

CANONICAL FORMS, HIGHER RANK NUMERICAL RANGES, TOTALLY ISOTROPIC SUBSPACES, AND MATRIX EQUATIONS

CHI-KWONG LI AND NUNG-SING SZE

(Communicated by N. Tomczak-Jaegermann)

ABSTRACT. The results on matrix canonical forms are used to give a complete description of the higher rank numerical range of matrices arising from the study of quantum error correction. It is shown that the set can be obtained as the intersection of closed half planes (of complex numbers). As a result, it is always a convex set in \mathbb{C} . Moreover, the higher rank numerical range of a normal matrix is a convex polygon determined by the eigenvalues. These two consequences confirm the conjectures of Choi et al. on the subject. In addition, the results are used to derive a formula for the optimal upper bound for the dimension of a totally isotropic subspace of a square matrix and to verify the solvability of certain matrix equations.

1. INTRODUCTION

Let M_n be the algebra of $n \times n$ complex matrices. In [3], the authors introduced the notion of the *rank- k numerical range* of $A \in M_n$ defined and denoted by

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

in connection with the study of quantum error correction; see [4]. Evidently, $\lambda \in \Lambda_k(A)$ if and only if there is a unitary matrix $U \in M_n$ such that U^*AU has λI_k as the leading principal submatrix. When $k = 1$, this concept reduces to the classical numerical range, which is well known to be convex by the Toeplitz-Hausdorff theorem; for example, see [8] for a simple proof. In [1] the authors conjectured that $\Lambda_k(A)$ is convex and reduced the convexity problem to the problem of showing that $0 \in \Lambda_k(T)$ for

$$T = \begin{pmatrix} I_k & X \\ Y & -I_k \end{pmatrix}$$

for arbitrary $X, Y \in M_k$. They further reduced this problem to the existence of a Hermitian matrix H satisfying the matrix equation

$$(1.1) \quad I_k + MH + HM^* - HPH = H$$

Received by the editors March 26, 2007.

2000 *Mathematics Subject Classification*. Primary 15A21, 15A24, 15A60, 15A90, 81P68.

Key words and phrases. Canonical forms, higher rank numerical range, convexity, totally isotropic subspace, matrix equations.

The research of Li was partially supported by an NSF grant and an HK RGC grant. He is an honorary professor of the University of Hong Kong.

©2008 American Mathematical Society
Reverts to public domain 28 years from publication

for arbitrary $M \in M_k$ and positive definite $P \in M_k$. In [12], the author observed that equation (1.1) can be rewritten as the continuous Riccati equation

$$(1.2) \quad HPH - H(M^* - I_k/2) - (M - I_k/2)H - I_k = 0_k,$$

and existing results on the Riccati equation will ensure its solvability; for example, see [7, Theorem 4]. This establishes the convexity of $\Lambda_k(A)$.

Denote by $\lambda_k(H)$ the k th largest eigenvalue of the Hermitian matrix $H \in M_n$. We will use results on canonical forms of complex square matrices to show that

$$\Lambda_k(A) = \bigcap_{\xi \in [0, 2\pi)} \{ \mu \in \mathbb{C} : e^{i\xi} \mu + e^{-i\xi} \bar{\mu} \leq \lambda_k(e^{i\xi} A + e^{-i\xi} A^*) \}.$$

Thus, $\Lambda_k(A)$ is the intersection of closed half planes on the complex plane and is therefore a convex set. Furthermore, specializing our result to normal matrices confirms the conjecture in [2] asserting that

$$\Lambda_k(A) = \bigcap_{1 \leq j_1 < \dots < j_{n-k+1} \leq n} \text{conv} \{ \lambda_{j_1}, \dots, \lambda_{j_{n-k+1}} \}$$

if $A \in M_n$ is a normal matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. In addition, from our results one can derive a formula for the optimal upper bound for the dimension of a totally isotropic subspace of a square matrix. As shown in [1], the convexity of the higher rank numerical range is closely related to the study of solvability of matrix equations. Following the idea in [1], we study the solvability of certain matrix equations including those of the form (1.1), (1.2) and

$$(1.3) \quad I_k + RZ + Z^*R^* - Z^*Z = 0_k$$

for a given $k \times k$ matrix R . In particular, it is shown that there is always a common solution Z satisfying a pair of equations of the form (1.3). In other words, given two matrices $R, S \in M_k$, the operator spheres

$$\{ Z : |Z - R^*| = \sqrt{I_k + RR^*} \} \quad \text{and} \quad \{ Z : |Z - S^*| = \sqrt{I_k + SS^*} \}$$

always have a nonempty intersection; here $|X|$ is the positive semidefinite square root of X^*X .

The following results on canonical forms of matrices will be used in our discussion; for example, see [11] and [6].

