Local Duality of Nonlinear Semidefinite Programming

Houduo Qi*

April 24, 2007; Revised November 1, 2007 and January 6, 2008

Abstract

In a recent paper [8], Chan and Sun reported for semidefinite programming (SDP) that the primal/dual constraint nondegeneracy is equivalent to the dual/primal strong second order sufficient condition (SSOSC). This result is responsible for a number of important results in stability analysis of SDP. In this paper, we study duality of this type in nonlinear semidefinite programming (NSDP). We introduce the dual SSOSC at a KKT triple of NSDP and study its various characterizations and relationships to the primal nondegeneracy. Although the dual SSOSC is nothing but the SSOSC for the Wolfe dual of the NSDP, it suggests new information for the primal NSDP. For example, it ensures that the inverse of the Hessian of the Lagrangian function exists at the KKT triple and the inverse is positive definite on some normal space. It also ensures the primal nondegeneracy. Some of our results generalize the corresponding classical duality results in nonlinear programming studied by Fujiwara, Han and Managsarian [13]. For the convex quadratic SDP (QSDP), we have complete characterizations for the primal and dual SSOSC. And our results reveal that the nearest correlation matrix problem satisfies not only the primal and dual SSOSC but also the primal and dual nondegeneracy at its solution, suggesting that it is a well-conditioned QSDP.

Key words: Nonlinear semidefinite programming, strong second-order sufficient condition, constraint nondegeneracy, nonsingularity.

AMS subject classification: 90C31, 65K10

1 Introduction

There have recently been significant advances in the stability analysis of nonlinear semidefinite programming (NSDP). Some of the results share a great deal of similarities to that in classical nonlinear programming (NLP), whereas some show essential differences. For detailed treatment of this topic, see the book by Bonnans and Shapiro [7] and the paper by Shapiro [28] for a good account on some similarities and differences of the stability analysis between NSDP and NLP.

It would need more space to review just a small part of the recent advances and their algorithmic implications to NSDP, see for example [3, 5, 8, 18, 31, 33, 34]. We are content with mentioning just a few that motivated our research. The first motivation comes from two recent papers of Sun [31, 8]. The paper on NSDP [31], which is strongly motivated by the book [7], establishes among many others the equivalence between the strong regularity of a KKT point of NSDP and the strong second-order sufficient condition (SSOSC) and the constraint

*School of Mathematics, The University of Southampton, Highfield, Southampton SO17 1BJ, UK. This author's research was partially supported by EPSRC Grant EP/D502535/1. E-mail: hdqi@soton.ac.uk

nondegeneracy [23, 24, 25] at a local optimal solution. It was further reported in Chan and Sun [8] that the SSOSC in the context of (linear) semidefinite programming (SDP) is equivalent to the dual nondegeneracy studied in [1] (known as the AHO nondegeneracy). We note that if the *strict complementarity* condition is assumed at an optimal solution of the primal SDP, such duality characterization of the SSOSC at that solution can also be obtained via [7, Thm. 5.91 and Thm. 5.85]. It is worth pointing out that the AHO nondegeneracy is actually the constraint nondegeneracy and the proof in [8] can be extended to prove that the SSOSC for the dual problem is equivalent to the primal nondegeneracy. One may ask to what extent this perfect duality result for SDP can be extended to NSDP?

The second motivation is from the algorithmic success in extending the primal-dual path-following method to convex quadratic semidefinite programming (QSDP) by Toh, Tütüncü, Todd [34] and Toh [33], where the dual nondegeneracy of QSDP was explored to ensure well-conditionedness of a sequence of matrices encountered in the path-following method. We note that Toh [33, Def. 3] used the dual nondegeneracy in SDP for the convex QSDP. We will see that the constraint nondegeneracy is weaker and is automatically satisfied by a good class of QSDP and some of the results in [33] due to the SDP dual nondegeneracy actually hold under the constraint nondegeneracy. Claims in this respect all come from the characterization of the dual constraint nondegeneracy in terms of the original QSDP.

The brief discussion above stimulates us to investigate the local duality of NSDP. It is no surprise that local duality of NLP has long been treated systematically, for example, in Wolfe [35], Luenberger [17], and Fujiwara, Han, and Mangasarian [13]. It is interesting to know what NLP local dualities have their NSDP counterparts and what do not have. This may be regarded as the third motivation to our current research.

To see why these questions are meaningful and nontrivial, we put them in more precise terms in conjunction with various problem formulations. Consider the NSDP:

$$(NSDP) \qquad \begin{array}{ll} \min_{x\in\mathcal{X}} & f(x) \\ \text{s.t.} & h(x) = 0 \\ & g(x) \in \mathcal{S}^n_+, \end{array} \tag{1}$$

where $f : \mathcal{X} \mapsto \mathrm{I\!R}$, $h : \mathcal{X} \mapsto \mathrm{I\!R}^m$, and $g : \mathcal{X} \mapsto \mathcal{S}^n$ are twice continuously differentiable and the second-order derivative of each function is locally Lipschitz continuous; $\mathcal{X}$ is a finite dimensional real space (endowed with a scalar product and the induced norm); $\mathcal{S}^n$ is the linear space of all the $n \times n$ real symmetric matrices (endowed with the standard trace inner product and the induced Frobenius norm); and $\mathcal{S}^n_+$ is the cone of all $n \times n$ positive semidefinite matrices in $\mathcal{S}^n$.

Let the Lagrangian function $L : \mathcal{X} \times \mathrm{I\!R}^m \times \mathcal{S}^n \mapsto \mathrm{I\!R}$ be

$$L(x, y, S) := f(x) - y^T h(x) - \langle S, g(x)\rangle,$$

where $\langle X, Y\rangle := \mathrm{Tr}(XY)$ for $X, Y \in \mathcal{S}^n$. Then the Wolfe dual [35] of NSDP (1) is

$$(NSDD) \qquad \begin{array}{ll} \max_{(x,y,S)} & L(x, y, S) \\ \text{s.t.} & \nabla_x L(x, y, S) = 0 \\ & S \in \mathcal{S}^n_+, \end{array} \tag{2}$$

where $\nabla_x L$ denotes the gradient of L with respect to its first argument x.

The quadratic SDP and its dual (QSDD) studied in [19, 34, 33] are just special cases of (1) and (2):

$$
\begin{array}{lll}
(QSDP) & \min & \frac{1}{2}\langle X, \mathcal{Q}(X)\rangle + \langle C, X\rangle \\
 & \text{s.t.} & \mathcal{A}(X) = b \\
 & & X \in \mathcal{S}^n_+,
\end{array}
\tag{3}
$$

and

$$
\begin{array}{lll}
(QSDD) & \max_{X,y,S} & -\frac{1}{2}\langle X, \mathcal{Q}(X)\rangle + b^T y \\
 & \text{s.t.} & -\mathcal{Q}(X) + \mathcal{A}^* y + S = C \\
 & & S \in \mathcal{S}^n_+,
\end{array}
\tag{4}
$$

where $\mathcal{Q} : \mathcal{S}^n \mapsto \mathcal{S}^n$ is a given self-adjoint linear operator in $\mathcal{S}^n$, $\mathcal{A} : \mathcal{S}^n \mapsto \mathbb{R}^m$ is a linear mapping and $\mathcal{A}^* : \mathbb{R}^m \mapsto \mathcal{S}^n$ denotes its adjoint. An interesting case is when $\mathcal{Q}$ is positive semidefinite (i.e, convex QSDP).

When there is no quadratic term, we have the standard semidefinite programming (SDP):

$$
\begin{array}{lll}
(SDP) & \min & \langle C, X\rangle \\
 & \text{s.t.} & \mathcal{A}(X) = b \\
 & & X \in \mathcal{S}^n_+,
\end{array}
\tag{5}
$$

and its dual

$$
\begin{array}{lll}
(SDD) & \max_{y,S} & b^T y \\
 & \text{s.t.} & \mathcal{A}^* y + S = C \\
 & & S \in \mathcal{S}^n_+.
\end{array}
\tag{6}
$$

Suppose $(\bar{x}, \bar{y}, \bar{S})$ is a KKT triple of NSDP (1), that is, it satisfies the KKT condition

$$
\left.
\begin{array}{r}
\nabla_x L(x, y, S) = 0 \\
h(x) = 0 \\
g(x) \in \mathcal{S}^n_+, \ S \in \mathcal{S}^n_+, \ \ \langle S, g(x)\rangle = 0
\end{array}
\right\} \ (KKT).
\tag{7}
$$

It is well known that the KKT condition is equivalent to the following generalized equation having a solution:

$$
0 \in \left[\begin{array}{c} \nabla_x L(x, y, S) \\ h(x) \\ -g(x) \end{array} \right] + \left[\begin{array}{c} N_{\mathcal{X}}(x) \\ N_{\mathbb{R}^m}(y) \\ N_{\mathcal{S}^n_-}(-S) \end{array} \right],
\tag{8}
$$

where $N_D(z)$ is the normal cone of D at $z \in D$ in an appropriate space. It is also well-known [11] that the KKT condition (7) can be equivalently reformulated as a system of equations:

$$
F(x, y, S) := \left[\begin{array}{l} \nabla_x L(x, y, S) \\ h(x) \\ g(x) - \Pi_{\mathcal{S}^n_+}(g(x) - S) \end{array} \right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right],
$$

where $\Pi_{\mathcal{S}^n_+}(X)$ denotes the orthogonal projection of a given matrix $X \in \mathcal{S}^n$ onto $\mathcal{S}^n_+$.

It is recently proved by Sun [31] that the *strong second-order sufficient condition* (SSOSC) at $(\bar{x}, \bar{y}, \bar{S})$ and the constraint nondegeneracy for NSDP (1) (see Sec. 2 for formal definitions) amount to the *strong regularity* of $(\bar{x}, \bar{y}, \bar{S})$ as a solution of the generalized equation (8) (under the assumption that $\bar{x}$ is a local solution of NSDP (1)). They each are also equivalent to the nonsingularity of every element in the generalized Jacobian $\partial F(\bar{x}, \bar{y}, \bar{S})$.

These equivalent characterizations are much simplified when applied to the linear SDP (5) and its dual (6). It is proved by Chan and Sun [8, Prop. 15] that

$$\text{The SSOSC for the primal SDP (5)} \iff \text{The dual nondengeneracy for the dual (6)} \tag{9}$$

and

$$\text{The SSOSC for the dual SDP (6)} \iff \text{The primal nondengeneracy for the primal (5).} \tag{10}$$

Here the primal/dual nondegeneracy means the constraint nondegeneracy for the primal/dual linear SDP problem.

This perfect duality for linear SDP between primal/dual nondegeneracy and the dual/primal SSOSC is responsible for a number of important characterizations of a strongly regular KKT triple. Unfortunately, there may exist no such a perfect duality for NSDP, even for the convex QSDP. To see this, let us reformulate SDP (5) as a convex QSDP:

$$\begin{array}{ll} \min & \frac{1}{2}\langle X, (\bar{X} \circledast \bar{S})(X)\rangle + \langle C, X\rangle \\ \text{s.t.} & \mathcal{A}(X) = b \\ & X \in \mathcal{S}^n_+, \end{array} \tag{11}$$

where $(U \circledast V)(X) := (UXV + V^T X^T U^T)/2$ is the symmetrized Kronecker product between dimensionally compatible matrices. The duality relationship (10) does not hold anymore for (11). We will see that the major reason is that $(\bar{X} \circledast \bar{S})$ is not nonsingular.

