STUDIA MATHEMATICA
Online First version

The joint numerical range of commuting matrices

by

PAN-SHUN LAU (Reno, NV), CHI-KWONG LI (Williamsburg, VA)
and YIU-TUNG POON (Ames, IA)

Abstract. It is shown that for $n \leq 3$ the joint numerical range of a family of commuting $n \times n$ complex matrices is always convex; for $n \geq 4$ there are two commuting matrices whose joint numerical range is not convex.

1. Introduction. Let $M_{m,n}$ be the set of $m \times n$ complex matrices. For $A \in M_{m,n}$, A^* (resp. A^t) stands for the conjugate transpose (resp. transpose) of A; for example, see [9, 10]. Denote by $\mathbb{C}^n$ (resp. $\mathbb{R}^n$) the set of column vectors with n complex (resp. real) entries. Let $M_n = M_{n,n}$ and M_n^m be the set of all m-tuples of $n \times n$ matrices. We identify $\mathbb{C}^n$ with $M_{n,1}$. For notational convenience, we will also say that $\mathbf{z} \in \mathbb{C}^n$ for a complex row vector $\mathbf{z} = (z_1, \dots, z_n)$. The *joint numerical range* of $\mathbf{A} = (A_1, \dots, A_m) \in M_n^m$ is defined by

$$W(\mathbf{A}) = \{(\mathbf{x}^* A_1 \mathbf{x}, \dots, \mathbf{x}^* A_m \mathbf{x}) : \mathbf{x} \in \mathbb{C}^n,\ \mathbf{x}^*\mathbf{x} = 1\} \subseteq \mathbb{C}^m.$$

When $m = 1$, it reduces to the classical numerical range $W(A_1)$ of $A_1 \in M_n$, which is a useful tool for studying matrices and operators; for example, see [10, Chapter 1]. The joint numerical range of m matrices is useful in studying the behavior of the family of matrices $\{A_1, \dots, A_m\} \subseteq M_n$, and has applications in many pure and applied areas. We refer the readers to the excellent survey [14] and the paper [15] on this subject.

When $m = 1$, the Toeplitz–Hausdorff theorem asserts that $W(A_1)$ is always convex. However, for $m \geq 2$, $W(A_1, \dots, A_m)$ may fail to be convex; see [11]. Many researchers have studied matrices with certain commutative properties that have convex joint numerical ranges; e.g., see [3, 4, 5, 6, 11, 12, 13]. In particular, Dash [5, Proposition 2.4] proved that $W(A_1, \dots, A_m)$ is always convex for any commuting family $\{A_1, \dots, A_m\} \subseteq M_2$ and raised the ques-

2020 *Mathematics Subject Classification*: Primary 15A60; Secondary 47A12, 47A20.
Key words and phrases: joint numerical range, convex set, commuting family.
Received 22 June 2020; revised 18 May 2022.
Published online *.

tion on the same result for $\{A_1,\dots,A_m\} \subseteq M_n$ with $n > 2$. In [13], Müller gave a simple example, which was incorporated in [15] with some improvements, of a commuting family $\{A_1, A_2, A_3\} \subseteq M_4$ such that $W(A_1, A_2, A_3)$ is not convex, and raised the question of whether $W(A_1, A_2)$ is convex if $A_1A_2 = A_2A_1$; see [13, Problem 2]. We will show that the answer is negative if A_1, A_2 is a commuting pair of matrices (or infinite-dimensional operators) with dimension at least 4. However, for a commuting pair of matrices $A_1, A_2 \in M_3$, $W(A_1, A_2)$ is always convex. We can then deduce from the results that $W(A_1,\dots,A_m)$ is always convex for any commuting family $\{A_1,\dots,A_m\} \subseteq M_3$.

Our paper is organized as follows. In Section 2, we present some preliminary results including a short proof of the convexity of $W(A_1,\dots,A_m)$ for every commuting family $\{A_1,\dots,A_m\} \subseteq M_2$. In Section 3, we present examples of commuting matrices (or infinite-dimensional operators) A_1, A_2 of dimension at least 4 such that $W(A_1, A_2)$ is not convex. We then state our main result that $W(A_1, A_2)$ is convex if $A_1, A_2 \in M_3$ commute, and we deduce that $W(A_1,\dots,A_m)$ is convex for any commuting family $\{A_1,\dots,A_m\} \subseteq M_3$. The rather involved proof of the main theorem on the convexity of $W(A_1, A_2)$ for a commuting pair $A_1, A_2 \in M_3$ will be given in Section 4.

2. Preliminaries and commuting families in M_2.

Let

$$\mathcal{H}_n = \{A \in M_n : A = A^*\}$$

be the real space of all $n \times n$ Hermitian matrices and I_n be the $n \times n$ identity matrix. We summarize some properties of joint numerical ranges which are useful for what follows. We refer the interested readers to [1, 8, 11].

PROPOSITION 2.1. *Let $\mathcal{F} = \{A_1,\dots,A_m\} \subseteq M_n$. Suppose the complex space spanned by $\{A_1,\dots,A_m\}$ has a basis $\{C_1,\dots,C_s\}$. Let $A_j = H_j + iG_j$, where $H_j, G_j \in \mathcal{H}_n$ for $j = 1,\dots,m$. Then:*

(a) $W(A_1,\dots,A_m) = W(U^*A_1U,\dots,U^*A_mU)$ *for any unitary* $U \in M_n$.
(b) $W(A_1,\dots,A_m) = W(A_1^t,\dots,A_m^t)$.
(c) $W(A_1,\dots,A_m)$ *is convex if and only if* $W(C_1,\dots,C_s)$ *is convex.*
(d) *The family $\mathcal{F}$ is commuting if and only if $\{C_1,\dots,C_s\}$ is commuting.*
(e) $W(A_1,\dots,A_m) \subseteq \mathbb{C}^m$ *can be identified with* $W(H_1, G_1,\dots,H_m, G_m) \subseteq \mathbb{R}^{2m}$.
(f) *For $n = 2$ and $H_1,\dots,H_m \in \mathcal{H}_2$, $W(H_1,\dots,H_m)$ is convex if and only if* $\operatorname{span}\{I_2, H_1,\dots,H_m\} \neq \mathcal{H}_2$.
(g) *Suppose $n \geq 3$ and $H_1,\dots,H_m \in \mathcal{H}_n$. If* $\operatorname{span}\{I_n, H_1,\dots,H_m\}$ *has dimension at most 4, then $W(H_1,\dots,H_m)$ is convex.*

Note that (c) and (f) are given in [8, Corollary 2.4 and Example 1] and (g) is given in [1, Corollary 1]. By (e), the study of convexity of $W(A_1,\dots,A_m)$

can be reduced to $W(H_1, G_1, \ldots, H_m, G_m)$ for Hermitian matrices H_1, G_1, ..., H_m, G_m. However, it is clear that the commutativity of $A_1, \ldots, A_m$ does not imply the commutativity of H_1, G_1, ..., H_m, G_m. In fact, if $\{H_1, G_1, \ldots, H_m, G_m\}$ is a commuting family, then $\{A_1, \ldots, A_m\}$ is a commuting family of normal matrices, and $W(A_1, \ldots, A_m)$ will be polyhedral, i.e., a convex hull of finitely many points in $\mathbb{C}^m$; see [5, Theorem 2.5]. It is clear that $(\mu_1, \ldots, \mu_m) \in W(A_1, \ldots, A_m)$ if and only if $(1, \mu_1, \ldots, \mu_m) \in W(I_n, A_1, \ldots, A_m)$ for any $(A_1, \ldots, A_m) \in M_n^m$. By Proposition 2.1, to study the convexity of $W(A_1, \ldots, A_m)$, one may focus on $W(C_1, \ldots, C_s)$ where $\{I_n, C_1, \ldots, C_s\}$ is a basis for the span of $\{I_n, A_1, \ldots, A_m\}$. It is well-known that if $\{A_1, \ldots, A_m\}$ is a commuting family of matrices then there is a unitary U such that U^*A_jU are in upper triangular form for all $j = 1, \ldots, m$; see [16]. Our proofs often use this property.

Denote by $\operatorname{conv} S$ and ∂S respectively the convex hull and the boundary of a set S in $\mathbb{R}^m$ or $\mathbb{C}^m$. The next result describes the intersection of support planes of $\operatorname{conv} W(A_1, \ldots, A_m)$ with $W(A_1, \ldots, A_m)$.