- I. *QR decomposition*: For every $A \in M_n$, there is a unitary matrix $Q \in M_n$ and an upper triangular matrix $R \in M_n$ such that $A = QR$.
- II. *CS decomposition*: For every unitary $U \in M_{2k}$, there are unitary matrices $V = V_1 \oplus V_2$ and $W = W_1 \oplus W_2$ with $V_1, V_2, W_1, W_2 \in M_k$ such that

$$VUW = \begin{pmatrix} C & \sqrt{I_k - C^2} \\ \sqrt{I_k - C^2} & -C \end{pmatrix},$$

where $C = \text{diag}(c_1, \dots, c_k)$ with $1 \geq c_1 \geq \dots \geq c_k \geq 0$.

- III. **-congruence canonical form*: For every $A \in M_n$, there is an invertible $S \in M_n$ such that S^*AS is a direct sum of the following three types of matrices:

- (1) $\Gamma_{2r}(\mu) = \begin{pmatrix} 0_r & I_r \\ J_r(\mu) & 0_r \end{pmatrix} \in M_{2r}$ with $|\mu| > 1$ for $\mu \in \mathbb{C}$, where $J_r(\mu)$ is the $r \times r$ upper triangular Jordan block with eigenvalue μ ;
- (2) $J_s(0)$, the $s \times s$ upper triangular Jordan block with eigenvalue zero;

- (3) $e^{i\xi}\Delta_t$ with $\xi \in [0, 2\pi)$, where Δ_t is the $t \times t$ matrix whose (p, q) entry equals 1 if $p + q = t + 1$, equals the imaginary unit i if $p + q = t + 2$ and equals 0 otherwise; in particular, $\Delta_1 = [1]$.

2. HIGHER RANK NUMERICAL RANGE

Definition 2.1. For $A \in M_n$, let $\Omega_k(A)$ be the set of $\mu \in \mathbb{C}$ such that for each $\xi \in [0, 2\pi)$, the Hermitian matrix $e^{i\xi}(A - \mu I_n) + e^{-i\xi}(A - \mu I_n)^*$ has at least k nonnegative eigenvalues. In particular, if $\lambda_k(H)$ denotes the k th largest eigenvalue of a Hermitian matrix $H \in M_n$, then

$$\Omega_k(A) = \bigcap_{\xi \in [0, 2\pi)} \{ \mu \in \mathbb{C} : e^{i\xi}\mu + e^{-i\xi}\bar{\mu} \leq \lambda_k(e^{i\xi}A + e^{-i\xi}A^*) \}.$$

When $k = 1$, it is well known that the classical numerical range $\Lambda_1(A)$ can be obtained by intersecting the closed half planes

$$\{ \mu \in \mathbb{C} : e^{i\xi}\mu + e^{-i\xi}\bar{\mu} \leq \lambda_1(e^{i\xi}A + e^{-i\xi}A^*) \}, \quad \xi \in [0, 2\pi).$$

We will show that $\Lambda_k(A) = \Omega_k(A)$, which extends the classical result. In particular, one can easily write a computer program to draw the boundary $\partial\Omega_k(A)$ of $\Omega_k(A)$, and it is clear that for $A \in M_n$, the convex curve $\partial\Omega_k(A)$ lies inside the convex curve $\partial\Omega_{k-1}(A)$ if $k > 1$.

If A is Hermitian, then we have the nested intervals

$$\Omega_1(A) \supseteq \Omega_2(A) \supseteq \Omega_3(A) \supseteq \dots$$

If A is normal with eigenvalues $\lambda_1, \dots, \lambda_n$, then

$$\Omega_k(A) = \bigcap_{1 \leq j_1 < \dots < j_{n-k+1} \leq n} \text{conv} \{ \lambda_{j_1}, \dots, \lambda_{j_{n-k+1}} \},$$

as described in [2, Section 3]. To see this, note that

$$\mu \notin \bigcap_{1 \leq j_1 < \dots < j_{n-k+1} \leq n} \text{conv} \{ \lambda_{j_1}, \dots, \lambda_{j_{n-k+1}} \}$$

if and only if there is a line passing through 0 such that $n - k + 1$ eigenvalues of $A - \mu I_n$ lie on one side of the open half plane determined by the line; equivalently, there is $\xi \in [0, 2\pi)$ such that $n - k + 1$ eigenvalues of $e^{i\xi}(A - \mu I_n) + e^{-i\xi}(A - \mu I_n)^*$ are negative.

Recall that we use $\lambda_k(H)$ to denote the k th largest eigenvalue of a Hermitian matrix $H \in M_n$. Our main theorem is the following.

Theorem 2.2. *Let $A \in M_n$. Then*

$$\Lambda_k(A) = \Omega_k(A) = \bigcap_{\xi \in [0, 2\pi)} \{ \mu \in \mathbb{C} : e^{i\xi}\mu + e^{-i\xi}\bar{\mu} \leq \lambda_k(e^{i\xi}A + e^{-i\xi}A^*) \}.$$

Since the intersection of half planes in \mathbb{C} is a convex set, the following corollary is immediate.

