Linear Algebra and its Applications 498 (2016) 160–180

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Preservers of unitary similarity functions on Lie products of matrices

Jianlian Cui[a], Chi-Kwong Li[b], Yiu-Tung Poon[c,*]

[a] Department of Mathematics, Tsinghua University, Beijing 100084, PR China
[b] Department of Mathematics, College of William & Mary, Williamsburg, VA 23187, USA
[c] Department of Mathematics, Iowa State University, Ames, IA 50011, USA

ARTICLE INFO

Article history:
Received 6 November 2014
Accepted 27 February 2015
Available online 23 March 2015
Submitted by P. Semrl

In memory of Professor Hans Schneider

MSC:
15A60
46B04

Keywords:
Lie product
Unitary similarity invariant function
Pseudo spectrum

ABSTRACT

Denote by M_n the set of $n \times n$ complex matrices. Let $f : M_n \to [0, \infty)$ be a continuous map such that $f(\mu UAU^*) = f(A)$ for any complex unit μ, $A \in M_n$ and unitary $U \in M_n$, $f(X) = 0$ if and only if $X = 0$ and the induced map $t \mapsto f(tX)$ is monotonically increasing on $[0, \infty)$ for any rank one nilpotent $X \in M_n$. Characterization is given for surjective maps ϕ on M_n satisfying $f(AB - BA) = f(\phi(A)\phi(B) - \phi(B)\phi(A))$. The general theorem is then used to deduce results on special cases when the function is the pseudo spectrum and the pseudo spectral radius.

* Corresponding author.
E-mail addresses: jcui@math.tsinghua.edu.cn (J. Cui), ckli@math.wm.edu (C.-K. Li), ytpoon@iastate.edu (Y.-T. Poon).

http://dx.doi.org/10.1016/j.laa.2015.02.036

1. Introduction

Let M_n be the set of $n \times n$ matrices. A function $f : M_n \to \mathbb{R}$ is a *radial unitary similarity invariant function* if

(P1) $f(\mu UAU^*) = f(A)$ for a complex unit μ, $A \in M_n$ and unitary $U \in M_n$.

In [11], the authors studied unitary similarity invariant functions that are norms on M_n, and determined the structure of maps $\phi : M_n \to M_n$ satisfying

$$f(AB - BA) = f(\phi(A)\phi(B) - \phi(B)\phi(A)) \quad \text{for all } A, B \in M_n. \tag{1.1}$$

In [11, Remark 2.7], it was pointed out that the result actually holds for more general unitary similarity invariant functions. However, no detail was given, and it is not straightforward to apply the results to a specific problem. For instance, it is unclear how one can apply the result to study preservers of pseudo spectrum of Lie product of matrices[1]; see the definition in Section 3. To fill this gap, we extend the result in [11] to continuous radial unitary similarity invariant functions $f : M_n \to \mathbb{R}$ satisfying the following properties.

(P2) For any $X \in M_n$ we have $f(X) = f(0_n)$ if and only if $X = 0_n$, the $n \times n$ zero matrix.

(P3) For any rank one nilpotent $X \in M_n$, the map $t \mapsto f(tX)$ on $[0, \infty)$ is strictly increasing.

For a function $f : M_n \to [0, \infty)$ satisfying (P1)–(P3), we show that if $\phi : M_n \to M_n$ is a surjective map satisfying (1.1), then there is a unitary $U \in M_n$ and a subset $\mathcal{N}_n$ of normal matrices in M_n such that ϕ has the form

$$\phi(A) = \begin{cases} \mu_A U A^\dagger U^* + \nu_A I_n & A \in M_n \setminus \mathcal{N}_n \\ \mu_A U (A^\dagger)^* U^* + \nu_A I_n & A \in \mathcal{N}_n, \end{cases}$$

where

(a) $\mu_A, \nu_A \in \mathbb{C}$ with $|\mu_A| = 1$, depending on A,
(b) $A^\dagger = A$, $\overline{A}$, A^t or A^*, and
(c) $\mathcal{N}_n$ depends on the given unitarily invariant function f.

[1] This is a question raised by Professor Molnar to the second and third authors at the 2014 Summer Conference of the Canadian Mathematics Society.

The proof of this result will be given in Section 2. In Section 3, we apply the main result to the case when f is the pseudo spectral radius, and then obtain the result for the case when f is the pseudo spectrum.

For other preserver problems on different types of products on matrices and operators, one may see [1–3,7,11,12] and their references.

2. Main theorem

In this section, we prove Theorem 2.1 extending the result in [11]. We use similar ideas in [11] with some intricate arguments to make the extension possible.

Theorem 2.1. *Let $f : M_n \to [0, \infty)$ be a function on M_n satisfying* (P1)–(P3). *Suppose $n \geq 3$, and $\phi : M_n \to M_n$ is a surjective map satisfying*

$$f([\phi(A), \phi(B)]) = f([A, B]).$$

Then there is a unitary matrix U and a subset $\mathcal{N}_n$ of normal matrices with non-collinear eigenvalues such that ϕ has the form

$$\phi(A) = \begin{cases} \mu_A U\psi(A)U^* + \nu_A I_n & A \in M_n \setminus \mathcal{N}_n \\ \mu_A U\psi(A)^* U^* + \nu_A I_n & A \in \mathcal{N}_n, \end{cases}$$

where $\mu_A, \nu_A \in \mathbb{C}$ with $|\mu_A| = 1$ depending on A, and ψ is one of the maps: $A \mapsto A$, $A \mapsto \overline{A}$, $A \mapsto A^t$ or $A \mapsto A^*$.

A bijective map P on M_n is said to be a *locally regular polynomial map* [14] if for every $A \in M_n$, there exists a polynomial $p_A(t)$ such that $P(A) = p_A(A)$ and A have the same commutant. To prove the above theorem, we need the following result from Šemrl [14].

