Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 400 (2005) 31–60
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Strengthened existence and uniqueness conditions for search directions in semidefinite programming

Levent Tunçel [1], Henry Wolkowicz [*,2]

Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Ont., Canada N2L 3G1

Received 30 July 2003; accepted 3 December 2004

Submitted by V. Mehrmann

Abstract

Primal–dual interior-point (p–d i-p) methods for Semidefinite Programming (SDP) are generally based on solving a system of matrix equations for a Newton type search direction for a symmetrization of the optimality conditions. These search directions followed the derivation of similar p–d i-p methods for linear programming (LP). Among these, a computationally interesting search direction is the AHO direction. However, in contrast to the LP case, existence and uniqueness of the AHO search direction is not guaranteed under the standard nondegeneracy assumptions. Two different sufficient conditions are known that guarantee the existence and uniqueness independent of the specific linear constraints. The first is given by Shida–Shindoh–Kojima and is based on the semidefiniteness of the symmetrization of the product SX at the current iterate. The second is a centrality condition given first by Monteiro–Zanjácomo and then improved by Monteiro–Todd.

In this paper, we revisit and strengthen both of the above mentioned sufficient conditions. We include characterizations for existence and uniqueness in the matrix equations arising

* Corresponding author. Tel.: +1 519 888 4567x5589.

E-mail addresses: ltuncel@math.uwaterloo.ca (L. Tunçel), hwolkowicz@uwaterloo.ca (H. Wolkowicz).

URL: http://www.orion.math.uwaterloo.ca/~hwolkowi/henry/reports/ABSTRACTS.html.

[1] Research supported by The Natural Sciences and Engineering Research Council of Canada and a PREA.

[2] Research supported by The Natural Sciences and Engineering Research Council of Canada.

doi:10.1016/j.laa.2004.12.010

from the linearization of the optimality conditions. As well, we present new results on the relationship between the Kronecker product and the *symmetric Kronecker product* that arise from these matrix equations. We conclude with a proof that the existence and uniqueness of the AHO direction is a generic property for every SDP problem and extend the results to the general Monteiro–Zhang family of search directions.

AMS classification: 90C22; 49K40; 65K05; 15A48; 15A24

Keywords: Semidefinite programming; Linear matrix equations; Existence and uniqueness; Kronecker product; Symmetric Kronecker product; Linear transformations

1. Introduction

Semidefinite Programming (SDP) has generated tremendous interest during the last 15 years, e.g. [31], both for the many applications and the mathematical elegance. Many of the early interesting algorithms were primal–dual interior-point (p–d i-p) methods based on solving a system of matrix equations for Newton type search directions. The search directions, after a symmetrization, followed the derivation of similar p–d i-p methods for linear programming (LP), mainly based on the hope to extend the computational and/or theoretical properties of successful p–d i-p methods for LP. However, in contrast to the LP case, existence and uniqueness of some of these search directions were not guaranteed under the standard nondegeneracy assumptions: (i) *the linear transformation $\mathscr{A}$ from the linear constraints is onto* and (ii) *Slater's constraint qualification* (*strict feasibility with respect to the cone constraints*) *holds for both primal and dual programs*. (We assume that condition (i) holds throughout this paper.)

There are many search directions proposed for p–d i-p methods for semidefinite programming. For an account up to 1997, see [11,26]. Some of the proposals describe a set of search directions such as Kojima, Shindoh and Hara [12] (KSH family), Zhang [32], Monteiro and Zhang [19] (denoted MZ family), Tunçel [29]. Another general approach to search directions is to compute the Gauss–Newton direction for the overdetermined optimality conditions, see Kruk et al. [13].

Among the earliest proposals were the HrvwKshM direction [9,12,15], NT direction [21], and the AHO direction [1]. For each value of the central path parameter (or the barrier parameter), $\mu > 0$, these search directions, under standard nondegeneracy assumptions, are found by solving a particular (symmetrized) linear system of matrix equations. The HrvwKshM and NT directions have the disadvantage that the linear system becomes ill-conditioned as μ approaches zero (it is singular for $\mu = 0$); while the linear system for the AHO direction is nonsingular at $\mu = 0$. On the other hand, the HrvwKshM and NT directions are well-defined for every pair of primal–dual interior-points, though the AHO direction is not.

Two different sufficient conditions that guarantee existence and uniqueness of the AHO direction are given in [24] (denoted SSK condition) and [18] (denoted MZ condition; this was improved later in [17]). Both of these conditions assume that the primal and dual matrices X, S are positive definite. Both conditions depend only on X, S and not on the data from the linear constraints of the SDP. Further conditions and proofs for the AHO direction are given in [28]; conditions for the Monteiro–Zhang family of search directions are given in e.g. [16], [27, Theorem 3.1] and [24].

In this paper we revisit the two sufficient conditions cited above. We provide several strengthened conditions as well as characterizations. In particular, we provide the strongest conditions using only the data matrices X, S; see, Theorems 3.9, 3.11 and 3.18. The results deal with (symmetric) matrix equations and, therefore, the Kronecker and the *symmetric Kronecker* products are involved. Though the Kronecker product has been extensively studied in the literature, this is not true for the (constrained version) symmetric Kronecker product. In Theorems 2.8 and 2.9, we present the new results that the positive semidefiniteness of the symmetric Kronecker product is equivalent to the positive semidefiniteness of the Kronecker product and that the eigenvalues have a special relationship. We include topological properties on the solutions, e.g. that the existence and uniqueness of the AHO direction is generic. We conclude by extending the results to the general Monteiro–Zhang family.

The paper is organized as follows. After introducing some notation, we outline the main results in Section 1.1. We include several illustrative examples. The Kronecker product is discussed in Section 2.2. The characterizations for existence and uniqueness are given in Section 3.1. The strengthened sufficient conditions of SSK type are given in Section 3.2.1. The strengthened sufficient conditions of MZ type are given in Sections 3.2.2 and 3.2.3. We include results on topological properties and genericity in Section 4 including the genericity results. We conclude by extending the results to the general Monteiro–Zhang family of search directions in Section 5.

1.1. Preliminaries; outline of results

1.1.1. Notation

We use the following standard notation: $\mathcal{S}^n$ is the space of $n \times n$ real symmetric matrices equipped with the trace inner-product for two matrices X, Y,

$$\langle X, Y\rangle := \operatorname{trace} X^{\mathrm{T}} Y;$$

$C \in \mathcal{S}^n$; the matrices $A_i \in \mathcal{S}^n$, $i = 1, \ldots, m$, are linearly independent; and $b \in \mathbb{R}^m$. The dimension of $\mathcal{S}^n$ is $t(n) := n(n+1)/2$. $\mathcal{M}^n$ denotes the space of $n \times n$ matrices over the reals. For $M \in \mathcal{M}^n$ with real eigenvalues,

$$\lambda_1(M) \geqslant \lambda_2(M) \geqslant \cdots \geqslant \lambda_n(M)$$

denote its eigenvalues. We also use $\lambda_{\max} := \lambda_1$ and $\lambda_{\min} := \lambda_n$. In the space $\mathcal{M}^n$, $\|\cdot\|$ denotes the operator 2-norm and $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenious norm (thus, $\|M\|_{\mathrm{F}} = \langle M, M\rangle^{1/2}$). We use $X \in \mathcal{S}^n_+$ or $X \succeq 0$ (respectively, $X \in \mathcal{S}^n_{++}$ or $X \succ 0$)

to denote symmetric positive semidefiniteness (and symmetric positive definiteness, respectively). Then $(X, y, S) \in \mathscr{S}^n_+ \times \mathbb{R}^m \times \mathscr{S}^n_+$ are the primal and dual variables for the following SDP and its dual.

$$
\text{(P)} \qquad \begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m, \\ & X \in \mathscr{S}^n_+, \end{array}
$$

$$
\text{(D)} \qquad \begin{array}{ll} \max & b^{\mathrm{T}} y \\ \text{s.t.} & \sum_{i=1}^m y_i A_i + S = C, \\ & S \in \mathscr{S}^n_+. \end{array}
$$

The linear constraints in the primal (P), respectively dual (D), can be written using the linear transformation notation

$$
\mathscr{A} X = b, \quad \mathscr{A}^* y + S = C, \quad X \succeq 0, S \succeq 0,
$$

where $\cdot^*$ denotes the adjoint transformation.

One of the first computationally interesting algorithms for SDP was based on the Alizadeh–Haeberly–Overton (AHO) search direction [1]. At a point $(X, y, S) \in \mathscr{S}^n_+ \times \mathbb{R}^m \times \mathscr{S}^n_+$ with centrality parameter $\sigma \in (0, 1]$ and barrier parameter $0 < \mu := \langle X, S\rangle / n$, this search direction (when it exists) is the solution $(\Delta X, \Delta y, \Delta S) \in \mathscr{S}^n \times \mathbb{R}^m \times \mathscr{S}^n$ of the (symmetrized linearized optimality conditions) linear system

$$
\begin{aligned}
&\sum_{i=1}^m (\Delta y)_i A_i + \Delta S = C - \sum_{i=1}^m y_i A_i - S, \\
&\langle A_i, \Delta X\rangle = b_i - \langle A_i, X\rangle, \quad i = 1, \ldots, m, \\
&X(\Delta S) + (\Delta S)X + S(\Delta X) + (\Delta X)S = 2\sigma\mu I - XS - SX.
\end{aligned} \tag{1.1}
$$

Let $H_P : \mathscr{M}^n \to \mathscr{S}^n$ denote the linear transformation

$$
H_P(M) := PMP^{-1} + P^{-\mathrm{T}} M^{\mathrm{T}} P^{\mathrm{T}},
$$

where P is any $n \times n$ nonsingular matrix. If we replace (1.1) by

$$
\begin{aligned}
&\sum_{i=1}^m (\Delta y)_i A_i + \Delta S = C - \sum_{i=1}^m y_i A_i - S, \\
&\langle A_i, \Delta X\rangle = b_i - \langle A_i, X\rangle, \quad i = 1, \ldots, m, \\
&H_P(X(\Delta S) + (\Delta X)S) = H_P(\sigma\mu I - XS),
\end{aligned} \tag{1.2}
$$

then the solutions $(\Delta X, \Delta y, \Delta S) \in \mathscr{S}^n \times \mathbb{R}^m \times \mathscr{S}^n$ are in the Monteiro–Zhang family of search directions. If we choose $P = I$, then we get the AHO direction. The choice $P = S^{\frac{1}{2}}$ gives the so-called HrvwKshM direction [9,12,15]; while any P with $P^{\mathrm{T}} P = X^{-\frac{1}{2}} (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{\frac{1}{2}} X^{-\frac{1}{2}}$, e.g., $P = (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{\frac{1}{4}} X^{-\frac{1}{2}}$ gives the NT direction [21].

In this paper we focus on conditions for the AHO direction. We then show in Section 5 that these conditions directly extend to the complete Monteiro–Zhang family of directions.

We now look at two classes of sufficient conditions that guarantee existence and uniqueness of solutions of (1.1).

1.1.2. Shida–Shindoh–Kojima (SSK) type sufficient conditions

It has been shown by Shida, Shindoh and Kojima [24] (see also [27]) that system (1.1) has a unique solution, and hence that the AHO search direction is well-defined, whenever

$$\text{(SSKcond)} \quad X, S \succ 0 \text{ and } XS + SX \succeq 0. \tag{1.3}$$

We strengthen the (SSKcond) condition, see Theorem 3.9: AHO is well-defined whenever

$$\begin{aligned} K := {} & [I \otimes (SX + XS)] + [(SX + XS) \otimes I] \\ & +2[X \otimes S] + 2[S \otimes X] \text{ is definite,} \end{aligned} \tag{1.4}$$

where $\otimes$ denotes the Kronecker product. We use K and derive the following sufficient condition for existence and uniqueness independent of $\mathscr{A}$, i.e. from (3.18), the sufficiency of (1.4) is implied by the sufficiency of the condition

$$L := K + TK + KT \text{ is semidefinite with rank } n(n+1)/2, \tag{1.5}$$

where T is the matrix representation of the transpose operator, see (2.10). (The matrix L in (1.5) has at least $n(n-1)/2$ zero eigenvalues, for any K. Therefore we can use the stable condition that rank $\geqslant n(n+1)/2$ implies (1.5).) These two new sufficient conditions have the advantage that they hold on an open set. Note that the set where existence and uniqueness holds is open (see Section 4 for related results).

