UPPER AND LOWER BOUNDS FOR NUMERICAL RADII OF BLOCK SHIFTS

HWA-LONG GAU AND PEI YUAN WU*

Dedicated to Professor Heydar Radjavi on his 80th birthday

Abstract. For any n-by-n matrix A of the form

$$\begin{bmatrix} 0 & A_1 & & \\ & 0 & \ddots & \\ & & \ddots & A_{k-1} \\ & & & 0 \end{bmatrix},$$

we consider two k-by-k matrices

$$A' = \begin{bmatrix} 0 & \|A_1\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_{k-1}\| \\ & & & 0 \end{bmatrix} \quad \text{and} \quad A'' = \begin{bmatrix} 0 & m(A_1) & & \\ & 0 & \ddots & \\ & & \ddots & m(A_{k-1}) \\ & & & 0 \end{bmatrix},$$

where $\|\cdot\|$ and $m(\cdot)$ denote the operator norm and minimum modulus of a matrix, respectively. It is shown that the numerical radii $w(\cdot)$ of A, A' and A'' are related by the inequalities $w(A'') \le w(A) \le w(A')$. We also determine exactly when either of the inequalities becomes an equality.
Keywords: Numerical radius, block shift, minimum modulus.
MSC(2010): Primary: 15A60; Secondary: 47A12.

1. Introduction

An n-by-n complex matrix A is call a *block shift* if it is of the form

$$\begin{bmatrix} 0 & A_1 & & \\ & 0 & \ddots & \\ & & \ddots & A_{k-1} \\ & & & 0 \end{bmatrix},$$

where the A_j's are in general rectangular matrices. In this paper, we obtain sharp upper and lower bounds for the numerical radius $w(A)$ of

*Corresponding author.

such an A. Recall that the *numerical radius* $w(B)$ of an n-by-n matrix B is the quantity

$$\max\{|\langle Bx, x\rangle| : x \in \mathbb{C}^n, \|x\| = 1\},$$

where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ denote the standard inner product and norm of vectors in $\mathbb{C}^n$, respectively. Note that $w(B)$ is the radius of the smallest circular disc centered at the origin which contains the *numerical range*

$$W(B) = \{\langle Bx, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$$

of B. For properties of the numerical range and numerical radius, the reader is referred to [3, Chapter 22] or [4, Chapter 1].

Note that if A is a block shift of the above form, then it is unitarily similar to $e^{i\theta}A$ for all real θ. Hence its numerical range is a closed circular disc centered at the origin with radius equal to its numerical radius. To estimate the latter, we consider two k-by-k scalar matrices

$$A' = \begin{bmatrix} 0 & \|A_1\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_{k-1}\| \\ & & & 0 \end{bmatrix} \quad \text{and} \quad A'' = \begin{bmatrix} 0 & m(A_1) & & \\ & 0 & \ddots & \\ & & \ddots & m(A_{k-1}) \\ & & & 0 \end{bmatrix},$$

where $\|A_j\|$ and $m(A_j)$, $1 \le j \le k-1$, are the operator norm and minimum modulus of A_j, respectively. Recall that the minimum modulus $m(B)$ of an m-by-n matrix B is, by definition, $\min\{\|Bx\| : x \in \mathbb{C}^n, \|x\| = 1\}$. In Sections 2 and 3 below, we show that $w(A'') \le w(A) \le w(A')$ always hold, and that, under the extra condition that the A_j's are all nonzero (resp., under $A_1 \dots A_{k-1} \ne 0$), $w(A) = w(A')$ (resp., $w(A) = w(A'')$) implies that A' (resp., A'') is a direct summand of A (cf. Theorems 2.1 and 3.1). Examples are given showing that the nonzero conditions on the A_j's are essential.

2. **Upper bound**

The main result of this section is the following theorem.