PROPOSITION 2.2. *Let $B_1, \ldots, B_r \in \mathcal{H}_n$ be Hermitian matrices. For every unit vector $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_r) \in \mathbb{R}^r$, let*

$$P_{\boldsymbol{\nu}} = \{\mathbf{b} \in \mathbb{R}^r : \mathbf{b}^*\boldsymbol{\nu} \le \lambda_1(\nu_1 B_1 + \cdots + \nu_r B_r)\},$$

where $\lambda_1(H)$ denotes the largest eigenvalue of $H \in \mathcal{H}_n$ and $\mathbf{b}^*\boldsymbol{\nu} = \sum_{i=1}^r b_i\nu_i$ for $\mathbf{b} = (b_1, \ldots, b_r) \in \mathbb{R}^r$. Then

$$\operatorname{conv} W(B_1, \ldots, B_r) = \bigcap \{P_{\boldsymbol{\nu}} : \boldsymbol{\nu} = (\nu_1, \ldots, \nu_r) \in \mathbb{R}^r,\ \boldsymbol{\nu}^*\boldsymbol{\nu} = 1\}.$$

Consequently,

$$\begin{aligned}&\partial P_{\boldsymbol{\nu}} \cap W(B_1, \ldots, B_r)\\ &\qquad = \{(\mathbf{x}^*B_1\mathbf{x}, \ldots, \mathbf{x}^*B_r\mathbf{x}) : \mathbf{x} \in \mathbb{C}^n,\ \mathbf{x}^*\mathbf{x} = 1,\ B_{\boldsymbol{\nu}}\mathbf{x} = \lambda_1(B_{\boldsymbol{\nu}})\mathbf{x}\},\end{aligned}$$

where $B_{\boldsymbol{\nu}} = \sum_{j=1}^r \nu_j B_j$. Moreover, $\partial P_{\boldsymbol{\nu}} \cap W(B_1, \ldots, B_r)$ is convex if and only if

$$W(X^*B_1X, \ldots, X^*B_rX)$$

is convex, where the columns of X form an orthonormal basis for the null space of $B_{\boldsymbol{\nu}} - \lambda_1(B_{\boldsymbol{\nu}})I_n$.

Proof. If $\mathbf{x} \in \mathbb{C}^n$ is a unit vector and

$$\mathbf{b} = (\mathbf{x}^*B_1\mathbf{x}, \ldots, \mathbf{x}^*B_r\mathbf{x}) \in W(B_1, \ldots, B_r),$$

then for any unit vector $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_r) \in \mathbb{R}^r$ we have

$$\mathbf{b}^*\boldsymbol{\nu} = \mathbf{x}^*\Big(\sum_{j=1}^r \nu_j B_j\Big)\mathbf{x} \le \lambda_1\Big(\sum_{j=1}^r \nu_j B_j\Big).$$

Thus, $W(B_1, \ldots, B_r) \subseteq P_{\boldsymbol{\nu}}$. As $P_{\boldsymbol{\nu}}$ is convex, $\operatorname{conv} W(B_1, \ldots, B_r) \subseteq P_{\boldsymbol{\nu}}$.

Conversely, suppose $\mathbf{b} = (b_1, \ldots, b_r) \notin \operatorname{conv} W(B_1, \ldots, B_r) \subseteq \mathbb{R}^r$. By the separation theorem, there exists a real unit vector $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_r) \in \mathbb{R}^r$ such that $\sum_{j=1}^r b_j\nu_j > \sum_{j=1}^r q_j\nu_j$ for all $(q_1, \ldots, q_r) \in W(B_1, \ldots, B_r)$, i.e., for every unit vector $\mathbf{x} \in \mathbb{C}^n$,

$$\sum_{j=1}^{r} b_j\nu_j > \sum_{j=1}^{r} \nu_j(\mathbf{x}^* B_j \mathbf{x}) = \mathbf{x}^*\Big(\sum_{j=1}^{r} \nu_j B_j\Big)\mathbf{x}.$$

So, $\sum_{j=1}^r b_j\nu_j > \lambda_1(\sum_{j=1}^r \nu_j B_j)$.

The last two assertions are clear. ∎

The following result is proven in [5, Proposition 2.4]. Recently, it was also given in [2, Theorem 2.2]. We give a short proof here for completeness.

Proposition 2.3. *For any commuting family $\mathcal{F} = \{A_1, \ldots, A_m\} \subseteq M_2$, $W(A_1, \ldots, A_m)$ is convex.*

Proof. To avoid trivial considerations, suppose $\mathcal{F}$ contains a nonscalar matrix $X \in M_2$. Applying a unitary similarity, we may assume that all matrices in $\mathcal{F}$ are in upper triangular form. Let $X_0 = X - \frac{\operatorname{tr} X}{2} I_2 = \left(\begin{smallmatrix} x_1 & x_2 \\ 0 & -x_1 \end{smallmatrix}\right)$. We claim that for every $Y \in \mathcal{F}$, $Y_0 = Y - \frac{\operatorname{tr} Y}{2} I_2 = \left(\begin{smallmatrix} y_1 & y_2 \\ 0 & -y_1 \end{smallmatrix}\right)$ is a multiple of X_0; see [7, Theorem II] for an alternative proof. Thus every A_j is a linear combination of I_2, $H_1 = (X_0 + X_0^*)/2$ and $H_2 = (X_0 - X_0^*)/(2i)$. By Proposition 2.1(f), $W(A_1, \ldots, A_m)$ is convex.

To prove our claim, note that X_0 commutes with Y_0, i.e., $x_1y_2 = x_2y_1$. Since X is nonscalar, either $x_1 \neq 0$ or $x_2 \neq 0$.

If $x_1 = 0$, then $x_2 \neq 0$ and $x_2y_1 = 0$. Thus $y_1 = 0$ and $Y_0 = (y_2/x_2)X_0$. Our claim follows.

If $x_1 \neq 0$, then $x_1y_2 = x_2y_1$ implies $Y_0 = (y_1/x_1)X_0$. Again, our claim follows. ∎

3. Convexity of commuting family of dimension at least 3. In [13], the author gave an elegant example of a commuting family $\{A_1, A_2, A_3\} \subseteq M_4$ with nonconvex $W(A_1, A_2, A_3)$. The following example illustrates that $W(A_1, A_2)$ may not be convex for a commuting pair $A_1, A_2 \in M_4$.

Example 3.1. Let $A_1 = H_1 + iG_1$ and $A_2 = A_1 + A_1^2 - A_1^3 - 12I_4 = H_2 + iG_2$ with

$$H_1 = \operatorname{diag}(2, 2, 1, 0),$$

$$G_1 = \begin{pmatrix} 1 & 0 & 2-i & -i \\ 0 & 0 & -1+i & 1-i \\ 2+i & -1-i & 0 & 0 \\ i & 1+i & 0 & 0 \end{pmatrix},$$

$$H_2 = \begin{pmatrix} 14 & -9-7i & 8-4i & -3i \\ -9+7i & 0 & 0 & 0 \\ 8+4i & 0 & 10 & -2-4i \\ 3i & 0 & -2+4i & -9 \end{pmatrix},$$

$$G_2 = \begin{pmatrix} 6 & -2-2i & 12-4i & -4-6i \\ -2+2i & 0 & -3+7i & 5-i \\ 12+4i & -3-7i & 5 & 1-2i \\ -4+6i & 5+i & 1+2i & 1 \end{pmatrix}.$$

Then $A_1A_2 = A_2A_1$. Note that for the unit vector $\boldsymbol{\nu} = (1,0,0,0)$, the matrix $A_{\boldsymbol{\nu}} = \nu_1 H_1 + \nu_2 G_1 + \nu_3 H_2 + \nu_4 G_1 = H_1$ has the largest eigenvalue 2, and the null space of $A_{\boldsymbol{\nu}} - 2I_4$ is spanned by the first two columns of I_4. Let $X \in M_{4,2}$ be the matrix formed by these two columns. It is easy to check that

$$\operatorname{span}\{X^*H_1X, X^*G_1X, X^*H_2X, X^*G_2X\} = \mathcal{H}_2.$$

By Proposition 2.1(e, f), $W(X^*A_1X, X^*A_2X)$ is not convex. Since

$$W(X^*A_1X, X^*A_2X) = \{(\mu_1,\mu_2) \in W(A_1,A_2) : \operatorname{Re}\mu_1 = 2\},$$

$W(A_1,A_2)$ is not convex.

Remark 3.2. For $n > 4$, one can extend the above example to $\tilde{A}_1 = A_1 \oplus 0_N$ and $\tilde{A}_2 = A_2 \oplus 0_N$, where $1 \le N \le \infty$. It is clear that $\tilde{A}_1\tilde{A}_2 = \tilde{A}_2\tilde{A}_1$ and $W(\tilde{A}_1, \tilde{A}_2)$ is not convex.

For commuting $A_1, A_2 \in M_3$, we have the following.

Theorem 3.3. *Suppose that $A_1, A_2 \in M_3$ commute. Then $W(A_1,A_2)$ is convex.*

The proof of the result is quite involved and technical. We will present it in the next section. From Theorem 3.3, we can deduce the following.

Theorem 3.4. *Let $\{A_1,\dots,A_m\} \subseteq M_3$ be a commuting family of matrices. Then the complex linear span of $\{I_3, A_1,\dots,A_m\}$ has dimension at most 3, and hence $W(A_1,\dots,A_m)$ is convex.*

Proof. We may assume that $A_1,\dots,A_m$ are in upper triangular form, and $\mathcal{F} = \{I_3, A_1,\dots,A_m\}$ is linearly independent. We are going to prove by contradiction that $m \le 2$. In the following, we will use $\operatorname{diag} A \in \mathbb{C}^n$ as the vector of diagonal entries of $A \in M_n$.