Theorem 2.2. *Suppose $n \geq 3$, and $\phi : M_n \to M_n$ is a bijective map satisfying*

$$[A, B] = 0_n \iff [\phi(A), \phi(B)] = 0_n.$$

Let Γ be the set of matrices A such that the Jordan form of A only has Jordan blocks of sizes 1 or 2. Then there is an invertible matrix S, an automorphism τ of the complex field and a regular locally polynomial map $A \to p_A(A)$ such that

$$\phi(A) = S(p_A(A_\tau^\dagger))S^{-1} \quad \text{for all } A \in \Gamma. \tag{2.1}$$

Here, X_τ is the matrix whose (i, j)-entry is $\tau(X_{ij})$, and $A^\dagger = A$ or A^t.

Our proof strategy is to show that $\phi(A)$ has the asserted form described in the theorem for a special class $\mathcal{C}_1$ of matrices A. Then we modify the map ϕ to ϕ_1 so that it will satisfy the same hypothesis of ϕ with the additional assumption that $\phi(X) = X$ for every $X \in \mathcal{C}_1$. Then we can set $B = \phi(A)$ for a certain matrix A not in $\mathcal{C}_1$ and use the condition that

$$f([A, X]) = f([\phi_1(A), \phi_1(X)]) = f([B, X]) \quad \text{for all } X \in \mathcal{C}_1$$

to show that $B = \phi_1(A)$ also has the asserted form. Thus, ϕ_1 has the asserted form for a larger class $\mathcal{C}_2$ of matrices. This process is repeated until we show that the modified map will fix every matrix after a finite number of steps.

In the next few lemmas, we will focus on the relations between a pair of matrices A and B such that

$$f([A, X]) = f([B, X]) \quad \text{for all } X \in \mathcal{C}$$

for a certain subset $\mathcal{C}$ of matrices.

Lemma 2.3. *Suppose $A \in M_n$ is a rank one nilpotent matrix. Then $A = xy^*$ for some nonzero orthogonal vectors, x and y. Furthermore, A is unitarily similar to $\|x\|\|y\|E_{21}$.*

Proof. Suppose $A \in M_n$ is a rank one matrix. Then $A = xy^*$ for some nonzero column vectors, x and $y \in \mathbb{C}^n$. If A is nilpotent, then $A^k = 0$ for some integer $k > 1$. Then we have

$$0 = \operatorname{tr} A^k = \operatorname{tr}(xy^*)^k = (y^*x)^k.$$

Therefore, x and y are orthogonal. Let U be a unitary matrix with $\dfrac{y}{\|y\|}$ and $\dfrac{x}{\|x\|}$ as the first and second columns respectively. Then $U^*AU = \|x\|\|y\|E_{21}$. $\square$

Denote by $\sigma(A)$ the spectrum of A and by $N(A)$ the null space of A.

Lemma 2.4. *For any two matrices A and B, if*

$$f([A, X]) = f([B, X]) \quad \textit{for all rank one } X \in M_n, \tag{2.2}$$

then there are $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$ such that one of the following holds with $\hat{A} = \mu A + \nu I_n$.

(a) *$\sigma(B) = \sigma(\hat{A})$ and for any $\lambda \in \sigma(\hat{A})$,*

$$N(B - \lambda I_n) = N(\hat{A} - \lambda I_n) \quad \textit{and} \quad N(B^t - \lambda I_n) = N(\hat{A}^t - \lambda I_n).$$

(b) *The eigenvalues of A are not collinear, $\sigma(B) = \overline{\sigma(\hat{A})}$ and for any $\lambda \in \sigma(\hat{A})$,*

$$N(B - \bar{\lambda} I_n) = N(\hat{A} - \lambda I_n) \quad \textit{and} \quad N(B^t - \bar{\lambda} I_n) = N(\hat{A}^t - \lambda I_n).$$

Proof. Note that for any rank one matrix $X = xy^t$, $[C, X] = 0$ if and only if x and y^t are the right and left eigenvectors of C corresponding to the same eigenvalue. To see this, as $[C, X] = (Cx)y^t - x(y^tC)$, then $[C, X] = 0$ if and only if $Cx = \lambda x$ and $y^tC = \lambda y^t$ for some $\lambda \in \mathbb{C}$.

Suppose A and B satisfy (2.2). By the above observation on rank one matrices and property (P2) of f, A and B must have the same set of left and right eigenvectors. Furthermore, x_1 and x_2 are the right eigenvectors of A corresponding to the same eigenvalue if and only if the two eigenvectors correspond to the same eigenvalue of B. Thus, the eigenvalues of A and B have the same geometric multiplicity.

Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of A with $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$ being the right and left eigenvectors. Also for each pair of eigenvectors x_i and y_i^t, let γ_i be the corresponding eigenvalue of B. Take $X_{ij} = x_i y_j^t$. Then $AX_{ij} = \lambda_i X_{ij}$ and $X_{ij}A = \lambda_j X_{ij}$. Using (P1), we see that for any $1 \le i, j \le n$,

$$f([A, X_{ij}]) = f(\lambda_i X_{ij} - \lambda_j X_{ij}) = f((\lambda_i - \lambda_j)X_{ij}) = f(|\lambda_i - \lambda_j| X_{ij}).$$

Similarly, $f([B, X_{ij}]) = f((\gamma_i - \gamma_j)X_{ij}) = f(|\gamma_i - \gamma_j| X_{ij})$.

By the fact that $f([A, X_{ij}]) = f([B, X_{ij}])$ and property (P3),

$$|\lambda_i - \lambda_j| = |\gamma_i - \gamma_j| \quad \text{for all } 1 \le i, j \le k.$$

As a result, there are $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$ such that either

(1) $\gamma_i = \mu\lambda_i + \nu$ for all $1 \le i \le k$; or
(2) the eigenvalues of A are non-collinear and $\bar{\gamma}_i = \mu\lambda_i + \nu$ for all $1 \le i \le k$.

Then the result follows with $\hat{A} = \mu A + \nu I_n$. $\square$

Lemma 2.5. *Suppose A and B commute and satisfy (2.2). If A has at least two distinct eigenvalues, then there are $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$ such that either*

(a) *$B = \mu A + \nu I_n$, or*
(b) *A is normal with non-collinear eigenvalues and $B = \mu A^* + \nu I_n$.*

Proof. As A and B commute, there is a unitary matrix U such that both U^*AU and U^*BU are upper triangular, see [9, Theorem 2.3.3]. Replacing (A, B) with (U^*AU, U^*BU), we may assume that A and B are upper triangular.

