonsingular. Therefore, one can employ Algorithm 1 using $\mathcal{U}_{\mathcal{I}}$ to solve this problem. However, for any point (x^0, ξ^0, Γ^0) such that

$$0 < \|(x^0, \xi^0, \Gamma^0) - (\bar{x}, \bar{\xi}, \bar{\Gamma})\| \leq 1, \quad x_1^0 + \Gamma_1^0 < 0, \quad x_2^0 + \Gamma_2^0 > 0 \text{ and } x_3^0 + \Gamma_3^0 < 0,$$

which is in proximity to $(\bar{x}, \bar{\xi}, \bar{\Gamma})$, the corresponding F is differentiable at (x^0, ξ^0, Γ^0) with a singular Jacobian $\nabla F(x^0, \xi^0, \Gamma^0)$. This suggests that the classical semismooth Newton method will encounter failure at such (x^0, ξ^0, Γ^0) . However, our proposed Algorithm 1 is capable of handling such scenarios through the correction. Taking

$$Z^0 = (x^0, \xi^0, \Gamma^0) = ((-\varepsilon, 1 + \varepsilon, 0)^T, (0, \varepsilon)^T, (0, 0, -\varepsilon)^T), \quad \varepsilon \in (0, 1/10),$$

as an example, Algorithm 1 (with $\delta = 1/5$) will first correct it to

$$\hat{Z}^0 = (\hat{x}^0, \hat{\xi}^0, \hat{\Gamma}^0) = ((-\varepsilon, 1 + \varepsilon, 0)^T, (0, \varepsilon)^T, (\varepsilon, 0, 0)^T).$$

Then the Newton step based on $\mathcal{U}_{\mathcal{I}}$ reads

$$(x^1, \xi^1, \Gamma^1) = \hat{Z}^0 - (\mathcal{U}_{\mathcal{I}}^0)^{-1} F(\hat{Z}^0) = ((0, 1, 0)^T, (0, 0)^T, (0, 0, 0)^T) = (\bar{x}, \bar{\xi}, \bar{\Gamma}),$$

which indicates that our algorithm converges to a KKT solution in only one iteration.

Remark 12 In the recent study [9], the authors introduced a Newton-type algorithm to solve the subgradient inclusion problem, achieving a superlinear convergence rate assuming metric regularity, but without requiring strong metric regularity. Even for tilt-stable local minimizers, their results improve on those of Mordukhovich and Sarabi [25]. However, in the context of optimization problems, the equivalence between the metric regularity and the strong metric regularity at local optimal solutions has been established for the nonlinear programming in [8], and for the nonlinear second-order cone programming and NLSQP in the recent preprints [4] and [3], respectively. Meanwhile, the conditions we impose for achieving a superlinear convergence rate, such as W-SRCQ and S-SOSC, do not even imply metric subregularity (see Example 5).

Remark 13 It is worth noting that the results obtained in Section 3 and Section 5 can be directly extended to the following generalized equation (GE):

$$0 \in \phi(x) + \nabla g(x)^T \mathcal{N}_D(g(x)), \quad (85)$$

where $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is twice continuously differentiable and D is the Cartesian product of a finite number of positive semidefinite cones and

zeros. Let $l(x, y) := \phi(x) + \nabla g(x)^T y$ for $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. Define the corresponding nonsmooth mapping $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ as

$$F(x, y) = \begin{bmatrix} l(x, y) \\ -g(x) + \Pi_D(g(x) + y) \end{bmatrix}.$$

For any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ (not necessarily a solution to (85)), define

$$\tilde{\mathcal{C}}(x, y) := \mathcal{T}_D(\Pi_D(g(x) + y)) \cap (g(x) + y - \Pi_D(g(x) + y))^\perp$$

and

$$\tilde{\sigma}(x, y) := \sigma(g(x) + y - \Pi_D(g(x) + y), \mathcal{T}_D^2(\Pi_D(g(x) + y), \nabla g(x)^T \Delta x)),$$

where $\Pi_D(\cdot)$ is the metric projection operator onto D , $\mathcal{T}_D(z)$ is the tangent cone to D at z defined in (6), and $\mathcal{T}_D^2(z, d)$ is the second order tangent set to D at z in the direction d defined in (9). The constraint nondegeneracy for the GE (85) at (x, y) is then defined as

$$\nabla g(x)^T \mathbb{R}^n + \text{lin}(\tilde{\mathcal{C}}(x, y)) = \mathbb{Y}. \quad (86)$$

