A COMPARISON OF THREE NONDEGENERACY CONDITIONS FOR SEMIDEFINITE PROGRAMS

Michael L. Flegel and Christian Kanzow

Preprint 258 April 2005

University of Würzburg
Institute of Applied Mathematics and Statistics
Am Hubland
97074 Würzburg
Germany

e-mail: flegel@mathematik.uni-wuerzburg.de
kanzow@mathematik.uni-wuerzburg.de

April 25, 2005

Abstract. Nondegeneracy assumptions are often needed in order to prove local fast convergence of suitable algorithms as well as in the sensitivity analysis for semidefinite programs. Here we investigate the precise relation between three nondegeneracy concepts introduced in the literature. The nondegeneracy conditions considered here are called KSS-, AHO-, and KN-nondegeneracy since they were first introduced by Kojima, Shida, and Shindoh [*Mathematical Programming*, 80 (1998), pp. 129–160], Alizadeh, Haeberly, and Overton [*Mathematical Programming*, 77 (1997), pp. 111–128], and Kanzow and Nagel [*SIAM Journal on Optimization*, to appear], respectively. While all three conditions are equivalent if strict complementarity holds at a solution of the semidefinite program, we show that KSS-nondegeneracy cannot hold without this assumption, whereas the other two nondegeneracy conditions are still equivalent even without strict complementarity. This result provides considerable new insight into both the AHO- and the KN-nondegeneracy conditions since the two corresponding definitions are completely different in nature.

Key Words. Semidefinite programs, Nondegeneracy conditions, Strict complementarity, Uniqueness of solutions.

1 Introduction

We are interested in the semidefinite program

$$\min C \bullet X \quad \text{s.t.} \quad A_i \bullet X = b_i \ \ (i = 1, \ldots, m), \ \ X \succeq 0$$

for given symmetric matrices $A_i, C \in \mathcal{S}^n$ and scalars $b_i \in \mathbb{R}\,(i = 1, \ldots, m)$. Under a Slater-type constraint qualification, this semidefinite program is equivalent to the following optimality conditions:

$$\begin{array}{rcl} \sum_{i=1}^{m} \lambda_i A_i + S &=& C, \\ A_i \bullet X &=& b_i \quad (i = 1, \ldots, m), \\ X \succeq 0, \ S \succeq 0, \ XS &=& 0. \end{array} \tag{1}$$

These optimality conditions are the basis of several algorithms for the solution of semidefinite programs. This includes both the class of interior-point methods (see, e.g., [3, 7, 13, 14, 16, 18, 21, 23]) and the class of smoothing-type methods (see, e.g., [4, 10, 11, 12, 19, 20, 22]). Of central importance for the local fast convergence of these methods (see, e.g., [3, 4, 6, 10, 12]) and for some sensitivity results for semidefinite programs (see, e.g., [2, 3, 5, 15, 17]) are some nondegeneracy conditions.

We are currently aware of three different nondegeneracy conditions, namely those introduced by

- Kojima, Shida, and Shindo [13]
- Alizadeh, Haeberly, and Overton [2]
- Kanzow and Nagel [12].

We will refer to these conditions as KSS-, AHO–, and KN-nondegeneracy, respectively. These nondegeneracy conditions, their implications, and their relation to each other are the subject of this paper.

To this end, we first state some preliminary results in Section 2. We then give a detailed investigation of the KSS-nondegeneracy condition in Section 3. In particular, we show that a KSS-nondegenerate solution of the optimality conditions (1) automatically satisfies strict complementarity. Moreover, the X- and S-components of a KSS-nondegenerate solution are unique. Note that we do not assume linear independence of the matrices A_i in order to prove these results. Hence they may be viewed as generalizations of some corresponding results given in [12].

We then consider the AHO-nondegeneracy condition in Section 4. We show that a KSS-nondegenerate solution of the optimality conditions (1) is automatically an AHO-nondegenerate solution, and that the converse is also true under strict complementarity. Hence the concepts of KSS-nondegeneracy and AHO-nondegeneracy are equivalent under strict complementarity. It seems that this result has been noted before (under the additional assumption that the A_i are linearly independent) by Haeberly since the authors of [13] give a corresponding note in their paper and refer to a private communication by Haeberly. However, we are not aware of an explicit reference where this result may be found.

We then turn to a discussion of the KN-nondegeneracy condition in Section 5. We admit that the definition of KN-nondegeneracy is not easy to understand, which has to do with the fact that it originates in a nonsmooth reformulation of the optimality conditions in [12]. In particular, the definition of KN-nondegeneracy is completely different from the definition of AHO-nondegeneracy. Nevertheless (and somewhat surprisingly), we show in Section 5 that these nondegeneracy conditions are equivalent. Hence we get a much simpler geometric view of the KN-nondegeneracy concept.

The notation used in the manuscript is standard for papers on semidefinite programs. In particular, we denote by $\mathbb{R}^n$ the n-dimensional real vector space, by $\mathbb{R}^{n\times n}$ the set of all real matrices of dimension $n \times n$, and by $\mathcal{S}^n$ the subset of all symmetric matrices of dimension $n \times n$. Moreover, $A \bullet B := \text{trace}(AB^T)$ for two (not necessarily symmetric) matrices $A, B \in \mathbb{R}^{n\times n}$ is the standard scalar product in $\mathbb{R}^{n\times n}$. This scalar product induces the Frobenius norm $\|A\|_F := (A \bullet A)^{1/2} = (\sum_{i,j=1}^n a_{ij}^2)^{1/2}$. The symbol $A \succeq 0$ means that A is symmetric positive semidefinite, and $A^{1/2}$ denotes the symmetric positive semidefinite square root of a given matrix $A \succeq 0$.

2 Preliminaries

The aim of this section is twofold: On the one hand, we want to introduce the notation that will be used throughout this paper. On the other hand, we state some preliminary results which will facilitate some of the subsequent proofs.

Let $(X^*, \lambda^*, S^*) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ be a solution of the optimality conditions (1). Then, in particular, we have $X^*S^* = 0$. Hence these two matrices commute. Therefore, they have a simultaneous spectral decomposition (see, e.g., [8]), i.e., there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ and diagonal matrices $D^X, D^S \in \mathbb{R}^{n\times n}$ such that

$$X^* = QD^XQ^T \quad \text{and} \quad S^* = QD^SQ^T. \tag{2}$$

Let $\lambda_i(X^*)$ and $\lambda_i(S^*)$ denote the eigenvalues of X^* and S^*, respectively. Then

$$\begin{aligned} & X^* \succeq 0,\ S^* \succeq 0,\ X^*S^* = 0 \\ \Longleftrightarrow\ & \lambda_i(X^*) \geq 0,\ \lambda_i(S^*) \geq 0,\ \lambda_i(X^*)\lambda_i(S^*) = 0\ \ \forall i = 1, \ldots, m. \end{aligned}$$

Hence the three index sets

$$\begin{aligned} \alpha &:= \big\{ i \,\big|\, \lambda_i(X^*) > 0, \lambda_i(S^*) = 0 \big\}, \\ \beta &:= \big\{ i \,\big|\, \lambda_i(X^*) = 0, \lambda_i(S^*) = 0 \big\}, \\ \gamma &:= \big\{ i \,\big|\, \lambda_i(X^*) = 0, \lambda_i(S^*) > 0 \big\} \end{aligned} \tag{3}$$

form a partition of the whole set $\{1, \ldots, n\}$. Note that

$$p := |\alpha| \text{ is the rank of } X^*, \quad \text{and} \quad q := |\gamma| \text{ is the rank of } S^*,$$

whereas β is the degenerate index set. In particular, the solution (X^*, λ^*, S^*) satisfies strict complementarity (i.e., $X^* + S^* \succ 0$) if and only if $\beta = \emptyset$. In this paper, however, we are mainly interested in the case $\beta \neq \emptyset$.

Using the three index sets α, β, and γ, we will, without loss of generality, partition the matrices Q, D^X, D^S from the simultaneous spectral decomposition (2) in the following way:

$$
\begin{aligned}
Q &= \begin{pmatrix} Q_\alpha & Q_\beta & Q_\gamma \end{pmatrix} \quad \text{with } Q_\alpha \in \mathbb{R}^{n\times|\alpha|}, Q_\beta \in \mathbb{R}^{n\times|\beta|}, Q_\gamma \in \mathbb{R}^{n\times|\gamma|}, \\
D^X &= \begin{pmatrix} D_\alpha^X & & \\ & 0 & \\ & & 0 \end{pmatrix} \quad \text{with } D_\alpha^X := \operatorname{diag}\big(\lambda_1(X^*), \dots, \lambda_p(X^*)\big) \in \mathbb{R}^{p\times p}, \\
D^S &= \begin{pmatrix} 0 & & \\ & 0 & \\ & & D_\gamma^S \end{pmatrix} \quad \text{with } D_\gamma^S := \operatorname{diag}\big(\lambda_{n-q+1}(S^*), \dots, \lambda_n(S^*)\big) \in \mathbb{R}^{q\times q}.
\end{aligned}
\tag{4}
$$

Furthermore, given two matrices $\Delta X, \Delta S \in \mathcal{S}^n$, we will frequently use the related matrices

$$\widetilde{\Delta X} := Q^T \Delta X Q \quad \text{and} \quad \widetilde{\Delta S} := Q^T \Delta S Q, \tag{5}$$

where Q denotes the orthogonal matrix from (2). We will partition these symmetric matrices in the following way:

$$\widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\beta} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}_{\alpha\beta}^T & \widetilde{\Delta X}_{\beta\beta} & \widetilde{\Delta X}_{\beta\gamma} \\ \widetilde{\Delta X}_{\alpha\gamma}^T & \widetilde{\Delta X}_{\beta\gamma}^T & \widetilde{\Delta X}_{\gamma\gamma} \end{pmatrix}, \quad \widetilde{\Delta S} = \begin{pmatrix} \widetilde{\Delta S}_{\alpha\alpha} & \widetilde{\Delta S}_{\alpha\beta} & \widetilde{\Delta S}_{\alpha\gamma} \\ \widetilde{\Delta S}_{\alpha\beta}^T & \widetilde{\Delta S}_{\beta\beta} & \widetilde{\Delta S}_{\beta\gamma} \\ \widetilde{\Delta S}_{\alpha\gamma}^T & \widetilde{\Delta S}_{\beta\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}. \tag{6}$$

Note that this notation simplifies considerably if $\beta = \emptyset$.

