SPECTRA, NORMS AND NUMERICAL RANGES OF GENERALIZED QUADRATIC OPERATORS

CHI-KWONG LI, YIU-TUNG POON, AND MASARU TOMINAGA

Abstract. A bounded linear operator acting on a Hilbert space is a generalized quadratic operator if it has an operator matrix of the form

$$\begin{bmatrix} aI & cT \\ dT^* & bI \end{bmatrix}.$$

It reduces to a quadratic operator if $d = 0$. In this paper, spectra, norms, and various kinds of numerical ranges of generalized quadratic operators are determined. Some operator inequalities are also obtained. In particular, it is shown that for a given generalized quadratic operator, the rank-k numerical range, the essential numerical range, and the q-numerical range are elliptical disks; the c-numerical range is a sum of elliptical disks. The Davis-Wielandt shell is the convex hull of a family of ellipsoids unless the underlying Hilbert space has dimension 2.

1. Introduction

Let $\mathcal{B}(\mathcal{H})$ be the algebra of bounded linear operators acting on a Hilbert space $\mathcal{H}$. We identify $\mathcal{B}(\mathcal{H})$ with M_n if $\mathcal{H}$ has dimension n. An operator $A \in \mathcal{B}(\mathcal{H})$ is a *generalized quadratic operators* if it has an operator matrix of the form

$$\begin{bmatrix} aI & cT \\ dT^* & bI \end{bmatrix}, \tag{1.1}$$

where T is an operator from $\mathcal{H}_2$ to $\mathcal{H}_1$, and a, b, c, d are complex numbers. [In the following discussion, we will not distinguish the operator and its operator matrix if there is no ambiguity.] When $d = 0$, such an operator A satisfies condition

$$(aI - A)(bI - A) = 0 \tag{1.2}$$

and is known as a *quadratic operator*. In fact, it is known that an operator A satisfies (1.2) if and only if it has an operator matrix of the form (1.1) with $d = 0$, by a suitable choice of orthonormal basis.

Motivated by theory and applied problems, there has been considerable interest in studying the norms and generalized numerical ranges (see the definition in later sections) of operators of the form (1.1) under the additional assumptions that (i) a, b, c, d are nonnegative, or (ii) $d = 0$; see [3, 13, 22, 27] and the references therein. In this paper, a complete description is given to the spectrum, the norm, and various types of generalized numerical ranges of an operator of the

2000 *Mathematics Subject Classification.* Primary 47A12, 15A60.

Key words and phrases. Spectra, norms, elliptical ranges, generalized numerical ranges, Davis-Wielandt shell.

Research of the first two authors were supported by USA NSF. The first author was also supported by the William and Mary Plumeri Award. He is an honorary professor of the University of Hong Kong and an honorary professor of the Taiyuan University of Technology.

form (1.1). In particular, the spectrum is a union of the spectrum of certain 2×2 matrices of the form

$$A_p = \begin{bmatrix} a & cp \\ dp & b \end{bmatrix}, \qquad p \geq 0.$$

Also the norm of A is the same as that of A_p with $p = \|T\|$; the closure of the numerical range (and also many generalized numerical ranges) is always an elliptical disk. Since quadratic operators have been studied in [22], **we always assume that** $cdT \neq 0$ in the following discussion.

Our paper is organized as follows. In Section 2, we obtain a different operator matrix for an generalized quadratic operator A. We then use the result to give a description of $\sigma(A)$, which is the spectrum of $A \in \mathcal{B}(\mathcal{H})$. In Section 3, we determine the numerical range, the matricial range, and the norm of generalized quadratic operators. Furthermore, we obtain some operator inequalities concerning generalized quadratic operators that extend some results of Furuta [13] and Garcia [14]. We then give the description of various generalized numerical ranges of A in Section 4. The results cover those in [3, 22, 27] and the reference therein. Additional remarks and further research are discussed in Section 5.

We will use the following notations in our discussion. For $S \subseteq \mathbb{C}$, we will use $\mathbf{int}(S)$, $\mathbf{cl}\,(S)$ and $\mathbf{conv}(S)$ to denote the relative interior, the closure and the convex hull of S, respectively. Note that in our discussion, it may happen that $S = \mathbf{conv}\{\mu_1, \mu_2\}$ is a line segment in $\mathbb{C}$ so that $\mathbf{int}(S) = S \setminus \{\mu_1, \mu_2\}$.

For $A \in \mathcal{B}(\mathcal{H})$, let $\ker A$ and range$A$ denote the null space and range space of A, respectively. Let $\mathcal{V}$ be a closed subspace of $\mathcal{H}$ and Q the embedding of $\mathcal{V}$ into $\mathcal{H}$. Then $B = Q^*AQ$ is the *compression* of A onto $\mathcal{V}$. More generally, A has a compression B if A has an operator matrix $\begin{bmatrix} B & * \\ * & * \end{bmatrix}$ with respect to an orthonormal basis; alternatively, there is a closed subspace $\mathcal{V}$ of $\mathcal{H}$ and $X : \mathcal{V} \to \mathcal{H}$ such that $X^*X = I_{\mathcal{V}}$ and $X^*AX = B$. Note that, in this case, $X(\mathcal{V})$ is closed and X^*AX is the compression of A on $X(\mathcal{V})$.

2. A different operator matrix representation and the spectrum

First, we obtain a different operator matrix for A of the form (1.1). The special form reduces to that of quadratic operators in [22, Theorem 1.1] if $d = 0$.

Theorem 2.1. *Let $A \in \mathcal{B}(\mathcal{H})$ be an operator with an operator matrix*

$$\begin{bmatrix} aI & cT \\ dT^* & bI \end{bmatrix} \qquad \text{with } cdT \neq 0.$$

Then $\mathcal{H}$ has a decomposition $\mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$ such that A has an operator matrix of the form

$$\begin{bmatrix} aI_r & cP \\ dP & bI_r \end{bmatrix} \oplus \gamma I_s \qquad \text{with } cdP \neq 0, \tag{2.1}$$

where $\gamma \in \{a, b\}$, $\dim \mathcal{H}_1 = r$, $\dim \mathcal{H}_2 = s$, and $P : \mathcal{H}_1 \to \mathcal{H}_1$ is a positive semidefinite operator, i.e., $\langle Px, x\rangle \geq 0$ for all $x \in \mathcal{H}_1$, with the additional condition that $\langle Px, x\rangle \neq 0$ for all nonzero

$x \in \mathcal{H}_1$ if $a = b$. Here I_s may be vacuous. Furthermore, we have decompositions $\mathcal{H}_1 = \oplus_{k=1}^n \mathcal{V}_k$ and $P = \oplus_{k=1}^n P_k$, where $n \le \infty$ and for each k, $P_k \in \mathcal{B}(\mathcal{V}_k)$ such that $\sigma(P_k)$ is an interval.

Proof. Given T on $\mathcal{H}_1$, there exist unitary operators U and V and a positive semidefinite operator P such that UTV is of one of the following form:

(a) P, if $\dim\ker T = \dim\ker T^*$.

(b) $\begin{bmatrix} P \\ 0 \end{bmatrix}$, if $\dim\ker T < \dim\ker T^*$.

(c) $\begin{bmatrix} P & 0 \end{bmatrix}$, if $\dim\ker T > \dim\ker T^*$.

Since $\ker T^* = (\mathrm{range}(T))^\perp$, (a) follows from the polar decomposition. (b) and (c) follow from (a). Applying this to (1.1), we have (2.1).

Let $\lambda_0 = \min\{\lambda : \lambda \in \sigma(P)\}$ and $\lambda_1 = \max\{\lambda : \lambda \in \sigma(P)\}$. Then $\lambda_0, \lambda_1 \in \sigma(P)$. If $\mu \notin \sigma(P)$ for some $\lambda_0 < \mu < \lambda_1$, then $P = Q \oplus R$, where $\sigma(Q) \subseteq [\lambda_0, \mu)$ and $\sigma(R) \subseteq (\mu, \lambda_1]$. Therefore, we can decompose P as $P = \oplus_{k=1}^\infty P_k$, where $\sigma(P_k)$ is an interval for each k. □

By the above theorem, we can focus on an operator A with an operator matrix of the form (2.1) with $cdP \neq 0$. In the following discussion, we will always identify the subspaces $\mathcal{H}_1 \oplus 0 \oplus 0$ and $0 \oplus \mathcal{H}_1 \oplus 0$ with $\mathcal{H}_1$. Also, the family of matrices

$$A_p = \begin{bmatrix} a & cp \\ dp & b \end{bmatrix}, \qquad p \ge 0, \tag{2.2}$$

will be very useful in our discussion.

Theorem 2.2. *Suppose $A \in \mathcal{B}(\mathcal{H})$ has an operator matrix of the form (2.1) as in Theorem 2.1 and A_p is defined as in (2.2). Then*

$$\sigma(A) = \cup_{p\in\sigma(P)} \sigma(A_p) \cup S = \cup_{p\in\sigma(P)} \left\{ \frac{1}{2} \left[(a+b) \pm \sqrt{(a-b)^2 + 4cdp^2} \right] \right\} \cup S\,,$$

where $S = \{\gamma\}$ if γI_s is non-trivial and $S = \emptyset$ otherwise.

Proof. For simplicity, we may assume that γI_s is vacuous. If $\alpha = \frac{(a+b)}{2}$ and $\beta = \frac{(a-b)}{2}$, then

$$A - \alpha I = \begin{bmatrix} \beta I & cP \\ dP & -\beta I \end{bmatrix}.$$

By direct computation, we have $(A - \alpha I)^2 = (\beta^2 I + cdP^2) \oplus (\beta^2 I + cdP^2)$ and

$$\begin{bmatrix} 0 & cI \\ -dI & 0 \end{bmatrix} (A - \alpha I) \begin{bmatrix} 0 & cI \\ -dI & 0 \end{bmatrix}^{-1} = -(A - \alpha I).$$

Thus, $\mu \in \sigma(A - \alpha I)$ if and only if $\mu^2 \in \sigma(\beta^2 I + cdP^2)$; $\mu \in \sigma(A - \alpha I)$ if and only if $-\mu \in \sigma(A - \alpha I)$. So the result follows. □

3. Numerical range, dilation, matricial range, and operator inequalities

Recall that the *numerical range* of $A \in \mathcal{B}(\mathcal{H})$ is defined by

$$W(A) = \{\langle Ax, x\rangle : x \in \mathcal{H}, \langle x, x\rangle = 1\};$$

see [15, 16, 17]. The numerical range is useful in studying matrices and operators. One of the basic properties of the numerical range is that $W(A)$ is always convex; for example, see [16]. In particular, we have the following result, e.g., see [18] and [17, Theorem 1.3.6].

Elliptical Range Theorem *If $A \in M_2$ has eigenvalues μ_1 and μ_2, then $W(A)$ is an elliptical disk with μ_1, μ_2 as foci and $\sqrt{\operatorname{tr}(A^*A) - |\mu_1|^2 - |\mu_2|^2}$ as the length of minor axis. Furthermore, if $\hat{A} = A - (\operatorname{tr} A)I/2$, then the lengths of minor and major axis of $W(A)$ are, respectively,*

$$\{\operatorname{tr}(\hat{A}^*\hat{A}) - 2|\det \hat{A}|\}^{1/2} \quad \text{and} \quad \{\operatorname{tr}(\hat{A}^*\hat{A}) + 2|\det \hat{A}|\}^{1/2}.$$

Using this theorem, one can deduce the convexity of the numerical range of a general operator; e.g., see [18]. It turns out that for an operator A of the form (2.1), $W(A)$ is also an elliptical disk with all the boundary points, two boundary points, or none of its boundary points as shown in the following.

Theorem 3.1. *Suppose $A \in \mathcal{B}(\mathcal{H})$ has an operator matrix of the form (2.1) described in Theorem 2.1. Let $\tilde{p} = \|P\|$, $\tilde{A} = \begin{bmatrix} a & c\tilde{p} \\ d\tilde{p} & b \end{bmatrix}$ so that $\tilde{A}$ has eigenvalues $\mu_\pm = \frac{1}{2}[(a+b) \pm \sqrt{(a-b)^2 + 4cd\tilde{p}^2}]$ and $W(\tilde{A})$ is the elliptical disk with foci μ_+, μ_- and minor axis of length*

$$\sqrt{|a|^2 + |b|^2 + \tilde{p}^2(|c|^2 + |d|^2) - |\mu_+|^2 - |\mu_-|^2}.$$

If $\|Px\| = \|P\|$ for some unit vector $x \in \mathcal{H}_1$, then

$$W(A) = W(\tilde{A}).$$

Otherwise, $W(A) = \mathbf{int}(W(\tilde{A})) \cup \{a, b\}$; more precisely, one of the following holds.

(1) *$|c| = |d|$ and $\bar{d}(a-b) = c(\bar{a} - \bar{b})$, both A and $\tilde{A}$ are normal, and*

$$W(A) = W(\tilde{A}) \setminus \sigma(\tilde{A}) = \mathbf{conv}\{\mu_+, \mu_-\} \setminus \{\mu_+, \mu_-\}.$$

(2) *$|c| = |d|$ and there is $\zeta \in (0, \pi)$ such that $\bar{d}(a-b) = e^{i2\zeta}c(\bar{a} - \bar{b}) \neq 0$, both numbers a, b lie on the boundary $\partial W(A)$ of $W(A)$, and*

$$W(A) = \mathbf{int}(W(\tilde{A})) \cup \{a, b\}.$$

(3) *$|c| \neq |d|$, and $W(A) = \mathbf{int}(W(\tilde{A}))$.*

To prove Theorem 3.1, we need the following lemma, which will also be useful for later discussion.

Lemma 3.2. *Let $A_p = \begin{bmatrix} a & cp \\ dp & b \end{bmatrix}$ for $p \geq 0$ so that $W(A_p)$ is the closed elliptical disk $\mathcal{E}(p)$ with foci $\mu_\pm = \frac{1}{2}[(a+b) \pm \sqrt{(a-b)^2 + 4cdp^2}]$ and minor axis of length*

$$\sqrt{|a|^2 + |b|^2 + p^2(|c|^2 + |d|^2) - |\mu_+|^2 - |\mu_-|^2}.$$

Suppose $p < q$. Then $W(A_p) \subseteq W(A_q)$.

