Linear Algebra and its Applications 440 (2014) 325–341

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Structures and numerical ranges of power partial isometries

Hwa-Long Gau [a,*,1], Pei Yuan Wu [b,2]

[a] Department of Mathematics, National Central University, Chung-Li 32001, Taiwan, ROC
[b] Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30010, Taiwan, ROC

ARTICLE INFO

Article history:
Received 10 October 2013
Accepted 12 November 2013
Available online 20 November 2013
Submitted by C.-K. Li

MSC:
15A99
15A60

Keywords:
Partial isometry
Power partial isometry
Numerical range
S_n-matrix
Jordan block

ABSTRACT

We derive a matrix model, under unitary similarity, of an n-by-n matrix A such that $A, A^2, \dots, A^k$ ($k \geqslant 1$) are all partial isometries, which generalizes the known fact that if A is a partial isometry, then it is unitarily similar to a matrix of the form $\begin{bmatrix} 0 & B \\ 0 & C \end{bmatrix}$ with $B^*B + C^*C = I$. Using this model, we show that if A has ascent k and $A, A^2, \dots, A^{k-1}$ are partial isometries, then the numerical range $W(A)$ of A is a circular disc centered at the origin if and only if A is unitarily similar to a direct sum of Jordan blocks whose largest size is k. As an application, this yields that, for any S_n-matrix A, $W(A)$ (resp., $W(A \otimes A)$) is a circular disc centered at the origin if and only if A is unitarily similar to the Jordan block J_n. Finally, examples are given to show that, for a general matrix A, the conditions that $W(A)$ and $W(A \otimes A)$ are circular discs at 0 are independent of each other.

1. Introduction

An n-by-n complex matrix A is a *partial isometry* if $\|Ax\| = \|x\|$ for any vector x in the orthogonal complement $(\ker A)^\perp$ in $\mathbb{C}^n$ of the kernel of A, where $\|\cdot\|$ denotes the standard norm in $\mathbb{C}^n$. The study of such matrices or, more generally, such operators on a Hilbert space dates back to 1962 [7]. Their general properties have since been summarized in [6, Chapter 15].

* Corresponding author.
E-mail addresses: hlgau@math.ncu.edu.tw (H.-L. Gau), pywu@math.nctu.edu.tw (P.Y. Wu).

1 Research supported by the National Science Council of the Republic of China under NSC-102-2115-M-008-007.
2 Research supported by the National Science Council of the Republic of China under NSC-102-2115-M-009-007 and by the MOE–ATU project.

0024-3795/$ – see front matter
http://dx.doi.org/10.1016/j.laa.2013.11.013

In this paper, we study matrices A such that, for some $k \geqslant 1$, the powers $A, A^2, \ldots, A^k$ are all partial isometries. In Section 2 below, we derive matrix models, under unitary similarity, of such a matrix (Theorems 2.2 and 2.4). They are generalizations of the known fact that A is a partial isometry if and only if it is unitarily similar to a matrix of the form $\left[\begin{smallmatrix} 0 & B \\ 0 & C \end{smallmatrix}\right]$ with $B^*B + C^*C = I$ (Lemma 2.1).

Recall that the *ascent* of a matrix, denoted by $a(A)$, is the smallest integer $k \geqslant 0$ for which $\ker A^k = \ker A^{k+1}$. It is easily seen that $a(A)$ is equal to the size of the largest Jordan block associated with the eigenvalue 0 in the Jordan form of A. We denote the n-by-n *Jordan block*

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

by J_n. The *numerical range* $W(A)$ of A is the subset $\{\langle Ax, x\rangle\colon\ x \in \mathbb{C}^n,\ \|x\| = 1\}$ of the complex plane $\mathbb{C}$, where $\langle\cdot,\cdot\rangle$ is the standard inner product in $\mathbb{C}^n$. It is known that $W(A)$ is a nonempty compact convex subset, and $W(J_n) = \{z \in \mathbb{C}\colon\ |z| \leqslant \cos(\pi/(n+1))\}$ (cf. [5, Proposition 1]). For other properties of the numerical range, the readers may consult [6, Chapter 22] or [10, Chapter 1].

Using the matrix model for power partial isometries, we show that if $a(A) = k \geqslant 2$ and $A, A^2, \ldots, A^{k-1}$ are all partial isometries, then the following are equivalent: (a) $W(A)$ is a circular disc centered at the origin, (b) A is unitarily similar to a direct sum $J_{k_1} \oplus J_{k_2} \oplus \cdots \oplus J_{k_\ell}$ with $k = k_1 \geqslant k_2 \geqslant \cdots \geqslant k_\ell \geqslant 1$, and (c) A has no unitary part and A^j is a partial isometry for all $j \geqslant 1$ (Theorem 2.6). An example is given, which shows that the number "$k-1$" in the above assumption is sharp (Example 2.7).

In Section 3, we consider the class of S_n-matrices. Recall that an n-by-n matrix A is of *class* S_n if A is a contraction ($\|A\| \equiv \max\{\|Ax\|\colon\ x \in \mathbb{C}^n,\ \|x\| = 1\} \leqslant 1$), its eigenvalues are all in $\mathbb{D}$ ($\equiv \{z \in \mathbb{C}\colon\ |z| < 1\}$), and it satisfies $\operatorname{rank}(I_n - A^*A) = 1$. Such matrices are the finite-dimensional versions of the *compression of the shift* $S(\phi)$, first studied by Sarason [11]. They also feature prominently in the Sz.-Nagy–Foiaş contraction theory [12]. It turns out that a hitherto unnoticed property of such matrices is that if A is of class S_n and k is its ascent, then $A, A^2, \ldots, A^k$ are all partial isometries. Thus the structure theorems in Section 2 are applicable to A or even to $A \otimes A$, the tensor product of A with itself. As a consequence, we obtain that, for an S_n-matrix A, the numerical range $W(A)$ (resp., $W(A \otimes A)$) is a circular disc centered at the origin if and only if A is unitarily similar to the Jordan block J_n (Theorem 3.3). The assertion concerning $W(A)$ is known before (cf. [13, Lemma 5]). Finally, we give examples to show that if A is a general matrix, then the conditions for the circularity (at the origin) of $W(A)$ and $W(A \otimes A)$ are independent of each other (Examples 3.5 and 3.6).

We use I_n and 0_n to denote the n-by-n identity and zero matrices, respectively. An identity or zero matrix with unspecified size is simply denoted by I or 0. For an n-by-n matrix A, $\operatorname{nullity} A$ is used for $\dim\ker A$, and $\operatorname{rank} A$ for its rank. The *real part* of A is $\operatorname{Re} A = (A + A^*)/2$. The *geometric* and *algebraic multiplicities* of an eigenvalue λ of A are $\operatorname{nullity}(A - \lambda I_n)$ and the multiplicity of the zero λ in the characteristic polynomial $\det(zI_n - A)$ of A, respectively. Note that $a(A)$ equals the smallest integer $k \geqslant 0$ for which the geometric and algebraic multiplicities of the eigenvalue 0 of A^k coincide. An n-by-n diagonal matrix with diagonal entries $a_1, \ldots, a_n$ is denoted by $\operatorname{diag}(a_1, \ldots, a_n)$.

2. Power partial isometries

We start with the following characterizations of partial isometries.

Lemma 2.1. *The following conditions are equivalent for an n-by-n matrix A:*

(a) *A is a partial isometry,*
(b) *A^*A is an (orthogonal) projection, and*
(c) *A is unitarily similar to a matrix of the form $\left[\begin{smallmatrix} 0 & B \\ 0 & C \end{smallmatrix}\right]$ with $B^*B + C^*C = I$.*

In this case, $\left[\begin{smallmatrix} 0 & B \\ 0 & C \end{smallmatrix}\right]$ acts on $\mathbb{C}^n = \ker A \oplus (\ker A)^{\perp}$.

Its easy proof is left to the readers.

The next theorem gives the matrix model, under unitary similarity, of a matrix A with $A, A^2, \dots, A^k$ $(1 \leqslant k \leqslant a(A))$ partial isometries.