Suppose to the contrary that $m > 2$. Then $\{\operatorname{diag} I_3, \operatorname{diag} A_1, \dots, \operatorname{diag} A_m\}$ is linearly dependent. Therefore, span $\mathcal{F}$ has a nonzero nilpotent. We may assume that A_1 is a nonzero nilpotent in span $\mathcal{F}$ of the largest rank. Consider the following cases:

CASE 1: $\operatorname{rank} A_1 = 2$. Then there is an invertible S such that $S^{-1}A_1S = J$ is the upper triangular Jordan block. Then for every $2 \le i \le m$, $A_1A_i = A_iA_1$ implies that $S^{-1}A_iS = a_iI_3 + b_iJ + c_iJ^2$ for some $a_i, b_i, c_i \in \mathbb{C}$. Since $\{I_3, A_1, \ldots, A_m\}$ is linearly independent, we have $m \le 2$, a contradiction.

CASE 2: $\operatorname{rank} A_1 = 1$. So, up to a nonzero multiple and a unitary similarity transform, we may assume that $A_1 = \begin{pmatrix} 0&0&1\\0&0&0\\0&0&0 \end{pmatrix}$. Then for every $2 \le i \le m$, the condition $A_1A_i = A_iA_1$ implies that A_i is in upper triangular form with the $(1,1)$-entry equal to the $(3,3)$-entry. We may then replace A_i by $A_i - \alpha_iI_3 - \beta_iA_1$ for some $\alpha_i, \beta_i \in \mathbb{C}$ and assume that

$$A_i = \begin{pmatrix} 0 & b_i & 0 \\ 0 & a_i & c_i \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{for some } a_i, b_i, c_i \in \mathbb{C},\ i = 2, \ldots, m.$$

If $a_i = 0$ for all $2 \le i \le m$, then $\operatorname{span}\{A_2, A_3\}$ would contain a nonzero nilpotent of rank 2, which contradicts the assumption that A_1 has the largest rank. Therefore, we may assume that $a_2 = 1$ and $a_3 = 0$. Then $A_2A_3 = A_3A_2$ implies that $b_3 = c_3 = 0$, which contradicts $\mathcal{F}$ being linearly independent.

This shows that $m \le 2$ and the convexity of $W(A_1, A_2)$ follows from Theorem 3.3. ■

4. Proof of Theorem 3.3. We divide the proof into two subsections. We will always assume that $A_1 = H_1 + iG_1$ and $A_2 = H_2 + iG_2$, where H_1, G_1, H_2, G_2 are Hermitian. In view of Proposition 2.1(g), we always assume that $\operatorname{span}\{I_n, H_1, G_1, H_2, G_2\}$ has dimension 5 to avoid trivial considerations.

4.1. $\operatorname{span}\{I_3, A_1, A_2\} \subseteq M_3$ **does not contain a nonzero nilpotent.** In this subsection, we assume that $\operatorname{span}\{I_3, A_1, A_2\} \subseteq M_3$ does not contain a nonzero nilpotent. Without loss of generality, by applying unitary similarity transforms and taking linear combinations of I_3, A_1, A_2, one can assume that

$$A_1 = \begin{pmatrix} 1 & u & w_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & 0 & w_2 \\ 0 & 0 & v \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{array}{l} \text{with } u, v \ge 0, \\ w_1 + uv + w_2 = 0. \end{array} \tag{4.1}$$

The reduction can be done as follows. Since A_1 and A_2 commute, we assume without loss of generality that both A_1, A_2 are in upper triangular form. Since $\operatorname{span}\{I_3, A_1, A_2\}$ does not contain a nonzero nilpotent matrix, $\{\operatorname{diag} I_3, \operatorname{diag} A_1, \operatorname{diag} A_2\} \subseteq \mathbb{C}^3$ is linearly independent. Replacing A_j by $\alpha_jA_1 + \beta_jA_2 + \gamma_jI_3$ with suitable $\alpha_j, \beta_j, \gamma_j \in \mathbb{C}$, $j = 1, 2$, we may assume that

$$A_1 = \begin{pmatrix} 1 & a_1 & a_2 \\ 0 & 0 & a_3 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & b_1 & b_2 \\ 0 & 0 & b_3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then we have

$$A_1A_2 = \begin{pmatrix} 0 & b_1 & a_2 + b_2 + a_1b_3 \\ 0 & 0 & a_3 \\ 0 & 0 & 0 \end{pmatrix} = A_2A_1 = \begin{pmatrix} 0 & 0 & a_3b_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore, $a_3 = b_1 = 0 = a_2 + b_2 + a_1b_3$. Replacing A_j by DA_jD^{-1} with a diagonal unitary matrix D, we may assume $a_1, b_3 \geq 0$, so that we get (4.1).

By Proposition 2.1, the convexity of $W(A_1, A_2)$ is equivalent to the convexity of the numerical range of (A_1, A_2) transformed into the form (4.1).

In the following, we will show that $W(A_1, A_2)$ is convex if $A_1, A_2 \in M_3$ are of the form in (4.1).

PROPOSITION 4.1. *Let $A_1, A_2 \in M_3$ be of the form* (4.1). *If* $(0, 0) \in \{(u, w_1), (v, w_2), (u, v)\}$*, then $W(A_1, A_2)$ is convex.*

Proof. If $u = w_1 = 0$, then set $(H_2, G_2) = (A_2 + A_2^*, i(A_2^* - A_2))/2$ and identify $W(A_1, A_2)$ with $W(A_1, H_2, G_2) \subseteq \mathbb{R}^3$, which is convex by Proposition 2.1(g).

If $v = w_2 = 0$, then set $(H_1, G_1) = (A_1 + A_1^*, i(A_1^* - A_1))/2$ and identify $W(A_1, A_2)$ with $W(H_1, G_1, A_2) \subseteq \mathbb{R}^3$, which is convex.

If $u = v = 0$, then $w_1 + w_2 = 0$. By the previous argument, $A_1 + A_2 = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, hence $W(A_1 + A_2, A_2)$ is convex, and so is $W(A_1, A_2)$. ∎

Next, we treat the case where $(0, 0) \notin \{(u, w_1), (v, w_2), (u, v)\}$. First, we show that $W(A_1, A_2)$ has convex boundary.

PROPOSITION 4.2. *Let $A_1, A_2 \in M_3$ be commuting matrices of the form* (4.1) *such that* $(0, 0) \notin \{(u, w_1), (v, w_2), (u, v)\}$. *Then $W(A_1, A_2)$ contains all of the boundary points of* $\operatorname{conv} W(A_1, A_2)$.

Proof. Suppose A_1 and A_2 satisfy the hypothesis, and $A_1 = H_1 + iG_1$, $A_2 = H_2 + iG_2$, where $H_1, H_2, G_1, G_2 \in \mathcal{H}_3$. For every unit vector $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3, \nu_4) \in \mathbb{R}^4$, let

$$P_{\boldsymbol{\nu}} = \Big\{(b_1, \ldots, b_4) \in \mathbb{R}^4 : \sum_{i=1}^{4} b_i\nu_i \leq \lambda_1(\nu_1H_1 + \nu_2G_1 + \nu_3H_2 + \nu_4G_2)\Big\}.$$

By Proposition 2.2 every boundary point of $\operatorname{conv} W(A_1, A_2)$ lies in $\partial P_{\boldsymbol{\nu}}$ for some $\boldsymbol{\nu} \in \mathbb{R}^4$, and

$$\partial P_{\boldsymbol{\nu}} \cap \operatorname{conv} W(A_1, A_2) = \operatorname{conv}(\partial P_{\boldsymbol{\nu}} \cap W(A_1, A_2)).$$

We will show that $\partial P_{\boldsymbol{\nu}} \cap \operatorname{conv} W(A_1, A_2) \subseteq W(A_1, A_2)$.

CASE 1. Suppose one of the following conditions holds:

(i) $uv = 0$,
(ii) $(w_1 - w_2)^2 = (uv)^2$, or
(iii) $|w_1|\sqrt{1+v^2} \neq |w_2|\sqrt{1+u^2}$.

In each of these cases, we will show that $\partial P_{\boldsymbol{\nu}} \cap \operatorname{conv} W(A_1, A_2)$ is a singleton lying in $W(A_1, A_2)$ for any unit vector $\boldsymbol{\nu}$.

Let $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3, \nu_4) \in \mathbb{R}^4$ be a unit vector. The matrix $B_{\boldsymbol{\nu}} = \nu_1 H_1 + \nu_2 G_1 + \nu_3 H_2 + \nu_4 G_2$ has the form

$$\begin{pmatrix} \nu_1 & \dfrac{u(\nu_1 - i\nu_2)}{2} & \dfrac{w_1(\nu_1 - i\nu_2) + w_2(\nu_3 - i\nu_4)}{2} \\ \dfrac{u(\nu_1 + i\nu_2)}{2} & 0 & \dfrac{v(\nu_3 - i\nu_4)}{2} \\ \dfrac{\overline{w}_1(\nu_1 + i\nu_2) + \overline{w}_2(\nu_3 + i\nu_4)}{2} & \dfrac{v(\nu_3 + i\nu_4)}{2} & \nu_3 \end{pmatrix}.$$

Let r be the multiplicity of $\lambda_1(B_{\boldsymbol{\nu}})$. Since $\{I_3, H_1, G_1, H_2, G_2\}$ is linearly independent, $r \leq 2$.

By Proposition 2.2, if $r = 1$, then $\partial P_{\boldsymbol{\nu}} \cap W(A_1, A_2)$ is singleton and equals $\partial P_{\boldsymbol{\nu}} \cap \operatorname{conv} W(A_1, A_2)$.