1.1.3. Monteiro–Zanjácomo (MZ) type sufficient conditions

A different, centrality type, sufficient condition for existence and uniqueness is given by Monteiro and Zanjácomo [18],

$$\text{(MZcond)} \quad X, S \succ 0 \text{ and } \|S^{\frac{1}{2}}XS^{\frac{1}{2}} - \nu I\| \leqslant \tfrac{1}{2}\nu \quad \text{for some } \nu > 0, \tag{1.6}$$

see Theorem 3.14. We note below that this latter condition is equivalent to the bound on the condition number

$$\text{(MZequivalent)} \quad X, S \succ 0 \text{ and } \gamma(S^{\frac{1}{2}}XS^{\frac{1}{2}}) \leqslant 3; \tag{1.7}$$

see Proposition 1.1.

Proposition 1.1. *Let $S, X \in \mathscr{S}^n$, $S \succ 0$ and $\alpha \in \mathbb{R}_+$. Then the condition $\|S^{\frac{1}{2}}XS^{\frac{1}{2}} - \nu I\| \leqslant \alpha\nu$ for some scalar $\nu > 0$, holds if and only if*

$$(1-\alpha)\nu \leqslant \lambda_{\min}(S^{\frac{1}{2}}XS^{\frac{1}{2}}) \quad \textit{and} \quad \lambda_{\max}(S^{\frac{1}{2}}XS^{\frac{1}{2}}) \leqslant (1+\alpha)\nu. \tag{1.8}$$

Proof. Let $S^{\frac{1}{2}}XS^{\frac{1}{2}} := PDP^{\mathrm{T}}$ for some diagonal matrix D and some unitary matrix P. Then D contains the eigenvalues of $S^{\frac{1}{2}}XS^{\frac{1}{2}}$. Moreover, $\|D - \nu I\| \leqslant \alpha\nu$ and the condition $|\lambda_i(S^{\frac{1}{2}}XS^{\frac{1}{2}}) - \nu| \leqslant \alpha\nu$ for all $i = 1, \ldots, n$ is equivalent to (1.8). □

Moreover, we strengthen these results. The constant $\frac{1}{2}$ in (1.6) was strengthened to $\frac{1}{\sqrt{2}}$ in [17]. (This implies that 3 in (1.7) increases to $3 + 2\sqrt{2}$.) We show that the techniques of Monteiro and Todd in (1.6) can be used to further improve the constant $\frac{1}{\sqrt{2}}$ to $\sqrt{3} - 1$. (And $3 + 2\sqrt{2}$ increases to $3 + 2\sqrt{3}$.) Furthermore, we show that the approach of Monteiro and Zanjácomo [17] can be used to further strengthen the constant in (1.6) and the equivalent condition number bound in (1.7), i.e. we get

$$\tfrac{4}{5} \text{ in (1.6), } 9 \text{ in (1.7), respectively,}$$

see Theorem 3.11. We also show that the constant $\frac{1}{2}$ in (1.6) cannot be improved beyond 0.9193, see Example 1.5. (A previous bound in this direction was 0.9837, due to Tseng [28].)

1.1.4. Comparison examples

However, the relationships between the different conditions are not clear. Some of these properties are seen in the following examples.

Example 1.2. From [27, p. 778], we see that $m := 1, n := 2$,

$$A_1 := \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}, \quad b := 3, \quad X := \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix}, \quad S := \begin{pmatrix} 1 & 0 \\ 0 & 11 \end{pmatrix}$$

provide an example where the AHO direction does not exist. Here $\mu := \frac{\langle X,S\rangle}{2} = 17$ and $\|S^{\frac{1}{2}}XS^{\frac{1}{2}} - \mu I\| = \sqrt{278}$, while the MZ result, Theorem 3.14, requires the RHS to be at most $\frac{17}{2}$. The condition number of $S^{\frac{1}{2}}XS^{\frac{1}{2}}$ is $\frac{17+\sqrt{278}}{17-\sqrt{278}} \cong 103.081$. Note that $XS + SX$ is not positive semidefinite.

Example 1.3. From [18], we see that

$$X := \begin{pmatrix} 8 & 1 \\ 1 & .5 \end{pmatrix}, \quad S := \begin{pmatrix} 1 & 0 \\ 0 & 16 \end{pmatrix}$$

provide an example where (MZcond) holds but (SSKcond) fails, i.e. $S^{\frac{1}{2}}XS^{\frac{1}{2}} = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$ has eigenvalues 4, 12 and condition number 3. But $XS + SX = \begin{pmatrix} 16 & 17 \\ 17 & 16 \end{pmatrix}$ is not positive semidefinite. (Though the new condition (1.4) holds.)

Example 1.4. From [18], we see that

$$X := \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}, \quad S := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

provide an example where (SSKcond) holds but (MZcond) fails, i.e. the condition number of $S^{\frac{1}{2}}XS^{\frac{1}{2}}$ is 10, while $XS + SX$ is positive semidefinite.

Example 1.5. Using a special parametrized grid search, we found the example $m := 1, n := 2$,

$$A_1 := \begin{pmatrix} -1 & a \\ a & d \end{pmatrix}, \quad a := -11.7078, \quad d := -40.2704,$$

$$X := Q^{\mathrm{T}} \begin{pmatrix} 30 & 0 \\ 0 & 1 \end{pmatrix} Q, \quad Q := \begin{pmatrix} \beta & -\sqrt{1-\beta^2} \\ \sqrt{1-\beta^2} & \beta \end{pmatrix},$$

with $\beta := -0.99$ and $S := \begin{pmatrix} 1 & 0 \\ 0 & 273 \end{pmatrix}$. The a, d were found using the solutions of quadratic polynomials based on the parameter $\theta := 273$ in S. We obtained the better condition number value 23.7911, though still greater than 9. Also, this example shows that the constant $\frac{1}{2}$ in (1.6) cannot be improved beyond 0.9193.

2. Linear algebra results

2.1. Subspaces and linear operators

Definition 2.1. Define the linear (Lyapunov) operator on $\mathscr{S}^n$:

$$L_U(V) := UV + VU \quad \text{for } U, V \in \mathscr{S}^n.$$

To investigate whether (1.1) has a unique solution, we choose to study the related condition on the corresponding homogeneous system in the variables $(\Delta X, \Delta y, \Delta S) \in \mathscr{S}^n \times \mathbb{R}^m \times \mathscr{S}^n$.

$$\begin{aligned} &\mathscr{A}^*(\Delta y) + (\Delta S) = 0, \\ &\mathscr{A}(\Delta X) = 0, \\ &L_S(\Delta X) + L_X(\Delta S) = 0. \end{aligned} \tag{2.1}$$

We single out the third equation

$$L_S(\Delta X) + L_X(\Delta S) = 0. \tag{2.2}$$

Equivalently, we need to establish nonsingularity of the linear operator with the block structure

$$\mathscr{K} := \begin{pmatrix} 0 & \mathscr{A}^* & I \\ \mathscr{A} & 0 & 0 \\ L_S & 0 & L_X \end{pmatrix}. \tag{2.3}$$

Define the one–one linear transformation $\mathscr{Z} : \mathbb{R}^{t(n)-m} \to \mathscr{S}^n$ so that the range space of $\mathscr{Z}$ is equal to the nullspace of $\mathscr{A}$, i.e.

$$\mathscr{R}(\mathscr{Z}) = \mathscr{N}(\mathscr{A}), \tag{2.4}$$

or equivalently

$$\mathscr{A}(\Delta X) = 0 \text{ iff } \Delta X = \mathscr{Z}(\Delta v) \text{ for some } \Delta v \in \mathbb{R}^{t(n)-m}.$$

Then, we can substitute for $\Delta X, \Delta S$ in (2.1) and see that nonsingularity of $\mathscr{K}$ is equivalent to nonsingularity of the linear operator (on $\mathbb{R}^{t(n)}$)

$$\bar{\mathscr{K}} := \left[L_S \mathscr{Z} \vdots - L_X \mathscr{A}^*\right], \tag{2.5}$$

i.e. $\bar{\mathscr{K}}$ acts on $\begin{pmatrix} \Delta v \\ \Delta y \end{pmatrix}$.

Lemma 2.2 (Also see Lemma 10.4.9 of [17]). *Suppose that $X \succ 0$. Then, the linear operator L_X is one–one, onto, self-adjoint, and positive definite.*

Proof. Let $U, V \in \mathscr{S}^n$. That

$$\langle XU + UX, V\rangle = \langle U, XV\rangle + \langle U, VX\rangle = \langle XV + VX, U\rangle$$

shows L_X is self-adjoint. That

$$\langle U, L_X U\rangle = 2\langle U, XU\rangle = 2\|UX^{\frac{1}{2}}\|_F^2 \geqslant 0$$

(the last inequality is strict, unless $U = 0$) shows that L_X is positive definite. Now, $L_X(U) = 0$ iff $U = 0$. Therefore, L_X is one-one and onto. □

2.2. *Kronecker and symmetric Kronecker algebra*

The two linear operators L_X, L_S have been studied in the literature in relation to matrix equations, e.g. [2], [10, Section 4.4] and [3,5,6]. The main tool is the Kronecker product. Some basic facts are given in [27, Appendix]. We also include the relevant definitions and properties that we need. (Further properties are given in e.g. [10,23,30].) We include new results on the *symmetric Kronecker product*.

It is well known that, for compatible matrices K, N, M, we can find the matrix representations of certain linear transformations on $\mathscr{M}^n$ (identified with $\mathbb{R}^{n^2}$) using the Kronecker product, i.e.

$$\operatorname{vec}(NKM^{\mathrm{T}}) = (M \otimes N)\operatorname{vec}(K), \tag{2.6}$$

where $\operatorname{vec}(K)$ is the vector formed from the columns of K and $\otimes$ denotes the *Kronecker product*. If λ, μ are the eigenvalues with corresponding eigenvectors v, w for M, N, respectively, then an eigenvalue/eigenvector pair is given in

$$(M \otimes N)(v \otimes w) = \lambda\mu(v \otimes w); \tag{2.7}$$

see e.g. [10, Theorem 4.2.12]. For compatible matrices,

$$(A \otimes B)(C \otimes D) = (AC \otimes BD); \quad (A \otimes B)^{\mathrm{T}} = (A^{\mathrm{T}} \otimes B^{\mathrm{T}}).$$

We have the following fact.

Lemma 2.3. *Let A, B, C be $n \times n$ matrices. Then*

$$2\langle A^{\mathrm{T}}, BC\rangle = [\mathrm{vec}(B)^{\mathrm{T}} | \mathrm{vec}(C)^{\mathrm{T}}] \left[\begin{array}{c|c} 0 & I \otimes A^{\mathrm{T}} \\ \hline I \otimes A & 0 \end{array}\right] \left[\begin{array}{c} \mathrm{vec}(B) \\ \hline \mathrm{vec}(C) \end{array}\right].$$

Proof. We compute

$$\begin{aligned}\langle A^{\mathrm{T}}, BC\rangle &= \mathrm{vec}(C)^{\mathrm{T}}\mathrm{vec}(ABI) \\ &= \mathrm{vec}(C)^{\mathrm{T}}(I \otimes A)\mathrm{vec}(B) \\ &= \mathrm{vec}(B)^{\mathrm{T}}(I \otimes A)^{\mathrm{T}}\mathrm{vec}(C) \\ &= \mathrm{vec}(B)^{\mathrm{T}}(I \otimes A^{\mathrm{T}})\mathrm{vec}(C).\end{aligned}$$

Therefore, using the second and the fourth equations, we obtain the desired result

$$2\langle A^{\mathrm{T}}, BC\rangle = \mathrm{vec}(C)^{\mathrm{T}}(I \otimes A)\mathrm{vec}(B) + \mathrm{vec}(B)^{\mathrm{T}}(I \otimes A^{\mathrm{T}})\mathrm{vec}(C). \qquad \square$$

Recall $t(n) = n(n+1)/2$. Let $\mathrm{s2vec}(X) \in \mathbb{R}^{t(n)}$ denote the isometry between $\mathcal{S}^n$ and $\mathbb{R}^{t(n)}$ that takes the lower triangular part of X columnwise and multiplies the strict lower triangular part by $\sqrt{2}$. Then the inner-product in $\mathcal{S}^n$ can be expressed as the vector inner-product $\langle X, Y\rangle = \mathrm{s2vec}(X)^{\mathrm{T}}\mathrm{s2vec}(Y)$. In addition, we can express the matrix representations of certain linear transformations on $\mathcal{S}^n$ (identified with $\mathbb{R}^{t(n)}$) using the *symmetric Kronecker product*. For $U \in \mathcal{S}^n$ we write, see [2,27],

$$(M \overset{s}{\otimes} N)\mathrm{s2vec}(U) := \mathrm{s2vec}\left(\tfrac{1}{2}(NUM^{\mathrm{T}} + MUN^{\mathrm{T}})\right).$$

(We have changed the notation used in [2].) If $M, N \in \mathcal{M}^n$, then the matrix $(M \overset{s}{\otimes} N) \in \mathcal{M}^{t(n)}$, while if $A, B \in \mathcal{S}^n$, then $(A \overset{s}{\otimes} B) \in \mathcal{S}^{t(n)}$. Note that by definition, $M \overset{s}{\otimes} N = N \overset{s}{\otimes} M$ and it is easily checked that $(M \overset{s}{\otimes} N)^{\mathrm{T}} = M^{\mathrm{T}} \overset{s}{\otimes} N^{\mathrm{T}}$.