Theorem 2.2. *Let A be an n-by-n matrix, $\ell \geqslant 1$, and $k = \min\{\ell, a(A)\}$. Then the following conditions are equivalent*:

(a) *$A, A^2, \dots, A^k$ are partial isometries,*

(b) *A is unitarily similar to a matrix of the form*

$$A' \equiv \begin{bmatrix} 0 & A_1 & & & \\ & 0 & \ddots & & \\ & & \ddots & A_{k-1} & \\ & & & 0 & B \\ & & & & C \end{bmatrix} \quad on\ \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_k} \oplus \mathbb{C}^m,$$

where the A_j's satisfy $A_j^* A_j = I_{n_{j+1}}$ for $1 \leqslant j \leqslant k-1$, and B and C satisfy $B^*B + C^*C = I_m$. In this case, $n_j = \operatorname{nullity} A$ if $j = 1$, $\operatorname{nullity} A^j - \operatorname{nullity} A^{j-1}$ if $2 \leqslant j \leqslant k$, and $m = \operatorname{rank} A^k$,

(c) *A is unitarily similar to a matrix of the form*

$$A'' \equiv \begin{bmatrix} 0 & I & & & \\ & 0 & \ddots & & \\ & & \ddots & I & \\ & & & 0 & B \\ & & & & C \end{bmatrix} \oplus (J_{k-1} \oplus \cdots \oplus J_{k-1}) \oplus \cdots \oplus (J_1 \oplus \cdots \oplus J_1)$$

$$on\ \mathbb{C}^n = \underbrace{\mathbb{C}^{n_k} \oplus \cdots \oplus \mathbb{C}^{n_k}}_{k} \oplus \mathbb{C}^m \oplus \underbrace{\mathbb{C}^{k-1} \oplus \cdots \oplus \mathbb{C}^{k-1}}_{n_{k-1}-n_k} \oplus \cdots \oplus \underbrace{\mathbb{C} \oplus \cdots \oplus \mathbb{C}}_{n_1-n_2},$$

where the n_j's, $1 \leqslant j \leqslant k$, and m are as in (b), *and B and C satisfy $B^*B + C^*C = I_m$.*

For the proof of Theorem 2.2, we need the following lemma.

Lemma 2.3. *Let $A = [A_{ij}]_{i,j=1}^n$ be a block matrix with $\|A\| \leqslant 1$, and let α be a nonempty subset of $\{1, 2, \dots, n\}$. If for some j_0, $1 \leqslant j_0 \leqslant n$, we have $\sum_{i \in \alpha} A_{ij_0}^* A_{ij_0} = I$, then $A_{ij_0} = 0$ for all i not in α.*

Proof. Since $\|A\| \leqslant 1$, we have $A^*A \leqslant I$. Thus the same is true for the (j_0, j_0)-block of A^*A, that is, $\sum_{i=1}^n A_{ij_0}^* A_{ij_0} \leqslant I$. Together with our assumption that $\sum_{i \in \alpha} A_{ij_0}^* A_{ij_0} = I$, this yields $\sum_{i \notin \alpha} A_{ij_0}^* A_{ij_0} \leqslant 0$. It follows immediately that $A_{ij_0} = 0$ for all i not in α. $\square$

Proof of Theorem 2.2. To prove (a) $\Rightarrow$ (b), let $H_1 = \ker A$, $H_j = \ker A^j \ominus \ker A^{j-1}$ for $2 \leqslant j \leqslant \ell$, and $H_{\ell+1} = \mathbb{C}^n \ominus \ker A^\ell$. Note that if $\ell > a(A)$, then at most $H_1, \dots, H_{k+1}$ are present. Hence A is unitarily similar to the block matrix $A' \equiv [A_{ij}]_{i,j=1}^{k+1}$ on $\mathbb{C}^n = H_1 \oplus \cdots \oplus H_{k+1}$. It is easily seen that $A_{ij} = 0$ for any $(i, j) \neq (k+1, k+1)$ with $1 \leqslant j \leqslant i \leqslant k+1$. For the brevity of notation, let $A_j = A_{j,j+1}$, $1 \leqslant j \leqslant k-1$, $B = A_{k,k+1}$, and $C = A_{k+1,k+1}$. We now check, by induction on j, that $A_j^* A_j = I_{n_{j+1}}$ for all j, and $A_{ij} = 0$ for $1 \leqslant i \leqslant j-2 \leqslant k-2$.

For $j = 1$, since A is a partial isometry, A^*A is an (orthogonal) projection by Lemma 2.1. We obviously have $A^*A = 0$ on $H_1 = \ker A$ and $A^*A = I$ on $H_1^\perp = H_2 \oplus \cdots \oplus H_{k+1}$. Thus $A'^*A' = 0 \oplus I \oplus \cdots \oplus I$ on $\mathbb{C}^n = H_1 \oplus H_2 \oplus \cdots \oplus H_{k+1}$. Since A'^*A' is of the form

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & A_1^*A_1 & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \cdots & * \end{bmatrix},$$

we conclude that $A_1^*A_1 = I$.

Next assume that, for some p $(2 \leqslant p < k)$, $A_j^*A_j = I$ for all j, $1 \leqslant j \leqslant p-1$, and all the blocks in A' which are above $A_1, \ldots, A_{p-1}$ are zero. We now check that $A_p^*A_p = I$ and all blocks above A_p are zero. Since A^p is a partial isometry, $A^{p*}A^p$ is an (orthogonal) projection with kernel equal to $H_1 \oplus \cdots \oplus H_p$. Thus $A'^{p*}A'^p = \underbrace{0 \oplus \cdots \oplus 0}_{p} \oplus \underbrace{I \oplus \cdots \oplus I}_{k-p+1}$. But from

$$A' = \begin{bmatrix} 0 & A_1 & 0 & \cdots & 0 & * & \cdots & * & * \\ & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ & & \ddots & \ddots & 0 & \vdots & & \vdots & \vdots \\ & & & \ddots & A_{p-1} & * & & \vdots & \vdots \\ & & & & 0 & A_p & \ddots & \vdots & \vdots \\ & & & & & 0 & \ddots & * & \vdots \\ & & & & & & \ddots & A_{k-1} & * \\ & & & & & & & 0 & B \\ & & & & & & & & C \end{bmatrix},$$

we have

$$A'^p = \begin{array}{c} \begin{array}{cc} \overbrace{}^{p} & \overbrace{\phantom{\prod_{j=1}^p A_j \quad * \quad \cdots \quad * \quad *}}^{k-p+1} \end{array} \\ \left[\begin{array}{ccccccc} 0 & \cdots & 0 & \prod_{j=1}^p A_j & * & \cdots & * & * \\ \cdot & & & 0 & \prod_{j=2}^{p+1} A_j & \ddots & \vdots & \vdots \\ \cdot & & & & \ddots & \ddots & * & \vdots \\ \cdot & & & & & \ddots & \prod_{j=k-p}^{k-1} A_j & * \\ \cdot & & & & & & 0 & * \\ \cdot & & & & & & \vdots & \vdots \\ \cdot & & & & & & 0 & BC^{p-1} \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & C^p \end{array}\right] \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} k-p \\ \left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} p+1 \end{array} \end{array}.$$

Thus the $(p+1, p+1)$-block of $A'^{p*}A'^p$ is $(\prod_{j=1}^p A_j)^*(\prod_{j=1}^p A_j) = A_p^*A_p$, which is equal to I from above. Lemma 2.3 then implies that all the blocks in A' which are above A_p are zero. Thus, by induction, the first k block columns of A' are of the asserted form.

Finally, we check that $B^*B + C^*C = I_m$. If this is the case, then all the blocks in A' above B and C are zero by Lemma 2.3 again and we will be done. As above, A'^{k-1} is of the form

$$\begin{bmatrix} 0 & \cdots & 0 & \prod_{j=1}^{k-1} A_j & D_1 \\ 0 & \cdots & 0 & 0 & D_2 \\ \vdots & & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & D_k \\ 0 & \cdots & 0 & 0 & C^{k-1} \end{bmatrix},$$

and the (orthogonal) projection ${A'}^{k-1*}{A'}^{k-1}$ equals $\underbrace{0\oplus\cdots\oplus 0}_{k-1}\oplus I\oplus I$ on $\mathbb{C}^n = H_1\oplus\cdots\oplus H_{k-1}\oplus H_k\oplus H_{k+1}$. Hence the $(k+1,k+1)$-block of ${A'}^{k-1*}{A'}^{k-1}$ is

$$\left(\sum_{j=1}^{k} D_j^* D_j\right) + C^{k-1*}C^{k-1}, \tag{1}$$

which is equal to I. Similarly,

$${A'}^k = {A'}^{k-1}A' = \begin{bmatrix} 0 & \cdots & 0 & (\prod_{j=1}^{k-1} A_j)B + D_1 C \\ 0 & \cdots & 0 & D_2 C \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & D_k C \\ 0 & \cdots & 0 & C^k \end{bmatrix}$$

and the $(k+1,k+1)$-block of ${A'}^{k*}{A'}^k$,

$$\begin{aligned} &B^*\left(\prod_{j=1}^{k-1} A_j\right)^*\left(\prod_{j=1}^{k-1} A_j\right)B + B^*\left(\prod_{j=1}^{k-1} A_j\right)^* D_1 C \\ &\quad + C^* D_1^*\left(\prod_{j=1}^{k-1} A_j\right)B + \left(\sum_{j=1}^{k} C^* D_j^* D_j C\right) + C^{k*}C^k, \end{aligned} \tag{2}$$

is also equal to I_m. We deduce from (1), (2) and $A_j^* A_j = I$ for $1\leqslant j\leqslant k-1$ that

$$B^*B + B^*\left(\prod_{j=1}^{k-1} A_j\right)^* D_1 C + C^* D_1^*\left(\prod_{j=1}^{k-1} A_j\right)B + C^*C = I_m. \tag{3}$$