The major purpose of this paper is to investigate the duality relationships of the type (9), (10) between NSDP (1) and its Wolfe dual (2). To achieve this purpose we first introduce the dual SSOSC, which generalizes its NLP counterpart studied in [13], but with a major difference of a *sigma* term now being attached to the SSOSC. There are a number of interesting characterizations of the dual SSOSC. For example, it is nothing but Sun's SSOSC applied to the dual problem (2) (Thm. 3.1). However, a necessary condition for the dual SSOSC indicates that the inverse of the Hessian of the Lagrangian function exists at the KKT triple $(\bar{x}, \bar{y}, \bar{S})$ and it is positive definite on the normal space (contained in the normal cone) to the feasible region at $\bar{x}$ (Prop. 3.3). A sufficient condition is that the Hessian is positive definite at $(\bar{x}, \bar{y}, \bar{S})$ and the primal nondegeneracy holds at $\bar{x}$ (Thm. 3.4).

These characterizations are significantly simplified for the convex case. For example, The fore-mentioned sufficient condition becomes necessary for the convex QSDP (3) (Thm. 3.10). Like SDP, for the convex QSDP we also have a similar result of the perfect duality:

$$\text{The SSOSC for the primal QSDP (3)} \iff \text{The dual nondegeneracy for the dual (4)} \tag{12}$$

and

$$\text{The SSOSC for the dual QSDP (4)} \iff \left\{ \begin{array}{l} \text{The primal nondegeneracy for the primal (3)} \\ \text{and } \mathcal{Q} \text{ is positive definite.} \end{array} \right. \tag{13}$$

The equivalence (12) has a nice application to the path-following method studied in [33] (see Prop. 3.8). It says that the positive definiteness of the submatrix $(\widetilde{\mathcal{Q}}_\rho)_{\alpha\alpha}$, which plays an important role in [33], is guaranteed under the weaker condition of the dual constraint nondegeneracy than the assumed SDP dual nondegeneracy in [33]. The equivalence (13) is very different from that in SDP (10). Those equivalence characterizations provide further insight into the efficiency of the path-following method studied in [34, 33] for convex QSDP including the well-known *nearest correlation matrix problem*, which satisfies not only the primal and dual SSOSC but also the primal and dual nondegeneracy according to our results.

For the convex QSDP, we also establish equivalences between a number of conditions for the strong regularity of a KKT triple. This includes the equivalence between the nonsingularity of $\partial_B F(\bar{X}, \bar{y}, \bar{S})$ and that of $\partial F(\bar{X}, \bar{y}, \bar{S})$, generalizing a surprising result of the equivalence between them for SDP [8].

The paper is organized as follows. We review some basic definitions in Section 2 concerning constraint nondegeneracy, primal and dual SSOSC and generalized Jacobians. The material in this part is kept minimal. One may refer to the cited references for more on those definitions. We present our major results in Section 3. We arrange our results for the nonconvex case and the convex case separately in order to distinguish the differences of duality results in the two contexts. The convex case is the most interesting one as it seems that the convex QSDP (3) is becoming more and more important as a class of problems in its own right [34, 33]. An example is also given to illustrate the point that perfect duality results may break down for nonconvex QSDP even with very good conditions. In Section 4, we conduct a brief similarity comparison between some obtained NSDP results and their counterparts in NLP. While some interesting explanation was given on existing results, a new finding (Prop. 4.2) states that the dual SOSC in NLP [13] actually assumes the *strict constraint qualification* as a priori. We conclude the paper in Section 5.

2 Constraint Nondegeneracy, Primal/Dual SSOSC and Generalized Jacobians

2.1 Constraint Nondegeneracy

Let us formally introduce the constraint nondegeneracy of a feasible set at a particular point. Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite dimensional real vector spaces each equipped with a scalar product and its induced norm. Let $G : \mathcal{X} \mapsto \mathcal{Y}$ be a continuously differentiable function. Denote by $\mathcal{J}_x G(x)$ the Jacobian of G at x and by $\nabla G(x)$ the transpose (or the adjoint when the Jacobian is viewed as an operator) of the Jacobian. Let K be a nonempty and closed convex set in $\mathcal{Y}$. G defines the following feasible set

$$G(x) \in K, \quad x \in \mathcal{X}. \tag{14}$$

Assume that $\bar{x} \in \mathcal{X}$ is a feasible point of (14). Let $T_K(G(\bar{x}))$ be the tangent cone of K at $G(\bar{x})$. We let $\operatorname{lin}(T_K(G(\bar{x})))$ be the largest linear space contained in $T_K(G(\bar{x}))$.

Definition 2.1 *[7, 4.172] The constraint nondegeneracy holds at $\bar{x}$ if*

$$\mathcal{J}_x G(\bar{x})\mathcal{X} + \operatorname{lin}(T_K(G(\bar{x}))) = \mathcal{Y}. \tag{15}$$

The well known Robinson constraint qualification [22], which is weaker than the constraint nondegeneracy, can be equivalently written as

$$\mathcal{J}_x G(\bar{x})\mathcal{X} + T_K(G(\bar{x})) = \mathcal{Y}.$$

The constraint nondegeneracy was introduced by Bonnans and Shapiro [6] for general optimization problems. It is also known as the transversality condition in nonlinear semidefinite context, see Shaprio and Fan [30]. For *polyhedral* set K, it coincides with the one used in Robinson [24].

It is useful later on to see that (15) takes on different forms when applied to the primal NSDP (1) and its dual (2). For NSDP (1), let

$$G(x) := \begin{pmatrix} h(x) \\ g(x) \end{pmatrix}, \qquad K := \begin{pmatrix} \{0\} \\ \mathcal{S}^n_+ \end{pmatrix}, \quad \text{and} \quad \mathcal{Y} := \begin{pmatrix} \mathbb{R}^m \\ \mathcal{S}^n \end{pmatrix}. \tag{16}$$

Then the constraint nondegeneracy holds at $\bar{x}$ if

$$\begin{bmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{bmatrix} \mathcal{X} + \begin{bmatrix} \{0\} \\ \mathrm{lin}\left(T_{\mathcal{S}^n_+}(g(\bar{x}))\right) \end{bmatrix} = \begin{bmatrix} \mathbb{R}^m \\ \mathcal{S}^n \end{bmatrix}. \tag{17}$$

For the dual problem (2), let

$$\tilde{h}(x,y,S) := \nabla_x L(x,y,S), \quad \tilde{g}(x,y,S) := S, \quad \mathcal{Z} := \mathcal{X} \times \mathbb{R}^m \times \mathcal{S}^n \tag{18}$$

and

$$G(x,y,S) := \begin{pmatrix} \tilde{h}(x,y,S) \\ \tilde{g}(x,y,S) \end{pmatrix}, \quad K := \begin{pmatrix} \{0\} \\ \mathcal{S}^n_+ \end{pmatrix}, \quad \text{and} \quad \mathcal{Y} := \begin{pmatrix} \mathcal{X} \\ \mathcal{S}^n \end{pmatrix}.$$

The constraint nondegeneracy for (2) holds at a feasible point $(\bar{x}, \bar{y}, \bar{S})$ if

$$\begin{bmatrix} \mathcal{J}_z \tilde{h}(\bar{x},\bar{y},\bar{S}) \\ \mathcal{J}_z \tilde{g}(\bar{x},\bar{y},\bar{S}) \end{bmatrix} \mathcal{Z} + \begin{bmatrix} \{0\} \\ \mathrm{lin}\left(T_{\mathcal{S}^n_+}(\bar{S})\right) \end{bmatrix} = \begin{bmatrix} \mathcal{X} \\ \mathcal{S}^n \end{bmatrix}, \tag{19}$$

where $z := (x,y,S) \in \mathcal{Z}$. Condition (17) and Condition (19) are often respectively referred to as the primal and dual nondegeneracy of the problem (1) and (2). The primal and dual nondegeneracy in SDP have been studied and used in [1, 15, 12, 34, 33].

2.2 Primal and Dual SSOSC

In this section, we introduce the (primal) SSOSC of Sun [31] for the primal problem (1) and the dual SSOSC for the dual problem (2). The dual SSOSC generalizes a corresponding concept of Fujiwara, Han and Mangasarian [13] from NLP to NSDP and is actually equivalent to Sun's SSOSC applied to the dual problem (2).

Let $(\bar{x}, \bar{y}, \bar{S})$ be a KKT triple satisfying (7). Let

$$\bar{A} := g(\bar{x}) - \bar{S}$$

and $\bar{A}$ has the following spectral decomposition

$$\bar{A} = P \Lambda P^T, \tag{20}$$

where Λ is the diagonal matrix of eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of $\bar{A}$ and P is a corresponding orthogonal matrix of orthonormal eigenvectors. For this eigenvector $\lambda \in \mathbb{R}^n$, define the corresponding symmetric matrix $U \in \mathcal{S}^n$ with entries

$$U_{ij} := \frac{\max\{\lambda_i, 0\} + \max\{\lambda_j, 0\}}{|\lambda_i| + |\lambda_j|}, \quad i,j = 1,\ldots,n, \tag{21}$$

where 0/0 is treated to be 1.

Define three index sets of positive, zero, and negative eigenvalues of $\bar{A}$, respectively, by

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

Write

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

with P_α being a submatrix of P containing the eigenvectors of positive eigenvalues, P_β the eigenvectors of zero eigenvalues and P_γ the eigenvectors of negative eigenvalues. Then we have

$$g(\bar{x}) = P \begin{bmatrix} \Lambda_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P^T, \quad \bar{S} = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\Lambda_\gamma \end{bmatrix} P^T.$$

It is also known [2, 1, 7] that

$$\mathrm{lin}\left(T_{\mathcal{S}^n_+}(g(\bar{x}))\right) = \left\{B \in \mathcal{S}^n \;\middle|\; [P_\beta, P_\gamma]^T B [P_\beta, P_\gamma] = 0\right\} \tag{22}$$

and

$$\mathrm{lin}\left(T_{\mathcal{S}^n_+}(\bar{S})\right) = \left\{B \in \mathcal{S}^n \;\middle|\; [P_\alpha, P_\beta]^T B [P_\alpha, P_\beta] = 0\right\}. \tag{23}$$

Recall G and K are defined in (16). The critical cone of the optimization problem (1) at $\bar{x}$ is defined by

$$C(\bar{x}) := \left\{d \in \mathcal{X} \;\middle|\; \nabla f(\bar{x})^T d \le 0 \;\text{ and }\; \mathcal{J}_x G(\bar{x}) d \in T_K(G(\bar{x}))\right\}.$$

Let $\mathrm{aff}(C(\bar{x}))$ denote the affine hull of $C(\bar{x})$. Since $0 \in C(\bar{x})$, $\mathrm{aff}(C(\bar{x})) = C(\bar{x}) - C(\bar{x})$, the linear space generated by $C(\bar{x})$. Define

$$\mathrm{app}(\bar{y}, \bar{S}) := \left\{d \in \mathcal{X}|\ \mathcal{J}_x h(\bar{x}) d = 0,\ [P_\beta, P_\gamma]^T (\mathcal{J}_x g(\bar{x}) d) P_\gamma = 0\right\}.$$

As pointed out in [31, Eq.38], $\mathrm{app}(\bar{y}, \bar{S})$ is an outer approximation to $\mathrm{aff}(C(\bar{x}))$. It was further pointed out by one referee that under the constraint nondegeneracy, $\mathrm{app}(\bar{y}, \bar{S}) = \mathrm{aff}(\bar{C})$ (see Cor. 2.3 and Eq.4.8 in [5]).