We can express $\overset{s}{\otimes}$ using $\otimes$; we consider $\mathcal{S}^n$ isomorphic to $\mathbb{R}^{t(n)}$. Let $Q \in \mathbb{R}^{t(n) \times n^2}$ be defined as follows:

$$Q_{(i,j),(k,l)} := \begin{cases} 1 & \text{if } i = j = k = l, \\ \frac{1}{\sqrt{2}} & \text{if } i = k \neq j = l \text{ or } i = l \neq j = k, \\ 0 & \text{otherwise;} \end{cases}$$

i.e. the columns of Q^{T} consist of $\mathrm{vec}(P_{ij})$, where $\{P_{ij} : 1 \leqslant i \leqslant j \leqslant n\}$ is the set of $t(n)$ orthonormal basis matrices of $\mathcal{S}^n$. Then $QQ^{\mathrm{T}} = I$, $Q^{\mathrm{T}}Q$ is the matrix representation of the orthogonal projection of $\mathcal{M}^n$ (as $\mathbb{R}^{n^2}$) onto $\mathcal{S}^n$ (as a subspace of $\mathbb{R}^{n^2}$). Let $u := Q^{\mathrm{T}}v$, $v \in \mathbb{R}^{t(n)}$, $U := \mathrm{Mat}(u)$, i.e. $U \in \mathcal{S}^n$ and $u = \mathrm{vec}(U)$. The quadratic form with $U \in \mathcal{S}^n$ is

$$\begin{aligned}\tfrac{1}{2}\langle U, NUM^{\mathrm{T}} + MUN^{\mathrm{T}}\rangle &= \tfrac{1}{2}u^{\mathrm{T}}[(N \otimes M) + (M \otimes N)]u \\ &= \tfrac{1}{2}v^{\mathrm{T}}Q[(N \otimes M) + (M \otimes N)]Q^{\mathrm{T}}v \\ &= \mathrm{s2vec}(U)^{\mathrm{T}}(M \overset{s}{\otimes} N)\mathrm{s2vec}(U) \\ &= \mathrm{s2vec}(U)^{\mathrm{T}}(M \overset{s}{\otimes} N)^{\mathrm{T}}\mathrm{s2vec}(U) \\ &= \mathrm{s2vec}(U)^{\mathrm{T}}(M^{\mathrm{T}} \overset{s}{\otimes} N^{\mathrm{T}})\mathrm{s2vec}(U) \\ &= \mathrm{s2vec}(U)^{\mathrm{T}}(N \overset{s}{\otimes} M)\mathrm{s2vec}(U). \end{aligned} \tag{2.8}$$

A similar derivation with $v_1, v_2 \in \mathbb{R}^{t(n)}$, $u_i := Q^{\mathrm{T}}v_i$, $U_i := \mathrm{Mat}(u_i)$ yields

$$M \overset{s}{\otimes} N = \tfrac{1}{2}Q(M \otimes N + N \otimes M)Q^{\mathrm{T}} = Q(M \otimes N)Q^{\mathrm{T}}. \tag{2.9}$$

We need a corresponding result to Lemma 2.3.

Lemma 2.4. *Let $A \in \mathscr{M}^n$ and $B, C \in \mathscr{S}^n$. Then*

$$2\langle A^{\mathrm{T}}, BC\rangle = [\mathrm{s2vec}(B)^{\mathrm{T}} | \mathrm{s2vec}(C)^{\mathrm{T}}] \left[\begin{array}{c|c} 0 & I \overset{s}{\otimes} A^{\mathrm{T}} \\ \hline I \overset{s}{\otimes} A & 0 \end{array}\right] \left[\begin{array}{c} \mathrm{s2vec}(B) \\ \hline \mathrm{s2vec}(C) \end{array}\right].$$

Proof. We compute

$$\begin{aligned}\langle A^{\mathrm{T}}, BC\rangle &= \tfrac{1}{2}\mathrm{trace}\, C(ABI^{\mathrm{T}} + IBA^{\mathrm{T}}) \\ &= \tfrac{1}{2}\mathrm{s2vec}(C)^{\mathrm{T}}\mathrm{s2vec}(ABI^{\mathrm{T}} + IBA^{\mathrm{T}}) \\ &= \mathrm{s2vec}(C)^{\mathrm{T}}(I \overset{s}{\otimes} A)\mathrm{s2vec}(B) \\ &= \mathrm{s2vec}(B)^{\mathrm{T}}(I \overset{s}{\otimes} A)^{\mathrm{T}}\mathrm{s2vec}(C) \\ &= \mathrm{s2vec}(B)^{\mathrm{T}}(I \overset{s}{\otimes} A^{\mathrm{T}})\mathrm{s2vec}(C). \end{aligned}$$

In the first equation we used $B, C \in \mathscr{S}^n$. Using the third and the fifth equations, we conclude

$$2\langle A^{\mathrm{T}}, BC\rangle = \mathrm{s2vec}(C)^{\mathrm{T}}(I \overset{s}{\otimes} A)\mathrm{s2vec}(B) + \mathrm{s2vec}(B)^{\mathrm{T}}(I \overset{s}{\otimes} A^{\mathrm{T}})\mathrm{s2vec}(C),$$

as claimed. □

Define the *transpose* operator $\mathscr{T} : \mathscr{M}^n \to \mathscr{M}^n$,

$$\mathscr{T}(M) := M^{\mathrm{T}} \quad \text{with matrix representation } T = (T_{kl}). \tag{2.10}$$

Since $\mathscr{T}(E_{ij}) = E_{ji}$, where E_{ij} is the zero matrix but for 1 in the *ij* position, we see that the $l = i + (j-1)n$ column of *T* is the all zero vector except for a 1 in the $k = j + (i-1)n$ position. Note that $\mathscr{T}$ is orthogonal, self-adjoint.

$$\langle \mathscr{T}(M), N\rangle = \operatorname{trace}(M^{\mathrm{T}})^{\mathrm{T}} N = \operatorname{trace} MN = \operatorname{trace}(M^{\mathrm{T}})N^{\mathrm{T}} = \langle M, \mathscr{T}(N)\rangle.$$

Thus the matrix representation satisfies $T = T^{\mathrm{T}}, T^2 = I$.

Instead of considering $\mathscr{S}^n$ isomorphic to $\mathbb{R}^{t(n)}$, we can consider it as a subspace of $\mathscr{M}^n$. Then for each $X \in \mathscr{M}^n$, the orthogonal projection onto $\mathscr{S}^n$ is $U = P_{\mathscr{S}^n} X = \frac{1}{2}(X + X^{\mathrm{T}})$. We can define a symmetric Kronecker product on $\mathscr{M}^n$ as

$$\begin{aligned}(M \overset{T}{\otimes} N)\operatorname{vec}(X) &:= \operatorname{vec}[P_{\mathscr{S}^n}(N P_{\mathscr{S}^n}(X) M^{\mathrm{T}})] \\ &= \operatorname{vec}\left(\tfrac{1}{2}\left[N \tfrac{1}{2}(X + X^{\mathrm{T}})M^{\mathrm{T}} + M \tfrac{1}{2}(X + X^{\mathrm{T}})N^{\mathrm{T}}\right]\right).\end{aligned}$$

The matrix $(M \overset{T}{\otimes} N) \in \mathcal{M}^{n^2}$ with rank the same as $M \overset{s}{\otimes} N$. In fact, the two symmetric Kronecker products agree on $\mathscr{S}^n$ and $\overset{T}{\otimes}$ is identically zero on $(\mathscr{S}^n)^{\perp}$, the space of skew-symmetric matrices in $\mathcal{M}^n$. Identify $x = \operatorname{vec}(X)$. We can expand to get an $n \times n$ matrix representation for the symmetric Kronecker product.

$$\begin{aligned}x^{\mathrm{T}}(M \overset{T}{\otimes} N)x &= \tfrac{1}{4} x^{\mathrm{T}} \operatorname{vec}\left(NXM^{\mathrm{T}} + NX^{\mathrm{T}}M^{\mathrm{T}} + MXN^{\mathrm{T}} + MX^{\mathrm{T}}N^{\mathrm{T}}\right) \\ &= \tfrac{1}{4} x^{\mathrm{T}}\left[(N \otimes M^{\mathrm{T}} + M \otimes N^{\mathrm{T}}) + \tfrac{1}{2}(N \otimes M^{\mathrm{T}} + M \otimes N^{\mathrm{T}})T \right. \\ &\qquad \left. + \tfrac{1}{2} T(N \otimes M^{\mathrm{T}} + M \otimes N^{\mathrm{T}})\right]x \\ &= \tfrac{1}{8} x^{\mathrm{T}}[(I + T)(N \otimes M^{\mathrm{T}} + M \otimes N^{\mathrm{T}}) \\ &\qquad + (N \otimes M^{\mathrm{T}} + M \otimes N^{\mathrm{T}})(I + T)]x.\end{aligned}$$

Thus $\overset{T}{\otimes}$ as an extension of $\overset{s}{\otimes}$ to all of $\mathscr{M}^n$ has the same $t(n)$ eigenvalues and eigenvectors when restricted to $\mathscr{S}^n \subset \mathscr{M}^n$ and has an additional $n(n-1)/2$ zero eigenvalues with eigenvectors corresponding to skew-symmetric matrices. Note that $\mathscr{P} := P_{\mathscr{S}^n}(N P_{\mathscr{S}^n}(\cdot) M^{\mathrm{T}})$ is a linear operator on $\mathscr{M}^n$ with invariant subspace $\mathscr{S}^n \subset \mathscr{M}^n$. In fact, $\mathscr{P}(S) = \lambda S$, $S \in \mathscr{S}^n$, implies that $P_{\mathscr{S}^n}(N(S)M^{\mathrm{T}}) = \lambda S$. This gives a relationship between the eigenvectors of $M \overset{s}{\otimes} N$ and the eigenvectors of $P_{\mathscr{S}^n}(N(\cdot)M^{\mathrm{T}})$. In summary, we can simply write

$$M \overset{T}{\otimes} N = \tfrac{1}{2} Q^{\mathrm{T}} Q(M \otimes N + N \otimes M) Q^{\mathrm{T}} Q = Q^{\mathrm{T}} Q(M \otimes N) Q^{\mathrm{T}} Q. \tag{2.11}$$

See Theorem 2.9.