Corollary 2.3. *Let $A \in M_n$. Then the rank- k numerical range $\Lambda_k(A)$ is convex.*

By the discussion on normal matrices before Theorem 2.2, we have the following corollary confirming the conjecture in [2].

Corollary 2.4. *Let $A \in M_n$ be a normal matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Then*

$$\Lambda_k(A) = \bigcap_{1 \leq j_1 < \dots < j_{n-k+1} \leq n} \text{conv} \{\lambda_{j_1}, \dots, \lambda_{j_{n-k+1}}\}.$$

To prove the theorem, we need the following lemma, which can be found in [1]. We give a short proof using the QR decomposition.

Lemma 2.5. *Let $A \in M_n$ and $1 \leq k \leq n$. Then $0 \in \Lambda_k(A)$ if and only if there is an invertible $S \in M_n$ such that S^*AS has 0_k as its leading $k \times k$ principal submatrix.*

Proof. The implication “ \Rightarrow ” is clear. Conversely, suppose there is an invertible $S \in M_n$ such that S^*AS has 0_k as the leading $k \times k$ principal submatrix. By the QR decomposition, $S = UR$, where U is unitary and R is upper triangular. Since R^{-1} is also in upper triangular form, we see that $U^*AU = (R^{-1})^*(S^*AS)R^{-1}$ also has 0_k as its leading principal submatrix. \square

We divide the proof of Theorem 2.2 into three lemmas. In particular, the construction in Lemmas 2.7 and 2.8 can be done explicitly using the results in [5, 6] and QR decomposition (which involves only the Gram-Schmidt process). Thus, for every $\mu \in \Lambda_k(A)$, one can construct a unitary matrix U such that U^*AU has μI_k as its leading principal submatrix.

Lemma 2.6. *Let $A \in M_n$. Then $\Lambda_k(A) \subseteq \Omega_k(A)$.*

Proof. Suppose $\mu \in \Lambda_k(A)$; equivalently, $0 \in \Lambda_k(B)$ for $B = A - \mu I_n$. Then $e^{i\xi}B + e^{-i\xi}B^*$ is unitarily similar to a matrix with 0_k as its leading principal submatrix. By the interlacing inequalities (for example, see [5]), $e^{i\xi}B + e^{-i\xi}B^*$ has at least k nonnegative eigenvalues. \square

Lemma 2.7. *Let $A \in M_n$ be normal. Then $\Omega_k(A) \subseteq \Lambda_k(A)$.*

Proof. Suppose $\mu \in \Omega_k(A)$. Let $B = A - \mu I_n$. Then for each $\xi \in [0, 2\pi)$, the Hermitian matrix $e^{i\xi}B + e^{-i\xi}B^*$ has at least k nonnegative eigenvalues. We show that $0 \in \Lambda_k(B)$.

We prove the result by induction on k . If $k = 1$, then the given condition ensures that 0 lies in the convex hull of the eigenvalues of B . Suppose $V \in M_n$ is unitary such that $V^*BV = \text{diag}(b_1, \dots, b_n)$ and p_1, \dots, p_n are nonnegative real numbers summing to 1 such that $\sum_{j=1}^n p_j b_j = 0$. Then $u = V(\sqrt{p_1}, \dots, \sqrt{p_n})^t$ is a unit vector such that $v^*Bv = 0$. Choose a unitary matrix $U \in M_n$ with u as the first column. Then U^*BU has zero as the $(1, 1)$ entry. So, the result holds for $k = 1$. (One can also use the convexity of the classical numerical range to get the conclusion. We include the argument so that the proof is independent of other convexity results.)

Assume that $k > 1$ and the result is valid for the rank- m numerical range of normal matrices whenever $m < k$. If B has an eigenvalue equal to 0, then there is a unitary $V \in M_n$ such that $V^*BV = [0] \oplus B_1$ so that $e^{i\xi}B_1 + e^{-i\xi}B_1^*$ has at least $k-1$ nonnegative eigenvalues for any $\xi \in [0, 2\pi)$. By the induction assumption, there is a unitary $U \in M_{n-1}$ such that U^*B_1U has 0_{k-1} as its leading principal submatrix. Then 0_k will be the leading principal submatrix of $([1] \oplus U)^*V^*BV([1] \oplus U)$. Thus, $0 \in \Lambda_k(B)$.

