Some Convexity Features Associated with Unitary Orbits

Man-Duen Choi, Chi-Kwong Li and Yiu-Tung Poon

Abstract

Let $\mathcal{H}_n$ be the real linear space of $n \times n$ complex Hermitian matrices. The unitary (similarity) orbit $\mathcal{U}(C)$ of $C \in \mathcal{H}_n$ is the collection of all matrices unitarily similar to C. We characterize those $C \in \mathcal{H}_n$ such that every matrix in the convex hull of $\mathcal{U}(C)$ can be written as the average of two matrices in $\mathcal{U}(C)$. The result is used to study spectral properties of submatrices of matrices in $\mathcal{U}(C)$, the convexity of images of $\mathcal{U}(C)$ under linear transformations, and some related questions concerning the joint C-numerical range of Hermitian matrices. Analogous results on real symmetric matrices are also discussed.

Keywords: Hermitian matrix, unitary orbit, eigenvalue, joint numerical range.
AMS Classifications: 15A60, 15A42.

1 Introduction

Let $\mathcal{H}_n$ be the real linear space of $n \times n$ complex Hermitian matrices and $\mathcal{U}_n$ the $n \times n$ unitary matrices. The *unitary* (*similarity*) *orbit* $\mathcal{U}(C)$ of $C \in \mathcal{H}_n$ is the collection of all matrices unitarily similar to C, i.e.,

$$\mathcal{U}(C) = \{U^*CU : U \in \mathbf{C}^{n\times n},\ U \in \mathcal{U}_n\}.$$

Evidently, $\mathcal{U}(C)$ can be viewed as the orbit of $C \in \mathcal{H}_n$ under the action of the unitary group $\mathcal{U}_n$ via similarity, and hence $\mathcal{U}(C)$ is a real homogeneous manifold with nice geometrical property (see e.g. [36]). In fact, $\mathcal{U}(C)$ is an interesting geometrical object even under the usual geometry. For instance, since all the matrices in $\mathcal{U}(C)$ has the same Frobenius norm as C, the set $\mathcal{U}(C)$ is part of the boundary of a strictly convex set. As a result, any three different points in $\mathcal{U}(C)$ are not collinear as noted in [35], and every point in $\mathcal{U}(C)$ is an extreme point of $\operatorname{conv}\mathcal{U}(C)$, the convex hull of $\mathcal{U}(C)$.

Suppose $C \in \mathcal{H}_n$ is not a scalar matrix. Then the affine span of $\mathcal{U}(C)$ is the set of matrices with trace equal to that of C (see e.g., [33]). Hence, $\operatorname{conv}\mathcal{U}(C)$ has (real) affine dimension n^2-1. It is known that every element in a convex set with (real) affine dimension n^2-1 can be written as the convex combination of n^2 extreme points. By a result in [3], every matrix X in $\operatorname{conv}\mathcal{U}(C)$ can be written as the *average* of n matrices in $\mathcal{U}(C)$. For some special $C \in \mathcal{H}_n$, the value n can be further reduced. For example, if C is a rank k orthogonal projection with $1 \le k < n$, then every matrix $X \in \operatorname{conv}\mathcal{U}(C)$ can be written as an average of $\max\{k+1, n-k+1\}$ matrices in $\mathcal{U}(C)$ (see [9, Theorem 3.5]). Such results have interesting applications including the proofs of many matrix and norm inequalities.

In Section 2, we characterize those matrices $C \in \mathcal{H}_n$ with the following special geometrical feature: every matrix $X \in \operatorname{conv}\mathcal{U}(C)$ can be written as the average of *two* matrices in $\mathcal{U}(C)$. Recall that $\mathcal{U}(C)$ is the set of extreme points of $\operatorname{conv}\mathcal{U}(C)$. In general, we say that a convex set is **mid-point convex** if every element in it is the mid-point of two extreme points. Clearly, every mid-point convex set is a convex set, and singletons are mid-point convex sets.

In $\mathbf{R}^2$, it is not hard to see that a compact convex set S is mid-point convex if and only if every boundary point of S is an extreme point. For higher dimensional linear spaces, characterizing mid-point convex sets is more difficult. In matrix spaces, an interesting example of mid-point convex set is the set S of $n \times n$ complex matrices with norm at most one; it is known that if X is such a matrix and if $X = PU$ for a positive semi-definite matrix P and $U \in \mathcal{U}_n$, then $X = (X_1 + X_2)/2$, where $X_1 = (P + i\sqrt{I - P^2})U$ and $X_2 = (P - i\sqrt{I - P^2})U$ are extreme points of S.

For $\operatorname{conv} \mathcal{U}(C)$, if $n = 2$ then every $\operatorname{conv} \mathcal{U}(C)$ is mid-point convex. To see this, observe that if C has eigenvalues $c_1 \geq c_2$, then every $X \in \operatorname{conv} \mathcal{U}(C)$ has eigenvalues $d_1 \geq d_2$ such that $c_1 \geq d_1 \geq d_2 \geq c_2$ and $c_1 + c_2 = d_1 + d_2$; see Proposition 1.1 below. Thus, there exists $U \in \mathcal{U}_n$ such that $X = U^* D U$, where D is the diagonal matrix with diagonal entries d_1 and d_2. Let $d = \sqrt{(c_1^2 + c_2^2 - d_1^2 - d_2^2)/2}$,

$$C_1 = U^* \begin{pmatrix} d_1 & d \\ d & d_2 \end{pmatrix} U \qquad \text{and} \qquad C_2 = U^* \begin{pmatrix} d_1 & -d \\ -d & d_2 \end{pmatrix} U.$$

Then $C_1, C_2 \in \mathcal{U}(C)$ and $X = (C_1 + C_2)/2$.

In general, we show in Theorem 2.1 that $\operatorname{conv} \mathcal{U}(C)$ is mid-point convex if and only if the eigenvalues of C form an arithmetic progression, i.e.,

$$C \text{ has eigenvalues } a,\ a+b,\ \ldots,\ a+(n-1)b, \qquad \text{for some } a, b \in \mathbf{R}\,. \tag{1}$$

Of course, this includes the trivial case when C is a scalar matrix, equivalently, $\mathcal{U}(C) = \operatorname{conv} \mathcal{U}(C)$ is a singleton. It is interesting to see that Hermitian matrices C satisfying (1) can be characterized by the mid-point convexity on $\operatorname{conv} \mathcal{U}(C)$, which is such a simple geometrical property. On the contrary, there does not seem to be simple algebraic ways to characterize such matrices.

In Section 3, we study the spectral properties (such as the possible eigenvalues or singular values) of submatrices of a given matrix in $C \in \mathcal{H}_n$ with prescribed eigenvalues. Such problems have been studied by other researchers; see [12, 27] and their references. Our approach is to consider the whole collection of submatrices of matrices in $\mathcal{U}(C)$ lying in certain rows and columns. For example, we consider the set $P_k(C)$ of $k \times k$ (leading) principal submatrices of matrices in $\mathcal{U}(C)$, where $k \leq n$. We give a complete description of $P_k(C)$ and $\operatorname{conv} P_k(C)$. Moreover, we characterize those $C \in \mathcal{H}_n$ for which $P_k(C)$ is convex. Also, we use the main result (Theorem 2.1) in Section 2 to study the set

$$S_k(C) = \left\{ X \in \mathbf{C}^{k \times (n-k)} : \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathcal{U}(C) \text{ for some } \ P \in \mathcal{H}_k,\ Q \in \mathcal{H}_{n-k} \right\}.$$

It turns out that a complete description of $S_k(C)$ is quite intricate[1]. We obtain complete description of $\operatorname{conv} S_k(C)$, and characterize those $C \in \mathcal{H}_n$ for which $S_k(C)$ is convex. It is shown in Theorem 3.3 that for a given $C \in \mathcal{H}_n$ with eigenvalues $c_1 \geq \cdots \geq c_n$, $S_k(C)$ is convex if and only if the sequences $(c_1, \ldots, c_p)$ and $(c_{n-p+1}, \ldots, c_n)$ are arithmetic progressions with

[1] A complete description of $S_k(C)$ has been obtained in [25] recently.

the same common difference, where $p = \min\{k, n-k\}$. We note that the study of $P_k(C)$ and $S_k(C)$ is related to the problem of completing a partial matrix of the form

$$\begin{pmatrix} D & ? \\ ? & ? \end{pmatrix} \qquad or \qquad \begin{pmatrix} ? & X \\ X^* & ? \end{pmatrix}$$

to a matrix in $\mathcal{U}(C)$.

In Section 4, we study the *C-numerical range* of $A = (A_1, \ldots, A_m) \in \mathcal{H}_n^m$ defined by

$$W_C(A) = \{(\operatorname{tr}(A_1 U^* C U), \ldots, \operatorname{tr}(A_m U^* C U)) : U \in \mathcal{U}_n\},$$

which can be viewed as the image of $\mathcal{U}(C)$ under the linear map $\phi : \mathcal{H}_n \to \mathbf{R}^m$ defined by

$$\phi(X) = (\operatorname{tr} A_1 X, \ldots, \operatorname{tr} A_m X).$$

If C is a rank one orthogonal projection, $W_C(A)$ reduces to the usual (joint) numerical range of $A = (A_1, \ldots, A_m)$, which has been studied extensively, see [1, 5, 6, 8, 10, 13, 19, 31] and their references. We will identify certain classes of $A = (A_1, \ldots, A_m)$ so that $W_C(A)$ is convex. It will be shown that $W_C(A)$ is convex if and only if $W_D(A) \subseteq W_C(A)$ whenever $D \in \operatorname{conv}\mathcal{U}(C)$. These extend the results of many other authors [1, 13, 15, 24, 26, 30, 37].