Analogously, the W-SRCQ for the GE (85) at (x, y) is defined by

$$\nabla g(x)^T \mathbb{R}^n + \text{aff}(\tilde{\mathcal{C}}(x, y)) = \mathbb{Y}.$$

We say the W-SOC for the GE (85) holds at (x, y) if and only if

$$\langle \Delta x, \nabla_x l(x, y) \Delta x \rangle - \tilde{\sigma}(x, y) \neq 0 \quad \forall \Delta x \in \text{appl}(x, y) \setminus \{0\}, \quad (87)$$

where

$$\text{appl}(x, y) := \left\{ \Delta x \mid \nabla g(x)^T \Delta x \in \text{lin}(\tilde{\mathcal{C}}(x, y)) \right\}.$$

Similarly, the S-SOC for the GE (85) at (x, y) is defined correspondingly by

$$\langle \Delta x, \nabla_x l(x, y) \Delta x \rangle - \tilde{\sigma}(x, y) \neq 0 \quad \forall \Delta x \in \text{app}(x, y) \setminus \{0\}, \quad (88)$$

where

$$\text{app}(x, y) := \left\{ \Delta x \mid \nabla g(x)^T \Delta x \in \text{aff}(\tilde{\mathcal{C}}(x, y)) \right\}.$$

Thus, the obtained results (e.g., Theorem 1 in Section 3 and Theorem 2 in Section 5) can be shown by using exactly the same manners. The noteworthy point is that, based on the proof of Theorem 1, the S-SOC of the generalized equation being “ $\neq 0$ ” instead of “ > 0 ” does not affect the validity of the conclusion.

In the recent work [12], a generalized semismooth* Newton-like method for solving generalized equations is proposed. When some appropriate regularity condition (see [12, Theorem 4.4] for details) is assumed, the method ensures local superlinear convergence. The proposed regularity conditions are shown to be weaker than the metric regularity. For (85) with D being a polyhedron, the regularity conditions hold under [12, Assumptions 1 and 2], as elaborated in [12, Section 5]. Moreover, as stated in [12, Remark 5.2], for a solution (\bar{x}, \bar{y}) to (85), [12, Assumption 1] is equivalent to the constraint nondegeneracy for the GE defined by (86). Additionally, when $D = \mathbb{R}^n$, it is stated in [12, Remark 5.10] that [12, Assumption 2] is equivalent to the S-SOC (88). However, it remains unclear under what circumstances the regularity conditions proposed in [12, Theorem 4.4] will hold for non-polyhedral sets, such as positive semidefinite cones. Interestingly, in [12, Example 5.13], their algorithm produces the same sequence as the classical semismooth Newton method using $\mathcal{U}_{\mathcal{I}}$ (i.e., Algorithm 1 without correction). For this particular case, the correction map $\mathcal{P}_\delta(\cdot)$, defined by (69), reduces to the identity map.

6 Numerical results

In this section, we present a series of examples to empirically validate our theoretical findings and demonstrate the fast local convergence of our proposed algorithm. In our numerical experiments, we consider a range of problems, including both convex and nonconvex nonlinear SDPs. It's worth noting that the generalized Jacobian regularity condition does not hold for all the test instances, thereby rendering classic convergence results for semismooth Newton methods inapplicable.

