Extremal Characterizations of the Schur Complement and Resulting Inequalities

Chi-Kwong Li and Roy Mathias*
Department of Mathematics
College of William & Mary
Williamsburg, VA 23187.
E-mail: ckli@math.wm.edu   mathias@math.wm.edu

October 26, 2004

Abstract

Let

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$$

be an $n \times n$ *positive semidefinite matrix, where* H_{11} *is* $k \times k$ *with* $1 \leq k < n$*. The generalized Schur complement of* H_{11} *in* H *is defined as*

$$S(H) = H_{22} - H_{12}^* H_{11}^\dagger H_{12},$$

where $H_{11}^\dagger$ *is the Moore-Penrose generalized inverse of* H_{11}*. It has the extremal characterizations*

$$S(H) = \max\left\{ W : H - \begin{pmatrix} 0_k & 0 \\ 0 & W \end{pmatrix} \geq 0, W \text{ is } (n-k) \times (n-k) \text{ Hermitian} \right\}$$

and

$$S(H) = \min \left\{ [Z|I_{n-k}] H [Z|I_{n-k}]^* : Z \text{ is } (n-k) \times k \right\}.$$

These characterizations are used to deduce many old and new inequalities for Schur complements of positive semidefinite matrices. In many cases, stronger statements and shorter proofs can be obtained using the extremal characterizations.

*Both authors were supported by NSF grants DMS-9504795 and DMS-9704534.

1 Introduction

Let

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$$

be an $n \times n$ positive semidefinite matrix, where H_{11} is $k \times k$ with $1 \leq k < n$. The *generalized Schur complement* of H_{11} in H is defined and denoted by

$$S(H) = H_{22} - H_{12}^* H_{11}^\dagger H_{12},$$

where $H_{11}^\dagger$ is the Moore-Penrose generalized inverse of H_{11}. If H_{11} is invertible, then $H_{11}^\dagger = H_{11}^{-1}$ and the concept reduces to the classical Schur complement. We state and prove our results for the generalized Schur complement since using our approach there is no extra work in doing so. We consider the Schur complement with respect to the leading $k \times k$ block in the interests of simplicity of notation. The results in this paper are valid for Schur complements with respect to principal submatrices corresponding to non-contiguous index sets since the positive definite partial order is invariant under symmetric permutation of the rows and columns of a matrix. More generally, one can consider the Schur complement with respect to any compression of the matrix to a k-dimensional subspace as is done in most operator theory literature on the subject (e.g., see [6, 10]).

We always assume, except in Section 4 where we specify otherwise, that $S(H)$ *is the Schur complement taken with respect to the leading* $k \times k$ *submatrix of* H.

We use $\Phi(X)$ to denote the $(n-k) \times (n-k)$ principal submatrix in the bottom right corner of a matrix X. One can check that if H is positive definite (and hence invertible) then

$$S(H)^{-1} = \Phi(H^{-1}). \tag{1.1}$$

The Schur complement has proved useful in many areas (see e.g., [11, 13] and their references), and there are many generalizations, motivated by theory as well as applications (see e.g., [6, 10, 11] and their references). Sophisticated techniques have been developed to deal with the different generalizations. For example, the authors of [6] and [10] used operator theoretic methods to treat a number of generalized Schur complements, and in [11] Carlson approached the topic by skilful matrix analytic techniques. It is not easy for non-experts in matrix or operator theory to understand the useful results on the Schur complement, nor their proofs. In this paper, we suggest studying the generalized Schur complement $S(H)$ using the following two extremal presentations:

$$S(H) = \max\left\{W : H - \begin{pmatrix} 0_k & 0 \\ 0 & W \end{pmatrix} \geq 0, W \text{ is } (n-k) \times (n-k) \text{ Hermitian}\right\} \tag{1.2}$$

and

$$S(H) = \min\left\{[Z|I_{n-k}]H[Z|I_{n-k}]^* : Z \text{ is } (n-k) \times k\right\}. \tag{1.3}$$

These extrema are with respect to the *Loewner partial order*; that is, for Hermitian matrices X, Y of the same order, $X \geq Y$ means that $X - Y$ is positive semidefinite.

There are a couple of advantages to our approach. First, the extremal representations are easily proved as we show in the next section. They do not involve deep theory nor skillful computation, and even may be familiar to many readers.

Second, the theorems are easily applied to deduce (old and new) results on Schur complement as we show in Sections 3 and 4. Our results are often stronger than those in the literature, and the proofs often shorter. It is actually somewhat surprising that these simple proofs have been overlooked by many researchers.

The reason that these characterizations are so useful in proving inequalities is that the Schur complement is a non-linear function of its argument, but in (1.2) and (1.3) we have presented two quasi-linear representations of the Schur Complement.

It may be difficult to obtain (operator or eigenvalue) inequalities involving the non-linear expression $S(H) = H_{22} - H_{12}^* H_{11}^\dagger H_{12}$, but it is much easier to handle matrices W for which

$$H - \begin{pmatrix} 0_k & 0 \\ 0 & W \end{pmatrix} \geq 0.$$

Ando [5, 7] has used the idea of quasi-linear representations very successfully to study different kinds of *means* on positive semi-definite matrices. Similar ideas will also be used in Section 3.