When formulating second-order necessary conditions for optimization problems involving the general constraints (14) where K may not be necessarily polyhedral, a widely known *sigma*-term plays an important role (see [7, Sec. 3.22], and Kawasaki [16] and Cominetti [10] for early development that leads to the *sigma*-term.) It was noted [7, P.177] that the σ-term vanishes when K is polyhedral. The term is often defined on the outer second order tangent set to K at a feasible point $\bar{x}$, denoted by $T^2_K(G(\bar{x}), \mathcal{J}_x G(\bar{x}) d)$ for some direction $d \in \mathcal{X}$ (see [7, Sec. 3.2.1]). In the nonlinear SDP context, when the Lagrangian multiplier $\bar{S}$ is available the term can be characterized by (see [27, P.313] and [7, P.487])

$$\sigma(\bar{S}, T_{\mathcal{S}^n_+}(g(\bar{x}), \mathcal{J}_x g(\bar{x}) d) = 2\langle \bar{S},\ (\mathcal{J}_x g(\bar{x})) g(\bar{x})^\dagger (\mathcal{J}_x g(\bar{x}) d) \rangle \quad \text{for } d \in C(\bar{x}),$$

where $g(\bar{x})^\dagger$ is the Moore-Penrose pseudo-inverse of $g(\bar{x})$. We note that for given $g(\bar{x})$ and $\bar{S}$, the σ-term is a quadratic function of $\mathcal{J}_x g(\bar{x}) d$ over $d \in C(\bar{x})$. This quadratic function was singled out and used by Sun [31] to derive his strong second order sufficient condition for NSDP (1). We restate this quadratic term below for easy reference.

Definition 2.2 *[31, Def. 2.1] For any given matrix $B \in \mathcal{S}^n$, define the linear-quadratic function $\Upsilon_B : \mathcal{S}^n \times \mathcal{S}^n \mapsto \mathbb{R}$, which is linear in the first argument and quadratic in the second argument, by*

$$\Upsilon_B(T, H) := 2\langle T, HB^\dagger H\rangle, \quad (T, H) \in \mathcal{S}^n \times \mathcal{S}^n.$$

It is easy to note that $\Upsilon_B(T, H) = 2\langle H, (T \circledast B^\dagger)(H)\rangle$. Moreover [31], $\Upsilon_{g(\bar{x})}(\bar{S}, H) \geq 0$ and $\Upsilon_{\bar{S}}(g(\bar{x}), H) \geq 0$ for any $H \in \mathcal{S}^n$ and

$$\Upsilon_{g(\bar{x})}(\bar{S}, H) > 0 \iff \Upsilon_{\bar{S}}(g(\bar{x}), H) > 0 \iff P_\alpha^T H P_\gamma \neq 0. \tag{24}$$

Now we are ready to state the SSOSC of Sun [31].

Definition 2.3 [1] *We say that the KKT triple $(\bar{x}, \bar{y}, \bar{S})$ of (1) satisfies the* strong second order sufficient condition *(SSOSC) if*

$$\langle d, \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})d\rangle + \Upsilon_{g(\bar{x})}(\bar{S}, \mathcal{J}_x g(\bar{x})d) > 0 \quad \forall\, 0 \neq d \in \mathrm{app}(\bar{y}, \bar{S}) \tag{25}$$

Definition 2.4 *We say that the KKT triple $(\bar{x}, \bar{y}, \bar{S})$ of (1) satisfies the* dual strong second-order sufficient condition *(dual SSOSC) if the Hessian $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ is nonsingular and*

$$\langle w, \left(\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})\right)^{-1} w\rangle + \Upsilon_{\bar{S}}(g(\bar{x}), H) > 0, \quad \forall\, 0 \neq (w, y, H) \in \mathrm{app}(\bar{x}), \tag{26}$$

where

$$\mathrm{app}(\bar{x}) := \left\{ (w, y, H) \in \mathcal{X} \times \mathbb{R}^m \times \mathcal{S}^n \,\middle|\, \begin{array}{l} w = \nabla h(\bar{x})y + \nabla g(\bar{x})H \\ P_\alpha^T H[P_\alpha, P_\beta] = 0 \end{array} \right\}.$$

We therefore refer to condition (25) as the *primal* SSOSC. The dual SSOSC generalizes a similar concept in [13] from NLP to NSDP.

2.3 Generalized Jacobians

Suppose $\Xi : \mathcal{O} \subseteq \mathcal{X} \mapsto \mathcal{Y}$ is a locally Lipschitz function on the open set $\mathcal{O}$. By Rademacher's theorem [26, Sec. 9.J], we know that Ξ is almost everywhere F(réchet)-differentiable in $\mathcal{O}$. Denote by $\mathcal{D}_\Xi$ the set of all points in $\mathcal{O}$ where Ξ is F-differentiable. Then Clarke's generalized Jacobian [9] of Ξ at (any) $y \in \mathcal{O}$ is defined by

$$\partial \Xi(y) := \mathrm{conv}\left\{\partial_B \Xi(y)\right\},$$

where "conv" denotes the convex hull and the B-subdifferential $\partial_B \Xi(y)$ is the set of the collection of all limits of Jacobians of Ξ near y, see Qi [21]:

$$\partial_B \Xi(y) := \left\{ V \,\middle|\, V = \lim_{k \to \infty} \mathcal{J}\Xi(y^k),\ y^k \to y \text{ and } y^k \in \mathcal{D}_\Xi \right\}.$$

The Lipschitz function that we are to encounter in this paper is the (orthogonal) projection operator $\Pi_{\mathcal{S}^n_+}(X)$ of a given matrix $X \in \mathcal{S}^n$ to the positive semidefinite cone. A lot of nonsmooth properties of this operator has been known, see [3, 4, 7, 32, 20]. But the following result of Sun [31, Prop. 2.2] is enough for our use

[1] The definition is slightly different from that of Sun [31], where the quantity in (25) runs over all possible Lagrangian multipliers $(\bar{y}, \bar{S})$. Therefore, it is there called the SSOSC at $\bar{x}$, whereas we define it at a specified KKT triple.

Proposition 2.5 *Suppose that $\bar{A} \in \mathcal{S}^n$ has the spectral decomposition as in (20). Then for any $V \in \partial_B \Pi_{\mathcal{S}^n_+}(\bar{A})$ (respectively, $\partial \Pi_{\mathcal{S}^n_+}(\bar{A})$), there exists a $V_{|\beta|} \in \partial_B \Pi_{\mathcal{S}^{|\beta|}_+}(0)$ (respectively, $\partial \Pi_{\mathcal{S}^{|\beta|}_+}(0)$) such that*

$$V(H) = P \begin{bmatrix} \widetilde{H}_{\alpha\alpha} & \widetilde{H}_{\alpha\beta} & U_{\alpha\gamma} \circ \widetilde{H}_{\alpha\gamma} \\ \widetilde{H}^T_{\alpha\beta} & V_{|\beta|}(\widetilde{H}_{\beta\beta}) & 0 \\ \widetilde{H}^T_{\alpha\gamma} \circ U^T_{\alpha\gamma} & 0 & 0 \end{bmatrix} P^T, \quad \forall\, H \in \mathcal{S}^n, \tag{27}$$

where $\widetilde{H} := P^T H P$, U is defined as in (21), and $\circ$ denotes the Hadamard product. Conversely, for any $V_{|\beta|} \in \partial_B \Pi_{\mathcal{S}^{|\beta|}_+}(0)$ (respectively, $\partial \Pi_{\mathcal{S}^{|\beta|}_+}(0)$), there exists a $V \in \partial_B \Pi_{\mathcal{S}^n_+}(\bar{A})$ (respectively, $\partial \Pi_{\mathcal{S}^n_+}(\bar{A})$) such that (27) holds.

3 Local Duality

3.1 The Nonconvex Case

In this section, we characterize the dual SSOSC and the dual nondegeneracy in terms of the primal problem (1). We may understand the dual SSOSC from at least three aspects. First, when specialized to SDP (5) (in this case, the Hessian matrix is 0, implying no w term in (26)), it is simply equivalent to the primal nondegeneracy of SDP[2].

The second way to understand the dual SSOSC is that it is actually Sun's primal SSOSC applied to the dual problem (2). This fact is proved in the next result.

Theorem 3.1 *Suppose $(\bar{x}, \bar{y}, \bar{S})$ is a KKT triple of the primal problem (1). Then*

(i) $(\bar{x}, \bar{y}, \bar{S})$ is a KKT point of the dual problem (2) with $\bar{t} = 0 \in \mathcal{X}$ and $\bar{T} = g(\bar{x}) \in \mathcal{S}^n_+$ as the corresponding Lagrangian multipliers; and

(ii) Sun's SSOSC for the dual problem (2) holds at $(\bar{x}, \bar{y}, \bar{S}, 0, g(\bar{x}))$ if and only if the dual SSOSC holds at $(\bar{x}, \bar{y}, \bar{S})$.

Proof. The proof is almost word-by-word extension of [13, Thm. 2.2] by noticing the following major facts to be used in the extension. In order to use Sun's SSOSC for the dual problem (2), the set $\mathrm{app}(\bar{t}, \bar{T})$ with $\bar{t} = 0$ and $\bar{T} = g(\bar{x})$ (see (i) of the theorem) has to be characterized. Denote $\mathcal{Z} := \mathcal{X} \times \mathrm{I\!R}^m \times \mathcal{S}^n$ and $d := (d_x, d_y, d_S) \in \mathcal{Z}$. By recalling that $\bar{A} = -(\bar{S} - g(\bar{x}))$ has spectral decomposition (20), we have

$$\mathrm{app}(\bar{t}, \bar{T}) := \left\{ d \in \mathcal{Z} \;\middle|\; \begin{array}{l} \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) d_x - \nabla h(\bar{x}) d_y - \nabla g(\bar{x}) d_S = 0 \\ P^T_\alpha d_S [P_\alpha, P_\beta] = 0 \end{array} \right\}.$$

Let $M : \mathcal{X} \times \mathrm{I\!R}^m \times \mathcal{S}^n \times \mathcal{X} \times \mathcal{S}^n \mapsto \mathrm{I\!R}$ be the Lagrangian function for the dual problem (2):

$$M(x, y, S, t, T) := L(x, y, S) + \langle t, \nabla_x L(x, y, S) \rangle + \langle S, T \rangle.$$

Then Sun's SSOSC for the dual problem (2) holds at $(\bar{x}, \bar{y}, \bar{S}, \bar{t}, \bar{T})$ if and only if

$$\langle d, \nabla_{zz} M(\bar{x}, \bar{y}, \bar{S}, \bar{t}, \bar{T}) d \rangle - \Upsilon_{\bar{S}}(\bar{T}, d_S) < 0 \;\; \forall\, 0 \neq d \in \mathrm{app}(\bar{\mathrm{t}}, \bar{\mathrm{T}}). \tag{28}$$

[2]This can be proved in a way similar to [8, Prop. 15], which states that the dual nondegeneracy is equivalent to the primal SSOSC in SDP.