As preliminary numerical results, we solve the Newton linear system involved in each iteration of Algorithm 1 exactly, i.e., the parameter η is set to be zero. Following Remark 10, we assess the accuracy of an approximate solution Z to the NLSDP (21) by measuring the residual $\|F(Z)\|$. For a given stopping tolerance $\epsilon > 0$, our algorithm terminates when $\|F(Z)\| < \epsilon$. Throughout our experiments, we fix $\epsilon = 10^{-10}$. All computational results are generated using MATLAB (version 9.10) running on a Mac mini (Apple M1, 16 G RAM).

While our focus has been on investigating the local convergence properties of our newly proposed semismooth Newton method, our numerical tests indicate that the algorithm is quite robust to the selection of the starting point. To facilitate reproducibility, we adhere to the following procedures for generating the starting point:

- (i) Fix the random seed and select the perturbation constant C_{perturb} . In our tests, we fix the random seed via `rng(2, 'twister')` and set $C_{\text{perturb}} = 10$ for examples 4, 5 and 6, and set $C_{\text{perturb}} = 1$ for example 7.
- (ii) Generate random noises $(\Delta x, \Delta \xi, \Delta \Gamma)$ with each entry following a standard normal distribution $N(0, 1)$, e.g.,

$$\Delta \Gamma = \text{randn}(n), \quad \Delta \Gamma = 0.5 * (\Delta \Gamma + \Delta \Gamma').$$

- (iii) Set the starting point \widehat{Z}^0 as

$$\widehat{Z}^0 = \bar{Z} + C_{\text{perturb}} \times \left(\frac{\Delta x}{\|\Delta x\|}, \frac{\Delta \xi}{\|\Delta \xi\|}, \frac{\Delta \Gamma}{\|\Delta \Gamma\|} \right),$$

where \bar{Z} is the known optimal solution to (21).

It is important to note that the perturbation constant C_{perturb} employed is not insignificant. Remarkably, varying the random seeds and the perturbation constant C_{perturb} results in comparable quadratic convergence behavior.

Example 4 Consider the following convex quadratic SDP problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|x_{11} - 1\|^2 + \frac{1}{2} \|x_{22} - x_{12} - x_{21}\|^2 \\ \text{s.t.} \quad & \langle E, X \rangle \leq 1, \\ & X \in \mathbb{S}_+^2, \end{aligned} \tag{89}$$

where $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \in \mathbb{S}^2$. It is not difficult to observe that $\bar{X} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is an optimal solution and $(\bar{\xi}, \bar{T}) = (0, 0)$ is the unique multiplier. For such $\bar{Z} = (\bar{X}, \bar{\xi}, \bar{T})$, one can check that the W-SOC and the constraint nondegeneracy hold at \bar{Z} , but the S-SOSC fails. Then, Theorem 1 implies the nonsingularity of \bar{U}_0 , which can also be verified directly by noting that the linear

system

$$\bar{\mathcal{U}}_0(\Delta Z) = \begin{bmatrix} \begin{bmatrix} \Delta x_{11} & -\Delta x_{22} + \Delta x_{12} + \Delta x_{21} \\ -\Delta x_{22} + \Delta x_{12} + \Delta x_{21} & \Delta x_{22} - \Delta x_{12} - \Delta x_{21} \end{bmatrix} + \Delta \xi E + \Delta \Gamma \\ \langle E, \Delta X \rangle \\ \begin{bmatrix} \Delta \Gamma_{11} & \Delta \Gamma_{12} \\ \Delta \Gamma_{21} & -\Delta x_{22} \end{bmatrix} \end{bmatrix} = 0$$

with $\Delta Z = (\Delta X, \Delta \xi, \Delta \Gamma) = \left(\begin{bmatrix} \Delta x_{11} & \Delta x_{12} \\ \Delta x_{21} & \Delta x_{22} \end{bmatrix}, \Delta \xi, \begin{bmatrix} \Delta \Gamma_{11} & \Delta \Gamma_{12} \\ \Delta \Gamma_{21} & \Delta \Gamma_{22} \end{bmatrix} \right) \in \mathbb{S}^2 \times \mathbb{R} \times \mathbb{S}^2$ has only trivial zero solution $\Delta Z = 0$. Thus, Algorithm 1 using \mathcal{U}_0 can be applied to solve (89). In the test, we set the parameter $\delta = 0.5$, which lies between 0 and the smallest nonzero magnitude eigenvalue of \bar{A} . We report the detailed performance of our algorithm in Figure 2 and Table 1. In the table,