2.2.1. Eigenvalue inequalities

Since $h \in \mathbb{R}^{t(n)}$, $\|h\| = 1 \Rightarrow \|Q^{\mathrm{T}} h\| = 1$, we see that (2.8) and the Rayleigh Principle yield some simple bounds on the largest and smallest eigenvalues of the symmetric Kronecker product, e.g. for every $A, B \in \mathscr{S}^n$ we have

$$\lambda_1(A \overset{s}{\otimes} B) \leqslant \lambda_1(A \otimes B) = \max_{i,j}(\lambda_i(A)\lambda_j(B))$$

and

$$\lambda_{t(n)}(A\overset{s}{\otimes}B) \geqslant \lambda_{n^2}(A \otimes B) = \min_{i,j}(\lambda_i(A)\lambda_j(B)).$$

We emphasize that the latter yields

$$(A \otimes B) \succeq 0 \quad \Rightarrow \quad (A\overset{s}{\otimes}B) \succeq 0, \tag{2.12}$$

but says nothing about the converse implication. Also let λ, μ be the eigenvalues of A, B with corresponding normalized eigenvectors a, b. Define $H := ab^{\mathrm{T}} + ba^{\mathrm{T}}$. Then

$$\begin{aligned}&\mathrm{s2vec}(H)^{\mathrm{T}}(A\overset{s}{\otimes}B)\mathrm{s2vec}(H)\\ &\quad = [1 + (a^{\mathrm{T}}b)^2]\lambda\mu + (a^{\mathrm{T}}b)^2\lambda\mu + (a^{\mathrm{T}}Ba)(b^{\mathrm{T}}Ab).\end{aligned}$$

Thus, a useful relation is

$$\frac{\mathrm{s2vec}(H)^{\mathrm{T}}(A\overset{s}{\otimes}B)\mathrm{s2vec}(H)}{\mathrm{trace}\, H^{\mathrm{T}}H} = \frac{1}{2}\lambda\mu + \frac{(a^{\mathrm{T}}b)^2\lambda\mu + (a^{\mathrm{T}}Ba)(b^{\mathrm{T}}Ab)}{2[1 + (a^{\mathrm{T}}b)^2]}. \tag{2.13}$$

For commuting matrices we can say more.

Corollary 2.5 [2, Lemma 7.2]. *Let $A, B \in \mathscr{S}$ such that A and B commute. Let λ_i, μ_i, v_i denote the eigenvalues and the corresponding (common) eigenvectors of A and B, respectively. Then, for $1 \leqslant i \leqslant j \leqslant n$, we get $\frac{1}{2}(\lambda_i\mu_j + \lambda_j\mu_i)$ as the eigenvalues and s2vec$(v_i v_j^{\mathrm{T}} + v_j v_i^{\mathrm{T}})$ as the corresponding eigenvectors of $(A\overset{s}{\otimes}B)$.*

Proof. Since A and B commute, we can use the same orthonormal system to describe their eigenspaces. Therefore, using direct computation (as in obtaining the identity (2.13)) we conclude the desired result, i.e. with $a = v_i, b = v_j$, the terms in the right hand side of (2.13) become $a^{\mathrm{T}}Ba = \mu_i, b^{\mathrm{T}}Bb = \lambda_j$. □

Corollary 2.6. *Let $A, B \in \mathscr{S}^n$ with eigenvalues λ_i, μ_j, respectively, and corresponding eigenvectors u_i, v_j, respectively. Then*

$$\left\{\mathrm{s2vec}(u_i u_i^{\mathrm{T}}), \mathrm{s2vec}(v_j v_j^{\mathrm{T}}) : \lambda_i = 0, \mu_j = 0\right\}$$

are eigenvectors of $(A\overset{s}{\otimes}B)$ corresponding to zero eigenvalues. Moreover, if two eigenvalues λ_s, μ_t have the same eigenvector v, then s2vec(vv^{T}) *is an eigenvector of $A\overset{s}{\otimes}B$.*

Proof. Direct computation. □

We can add to the simple bounds above using the relation (2.13).

Theorem 2.7. *Let* $A, B \succeq 0$. *Then*

$$\tfrac{1}{2}[\lambda_1(A)\lambda_1(B) + \lambda_n(A)\lambda_n(B)] \leqslant \lambda_1(A \overset{s}{\otimes} B) \leqslant \lambda_1(A)\lambda_1(B) = \lambda_1(A \otimes B)$$

and

$$\lambda_{n^2}(A \otimes B) = \lambda_n(A)\lambda_n(B) \leqslant \lambda_{t(n)}(A \overset{s}{\otimes} B) \leqslant \tfrac{1}{2}[\lambda_1(A)\lambda_1(B) + \lambda_n(A)\lambda_n(B)].$$

Proof. First consider the (pseudoconvex/pseudoconcave) function

$$f(t) = \frac{\alpha t + \beta}{2(1+t)}, \quad t \in [0, 1].$$

Then $f'(t) = 2(\alpha - \beta)/(\ldots)^2$, i.e. the function is nondecreasing (resp. nonincreasing) if $\alpha \geqslant \beta$ (resp. $\alpha \leqslant \beta$).

Choosing $\lambda := \lambda_1(A)$, $\mu := \lambda_1(B)$ in (2.13), and using $a^{\mathrm{T}}Ba \geqslant \lambda_n(B)$, $b^{\mathrm{T}}Ab \geqslant \lambda_n(A)$ and $(a^{\mathrm{T}}b)^2 \in [0, 1]$, we get $\alpha = \lambda\mu \geqslant \beta = (a^{\mathrm{T}}Ba)(b^{\mathrm{T}}Ab)$. The minimum value in (2.13) is therefore attained at $t = (a^{\mathrm{T}}b)^2 = 0$.

$$\lambda_1(A \overset{s}{\otimes} B) \geqslant \tfrac{1}{2}[\lambda_1(A)\lambda_1(B) + \lambda_n(A)\lambda_n(B)].$$

Similarly, choosing $\lambda := \lambda_n(A)$, $\mu := \lambda_n(B)$ in (2.13), and using $a^{\mathrm{T}}Ba \leqslant \lambda_1(B)$, $b^{\mathrm{T}}Ab \leqslant \lambda_1(A)$ and $(a^{\mathrm{T}}b)^2 \in [0, 1]$, we obtain $\alpha \leqslant \beta$ and the maximum value in (2.13) is attained at $t = (a^{\mathrm{T}}b)^2 = 0$.

$$\lambda_{t(n)}(A \overset{s}{\otimes} B) \leqslant \tfrac{1}{2}[\lambda_1(A)\lambda_1(B) + \lambda_n(A)\lambda_n(B)].$$

The remaining bounds were already established. □

The following theorem establishes the converse implication in (2.12), i.e. semidefiniteness of the smaller matrix implies semidefiniteness of the larger matrix.

Theorem 2.8. *Let* $A, B \in \mathscr{S}^n$. *Then*:

$$A \overset{s}{\otimes} B \succeq 0 \iff A \otimes B \succeq 0;$$

$$A \overset{s}{\otimes} B \succ 0 \iff A \otimes B \succ 0.$$

Proof. If the following Kronecker product $A \otimes B \succeq 0$ (resp. $\succ 0$), then necessarily the (restricted) symmetric Kronecker product $A \overset{s}{\otimes} B \succeq 0$ (resp. $\succ 0$). (This was already established in (2.12); see also (2.9).)

To show the converse, we first consider the $\succeq 0$ result. Assume that $A \overset{s}{\otimes} B \succeq 0$. Since

$$\mathrm{s2vec}(vv^{\mathrm{T}})^{\mathrm{T}}(A \overset{s}{\otimes} B)\mathrm{s2vec}(vv^{\mathrm{T}}) = (v^{\mathrm{T}}Av)(v^{\mathrm{T}}Bv) \geqslant 0 \quad \forall v \in \mathbb{R}^n,$$

we get

$$(v^{\mathrm{T}}Av) > 0 \quad \Rightarrow \quad (v^{\mathrm{T}}Bv) \geqslant 0; \tag{2.14}$$

$$(v^{\mathrm{T}}Bv) > 0 \quad \Rightarrow \quad (v^{\mathrm{T}}Av) \geqslant 0. \tag{2.15}$$

Now suppose

$$\lambda_{\min}(A \otimes B) < 0. \tag{2.16}$$

Without loss of generality, we can assume that $\lambda_{\min}(A \otimes B) = \lambda_{\min}(A)\lambda_{\max}(B)$, i.e. we assume that

$$\lambda := \lambda_{\min}(A) < 0 < \mu := \lambda_{\max}(B), \tag{2.17}$$

with corresponding normalized eigenvectors a, b, respectively. Note that $a^{\mathrm{T}}Aa < 0$ implies $a^{\mathrm{T}}Ba \leqslant 0$ and $b^{\mathrm{T}}Bb > 0$ implies $b^{\mathrm{T}}Ab \geqslant 0$, by (2.15), i.e. $(b^{\mathrm{T}}Ab)(a^{\mathrm{T}}Ba) \leqslant 0$. Therefore (2.13) implies that

$$\langle B, (ab^{\mathrm{T}} + ba^{\mathrm{T}})A(ab^{\mathrm{T}} + ba^{\mathrm{T}})\rangle < 0.$$

This contradicts the assumption $A \overset{s}{\otimes} B \succeq 0$, i.e. (2.16) fails.

Sufficiency for the $\succ 0$ result follows similarly. We assume $A \overset{s}{\otimes} B \succ 0$. The right-hand sides in (2.14) and (2.15) both become > 0. If we assume that (2.16) holds, then we get the desired contradiction. If we assume that (2.16) holds with equality, $= 0$, and (2.17) is changed to $\lambda = 0 < \mu$, then (2.13) now implies that

$$\left\langle B, (ab^{\mathrm{T}} + ba^{\mathrm{T}})A(ab^{\mathrm{T}} + ba^{\mathrm{T}})\right\rangle \leqslant 0,$$

a contradiction to the assumption $A \overset{s}{\otimes} B \succ 0$. □

Theorem 2.9. *Let $M, N \in \mathscr{M}^n$. For $u \in \mathbb{R}^{t(n)}$, we have the eigenpair relationship*

$$(M \overset{s}{\otimes} N)u = \lambda u \quad \Longrightarrow \quad \tfrac{1}{2}(M \otimes N + N \otimes M)(Q^{\mathrm{T}}u) = \lambda(Q^{\mathrm{T}}u). \tag{2.18}$$

Proof. Directly follows from the identity (2.9). □

The last theorem is quite powerful in that it establishes that each eigenspace of $(M \otimes N + N \otimes M)$ always decomposes as the direct sum of symmetric and skew-symmetric matrices (viewed in $\mathbb{R}^{n^2}$).

In addition to the above, a tight, interlacing type of relationship seems to exist between the eigenvalues of $\frac{1}{2}(A \otimes B + B \otimes A)$ corresponding to the skew-symmetric eigenvectors and the eigenvalues of $A \overset{s}{\otimes} B$; see Conjecture 2.12. We elaborate below.

Note that $u \in \mathbb{R}^{n^2}$ corresponds to a symmetric matrix iff $Tu = u$ and it corresponds to a skew-symmetric matrix iff $Tu = -u$. Suppose $u, w \in \mathbb{R}^{n^2}$ such that $Tu = u$ and $Tw = -w$. Then

$$\tfrac{1}{2}u^{\mathrm{T}}[(A \otimes B) + (B \otimes A)]u = u^{\mathrm{T}}(A \otimes B)u,$$

and

$$\tfrac{1}{2}w^{\mathrm{T}}[(A \otimes B) + (B \otimes A)]w = w^{\mathrm{T}}(A \otimes B)w.$$

We conjecture:

Conjecture 2.10. *Let $A, B \in \mathscr{S}^n$. Then*

$$\min_{Tu=u} \frac{u^{\mathrm{T}}(A \otimes B)u}{u^{\mathrm{T}}u} \leqslant \min_{Tw=-w} \frac{w^{\mathrm{T}}(A \otimes B)w}{w^{\mathrm{T}}w}$$

and

$$\max_{Tu=u} \frac{u^{\mathrm{T}}(A \otimes B)u}{u^{\mathrm{T}}u} \geqslant \max_{Tw=-w} \frac{w^{\mathrm{T}}(A \otimes B)w}{w^{\mathrm{T}}w}.$$

If the above conjecture is true, then we will have

$$\lambda_1(A \overset{s}{\otimes} B) = \tfrac{1}{2}\lambda_1(A \otimes B + B \otimes A)$$

and

$$\lambda_{t(n)}(A \overset{s}{\otimes} B) = \tfrac{1}{2}\lambda_{n^2}(A \otimes B + B \otimes A).$$

Another way of expressing this conjecture is as follows:

Conjecture 2.11. *Let $A, B \in \mathscr{S}^n$. Then*

$$\min_{U \in \mathscr{S}^n, \|U\|_{\mathrm{F}}=1} \operatorname{trace}(BUAU) \leqslant \min_{W \in \tilde{\mathscr{S}}^n, \|W\|_{\mathrm{F}}=1} \operatorname{trace}(BWAW^{\mathrm{T}})$$

and

$$\max_{U \in \mathscr{S}^n, \|U\|_{\mathrm{F}}=1} \operatorname{trace}(BUAU) \geqslant \max_{W \in \tilde{\mathscr{S}}^n, \|W\|_{\mathrm{F}}=1} \operatorname{trace}(BWAW^{\mathrm{T}}),$$

where $\tilde{\mathscr{S}}^n$ denotes the space of $n \times n$ skew-symmetric matrices.