Now, assume that B is invertible. Then $k \leq n/2$. Suppose there is a pair of eigenvalues of B , say λ_1 and λ_2 , satisfying $\lambda_1/|\lambda_1| = e^{i\theta}$ and $\lambda_2/|\lambda_2| = -e^{i\theta} = e^{i\theta+\pi}$

for some $\theta \in [0, 2\pi)$. Then there is a unitary $V \in M_n$ such that $V^*BV = B_1 \oplus B_2$ with $B_1 = \text{diag}(\lambda_1, \lambda_2)$. Note that for each $\xi \in [0, 2\pi)$, $e^{i\xi}B_1 + e^{-i\xi}B_1^*$ has at least 1 nonnegative eigenvalue and $e^{i\xi}B_2 + e^{-i\xi}B_2^*$ has at least $k - 1$ nonnegative eigenvalues. By the induction assumption, there are unitary $U_1 \in M_2$ and $U_2 \in M_{n-2}$ such that $U_1^*B_1U_1$ and $U_2^*B_2U_2$ have 0_1 and 0_{k-1} as their leading principal submatrices, respectively. Let $U = U_1 \oplus U_2$. Then 0_k will be a principal submatrix of U^*V^*BVU lying in rows and columns $1, 3, 4, \dots, k + 1$. Thus, $0 \in \Lambda_k(B)$.

Continue to assume that B is invertible; assume in addition that no pair of eigenvalues of B has arguments θ and $\theta + \pi$.

Claim. *There is an invertible $S \in M_n$ such that S^*BS has 0_k as the leading principal submatrix.*

Once the claim is proved, we see that $0 \in \Lambda_k(B)$ by Lemma 2.5, and the induction proof will be complete.

To prove the claim, let $\xi \in [0, 2\pi)$ be such that $e^{i\xi}B + e^{-i\xi}B^*$ has the smallest number of nonnegative eigenvalues, say, k' . Then $k' \geq k$. We may assume that $k = k'$. Furthermore, we may assume that $\xi = 0$; otherwise, replace B by $e^{i\xi}B$. Apply a $*$ -congruence to B and assume that $B = H + iG$ such that $H = I_k \oplus -I_{n-k}$ and $G = \text{diag}(g_1, \dots, g_n)$ with $g_1 \geq \dots \geq g_k$ and $g_{k+1} \geq \dots \geq g_n$. Note that the given assumption on B ensures that

- (i) for every straight line passing through the origin, there are at least k eigenvalues of B lying in each of the closed half planes determined by the line, and
- (ii) there is no pair of eigenvalues of B having arguments θ and $\theta + \pi$.

We claim that $-g_n > g_1$. Otherwise, $-g_n \leq g_1$. Since condition (ii) holds, we see that $-g_n < g_1$. Moreover, the line \mathcal{L} passing through 0 and the eigenvalue $1 + ig_1$ of B will divide the plane into two parts so that k of the eigenvalues of B , namely, $1 + ig_1, \dots, 1 + ig_k$, lie below \mathcal{L} and all other eigenvalues lie in the open half plane above \mathcal{L} . We may then rotate \mathcal{L} in the clockwise direction by a very small angle so that at most $k - 1$ of the eigenvalues of B , namely, $1 + ig_2, \dots, 1 + ig_k$, will lie on the closed half plane below the resulting line, contradicting condition (i).

Similarly, we can argue that $-g_{n-1} > g_2$. Otherwise, $-g_{n-1} < g_2$, and we can rotate the line passing through 0 and $1 + ig_2$ in the clockwise direction by a very small angle so that at most $k - 1$ eigenvalues of B , namely, $1 + ig_3, \dots, 1 + ig_k$ and $-1 + ig_n$, will lie on the closed half plane below the resulting line, contradicting condition (i).

Repeating this argument, we see that

$$-g_n > g_1, \quad -g_{n-1} > g_2, \quad \dots, \quad -g_{n-k+2} > g_{k-1}, \quad -g_{n-k+1} > g_k.$$

We can use a similar argument to show that

$$-g_{k+1} < g_k, \quad -g_{k+2} < g_{k-1}, \quad \dots, \quad -g_{2k-1} < g_2, \quad -g_{2k} < g_1.$$

By [5, Theorem 1], there is a unitary $V \in M_{n-k}$ such that

$$V^*(\text{diag}(g_{k+1}, \dots, g_n))V = \begin{pmatrix} -D & * \\ * & * \end{pmatrix} \quad \text{with } D = \text{diag}(g_1, g_2, \dots, g_k).$$

Thus, the leading $2k \times 2k$ submatrix of $(I_k \oplus V)^*B(I_k \oplus V)$ equals

$$(I_k + iD) \oplus (-I_k - iD).$$

Let

$$W = \frac{1}{\sqrt{2}} \begin{pmatrix} I_k & I_k \\ I_k & -I_k \end{pmatrix} \oplus I_{n-2k}.$$

Then the leading $2k \times 2k$ submatrix of $W^*(I_k \oplus V)^*B(I_k \oplus V)W$ equals

$$\begin{pmatrix} 0_k & I + iD \\ I + iD & 0_k \end{pmatrix}.$$

So, the claim holds. □

Lemma 2.8. *For any matrix $A \in M_n$, we have $\Omega_k(A) \subseteq \Lambda_k(A)$.*

Proof. Suppose $A \in M_n$ and $\mu \in \Omega_k(A)$. Let $S \in M_n$ be such that $S^*(A - \mu I_n)S$ is a direct sum of the following matrices as defined in Section 1 (III):

- (a) $\Gamma_{2r_1}(\mu_1), \dots, \Gamma_{2r_u}(\mu_u)$;
- (b) $J_{s_1}(0), \dots, J_{s_v}(0)$, where s_1, \dots, s_p are odd and s_{p+1}, \dots, s_v are even;
- (c) $e^{i\xi_1} \Delta_{t_1}, \dots, e^{i\xi_w} \Delta_{t_w}$, where t_1, \dots, t_q are odd and t_{q+1}, \dots, t_w are even.