(1) *If $|c| = |d|$ and $\bar{d}(a-b) = c(\bar{a}-\bar{b})$, then $W(A_p) = \mathbf{conv}\sigma(A_p)$ and $W(A_q) = \mathbf{conv}\sigma(A_q)$ are line segments such that $W(A_p)$ is a subset of the relative interior of $W(A_q)$.*

(2) *If $|c| = |d|$ and there is $\zeta \in (0,\pi)$ such that $\bar{d}(a-b) = e^{i2\zeta}c(\bar{a}-\bar{b}) \neq 0$, then $\{a,b\} = \partial W(A_p) \cap \partial W(A_q)$, and $W(A_p) \subseteq \mathbf{int}(W(A_q)) \cup \{a,b\}$.*

(3) *If $|c| \neq |d|$, then $W(A_p) \subseteq \mathbf{int} W(A_q)$.*

Proof. Let $\mathcal{E}(p) = W(A_p)$. Then all $\mathcal{E}(p)$ has the same center $\alpha = (a+b)/2$. Suppose $\beta = (a-b)/2$. Denote by $\lambda_1(X)$ the largest eigenvalue of a self-adjoint matrix X. Then

$$W(A_p) = \bigcap_{\xi \in [0,2\pi)} \Pi_\xi(A_p),$$

where

$$\Pi_\xi(A_p) = \{\mu \in \mathbb{C} : e^{i\xi}\mu + e^{-i\xi}\bar{\mu} \leq \lambda_1(e^{i\xi}A_p + e^{-i\xi}A_p^*)\}$$

is a half space in $\mathbb{C}$. Since

$$\lambda_1(e^{i\xi}A_p + e^{-i\xi}A_p^*) = e^{i\xi}\alpha + e^{-i\xi}\bar{\alpha} + \sqrt{|e^{i\xi}\beta + e^{-i\xi}\bar{\beta}|^2 + p^2|e^{i\xi}c + e^{-i\xi}d|^2}$$

is an increasing function of p, we see that $\Pi_\xi(A_p) \subseteq \Pi_\xi(A_q)$ and hence $W(A_p) \subseteq W(A_q)$ if $p \leq q$.

Case 1. Suppose a, b, c, d satisfy condition (1). Then A_p is normal and $A_p = \alpha I_2 + B_p$, where $W(B_p) = \mathbf{conv}\{\pm\sqrt{-\det(B_p)}\}$ is a line segment of length $2\sqrt{|\beta|^2 + p^2|c|^2} = 2\sqrt{|\beta|^2 + p^2|d|^2}$. Thus, the conclusion of (1) holds.

Case 2. Suppose a, b, c, d satisfy condition (2). Then $A_p = \alpha I_2 + \beta B_p$ with

$$e^{i\zeta}B_p = \begin{bmatrix} e^{i\zeta} & \delta p \\ \bar{\delta} p & -e^{i\zeta} \end{bmatrix} \qquad \delta = e^{i\zeta}\frac{2c}{a-b} = e^{-i\zeta}\frac{2\bar{d}}{\bar{a}-\bar{b}}.$$

Using the Elliptical Range Theorem, one readily checks that $W(e^{i\zeta}B_p)$ is a nondegenerate elliptical disk. Since $B_p = \begin{bmatrix} 1 & \delta p e^{-i\zeta} \\ \bar{\delta} p e^{-i\zeta} & -1 \end{bmatrix}$,

$$e^{i\xi}B_p + e^{-i\xi}B_p^* = 2\begin{bmatrix} \cos\xi & \delta p \cos(\xi-\zeta) \\ \bar{\delta} p \cos(\xi-\zeta) & -\cos\xi \end{bmatrix},$$

we have

$$\lambda_1(e^{i\xi}B_p + e^{-i\xi}B_p^*) = 2\sqrt{\cos^2\xi + |\delta|^2 p^2 \cos^2(\xi-\zeta)} \geq \pm 2\cos\xi = \pm\left(e^{i\xi} + e^{-i\xi}\right),$$

where equality holds only for $\xi = \zeta \pm \pi/2$. Therefore, 1 and -1 are on the boundary of $W(B_p)$ and $\lambda_1(e^{i\xi}B_p + e^{-i\xi}B_p^*)$ is a strictly increasing function for $p \geq 0$, except for $\xi = \zeta \pm \pi/2$. From this, we get the conclusion of (2).

Case 3. Suppose a, b, c, d do not satisfy the conditions in (1) or (2). Since $|c| \neq |d|$, for every $\xi \in [0, 2\pi)$,

$$\lambda_1(e^{i\xi}A_p + e^{-i\xi}A_p^*) = e^{i\xi}\alpha + e^{-i\xi}\bar{\alpha} + \sqrt{|e^{i\xi}\beta + e^{-i\xi}\bar{\beta}|^2 + p^2|e^{i\xi}c + e^{-i\xi}\bar{d}|^2}$$

is a strictly increasing function for $p \geq 0$. Thus, condition (3) holds. $\square$

Proof of Theorem 3.1

Since $W(X \oplus Y) = \mathbf{conv}\{W(X) \cup W(Y)\} = W(X)$ if $W(Y) \subseteq W(X)$, we may assume that γI_s is vacuous.

Suppose $x \in \mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1$ is a unit vector and $\mu = \langle Ax, x\rangle \in W(A)$. Let $x = \begin{bmatrix} \cos\theta x_1 \\ \sin\theta x_2 \end{bmatrix}$ for some unit vectors $x_1, x_2 \in \mathcal{H}_1$. Let $\langle Px_1, x_2\rangle = pe^{i\phi}$ with $p \in [0, \tilde{p}]$ and $\phi \in [0, 2\pi)$. Then

$$\mu = [\cos\theta, e^{-i\phi}\sin\theta]A_p \begin{bmatrix} \cos\theta \\ e^{i\phi}\sin\theta \end{bmatrix} \in W(A_p) \subseteq W(\tilde{A})$$

by Lemma 3.2.

If there is a unit vector $x \in \mathcal{H}_1$ such that $\|P\| = \|Px\|$, then

$$\|P\|^2 = \langle P^2x, x\rangle \le \|P^2x\|\|x\| \le \|P^2\| = \|P\|^2 .$$

Thus, $P^2x = \|P\|^2 x$ and hence $Px = \|P\|x$ as P is positive semi-definite. Then the operator matrix of A with respect to $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_0^{\perp}$, where

$$\mathcal{H}_0 = \operatorname{span}\left\{ \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x \end{bmatrix} \right\}$$

has the form $\tilde{A} \oplus \hat{A} \in \mathcal{B}(\mathcal{H})$. Thus, $W(\tilde{A}) \subseteq W(A)$, and the equality holds.

Suppose there is no unit vector $z \in \mathcal{H}_1$ such that $\|P\| = \|Pz\|$. Then for any unit vector $x \in \mathcal{H}$, let $x = \begin{bmatrix} \cos\theta x_1 \\ \sin\theta x_2 \end{bmatrix}$ for some unit vectors $x_1, x_2 \in \mathcal{H}_1$. If $\langle Px_1, x_2\rangle = pe^{i\phi}$ with $p \in [0, \tilde{p}]$ and $\phi \in [0, 2\pi)$, then $p < \tilde{p}$. By Lemma 3.2, we see that $\mu \in \mathbf{int}(W(\tilde{A}))$ if (a) or (c) holds, and $\mu \in \mathbf{int}(W(\tilde{A})) \cup \{a, b\}$ if (b) holds.

To prove the reverse set equalities, note that there is a sequence of unit vectors $\{x_m\}$ in $\mathcal{H}_1$ such that $\langle Px_m, x_m\rangle = p_m$ converges to $\tilde{p}$. Then the compression of A on the subspace

$$\mathcal{V}_m = \operatorname{span}\left\{ \begin{bmatrix} x_m \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x_m \end{bmatrix} \right\} \subseteq \mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1$$

has the form A_{p_m}. Since $W(A_{p_m}) \to W(\tilde{A})$, we see that $\mathbf{int}(W(\tilde{A})) \subseteq W(A)$. It is also clear that $\{a, b\} \subseteq W(A)$. Thus, the set equalities in (1) – (3) hold. $\square$

By Theorem 3.1 and [7, Theorem 2.1], we have the following corollary.

Corollary 3.3. *Suppose A and $\tilde{A}$ satisfy the condition in Theorem 3.1. Then A has a dilation of the form $I \otimes \tilde{A}$.*

The n-th *matricial range* $W^n(L)$ of L is the set of $n \times n$ matrices of the form $\phi(L)$, where $\phi : \mathcal{B}(\mathcal{H}) \to M_n$ is a unital completely positive map.

Theorem 3.4. *Suppose $A \in \mathcal{B}(\mathcal{H})$ and $\tilde{A} \in M_2$ satisfy the hypothesis of Theorem 3.1. Then for every $n \ge 1$, $W^n(A)$ consists of all $n \times n$ matrices B with $W(B) \subseteq \overline{W(A)} = W(\tilde{A})$.*

Proof. By [7, Theorem 2.1], an $n \times n$ matrix $B \in W^n(\tilde{A})$ if and only if $W(B) \subseteq W(\tilde{A})$. The proof of the result is similar to the proof of Theorem 3.1 in [29]. $\square$

We will consider the other kinds of numerical ranges for generalized quadratic operators in Section 4. We consider some operator inequalities in the following.

Denote by $\rho(A) = \max\{|\mu| : \mu \in \sigma(A)\}$ and $w(A) = \sup\{|\mu| : \mu \in W(A)\}$ the *spectral radius* and *numerical radius* of $A \in \mathcal{B}(\mathcal{H})$. It follows readily from Theorems 2.2 and 3.1, that

$$\rho(A) = \rho(\tilde{A}) \quad \text{and} \quad w(A) = w(\tilde{A})$$

if A and $\tilde{A}$ are defined as in Theorem 3.1. Since A has a dilation of the form $I \otimes \tilde{A}$ by Corollary 3.3, we have $\|A\| \leq \|\tilde{A}\|$. As shown in the proof of Theorem 3.1, there is a sequence of two dimensional subspaces $\{\mathcal{V}_m\}$ such that the compression of A on $\mathcal{V}_m$ is A_{p_m}, which converges to $\tilde{A}$. Thus, we have $\|A\| = \|\tilde{A}\|$. Suppose $\tilde{A}$ has singular values $s_1 \geq s_2$. Then $\|\tilde{A}\| = s_1$, $\mathrm{tr}\,(\tilde{A}^*\tilde{A}) = s_1^2 + s_2^2$ and $|\det(\tilde{A})| = s_1 s_2$. Hence, for $\tilde{p} = \|P\|$,

$$\begin{aligned} \|\tilde{A}\| &= \frac{1}{2}\left\{\sqrt{\mathrm{tr}\,(\tilde{A}^*\tilde{A}) + 2|\det(\tilde{A})|} + \sqrt{\mathrm{tr}\,(\tilde{A}^*\tilde{A}) - 2|\det(\tilde{A})|}\right\} \\ &= \frac{1}{2}\left\{\sqrt{|a|^2 + |b|^2 + (|c|^2 + |d|^2)\tilde{p}^2 + 2|ab - cd\tilde{p}^2|}\right. \\ &\qquad \left. + \sqrt{|a|^2 + |b|^2 + (|c|^2 + |d|^2)\tilde{p}^2 - 2|ab - cd\tilde{p}^2|}\right\}. \end{aligned}$$

By the fact that s_1^2 is the larger zero of $\det(\lambda I - \tilde{A}^*\tilde{A})$ and that $\det(\tilde{A}^*\tilde{A}) = |\det(\tilde{A})|^2$, we have

$$\begin{aligned} \|\tilde{A}\| &= \frac{1}{\sqrt{2}}\left\{\sqrt{\mathrm{tr}\,(\tilde{A}^*\tilde{A}) + \sqrt{[\mathrm{tr}\,(\tilde{A}^*\tilde{A})]^2 - 4|\det(\tilde{A})|^2}}\right\} \\ &= \frac{1}{\sqrt{2}}\sqrt{|a|^2 + |b|^2 + (|c|^2 + |d|^2)\tilde{p}^2 + \sqrt{(|a|^2 + |b|^2 + (|d|^2 + |c|^2)\tilde{p}^2)^2 - 4|ab - cd\tilde{p}^2|^2}} \\ &= \frac{1}{\sqrt{2}}\sqrt{|a|^2 + |b|^2 + (|c|^2 + |d|^2)\tilde{p}^2 + \sqrt{(|a|^2 - |b|^2 + (|d|^2 - |c|^2)\tilde{p}^2)^2 + 4|a\bar{c} + \bar{b}d|^2\tilde{p}^2}}. \end{aligned}$$

We summarize the above discussion in the following corollary, which also covers the result of Furuta [13] on $w(A)$ for A of the form (1.1) for $a, b, c, d \geq 0$.