Theorem 2.1. *Let*

$$A = \begin{bmatrix} 0 & A_1 & & \\ & 0 & \ddots & \\ & & \ddots & A_{k-1} \\ & & & 0 \end{bmatrix} \quad \text{on } \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_k} \tag{2.1}$$

be an n-by-n block shift, where A_j is an n_j-by-n_{j+1} matrix for $1 \leq j \leq k-1$, and let

$$A' = \begin{bmatrix} 0 & \|A_1\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_{k-1}\| \\ & & & 0 \end{bmatrix} \quad \text{on} \quad \mathbb{C}^k.$$

Then (a) $w(A) \leq w(A')$, *and* (b) *under the assumption that $A_j \neq 0$ for all j, $w(A) = w(A')$ if and only if A is unitarily similar to $A' \oplus B$, where B is a block shift with $w(B) \leq w(A')$.*

Proof. (a) Let $x = [x_1 \ \ldots \ x_k]^T$ be a unit vector in $\mathbb{C}^n$ such that $|\langle Ax, x\rangle| = w(A)$. Hence

$$\begin{aligned} w(A) &= \left| \left\langle \begin{bmatrix} 0 & A_1 & & \\ & 0 & \ddots & \\ & & \ddots & A_{k-1} \\ & & & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \right\rangle \right| \\ &= \Big| \sum_{j=1}^{k-1} \langle A_j x_{j+1}, x_j \rangle \Big| \\ &\leq \sum_{j=1}^{k-1} |\langle A_j x_{j+1}, x_j \rangle| \\ &\leq \sum_{j=1}^{k-1} \|A_j\| \|x_{j+1}\| \|x_j\| \qquad (2.2) \\ &= \left\langle \begin{bmatrix} 0 & \|A_1\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_{k-1}\| \\ & & & 0 \end{bmatrix} \begin{bmatrix} \|x_1\| \\ \vdots \\ \|x_k\| \end{bmatrix}, \begin{bmatrix} \|x_1\| \\ \vdots \\ \|x_k\| \end{bmatrix} \right\rangle \\ &\leq w(A'), \qquad (2.3) \end{aligned}$$

where the last inequality follows from the fact that $[\|x_1\| \ \ldots \ \|x_k\|]^T$ is a unit vector in $\mathbb{C}^k$.

(b) Assume that $A_j \neq 0$ for all j, and that $w(A) = w(A')$. Then we have equalities throughout the chain of inequalities in (a). Since A' is an (entrywise) nonnegative matrix with irreducible real part, the equality in (2.3) yields, by [5, Proposition 3.3], that $x_j \neq 0$ for all j.

Let $\widehat{x}_j = [0 \ \ldots \ 0 \ \underset{j\text{th}}{x_j} \ 0 \ \ldots \ 0]^T$ for $1 \leq j \leq k$, and let K be the subspace of $\mathbb{C}^n$ spanned by the $\widehat{x}_j$'s. The equality in (2.2) implies that

$$|\langle A_j x_{j+1}, x_j\rangle| = \|A_j x_{j+1}\|\|x_j\| = \|A_j\|\|x_{j+1}\|\|x_j\|. \tag{2.4}$$

Hence $A_j x_{j+1} = a_j x_j$ for some scalar a_j. Therefore, $A\widehat{x}_1 = 0$ and

$$A\widehat{x}_j = [0 \ \ldots \ 0 \ \underset{(j-1)\text{st}}{A_{j-1}x_j} \ 0 \ \ldots \ 0]^T = [0 \ \ldots \ 0 \ \underset{(j-1)\text{st}}{a_{j-1}x_{j-1}} \ 0 \ \ldots \ 0]^T = a_{j-1}\widehat{x}_{j-1}$$

is in K for all j, $2 \leq j \leq k$. This shows that $AK \subseteq K$.