As A and B satisfy (2.2), Lemma 2.4 holds. Suppose Lemma 2.4(a) holds with $\hat{A} = \mu A + \nu I_n$. Notice that $\sigma(B) = \sigma(\hat{A})$ and

$$f([\hat{A}, X]) = f([\mu A + \nu I_n, X]) = f([B, X]) \quad \text{for all rank one } X \in M_n.$$

Suppose λ is an eigenvalue of $\hat{A}$ and $y \in N(\hat{A}^t - \lambda I_n)$. For any $z \in \mathbb{C}^n$, let $Z = zy^t$. Then $Z\hat{A} = \lambda Z$ and $[\hat{A}, Z] = (\hat{A} - \lambda I_n)Z$. Note that $(\hat{A} - \lambda I_n)Z$ has rank at most one and $\mathrm{tr}((\hat{A}-\lambda I_n)Z) = \mathrm{tr}([\hat{A}, Z]) = 0$, so $(\hat{A}-\lambda I_n)Z$ is unitarily similar to $\|(\hat{A}-\lambda I_n)z\|\|y^t\|E_{12}$. Thus,

$$f([\hat{A}, Z]) = f(\|(\hat{A} - \lambda I_n)z\| \, \|y^t\|E_{12}).$$

Similarly, $f([B, Z]) = f(\|(B - \lambda I_n)z\| \, \|y^t\|E_{12})$. Hence, by (P1) and (P3),

$$\|(\hat{A} - \lambda I_n)z\| = \|(B - \lambda I_n)z\| \quad \text{for all } z \in \mathbb{C}^n \text{ and } \lambda \in \sigma(\hat{A}).$$

As a result,

$$\begin{aligned} z^*\hat{A}^*\hat{A}z - 2\,\mathrm{Re}(\bar{\lambda}z^*\hat{A}z) + |\lambda|^2 z^*z &= \|(\hat{A} - \lambda I_n)z\|^2 \\ &= \|(B - \lambda I_n)z\|^2 = z^*B^*Bz - 2\,\mathrm{Re}(\bar{\lambda}z^*Bz) + |\lambda|^2 z^*z. \end{aligned}$$

This implies that

$$2\,\mathrm{Re}(\bar{\lambda}z^*(\hat{A} - B)z) = z^*(\hat{A}^*\hat{A} - B^*B)z \quad \text{for all } z \in \mathbb{C}^n \text{ and } \lambda \in \sigma(\hat{A}).$$

As A has at least two distinct eigenvalues, so does $\hat{A}$. Taking any $\lambda, \gamma \in \sigma(\hat{A})$ with $\lambda \neq \gamma$, we have

$$2\,\mathrm{Re}(\bar{\lambda}z^*(\hat{A} - B)z) = z^*(\hat{A}^*\hat{A} - B^*B)z = 2\,\mathrm{Re}(\bar{\gamma}z^*(\hat{A} - B)z).$$

Thus, $W((\overline{\lambda - \gamma})(\hat{A} - B)) \subseteq i\mathbb{R}$, where $W(X)$ is the numerical range of X.

Then $(\overline{\lambda - \gamma})(\hat{A} - B)$ is a skew-Hermitian matrix [8]. Since both $\hat{A}$ and B are upper triangular, they must be diagonal matrices. Now for any $1 \leq i \leq n$, $b_{ii} \in \sigma(B) = \sigma(\hat{A})$. Then

$$0 = \|(B - b_{ii}I_n)e_i\| = \|(\hat{A} - b_{ii}I_n)e_i\| = \|(B - b_{ii}I_n)e_i + (\hat{A} - B)e_i\|\|(\hat{A} - B)e_i\|.$$

Thus, $(\hat{A} - B)e_i = 0$ for all $1 \leq i \leq n$ and hence $B = \hat{A}$.

Now suppose Lemma 2.4(b) holds. Then by a similar argument, we can show that

$$\|(\hat{A} - \lambda I_n)z\| = \|(B - \bar{\lambda}I_n)z\| \quad \text{for all } \lambda \in \sigma(\hat{A}) \text{ and } z \in \mathbb{C}^n \tag{2.3}$$

and so $(\overline{\lambda - \gamma})\hat{A} - (\lambda - \gamma)B$ is a skew-Hermitian matrix. It follows that $(\overline{\lambda - \gamma})T_A - (\lambda - \gamma)T_B = 0$, or equivalently, $T_B = \frac{\overline{\lambda-\gamma}}{\lambda-\gamma}T_A$, where T_A and T_B are the strictly upper triangular parts of A and B. Now as the eigenvalues of A and hence $\hat{A}$ are not collinear, we can always find another $\omega \in \sigma(\hat{A})$ such that $\frac{\overline{\lambda-\omega}}{\lambda-\omega} \neq \frac{\overline{\lambda-\gamma}}{\lambda-\gamma}$. Then the above equation is possible only if $T_A = T_B = 0$. In this case, A and B are both diagonal and hence normal. Then (2.3) implies that $\hat{A} = \overline{B}$. $\square$

From Lemma 2.5, we have the following consequence for diagonalizable matrices.

Corollary 2.6. *Suppose A and B satisfy (2.2) and A is diagonalizable. Then there are $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$ such that*

(a) $B = \mu A + \nu I_n$, *or*
(b) *A is normal with non-collinear eigenvalues and $B = \mu A^* + \nu I_n$.*

Proof. Suppose A is diagonalizable. Then $A = SDS^{-1}$ for some invertible S and diagonal D. By Lemma 2.4, $B = S(\mu D + \nu I_n)S^{-1}$ or $B = S(\mu \overline{D} + \nu I_n)S^{-1}$. If A has only one eigenvalue, then A is a scalar matrix and so is B. Then the result follows. Suppose A has at least two eigenvalues. As A and B commute, the result now follows by Lemma 2.5. □

Lemma 2.7. *For any two matrices A and B, if*

$$f([A, X]) = f([B, X]) \quad \textit{for all } X \in M_n, \tag{2.4}$$

then there are $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$ such that either

(a) $B = \mu A + \nu I_n$, *or*
(b) *A is normal with non-collinear eigenvalues and $B = \mu A^* + \nu I_n$.*

Proof. Suppose A and B satisfy (2.4). Then, putting $X = B$ in (2.4), it follows from (P2) that A and B commute. If A has at least two eigenvalues, then the result follows from Lemma 2.5.