Corollary 3.5. *Suppose A and $\tilde{A}$ satisfy the hypothesis of Theorem 3.1. Then $\rho(A) = \rho(\tilde{A})$, $w(A) = w(\tilde{A})$, and $\|A\| = \|\tilde{A}\|$. In particular, if $a, b \in \mathbb{R}$ and $c, d \in \mathbb{C}$ satisfy $cd \geq 0$, then $\mathbf{cl}\,(W(A)) = W(\tilde{A})$ is symmetric about the real axis, and*

$$\begin{aligned} w(A) &= w((A + A^*)/2) = w(\tilde{A}) = w((\tilde{A} + \tilde{A}^*)/2) = \rho((\tilde{A} + \tilde{A}^*)/2) \\ &= \frac{1}{2}\left\{|a + b| + \sqrt{(a - b)^2 + (|c| + |d|)^2\|P\|^2}\right\} \end{aligned}$$

and

$$\|A\| = \|\tilde{A}\| = \frac{1}{2}\left\{\sqrt{(a + b)^2 + (|b| - |c|)^2\|P\|^2} + \sqrt{(a - b)^2 + (|b| + |c|)^2\|P\|^2}\right\}.$$

Proof. The first assertion follows readily from Theorem 3.1. Suppose $a, b \in \mathbb{R}$ and $c, d \in \mathbb{C}$ with $cd \geq 0$. Then there is a diagonal unitary matrix $D = \mathrm{diag}\,(1, \mu)$ such that $D^*\tilde{A}D = \begin{bmatrix} a & |c|\|P\| \\ |d|\|P\| & b \end{bmatrix}$. It is then easy to get the equalities. $\square$

Corollary 3.6. *Let A_i be self-adjoint operators on $\mathcal{H}_i$ with $\sigma(A_i) \subseteq [m, M]$ for $i = 1, 2$, and let T be an operator from $\mathcal{H}_2$ to $\mathcal{H}_1$. Then*

$$w\left(\begin{bmatrix} A_1 & T \\ T^* & -A_2 \end{bmatrix}\right) \le \frac{1}{2}(M-m) + \frac{1}{2}\sqrt{(M+m)^2 + 4\|T\|^2}. \tag{3.1}$$

Proof. For two self-adjoint operators $X, Y \in \mathcal{B}(\mathcal{H})$, we write $X \le Y$ if $Y - X$ is positive semidefinite. Since $mI \le A_i \le MI$ for $i = 1, 2$, we have

$$\begin{bmatrix} mI & T \\ T^* & -MI \end{bmatrix} \le \begin{bmatrix} A_1 & T \\ T^* & -A_2 \end{bmatrix} \le \begin{bmatrix} MI & T \\ T^* & -mI \end{bmatrix}.$$

By Theorem 3.1,

$$\left\|\begin{bmatrix} mI & T \\ T^* & -MI \end{bmatrix}\right\| = \left\|\begin{bmatrix} MI & T \\ T^* & -mI \end{bmatrix}\right\| = \frac{1}{2}(M-m) + \frac{1}{2}\sqrt{(M+m)^2 + 4\|T\|^2}.$$

The desired inequality (3.1) holds. $\square$

Note that if $X, Y \in \mathcal{B}(\mathcal{H})$, then we have the unitary similarity relations

$$\begin{aligned} \begin{bmatrix} S+iY & 0 \\ 0 & X-iY \end{bmatrix} &= \frac{1}{\sqrt{2}}\begin{bmatrix} I & iI \\ iI & I \end{bmatrix}\begin{bmatrix} X & -Y \\ Y & X \end{bmatrix}\begin{bmatrix} I & -iI \\ -iI & I \end{bmatrix}\frac{1}{\sqrt{2}} \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix} I & I \\ -I & I \end{bmatrix}\begin{bmatrix} X & iY \\ iY & X \end{bmatrix}\begin{bmatrix} I & -I \\ I & I \end{bmatrix}\frac{1}{\sqrt{2}}. \end{aligned}$$

Thus,

$$\max\{\|X+iY\|, \|X-iY\|\} = \left\|\begin{bmatrix} X & -Y \\ Y & X \end{bmatrix}\right\| = \left\|\begin{bmatrix} X & iY \\ iY & X \end{bmatrix}\right\|.$$

Consequently, if $X, Y \in \mathcal{B}(\mathcal{H})$ are self-adjoint with $\sigma(X) \subseteq [m, M]$, then using Corollary 3.6, we have

$$\begin{aligned} \|X+iY\| &= \|X-iY\| = \left\|\begin{bmatrix} X & iY \\ iY & X \end{bmatrix}\right\| = \left\|\begin{bmatrix} X & -Y \\ Y & X \end{bmatrix}\right\| = \left\|\begin{bmatrix} X & Y \\ -Y & X \end{bmatrix}\right\| \\ &\le \frac{1}{2}(M-m) + \frac{1}{2}\sqrt{(M+m)^2 + 4\|Y\|^2}. \end{aligned}$$

This covers a result in [14]. Of course, one can deduce the optimal norm bounds of the sum of normal operators using the results in [7], and the bounds of the sum of two general operators using the results in [8].

4. Generalized numerical ranges

Motivated by theoretical study and applications, there has been many generalizations of the numerical range such as the the q-numerical range, k-numerical range, the c-numerical range, and the essential numerical range; for example, see [2, 10, 11, 16, 17, 19, 24, 30] and their references. Recently, researchers have studied the higher rank numerical range in connection to quantum error correction; see [4, 5, 6, 21, 23] and Section 2.1. Each of these generalizations encodes certain specific information of the operator that leads to interesting applications. To advance the study of these generalized numerical ranges, it is useful to have concrete descriptions of the numerical ranges of certain operators. In most cases, it is relatively easy to solve the problem for self-adjoint or normal operators. The task is more challenging for general operators. In the following, we consider A of the form (2.1). Note that some of the proofs are modifications of those in [22], but some other require different treatments. Moreover, some of the results actually are different. We will include remarks indicating these different situations.

4.1. **q-numerical range.** For $q \in [0,1]$, the *q-numerical range* of A is the set

$$W_q(A) = \{\langle Ax, y\rangle : x, y \in \mathcal{H}, \langle x, x\rangle = \langle y, y\rangle = 1, \langle x, y\rangle = q\}.$$

It is known [19, 28] that

$$W_q(A) = \left\{ q\langle Ax, x\rangle + \sqrt{1-q^2}\langle Ax, y\rangle : \exists \text{ orthonormal } \{x,\ y\} \subseteq \mathcal{H} \right\}, \tag{4.1}$$

and also

$$W_q(A) = \left\{ q\mu + \sqrt{1-q^2}\nu : \exists\ x \in \mathcal{H} \text{ with } \|x\| = 1,\ \mu = \langle Ax, x\rangle,\ |\mu|^2 + |\nu|^2 \le \|Ax\|^2 \right\}. \tag{4.2}$$

If $q = 1$, $W_q(A) = W(A)$. For $0 \le q < 1$, we have the following description of $W_q(A)$ for a generalized quadratic operator $A \in \mathcal{B}(\mathcal{H})$. In particular, $W_q(A)$ will always be an open or closed elliptical disk, which may degenerate to a line segment or a point.

Theorem 4.1. *Suppose A and $\tilde{A}$ satisfy the condition in Theorem 3.1. For any $q \in [0,1)$, if there is a unit vector $z \in \mathcal{H}_1$ such that $\|Pz\| = \|P\|$, then $W_q(A) = W_q(\tilde{A})$; otherwise $W_q(A) = \mathbf{int}\left(W_q(\tilde{A})\right)$.*

We need the following lemma.

Lemma 4.2. *Let A_p be defined as in (2.2). If $p < q$, then for any unit vector $x \in \mathbb{C}^2$ there is a unit vector $y \in \mathbb{C}^2$ such that $\langle A_p x, x\rangle = \langle A_q y, y\rangle$ and $\|A_p x\| < \|A_q y\|$.*

Proof. Choose a unit vector y orthogonal to x such that $A_p x = \mu_1 x + \nu_1 y$. Let $U = [x|y]$. Then $U^* A_p U$ has the form

$$\hat{A}_p = \begin{bmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{bmatrix}$$

with $\langle A_p x, x\rangle = \mu_1$ and $\|A_p x\|^2 = |\mu_1|^2 + |\nu_1|^2$. Since $W(A_p) \subseteq W(A_q)$ by Lemma 3.2 and $\operatorname{tr} A_p = \operatorname{tr} A_q$, we see that A_q is unitarily similar to a matrix of the form

$$\hat{A}_q = \begin{bmatrix} \mu_1 & \hat{\mu}_2 \\ \hat{\nu}_1 & \nu_2 \end{bmatrix}.$$

Since $X \in M_2$ is unitarily similar to X^t, we may assume that $|\hat{\nu}_1| \ge |\hat{\mu}_2|$. Note that

$$\begin{aligned} & |\hat{\nu}_1|^2 + |\hat{\mu}_2|^2 - |\nu_1|^2 - |\mu_2|^2 = \operatorname{tr}(\hat{A}_q^* \hat{A}_q - \hat{A}_p^* \hat{A}_p) \\ = \ & \operatorname{tr}(A_q^* A_q - A_p^* A_p) = (|c|^2 + |d|^2)(|q|^2 - |p|^2) \ge 0, \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} \left|\,|\hat{\nu}_1\hat{\mu}_2| - |\nu_1\mu_2|\,\right| \le |\hat{\nu}_1\hat{\mu}_2 - \nu_1\mu_2| = |\det(\hat{A}_p) - \det(\hat{A}_q)| \\ = \quad |\det(A_p) - \det(A_q)| = |cd|(|q|^2 - |p|^2). \end{aligned}$$

Consequently,

$$(|\hat{\nu}_1| + |\hat{\mu}_2|)^2 - (|\nu_1| + |\mu_2|)^2 \ge (|c| - |d|)^2(|q|^2 - |p|^2) \ge 0$$

and

$$(|\hat{\nu}_1| - |\hat{\mu}_2|)^2 - (|\nu_1| - |\mu_2|)^2 \ge (|c| - |d|)^2(|q|^2 - |p|^2) \ge 0\,.$$

So we have

$$|\hat{\nu}_1| + |\hat{\mu}_2| \ge |\nu_1| + |\mu_2| \quad \text{and} \quad |\hat{\nu}_1| - |\hat{\mu}_2| \ge ||\nu_1| - |\mu_2|| \ge |\nu_1| - |\mu_2|$$

which implies that $|\hat{\nu}_1| \ge |\nu_1|$. From the proof, we can see that if $|\hat{\nu}_1| = |\nu_1|$, then we must have $|c| = |d|$ and $|\hat{\mu}_2| = |\mu_2|$. Then the left hand side of (4.3) is 0, a contradiction.

Therefore, we must have $|\hat{\nu}_1| > |\nu_1|$ and the result follows. □

Proof of Theorem 4.1

By Corollary 3.3, A has a dilation of the form $I \otimes \tilde{A}$. It is then easy to check that

$$W_q(A) \subseteq W_q(I \otimes \tilde{A}) = W_q(\tilde{A}).$$

Let $\{z_m\}$ be a sequence of unit vectors in $\mathcal{H}_1$ such that $\langle Pz_m, z_m\rangle = p_m \to \|P\|$. Then the compression of A on the subspace $\mathcal{V}_m = \text{span}\left\{\begin{bmatrix} z_m \\ 0\end{bmatrix}, \begin{bmatrix} 0 \\ z_m\end{bmatrix}\right\}$ equals A_{p_m} as defined in (2.2). Thus, $W_q(A_{p_m}) \subseteq W(A)$ for all m. Since $A_{p_m} \to \tilde{A}$, we see that $\mathbf{int}(W_q(\tilde{A})) \subseteq W_q(A)$. Suppose there is a unit vector $z \in \mathcal{H}_1$ such that $\|Pz\| = \|P\|$. We may assume that $z_m = z$ for each m so that $W_q(A) = W_q(\tilde{A})$.

Suppose there is no unit vector $z \in \mathcal{H}_1$ such that $\|Pz\| = \|P\|$. For any unit vectors $x, y \in \mathcal{H}$ with $\langle x, y\rangle = q$, we show that $\zeta = \langle Ax, y\rangle \in \mathbf{int}(W_q(\tilde{A}))$ in the following. To prove our claim, let $x = \begin{bmatrix} \alpha_1 u_1 \\ \alpha_2 u_2 \\ \alpha_3 u_3 \end{bmatrix}$, $y = \begin{bmatrix} \beta_1 u_1 + \gamma_1 v_1 \\ \beta_2 u_1 + \gamma_2 v_2 \\ \beta_3 u_3 + \gamma_3 v_3 \end{bmatrix} \in \mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$ such that $u_1, u_2, v_1, v_2 \in \mathcal{H}_1$, $u_3, v_3 \in \mathcal{H}_2$ are unit vectors, $u_i \perp v_i$ for $i = 1, 2, 3$, and $(\alpha_1, \alpha_2, \alpha_3)^t$, $(\beta_1, \beta_2, \beta_3)^t$, $(\gamma_1, \gamma_2, \gamma_3)^t \in \mathbb{C}^3$. Then the compression of A on

$$\mathcal{V} = \text{span}\left\{ \begin{bmatrix} u_1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ u_2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ u_3 \end{bmatrix}, \begin{bmatrix} v_1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ v_2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ v_3 \end{bmatrix} \right\}$$

has the form

$$B = \begin{bmatrix} aI_2 & cP_0 & 0 \\ dP_0^* & bI_2 & 0 \\ 0 & 0 & \gamma I_2 \end{bmatrix},$$

where $P_0 \in M_2$ satisfies $\|P_0\| < \|P\|$. Here note that the $(3,3)$ entry may be vacuous. Then B is unitarily similar to

$$\tilde{B} = \begin{bmatrix} aI_2 & cD_0 & 0 \\ dD_0 & bI_2 & 0 \\ 0 & 0 & \gamma I_2 \end{bmatrix}, \text{ where } D_0 = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \text{ with } d_2 \leq d_1 < \|P\|$$

Since $W(\tilde{B}) \subseteq W(A_{d_1})$, $\tilde{B}$ has a dilation $I \otimes A_{d_1}$. Therefore, $W_q(\tilde{B}) \subseteq W_q(I \otimes A_{d_1}) = W_q(A_{d_1})$ Since $\zeta \in W_q(\tilde{B}) \subseteq W_q(A_{d_1})$, A_{d_1} is unitarily similar to a matrix of the form

$$\hat{A} = \begin{bmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{bmatrix}$$

such that $\zeta = q\mu_1 + \sqrt{1-q^2}\nu_1$, where $\mu_1 = \langle A_{d_1}x', x'\rangle$ and $|\nu_1|^2 \leq \|A_{d_1}x_1\|^2 - |\mu_1|^2$ for some unit vector $x' \in \mathbb{C}^2$. By Lemma 4.2, there is a unit vector $y' \in \mathbb{C}^2$ such that $\langle \tilde{A}y', y'\rangle = \mu_1$ and $\|\tilde{A}y'\| > \|A_{d_1}x'\|$. Hence, by (4.1) and (4.2),

$$\zeta \in \{q\mu_1 + \sqrt{1-q^2}\nu : |\mu_1|^2 + |\nu|^2 < \|\tilde{A}y'\|^2\} \subseteq \mathbf{int}(W_q(\tilde{A})). \qquad \square$$

4.2. **Rank-k numerical ranges and essential numerical ranges.** For a positive integer k, define the *rank-k numerical range* of $A \in \mathcal{B}(\mathcal{H})$ by

$$\Lambda_k(A) = \{\lambda \in \mathbb{C} : PAP = \lambda P \text{ for some rank-}k \text{ orthogonal projection } P\}.$$

This generalized numerical range is motivated by the study of quantum error correction; see [4, 5, 6].