The second representation resembles the well known Courant-Fischer characterization of extremal eigenvalue for the mimimum eigenvalue $\lambda_{\min}(G)$ for a Hermitian matrix G, namely,

$$\lambda_{\min}(G) = \min\{x^* G x : x \in \mathbf{C}^n,\ x^* x = 1\}.$$

This characterization has been extended to more general min-max principles, which lead to many interesting eigenvalue inequalities (see e.g. [18, 19]). In view of this, it is not surprising to see that our second extremal representation of $S(H)$ is particularly useful in studying eigenvalue inequalities (cf. Section 4).

Remember that if H is positive definite then $\Phi(H^{-1}) = S(H)^{-1}$ and so (1.2) and (1.3) can be used to study principal submatrices of H^{-1} as we do in Sections 3 and 4.

2 Extremal Characterizations

In this section, we give short proofs for the extremal characterizations of the generalized Schur complement $S(H)$.

Theorem 2.1 *Let H be an $n \times n$ positive semidefinite matrix. Then*

$$S(H) = \max\left\{ W : H \geq \begin{pmatrix} 0_k & 0 \\ 0 & W \end{pmatrix}, W \text{ is } (n-k)\times(n-k) \text{ Hermitian} \right\}.$$

Proof. Let $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$ be positive semidefinite, and let W be $(n-k) \times (n-k)$ Hermitian. Set $\tilde{W} = \begin{pmatrix} 0_k & 0 \\ 0 & W \end{pmatrix}$ and $T = \begin{pmatrix} I_k & 0 \\ -H_{12}^* H_{11}^\dagger & I_{n-k} \end{pmatrix}$. Since H is positive semidefinite we have $H_{12} = H_{11}^{1/2} X$ for some matrix X, and since H_{11} is positive semidefinite,

$(H_{11}^{1/2})^\dagger = (H_{11}^\dagger)^{1/2}$. This implies that $H_{11}H_{11}^\dagger H_{12} = H_{12}$ and so we have

$$T(H - \tilde{W})T^* = \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} - W - H_{12}^* H_{11}^\dagger H_{12} \end{pmatrix}.$$

We have $H \geq \tilde{W}$ if and only if the matrix on the right hand side is positive semidefinite and this occurs if and only if $S(H) - W \geq 0$. The result follows. □

The characterization in Theorem 2.1 has been observed a number of times before. It is the third of six equivalent generalizations of the Schur complement given by Butler and Morley [10, page 261], who attribute it to M. G. Krein. It was independently proved by Anderson who called $0_k \oplus S(H)$ the shorted operator [2, Theorem 1].

The next characterization of the generalized Schur complement was inspired by an observation in [22, Proof of Theorem 1]. It has also been observed by Ando [7, equation (5) p. 16], who attributes this result also to M.G. Krein.

In [21] Krein did not explicitly state Theorem 2.1, but it is implicit in his Theorem 1 together with the discussion on page 436. He established the extremal representation in our Theorem 2.2 in the process of proving his Theorem 1.

Theorem 2.2 *Let* $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$ *be an* $n \times n$ *positive semidefinite matrix, where* H_{11} *is* $k \times k$*. Then*

$$[Z|I_{n-k}]H[Z|I_{n-k}]^* \geq S(H) \tag{2.1}$$

for any $(n-k) \times k$ *matrix* Z*. Equality holds in (2.1) if and only if*

$$(Z + H_{12}^* H_{11}^\dagger)H_{11} = 0. \tag{2.2}$$

Consequently,

$$S(H) = \min\{[Z|I_{n-k}]H[Z|I_{n-k}]^* : Z \text{ is } (n-k) \times k\}.$$

Proof. Observe that

$$[Z|I_{n-k}]H[Z|I_{n-k}]^* = S(H) + (Z + H_{12}^* H_{11}^\dagger)H_{11}(H_{11}^\dagger H_{12} + Z^*).$$

Thus

$$[Z|I_{n-k}]H[Z|I_{n-k}]^* \geq S(H),$$

and the equality holds if and only if

$$(Z + H_{12}^* H_{11}^\dagger)H_{11}(H_{11}^\dagger H_{12} + Z^*) = 0,$$

which is equivalent to

$$(Z + H_{12}^* H_{11}^\dagger)H_{11} = 0$$

as asserted. □

Note that one can also obtain (2.1) from Theorem 2.1 by conjugating the inequality

$$H \geq \begin{pmatrix} 0_k & 0 \\ 0 & S(H) \end{pmatrix}$$

by $[Z|I_{n-k}]$.

Notice that if H_{11} is invertible then the proof of Theorem 2.1 is somewhat simpler, and the condition (2.2) becomes just $Z = -H_{12}^* H_{11}^{-1}$.

In the next sections we show that despite the simplicity of these two characterizations of the Schur Complement, they easily yield many operator and eigenvalue inequalities.

3 Operator inequalities

In this section, we consider operator inequalities on Hermitian matrices X and Y under the Loewner ordering. We shall show that many operator inequalities can be obtained efficiently using the extremal characterizations in Section 1. Some of the results have been considered by other authors (e.g., see [23, 28, 29]) for positive definite matrices. Using our approach, we can often relax the invertibility assumption. Furthermore, most of our proofs are valid for infinite dimensional operators.