Iteration k	$\ F(Z^k)\ $	$\ Z^k - \bar{Z}\ $	$\sigma_{\min}(\mathcal{U}^k)$
0	2.51e+01	1.43e+01	4.03e-01
1	7.50e-01	9.36e-01	3.07e-01
2	4.19e-02	2.10e-02	3.04e-01
3	3.33e-05	2.88e-05	3.04e-01
4	2.06e-14	1.80e-14	3.04e-01

Table 1: Convergence details of Algorithm 1 for solving (89).

we report residuals $\|F(Z^k)\|$ and $\|Z^k - \bar{Z}\|$, and the minimum singular value of \mathcal{U}^k , denoted by $\sigma_{\min}(\mathcal{U}^k)$. The detailed results in Table 1 demonstrate the quadratic convergence of Algorithm 1, the locally uniform boundedness of $(\mathcal{U}^k)^{-1}$, and the relation $F(Z^k) = \Theta(\|Z^k - \bar{Z}\|)$ asserted in Remark 10.

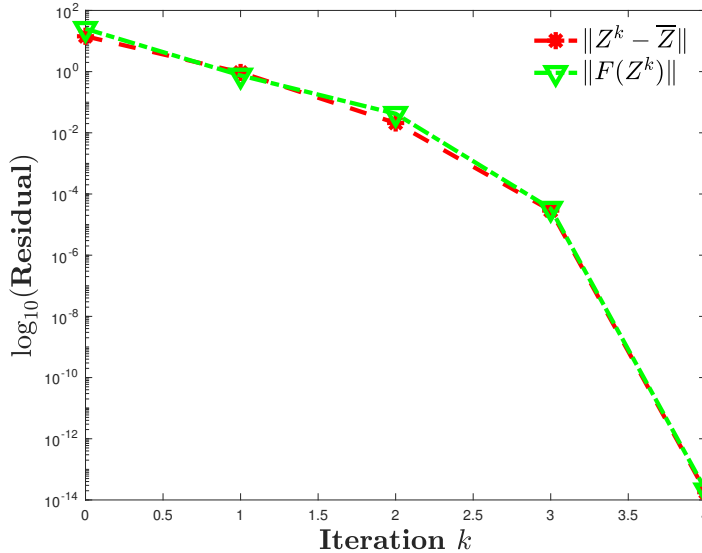


Fig. 2: Quadratic convergence of Algorithm 1 for solving (89).

The following example illustrates the efficiency of our algorithm in solving nonlinear SDPs even in the absence of calmness, as defined in [33, Definition 9(30)].

Example 5 (Example 3 in [7]; see also [42]) Let $B = \begin{bmatrix} 3/2 & -2 \\ -2 & 3 \end{bmatrix}$ and $b = B^{-1/2} \begin{bmatrix} 5/2 \\ -1 \end{bmatrix}$. Consider the following convex quadratic SDP problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|x + b\|^2 + t \\ \text{s.t.} \quad & \mathcal{A}^*x + tE + I + \Delta \in \mathbb{S}_+^2, \\ & t \geq 0, \end{aligned} \tag{90}$$

and its dual

$$\begin{aligned} \max \quad & -\frac{1}{2} \|\mathcal{A}Y - b\|^2 - \langle I + \Delta, Y \rangle \\ \text{s.t.} \quad & \langle E, Y \rangle \leq 1, \\ & Y \in \mathbb{S}_+^2, \end{aligned} \tag{91}$$