We show that $r = 2$ is impossible under any one of the assumptions (i), (ii) or (iii). Assume to the contrary that $r = 2$. As $(0,0) \notin \{(u,v), (u, w_1), (v, w_2)\}$, we see that $\lambda_1(B_{\boldsymbol{\nu}}) \neq 0$. Since the $(2,2)$-entry of $B_{\boldsymbol{\nu}}$ is 0, we have $\lambda_1(B_{\boldsymbol{\nu}}) = \lambda_2(B_{\boldsymbol{\nu}}) > 0 \geq \lambda_3(B_{\boldsymbol{\nu}})$. Thus, there is a nonzero real vector (a, b, c, d) such that

$$R = I_3 + aH_1 + bG_1 + cH_2 + dG_2 = \mathbf{z}\mathbf{z}^* \tag{4.2}$$

for some nonzero $\mathbf{z} \in \mathbb{C}^3$, so that R is a rank 1 positive semidefinite matrix. Let $\mathcal{S}$ be the set of all nonzero real vectors (a, b, c, d) such that R is a rank 1 positive semidefinite matrix. We are going to show that $S = \emptyset$, thus arriving at a contradiction. In such a case, we may assume that

$$\mathbf{z} = (z_1, z_2, z_3) = (\sqrt{a+1}\, e^{i\theta_1}, 1, \sqrt{c+1}\, e^{i\theta_2})$$

with

$$\sqrt{a+1}\, e^{i\theta_1} = z_1 = z_1 \bar{z}_2 = \frac{u}{2}(a - ib), \quad \sqrt{c+1}\, e^{-i\theta_2} = \bar{z}_3 = z_2 \bar{z}_3 = \frac{v}{2}(c - id),$$

and

$$\begin{aligned} \frac{uv(a-ib)(c-id)}{4} &= \sqrt{(a+1)(c+1)}\, e^{i(\theta_1 - \theta_2)} \\ &= z_1 \bar{z}_3 = \frac{w_1(a-ib) + w_2(c-id)}{2}. \end{aligned}$$

The matrix R given by (4.2) then has the form

$$R = \begin{pmatrix} a+1 & u(a-ib)/2 & uv(a-ib)(c-id)/4 \\ u(a+ib)/2 & 1 & v(c-id)/2 \\ uv(a+ib)(c+id)/4 & v(c+id)/2 & c+1 \end{pmatrix}. \tag{4.3}$$

Since R has rank 1, we have $(a,b) \neq (0,0)$ and $(c,d) \neq (0,0)$. If any one of the assumptions (i), (ii) or (iii) holds, we are going to derive a contradiction.

Suppose that (i) holds, i.e., $uv = 0$. Recall from (4.1) that u and v are nonnegative. Since $(u,v) \neq (0,0)$, we assume $u = 0 < v$ or $v = 0 < u$. Let $u = 0$ and $v > 0$. Since $(u, w_1) \neq (0,0)$, we may replace (A_1, A_2) by (D^*A_1D, D^*A_2D) for some suitable diagonal unitary matrix D and assume that $-w_2 = w_1 > 0$. Suppose there is a real vector (a,b,c,d) such that R given by (4.2) is a rank 1 positive semidefinite matrix of the form (4.3). Since the $(1,2)$-entry is zero, we see that $a = -1$. Now, $-w_2 = w_1 > 0$ and the $(1,3)$-entry of R is $w_1((a-ib)-(c-id)) = 0$. Thus $c = a = -1$ and $b = d$. As a result, the $(3,3)$-entry of R is zero and so must be the $(2,3)$-entry. Hence, $v = 0$, which is a contradiction. Similarly, we can show that for $u > 0$ and $v = 0$, $\mathcal{S} = \emptyset$.

Suppose now that $u, v > 0$. As the matrix R in (4.3) is rank 1, we have $4(a+1)/u^2 = (a^2+b^2)$ and $4(c+1)/v^2 = (c^2+d^2)$. Therefore,

$$a+ib \in \mathcal{E}_u := \{x+iy : (x-2/u^2)^2 + y^2 = 4(1/u^2 + 1/u^4)\}, \tag{4.4}$$

$$c+id \in \mathcal{E}_v := \{x+iy : (x-2/v^2)^2 + y^2 = 4(1/v^2 + 1/v^4)\}. \tag{4.5}$$

Since $w_1+uv+w_2 = 0$, we may let $w_1 = -uv(1-\xi)/2$, $w_2 = -uv(1+\xi)/2$ for some $\xi \in \mathbb{C}$. As $\xi = (w_1-w_2)/uv$, assumption (ii) holds, i.e., $(w_1-w_2)^2 = (uv)^2$, if and only if $\xi = \pm 1$. Now the $(1,3)$-entry of R becomes

$$\begin{aligned} uv(a-ib)(c-id)/4 &= [w_1(a-ib) + w_2(c-id)]/2 \\ &= -[uv(a-ib)(1-\xi) + uv(c-id)(1+\xi)]/4. \end{aligned}$$

Thus,

$$(a-ib)(c-id) = (\xi-1)(a-ib) - (\xi+1)(c-id). \tag{4.6}$$

If $\xi = 1$, then we have $(a-ib)(c-id) = -2(c-id)$ so that $a-ib = -2$. Thus, the $(1,1)$-entry of R is -1, which is impossible. Similarly, if $\xi = -1$, then the $(3,3)$-entry of R is -1, which is impossible.

Suppose $\xi \neq \pm 1$ and (iii) holds. Substituting $w_1 = -uv(1-\xi)/2$, $w_2 = -uv(1+\xi)/2$, we have

$$|1-\xi|\sqrt{1+v^2} \neq |1+\xi|\sqrt{1+u^2}. \tag{4.7}$$

Since $a-ib, c-id \neq 0$, (4.6) is equivalent to

$$\frac{1+\xi}{a-ib} + \frac{1-\xi}{c-id} + 1 = 0. \tag{4.8}$$

Note that $\mu \in \mathbb{C}$ lies on a circle with center $\mu_0 \geq 0$ and radius $r > \mu_0$ if and only if

$$0 = (\mu - \mu_0)(\bar{\mu} - \mu_0) - r^2 = \mu\bar{\mu} - (\mu_0\bar{\mu} + \mu_0\mu) + (\mu_0^2 - r^2).$$

Dividing by $\mu\bar{\mu}(\mu_0^2 - r^2)$, we have

$$(\mu\bar{\mu})^{-1} - \left(\frac{\mu_0}{\mu_0^2 - r^2}\mu^{-1} + \frac{\mu_0}{\mu_0^2 - r^2}\bar{\mu}^{-1}\right) = -\frac{1}{\mu_0^2 - r^2},$$

equivalently,

$$\left(\mu^{-1} - \frac{\mu_0}{\mu_0^2 - r^2}\right)\left(\bar{\mu}^{-1} - \frac{\mu_0}{\mu_0^2 - r^2}\right) = \frac{\mu_0^2}{(\mu_0^2 - r^2)^2} - \frac{1}{\mu_0^2 - r^2} = \frac{r^2}{(\mu_0^2 - r^2)^2}.$$

Applying this to the circles $\mathcal{E}_u$ and $\mathcal{E}_v$, we see that

$$\mathcal{E}_u^{-1} = \{1/\mu : \mu \in \mathcal{E}_u\} = \{-1/2 + 1/2\sqrt{1+u^2}\, e^{i\theta} : t \in [0, 2\pi)\},$$
$$\mathcal{E}_v^{-1} = \{1/\mu : \mu \in \mathcal{E}_v\} = \{-1/2 + 1/2\sqrt{1+v^2}\, e^{i\theta} : t \in [0, 2\pi)\}.$$

Since $c - id \in \mathcal{E}_v$ is nonzero, (4.8) yields

$$\text{(4.9)} \qquad \frac{1}{c - id} = \frac{1}{\xi - 1} + \frac{\xi + 1}{(\xi - 1)(a - ib)} \in \tilde{\mathcal{E}}_u \cap \mathcal{E}_v^{-1},$$

where

$$\begin{aligned}\tilde{\mathcal{E}}_u &= \left\{\frac{1}{\xi - 1} + \frac{(\xi+1)}{2(\xi-1)}(-1 + \sqrt{1+u^2}\, e^{i\theta}) : \theta \in [0, 2\pi)\right\} \\ &= \left\{-\frac{1}{2} + \frac{(\xi+1)}{2(\xi-1)}\sqrt{1+u^2}\, e^{i\theta} : \theta \in [0, 2\pi)\right\}.\end{aligned}$$

By (4.7), $\tilde{\mathcal{E}}_u \cap \mathcal{E}_v^{-1} = \emptyset$, a contradiction to (4.9). Thus the proof in Case 1 is complete.

Case 2. Suppose conditions (i), (ii) and (iii) in Case 1 do not hold. Then $|w_1|\sqrt{1+v^2} = |w_2|\sqrt{1+u^2}$. If $m \in \mathbb{N}$, then $B_m = A_1 + E_{13}/m$ and $C_m = A_2 - E_{13}/m$ are commuting matrices in M_3 with (1,3)-entries $w_1 + 1/m$ and $w_2 - 1/m$, respectively. We are going to show that

$$\text{(4.10)} \qquad |w_1 + 1/m|\sqrt{1+v^2} = |w_2 - 1/m|\sqrt{1+u^2}$$

for at most one m.