Note that if $B = I$, then the conjecture is easy to prove using the well-known trace inequality. For the first inequality, let $U := uu^{\mathrm{T}}$, where $u \in \mathbb{R}^n$ is the eigenvector of A corresponding to $\lambda_n(A)$; for the second inequality, let $U := uu^{\mathrm{T}}$ where $u \in \mathbb{R}^n$ is the eigenvector of A corresponding to $\lambda_1(A)$.

In fact, a stronger conjecture may be true:

Conjecture 2.12. *Let $A, B \in \mathscr{S}^n$. Also let $w \in \mathbb{R}^{n^2}$ such that $Tw = -w$ and w is the eigenvector of $\frac{1}{2}[(A \otimes B) + (B \otimes A)]$ corresponding to its kth largest eigenvalue. Then λ_{k-1} and λ_{k+1} of the matrix are well-defined and they are determined by some $u, v \in \mathbb{R}^{n^2}$ such that $Tu = u$ and $Tv = v$.*

The statement of the above conjecture is clear when all eigenvalues are distinct. In case of ties, they are to be broken (that is, the eigenvalues are numbered) in favor of the conjecture. If true, the last conjecture 2.12 implies the other two equivalent conjectures, 2.10 and 2.11.

3. Conditions for existence and uniqueness

3.1. Characterizations for existence and uniqueness

3.1.1. Conditions using a subspace

Definition 3.1. Let $\mathscr{A} : \mathscr{S}^n \mapsto \mathbb{R}^m$ be surjective. Define

$$\text{AHO}(\mathscr{A}) := \big\{(X, S) \in \mathscr{S}^n \times \mathscr{S}^n : \text{the system (2.1), determined by } (X, S) \text{ and } \mathscr{A}, \text{ has a unique solution}\big\}.$$

For convenience, we define the complement of $\text{AHO}(\mathscr{A})$:

$$\overline{\text{AHO}}(\mathscr{A}) := \big\{(X, S) \in \mathscr{S}^n \times \mathscr{S}^n : (X, S) \notin \text{AHO}(\mathscr{A})\big\}.$$

If $\mathscr{L} = \mathscr{N}(\mathscr{A})$, we also use the alternative notation

$$\text{AHO}(\mathscr{L}) := \text{AHO}(\mathscr{A}), \quad \overline{\text{AHO}}(\mathscr{L}) := \overline{\text{AHO}}(\mathscr{A}).$$

The following is one characterization for existence and uniqueness in (1.1).

Proposition 3.2. *Let $(X, S) \in \mathscr{S}^n \times \mathscr{S}^n$ and a subspace $\mathscr{L} \subset \mathscr{S}^n$ be given. Then $(X, S) \in \overline{\text{AHO}}(\mathscr{L})$ if and only if*

$$L_X(\mathscr{L}^\perp) \cap L_S(\mathscr{L}) \neq \{0\}. \tag{3.1}$$

Moreover, if L_X is invertible (similar result if L_S is invertible), then (3.1) holds if and only if

$$L_X^{-1} L_S(\Delta X) \in \mathscr{L}^\perp \quad \text{for some } 0 \neq \Delta X \in \mathscr{L}. \tag{3.2}$$

Equivalently, for $\mathscr{L} = \mathscr{N}(\mathscr{A})$, $(X, S) \in \overline{\text{AHO}}(\mathscr{L})$ if and only if $\exists \Delta X \in \mathscr{N}(\mathscr{A}) \setminus \{0\}$ such that $\Delta S = L_X^{-1} L_S(\Delta X) \in \mathscr{R}(\mathscr{A}^*)$.

Proof. It suffices to note that the system (2.1) can be equivalently written as

$$\boxed{\begin{array}{rcl} L_X(\Delta S) & = & -L_S(\Delta X) \\ \Delta X \in \mathscr{L}, & & \Delta S \in \mathscr{L}^\perp. \end{array}} \quad \square \tag{3.3}$$

The following fact was already established in Section 2.1.

Corollary 3.3. *Suppose that* $X, S, \mathscr{A}, \mathscr{L}$ *are defined as in Proposition* 3.2 *and* (2.4). *Then* $(X, S) \in AHO(\mathscr{A})$ *if and only if the linear operator* $\bar{\mathscr{K}} = [L_S\mathscr{L} \vdots - L_X\mathscr{A}^*]$ *is nonsingular.*

3.1.2. Conditions using an indefinite trust region subproblem

To obtain further characterizations for existence and uniqueness in (1.1), we use the following properties.

Lemma 3.4. *Suppose that* $(\Delta X, \Delta y, \Delta S)$ *solves* (2.1) *for a given pair* $X \succ 0$, $S \succ 0$. *Then the following hold*:

1. $\langle \Delta X, \Delta S \rangle = 0$.
2. *The matrix*
$$M := X(\Delta S) + (\Delta X)S \tag{3.4}$$
is skew-symmetric.
3. $$M = X(\Delta S) + (\Delta X)S = 0 \quad \text{if and only if} \quad \Delta X = \Delta S = 0, \quad \Delta y = 0. \tag{3.5}$$

Proof

1. This can be viewed simply as $\Delta X \in \mathscr{N}(\mathscr{A})$ is orthogonal to $\Delta S \in \mathscr{R}(\mathscr{A}^*)$.
2. This follows from (2.2), the third equation in (2.1).
3. Sufficiency is clear. From the assumption we get $\Delta S = -X^{-1}(\Delta X)S$. We now compute the weighted norm of ΔX with this equation using part 1 above.
$$\begin{aligned} 0 = \langle \Delta X, \Delta S \rangle &= -\langle \Delta X, X^{-1}(\Delta X)S \rangle \\ &= -\langle X^{-\frac{1}{2}}(\Delta X)S^{\frac{1}{2}}, X^{-\frac{1}{2}}(\Delta X)S^{\frac{1}{2}} \rangle \\ &= -\|X^{-\frac{1}{2}}(\Delta X)S^{\frac{1}{2}}\|_F^2. \end{aligned}$$
This equation holds if and only if $\Delta X = 0$. Using the equation for ΔS, it follows that $\Delta S = 0$ as well. That $\Delta y = 0$ follows from the second equation in (2.1). □

Lemma 3.4 yields a second characterization of existence and uniqueness.

Proposition 3.5. *Under the hypothesis in Proposition* 3.2 *and the assumptions* $X \succ 0$, $S \succ 0$, *we have* $(X, S) \in \text{AHO}(\mathscr{L})$ *if and only if*
$$(2.1) \quad \Rightarrow \quad M = X(\Delta S) + (\Delta X)S = 0 \tag{3.6}$$
if and only if

$$\begin{array}{rll} 0 = \mu^* := \min & \operatorname{trace}\{[X(\Delta S) + (\Delta X)S]^2\} & (= \operatorname{trace}(M^2)) \\ \text{s.t.} & (2.1). & \end{array} \tag{3.7}$$

Proof. The first characterization (3.6) is clear from Lemma 3.4. Necessity in the second characterization is also clear. Now, suppose that $(\Delta X, \Delta y, \Delta S)$ is a solution of the homogeneous system (2.1). Then Lemma 3.4 Part 2 implies that $M^2 \preceq 0$ and $\operatorname{trace}(M^2) \leqslant 0$. Therefore, $\mu^* = 0$ implies that $\operatorname{trace}(M^2) = 0$, i.e. $0 = \operatorname{trace}(M^2) = \operatorname{trace}(-MM^{\mathrm{T}})$. This further implies that $MM^{\mathrm{T}} = 0$ or $M = 0$. The conclusion now follows from Lemma 3.4 Part 3. □

Corollary 3.6. *Under the hypothesis of Proposition 3.5, we get*

$$\begin{cases} 0 = \mu^* := \min & \operatorname{trace}[(X(\Delta S) + (\Delta X)S)^2] \\ \qquad\qquad \text{s.t.} & \langle \Delta X, \Delta S \rangle = 0 \\ & \|L_S(\Delta X) + L_X(\Delta S)\|^2 = 0 \end{cases} \tag{3.8}$$

if and only if

$$(2.1) \text{ implies } \Delta X = \Delta S = 0, \text{ independent of } \mathscr{A}. \tag{3.9}$$

Proof. The result follows since the feasible set in the quadratic program is smaller in Proposition 3.5, i.e. it depends on the specific linear transformation $\mathscr{A}$. □

Corollary 3.7

$$\begin{cases} 0 = \mu^* := \min & \operatorname{trace}[(X\mathscr{A}^*(\Delta y) + \mathscr{Z}(\Delta v)S)^2] \\ \qquad\qquad \text{s.t.} & \|L_S(\mathscr{Z}(\Delta v)) + L_X(\mathscr{A}^*(\Delta y))\|^2 = 0 \end{cases} \tag{3.10}$$

if and only if

$$(2.1) \text{ implies } \Delta X = \Delta S = 0. \tag{3.11}$$

Proof. The result follows from Proposition 3.5 and the definitions of the transformations $\mathscr{A}, \mathscr{Z}$. □

3.2. *Sufficient conditions independent of $\mathscr{A}$*

We present: (i) characterizations of existence and uniqueness; (ii) strengthened conditions of the SSK type; and strengthened conditions of the MZ type.

Corollary 3.8. *Suppose that $X, S, \mathscr{L}, \mathscr{A}$ are defined as in Proposition 3.2. If L_X is invertible (similar result if L_S is invertible), then*

$$\left(L_X^{-1}L_S + (L_X^{-1}L_S)^*\right) \text{ is definite } \Rightarrow (X, S) \in \mathrm{AHO}(\mathscr{A}). \tag{3.12}$$

Proof. Note that (3.2) holds only if there exists $0 \neq \Delta X \in \mathscr{L}$ such that

$$\left\langle \Delta X, L_X^{-1} L_S(\Delta X) \right\rangle = 0.$$

This cannot happen if the quadratic form is definite. □

We now present sufficient conditions using the Kronecker product and then apply this to get a strengthened SSK condition. These conditions do not take into account the constraint transformation $\mathscr{A}$ but do use the orthogonality between ΔS and ΔX.

Theorem 3.9. *Suppose that L_X is invertible (similar results if L_S is invertible) and $\mathscr{T}$, T are defined as above in (2.10). Let*

$$K := [I \otimes (SX + XS)] + [(SX + XS) \otimes I] + 2[X \otimes S] + 2[S \otimes X]. \tag{3.13}$$

Then

the condition (3.14)
implies
the equivalent conditions (3.15), (3.16), (3.17), (3.18)
which imply
$(X, S) \in \mathrm{AHO}(\mathscr{L})$ *for all subspaces* $\mathscr{L} \subseteq \mathscr{S}^n$.