We first present a result that follows immediately from the representation in Theorem 2.2.

Theorem 3.1 *Let* $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{pmatrix}$ *and* $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^* & B_{22} \end{pmatrix}$ *be* $n \times n$ *positive semidefinite matrices such that* A_{11} *and* B_{11} *are* $k \times k$*.*

(a) *We have*

$$S(A+B) \geq S(A) + S(B). \tag{3.1}$$

Equality holds in (3.1) if and only if there exists an $(n-k) \times k$ *matrix* Z *such that*

$$(Z + A_{12}^* A_{11}^\dagger) A_{11} = (Z + B_{12}^* B_{11}^\dagger) B_{11} = 0. \tag{3.2}$$

(b) *If* $A \geq B$*, then*

$$S(A) \geq S(B). \tag{3.3}$$

Equality holds in (3.3) if and only if there exists an $(n-k) \times k$ *matrix* Z *satisfying (3.2) and*

$$[Z|I_{n-k}](A - B) = 0. \tag{3.4}$$

This theorem states that the Schur complement is concave and monotone on the set of positive semidefinite matrices. It has been considered by other authors (see [16], [14, Theorem 1] and [28, Remark 3.1]) for positive definite matrices. In particular, if A and B are positive definite, then the condition (3.2) reduces to

$$Z = -A_{12}^* A_{11}^{-1} = -B_{12}^* B_{11}^{-1}.$$

This was proved by Fiedler and Markham using some rather detailed analysis [14]. Marshall and Olkin deduce the concavity by a more complicated differentiation argument [24, Chapter 16, E.7.h]. Ando [7, Theorem 2.4] and Anderson [2] deduce monotonicity and concavity by an argument similar to ours.

Proof. (a) By Theorem 2.2, there exists an $(n-k)\times k$ matrix Z satisfying

$$\begin{aligned} S(A+B) &= [Z|I_{n-k}](A+B)[Z|I_{n-k}]^* \\ &= [Z|I_{n-k}]A[Z|I_{n-k}]^* + [Z|I_{n-k}]B[Z|I_{n-k}]^* \\ &\geq S(A)+S(B). \end{aligned} \tag{3.5}$$

If the equality holds in (3.5), then (3.2) holds by Theorem 2.2.

Conversely, if there exists an $(n-k)\times k$ matrix Z such that (3.2) holds, then

$$\begin{aligned} S(A)+S(B) &= [Z|I_{n-k}]A[Z|I_{n-k}]^* + [Z|I_{n-k}]B[Z|I_{n-k}]^* \\ &= [Z|I_{n-k}](A+B)[Z|I_{n-k}]^* \\ &\geq S(A+B). \end{aligned} \tag{3.6}$$

We have shown that $S(A+B)\geq S(A)+S(B)$, and so we must have equality in (3.6), as desired.

To prove (b), suppose $A\geq B$. Let Z be an $(n-k)\times k$ matrix satisfying $(Z+A_{12}^*A_{11}^\dagger)A_{11} = 0$ so that $[Z|I_k]A[Z|I_k]^* = S(A)$ by Theorem 2.2. Using Theorem 2.2 for the second inequality, we have

$$S(A) = [Z|I_k]A[Z|I_k]^* \geq [Z|I_k]B[Z|I_k]^* \geq S(B). \tag{3.7}$$

To analyse the equality in (3.3), note that the second inequality holds in (3.7) if and only if Z satisfies $(Z+B_{12}^*B_{11}^\dagger)B_{11}=0$, and the first inequality holds in (3.7) if and only if

$$[Z|I_{n-k}](A-B)[Z|I_{n-k}]^*=0.$$

Since $A-B\geq 0$, the last condition is equivalent to (3.4). □

The inequalities (3.1) and (3.3) can be proved even more easily from Theorem 2.1. We have used Theorem 2.2 here since want to derive conditions for equality also.

One can combine Theorem 2.1 with various operations that preserve the positive definite order to derive a variety of operator inequalities. Here are some.

Theorem 3.2 *Let f be matrix monotone on $[0,\infty)$ and such that $f(0)\geq 0$. Then for any $n\times n$ positive semidefinite matrix A*

$$f(\Phi(A)) \geq \Phi(f(A)) \geq S(f(A)) \geq f(S(A)).$$

Proof. First consider the left most inequality. It is known that a function that is operator monotone on $[0,\infty)$ is necessarily operator concave on $[0,\infty)$, see e.g., [8, Theorem V.2.5] or [15] for the basic result, and see [25, Theorem 2.1] for a simple extension to monotone matrix functions of finite order. Since f is matrix concave it can be shown that

$$f(\Phi(A)) \geq \Phi(f(A)).$$

See [15, Lemma 1] or [8, Theorem V.2.3 (ii)] for the details.

Alternatively, one can deduce the leftmost inequality directly from [5, Theorem 4] since $\Phi(X)$ is a normalized positive liner map.