With those facts in mind, we can extend the proof from [13, Thm.2.2] to the semidefinite programming case and we omit the details. □

The dual SSOSC prerequisites existence of the inverse of the Hessian of the Lagrangian function at $(\bar{x}, \bar{y}, \bar{S})$. The next result shows that the inverse is positive definite on a normal space of the feasible set at $\bar{x}$. First let us define

$$\mathcal{T}_{\bar{x}} := \left\{ d \in \mathcal{X} \;\middle|\; \mathcal{J}_x h(\bar{x}) d = 0, \text{ and } \mathcal{J}_x g(\bar{x}) d \in \text{lin}\left(T_{\mathcal{S}^n_+}(g(\bar{x})) \right) \right\} \tag{29}$$

and

$$\mathcal{N}_{\bar{x}} := \left\{ w \in \mathcal{X} \;\middle|\; \begin{array}{l} w = \nabla h(\bar{x}) y + \nabla g(\bar{x}) H \\ P_\alpha^T H = 0 \end{array} \right\}.$$

Then we have

Lemma 3.2 *$\mathcal{N}_{\bar{x}} = \mathcal{T}_{\bar{x}}^\perp$, where $\mathcal{T}_{\bar{x}}^\perp$ is the orthogonal complement of $\mathcal{T}_{\bar{x}}$.*

Proof [3] Let G and K be defined as in (16). We first note that

$$\left[\text{lin}\left(T_{\mathcal{S}^n_+}(g(\bar{x})) \right) \right]^\perp = \left\{ H \in \mathcal{S}^n \;\middle|\; P_\alpha^T H = 0 \right\},$$

which implies by the definition of $\mathcal{N}_{\bar{x}}$

$$\begin{aligned} \mathcal{N}_{\bar{x}} &= \left\{ J_x G(\bar{x})^T \left(\text{lin } T_K(G(\bar{x})) \right)^\perp \right\} \\ &= \left\{ v \;\middle|\; (J_x G(\bar{x}))^{-T} v \in \left(\text{lin } T_K(G(\bar{x})) \right)^\perp \right\}. \end{aligned}$$

We further note the following known fact in linear algebra: Let $A : \mathcal{X} \mapsto \mathcal{Y}$ be a linear mapping from one finite dimensional space $\mathcal{X}$ to another $\mathcal{Y}$ and $V \subset \mathcal{Y}$ is a subspace, it holds

$$\left(A^T V^\perp \right)^\perp = A^{-1}(V) := \left\{ v \in \mathcal{X} \;\middle|\; Av \in V \right\}. \tag{30}$$

Now let $V = \left(\text{lin}\,(T_K(G(\bar{x}))) \right)^\perp$ and $A = (\mathcal{J}_x G(\bar{x}))^{-T}$ in (30), we obtain that

$$\begin{aligned} \mathcal{N}_{\bar{x}} &= \left(A^T V^T \right)^\perp \\ &= \left\{ (\mathcal{J}_x G(\bar{x}))^{-1} \text{lin}\,(T_K(G(\bar{x}))) \right\}^\perp \\ &= \left\{ d \in \mathcal{X} \;\middle|\; \mathcal{J}_x G(\bar{x}) d \in \text{lin}\,(T_K(G(\bar{x}))) \right\}^\perp = \mathcal{T}_{\bar{x}}. \end{aligned}$$

□

The following result says that the dual SSOSC implies the positive definiteness of the inverse of the Hessian on the normal space $\mathcal{N}_{\bar{x}}$, generalizing a similar result [13, Eq.(2.9)] in NLP. The dual SSOSC also implies the primal nondegeneracy. This result provides the third aspect of the dual SSOSC.

[3] The author thanks one of the referees for providing this compact proof, replacing the original lengthy one.

Proposition 3.3 *Suppose $(\bar{x}, \bar{y}, \bar{S})$ is a KKT triple of the primal problem (1) and the dual SSOSC holds at $(\bar{x}, \bar{y}, \bar{S})$. Then the following hold.*

(i) The inverse of the Hessian $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}))$ exists and

$$\langle w, \left(\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})\right)^{-1} w\rangle > 0 \quad \forall\, 0 \neq w \in \mathcal{N}_{\bar{x}}.$$

(ii) The primal nondegeneracy holds at $\bar{x}$.

Proof. (i) The existence of the inverse of the Hessian matrix at $(\bar{x}, \bar{y}, \bar{S})$ is by definition of the dual SSOSC. Another way to look at $\mathcal{N}_{\bar{x}}$ is that

$$\mathcal{N}_{\bar{x}} = \left\{ w \in \mathcal{X} \;\middle|\; \begin{array}{l} \exists\, (y, H) \in \mathbb{R}^m \times \mathcal{S}^n \text{ such that} \\ (w, y, H) \in \operatorname{app}(\bar{x}) \text{ and } P_\alpha^T H P_\gamma = 0 \end{array} \right\}.$$

Therefore, for any $0 \neq w \in \mathcal{N}_{\bar{x}}$, there exists a pair $(y, H) \in \mathbb{R}^m \times \mathcal{S}^n$ such that $(w, y, H) \in \operatorname{app}(\bar{x})$ and $P_\alpha^T H P_\gamma = 0$, which implies $\Upsilon_{\bar{S}}(g(\bar{x}), H) = 0$ because of (24). The dual SSOSC guarantees

$$\langle w, \left(\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})\right)^{-1} w\rangle > 0.$$

That is, $\nabla_{xx} \left(L(\bar{x}, \bar{y}, \bar{S})\right)^{-1}$ is positive definite on the normal space $\mathcal{N}_{\bar{x}}$.

(ii) Suppose that the primal nondegeneracy does not hold at $\bar{x}$. Then

$$\left\{ \begin{bmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{bmatrix} \mathcal{X} \right\}^\perp \cap \begin{bmatrix} \{0\} \\ \operatorname{lin}\left(T_{\mathcal{S}_+^n}(g(\bar{x}))\right) \end{bmatrix}^\perp \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{31}$$

which means there exists $0 \neq (y, H) \in \mathbb{R}^m \times \mathcal{S}^n$ such that it belongs to the left-hand side of (31), i.e.

$$\langle y, \mathcal{J}_x h(\bar{x}) d\rangle + \langle H, \mathcal{J}_x g(\bar{x}) d\rangle = 0 \quad \forall\, d \in \mathcal{X} \tag{32}$$

and

$$\langle H, S\rangle = 0 \quad \forall\, S \in \operatorname{lin}\left(\mathcal{T}_{\mathcal{S}_+^n}(g(\bar{x}))\right). \tag{33}$$

Then (32) and (33) together with (22) imply that

$$\nabla h(\bar{x}) y + \nabla g(\bar{x}) H = 0 \quad \text{and} \quad P_\alpha^T H = 0. \tag{34}$$

Then $(w = 0, y, H) \in \operatorname{app}(\bar{x})$ and $\Upsilon_{\bar{S}}(g(\bar{x}), H) = 0$ because of (24) and $P_\alpha^T H P_\gamma = 0$. This means

$$\langle w, \left(\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})\right)^{-1} w\rangle + \Upsilon_{\bar{S}}(g(\bar{x}), H) = 0$$

for $0 \neq (w, y, H) \in \operatorname{app}(\bar{x})$, contradicting the dual SSOSC. Hence the primal nondegeneracy must hold at $\bar{x}$ if the dual SSOSC holds at $(\bar{x}, \bar{y}, \bar{S})$. $\square$

As pointed out in [31] and easily verified by itself, the positive definiteness of the Hessian $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ is sufficient to ensure the primal SSOSC. But it is not sufficient to ensure the dual SSOSC. Nevertheless, the next result shows that if in addition the primal nondegeneracy holds at $\bar{x}$, then it is sufficient.

Theorem 3.4 *Suppose $(\bar{x}, \bar{y}, \bar{S})$ is a KKT triple of the primal problem (1). Assume that the primal nondegeneracy holds at $\bar{x}$. If $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ is positive definite then the dual SSOSC holds at $(\bar{x}, \bar{y}, \bar{S})$.*

Proof. We only need to consider those points $(w, y, H) \in \mathrm{app}(\bar{x})$ with $(y, H) \neq 0$. If $w \neq 0$, then the positive definiteness of $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ implies condition (26) because the quadratic term Υ is always nonnegative. So we consider the case $w = 0$, which means

$$\left.\begin{aligned} \nabla h(\bar{x}) y + \nabla g(\bar{x}) H &= 0 \\ P_\alpha^T H [P_\alpha, P_\beta] &= 0 \end{aligned}\right\} \tag{35}$$

If $P_\alpha^T H P_\gamma \neq 0$, then $\Upsilon_{\bar{S}}(g(\bar{x}), H) > 0$, which together with $w = 0$ implies the condition (26).

We shall prove that (35) with the condition

$$P_\alpha^T H P_\gamma = 0 \tag{36}$$

implies $(y, H) = 0$, contradicting the choice $(y, H) \neq 0$. This contradiction means either $w \neq 0$ or if $w = 0$ then $P_\alpha^T H P_\gamma \neq 0$. As proved above, the dual SSOSC holds for those two cases.

Now we prove that (35) and (36) imply $(y, H) = 0$. Note that the primal nondegeneracy (17) holds at $\bar{x}$. This equivalently implies

$$\left\{ \begin{bmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{bmatrix} \mathcal{X} \right\}^\perp \cap \begin{bmatrix} \{0\} \\ \mathrm{lin}\left(T_{\mathcal{S}^n_+}(g(\bar{x})) \right) \end{bmatrix}^\perp = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{37}$$

Let (y, H) belong to the left-hand side of (37). Then the primal nondegeneracy means that the linear system (34) has $(y, H) = (0, 0)$ as its unique solution. This is equivalent to say that (35) and (36) have $(y, H) = 0$ as its only solution. This finishes the proof. $\square$

Since $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ is assumed to be positive definite in Thm. 3.4, one may be tempted to think whether a stronger result holds: $w^T \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) w > 0$ for all $0 \neq (w, y, H) \in \mathrm{app}(\bar{x})$. The answer is no. The major reason is that the linear equations (35) in (y, H) may have *nonzero* solutions, leading to existence of such a point $0 \neq (w, y, H) \in \mathrm{app}(\bar{x})$ with $w = 0$.