where $\mathcal{A}Y = B^{1/2}\text{Diag}(Y)$ for all $Y \in \mathbb{S}^2$, $\mathcal{A}^*x = \text{Diag}(B^{1/2}x)$ for all $x \in \mathbb{R}^2$ and $\Delta = \text{Diag}(-\varepsilon, \varepsilon)$ with $\varepsilon \geq 0$. As shown in [7], the solution map of the KKT system associated with problems (90) and (91), denoted by S_{KKT} in [7], is not calm with respect to the perturbation ε . When $\varepsilon = 0$, it is not difficult to see that $\bar{Y} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is the unique optimal solution to (91) with the unique multiplier $(0, 0) \in \mathbb{R} \times \mathbb{S}^2$. Then, one can check that the W-SOC and constraint nondegeneracy hold at $\bar{Z} = (\bar{Y}, 0, 0)$, but the S-SOSC fails. From Theorem 1 and Proposition 5, we know that $\bar{\mathcal{U}}_0$ is nonsingular but $\bar{\mathcal{U}}_{\mathcal{I}}$ is singular. Thus, we can apply Algorithm 1 with \mathcal{U}_0 to solve (91). In the test, we set parameter $\delta = 0.5$. The convergence details of our algorithm are reported in Table 2 and Figure 3. Again, the quadratic convergence of our algorithm is confirmed.

Iteration k	$\ F(Z^k)\ $	$\ Z^k - \bar{Z}\ $	$\sigma_{\min}(\mathcal{U}^k)$
0	1.10e+01	1.43e+01	4.16e-02
1	6.26e-01	1.22e+00	1.48e-01
2	8.23e-01	5.12e-01	1.38e-01
3	1.23e-01	2.33e-01	3.70e-01
4	1.46e-02	3.02e-02	2.85e-01
5	2.13e-04	4.51e-04	2.76e-01
6	5.08e-08	1.08e-07	2.76e-01
7	2.89e-15	6.05e-15	2.76e-01

Table 2: Convergence details of Algorithm 1 for solving (91).

In the next two examples, we evaluate the performance of our algorithm for solving medium size nonlinear SDPs.

Example 6 Consider the following convex quadratic SDP problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|X_{11} - I\|^2 + \frac{1}{2} \|X_{12}\|^2 + \frac{1}{2} \|X_{21}\|^2 \\ \text{s.t.} \quad & X \in \mathbb{S}_+^n, \end{aligned} \tag{92}$$

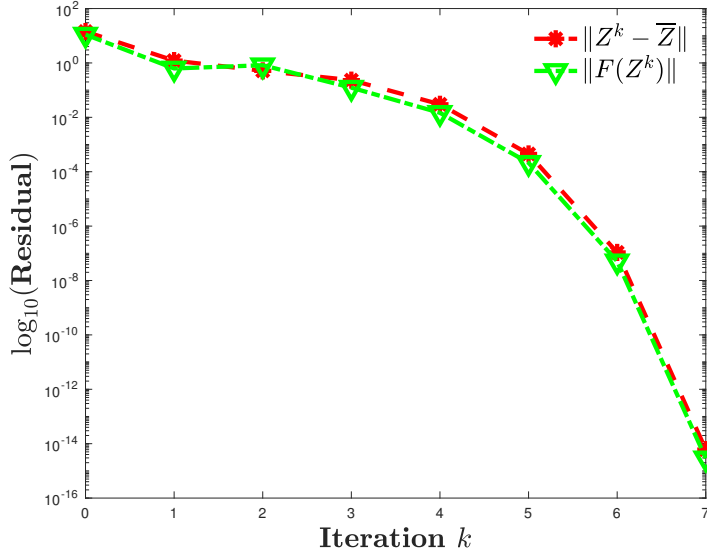


Fig. 3: Quadratic convergence of Algorithm 1 for solving (91).