Note that (4.10) holds if and only if

$$\begin{aligned}\text{(4.11)} \qquad &(mw_1 + 1)(m\overline{w}_1 + 1)(1 + v^2) = (mw_2 - 1)(m\overline{w}_2 - 1)(1 + u^2) \\ &\iff 2((\operatorname{Re} w_1)(1+v^2) + (\operatorname{Re} w_2)(1+u^2))m + (v^2 - u^2) = 0.\end{aligned}$$

If (4.11) holds for more than one m, then $v^2 = u^2$ and $\operatorname{Re} w_1 = -\operatorname{Re} w_2$. Then it follows from $u, v \geq 0$ and $w_1 + uv + w_2 = 0$ in (4.1) that $uv = 0$ and (i) holds, a contradiction.

So there exists m_0 such that $|w_1 + 1/m|\sqrt{1+v^2} \neq |w_2 - 1/m|\sqrt{1+u^2}$ for all $m \geq m_0$. By Case 1, $\partial \operatorname{conv} W(B_m, C_m) \subseteq W(B_m, C_m)$. Now, every boundary point $(\mu_1, \mu_2) \in \operatorname{conv} W(A_1, A_2)$ is the limit of a sequence of points $\{(\mu_1(m), \mu_2(m)) : m \geq m_0\}$ with $(\mu_1(m), \mu_2(m)) \in \partial(\operatorname{conv} W(B_m, C_m)) \subseteq W(B_m, C_m)$. Note that $W(B_m, C_m) \to W(A_1, A_2)$ as $m \to \infty$ in the Hausdorff metric on compact subsets of $\mathbb{R}^2$. We have $(\mu_1, \mu_2) \in W(A_1, A_2)$. Hence, $\partial(\operatorname{conv} W(A_1, A_2)) \subseteq W(A_1, A_2)$. This finishes the proof in Case 2, and thus also the proof of Proposition 4.2. ∎

Let $\mu_1 \in W(A_1)$ and $W(\mu_1, A_2) = \{\mu : (\mu_1, \mu) \in W(A_1, A_2)\}$. Now, we know that $W(A_1, A_2)$ has convex boundary if $A_1, A_2 \in M_3$ commute. Therefore, to prove that $W(A_1, A_2)$ is convex, we only need to show that $W(\mu_1, A_2)$ is simply connected for every $\mu_1 \in W(A_1)$.

To prove the latter property, we will show that

$$W(\mu_1, A_2) = \{\mu : (\mu_1, \mu) \in \operatorname{conv} W(A_1, A_2)\}.$$

To this end, using linear combinations, unitary similarity and transposition of matrices, we find that the matrices A_1, A_2 in (4.1) can be transformed as

$$A_1 = E_{11} + aE_{12}, \ A_2 = \begin{pmatrix} -a \\ 1 \\ b \end{pmatrix} (0\ 1\ \xi) \quad \text{where } a > 0,\ b \geq 0,\ \xi \in \mathbb{C}. \tag{4.12}$$

To prove this, observe that if $w_1 = 0$, then $w_2 = -uv$ and we can replace A_2 with

$$I_3 - (A_1 + A_2) = \begin{pmatrix} 0 & -u & uv \\ 0 & 1 & -v \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -u \\ 1 \\ 0 \end{pmatrix} (0\ 1\ -v).$$

If $w_2 = 0$, replace (A_1, A_2) with (TA_2^tT, TA_1^tT), where $T = E_{13} + E_{22} + E_{31}$. We have

$$TA_2^tT = \begin{pmatrix} 1 & v & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad TA_1^tT = \begin{pmatrix} 0 & 0 & -uv \\ 0 & 0 & u \\ 0 & 0 & 1 \end{pmatrix}.$$

Then we can proceed as in the above case for $w_1 = 0$.

Suppose $w_1, w_2 \neq 0$. Let $a = \sqrt{u^2 + |w_1|^2}$ and $U = (1) \oplus \frac{1}{a}\left(\begin{smallmatrix} u & w_1 \\ \overline{w}_1 & -u \end{smallmatrix}\right)$ be unitary. Then

$$U^*A_1U = \begin{pmatrix} 1 & a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad U^*A_2U = \gamma \begin{pmatrix} 0 & -a & -ac \\ 0 & 1 & c \\ 0 & b & bc \end{pmatrix},$$

where $\gamma = -(\overline{w}_1 w_2)/a^2$, $b = (u - v\overline{w}_1)/w_2$ and $c = -u/\overline{w}_1$. Let $b = |b|e^{i\theta}$

and $D = \operatorname{diag}(1,1,e^{i\theta})$. Replace (A_1, A_2) with $\big(D^*U^*A_1UD, \frac{1}{\gamma}D^*U^*A_2UD\big)$. Direct calculation gives

$$DU^*A_1UD^* = \begin{pmatrix} 1 & a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \frac{1}{\gamma}DU^*A_2UD^* = \begin{pmatrix} -a \\ 1 \\ |b| \end{pmatrix} (0\ 1\ \xi),$$

where $\xi = ce^{i\theta}$. If $\xi = 0 = b$, then $A_1 + A_2 = \operatorname{diag}(1,1,0)$ is Hermitian. By Proposition 2.1(g), $W(A_1, A_1 + A_2)$ is convex and hence $W(A_1, A_2)$ is also convex. So, we now assume that $(b, \xi) \neq (0,0)$.

Recall that a set $\mathcal{S}$ in $\mathbb{R}^n$ or $\mathbb{C}^n$ is *star-shaped* with *star center* $s_0 \in \mathcal{S}$ if $ts_0 + (1-t)s \in \mathcal{S}$ for all $t \in [0,1]$ and $s \in \mathcal{S}$. We have the following.

Proposition 4.3. *Suppose that A_1 and A_2 are as in* (4.1). *For every $\mu_1 \in W(A_1)$, the set*

$$W(\mu_1, A_2) = \{\mu : (\mu_1, \mu) \in W(A_1, A_2)\}$$

is star-shaped. Consequently,

$$W(\mu_1, A_2) = \{\mu : (\mu_1, \mu) \in \operatorname{conv} W(A_1, A_2)\},$$

and $W(A_1, A_2)$ is convex.

Proof. Without loss of generality, we may assume that A_1 and A_2 are of the form (4.12). Suppose $\mu_1 \in W(A_1)$. We are going to show that $W(\mu_1, A_2)$ is star-shaped with star center $1 - \mu_1$.

Let $\nu \in \mathbb{C}^3$ be a unit vector such that $\boldsymbol{\nu}^* A_1 \boldsymbol{\nu} = \mu_1$. By replacing $\boldsymbol{\nu}$ with $\tilde{\boldsymbol{\nu}} = e^{i\theta}\boldsymbol{\nu}$ for some $\theta \in \mathbb{R}$, we may assume that the first entry of $\boldsymbol{\nu}$ is nonnegative. Let

$$S = \{(p_1, p_2e^{i\theta}, p_3e^{i\phi})^t : \theta, \phi \in [0, 2\pi),\ p_1, p_2, p_3 \geq 0,\ p_1^2 + p_2^2 + p_3^2 = 1\}.$$

If $\boldsymbol{\nu} = (0, p_2e^{i\theta}, p_3e^{i\phi})^t \in S$, we have $\mu_1 = \boldsymbol{\nu}^* A_1 \boldsymbol{\nu} = 0$. Moreover

$$\boldsymbol{\nu}^* A_2 \boldsymbol{\nu} \in W\left(\begin{pmatrix} 1 & \xi \\ b & b\xi \end{pmatrix}\right) \subseteq W(0, A_2).$$

Since $W\big(\big(\begin{smallmatrix} 1 & \xi \\ b & b\xi \end{smallmatrix}\big)\big)$ is convex, and it contains the point $\{1\}$, we can see that $t + (1-t)\boldsymbol{\nu}^* A_2 \boldsymbol{\nu} \in W(0, A_2)$ for all $t \in [0,1]$. Now assume $\nu \in S$ with $\boldsymbol{\nu}^* A_1 \boldsymbol{\nu} = \mu_1$ and $p_1 > 0$. Then

$$\mu_1 = p_1^2 + ap_1p_2e^{i\theta}, \quad \text{i.e.,} \quad p_2e^{i\theta} = \frac{\mu_1 - p_1^2}{ap_1},$$

and

$$\begin{aligned} 1 - p_3^2 = p_1^2 + p_2^2 = p_1^2 + \left|\frac{\mu_1 - p_1^2}{ap_1}\right|^2 &= \frac{a^2p_1^4 + |\mu_1 - p_1^2|^2}{a^2p_1^2} \\ &= \frac{(a^2+1)p_1^4 + |\mu_1|^2 - 2(\operatorname{Re}\mu_1)p_1^2}{a^2p_1^2}. \end{aligned}$$

Therefore, we have

$$-a^2p_1^2p_3^2 = (a^2+1)p_1^4 - (2\,\mathrm{Re}\,\mu_1 + a^2)p_1^2 + |\mu_1|^2 \le 0. \tag{4.13}$$