In Section 5, we discuss results for symmetric matrices analogous to those on complex Hermitian matrices in Sections 2 – 4. In particular, using a result of Fulton [14], we can transfer results on complex Hermitian matrices to real symmetric matrices.

In our discussion, we let $\{E_{11}, E_{12}, \ldots, E_{nn}\}$ be the standard basis for the algebra of $n \times n$ complex matrices. Denote by $\lambda(X)$ the vector of eigenvalues of $X \in \mathcal{H}_n$ with entries arranged in descending order. For $C \in \mathcal{H}_n$, $\operatorname{diag}(C) \in \mathbf{R}^n$ denotes the diagonal of C. If $c = (c_1, \cdots, c_n) \in \mathbf{R}^n$, then $[c]$ denotes the diagonal matrix with $c_1, \cdots, c_n$ along the diagonal. Given two vectors x and y in $\mathbf{R}^n$, we say that x is *weakly majorized* by y, denoted by $x \prec_w y$, if the sum of the k largest entries of x is not larger than that of y for $k = 1, \ldots, n$. If in addition that the sum of the entries of x is the same as that of y, we say that x is *majorized* by y, denoted by $x \prec y$; see [28] for the general background and basic properties of majorization. The following result is useful in our discussion.

Proposition 1.1 [17, 29, 30, 32, 34] *Let $C \in \mathcal{H}_n$ and $d_1, \ldots, d_n \in \mathbf{R}$. The following conditions are equivalent.*

(a) *There is a matrix in $\mathcal{U}(C)$ with diagonal entries $d_1, \ldots, d_n$.*

(b) *There is a matrix in $\operatorname{conv}\mathcal{U}(C)$ with diagonal entries $d_1, \ldots, d_n$.*

(c) *There is a matrix in $\operatorname{conv}\mathcal{U}(C)$ with eigenvalues $d_1, \ldots, d_n$.*

(d) $(d_1, \ldots, d_n) \prec \lambda(C)$.

Moreover, the set of vectors $(d_1, \ldots, d_n)$ satisfying any one (and hence all) of the above conditions is the convex hull of vectors of the form $(c_{i_1}, \ldots, c_{i_n})$, where $(i_1, \ldots, i_n)$ is a permutation of $(1, \ldots, n)$.

Proposition 1.1 has many applications in the study of inequalities and convexity problems. Note that if $C \in \mathcal{H}_n$ and $A_j = E_{jj}$ for $j \in \{1, \ldots, n\}$, then $(d_1, \ldots, d_n) \in W_C(A_1, \ldots, A_n)$ if and only if $d_1, \ldots, d_n$ are the diagonal entries of a matrix in $\mathcal{U}(C)$; so Proposition 1.1 gives a complete description of $W_C(A_1, \ldots, A_n)$. Also, Proposition 1.1 gives the answer to the problem of completing a partial matrix of the form

$$\begin{pmatrix} d_1 & ? & ? \\ ? & \ddots & ? \\ ? & ? & d_n \end{pmatrix}$$

to a matrix in $\mathcal{U}(C)$.

2 conv $\mathcal{U}(C)$ with the mid-point convex property

In this section, we characterize those matrices $C \in \mathcal{H}_n$ so that conv $\mathcal{U}(C)$ has the mid-point convex property, i.e., every matrix in conv $\mathcal{U}(C)$ can be written as the average of two matrices in $\mathcal{U}(C)$. It turns out that these are the same as those $C \in \mathcal{H}_n$ satisfying the much weaker condition that every matrix in conv $\mathcal{U}(C)$ can be written as a convex combination (not necessary the average) of two matrices in $\mathcal{U}(C)$, and they are the matrices of form (1).

Theorem 2.1 *The following conditions are equivalent for a given $C \in \mathcal{H}_n$:*

(a) $\mathrm{conv}\,\mathcal{U}(C) = \left\{\frac{1}{2}(C_1 + C_2) : C_1, C_2 \in \mathcal{U}(C)\right\}$.

(b) $\mathrm{conv}\,\mathcal{U}(C) = \{tC_1 + (1-t)C_2 : C_1, C_2 \in \mathcal{U}(C), 0 \le t \le 1\}$.

(c) *C has eigenvalues $a, a+b, \ldots, a+(n-1)b$ for some $a, b \in \mathbf{R}$, i.e., the eigenvalues of C form an arithmetic progression.*

To prove Theorem 2.1, we need some notations and a recent result concerning the eigenvalues of $X, Y, Z \in \mathcal{H}_n$ with $X = Y + Z$. Let

$$\mathcal{T}_1^n = \{((r), (s), (t)) : 1 \le r, s, t \le 1, \quad r + 1 = s + t\}.$$

For $m > 1$, let $\mathcal{T}_m^n$ to be the set of (R, S, T), where $R = (r_1, \ldots, r_m)$, $S = (s_1, \ldots, s_m)$ and $T = (t_1, \ldots, t_m)$ are three subsequences of $(1, \ldots, n)$ satisfying the following conditions:

(1) $\sum_{j=1}^m r_j + m(m+1)/2 = \sum_{j=1}^m (s_j + t_j)$.

(2) For all $1 \le k < m$ and $(U, V, W) \in \mathcal{T}_k^n$, where $U = (u_1, \ldots, u_k)$, $V = (v_1, \ldots, v_k)$ and $W = (w_1, \ldots, w_k)$ with $k < m$, we have

$$\sum_{j=1}^k r_{u_j} + k(k+1)/2 \ge \sum_{j=1}^k (s_{v_j} + t_{w_j}).$$

With these notations, we can state the following theorem, see [14, 20, 21] and also [7, 11, 16, 18, 38].

Eigenvalue Inequality Theorem *The necessary and sufficient condition for the real numbers $a_1 \geq \cdots \geq a_n$, $b_1 \geq \cdots \geq b_n$ and $c_1 \geq \cdots \geq c_n$ to be the eigenvalues of three matrices $X, Y, Z \in \mathcal{H}_n$ such that $X = Y + Z$ is:*

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j + \sum_{j=1}^{n} c_j$$

and for every $1 \leq m < n$, and $(R, S, T) \in \mathcal{T}_m^n$,

$$\sum_{r\in R} a_r \leq \sum_{s\in S} b_s + \sum_{t\in T} c_t.$$

The following lemma will be useful in the proof of Theorem 2.1 and Corollary 2.3

Lemma 2.2 *Let $b = (b_1, \ldots, b_n), c = (c_1, \ldots, c_n) \in \mathbf{R}^n$ with $b_1 \geq \cdots \geq b_n$ and $c_1 \geq \cdots \geq c_n$. If there exist $B_1, B_2 \in \mathcal{U}([b])$ and $C_1, C_2 \in \mathcal{U}([c])$ such that*

$$B_1 + C_1 = rI_n \quad \text{and} \quad B_2 + C_2 = [b_1 + c_1] \oplus s\, I_{n-1}$$

for some r, $s \in \mathbf{R}$, then the sequences $(b_1, \ldots, b_n)$ and $(c_1, \ldots, c_n)$ are arithmetic progressions with the same common difference.

Proof. Since $B_1 + C_1 = rI_n$, we have $b_i = r - c_{n+1-i}$ for $1 \leq i \leq n$, and hence

$$b_i - b_{i+1} = c_{n-i} - c_{n+1-i} \text{ for } 1 \leq i \leq n-1\,. \tag{2}$$

If $B_2 + C_2 = [b_1 + c_1] \oplus s\, I_{n-1}$, then (see [22, Lemma 4.1]) $B_2 = [b_1] \oplus B_3$ and $C_2 = [c_1] \oplus C_3$, where $\lambda(B_3) = (b_2, \ldots, b_n)$ and $\lambda(C_3) = (c_2, \ldots, c_n)$. Applying the above argument to $sI_{n-1} = B_3 + C_3$, we have

$$b_{i+1} - b_{i+2} = c_{n-i} - c_{n+1-i} \text{ for } 1 \leq i \leq n-2. \tag{3}$$

The result now follows from (2) and (3). □

Now, we are ready to present the

Proof of Theorem 2.1. Suppose $C \in \mathcal{H}_n$ has eigenvalues $c_1 \geq \cdots \geq c_n$.

$(a) \Rightarrow (b)$ follows from

$$\begin{aligned} & \left\{\frac{1}{2}(C_1 + C_2) : C_1,\, C_2 \in \mathcal{U}(C)\right\} \\ \subseteq\ & \{tC_1 + (1-t)C_2 : C_1,\, C_2 \in \mathcal{U}(C),\, 0 \leq t \leq 1\} \\ \subseteq\ & \operatorname{conv} \mathcal{U}(C). \end{aligned}$$

Now, suppose (b) holds. We may assume that $n \geq 3$ and $c_2 > c_n$. Let $r = (c_1 + \cdots + c_n)/n$ and $s = (c_2 + \cdots + c_n)/(n-1)$. Then rI_n, $[c_1] \oplus sI_{n-1} \in \operatorname{conv} \mathcal{U}(C)$. So, there exist $C_i \in \mathcal{U}(C)$ $i = 1, 2, 3, 4$ and $0 \leq t_j \leq 1$, $j = 1, 2$ such that

$$\begin{aligned} \text{i) } t_1 C_1 + (1-t_1) C_2 &= rI_n \\ \text{ii) } t_2 C_3 + (1-t_2) C_4 &= [c_1] \oplus sI_{n-1} = [t_2 c_1] \oplus (t_2 s) I_{n-1} + [(1-t_2)c_1] \oplus (1-t_2) s I_{n-1} \end{aligned}$$

Applying the argument in Lemma 2.2 to i), we have for $1 \leq i \leq n-1$

$$t_1(c_i - c_{i+1}) = (1-t_1)(c_{n-i} - c_{c+1-i}) \Rightarrow t_1(c_1 - c_n) = (1-t_1)(c_1 - c_n) \Rightarrow t_1 = 1/2\,.$$

Similarly, we have $t_2 = 1/2$. Thus $C_1 + C_2 = 2rI_n$ and $C_3 + C_4 = 2([c_1] \oplus sI_{n-1})$. Hence, (c) follows from Lemma 2.2.