To complete the proof, we need only show that $(\prod_{j=1}^{k-1} A_j)^* D_1 = 0$. Indeed, since $(\prod_{j=1}^{k-1} A_j)^*(\prod_{j=1}^{k-1} A_j) = I_{n_k}$, there is an n_1-by-n_1 unitary matrix U such that $U^*(\prod_{j=1}^{k-1} A_j) = \left[\begin{smallmatrix} I_{n_k} \\ 0 \end{smallmatrix}\right]$. Then $V \equiv U\oplus \underbrace{I\oplus\cdots\oplus I}_{k}$ is unitary and

$$V^*{A'}^{k-1}V = \begin{bmatrix} 0 & \cdots & 0 & U^*(\prod_{j=1}^{k-1} A_j) & U^*D_1 \\ 0 & \cdots & 0 & 0 & D_2 \\ \vdots & & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & D_k \\ 0 & \cdots & 0 & 0 & C^{k-1} \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & \left[\begin{smallmatrix} I_{n_k} \\ 0 \end{smallmatrix}\right] & \left[\begin{smallmatrix} 0 \\ D_1' \end{smallmatrix}\right] \\ 0 & \cdots & 0 & 0 & D_2 \\ \vdots & & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & D_k \\ 0 & \cdots & 0 & 0 & C^{k-1} \end{bmatrix}.$$

Hence

$$\left(\prod_{j=1}^{k-1} A_j\right)^* D_1 = [I_{n_k}\ 0]U^*U\begin{bmatrix} 0 \\ D_1' \end{bmatrix} = [I_{n_k}\ 0]\begin{bmatrix} 0 \\ D_1' \end{bmatrix} = 0$$

as asserted. We conclude from (3) that $B^*B + C^*C = I_m$. Moreover, the sizes of the blocks in A' are as asserted from our construction. This proves (a) $\Rightarrow$ (b).

Next we prove (b) $\Rightarrow$ (c). Let A' be as in (b), and let $n_1,\ldots,n_k,m$ be the sizes of the diagonal blocks of A'. Since $A_j^*A_j = I_{n_{j+1}}$ for all j, $1\leqslant j\leqslant k-1$, we have $n_1\geqslant n_2\geqslant\cdots\geqslant n_k$. Also, from $A_{k-1}^*A_{k-1} = I_{n_k}$, we deduce that there is a unitary matrix U_{k-1} of size n_{k-1} such that

$U_{k-1}^*A_{k-1} = \left[\begin{smallmatrix} I_{n_k} \\ 0 \end{smallmatrix}\right]$. Similarly, since $(A_{k-2}U_{k-1})^*(A_{k-2}U_{k-1}) = I_{n_{k-1}}$, there is a unitary U_{k-2} of size n_{k-2} such that $U_{k-2}^*(A_{k-2}U_{k-1}) = \left[\begin{smallmatrix} I_{n_{k-1}} \\ 0 \end{smallmatrix}\right]$. Proceeding inductively, we obtain a unitary U_j of size n_j satisfying $U_j^*(A_jU_{j+1}) = \left[\begin{smallmatrix} I_{n_{j+1}} \\ 0 \end{smallmatrix}\right]$ for each j, $1 \leqslant j \leqslant k-3$. If $U = U_1 \oplus \cdots \oplus U_{k-1} \oplus I_{n_k} \oplus I_m$, then

$$U^*A'U = \begin{bmatrix} 0 & U_1^*A_1U_2 & & & & \\ & 0 & \ddots & & & \\ & & \ddots & U_{k-2}^*A_{k-2}U_{k-1} & & \\ & & & 0 & U_{k-1}^*A_{k-1} & \\ & & & & 0 & B \\ & & & & & C \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \left[\begin{smallmatrix} I_{n_2} \\ 0 \end{smallmatrix}\right] & & & & \\ & 0 & \ddots & & & \\ & & \ddots & \left[\begin{smallmatrix} I_{n_{k-1}} \\ 0 \end{smallmatrix}\right] & & \\ & & & 0 & \left[\begin{smallmatrix} I_{n_k} \\ 0 \end{smallmatrix}\right] & \\ & & & & 0 & B \\ & & & & & C \end{bmatrix}.$$

Note that this last matrix is unitarily similar to the one asserted in (c).

To prove (c) $\Rightarrow$ (a), we may assume that

$$A'' = \begin{bmatrix} 0 & I & & & \\ & 0 & \ddots & & \\ & & \ddots & I & \\ & & & 0 & B \\ & & & & C \end{bmatrix}$$

with $B^*B + C^*C = I_m$. This is because powers of any Jordan block are all partial isometries and the direct sums of partial isometries are again partial isometries. Simple computations show that

$$A''^{\,j} = \begin{array}{c} \begin{array}{cc} \overbrace{\hphantom{0\ \ \cdots\ \ 0}}^{j} & \overbrace{\hphantom{I\ \ 0\ \ \cdots\ \ 0\ \ \ 0}}^{k-j+1} \end{array} \\ \left[\begin{array}{cccccccc} 0 & \cdots & 0 & I & 0 & \cdots & 0 & 0 \\ & \cdot & & 0 & \ddots & \ddots & \vdots & \vdots \\ & & \cdot & & \ddots & \ddots & 0 & \vdots \\ & & & \cdot & & \ddots & I & 0 \\ & & & & \cdot & & 0 & B \\ & & & & & \cdot & \vdots & \vdots \\ & & & & & & 0 & BC^{j-1} \\ & & & & & & & C^j \end{array}\right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array}}\right\} k-j \\ \left.\vphantom{\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array}}\right\} j+1 \end{array}$$

and $A''^{\,j*}A''^{\,j} = \underbrace{0 \oplus \cdots \oplus 0}_{j} \oplus \underbrace{I \oplus \cdots \oplus I}_{k-j} \oplus D$, where $D = (\sum_{s=0}^{j-1} C^{s*}B^*BC^s) + C^{j*}C^j$ for each j, $1 \leqslant j \leqslant k$. From $B^*B + C^*C = I_m$, we deduce that

$$D = B^*B + \left(\sum_{s=1}^{j-2} C^{s*}B^*BC^s\right) + C^{j-1*}\big(B^*B + C^*C\big)C^{j-1}$$

$$\begin{aligned}
&= B^*B + \left(\sum_{s=1}^{j-2} C^{s*}B^*BC^s\right) + C^{j-1*}C^{j-1}\\
&= B^*B + \left(\sum_{s=1}^{j-3} C^{s*}B^*BC^s\right) + C^{j-2*}\left(B^*B + C^*C\right)C^{j-2}\\
&= \cdots\\
&= B^*B + C^*C\\
&= I_m.
\end{aligned}$$

Hence $A''^{j*}A''^{j} = 0 \oplus I$, which implies that A''^{j} is a partial isometry by Lemma 2.1 for all j, $1 \leqslant j \leqslant k$. This proves (c) $\Rightarrow$ (a). □

A consequence of Theorem 2.2 is the following.

Theorem 2.4. *Let A be an n-by-n matrix and $\ell > a(A)$. Then the following conditions are equivalent:*

(a) *$A, A^2, \ldots, A^\ell$ are partial isometries,*
(b) *A is unitarily similar to a matrix of the form $U \oplus J_{k_1} \oplus \cdots \oplus J_{k_m}$, where U is unitary, which may be absent, $m \geqslant 0$, and $a(A) = k_1 \geqslant \cdots \geqslant k_m \geqslant 1$, and*
(c) *A^j is a partial isometry for all $j \geqslant 1$.*

The equivalence of (b) and (c) here is the finite-dimensional version of a result of Halmos and Wallen [8, Theorem].

Proof of Theorem 2.4. Since $\ell > k \equiv a(A)$, Theorem 2.2(a) $\Rightarrow$ (b) says that A is unitarily similar to a matrix of the form

$$A' \equiv \begin{bmatrix} 0_{n_1} & A_1 & & & \\ & 0_{n_2} & \ddots & & \\ & & \ddots & A_{k-1} & \\ & & & 0_{n_k} & B \\ & & & & C \end{bmatrix} \quad \text{on } \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_k} \oplus \mathbb{C}^m$$

with the A_j's, B and C satisfying the properties asserted therein. As k is the ascent of A, nullity A^k equals the algebraic multiplicity of eigenvalue 0 of A. Since nullity $A^k = \text{nullity } A'^k = \sum_{j=1}^k n_j$, it is seen from the structure of A' that the eigenvalue 0 appears fully in the diagonal 0_{n_j}'s. This shows that 0 cannot be an eigenvalue of C or C is invertible.