where $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ is the block representation with $X_{11} \in \mathbb{R}^{l_1 \times l_1}$ and $X_{22} \in \mathbb{R}^{l_2 \times l_2}$. We know that $\bar{X} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ is an optimal solution and $\bar{\Gamma} = 0$ is the unique multiplier. One can check that the W-SOC and the constraint nondegeneracy holds at $\bar{Z} = (\bar{X}, \bar{\Gamma})$, but the SOSC fails since \bar{X} is not isolated. Then, we have from Theorem 1 that \bar{U}_0 is nonsingular and Algorithm 1 with \mathcal{U}_0 can be applied to solve (92). Here, the test instance is constructed by setting $l_1 = 60$, $l_2 = 40$, and the parameter δ is set to 0.5. Table Figure 4 shows our algorithm converges rapidly.

Iteration k	$\ F(Z^k)\ $	$\ Z^k - \bar{Z}\ $	$\sigma_{\min}(\mathcal{U}^k)$
0	8.39e-01	7.06e-01	6.02e-01
1	3.54e-02	3.54e-02	6.18e-01
2	1.85e-14	1.75e-14	6.18e-01

Table 3: Convergence details of Algorithm 1 for solving (92).

Example 7 Recall problem (18) in Example 1. Note that problem (18) is nonconvex, and $\bar{X} = 0$ is a local optimal solution with unbounded Lagrange multiplier set $M(\bar{X})$. It can be verified that the S-SOSC holds at $\bar{Z} = (\bar{X}, \bar{\xi}_{12}, \bar{\xi}_{22}, \bar{\Gamma}) = (0, 0, 0, 0) \in \mathbb{S}^n \times \mathbb{R}^{l_1 \times l_2} \times \mathbb{R}^{l_2 \times l_2} \times \mathbb{S}^n$ and the W-SRCQ holds at \bar{X} with respect to $(\bar{\xi}_{12}, \bar{\xi}_{22}, \bar{\Gamma})$. Then, it is known from Theorem 1 that $\bar{U}_{\mathcal{I}}$ is nonsingular. Thus, Algorithm 1 with $\mathcal{U}_{\mathcal{I}}$ can be applied to solve (18). In our test, we set the parameter $\delta = 0.5$, and $l_1 = 60$ and $l_2 = 40$. The detailed performance of our algorithm is reported in Table 4. As one can observe, our algorithm converges in one iteration from a randomly generated starting point with initial distance $\|Z^0 - \bar{Z}\| = 1.73$.

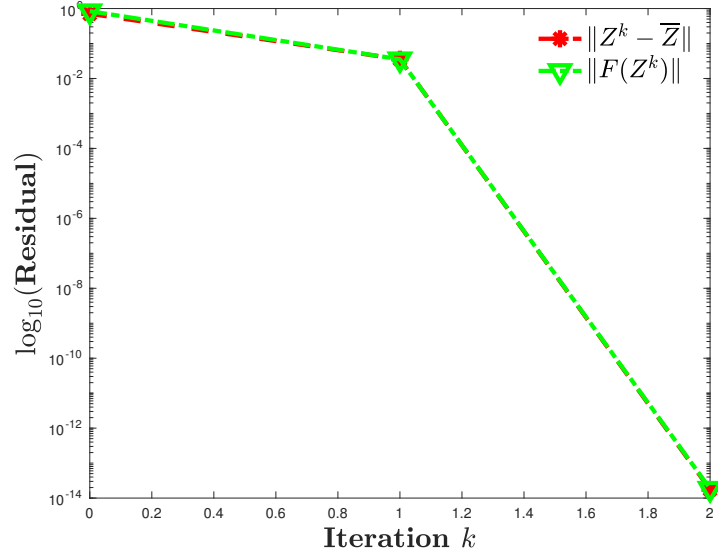


Fig. 4: Quadratic convergence of Algorithm 1 for solving (92).