Finally, suppose (c) holds. Replacing C by $\alpha C + \beta I$ for some suitable $\alpha, \beta \in \mathbf{R}$, we may assume that $c_j = -j$ for $1 \leq j \leq n$. Let $B \in \operatorname{conv} \mathcal{U}(C)$. We may assume that $B = \operatorname{diag}(b_1, \cdots, b_n)$ such that $b_1 \geq \cdots \geq b_n$ and $(b_1, \cdots, b_n) \prec (c_1, \cdots, c_n) = (-1, \cdots, -n)$. Let $D = \operatorname{diag}(d_1, \cdots, d_n)$ where $d_j = 2b_j$ for $1 \leq j \leq n$. Then $(d_1, \cdots, d_n) \prec (-2, \cdots, -2n)$. We are going to show that for $1 \leq k_1 < \cdots < k_m \leq n$ and $1 \leq m$, we have

$$d_{k_1} + \cdots + d_{k_m} \leq -(k_1 + \cdots + k_m + m(m+1)/2). \tag{4}$$

Suppose (4) holds. Then for any $1 \leq i_1 < \cdots < i_m \leq n$, $1 \leq j_1 < \cdots < j_m \leq n$ and $1 \leq k_1 < \cdots < k_m \leq n$ satisfying $i_1 + \cdots + i_m + j_1 + \cdots + j_m = k_1 + \cdots + k_m + m(m+1)/2$, we have

$$d_{k_1} + \cdots + d_{k_m} \leq c_{i_1} + \cdots + c_{i_m} + c_{j_1} + \cdots + c_{j_m}.$$

Also, we have $\sum_{i=1}^n d_i = -n(n+1) = 2\sum_{i=1}^n c_i$. By the Eigenvalue Inequality Theorem, there exist C_1, $C_2 \in \mathcal{U}(C)$ such that

$$2B = D = C_1 + C_2 \Rightarrow B = (C_1 + C_2)/2.$$

Given $1 \leq k_1 < \cdots < k_m \leq n$, define $k_0 = 0$. We are going to prove (4) by induction on $k_m - m$.

If $k_m - m = 0$, then $k_i = i$ for $1 \leq i \leq m$. Since $(d_1, \cdots, d_n) \prec (-2, \cdots, -2n)$, we have

$$\begin{aligned} & d_{k_1} + \cdots + d_{k_m} \\ = \;& d_1 + \cdots + d_m \\ \leq \;& -(2 + \cdots + 2m) \\ = \;& -m(m+1) \\ = \;& -(1 + \cdots + m + m(m+1)/2) \\ = \;& -(k_1 + \cdots + k_m + m(m+1)/2). \end{aligned}$$

Suppose (3) holds whenever $k_m - m < p$, where $p > 0$. Let $1 \leq k_1 < \cdots < k_m \leq n$ be a sequence such that $k_m - m = p$. Choose the largest i such that $k_{i-1} + 1 < k_i$. Let $r = m - i$. Then

$$k_{i+j} = j + k_i, \qquad j = 1, \ldots, r. \tag{5}$$

Note that the sequence $(k_1, \ldots, k_{i-1})$ has k_{i-1} terms and

$$k_{i-1} - i - 1 < k_i - i = k_m - m = p.$$

By induction assumption, we have

$$d_{k_1} + \cdots + d_{k_{i-1}} \leq -(k_1 + \cdots + k_{i-1} + i(i-1)/2)$$

Also, the sequence $(k_1, \ldots, k_{i-1}, k_i - 1, k_i, \ldots, k_m)$ has $(m+1)$ terms, and $k_m - (m+1) = p - 1$. By induction assumption, we have

$$\begin{aligned} & d_{k_1} + \cdots + d_{k_{i-1}} + d_{k_i - 1} + d_{k_i} + \cdots + d_{k_m} \\ \leq\ & -(k_1 + \cdots + k_{i-1} + (k_i - 1) + k_i + \cdots + k_m + (m+1)(m+2)/2). \end{aligned}$$

Hence,

$$\begin{aligned} & (r+2)(d_{k_1} + \cdots + d_{k_{i-1}} + d_{k_i} + \cdots + d_{k_m}) \\ \leq\ & (r+1)(d_{k_1} + \cdots + d_{k_{i-1}} + d_{k_i - 1} + d_{k_i} + \cdots + d_{k_m}) + (d_{k_1} + \cdots + d_{k_{i-1}}) \\ \leq\ & -(r+1)(k_1 + \cdots + k_{i-1} + (k_i - 1) + k_i + \cdots + k_m + (m+1)(m+2)/2) \\ & -(k_1 + \cdots + k_{i-1} + i(i-1)/2) \\ =\ & -(r+2)(k_1 + \cdots + k_m + m(m+1)/2) - (r+1)[(k_i - 1) + (m+1)] \\ & (i + \ldots + m) + (k_i + \cdots + k_m) \\ =\ & -(r+2)(k_1 + \cdots + k_m + m(m+1)/2), \end{aligned}$$

where the last equality follows from (5). The proof is complete. □

By private communication, O. Azenhas informed us that part of the proof of Theorem 2.1 can be done using the idea in [2] and some combinatorial arguments involving the manipulation of Young diagrams.

By Theorem 2.1 and the Eigenvalue Inequality Theorem, we have the following corollary.

Corollary 2.3 *Let $b = (b_1, \ldots, b_n), c = (c_1, \ldots, c_n) \in \mathbf{R}^n$ with $b_1 \geq \cdots \geq b_n$ and $c_1 \geq \cdots \geq c_n$. The following conditions are equivalent.*

(a) *The sequences $(b_1, \ldots, b_n)$ and $(c_1, \ldots, c_n)$ are arithmetic progressions with the same common difference.*

(b) *$\mathcal{U}([b]) + \mathcal{U}([c]) = \{B + C : B \in \mathcal{U}([b]),\ C \in \mathcal{U}([c])\}$ is convex.*

(c) *Whenever $A \in \mathcal{H}_n$ satisfies $\lambda(A) \prec b + c$, there exist $B \in \mathcal{U}([b])$, $C \in \mathcal{U}([c])$ such that $A = B + C$.*

Proof. Let $\mathcal{A} = \{A \in \mathcal{H}_n : \lambda(A) \prec b + c\}$. Suppose $B \in \mathcal{U}([b])$, $C \in \mathcal{U}([c])$. Then there exists $U \in \mathcal{U}_n$ such that $U(B+C)U^* = [\lambda(B+C)]$. So, by Proposition 1.1, we have

$$\lambda(B+C) = \operatorname{diag} UBU^* + \operatorname{diag} UCU^* \prec b + c\,.$$

Hence, one can use the fact that

$$\mathcal{U}([b]) + \mathcal{U}([c]) \subseteq \mathcal{A} = \operatorname{conv}(\mathcal{U}([b+c])) \subseteq \operatorname{conv}(\mathcal{U}([b]) + \mathcal{U}([c]))$$

to conclude that

$$\mathcal{U}([b]) + \mathcal{U}([c]) \subseteq \mathcal{A} = \operatorname{conv}(\mathcal{U}([b]) + \mathcal{U}([c]))\,.$$

Therefore, (b) and (c) are equivalent.

Let $v = (1, \ldots, 1) \in \mathbf{R}^n$. If (a) holds, then the sequence $a = b - b_n v = c - c_n v$ is in arithmetic progression. By Theorem 2.1, we have

$$\begin{aligned} \mathcal{U}([b]) + \mathcal{U}([c]) = &\ \mathcal{U}([a + b_n v]) + \mathcal{U}([a + c_n v]) \\ = &\ (\mathcal{U}([a]) + \mathcal{U}([a])) + (b_n + c_n)I_n \\ = &\ 2\left[\frac{1}{2}(\mathcal{U}([a]) + \mathcal{U}([a]))\right] + (b_n + c_n)I_n \\ = &\ 2\operatorname{conv}(\mathcal{U}([a])) + (b_n + c_n)I_n \end{aligned}$$

is convex. This proves (a) $\Rightarrow$ (b).

Conversely, suppose (b) holds. Let

$$r = \frac{(b_1 + c_1 + \cdots + b_n + c_n)}{n}, \quad s = \frac{(b_2 + \cdots + b_n)}{(n-1)} \text{ and } t = \frac{(c_2 + \cdots + c_n)}{(n-1)}\,.$$

Then rI_n, $([b_1] \oplus sI_{n-1} + [c_1] \oplus tI_{n-1}) \in \operatorname{conv}(\mathcal{U}([b]) + \mathcal{U}([c])) = \mathcal{U}([b]) + \mathcal{U}([c])$ and (a) follows from Lemma 2.2. □

3 Submatrices of matrices in $\mathcal{U}(C)$

We begin with the study of the set $P_k(C)$ of $k \times k$ (leading) principal submatrices of matrices in $\mathcal{U}(C)$.