Suppose A has only one eigenvalue, say λ. Then by Lemma 2.4, B has one eigenvalue only, say γ. Write $A = SJS^{-1} + \lambda I_n$, where S is invertible and $J = J_{n_1} \oplus \cdots \oplus J_{n_s}$ is the Jordan form of A with $n_1 \geq \cdots \geq n_s$. Now as A and B satisfy (2.4), A and B have the same set of commuting matrices. Then $B = Sp(J)S^{-1} + \gamma I_n$ for some polynomial p of degree at most $m = n_1 - 1$ with $p(0) = 0$.

By a similar argument as in Lemma 2.5, we can show that

$$\|(B - \gamma I_n)z\| = \|(A - \lambda I_n)z\| \quad \text{for all } z \in \mathbb{C}^n.$$

Then there is a unitary matrix W such that

$$Sp(J)S^{-1} = (B - \gamma I_n) = W(A - \lambda I_n) = WSJS^{-1}.$$

Write $S = UT$ for unitary U and upper triangular T, $V = U^*WU$ and $p(x) = \sum_{i=1}^{m} c_i x^i$. Then we have

$$Tp(J)T^{-1} = VTJT^{-1}. \tag{2.5}$$

Notice that both $Tp(J)T^{-1}$ and TJT^{-1} are strictly upper triangular. Furthermore, the first $n_1 - 1$ entries in the super-diagonal of $Tp(J)T^{-1}$ are equal to c_1 times the corresponding $n_1 - 1$ super-diagonal entries of TJT^{-1}.

As V is unitary, we must have $|c_1| = 1$ and $V = c_1 I_{n_1-1} \oplus V_1$ for some unitary $V_1 \in M_{n-n_1+1}$. Now comparing the leading $n_1 \times n_1$ principal submatrices on both sides in (2.5), we have

$$T_1 p(J_{n_1}) T_1^{-1} = (c_1 I_{n_1-1} \oplus [v_{n_1,n_1}]) T_1 J_{n_1} T_1^{-1} = c_1 T_1 J_{n_1} T_1^{-1},$$

where T_1 is the $n_1 \times n_1$ principal submatrix of T. Therefore, $T_1 \left(\sum_{i=2}^{m} c_i J_{n_1}^i\right) T_1^{-1} = 0$ and so $\sum_{i=2}^{m} c_i J_{n_1}^i = 0$. Hence, $c_2 = \cdots = c_m = 0$. Then $p(x) = c_1 x$ and so $B = c_1 A + (\gamma - c_1\lambda) I_n$. $\square$

We are now ready to present the following.

Proof of Theorem 2.1. *First we assume that ϕ is bijective.* Suppose ϕ is a bijective map satisfying

$$f([A, B]) = f([\phi(A), \phi(B)]) \quad \text{for all } A, B \in M_n.$$

Because $f(X) = f(0)$ if and only if $X = 0$ by (P2), we see that $[A, B] = 0$ if and only if $[\phi(A), \phi(B)] = 0$. We can apply Theorem 2.2 and conclude that ϕ has the form (2.1) with $A^\dagger = A$ or A^t. In particular, for any rank one matrix $R \in M_n$, there are $\mu_R, \nu_R \in \mathbb{C}$ such that

$$\phi(R) = S(\mu_R R_\tau^\dagger + \nu_R I_n) S^{-1}.$$

Suppose $\mu_R = |\mu_R| e^{i\theta_R}$. By replacing $\phi(R)$ with $e^{-i\theta_R}(\phi(R) - \nu_R I_n)$, we may assume that $\mu_R > 0$ and $\nu_R = 0$.

Here we consider only the case when $A^\dagger = A$. The case when $A^\dagger = A^t$ is similar. Fix an orthonormal basis $\{x_1, \ldots, x_n\}$ and define $X_{ij} = x_i x_j^*$. Take $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$ and let $A = \sum_{j=1}^{n} \alpha_j X_{j1}$. For $k = 2, \ldots, n$,

$$f(\mu_A \mu_{X_{kk}} \tau(\alpha_k) S(X_{k1})_\tau S^{-1}) = f([\phi(A), \phi(X_{kk})]) = f([A, X_{kk}]) = f(\alpha_k X_{k1}). \quad (2.6)$$

In particular, if $Z = \mu_A \mu_{X_{22}} S(X_{21})_\tau S^{-1}$, then

$$f(\tau(\alpha) Z) = f(\alpha X_{21}) \quad \text{for all } \alpha \in \mathbb{C}.$$

Suppose τ is neither the identity map $\lambda \to \lambda$ nor the conjugate map $\lambda \to \bar{\lambda}$. By [10, Theorem 1], the set $\tau([0, 1])$ is an unbounded subset of $\mathbb{C}$. Thus, there exists $\alpha \in [0, 1]$ such that $|\tau(\alpha)| > |\tau(2)|$. But then by (P1) and (P3), we have

$$f(2X_{21}) = f(\tau(2)Z) = f(|\tau(2)|Z) < f(|\tau(\alpha)|Z) = f(\tau(\alpha)Z) = f(\alpha X_{21}) < f(2X_{21}),$$

which is a contradiction. Thus, τ is either the identity map or the conjugate map.