1. $$K = [I \otimes (SX + XS)] + [(SX + XS) \otimes I] + 2[X \otimes S] + 2[S \otimes X] \text{ is definite} \tag{3.14}$$

⇓ ⇓

2. (a)
$$[I \overset{s}{\otimes} (SX + XS)] + 2[X \overset{s}{\otimes} S] \text{ is definite} \tag{3.15}$$
(b)
$$L_X^{-1} L_S + L_S L_X^{-1} \text{ is definite} \tag{3.16}$$
(c)
$$L_S L_X + L_X L_S \text{ is definite} \tag{3.17}$$
(d)
$$2K + TK + KT \text{ is semidefinite with rank } t(n) \tag{3.18}$$

⇓ ⇓

3.
$$(X, S) \in \mathrm{AHO}(\mathscr{L}) \quad \text{for all subspaces } \mathscr{L} \in \mathscr{S}^n. \tag{3.19}$$

Proof. That (3.16) implies (3.19) follows from Corollary 3.8. The equivalence of (3.16) and (3.17) follows from the congruence by the nonsingular self-adjoint operator L_X.

Now, we prove that $(L_S L_X + L_X L_S)$ is definite if and only if $\frac{1}{2}[I \overset{s}{\otimes} (SX + XS)] + [X \overset{s}{\otimes} S]$ is. Let $U \in \mathscr{S}^n$, then

$$\begin{aligned}\langle U, (L_S L_X + L_X L_S)(U)\rangle &= \langle U, [L_S(XU + UX) + L_X(SU + US)]\rangle \\ &= 2\langle U, SUX + XUS\rangle \\ &\quad + \langle U, [(XS + SX)UI + IU(SX + XS)]\rangle \\ &= 4[\text{s2vec}(U)]^{\mathrm{T}}(X \overset{s}{\otimes} S)\text{s2vec}(U) \\ &\quad + 2[\text{s2vec}(U)]^{\mathrm{T}}(I \overset{s}{\otimes} (XS + SX))\text{s2vec}(U).\end{aligned}$$

Therefore, the equivalence of (3.15) and (3.17) follows.

With $U = U^{\mathrm{T}} = \frac{1}{2}(V + V^{\mathrm{T}})$, $V \in \mathscr{M}^n$, the quadratic form under consideration is

$$\begin{aligned}&4\langle U, (L_S L_X + L_X L_S)(U)\rangle \\ &\quad = \langle V + V^{\mathrm{T}}, (L_S L_X + L_X L_S)(V + V^{\mathrm{T}})\rangle \\ &\quad = 2\langle V, (XS + SX)V\rangle + 2\langle V, V(XS + SX)\rangle \\ &\qquad + 4\langle V, XVS\rangle + 4\langle V, SVX\rangle \\ &\qquad + \langle V^{\mathrm{T}}, (XS + SX)V\rangle + \langle V, (XS + SX)V^{\mathrm{T}}\rangle \\ &\qquad + \langle V^{\mathrm{T}}, V(XS + SX)\rangle + \langle V, V^{\mathrm{T}}(XS + SX)\rangle \\ &\qquad 2\langle V^{\mathrm{T}}, XVS\rangle + 2\langle V, XV^{\mathrm{T}}S\rangle + 2\langle V^{\mathrm{T}}, SVX\rangle + 2\langle V, SV^{\mathrm{T}}X\rangle. \end{aligned} \tag{3.20}$$

We have shown that

$$4\langle U, (L_S L_X + L_X L_S)(U)\rangle = \langle V, (2K + TK + KT)V\rangle, \tag{3.21}$$

if $U = \frac{1}{2}(V + V^{\mathrm{T}})$. We conclude that (3.14) implies (3.17), since we have ignored the restriction of V to symmetric matrices in the Kronecker product. Let $V \in \mathscr{M}^n$ be skew-symmetric. Also let $v := \text{vec}(V)$. Then $Tv = -v$. Note that

$$v^{\mathrm{T}}(2K + TK + KT)v = 2v^{\mathrm{T}}Kv - 2v^{\mathrm{T}}Kv = 0.$$

Therefore, the rank of $(K + TK + KT)$ is always upperbounded by $t(n)$. Using (3.21), we conclude that (3.17) and (3.18) are equivalent. □

Remark 3.10. The equivalent condition (3.18) raises the question of finding a simpler expression for the eigenvalues of the sum involving *TK*.

The above conditions involve a similarity scaling and numerical range, see e.g. [4,14,22]. We now present sufficient conditions for existence and uniqueness using centrality and condition number measures. This strengthens the MZ result.

Theorem 3.11. *Suppose that* $S, X \succ 0$. *Then the equivalent conditions,* (3.22) *and* (3.23), *imply* $(X, S) \in \mathrm{AHO}(\mathscr{L})$ *for every* $\mathscr{L} \subseteq \mathscr{S}^n$:

1.
$$\|S^{\frac{1}{2}} X S^{\frac{1}{2}} - \nu I\| \leqslant \tfrac{4}{5}\nu \quad \textit{for some scalar } \nu > 0. \tag{3.22}$$

2. *The condition number*
$$\gamma\left(S^{\frac{1}{2}} X S^{\frac{1}{2}}\right) \leqslant 9. \tag{3.23}$$

Proof. Let
$$D^2 := S^{-1/2}(S^{1/2} X S^{1/2})^{1/2} S^{-1/2}, \tag{3.24}$$
and
$$V := DSD = D^{-1} X D^{-1}. \tag{3.25}$$
Note that the eigenvalues of V^2 and $S^{1/2} X S^{1/2}$ are the same. We define
$$\overline{\Delta X} := D^{-1}(\Delta X) D^{-1} \quad \text{and} \quad \overline{\Delta S} := D(\Delta S) D,$$
where $\Delta X, \Delta S \in \mathscr{S}^n$ satisfy the constraints of (3.8). Note that
$$\langle \overline{\Delta X}, \overline{\Delta S} \rangle = \langle \Delta X, \Delta S \rangle = 0.$$
By Corollary 3.6, we only need to show that
$$0 \leqslant \min\left\{\mathrm{trace}(M^2) : \langle \overline{\Delta X}, \overline{\Delta S} \rangle = 0\right\}. \tag{3.26}$$

$$\begin{aligned}\mathrm{trace}(M^2) &= \mathrm{trace}\{[X(\Delta S) + (\Delta X)S]^2\} \\ &= \mathrm{trace}[V(\overline{\Delta S})V(\overline{\Delta S})] + \mathrm{trace}[V(\overline{\Delta X})V(\overline{\Delta X})] \\ &\quad + 2\mathrm{trace}[(\overline{\Delta X})V^2(\overline{\Delta S})] \\ &= \mathrm{trace}[V(\overline{\Delta S})V(\overline{\Delta S})] + \mathrm{trace}[V(\overline{\Delta X})V(\overline{\Delta X})] \\ &\quad + 2\mathrm{trace}[(V^2 - \mu I)(\overline{\Delta X})(\overline{\Delta S})].\end{aligned}$$

The parameter μ plays the role of a Lagrange multiplier for (3.26), see [20,25]. Let $s := \mathrm{s2vec}(\overline{\Delta S})$, $x := \mathrm{s2vec}(\overline{\Delta X})$. Then using Lemma 2.4 yields

$$\begin{aligned}\mathrm{trace}(M^2) = \begin{pmatrix} x \\ s \end{pmatrix}^{\mathrm{T}} \Bigg\{ &\begin{pmatrix} V \overset{s}{\otimes} V & 0 \\ 0 & V \overset{s}{\otimes} V \end{pmatrix} \\ &+ \begin{pmatrix} 0 & I \overset{s}{\otimes} (V^2 - \mu I) \\ I \overset{s}{\otimes} (V^2 - \mu I) & 0 \end{pmatrix} \Bigg\} \begin{pmatrix} x \\ s \end{pmatrix}.\end{aligned}$$

Therefore, it suffices to find conditions which guarantee that

$$\begin{pmatrix} V\overset{s}{\otimes}V & I\overset{s}{\otimes}(V^2-\mu I) \\ I\overset{s}{\otimes}(V^2-\mu I) & V\overset{s}{\otimes}V \end{pmatrix} \succeq 0 \quad \text{for some } \mu. \tag{3.27}$$

We can take the Schur complement in (3.27) to get the equivalent condition

$$V\overset{s}{\otimes}V - (I\overset{s}{\otimes}(V^2-\mu I))(V\overset{s}{\otimes}V)^{-1}(I\overset{s}{\otimes}(V^2-\mu I)) \succeq 0. \tag{3.28}$$

Note that the matrices V, $(V^2-\mu I)$ and I all commute. Thus, using Corollary 2.5, we can diagonalize all four terms in (3.28) which yields the following result for the eigenvalues λ_i of V:

$$4\lambda_i^2\lambda_j^2 - \left(\lambda_i^2+\lambda_j^2-2\mu\right)^2 \geqslant 0 \quad \forall i, j. \tag{3.29}$$

Equivalently,

$$2\lambda_i\lambda_j \geqslant \left|\lambda_i^2+\lambda_j^2-2\mu\right| \quad \forall i, j. \tag{3.30}$$

Equivalently,

$$\tfrac{1}{2}(\lambda_i-\lambda_j)^2 \leqslant \mu \leqslant \tfrac{1}{2}(\lambda_i+\lambda_j)^2 \quad \forall i, j. \tag{3.31}$$

Such μ exists iff $(\lambda_{\max}-\lambda_{\min})^2 \leqslant 4\lambda_{\min}^2$. Equivalently, we see $\gamma(V) \leqslant 3$ or $\gamma^2(V) \leqslant 9$. Since the condition numbers of V^2 and $S^{1/2}XS^{1/2}$ are the same, using Proposition 1.1, we conclude the desired result. □

Remark 3.12. Note that in the above proof, we used the skew-symmetry condition early in the proof, and then threw it away. In fact, we can substitute

$$s = \text{s2vec}(\overline{\Delta S}) = -(D\overset{s}{\otimes}D)\text{s2vec}(L_X^{-1}L_S(D(\overline{\Delta X})D)).$$

This should lead to stronger results. Also, we could have used the usual Kronecker product and employed Lemma 2.3 instead of the symmetric Kronecker product and Lemma 2.4. This alternative proof leads to the bounds $\frac{3}{5}$ for α and 4 for the condition number. Using the same proof technique with the symmetric Kronecker product not only made the proof more elegant but it also improved the bounds. Finally, a related but weaker bound was obtained by Monteiro and Todd [17], using different techniques. We slightly improve their result in Theorem 3.18.

3.2.1. Shida–Shindoh–Kojima (SSK) sufficient condition

From Corollary 3.8 and Theorem 3.9, we get the Shida, Shindoh and Kojima (SSK) sufficient condition [24], and see why this condition is so 'weak', i.e. it does not use the positive definiteness of $X \otimes S$ or $S \otimes X$.

Corollary 3.13. *Suppose that $X, S, \mathscr{L}, \mathscr{A}$ are defined as in Proposition 3.2 and $X, S \succ 0$. Then each of the conditions in Theorem 3.9 is implied by the Shida et al. (SSK) hypothesis*

$$XS + SX \succeq 0.$$

Proof. The results follow immediately from Theorem 3.9. □

3.2.2. Monteiro–Zanjácomo (MZ) sufficient condition

In this section we present the theorem from [18] that provides a sufficient condition for the above system (1.1) to have a unique solution. The result follows from the strengthened version Theorem 3.11. We include a note on a typographical error in [18].

Theorem 3.14 (Monteiro–Zanjácomo [18]). *If $(X, S, y) \in \mathscr{S}^n_+ \times \mathscr{S}^n_+ \times \mathbb{R}^m$ is such that*

$$\|S^{\frac{1}{2}} X S^{\frac{1}{2}} - \nu I\| \leqslant \tfrac{1}{2}\nu \quad \textit{for some scalar } \nu > 0, \tag{3.32}$$

then system (1.1) *has a unique solution*.

Proof. The proof follows from the strengthened version in Theorem 3.11. Note that the last expressions in the proof in [18] should read

$$\begin{array}{rcl} & \vdots & \\ \cdots & \geqslant & \|\widehat{\Delta S}\|_F^2 + \frac{1}{4}\nu^2\|\widehat{\Delta X}\|_F^2 - 2\frac{1}{2}\nu\|\widehat{\Delta S}\|_F\|\widehat{\Delta X}\|_F \\ & = & \left(\|\widehat{\Delta S}\|_F - \frac{1}{2}\nu\|\widehat{\Delta X}\|_F\right)^2 \\ & \geqslant & 0. \qquad \square \end{array} \tag{3.33}$$

3.2.3. Monteiro–Todd (MT) sufficient condition

The previously stated MZ result was improved by Monteiro and Todd [17] using different techniques.