We next prove that $A^*K \subseteq K$. Indeed, we have $A^*\widehat{x}_j = [0 \ \ldots \ 0 \ \underset{(j+1)\text{st}}{A_j^* x_j} \ 0 \ \ldots \ 0]^T$ for $1 \leq j \leq k-1$. Since

$$|a_j|\|x_j\|^2 = \|a_j x_j\|\|x_j\| = \|A_j x_{j+1}\|\|x_j\| = \|A_j\|\|x_{j+1}\|\|x_j\|$$

by (2.4), the nonzeroness of the A_j's and x_j's yields the same for the a_j's. Letting $B_j = A_j/\|A_j\|$ and $y_j = (\|A_j\|/a_j)x_{j+1}$, we have $B_j y_j = (1/a_j)A_j x_{j+1} = x_j$ with $\|B_j\| = 1$ and

$$\|y_j\| = \frac{\|A_j\|}{|a_j|}\|x_{j+1}\| = \frac{\|A_j x_{j+1}\|}{|a_j|} = \|x_j\|$$

by (2.4). It follows from an extended lemma of Riesz and Sz.-Nagy that $B_j^* x_j = y_j$ (cf. [7, p. 215]). Therefore, we have $A_j^* x_j = (\|A_j\|^2/a_j)x_{j+1}$, which shows that $A_j^*\widehat{x}_j = (\|A_j\|^2/a_j)\widehat{x}_{j+1}$ is in K for $1 \leq j \leq k-1$. Moreover, we also have $A^*\widehat{x}_k = 0$. Thus $A^*K \subseteq K$ as asserted.

Since $\{\widehat{x}_j/\|x_j\|\}_{j=1}^k$ is an orthonormal basis of K, $A(\widehat{x}_1/\|x_1\|) = 0$, and

$$\begin{aligned} A(\frac{\widehat{x}_j}{\|x_j\|}) &= \frac{a_{j-1}\|x_{j-1}\|}{\|x_j\|}\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} = \frac{a_{j-1}}{|a_{j-1}|}\frac{\|a_{j-1}x_{j-1}\|}{\|x_j\|}\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} \\ &= \frac{a_{j-1}}{|a_{j-1}|}\frac{\|A_{j-1}x_j\|}{\|x_j\|}\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} = \frac{a_{j-1}}{|a_{j-1}|}\|A_{j-1}\|\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} \end{aligned}$$

for $2 \leq j \leq k$ by (2.4), we derive that the restriction $A|K$ is unitarily similar to A'. Thus A is unitarily similar to $A' \oplus (A|K^\perp)$. We now show that $A|K^\perp$ is also unitarily similar to a block shift. Indeed, let $\widehat{H}_j = 0 \oplus \cdots \oplus 0 \oplus \underset{j\text{th}}{\mathbb{C}^{n_j}} \oplus 0 \oplus \cdots \oplus 0$, $K_j = \mathbb{C}^{n_j} \ominus \bigvee\{x_j\}$, and $\widehat{K}_j = 0 \oplus \cdots \oplus 0 \oplus \underset{j\text{th}}{K_j} \oplus 0 \oplus \cdots \oplus 0$ for $1 \leq j \leq k$. Then $K^\perp = K_1 \oplus \cdots \oplus K_k$. Since $A\widehat{H}_{j+1} \subseteq \widehat{H}_j$ and $A^*\widehat{x}_j \in \bigvee\{\widehat{x}_{j+1}\}$ from before, we have $A\widehat{K}_{j+1} \subseteq \widehat{K}_j$ for $1 \leq j \leq k-1$. Moreover, $A\widehat{H}_k = \{0\}$ implies that $A\widehat{K}_k = \{0\}$. We conclude that $B \equiv A|K^\perp$ is unitarily similar to a block shift with

$w(B) \leq w(A) = w(A')$. This proves one direction of (b). The converse is trivial. □

Corollary 2.2. *Let A be an n-by-n block shift as in (2.1), and let $M = \max_j \|A_j\|$. Then*

(a) $w(A) \leq M \cdot \cos(\pi/(k+1))$, *and*

(b) $w(A) = M \cdot \cos(\pi/(k+1))$ *if and only if A is unitarily similar to $(M \cdot J_k) \oplus B$, where B is a block shift with $w(B) \leq M \cdot \cos(\pi/(k+1))$.*

Here J_k denotes the k-by-k *Jordan block*

$$\begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

whose numerical range is known to be $\{z \in \mathbb{C} : |z| \leq \cos(\pi/(k+1))\}$ (cf. [6]).