A simple computation yields that

$$A'^{k+1} = \begin{bmatrix} 0 & \cdots & 0 & (\prod_{j=1}^{k-1} A_j)BC \\ 0 & \cdots & 0 & (\prod_{j=2}^{k-1} A_j)BC^2 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & A_{k-1}BC^{k-1} \\ 0 & \cdots & 0 & BC^k \\ 0 & \cdots & 0 & C^{k+1} \end{bmatrix}$$

and

$$A'^{k+1*}A'^{k+1} = 0_{n_1} \oplus \cdots \oplus 0_{n_k} \oplus D, \tag{4}$$

where, after simplification by using $A_j^* A_j = I_{n_{j+1}}$ for $1 \leqslant j \leqslant k-1$, $D = (\sum_{j=1}^{k} C^{j^*} B^* B C^j) + C^{k+1^*} C^{k+1}$. As A'^{k+1} is a partial isometry, $A'^{k+1^*} A'^{k+1}$ is a projection by Lemma 2.1. Moreover, we also have

$$\text{nullity } A'^{k+1^*} A'^{k+1} = \text{nullity } A'^{k+1} = \text{nullity } A'^{k} = \sum_{j=1}^{k} n_j,$$

where the second equality holds because of $k = a(A')$. Thus we obtain from (4) that $D = I_m$. Therefore,

$$\begin{aligned} I_m = D &= \left(\sum_{j=1}^{k} C^{j^*} B^* B C^j \right) + C^{k+1^*} C^{k+1} \\ &= \left(\sum_{j=1}^{k-1} C^{j^*} B^* B C^j \right) + C^{k^*} \left(B^* B + C^* C \right) C^k \\ &= \left(\sum_{j=1}^{k-1} C^{j^*} B^* B C^j \right) + C^{k^*} C^k \\ &= \cdots \\ &= C^* \left(B^* B + C^* C \right) C \\ &= C^* C. \end{aligned}$$

This shows that C is unitary and hence $B = 0$ (from $B^* B + C^* C = I_m$). Thus A' is unitarily similar to the asserted form in (b). This completes the proof of (a) $\Rightarrow$ (b). The implications (b) $\Rightarrow$ (c) and (c) $\Rightarrow$ (a) are trivial. $\square$

At this juncture, it seems appropriate to define the *power partial isometry index* $p(\cdot)$ for any matrix A:

$$p(A) \equiv \sup\{k \geqslant 0\colon\ I, A, A^2, \ldots, A^k \text{ are all partial isometries}\}.$$

An easy corollary of Theorem 2.4 is the following estimate for $p(A)$.

Corollary 2.5. *If A is an n-by-n matrix, then $0 \leqslant p(A) \leqslant a(A)$ or $p(A) = \infty$. In particular, we have* (a) *$0 \leqslant p(A) \leqslant n-1$ or $p(A) = \infty$, and* (b) *$p(A) = n-1$ if and only if A is unitarily similar to a matrix of the form*

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & 1 & \\ & & & 0 & a \\ & & & & b \end{bmatrix} \tag{5}$$

with $|a|^2 + |b|^2 = 1$ and $a, b \neq 0$.

Proof. The first assertion follows from Theorem 2.4. If $p(A) = n$, then $a(A) = n$, which implies that the Jordan form of A is J_n. Thus $p(A) = \infty$, a contradiction. This proves (a) of the second assertion.

As for (b), if $p(A) = n-1$, then $a(A) = n$ will lead to a contradiction as above. Thus we must have $a(A) = n-1$. Theorem 2.2 implies that A is unitarily similar to a matrix of the form (5) with $|a|^2 + |b|^2 = 1$. Since either $a = 0$ or $b = 0$ will lead to the contradictory $p(A) = \infty$, we have thus proven one direction of (b). The converse follows easily from Theorem 2.2 and the arguments in the preceding paragraph. $\square$

The next theorem gives conditions for which $p(A) \geqslant a(A) - 1$ implies that A is unitarily similar to a direct sum of Jordan blocks.

Theorem 2.6. *Let A be an n-by-n matrix with $a(A) \geqslant 2$ and $p(A) \geqslant a(A) - 1$. Then the following conditions are equivalent*:

(a) *$W(A)$ is a circular disc centered at the origin,*
(b) *A is unitarily similar to a direct sum of Jordan blocks,*
(c) *A has no unitary part and A^j is a partial isometry for all $j \geqslant 1$, and*
(d) *A has no unitary part and $A, A^2, \ldots, A^\ell$ are partial isometries for some $\ell > a(A)$.*

In this case, $W(A) = \{z \in \mathbb{C}\colon |z| \leqslant \cos(\pi/(a(A)+1))\}$ and $p(A) = \infty$.

Here a matrix is said to have *no unitary part* if it is not unitarily similar to one with a unitary summand.

Note that, in the preceding theorem, the condition $p(A) \geqslant a(A) - 1$ cannot be replaced by the weaker $p(A) \geqslant a(A) - 2$. This is seen by the next example.

Example 2.7. If $A = J_3 \oplus \begin{bmatrix} 0 & (1-|\lambda|^2)^{1/2} \\ 0 & \lambda \end{bmatrix}$, where $0 < |\lambda| \leqslant \sqrt{2} - 1$, then $a(A) = 3$ and $W(A) = \{z \in \mathbb{C}\colon |z| \leqslant \sqrt{2}/2\}$. Since A is a partial isometry while A^2 is not, we have $p(A) = 1$. Note that A has a nonzero eigenvalue. Hence it is not unitarily similar to any direct sum of Jordan blocks.

The proof of Theorem 2.6 depends on the following series of lemmas, the first of which is a generalization of [14, Theorem 1].

Lemma 2.8. *Let*

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

where the A_j's satisfy $A_j^* A_j = I_{n_{j+1}}$, $1 \leqslant j \leqslant k-1$. If $W(A)$ is a circular disc centered at the origin with radius r larger than $\cos(\pi/(k+1))$, then C is not invertible.

Proof. Since $W(A) = \{z \in \mathbb{C}\colon |z| \leqslant r\}$, r is the maximum eigenvalue of $\mathrm{Re}(e^{i\theta}A)$ and hence $\det(rI_n - \mathrm{Re}(e^{i\theta}A)) = 0$ for all real θ. We have

$$\begin{aligned} 0 &= \det \begin{bmatrix} rI_{n_1} & -(e^{i\theta}/2)A_1 & & & \\ -(e^{-i\theta}/2)A_1^* & rI_{n_2} & \ddots & & \\ & \ddots & \ddots & -(e^{i\theta}/2)A_{k-1} & \\ & & -(e^{-i\theta}/2)A_{k-1}^* & rI_{n_k} & -(e^{i\theta}/2)B \\ & & & -(e^{-i\theta}/2)B^* & rI_m - \mathrm{Re}(e^{i\theta}C) \end{bmatrix} \\ &= \det D_k(\theta) \cdot \det\bigl(E(\theta) - F(\theta)\bigr), \end{aligned} \tag{6}$$

where

$$D_k(\theta)=\begin{bmatrix} rI_{n_1} & -(e^{i\theta}/2)A_1 & & & \\ -(e^{-i\theta}/2)A_1^* & rI_{n_2} & \ddots & & \\ & \ddots & \ddots & -(e^{i\theta}/2)A_{k-1} \\ & & -(e^{-i\theta}/2)A_{k-1}^* & rI_{n_k} \end{bmatrix},$$

$$E(\theta)=rI_m-\operatorname{Re}\bigl(e^{i\theta}C\bigr), \tag{7}$$

and

$$F(\theta)=\bigl[0\ \ldots\ 0\ -\bigl(e^{-i\theta}/2\bigr)B^*\bigr]D_k(\theta)^{-1}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ -(e^{i\theta}/2)B \end{bmatrix}, \tag{8}$$

by using the Schur complement of $D_k(\theta)$ in the matrix in (6) (cf. [9, p. 22]). Note that here the invertibility of $D_k(\theta)$ follows from the facts that $D_k(\theta)$ is unitarily similar to $rI-\operatorname{Re} J$, where $J=(\sum_{j=1}^{n_k}\oplus J_k)\oplus(\sum_{j=1}^{n_{k-1}-n_k}\oplus J_{k-1})\oplus\cdots\oplus(\sum_{j=1}^{n_1-n_2}\oplus J_1)$ (cf. the proof of Theorem 2.2(b) $\Rightarrow$ (c)), and r $(>\cos(\pi/(k+1)))$ is not an eigenvalue of $\operatorname{Re} J$. Moreover, the (k,k)-block of $D_k(\theta)^{-1}$ is independent of the value of θ. Thus the same is true for the entries of $F(\theta)$. Under a unitary similarity, we may assume that $C=[c_{ij}]_{i,j=1}^m$ is upper triangular with $c_{ij}=0$ for all $i>j$. Let $F(\theta)=[b_{ij}]_{i,j=1}^m$ and $E(\theta)-F(\theta)=[d_{ij}(\theta)]_{i,j=1}^m$. Then

$$d_{ij}(\theta)=\begin{cases} r-\operatorname{Re}(e^{i\theta}c_{jj})-b_{jj} & \text{if } i=j, \\ -(e^{i\theta}/2)c_{ij}-b_{ij} & \text{if } i<j, \\ -(e^{-i\theta}/2)\bar{c}_{ji}-b_{ij} & \text{if } i>j. \end{cases}$$