In order to reformulate linear systems involving matrix variables in the usual matrix-vector format, we need to transfrom matrices into vectors. For a general (not necessarily symmetric) matrix $A \in \mathbb{R}^{n\times n}$, this can be done by using the mapping $\operatorname{vec} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ defined by

$$\operatorname{vec}(A) := \big(a_{11}, a_{21}, \dots, a_{n1}, a_{12}, a_{22}, \dots, a_{n2}, \dots, a_{nn}\big)^T \in \mathbb{R}^{n^2},$$

i.e., vec stacks the columns of A into a vector of length n^2, see [9]. For a symmetric matrix, we are not interested in all entries of A. It suffices to consider the lower triangular part of A, and the corresponding transformation can be done using the mapping $\operatorname{svec} : \mathcal{S}^n \to \mathbb{R}^{\frac{n(n+1)}{2}}$ defined by

$$\operatorname{svec}(A) := \big(a_{11}, \sqrt{2}a_{21}, \dots, \sqrt{2}a_{n1}, a_{22}, \sqrt{2}a_{32}, \dots, \sqrt{2}a_{n2}, \dots, a_{nn}\big)^T \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

The reason for the factor $\sqrt{2}$ in front of the nondiagonal elements is due to the fact that this is consistent with the inner product, i.e.,

$$A \bullet B = \operatorname{svec}(A)^T \operatorname{svec}(B) \quad \forall A, B \in \mathcal{S}^n. \tag{7}$$

Having introduced vec and svec, the next question is how an ordinary matrix product can be expressed in terms of vec and svec. To this end, we define the *Kronecker product* of two (not necessarily symmetric) matrices $G, K \in \mathbb{R}^{n\times n}$ by

$$G \otimes K := \big[g_{ij}K\big] = \begin{pmatrix} g_{11}K & \cdots & g_{1n}K \\ \vdots & \ddots & \vdots \\ g_{n1}K & \cdots & g_{nn}K \end{pmatrix} \in \mathbb{R}^{n^2\times n^2},$$

see [9]. Then it can easily be verified that

$$\big(G \otimes K\big)\,\mathrm{vec}(H) = \mathrm{vec}\,\big(KHG^T\big) \qquad (H \in \mathbb{R}^{n\times n}).$$

Similarly, we define the *symmetric Kronecker product* by

$$\big(G \otimes_s K\big)\,\mathrm{svec}(H) := \frac{1}{2}\,\mathrm{svec}\,\big(KHG^T + GHK^T\big) \quad (H \in \mathcal{S}^n). \tag{8}$$

Some properties of the symmetric Kronecker product are summarized in the following result. The proofs of these properties are elementary and can be found in [3, 21].

Lemma 2.1 *The symmetric Kronecker product $\otimes_s$ defined by* (8) *has the following properties:*

(a) $G \otimes_s K = K \otimes_s G$.

(b) $\big(G \otimes_s K\big)^T = G^T \otimes_s K^T$.

(c) $\big(G \otimes_s K\big)\big(H \otimes_s L\big) = \frac{1}{2}\big(GH \otimes_s KL + GL \otimes_s KH\big)$.

(d) Let G and K be two commuting symmetric matrices with eigenvalues $\lambda_i\,(i = 1,\dots,n)$ and $\mu_j\,(j = 1,\dots,n)$, respectively, Then the $n(n+1)/2$ scalars $\frac{1}{2}\big(\lambda_i\mu_j + \lambda_j\mu_i\big)\,(1 \le j \le i \le n)$ are the eigenvalues of $G \otimes_s K$.

We are now in the position to state the following result. A similar statement may be found in [6, Lemma 2.3], where, however, strict complementarity is assumed. Also [14, Proof of Lemma 6.2] presents a related result, but again under the additional assumption that strict complementarity holds. In the form stated here, this result has implicitly been used in [12, Proof of Lemma 4.2].

Lemma 2.2 *Let $(X^*, \lambda^*, S^*) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ be a solution of the optimality conditions* (1)*. Then the following two statements are equivalent for any two matrices $(\Delta X, \Delta S) \in \mathcal{S}^n \times \mathcal{S}^n$:*

(a) $X^*\Delta S + \Delta X S^* = 0$.

(b) $\big(I \otimes_s X^*\big)\,\mathrm{svec}(\Delta S) + \big(I \otimes_s S^*\big)\,\mathrm{svec}(\Delta X) = 0$.

Proof. (a) $\Longrightarrow$ (b): First assume that $X^*\Delta S + \Delta X S^* = 0$. Taking the transpose gives $\Delta S X^* + S^*\Delta X = 0$. Adding these two equations, we obtain

$$X^*\Delta S + \Delta S X^* + S^*\Delta X + \Delta X S^* = 0. \tag{9}$$

Applying $\frac{1}{2}$svec on both sides and using (8), we obtain statement (b).

(b) $\Longrightarrow$ (a): In view of (8), statement (b) is equivalent to (9). Let us consider the simultaneous spectral decomposition (2) of X^* and S^*. Then (9) may be rewritten as

$$QD^XQ^T\Delta S + \Delta S QD^XQ^T + QD^SQ^T\Delta X + \Delta X QD^SQ^T = 0.$$

Premultiplying this equation by Q^T, postmultiplying it by Q, and introducing the matrices $\widetilde{\Delta X}, \widetilde{\Delta S}$ from (5), we obtain

$$D^X \widetilde{\Delta S} + \widetilde{\Delta S} D^X + D^S \widetilde{\Delta X} + \widetilde{\Delta X} D^S = 0. \tag{10}$$

Now let us use the partitionings from (6) (for $\widetilde{\Delta X}, \widetilde{\Delta S}$) and (4) (for Q, D^X, D^S). Then an easy calculation shows that (10) becomes

$$\begin{pmatrix} D^X_\alpha \widetilde{\Delta S}_{\alpha\alpha} + \widetilde{\Delta S}_{\alpha\alpha} D^X_\alpha & D^X_\alpha \widetilde{\Delta S}_{\alpha\beta} & D^X_\alpha \widetilde{\Delta S}_{\alpha\gamma} + \widetilde{\Delta X}_{\alpha\gamma} D^S_\gamma \\ \widetilde{\Delta S}^T_{\alpha\beta} D^X_\alpha & 0 & \widetilde{\Delta X}_{\beta\gamma} D^S_\gamma \\ \widetilde{\Delta S}^T_{\alpha\gamma} D^X_\alpha + D^S_\gamma \widetilde{\Delta X}^T_{\alpha\gamma} & D^S_\gamma \widetilde{\Delta X}^T_{\beta\gamma} & D^S_\gamma \widetilde{\Delta X}_{\gamma\gamma} + \widetilde{\Delta X}_{\gamma\gamma} D^S_\gamma \end{pmatrix} = 0. \tag{11}$$

Since D^X_α and D^S_γ are positive definite, this implies $\widetilde{\Delta S}_{\alpha\beta} = 0$ and $\widetilde{\Delta X}_{\beta\gamma} = 0$. Moreover, $D^X_\alpha \widetilde{\Delta S}_{\alpha\alpha} + \widetilde{\Delta S}_{\alpha\alpha} D^X_\alpha = 0$ and $D^S_\gamma \widetilde{\Delta X}_{\gamma\gamma} + \widetilde{\Delta X}_{\gamma\gamma} D^S_\gamma = 0$ are two homogeneous Lyapunov equations, so that the positive definiteness of D^X_α and D^S_γ also imply $\widetilde{\Delta S}_{\alpha\alpha} = 0$ and $\widetilde{\Delta X}_{\gamma\gamma} = 0$, see, e.g., [9]. Hence the matrices $\widetilde{\Delta X}$ and $\widetilde{\Delta S}$ simplify to

$$\widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\beta} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}^T_{\alpha\beta} & \widetilde{\Delta X}_{\beta\beta} & 0 \\ \widetilde{\Delta X}^T_{\alpha\gamma} & 0 & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{\Delta S} = \begin{pmatrix} 0 & 0 & \widetilde{\Delta S}_{\alpha\gamma} \\ 0 & \widetilde{\Delta S}_{\beta\beta} & \widetilde{\Delta S}_{\beta\gamma} \\ \widetilde{\Delta S}^T_{\alpha\gamma} & \widetilde{\Delta S}^T_{\beta\gamma} & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}.$$

Exploiting this structure and using (4), (11), a direct computation shows that $D^X \widetilde{\Delta S} + \widetilde{\Delta X} D^S = 0$. Pre- and postmultiplication with Q and Q^T, respectively, and recalling the definitions (2), (5), we directly obtain statement (a). □

We next introduce the matrix

$$\mathcal{A} := \big(\text{svec}(A_1), \dots, \text{svec}(A_m)\big)^T \in \mathbb{R}^{m \times \frac{n(n+1)}{2}}, \tag{12}$$

i.e., the rows of $\mathcal{A}$ are given by the matrices A_i viewed as long vectors. Using this notation, we obviously have

$$\begin{array}{rcl} \sum_{i=1}^m \Delta\lambda_i A_i + \Delta S = 0 & \iff & \mathcal{A}^T \Delta\lambda + \text{svec}(\Delta S) = 0, \\ A_i \bullet \Delta X = 0 \;\; \forall i = 1, \dots, m & \iff & \mathcal{A}\,\text{svec}(\Delta X) = 0 \end{array} \tag{13}$$

for any triple $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$. In our subsequent results, we usually prefer the matrix-vector form of these two equivalent statements.

Let

$$r := \text{rank}(\mathcal{A}),$$

so that the dimension of the space $\text{span}\{A_1, \dots, A_m\}$ is equal to r. Note that we do not assume anywhere in this paper that the matrix $\mathcal{A}$ has full rank, i.e., we do not assume that the matrices $A_1, \dots, A_m$ are linearly independent. Hence, the following result will be quite useful for our later analysis.

Lemma 2.3 *Let* $(X^*, \lambda^*, S^*) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ *be a solution of the optimality conditions* (1). *Furthermore, let us define*

$$\mathcal{G} := I \otimes_s S^* \quad \text{and} \quad \mathcal{H} := I \otimes_s X^*. \tag{14}$$

Then the following two statements are equivalent:

(a) There exists a triple $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ *satisfying the system*

$$\mathcal{A}^T \Delta\lambda + \operatorname{svec}(\Delta S) = 0, \;\; \mathcal{A}\operatorname{svec}(\Delta X) = 0, \;\; \mathcal{G}\operatorname{svec}(\Delta X) + \mathcal{H}\operatorname{svec}(\Delta S) = 0. \tag{15}$$

(b) For any submatrix $\bar{\mathcal{A}}$ *of* $\mathcal{A}$ *consisting of* r *linearly independent rows of* $\mathcal{A}$*, there exists a vector* $\overline{\Delta\lambda} \in \mathbb{R}^r$ *such that* $(\Delta X, \overline{\Delta\lambda}, \Delta S)$ *satisfies the system*

$$\bar{\mathcal{A}}^T \overline{\Delta\lambda} + \operatorname{svec}(\Delta S) = 0, \;\; \bar{\mathcal{A}}\operatorname{svec}(\Delta X) = 0, \;\; \mathcal{G}\operatorname{svec}(\Delta X) + \mathcal{H}\operatorname{svec}(\Delta S) = 0. \tag{16}$$

Proof. (a) $\Longrightarrow$ (b): Let $(\Delta X, \Delta\lambda, \Delta S)$ satisfy (15). Furthermore, let $\bar{\mathcal{A}}$ be a submatrix of $\mathcal{A}$ consisting of r linearly independent rows. Let $J \subseteq \{1, \dots, m\}$ be the corresponding index set such $\bar{\mathcal{A}}$ consists of the rows from $\mathcal{A}$ belonging to the index set J. Since the matrices $A_i (i \notin J)$ are linearly dependent from the set $\{A_i \,|\, i \in J\}$, there exists a vector $\overline{\Delta\lambda} \in \mathbb{R}^r$ such that $\sum_{i=1}^m \Delta\lambda_i A_i = \sum_{i \in J} \overline{\Delta\lambda}_i A_i$. Using the same argument, we see that

$$A_i \bullet \Delta X = 0 \;\; (i = 1, \dots, m) \Longleftrightarrow A_i \bullet \Delta X = 0 \;\; (i \in J). \tag{17}$$

Hence the triple $(\Delta X, \overline{\Delta\lambda}, \Delta S)$ satisfies (16).