Iteration k	$\ F(Z^k)\ $	$\ Z^k - \bar{Z}\ $	$\sigma_{\min}(\mathcal{U}^k)$
0	1.85e+00	1.73e+00	3.27e-01
1	2.98e-15	4.18e-15	3.27e-01

Table 4: Convergence details of Algorithm 1 for solving (18).

7 Conclusions

In this paper, we explore the weak second-order condition (W-SOC) and the weak strict Robinson constraint qualification (W-SRCQ) within the framework of general optimization problems (1). We establish the sufficient and necessary conditions for the existence of a nonsingular element in the generalized Jacobians of the KKT residual map F at a given KKT point. Moreover, we uncover a profound primal-dual connection between the W-SOC and the W-SRCQ for the convex quadratic SDP. Drawing upon these theoretical advancements, we propose a novel semismooth Newton method with a correction step, ensuring the local convergence and yielding a (quadratic) superlinear convergence rate without relying on the classic generalized Jacobian regularity. There are still many issues that deserve further exploration. For example, addressing the globalization of the proposed algorithm and developing highly efficient implementations present significant challenges. Additionally, extending our theoretical findings to broader sets is highly desirable. Our results could potentially be extended to variational inequalities with semidefinite constraints and optimization problems with matrix cone constraints, as defined in [6]. Furthermore, exploring the generalization of our results to more general sets, such as \mathcal{C}^2 -cone reducible sets, would be another challenging topic.

Acknowledgements We would like to thank the referees as well as the editors for their constructive comments that have helped to improve the quality of the paper.