Now consider the middle inequality. Since A is positive semidefinite and $f(0) \geq 0$ it follows that $f(A)$ is positive semidefinite. We know that $\Phi(M) \geq S(M)$ for any positive semidefinite matrix M, and so we have the second inequality.

Finally, consider the rightmost inequality. Since f satisfies the hypotheses we have

$$f(A) \geq f\left(\begin{pmatrix} 0_k & 0 \\ 0 & S(A) \end{pmatrix}\right) = \begin{pmatrix} f(0_k) & 0 \\ 0 & f(S(A)) \end{pmatrix} \geq \begin{pmatrix} 0_k & 0 \\ 0 & f(S(A)) \end{pmatrix}.$$

By Theorem 2.1, we have

$$S(f(A)) \geq f(S(A))$$

as desired. □

There are many functions that are matrix monotone–perhaps the best known ones are $f(t) = t^p$, with $0 \leq p \leq 1$, and we will apply the result with these functions below. The function $f(t) = \log(1 + t)$ also satisfies the hypotheses of Theorem 3.2. See [19, §6.6] or [8, Chapter 5] for a useful survey of results on monotone matrix functions, some examples, and further references to the literature. Applying Theorem 3.2 in clever ways, one may get different inequalities as shown in the following corollaries.

Corollary 3.3 *Let A be an $n \times n$ positive semidefinite matrix.*

(a) *If $p \geq 1$, then*

$$[\Phi(A^p)]^{1/p} \geq \Phi(A) \geq S(A) \geq [S(A^p)]^{1/p}. \tag{3.8}$$

(b) *If A is invertible and $p \leq -1$, then*

$$[\Phi(A^p)]^{1/p} \leq S(A) \leq \Phi(A) \leq [S(A^p)]^{1/p}. \tag{3.9}$$

Proof. Applying Theorem 3.2 with $f(t) = t^{1/p}$, $p \geq 1$ to the matrix A^p gives (3.8).

If A is invertible, one can replace A by A^{-1} in (3.8), take inverses of all the terms, there by reversing the inequalities, and obtain the inequality (3.9). □

Wang and Zhang [28, Theorem 3] obtained the right hand inequalities in (3.8) and (3.9) with the added restriction that p be an integer.

Another operation that preserves the positive semidefinite partial order is that of taking a Schur product with a positive semidefinite matrix. This idea yields the next result. Let $A \circ B$ denote the Schur product, or entry-wise product, of two matrices A and B of the same order.

Corollary 3.4 *Let A and B be $n \times n$ positive semidefinite matrices. Then*

$$S(A \circ B) \geq S(A) \circ \Phi(B) \geq S(A) \circ S(B). \tag{3.10}$$

Proof. The Schur Product Theorem (e.g., [18, Theorem 7.5.3]) implies that the Schur product with a positive definite matrix preserves the positive semidefinite partial order. Thus

$$A \circ B \geq \begin{pmatrix} 0_k & 0 \\ 0 & S(A) \end{pmatrix} \circ B = \begin{pmatrix} 0_k & 0 \\ 0 & S(A) \circ \Phi(B) \end{pmatrix}. \tag{3.11}$$

Theorem 2.1 now implies the left hand inequality in the corollary. The right hand inequality follows from another application of the Schur Product Theorem since $\Phi(B) \geq S(B)$. □

Wang and Zhang [28, Theorem 2] (see also the remark following their theorem) used a similar, but longer and more computational, proof to obtain the same inequality. Markham and Smith [23, Theorem 1.2] have recently obtained

$$S(A \circ B) \geq S(A) \circ S(B), \tag{3.12}$$

which is weaker in that it does not contain the middle term in (3.10). They showed that the equality in (3.12) holds if and only if $A = A_{11} \oplus A_{22}$ and $B = B_{11} \oplus B_{22}$, where A_{11} and B_{22} are $k \times k$.

Note that one can use $S(A \circ B) \geq S(A) \circ \Phi(B)$ to give a simple proof of Oppenheim's inequality for positive definite matrices:

$$\det(A \circ B) \geq \det(A) \prod_{i=1}^{n} b_{ii}.$$

This is essentially the approach taken in [18, proof of Theorem 7.8.6].

The next proposition may appear artificial, but it has a lot of applications, especially to the study of different kinds of means of positive semidefinite matrices, as we will see. We omit the proof as it is an application of Theorem 2.1, and is similar to the proofs of Theorem 3.2 and Corollary 3.4. Let $\mathcal{H}_n$ denote the cone of $n \times n$ positive semidefinte matrices.