Hence $p(\theta)\equiv\det(E(\theta)-F(\theta))$ is a trigonometric polynomial of degree at most m, say, $p(\theta)=\sum_{j=-m}^m a_je^{ij\theta}$. Since $\det(rI_m-\operatorname{Re}(e^{i\theta}A))=0$ and $\det D_k(\theta)\neq 0$, we obtain from (6) that $p(\theta)=0$ for all real θ. This implies that $a_j=0$ for all j. In particular, $a_m=(-1)^m\prod_{j=1}^m(c_{jj}/2)=0$ from the above description of the $d_{ij}(\theta)$'s. This yields that $c_{jj}=0$ for some j or C is not invertible. □

The next lemma is to be used in the proof of Lemma 2.10.

Lemma 2.9. *Let $A=\begin{bmatrix} 0_p & B \\ 0 & C\end{bmatrix}$ be an n-by-n matrix, and let $B=[b_{ij}]_{i=1,j=1}^{p,n-p}$ and $C=[c_{ij}]_{i,j=1}^{n-p}$ with $c_{ij}=0$ for all $i>j$. If the geometric and algebraic multiplicities of the eigenvalue 0 of A are equal to each other and $c_{11}=0$, then $b_{i1}=0$ for all i, $1\leqslant i\leqslant p$.*

Proof. Let e_j denote the jth standard unit vector $[0\ \ldots\ 0\ \underset{j\text{th}}{1}\ 0\ \ldots\ 0]^T$, $1\leqslant j\leqslant n$. Then $e_1,\ldots,e_p$ are all in $\ker A$. Since $c_{11}=0$, we have $Ae_{p+1}=b_{11}e_1+\cdots+b_{p1}e_p$, which is also in $\ker A$. Thus $A^2e_{p+1}=0$ or $e_{p+1}\in\ker A^2$. Our assumption on the multiplicities of 0 implies that $\ker A=\ker A^2=\cdots$. Hence we obtain $e_{p+1}\in\ker A$ or $Ae_{p+1}=0$, which yields that $b_{i1}=0$ for all i, $1\leqslant i\leqslant p$. □

The following lemma is the main tool in proving, under the condition of circular $W(A)$, that $p(A)\geqslant a(A)-1$ yields $p(A)\geqslant a(A)$.

Lemma 2.10. *Let*

$$A = \begin{bmatrix} 0 & A_1 & & & \\ & 0 & \ddots & & \\ & & \ddots & A_{k-2} & \\ & & & 0 & B \\ & & & & C \end{bmatrix} \quad \text{on } \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_{k-1}} \oplus \mathbb{C}^m,$$

where $k = a(A)$ $(\geqslant 2)$, *the* A_j*'s satisfy* $A_j^* A_j = I_{n_{j+1}}$, $1 \leqslant j \leqslant k-2$, *and* $B = \left[\begin{smallmatrix} I_p & 0 \\ 0 & B_1 \end{smallmatrix}\right]$ *and* $C = \left[\begin{smallmatrix} 0_p & C_1 \\ 0 & C_2 \end{smallmatrix}\right]$ $(1 \leqslant p \leqslant \min\{n_{k-1}, m\})$ *satisfy* $B^*B + C^*C = I_m$. *If* $W(A)$ *is a circular disc centered at the origin with radius* r *larger than* $\cos(\pi/(k+1))$, *then* A *is unitarily similar to a matrix of the form*

$$\begin{bmatrix} 0 & A_1' & & & \\ & 0 & \ddots & & \\ & & \ddots & A_{k-1}' & \\ & & & 0 & B' \\ & & & & C' \end{bmatrix} \quad \text{on } \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_{k-1}} \oplus \mathbb{C}^q \oplus \mathbb{C}^{m-q},$$

where $q = \min\{n_{k-1}, m\}$, *the* A_j'*'s satisfy* ${A_j'}^* A_j' = I_{n_{j+1}}$, $1 \leqslant j \leqslant k-2$, ${A_{k-1}'}^* A_{k-1}' = I_q$, *and* B' *and* C' *satisfy* $B'^*B' + C'^*C' = I_{m-q}$.

Proof. Since $W(A) = \{z \in \mathbb{C}\colon |z| \leqslant r\}$, we have $\det(rI_n - \operatorname{Re}(e^{i\theta}A)) = 0$ for all real θ. As in the proof of Lemma 2.8, we have the factorization $\det(rI_n - \operatorname{Re}(e^{i\theta}A)) = \det D_{k-1}(\theta) \cdot \det(E(\theta) - F(\theta))$, where $D_{k-1}(\theta)$, $E(\theta)$ and $F(\theta)$ are as in (7) and (8) with $D_k(\theta)^{-1}$ in the expression of $F(\theta)$ there replaced by $D_{k-1}(\theta)^{-1}$. Since $D_{k-1}(\theta)$ is unitarily similar to $rI - \operatorname{Re} J$, where $J = (\sum_{j=1}^{n_{k-1}} \oplus J_{k-1}) \oplus (\sum_{j=1}^{n_{k-2}-n_{k-1}} \oplus J_{k-2}) \oplus \cdots \oplus (\sum_{j=1}^{n_1-n_2} \oplus J_1)$ and the $(k-1, k-1)$-entry of $(rI_{k-1} - \operatorname{Re} J_{k-1})^{-1}$ is $a \equiv \det(rI_{k-2} - \operatorname{Re} J_{k-2})/\det(rI_{k-1} - \operatorname{Re} J_{k-1})$, the $(k-1, k-1)$-block of $D_{k-1}(\theta)^{-1}$ is given by $aI_{n_{k-1}}$. Hence we have $F(\theta) = (a/4)B^*B$. As before, from $\det D_{k-1}(\theta) \neq 0$, we obtain $\det(E(\theta) - F(\theta)) = 0$. Thus

$$\begin{aligned} 0 &= \det\bigl(E(\theta) - F(\theta)\bigr) \\ &= \det\left(rI_m - \begin{bmatrix} 0_p & (e^{i\theta}/2)C_1 \\ (e^{-i\theta}/2)C_1^* & \operatorname{Re}(e^{i\theta}C_2) \end{bmatrix} - \frac{a}{4}\begin{bmatrix} I_p & 0 \\ 0 & B_1^*B_1 \end{bmatrix}\right) \\ &= \det\begin{bmatrix} (r - (a/4))I_p & -(e^{i\theta}/2)C_1 \\ -(e^{-i\theta}/2)C_1^* & rI_{m-p} - \operatorname{Re}(e^{i\theta}C_2) - (a/4)B_1^*B_1 \end{bmatrix}. \end{aligned} \tag{9}$$