Theorem 3.1 *Let $1 \le k \le n$. Suppose $C \in \mathcal{H}_n$ have eigenvalues $c_1 \ge \cdots \ge c_n$, and $P_k(C)$ be the set of $k \times k$ (leading) principal submatrices of matrices in $\mathcal{U}(C)$. Let $D \in \mathcal{H}_k$ have eigenvalues $d_1 \ge \cdots \ge d_k$. Then*

(a) *$D \in P_k(C)$ if and only if*

$$c_{n-k+j} \le d_j \le c_j, \qquad 1 \le j \le k;$$

(b) *$D \in \operatorname{conv} P_k(C)$ if and only if there is a matrix $\tilde{D} \in \operatorname{conv} P_k(C)$ with diagonal entries $d_1, \ldots, d_k$, equivalently,*

$$c_n + \cdots + c_{n-m+1} \le d_{i_1} + \cdots + d_{i_m} \le c_1 + \cdots + c_m \tag{6}$$

whenever $1 \le m \le k$ and $1 \le i_1 < \cdots < i_m \le k$;

(c) *D is an extreme point of* $\operatorname{conv} P_k(C)$ *if and only if*

$$(d_1, \ldots, d_k) = (c_1, \ldots, c_p, c_{n-q+1}, \ldots, c_n)$$

for some $0 \le p, q \le k$ *such that* $p + q = k$.

Consequently, the set of $k \times k$ *principal submatrices of matrices in* $\mathcal{U}(C)$ *is convex if and only if* $c_1 = c_k$ *and* $c_{n-k+1} = c_n$*; in such case,*

$$\operatorname{conv} P_k(C) = \{P \in \mathcal{H}_k : P \text{ has eigenvalues in the interval } [c_n, c_1]\}.$$

Proof. Part (a) follows from the result in [12].

To prove (b), note that by Proposition 1.1, $D \in \operatorname{conv} P_k(C)$ has eigenvalues $d_1, \ldots, d_k$ if and only if there is a matrix $\tilde{D} \in \operatorname{conv} P_k(C)$ with diagonal entries $d_1, \ldots, d_k$. By the arguments in the proof of [4, Theorem 2.4], these equivalent conditions hold if and only if (6) holds whenever $1 \le m \le k$ and $1 \le i_1 < \cdots < i_m \le k$.

To prove (c), note that if (6) holds whenever $1 \le m \le k$ and $1 \le i_1 < \cdots < i_m \le k$, then for $d_{k+1} = \cdots = d_n = [(\sum_{j=1}^n c_j) - (\sum_{j=1}^k d_j)]/(n-k)$ we have

$$(d_1, \ldots, d_n) \prec (c_1, \ldots, c_n),$$

see the proof of [4, Theorem 2.4]. By Proposition 1.1, $(d_1, \ldots, d_n)$ is a convex combination of the vectors of the form $(c_{i_1}, \ldots, c_{i_n})$, where $(i_1, \ldots, i_n)$ is a permutation of $(1, \ldots, n)$. Hence, the $k \times k$ diagonal matrix $[(d_1, \ldots, d_k)]$ is a convex combination of the diagonal matrices of the form $[(c_{i_1}, \ldots, c_{i_k})]$. Thus, every matrix $D \in \mathcal{H}_k$ with eigenvalues $d_1, \ldots, d_k$ is a convex combination of matrices in $\mathcal{H}_k$ with eigenvalues $c_{i_1}, \ldots, c_{i_k}$. Hence, if C_0 is an extreme point of $\operatorname{conv} P_k(C)$, then C_0 has eigenvalues $c_{i_1}, \ldots, c_{i_k}$, where $\{i_1, \ldots, i_k\}$ is a k-element subset of $\{1, \ldots, n\}$. To complete our proof, we show that $\{c_{i_1}, \ldots, c_{i_k}\}$ must be of the form $\{c_1, \ldots, c_p, c_{n-q+1}, \ldots, c_n\}$ for some $p, q \ge 0$ with $p + q = n$. If it is not true and if $i_1, \ldots, i_k$ are rearranged so that $1 \le i_1 < \cdots < i_k \le n$, then there exists r such that $c_r > c_{i_r}$ and $(c_{i_r}, \ldots, c_{i_k}) \ne (c_{n-k+r}, \ldots, c_n)$. Thus $r \notin \{i_1, \ldots, i_k\}$ and there exists $s > i_r$ so that $s \notin \{i_r, \ldots, i_k\}$ and $c_{i_r} > c_s$. But then we may obtain C_1 (respectively, C_2) from C_0 by replacing the diagonal entry c_{i_r} by c_r (respectively, c_s) so that C_0 is a convex combination of the matrices C_1 and C_2, contradicting the fact that C_0 is an extreme point. □

Next, we turn to the set

$$S_k(C) = \left\{X \in \mathbf{C}^{k\times(n-k)} : \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathcal{U}(C) \text{ for some } P \in \mathcal{H}_k,\ Q \in \mathcal{H}_{n-k}\right\}$$

for a given $C \in \mathcal{H}_n$. Since $\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix}$ is permutationally similar to $\begin{pmatrix} Q & X^* \\ X & P \end{pmatrix}$, we see that

$$S_k(C) = \{X^* : X \in S_{n-k}(C)\}.$$

Thus, we can focus on the cases for $1 \le k \le n/2$. This assumption and the notation $S_k(C)$ will be used throughout this section, and the vector of singular values of $X \in S_k(C)$ will be denoted by

$$s(X) = (s_1, \ldots, s_k), \quad \text{with } s_1 \ge \cdots \ge s_k.$$

Proposition 3.2 *Suppose $C \in \mathcal{H}_n$ has eigenvalues $c_1 \geq \cdots \geq c_n$, and $1 \leq k \leq n/2$.*

(a) *$X \in S_k(C)$ if and only if $UXV \in S_k(C)$ for any unitary matrices U and V of sizes k and $(n-k)$ respectively.*

(b) *$X \in \operatorname{conv} S_k(C)$ if and only if*

$$s(X) \prec_w (c_1 - c_n, c_2 - c_{n-1}, \ldots, c_k - c_{n-k+1})/2. \tag{7}$$

(c) *X is an extreme point of $\operatorname{conv} S_k(C)$ if and only if*

$$s(X) = (c_1 - c_n, \ldots, c_k - c_{n-k+1})/2. \tag{8}$$

Proof. (a) Suppose X is the $(1,2)$ block of $C_0 \in \mathcal{U}(C)$. Then UXV is the off-diagonal block of $C_1 = (U \oplus V^*)C_0(U^* \oplus V) \in \mathcal{U}(C)$. The converse can be proved similarly.

To prove (b) and (c), let T be the set of $k \times (n-k)$ matrices X satisfying (8). By the result in [34], $\operatorname{conv} T$ is the collection of $k \times (n-k)$ matrices X satisfying (7). Thus, condition (b) can be restated as $\operatorname{conv} S_k(C) = \operatorname{conv} T$.

We claim that $T \subseteq S_k(C)$. To see this, let $P = \operatorname{diag}(c_1 + c_n, c_2 + c_{n-1}, \ldots, c_k + c_{n-k+1})/2$, $Q = P \oplus \operatorname{diag}(c_{k+1}, \ldots, c_{n-k})$, and X be the $k \times (n-k)$ matrix with (j,j) entry equal to $(c_j - c_{n-j+1})/2$ for $j = 1, \ldots, k$. Then

$$C_0 = \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathcal{U}(C).$$

By condition (a), we see that for every unitary matrix U and V of sizes k and $(n-k)$ respectively, UXV belongs to $S_k(C)$. Thus, $T \subseteq S_k(C)$.

By the result in [23], we have $S_k(C) \subseteq \operatorname{conv} T$. Hence,

$$T \subseteq S_k(C) \subseteq \operatorname{conv} T,$$

and T contains all the extreme points of $\operatorname{conv} S_k(C)$. Since every elements in T has the same Frobenius norm, none of $X \in T$ is the convex combination of other elements in T. So, T is the set of extreme points of $\operatorname{conv} T$. □

In general, it is difficult to give a complete description of $S_k(C)$ except for special C or special values of k. For example, if $k = 1$, then $X \in S_1(C)$ if and only if $s_1(X) \leq (\lambda_1(C) - \lambda_n(C))/2$; if C is a rank p orthogonal projection, then $X \in S_k(C)$ if and only if $s_1(X) \leq 1/2$ and $\operatorname{rank}(X) \leq p$. Nonetheless, we have the following characterization of $C \in \mathcal{H}_n$ for which $S_k(C)$ is convex.

Theorem 3.3 *Suppose $C = \operatorname{diag}(c_1, \cdots, c_n)$ with $c_1 \geq \cdots \geq c_n$, and $1 \leq k \leq n/2$. The following conditions are equivalent.*

(a) *$S_k(C)$ is convex, i.e.,*

$$S_k(C) = \{X \in \mathbf{C}^{k \times (n-k)} : s(X) \prec_w (c_1 - c_n, \ldots, c_k - c_{n-k+1})/2\}.$$

(b) *The sequences* $(c_1, \cdots, c_k)$ *and* $(c_{n-k+1}, \cdots, c_n)$ *are arithmetic progressions with the same common difference.*

The proof of Theorem 3.3 is divided into several lemmas. In the rest of this section, we always assume that k, C and $S_k(C) = S$ (for simplicity) satisfy the hypotheses of Theorem 3.3.

The next two lemmas deal with the implication (a) $\Rightarrow$ (b) in Theorem 3.3.