To describe some basic results of $\Lambda_k(A)$, we need the following notation. Let $H \in \mathcal{B}(\mathcal{H})$ be a self-adjoint operator. If $\dim \mathcal{H} = n$, denote by $\lambda_1(H) \geq \cdots \geq \lambda_n(H)$ the eigenvalues of H. If $\mathcal{H}$ is infinite dimensional, define

$$\lambda_m(H) = \sup\{\lambda_m(X^*HX) : X \text{ is an isometry from } \mathbb{C}^m \text{ to } \mathcal{H}\}.$$

It is known (see [26]) and not hard to verify that $\lambda_m(H)$ of an infinite dimensional operator H can be determined as follows. Let

$$\sigma_e(A) = \cap\{\sigma(A+F) : F \in \mathcal{B}(\mathcal{H}) \text{ has finite rank}\}$$

be the *essential spectrum* of $A \in \mathcal{B}(\mathcal{H})$, and let

$$\lambda_\infty(H) = \sup \sigma_e(H),$$

which also equals the supremum of the set

$$\sigma(H) \setminus \{\mu \in \mathbb{C} : H - \mu I \text{ has a non-trivial finite dimensional null space}\}.$$

Then $\mathcal{S} = \sigma(H) \cap (\lambda_\infty(H), \infty)$ has only isolated points, and we can arrange the elements in descending order, say, $\lambda_1 \geq \lambda_2 \geq \cdots$ counting multiplicities, i.e., each element repeats according to the dimension of its eigenspace. If $\mathcal{S}$ is infinite, then $\lambda_j(H) = \lambda_j$ for each positive integer j. If $\mathcal{S}$ has m elements, then $\lambda_j(H) = \lambda_j$ for $j = 1, \ldots, m$, and $\lambda_j(H) = \lambda_\infty(H)$ for $j > m$.

Let

$$\Omega_k(A) = \bigcap_{\xi\in[0,2\pi)} \left\{\mu \in \mathbb{C} : \mathrm{Re}\,(e^{i\xi}\mu) \le \lambda_k\left(\mathrm{Re}\,(e^{i\xi}A)\right)\right\},$$

where $\mathrm{Re}\,(B) = (B + B^*)/2$ is the real part of B. It was shown in [20] that

$$\mathbf{int}(\Omega_k(A)) \subseteq \Lambda_k(A) \subseteq \Omega_k(A) = \mathbf{cl}\,(\Lambda_k(A)).$$

In particular, $\Lambda_k(A) = \Omega_k(A)$ if $A \in M_n$; see also [23].

The rank-k numerical range of a generalized quadratic operator A of the form (2.1) can be an empty set, a singleton, a line segment or an elliptical disk with all or part of its boundary. The following theorem gives the precise description of the set using Theorem 3.1 and Lemma 3.2.

Theorem 4.3. *Suppose $A \in \mathcal{B}(\mathcal{H})$ satisfies the hypothesis of Theorem 3.1. Let k be a positive integer not larger than $\dim\mathcal{H}$. Suppose $\mathcal{E}$ is the closed elliptical disk*

$$\mathcal{E} = W\left(\begin{bmatrix} a & c\lambda_k(P) \\ d\lambda_k(P) & b\end{bmatrix}\right)$$

with foci $\mu_\pm = \frac{1}{2}[(a+b) \pm \sqrt{(a-b)^2 + 4cd\lambda_k(P)^2}]$ and minor axis of length

$$\sqrt{|a|^2 + |b|^2 + \lambda_k(P)^2(|c|^2 + |d|^2) - |\mu_+|^2 - |\mu_-|^2}.$$

(a) *If $r + s < k$, then $\Lambda_k(A) = \emptyset$.*
(b) *If $r < k \le r + s$, then $\Lambda_k(A) = \{\gamma\}$.*
(c) *Suppose $k \le r$. If $P : \mathcal{H}_1 \to \mathcal{H}_1$ has a compression $\mathrm{diag}\,(p_1, \dots, p_k)$ with $p_1 \ge \cdots \ge p_k = \lambda_k(P)$, then $\Lambda_k(A) = \mathcal{E}$. Otherwise, $\Lambda_k(A) = \mathbf{int}(\mathcal{E}) \cup \{a, b\}$; more precisely, one of the following holds.*
   (1) *If $|c| = |d|$ and $\bar{d}(a - b) = c(\bar{a} - \bar{b})$, then $\mathcal{E} = \mathbf{conv}\{\mu_+, \mu_-\}$ is a line segment and $\Lambda_k(A) = \mathcal{E} \setminus \{\mu_+, \mu_-\}$.*
   (2) *If $|c| = |d|$ and there is $\zeta \in (0, \pi)$ such that $\bar{d}(a - b) = e^{i2\zeta}c(\bar{a} - \bar{b}) \neq 0$, then $a, b \in \partial\Lambda_k(A)$ and $\Lambda_k(A) = \mathbf{int}(\mathcal{E}) \cup \{a, b\}$.*
   (3) *If $|c| \neq |d|$, then $\Lambda_k(A) = \mathbf{int}(\mathcal{E})$.*

Remark 4.4. In (c) of Theorem 4.3, it is not hard to show that another equivalent condition for $\Lambda_k(A) = \mathcal{E}$ is that

$$P : \mathcal{H}_1 \to \mathcal{H}_1 \text{ has a compression } \mathrm{diag}\,(\lambda_1(P), \dots, \lambda_k(P)),$$

and therefore $\Lambda_\ell(A)$ is an elliptical disk with foci $\mu_\pm = \frac{1}{2}[(a+b) \pm \sqrt{(a-b)^2 + 4cd\lambda_\ell(P)^2}]$ and minor axis of length

$$\sqrt{|a|^2 + |b|^2 + \lambda_\ell(P)^2(|c|^2 + |d|^2) - |\mu_+|^2 - |\mu_-|^2}.$$

for any $\ell \in \{1, \dots, k\}$. Also if $\lambda_k(P) = 0$, then $\mathcal{E}$ becomes the line segment joining a and b, i.e., $\mathcal{E} = \mathbf{conv}\{a, b\}$. In this case, $\Lambda_k(A)$ equals $\mathbf{conv}\{a, b\}$.

The proof of Theorem 4.3 is very similar to Theorem 2.1 in [22]. First, we state Lemma 2.4 in [22], and then prove a lemma analogous to Lemma 2.5 in [22].

Lemma 4.5. *Suppose P is a positive semidefinite operator in $\mathcal{B}(\mathcal{H}_1)$ with $\dim(\mathcal{H}_1) \geq k$. For any $\varepsilon > 0$, there exist $p_1, \ldots, p_k \in [0, \infty)$ with $\lambda_j(P) - \varepsilon < p_j \leq \lambda_j(P)$ for $j = 1, \ldots, k$, such that P has a compression of the form $\operatorname{diag}(p_1, \ldots, p_k)$.*

Lemma 4.6. *Let $A \in \mathcal{B}(\mathcal{H})$ be an operator of the form described in Theorem 3.1 with the additional assumption that $r = \infty$. Suppose $\mathcal{V}_1$ is a k-dimensional subspace of $\mathcal{H}$. Then there is a $(4k + \ell)$-dimensional subspace $\mathcal{V}_2$ of $\mathcal{H}$ containing $\mathcal{V}_1$ with $\ell = \min\{s, k\}$ such that the compression of A on $\mathcal{V}_2$ has the form*

$$\begin{bmatrix} aI_{2k} & cP' \\ dP' & bI_{2k} \end{bmatrix} \oplus \gamma I_\ell,$$

where $P' = \operatorname{diag}(p_1, \ldots, p_{2k})$ is a compression of P, with $p_1 \geq \cdots \geq p_{2k}$ and $p_i \leq \lambda_i(P)$ for $1 \leq i \leq 2k$.

Proof. Suppose A has the form described in Theorem 3.1, with respect to the decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$ and $\dim \mathcal{H}_1 = r = \infty$. Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be k-dimensional subspaces of $\mathcal{H}_1$ such that $\mathcal{K}_1 \oplus 0 \oplus 0$ and $0 \oplus \mathcal{K}_2 \oplus 0$ contain the orthogonal projections of $\mathcal{V}_1$ on $\mathcal{H}_1 \oplus 0 \oplus 0$ and $0 \oplus \mathcal{H}_1 \oplus 0$, respectively. Also let $\mathcal{K}_3$ be a ℓ-dimensional subspace of $\mathcal{H}_2$, with $\ell = \min\{s, k\}$, such that $0 \oplus 0 \oplus \mathcal{K}_3$ contains the orthogonal projection of $\mathcal{V}_1$ on $0 \oplus 0 \oplus \mathcal{H}_2$. Clearly, $\mathcal{K}_1 \oplus \mathcal{K}_2 \oplus \mathcal{K}_3$ contains $\mathcal{V}_1$. Take a $2k$-dimensional subspace $\mathcal{K}$ of $\mathcal{H}_1$ containing $\mathcal{K}_1 + \mathcal{K}_2$ and $\mathcal{V}_2 = \mathcal{K} \oplus \mathcal{K} \oplus \mathcal{K}_3$. Then $\mathcal{V}_2$ also contains $\mathcal{V}_1$. Let $S : \mathcal{V}_2 \hookrightarrow \mathcal{H}$ be the imbedding of $\mathcal{V}_2$ into $\mathcal{H}$. Then S^*AS has an operator matrix of the form

$$\begin{bmatrix} aI_{2k} & cX^*PX \\ dX^*PX & bI_{2k} \end{bmatrix} \oplus \gamma I_\ell,$$

where X is the imbedding of $\mathcal{K}$ into $\mathcal{H}_1$. Furthermore, we can find a unitary operator U such that

$$U^*X^*PXU = \operatorname{diag}(p_1, \ldots, p_{2k}) \qquad \text{with} \qquad p_1 \geq \cdots \geq p_{2k}.$$

Let $R = S(U \oplus U \oplus I_\ell)$. Then $R^*R = I_{4k+\ell}$ and R^*AR has the asserted form. □

Proof of Theorem 4.3.

We first consider the finite dimensional case. Let $n = \dim \mathcal{H} = 2r + s$. Assume that P has eigenvalues $p_1 \geq \cdots \geq p_r \geq 0$. By Theorem 2.1, A is unitarily similar to

$$B_1 \oplus B_2 \oplus \cdots \oplus B_r \oplus \gamma I_s,$$

where $B_j = A_{p_j}$, where A_p is defined as in (2.2).

We note that if $a = b$, then our assumption ensures that $p_j \neq 0$. Therefore, B_j is never a scalar matrix and $\Omega_2(B_j) = \emptyset$.

By the argument in the preceding paragraph,

$$\lambda_1\left(\operatorname{Re}(e^{i\xi}B_1)\right) \geq \cdots \geq \lambda_1\left(\operatorname{Re}(e^{i\xi}B_r)\right)$$
$$\geq \operatorname{Re}(e^{i\xi}\gamma) \geq \lambda_2\left(\operatorname{Re}(e^{i\xi}B_r)\right) \geq \cdots \geq \lambda_2\left(\operatorname{Re}(e^{i\xi}B_1)\right).$$

Then $\lambda_k\left(\mathrm{Re}\,(e^{i\xi}A)\right)$ equals

$$\begin{cases} \lambda_1\left(\mathrm{Re}\,(e^{i\xi}B_k)\right) & \text{if } k \le r, \\ \mathrm{Re}\,(e^{i\xi}\gamma) & \text{if } r < k \le r+s, \\ \lambda_2\left(\mathrm{Re}\,(e^{i\xi}B_{n-k+1})\right) & \text{if } r+s < k \le n. \end{cases}$$

Recall that $\mu \in \Omega_k(A)$ if and only if $\mathrm{Re}\,(e^{i\xi}\mu) \le \lambda_k\left(\mathrm{Re}\,(e^{i\xi}A)\right)$ for all $\xi \in [0, 2\pi)$. We have

$$\Lambda_k(A) = \Omega_k(A) = \begin{cases} \Omega_1(B_k) & \text{if } k \le r, \\ \{\gamma\} & \text{if } r < k \le r+s, \\ \Omega_2(B_{n-k+1}) = \emptyset & \text{if } r+s < k \le n\,. \end{cases}$$

Then the assertion holds when $k > r$. If $k \le r$, then

$$\Lambda_k(A) = \Omega_k(A) = \Omega_1(B_k) = \Lambda_1(B_k) = \mathcal{E}(p_k).$$

Thus, the result holds in the finite dimensional case.

Next, suppose $\mathcal{H}$ is an infinite dimensional Hilbert space. If $r < k$, then $\Omega_k(A) = \{\gamma\}$ and hence $\Lambda_k(A) = \{\gamma\}$.

If $r \ge k$ is finite or $\lambda_k(P) = 0$, then $P : \mathcal{H}_1 \to \mathcal{H}_1$ has a compression $\mathrm{diag}\,(\lambda_1(P), \ldots, \lambda_k(P))$. Let

$$\tilde{A} = B_1 \oplus \cdots \oplus B_k \in M_{2k}$$

with $B_j = \begin{bmatrix} a & c\lambda_j(P) \\ d\lambda_j(P) & b \end{bmatrix}$ for $j = 1, \ldots, k$. Notice that $\tilde{A}$ is a compression of A and

$$\lambda_k\left(\mathrm{Re}\,\left(e^{i\xi}A\right)\right) = \lambda_k\left(\mathrm{Re}\,\left(e^{i\xi}\tilde{A}\right)\right) \qquad \text{for all } \xi \in [0, 2\pi).$$

Hence,

$$\Lambda_k(\tilde{A}) \subseteq \Lambda_k(A) \subseteq \Omega_k(A) = \Omega_k(\tilde{A}) = \Lambda_k(\tilde{A}).$$

Thus, $\Lambda_k(A) = \Lambda_k(\tilde{A})$ so that the result follows from the finite dimensional case.