References

1. Bonnans, J.F., Shapiro, A.: *Perturbation Analysis of Optimization Problems*. Springer Verlag (2000)
2. Chan, Z.X., Sun, D.: Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming. *SIAM Journal on optimization* **19**, 370–396 (2008)
3. Chen, L., Chen, R., Sun, D., Zhang, L.: Equivalent characterizations of the Aubin property for nonlinear semidefinite programming. arXiv preprint arXiv:2408.08232 (2024)
4. Chen, L., Chen, R., Sun, D., Zhu, J.: Aubin property and strong regularity are equivalent for nonlinear second-order cone programming. arXiv preprint arXiv:2406.13798 (2024)
5. Clarke, F.H.: *Optimization and Nonsmooth Analysis*. John Wiley & Sons, Inc., New York (1983)
6. Ding, C., Sun, D., Toh, K.C.: An introduction to a class of matrix cone programming. *Mathematical Programming* **144**, 141–179 (2014)
7. Ding, C., Sun, D., Zhang, L.: Characterization of the robust isolated calmness for a class of conic programming problems. *SIAM Journal on Optimization* **27**, 67–90 (2017)
8. Dontchev, A.L., Rockafellar, R.T.: Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM Journal on Optimization* **6**, 1087–1105 (1996)
9. Duy Khanh, P., Mordukhovich, B., Phat, V.T.: A generalized Newton method for subgradient systems. *Mathematics of Operations Research* **48**, 1811–1845 (2023)
10. Eberhard, A., Hill, R., Ralph, D., Glover, B.M., Qi, L., Sun, D.: A survey of some nonsmooth equations and smoothing Newton methods. *Progress in optimization: contributions from Australasia* pp. 121–146 (1999)
11. Facchinei, F., Pang, J.S.: *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer (2003)
12. Gfrerer, H., Outrata, J.V.: On a semismooth* Newton method for solving generalized equations. *SIAM Journal on Optimization* **31**, 489–517 (2021)
13. Gowda, M.S.: Inverse and implicit function theorems for H-differentiable and semismooth functions. *Optimization Methods and Software* **19**, 443–461 (2004)
14. Han, D., Sun, D., Zhang, L.: Linear rate convergence of the alternating direction method of multipliers for convex composite programming. *Mathematics of Operations Research* **43**, 622–637 (2018)
15. Hintermüller, M.: *Semismooth Newton methods and applications*. Department of Mathematics, Humboldt-University of Berlin (2010)
16. Izmailov, A.F., Solodov, M.V.: Karush-Kuhn-Tucker systems: regularity conditions, error bounds and a class of Newton-type methods. *Mathematical Programming* **95**, 631–650 (2003)
17. Klatte, D., Kummer, B.: *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, vol. 60. Springer Science & Business Media (2002)
18. Kummer, B.: Newton’s method based on generalized derivatives for nonsmooth functions: convergence analysis. In: *Advances in Optimization: Proceedings of the 6th French-German Colloquium on Optimization Held at Lambrecht, FRG, June 2–8, 1991*, pp. 171–194. Springer (1992)
19. Lancaster, P.: On eigenvalues of matrices dependent on a parameter. *Numerische Mathematik* **6**, 377–387 (1964)
20. Li, X., Sun, D., Toh, K.C.: A highly efficient semismooth Newton augmented Lagrangian method for solving LASSO problems. *SIAM Journal on Optimization* **28**, 433–458 (2018)
21. Li, X., Sun, D., Toh, K.C.: QSDPNAL: A two-phase augmented Lagrangian method for convex quadratic semidefinite programming. *Mathematical Programming Computation* **10**, 703–743 (2018)
22. Meng, F., Sun, D., Zhao, G.: Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization. *Mathematical programming* **104**, 561–581 (2005)
23. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. *SIAM Journal on Control and Optimization* **15**, 959–972 (1977)
24. Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation II: Applications*, vol. 331. Springer (2006)
25. Mordukhovich, B.S., Sarabi, M.E.: Generalized Newton algorithms for tilt-stable minimizers in nonsmooth optimization. *SIAM Journal on Optimization* **31**, 1184–1214 (2021)
26. Pang, J.S., Sun, D., Sun, J.: Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems. *Mathematics of Operations Research* **28**, 39–63 (2003)
27. Qi, H., Sun, D.: A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM journal on matrix analysis and applications* **28**, 360–385 (2006)
28. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. *Mathematics of operations research* **18**, 227–244 (1993)
29. Qi, L., Sun, J.: A nonsmooth version of Newton’s method. *Mathematical programming* **58**, 353–367 (1993)
30. Rademacher, H.: Über partielle und totale differenzierbarkeit von funktionen mehrerer variablen und über die transformation der doppelintegrale. *Mathematische Annalen* **79**, 340–359 (1919)
31. Robinson, S.M.: Strongly regular generalized equations. *Mathematics of Operations Research* **5**, 43–62 (1980)
32. Rockafellar, R.T.: *Convex Analysis*, vol. 11. Princeton university press (1970)

33. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis, vol. 317. Springer Science & Business Media (1997)
34. Sun, D.: The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. *Mathematics of Operations Research* **31**, 761–776 (2006)
35. Sun, D., Sun, J.: Semismooth matrix-valued functions. *Mathematics of Operations Research* **27**, 150–169 (2002)
36. Sun, D., Sun, J.: Strong semismoothness of eigenvalues of symmetric matrices and its application to inverse eigenvalue problems. *SIAM Journal on Numerical Analysis* **40**, 2352–2367 (2002)
37. Tibshirani, R.: Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **58**, 267–288 (1996)
38. Ulbrich, M.: Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces. SIAM (2011)
39. Wolfe, P.: A duality theorem for non-linear programming. *Quarterly of Applied Mathematics* **19**, 239–244 (1961)
40. Yang, L., Sun, D., Toh, K.C.: SDPNAL+: a majorized semismooth Newton-CG augmented Lagrangian method for semidefinite programming with nonnegative constraints. *Mathematical Programming Computation* **7**, 331–366 (2015)
41. Zhao, X.Y., Sun, D., Toh, K.C.: A Newton-CG augmented Lagrangian method for semidefinite programming. *SIAM Journal on Optimization* **20**, 1737–1765 (2010)
42. Zhou, Z., So, A.M.C.: A unified approach to error bounds for structured convex optimization problems. *Mathematical Programming* **165**, 689–728 (2017)