Let $B = S^*(A - \mu I_n)S$. For each $\xi \in [0, 2\pi)$, consider $e^{i\xi}B + e^{-i\xi}B^*$. Each type (a) direct summand has the form $e^{i\xi}\Gamma_{2r_j}(\mu_j) + e^{-i\xi}\Gamma_{2r_j}(\mu_j)^*$, which will contribute r_j nonnegative (positive) eigenvalues to $e^{i\xi}B + e^{-i\xi}B^*$. Consequently, these summands will contribute a total of $\sum_{j=1}^u r_j$ nonnegative eigenvalues to $e^{i\xi}B + e^{-i\xi}B^*$.

Each type (b) direct summand has the form $e^{i\xi}J_{s_j}(0) \oplus e^{-i\xi}J_{s_j}(0)^*$, which will contribute $[(s_j + 1)/2]$ nonnegative eigenvalues to $e^{i\xi}B + e^{-i\xi}B^*$, where $[x]$ denotes the integral part of the real number x . Consequently, these summands will contribute a total of $\frac{1}{2} \left(\sum_{j=1}^v s_j + p \right)$ nonnegative eigenvalues to $e^{i\xi}B + e^{-i\xi}B^*$.

Each type (c) direct summand has the form

$$(2.1) \quad e^{i(\xi+\xi_j)} \Delta_{t_j} + e^{-i(\xi+\xi_j)} \Delta_{t_j}^* = \begin{pmatrix} 0 & & a_j \\ & \cdot & b_j \\ & \cdot & \cdot \\ & \cdot & \cdot \\ a_j & b_j & 0 \end{pmatrix}$$

with $a_j = \cos(\xi + \xi_j)$ and $b_j = -\sin(\xi + \xi_j)$. Suppose t_j is even. Since there is a $0_{t_j/2}$ leading principal submatrix, the matrix has at least $t_j/2$ nonnegative eigenvalues. If ξ is chosen so that $a_j \neq 0$, then there will be exactly $t_j/2$ nonnegative (positive) eigenvalues. Thus, the matrix in (2.1) will contribute $t_j/2$ nonnegative eigenvalues to $e^{i\xi}B + e^{-i\xi}B^*$. Suppose t_j is odd. Then $e^{i(\xi+\xi_j)} \Delta_{t_j} + e^{-i(\xi+\xi_j)} \Delta_{t_j}^*$ is congruent to $[e^{i(\xi+\xi_j)} + e^{-i(\xi+\xi_j)}] \oplus D_j$ such that D_j has $(t_j - 1)/2$ nonnegative eigenvalues. Consequently, if ξ is chosen so that $a_j \neq 0$ in (2.1) whenever t_j is even, then these summands will contribute a total of $\frac{1}{2} \left(\sum_{j=1}^w t_j - q \right) + \ell(\xi)$ nonnegative eigenvalues to $e^{i\xi}B + e^{-i\xi}B^*$, where $\ell(\xi)$ is the number of nonnegative eigenvalues of $e^{i\xi}N + e^{-i\xi}N^*$ with

$$N = \text{diag}(e^{i\xi_1}, \dots, e^{i\xi_q}).$$

Denote by $\nu(H)$ the number of nonnegative eigenvalues of the Hermitian matrix H , and let $\ell' = \min\{\nu(e^{i\xi}N + e^{-i\xi}N^*) : \xi \in [0, 2\pi)\}$. Then there are infinitely many

choices of ξ which attain ℓ' . So, we may choose ξ to attain ℓ' with the additional assumption that $a_j \neq 0$ in (2.1) whenever t_j is even. Let

$$\begin{aligned} k' &= \sum_{j=1}^u r_j + \frac{1}{2} \left(\sum_{j=1}^v s_j + p \right) + \frac{1}{2} \left(\sum_{j=1}^w t_j - q \right) + \ell' \\ &= \min\{\nu(e^{i\xi}B + e^{-i\xi}B^*) : \xi \in [0, 2\pi)\}. \end{aligned}$$

Then $k' \geq k$. Hence, the conclusion that $0 \in \Lambda_k(A - \mu I_n)$ will follow once we show that $0 \in \Lambda_{k'}(A - \mu I_n)$.

By our assumption, $S^*(A - \mu I_n)S$ is a direct sum of the matrices listed in (a)–(c). For each direct summand Γ_{2r_j} in (a), it is clear that the leading principal submatrix is 0_{r_j} . Thus, these direct summands contain a zero principal submatrix of dimension $\sum_{j=1}^u r_u$.