Lemma 3.4 *Let h be a positive integer such that $h \le n/2$, and let $s = \sum_{i=1}^{h}(c_i - c_{n-i+1})/(2h)$. Suppose there is an $X = (x_{pq}) \in S$ such that $x_{11} = \cdots = x_{hh} = s$ and $x_{pq} = 0$ for all other (p, q). Then*

$$c_1 - c_{n-h+1} = c_2 - c_{n-h+2} = \cdots = c_{h-1} - c_{n-1} = c_h - c_n = 2s.$$

Proof. Suppose $C_0 = \begin{pmatrix} P_0 & X \\ X^* & Q_0 \end{pmatrix} \in \mathcal{U}(C)$, where $X = (x_{pq}) \in S$ satisfies the hypothesis of the lemma. Then the $2h \times 2h$ principal submatrix of C_0 lying in the rows and columns with indices $1, \ldots, h, k+1, \ldots, k+h$, is of the form $\begin{pmatrix} P & sI_h \\ sI_h & Q \end{pmatrix}$. Since

$$\begin{aligned} \tilde{C}_0 &= \frac{1}{\sqrt{2}} \begin{pmatrix} I_h & I_h \\ I_h & -I_h \end{pmatrix} \begin{pmatrix} P & sI_h \\ sI_h & Q \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} I_h & I_h \\ I_h & -I_h \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} P + Q + 2sI & P - Q \\ P - Q & P + Q - 2sI \end{pmatrix}, \end{aligned}$$

the matrix C_0 is unitarily similar to a matrix of the form $\begin{pmatrix} \tilde{C}_0 & * \\ * & * \end{pmatrix} \in \mathcal{U}(C)$. Suppose $(P+Q)/2$ has eigenvalues $\gamma_1 \ge \cdots \ge \gamma_h$. Then $(P + Q + 2sI)/2$ and $(P + Q - 2sI)/2$ have eigenvalues $\gamma_1 + s \ge \cdots \ge \gamma_h + s$ and $\gamma_1 - s \ge \cdots \ge \gamma_h - s$, respectively. Since $(P+Q+2sI)/2$ and $(P + Q - 2sI)/2$ are principal submatrices of a matrix in $\mathcal{U}(C)$, we have

$$c_{n-j+1} \le \gamma_{h-j+1} - s \quad \text{and} \quad \gamma_j + s \le c_j, \quad j = 1, \ldots, h, \tag{9}$$

by Theorem 3.1. Thus,

$$\sum_{j=1}^{h}(c_j - c_{n-j+1}) = 2hs = \sum_{j=1}^{h}(\gamma_j + s) - \sum_{j=1}^{h}(\gamma_j - s) \le \sum_{j=1}^{h}(c_j - c_{n-j+1}).$$

It follows that the inequalities in (9) are equalities. Hence $(c_1, \ldots, c_h) = (\gamma_1 + s, \ldots, \gamma_h + s)$, $(c_{n-h+1}, \ldots, c_n) = (\gamma_1 - s, \ldots, \gamma_h - s)$, and the result follows. □

Lemma 3.5 *The implication* (a) $\Rightarrow$ (b) *in Theorem 3.3 holds.*

Proof. Suppose Condition (a) of Theorem 3.3 holds. Condition (b) is trivial if $k = 1$. We therefore assume that $k \geq 2$. Let $h = k-1$ or k. Set $s_j = (c_j - c_{n-j+1})/2$ for $1 \leq j \leq h$, and $s = (\sum_{j=1}^h s_j)/h$. By Condition (a) of Theorem 3.3, the $k \times (n-k)$ matrix X with (p,q) entry equal to

$$x_{pq} = \begin{cases} s & \text{if } 1 \leq p = q \leq h, \\ 0 & \text{otherwise}, \end{cases}$$

belongs to S. By Lemma 3.4, we have

$$c_1 - c_{n-h+1} = c_2 - c_{n-h+2} = \cdots = c_{h-1} - c_{n-1} = c_h - c_n.$$

For $h = k$, we have

$$c_j - c_{j+1} = c_{n-k+j} - c_{n-k+j+1}, \quad j = 1, \ldots, k-1. \tag{10}$$

For $h = k-1$, we have

$$c_j - c_{j+1} = c_{n-k+j+1} - c_{n-k+j+2}, \quad j = 1, \ldots, k-2. \tag{11}$$

As a result,

$$c_{n-k+j} - c_{n-k+j+1} = c_{n-k+j+1} - c_{n-k+j+2}, \quad j = 1, \ldots, k-2,$$

and hence $c_{n-k+1}, \ldots, c_n$ is in arithmetic progression. By (10), we see that $c_1, \ldots, c_k$ is also in arithmetic progression with the same common difference. □

The next three lemmas deal with the implication (b) $\Rightarrow$ (a) of Theorem 3.3. We first use Theorem 2.1 to prove a special case of the implication in the following lemma.

Lemma 3.6 *Suppose $n = 2k$, and Condition* (b) *of Theorem 3.3 holds. If $t \in [0,1]$ and $D \in \mathcal{H}_n$ with $\lambda(D) \prec t(c_1 - c_{2k}, \cdots, c_k - c_{k+1})/2$, then $D \in S$.*

Proof. Without loss of generality, we may assume that $c_1 \geq \cdots \geq c_k \geq 0$ and $c_{2k-i+1} = -c_i$ for $1 \leq i \leq k$. Then $(c_1 - c_{2k}, \cdots, c_k - c_{k+1})/2 = (c_1, \cdots, c_k)$. Suppose $\lambda(D) = (d_1, \cdots, d_k) \prec t(c_1, \cdots, c_k)$. Let $A = \text{diag}\,(c_1, \cdots, c_k)$. Suppose $0 \leq t \leq 1$. By Proposition 1.1, $D \in \text{conv}\,\mathcal{U}(tA)$. Since $(c_1, \cdots, c_k)$ is an arithmetic progression, by Theorem 2.1, there exists $A_1,\ A_2 \in \mathcal{U}(A)$ such that $D = t(A_1 + A_2)/2$. Let $C = \text{diag}\,(c_1, \cdots, c_{2k})$. Then $A_1 \oplus (-A_2) \in \mathcal{U}(C)$. We have

$$\frac{1}{\sqrt{2}}\begin{pmatrix} I_k & I_k \\ I_k & -I_k \end{pmatrix}\begin{pmatrix} A_1 & 0_k \\ 0_k & -A_2 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} I_k & I_k \\ I_k & -I_k \end{pmatrix} = \frac{1}{2}\begin{pmatrix} A_1 - A_2 & A_1 + A_2 \\ A_1 + A_2 & A_1 - A_2 \end{pmatrix} \in \mathcal{U}(C).$$

Choose $0 \leq \theta \leq \pi/4$ such that $\cos 2\theta = t$. Then

$$\begin{aligned} & \frac{1}{2}\begin{pmatrix} \cos\theta I_k & \sin\theta I_k \\ -\sin\theta I_k & \cos\theta I_k \end{pmatrix}\begin{pmatrix} A_1 - A_2 & A_1 + A_2 \\ A_1 + A_2 & A_1 - A_2 \end{pmatrix}\begin{pmatrix} \cos\theta I_k & -\sin\theta I_k \\ \sin\theta I_k & \cos\theta I_k \end{pmatrix} \\ = & \begin{pmatrix} (A_1 - A_2)/2 + \sin 2\theta (A_1 + A_2)/2 & D \\ D & (A_1 - A_2)/2 - \sin 2\theta (A_1 + A_2)/2 \end{pmatrix} \in \mathcal{U}(C). \end{aligned}$$

Hence, $D \in S$. □

Lemma 3.7 *Suppose* $c_1 \geq \cdots \geq c_k \geq 0$, $d_1 \geq \cdots \geq d_k \geq 0$ *and* $\mathbf{d} = (d_1, \cdots, d_k) \prec_w (c_1, \cdots, c_k) = \mathbf{c}$. *Then we can write* $\mathbf{d} = (\mathbf{d_1}, \cdots, \mathbf{d_r})$, $\mathbf{c} = (\mathbf{c_1}, \cdots, \mathbf{c_r})$, *with* $\mathbf{d_i}, \mathbf{c_i} \in \mathbf{R}^{k_i}$, $\sum_{i=1}^r k_i = k$ *and* $\mathbf{d_i} \prec t_i \mathbf{c_i}$ *for some* $t_i \in [0,1]$, $1 \leq i \leq r$.

Proof. We prove the lemma by induction on k. The result clearly holds for $k = 1$. Since $(d_1, \cdots, d_k) \prec_w (c_1, \cdots, c_k)$, we have

$$\sum_{i=1}^{j} d_i \leq \sum_{i=1}^{j} c_i \quad \text{for } 1 \leq j \leq k. \tag{12}$$

If the equality sign holds for some $1 \leq j \leq k$, then the result follows. So, we may assume that

$$\sum_{i=1}^{j} d_i < \sum_{i=1}^{j} c_i \quad \text{for } 1 \leq j \leq k.$$

So for some $0 \leq t < 1$, we have

$$\sum_{i=1}^{j} d_i \leq \sum_{i=1}^{j} t c_i \quad \text{for } 1 \leq j \leq k\,,$$

with equality holds for at least one j and the result follows. □

We complete the proof of Theorem 3.3 with the following lemma.

Lemma 3.8 *The implication* (b) $\Rightarrow$ (a) *in Theorem 3.3 holds.*

Proof. Suppose Condition (b) of Theorem 3.3 holds. By the result in [23], if $X \in S$, then (7) holds. We have to prove the converse.