(b) $\Longrightarrow$ (a): Let $(\Delta X, \overline{\Delta\lambda}, \Delta S)$ satisfy (16), and let J be the index set having the same meaning as in the first part of the proof. Then define a vector $\Delta\lambda \in \mathbb{R}^m$ in such a way that $\Delta\lambda_i = 0$ for $i \notin J$, whereas $\Delta\lambda_i$ is equal to the corresponding component of $\overline{\Delta\lambda}$ for all $i \in J$. Then we have $\mathcal{A}^T \Delta\lambda + \operatorname{svec}(\Delta S) = \bar{\mathcal{A}}^T \overline{\Delta\lambda} + \operatorname{svec}(\Delta S) = 0$. Taking into account (17), we see that $(\Delta X, \Delta\lambda, \Delta S)$ satisfies (15). $\square$

3 The KSS-Nondegeneracy Condition

We begin with the definition of a nondegeneracy assumption introduced by Kojima, Shida, and Shindoh [13] and which we therefore call KSS-nondegeneracy in order to distinguish it from other nondegeneracy conditions to be introduced in subsequent sections.

Definition 3.1 *A solution* (X^*, λ^*, S^*) *of the optimality conditions* (1) *is called* KSS-nondegenerate *if the following implication holds for any triple* $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$*:*

$$\left.\begin{array}{rcl} \sum_{i=1}^m \Delta\lambda_i A_i + \Delta S &=& 0, \\ A_i \bullet \Delta X &=& 0 \;\; (i = 1, \dots, m), \\ X^* \Delta S + \Delta X S^* &=& 0. \end{array}\right\} \Longrightarrow \left\{ \begin{array}{l} \Delta X = 0, \\ \Delta S = 0. \end{array}\right.$$

Recalling the definition of $\mathcal{A}$ from (12), using (13) and applying Lemma 2.2, we obtain the following reformulation of a KSS-nondegenerate solution in the more standard matrix-vector notation.

Lemma 3.2 *Let (X^*, λ^*, S^*) be a solution of the optimality conditions (1). Then (X^*, λ^*, S^*) is KSS-nondegenerate if and only if the following implication holds for any triple $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$:*

$$\left.\begin{array}{l} \mathcal{A}^T \Delta\lambda + \mathrm{svec}(\Delta S) = 0, \\ \mathcal{A}\,\mathrm{svec}(\Delta X) = 0, \\ \big(I \otimes_s X^*\big)\,\mathrm{svec}(\Delta S) + \big(I \otimes_s S^*\big)\,\mathrm{svec}(\Delta X) = 0 \end{array}\right\} \Longrightarrow \left\{ \begin{array}{l} \mathrm{svec}(\Delta X) = 0, \\ \mathrm{svec}(\Delta S) = 0. \end{array}\right.$$

Another equivalent formulation of the KSS-nondegeneracy condition is given in our next result. Note that this result is similar to a characterization obtained in [6, 15] (under the linear independence of the matrices A_i) for an AHO-nondegenerate solution, cf. the corresponding notes at the end of the next section.

Lemma 3.3 *Let (X^*, λ^*, S^*) be a solution of the optimality conditions (1). Furthermore, let $r := \mathrm{rank}(\mathcal{A})$ be the rank of the matrix $\mathcal{A}$ from (12). Then (X^*, λ^*, S^*) is KSS-nondegenerate if and only if, for any submatrix $\bar{\mathcal{A}}$ consisting of r linearly independent rows of $\mathcal{A}$, the matrix*

$$M_{\bar{\mathcal{A}}} := \begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{I} \\ \bar{\mathcal{A}} & 0 & 0 \\ \mathcal{G} & 0 & \mathcal{H} \end{pmatrix} \tag{18}$$

is nonsingular, where $\mathcal{G}$ and $\mathcal{H}$ are defined in (14).

Proof. First assume that (X^*, λ^*, S^*) is KSS-nondegenerate. Let $\bar{\mathcal{A}}$ be an arbitrary submatrix of $\mathcal{A}$ consisting of r linearly independent rows of $\mathcal{A}$. Let $(\Delta X, \overline{\Delta\lambda}, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^r \times \mathcal{S}^n$ be any triple such that

$$\begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{I} \\ \bar{\mathcal{A}} & 0 & 0 \\ \mathcal{G} & 0 & \mathcal{H} \end{pmatrix} \begin{pmatrix} \mathrm{svec}(\Delta X) \\ \overline{\Delta\lambda} \\ \mathrm{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{19}$$

In view of Lemma 2.3, this is equivalent to the existence of $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ satisfying

$$\begin{aligned} \mathcal{A}^T \Delta\lambda + \mathrm{svec}(\Delta S) &= 0, \\ \mathcal{A}\,\mathrm{svec}(\Delta X) &= 0, \\ \mathcal{G}\,\mathrm{svec}(\Delta X) + \mathcal{H}\,\mathrm{svec}(\Delta S) &= 0. \end{aligned}$$

Since (X^*, λ^*, S^*) is KSS-nondegenerate by assumption, we obtain that $\mathrm{svec}(\Delta X) = 0$ and $\mathrm{svec}(\Delta S) = 0$ from Lemma 3.2. This, in turn, implies $\overline{\Delta\lambda} = 0$ in view of (19) since the rows of $\bar{\mathcal{A}}$ are linearly independent by construction. Hence the matrix in (19) is nonsingular.

Conversely, assume that the matrices $M_{\bar{\mathcal{A}}}$ are nonsingular for all submatrices $\bar{\mathcal{A}}$ with r linearly independent rows from $\mathcal{A}$. Let $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ be any triple satisfying

$$\begin{aligned} \sum_{i=1}^m \Delta\lambda_i A_i + \Delta S &= 0, \\ A_i \bullet \Delta X &= 0 \quad \forall i = 1, \ldots, m, \\ X^* \Delta S + \Delta X S^* &= 0. \end{aligned}$$

In view of (13) and Lemma 2.2, this is equivalent to

$$\begin{aligned} \mathcal{A}^T \Delta\lambda + \operatorname{svec}(\Delta S) &= 0, \\ \mathcal{A} \operatorname{svec}(\Delta X) &= 0, \\ \mathcal{G} \operatorname{svec}(\Delta X) + \mathcal{H} \operatorname{svec}(\Delta S) &= 0. \end{aligned}$$

Using Lemma 2.3, this may be rewritten as in (16) for a suitable vector $\overline{\Delta\lambda} \in \mathbb{R}^r$. However, the system (16) is equivalent to the triple $(\Delta X, \overline{\Delta\lambda}, \Delta S)$ satisfying

$$\begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{I} \\ \bar{\mathcal{A}} & 0 & 0 \\ \mathcal{G} & 0 & \mathcal{H} \end{pmatrix} \begin{pmatrix} \operatorname{svec}(\Delta X) \\ \overline{\Delta\lambda} \\ \operatorname{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since the matrix of this homogeneous linear system is precisely the matrix $M_{\bar{\mathcal{A}}}$ which is nonsingular by assumption, we obtain $\Delta X = 0$ and $\Delta S = 0$. This shows that (X^*, λ^*, S^*) is KSS-nondegenerate. □

We are now in the position to state our first main result that a KSS-nondegenerate solution is automatically a strictly complementary solution. Under the additional assumption that the matrices $A_1, \ldots, A_m$ are linearly independent, this result may be found in [12, Theorem A.1].

Theorem 3.4 *Let (X^*, λ^*, S^*) be a KSS-nondegenerate solution of the optimality conditions (1). Then $X^* + S^* \succ 0$, i.e., the solution is strictly complementary.*

Proof. Let $r := \operatorname{rank}(\mathcal{A})$ and $\bar{\mathcal{A}}$ be an arbitrary submatrix of A consisting of r linearly independent rows of $\mathcal{A}$. In view of Lemma 3.3, the matrix $M_{\bar{\mathcal{A}}}$ from (18) is nonsingular. We now follow the technique of proof of [12, Theorem A.1]: Since X^* and S^* commute, Lemma 2.1 shows that $\mathcal{G}$ and $\mathcal{H}$ commute, too. Hence there is a simultaneous spectral decomposition, i.e., there is an orthogonal matrix $\mathcal{U} \in \mathbb{R}^{\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}}$ and diagonal matrices $\mathcal{D}^{\mathcal{G}}, \mathcal{D}^{\mathcal{H}} \in \mathbb{R}^{\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}}$ such that

$$\mathcal{G} = \mathcal{U}\mathcal{D}^{\mathcal{G}}\mathcal{U}^T \quad \text{and} \quad \mathcal{H} = \mathcal{U}\mathcal{D}^{\mathcal{H}}\mathcal{U}^T.$$

We therefore have

$$M_{\bar{\mathcal{A}}} = \begin{pmatrix} \mathcal{I} & 0 & 0 \\ 0 & \mathcal{I} & 0 \\ 0 & 0 & \mathcal{U} \end{pmatrix} \begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{U} \\ \bar{\mathcal{A}}\mathcal{U} & 0 & 0 \\ \mathcal{D}^{\mathcal{G}} & 0 & \mathcal{D}^{\mathcal{H}} \end{pmatrix} \begin{pmatrix} \mathcal{U}^T & 0 & 0 \\ 0 & \mathcal{I} & 0 \\ 0 & 0 & \mathcal{U}^T \end{pmatrix}.$$

Since $M_{\bar{\mathcal{A}}}$ is nonsingular, it follows that

$$\begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{U} \\ \bar{\mathcal{A}}\mathcal{U} & 0 & 0 \\ \mathcal{D}^{\mathcal{G}} & 0 & \mathcal{D}^{\mathcal{H}} \end{pmatrix} \tag{20}$$

is also nonsingular.