Now we present a characterization of the dual nondegeneracy, which has significant implications in the convex case. Recall the definitions of $\tilde{h}$ and $\tilde{g}$ defined in (18). It is easy to calculate that

$$\mathcal{J}_z \tilde{h}(\bar{x}, \bar{y}, \bar{S}) = \left[\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}), \ -\nabla h(\bar{x}), \ -\nabla g(\bar{x}) \right]$$

and

$$\mathcal{J}_z \tilde{g}(\bar{x}, \bar{y}, \bar{S}) = [0, \ 0, \ I].$$

Then the dual nondegeneracy (19) for the dual problem (2) means

$$\begin{bmatrix} \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) & -\nabla h(\bar{x}) & -\nabla g(\bar{x}) \\ 0 & 0 & I \end{bmatrix} \mathcal{Z} + \begin{bmatrix} \{0\} \\ \mathrm{lin}\left(T_{\mathcal{S}^n_+}(\bar{S}) \right) \end{bmatrix} = \begin{bmatrix} \mathcal{X} \\ \mathcal{S}^n \end{bmatrix}. \tag{38}$$

Due to the special structure of the matrix in (38), we see that (38) is satisfied if $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ is nonsingular. Anyway, we have the following characterization, part (i) of which strengthens this observation.

Proposition 3.5 *Suppose $(\bar{x}, \bar{y}, \bar{S})$ is a KKT triple of the primal problem (1). The following hold.*

(i) The dual nondegeneracy (19) for the dual problem (2) holds at $(\bar{x}, \bar{y}, \bar{S})$ if and only if the linear system of variable d_x

$$\left.\begin{array}{rl} \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) d_x & = 0 \\ \mathcal{J}_x h(\bar{x}) d_x & = 0 \\ P_\gamma^T \left(\mathcal{J}_x g(\bar{x}) d_x \right) & = 0 \end{array}\right\} \tag{39}$$

has $d_x = 0$ as its unique solution.

(ii) Suppose further that $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ is positive semidefinite plus on $\mathrm{app}(\bar{y}, \bar{S})$, i.e.,

$$\left\{ \begin{array}{l} 0 \neq d_x \in \mathrm{app}(\bar{y}, \bar{S}) \implies \langle d_x, \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) d_x \rangle \geq 0 \\ \langle d_x, \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) d_x \rangle = 0 \implies \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) d_x = 0. \end{array} \right.$$

Then the dual nondegeneracy (19) for the dual problem (2) holds at $(\bar{x}, \bar{y}, \bar{S})$ if and only if the primal SSOSC holds at $\bar{x}$.

Proof. (i) The dual nondegeneracy condition (38) holds at $(\bar{x}, \bar{y}, \bar{S})$ if and only if

$$\left\{ \left[\begin{array}{lll} \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) & -\nabla h(\bar{x}) & -\nabla g(\bar{x}) \\ 0 & 0 & I \end{array} \right] \mathcal{Z} \right\}^{\perp} \cap \left[\begin{array}{c} \{0\} \\ \mathrm{lin}\left(T_{\mathcal{S}^n_+}(\bar{S}) \right) \end{array} \right]^{\perp} = \left[\begin{array}{c} 0 \\ 0 \end{array} \right]. \tag{40}$$

On the one hand,

$$(d_x, d_S) \in \left\{ \left[\begin{array}{lll} \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) & -\nabla h(\bar{x}) & -\nabla g(\bar{x}) \\ 0 & 0 & I \end{array} \right] \mathcal{Z} \right\}^{\perp}$$

if and only if

$$\langle d_x, \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) x \rangle - \langle \mathcal{J}_x h(\bar{x}) d_x, y \rangle + \langle -\mathcal{J}_x g(\bar{x}) d_x + d_S, S \rangle = 0 \ \ \forall \, (x, y, S) \in \mathcal{Z}.$$

This holds if and only if

$$\left.\begin{array}{rl} \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) d_x & = 0 \\ \mathcal{J}_x h(\bar{x}) d_x & = 0 \\ d_S - \mathcal{J}_x g(\bar{x}) d_x & = 0. \end{array}\right\} \tag{41}$$

On the other hand,

$$(d_x, d_S) \in \left[\begin{array}{c} \{0\} \\ \mathrm{lin}\left(T_{\mathcal{S}^n_+}(\bar{S}) \right) \end{array} \right]^{\perp}$$

if and only if

$$\langle d_S, Y \rangle = \langle P^T d_S P, P^T Y P \rangle = 0, \quad \forall \, Y \in \mathrm{lin}\left(T_{\mathcal{S}^n_+}(\bar{S}) \right). \tag{42}$$

Taking into consideration of the structure of $\mathrm{lin}\left(T_{\mathcal{S}^n_+}(\bar{S}) \right)$ in (23), (42) holds if and only if

$$P_\gamma^T d_S = 0. \tag{43}$$

Combining (40), (41) and (43) together implies that the linear system (39) has $d_x = 0$ as its unique solution.

(ii) For the necessary part it is enough to show that for any $0 \neq d_x \in \mathrm{app}(\bar{y}, \bar{S})$, the two quantities

$$\langle d_x, \nabla_{xx} L(\bar{x}, \bar{y}, \bar{S}) d_x \rangle \ \text{ and } \ \Upsilon_{g(\bar{x})}(\bar{S}, \mathcal{J}_x g(\bar{x}) d_x)$$

cannot be zero at the same time. Otherwise the fact that the Hessian is positive semidefinite plus on $\mathrm{app}(\bar{y}, \bar{S})$, together with (24) implies that the linear system (39) has $d_x \neq 0$ as its solution, contradicting the result in (i). Hence, the primal SSOSC must hold as one of two quantities must be positive and the other is nonnegative. The sufficient part is obvious by (i). $\square$

We note that the dual SSOSC at $(\bar{x}, \bar{y}, \bar{S})$ implies the nonsingularity of Hessian matrix $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ and that the nonsingularity of $\nabla_{xx} L(\bar{x}, \bar{y}, \bar{S})$ is sufficient to ensure the dual non-degeneracy (38). These results and Prop. 3.3 (ii) yield

Corollary 3.6 *Suppose $(\bar{x}, \bar{y}, \bar{S})$ is a KKT triple of the primal problem (1) and the dual SSOSC holds at $(\bar{x}, \bar{y}, \bar{S})$. Then both the primal nondegeneracy (for the primal problem (1) at $\bar{x}$) and the dual nondegeneracy (for the dual problem (2) at $(\bar{x}, \bar{y}, \bar{S})$) hold.*

3.2 The Convex Case

Some of the results for the general nonlinear case can be strengthened significantly in the convex case, where f, g, h are all convex functions. We present those strengthened results only for the convex QSDP due to its significance [34, 33].

Our first result in this section is the direct consequence of the characterization of the dual nondegeneracy in Prop. 3.5. Note for QSDP (3) that calculations are much simplified because

$$\mathcal{X} = \mathcal{S}^n, \ \nabla_{xx} L(x, y, S) = \mathcal{Q}, \ \mathcal{J}_x h(x) = \mathcal{A}^*, \ \text{and} \ \mathcal{J}_x g(x) = I.$$

and that the Hessian matrix $\mathcal{Q}$ is automatically positive semidefinite plus on $\mathrm{app}(\bar{y}, \bar{S})$.

Theorem 3.7 *(Corollary of Prop. 3.5) Suppose $(\bar{X}, \bar{y}, \bar{S})$ is a KKT triple of the convex QSDP (3). Then the dual nondegeneracy for the dual problem (4) holds at $(\bar{X}, \bar{y}, \bar{S})$ if and only if the primal SSOSC for the primal problem (3) holds at $(\bar{X}, \bar{y}, \bar{S})$.*

We illustrate an application of this result to the path-following method studied in Toh [33] for QSDP (3). Suppose $(\bar{X}, \bar{y}, \bar{S})$ is a KKT triple for the convex QSDP (3). In [33, Def. 3.1], $\bar{S}$ is said to be *dual nondegeneracy* if [4]

$$\mathcal{A}^* \mathbb{R}^m + \mathrm{lin}\, T_{\mathcal{S}^n_+}(\bar{S}) = \mathcal{S}^n. \tag{44}$$

If there exists $\hat{y} \in \mathbb{R}^m$ such that $(\hat{y}, \bar{S})$ is a feasible point of the dual SDP problem (6), then condition (44) is actually the constraint nondegeneracy of (6) at $(\hat{y}, \bar{S})$.

The constraint nondegeneracy for the dual QSDP (4) at $(\bar{X}, \bar{y}, \bar{S})$ takes the following form

$$-\mathcal{Q}(\mathcal{S}^n) + \mathcal{A}^* \mathbb{R}^m + \mathrm{lin}\, T_{\mathcal{S}^n_+}(\bar{S}) = \mathcal{S}^n. \tag{45}$$

[4] It is defined via the dual nondegeneracy of [1], which is equivalent to (44), see [8, Def. 9]

Obviously, condition (44) is stronger than the constraint nondegeneracy (45). Define

$$\mathcal{Q}_\rho := \mathcal{Q} + \rho \mathcal{A}^*\mathcal{A} \quad \text{and} \quad \mathcal{P}_\alpha := P_\alpha \circledast P_\alpha^T,$$

where $\rho > 0$ is a fixed parameter, and P satisfies the spectral decomposition (20). We further define

$$\left(\widetilde{\mathcal{Q}}_\rho\right)_{\alpha\alpha} := \mathcal{P}_\alpha^T \mathcal{Q}_\rho \mathcal{P}_\alpha.$$

The positive definiteness of $\left(\widetilde{\mathcal{Q}}_\rho\right)_{\alpha\alpha}$ plays a very important role in the analysis of the path-following method in [33]. It is obvious that if $\mathcal{Q}$ is positive definite itself, so is $\left(\widetilde{\mathcal{Q}}_\rho\right)_{\alpha\alpha}$. Another sufficient condition is that $(\bar{X}, \bar{y}, \bar{S})$ satisfies condition (44) [33, Remark 3.1]. We now show that this condition can be weakened to the dual constraint nondegeneracy (45).