Suppose r is infinite and $\lambda_k(P) > 0$. We prove that (c) holds with $\mathcal{E} = \mathcal{E}(\lambda_k(P))$. Let μ be an interior point of $\mathcal{E}$. Then there exists $\varepsilon > 0$ such that $\mu \in \mathcal{E}(\lambda_k(P) - \varepsilon)$. By Lemma 4.5, there exist a k-dimensional subspace $\mathcal{V}$ of $\mathcal{H}$ and $X : \mathcal{V} \to \mathcal{H}_1$ satisfying $X^*X = I_k$ and

$$\lambda_k(X^*PX) > \lambda_k(P) - \varepsilon.$$

Let $Z = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \oplus I_s$. Then we have $Z^*AZ = \begin{bmatrix} aI_k & cX^*PX \\ dX^*PX & bI_k \end{bmatrix} \oplus \gamma I_s$ and

$$\mu \in \mathcal{E}(\lambda_k(P) - \varepsilon) \subseteq \Lambda_k(Z^*AZ) \subseteq \Lambda_k(A).$$

Conversely, suppose $\mu \in \Lambda_k(A)$. Then there exist a k-dimensional subspace $\mathcal{V}_1$ of $\mathcal{H}$ and $X : \mathcal{V}_1 \to \mathcal{H}$ such that $X^*X = I_{\mathcal{V}_1}$ and $X^*AX = \mu I_{\mathcal{V}_1}$. By Lemma 4.6, there is a $(4k+\ell)$-dimensional subspace $\mathcal{V}_2$ containing $\mathcal{V}_1$ such that the compression of A on $\mathcal{V}_2$ has an operator matrix

$$A' = \begin{bmatrix} aI_{2k} & cP' \\ dP' & bI_{2k} \end{bmatrix} \oplus \gamma I_\ell \in M_{4k+\ell},$$

where $P' = \operatorname{diag}(p_1, \ldots, p_{2k})$ is a $2k$-dimensional compression of P, with $p_1 \geq \cdots \geq p_{2k}$ and $p_i \leq \lambda_i(P)$ for $1 \leq i \leq 2k$. By the result in the finite dimensional case, we have

$$\mu \in \Lambda_k(A') = \mathcal{E}(\lambda_k(P')) \subseteq \mathcal{E}(\lambda_k(P)).$$

So, we have shown that

$$\mathbf{int}(\mathcal{E}(\lambda_k(P))) \subseteq \Lambda_k(A) \subseteq \mathcal{E}(\lambda_k(P)).$$

Also, if P has a k-dimensional compression $\operatorname{diag}(p_1, \ldots, p_k)$ with $p_k = \lambda_k(P)$ and $\Lambda_k(A) = \mathcal{E}(\lambda_k(P))$. Otherwise, $\Lambda_k(A)$ can only contain points in the relative interior of $\mathcal{E}(\lambda_k(P))$ unless (2) holds so that $a, b \in \Lambda_k(A) \cap \partial\mathcal{E}(\lambda_k(P))$. The proof is complete. □

For an infinite dimensional operator A, one can extend the definition of rank-k numerical range to $\Lambda_\infty(A)$ defined as the set of scalars $\lambda \in \mathbb{C}$ such that $PAP = \lambda P$ for an infinite rank orthogonal projection P on $\mathcal{H}$, see [20, 25]. Evidently, $\Lambda_\infty(A)$ consists of those $\lambda \in \mathbb{C}$ for which there exists an infinite orthonormal set $\{x_i \in \mathcal{H} : i \geq 1\}$ such that $\langle Ax_i, x_j\rangle = \delta_{ij}\lambda$ for all $i, j \geq 1$. It is shown in [20] that

$$\Lambda_\infty(A) = \bigcap_{k\geq 1} \Lambda_k(A) = \bigcap\{W(A + F) : F \in \mathcal{B}(\mathcal{H}) \text{ has a finite rank}\}.$$

Recall that $\lambda_\infty(H)$ is the supremum of the set

$$\sigma(H) \setminus \{\mu \in \mathbb{C} : H - \mu I \text{ has a non-trivial finite dimensional null space}\}.$$

One can extend the definition of $\Omega_k(A)$ to

$$\Omega_\infty(A) = \bigcap_{k\geq 1} \Omega_k(A).$$

By Theorem 5.1 in [20] (see also [1, Theorem 4]),

$$\Omega_\infty(A) = \bigcap\{\mathbf{cl}\,(W(A + F)) : F \in \mathcal{B}(\mathcal{H}) \text{ has a finite rank}\}$$

is the essential numerical range $W_e(A)$ of A; $\Omega_\infty(A) = \mathbf{cl}\,(\Lambda_\infty(A))$ if and only if $\Lambda_\infty(A)$ is non-empty.

By Theorem 4.3, we have the following corollary, which gives a complete description of $\Lambda_\infty(A)$ and the essential numerical range of a quadratic operator A. It turns out that each of them can be a singleton, a line segment or an elliptical disk.

Corollary 4.7. *Suppose $A \in \mathcal{B}(\mathcal{H})$ is an infinite dimensional generalized quadratic operator with operator matrix in the form in Theorem 3.1. Suppose*

$$\mathcal{E} = W\left(\begin{bmatrix} a & c\lambda_\infty(P) \\ d\lambda_\infty(P) & b \end{bmatrix}\right)$$

with foci $\mu_\pm = \frac{1}{2}[(a + b) \pm \sqrt{(a - b)^2 + 4cd\lambda_\infty(P)^2}]$ and minor axis of length

$$\sqrt{|a|^2 + |b|^2 + \lambda_\infty(P)^2(|c|^2 + |d|^2) - |\mu_+|^2 - |\mu_-|^2}.$$

(a) *If $r < \infty$, then $\Lambda_\infty(A) = \{\gamma\}$.*

(b) *Suppose $r=\infty$. If $\sigma(P)\cap(\lambda_\infty(P),\infty)$ is infinite, equivalently, $P-\lambda_\infty(P)I$ has an infinite dimensional null space, then $\Lambda_\infty(A)=\mathcal{E}$. Otherwise, $\Lambda_\infty(A)=\mathbf{int}(\mathcal{E})\cup\{a,b\}$; more precisely, one of the following holds.*

(1) *If $|c|=|d|$ and $\bar d(a-b)=c(\bar a-\bar b)$, then $\mathcal{E}=\mathbf{conv}\{\mu_+,\,\mu_-\}$ is a line segment and $\Lambda_\infty(A)=\mathcal{E}\setminus\{\mu_+,\,\mu_-\}$.*

(2) *If $|c|=|d|$ and there is $\zeta\in(0,\pi)$ such that $\bar d(a-b)=e^{i2\zeta}c(\bar a-\bar b)\neq 0$, then $a,b\in\partial\Lambda_k(A)$ and $\Lambda_\infty(A)=\mathbf{int}(\mathcal{E})\cup\{a,b\}$.*

(3) *If $|c|\neq|d|$, then $\Lambda_\infty(A)=\mathbf{int}(\mathcal{E})$.*

Consequently, $W_e(A)=\Omega_\infty(A)=\mathbf{cl}\,(\Lambda_\infty(A))$ is a singleton, a line segment or a closed elliptical disk.

4.3. **c-numerical ranges.** For $c=(c_1,\dots,c_k)$ with $c_1\ge\cdots\ge c_k$ and $k\le\dim\mathcal{H}$, the *c-numerical range* of A is

$$W_c(A)=\left\{\sum_{j=1}^k c_j\langle Ax_j,x_j\rangle:\{x_1,\dots,x_k\}\subseteq\mathcal{H}\text{ is an orthonormal set}\right\}.$$

If $(c_1,\dots,c_k)=(1,\dots,1)$, then $W_c(A)$ reduces to the *k-numerical range*; see [16].

Suppose $A\in M_2$ has eigenvalues λ_1 and λ_2, and $c=(c_1,c_2)$. Then

$$W_c(A)=(c_1-c_2)W(A)+c_2\mathrm{tr}\,A=W((c_1-c_2)A+(c_2\mathrm{tr}\,A)I_2)$$

is the elliptical disk with foci $c_1\lambda_1+c_2\lambda_2$ and $c_2\lambda_1+c_1\lambda_2$, and the lengths of minor and major axis of $W_c(A)$ are, respectively,

$$\{\mathrm{tr}\,(\hat A^*\hat A)-2|\det\hat A|\}^{1/2}\qquad\text{and}\qquad\{\mathrm{tr}\,(\hat A^*\hat A)+2|\det\hat A|\}^{1/2},$$

where $\hat A=\dfrac{(c_1-c_2)}{2}\,(A-2(\mathrm{tr}\,A)I_2)$.

For a self-adjoint operator $H\in B(\mathcal{H})$, we have

$$\mathbf{cl}\,(W_c(H))=[m_c(H),M_c(H)],$$

where

$$m_c(H)=\inf\left\{-\sum_{j=1}^{\ell}c_j\lambda_j(-H)+\sum_{j=1}^{k-\ell}c_{k-j+1}\lambda_j(H):0\le\ell\le k\right\}$$

and

$$M_c(H)=\sup\left\{\sum_{j=1}^{\ell}c_j\lambda_j(H)-\sum_{j=1}^{k-\ell}c_{k-j+1}\lambda_j(-H):0\le\ell\le k\right\}.$$

For a general operator $A\in\mathcal{B}(\mathcal{H})$, we have

$$\mathbf{cl}\,(W_c(A))=\bigcap_{t\in[0,2\pi)}\left\{\mu\in\mathbb{C}:\mathrm{Re}\,\left(e^{it}\mu\right)\le M_c\left(\mathrm{Re}\,(e^{it}A)\right)\right\}.\tag{4.4}$$

For a generalized quadratic operator $A \in \mathcal{B}(\mathcal{H})$, it is easy to determine $\lambda_m\left(\mathrm{Re}\,(e^{it}A)\right)$. Thus, it is not hard to determine $W_c(A)$ using (4.4). It turns out that $\mathbf{cl}\,(W_c(A))$ can always be expressed as the sum of a finite number of elliptical disks, namely,

$$\mathbf{cl}\,(W_c(A)) = W(C_1) + \cdots + W(C_t) + d$$

for some constant $d \in \mathbb{C}$ and $C_1, \ldots, C_t \in M_2$ with $t \leq k$.

To simplify the statement of our results, we will impose the following assumption on the vector $c = (c_1, \ldots, c_k)$:

$$\begin{aligned} &c_1 \geq \cdots \geq c_k \quad \text{with} \quad c_{m+1} = 0, \quad \text{where} \\ &\dim \mathcal{H} = \infty > k = 2m \quad \text{or} \quad \dim \mathcal{H} = k \in \{2m, 2m+1\}. \end{aligned} \tag{4.5}$$

Note that it is easy to reduce the general case to the study of the special vector c with assumption (4.5). In the infinite dimensional case, this can be achieved by adding zeros to the vector $c = (c_1, \ldots, c_k)$. In the finite dimensional case, we can first assume that $k = \dim \mathcal{H}$ by adding zeros to the vector c, and then replace c with $\hat{c} = c - c_{m+1}(1, \ldots, 1)$. One can then use the fact that $W_c(A) = W_{\hat{c}}(A) + c_{m+1}\mathrm{tr}\, A$ to determine the shape of $W_c(A)$. Note also that the advantage of this assumption on c is that the supremum in the definition of $M_c(H)$ is always attained at $\ell = m$.

For notational convenience, we assume that the generalized quadratic operator A in the form (2.2) has $(1,2)$ block equal to P instead of cP in the following theorem.

Theorem 4.8. *Let $A \in \mathcal{B}(\mathcal{H})$ be a generalized quadratic operator with an operator matrix in the form*

$$\begin{bmatrix} aI_r & P \\ dP & bI_r \end{bmatrix} \oplus \gamma I_s \qquad \text{with } \ dP \neq 0. \tag{4.6}$$

Suppose $c = (c_1, \ldots, c_k)$ satisfies (4.5) and $t = \min\{m, r\}$. Let

$$\mathcal{E} = W(C_1) + \cdots + W(C_t) + \gamma \sum_{j=t+1}^{k-t} c_j,$$

where

$$C_j = (c_j - c_{k-j+1}) \begin{bmatrix} a & \lambda_j(P) \\ d\lambda_j(P) & b \end{bmatrix} + c_{k-j+1}(a+b)I_2, \qquad j = 1, \ldots, t.$$

If $P : \mathcal{H}_1 \to \mathcal{H}_1$ has a compression $\mathrm{diag}\,(\lambda_1(P), \ldots, \lambda_t(P))$, then $W_c(A) = \mathcal{E}$. Otherwise,

$$W_c(A) = \mathbf{int}(\mathcal{E}) \cup \{\mathbf{q}_1, \mathbf{q}_2\},$$

where

$$\mathbf{q}_1 = \sum_{j=1}^{t}(ac_j - bc_{k-j+1}) + \gamma \sum_{j=t+1}^{k-\ell} c_j, \quad \mathbf{q}_2 = \sum_{j=1}^{t}(bc_j - ac_{k-j+1}) + \gamma \sum_{j=t+1}^{k-\ell} c_j;$$

more precisely, one of the following holds.

(1) *If* $|d| = 1$ *and* $\bar{d}(a-b) = (\bar{a}-\bar{b})$, *then* $\mathcal{E}$ *is a line segment and* $W_c(A)$ *is the relative interior of* $\mathcal{E}$.
(2) *If* $|d| = 1$ *and there is* $\zeta \in (0,\pi)$ *such that* $\bar{d}(a-b) = e^{i2\zeta}(\bar{a}-\bar{b}) \neq 0$, *then* $W_c(A)$ *consists of the interior of the non-degenerate elliptical disk* $\mathcal{E}$ *and its two boundary points* $\mathbf{q}_1$ *and* $\mathbf{q}_2$.
(3) *If* $|d| \neq 1$, *then* $\Lambda_k(A) = \mathbf{int}(\mathcal{E})$.