Proposition 3.5 *Let* $f : \overbrace{\mathcal{H}_n^+ \times \cdots \times \mathcal{H}_n^+}^{p} \to \mathcal{H}_n^+$ *and* $g : \overbrace{\mathcal{H}_{n-k}^+ \times \cdots \times \mathcal{H}_{n-k}^+}^{p} \to \mathcal{H}_{n-k}^+$ *be such that*

$$f\left(\begin{pmatrix} 0_k & 0 \\ 0 & A_1 \end{pmatrix}, \ldots, \begin{pmatrix} 0_k & 0 \\ 0 & A_p \end{pmatrix}\right) = \begin{pmatrix} 0_k & 0 \\ 0 & g(A_1, \ldots, A_p) \end{pmatrix}$$

where $A_i \in \mathcal{H}_{n-k}$. *Suppose also that* f *is monotone in the sense that if* $B_i \geq C_i \geq 0$, *for* $i = 1, \ldots, p$, *then*

$$f(B_1, \ldots, B_p) \geq f(C_1, \ldots, C_p).$$

Then for any $H_i \geq 0$ *we have*

$$S(f(H_1, \ldots, H_p)) \geq g(S(H_1), \ldots, S(H_p)). \tag{3.13}$$

Taking $p = 1$ and $f(H) = H \circ B$ and $g(K) = K \circ \Phi(B)$ yields the left hand inequality in (3.10). Taking $p = 2$ and $f(H_1, H_2) = H_1 + H_2$ yields Theorem 3.1 (b). If we had taken

$f(H_1, H_2) = (H_1 + H_2)/2$, then we would get an inequality involving the arithmetic mean and the Schur complement:

$$S((H_1 + H_2)/2) \geq (S(H_1) + S(H_2))/2.$$

One could also take $p = 2$, and f to be either the *harmonic mean* defined by

$$H_1!H_2 \equiv 2(H_1^{-1} + H_2^{-1})^{-1} \tag{3.14}$$

when H_1 and H_2 are both positive definite and by $\lim_{\epsilon\downarrow 0}(H_1 + \epsilon I)!(H_2 + \epsilon I)$ otherwise, or the *geometric mean* defined by

$$H_1\#H_2 \equiv H_1^{1/2}(H_1^{-1/2} H_2 H_1^{-1/2})^{1/2} H_1^{1/2}. \tag{3.15}$$

Note that $(A!B)/2$, half the harmonic mean, is the same as the parallel sum of A and B which was introduced by Anderson and Duffin [3]. It is easy to see that the harmonic mean is symmetric and monotone. The geometric mean is monotone in H_2, because of the monotonicity of the square root. It turns out that it is also symmetric in H_1 and H_2; and so by symmetry, it is also monotone in H_1. Consequently, Proposition 3.5 yields:

Corollary 3.6 *Let A and B be $n \times n$ positive semidefinite. Then*

$$S(A)!S(B) \leq S(A!B), \tag{3.16}$$

and

$$S(A)\#S(B) \leq S(A\#B). \tag{3.17}$$

Corollary 3.6 can also be proved using the following extremal characterizations of the harmonic and geometric means of positive semidefinte matrices A and B (see [5]):

$$A!B = \max\left\{C : C = C^*, \begin{pmatrix} A - C & -C \\ -C & B - C \end{pmatrix} \geq 0\right\},$$

$$A\#B = \max\left\{C : C = C^*, \begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0\right\}.$$

There are many other means defined for p-tuples of positive semidefinite matrices (e.g., see [4] and its references). All of these satisfy the monotonicity property required in Proposition 3.5, and so for these means also we have "mean of the Schur complements is less than or equal to Schur complement of the mean".

4 Eigenvalue inequalities

It is inconvenient that H and $S(H)$ are of different orders and act on different spaces. For this reason, we define $\tilde{H} = 0_k \oplus S(H)$, a matrix of the same order as H. This definition has been used in [2, 6, 10].

Whenever an $n \times n$ matrix H has only real eigenvalues we shall order them $\lambda_1(H) \geq \cdots \geq \lambda_n(H)$. We deduce eigenvalue inequalities as corollaries of operator inequalitites using the well known monotonicity principle: *if* $A \geq B$, *then* $\lambda_j(A) \geq \lambda_j(B)$ *for all* j (see e.g. [19, Chapter 3]). For an $m \times n$ complex matrix X, let $\sigma_i(X) = \sqrt{\lambda_i(X^*X)}$ denote the ith singular value for $i = 1, \ldots, k$, where $k = \min\{m, n\}$, and let $\sigma(X) = (\sigma_1(X), \ldots, \sigma_k(X))$ be the vector of singular values of X.

There are many inequalities on the eigenvalues of a sum of Hermitian matrices. Suppose that the $n \times n$ Hermitian matrices A, B, C satisfy $A + B = C$. Then the possible eigenvalues of A, B, C are completely characterized by inequalities of the form

$$\sum_{r\in R} \lambda_r(A) + \sum_{s\in S} \lambda_s(B) \geq \sum_{t\in T} \lambda_t(C),$$

for suitable subsets $R, S, T \subseteq \{1, \ldots, n\}$ of the same sizes. Moreover, since $\operatorname{tr} A + \operatorname{tr} B = \operatorname{tr} C$, we have

$$\sum_{r\in R'} \lambda_r(A) + \sum_{s\in S'} \lambda_s(B) \leq \sum_{t\in T'} \lambda_t(C),$$

where R', S' and T' denote the complements of R, S and T in $\{1, \ldots, n\}$. Sometimes, it is convenient and useful to state the inequalities in this direction. See [17, 20] (and also [9]) for the details. In the interests of simplicity we shall use restrict ourselves to the special case

$$R = \{i_1, \ldots, i_m\}, \quad S = \{j_1, \ldots, j_m\}, \quad T = \{i_1 + j_1 - 1, \ldots, i_m + j_m - m\}, \tag{4.1}$$

where $1 \leq i_1 < \cdots < i_m \leq n$ and $1 \leq j_1 < \cdots < j_m \leq n$. This case was originally proved by Thompson, and a simplified proof was given in [26]. Though we restrict ourselves to this simple case the inequalities we deduce are already more general than those in the literature.