By the above calculation, $\boldsymbol{\nu} \in S$ with positive first entry and $\boldsymbol{\nu}^*A_1\boldsymbol{\nu} = \mu_1$ if and only if $\boldsymbol{\nu} = (p_1, (\mu_1/p_1 - p_1)/a, p_3e^{i\phi})$ for $p_1 > 0$ satisfying (4.13), $\phi \in [0, 2\pi)$ and $p_3 = \sqrt{1 - p_1^2 - |(\mu_1/p_1 - p_1)/a|^2}$. Now

$$\begin{aligned}
\boldsymbol{\nu}^*A_2\boldsymbol{\nu} &= \begin{pmatrix} p_1 & (\bar\mu_1/p_1 - p_1)/a & p_3e^{-i\phi}\end{pmatrix} \begin{pmatrix} 0 & -a & -a\xi \\ 0 & 1 & \xi \\ 0 & b & b\xi \end{pmatrix} \begin{pmatrix} p_1 \\ (\mu_1/p_1 - p_1)/a \\ p_3e^{i\phi} \end{pmatrix} \\
&= (-ap_1 + (\bar\mu_1/p_1 - p_1)/a + bp_3e^{-i\phi})((\mu_1/p_1 - p_1)/a + \xi p_3e^{i\phi}) \\
&= p_1^2 - \mu_1 + |\mu_1/p_1 - p_1|^2/a^2 + b\xi p_3^2 \\
&\quad + p_3\big\{(-ap_1 + (\bar\mu_1/p_1 - p_1)/a)\xi e^{i\phi} + (\mu_1/p_1 - p_1)(b/a)e^{-i\phi}\big\} \\
&= 1 - \mu_1 + (b\xi - 1)p_3^2 \\
&\quad + p_3\big\{(-ap_1 + (\bar\mu_1/p_1 - p_1)/a)\xi e^{i\phi} + (\mu_1/p_1 - p_1)(b/a)e^{-i\phi}\big\}.
\end{aligned}$$

For a fixed $p_1 > 0$, if we let ϕ vary in $[0, 2\pi)$, we see that $\boldsymbol{\nu}^*A_2\boldsymbol{\nu}$ generates all the points of an ellipse denoted by $\mathcal{E}(p_1)$. Hence, $\mathcal{E}(p_1) \subseteq W(\mu_1, A_2)$. For a fixed $\mu_1 \in W(A_1)$, let p_u and p_ℓ be the largest and smallest nonnegative values of p_1 respectively for which the inequality

$$(a^2+1)p_1^4 - (2\,\mathrm{Re}\,\mu_1 + a^2)p_1^2 + |\mu_1|^2 \le 0$$

in (4.13) is satisfied. Then

$$W(\mu_1, A_2) = \bigcup_{p\in[p_\ell, p_u]} \mathcal{E}(p).$$

Here we denote $\mathcal{E}(0) = W\big(\big(\begin{smallmatrix} 1 & \xi \\ b & b\xi \end{smallmatrix}\big)\big)$. Next we show that every point inside the ellipse $\mathcal{E}(p)$ also lies in $W(\mu_1, A_2)$. As $\mu_1 \in W(A_1) = W(A_0)$ with $A_0 = \big(\begin{smallmatrix} 1 & a \\ 0 & 0 \end{smallmatrix}\big)$, there is a unit vector $\tilde{\boldsymbol{\nu}} = (\tilde p, \nu_2) \in \mathbb{C}^2$ with $\tilde p \ge 0$ such that $\tilde{\boldsymbol{\nu}}^*A_0\tilde{\boldsymbol{\nu}} = \mu_1$. Thus, with $\boldsymbol{\nu} = (\tilde p, \nu_2, 0) \in \mathbb{C}^3$ we have $\boldsymbol{\nu}^*A_1\boldsymbol{\nu} = \mu_1$. The corresponding ellipse $\mathcal{E}(\tilde p) = \{1 - \mu_1\}$ is a singleton as $p_3 = 0$. For every $p_1 \in [p_\ell, p_u]$, we may let p_1 change continuously to $\tilde p$. Recall that $\boldsymbol{\nu} = (p_1, (\mu_1/p_1 - p_1)/a, p_3e^{i\phi})$. As the entries of $\boldsymbol{\nu}$ are continuous functions in $p_1 > 0$, the ellipse $\mathcal{E}(p_1)$ will deform continuously to the singleton $\mathcal{E}(\tilde p)$ in the set $W(\mu_1, A_2)$. Hence, by continuity all the points inside the ellipse $\mathcal{E}(p_1)$ also lie in $W(\mu_1, A_2)$, i.e.,

$$W(\mu_1, A_2) = \bigcup_{p\in[p_\ell, p_u]} \mathcal{E}(p) = \bigcup_{p\in[p_\ell, p_u]} \bar{\mathcal{E}}(p), \tag{4.14}$$

where $\overline{\mathcal{E}}(p)$ is the elliptical disk with $\mathcal{E}(p)$ as boundary.

We will show that $\bigcup_{p\in[p_\ell,p_u]}\bar{\mathcal{E}}(p)$ is star-shaped with star center $1-\mu_1$. Expressing p_3 as a function of p_1 from (4.13), we see that p_3 attains the maximum value

$$\hat{p}_3=\sqrt{1-p_1^2-|(\mu_1/p_1-p_1)/a|^2}=\frac{\sqrt{a^2+2(\operatorname{Re}\mu_1-\sqrt{1+a^2}|\mu_1|)}}{a}$$

when $p_1=\hat{p}=\sqrt{|\mu_1|/\sqrt{1+a^2}}$. In general, for each choice of $p_3\in[0,\hat{p}_3]$, there are $p_1^-\in[p_\ell,\hat{p}]$ and $p_1^+\in[\hat{p},p_u]$ satisfying the equality in (4.13). For every $0<r<1$ and $p_3\in[0,\hat{p}_3]$, set $\tilde{p}_3=rp_3$ and let $\tilde{p}_1^-\in[p_\ell,\hat{p}]$ and $\tilde{p}_1^+\in[\hat{p},p_u]$ satisfy (4.13) for p_3. With some intricate arguments presented in the Appendix, we will show that

(I) If $|\xi|^2(1+a^2)\geq b^2$, then $\overline{\mathcal{E}}(p_1^-)\subseteq\overline{\mathcal{E}}(p_1^+)$, and for every $\mu_2\in\overline{\mathcal{E}}(p_1^+)$, $(1-r^2)(1-\mu_1)+r^2\mu_2\in\overline{\mathcal{E}}(\tilde{p}_1^+)$.

(II) If $|\xi|^2(1+a^2)\leq b^2$, then $\overline{\mathcal{E}}(p_1^+)\subseteq\overline{\mathcal{E}}(p_1^-)$, and for every $\mu_2\in\overline{\mathcal{E}}(p_1^-)$, $(1-r^2)(1-\mu_1)+r^2\mu_2\in\overline{\mathcal{E}}(\tilde{p}_1^-)$.

Once (I) and (II) are proved, by (4.14) we see that $W(\mu_1,A_2)$ is star-shaped with $1-\mu_1$ as a star center, i.e., for any $\mu_2\in W(\mu_1,A_2)$ and $t\in[0,1]$,

$$t\mu_2+(1-t)(1-\mu_1)\in W(\mu_1,A_2).$$

Let $S=\{\mu:(\mu_1,\mu)\in\operatorname{conv}W(A_1,A_2)\}$. We have $W(\mu_1,A_2)\subseteq S$. Note that $S\subseteq\mathbb{C}$ is convex and compact. By Proposition 4.2,

$$\begin{aligned}\partial S&\subseteq\{\mu:(\mu_1,\mu)\in\partial\left(\operatorname{conv}W(A_1,A_2)\right)\}\\&\subseteq\{\mu:(\mu_1,\mu)\in W(A_1,A_2)\}=W(\mu_1,A_2).\end{aligned}$$

The star-shapedness of $W(\mu_1,A_2)$ implies that this set is simply connected. Therefore, $S\subseteq W(\mu_1,A_2)$. Hence, $S=W(\mu_1,A_2)$.

Now, we can show that $W(A_1,A_2)$ is convex as follows. Suppose (x_1,y_1), $(x_2,y_2)\in W(A_1,A_2)$, $t\in[0,1]$ and $(\mu_1,\mu_2)=t(x_1,y_1)+(1-t)(x_2,y_2)$. Then $(\mu_1,\mu_2)\in\operatorname{conv}W(A_1,A_2)$. We have $\mu_2\in\{\mu:(\mu_1,\mu)\in\operatorname{conv}W(A_1,A_2)\}=W(\mu_1,A_2)$. Thus, $(\mu_1,\mu_2)\in W(A_1,A_2)$. So, $W(A_1,A_2)$ is convex. ∎

4.2. span $\{I_3,A_1,A_2\}\subseteq M_3$ **contains a nonzero nilpotent.** Here we present the proof of Theorem 3.3 when span $\{I_3,A_1,A_2\}$ contains a nonzero nilpotent matrix. We may assume that $\{I_3,A_1,A_2\}$ is linearly independent and A_1 is nilpotent.