Now suppose that the solution (X^*, λ^*, S^*) is not strictly complementary. Then there exists a component $s \in \beta$ such that $\lambda_s(X^*) = 0$ and $\lambda_s(S^*) = 0$. Since $\mathcal{D}^{\mathcal{G}}$ and $\mathcal{D}^{\mathcal{H}}$ are diagonal matrices with the eigenvalues of $\mathcal{H}$ and $\mathcal{G}$ on their diagonals, it follows from Lemma 2.1 that

$$\begin{aligned} \mathcal{D}^{\mathcal{G}} &= \frac{1}{2}\operatorname{diag}\big(\ldots, \lambda_j(S^*) + \lambda_i(S^*), \ldots\big)_{1\le j\le i\le n}, \\ \mathcal{D}^{\mathcal{H}} &= \frac{1}{2}\operatorname{diag}\big(\ldots, \lambda_j(X^*) + \lambda_i(X^*), \ldots\big)_{1\le j\le i\le n}. \end{aligned}$$

In particular, taking the pair $(j, i) = (s, s)$, we see that $\mathcal{D}^{\mathcal{G}}$ and $\mathcal{D}^{\mathcal{H}}$ have a zero entry in the same diagonal element. This implies that the matrix from (20) has a zero row, a contradiction to the nonsingularity of this matrix. $\square$

We next wish to show that the X- and S-components of a KSS-nondegenerate solution (X^*, λ^*, S^*) are unique, see also Remark 5.7 for a generalization and the relation of this statement to existing results.

Before stating our next result, however, we first recall some facts about duality for semidefinite programs. To this end, let (X^*, λ^*, S^*) be a solution of the optimality conditions (1). In particular, we then have $X^* \succeq 0, S^* \succeq 0$, and [1, Lemma 2.9] hence implies the equivalence

$$X^*S^* = 0 \Longleftrightarrow X^* \bullet S^* = 0.$$

Consequently, we have

$$0 = X^* \bullet S^* = X^* \bullet \big(C - \sum_{i=1}^{m} \lambda_i^* A_i\big) = C \bullet X^* - b^T\lambda^*,$$

i.e., there is no duality gap between the primal and dual objective functions. Now let $(X^{**}, \lambda^{**}, S^{**})$ be another solution of the optimality conditions (1). Then we have $C \bullet X^* = C \bullet X^{**}$ and $b^T\lambda^* = b^T\lambda^{**}$ by convexity. Using the previous argument, we therefore get

$$0 = C \bullet X^{**} - b^T\lambda^* = X^{**} \bullet S^* \quad \text{and} \quad 0 = C \bullet X^* - b^T\lambda^{**} = X^* \bullet S^{**}.$$

Since the matrices X^*, X^{**}, S^*, and S^{**} are positive semidefinite, this is equivalent to

$$X^*S^{**} = 0 \quad \text{and} \quad X^{**}S^* = 0 \tag{21}$$

This fact will be used in the following result.

Theorem 3.5 *Suppose that (X^*, λ^*, S^*) is a KSS-nondegenerate solution of the optimality conditions* (1). *Then the X^*- and S^*-components are unique.*

Proof. Let $r := \text{rank}(\mathcal{A})$ and let $\bar{\mathcal{A}}$ be an arbitrary submatrix of $\mathcal{A}$ consisting of r linearly independent rows. Furthermore, let $M_{\bar{\mathcal{A}}}$ be the corresponding matrix from (18).

Now let $(X^{**}, \lambda^{**}, S^{**})$ be another solution of the optimality conditions (1). Then define

$$(\Delta X, \Delta\lambda, \Delta S) := \big(X^* - X^{**}, \lambda^* - \lambda^{**}, S^* - S^{**}\big),$$

and let $\overline{\Delta\lambda} \in \mathbb{R}^r$ be a vector such that $\mathcal{A}^T \Delta\lambda = \bar{\mathcal{A}}^T \overline{\Delta\lambda}$ (note that such a vector always exists). Then it is easy to see that we have $\bar{\mathcal{A}}^T \overline{\Delta\lambda} + \text{svec}(\Delta S) = 0$ and $\bar{\mathcal{A}}\,\text{svec}(\Delta X) = 0$. Moreover, using (8) and the matrices $\mathcal{G}$ and $\mathcal{H}$ from (14), we obtain

$$\begin{aligned} \mathcal{G}\,\text{svec}(\Delta X) + \mathcal{H}\,\text{svec}(\Delta S) &= \big(I \otimes_s S^*\big)\,\text{svec}(\Delta X) + \big(I \otimes_s X^*\big)\,\text{svec}(\Delta S) \\ &= \frac{1}{2}\,\text{svec}\,\big(S^*\Delta X + \Delta X S^*\big) + \frac{1}{2}\,\text{svec}\,\big(X^*\Delta S + \Delta S X^*\big) \;=\; 0 \end{aligned}$$

since $X^*S^* = 0 = S^*X^*$ in view of the optimality conditions, and $X^*S^{**} = 0 = S^*X^{**}$ because of (21). Hence we have

$$M_{\bar{\mathcal{A}}} \begin{pmatrix} \text{svec}(\Delta X) \\ \overline{\Delta\lambda} \\ \text{svec}(\Delta S) \end{pmatrix} = 0.$$

Lemma 3.3 therefore implies $\Delta X = 0$ and $\Delta S = 0$. Hence the X^*- and S^*-components of the solution (X^*, λ^*, S^*) are unique. □

We next state that the converse of Theorems 3.4 and 3.5 is also true, i.e., the uniqueness of the solution with respect to the X- and S-components together with the strict complementarity assumption implies that the solution is KSS-nondegenerate.

In order to prove this result, we need to introduce some more notation. To this end, let (X^*, λ^*, S^*) be a (fixed) solution of the optimality conditions (1), and let (2) be a corresponding simultaneous spectral decomposition. Furthermore, recall that

$$p = \text{rank}(X^*) \quad \text{and} \quad q = \text{rank}(S^*),$$

so that $p + q \le n$ in general, whereas we have $p + q = n$ if and only if strict complementarity holds. We now define three important subsets (see [2]):

$$\begin{aligned} \mathcal{N} &:= \big\{Y \in \mathcal{S}^n \,\big|\, A_i \bullet Y = 0 \;\forall i = 1, \dots, m\big\}, \\ \mathcal{T}_X &:= \left\{Q \begin{pmatrix} U & V \\ V^T & 0 \end{pmatrix} Q^T \,\middle|\, U \in \mathcal{S}^p, V \in \mathbb{R}^{p\times(n-p)}\right\}, \\ \mathcal{T}_S &:= \left\{Q \begin{pmatrix} 0 & V \\ V^T & W \end{pmatrix} Q^T \,\middle|\, V \in \mathbb{R}^{(n-q)\times q}, W \in \mathcal{S}^q\right\}. \end{aligned}$$

The corresponding orthogonal subspaces (with respect to the inner product $\bullet$ in $\mathcal{S}^n$) are given by

$$\mathcal{N}^\perp \;=\; \text{span}\,\big\{A_1, \dots, A_m\big\},$$

$$\mathcal{T}_X^\perp = \Big\{ Q \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix} Q^T \Big| W \in \mathcal{S}^{n-p} \Big\},$$
$$\mathcal{T}_S^\perp = \Big\{ Q \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} Q^T \Big| U \in \mathcal{S}^{n-q} \Big\}.$$

These subspaces will play an important role in our next section. Some of them, however, are also used in the following preliminary result.

Lemma 3.6 *Let (X^*, λ^*, S^*) be a strictly complementary solution of the optimality conditions* (1)*, and suppose that $(\Delta X, \Delta S) \in \mathcal{S}^n \times \mathcal{S}^n$ are given such that*

$$\Delta X \bullet \Delta S = 0 \quad \textit{and} \quad X^* \Delta S + \Delta X S^* = 0. \tag{22}$$

Then we have $\Delta X \in \mathcal{T}_S^\perp$ and $\Delta S \in \mathcal{T}_X^\perp$.

Proof. Using the simultaneous spectral decomposition (2), we may rewrite the second condition in (22) as $QD^X Q^T \Delta S + \Delta X Q D^S Q^T = 0$. Pre- and postmultiplying this matrix with Q^T and Q, respectively, and using the matrices $\widetilde{\Delta X}, \widetilde{\Delta S}$ from (5), we obtain

$$D^X \widetilde{\Delta S} + \widetilde{\Delta X} D^S = 0. \tag{23}$$

Using the partitions from (4), (6), and recalling that $\beta = \emptyset$, we may rewrite (23) as

$$\begin{aligned} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} D_\alpha^X & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widetilde{\Delta S}_{\alpha\alpha} & \widetilde{\Delta S}_{\alpha\gamma} \\ \widetilde{\Delta S}_{\alpha\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix} + \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}_{\alpha\gamma}^T & \widetilde{\Delta X}_{\gamma\gamma} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & D_\gamma^S \end{pmatrix} \\ &= \begin{pmatrix} D_\alpha^X \widetilde{\Delta S}_{\alpha\alpha} & D_\alpha^X \widetilde{\Delta S}_{\alpha\gamma} + \widetilde{\Delta X}_{\alpha\gamma} D_\gamma^S \\ 0 & \widetilde{\Delta X}_{\gamma\gamma} D_\gamma^S \end{pmatrix}. \end{aligned}$$

Since D_α^X, D_γ^S are positive definite diagonal matrices, this implies

$$\widetilde{\Delta S}_{\alpha\alpha} = 0, \quad \widetilde{\Delta X}_{\gamma\gamma} = 0, \quad \text{and} \quad D_\alpha^X \widetilde{\Delta S}_{\alpha\gamma} + \widetilde{\Delta X}_{\alpha\gamma} D_\gamma^S = 0. \tag{24}$$

We therefore have

$$\widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}_{\alpha\gamma}^T & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{\Delta S} = \begin{pmatrix} 0 & \widetilde{\Delta S}_{\alpha\gamma} \\ \widetilde{\Delta S}_{\alpha\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}.$$

Since $\Delta X \bullet \Delta S = 0$ implies $\widetilde{\Delta X} \bullet \widetilde{\Delta S} = 0$, we further obtain

$$\begin{aligned} 0 &= \operatorname{trace}\big(\widetilde{\Delta X}\widetilde{\Delta S}\big) \\ &= \operatorname{trace}\left[\begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}_{\alpha\gamma}^T & 0 \end{pmatrix} \begin{pmatrix} 0 & \widetilde{\Delta S}_{\alpha\gamma} \\ \widetilde{\Delta S}_{\alpha\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix} \right] \\ &= \operatorname{trace}\big(\widetilde{\Delta X}_{\alpha\gamma} \widetilde{\Delta S}_{\alpha\gamma}^T\big) + \operatorname{trace}\big(\widetilde{\Delta X}_{\alpha\gamma}^T \widetilde{\Delta S}_{\alpha\gamma}\big). \end{aligned}$$