Theorem 3.15 (Monteiro–Todd [17]). *If $(X, S, y) \in \mathscr{S}^n_+ \times \mathscr{S}^n_+ \times \mathbb{R}^m$ is such that*

$$\|S^{\frac{1}{2}} X S^{\frac{1}{2}} - \nu I\| < \frac{1}{\sqrt{2}}\nu \quad \textit{for some scalar } \nu > 0,$$

then system (1.1) *has a unique solution*.

Here, we extend the techniques of MT by utilizing the primal–dual symmetry property of the AHO direction. As a result, we slightly improve the corresponding MT bound. We need the following elementary lemmas:

Lemma 3.16. *Let $A \in \mathscr{S}^n$ and $B \in \mathscr{M}^n$. Then*

1.

$$\|AB\|_F \leqslant \|A\|\|B\|_F. \tag{3.34}$$

2.
$$\|B^{\mathrm{T}}AB\|_{\mathrm{F}} \geqslant \lambda_{\min}(B^{\mathrm{T}}B)\|A\|_{\mathrm{F}}. \tag{3.35}$$

Lemma 3.17 [17]. *Let $M \in \mathscr{M}^n$ such that (PMP^{-1}) is skew-symmetric for some nonsingular $P \in \mathscr{M}^n$. Suppose $M = A + B$, where $A \in \mathscr{S}^n$. Then $\|A\|_{\mathrm{F}} \leqslant \sqrt{2}\|B\|_{\mathrm{F}}$.*

The next theorem improves the constants $\frac{1}{\sqrt{2}}$ and $(3 + 2\sqrt{2})$ (resp.) of [17] to $(\sqrt{3} - 1)$ and $(3 + 2\sqrt{3})$ (resp.). The proof of the next theorem follows the ideas of Monteiro and Todd [17] for primal and dual scalings and then combines them in a symmetric way.

Theorem 3.18. *Suppose that $S, X \succ 0$. Then the following equivalent conditions imply that $(X, S) \in \mathrm{AHO}(\mathscr{L})$, for every $\mathscr{L} \subseteq \mathscr{S}^n$:*

1.
$$\|S^{1/2}XS^{1/2} - \nu I\| < (\sqrt{3} - 1)\nu \quad \text{for some scalar } \nu > 0. \tag{3.36}$$

2. *The condition number*
$$\gamma(S^{1/2}XS^{1/2}) < 3 + 2\sqrt{3}. \tag{3.37}$$

Proof. Suppose $(U, \Delta y, W) \in \mathscr{S}^n \times \mathbb{R}^m \times \mathscr{S}^n$ solves the system (2.1). Let $M := US + XW$. Then M is skew-symmetric. We will apply Lemma 3.17 to $X^{-1/2}MX^{1/2}$. Let

$$A := \nu X^{-1/2}UX^{-1/2} + X^{1/2}WX^{1/2} \quad (A \in \mathscr{S}^n),$$

$$B := X^{-1/2}UX^{-1/2}(X^{1/2}SX^{1/2} - \nu I).$$

Note that

$$\langle X^{-1/2}UX^{-1/2}, X^{1/2}WX^{1/2}\rangle = \langle U, W\rangle = 0$$

implies

$$\|A\|_{\mathrm{F}}^2 = \nu^2\|X^{-1/2}UX^{-1/2}\|_{\mathrm{F}}^2 + \|X^{1/2}WX^{1/2}\|_{\mathrm{F}}^2.$$

Lemmas 3.17 and 3.16 Part 1 imply

$$\|A\|_{\mathrm{F}}^2 \leqslant 2\|B\|_{\mathrm{F}}^2 \leqslant 2\|X^{-1/2}UX^{-1/2}\|_{\mathrm{F}}^2\|X^{1/2}SX^{1/2} - \nu I\|^2.$$

Therefore,

$$(\nu^2 - 2\|X^{1/2}SX^{1/2} - \nu I\|^2)\|X^{-1/2}UX^{-1/2}\|_{\mathrm{F}}^2 \leqslant -\|X^{1/2}WX^{1/2}\|_{\mathrm{F}}^2. \tag{3.38}$$

Similarly, using the above MT idea with $S^{1/2}MS^{-1/2} = S^{1/2}US^{1/2} + S^{1/2}XWS^{-1/2}$, we apply Lemma 3.17 to $S^{1/2}MS^{-1/2}$. Let

$$A := \nu S^{-1/2}WS^{-1/2} + S^{1/2}US^{1/2} \quad (A \in \mathscr{S}^n),$$

$$B := (S^{1/2}XS^{1/2} - \nu I)S^{-1/2}WS^{-1/2}.$$

Using $\langle U, W\rangle = 0$, Lemma 3.17 and Lemma 3.16 Part 1, we obtain

$$(\nu^2 - 2\|S^{1/2}XS^{1/2} - \nu I\|^2)\|S^{-1/2}WS^{-1/2}\|_F^2 \leqslant -\|S^{1/2}US^{1/2}\|_F^2. \tag{3.39}$$

Since $(S^{1/2}XS^{1/2})$ and $(X^{1/2}SX^{1/2})$ have the same eigenvalues, combining the relations (3.38) and (3.39) we conclude

$$\begin{aligned}&(\nu^2 - 2\|S^{1/2}XS^{1/2} - \nu I\|^2)(\|X^{-1/2}UX^{-1/2}\|_F^2 + \|S^{-1/2}WS^{-1/2}\|_F^2)\\ &\quad\leqslant -(\|X^{1/2}WX^{1/2}\|_F^2 + \|S^{1/2}US^{1/2}\|_F^2).\end{aligned} \tag{3.40}$$

Using Lemma 3.16 Part 2,

$$\begin{aligned}\|X^{1/2}WX^{1/2}\|_F^2 &= \|X^{1/2}S^{1/2}(S^{-1/2}WS^{-1/2})S^{1/2}X^{1/2}\|_F^2\\ &\geqslant [\lambda_{\min}(X^{1/2}SX^{1/2})]^2\|S^{-1/2}WS^{-1/2}\|_F^2.\end{aligned}$$

Similarly,

$$\|S^{1/2}US^{1/2}\|_F^2 \geqslant [\lambda_{\min}(S^{1/2}XS^{1/2})]^2\|X^{-1/2}UX^{-1/2}\|_F^2.$$

Thus, (3.40) yields that every solution $(U, \Delta y, W)$ of the system (2.1) satisfies

$$\begin{aligned}&\left(\nu^2 - 2\|S^{1/2}XS^{1/2} - \nu I\|^2 + [\lambda_{\min}(S^{1/2}XS^{1/2})]^2\right)\\ &\quad\times\left(\|X^{-1/2}UX^{-1/2}\|_F^2 + \|S^{-1/2}WS^{-1/2}\|_F^2\right) \leqslant 0.\end{aligned}$$

Let $a := \frac{1}{\nu}\|S^{1/2}XS^{1/2} - \nu I\|$ with $\nu = \frac{\lambda_{\min}(S^{1/2}XS^{1/2})+\lambda_{\max}(S^{1/2}XS^{1/2})}{2}$. Then $\lambda_{\min}(S^{1/2}XS^{1/2}) = (1-a)\nu$. Therefore, the above inequality implies that $(0, 0, 0) \in \mathscr{S}^n \times \mathbb{R}^m \times \mathscr{S}^n$ is the unique solution of (2.1) if

$$\nu^2 - 2a^2\nu^2 + (1-a)^2\nu^2 > 0.$$

Since $\nu^2 > 0$, we analyze the quadratic inequality in a and find that if $a < (\sqrt{3} - 1)$ then $(X, S) \in \mathrm{AHO}(\mathscr{L})$. The value $(\sqrt{3} - 1)$ of a corresponds to the bound $(3 + 2\sqrt{3})$ on the condition number of $(S^{1/2}XS^{1/2})$ by Proposition 1.1. □

3.3. *Sufficient conditions dependent on $\mathscr{A}$*

The following is similar to Theorem 3.1 from [27].

Corollary 3.19. *Suppose that $X, S, \mathscr{L}, \mathscr{A}$ are defined as in Proposition 3.2. If L_X is invertible then*

$$\mathscr{Z}^*\left(L_X^{-1}L_S + (L_X^{-1}L_S)^*\right)\mathscr{Z} \succ 0 \;\Rightarrow\; (X, S) \in \mathrm{AHO}(\mathscr{A}). \tag{3.41}$$

If L_S is invertible, then

$$\mathscr{A}^*\left(L_S^{-1}L_X + (L_S^{-1}L_X)^*\right)\mathscr{A} \succ 0 \;\Rightarrow\; (X, S) \in \mathrm{AHO}(\mathscr{A}). \tag{3.42}$$

Proof. Note that (3.2) holds only if there exists $0 \neq \Delta X \in \mathscr{L}$ such that

$$\left\langle \Delta X, L_X^{-1} L_S(\Delta X) \right\rangle = 0.$$

We use $\Delta X = \mathscr{Z}(\Delta v)$. (The second result follows similarly.) □

We use the notation $\mathscr{Z}(\Delta v) = \sum_{i=m+1}^{t(n)} (\Delta v)_i A_i$. Since $\mathscr{R}(\mathscr{A}^*) \perp \mathscr{N}(\mathscr{A})$, we have

$$\langle A_j, A_i \rangle = 0 \quad \forall i = 1, \ldots, m, \quad \forall j = m+1, \ldots, t(n).$$

If $S \succ 0$ then $L_S \succ 0$ and $L_S \mathscr{Z}$ is one–one. (Similarly for X, L_X and $L_X \mathscr{A}^*$.) Therefore, nonsingularity of $\bar{\mathscr{K}}$ depends on the relationship of the range spaces of the two linear transformations

$$\mathscr{R}_X := \mathscr{R}(L_X \mathscr{A}^*), \quad \mathscr{R}_S := \mathscr{R}(L_S \mathscr{Z}),$$

i.e. the rotations of the range spaces of $\mathscr{Z}$, $\mathscr{A}^*$ or equivalently of the orthogonal spaces $\mathscr{L}$, $\mathscr{L}^\perp$. Equivalently, we need to study the *scaled* matrices

$$XA_i + A_i X, i = 1, \ldots, m, \quad SA_i + A_i S, i = m+1, \ldots, t(n).$$

4. Topological properties

We have seen the following interesting conundrum. The three best known search directions are the AHO, HrvwKshM and NT directions. Under standard nondegeneracy assumptions, all three directions are well-defined in a neighbourhood of the central path for $X, S \succ 0$. The linear systems for both HrvwKshM and NT become increasingly ill-conditioned as the barrier parameter $\mu \downarrow 0$, i.e. when X, S approach the optimum. This is not the case for AHO *near* the central path. However, both HrvwKshM and NT are well defined for all $X, S \succ 0$, while this is *not* the case for AHO, i.e. the linearized system for the AHO direction can become ill-conditioned unless the iterates stay close to the central path. Nevertheless, we now show that AHO is generically well-defined.

Theorem 4.1. *For every subspace* $\mathscr{L} \in \mathscr{S}^n$, *the dimension of* AHO($\mathscr{L}$) *is full*; *i.e.*,

$$\dim(\mathrm{AHO}(\mathscr{L})) = n(n+1).$$

Moreover,

$$\overline{\mathrm{AHO}}(\mathscr{L}) \textit{ is a set of measure zero in } \mathscr{S}^n \times \mathscr{S}^n.$$

In the above sense, *the AHO direction is generically well-defined*.