For each direct summand $J_{s_j}(0)$ in (b), the principal submatrix lying in rows and columns indexed by odd numbers is a zero principal submatrix. Thus, these direct summands contain a zero principal submatrix of dimension $(\sum_{j=1}^v s_j + p)/2$.

For each direct summand $e^{i\xi_j} \Delta_{t_j}$ in (c), if t_j is even, then the leading principal submatrix is $0_{t_j/2}$; if t_j is odd, then the leading principal submatrix is $0_{(t_j-1)/2}$. Thus, these direct summands contain a zero principal submatrix of dimension $(\sum_{j=1}^w t_j - q)/2$. Moreover, these direct summands are permutationally similar to a matrix T with $0_t \oplus N$ as the $(t+q) \times (t+q)$ leading principal submatrix, where $t = (\sum_{j=1}^w t_j - q)/2$ and $N = \text{diag}(e^{i\xi_1}, \dots, e^{i\xi_q})$. By Lemma 2.7, $0 \in \Lambda_{\ell'}(N)$. Thus, there is a unitary matrix $V \in M_q$ such that V^*NV has $0_{\ell'}$ as the principal submatrix. Then $(I_t \oplus V \oplus I_{n-t-\ell'})^*T(I_t \oplus V \oplus I_{n-t-\ell'})$ has $0_{t+\ell'}$ as the leading principal submatrix.

Now combining all these zero principal submatrices yields a zero principal submatrix of dimension

$$(2.2) \quad k' = \sum_{j=1}^u r_j + \frac{1}{2} \left(\sum_{j=1}^v s_j + p \right) + \frac{1}{2} \left(\sum_{j=1}^w t_j - q \right) + \ell'.$$

The result follows. □

3. TOTALLY ISOTROPIC SUBSPACES AND MATRIX EQUATIONS

Let $A \in M_n$. A subspace \mathbf{V} of \mathbb{C}^n is a *totally isotropic subspace* of A if $x^*Ay = 0$ for any $x, y \in \mathbf{V}$. Note that $U \in M_n$ is unitary such that the first k columns of U form a totally isotropic subspace of A if and only if U^*AU has 0_k as its leading principal submatrix. One can also write $A = H + iG$ and discuss the totally isotropic subspace of the Hermitian matrix pair (H, G) , i.e., a subspace \mathbf{V} of \mathbb{C}^n such that $x^*Hy = 0 = x^*Gy$ for all $x, y \in \mathbf{V}$. It is clear that $A \in M_n$ has a totally isotropic subspace of dimension k if and only if $0 \in \Lambda_k(A)$. By Theorem 2.2, we have the following.

Theorem 3.1. *Let $A \in M_n$. Denote by $\nu(H)$ the number of nonnegative eigenvalues of the Hermitian matrix H . Then*

$$\begin{aligned} &\min\{\nu(e^{i\xi}A + e^{-i\xi}A^*) : \xi \in [0, 2\pi)\} \\ &= \max\{\dim \mathbf{V} : \mathbf{V} \text{ is a totally isotropic subspace of } A\}. \end{aligned}$$

Note that the quantity $\min\{\nu(e^{i\xi}A + e^{-i\xi}A^*) : \xi \in [0, 2\pi)\}$ is equal to k' in (2.2), where the quantities $r_1, \dots, r_u, s_1, \dots, s_v$, etc., are determined by the canonical form of A under $*$ -congruence as in the proof of Lemma 2.8 by putting $B = A - 0I$. By the result in [6], one can obtain the canonical form S^*AS by a finite algorithm using exact arithmetic.

The authors of [1] showed that the study of the convexity of the higher rank numerical range can be reduced to verifying the following lemma, which follows readily from Corollary 2.3.

Lemma 3.2. *Let $A = \begin{pmatrix} I_k & X \\ Y & -I_k \end{pmatrix} \in M_{2k}$, where $X, Y \in M_k$. Then there is a unitary $U \in M_{2k}$ such that U^*AU has 0_k as the leading $k \times k$ principal submatrix.*

Proof. Since $1, -1 \in \Lambda_k(A)$, we see that $0 \in \Lambda_k(A)$ by Corollary 2.3. \square

In [1], it was shown that the existence of U in Lemma 3.2 is equivalent to the solvability of some matrix equations; see [1, Theorem 2.12]. In the next theorem, we will use Lemma 3.2 and the CS decomposition of matrices to prove the solvability of a number of matrix equations and systems of matrix equations. The equations in (a), (d), (f) have been considered in [1]. We give slightly different proofs of them.

We consider also other matrix equations. In particular, assertion (c) of the theorem can be restated as

$$\{Z : |Z - R^*| = \sqrt{I_k + RR^*}\} \cap \{Z : |Z - S^*| = \sqrt{I_k + SS^*}\} \neq \emptyset.$$

One can use the results in [5, 6] and QR decomposition to construct the unitary matrix U in Lemma 3.2. As a result, one can give an explicit construction of the solutions of the matrix equations (a)–(d) following our proof.