Solving the last equation in (24) for $\widetilde{\Delta X}_{\alpha\gamma}$ and substituting into the above expression, we get

$$\begin{aligned}
0 &= \operatorname{trace}\big(D_\alpha^X \widetilde{\Delta S}_{\alpha\gamma}(D_\gamma^S)^{-1}\widetilde{\Delta S}_{\alpha\gamma}^T\big) + \operatorname{trace}\big((D_\gamma^S)^{-1}\widetilde{\Delta S}_{\alpha\gamma}^T D_\alpha^X \widetilde{\Delta S}_{\alpha\gamma}\big) \\
&= \operatorname{trace}\big((D_\alpha^X)^{1/2}\widetilde{\Delta S}_{\alpha\gamma}(D_\gamma^S)^{-1/2}(D_\gamma^S)^{-1/2}\widetilde{\Delta S}_{\alpha\gamma}^T(D_\alpha^X)^{1/2}\big) \\
&\quad + \operatorname{trace}\big((D_\gamma^S)^{-1/2}\widetilde{\Delta S}_{\alpha\gamma}^T(D_\alpha^X)^{1/2}(D_\alpha^X)^{1/2}\widetilde{\Delta S}_{\alpha\gamma}(D_\gamma^S)^{-1/2}\big) \\
&= \operatorname{trace}\big(\big[(D_\alpha^X)^{1/2}\widetilde{\Delta S}_{\alpha\gamma}(D_\gamma^S)^{-1/2}\big]\big[(D_\alpha^X)^{1/2}\widetilde{\Delta S}_{\alpha\gamma}(D_\gamma^S)^{-1/2}\big]^T\big) \\
&\quad + \operatorname{trace}\big(\big[(D_\gamma^S)^{-1/2}\widetilde{\Delta S}_{\alpha\gamma}^T(D_\alpha^X)^{1/2}\big]\big[(D_\gamma^S)^{-1/2}\widetilde{\Delta S}_{\alpha\gamma}^T(D_\alpha^X)^{1/2}\big]^T\big) \\
&= \big\|(D_\alpha^X)^{1/2}\widetilde{\Delta S}_{\alpha\gamma}(D_\gamma^S)^{-1/2}\big\|_F^2 + \big\|(D_\gamma^S)^{-1/2}\widetilde{\Delta S}_{\alpha\gamma}^T(D_\alpha^X)^{1/2}\big\|_F^2.
\end{aligned}$$

Hence we get $(D_\alpha^X)^{1/2}\widetilde{\Delta S}_{\alpha\gamma}(D_\gamma^S)^{-1/2} = 0$ and therefore $\widetilde{\Delta S}_{\alpha\gamma} = 0$. This, in turn, gives $\widetilde{\Delta X}_{\alpha\gamma} = 0$. Consequently, we have

$$\widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{\Delta S} = \begin{pmatrix} 0 & 0 \\ 0 & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}.$$

Using (5), we therefore get

$$\begin{aligned}
\Delta X &= Q\widetilde{\Delta X}Q^T = Q\begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & 0 \\ 0 & 0 \end{pmatrix}Q^T \in \mathcal{T}_S^\perp, \\
\Delta S &= Q\widetilde{\Delta S}Q^T = Q\begin{pmatrix} 0 & 0 \\ 0 & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}Q^T \in \mathcal{T}_X^\perp.
\end{aligned}$$

This completes the proof. □

We are now in the position to prove the converse of Theorems 3.4 and 3.5.

Theorem 3.7 *Suppose that* (X^*, λ^*, S^*) *is a solution of the optimality conditions* (1) *such that strict complementarity is satisfied and the* X^**- and* S^**-components are unique. Then* (X^*, λ^*, S^*) *is KSS-nondegenerate.*

Proof. Let $r := \operatorname{rank}(\mathcal{A})$ and $\bar{\mathcal{A}}$ be a submatrix consisting of r linearly independent rows of $\mathcal{A}$. In view of Lemma 3.3, we have to show that the matrix $M_{\bar{\mathcal{A}}}$ from (18) is nonsingular.

Assume this matrix is singular. Then there is a nonzero triple $(\Delta X, \overline{\Delta\lambda}, \Delta S)$ satisfying

$$\begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{I} \\ \bar{\mathcal{A}} & 0 & 0 \\ \mathcal{G} & 0 & \mathcal{H} \end{pmatrix}\begin{pmatrix} \operatorname{svec}(\Delta X) \\ \overline{\Delta\lambda} \\ \operatorname{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Note that $(\Delta X, \overline{\Delta\lambda}, \Delta S) \neq (0,0,0)$ implies $(\Delta X, \Delta S) \neq (0,0)$ since the matrix $\bar{\mathcal{A}}$ has full rank.

Using Lemma 2.3, we see that there exists a vector $\Delta\lambda \in \mathbb{R}^m$ such that the triple $(\Delta X, \Delta\lambda, \Delta S)$ satisfies

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ \mathcal{G} & 0 & \mathcal{H} \end{pmatrix} \begin{pmatrix} \text{svec}(\Delta X) \\ \Delta\lambda \\ \text{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{25}$$

Note that the first two block lines of (25) are equivalent to $\sum_{i=1}^m \Delta\lambda_i A_i + \Delta S = 0$ and $A_i \bullet \Delta X = 0$ $(i = 1, \dots, m)$. This means that $\Delta X \in \mathcal{N}$ and $\Delta S \in \mathcal{N}^\perp$. In particular, we therefore have

$$\Delta X \bullet \Delta S = 0. \tag{26}$$

Now it is easy to see that $\big(X^* + \Delta X, \lambda^* + \Delta\lambda, S^* + \Delta S\big)$ satisfies

$$\mathcal{A}^T(\lambda^* + \Delta\lambda) + \text{svec}(S^* + \Delta S) = \text{svec}(C), \quad \mathcal{A}\,\text{svec}(X^* + \Delta X) = b.$$

Furthermore, from the last line in (25) and Lemma 2.2, we obtain

$$\begin{aligned} &\mathcal{G}\,\text{svec}(\Delta X) + \mathcal{H}\,\text{svec}(\Delta S) = 0 \\ \iff\; & \big(I \otimes_s S^*\big)\,\text{svec}(\Delta X) + \big(I \otimes_s X^*\big)\,\text{svec}(\Delta S) = 0 \\ \iff\; & X^*\Delta S + \Delta X S^* = 0. \end{aligned} \tag{27}$$

Using (26) and Lemma 3.6, we therefore get

$$\Delta X = Q \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & 0 \\ 0 & 0 \end{pmatrix} Q^T \quad \text{and} \quad \Delta S = Q \begin{pmatrix} 0 & 0 \\ 0 & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix} Q^T$$

for certain matrices $\widetilde{\Delta X}_{\alpha\alpha} \in \mathcal{S}^{p\times p}$ and $\widetilde{\Delta S}_{\gamma\gamma} \in \mathcal{S}^{q\times q}$ (recall that $p + q = n$ because of strict complementarity). This implies

$$X^* + \Delta X = Q \begin{pmatrix} D_\alpha^X + \widetilde{\Delta X}_{\alpha\alpha} & 0 \\ 0 & 0 \end{pmatrix} Q^T \succeq 0$$

for all ΔX sufficiently small satisfying (25), and

$$S^* + \Delta S = Q \begin{pmatrix} 0 & 0 \\ 0 & D_\gamma^S + \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix} Q^T \succeq 0$$

for all ΔS sufficiently small satisfying (25). Moreover, we have

$$(X^* + \Delta X)(S^* + \Delta S) = Q \begin{pmatrix} D_\alpha^X + \widetilde{\Delta X}_{\alpha\alpha} & 0 \\ 0 & 0 \end{pmatrix} Q^T Q \begin{pmatrix} 0 & 0 \\ 0 & D_\gamma^S + \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix} Q^T = 0.$$

Hence $\big(X^* + \Delta X, \lambda^* + \Delta\lambda, S^* + \Delta S\big)$ is also a solution of the optimality conditions (1) for any sufficiently small $\big(\Delta X, \Delta\lambda, \Delta S\big)$ satisfying (25). Since $(\Delta X, \Delta S) \neq 0$, this contradicts our assumption that the X^*- and S^*-components of the solution (X^*, λ^*, S^*) are unique. □

4 The AHO-Nondegeneracy Condition

We begin by stating the nondegeneracy condition introduced by Alizadeh, Haeberly and Overton [2] and which we therefore call AHO-nondegeneracy.

Definition 4.1 *A solution* (X^*, λ^*, S^*) *of the optimality conditions* (1) *is called* AHO-nondegenerate *if* $\mathcal{N}^\perp \cap \mathcal{T}_X^\perp = \{0\}$ *and* $\mathcal{N} \cap \mathcal{T}_S^\perp = \{0\}$.

Note that our definition of an AHO-nondegenerate solution corresponds to what is called primal and dual nondegenerate in [2]. Furthermore, recall that the sets $\mathcal{N}, \mathcal{T}_X, \mathcal{T}_S$ as well as their orthogonal subspaces were introduced in the previous section.

We first state the following result.

Lemma 4.2 *Let* (X^*, λ^*, S^*) *be a solution of the optimality conditions, and suppose that* $\Delta S \in \mathcal{N}^\perp \cap \mathcal{T}_X^\perp$ *and* $\Delta X \in \mathcal{N} \cap \mathcal{T}_S^\perp$. *Then there exists a* $\Delta\lambda \in \mathbb{R}^m$ *such that*

$$\sum_{i=1}^m \Delta\lambda_i A_i + \Delta S = 0, \quad A_i \bullet \Delta X = 0 \ \ (i = 1, \ldots, m), \quad X^* \Delta S + \Delta X S^* = 0. \tag{28}$$

Proof. Since $\Delta S \in \mathcal{N}^\perp$, there exist $\Delta\lambda_i \in \mathbb{R}$ such that $\Delta S = -\sum_{i=1}^m \Delta\lambda_i A_i$. Furthermore, $\Delta X \in \mathcal{N}$ implies $A_i \bullet \Delta X = 0$ for all $i = 1, \ldots, m$. Moreover, it follows from $\Delta S \in \mathcal{T}_X^\perp$ and $\Delta X \in \mathcal{T}_S^\perp$ that

$$\Delta S = Q \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix} Q^T \quad \text{and} \quad \Delta X = Q \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} Q^T$$

for some symmetric matrices $W \in \mathcal{S}^{n-p}$ and $U \in \mathcal{S}^{n-q}$, respectively. Using the spectral decomposition from (2) and the notation from (4), we therefore get

$$\begin{aligned} Q^T X^* \Delta S Q + Q^T \Delta X S^* Q &= \big(Q^T X^* Q\big)\big(Q^T \Delta S Q\big) + \big(Q^T \Delta X Q\big)\big(Q^T S^* Q\big) \\ &= D^X \big(Q^T \Delta S Q\big) + \big(Q^T \Delta X Q\big) D^S \\ &= \begin{pmatrix} D_\alpha^X & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix} + \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & D_\gamma^S \end{pmatrix} = 0. \end{aligned}$$

If we premultiply this equation with Q and postmultiply it with Q^T, we finally obtain $X^* \Delta S + \Delta X S^* = 0$. □

Note that we may choose $\Delta\lambda_i = 0$ in the previous result for all indices $i \in \{1, \ldots, m\}$ belonging to linearly dependent rows of the matrix $\mathcal{A}$.