Proof of Corollary 2.2. (a) is an easy consequence of Theorem 2.1 (a) and [8, Lemma 5 (1)] while (b) follows from Theorem 2.1 (b) and [8, Lemma 5 (2)]. □

We remark that the assertion in Theorem 2.1 (b) still holds for $n \leq 5$ even without the nonzero assumption on the A_j's. This can be proven via a case-by-case verification by invoking, in most cases, the known result on the numerical ranges of square-zero matrices (cf. [9, Theorem 2.1]), which we omit. This is no longer the case for $n \geq 6$. Here we give a counterexample for $n = 6$.

Example 2.3. Let

$$A = \begin{bmatrix} 0 & \sqrt{2} & & & & \\ & 0 & 0 & & & \\ & & 0 & 1 & 0 & \\ & & & 0 & 0 & 0 \\ & & & 0 & 0 & 1 \\ & & & & & 0 \end{bmatrix}$$

with $A_1 = [\sqrt{2}]$, $A_2 = [0]$, $A_3 = [1\ 0]$ and $A_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Then

$$A' = \begin{bmatrix} 0 & \sqrt{2} & & & \\ & 0 & 0 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix},$$

and A and A' are unitarily similar to

$$\begin{bmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{bmatrix},$$

respectively. Hence $w(A) = w(A') = \sqrt{2}/2$, but A' is not a direct summand of A. To see the latter, note that $\ker A \cap \ker A^* = \{0\}$. Hence A cannot have the 1-by-1 zero matrix $[0]$ as a direct summand, and thus A cannot be unitarily similar to $A' \oplus [0]$, or A' is not a direct summand of A.

3. **Lower bound**

Here is the main result of this section.

Theorem 3.1. *Let A be an n-by-n block shift as in (2.1), and let*

$$A'' = \begin{bmatrix} 0 & m(A_1) & & \\ & 0 & \ddots & \\ & & \ddots & m(A_{k-1}) \\ & & & 0 \end{bmatrix} \quad \text{on} \quad \mathbb{C}^k.$$

Then (a) $w(A) \geq w(A'')$, *and* (b) *under the assumption that $A_1 \dots A_{k-1} \neq 0$, $w(A) = w(A'')$ if and only if A is unitarily similar to $A'' \oplus C$, where C is a block shift with $w(C) \leq w(A'')$.*

Our first lemma gives some basic properties of the minimum modulus of a rectangular matrix. For a square matrix (or, for that matter, an operator on a possibly infinite-dimensional Hilbert space), these appeared in [2, Theorem 1].

Lemma 3.2. *Let A be an m-by-n matrix. Then*

(a) *$m(A) > 0$ if and only if A is left invertible, and*

(b) *$m(A)$ equals the minimum singular value of A. In particular, if $m < n$, then $m(A) = 0$.*

Proof. (a) Note that $m(A) > 0$ means that there is a $c > 0$ such that $\|Ax\| \geq c\|x\|$ for all x in $\mathbb{C}^n$, which is equivalent to the well-definedness of the linear transformation $Ax \mapsto x$ from the range of A to $\mathbb{C}^n$, or to the left-invertibility of A.

(b) Consider the polar decomposition of A: $A = V(A^*A)^{1/2}$, where V is an m-by-n partial isometry with $\ker V = \ker A$ (cf. [3, Problem 134]). Then

$$\begin{aligned} m(A) &= \min\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\} \\ &= \min\{\|V(A^*A)^{1/2}x\| : x \in \mathbb{C}^n, \|x\| = 1\} \\ &= \min\{\|(A^*A)^{1/2}x\| : x \in \mathbb{C}^n, \|x\| = 1\} \\ &= \text{minimum eigenvalue of } (A^*A)^{1/2} \\ &= \text{minimum singular value of } A. \end{aligned}$$

□

To prove Theorem 3.1 (b), we need another lemma to get around the restriction $A_1 \dots A_{k-1} \neq 0$.