Furthermore, as $f([X_{32}, X_{22}]) = f(X_{32}) = f([X_{32}, X_{33}])$,

$$\begin{aligned} f(\mu_{X_{32}}\mu_{X_{22}} S(X_{32})_\tau S^{-1}) &= f([\phi(X_{32}), \phi(X_{22})]) \\ &= f([\phi(X_{32}), \phi(X_{33})]) = f(\mu_{X_{32}}\mu_{X_{33}} S(X_{32})_\tau S^{-1}). \end{aligned}$$

Thus, $\mu_{X_{22}} = \mu_{X_{33}}$ by (P3). By (2.6) and the fact that $f(\xi X_{21}) = f(\xi X_{31})$ for all $\xi \in \mathbb{C}$, we have

$$f(S(X_{21})_\tau S^{-1}) = f(S(X_{31})_\tau S^{-1}).$$

We now claim that S is a multiple of some unitary matrix. If not, then there is a pair of orthonormal vectors y_2, y_3 such that $\|Sy_2\| \neq \|Sy_3\|$. Extend y_2, y_3 to an orthonormal basis $\{y_1, y_2, y_3, \dots, y_n\}$ and let $x_j = (y_j)_{\tau^{-1}}$. Then $\{x_1, \dots, x_n\}$ also forms an orthonormal basis. By the above study, we have

$$f(\|Sy_2\|\|y_1^* S^{-1}\|\, E_{12}) = f(S(X_{21})_\tau S^{-1}) = f(S(X_{31})_\tau S^{-1}) = f(\|Sy_3\|\|y_1^* S^{-1}\|\, E_{12}),$$

which contradicts that $\|Sy_2\| \neq \|Sy_3\|$. Thus, S is a multiple of some unitary matrix. By absorbing the constant term, we may assume that S is unitary. Now for any rank one matrices R and S,

$$f([R, S]) = f([\phi(R), \phi(S)]) = f(\mu_R\mu_S[R_\tau, S_\tau]).$$

By (P1), $f([R, S]) = f([R_\tau, S_\tau])$ whenever $[R, S]$ is a rank one nilpotent matrix, and hence $\mu_R\mu_S = 1$ in this case.

Now for any rank one matrix A, we can always find two other rank one matrices B and C such that $[A, B]$, $[A, C]$ and $[B, C]$ are all rank one nilpotents. Then we must have $\mu_A\mu_B = \mu_A\mu_C = \mu_B\mu_C = 1$. As all μ_A, μ_B, μ_C are positive real numbers, the equality is possible only when $\mu_A = \mu_B = \mu_C = 1$. Then we have $\phi(A) = SA_\tau S^{-1} = SA_\tau S^*$ for all rank one A.

By replacing ϕ with the map $A \mapsto S^*\phi(A)S$, we may assume that $\phi(X) = X^+$ for all rank one matrices X, where $X^+ = X$, $\overline{X}$, X^t or X^*. Then

$$f([A, B]) = f([\phi(A), \phi(B)]) = f([A^+, B^+]) = f([A, B]^+)$$

for all rank one $A, B \in M_n$. Notice that the set

$$\{X : X = [A, B] \text{ for some rank one } A \text{ and } B\}$$

contains the set of trace zero non-nilpotent matrices with rank at most two and so is dense in the set of trace zero matrices with rank at most two. Thus, by continuity of f we see that

$$f(X) = f(X^+) \quad \text{for all trace zero matrices } X \text{ with rank at most two.}$$

Now define $\Phi : M_n \to M_n$ by $A \mapsto \phi(A)^+$. Then $\Phi(X) = X$ for all rank one matrices X. For any $A \in M_n$ and rank one matrix $X \in M_n$, as $[A, X]$ is a trace zero matrix with rank at most two,

$$f([A,X]) = f([\phi(A), \phi(X)]) = f([\phi(A), X^+]) = f([\phi(A)^+, X]) = f([\Phi(A), X]).$$

Thus, $f([A,X]) = f([\Phi(A), X])$ for all rank one X. Then Corollary 2.6 implies that $\Phi(A) = \mu_A A + \nu_A I_n$ or $\Phi(A) = \mu_A A^* + \nu_A I_n$ for all diagonalizable matrices A and the latter case happens only when A is normal with non-collinear eigenvalues.

After absorbing the constants μ_A and ν_A, we may assume that $\Phi(X) = X$ for all non-normal diagonalizable matrices X. Then

$$f([A,B]) = f([\phi(A), \phi(B)]) = f([\Phi(A), \Phi(B)]^+) = f([A,B]^+)$$

for all non-normal diagonalizable matrices A and B. Since the set of all non-normal diagonalizable matrices is dense in M_n, we see that $f([A,B]) = f([A,B]^+)$ for all $A, B \in M_n$. Then for any $A \in M_n$,

$$f([A,X]) = f([\phi(A), \phi(X)]) = f([\Phi(A), \Phi(X)]^+) = f([\Phi(A), X])$$

for all non-normal diagonalizable matrices X, and so $f([A,X]) = f([\Phi(A), X])$ for all $X \in M_n$ by the continuity of f. Now the result follows by Lemma 2.7.

Finally, we show that one only needs the surjective assumption on ϕ. For any $A, B \in M_n$, we say $A \sim B$ if

$$f([A,X]) = f([B,X]) \quad \text{for all } X \in M_n.$$

Clearly, $\sim$ is an equivalence relation and for each $A \in M_n$, denote by $S_A = \{B : B \sim A\}$ the equivalence class of A. By Lemma 2.7, either

(I) S_A is the set of matrices of the form $\mu A + \nu I$ for some $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$, or
(II) A is normal and $A \sim A^*$, S_A is the set of matrices of the form $\mu A + \nu I$ or $\mu A^* + \nu I$ for some $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$.

Pick a representative for each equivalence class and write $\mathcal{A}$ for the set of these representatives. Since ϕ is surjective, S_A and $\phi^{-1}(S_A)$ have the same cardinality c for every $A \in \mathcal{A}$. Thus there exists a map $\psi : M_n \to M_n$ which maps $\phi^{-1}(S_A)$ bijectively onto S_A for each $A \in \mathcal{A}$. Clearly ψ is bijective and $\psi(A) \sim \phi(A)$ for all $A \in M_n$. Then, for any $A, B \in M_n$,

$$f([A,B]) = f([\phi(A), \phi(B)]) = f([\psi(A), \phi(B)]) = f([\psi(A), \psi(B)]).$$

That is, ψ is bijective map satisfying (2.2). By the proof of Theorem 2.1 with bijective ϕ in the previous paragraphs, ψ has the desired form and hence so does ϕ, as $\psi(A) \sim \phi(A)$ implies $\phi(A) = \mu\psi(A) + \nu I$ or $\phi(A) = \mu\psi(A)^* + \nu I$ when $\psi(A)^*$ is normal and $\psi(A)^* \sim \psi(A)$. □

Remark. Using the argument in the last part of the proof on the replacement of the bijective assumption by the surjective assumption on ϕ, one may further weaken the surjective assumption on ϕ by any one of the following (weaker) assumptions on the following modified map $\tilde{\phi}$ defined by

$$\tilde{\phi}(X) = \phi(X) - \mathrm{tr}(\phi(X))I/n$$

on the set M_n^0 of trace zero matrices in M_n.