Theorem 4.3 *Let* (X^*, λ^*, S^*) *be a KSS-nondegenerate solution of the optimality conditions* (1). *Then* (X^*, λ^*, S^*) *is also AHO-nondegenerate.*

Proof. Let $\Delta S \in \mathcal{N}^\perp \cap \mathcal{T}_X^\perp$ and $\Delta X \in \mathcal{N} \cap \mathcal{T}_S^\perp$ be given. In view of Lemma 4.2, it follows that there exists $\Delta\lambda \in \mathbb{R}^m$ such that (28) holds. Since (X^*, λ^*, S^*) is KSS-nondegenerate by assumption, we obtain $\Delta X = 0$ and $\Delta S = 0$. This shows that $\mathcal{N}^\perp \cap \mathcal{T}_X^\perp = \{0\}$ and

$\mathcal{N} \cap \mathcal{T}_S^\perp = \{0\}$, i.e., (X^*, λ^*, S^*) is AHO-nondegenerate. □

Note that Theorem 4.3 holds without strict complementarity. The converse is not true in general, as can be seen from an example in [2, page 117] where the authors give an example of a semidefinite program whose solution satisfies AHO-nondegeneracy but not strict complementarity. Hence KSS-nondegeneracy cannot hold in view of Theorem 3.4.

However, in our next result we show that the converse of Theorem 4.3 holds if we assume, in addition, that strict complementarity is satisfied.

Theorem 4.4 *Let (X^*, λ^*, S^*) be a strictly complementary and AHO-nondegenerate solution of the optimality conditions* (1)*. Then (X^*, λ^*, S^*) is also KSS-nondegenerate.*

Proof. Let $(\Delta X, \Delta\lambda, \Delta S)$ be any triple satisfying the conditions

$$\begin{aligned} \sum_{i=1}^m \Delta\lambda_i A_i + \Delta S &= 0, && (29)\\ A_i \bullet \Delta X &= 0 \quad \forall i = 1, \ldots, m, && (30)\\ X^*\Delta S + \Delta X S^* &= 0. && (31) \end{aligned}$$

From (29) and (30), we immediately obtain

$$\Delta X \in \mathcal{N} \quad \text{and} \quad \Delta S \in \mathcal{N}^\perp. \tag{32}$$

In particular, this shows that

$$\Delta X \bullet \Delta S = 0. \tag{33}$$

Using (31), (33), and strict complementarity of the solution, we obtain from Lemma 3.6 that $\Delta X \in \mathcal{T}_S^\perp$ and $\Delta S \in \mathcal{T}_X^\perp$. Together with (32), we therefore get $\Delta X = 0$ and $\Delta S = 0$ from AHO-nondegeneracy. Hence (X^*, λ^*, S^*) is a KSS-nondegenerate solution. □

Using Theorems 3.4, 4.3, and 4.4, we, in particular, obtain the following consequence.

Corollary 4.5 *Let (X^*, λ^*, S^*) be a solution of the optimality conditions* (1)*. Then (X^*, λ^*, S^*) is KSS-nondegenerate if and only if it is strictly complementary and AHO-nondegenerate.*

In particular, it follows that KSS- and AHO-nondegeneracy are equivalent if the solution (X^*, λ^*, S^*) of the optimality conditions (1) satisfies strict complementary. According to Kojima et al. [13], this result has been noted before by Haeberly in a private communication to the authors of [13].

Furthermore, we note that both Haeberly [6] and Miller [15] are able to characterize strict complementarity and AHO-nondegeneracy in terms of the nonsingularity of a certain Jacobian matrix arising within the framework of interior-point methods. In view of Corollary 4.5, this is essentially the statement given in Lemma 3.3 except that both papers [6, 15] assume, in addition, that the matrices A_i are linearly independent.

5 The KN-Nondegeneracy Condition

Before we state the definition of nondegeneracy as introduced in [12], we begin with a short discussion in order to get a better understanding of this nondegeneracy concept which was used in [12] in order to prove local quadratic convergence of a nonsmooth Newton method for semidefinite programs without strict complementarity.

Recall from Lemma 3.2 that the KSS-nondegeneracy of a solution (X^*, λ^*, S^*) is equivalent to the implication

$$\left.\begin{array}{l} \mathcal{A}^T \Delta\lambda + \operatorname{svec}(\Delta S) = 0, \\ \mathcal{A}\operatorname{svec}(\Delta X) = 0, \\ \big(I \otimes_s X^*\big)\operatorname{svec}(\Delta S) + \big(I \otimes_s S^*\big)\operatorname{svec}(\Delta X) = 0 \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \operatorname{svec}(\Delta X) = 0, \\ \operatorname{svec}(\Delta S) = 0. \end{array}\right.$$

Using the simultaneous spectral decomposition (2), defining the orthogonal matrix

$$\mathcal{V}^* := Q \otimes_s Q$$

and exploiting standard properties of the symmetric Kronecker product (see Lemma 2.1), it follows that the KSS-nondegeneracy condition may further be rewritten as

$$\left.\begin{array}{l} \mathcal{A}^T \Delta\lambda + \operatorname{svec}(\Delta S) = 0 \\ \mathcal{A}\operatorname{svec}(\Delta X) = 0, \\ \mathcal{V}^*(I \otimes_s D^S)(\mathcal{V}^*)^T\operatorname{svec}(\Delta X) + \mathcal{V}^*(I \otimes_s D^X)(\mathcal{V}^*)^T\operatorname{svec}(\Delta S) = 0 \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \operatorname{svec}(\Delta X) = 0, \\ \operatorname{svec}(\Delta S) = 0. \end{array}\right.$$

Now define $\Sigma_- := I \otimes_s D^S$ and $\Sigma_+ := I \otimes_s D^X$. Then KSS-nondegeneracy is equivalent to

$$\left.\begin{array}{l} \mathcal{A}^T \Delta\lambda + \operatorname{svec}(\Delta S) = 0 \\ \mathcal{A}\operatorname{svec}(\Delta X) = 0, \\ \mathcal{V}^*\Sigma_-(\mathcal{V}^*)^T\operatorname{svec}(\Delta X) + \mathcal{V}^*\Sigma_+(\mathcal{V}^*)^T\operatorname{svec}(\Delta S) = 0 \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \operatorname{svec}(\Delta X) = 0, \\ \operatorname{svec}(\Delta S) = 0, \end{array}\right.$$

and an easy calculation shows that Σ_- and Σ_+ are diagonal matrices with nonnegative diagonal elements σ_{ij}^- and σ_{ij}^+, respectively. Moreover, these diagonal entries satisfy the conditions

$$\begin{aligned} \sigma_{ij}^- &= 0 \quad \text{if } (i,j) \in (\alpha \times \alpha) \cup (\alpha \times \beta) \cup (\beta \times \alpha), \\ \sigma_{ij}^+ &= 0 \quad \text{if } (i,j) \in (\beta \times \gamma) \cup (\gamma \times \beta) \cup (\gamma \times \gamma) \end{aligned}$$

as well as

$$\sigma_{ij}^- = 0, \sigma_{ij}^+ = 0 \quad \forall (i,j) \in \beta \times \beta, \tag{34}$$

whereas all other entries $\sigma_{ij}^-, \sigma_{ij}^+$ are positive scalars. Although only the sign (not the precise value) of the elements of Σ_- and Σ_+ are important, property (34) causes some singularity problems (cf. the proof of Theorem 5.5). This is precisely the point we try to overcome using the KN-nondegeneracy condition.

Definition 5.1 *A solution* (X^*, λ^*, S^*) *of the optimality conditions* (1) *is called* KN-nondegenerate *if, for all diagonal matrices*

$$\Sigma_- = \operatorname{diag}\big(\ldots, \sigma_{ij}^-, \ldots\big)_{1 \le j \le i \le n}, \quad \Sigma_+ = \operatorname{diag}\big(\ldots, \sigma_{ij}^+, \ldots\big)_{1 \le j \le i \le n} \tag{35}$$

satisfying

$$\begin{aligned} &\sigma_{ij}^- \geq 0, \sigma_{ij}^+ \geq 0, \sigma_{ij}^- + \sigma_{ij}^+ > 0 \quad \textit{if } (i,j) \in \beta \times \beta, \\ &\sigma_{ij}^- = 0 \quad \textit{if } (i,j) \in (\alpha \times \alpha) \cup (\alpha \times \beta) \cup (\beta \times \alpha), \\ &\sigma_{ij}^+ = 0 \quad \textit{if } (i,j) \in (\beta \times \gamma) \cup (\gamma \times \beta) \cup (\gamma \times \gamma), \\ &\sigma_{ij}^- > 0, \sigma_{ij}^+ > 0 \quad \textit{for all other pairs } (i,j), \end{aligned} \tag{36}$$

the following implication holds for any triple $\big(\Delta X, \Delta\lambda, \Delta S\big) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$*:*

$$\left.\begin{array}{l} \mathcal{A}^T \Delta\lambda + \operatorname{svec}(\Delta S) = 0 \\ \mathcal{A} \operatorname{svec}(\Delta X) = 0, \\ \mathcal{V}^* \Sigma_- (\mathcal{V}^*)^T \operatorname{svec}(\Delta X) + \mathcal{V}^* \Sigma_+ (\mathcal{V}^*)^T \operatorname{svec}(\Delta S) = 0 \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \operatorname{svec}(\Delta X) = 0, \\ \operatorname{svec}(\Delta S) = 0. \end{array}\right.$$

The only difference between KSS-nondegeneracy and KN-nondegeneracy is that we now require that at least one of the two elements σ_{ij}^- or σ_{ij}^+ is positive for all components $(i,j) \in \beta \times \beta$, hence the sum of these elements is positive.

In particular, it follows from our previous discussion that the definition of KN-nondegeneracy coincides with KSS-nondegeneracy if $\beta = \emptyset$. Hence, in this case, it also coincides with AHO-nondegeneracy due to Corollary 4.5. We formally summarize this observation in the following result.

Theorem 5.2 *Let* (X^*, λ^*, S^*) *be a strictly complementary solution of the optimality conditions* (1)*. Then* (X^*, λ^*, S^*) *is KN-nondegenerate if and only if it is KSS-nondegenerate or, equivalently, AHO-nondegenerate.*

While KSS-nondegeneracy automatically implies strict complementarity, we will show in our subsequent analysis that the concept of KN-nondegeneracy is also useful in the absense of strict complementarity. To this end, we begin with an equivalent formulation of KN-nondegeneracy which is similar to Lemma 3.3 for a KSS-nondegenerate solution.

Lemma 5.3 *Let* (X^*, λ^*, S^*) *be a solution of the optimality conditions* (1)*. Furthermore, let* $r := \operatorname{rank}(\mathcal{A})$ *be the rank of the matrix* $\mathcal{A}$ *from* (12)*. Then* (X^*, λ^*, S^*) *is KN-nondegenerate if and only if, for any submatrix* $\bar{\mathcal{A}}$ *consisting of* r *linearly independent rows of* $\mathcal{A}$ *and any diagonal matrices* Σ_-, Σ_+ *satisfying* (35), (36)*, the matrix*

$$M_{\bar{\mathcal{A}},\Sigma_-,\Sigma_+} := \begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{I} \\ \bar{\mathcal{A}} & 0 & 0 \\ \mathcal{V}^* \Sigma_- \mathcal{V}^{*T} & 0 & \mathcal{V}^* \Sigma_+ \mathcal{V}^{*T} \end{pmatrix} \tag{37}$$

is nonsingular.