In fact, such inequalities with suitable choices of R, S, T, completely characterize the relation between the eigenvalues of three Hermitian matrices A, B, C satisfying $A + B = C$ Similarly, for general $n \times n$ matrices X, Y, Z such that $XY = Z$, there are singular value inequalities (see e.g. [27]) of the form

$$\prod_{r\in R} \sigma_r(X) \prod_{s\in S} \sigma_s(Y) \geq \prod_{t\in T} \sigma_t(Z),$$

with R, S, T as in (4.1), and for even more general R, S, T. Using these types of inequalities and the results in the previous sections, one can obtain many eigenvalue inequalities involving Schur complements.

It is always desirable to obtain a family of basic inequalities of the form

$$\prod_{j=1}^{m} x_j \leq \prod_{j=1}^{m} y_j \qquad \text{or} \qquad \sum_{j=1}^{m} x_j \leq \sum_{j=1}^{m} y_j$$

for positive numbers $x_1 \geq \cdots \geq x_n$ and $y_1 \geq \cdots \geq y_n$. Then one can apply the theory of majorization to conclude that

$$f(x_1, \ldots, x_n) \leq f(y_1, \ldots, y_n)$$

whenever f is a Schur convex functions (see [24] for a thorough treatment of majorization and Schur convexity). Common examples of Schur convex functions include the ℓ_p norms defined by

$$f(x_1, \ldots, x_n) = \left\{ \sum_{j=1}^{p} x_j^p \right\}^{1/p}$$

with $p \geq 1$, and elementary symmetric functions

$$E_m(x_1, \ldots, n) = \sum_{1 \leq i_1 < \cdots < i_m \leq n} x_{i_1} \ldots, x_{i_m}$$

with $1 \leq m \leq n$.

In the interests of the simplicity of the statement of our results, we shall use only the inequalities in [26] and [27]. See [20] for the most general inequalities. Special cases of our results reduce to those of others. For easy comparisons with the results of others, and for simplicity of notation, we *assume that $S(A)$ is the Schur complement of the leading $(n-k) \times (n-k)$ leading principal submatrix in Theorems 4.1 and 4.2.*

Theorem 4.1 *Let A and B be $n \times n$ positive semidefinite matrices. Let $1 \leq i_1 < \cdots < i_m \leq k$ and $1 \leq j_1 < \cdots < j_m \leq k$ such that $i_1 + j_1 \geq k - m + 2$, then*

$$\sum_{s=1}^{m} \lambda_{i_s + j_s + m - k - s}(S(A+B)) \geq \sum_{s=1}^{m} \lambda_{i_s}(S(A)) + \sum_{s=1}^{m} \lambda_{j_s}(S(B)).$$

Proof. From Theorem 3.1, or directly from Theorem 2.1 or even Theorem 2.2, it follows that

$$S(A+B) \geq S(A) + S(B),$$

and hence

$$\lambda_i(S(A+B)) \geq \lambda_i(S(A) + S(B)), \quad i = 1, \ldots, k.$$

The desired bound follows from this and [26, Theorem 1], which is an inequality relating the sum of certain eigenvalues of the sum of two Hermitian matrices and the sums of the eigenvalues of the individual matrices. □

If $(i_1, \ldots, i_m) = (k-m+1, k-m+2, \ldots, k)$, then for any $1 \leq j_1 < \cdots < j_m \leq k$ we have

$$\sum_{s=1}^{m} \lambda_{j_s}(S(A+B)) \geq \sum_{s=1}^{m} \lambda_{k-s+1}(S(A)) + \sum_{s=1}^{m} \lambda_{j_s}(S(B)).$$

Similarly, we have

$$\sum_{s=1}^{m} \lambda_{j_s}(S(A+B)) \geq \sum_{s=1}^{m} \lambda_{j_s}(S(A)) + \sum_{s=1}^{m} \lambda_{k-s+1}(S(B)).$$

Combining these, we obtain [22, Theorem 5]. Using a similar substitution, we see that the next theorem generalizes the results in [22, Theorems 2–4].

Theorem 4.2 *Suppose A and B are $n \times n$ complex matrices such that A is positive semidefinite. Then for any $1 \leq i_1 < \cdots < i_m \leq k$ and $1 \leq j_1 < \cdots < j_m \leq n$, we have*

(a) *If $i_m + j_m \leq m + k$, then*

$$\prod_{s=1}^{m} \lambda_{i_s+j_s-s}(S(BAB^*)) \leq \prod_{s=1}^{m} [\lambda_{i_s}(S(BB^*))\lambda_{j_s}(A)].$$

(b) *If $i_1 + j_1 \geq k - m + 2$, then*

$$\prod_{s=1}^{m} \lambda_{i_s+j_s+m-k-s}(S(BAB^*)) \geq \prod_{s=1}^{m} [\lambda_{i_s}(S(BB^*))\lambda_{j_s}(A)].$$

Proof. Let Z be such that $S(BB^*) = [Z|I_k]BB^*[Z|I_k]^*$. Set $C = B^*[Z|I_k]^*$. Then the squares of the k largest singular values of C^* are the eigenvalues of $S(BB^*)$, and the squares of the k largest singular values of $A^{1/2}C$ are the eigenvalues of C^*AC.