Proof. Suppose $\dim \mathcal{H} = n$ is finite. So we have $k = n$ and $r \leq m$. Notice that A is unitarily similar to

$$B_1 \oplus \cdots \oplus B_r \oplus \gamma I_s,$$

where $B_j = \begin{bmatrix} a & \lambda_j(P) \\ d\lambda_j(P) & b \end{bmatrix}$ for $j = 1, \ldots, r$. By the argument in the proof of Theorem 4.3, we have

$$\lambda_j\left(\mathrm{Re}\,(e^{i\xi}A)\right) = \begin{cases} \lambda_1\left(\mathrm{Re}\,(e^{i\xi}B_j)\right) & \text{if } j \leq r, \\ \mathrm{Re}\,(e^{i\xi}\gamma) & \text{if } r < j \leq r+s, \\ \lambda_2\left(\mathrm{Re}\,(e^{i\xi}B_{n-j+1})\right) & \text{if } r+s < j \leq n. \end{cases}$$

Under assumption (4.5) and $k = n$, we have

$$M_c\left(\mathrm{Re}\,(e^{i\xi}A)\right) = \sum_{j=1}^{n} c_j\lambda_j\left(\mathrm{Re}\,(e^{i\xi}A)\right).$$

On the other hand,

$$\mathrm{Re}\,(e^{i\xi}\gamma)\sum_{j=r+1}^{n-r} c_j = \sum_{j=r+1}^{n-r} c_j\lambda_j\left(\mathrm{Re}\,(e^{i\xi}A)\right)$$

and

$$\begin{aligned} &\sum_{j=1}^{r} M_{(c_j,c_{n-j+1})}\left(\mathrm{Re}\,(e^{i\xi}B_j)\right) \\ =\ &\sum_{j=1}^{r}\left[c_j\lambda_1\left(\mathrm{Re}\,(e^{i\xi}B_j)\right) + c_{n-j+1}\lambda_2\left(\mathrm{Re}\,(e^{i\xi}B_j)\right)\right] \\ =\ &\sum_{j=1}^{r} c_j\lambda_j\left(\mathrm{Re}\,(e^{i\xi}A)\right) + \sum_{j=n-r+1}^{n} c_j\lambda_j\left(\mathrm{Re}\,(e^{i\xi}A)\right). \end{aligned}$$

Thus, $M_c\left(\mathrm{Re}\,(e^{i\xi}A)\right)$ equals

$$\sum_{j=1}^{r} M_{(c_j,c_{n-j+1})}\left(\mathrm{Re}\,(e^{i\xi}B_j)\right) + \mathrm{Re}\,(e^{i\xi}\gamma)\sum_{j=r+1}^{n-r} c_j.$$

By (4.4) and the above equation, the two compact convex sets

$$W_c(A) \quad \text{and} \quad W_{(c_1,c_n)}(B_1) + \cdots + W_{(c_r,c_{n-r+1})}(B_r) + \gamma \sum_{j=r+1}^{n-r} c_j$$

always share the same support line in each direction. Thus, the two sets are the same. Since $W_{(c_j,c_{n-r+j})}(B_j) = W(C_j)$ for $j = 1, \ldots, r$, it follows that

$$W_c(A) = W(C_1) + \cdots + W(C_r) + \gamma \sum_{j=r+1}^{n-r} c_j.$$

Next, suppose $\dim \mathcal{H}$ is infinite. Suppose r is finite or $\lambda_m(P) = 0$. Let $t = \min\{m, r\}$. Then $P : \mathcal{H}_1 \to \mathcal{H}_1$ has a compression $\operatorname{diag}(\lambda_1(P), \ldots, \lambda_t(P))$. Take

$$\tilde{A} = B_1 \oplus \cdots \oplus B_t \oplus \gamma I_{k-2t} \in M_k$$

with $B_j = \begin{bmatrix} a & \lambda_j(P) \\ d\lambda_j(P) & b \end{bmatrix}$ for $j = 1, \ldots, t$. Then we have $\lambda_j(\operatorname{Re}(e^{i\xi}A)) = \lambda_j(\operatorname{Re}(e^{i\xi}\tilde{A}))$ for each $\xi \in [0, 2\pi)$ and $j = 1, \ldots, m$. Thus, $M_c(\operatorname{Re}(e^{i\xi}A)) = M_c(\operatorname{Re}(e^{i\xi}\tilde{A}))$ for all $\xi \in [0, 2\pi)$ and so $W_c(A) = W_c(\tilde{A})$. The result follows from the finite dimensional case.

Suppose r is infinite and $\lambda_m(P) > 0$. For $\mu_1 \geq \cdots \geq \mu_m > 0$, let

$$\mathcal{E}(\mu_1, \ldots, \mu_m) = W_{(c_1,c_k)}\left(\begin{bmatrix} a & \mu_1 \\ d\mu_1 & b \end{bmatrix}\right) + \cdots + W_{(c_m,c_{k-m+1})}\left(\begin{bmatrix} a & \mu_m \\ d\mu_m & b \end{bmatrix}\right).$$

Notice that $\mathcal{E}(\lambda_1(P), \cdots \lambda_m(P)) = W(C_1) + \cdots + W(C_m)$.

By Lemma 4.5, there exist an m-dimensional subspace $\mathcal{V}$ of $\mathcal{H}$ and $X : \mathcal{V} \to \mathcal{H}_1$ satisfying $X^*X = I_m$ and $X^*PX = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$ with $\lambda_j(P) - \varepsilon < \lambda_j \leq \lambda_j(P)$ for $j = 1, \ldots, m$. Let $Z = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \oplus I_s$. Then Z^*AZ is unitary similar to

$$\begin{bmatrix} a & \lambda_1 \\ d\lambda_1 & b \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} a & \lambda_m \\ d\lambda_m & b \end{bmatrix} \oplus \gamma I_s.$$

Note that $W_c(B) \subseteq W_c(A)$ if B is a compression of A. Applying the result for finite $r = m$, we have

$$\mathcal{E}(\lambda_1, \ldots, \lambda_m) = W_c(Z^*AZ) \subseteq W_c(A).$$

As $\lambda_j \to \lambda_j(P)$ and hence $\begin{bmatrix} a & \lambda_j \\ d\lambda_j & b \end{bmatrix} \to \begin{bmatrix} a & \lambda_j(P) \\ d\lambda_j(P) & b \end{bmatrix}$ when $\varepsilon \to 0$, we see that all the interior points of $\mathcal{E}(\lambda_1(P), \ldots, \lambda_m(P))$ lie in $W_c(A)$.

Conversely, suppose $\mu \in W_c(A)$. Then there exist a k-dimensional subspace $\mathcal{V}_1$ of $\mathcal{H}$ and $X : \mathcal{V}_1 \to \mathcal{H}$ such that $X^*X = I_k$ and $\mu \in W_c(X^*AX)$. By Lemma 4.6, there are a $(4k + \ell)$-dimensional subspace $\mathcal{V}_2$, containing $\mathcal{V}_1$ and $Y : \mathcal{V}_2 \to \mathcal{H}$ such that $Y^*Y = I_{\mathcal{V}_2}$ and Y^*AY has an operator matrix

$$\begin{bmatrix} aI_{2k} & P' \\ dP' & bI_{2k} \end{bmatrix} \oplus \gamma I_\ell \in M_{4k+\ell},$$

where $P' = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_{2k})$ is a $2k$-dimensional compression of P, with $\lambda_1 \geq \cdots \geq \lambda_{2k}$ and $\lambda_i \leq \lambda_i(P)$ for $1 \leq i \leq 2k$. Since X^*AX is a compression of Y^*AY, we have $\mu \in W_c(X^*AX) \subseteq W_c(Y^*AY)$. By the finite dimensional result, we have

$$\mu \in W_c(Y^*AY) = \mathcal{E}(\lambda_1, \ldots, \lambda_p) \subseteq \mathcal{E}(\lambda_1(P), \ldots, \lambda_p(P)).$$

So, we have shown that

$$\mathbf{int}(\mathcal{E}(\lambda_1(P), \ldots, \lambda_p(P))) \subseteq W_c(A) \subseteq \mathcal{E}(\lambda_1(P), \ldots, \lambda_p(P)).$$

If P has a p-dimensional compression $\tilde{P} = \mathrm{diag}\,(\lambda_1(P), \ldots, \lambda_p(P))$, then A is a compression of the form

$$\begin{bmatrix} I_p & \tilde{P} \\ d\tilde{P} & I_p \end{bmatrix}.$$

Thus, $\mathcal{E}(\lambda_1(P), \ldots, \lambda_p(P)) \subseteq W_c(A)$. Hence, $W_c(A) = \mathcal{E}$.

Suppose P does not have an compression of the above form. Then the $W_c(A)$ can only include boundary points of $\mathcal{E}$ if A satisfies condition (2) in Theorem 3.1, i.e., condition (2) in Lemma 3.2. Moreover, the two boundary points of $\mathcal{E}$ included in $W_c(A)$ will be the points described in (2). □

In Theorem 4.8, if $\lambda_j(P) = 0$ for some $j \leq t$, then $W(C_j) + \cdots + W(C_t)$ becomes a line segment joining

$$a\sum_{i=j}^{t} c_i + b\sum_{i=k-t+1}^{k-m+1} c_i \quad \text{and} \quad b\sum_{i=j}^{t} c_i + a\sum_{i=k-t+1}^{k-m+1} c_i.$$

Thus, $W_c(A)$ is a sum of $j-1$ nondegenerate elliptical disks with one line segment. Therefore, we have the following corollary.

Corollary 4.9. *Let $c = (c_1, \ldots, c_k)$ and $A \in \mathcal{B}(\mathcal{H})$ satisfy the hypotheses of Theorem 4.8. Then the boundary of* $\mathbf{cl}\,(W_c(A))$ *is differentiable. If $\lambda_j(P) = 0$ for some $j \leq \min\{m, r\}$, then there are exactly two flat portions on the boundary. Otherwise, there is no flat portion on the boundary.*

Corollary 4.10. *Suppose A is a generalized quadratic operator with an operator matrix in the form described in Theorem 3.1. Let $t = \min\{k, r\}$ and*

$$B_j = \begin{bmatrix} a & c\lambda_j(P) \\ d\lambda_j(P) & b \end{bmatrix} \qquad j = 1, \ldots, t.$$

Let

$$\mathcal{E} = W(B_1) + \cdots + W(B_t) + (k-t)\gamma.$$

(a) *Suppose $k \leq r+s$. If $P : \mathcal{H}_1 \to \mathcal{H}_1$ has a compression* $\mathrm{diag}\,(\lambda_1(P), \ldots, \lambda_t(P))$*, then $W_k(A) = \mathcal{E}$. Otherwise, $W_k(A) = \mathbf{int}(\mathcal{E}) \cup \{ta + (k-t)\gamma, tb + (k-t)\gamma\}$.*
(b) *If $k > r+s$, then $W_k(A)$ equals*

$$W(B_1) + \cdots + W(B_{2r+s-k}) + (k-r-s)(a+b) + s\gamma.$$

4.4. **Davis-Wielandt shell.** Define the *Davis-Wielandt shell* of A by

$$DW(A) = \{(\langle Ax, x\rangle, \langle Ax, Ax\rangle) : x \in \mathcal{H}, \langle x, x\rangle = 1\};$$

see [10, 11, 30]. Evidently, the projection of the set $DW(A)$ on the first co-ordinate is $W(A)$. So, $DW(A)$ captures more information about the operator A than $W(A)$. For example, in the finite dimensional case, normality of operators can be completely determined by the geometrical shape of their Davis-Wielandt shells, namely, $A \in M_n$ is normal if and only if $DW(A)$ is a polyhedron in $\mathbb{C} \times \mathbb{R}$ identified with $\mathbb{R}^3$. Suppose $A \in \mathcal{B}(\mathcal{H})$. It is known that if $\dim \mathcal{H} \geq 3$ then $DW(A)$ is always convex. Suppose $A = \begin{bmatrix} a & c \\ d & b \end{bmatrix} \in M_2$ is non-scalar and has eigenvalues λ_1, λ_2. If A is normal, then $DW(A)$ degenerates to the line segment joining the points $(\lambda_1, |\lambda_1|^2)$ and $(\lambda_2, |\lambda_2|^2)$; otherwise, $DW(A)$ is an $DW(A)$ is an ellipsoid (without its interior) centered at $((\lambda_1 + \lambda_2)/2, \operatorname{tr}(A^*A)/2)$.

Suppose $\dim \mathcal{H} \geq 3$. In [22], a complete description of $DW(A)$ for a quadratic operator A was given. For generalized quadratic operators, we have the following.

Theorem 4.11. *Suppose* $\dim \mathcal{H} \geq 3$ *and* $A \in \mathcal{B}(\mathcal{H})$ *is a generalized quadratic operator with operator matrix in the form in Theorem 3.1 and let* A_p *be defined as in* (2.2). *Let* $S = \{(\gamma, |\gamma|^2)\}$ *if* γI_s *is non-trivial and* $S = \emptyset$ *otherwise. Suppose* $\sigma(P) = \sigma_1(P) \cup \sigma_2(P)$, *where* $\sigma_1(P)$ *is the set of eigenvalues of* P. *Then*

$$DW(A) = \mathbf{conv}\left[\bigcup_{p\in\sigma_1(P)} DW(A_p) \cup \bigcup_{p\in\sigma_2(P)} \mathbf{int}(DW(A_p)) \cup S\right] \cup \mathcal{L}, \tag{4.7}$$

where

(i) $\mathcal{L} = \{(\mu, |\mu|^2 + \eta^2) : \mu \in \{a, b\}, \eta \in W(|c|P \oplus |d|P)\} \subseteq \partial W(A)$ *if* $a, b \in \partial DW(A)$, *i.e., condition* (2) *in Theorem* 3.1 *holds, and*

(ii) $\mathcal{L} = \emptyset$ *otherwise, i.e., condition* (1) *or* (3) *in Theorem* 3.1 *holds.*

Consequently,

$$\mathbf{cl}\,(DW(A)) = \mathbf{cl}\left(\mathbf{conv}\left[\bigcup_{p\in\sigma(P)} DW(A_p) \cup S\right]\right). \tag{4.8}$$

As pointed out by the referee, for a generalized quadratic operator A, the q-numerical range is just an elliptical disk (with all or without any boundary points), see Theorem 4.1, whereas the description of the Davis Wielandt shell is more involved. This shows the subtlety of the study.

To prove Theorem 4.11, we need the following lemmas.