Proof. First assume that (X^*, λ^*, S^*) is KN-nondegenerate. Let $\bar{\mathcal{A}}$ be an arbitrary submatrix consisting of r linearly independent rows, and let Σ_-, Σ_+ be two diagonal matrices satisfying (35), (36). Let $(\Delta X, \overline{\Delta\lambda}, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^r \times \mathcal{S}^n$ be any triple such that

$$\begin{pmatrix} 0 & \bar{\mathcal{A}}^T & \mathcal{I} \\ \bar{\mathcal{A}} & 0 & 0 \\ \mathcal{V}^* \Sigma_- \mathcal{V}^{*T} & 0 & \mathcal{V}^* \Sigma_+ \mathcal{V}^{*T} \end{pmatrix} \begin{pmatrix} \operatorname{svec}(\Delta X) \\ \overline{\Delta\lambda} \\ \operatorname{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{38}$$

Adding some zero components to $\overline{\Delta\lambda}$, we obtain a vector $\Delta\lambda$ such that the triple $(\Delta X, \Delta\lambda, \Delta S)$ satisfies

$$\begin{aligned} \mathcal{A}^T\Delta\lambda + \operatorname{svec}(\Delta S) &= 0, & (39)\\ \mathcal{A}\operatorname{svec}(\Delta X) &= 0, & (40)\\ \mathcal{V}^*\Sigma_-\mathcal{V}^{*T}\operatorname{svec}(\Delta X) + \mathcal{V}^*\Sigma_+\mathcal{V}^{*T}\operatorname{svec}(\Delta S) &= 0. & (41) \end{aligned}$$

Since (X^*, λ^*, S^*) is KN-nondegenerate, we get $\Delta X = 0$ and $\Delta S = 0$. This, in turn, implies $\overline{\Delta\lambda} = 0$ in view of (38) and the full rank of the matrix $\bar{\mathcal{A}}$. Hence the matrix in (38) is nonsingular.

Conversely, assume that the matrices $M_{\bar{\mathcal{A}},\Sigma_-,\Sigma_+}$ are nonsingular for all submatrices $\bar{\mathcal{A}}$ with r linearly independent rows from $\mathcal{A}$ and for all diagonal matrices Σ_-, Σ_+ satisfying (35), (36). Let $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ be any triple satisfying (39)–(41). Let $\overline{\Delta\lambda} \in \mathbb{R}^r$ be a vector such that $\mathcal{A}^T\Delta\lambda = \bar{\mathcal{A}}^T\overline{\Delta\lambda}$. Then (39)–(41) may be rewritten as

$$M_{\bar{\mathcal{A}},\Sigma_-,\Sigma_+}\begin{pmatrix} \operatorname{svec}(\Delta X) \\ \overline{\Delta\lambda} \\ \operatorname{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since $M_{\bar{\mathcal{A}},\Sigma_-,\Sigma_+}$ is nonsingular, this implies $\Delta X = 0$ and $\Delta S = 0$. Hence (X^*, λ^*, S^*) is KN-nondegenerate. □

In view of our introductory remarks to this section, it is not surprising that KN-nondegeneracy is implied by KSS-nondegeneracy. In the following, we want to prove the stronger result that AHO-nondegeneracy implies KN-nondegeneracy. To this end, we first state a simple lemma.

Lemma 5.4 *Let $A, B \in \mathbb{R}^{k\times l}$ be (not necessarily square) matrices with elements a_{ij} and b_{ij}, respectively. Then the following statements hold:*

(a) If $a_{ij}b_{ij} \le 0$ for all $i \in \{1, \ldots, k\}$ and all $j \in \{1, \ldots, l\}$, then $\operatorname{trace}(AB^T) \le 0$.

(b) If $a_{ij} = -\sigma_{ij}b_{ij}$ for all i, j with some positive scalars σ_{ij}, then $\operatorname{trace}(AB^T) \le 0$, *and equality holds if and only if $B = 0$ (which, in turn, is equivalent to $A = 0$).*

Proof. Let $C := AB^T \in \mathbb{R}^{k\times k}$ and write $C = (c_{ij})$. Then we have $c_{ij} = \sum_{m=1}^{l} a_{im}b_{jm}$ for all $i, j = 1, \ldots, k$. Therefore, under the assumption of part (a), we obtain

$$\operatorname{trace}(AB^T) = \operatorname{trace}(C) = \sum_{i=1}^{k} c_{ii} = \sum_{i=1}^{k}\sum_{m=1}^{l} \underbrace{a_{im}b_{im}}_{\le 0} \le 0.$$

Similarly, under the assumption of part (b), we get

$$\operatorname{trace}(AB^T) = \sum_{i=1}^{k}\sum_{m=1}^{l} a_{im}b_{im} = -\sum_{i=1}^{k}\sum_{m=1}^{l} \underbrace{\sigma_{im}}_{>0} b_{im}^2 \le 0,$$

and equality holds if and only if $B = 0$. $\square$

We are now in the position to show that AHO-nondegeneracy implies KN-nondegeneracy. Note that we do not assume strict complementarity in this result.

Theorem 5.5 *Let (X^*, λ^*, S^*) be an AHO-nondegenerate solution of the optimality conditions* (1)*. Then (X^*, λ^*, S^*) is also KN-nondegenerate.*

Proof. Let Σ_- and Σ_+ be two arbitrary diagonal matrices satisfying (35), (36). Furthermore, let $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ be any triple satisfying (39)–(41). Note that (39), (40) imply $\Delta X \in \mathcal{N}$ and $\Delta S \in \mathcal{N}^\perp$. In particular, we therefore have $\Delta X \bullet \Delta S = 0$. Furthermore, since $\mathcal{V}^* = Q \otimes_s Q$ is nonsingular, (41) is equivalent to $\Sigma_-(\mathcal{V}^*)^T \operatorname{svec}(\Delta X) + \Sigma_+(\mathcal{V}^*)^T \operatorname{svec}(\Delta S) = 0$. In view of Lemma 2.1 and (8), we have

$$(\mathcal{V}^*)^T \operatorname{svec}(\Delta X) = \left(Q^T \otimes_s Q^T\right) \operatorname{svec}(\Delta X) = \operatorname{svec}(Q^T \Delta X Q)$$

and, similarly, $(\mathcal{V}^*)^T \operatorname{svec}(\Delta S) = \operatorname{svec}(Q^T \Delta S Q)$. Introducing the matrices $\widetilde{\Delta X}$ and $\widetilde{\Delta S}$ as in (5), we get $\Sigma_- \operatorname{svec}(\widetilde{\Delta X}) + \Sigma_+ \operatorname{svec}(\widetilde{\Delta S}) = 0$. Componentwise, this may be rewritten as

$$\sigma_{ij}^- \widetilde{\Delta X}_{ij} + \sigma_{ij}^+ \widetilde{\Delta S}_{ij} = 0 \quad \forall 1 \leq j \leq i \leq n. \tag{42}$$

Taking into account properties (35), (36) of the two diagonal matrices Σ_-, Σ_+, we obtain

$$\widetilde{\Delta S}_{ij} = 0 \quad \forall (i,j) \in (\alpha \times \alpha) \cup (\alpha \times \beta) \cup (\beta \times \alpha), \tag{43}$$

$$\widetilde{\Delta X}_{ij} = 0 \quad \forall (i,j) \in (\beta \times \gamma) \cup (\gamma \times \beta) \cup (\gamma \times \gamma). \tag{44}$$

Furthermore, since $\sigma_{ij}^-, \sigma_{ij}^+ \geq 0, \sigma_{ij}^- + \sigma_{ij}^+ > 0$ for all $(i,j) \in \beta \times \beta$, it follows from (42) that

$$\widetilde{\Delta X}_{ij} \cdot \widetilde{\Delta S}_{ij} \leq 0 \quad \forall (i,j) \in \beta \times \beta. \tag{45}$$

Let us partition the matrices $\widetilde{\Delta X}, \widetilde{\Delta S}$ as in (6). Then we obtain from (43), (44) that

$$\widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\beta} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}_{\alpha\beta}^T & \widetilde{\Delta X}_{\beta\beta} & 0 \\ \widetilde{\Delta X}_{\alpha\gamma}^T & 0 & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{\Delta S} = \begin{pmatrix} 0 & 0 & \widetilde{\Delta S}_{\alpha\gamma} \\ 0 & \widetilde{\Delta S}_{\beta\beta} & \widetilde{\Delta S}_{\beta\gamma} \\ \widetilde{\Delta S}_{\alpha\gamma}^T & \widetilde{\Delta S}_{\beta\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}.$$

Since $\Delta X \bullet \Delta S = 0$ and, therefore, $\widetilde{\Delta X} \bullet \widetilde{\Delta S} = 0$, it follows from the previous representations of $\widetilde{\Delta X}$ and $\widetilde{\Delta S}$ that

$$0 = \widetilde{\Delta X} \bullet \widetilde{\Delta S} = \operatorname{trace}\left(\widetilde{\Delta X}_{\alpha\gamma} \widetilde{\Delta S}_{\alpha\gamma}^T\right) + \operatorname{trace}\left(\widetilde{\Delta X}_{\beta\beta} \widetilde{\Delta S}_{\beta\beta}\right) + \operatorname{trace}\left(\widetilde{\Delta X}_{\alpha\gamma}^T \widetilde{\Delta S}_{\alpha\gamma}\right). \tag{46}$$

Since

$$\widetilde{\Delta X}_{ij} = -\frac{\sigma_{ij}^+}{\sigma_{ij}^-} \widetilde{\Delta S}_{ij} \quad \forall (i,j) \in \alpha \times \gamma \tag{47}$$

and $\sigma_{ij}^-, \sigma_{ij}^+ > 0$ for all $(i,j) \in \alpha \times \gamma$, we obtain from Lemma 5.4, (45), and (47) that

$$\text{trace}\left(\widetilde{\Delta X}_{\alpha\gamma}\widetilde{\Delta S}_{\alpha\gamma}^T\right) \le 0, \ \ \text{trace}\left(\widetilde{\Delta X}_{\beta\beta}\widetilde{\Delta S}_{\beta\beta}\right) \le 0, \ \ \text{and} \ \ \text{trace}\left(\widetilde{\Delta X}_{\alpha\gamma}^T\widetilde{\Delta S}_{\alpha\gamma}\right) \le 0.$$

Consequently, we have $\text{trace}\left(\widetilde{\Delta X}_{\alpha\gamma}\widetilde{\Delta S}_{\alpha\gamma}^T\right) = 0$ in view of (46) and therefore $\widetilde{\Delta X}_{\alpha\gamma} = 0$ and $\widetilde{\Delta S}_{\alpha\gamma} = 0$ because of (47) and Lemma 5.4. Hence

$$\widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\beta} & 0 \\ \widetilde{\Delta X}_{\alpha\beta}^T & \widetilde{\Delta X}_{\beta\beta} & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{\Delta S} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \widetilde{\Delta S}_{\beta\beta} & \widetilde{\Delta S}_{\beta\gamma} \\ 0 & \widetilde{\Delta S}_{\beta\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}.$$

Using (5), this implies $\Delta X = Q\widetilde{\Delta X}Q^T \in \mathcal{T}_S^\perp$ and $\Delta S = Q\widetilde{\Delta S}Q^T \in \mathcal{T}_X^\perp$. Thus we have $\Delta X \in \mathcal{N} \cap \mathcal{T}_S^\perp$ and $\Delta S \in \mathcal{N}^\perp \cap \mathcal{T}_X^\perp$. Since (X^*, λ^*, S^*) is an AHO-nondegenerate solution by assumption, it follows that $\Delta X = 0$ and $\Delta S = 0$. This shows that (X^*, λ^*, S^*) is KN-nondegenerate. □

Somewhat to the surprise of the authors, it turned out that the reverse of Theorem 5.5 is also true. We state the resulting equivalence in the following theorem.