It is easy to check that solvability of the equations in the next theorem is equivalent to the existence of a unitary $U \in M_{2k}$ satisfying the conclusion of Lemma 3.2.

As suggested by Professor T. Ando, it is interesting and inspiring to consider the scalar case of the statements and the proofs of the equations in the theorem.

Theorem 3.3. *Let $R, S, P, C \in M_k$ such that P is positive definite, C is a strict contraction, and $\gamma \in \mathbb{R}$.*

(a) *There is a $Z \in M_k$ such that*

$$I_k + RZ + Z^*S^* - Z^*Z = 0_k.$$

(b) *There is a $Z \in M_k$ such that*

$$I_k + RZ + Z^*R^* - Z^*Z = 0_k \quad \text{and} \quad SZ + Z^*S^* = 0_k.$$

(c) *There is a $Z \in M_k$ such that*

$$I_k + RZ + Z^*R^* - Z^*Z = 0_k \quad \text{and} \quad I_k + SZ + Z^*S^* - Z^*Z = 0_k.$$

(d) *There is a Hermitian $H \in M_k$ such that*

$$I_k + RH + HR^* - HPH = \gamma H.$$

(e) *There is a unitary $U \in M_k$ such that*

$$SR^* + RS^* = SU\sqrt{I_k + RR^*} + \sqrt{I_k + RR^*}U^*S^*.$$

(f) *There is a unitary $U \in M_k$ and a Hermitian $H \in M_k$ such that*

$$C = U + PH.$$

Proof. Consider the equation in (a). By Lemma 3.2, there is a unitary $U \in M_{2k}$ such that

$$U^*BU = \begin{pmatrix} 0_k & * \\ * & * \end{pmatrix} \quad \text{with} \quad B = \begin{pmatrix} I_k & R \\ S^* & -I_k \end{pmatrix}.$$

By the CS decomposition, there are unitary matrices $V = V_1 \oplus V_2$, $W = W_1 \oplus W_2 \in M_{2k}$ with $V_1, V_2, W_1, W_2 \in M_k$ such that

$$(V_1 \oplus V_2)U(W_1 \oplus W_2) = \begin{pmatrix} C & \sqrt{I_k - C^2} \\ \sqrt{I_k - C^2} & -C \end{pmatrix},$$

where $C = \text{diag}(c_1, \dots, c_k)$ with $1 \geq c_1 \geq \dots \geq c_k \geq 0$. Then $W(U^*BU)W^*$ also has 0_k as the leading $k \times k$ principal submatrix. Equivalently,

$$\begin{aligned} 0_k &= (C\sqrt{I_k - C^2})V^*BV \begin{pmatrix} C \\ \sqrt{I_k - C^2} \end{pmatrix} \\ &= C^2 + CV_1^*RV_2\sqrt{I_k - C^2} + \sqrt{I_k - C^2}V_2^*S^*V_1C - (I_k - C^2). \end{aligned}$$

Evidently, $c_k > 0$. Otherwise, the (k, k) entry of the above matrix is -1 . Thus, we can multiply the above equation by V_1C^{-1} on the left and $C^{-1}V_1^*$ on the right to get

$$0_k = I_k + RZ + Z^*S^* - Z^*Z \quad \text{with} \quad Z = V_2\sqrt{I_k - C^2}(C^{-1}V_1^*).$$

To prove (b), let $\tilde{R} = R + S$ and $\tilde{S} = R - S$. By (a), there is $Z \in M_k$ such that

$$I_k + \tilde{R}Z + Z^*\tilde{S}^* - Z^*Z = 0_k.$$

Taking the Hermitian part and skew-Hermitian part of the above equation, we get the two equations in (b).

To prove (c), let $\tilde{S} = S - R$. By (b) there is $Z \in M_k$ such that

$$I_k + RZ + Z^*R^* - Z^*Z = 0_k \quad \text{and} \quad \tilde{S}Z + Z^*\tilde{S}^* = 0_k.$$

Adding the two equations, we get $I_k + SZ + Z^*S^* - Z^*Z = 0_k$.

To prove (d), we may assume that $\gamma = 0$. Otherwise, replace R by $R - \gamma I_k/2$. Let $S = iP^{-1/2}$ and $\tilde{R} = RP^{-1/2}$. By (b) there is $Z \in M_k$ so that

$$I_k + \tilde{R}Z + Z^*\tilde{R}^* - Z^*Z = 0_k \quad \text{and} \quad SZ + Z^*S^* = 0_k.$$

The second equation implies that $Z = P^{1/2}H$ for some Hermitian H . Putting $Z = P^{1/2}H$ in the first equation, we have $I_k + RH + HR^* - HPH = 0_k$ as asserted.