Proposition 3.8 *Suppose $(\bar{X}, \bar{y}, \bar{S})$ is a KKT triple of the convex QSDP (3). If the dual constraint nondegeneracy (45) holds at $(\bar{X}, \bar{y}, \bar{S})$, then $\left(\widetilde{\mathcal{Q}}_\rho\right)_{\alpha\alpha}$ is positive definite on $\mathcal{S}^r$, where $r = |\alpha|$, the number of positive eigenvalues of $\bar{X}$.*

Proof. We first note that

$$\left(\widetilde{\mathcal{Q}}_\rho\right)_{\alpha\alpha} = \mathcal{P}_\alpha^T \mathcal{Q}_\rho \mathcal{P}_\alpha = \mathcal{P}_\alpha^T \mathcal{Q} \mathcal{P}_\alpha + \rho \mathcal{P}_\alpha^T \mathcal{A}^*\mathcal{A}\mathcal{P}_\alpha.$$

For any $0 \neq U \in \mathcal{S}^r$ with $\mathcal{A}(\mathcal{P}_\alpha(U)) \neq 0$, we obviously have

$$\langle U, \left(\widetilde{\mathcal{Q}}_\rho\right)_{\alpha\alpha}(U)\rangle \geq \rho \|\mathcal{A}(P_\alpha U P_\alpha^T)\|^2 > 0. \tag{46}$$

Therefore, we only need to consider the case $\mathcal{A}(\mathcal{P}_\alpha(U)) = 0$. Let

$$B := [P_\alpha, P_\beta, P_\alpha] \begin{bmatrix} U & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_\alpha^T \\ P_\beta^T \\ P_\gamma^T \end{bmatrix} = P_\alpha U P_\alpha^T = \mathcal{P}_\alpha(U).$$

Apparently, we have

$$P^T B P = \begin{bmatrix} U & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

which means

$$B[P_\beta, P_\gamma] = 0 \quad \text{and} \quad \mathcal{A}(B) = 0.$$

This is sufficient for $B \in \operatorname{app}(\bar{y}, \bar{S})$. We also note that $P_\alpha^T B P_\gamma = 0$, which implies $\Upsilon_{\bar{X}}(\bar{S}, B) = 0$ by (24). The primal SSOSC guarantees that

$$\begin{aligned} 0 &< \langle B, \mathcal{Q}(B)\rangle + \Upsilon_{\bar{X}}(\bar{S}, B) \\ &= \langle B, \mathcal{Q}(B)\rangle \\ &= \langle P_\alpha U P_\alpha^T, \mathcal{Q}(P_\alpha U P_\alpha^T)\rangle \\ &= \langle U, \mathcal{P}_\alpha^T \mathcal{Q} \mathcal{P}_\alpha(U)\rangle \\ &= \langle U, \left(\widetilde{\mathcal{Q}}_\rho\right)_{\alpha\alpha}(U)\rangle. \end{aligned}$$

This, together with (46), proves that $(\widetilde{\mathcal{Q}}_\rho)_{\alpha\alpha}$ is positive definite on $\mathcal{S}^r$ under the primal SSOSC. However, the latter condition is equivalent to the dual constraint nondegeneracy at $(\bar{X}, \bar{y}, \bar{S})$ by Theorem 3.7 $\square$

We note that many results in [33] are based on the positive definiteness of $(\widetilde{\mathcal{Q}}_\rho)_{\alpha\alpha}$ and the strict complementarity condition (i.e. $\bar{X} + \bar{S}$ is positive definite). The following example shows that it may happen that the dual constraint nondegeneracy (45) holds (and hence $(\widetilde{\mathcal{Q}}_\rho)_{\alpha\alpha}$ is positive definite), but neither the SDP dual nondegeneracy (44) nor the strict complementarity condition holds at $(\bar{X}, \bar{y}, \bar{S})$.

Example 3.9 *Consider the convex QSDP* [5]*,*

$$\begin{array}{ll} \min & \frac{1}{2}\|H \circ (X - G)\|^2 \\ s.t. & diag(X) = e \\ & X \in \mathcal{S}^3_+, \end{array}$$

where e is the vector of all ones and

$$H = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad G = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

The optimal solution is

$$\bar{X} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ with the unique Lagrangian multipliers } \bar{y} = \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} \text{ and } \bar{S} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

It is easy to calculate that

$$\mathrm{app}(\bar{y}, \bar{S}) = \left\{ B = \begin{bmatrix} 0 & \tau & \tau \\ \tau & 0 & 0 \\ \tau & 0 & 0 \end{bmatrix} \;\middle|\; \tau \in \mathbb{R} \right\}$$

The primal SSOSC condition holds as

$$\langle B, \mathcal{Q}(B) \rangle = 2\tau^2 > 0 \quad \forall\, \tau \neq 0.$$

Therefore, $(\widetilde{\mathcal{Q}}_\rho)_{\alpha\alpha}$ is positive definite according to Prop. 3.8. However, both the SDP dual nondegeneracy (44) and the strict complementarity condition failed to hold.

Theorem 3.4 says that the positive definiteness of the Hessian and the primal nondegeneracy are sufficient for the dual SSOSC. The converse is also true for the convex QSDP.

Theorem 3.10 *(The converse of Thm. 3.4) Suppose $(\bar{X}, \bar{y}, \bar{S})$ is a KKT triple of the convex QSDP (3). Then the dual SSOSC holds at $(\bar{X}, \bar{y}, \bar{S})$ if and only if $\mathcal{Q}$ is positive definite and the primal nondegeneracy for QSDP (3) holds at $\bar{X}$.*

[5]This type of the problem is called the *nearest correlation matrix problem under the H-weight* in [14].

Proof. The sufficient part has been proved in Theorem 3.4, we only need to show the necessity. Suppose the dual SSOSC holds at $(\bar{X}, \bar{y}, \bar{S})$. $\mathcal{Q}$ is nonsingular. It is hence positive definite because it is positive semidefinite. The primal nondegeneracy follows from Prop. 3.3 (ii). $\square$

The above results provide us with a good understanding of a class of convex quadratic QSDP called the *nearest correlation matrix problem*, a problem from finance and formally studied by Higham [14]:

$$\begin{array}{ll} \min & \frac{1}{2}\|X - G\|^2 \\ \text{s.t.} & \operatorname{diag}(X) = e \\ & X \in \mathcal{S}^n_+, \end{array} \tag{47}$$

where $G \in \mathcal{S}^n$ is given. The constraints in (47) define the set of all correlation matrices. The problem is simply to seek the nearest correlation matrix to G under the Frobenius norm. This problem has been cast as a convex QSDP in [34, 33]. For this case, $\mathcal{Q} \equiv I$.

Let $\bar{X}$ be the unique solution of (47) and $(\bar{y}, \bar{S}) \in \mathbb{R}^n \times \mathcal{S}^n_+$ be the corresponding Lagrangian multipliers (there are only one pair of multipliers). It is known [34] that the primal nondegeneracy holds at $\bar{X}$. The primal SSOSC obviously holds at $(\bar{X}, \bar{y}, \bar{S})$ because $\mathcal{Q} = I$. This further implies by Theorem 3.10 that the dual SSOSC also holds at $(\bar{X}, \bar{y}, \bar{S})$. Furthermore, by Theorem 3.7 we know that the dual nondegeneracy also automatically holds. We summarize these very strong claims in the following result.

Corollary 3.11 *For the nearest correlation matrix problem, it is not only that both the primal and dual nondegeneracy hold, but also that both the primal SSOSC and the dual SSOSC hold at its unique solution.*

The dual SSOSC may not hold for nonconvex QSDP even if the primal SSOSC, primal and dual constraint nondegeneracy hold and $\mathcal{Q}$ is nonsingular. The following example illustrates this point.

Example 3.12 *(non-convex QSDP) Consider the problem in $\mathcal{S}^2$:*

$$\begin{array}{ll} \min & \frac{1}{2}\langle X, (I \circledast C_1)(X)\rangle + \frac{1}{2}\langle X, (I \circledast C_2)(X)\rangle \\ s.t. & diag(X) = e \\ & X \in \mathcal{S}^2_+, \end{array}$$

where

$$C_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}.$$

It is easy to see the problem is equivalent to

$$\begin{array}{ll} \min & \frac{1}{2}(X_{11}^2 - X_{22}^2) - \frac{1}{2}(X_{12}^2 + X_{22}^2) \\ s.t. & X_{11} = 1, \ X_{22} = 1 \\ & X \in \mathcal{S}^2_+. \end{array}$$

There are two optimal solutions, namely

$$\bar{X} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \textit{with Lagrangian multipliers} \ \ \bar{y} = \begin{bmatrix} -\frac{1}{2} \\ \frac{5}{2} \end{bmatrix}, \quad \bar{S} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

and

$$\bar{X} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \textit{with Lagrangian multipliers} \;\; \bar{y} = \begin{bmatrix} -\frac{1}{2} \\ \frac{5}{2} \end{bmatrix}, \quad \bar{S} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

For both solutions, the primal SSOSC (because $\mathrm{app}(\bar{y}, \bar{S}) = \{0\}$*), the primal nondegeneracy (because of the correlation constraint), and the dual nondegeneracy (because of the nonsingularity of* $\mathcal{Q}$ *as we see below) hold. Furthermore,* $\mathcal{Q}$ *is nonsingular because*

$$\mathcal{Q}(X) = \begin{bmatrix} X_{11} & -\frac{1}{2}X_{12} \\ -\frac{1}{2}X_{12} & -2X_{22} \end{bmatrix} = M \circ X \quad \textit{with} \;\; M := \begin{bmatrix} 1 & -1/2 \\ -1/2 & -2 \end{bmatrix}.$$

Therefore, the inverse of $\mathcal{Q}$ *is*

$$\mathcal{Q}^{-1}(X) = \begin{bmatrix} 1 & -2 \\ -2 & -1/2 \end{bmatrix} \circ X.$$

For the first solution, the eigenvectors are

$$P_\alpha = \frac{\sqrt{2}}{2}[1, 1]^T, \quad \beta = \emptyset, \quad \textit{and} \;\; P_\gamma = \frac{\sqrt{2}}{2}[1, -1]^T.$$

The associated set $\mathrm{app}(\bar{X})$ *has the characterization:*

$$\mathrm{app}(\bar{X}) = \left\{ (W, y, H) \in \mathcal{S}^2 \times \mathrm{I\!R}^2 \times \mathcal{S}^2 \;\middle|\; \begin{array}{l} W = \mathit{Diag}(y) + H \\ H_{11} + H_{22} + 2H_{12} = 0 \end{array} \right\},$$

where $\mathit{Diag}(y)$ *is the diagonal matrix given by* y*. Choose*

$$y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad H = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \textit{and} \;\; W = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then $0 \neq (W, y, H) \in \mathrm{app}(\bar{X})$ *and* $P_\alpha^T H P_\gamma = 0$*, implying* $\Upsilon_{\bar{S}}(\bar{X}, H) = 0$*. We can calculate that*

$$\langle W, \; \mathcal{Q}^{-1}(W) \rangle = -4 < 0.$$

Therefore, the dual SSOSC does not hold at the first solution. We may verify through similar calculation that the dual SSOSC also fails to hold at the second solution.

Our last main result is on characterization of nonsingularity of $\partial_B F(\bar{X}, \bar{y}, \bar{S})$, where F for the QSDP takes the form

$$F(X, y, S) = \begin{bmatrix} \mathcal{Q}(X) + \mathcal{A}^*(y) - S \\ \mathcal{A}(x) - b \\ X - \Pi_{\mathcal{S}^n_+}(X - S) \end{bmatrix}.$$

It follows from [8, Lem. 1] that $W \in \partial_B F(\bar{X}, \bar{y}, \bar{S})$ if and only if there exists a $V \in \partial_B \Pi_{\mathcal{S}^n_+}(\bar{A})$ such that

$$W(\Delta X, \Delta y, \Delta S) = \begin{bmatrix} \mathcal{Q}(\Delta X) + \mathcal{A}^*(\Delta y) - \Delta S \\ \mathcal{A}(\Delta X) \\ \Delta X - V(\Delta X - \Delta S) \end{bmatrix} \quad \forall \, (\Delta X, \Delta y, \Delta S) \in \mathcal{S}^n \times \mathrm{I\!R}^m \times \mathcal{S}^n. \tag{48}$$

As observed in [8] there are two special choices of V. One corresponds to $V_{|\beta|} = 0 \in \mathcal{S}^{|\beta|}$ and the other corresponds to $V_{|\beta|} = I \in \mathcal{S}^{|\beta|}$ in (27). They are respectively denoted by V^0 and $V^{\mathcal{I}}$. Then V^0 and $V^{\mathcal{I}}$ give rise to two generalized Jacobians defined in (48) and they are correspondingly denoted by W^0 and $W^{\mathcal{I}}$.