(a) The map $\tilde{\phi} : M_n^0 \to M_n^0$ is surjective.
(b) For any $A \in M_n^0$ the range of $\tilde{\phi}$ contains a matrix of the form $e^{it}A$ for some $t \in [0, 2\pi)$.

3. Pseudo spectrum and pseudo spectral radius

In this section, we use Theorem 2.1 to study maps preserving the pseudo spectral radius (see the definitions below) of the Lie product of matrices. Then we further deduce the result for maps preserving the pseudo spectrum. As one shall see, with considerable effort, one will be able to get more specific structure of the preserving maps.

For $\varepsilon > 0$, define the ε-pseudospectrum $\sigma_\varepsilon(A)$ of $A \in M_n$ as

$$\sigma_\varepsilon(A) = \{z \in \sigma(A + E) : E \in M_n,\ \|E\| < \varepsilon\} = \{z \in \mathbb{C} : s_n(A - zI_n) < \varepsilon\},$$

where $s_1(X) \geq \cdots \geq s_n(X)$ denote the singular values of $X \in M_n$, and the ε-pseudospectral radius $r_\varepsilon(A)$ of $A \in M_n$ as

$$r_\varepsilon(A) = \sup\{|\mu| : \mu \in \sigma_\varepsilon(A)\}.$$

Note that the pseudo spectral radius is useful in studying the stability of matrices under perturbations, and there are efficient algorithms for its computation; see, for example, [6] and its references. Preservers of pseudo spectrum have been considered for several types of products in [4] (see also [5]). Here we characterize the preservers of pseudo spectral radius and pseudo spectrum for Lie products. We first prove the following.

Theorem 3.1. *Suppose $n \geq 3$ and $\varepsilon > 0$. Then a surjective map $\phi : M_n \to M_n$ satisfying*

$$r_\varepsilon([A, B]) = r_\varepsilon([\phi(A), \phi(B)]) \qquad \textit{for all } A, B \in M_n$$

if and only if there is a unitary $U \in M_n$ such that

$$\phi(A) = \mu_A U \psi(A) U^* + \nu_A I_n \qquad \text{for all } A \in M_n,$$

where $\mu_A, \nu_A \in \mathbb{C}$ with $|\mu_A| = 1$, depending on A, and ψ is one of the following maps: $A \mapsto A$, $A \mapsto \overline{A}$, $A \mapsto A^t$ or $A \mapsto A^*$.

Proof. The sufficiency can be readily checked. To prove the necessity, let $f(A) = r_\varepsilon(A)$ for $A \in M_n$. It is clear that f is a continuous map satisfying (P1) and (P2). Suppose X is a rank one nilpotent matrix. It follows from Proposition 2.4 in [5] that $r_\varepsilon(X) = \sqrt{\varepsilon^2 + \|X\|\varepsilon}$. Hence, (P3) is also satisfied. So, we can apply Theorem 2.1 and conclude that ϕ has the form in Theorem 2.1. To get the desired conclusion, we need to show that the set $\mathcal{N}$ is empty. Assume not, and there is $A \in \mathcal{N}$. Since A is normal with non-collinear eigenvalues, there is a unitary V and $\gamma, \xi \in \mathbb{C}$ such that

$$V(\psi(A) - \xi I)V^* = \gamma \operatorname{diag}(1, \mu, 0, \mu_4, \ldots, \mu_n),$$

where $\mu \notin \mathbb{R}$. Let $B \in M_n$ be such that

$$\tilde{B} = V\psi(B)V^* = \begin{bmatrix} 0 & 1 & 0 \\ a & 0 & b \\ 0 & c & 0 \end{bmatrix} \oplus O_{n-3},$$

where $a = (1-\bar{\mu})/(1-\mu)$, $b > 0$ and $c = b\bar{\mu}/\mu$. Then

$$\tilde{B}\tilde{B}^* = \begin{bmatrix} 1 & 0 & \bar{c} \\ 0 & |a|^2 + |b|^2 & 0 \\ c & 0 & |c|^2 \end{bmatrix} \quad \text{and} \quad \tilde{B}^*\tilde{B} = \begin{bmatrix} |a|^2 & 0 & \bar{a}b \\ 0 & 1 + |c|^2 & 0 \\ \bar{b}a & 0 & |b|^2 \end{bmatrix}$$

and we can choose $b > 0$ so that $\tilde{B}$ is not normal, and neither is B. As a result, $\phi(B) = \mu_B U \psi(B) U^* + \nu_B I$.

Now,

$$C_1 = V[\psi(A), \psi(B)]V^* = \gamma \begin{bmatrix} 0 & 1-\mu & 0 \\ \bar{\mu} - 1 & 0 & b\mu \\ 0 & -b\bar{\mu} & 0 \end{bmatrix} \oplus O_{n-3}$$

is normal with eigenvalues $s_\pm = \pm\gamma\sqrt{|1-\mu|^2 + b^2|\mu|^2}$ so that

$$r_\varepsilon([A, B]) = r_\varepsilon([\psi(A), \psi(B)]) = |\gamma|\sqrt{|1-\mu|^2 + b^2|\mu|^2} + \varepsilon.$$

However, $[\phi(A), \phi(B)]$ is unitarily similar to

$$C_2 = \mu_A \mu_B \bar{\gamma} \begin{bmatrix} 0 & 1-\mu & 0 \\ (1-\bar{\mu})^2/(\mu - 1) & 0 & b\bar{\mu} \\ 0 & -b\bar{\mu}^2/\mu & 0 \end{bmatrix} \oplus O_{n-3}.$$