Lemma 4.12. *Let* A_p *be defined as in* (2.2). *If* $T \in \mathcal{B}(\mathcal{H})$ *such that* $W(T) \subseteq W(A_p)$, *then for any unit vector* $x \in \mathcal{H}$ *there is a unit vector* $y \in \mathbb{C}^2$ *such that* $\langle Tx, x\rangle = \langle A_p y, y\rangle$ *and* $\|Tx\| \leq \|A_p y\|$.

Proof. Suppose $W(T) \subseteq W(A_p)$. By Corollary 3.3, there exists an isometry U such that $T = U^*(I \otimes A_p)U$. For any unit vector $x \in \mathcal{H}$, if $z = Ux$, then $\mu = \langle Tx, x\rangle = \langle (I \otimes A_p)z, z\rangle$ and $\|Tx\|^2 \le \|(I \otimes A_p)z\|^2 = \nu$. Since $(\mu, \nu) \in DW(I \otimes A_p) \subseteq \mathbf{conv} DW(A_p)$ and $DW(A)$ is either a line segment or an ellipsoid (without its interior), there exists $(\mu, \hat{\nu}) \in DW(A_p)$ such that $\hat{\nu} \ge \nu$. The result follows. $\square$

Lemma 4.13. *Suppose*

$$B_k = \begin{bmatrix} aI & cP_k \\ dP_k & bI \end{bmatrix} \in \mathcal{B}(\mathcal{V}_k \oplus \mathcal{V}_k), \tag{4.9}$$

where P_k is one of the summand in $P = \oplus_{k=1}^n P_k$ described in Theorem 2.1. Then

$$DW(B_k) \subseteq \mathbf{conv}\left[\cup_{p\in\sigma(P_k)} DW(A_p)\right]. \tag{4.10}$$

Furthermore, if $(\mu, \nu) \in DW(B_k)$ is an extreme point of $DW(B_k)$, then $(\mu, \nu) \in DW(A_p)$ for some eigenvalue p of P_k.

Proof. First, we prove (4.10). The result is clear if P_k is a singleton. So, we assume that P_k is a non-degenerate interval. Hence, $\mathcal{V}_k$ is infinite dimensional.

Suppose $(\mu, \nu) \in DW(B_k)$. Then there exists a unit vector $x \in \mathcal{V}_k \oplus \mathcal{V}_k$, such that $(\mu, \nu) = (\langle B_k x, x\rangle, \|B_k x\|^2)$. Let $x = \begin{bmatrix} \cos\theta x_1 \\ \sin\theta x_2 \end{bmatrix}$ for some unit vectors, $x_1,\ x_2 \in \mathcal{V}_k$. Let $\{u_1, u_2\} \in \mathcal{V}_k$ be an orthogonal normal family such that $x_1, x_2 \in \operatorname{span}\{u_1, u_2\}$. Let

$$U = [u_1\ u_2] : \mathbb{C}^2 \to \mathcal{V}_k \quad \text{and } X = \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} : \mathbb{C}^4 \to \mathcal{V}_k \oplus \mathcal{V}_k.$$

Then $X^*X = I_4$ and

$$C_k = X^* B_k X = \begin{bmatrix} aI_2 & cU^*P_kU \\ dU^*P_kU & bI_2 \end{bmatrix} \in M_4$$

is a compression of A. Clearly, $\sigma(U^*P_kU) \subseteq \sigma(P_k)$. Now, suppose $v_1, v_2 \in \mathbb{C}^2$ are unit vectors such that $x_1 = Uv_1$ and $x_2 = Uv_2$. Then for $v = \begin{bmatrix} \cos\theta v_1 \\ \sin\theta v_2 \end{bmatrix}$, we have

$$\langle C_k v, v\rangle = \langle B_k x, x\rangle = \mu.$$

Furthermore, if $\tilde{\nu} = \|C_k v\|^2$, then $(\mu, \nu) \in DW(C_k)$ with

$$\begin{aligned}
\nu - \tilde{\nu} &= \langle B_k^* B_k x, x\rangle - \langle C_k^* C_k x, x\rangle \\
&= |d|^2(v_1^* U^* P_k^2 U v_1 - v_1^*(U^*P_kU)^2 v_1) + |c|^2(v_2^* U^* P_k^2 U v_2 - v_2^*(U^*P_kU)^2 v_2) \\
&= |d|^2(\|P_k x_1\|^2 - \|U^* P_k x_1\|^2) + |c|^2(\|P_k x_2\|^2 - \|U^* P_k x_2\|^2) \\
&\ge 0.
\end{aligned}$$

Let $p = \max\sigma(P_k)$. Then by Lemma 3.2, $W(B_k) \subseteq W(A_p)$. By Lemma 4.12, there is $\eta \ge \nu$ such that $(\mu, \eta) \in DW(A_p)$. Since C_k is unitarily similar to $A_{q_1} \oplus A_{q_2}$ for some $q_1,\ q_2 \in \sigma(P_k)$, it follows that $(\mu, \eta) \in DW(A_p)$, $(\mu, \tilde{\nu}) \in \mathbf{conv}\,[DW(A_{q_1}) \cup DW(A_{q_2})]$, and

$$(\mu, \nu) \in \mathbf{conv}\{(\mu, \eta), (\mu, \tilde{\nu})\} \subseteq \mathbf{conv}\left[\cup_{q\in\sigma(P_k)} DW(A_q)\right].$$

Suppose $(\mu, \nu) \in DW(B_k)$ is an extreme point of $DW(B_k)$. Following the argument in the previous paragraph, we see that

$$(1)\ \nu = \tilde{\nu}, \quad \text{or} \quad (2)\ \eta = \nu \text{ and } p \in \{q_1, q_2\}.$$

Suppose (1) holds. Let $\mathcal{S}$ be the subspace spanned by $\{u_1, u_2\}$. Analyzing the equality condition for $\nu - \tilde{\nu} \geq 0$, we see that both $P_k x_1$ and $P_k x_2$ lie in $\mathcal{S}$. So $\mathcal{S}$ is a reducing subspace of P_k, and consequently, a reducing subspace of P. We can choose u_1 and u_2 to be eigenvectors of P corresponding to eigenvalues p_1 and p_2 respectively. Then C_k is a direct summand of A and unitarily similar to $A_{p_1} \oplus A_{p_2}$. Since (μ, ν) is an extreme point of $\mathbf{cl}\,(DW(A))$, it must lie in $DW\,(A_{p_i})$ for some $i = 1, 2$.

Suppose (2) holds, and assume that $p = q_1$. In particular, we may assume that $A_p = \begin{bmatrix} a & cp \\ dp & b \end{bmatrix}$ is a principal submatrix of C_k, and P is unitarily similar to $[p] \oplus \hat{P}$. Thus, p is an eigenvalue of P and $(\mu, \nu) \in DW(A_p)$. □

Proof of Theorem 4.11.

We will first prove the equality (4.8) and then use the result to prove (4.7).

Step 1. We prove the inclusion "$\supseteq$" of (4.8). Suppose $(\mu, \nu) \in DW(A_p)$, with $p \in \sigma(P)$. Then there is a sequence of unit vectors $\{x_m\}$ in $\mathcal{H}_1$ such that $\|Px_m - px_m\| \to 0$ so that $\langle Px_m, x_m \rangle = p_m \to p$. Thus, for each m, P is unitarily similar to

$$P_m = \begin{bmatrix} p_m & \delta_m & 0 \\ \delta_m & * & * \\ 0 & * & * \end{bmatrix},$$

where $\delta_m \geq 0$ and $\delta_m \to 0$. Hence, A is untarily similar to

$$T_m = \begin{bmatrix} aI_r & cP_m \\ dP_m & bI_r \end{bmatrix} \oplus \gamma I_s.$$

Then, for any unit vector $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathbb{C}^2$, we can extend it to $\tilde{u}_1 = \begin{bmatrix} u_1 \\ 0 \end{bmatrix}, \tilde{u}_2 = \begin{bmatrix} u_2 \\ 0 \end{bmatrix} \in \mathcal{H}_1$ so that for $\tilde{u} = \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \\ 0 \end{bmatrix} \in \mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$. It follows that

$$(\langle T_m \tilde{u}, \tilde{u} \rangle, \|T_m \tilde{u}\|^2) \to (\langle A_p u, u \rangle, \|Au\|^2).$$

Thus, $\mathbf{cl}\, DW(A) \supseteq DW(A_p)$. By convexity of $DW(A)$, we get the inclusion "$\supseteq$" of (4.8).

Step 2. We prove the inclusion "$\subseteq$" of (4.8). Since $DW(X \oplus Y) = \mathbf{conv}\{DW(X) \cup DW(Y)\}$, we may assume that γI_s is vacuous. So, it suffices to show that

$$DW(B_k) \subseteq \mathbf{conv}\left[\cup_{p \in \sigma(P_k)} DW(A_p)\right],$$

which follows from Lemma 4.13.

By Step 1 and Step 2, we get the equality (4.8). Next, we turn to the equality (4.7).

Step 3. We prove the inclusion "$\supseteq$" of (4.7). By the description of $\mathbf{cl}\,(DW(A))$ and the fact that $DW(A)$ has non-empty relative interior as $cdP \neq 0$, we see that

$$DW(A) \supseteq \mathbf{conv}\left[\bigcup_{p\in\sigma_1(P)} DW(A_p) \cup \bigcup_{p\in\sigma_2(P)} \mathbf{int}(DW(A_p)) \cup S\right].$$

Suppose condition (2) of Theorem 3.1 holds. We will show that $\mathcal{L} \subseteq DW(A)$. Note that one can apply the construction in Step 1 to get the sequence of operators $\{T_m\}$ so that

$$\{p_m\} \to \hat{p} \in \{\max\sigma(P), \min\sigma(P)\}.$$

Let $e_1 = \begin{bmatrix}1\\0\end{bmatrix} \in \mathcal{H}_1$. Then for $v_1 = \begin{bmatrix}e_1\\0\\0\end{bmatrix}, v_2 = \begin{bmatrix}0\\e_1\\0\end{bmatrix} \in \mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$, we have

$$(\langle T_m v_1, v_1\rangle, \|T_m v_1\|^2) \to (a, |a|^2 + |d\hat{p}|^2) \quad \text{and} \quad (\langle T_m v_2, v_2\rangle, \|T_m v_2\|^2) \to (b, |b|^2 + |c\hat{p}|^2).$$

Note that there is a unitary $U = (u_{ij}) \in M_2$ such that

$$U\begin{bmatrix} a & cp_m \\ dp_m & b\end{bmatrix}U^* = \begin{bmatrix} b & cp_m \\ dp_m & a\end{bmatrix}.$$

Extend U to

$$V = \begin{bmatrix} u_{11} & 0 & u_{12} & 0 \\ 0 & I_{r-1} & 0 & 0 \\ u_{21} & 0 & u_{22} & 0 \\ 0 & 0 & 0 & I_{r-1}\end{bmatrix} \oplus I_s \in \mathcal{B}(\mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2).$$

Then

$$\hat{T}_m = VT_mV^* = \begin{bmatrix} b & du_{12}\delta_m & 0 & cp_m & cu_{11}\delta_m & 0 \\ c\bar{u}_{12}\delta_m & a & 0 & cu_{22}\delta_m & * & * \\ 0 & 0 & aI_{r-2} & 0 & * & * \\ dp_m & du_{22}\delta_m & 0 & a & cu_{21}\delta_m & 0 \\ d\bar{u}_{11}\delta_m & * & * & d\bar{u}_{21}\delta_m & b & 0 \\ 0 & * & * & 0 & 0 & bI_{r-2}\end{bmatrix} \oplus \gamma I_s.$$

For $e_1 = \begin{bmatrix}1\\0\end{bmatrix} \in \mathcal{H}_1$, and $v_1 = \begin{bmatrix}e_1\\0\\0\end{bmatrix}, v_2 = \begin{bmatrix}0\\e_1\\0\end{bmatrix} \in \mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$, we have

$$(\langle \hat{T}_m v_1, v_1\rangle, \|\hat{T}_m v_1\|^2) \to (b, |b|^2 + |d\hat{p}|^2) \quad \text{and} \quad (\langle \hat{T}_m v_2, v_2\rangle, \|\hat{T}_m v_2\|^2) \to (a, |a|^2 + |c\hat{p}|^2).$$

Hence, $(\mu, |\mu|^2 + \eta^2) \in \mathbf{cl}\,(DW(A))$, for $\mu \in \{a, b\}$ and $\eta \in \{\max\sigma(P), \min\sigma(P)\}$. By the convexity of $DW(A)$, the relative interior of $\mathcal{L} \subseteq DW(A)$. Moreover, if $\hat{p} \in \{\max\sigma(A), \min\sigma(A)\}$ and $\hat{p} \in W(P)$, then there is a unit vector $u \in \mathcal{H}_1$ such that $\|P\| = \langle Pu, u\rangle = \|Pu\|\|u\| \le \|P\|$. As a result, A is unitarily similar to $A_{\hat{p}} \oplus \hat{A}$. In particular, $DW(A_{\hat{p}}) \subseteq DW(A)$ so that $(\mu, |\mu|^2 + \hat{p}^2) \in DW(A)$. Thus, we see that $\mathcal{L} \subseteq DW(A)$.

Step 4. We prove the inclusion "$\subseteq$" of (4.7). Suppose $(\mu_0, \nu_0) \in \partial DW(A) \cap DW(A)$. Then there is a support plane $\mathbf{P}$ of the convex set $\mathbf{cl}\,(DW(A))$ passing through $(\mu_0, \nu_0) \subseteq \mathbb{C} \times \mathbb{R} \sim \mathbb{R}^3$.