We claim that $r \neq a/4$. Indeed, since $\det(rI_k - \operatorname{Re} J_k) = r\det(rI_{k-1} - \operatorname{Re} J_{k-1}) - (1/4)\det(rI_{k-2} - \operatorname{Re} J_{k-2})$, we have $\det(rI_k - \operatorname{Re} J_k)/\det(rI_{k-1} - \operatorname{Re} J_{k-1}) = r - (a/4)$. Therefore, $r = a/4$ if and only if $\det(rI_k - \operatorname{Re} J_k) = 0$. The latter would imply $r \leqslant \cos(\pi/(k+1))$ contradicting our assumption that $r > \cos(\pi/(k+1))$. Hence $r \neq a/4$ as asserted. Using the Schur complement, we infer from (9) that

$$p(\theta) \equiv \det\left(rI_{m-p} - \operatorname{Re}\bigl(e^{i\theta}C_2\bigr) - \frac{a}{4}B_1^*B_1 - \frac{1}{r - (a/4)} \cdot \frac{1}{4}C_1^*C_1\right) = 0$$

for all real θ. As $p(\theta)$ is a trigonometric polynomial of degree at most $m-p$, say, $p(\theta) = \sum_{j=-(m-p)}^{m-p} a_j e^{ij\theta}$, this implies that $a_j = 0$ for all j. After a unitary similarity, we may assume that $C_2 = [c_{ij}]_{i,j=1}^{m-p}$ with $c_{ij} = 0$ for all $i > j$. Hence $a_{m-p} = (1/2^{m-p})c_{11} \cdots c_{m-p,m-p} = 0$. Thus $c_{jj} = 0$ for some j. We may assume that $c_{11} = 0$. Note that

$$A^k = \begin{bmatrix} 0 & \cdots & 0 & (\prod_{j=1}^{k-2} A_j)BC \\ 0 & \cdots & 0 & (\prod_{j=2}^{k-2} A_j)BC^2 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & A_{k-2}BC^{k-2} \\ 0 & \cdots & 0 & BC^{k-1} \\ 0 & \cdots & 0 & C^k \end{bmatrix},$$

$$BC^j = \begin{bmatrix} I_p & 0 \\ 0 & B_1 \end{bmatrix} \begin{bmatrix} 0_p & C_1C_2^{j-1} \\ 0 & C_2^j \end{bmatrix} = \begin{bmatrix} 0_p & C_1C_2^{j-1} \\ 0 & B_1C_2^j \end{bmatrix}, \quad 1 \leqslant j \leqslant k-1,$$

and

$$C^k = \begin{bmatrix} 0_p & C_1C_2^{k-1} \\ 0 & C_2^k \end{bmatrix}.$$

Since the first column of C_2 is zero, the same is true for the $(p+1)$st columns of $(\prod_{j=t}^{k-2} A_j)BC^t$ $(2 \leqslant t \leqslant k-2)$, BC^{k-1} and C^k. As for $(\prod_{j=1}^{k-2} A_j)BC$, we need Lemma 2.9. Because $k = a(A)$, the geometric and algebraic multiplicities of the eigenvalue 0 of A^k coincide. Hence we may apply Lemma 2.9 to A^k to infer that the $((\sum_{j=1}^{k-1} n_j) + p + 1)$st column of A^k is zero. In particular, since $\ker(\prod_{j=1}^{k-2} A_j) = \{0\}$, the $(p+1)$st column of $BC = \left[\begin{smallmatrix} 0_p & C_1 \\ 0 & B_1C_2 \end{smallmatrix}\right]$ is zero and thus the first column of C_1 is zero. Together with the zero first column of C_2, this yields $C = \left[\begin{smallmatrix} 0_{p+1} & C_1^{(1)} \\ 0 & C_2^{(1)} \end{smallmatrix}\right]$. As

$$\begin{aligned} I_m = B^*B + C^*C &= \begin{bmatrix} I_p & 0 \\ 0 & B_1^* \end{bmatrix} \begin{bmatrix} I_p & 0 \\ 0 & B_1 \end{bmatrix} + \begin{bmatrix} 0_{p+1} & 0 \\ C_1^{(1)*} & C_2^{(1)*} \end{bmatrix} \begin{bmatrix} 0_{p+1} & C_1^{(1)} \\ 0 & C_2^{(1)} \end{bmatrix} \\ &= \begin{bmatrix} I_p & 0 \\ 0 & B_1^*B_1 \end{bmatrix} + \begin{bmatrix} 0_{p+1} & 0 \\ 0 & C_1^{(1)*}C_1^{(1)} + C_2^{(1)*}C_2^{(1)} \end{bmatrix}, \end{aligned}$$

we infer that the first column of B_1 is a unit vector. After another unitary similarity, we may further assume that

$$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & B_1^{(1)} \end{bmatrix} \quad \text{or} \quad B = \begin{bmatrix} I_{p+1} & 0 \\ 0 & B_1^{(1)} \end{bmatrix}.$$

Applying the above arguments again, we have

$$C = \begin{bmatrix} 0_{p+2} & C_1^{(2)} \\ 0 & C_2^{(2)} \end{bmatrix}, \qquad B_1^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & B_1^{(2)} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} I_{p+2} & 0 \\ 0 & B_1^{(2)} \end{bmatrix}.$$

Continuing this process yields that

$$\text{(i)} \quad C = \begin{bmatrix} 0_{n_{k-1}} & C_1' \\ 0 & C_2' \end{bmatrix} \quad \text{and} \quad B = [I_{n_{k-1}}\ 0] \quad \text{if } n_{k-1} < m,$$

and

$$\text{(ii)} \quad C = 0_m \quad \text{and} \quad B = \begin{bmatrix} I_m \\ 0 \end{bmatrix} \quad \text{if } n_{k-1} \geqslant m.$$

Finally, let $A_j' = A_j$ for $1 \leqslant j \leqslant k-2$. In case (i), let $A_{k-1}' = I_{n_{k-1}}$, $B' = C_1'$ and $C' = C_2'$. Since

$$I_m = B^*B + C^*C = \begin{bmatrix} I_{n_{k-1}} & 0 \\ 0 & C_1'^*C_1' + C_2'^*C_2' \end{bmatrix},$$

we have $B'^*B' + C'^*C' = C_1'^*C_1' + C_2'^*C_2' = I_{m-n_{k-1}}$. On the other hand, for case (ii), let $A'_{k-1} = \left[\begin{smallmatrix} I_m \\ 0 \end{smallmatrix}\right]$. In this case, B' and C' are absent. □

A consequence of the previous results is the following.

Proposition 2.11. *If A is an n-by-n matrix with $W(A)$ a circular disc centered at the origin and $p(A) \geqslant a(A) - 1$, then $p(A) = a(A)$ or ∞.*

Proof. Let $k = a(A)$. The assumption $p(A) \geqslant a(A) - 1$ says that $A, A^2, \ldots, A^{k-1}$ are all partial isometries. In particular, we have $A^{k-1} = 0$ or $\|A^{k-1}\| = 1$. In the former case, $p(A)$ equals ∞. Hence we may assume that $\|A^{k-1}\| = 1$ and thus also $\|A\| = 1$. By [1, Theorem 2.10], we have $w(A) \geqslant \cos(\pi/(k+1))$. Two cases are considered separately:

(i) $w(A) = \cos(\pi/(k+1))$. In this case, [1, Theorem 2.10] yields that A is unitarily similar to a matrix of the form $J_k \oplus A_1$ with $\|A_1\| \leqslant 1$ and $w(A_1) \leqslant \cos(\pi/(k+1))$. Since A_1^{k-1} is also a partial isometry, we may assume as before that $\|A_1^{k-1}\| = 1$ and thus also $\|A_1\| = 1$. Now applying [1, Theorem 2.10] again to A_1 yields that $w(A_1) = \cos(\pi/(k+1))$ and A_1 is unitarily similar to $J_k \oplus A_2$ with $\|A_2\| \leqslant 1$ and $w(A_2) \leqslant \cos(\pi/(k+1))$. Continuing this process, we obtain that either $p(A) = \infty$ or A is unitarily similar to a direct sum of copies of J_k. In the latter case, we again have $p(A) = \infty$.

(ii) $w(A) > \cos(\pi/(k+1))$. Since $A, A^2, \ldots, A^{k-1}$ are partial isometries, Theorem 2.2 yields the unitary similarity of A to a matrix of the form

$$\begin{bmatrix} 0 & A_1 & & & \\ & 0 & \ddots & & \\ & & \ddots & A_{k-2} & \\ & & & 0 & B \\ & & & & C \end{bmatrix} \quad \text{on } \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_{k-1}} \oplus \mathbb{C}^m$$

with $A_j^* A_j = I_{n_{j+1}}$, $1 \leqslant j \leqslant k-2$, and $B^*B + C^*C = I_m$. By Lemma 2.8, C is not invertible. We may assume, after a unitary similarity, that B and C are of the forms $\left[\begin{smallmatrix} 1 & 0 \\ 0 & B_1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & C_1 \\ 0 & C_2 \end{smallmatrix}\right]$, where B_1, C_1 and C_2 are $(n_{k-1}-1)$-by-$(m-1)$, 1-by-$(m-1)$ and $(m-1)$-by-$(m-1)$ matrices, respectively. Using Lemma 2.10, we obtain the unitary similarity of A to a matrix of the form in Theorem 2.2(b). Thus, by Theorem 2.2 again, $A, A^2, \ldots, A^k$ are partial isometries. Hence $p(A) \geqslant k = a(A)$. Our assertion then follows from Corollary 2.5. □

Note that, in the preceding proposition, the number "$a(A) - 1$" is sharp as was seen from Example 2.7.

We are now ready to prove Theorem 2.6.