First consider the case when $n = 2k$. Without loss of generality, we may assume that $c_1 \geq \cdots \geq c_k \geq 0$ and $c_{2k-j+1} = -c_j$ for $1 \leq j \leq k$. Then $(c_1 - c_{2k}, \cdots, c_k - c_{k+1})/2 = (c_1, \cdots, c_k)$. By Proposition 3.2 (a), we may assume that $X = \operatorname{diag}(d_1, \cdots, d_k)$, where $s(X) = (d_1, \ldots, d_k)$. By Lemma 3.7, we can write $\mathbf{d} = (\mathbf{d_1}, \cdots, \mathbf{d_r})$, $\mathbf{c} = (\mathbf{c_1}, \cdots, \mathbf{c_r})$, with $\mathbf{d_i}, \mathbf{c_i} \in \mathbf{R}^{k_i}$, $\sum_{i=1}^r k_i = k$ and $\mathbf{d_i} \prec t_i \mathbf{c_i}$ for some $0 \leq t_i \leq 1$, $1 \leq i \leq r$. Let $D_i = \operatorname{diag}(\mathbf{d_i})$. By Lemma 3.6, for each $1 \leq i \leq r$ we have $X_i, Y_i \in \mathcal{H}_{k_i}$ such that

$$\begin{pmatrix} X_1 & D_1 \\ D_1 & Y_1 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} X_r & D_r \\ D_r & Y_r \end{pmatrix}$$

has eigenvalues $c_1, \cdots, c_{2k}$.

Finally, suppose $n > 2k$. Again, by Proposition 3.2 (a), we may assume that X is an $k \times (n-k)$ matrix with (j,j) entry equal to d_j for $j = 1, \ldots, k$, and all other entries equal to zero. One can then use the construction in the previous case to obtain a matrix

$$\tilde{C} = \begin{pmatrix} P & \tilde{X} \\ \tilde{X}^* & Q \end{pmatrix} \in \mathcal{H}_{2k},$$

where $\tilde{X} = \operatorname{diag}(d_1, \ldots, d_k)$, with eigenvalues $c_1 \geq \cdots \geq c_k$ and $c_{n-k} \geq \cdots \geq c_n$. Then the matrix $\tilde{C} \oplus \operatorname{diag}(c_{k+1}, \ldots, c_{n-k}) \in \mathcal{U}(C)$ will have X as the off-diagonal block. □

4 The C-numerical Ranges

In this section, we use the results in the previous sections to study the convexity of the C-numerical range of $A = (A_1, \dots, A_m) \in \mathcal{H}_n^m$ and some related problems. We begin with the following observation, which is easy to verify.

Proposition 4.1 *Let $C \in \mathcal{H}_n$ and $A = (A_1, \dots, A_m) \in \mathcal{H}_n^m$.*

(a) $W_C(A) = W_C(V^*A_1V, \dots, V^*A_mV)$ *for any unitary V.*

(b) $W_C(A)$ *is convex if and only if* $W_C(I, A_1, \dots, A_m)$ *is convex.*

(c) *Suppose $W_C(A)$ is convex and $B_j \in \operatorname{span}\{I, A_1, \dots, A_m\}$ for $1 \le j \le s$. Then $W_C(B_1, \dots, B_s)$ is convex.*

In general, for $n \ge 3$, the convexity of $W_C(A)$ for arbitrary $C \in \mathcal{H}_n$ and $A \in \mathcal{H}_n^m$ is only guaranteed when $m = 3$, as shown in [1]. In particular, there exists $A \in \mathcal{H}_n^4$ such that $W_C(A)$ is not convex if C is a rank one orthogonal projection. In [24], it was shown that if C is a rank k orthogonal projection, then for any $A_1, A_2, A_3 \in \mathcal{H}_n$ such that $\{I, A_1, A_2, A_3\}$ is linearly independent, one can always find an $A_4 \in \mathcal{H}_n$ so that $W_C(A_1, \dots, A_4)$ is not convex. We can extend the result to general $C \in \mathcal{H}_n$. Note that for $n = 2$, it is known (see e.g. [26]) that $W_C(A)$ is convex for $A = (A_1, \dots, A_m) \in \mathcal{H}_n^m$ if and only if the linear span of $\{I, A_1, \dots, A_m\}$ has dimension not larger than 3. We shall focus on the case when $n \ge 3$.

Theorem 4.2 *Let $n \ge 3$ and let $C \in \mathcal{H}_n$ be nonscalar. If $A_1, A_2, A_3 \in \mathcal{H}_n$ are such that $\{I, A_1, A_2, A_3\}$ is linearly independent, then there exists A_4 such that $W_C(A_1, \dots, A_4)$ is not convex.*

Proof. Suppose $C = \operatorname{diag}(c_1, \dots, c_n)$ with $c_1 \ge \dots \ge c_n$. Let k be the largest integer such that $c_k > c_{k+1}$. Since we assume that C is nonscalar, we have $k < n$. Furthermore, since $W_C(A)$ is convex if and only if $W_D(A)$ if convex for $D = C - c_{k+1}I$, we may assume that $c_{k+1} = \dots = c_n = 0$.

Suppose $A_1, A_2, A_3 \in \mathcal{H}_n$ are such that $\{I, A_1, A_2, A_3\}$ is linearly independent. By [24, Theorem 4.1], there exists $X \in \mathbf{C}^{n\times 2}$ such that $X^*X = I_2$ and $\{I_2, X^*A_1X, X^*A_2X, X^*A_3X\}$ is a basis for $\mathcal{H}_2$. Let U be a unitary matrix so that the columns of X are the kth and $(k+1)$st columns of U. Define $A_4 = U(\operatorname{diag}(k, k-1, \dots, 2) \oplus I_2 \oplus 0_{n-k-1})U^*$. We are going to prove that $W_C(A_1, \dots, A_4)$ is not convex.

By Proposition 4.1, $W(A_1, \dots, A_4) = W(U^*A_1U, \dots, U^*A_4U)$. Let $A[k, k+1]$ denote the 2×2 principal submatrix of $A \in \mathcal{H}_n$ lying in the kth and $(k+1)$st rows and columns. Then $\{U^*A_jU[k, k+1] : 1 \le j \le 4\}$ is a basis for $\mathcal{H}_2$ by our construction. Taking a suitable linear combinations of the matrices $U^*A_1U, \dots, U^*A_4U$, we can get $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3 \in \mathcal{H}_n$ so that $\tilde{A}_j[k, k+1] = B_j$ for $j = 1, 2, 3$, where

$$B_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ B_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$

and $\tilde{A}_4 = U^*A_4U = \operatorname{diag}(k, k-1, \ldots, 2) \oplus I_2 \oplus 0_{n-k-1}$. Let $\tilde{A} = (\tilde{A}_1, \ldots, \tilde{A}_4)$. We claim that the set $W(\tilde{A}) \cap \{(a_1, a_2, a_3, \alpha) : a_1, a_2, a_3 \in \mathbf{R}\}$ is not convex, where

$$\alpha = kc_1 + (k-1)c_2 + \cdots + 2c_{k-1} + c_k.$$

If this claim is proved, then we see that $W(\tilde{A})$ is not convex and neither is $W(A_1, \ldots, A_4)$ by Proposition 4.1.

To prove our claim, recall that $C = \operatorname{diag}(c_1, \ldots, c_n)$ with $c_1 \geq \cdots \geq c_k > 0 = c_{k+1} = \cdots = c_n$ by our assumption. Suppose V is unitary so that $\operatorname{tr} CV^*\tilde{A}_4V = \alpha$. Let $d_1, \ldots, d_{k+1}$ be the first $k+1$ diagonal entries of VCV^*. Since $(d_1, \ldots, d_{k+1}) \prec_w (c_1, \ldots, c_{k+1})$, we have

$$\begin{aligned} \alpha &= kd_1 + (k-1)d_2 + \cdots + 2d_{k-1} + d_k + d_{k+1} \\ &= d_1 + (d_1 + d_2) + \cdots + \left(\sum_{j=1}^{k-1} d_j\right) + \left(\sum_{j=1}^{k+1} d_j\right) \\ &\leq c_1 + (c_1 + c_2) + \cdots + \left(\sum_{j=1}^{k-1} c_j\right) + \left(\sum_{j=1}^{k+1} c_j\right) \\ &= \alpha. \end{aligned}$$

Thus, $\sum_{j=1}^m d_j = \sum_{j=1}^m c_j$ for $m = 1, \ldots, k-1$, and for $m = k+1$. By [22, Lemma 4.1], $VCV^* = \operatorname{diag}(c_1, \ldots, c_{k-1}) \oplus c_kP \oplus 0_{n-k-1}$ for a rank one orthogonal projection $P \in \mathcal{H}_2$. Hence, for $j = 1, 2, 3$, if $\tilde{A}_j = (a_{pq}^{(j)})$ and $\alpha_j = \sum_{s=1}^{k-1} c_s a_{ss}^{(j)}$, then

$$\operatorname{tr} CV^*\tilde{A}_jV = \operatorname{tr} VCV^*\tilde{A}_j = \alpha_j + c_k \operatorname{tr} PB_j.$$

It follows that

$$\begin{aligned} & W_C(\tilde{A}) \cap \{(a_1, a_2, a_3, \alpha) : a_1, a_2, a_3 \in \mathbf{R}\} \\ = \; & \{(\alpha_1, \alpha_2, \alpha_3, \alpha) + c_k(\operatorname{tr} PB_1, \operatorname{tr} PB_2, \operatorname{tr} PB_3, 0) : P = vv^*, \ v \in \mathbf{C}^2, \ v^*v = 1\} \\ = \; & \{(\alpha_1, \alpha_2, \alpha_3, \alpha) + c_k(v^*B_1v, v^*B_2v, v^*B_3v, 0) : v \in \mathbf{C}^2, \ v^*v = 1\} \\ = \; & \{(\alpha_1, \alpha_2, \alpha_3, \alpha) + c_k(b_1, b_2, b_3, 0) : b_1^2 + b_2^2 + b_3^2 = 1\}, \end{aligned}$$

which is not convex as desired. □

Despite the negative result in Theorem 4.2, there has been considerable interest in studying those $A \in \mathcal{H}_n^m$ so that $W_C(A)$ is convex for a given $C \in \mathcal{H}_n$. In fact, the convexity of $W_C(A)$ is equivalent to several other properties of $A \in \mathcal{H}_n^m$ pursued by researchers in other areas (see [1, 13, 15, 24, 26] and their references). We summarize them in the following proposition.