Theorem 5.6 *Let (X^*, λ^*, S^*) be a solution of the optimality conditions* (1). *Then (X^*, λ^*, S^*) is AHO-nondegenerate if and only if it is KN-nondegenerate.*

Proof. In view of Theorem 5.5, we only have to show that KN-nondegeneracy implies AHO-nondegeneracy. Therefore, let (X^*, λ^*, S^*) be a KN-nondegenerate solution of (1). We proceed by contradiction and assume that (X^*, λ^*, S^*) is not AHO-nondegenerate. Then there exists a pair $(\Delta X, \Delta S) \neq (0,0)$ such that $\Delta X \in \mathcal{N} \cap \mathcal{T}_S^\perp$ and $\Delta S \in \mathcal{N}^\perp \cap \mathcal{T}_X^\perp$. Hence we have $\Delta X \neq 0$ or $\Delta S \neq 0$.

First consider the case where $\Delta X \neq 0$. Then $(\overline{\Delta X}, \overline{\Delta S}) := (\Delta X, 0)$ is also a nontrivial pair of matrices with $\overline{\Delta X} \in \mathcal{N} \cap \mathcal{T}_S^\perp$ and $\overline{\Delta S} \in \mathcal{N}^\perp \cap \mathcal{T}_X^\perp$. Let us define

$$\widetilde{\Delta X} := Q^T\overline{\Delta X}Q = Q^T \Delta X Q \quad \text{and} \quad \widetilde{\Delta S} := Q^T\overline{\Delta S}Q = 0,$$

and partition $\widetilde{\Delta X}$ as in (6). Since $\overline{\Delta X} = \Delta X \in \mathcal{T}_S^\perp$, we have

$$\widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\beta} & 0 \\ \widetilde{\Delta X}_{\alpha\beta}^T & \widetilde{\Delta X}_{\beta\beta} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We now introduce two diagonal matrices Σ_-, Σ_+ with elements $\sigma_{ij}^-, \sigma_{ij}^+ \, (1 \le j \le i \le n)$, respectively, being defined by

$$\sigma_{ij}^- \ := \ \begin{cases} 0, & \text{if } (i,j) \in (\alpha \times \alpha) \cup (\alpha \times \beta) \cup (\beta \times \alpha) \cup (\beta \times \beta), \\ 1, & \text{if } (i,j) \in (\beta \times \gamma) \cup (\gamma \times \beta) \cup (\gamma \times \gamma), \\ 1, & \text{if } (i,j) \in (\alpha \times \gamma) \cup (\gamma \times \alpha), \end{cases}$$

$$\sigma_{ij}^{+} \;:=\; \begin{cases} 1, & \text{if } (i,j) \in (\alpha\times\alpha)\cup(\alpha\times\beta)\cup(\beta\times\alpha)\cup(\beta\times\beta), \\ 0, & \text{if } (i,j) \in (\beta\times\gamma)\cup(\gamma\times\beta)\cup(\gamma\times\gamma), \\ 1, & \text{if } (i,j) \in (\alpha\times\gamma)\cup(\gamma\times\alpha). \end{cases}$$

Then Σ_-, Σ_+ satisfy (36), and it follows immediately that we have

$$\sigma_{ij}^{-}\widetilde{\Delta X}_{ij} + \sigma_{ij}^{+}\widetilde{\Delta S}_{ij} = 0 \quad \forall 1 \le j \le i \le n$$

or, equivalently,

$$\Sigma_- \operatorname{svec}(\widetilde{\Delta X}) + \Sigma_+ \operatorname{svec}(\widetilde{\Delta S}) = 0.$$

As in the proof of Theorem 5.5, it is not difficult to see that this may be rewritten as

$$(\mathcal{V}^*)\Sigma_-(\mathcal{V}^*)^T \operatorname{svec}(\overline{\Delta X}) + (\mathcal{V}^*)\Sigma_+(\mathcal{V}^*)^T \operatorname{svec}(\overline{\Delta S}) = 0. \tag{48}$$

Since we also have $\overline{\Delta X} = \Delta X \in \mathcal{N}$ and $\overline{\Delta S} = 0 \in \mathcal{N}^\perp$, there exists a vector $\Delta\lambda \in \mathbb{R}^m$ such that

$$\sum_{i=1}^{m} \Delta\lambda_i A_i + \overline{\Delta S} = 0 \quad \text{and} \quad A_i \bullet \overline{\Delta X} = 0 \;\; \forall i = 1, \ldots, m. \tag{49}$$

However, (X^*, λ^*, S^*) is KN-nondegenerate, hence we immediately obtain from (48) and (49) that $\Delta X = \overline{\Delta X} = 0$, contradicting our assumption that $\Delta X \neq 0$.

In a similar way, we can derive a contradiction in the case where $\Delta S \neq 0$ by taking the nontrivial pair $(\overline{\Delta X}, \overline{\Delta S}) := (0, \Delta S)$ instead of $(\overline{\Delta X}, \overline{\Delta S}) := (\Delta X, 0)$, and by using suitably defined matrices Σ_-, Σ_+. Hence, in either case, we get a contradiction, completing the proof. □

Theorem 5.6 implies, for example, that the nonsmooth Newton method from [12] for the solution of semidefinite programs is locally quadratically convergent under the AHO-nondegeneracy condition.

We close this section with the following remark.

Remark 5.7 Following the discussion in [12], it is possible to show that the X^*- and S^*-components of a KN-nondegenerate solution (X^*, λ^*, S^*) of the optimality conditions (1) are unique. In view of our previous results, this implies that the X^*- and S^*-components of a KSS- or AHO-nondegenerate solution is also unique. In this way, we reobtain the statement of Theorem 3.5. Moreover, under the additional assumption that the matrices A_i are linearly independent, we also obtain the uniqueness statements given in [2] for AHO-nondegenerate solutions.

6 Final Remarks

This paper gives an in-depth treatment of three nondegeneracy concepts related to semidefinite programs. It shows that all three concepts are identical under strict complementarity, whereas one of them cannot hold without this assumption and the other two are still equivalent in the degenerate case. The latter result is somewhat surprising since the more geometric definition of the nondegeneracy condition from [2] is completely different from the more algebraic definition of the nondegeneracy concept from [12]. Nevertheless, this equivalence is quite useful for a better understanding of the nondegeneracy condition from [12]. We also feel that it should be possible to exploit results from nonsmooth analysis (which is the basic background where the definition in [12] comes from) in order to obtain some additional properties and other sufficient conditions for the two nondegeneracy conditions from [2, 12] to hold. We leave this as an open question for our future research.

References

[1] F. Alizadeh: *Interior point methods in semidefinite programming with applications to combinatorial optimization.* SIAM Journal on Optimization 5, 1995, pp. 12–51.

[2] F. Alizadeh, J.-P. Haeberly, and M.L. Overton: *Complementarity and nondegeneracy in semidefinite programming.* Mathematical Programming 77, 1997, pp. 111–128.

[3] F. Alizadeh, J.-P. Haeberly, and M.L. Overton: *Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results.* SIAM Journal on Optimization 8, 1998, pp. 746–768.

[4] X. Chen and P. Tseng: *Non-interior continuation methods for solving semidefinite complementarity problems.* Mathematical Programming 95, 2003, pp. 431–474.

[5] R.W. Freund and F. Jarre: *A sensitivity result for semidefinite programs.* Operations Research Letters 32, 2004, pp. 126–132.

[6] J.P. Haeberly: *Remarks on nondegeneracy in mixed semidefinite-quadratic programming.* Technical Report, Department of Mathematics, Fordham University, September 1998.

[7] C. Helmberg, F. Rendl, R. Vanderbei, and H. Wolkowicz: *An interior-point method for semidefinite programming*, SIAM Journal on Optimization 6, 1996, pp. 342–361.

[8] R.A. Horn and C.R. Johnson: *Matrix Analysis.* Cambridge University Press, 1985.

[9] R.A. Horn and C.R. Johnson: *Topics in Matrix Analysis.* Cambridge University Press, 1991.

[10] C. Kanzow and C. Nagel: *Semidefinite programs: New search directions smoothing-type methods, and numerical results.* SIAM Journal on Optimization 13, 2003, pp. 1–23 (Corrigendum in SIAM Journal on Optimization 14, 2004, pp. 936–937).

[11] C. Kanzow and C. Nagel: *Some structural properties of a Newton-type method for semidefinite programs.* Journal of Optimization Theory and Applications 122, 2004, pp. 219–226.

[12] C. Kanzow and C. Nagel: *Quadratic convergence of a nonsmooth Newton-type method for semidefinite programs without strict complementarity.* SIAM Journal on Optimization, to appear.

[13] M. Kojima, M. Shida, and S. Shindoh: *Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs.* Mathematical Programming 80, 1998, pp. 129–160.

[14] M. Kojima, M. Shida, and S. Shindoh: *A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberly-Overton search direction.* SIAM Journal on Optimization 9, 1999, pp. 444–465.

[15] S.A. Miller: *Sensitivity of solutions to semidefinite programs.* Technical Report CCEC-97-0519, Center for Control Engineering and Computation, University of California, Santa Barbara, CA, May 1997.

[16] R.D.C. Monteiro: *Polynomial convergence of primal-dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions.* SIAM Journal on Optimization 8, 1998, pp. 797–812.

[17] M.V. Nayakkankuppam and M.L. Overton: *Conditioning of semidefinite programs.* Mathematical Programming 85, 1999, pp. 525–540.

[18] Y. Nesterov and M.J. Todd: *Primal-dual interior-point methods for self-scaled cones.* SIAM Journal on Optimization 8, 1998, pp. 324–364.

[19] D. Sun and J. Sun: *Semismooth matrix valued functions.* Mathematics of Operations Research 27, 2002, pp. 150–169.

[20] J. Sun, D. Sun, and L. Qi: *A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems.* SIAM Journal on Optimization 14, 2004, pp. 783–806.

[21] M.J. Todd, K.C. Toh, and R.H. Tütüncü: *On the Nesterov-Todd direction in semidefinite programming.* SIAM Journal on Optimization 8, 1998, pp. 769–796.

[22] P. Tseng: *Merit functions for semi-definite complementarity problems.* Mathematical Programming 83, 1998, pp. 159–185.

[23] Y. Zhang: *On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming.* SIAM Journal on Optimization 8, 1998, pp. 365–386.