Proof of Theorem 2.6. The implications (b) $\Rightarrow$ (c) and (c) $\Rightarrow$ (d) are trivial. On the other hand, (d) $\Rightarrow$ (a) follows from Theorem 2.4. Hence we need only prove (a) $\Rightarrow$ (b). Let $k = a(A)$. By Proposition 2.11, $A, A^2, \ldots, A^k$ are partial isometries. Thus A is unitarily similar to the matrix A' in Theorem 2.2(b). Since k is the ascent of A, the geometric multiplicity of A^k, that is, nullity A^k is equal to the algebraic multiplicity of eigenvalue 0 of A. As proven in (a) $\Rightarrow$ (b) of Theorem 2.2, nullity $A^k = \sum_{j=1}^k n_j$. We infer from the structure of A' that 0 cannot be an eigenvalue of C. On the other hand, applying Lemma 2.8 to A' yields the noninvertibility of C. This leads to a contradiction. Thus B and C won't appear in A' and, therefore, A', together with A, is unitarily similar to a direct sum of Jordan blocks by Theorem 2.2(c). This proves (b). □

3. S_n-matrices

In this section, we apply the results in Section 2 to the class of S_n-matrices. This we start with the following.

Proposition 3.1. *Let A be a noninvertible S_n-matrix. Then*

(a) *$a(A)$ equals the algebraic multiplicity of the eigenvalue 0 of A,*
(b) *$p(A) = a(A)$ or ∞,*
(c) *$p(A) = \infty$ if and only if A is unitarily similar to J_n, and*
(d) *$\operatorname{rank} A^j = n - j$ for $1 \leqslant j \leqslant a(A)$.*

Proof. Let $k = a(A)$.

(a) It is known that, for any eigenvalue λ of A, there is exactly one associated block, say, $\lambda I_\ell + J_\ell$ in the Jordan form of A. In particular, for $\lambda = 0$, both $a(A)$ and the algebraic multiplicity of 0 are equal to the size ℓ of its associated Jordan block J_ℓ.

(b) By [2, Corollary 1.3], A is unitarily similar to a matrix of the form $A' \equiv \left[\begin{smallmatrix} J_k & B \\ 0 & C \end{smallmatrix}\right]$, where $B = \left[\begin{smallmatrix} 0 \\ b \end{smallmatrix}\right]$ is a k-by-$(n-k)$ matrix with b a row vector of $n-k$ components, and C is an invertible $(n-k)$-by-$(n-k)$ upper-triangular matrix. Since $\operatorname{rank}(I_n - A^*A) = 1$, we infer from

$$I_n - A'^*A' = \begin{bmatrix} I_k & 0 \\ 0 & I_{n-k} \end{bmatrix} - \begin{bmatrix} J_k^* & 0 \\ B^* & C^* \end{bmatrix} \begin{bmatrix} J_k & B \\ 0 & C \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} & 0 \\ 0 & I_{n-k} - (B^*B + C^*C) \end{bmatrix}$$

that $B^*B + C^*C = I_{n-k}$. As A' can also be expressed as

$$\begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & \ddots & & \vdots \\ & & \ddots & 1 & 0 \\ & & & 0 & b \\ & & & & C \end{bmatrix} \quad \text{on } \mathbb{C}^n = \underbrace{\mathbb{C} \oplus \cdots \oplus \mathbb{C}}_{k} \oplus \mathbb{C}^{n-k}$$

with $b^*b + C^*C = I_{n-k}$, Theorem 2.2 can be invoked to conclude that $A, A^2, \ldots, A^k$ are partial isometries. Thus $p(A) \geqslant k$. It follows from Corollary 2.5 that $p(A) = k$ or ∞.

(c) If $p(A) = \infty$, then the unitary similarity of A and J_n is an easy consequence of Theorem 2.4 and the fact that A is irreducible (in the sense that it is not unitarily similar to the direct sum of two other matrices). The converse is trivial.

(d) As in the proof of (b), A is unitarily similar to $A' = \left[\begin{smallmatrix} J_k & B \\ 0 & C \end{smallmatrix}\right]$, where $B = \left[\begin{smallmatrix} 0 \\ b \end{smallmatrix}\right]$ and C is invertible. Then A^j is unitarily similar to

$$A'^j = \left[\begin{array}{cccccccc|c} \overbrace{0 & \cdots & 0}^{j} & \overbrace{1 & 0 & \cdots & 0}^{k-j} & \\ & \cdot & 0 & \ddots & \ddots & \vdots & & & \\ & & \cdot & & \ddots & \ddots & 0 & & 0 \\ & & & \cdot & & & \ddots & 1 & \\ & & & & \cdot & & & 0 & \\ & & & & & \cdot & & \vdots & B_j \\ & & & & & & & 0 & \\ \hline & & & & 0 & & & & C^j \end{array}\right] \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} k-j \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} j \\ \\ \end{array}.$$

for some j-by-$(n-k)$ matrix B_j. Since the first $k-j$ rows and the last $n-k$ rows of A'^j are linearly independent, we infer that $\operatorname{rank} A^j = \operatorname{rank} A'^j = (k-j) + (n-k) = n-j$ for $1 \leqslant j \leqslant k$. $\square$

The next corollary complements Corollary 2.5: it shows that any allowable value for $p(A)$ can actually be attained by some matrix A.

Corollary 3.2. *For any integers n and j satisfying $0 \leqslant j \leqslant n-1$, there is an n-by-n matrix A with $p(A) = j$.*

Proof. If $j = 0$ and $n = 1$, let $A = [1/2]$. Otherwise, let A be a noninvertible S_n-matrix with the algebraic multiplicity of its eigenvalue 0 equal to j (cf. [2, Corollary 1.3]). Then $p(A) = a(A) = j$ by Proposition 3.1. $\square$

We remark that the more refined question whether, for any nonnegative integers j, k and n satisfying $0 \leqslant j \leqslant \min\{k, n-1\}$, there is an n-by-n matrix A with $p(A) = j$ and $a(A) = k$ will be addressed in [4]. It turns out not to be the case.

For an n-by-n matrix $A = [a_{ij}]_{i,j=1}^n$ and an m-by-m matrix B, their *tensor product* (or *Kronecker product*) $A \otimes B$ is the (nm)-by-(nm) matrix

$$\begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix}.$$

Basic properties of tensor products can be found in [10, Chapter 4]. Our main concern here is when $W(A)$ and $W(A \otimes A)$ are circular discs (centered at the origin). Problems of this nature have also been considered in [1]. The main result of this section is the following theorem.

Theorem 3.3. *Let A be an S_n-matrix. Then the following conditions are equivalent*:

(a) *$W(A)$ is a circular disc centered at the origin,*
(b) *$W(A \otimes A)$ is a circular disc centered at the origin, and*
(c) *A is unitarily similar to J_n.*

In preparation for its proof, we need the next lemma.

Lemma 3.4. *Let A and B be n-by-n and m-by-m nonzero matrices, respectively.*

(a)

$$a(A \otimes B) = \begin{cases} \min\{a(A), a(B)\} & \text{if } a(A), a(B) \geqslant 1, \\ a(A) & \text{if } a(B) = 0, \\ a(B) & \text{if } a(A) = 0. \end{cases}$$

(b) *If A and B are partial isometries, then so is $A \otimes B$. The converse is false.*
(c) *Assume that A and B are (nonzero) contractions. Then A and B are partial isometries if and only if $A \otimes B$ is a partial isometry.*
(d) *If A and B are (nonzero) contractions, then $p(A \otimes B) = \min\{p(A), p(B)\}$.*
(e) *A is a partial isometry if and only if $A \otimes A$ is. Thus, in particular, $p(A \otimes A) = p(A)$.*

The proof makes use of the facts that (i) if A (resp., B) is similar to A' (resp., B'), then $A \otimes B$ is similar to $A' \otimes B'$, and (ii) if the eigenvalues of A (resp., B) are a_i, $1 \leqslant i \leqslant n$ (resp., b_j, $1 \leqslant j \leqslant m$), then the eigenvalues of $A \otimes B$ are $a_i b_j$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$, counting algebraic multiplicities (cf. [10, Theorem 4.2.12]).

Proof of Lemma 3.4. (a) Let $k_1 = a(A)$ and $k_2 = a(B)$, and assume that $2 \leqslant k_1 \leqslant k_2$. Let J_{k_1} (resp., J_{k_2}) be a Jordan block in the Jordan form of A (resp., B). Since

$$(J_{k_1} \otimes J_{k_2})^{k_1} = J_{k_1}^{k_1} \otimes J_{k_2}^{k_1} = 0_{k_1} \otimes J_{k_2}^{k_1} = 0_{k_1 k_2}$$

and

$$(J_{k_1} \otimes J_{k_2})^{k_1-1} = J_{k_1}^{k_1-1} \otimes J_{k_2}^{k_1-1} \neq 0_{k_1 k_2},$$

the size of the largest Jordan block in the Jordan form of $A \otimes B$ is k_1. This shows that $a(A \otimes B) = k_1 = \min\{a(A), a(B)\}$. The other cases can be proven even easier.