Lemma 3.3. *If A_j is an n_j-by-n_{j+1} matrix, $1 \leq j \leq k-1$, such that $A_1 \dots A_{k-1} = 0$, then for any $\varepsilon > 0$, there are n_j-by-n_{j+1} matrices B_j such that $\|B_j - A_j\| < \varepsilon$ for all j and $B_1 \dots B_{k-1} \neq 0$.*

Proof. This is proven by induction on k. The case of $k = 2$ is trivial. We now assume that $k = 3$ and $A_1A_2 = 0$. Consider the following four cases separately:

(i) $A_1 = 0$ and $A_2 = 0$. Let B_1 (resp., B_2) be the n_1-by-n_2 (resp., n_2-by-n_3) matrix with its $(1, 1)$-entry equal to $\varepsilon/2$ and all other entries 0. Then B_1B_2 has the $(1, 1)$-entry $\varepsilon^2/4$, and hence is nonzero.

(ii) $A_1 \neq 0$ and $A_2 = 0$. Assume that a_{ij}, the (i, j)-entry of A_1, is nonzero. Let $B_1 = A_1$ and let B_2 be the n_2-by-n_3 matrix with its $(j, 1)$-entry equal to $\varepsilon/2$ and all others 0. Then the $(i, 1)$-entry of B_1B_2 is $a_{ij}\varepsilon/2$, which is nonzero. Hence $B_1B_2 \neq 0$.

(iii) $A_1 = 0$ and $A_2 \neq 0$. By symmetry, this case can be dealt with as in (ii).

(iv) $A_1, A_2 \neq 0$. Assume that x_i^T, the ith row of A_1, and y_j, the jth column of A_2, are nonzero. Since $x_i^T y_j = 0$, we may perturb y_j slightly to a column vector z_j such that $x_i^T z_j \neq 0$. Let $B_1 = A_1$ and B_2 be obtained from A_2 by replacing its y_j by z_j. Then $B_1B_2 \neq 0$.

Note that in (ii) and (iv) above, we have actually shown that if $A_1A_2 = 0$ and $A_1 \neq 0$, then for any $\varepsilon > 0$ there is a matrix B_2 such that $\|B_2 - A_2\| < \varepsilon$ and $A_1B_2 \neq 0$. This will be used in the induction process below.

Assume that our assertion is true for $k-2$ and that $A_1 \ldots A_{k-1} = 0$. If $A_1 \ldots A_{k-2} = 0$, then the induction hypothesis implies, for each $\varepsilon > 0$, the existence of matrices $B_1, \ldots, B_{k-2}$ such that $\|B_j - A_j\| < \varepsilon$ for $1 \leq j \leq k-2$ and $B_1 \ldots B_{k-2} \neq 0$. If $(B_1 \ldots B_{k-2})A_{k-1} \neq 0$, then simply let $B_{k-1} = A_{k-1}$; otherwise, from (ii) and (iv) above, there is a matrix B_{k-1} such that $\|B_{k-1} - A_{k-1}\| < \varepsilon$ and $(B_1 \ldots B_{k-2})B_{k-1} \neq 0$. On the other hand, if $A_1 \ldots A_{k-2} \neq 0$, then, since $(A_1 \ldots A_{k-2})A_{k-1} = 0$, (ii) and (iv) above yields a matrix B_{k-1} such that $\|B_{k-1} - A_{k-1}\| < \varepsilon$ and $(A_1 \ldots A_{k-2})B_{k-1} \neq 0$. Letting $B_j = A_j$ for $1 \leq j \leq k-2$ proves our assertion. □