So, there exist real numbers f, g, h, λ such that

$$\lambda = \mathrm{Re}\,((f+ig)\mu_0) + h\nu_0 = \max\{\mathrm{Re}\,((f+ig)\mu) + h\nu : (\mu,\nu) \in DW(A)\}.$$

Now, if $\mathbf{P}$ is a support plane for $DW(A_p)$, then there is a unit vector $x \in \mathbb{C}^2$ such that

$$\lambda = \mathrm{Re}\,((f+ig)(x^*A_px)) + hx^*A_p^*A_px \geq \mathrm{Re}\,((f+ig)(y^*A_py)) + hy^*A_p^*A_py, \quad y \in \mathbb{C}^2, y^*y = 1.$$

Thus, λ is the largest eigenvalue of

$$L_p = \mathrm{Re}\,((f+ig)A_p) + hA_p^*A_p = \begin{bmatrix} a_{11} + h|d|^2p^2 & a_{12}p \\ a_{21}p & a_{22} + h|c|^2p^2 \end{bmatrix}$$

and hence

$$0 = \det(\lambda I - L_p) = |hcd|^2p^4 + t_1p^2 + t_2, \tag{4.11}$$

where $t_1 = -h[(\lambda - a_{11})|c|^2 + (\lambda - a_{22})|d|^2] - a_{12}a_{21}, t_2 = (\lambda - a_{11})(\lambda - a_{22}) \in \mathbb{R}$. If $h \neq 0$, then $p^2 = [-t_1 \pm \sqrt{t_1^2 - 4t_2|hcd|^2}]/2|hcd|^2$, and there are at most two distinct $p \geq 0$ satisfying (4.11). If $h = 0$ and $t_1 \neq 0$, then $p^2 = -t_2/t_1$ and there is at most one $p \geq 0$ satisfying (4.11). If $h = t_1 = 0$ and $t_2 \neq 0$, then there is no $p \geq 0$ satisfying (4.11). Finally, if $h = t_1 = t_2 = 0$, we see that (4.11) holds for all $p \geq 0$.

Note that for any unit vector $x \in \mathcal{H}$, we can decomposed x according to the range space of B_k. It follows that every element $(\mu,\nu) = (\langle Ax, x\rangle, \|Ax\|^2)$ in $DW(A)$ can be written as $(\mu,\nu) = \sum_k t_k(\mu_k,\nu_k)$ with $(\mu_k,\nu_k) \in DW(B_k)$, where the sum could be infinite. In any event, if $(\mu_0,\nu_0) \in \partial DW(A) \cap \mathbf{P}$, where $\mathbf{P}$ is a support plane of $DW(A)$, then (μ_0,ν_0) is a convex combination of elements in $DW(A_p) \cap \mathbf{P}$ for some $p \in \sigma(A)$ such that $DW(A_p) \cap P \neq \emptyset$. By the previous discussion, we have the following three cases.

Case 1 If the support plane $\mathbf{P}$ of $DW(A)$ has intersection with $DW(A_p)$ for only one $p \in \sigma(P)$ and if $p \in \sigma(P_k)$, then we have $\mathbf{P} \cap DW(B_k) = \{(\mu_0,\nu_0)\}$. Hence, (μ_0,ν_0) is an extreme point of $DW(B_k)$. By Lemma 4.13, p is an eigenvalue of P_k.

Case 2 If the support plane $\mathbf{P}$ of $DW(A)$ has intersection with $DW(A_p)$ and $DW(A_q)$ for two different $p, q \in \sigma(P)$, then $\mathbf{P} \cap DW(A) = \mathbf{conv}\,[(\mathbf{P} \cap DW(A_p)) \cup (\mathbf{P} \cap DW(A_q))]$. If (μ_0,ν_0) is one of the two points in the set $(\mathbf{P} \cap DW(A_p)) \cup (\mathbf{P} \cap DW(A_q))$, then (μ_0,ν_0) is the extreme point of $DW(B_k)$ for some B_k defined as in Lemma 4.13. If (μ_0,ν_0) is a non-trivial combination of the two points in the set $(\mathbf{P} \cap DW(A_p)) \cup (\mathbf{P} \cap DW(A_q))$, then the two points must lie in $DW(A)$. Moreover, they must be extreme points of $DW(B_k)$ and $DW(B_\ell)$, where B_k and B_ℓ defined as in Lemma 4.13 such that $p \in \sigma(P_k)$ and $q \in \sigma(P_\ell)$. By Lemma 4.13, p, q are eigenvalues of P_k and P_ℓ, respectively.

In both Cases 1 and 2, we conclude that $(\mu_0,\nu_0) \in \mathbf{conv}\left[\bigcup_{p\in\sigma_1(P)} DW(A_p)\right]$.

Case 3 If $(\mu_0,\nu_0) \in DW(A_p)$ for all $p \geq 0$, then $h = 0$ so that $\mu_0 \in \cap_{p\geq 0} W(A_p)$. It follows that condition (2) of Lemma 3.2 holds, and $\mu_0 \in \{a, b\}$. By the argument in Step 3, we have $(\mu_0,\nu_0) \in \mathcal{L} \subseteq \mathbf{conv}\left[\bigcup_{p\in\sigma_1(P)} DW(A_p)\right]$. The result follows. □

5. Additional remarks and further research

We may extend most of our results to $A \in \mathcal{B}(\mathcal{H})$ of the form in (2.1) without requiring that $\gamma \in \{a,b\}$, using the simple fact that $\sigma(X \oplus Y) = \sigma(X) \cup \sigma(Y)$, $W(X \oplus Y) = \mathbf{conv}\{W(X) \cup W(Y)\}$, $\|X \oplus Y\| = \max\{\|X\|, \|Y\|\}$, $DW(X \oplus Y) = \mathbf{conv}\{DW(X) \cup DW(Y)\}$, etc.

We may consider $A \in \mathcal{B}(\mathcal{H})$ with an operator matrix of the form

$$\begin{bmatrix} aI & R \\ S & bI \end{bmatrix} \tag{5.1}$$

so that RS and SR are normal. If A is a Hilbert-Schmidt operator, i.e., A^*A has a bounded trace, then $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2$ and A is unitarily similar to

$$\begin{bmatrix} aI_r & D_1 \\ D_2 & bI_r \end{bmatrix} \oplus \gamma I_s, \tag{5.2}$$

where $D_1, D_2 \in \mathcal{B}(\mathcal{H}_1)$ are commuting normal operators, and $\gamma \in \{a,b\}$.

To see this, we may first find U, V such that URV has the form (a), (b), or (c) as in the proof of Theorem 2.1, we see that A is unitarily similar to

$$\begin{bmatrix} aI_r & P \\ Q & bI_r \end{bmatrix} \oplus \gamma I_s$$

where $P, Q \in \mathcal{B}(\mathcal{H}_1)$ with $\gamma \in \{a,b\}$. We may further assume that $P = P_1 \oplus 0$ so that $\ker(P_1) = \{0\}$. Note that P_1 is compact and has discrete eigenvalues. Thus, we may assume that P_1 is a direct sum of $p_j I_{r_j}$ such that $p_1 > p_2 > \cdots$. Note that PQ and QP are normal. Thus, $Q = Q_1 \oplus 0$ so that P_1Q_1 and Q_1P_1 are normal. Now, decompose Q_1 according to the decomposition of P_1. Using the equalities $P_1Q_1Q_1^*P_1 = Q_1^*P_1P_1Q_1$ and $Q_1P_1P_1Q_1^* = P_1Q_1^*Q_1P$, we see that Q_1 is also a direct sum of $\hat{Q}_j$ and each $\hat{Q}_j$ is normal. We get the decomposition (5.2).

With the decomposition (5.2), we easily show that $\sigma(A)$, $W(A)$, $\rho(A)$, $w(A)$, $\|A\|$, etc. are completely determined by matrices of the form

$$A(u,v) = \begin{bmatrix} a & u \\ v & b \end{bmatrix}, \qquad (u,v) \in \sigma(D_1, D_2),$$

where $\sigma(D_1, D_2)$ is the joint spectrum of (D_1, D_2) defined as the collection of $(u,v) \in \mathbb{C} \times \mathbb{C}$ such that $\|Rx_m - ux_m\| + \|Sx_m - vx_m\| \to 0$ for a sequence of unit vectors $\{x_m\} \subseteq \mathcal{H}_1$. For instance, one has

$$\mathbf{cl}\,(W_q(A)) = \mathbf{conv}\left[\mathbf{cl}\,\left(\cup_{(u,v)\in\sigma(D_1,D_2)} W_q(A(u,v))\right)\right], \quad q \in [0,1],$$

and

$$\mathbf{cl}\,(DW(A)) = \mathbf{conv}\left[\mathbf{cl}\,\left(\cup_{(u,v)\in\sigma(D_1,D_2)} DW(A(u,v))\right)\right].$$

If there is $(\tilde{u}, \tilde{v}) \in \sigma(D_1, D_2)$ such that

$$W(A(u,v)) \subseteq W(A(\tilde{u},\tilde{v})) \quad \text{for all } (u,v) \in \sigma(D_1,D_2),$$

then we have $\mathbf{cl}\,(W(A)) = W(A(\tilde{u},\tilde{v}))$, $\mathbf{cl}\,(W_q(A)) = W(A_q(\tilde{u},\tilde{v}))$, A has a dilation of the form $I \otimes A(\tilde{u},\tilde{v})$, $\|A\| = \|A(\tilde{u},\tilde{v})\|$, and the matricial range result holds, namely, $B \in W^n(A)$ if and only if $W(B) \subseteq W(A(\tilde{u}, \tilde{v}))$. Furthermore, if A is unitarily similar to a direct sum of $A(u_j, v_j)$ for $j = 1, \ldots, m$, and $\hat{A}$ such that

$$W(\hat{A}) \subseteq W(A(u_m, v_m)) \subseteq W(A(u_{m-1}, v_{m-1})) \subseteq \cdots \subseteq W(A(u_1, v_1)),$$

then we can obtain results for the rank-k numerical range, essential numerical range (if $m = \infty$), c-numerical range.

It would be nice to show that the decomposition (5.2) holds also for general operators with an operator matrix of the form (5.1).

References

1. J. Anderson and J.G. Stampfli, Commutators and compressions, Israel J. Math. 10 (1971), 433-441.
2. Y.H. Au-Yeung and N.K. Tsing, A conjecture of Marcus on the generalized numerical range, Linear and Multilinear Algebra 14 (1983), 235-239.
3. M.T. Chien, S.H. Tso, and P.Y. Wu, Higher-dimensional numerical ranges of quadratic operators, J. Operator Theory 49 (2003), no. 1, 153–171.
4. M.D. Choi, J.A. Holbrook, D. W. Kribs, and K. Życzkowski, Higher-rank numerical ranges of unitary and normal matrices, Operators and Matrices 1 (2007), 409-426.
5. M.D. Choi, D. W. Kribs, and K. Życzkowski, Higher-rank numerical ranges and compression problems, Linear Algebra Appl., 418 (2006), 828–839.
6. M.D. Choi, D. W. Kribs, and K. Życzkowski, Quantum error correcting codes from the compression formalism, Rep. Math. Phys., 58 (2006), 77–91.
7. M.D. Choi and C.K. Li, Numerical ranges and dilations, Linear and Multilinear Algebra 47(2000), 35–48.
8. M.D. Choi and C.K. Li, Norm bounds for summation of two normal matrices, Linear Algebra Appl. 379 (2004), 137-157.
9. M.D. Choi and C.K. Li, The Ultimate Estimate of the Upper Norm Bound for the Summation of Operators, Journal of Functional Analysis 232 (2006), 455-476.
10. C. Davis, The shell of a Hilbert-space operator, Acta Sci. Math.(Szeged) 29 (1968), 69-86.
11. C. Davis, The shell of a Hilbert-space operator. II, Acta Sci. Math. (Szeged) 31 (1970), 301-318.
12. T. Furuta, *Invitation to linear operators*, Taylor & Francis, London, 2001.
13. T. Furuta, Applications of polar decompositions of idempotent and 2-nilpotent operators, Linear and Multilinear Algebra 56(2008), 69–79.
14. S.R. Garcia, Approximate antilinear eigenvalue problems and related inequalities, Proc. Amer. Math. Soc. 136 (2008), 171–179.
15. K. E. Gustafson and D. K. M. Rao *Numerical range: The field of values of linear operators and matrices*, Universitext, Springer-Verlag, New York, 1997.
16. P. Halmos, A Hilbert Space Problem Book, Second edition, Graduate Texts in Mathematics, 19, Encyclopedia of Mathematics and its Applications, Spring-Verlag, New York, 1982.
17. R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
18. C.K. Li, A simple proof of the elliptical range theorem, Proc. Amer. Math. Soc. 124 (1996), 1985-1986.
19. C.K. Li and H. Nakazato, Some results on the q-numerical ranges, Linear and Multilinear Algebra 43 (1998), 385-410.
20. C.K. Li, Y.T. Poon and N.S. Sze, Higher rank numerical ranges and low rank perturbation of quantum channels, J. Math. Anal. Appl. 348 (2008), 843-855.
21. C.K. Li, Y.T. Poon and N.S. Sze, Condition for the higher rank numerical range to be non-empty, Linear and Multilinear Algebra 57 (2009), 365-368.
22. C.K. Li, Y.T. Poon and N.S. Sze, Elliptical range theorems for generalized numerical ranges of quadratic operators, Rocky Mountain J. Math., to appear.
   Preprint available at `http://www.people.wm.edu/~cklixx/quadra.pdf`.

23. C.K. Li and N.S. Sze, Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations, Proc. Amer. Math. Soc. 136 (2008), 3013–3023.
24. M. Marcus, Some combinatorial aspects of numerical range, Second International Conference on Combinatorial Mathematics, Ann. New York Acad. Sci., 319 (1979), 368–376.
25. R.A. Martinez-Avendano, Higher-rank numerical range of operators in infinite-dimensional Hilbert space, Operators and Matrices 2 (2008), 249–264.
26. A. Pietsch, Eigenvalues and s-numbers, Cambridge Univ. Press, Cambridge, 1987.
27. L. Rodman and I. Spitkovsky, On generalized numerical ranges of quadratic operators, Operator theory: Advances and Applications 179 (2008), 241–256.
28. N.K. Tsing, The constrained bilinear form and the C-numerical range, Linear Algebra Appl. 56 (1984), 195-206.
29. S.-H. Tso and P. Y. Wu, Matricial ranges of quadratic operators, Rocky Mountain J. Math. 29 (1999), 1139-1152.
30. H. Wielandt, On eigenvalues of sums of normal matrices, Pacific J. Math. 5 (1955), 633–638.

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

Department of Mathematics, Iowa State University, Ames, IA 50051
E-mail address: ytpoon@iastate.edu

Department of Information Systems and Management, Faculty of Applied Information Science, Hiroshima Institute of Technology, 2-1-1 Miyake, Saeki-Ku, Hiroshima 731-5193, Japan
E-mail address: m.tominaga.3n@it-hiroshima.ac.jp