(b) This is a consequence of the equivalence of (a) and (b) in Lemma 2.1 as A^*A and B^*B are projections, which implies the same for $(A \otimes B)^*(A \otimes B)$. The converse is false as seen by the example of $A = [2]$ and $B = [1/2]$.

(c) If $A \otimes B$ is a partial isometry, then $(A \otimes B)^*(A \otimes B) = (A^*A) \otimes (B^*B)$ is a projection by Lemma 2.1. Since the positive semidefinite A^*A and B^*B are both contractions, their eigenvalues a_i, $1 \leqslant i \leqslant n$, and b_j, $1 \leqslant j \leqslant m$, are such that $0 \leqslant a_i, b_j \leqslant 1$ for all i and j. As the eigenvalues of $(A^*A) \otimes (B^*B)$, the products $a_i b_j$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$, can only be 0 and 1. Thus the same is true for the a_i's and b_j's. It follows that A^*A and B^*B are projections. Therefore, A and B are partial isometries.

(d) This follows from (c) immediately.

(e) If $A \otimes A$ is a partial isometry, then $(A \otimes A)^*(A \otimes A) = (A^*A) \otimes (A^*A)$ is a projection with eigenvalues 0 and 1. But its eigenvalues are also given by $a_i a_j$, $1 \leqslant i, j \leqslant n$, where the a_i's are eigenvalues of A^*A. If any a_i is nonzero and not equal to 1, then the same is true for a_i^2, which is a contradiction. Hence all the a_i's are either 0 or 1. It follows that A^*A is a projection and A is a partial isometry. The converse was proven in (c). $\square$

Finally, we are ready to prove Theorem 3.3.

Proof of Theorem 3.3. To prove (a) $\Rightarrow$ (c) (resp., (b) $\Rightarrow$ (c)), note that the center of the circular $W(A)$ (resp., $W(A \otimes A)$) must be an eigenvalue of A (resp., $A \otimes A$) (cf. [3, Theorem]). In particular, this says that A (resp., $A \otimes A$) is noninvertible. Since the eigenvalues of $A \otimes A$ are $a_i a_j$, $1 \leqslant i, j \leqslant n$, where the a_i's are the eigenvalues of A (cf. [10, Theorem 4.2.12]), the noninvertibility of $A \otimes A$ also implies that of A. Hence $p(A) = a(A)$ or ∞ by Proposition 3.1(b). If $p(A) = \infty$, then we have already had (c) by Proposition 3.1(c). Thus we may assume that $p(A) = a(A)$. In this case, we also have

$$p(A \otimes A) = p(A) = a(A) = a(A \otimes A)$$

by Lemma 3.4(d) (or (e)) and (a). Applying Theorem 2.6, we obtain the unitary similarity of A (resp., $A \otimes A$) to a direct sum of Jordan blocks. It follows that the only eigenvalue of A (resp., $A \otimes A$ and hence of A) is 0. Hence A is unitarily similar to J_n, that is, (c) holds.

The implication (c) $\Rightarrow$ (a) is trivial since, under (c), we have $W(A) = \{z \in \mathbb{C}: |z| \leqslant \cos(\pi/(n+1))\}$. For (c) $\Rightarrow$ (b), note that (c) implies that A is unitarily similar to $e^{i\theta}A$ for all real θ. Hence $A \otimes A$ is unitarily similar to $e^{i\theta}(A \otimes A)$ for real θ. Thus $W(A \otimes A)$ is a circular disc centered at the origin. This also follows from [1, Proposition 2.8]. $\square$

We remark that the equivalence of (a) and (c) in Theorem 3.3 was shown before in [13, Lemma 5] by a completely different proof.

We end this section with two examples. The examples show that, in contrast to the case of S_n-matrices, the conditions of $W(A)$ and $W(A \otimes A)$ being circular discs centered at the origin are independent of each other for a general matrix A.

Example 3.5. Let $A = [\lambda] \oplus J_2$, where $1/2 < |\lambda| \leqslant 1/\sqrt{2}$. Then

$$W(A \otimes A) = W\left(\left[\lambda^2\right] \oplus \lambda J_2 \oplus \lambda J_2 \oplus \begin{bmatrix} 0_2 & J_2 \\ 0 & 0_2 \end{bmatrix}\right) = \left\{z \in \mathbb{C}: |z| \leqslant \frac{1}{2}\right\},$$

but $W(A)$, being the convex hull of $\{\lambda\} \cup \{z \in \mathbb{C}: |z| \leqslant 1/2\}$, is obviously not a circular disc.

Example 3.6. Let

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

Then, for any real θ,

$$\operatorname{Re}\left(e^{i\theta}A\right)=\frac{1}{2}\begin{bmatrix} 0 & -\sqrt{2}e^{i\theta} & e^{i\theta} \\ -\sqrt{2}e^{-i\theta} & 0 & e^{i\theta} \\ e^{-i\theta} & e^{-i\theta} & \sqrt{2}\cos\theta \end{bmatrix},$$

whose maximum eigenvalue can be computed to be always equal to 1. Hence $W(A)=\overline{\mathbb{D}}$. On the other hand, a long and tedious computation shows that the characteristic polynomial $p(z)\equiv\det(zI_9-2\operatorname{Re}(A\otimes A))$ of $2\operatorname{Re}(A\otimes A)$ can be factored as

$$z^2\left(z^2-3\right)\left(z^5-z^4-17z^3+17z^2+46z-48\right). \tag{10}$$

Assume that $W(A\otimes A)=\{z\in\mathbb{C}\colon\ |z|\leqslant\sqrt{r}/2\}$ for some $r>0$. Then the maximum and minimum eigenvalues of $2\operatorname{Re}(A\otimes A)$ are $\sqrt{r}$ and $-\sqrt{r}$, respectively. Note that $p(2)=-8<0$ and $p(\infty)=\infty$ imply that p has a zero larger than 2. Hence $r\neq 3$. Similarly, we have $r\neq -3$. Since both $\sqrt{r}$ and $-\sqrt{r}$ are zeros of p, we also have

$$\begin{aligned} p(z)&=z^2\left(z^2-3\right)\left(z^2-r\right)\left(z^3+az^2+bz+c\right)\\ &=z^2\left(z^2-3\right)\left(z^5+az^4+(b-r)z^3+(c-ar)z^2-brz-cr\right) \end{aligned} \tag{11}$$

for some real a, b and c. Comparing the coefficients of the last factors in (10) and (11) yields that $a=-1$, $b-r=-17$, $c-ar=17$, $br=-46$ and $cr=48$. From these, we deduce that $c+r=17$ and hence $b=-c$. This leads to $-46=br=-cr$, which contradicts $cr=48$. Thus $W(A\otimes A)$ cannot be a circular disc at 0.

The matrix A in the preceding example was also considered in [1, Example 3.4] for another purpose.

References

[1] H.-L. Gau, K.-Z. Wang, P.Y. Wu, Numerical radii for tensor products of matrices, Linear Multilinear Algebra (2013), http://dx.doi.org/10.1080/03081087.2013.839669.
[2] H.-L. Gau, P.Y. Wu, Lucas' theorem refined, Linear Multilinear Algebra 45 (1999) 359–373.
[3] H.-L. Gau, P.Y. Wu, Condition for the numerical range to contain an elliptic disc, Linear Algebra Appl. 364 (2003) 213–222.
[4] H.-L. Gau, P.Y. Wu, Power partial isometry index and ascent of a finite matrix, arXiv:1311.2361.
[5] U. Haagerup, P. de la Harpe, The numerical radius of a nilpotent operator on a Hilbert space, Proc. Amer. Math. Soc. 115 (1992) 371–379.
[6] P.R. Halmos, A Hilbert Space Problem Book, 2nd ed., Springer, New York, 1982.
[7] P.R. Halmos, J.E. McLaughlin, Partial isometries, Pacific J. Math. 13 (1962) 585–596.
[8] P.R. Halmos, L.J. Wallen, Powers of partial isometries, J. Math. Mech. 19 (1970) 657–663.
[9] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[10] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[11] D. Sarason, Generalized interpolation in H^∞, Trans. Amer. Math. Soc. 127 (1967) 179–203.
[12] B. Sz.-Nagy, C. Foiaş, H. Bercovici, L. Kérchy, Harmonic Analysis of Operators on Hilbert Space, 2nd ed., Springer, New York, 2010.
[13] P.Y. Wu, A numerical range characterization of Jordan blocks, Linear Multilinear Algebra 43 (1998) 351–361.
[14] P.Y. Wu, Numerical ranges as circular discs, Appl. Math. Lett. 24 (2011) 2115–2117.