Proposition 4.3 *Let $C \in \mathcal{H}_n$ have eigenvalues $c_1 \geq \cdots \geq c_n$. Suppose $A = (A_1, \ldots, A_m) \in \mathcal{H}_n^m$. The following conditions are equivalent.*

(a) *The set $W_C(A)$ is convex.*

(b) *If* $\phi : \mathcal{H}_n \to \mathbf{R}^m$ *is defined by* $\phi(X) = (\operatorname{tr} A_1X, \ldots, \operatorname{tr} A_mX)$ *then* $\phi(\mathcal{U}(C))$ *is convex.*

(c) *For any* $D \in \mathcal{H}_n$ *with* $\lambda(D) \prec \lambda(C)$*, the inclusion relation* $W_D(A) \subseteq W_C(A)$ *holds.*

(d) *For any* $D \in \mathcal{H}_n$ *with* $\lambda(D) \prec \lambda(C)$*, there exists* $C_0 \in \mathcal{U}(C)$ *such that* $\operatorname{tr} C_0A_j = \operatorname{tr} DA_j$ *for all* $j = 1, \ldots, m$*.*

Now, we can apply the results in the previous sections to get information about the convexity of $W_C(A)$, and hence for other equivalent conditions in Proposition 4.3. First of all, using Theorem 3.3, we have the following.

Proposition 4.4 *Suppose* $1 \le k \le n/2$*. Let* $C \in \mathcal{H}_n$ *have eigenvalues* $c_1 \ge \cdots \ge c_n$ *be such that the sequences* $(c_1, \cdots, c_k)$ *and* $(c_{n-k+1}, \cdots, c_n)$ *are arithmetic progressions with the same common difference. Suppose* $A = (A_1, \ldots, A_m) \in \mathcal{H}_n^m$ *is such that*

$$A_j = \alpha_j I + \begin{pmatrix} 0_k & Y_j \\ Y_j^* & 0_{n-k} \end{pmatrix}, \quad j = 1, \ldots, m.$$

Then Conditions (a) – (d) *of Proposition 4.3 hold.*

Proof. It suffices to prove one of the conditions in Proposition 4.3 holds. We consider Condition (b) and define the linear transformation $\phi : \mathcal{H}_n \to \mathbf{R}^m$ by

$$\phi(Z) = (\operatorname{tr} A_1Z, \ldots, \operatorname{tr} A_mZ).$$

Then $\phi(\mathcal{U}(C)) = \phi(\tilde{S}) + (\operatorname{tr} C)(\alpha_1, \ldots, \alpha_m)$ with

$$\tilde{S} = \left\{ \begin{pmatrix} 0_k & X \\ X^* & 0_{n-k} \end{pmatrix} : X \in S_k(C) \right\},$$

where $S_k(C)$ is defined as in Theorem 3.3. By the assumption on C and Theorem 3.3, the set $\tilde{S}$ is convex. Hence $\phi(\mathcal{U}(C)) = \phi(\tilde{S}) + (\operatorname{tr} C)(\alpha_1, \ldots, \alpha_m)$ is convex. □

Similarly, by the convexity result on the set of leading $k \times k$ principal submatrices of matrices in $\mathcal{U}(C)$ (see [27] and Proposition 3.1), we have the following.

Proposition 4.5 *Let* $1 \le k \le n$ *and let* C *have eigenvalues* $c_1 \ge \cdots \ge c_n$ *such that* $c_1 = c_k$ *and* $c_{n-k+1} = c_n$*. Suppose* $A = (A_1, \ldots, A_m) \in \mathcal{H}_n^m$ *is such that*

$$A_j = \alpha_j I + (P_j \oplus 0_{n-k}), \qquad j = 1, \ldots, m.$$

Then Conditions (a) – (d) *of Proposition 4.3 hold.*

Note that by Proposition 1.1, the set of vectors of diagonal entries of matrices in $\mathcal{U}(C)$ is convex. We have the following consequence.

Proposition 4.6 *Let* $C \in \mathcal{H}_n$ *and* $A = (A_1, \ldots, A_m) \in \mathcal{H}_n^m$ *be an* m*-tuple of diagonal matrices. Then Conditions* (a) – (d) *of Proposition 4.3 hold.*

By Theorem 4.2, if $C \in \mathcal{H}_n$ is non-scalar, then there exists $A \in \mathcal{H}_n^4$ so that $W_C(A)$ is not convex. In Proposition 4.6, we can find $A \in \mathcal{H}_n^n$ with linearly independent components so that $W_C(A)$ is convex for any $C \in \mathcal{H}_n$. In Propositions 4.4 and 4.5, we can find $A \in \mathcal{H}_n^m$ with $m = k(n-k)+1$ and k^2+1, respectively, so that A has linearly independent components and $W_C(A)$ is convex for some special choices of $C \in \mathcal{H}_n$. It is of interest to determine whether we can add more linearly independent components to those $A \in \mathcal{H}_n^m$ and still get a convex $W_C(A)$. Using the idea in the proof of [24, Theorem 2.3], we can show that if $C \in \mathcal{H}_n$ has distinct eigenvalues $\gamma_1 > \cdots > \gamma_k$ with multiplicities $n_1, \ldots, n_k$, respectively, then the (real) homogeneous manifold $\mathcal{U}(C)$ has dimension

$$r = n^2 - (n_1^2 + \cdots + n_k^2);$$

and if $\{A_1, \ldots, A_m\} \subseteq \mathcal{H}_n$ is linearly independent such that $W_C(A_1, \ldots, A_m)$ is convex, then

$$m \le r + 1.$$

If C is a rank k orthogonal projection with $k \le n/2$, then one can actually have linearly independent $A_1, \ldots, A_m \in \mathcal{H}_n$ with $m = r+1$ such that $W_C(A_1, \ldots, A_m)$ is convex (see [24]). Nevertheless, it is unknown whether the bound can be improved for other cases. In general, for a given $C \in \mathcal{H}_n$, it is interesting to determine the size of a maximal linearly independent set $\{A_1, \ldots, A_m\} \subseteq \mathcal{H}_n$ so that $W_C(A_1, \ldots, A_m)$ is convex.

Even if $W_C(A)$ is not convex, there are motivations from applications (see [13, 31]) to study the minimum number p such that every $v \in \operatorname{conv} W_C(A) \subseteq \mathbf{R}^m$ can be written as the convex combination of p elements in $W_C(A)$. Let us regard $W_C(A)$ as $\phi(\mathcal{U}(C))$ for a linear map $\phi : \mathcal{H}_n \to \mathbf{R}^m$. Using the results in [3] and Theorem 2.1, we have the following proposition.

Proposition 4.7 *Suppose $C \in \mathcal{H}_n$ and $A \in \mathcal{H}_n^m$. Then every point in* $\operatorname{conv} W_C(A)$ *can be written as the average of at most p elements in $W_C(A)$, where $p = \min\{n, m+1\}$. If C is a rank k orthogonal projection with $1 \le k < n$, then $p = \min\{q, m+1\}$ with $q = \max\{k+1, n-k+1\}$. If C satisfies any one (and hence all) of Conditions* (a) – (c) *in Theorem 2.1, then $p = 2$.*

5 Results on real symmetric matrices

The problems considered in the previous sections admit formulations for real symmetric matrices with $\mathcal{U}(C)$ replaced by $\mathcal{O}(C)$, the set of real symmetric matrices having the same eigenvalues as C. In fact, one often deals with real symmetric matrices instead of complex Hermitian matrices in problems arising in applications. For instance, in the computation of structured singular values that motivates the study of the convexity of $W_C(A)$, the matrices $C, A_1, \ldots, A_m$ under consideration are often real (see [10, 13]). Actually, in this case, one would consider

$$W_C^{\mathbf{R}}(A_1, \ldots, A_m) = \{(\operatorname{tr} A_1 X, \ldots, \operatorname{tr} A_m X) : X \in \mathcal{O}(C)\}.$$

In the following, we discuss results on real matrices analogous to those on complex Hermitian matrices in the previous sections. First, we present the following result of Fulton [14, Theorem 20] that allows us to "transfer" results on complex Hermitian matrices to real symmetric matrices.

Proposition 5.1 *Let C be a nonscalar $n \times n$ real symmetric matrix. If there exist complex Hermitian matrices $X_1, \cdots, X_m$ such that $C = X_1 + \cdots + X_m$, then there exist real symmetric matrices $Y_1, \ldots, Y_m$ so that $C = Y_1 + \cdots + Y_m$ and $Y_j \in \mathcal{U}(X_j)$ for each $j = 1, \ldots, m$.*

Note that every matrix $X \in \operatorname{conv} \mathcal{O}(C)$ can be written as the average of n complex Hermitian matrices (see [3]). By Proposition 5.1, we have the following corollary.