To prove (e), note that the first equation in (b) can be written as $(Z^* - R)(Z - R^*) = I_k + RR^*$. Thus, its solution has the form $Z = R^* - U\sqrt{I_k + RR^*}$ for some unitary $U \in M_k$. Substituting this into the second equation in (b), we get the desired conclusion.

Finally, to prove (f), let $S = iTP^{-1}$ and $R = TC^*$ with $T = (I_k - C^*C)^{-1/2}$. Then $\sqrt{I_k + RR^*} = T$. By (e), there is a unitary $U \in M_k$ such that

$$(SR^* - SU\sqrt{I_k + RR^*}) + (RS^* - \sqrt{I_k + RR^*}U^*S^*) = 0_k.$$

Hence, $iTP^{-1}(C - U)T = SR^* - SU\sqrt{I_k + RR^*} = iK$ for some Hermitian K . Take $H = T^{-1}KT^{-1}$. The result follows. \square

4. INFINITE DIMENSIONAL OPERATORS AND RELATED RESULTS

One can easily extend the definition of $\Lambda_k(A)$ to a bounded linear operator A acting on infinite dimensional Hilbert spaces \mathcal{H} ; for example, see [12]. Results on $\Lambda_k(A)$ for infinite dimensional operators have been obtained in [10], including Theorem 4.1 below. For a selfadjoint operator H , we let

$$\lambda_k(H) = \sup\{\lambda_k(X^*HX) : X : \mathbb{C}^k \rightarrow \mathcal{H}, X^*X = I_k\}.$$

Theorem 4.1. *Let A be a bounded linear operator acting on an infinite dimensional Hilbert space \mathcal{H} . Then $\Lambda_k(A)$ is convex and its closure equals*

$$\bigcap_{\xi \in [0, 2\pi)} \{\mu \in \mathbb{C} : e^{i\xi}\mu + e^{-i\xi}\bar{\mu} \leq \lambda_k(e^{i\xi}A + e^{-i\xi}A^*)\}.$$

An open question in [2] concerns the lower bound of $\dim \mathcal{H}$ which ensures that $\Lambda_k(A)$ is nonempty for every bounded linear operator A acting on \mathcal{H} . The following result was proved in [9] and answers the above question.

Theorem 4.2. *Let \mathcal{H} be a Hilbert space, and let k be a positive integer. Then $\Lambda_k(A) \neq \emptyset$ for every bounded linear operator A acting on \mathcal{H} if and only if $\dim \mathcal{H} > 3k - 3$.*

ACKNOWLEDGMENTS

We would like to thank the authors of [1] and [12] for sending us their preprints. We also thank Professors T. Ando, V. Bolotnikov, M.-D. Choi, R. A. Horn, Y. T. Poon and L. Rodman for some helpful correspondence and discussions.

REFERENCES

1. M.-D. Choi, M. Giesinger, J. A. Holbrook, and D. W. Kribs, Geometry of higher-rank numerical ranges, *Linear and Multilinear Algebra* 56 (2008), 53-64. MR2378301
2. 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. MR2344684
3. M.-D. Choi, D. W. Kribs, and K. Życzkowski, Higher-rank numerical ranges and compression problems, *Linear Algebra Appl.* 418 (2006), 828-839. MR2260232 (2007k:15041)
4. M.-D. Choi, D. W. Kribs, and K. Życzkowski, Quantum error correcting codes from the compression formalism, *Rep. Math. Phys.* 58 (2006), 77-91. MR2273568 (2007h:81039)
5. K. Fan and G. Pall, Imbedding conditions for Hermitian and normal matrices, *Canad. J. Math.* 9 (1957), 298-304. MR0085216 (19:6e)
6. R. A. Horn and V. V. Sergeichuk, Canonical forms for complex matrix congruence and *-congruence, *Linear Algebra Appl.* 416 (2006), 1010-1032. MR2242477 (2007c:15018)
7. P. Lancaster and L. Rodman, *Algebraic Riccati equations*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1995. MR1367089 (97b:93003)
8. C.-K. Li, A simple proof of the elliptical range theorem, *Proc. Amer. Math. Soc.* 124 (1996), 1985-1986. MR1322932 (96i:15026)
9. 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*, to appear.
10. C.-K. Li, Y. T. Poon and N.-S. Sze, Higher rank numerical ranges and low rank perturbations of quantum channels, preprint. <http://arxiv.org/abs/0710.2898>

11. G. W. Stewart and J.-G. Sun, Matrix Perturbation Theory, Academic Press, New York, 1990. MR1061154 (92a:65017)
12. H. Woerdeman, The higher rank numerical range is convex, Linear and Multilinear Algebra 56 (2008), 65-67. MR2378302

DEPARTMENT OF MATHEMATICS, COLLEGE OF WILLIAM & MARY, WILLIAMSBURG, VIRGINIA
23185

E-mail address: `ckli@math.wm.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CONNECTICUT, STORRS, CONNECTICUT 06269

E-mail address: `sze@math.uconn.edu`