The next result generalizes [8, Prop. 17] from SDP to QSDP.

Proposition 3.13 *Let $(\bar{X}, \bar{y}, \bar{S})$ be a KKT triple of the convex QSDP (3). The following hold.*

(i) The primal nondegeneracy holds for QSDP at $\bar{X}$ if W^0 is nonsingular.

(ii) The dual nondegeneracy holds for QSDP at $(\bar{X}, \bar{y}, \bar{S})$ if $W^{\mathcal{I}}$ is nonsingular.

Proof. (i) can be proved by slightly modifying the first part of the proof [8, Prop. 17]. So we only need to show (ii). By Theorem 3.7, it is enough to prove that the primal SSOSC holds at $(\bar{X}, \bar{y}, \bar{S})$ if $W^{\mathcal{I}}$ is nonsingular.

Let $0 \neq \Delta X \in \mathrm{app}(\bar{y}, \bar{S})$. This means ΔX satisfies

$$\mathcal{A}(\Delta X) = 0 \quad \text{and} \quad P_\gamma^T(\Delta X)[P_\beta, P_\gamma] = 0. \tag{49}$$

Suppose first $P_\gamma^T(\Delta X)P_\alpha \neq 0$. Then $\Upsilon_{\bar{X}}(\bar{S}, \Delta X) > 0$ due to (24) and therefore

$$\langle \Delta X, \mathcal{Q}(\Delta X)\rangle + \Upsilon_{\bar{X}}(\bar{S}, \Delta X) > 0 \tag{50}$$

because $\mathcal{Q}$ is positive semidefinite. Hence the primal SSOSC holds. Now suppose that

$$P_\gamma^T(\Delta X)P_\alpha = 0. \tag{51}$$

We will show that $\mathcal{Q}(\Delta X) \neq 0$, which implies (50) holds as the first term in (50) is positive for this case.

We prove it by assuming that $\mathcal{Q}(\Delta X) = 0$. We will get a contradiction. We note that it is proved in [8, Eq. 65] that conditions (49) and (51) by making use of (27) imply

$$V^{\mathcal{I}}(\Delta X) = \Delta X.$$

Therefore, for $(\Delta y, \Delta S) = (0, 0) \in \mathbb{R}^m \times \mathcal{S}^n$ we have

$$W^{\mathcal{I}}(\Delta X, \Delta y, \Delta S) = \begin{bmatrix} \mathcal{Q}(\Delta X) + \mathcal{A}^*(\Delta y) - \Delta S \\ \mathcal{A}(\Delta X) \\ \Delta X - V^{\mathcal{I}}(\Delta X - \Delta S) \end{bmatrix} = 0,$$

which implies $W^{\mathcal{I}}$ is singular as $\Delta X \neq 0$. This contradiction shows that the dual nondegeneracy holds at $(\bar{X}, \bar{y}, \bar{S})$. $\square$

The following result states some equivalent conditions for the strong regularity of a KKT triple for the convex QSDP (3). It also extends the surprising result of Chan and Sun [8] from SDP to QSDP of equivalence between the nonsingularity of $\partial_B F(\bar{X}, \bar{y}, \bar{S})$ and the nonsingularity of $\partial F(\bar{X}, \bar{y}, \bar{S})$

Theorem 3.14 *Let $(\bar{X}, \bar{y}, \bar{S})$ be a KKT triple of the convex QSDP (3). Then the following conditions are equivalent.*

(i) $(\bar{X}, \bar{y}, \bar{S})$ *is strongly regular of the generalized equation (8).*

(ii) The primal SSOSC holds at $(\bar{X}, \bar{y}, \bar{S})$ *and the primal nondegeneracy holds at* $\bar{X}$.

(iii) Any element in $\partial F(\bar{X}, \bar{y}, \bar{S})$ *is nonsingular.*

(iv) The primal nondegeneracy holds at $\bar{X}$ *and the dual nondegeneracy holds at* $(\bar{X}, \bar{y}, \bar{S})$.

(v) Any element in $\partial_B F(\bar{X}, \bar{y}, \bar{S})$ *is nonsingular.*

Proof. The equivalence between (i), (ii) and (iii) has been proved for the general nonlinear SDP in [31, Thm. 4.1]. It is proved in Theorem 3.7 that the dual constraint nondegeneracy is equivalent to the primal SSOSC. Hence (ii) is equivalent to (iv). Obviously (ii) implies (v). Prop. 3.13 implies (v) $\Longrightarrow$ (iv). Therefore, (ii) $\Longleftrightarrow$ (v) $\square$

We finish this section with the remark that the smoothing Newton method analyzed in [8] in the context of SDP can be similarly studied for the general convex NSDP (1), where f, g, and h are all convex functions of x. All results there have their counterparts for the convex NSDP including the convex QSDP.

4 Comparison with NLP

It is interesting to compare Prop. 3.3(i) with the similar result [13, Eq.2.9] in NLP. The comparison also leads to a new result in NLP which corresponds to Prop. 3.3(ii). The new result Prop. 4.2 says that the dual SOSC of [13] implies the *strict constraint qualification* (see [7, Def. 4.46] for its definition), a necessary condition not known in literature. In NLP, the strict constraint qualification is equivalent to the uniqueness of the Lagrangian multiplier. In contrast, our dual SSOSC implies the *constraint nondegeneracy*, which also implies the uniqueness of the Lagrangian multiplier.

Consider the NLP

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h(x) = 0, \\ & g(x) \le 0, \end{array} \tag{52}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$, $h : \mathbb{R}^n \mapsto \mathbb{R}^m$, and $g : \mathbb{R}^n \mapsto \mathbb{R}^\ell$ are twice continuously differentiable functions on $\mathbb{R}^n$. Suppose $(\bar{x}, \bar{u}, \bar{v})$ is a KKT triple of NLP (52). We first note that the tangent cone T in [13, Eq.2.4] is actually the critical cone $C(\bar{x})$ of (52) at $\bar{x}$, which can be written as (see [7, 5.77])

$$C(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \mathcal{J}_x h(\bar{x}) d = 0, \mathcal{J}_x g_i(\bar{x}) d = 0, \ i \in I_+(\bar{x}, \bar{v}), \ \mathcal{J}_x g_i(\bar{x}) d \le 0, \ i \in I_0(\bar{x}, \bar{v}) \right\}, \tag{53}$$

where

$$I_+(\bar{x}, \bar{v}) := \{ i \mid g_i(\bar{x}) = 0, \ \bar{v}_i > 0 \}, \quad \text{and} \quad I_0(\bar{x}, \bar{v}) := \{ i \mid g_i(\bar{x}) = 0, \ \bar{v}_i = 0 \}.$$

Then [13, Eq.2.9] means that the dual second-order sufficent condition defined in [13, Def.2.1] requires the positive definiteness of the inverse of the Hessian matrix, $\nabla_{xx} L(\bar{x}, \bar{u}, \bar{v})^{-1}$, on the polar cone $C(\bar{x})^- := \{ p \in \mathbb{R}^n \mid \langle p, d \rangle \le 0, d \in C(\bar{x}) \}$.

When specialized to NLP (52), the set $\mathcal{T}_{\bar{x}}$ in (29) reduces to

$$\mathcal{T}_{\bar{x}} = \{d \in \mathbb{R}^n \mid \mathcal{J}_x h(\bar{x})d = 0,\ \mathcal{J}_x g_i(\bar{x})d = 0,\ i \in I(\bar{x})\}\,, \tag{54}$$

where $I(\bar{x}) = \{i \mid g_i(\bar{x}) = 0\}$. The dual SSOSC in Def. 2.4 when applied to NLP (52) implies that the inverse of the Hessian matrix, $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$, is positive definite on the complementary space $\mathcal{T}_{\bar{x}}^{\perp}$. Apparently $\mathcal{T}_{\bar{x}} \subset C(\bar{x})$ by (53) and (54), implying $\mathcal{T}_{\bar{x}}^{\perp} \supset C(\bar{x})^{-}$ due to the fact of the same set of linear functions defining $C(\bar{x})$ and $\mathcal{T}_{\bar{x}}$.

In short, when specialized to NLP (52) our dual SSOSC implies the positive definiteness of $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$ on $(\text{lin}\, C(\bar{x}))^{\perp} \setminus \{0\}$ (note that $\mathcal{T}_{\bar{x}} = \text{lin}\, C(\bar{x})$), whereas the dual SOSC in [13] implies the positive definiteness of $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$ on $\mathcal{W}$, where

$$\mathcal{W} := \left\{ w = \nabla h(\bar{x})y + \nabla g(\bar{x})p \;\middle|\; p_i = 0,\ i \in \bar{I}(\bar{x}),\ p_i \geq 0,\ i \in I_0(\bar{x},\bar{v}),\ (y,p) \neq 0 \right\}, \tag{55}$$

and $\bar{I}(\bar{x}) = \{1,\ldots,\ell\} \setminus I(\bar{x})$. We note that $\mathcal{W} \cup \{0\} = C(\bar{x})^{-}$.

We note that the constraint nondegeneracy of Def. 2.1 corresponds to the linear independence assumption in NLP [29, Example 2.1]. The comparison above and direct application of Prop. 3.3(ii) to NLP imply the following corollary, where we rephrase the dual SSOSC in the context of NLP. The result can also be proved directly.

Corollary 4.1 *Suppose $(\bar{x},\bar{u},\bar{v})$ is a KKT triple of NLP (52). If the inverse of the Hessian matrix (i.e., $\nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}$) exists and*

$$\langle w, \nabla_{xx}L(\bar{x},\bar{u},\bar{v})^{-1}(w)\rangle > 0, \quad \forall\ 0 \neq (w,y,p) \text{ satisfying } w = \nabla h(\bar{x})y + \nabla g_{I(\bar{x})}(\bar{x})p,$$

then the collection of the gradient vectors of the active constraints $\{\nabla h_i(\bar{x}),\ i = 1,\ldots,m$ and $\nabla g_i(\bar{x}),\ i \in I(\bar{x})\}$ is linearly independent. Here $g_{I(\bar{x})} := (g_i)_{i\in I(\bar{x})}$.

The above corollary means that the dual SSOSC implies the linear independence in NLP. In contrast, the dual SOSC of [13] implies a weaker condition, as stated in the following result. This result seems to have not appeared in [13] or elsewhere and hence a proof is included.