Corollary 5.2 *Let C be an $n \times n$ nonscalar real symmetric matrix. Then every $X \in \operatorname{conv} \mathcal{O}(C)$ can be written as the average of n matrices in $\mathcal{O}(C)$.*

In general, it is not hard to write a matrix in $\operatorname{conv} \mathcal{O}(C)$ as the *convex* combination of n matrices in $\mathcal{O}(C)$. In fact, for any $X = Q^tDQ \in \operatorname{conv} \mathcal{O}(C)$, where D is diagonal and Q is orthogonal, the vector of diagonal entries of D is majorized by $(c_1, \ldots, c_n)$. Thus (see e.g. [28]), D can be written as a convex combination of diagonal matrices in $\mathcal{O}(C)$. Since the diagonal matrices in $\mathcal{O}(C)$ have the same trace, the affine dimension of their convex hull is $n-1$. Hence, D (respectively, Q^tDQ) is a convex combination of n diagonal matrices $C_1, \ldots, C_n$ (respectively, $Q^tC_1Q, \ldots, Q^tC_nQ$) in $\mathcal{O}(C)$. Also, it is easy to write a matrix in $\operatorname{conv} \mathcal{O}(C)$ as the average of 2^{n-1} matrices in $\mathcal{O}(C)$ (see e.g. [3]). However, there does not seem to be an easy way to write a real symmetric matrix in $\operatorname{conv} \mathcal{O}(C)$ as the average of n matrices in $\mathcal{O}(C)$. It would be nice to find such a construction.

Next, we turn to the results in Sections 2 – 4. The results in Sections 2 and 3 are valid for the real symmetric case. One can use our proofs together with Proposition 5.1 and the real symmetric version of the Eigenvalue Inequality Theorem in [14, Theorem 3]. For those results in Section 4, some modifications are needed. First, the real symmetric version of Proposition 4.1 is valid for $W_C^{\mathbf{R}}(A)$; of course, we require V to be orthogonal in Condition (a). The real symmetric version of Theorem 4.2 is:

Theorem 5.3 *Suppose C is an $n \times n$ real symmetric matrix, where $n \geq 3$. Then for any $n \times n$ real symmetric matrices A_1, A_2 such that $\{I, A_1, A_2\}$ is linearly independent, there exists A_3 such that $W_C^{\mathbf{R}}(A_1, A_2, A_3)$ is not convex.*

The proof is similar to that of Theorem 4.2. One needs to restrict the choice of unitary matrices to real orthogonal matrices, and therefore the matrix $\tilde{A}_3$ in the proof will not appear.

For a given real symmetric C, one can also obtain an upper bound for the size of a linearly independent set $\{A_1, \ldots, A_m\}$ of real symmetric matrices such that $W_C^{\mathbf{R}}(A_1, \ldots, A_m)$ (see [24]). In such case, if C is a real symmetric matrix having distinct eigenvalues $\gamma_1 > \cdots > \gamma_k$ with multiplicities $n_1, \ldots, n_k$, respectively, then the (real) homogeneous manifold $\mathcal{U}(C)$ has dimension

$$r = \{n(n-1) - [n_1(n_1-1) + \cdots + n_k(n_k-1)]\}/2;$$

and if $\{A_1, \ldots, A_m\}$ is a linearly independent set of real symmetric matrices such that $W_C^{\mathbf{R}}(A_1, \ldots, A_m)$ is convex, then

$$m \leq r + 1.$$

Finally, Proposition 4.7 is valid for real symmetric matrices with the same proof.

Note that there are other types of joint numerical ranges and matrix completion problems associated with matrix orbits associated with other group actions (see [24, 26]). Our techniques can be used to obtain similar results for some of these variations.

Acknowledgement

Research of the first author was partially supported by an NSERC Grant.

Research of the second author was partially supported by an NSF grant and a faculty research grant of the College of William & Mary in the academic year 1998-99, when he was visiting the University of Toronto. He would like to thank the staff of the University of Toronto for their warm hospitality.

References

[1] Y.H. Au-Yeung and N.K. Tsing, An extension of the Hausdorff-Toeplitz theorem on the numerical range, *Proc. Amer. Math. Soc.* 89 (1983), 215–218.

[2] O. Azenhas, The admissible interval for the invariant factors of a product of matrices, preprint, 1998.

[3] R. Bhatia, M.D. Choi and C. Davis, Comparing a matrix to its off-diagonal part, *Operator Theory: Advances and Applications* 40 (1989), 151-164.

[4] N. Bebiano, C.K. Li and J. da Providencia, Product of diagonal elements of matrices, *Linear Algebra Appl.* 178 (1993), 185-200.

[5] P. Binding, D. Farenick and C.K. Li, A dilation and norm in several variable operator theory, *Canad. J. Math.* 47 (1995), 449-461.

[6] P. Binding and C.K. Li, Joint ranges of Hermitian matrices and simultaneous diagonalization, *Linear Algebra Appl.* 151 (1991), 157-168.

[7] A.S. Buch, The saturation conjecture (after A. Knutson and T.Tao), Enseignement Math. (2) 46 (2000), 43-60 .

[8] M. Cho and M. Takaguchi, Some classes of commuting m-tuples of operators, *Studia Math.* 80 (1984), 245-259.

[9] M.D. Choi and P.Y. Wu, Convex combinations of projections, *Linear Algebra Appl.* 136 (1990), 25-42.

[10] J.C. Doyle, Analysis of feedback systems with structured uncertainties, *Proc. IEEE-D* 129, 6 (1982), 242-250.

[11] J. Day, W. So and R.C. Thompson, The spectrum of a Hermitian matrix sum, *Linear Alg. Appl.* 280 (1998), 289-332.

[12] K. Fan and G. Pall, Imbedding conditions for Hermitian and normal matrices, *Canad. J. Math.* 9 (1957), 298-304.

[13] M. Fan and A. Tits, m-form numerical range and the computation of the structured singular values, *IEEE Trans. Automat. Control* AC-33 (1988), 284-289.

[14] W. Fulton, Eigenvalues of sums of Hermitian matrices (after A. Klyachko), Séminaire Bourbaki 845, June 1998, Astérisque 252 (1998), 255-269.

[15] M. Goldberg and E.G. Straus, Elementary inclusion relations for generalized numerical ranges, *Linear Algebra Appl.* 18 (1977), 1-24.

[16] U. Helmke and J. Rosenthal, Eigenvalue inequalities and Schubert calculus, *Math. Nachr.* 171 (1995), 207-225.

[17] A. Horn, Doubly Stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.* 76 (1954), 620-630.

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

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

[20] A.A. Klyachko, Stable bundles, representation theory and Hermitian operators, *Selecta Math. (N.S.)* 4 (1998), 419-445.

[21] A. Knutson and T. Tao, The honeycomb model of the Berenstein-Zelevinsky polytope I: Klyacho's saturation conjecture, lanl preprint, 1998.

[22] C.K. Li, Matrices with some extremal properties, *Linear Algebra Appl.* 101 (1988), 255-267.

[23] C.K. Li and R. Mathias, Inequalities on the singular values of an off-diagonal block of a Hermitian matrix, *J. of Inequalities and Applications* 3 (1999), 137-142.

[24] C.K. Li and Y.T. Poon, Convexity of the kth joint numerical range, *SIAM J. Matrix Analysis Appl.* 21 (1999), 668-678.

[25] C.K. Li and Y.T. Poon, Off diagonal submatrices of a Hermitian matrix, submitted. Preprint available at http://www.resnet.wm.edu/~cklixx/poon-9.pdf.

[26] C.K. Li and T.Y. Tam, Numerical ranges arising from simple Lie algebras, *Canad. J. Math.* 52 (2000), 141-171.

[27] C.K. Li and N.K. Tsing, On the k^{th} matrix numerical range, *Linear and Multilinear Algebra* 28 (1991), 229-239.

[28] A.W. Marshall and I. Olkin, *Inequalities: The Theory of Majorizations and Its Applications,* Academic Press, 1979.

[29] L. Mirsky, Matrices with prescribed characteristic roots and diagonal elements, *J. London Math. Soc.* 33 (1958), 14-21.

[30] Y.T. Poon, Another proof of a result of Westwick, *Linear and Multilinear Algebra* 9 (1980), 181-186.

[31] Y.T. Poon, On the convex hull of the multiform numerical range, *Linear and Multilinear Algebra* 37 (1994), 221-224.

[32] I. Schur, *Über eine Klasse von Mittelbildungen mit Anwendungen die Determinanten-Theorie Sitzungsber. Berlin. Math. Gesellschaft* 22 (1923), 9-20. (See also, *Issai Schur Collected Works* (A. Brauer and H. Rohrbach, eds.) Vol. II. pp. 416-427. Springer-Verlag, Berlin, 1973.)

[33] B.S. Tam, A simple proof of Goldberg-Straus theorem on numerical radii, *Glasgow Math. J.* 28 (1986), 139-141.

[34] R.C. Thompson, Singular values, diagonal elements, and convexity, *SIAM J. Appl. Math.* 32 (1977), 39-63.

[35] R.C. Thompson, Research problem: The matrix numerical range, *Linear and Multilinear Algebra* 21 (1987), 321-323.

[36] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups,* Scott. Foresman, London, 1971.

[37] R. Westwick, A theorem on numerical ranges, *Linear and Multilinear Algebra* 2 (1975), 311-315.

[38] A. Zelevinsky, Littlewood-Richardson semigroups, MSRI preprint 1997-044.

Department of Mathematics, University of Toronto, Ontario, Canada M5S 3G3
(choi@math.toronto.edu)

Department of Mathematics, College of William & Mary, Williamsburg, Virginia 23187
(ckli@math.wm.edu)

Department of Mathematics, Iowa State University, Ames, Iowa 50011
(ytpoon@iastate.edu)