Similar to the case considered in Subsection 4.1, we can apply linear combinations and unitary similarity transforms to change A_1,A_2 to a simpler form. First, we show that one may assume that A_1 is rank 1. Suppose A_1 is rank 2. Then there is an invertible S such that $S^{-1}A_1S=J$ is the upper triangular Jordan block. Then $A_1A_2=A_2A_1$ implies that $S^{-1}A_2S=aI_3+bJ+cJ^2$. We may replace A_2 by $A_2-aI_3-bA_1$. Then A_2 is a rank 1 nilpotent. We may then interchange the roles of A_1 and A_2. Now, A_1 is a rank 1 nilpotent matrix in $\operatorname{span}\{I_3, A_1, A_2\}$. So, up to a nonzero multiple and a unitary similarity transform, we may assume that $A_1 = E_{13}$, where as before $\{E_{ij} : i, j = 1, 2, 3\}$ is the standard basis of M_3. The condition $A_1A_2 = A_2A_1$ implies that A_2 is in upper triangular form with the $(1,1)$-entry equal to the $(3,3)$-entry. We may then replace A_2 by $A_2 - \gamma_1 I_3 - \gamma_2 A_1$ and assume that

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & b & 0 \\ 0 & a & c \\ 0 & 0 & 0 \end{pmatrix}.$$

If necessary, we may also replace (A_1, A_2) with (DA_1^tD, DA_2^tD), where $D = E_{13} + E_{22} + E_{31}$, and assume that $|b| \geq |c|$.

If $b = 0$, then we may assume that $A_2 = E_{22}$. By Proposition 2.1(g, e),

$$W(A_1, A_2) \cong W\left(\frac{(E_{13} + E_{31})}{2}, \frac{i(E_{13} - E_{31})}{2}, E_{22}\right)$$

is convex.

If $b \neq 0$, let $\zeta = |a/b|$ and $\xi = |c/b|$. Suppose $a/b = \zeta e^{i\theta}$ and $c/b = \xi e^{i\phi}$, $\theta, \phi \in [0, 2\pi)$. Let $U = \operatorname{diag}\left(1, e^{i\theta}, e^{i(2\theta-\phi)}\right)$. Replacing (A_1, A_2) with $(e^{i(\phi-2\theta)}U^*A_1U, e^{-i\theta}U^*A_2U/b)$, we have $(A_1, A_2) = (E_{13}, \zeta E_{22} + E_{12} + \xi E_{23})$, where $\zeta \geq 0$ and $\xi \in [0, 1]$.

Let $P_m = E_{11}/m$ and $Q_m = (E_{22} - E_{32})/m$ for $m \in \mathbb{N}$. Then

$$A_1 + P_m = \begin{pmatrix} 1/m & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 + Q_m = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \zeta + 1/m & \xi \\ 0 & -1/m & 0 \end{pmatrix}$$

commute. Moreover,

$$aI_3 + b(A_1 + P_m) + c(A_2 + Q_m) = \begin{pmatrix} a + b/m & c & b \\ 0 & a + c(\zeta + 1/m) & c\xi \\ 0 & -c/m & a \end{pmatrix}$$

is nilpotent if and only if

$$a + b/m = 0,\ 2a + c(\zeta + 1/m) = 0 \quad \text{and} \quad a^2 + ac(\zeta + 1/m) + c^2\xi/m = 0.$$

From the last two equations, if $\zeta + 1/m \neq 0$, then

$$\frac{a}{c} = \frac{-(\zeta + 1/m)}{2}, \qquad 0 = \left(\frac{a}{c}\right)^2 + \frac{a}{c}\left(\zeta + \frac{1}{m}\right) + \frac{\xi}{m} = \frac{\xi}{m} - \frac{1}{4}\left(\zeta + \frac{1}{m}\right)^2,$$

which can be true for at most two choices of m. Hence, except for finitely many values of m, the linear span of the set $\{I_3, A_1 + P_m, A_2 + Q_m\}$ contains no nonzero nilpotent and $W(A_1+P_m, A_2+Q_m)$ is convex by Proposition 4.3 in Subsection 4.1.

Suppose L is the line segment joining $(x^*A_1x, x^*A_2x), (y^*A_1y, y^*A_2y) \in W(A_1, A_2)$. Let L_m be the line segment joining $(x^*(A_1+P_m)x, x^*(A_2+Q_m)x)$ and $(y^*(A_1+P_m)y, y^*(A_2+Q_m)y)$. Clearly, the endpoints of the line segments L_m converge to those of L. Thus, $L_m \to L$ in the Hausdorff metric as $m \to \infty$. Note that $L_m \subseteq W(A_1 + P_m, A_2 + Q_m)$ because $W(A_1 + P_m, A_2 + Q_m)$ is convex by Proposition 4.3. Since $W(A_1 + B_m, A_2 + Q_m) \to W(A_1, A_2)$ in the Hausdorff metric as $m \to \infty$, we infer that $L_m \to L$ as $m \to \infty$, so that $L \subseteq W(A_1, A_2)$, and therefore $W(A_1, A_2)$ is convex. ∎

Appendix: Proof of (I) and (II). We use the notation introduced in Section 4.2. For every $q \in [p_\ell, p_u]$, let

$$C_q = \begin{pmatrix} 0 & \xi((\bar{\mu}_1/q - q)/a - aq) \\ b(\mu_1/q - q)/a & 0 \end{pmatrix}.$$

If $q \in [p_\ell, p_u]$ and $q_3^2 = 1 - q^2 - |(\mu_1/q - q)/a|^2$, then

$$\bar{\mathcal{E}}(q) = 1 - \mu_1 + (b\xi - 1)q_3^2 + 2q_3 W(C_q).$$

It is clear that $W(C_{p_1^-}) \subseteq W(C_{p_1^+})$ if and only if $\bar{\mathcal{E}}(p_1^-) \subseteq \bar{\mathcal{E}}(p_1^+)$. For every $0 < r < 1$ and $\mu_2 \in \overline{\mathcal{E}}(p_1^+)$, we have

$$(1 - r^2)(1 - \mu_1) + r^2\mu_2 \in 1 - \mu_1 + (b\xi - 1)(rq_3)^2 + 2(rq_3)W(rC_{p_1^+}).$$

Let $\tilde{p}_3 = rq_3$. Thus, to prove (I), it suffices to show that

$$W(rC_{p_1^-}) \subseteq W(rC_{p_1^+}) \subseteq W(C_{\tilde{p}_1^+}), \tag{A.1}$$

By Proposition 2.2, the inclusions (A.1) are equivalent to

$$r\lambda_1(e^{i\theta}C_{p_1^-} + e^{-i\theta}C^*_{p_1^-}) \le r\lambda_1(e^{i\theta}C_{p_1^+} + e^{-i\theta}C^*_{p_1^+}) \le \lambda_1(e^{i\theta}C_{\tilde{p}_1^+} + e^{-i\theta}C^*_{\tilde{p}_1^+}),$$

for every $\theta \in [0, 2\pi)$.

Note that

$$\lambda_1(e^{i\theta}C_q + e^{-i\theta}C_q^*) = \sqrt{|\det(e^{i\theta}C_q + e^{-i\theta}C_q^*)|}.$$

Hence, it suffices to show that for every $\theta \in [0, 2\pi)$,

$$\begin{aligned} r^2|\det(e^{i\theta}C_{p_1^-} + e^{-i\theta}C^*_{p_1^-})| &\le r^2|\det(e^{i\theta}C_{p_1^+} + e^{-i\theta}C^*_{p_1^+})| \\ &\le |\det(e^{i\theta}C_{\tilde{p}_1^+} + e^{-i\theta}C^*_{\tilde{p}_1^+})|. \end{aligned} \tag{A.2}$$

For every $q \in [p_\ell, p_u]$ and $q_3^2 = 1 - q^2 - |(\mu_1/q - q)/a|^2$, we have

$$\begin{aligned} &|\det(e^{i\theta}C_q + e^{-i\theta}C_q^*)| \\ &= |e^{i\theta}\xi(-aq + (\bar{\mu}_1/q - q)/a) + e^{-i\theta}b(\bar{\mu}_1/q - q)/a|^2 \end{aligned}$$

$$
\begin{aligned}
&= |\xi|^2|(\bar\mu_1/q-q)/a-aq|^2+b^2|(\bar\mu_1/q-q)/a|^2\\
&\quad +2\operatorname{Re}(e^{2i\theta}\xi b(-aq+(\bar\mu_1/q-q)/a)(\mu_1/q-q)/a)\\
&= |\xi|^2(|(\bar\mu_1/q-q)/a|^2+a^2q^2-2\operatorname{Re}(\bar\mu_1-q^2))+b^2|(\bar\mu_1/q-q)/a|^2\\
&\quad +2\operatorname{Re}(e^{2i\theta}\xi b(-aq+(\bar\mu_1/q-q)/a)(\mu_1/q-q)/a)\\
&= (|\xi|^2(1+a^2)-b^2)q^2+(|\xi|^2+b^2)(1-q_3^2)\\
&\quad -2\operatorname{Re}(|\xi|^2\bar\mu_1+e^{2i\theta}\xi b(1-\mu_1-q_3^2)).
\end{aligned}
$$