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. (a) First assume that $A_1 \ldots A_{k-1} \neq 0$. Since A'' is an (entrywise) nonnegative matrix, there is a unit vector $y = [y_1 \ \ldots \ y_k]^T$ in $\mathbb{C}^k$ with $y_j \geq 0$ for all j such that $\langle A''y, y\rangle = w(A'')$ (cf. [5, Proposition 3.3]). Let u be a unit vector in $\mathbb{C}^{n_k}$ such that $A_1 \ldots A_{k-1}u \neq 0$, let $x_j = A_jA_{j+1} \ldots A_{k-1}u/\|A_jA_{j+1} \ldots A_{k-1}u\|$ for $1 \leq j \leq k-1$ and $x_k = u$, and let $v = [y_1x_1 \ \ldots \ y_kx_k]^T$. Then v is a unit vector in $\mathbb{C}^n$ since

$$\|v\| = \left(|y_1|^2\|x_1\|^2 + \cdots + |y_k|^2\|x_k\|^2\right)^{1/2} = \left(|y_1|^2 + \cdots + |y_k|^2\right)^{1/2} = 1.$$

Moreover,

$$\langle Av, v\rangle = \sum_{j=1}^{k-1}\langle A_j(y_{j+1}x_{j+1}), y_jx_j\rangle = \sum_{j=1}^{k-1} y_{j+1}y_j\langle A_jx_{j+1}, x_j\rangle.$$

Note that

$$\langle A_jx_{j+1}, x_j\rangle = \left\langle \frac{A_jA_{j+1} \ldots A_{k-1}u}{\|A_{j+1} \ldots A_{k-1}u\|}, \frac{A_j \ldots A_{k-1}u}{\|A_j \ldots A_{k-1}u\|}\right\rangle$$
$$= \frac{\|A_j \ldots A_{k-1}u\|}{\|A_{j+1} \ldots A_{k-1}u\|} \geq m(A_j).$$

Hence

$$\langle Av, v\rangle \geq \sum_{j=1}^{k-1} y_{j+1}y_jm(A_j) = \langle A''y, y\rangle = w(A''). \tag{3.1}$$

It follows that $w(A) \geq w(A'')$ as asserted.

Now if $A_1 \dots A_{k-1} = 0$, then, for any $\varepsilon > 0$, let $B_1, \dots, B_{k-1}$ be as in Lemma 3.3, and let

$$B = \begin{bmatrix} 0 & B_1 & & \\ & 0 & \ddots & \\ & & \ddots & B_{k-1} \\ & & & 0 \end{bmatrix} \text{ on } \mathbb{C}^n \text{ and } B'' = \begin{bmatrix} 0 & m(B_1) & & \\ & 0 & \ddots & \\ & & \ddots & m(B_{k-1}) \\ & & & 0 \end{bmatrix} \text{ on } \mathbb{C}^k.$$

From the first half of the proof, we have $w(B) \geq w(B'')$. Since

$$\|A'' - B''\| = \max_j |m(A_j) - m(B_j)| \leq \max_j \|A_j - B_j\| < \varepsilon$$

(cf. [10, Lemma 2.2 (1)]), we infer from the continuity of $w(\cdot)$ that $w(A) \geq w(A'')$ (cf. [3, Problem 220]). This completes the proof of (a).

(b) Assume that $A_1 \dots A_{k-1} \neq 0$ and $w(A) = w(A'')$. Let $y = [y_1 \ \dots \ y_k]^T \in \mathbb{C}^k$, $u \in \mathbb{C}^{n_k}$, $x_j \in \mathbb{C}^{n_j}$ for $1 \leq j \leq k$, and $v \in \mathbb{C}^n$ be as in the first half of the proof of (a). Let $\widehat{x}_j = [0 \ \dots \ 0 \ \underset{j\text{th}}{x_j} \ 0 \ \dots \ 0]^T$ for $1 \leq j \leq k$, and let K be the subspace of $\mathbb{C}^n$ spanned by the $\widehat{x}_j$'s. Since $A\widehat{x}_1 = 0$ and
(3.2)

$$A\widehat{x}_j = \big[0 \ \dots \ 0 \ \underset{(j-1)\text{st}}{\frac{A_{j-1}A_j \dots A_{k-1}u}{\|A_j \dots A_{k-1}u\|}} \ 0 \ \dots \ 0]^T = \frac{\|A_{j-1} \dots A_{k-1}u\|}{\|A_j \dots A_{k-1}u\|}\widehat{x}_{j-1}, \ 2 \leq j \leq k,$$

we obtain $AK \subseteq K$.