Proposition 4.2 *Consider the nonlinear programming problem (52). Let $(\bar{x},\bar{u},\bar{v})$ be its KKT triple. Suppose that the dual second-order sufficient condition in [13] is satisfied at $(\bar{x},\bar{u},\bar{v})$. Then it holds*

$$\begin{bmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{bmatrix} \mathbb{R}^n + \begin{bmatrix} \{0\} \\ T_{\mathbb{R}^\ell_-}(g(\bar{x})) \cap \bar{v}^T \end{bmatrix} = \begin{bmatrix} \mathbb{R}^m \\ \mathbb{R}^\ell \end{bmatrix}. \tag{56}$$

Proof. First, we note that (56) is equivalent to

$$\begin{bmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{bmatrix} \mathbb{R}^n + \begin{bmatrix} \{0\} \\ C_0 \end{bmatrix} = \begin{bmatrix} \mathbb{R}^m \\ \mathbb{R}^\ell \end{bmatrix}, \tag{57}$$

where $C_0 := \left\{p \in \mathbb{R}^\ell \mid p_i \leq 0,\ i \in I_0(\bar{x},\bar{v}),\ \text{and } p_i = 0,\ i \in I_+(\bar{x},\bar{v})\right\}$. This is because

$$T_{\mathbb{R}^\ell_-}(g(\bar{x})) = \left\{p \in \mathbb{R}^\ell \mid p_i \leq 0,\ i \in I(\bar{x})\right\}$$

and

$$\bar{v}^{\perp} = \left\{ p \in \mathbb{R}^{\ell} \;\Big|\; \sum_{i \in I_+(\bar{x},\bar{v})} p_i \bar{v}_i = 0 \right\}.$$

Notice that $\bar{v}_i > 0$ for $i \in I_+(\bar{x}, \bar{v})$, the intersection of the two sets is obviously C_0.

Now suppose (57) does not hold at $\bar{x}$. Certainly, C_0 is a cone. Then by duality results for cones (e.g., [7, Eq.2.31]), we have

$$\left\{ \begin{bmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{bmatrix} \mathbb{R}^n \right\}^{\perp} \cap \begin{bmatrix} \mathbb{R}^m \\ C_0^- \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{58}$$

which means that there exists $0 \neq (y, p) \in \mathbb{R}^m \times \mathbb{R}^{\ell}$ such that it belongs to the left-hand side of (58), i.e.,

$$p \in C_0^- \qquad \text{and} \qquad \langle y, \mathcal{J}_x h(\bar{x}) d\rangle + \langle p, \mathcal{J}_x g(\bar{x}) d\rangle = 0, \;\; \forall\, d \in \mathbb{R}^n,$$

which implies

$$w := \nabla h(\bar{x}) y + \nabla g(\bar{x}) p = 0 \;\text{ and }\; C_0^- = \left\{ p \in \mathbb{R}^{\ell} \mid p_i \geq 0, \; i \in I_0(\bar{x}, \bar{v}), \text{ and } p_i = 0, \; i \in \bar{I}(\bar{x}) \right\}.$$

This means that $0 = w \in \mathcal{W}$ defined in (55) with the corresponding $(y, p) \neq 0$, contradicting the fact that $\nabla L(\bar{x}, \bar{u}, \bar{v})^{-1}$ is positive definite on such points according to [13, Eq.2.9]. This contradiction establishes (57) $\square$

There are a few interesting remarks to make about (56). First, it is actually the *strict constraint qualification* applied to NLP (52). Furthermore, it has a straightforward extension to NSDP (1) and takes the following form.

$$\begin{bmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{bmatrix} \mathcal{X} + \begin{bmatrix} \{0\} \\ T_{\mathcal{S}^n_+}(g(\bar{x})) \cap \bar{S}^{\perp} \end{bmatrix} = \begin{bmatrix} \mathbb{R}^m \\ \mathbb{R}^{\ell} \end{bmatrix}, \tag{59}$$

where $(\bar{x}, \bar{y}, \bar{S})$ is a KKT triple for NSDP (1). The strict constraint qualification (59) ensures the uniqueness of the Lagrangian multipliers of NSDP (1), see [7, Prop. 5.86] and [31, Prop. 3.1].

The second remark is that (56) is equivalent to the qualification [13, (2.10)], which was used there to derive the dual SOSC under the positive definiteness of the Hessian of the Lagrangian. For reasons why they are equivalent, see [7, Remark 4.49]. The last remark we want to make is that the primal constraint nondegeneracy in Thm. 3.4 cannot be replaced by the strict constraint qualification (59), because under this weaker condition, linear equations in (35) may have nonzero solutions. In contrast, the strict constraint qualification and the positive definiteness of the Hessian matrix are sufficient in ensuring the dual SOSC in NLP [13, Thm. 2.3]. While the difference between the NSDP and the NLP in this case once again illustrates the importance of the *sigma*-term involved in second-order optimality conditions in NSDP, we would like to point out that the SSOSC used in this paper is slightly stronger than the SOSC used in [13].

5 Conclusion

In SDP, there exists a perfect duality between the primal/dual nondegeneracy and the dual/primal strong second-order sufficient condition (SSOSC) [8]. In this paper, we investigate to what extent such a perfect duality exists for nonlinear SDP.

For the convex QSDP, the primal SSOSC is nothing more than the dual nondegeneracy (as is in SDP). But, the dual SSOSC is a little more than just being the primal nondegeneracy. It also implies the strict convexity of the problem (i.e., $\mathcal{Q}$ is positive definite). We also show that the primal and dual nondegeneracy are already enough to ensure the equivalence between the nonsingularity of $\partial_B F$ and that of ∂F at its unique KKT triple of the convex QSDP. Above all, the nearest correlation matrix problem satisfies not only the primal and dual SSOSC but also the primal and dual nondegeneracy at its unique solution.

For the general nonlinear SDP, the dual SSOSC implies both the primal and the dual nondegeneracy. Moreover, it ensures that the inverse of the Hessian of the Lagrangian function exists at the concerned KKT triple and the inverse is positive definite on the normal space, generalizing some results in nonlinear programming [13]. Various relationships between the primal SSOSC, the dual SSOSC and the primal and the dual nondegeneracy are studied.

Acknowledgement. The author is grateful to the associate editor and the two referees for their constructive comments, which not only clarify confusing usage of the constraint nondegeneracy and the *sigma*-term in the first version of the paper, but also suggest interesting ways to address some raised issues concerning the comparison of some results in NSDP and NLP. This has led to the addition of Section 4.

References

[1] F. Alizadeh, J.-P. A. Haeberly, and M.L. Overton, *Complementarity and nondegenracy in semidefinite programming*, Math. Programming 77 (1997), pp. 111-128.

[2] V.I. Arnold, *On matrices depending on parameters*, Russian Mathematical Surveys 26 (1971), pp. 29–43.

[3] J.F. Bonnans, R. Cominetti and A. Shapiro, *Sensitivity analysis of optimization problems under second order regularity constraints*, Math. Oper. Res. **23** (1998) 803–832.

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

[5] J.F. Bonnans and H. Ramírez C., *Strong regularity of semidefinite programming problems*, Techical Report, Department of Mathematical Engineering, Universida de Chile, 2005.

[6] J.F. Bonnans and A. Shapiro, *Nondegeneracy and quantitative satbility of parameterized optimization problems with multiple solutions*, SIAM J. Optimization 8 (1998), pp. 940–946.

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

[8] Z.X. Chan and D. Sun, *Constraint nondegeneracy, strong regularity and nonsingulairty in semidefinite programming.* To appear in SIAM J. Optim..

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

[10] R. Cominetti, *Metric regularity, tangent sets and second order optimality conditions*, J. of Applied Mathematics and Optimization 21 (1990), pp. 265–287.

[11] B.C. Eaves, *On the basic theorem for complemenarity*, Math. Programming 1 (1971), pp. 68–75.

[12] M.L. Flegel and C. Kanzow, *Equivalence of two nondegeneracy conditions for semidefinite programs*, Journal of Optimization Theory and Applications 135 (2007), pp. 381-397.

[13] O. Fujiwara, S.-P. Han, and O.L. Mangasarian, *Local dulaity of nonlinear programs*, SIAM J. Control and Optimization 22 (1984), pp. 162–169.

[14] N.J. Higham, *Computing the nearest correlation matrix – a problem from finance*, IMA J. Numer. Analysis 22 (2002), pp. 329–343.

[15] C. Kanzow and C. Nagel, *Semidefinite programs: new search directions, smoothing-type methods, and numerical results*, SIAM J. Optimization 13 (2002), pp. 1–23.

[16] H. Kawasaki, *An envelope-like effect of infinitely many inequality constraints on second order necessary conditions for minimization problems*, Math. Programming 41 (1988), pp. 73–96.

[17] D.G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, 1973.

[18] T. Matsumoto, *An algebraic condition equivalent to strong stability of stationary solutions of nonlinear positive semidefinite programms*, SIAM J. Optimization 16 (2005), pp. 452–470.

[19] J.W. Nie and Y.X. Yuan, *A predictor-corrector algorithm for QSDP combining Dikin-type and Newton centering steps*, Annals of Operations Research 103 (2001), pp. 115–133.

[20] J.-S. Pang, D. Sun and J. Sun, *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems*, Math. Oper. Res. 28 (2003), pp. 39–63.

[21] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Math. Oper. Res. 18 (1993), pp. 227–244.

[22] S.M. Robinson, *First order conditions for general nonlinear optimization*, SIAM J. Appl. Math. 30 (1976), pp. 597–607.

[23] S.M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res. 5 (1980), pp. 43–62.

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

[25] S.M. Robinson, *Constriant nondegeneracy in variational analysis*, Math. Oper. Res. 28 (2003), pp. 201–232.

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

[27] A. Shapiro, *First and second order analysis of nonlinear semidefinite programms*, Math. Programming 77 (1997), pp. 301–320.

[28] A. Shapiro, *Duality, optimality conditions and perturbation analysis*, in: Semidefinite Programming and Applications Handbook, R. Saigal and L. Vandenberghe and H.Wolkowicz, Eds., KluwerAcademic Publishers, Boston, pp. 67-92, 2000.

[29] A. Shapiro, *Sensitivity anlaysis of generalized equations*, J. Math. Sci. 115 (2003), pp. 2554–2565.

[30] A. Shapiro and M.K.H. Fan, *On eigenvalue optimization*, SIAM J. Optimization 5 (1995), pp. 552–569.

[31] D. Sun, *The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*, Math. Oper. Res. 31 (2006), pp. 761–776.

[32] D. Sun and J. Sun, *Semismooth matrix valued functions*, Math. Oper. Res. 27 (2002), pp. 150–169.

[33] K.C. Toh, *An inexact primal-dual path-following algorithm for convex quadratic SDP*. Math. Programming 112 (2008), 221–254.

[34] K.C. Toh, R.H. Tütüncü, and M.J. Todd, *Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems*, Pacific J. Optimization 3 (2007), pp. 135–164.

[35] P. Wolfe, *A duality theorem for nonlionear programming*, Quart. Appl. Math. 19 (1961), pp. 239–244.