One readily checks that the matrix C_2 is normal if and only if μ is pure imaginary. In all other cases, there is a unitary $R \in M_n$ obtained from I_n by changing the $(1,1), (1,3), (3,1), (3,3)$ entries so that

$$RC_2R^* = \bar{\gamma} \begin{bmatrix} 0 & c_1 & 0 \\ c_2 & 0 & c_3 \\ 0 & 0 & 0 \end{bmatrix} \oplus O_{n-3}.$$

If C_2 has singular values $s_1 \geq s_2$, then

$$|\gamma|^2(|c_1|^2 + |c_2|^2 + |c_3|^2) = \operatorname{tr}(C_2C_2^*) = \operatorname{tr}(C_1C_1^*) = |\gamma|^2(s_+^2 + s_-^2).$$

Because C_2 is not normal, $s_1 < s_+$, we see that $s_2 > s_-$. Then for any $z \in \mathbb{C}$, if $\tilde{C} - zI$ has singular values $s_1(z) \geq s_2(z)$, then

$$s_1(z)^2 + s_2(z)^2 = 2|z|^2 + |c_1|^2 + |c_2|^2 + |c_3|^2 2|z|^2 + s_+^2 + s_-^2 = s_+(z)^2 + s_-(z)^2,$$

where $s_+(z) \geq s_-(z)$ are the singular values of $C_1 - zI$. Again, because $C_2 - zI$ is not normal, we see that $s_+(z) > s_1(z) \geq s_2(z) > s_-(z)$. It follows that $s_+(z) > s_-(z)$ for any $z \in \mathbb{C}$ with $|z| \leq |\gamma|\sqrt{|1+\mu|^2 + b^2|\mu|^2} + \varepsilon$. Thus,

$$\max\{z \in \mathbb{C} : s_2(C_2 - zI) \leq \varepsilon\} < \max\{z \in \mathbb{C} : s_2(C_1 - zI) \leq \varepsilon\}.$$

So, if a normal matrix A has three collinear eigenvalues $\gamma + \nu$, $\gamma\mu + \nu$, ν, where μ is not real and $\mu \neq \pm i$, then $A \notin \mathcal{N}$. Clearly, if $A \in \mathcal{N}$ has eigenvalues of the form $\gamma + \nu$, $\gamma + i\nu$, γ, then $\psi(A)^*$ can be viewed as a multiple of $\psi(A)$. Thus, we may assume that $A \notin \mathcal{N}$ by adjusting μ_A and ν_A. The result follows. $\square$

We will use the above theorem to determine the structure of preservers of the pseudo spectrum of Lie product of matrices. To achieve this, we need a characterization of normal matrices A with two distinct eigenvalues: there exists $b \in \mathbb{C}$ such that $A - bI$ is a nonzero multiple of a rank k orthogonal projection P with $1 \leq k < n$; see Proposition 3.3 below. The proof depends on the following lemma.

Lemma 3.2. *Suppose $C = C_1 \oplus O_{n-3}$, where $C_1 \in M_3$ has rank ≤ 2 and $\operatorname{tr} C_1 = 0$. Then for every $\varepsilon > 0$, $\sigma_\varepsilon(C) = \sigma_\varepsilon(C_1)$. Furthermore, suppose for $t \in \mathbb{R}$,*

$$f(\lambda, t) = \det(\lambda I_3 - (C_1 - tI_3)^*(C_1 - tI_3)) = \lambda^3 + p_2(t)\lambda^2 + p_1(t)\lambda + p_0(t)$$

where $p_1(t) = q_1(t) + at$ with $a \neq 0$ and $p_0(t)$, $q_1(t)$, $p_2(t)$ contains only even powers of t. Then $\sigma_\varepsilon(C) \neq -\sigma_\varepsilon(C)$.

Proof. Since rank $C_1 \leq 2$, $0 \in \sigma(C_1)$. Therefore, $\sigma_\varepsilon(C) = \sigma_\varepsilon(C_1) \cup \sigma_\varepsilon(0_{n-k}) = \sigma_\varepsilon(C_1)$.

Note that for each $t \in \mathbb{R}$, $f(\lambda, t)$ is a cubic polynomial in λ with three non-negative real roots $\lambda_1(t) \geq \lambda_2(t) \geq \lambda_3(t) \geq 0$ and $s_{\min}(C_1 - tI_3) = \sqrt{\lambda_3(t)}$.

Without loss of generality, we may assume that $a < 0$. Given $\varepsilon > 0$, $t \in \sigma_\varepsilon(C_1) \cap \mathbb{R}$ if and only if $\lambda_3(t) < \varepsilon^2$. Since $\lambda_3(0) = 0$ and $\lim_{t\to\infty} \lambda_3(t) = \infty$, there exists $t_0 > 0$ such that $\lambda_3(t_0) = \varepsilon^2$. We have $t_0 \notin \sigma_\varepsilon(C)$ and $f(\varepsilon^2, t_0) = 0$. But then

$$f(\varepsilon^2, -t_0) = f(\varepsilon^2, t_0) - 2at_0\varepsilon^2 > 0.$$

Thus, $\lambda_3(-t_0) < \varepsilon^2$ implying that $-t_0 \in \sigma_\varepsilon(C)$. So, $t_0 \in -\sigma_\varepsilon(C)$, and thus $\sigma_\varepsilon(C) \neq -\sigma_\varepsilon(C)$. $\square$

Proposition 3.3. *Let $n \geq 3$ and $A \in M_n$. The following conditions are equivalent.*

(a) *A is a normal matrix with at most two distinct eigenvalues.*
(b) *$\sigma_\varepsilon([A, B]) = -\sigma_\varepsilon([A, B])$ for all $B \in M_n$.*
(c) *$\sigma_\varepsilon([A, B]) = -\sigma_\varepsilon([A, B])$ for all rank one nilpotent $B \in M_n$.*

Proof. Suppose (a) holds. Then there is a unitary V and $\nu \in \mathbb{C}$ such that $VAV^* - \nu I = \lambda J$ with $J = I_k \oplus -I_{n-k}$. Then for any $B \in M_n$ such that $VBV^* = (B_{ij})_{1 \leq i,j \leq 2}$ with $B_{11} \in M_k$, $B_{22} \in B_{22}$, we have

$$C = V[A, B]V^* = 2\lambda \begin{bmatrix} O_k & B_{12} \\ -B_{21} & O_{n-k} \end{bmatrix}$$

satisfies $-C = JCJ^*$. Thus,

$$\sigma_\varepsilon([A, B]) = \sigma_\varepsilon([VAV^*, VBV^*]) = \sigma_\varepsilon(-J[A, B]J^*) = \sigma_\varepsilon(-[A, B]).$$

So, condition (b) holds.