We next show that $A^*K \subseteq K$. Indeed, since $w(A) = w(A'')$, we have an equality in (3.1), which yields that $\|A_j \dots A_{k-1}u\|/\|A_{j+1} \dots A_{k-1}u\| = m(A_j)$ for all j, $1 \leq j \leq k-1$. This is because $A_j \neq 0$ for all j and thus A'' is a nonnegative matrix with irreducible real part, from which we infer that $y_j > 0$ for all j (cf. [5, Proposition 3.3]). Since the x_j's are unit vectors satisfying $\|A_j x_{j+1}\| = m(A_j)$, we have $\langle (A_j^*A_j - m(A_j)^2 I_{n_{j+1}})x_{j+1}, x_{j+1}\rangle = 0$, $1 \leq j \leq k-1$. From $A_j^*A_j \geq m(A_j)^2 I_{n_{j+1}}$, we infer that $A_j^*A_j x_{j+1} = m(A_j)^2 x_{j+1}$ and hence $A_j^*A_jA_{j+1} \dots A_{k-1}u = m(A_j)^2 A_{j+1} \dots A_{k-1}u$. This shows that $A_j^* x_j$ is a multiple of x_{j+1} and thus $A^*\widehat{x}_j$ is a multiple of $\widehat{x}_{j+1}$ for $1 \leq j \leq k-1$. Therefore, $A^*\widehat{x}_j$ is in K for all j, $1 \leq j \leq k-1$. Together with $A^*\widehat{x}_k = 0$, these imply that $A^*K \subseteq K$. Hence A is unitarily similar to $(A|K) \oplus (A|K^\perp)$. Since $A\widehat{x}_1 = 0$ and $A\widehat{x}_j = m(A_{j-1})\widehat{x}_{j-1}$, $2 \leq j \leq k$, from (3.2), we have the unitary similarity of $A|K$ and A''. On the other hand, the unitary similarity of $A|K^\perp$ to a block shift follows as in the last part of the proof of Theorem 2.1 (b). □

Corollary 3.4. *Let A be an n-by-n block shift as in (2.1), and let $m = \min_j m(A_j)$. Then*

(a) $w(A) \geq m \cdot \cos(\pi/(k+1))$, *and*

(b) $w(A) = m \cdot \cos(\pi/(k+1))$ *if and only if A is unitarily similar to $(m \cdot J_k) \oplus C$, where C is a block shift with $w(C) \leq m \cdot \cos(\pi/(k+1))$.*

This can be proven as Corollary 2.2 by using Theorem 3.1 and [8, Lemma 5].

Analogous to the situation in Section 2, Theorem 3.1 (b) remains true for $n \leq 3$ without the assumption $A_1 \dots A_{k-1} \neq 0$. This is no longer the case for $n \geq 4$. A counterexample for $n = 4$ is given below.

Example 3.5. Let

$$A = \begin{bmatrix} 0 & 1 & 1 & \\ & 0 & 0 & 1 \\ & 0 & 0 & -1 \\ & & & 0 \end{bmatrix}$$

with $A_1 = [1\ 1]$ and $A_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. In this case, $A_1 A_2 = [0]$ and

$$A'' = \begin{bmatrix} 0 & 0 & \\ & 0 & \sqrt{2} \\ & & 0 \end{bmatrix}.$$

Since $A^2 = 0$, we have $w(A) = \|A\|/2 = \sqrt{2}/2$ (cf. [9, Theorem 2.1]). On the other hand, we also have $w(A'') = \sqrt{2}/2$. But A'' is not a direct summand of A. This is because if it is, then A will be unitarily similar to $A'' \oplus [0]$, which is impossible since $\ker A \cap \ker A^* = \{0\}$.