To prove (a), suppose $1 \leq i_1 < \cdots < i_m \leq k$ and $1 \leq j_1 < \cdots < j_m \leq n$ satisfy $i_m + j_m \leq m + k$. By Theorem 2.2

$$S(BAB^*) \leq [Z|I_k]BAB^*[Z|I_k]^* = C^*AC.$$

Using this for the first inequality and [27, Theorem 1] for the second, we have and so

$$\begin{aligned}
\prod_{s=1}^{m} \lambda_{i_s+j_s-s}(S(BAB^*)) &\leq \prod_{s=1}^{m} \lambda_{i_s+j_s-s}(C^*AC) \\
&= \prod_{s=1}^{m} \lambda_{i_s+j_s-s}((A^{1/2}C)^*A^{1/2}C) \\
&= \prod_{s=1}^{m} \sigma^2_{i_s+j_s-s}(A^{1/2}C). \\
&\leq \prod_{s=1}^{m} [\sigma^2_{i_s}(C)\sigma^2_{j_s}(A^{1/2})] \\
&= \prod_{s=1}^{m} [\lambda_{i_s}(S(BB^*))\lambda_{j_s}(A)].
\end{aligned}$$

The proof of (b) is similar. □

Using the result in [27, p.109], we see that $\prod$ can be replaced by $\sum$ in the above theorem.

Corollary 4.3 *Suppose A and B are $n \times n$ complex matrices such that A is positive semidefinite. Then for any $1 \leq i_1 < \cdots < i_m \leq k$ and $1 \leq j_1 < \cdots < j_m \leq n$, we have*

(a) *If $i_m + j_m \leq m + k$, then*

$$\sum_{s=1}^{m} \lambda_{i_s+j_s-s}(S(BAB^*)) \leq \sum_{s=1}^{m} [\lambda_{i_s}(S(BB^*))\lambda_{j_s}(A)].$$

(b) *If* $i_1 + j_1 \ge k - m + 2$, *then*

$$\sum_{s=1}^{m} \lambda_{i_s+j_s+m-k-s}(S(BAB^*)) \ge \sum_{s=1}^{m} [\lambda_{i_s}(S(BB^*))\lambda_{j_s}(A)].$$

Note that if A and B are positive semidefinite, then $A \circ B$ is also positive semidefinite, and AB has nonnegative eigenvalues. We have the following result.

Theorem 4.4 *Let A and B be $n \times n$ positive semidefinite matrices. Let $\Phi(A)$ and $\Phi(B)$ denote the $(n-k)\times(n-k)$ principal submatrices in the bottom right of A and B respectively.*

(a) *For* $1 \le i \le n-k$, $\lambda_i(S(A \circ B)) \ge \lambda_i(S(A) \circ \Phi(B)) \ge \lambda_i(S(A) \circ S(B))$.

(b) *For* $1 \le i \le n-k$, $\lambda_i(\Phi(A)\Phi(B)) \ge \lambda_i(\Phi(A)S(B)) \ge \lambda_i(S(A)S(B))$.

(c) *For* $1 \le i \le n-k$, $\lambda_i(AB) \ge \lambda_i(\Phi(A)S(B)) \ge \lambda_{k+i}(AB)$.

The bounds on $\lambda_i(\Phi(A)S(B))$ in (b) and (c) are not comparable as we now show. Let

$$W = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 & 1 \\ 1 & 1.1 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & -1 \\ -1 & 1.1 \end{pmatrix}.$$

Then $\lambda_1(WX) < \lambda_1(\Phi(W)\Phi(X))$ but $\lambda_1(WW) > \lambda_1(\Phi(W)\Phi(W))$, and $\lambda_1(S(Y)S(Y)) > \lambda_2(YY)$, but $\lambda_1(S(Y)S(Z)) < \lambda_2(YZ)$.

Proof. The inequalitites in (a) follow immediately from Corollary 3.4 and the the monotonicity principle.

Note that if X, Y, Z are positive semidefinite and $X \ge Y$, then

$$\lambda_i(XZ) = \lambda_i(Z^{1/2}XZ^{1/2}) \ge \lambda_i(Z^{1/2}YZ^{1/2}) = \lambda_i(YZ) \tag{4.2}$$

for all i. The inequalities in (b) now follow immediately from $\Phi(X) \ge S(X)$ for positive semidefinite X.