Proof. The fact that $\dim(\mathrm{AHO}(\mathscr{L})) = n(n+1)$ follows from Theorem 3.11. Note that $\overline{\mathrm{AHO}}(\mathscr{L})$ and therefore AHO($\mathscr{L}$) are both semi-algebraic sets. Since semi-algebraic sets without interior are of measure zero (see Theorem 8.10 in [7]), it suffices to prove that the interior of $\overline{\mathrm{AHO}}(\mathscr{L})$ in $\mathscr{S}^n \times \mathscr{S}^n$ is empty. Note that $\bar{\mathscr{K}}(X,S) := [L_S\mathscr{Z}\vdots - L_X\mathscr{A}^*]$ is linear in X and S. Therefore, $\det(\bar{\mathscr{K}}(X,S)) : \mathscr{S}^n \times \mathscr{S}^n \mapsto \mathscr{R}$ is an analytic function. If this function is zero on an open subset of $\mathscr{S}^n \times \mathscr{S}^n$, then, since the function is analytic, it has to be identically zero on all of $\mathscr{S}^n \times \mathscr{S}^n$ (see, for instance, [8, p. 240]). But we know that there are full-dimensional connected subsets of $\mathscr{S}^n \times \mathscr{S}^n$ over which $\det(\bar{\mathscr{K}}(X,S)) \neq 0$. Therefore, the interior of $\overline{\mathrm{AHO}}(\mathscr{L})$ is empty for every $\mathscr{L} \subseteq \mathscr{S}^n$ and $\overline{\mathrm{AHO}}(\mathscr{L})$ is of measure zero. □

Remark 4.2. It also easily follows from the above proof that $\mathrm{AHO}(\mathscr{L})$ is an open subset of $\mathscr{S}^n \times \mathscr{S}^n$. We can also give an alternative (but related) approach to proving the last theorem (via proving that the interior of $\overline{\mathrm{AHO}}(\mathscr{L})$ is empty). Let $(\overline{X}, \overline{S}) \in \overline{\mathrm{AHO}}(\mathscr{L})$. Let $W \in \mathscr{S}^n_{++}$ (negative definite also works). Then by Lemma 2.2, $L_W(\mathscr{S}^n) = \mathscr{S}^n$. Therefore, there exists $\overline{\epsilon} > 0$ such that for every $\epsilon \in [-\overline{\epsilon}, \overline{\epsilon}] \setminus \{0\}$,

$$L_{\overline{X}+\epsilon W}(\mathscr{L}^\perp) \cap L_{\overline{S}+\epsilon W}(\mathscr{L}) = \{0\}.$$

(Another way of seeing this is to consider $\bar{\mathscr{K}}$. What we are doing here is equivalent to replacing $\bar{\mathscr{K}}$ by $(\bar{\mathscr{K}} + \epsilon I)$; in our case, we are using a nonsingular matrix instead of I. Clearly, for all sufficiently small $\epsilon \neq 0$, the matrix $(\bar{\mathscr{K}} + \epsilon I)$ is nonsingular.) Therefore, there exists $\overline{\epsilon} > 0$ such that for every $\epsilon \in [-\overline{\epsilon}, \overline{\epsilon}] \setminus \{0\}$,

$$(\overline{X} + \epsilon W, \overline{S} + \epsilon W) \in \mathrm{AHO}(\mathscr{L}).$$

5. Extension to the Monteiro–Zhang family

We now present an argument extending our $\gamma \leqslant 9$ result to all search directions in the Monteiro–Zhang family. A simple argument is given by Monteiro [16] to show that MZ result extends to the whole Monteiro–Zhang family (also see [17] for the extension of the MT result). Monteiro's arguments are also applicable to our improvements. However, below we give more mechanical computations showing that the proof techniques also extend.

Recall that $H_P : \mathscr{M}^n \to \mathscr{S}^n$ is the linear transformation

$$H_P(M) := PMP^{-1} + P^{-\mathrm{T}}M^{\mathrm{T}}P^{\mathrm{T}},$$

where P is any $n \times n$ nonsingular matrix. Then in the last equation group (in finding the search direction) apply $H_P(\cdot)$ to both sides and then solve for ΔX and ΔS.

When we write the corresponding homogeneous system, we have

$$P(\Delta X)SP^{-1} + PX(\Delta S)P^{-1} + P^{-\mathrm{T}}S(\Delta X)P^{\mathrm{T}} + P^{-\mathrm{T}}X(\Delta S)P^{\mathrm{T}} = 0.$$

Note that this equation is equivalent to

$$[P(\Delta X)SP^{-1} + PX(\Delta S)P^{-1}] \text{ is skew-symmetric.}$$

Therefore, as in our current arguments for the special case $P = I$, to prove that the direction is well-defined, it suffices to prove that

$$\text{trace}\left\{[P(\Delta X)SP^{-1} + PX(\Delta S)P^{-1}]^2\right\} \geqslant 0.$$

Now, we evaluate the above expression and find that it is equal to:

$$\begin{aligned}&\text{trace}[P(\Delta X)S(\Delta X)SP^{-1}] + \text{trace}[P(\Delta X)SX(\Delta S)P^{-1}]\\ &\quad + \text{trace}[PX(\Delta S)(\Delta X)SP^{-1}] + \text{trace}[PX(\Delta S)X(\Delta S)P^{-1}].\end{aligned}$$

All P's under the trace cancel and we are back to the special case with $P := I$. Therefore, the rest of the current proof of Theorem 3.11 applies and *all search directions in the MZ-family are well-defined if* $\gamma(S^{1/2}XS^{1/2}) = \gamma(V^2) \leqslant 9$.

For the SSK condition, note that $X \mapsto PXP^{\mathrm{T}}$ and $S \mapsto P^{-\mathrm{T}}SP^{-1}$. So,

$$XS + SX \mapsto PXSP^{-1} + P^{-\mathrm{T}}SXP^{\mathrm{T}}.$$

For instance,

$$K \mapsto \begin{cases}[I \otimes (PXSP^{-1} + P^{-\mathrm{T}}SXP^{\mathrm{T}})] + [(PXSP^{-1} + P^{-\mathrm{T}}SXP^{\mathrm{T}}) \otimes I]\\ \quad + 2[(P \otimes P^{-\mathrm{T}})(X \otimes S)(P^{\mathrm{T}} \otimes P^{-1})]\\ \quad + 2[(P^{-\mathrm{T}} \otimes P)(S \otimes X)(P^{-1} \otimes P^{\mathrm{T}})].\end{cases}$$

Again the positive definiteness of the last two terms is clearly implied by $X, S \succ 0$. Therefore, we still have a strengthening of the extension of the SSK result.

Acknowledgment

The authors thank: Kathrin Schäcke, Department of Combinatorics, University of Waterloo, for her work and many helpful conversations; Mike Todd, Cornell University, for pointing out the reference [17]; and we are grateful to Christoph Helmberg for pointing out that an error in the last part of the proof of Theorem 3.11 leads to the improved bound $\frac{4}{5}$. (Our original proof, though it used symmetric Kronecker products, used conditions on the eigenvalues implied only by the usual Kronecker product.) We also thank an anonymous referee for his/her very constructive criticism of the original proof of Theorem 3.9 and the material on Kronecker products.

References

[1] F. Alizadeh, J.-P.A. Haeberly, M.L. Overton, A new primal–dual interior-point method for semidefinite programming, in: J.G. Lewis, (Ed.), Proceedings of the Fifth SIAM Conference on Applied Linear Algebra, SIAM, 1994, pp. 113–117.

[2] F. Alizadeh, J.-P.A. Haeberly, M.L. Overton, Primal–dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results, SIAM J. Optim. 8 (1998) 746–768.

[3] V. Amoia, G. Demicheli, M. Santomauro, Computer-oriented formulation of transition-rate matrices via Kronecker algebra, IEEE Trans. Reliability R-30 (2) (1981) 123–132.

[4] V. Balakrishnan, S. Boyd, Existence and uniqueness of optimal matrix scalings, SIAM J. Matrix Anal. Appl. 16 (1) (1995) 29–39.
[5] J.W. Brewer, Kronecker products and matrix calculus in system theory, IEEE Trans. Circuits and Systems 25 (9) (1978) 772–781., Special issue on the mathematical foundations of system theory.
[6] J.W. Brewer, Correction to: "Kronecker products and matrix calculus in system theory", IEEE Trans. Circuits and Systems 26 (5) (1979) 360 (IEEE Trans. Circuits and Systems 25 (9) (1978) 772–781).
[7] H. Delfs, M. Knebusch, Semialgebraic topology over a real closed field. II. Basic theory of semialgebraic spaces, Math. Z. 178 (2) (1981) 175–213.
[8] H. Federer, Geometric measure theory. Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.
[9] C. Helmberg, F. Rendl, R.J. Vanderbei, H. Wolkowicz, An interior-point method for semidefinite programming, SIAM J. Optim. 6 (2) (1996) 342–361.
[10] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1994, Corrected reprint of the 1991 original.
[11] M. Kojima, M. Shida, S. Shindoh, Search directions in the SDP and the monotone SDLCP: generalization and inexact computation, Math. Programming 85 (1999) 51–80.
[12] M. Kojima, S. Shindoh, S. Hara, Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices, SIAM J. Optim. 7 (1) (1997) 86–125.
[13] S. Kruk, M. Muramatsu, F. Rendl, R.J. Vanderbei, H. Wolkowicz, The Gauss–Newton direction in linear and semidefinite programming, Optim. Methods Software 15 (1) (2001) 1–27.
[14] C.K. Li, C-numerical ranges and C-numerical radii, Linear and Multilinear Algebra 37 (1–3) (1994) 51–82, Special Issue: The numerical range and numerical radius.
[15] R.D.C. Monteiro, Primal–dual path-following algorithms for semidefinite programming, SIAM J. Optim. 7 (3) (1997) 663–678.
[16] R.D.C. Monteiro, Polynomial convergence of primal–dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions, SIAM J. Optim. 8 (1998) 797–812.
[17] R.D.C. Monteiro, M.J. Todd, Path-following methods, in: Handbook of Semidefinite Programming, Kluwer Academic Publishers, Boston, MA, 2000, pp. 267–306.
[18] R.D.C. Monteiro, P.R. Zanjácomo, A note on the existence of the Alizadeh–Haeberly–Overton direction for semidefinite programming, Math. Programming 78 (3, Ser. A) (1997) 393–396.
[19] R.D.C. Monteiro, Y. Zhang, A unified analysis for a class of long-step primal–dual path-following interior-point algorithms for semidefinite programming, Math. Programming 81 (3, Ser. A) (1998) 281–299.
[20] J.J. Moré, Generalizations of the trust region problem, Optim. Methods Software 2 (1993) 189–209.
[21] Y.E. Nesterov, M.J. Todd, Self-scaled barriers and interior-point methods for convex programming, Math. Oper. Res. 22 (1) (1997) 1–42.
[22] B.D. Saunders, H. Schneider, Cones graphs and optimal scalings of matrices, Linear and Multilinear Algebra 8 (2) (1979/80) 121–135.
[23] K. Schaecke, Essay on: The Kronecker product, Master's thesis, University of Waterloo, 2004.
[24] M. Shida, S. Shindoh, M. Kojima, Existence and uniqueness of search directions in interior-point algorithms for the SDP and monotone SDLCP, SIAM J. Optim. 8 (1998) 387–396.
[25] R. Stern, H. Wolkowicz, Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations, SIAM J. Optim. 5 (2) (1995) 286–313.
[26] M.J. Todd, A study of search directions in primal–dual interior-point methods for semidefinite programming, Optim. Methods Software 11&12 (1999) 1–46.
[27] M.J. Todd, K.C. Toh, R.H. Tütüncü, On the Nesterov–Todd direction in semidefinite programming, SIAM J. Optim. 8 (3) (1998) 769–796.
[28] P. Tseng, Further notes on the existence of search directions in interior-point methods for the monotone semi-definite LCP, Technical Report Aug-96, University of Washington, Seattle, WA, 1996.

[29] L. Tunçel, Generalization of primal–dual interior-point methods to convex optimization problems in conic form, Foundat. Comput. Math. 1 (3) (2001) 229–254.
[30] D.A. Turkington, Matrix Calculus and Zero–one Matrices, Cambridge University Press, New York, 2002, Statistical and econometric applications.
[31] H. Wolkowicz, R. Saigal, L. Vandenberghe (Eds.), Handbook Of Semidefinite Programming: Theory, Algorithms, and Application, Kluwer Academic Publishers, Boston, MA, 2000, pp. xxvi+6542.
[32] Y. Zhang, On extending some primal–dual interior-point algorithms from linear programming to semidefinite programming, SIAM J. Optim. 8 (1998) 365–386.