A larger parameter than the minimum modulus of an m-by-n matrix A is its *reduced minimum modulus* $\gamma(A)$ defined by

$$\gamma(A) = \begin{cases} \min\{\|Ax\| : x \in \mathbb{C}^n, x \perp \ker A, \|x\| = 1\} & \text{if } A \neq 0, \\ 0 & \text{if } A = 0. \end{cases}$$

A general reference for $\gamma(A)$ (when A is an operator on a possibly infinite-dimensional Hilbert space) is [1]. For an n-by-n block shift A of the form (2.1), consider the k-by-k matrix

$$A''' = \begin{bmatrix} 0 & \gamma(A_1) & & \\ & 0 & \ddots & \\ & & \ddots & \gamma(A_{k-1}) \\ & & & 0 \end{bmatrix}.$$

We may expect to have $w(A''')$ as a lower bound for $w(A)$ under some extra conditions on A. The next theorem shows that this is indeed the case for small values of k.

Theorem 3.6. (i) *Let* $A = \begin{bmatrix} 0 & A_1 \\ 0 & 0 \end{bmatrix}$ *and* $A''' = \begin{bmatrix} 0 & \gamma(A_1) \\ 0 & 0 \end{bmatrix}$. *Then*

(a) $w(A) \geq w(A''')$, *and*

(b) $w(A) = w(A''')$ *if and only if* A *is unitarily similar to* $A''' \oplus \cdots \oplus A''' \oplus 0$.

(ii) *Let*

$$A = \begin{bmatrix} 0 & A_1 & \\ & 0 & A_2 \\ & & 0 \end{bmatrix} \text{ on } \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2} \oplus \mathbb{C}^{n_3} \text{ and } A''' = \begin{bmatrix} 0 & \gamma(A_1) & \\ & 0 & \gamma(A_2) \\ & & 0 \end{bmatrix} \text{ on } \mathbb{C}^3.$$

Assume that $\operatorname{rank} A_1 + \operatorname{rank} A_2 > n_2$. *Then*

(a) $w(A) \geq w(A''')$, *and*

(b) $w(A) = w(A''')$ *if and only if* A *is unitarily similar to* $A''' \oplus C$, *where* C *is a block shift with* $w(C) \leq w(A''')$.

The proof is quite similar to the one for Theorem 3.1, which we omit. For larger values of k, the extra conditions on the A_j's are two cumbersome to be of any practical use.

Acknowledgments

The two authors acknowledge supports from the Ministry of Science and Technology of the Republic of China under projects MOST 103-2115-M-008-006 and NSC 102-2115-M-009-007, respectively. The second author was also supported by the MOE-ATU.

References

[1] C. Apostol, The reduced minimum modulus, *Michigan Math. J.*, **32** (1985), 279–294.

[2] R. Bouldin, The essential minimum modulus, *Indiana Univ. Math. J.*, **30** (1981), 513–517.

[3] P.R. Halmos, *A Hilbert Space Problem Book*, 2nd edition, Springer, New York, 1982.

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

[5] C.K. Li, B.-S. Tam and P.Y. Wu, The numerical range of a nonnegative matrix, *Linear Algebra Appl.*, **350** (2002), 1–23.

[6] M. Marcus and B.N. Shure, The numerical range of certain $0, 1$-matrices, *Linear Multilinear Algebra*, **7** (1979), 111–120.
[7] D. Petz, Application of a lemma of Riesz and Sz.-Nagy, *Acta Sci. Math.* (*Szeged*), **57** (1993), 215–222.
[8] M.C. Tsai and P.Y. Wu, Numerical ranges of weighted shift matrices, *Linear Algebra Appl.*, **435** (2011), 243–254.
[9] S.-H. Tso and P.Y. Wu, Matricial ranges of quadratic operators, *Rocky Mountain J. Math.*, **29** (1999), 1139–1152.
[10] P.Y. Wu, Approximation by partial isometries, *Proc. Edinburgh Math. Soc.*, **29** (1986), 255–261.

(Hwa-Long Gau) Department of Mathematics, National Central University, Chungli 32001, Taiwan
E-mail address: hlgau@math.ncu.edu.tw

(Pei Yuan Wu) Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30010, Taiwan
E-mail address: pywu@math.nctu.edu.tw