For (c) note that

$$A\tilde{B} = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{pmatrix} \begin{pmatrix} 0_k & 0 \\ 0 & S(B) \end{pmatrix} = \begin{pmatrix} 0_k & A_{12}S(B) \\ 0 & A_{22}S(B) \end{pmatrix},$$

where $\tilde{X} = 0_{n-k} \oplus S(X)$. Hence $\lambda_i(AB) \ge \lambda_i(A\tilde{B}) = \lambda_i(A_{22}S(B)) = \lambda_i(\Phi(A)S(B))$ which is the first inequality in (c). For the second inequality, first suppose A and B are invertible. We have just shown

$$\lambda_j(B^{-1}A^{-1}) \ge \lambda_j(\Phi(B^{-1})S(A^{-1})). \tag{4.3}$$

For a $p \times p$ matrix X with positive eigenvalues $\lambda_j(X) = \lambda_{p-j+1}^{-1}(X^{-1})$. Also, from (1.1) it follows that $\Phi(B^{-1})^{-1} = S(B)$ and that $S(A^{-1})^{-1} = \Phi(A)$ and hence (4.3) is equivalent to

$$\lambda_{n-j+1}^{-1}(AB) \ge \lambda_{n-k-j+1}^{-1}(\Phi(A)S(B)).$$

Now put $j = n - k + 1 - i$ and take inverses. This yields the asserted inequality, at least in the case that A and B are invertible. By continuity, we get the result for singular matrices as well. □

S. Fallat informed us that the special case of (c) when $k = 1$ has been observed independently by C. Johnson and T. Markham (unpublished).

Acknowledgment We thank T. Ando, T. Markham and F. Zhang for some helpful suggestions and for supplying the preprints [7], [23] and [29]. We also thank I. Spitkovsky for his help with the paper [21].

References

[1] A. Albert. Conditions for positive and nonnegative definiteness in terms of pseudoinverse. *SIAM J. Appl. Math.*, 17:434–440, 1969.

[2] W. N. Anderson. Shorted operators. *SIAM J. Appl. Math.*, 20:520–525, 1969.

[3] W. N. Anderson and R. J. Duffin. Series and parallel addition of matrices. *J. Math. Anal. Appl.*, 26:576–594, 1969.

[4] W.N. Anderson, Fr. and T.D. Morley. Symmetric functions means of positive operators. *Linear Algebra Appl.,* 60:129-143, 1984.

[5] T. Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. *Linear Algebra Appl.*, 26:203–241, 1979.

[6] T. Ando. Generalized Schur complements. *Linear Algebra Appl.*, 27:173–186, 1979.

[7] T. Ando. Operator-theoretic methods for matrix inequalities. *Unpublished monograph*, 1998.

[8] R. Bhatia. *Matrix Analysis.* Springer-Verlag, New York, 1997.

[9] A.S. Buch. The saturation conjecture, preprint. Available at http://xxx.lanl.gov, math.CO/9810180.

[10] C.A. Butler and T.D. Morley. Six generalized Schur complements. *Linear Algebra Appl.*, 106:259–269, 1988.

[11] D. Carlson. What are Schur complements, anyway? *Linear Algebra Appl.*, 74:257–275 (1986).

[12] D. Carlson, E. Haynsworth, and T. Markham. A generalization of the Schur complement by means of the Moore-Penrose inverse. *SIAM J. Appl. Math.*, 26:169–179, 1974.

[13] R.W. Cottle. Manifestations of the Schur complement. *Linear Algebra Appl.*, 8:189–211 (1974).

[14] M. Fiedler and T. Markham. Some results on the Bergstrom and Minkowskii inequalities. *Linear Algebra Appl.*, 232:199–211, 1996.

[15] S. Friedland and M. Katz. On a matrix inequality. *Linear Algebra Appl.*, 85:185–190, 1987.

[16] E.V. Haynsworth. Applications of an inequality for the Schur complement. *Proc. Amer. Math. Soc.*, 24:512–516, 1970.

[17] A. Horn. Eigenvalues of sums of Hermitian matrices. *Pacific J. Math.,* 12:225-241, (1962),.

[18] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, New York, 1985.

[19] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, New York, 1991.

[20] A.A. Klyachko, Stable bundles, representation theory and Hermitian operators. *Selecta Math.* (N.S.), 3:419-445, 1998.

[21] M.G. Krein, The theory of self adjoint extensions of semibounded Hermitian transformations and its applications I. *Mat. Sb.,* 20:431-495, 1947. (Russian with English summary.)

[22] J. Liu and L. Zhu. A minimum principle and estimates of eigenvalues for Schur complements positive semidefinite matrices. *Linear Algebra Appl.*, 265:123–145, 1997.

[23] T. Markham and R. Smith. A Schur complement inequality for certain P-matrices. *preprint*, 1998.

[24] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications.* Academic Press, London, 1979.

[25] R. Mathias. Concavity of monotone matrix functions of finite order. *Linear Multilinear Algebra*, 27:129–138, 1990.

[26] R. C. Thompson and L. J. Freede. On the eigenvalues of sums of Hermitian matrices. *Linear Algebra Appl.*, 4:369–376, 1971.

[27] R. C. Thompson and S. Therianos. On the singular values of a matrix product-I, II, III. *Scripta Math.*, 29:99–123, 1973.

[28] B.-Y. Wang and F. Zhang. Schur complements and matrix inequalities of Hadamard products. *Linear Multilinear Algebra*, 43:315–326, 1997.

[29] B.-Y. Wang, X. Zhang, and F. Zhang. Some inequalities on generalized Schur complements. *preprint*, 1998.