As

$$1-(p_1^-)^2-|(\mu_1/p_1^- - p_1^-)/a|^2 = 1-(p_1^+)^2-|(\mu_1/p_1^+ - p_1^+)/a|^2 = p_3^2,$$

the first inequality in (A.2) follows from $|\xi|^2(1+a^2)-b^2\ge 0$ and $p_1^+\ge p_1^-$. Now

$$
\begin{aligned}
&\det|e^{i\theta}C_{\tilde p_1^+}+e^{-i\theta}C^*_{\tilde p_1^+}|-r^2|\det(e^{i\theta}C_{p_1^+}+e^{-i\theta}C^*_{p_1^+})|\\
&= (|\xi|^2(1+a^2)-b^2)((\tilde p_1^+)^2-r^2(p_1^+)^2)+(1-r^2)(|\xi|^2+b^2)\\
&\quad -2(1-r^2)\operatorname{Re}(|\xi|^2\bar\mu_1+e^{2i\theta}\xi b(1-\mu_1))\\
&\ge (|\xi|^2(1+a^2)-b^2)(\tilde p_1^+)^2+(|\xi|^2+b^2)-2(|\xi|^2\operatorname{Re}\bar\mu_1+|\xi b(1-\bar\mu_1)|)\\
&\quad -r^2\big((|\xi|^2(1+a^2)-b^2)(\tilde p_1^+)^2+(|\xi|^2+b^2)-2(|\xi|^2\operatorname{Re}\bar\mu_1+|\xi b(1-\bar\mu_1)|)\big).
\end{aligned}
$$

For every $y\in[0,\hat p_3^2]$, let

$$(q_y^+)^2=\frac{2\operatorname{Re}\mu_1+a^2(1-y)+\sqrt{(2\operatorname{Re}\mu_1+a^2(1-y))^2-4(a^2+1)|\mu_1|^2}}{2(1+a^2)}.$$

It is not hard to see that $q_y^+\in[\hat p,p_u]$ satisfies the left-hand equality of (4.13) with $p_3=\sqrt{y}$, i.e.,

$$-a^2(q_y^+)^2y=(a^2+1)(q_y^+)^4-(2\operatorname{Re}\mu_1+a^2)(q_y^+)^2+|\mu_1|^2.$$

Define the function $M:[0,\hat p_3^2]\to\mathbb{R}$ by

$$
\begin{aligned}
M(y)=(|\xi|^2(1+a^2)-b^2)(q_y^+)^2+(|\xi|^2+b^2)\\
-2(|\xi|^2\operatorname{Re}\bar\mu_1+|\xi b(1-\bar\mu_1)|).
\end{aligned}
$$

For $y=0$, we have $(1+a^2)(q_0^+)^4-(2\operatorname{Re}\mu_1+a^2)(q_0^+)^2+|\mu_1|^2=0$ and

$$M(0)=\frac{|1-\mu_1|^2|\xi|^2}{1-(q_0^+)^2}-2b|\xi|\,|1-\bar\mu_1|+b^2(1-(q_0^+)^2)\ge 0.$$

We will show that M is concave so that

$$
\begin{aligned}
&|\det(e^{i\theta}C_{\tilde p_1^+}+e^{-i\theta}C^*_{\tilde p_1^+})|-r^2|\det(e^{i\theta}C_{p_1^+}+e^{-i\theta}C^*_{p_1^+})|\\
&\qquad\qquad \ge M(r^2p_3^2)-r^2M(p_3^2)\ge(1-r^2)M(0)\ge 0.
\end{aligned}
$$

Noting that $|\xi|^2(1+a^2)-b^2\ge 0$, we have

$$\begin{aligned}\frac{d^2M}{dy^2} &= (|\xi|^2(1+a^2)-b^2)((q_y^+)^2)'' \\ &= \frac{|\xi|^2(1+a^2)-b^2}{2(a^2+1)}\Big(\sqrt{(2\operatorname{Re}\mu_1+a^2(1-y))^2-4(1+a^2)|\mu_1|^2}\Big)'' \\ &= \frac{-\big(|\xi|^2(1+a^2)-b^2\big)\big(4a^4\left(a^2+1\right)\big)|\mu_1|^2}{2(a^2+1)\big((2\operatorname{Re}\mu_1+a^2\left(1-y^2\right))^2-4(a^2+1)|\mu_1|^2\big)^{3/2}}\le 0.\end{aligned}$$

Hence M is concave.

The proof of (II) is similar, and we just give a sketch. It suffices to show that for every $\theta\in[0,2\pi)$,

$$\begin{aligned} r^2|\det(e^{i\theta}C_{p_1^+}+e^{-i\theta}C^*_{p_1^+})| &\le r^2|\det(e^{i\theta}C_{p_1^-}+e^{-i\theta}C^*_{p_1^-})| \\ &\le |\det(e^{i\theta}C_{\tilde p_1^-}+e^{-i\theta}C^*_{\tilde p_1^-})|. \end{aligned} \tag{A.3}$$

Recall that

$$\begin{aligned}|\det(e^{i\theta}C_q+e^{-i\theta}C_q^*)| &= (|\xi|^2(1+a^2)-b^2)q^2+(|\xi|^2+b^2)(1-q_3^2) \\ &\quad -2\operatorname{Re}\big(|\xi|^2\bar\mu_1+e^{2i\theta}\xi b(1-\mu_1-q_3^2)\big).\end{aligned}$$

So, the first inequality in (A.3) follows from $|\xi|^2(1+a^2)\le b^2$ and $p_1^-\le p_1^+$. The second inequality will follow from the concavity of

$\tilde M(y) = (|\xi|^2(1+a^2)-b^2)(q_y^-)^2+(|\xi|^2+b^2)-2(|\xi|^2\operatorname{Re}\bar\mu_1+|\xi b(1-\bar\mu_1)|)$,

where

$$(q_y^-)^2=\frac{2\operatorname{Re}\mu_1+a^2(1-y)-\sqrt{(2\operatorname{Re}\mu_1+a^2(1-y))^2-4(a^2+1)|\mu_1|^2}}{2(1+a^2)}.$$

Since $|\xi|^2(1+a^2)-b^2\le 0$, we have

$$\begin{aligned}\frac{d^2\tilde M}{dy^2} &= (|\xi|^2(1+a^2)-b^2)((q_y^-)^2)'' \\ &= \frac{|\xi|^2(1+a^2)-b^2}{2(a^2+1)}\Big(-\sqrt{(2\operatorname{Re}\mu_1+a^2(1-y))^2-4(1+a^2)|\mu_1|^2}\Big)''\le 0.\end{aligned}$$

Thus (II) holds. ∎

REMARK A.1. It is worth pointing out that our proofs use some continuity arguments and a simple idea of homotopy (in deforming ellipses inside the numerical range of a certain matrix). In particular, intricate linear-algebraic arguments are used. It would be nice if a less computational proof could be found.

Acknowledgments. The authors would like to thank the referee for some helpful comments leading to an improvement of the presentation of the paper.

Li is an affiliate member of the Institute for Quantum Computing. His research was partially supported by the Simons Foundation grant 851334.

References

[1] Y. H. Au-Yeung and Y. T. Poon, *A remark on the convexity and positive definiteness concerning Hermitian matrices*, Southeast Asian Bull. Math. 3 (1979), 85–92.

[2] H. Baklouti, K. Feki and O. A. M. Sid Ahmed, *Joint numerical ranges of operators in semi-Hilbertian spaces*, Linear Algebra Appl. 555 (2018), 266–284.

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

[4] V. Bolotnikov and L. Rodman, *Normal forms and joint numerical ranges of doubly commuting matrices*, Linear Algebra Appl. 301 (1999), 187–194.

[5] A. T. Dash, *Joint numerical range*, Glasnik Mat. Ser. III 7(27) (1972), 75–81.

[6] A. T. Dash, *Tensor products and joint numerical range*, Proc. Amer. Math. Soc. 40 (1973), 521–526.

[7] C. J. Eliezer, *A note on group commutators of* 2×2 *matrices*, Amer. Math. Monthly 75 (1968), 1090–1091.

[8] E. Gutkin, E. A. Jonckheere and M. Karow, *Convexity of the joint numerical range: topological and differential geometric viewpoints*, Linear Algebra Appl. 376 (2004), 143–171.

[9] P. R. Halmos, *A Hilbert Space Problem Book*, 2nd ed., Grad. Texts in Math. 19, Springer, New York, 1982.

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

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

[12] C. K. Li, Y. T. Poon and Y. Wang, *Joint numerical ranges and commutativity of matrices*, J. Math. Anal. Appl. 491 (2020), no. 1, art. 124310, 12 pp.

[13] V. Müller, *Joint numerical range of commuting tuples is not convex*, preprint 64-2019, Inst. Math., Czech Acad. Sci., 2019; http://www.math.cas.cz/fichier/preprints/IM_20191029093220_33.pdf.

[14] V. Müller and Yu. Tomilov, *Joint numerical ranges: recent advances and applications* (minicourse), Concr. Oper. 7 (2020), 133–154.

[15] V. Müller and Yu. Tomilov, *In search of convexity: diagonals and numerical ranges*, Bull. London Math. Soc. 53 (2021), 1016–1029.

[16] M. Newman, *Two classical theorems on commuting matrices*, J. Res. Nat. Bur. Standards Sect. B 71B (1967), 69–71.

Pan-Shun Lau
Department of Mathematics
University of Nevada Reno
Reno, NV 89557, USA
E-mail: plau@unr.edu

Chi-Kwong Li
Department of Mathematics
College of William & Mary
Williamsburg, VA 23185, USA
E-mail: ckli@math.wm.edu

Yiu-Tung Poon
Department of Mathematics
Iowa State University
Ames, IA 50011, USA
E-mail: ytpoon@iastate.edu