The implication (b) $\Rightarrow$ (c) is clear. To prove (c) $\Rightarrow$ (a), we consider the contra-positive. Assume (a) is not true. We consider 2 cases.

Case 1. Suppose A is normal with more than two distinct eigenvalues. We may assume that $A = \operatorname{diag}(a, b, c) \oplus A_2$ such that a, b and c are distinct. If $\operatorname{Re}((b-a)\overline{(c-a)}) \leq 0$, then we have $\operatorname{Re}((b-c)\overline{a-c}) = \operatorname{Re}((b-a+a-c)\overline{a-c}) = |a-c|^2 - \operatorname{Re}((b-a)\overline{(c-a)}) > 0$. Thus, we may assume that $\operatorname{Re}((b-a)\overline{(c-a)}) > 0$ which implies that

$$\begin{aligned} &|2a - (b+c)|^2 = |(b-a) + (c-a)|^2 > |b-a|^2 + |c-a|^2 > |b-c|^2 \\ &\Rightarrow \quad \left| a - \frac{b+c}{2} \right| > \frac{|b-c|}{2}. \end{aligned}$$

Thus, by replacing A with $\frac{2}{(b-c)}\left(A-\frac{(b+c)}{2}I\right)$, we may assume that $A = \operatorname{diag}(a,1,-1) \oplus A_2$ such that $|a| > 1$. Consider the rank one nilpotent $X = \begin{bmatrix} 0 & -\sqrt{2} & \sqrt{2} \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \oplus 0_{n-3}$. We have $[A, X] = C \oplus 0_{n-3}$, where $C = \begin{bmatrix} 0 & \sqrt{2}(1-a) & \sqrt{2}(1+a) \\ 0 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$. Then

$$\det(\lambda I_3 - (C - tI_3)^*(C - tI_3)) = \lambda^3 + p_2(t)\lambda^2 + p_1(t)\lambda + p_0(t),$$

where

$$\begin{aligned} p_2(t) &= -3t^2 - 4|a|^2 - 12, \\ p_1(t) &= 3t^4 + 4\left(1+|a|^2\right)t^2 + 16\left(1-|a|^2\right)t + 16\left(2+|a|^2\right), \\ p_0(t) &= -t^6 + 8t^4 - 16t^2. \end{aligned}$$

Since $|a| > 1$, the condition in Lemma 3.2 is satisfied. Therefore, $\sigma_\varepsilon(C) \neq -\sigma_\varepsilon(C)$.

Case 2. Assume that A is not normal. We may assume that $A = (a_{ij})$ is in upper triangular form such that the $(1,2)$ entry is nonzero; see [13, Lemma 1]. We may replace A by $A - a_{33}I$ and assume that $A = (A_{ij})$ with $A_{22} \in M_{n-3}$, $A_{21} = O$, and

$$A_{11} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & 0 \end{bmatrix}.$$

Subcase (2.a). Suppose not both $[a_{13}, \dots, a_{1n}]$ and $[a_{23}, \dots, a_{2n}]$ are zero. Then there is a unitary $U = U_1 \oplus U_2$ with $U_1 \in M_2$ such that $UAU^* = \tilde{A} = (\tilde{a}_{ij})$, where the second row of $\tilde{A}$ equals $[\tilde{a}_{21}, \tilde{a}_{22}, \tilde{a}_{23}, 0, \dots, 0]$ with $\tilde{a}_{21} \in \mathbb{R}$ and $\tilde{a}_{21} \neq 0$ and $\tilde{a}_{23} \neq 0$. Let $B = E_{12}$. Then

$$C = [\tilde{A}, B] = \begin{bmatrix} -\tilde{a}_{21} & \tilde{a}_{11} - \tilde{a}_{22} & -\tilde{a}_{23} \\ 0 & \tilde{a}_{21} & 0 \\ 0 & 0 & 0 \end{bmatrix} \oplus O_{n-3}.$$

Then

$$\det(\lambda I_3 - (C - tI_3)^*(C - tI_3)) = \lambda^3 + p_2(t)\lambda^2 + p_1(t)\lambda + p_0(t),$$

where

$$\begin{aligned} p_2(t) &= -3t^2 - |\tilde{a}_{22} - \tilde{a}_{11}|^2 - |\tilde{a}_{23}|^2 - 2\tilde{a}_{21}^2, \\ p_1(t) &= 3t^4 + \left(|\tilde{a}_{22} - \tilde{a}_{11}|^2 + |\tilde{a}_{23}|^2\right)t^2 - 2\tilde{a}_{21}|\tilde{a}_{23}|^2 t + \tilde{a}_{21}^2\left(\tilde{a}_{21}^2 + |\tilde{a}_{23}|^2\right), \\ p_0(t) &= -t^6 + 2\tilde{a}_{21}^2 t^4 - \tilde{a}_{21}^4 t^2. \end{aligned}$$

Since a_{21} and $\tilde{a}_{23} \neq 0$, the condition in Lemma 3.2 is satisfied. Therefore, $\sigma_\varepsilon(C) \neq -\sigma_\varepsilon(C)$.

Subcase (2.b). Suppose both $[a_{13}, \ldots, a_{1n}]$ and $[a_{23}, \ldots, a_{2n}]$ are zero.

i) If $a_{11} = a_{22} = 0$, then we may assume that $a_{12} = 1$. Let

$$B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix} \oplus O_{n-3} \quad \text{so that} \quad C = [A, B] = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \oplus O_{n-3}.$$

Then

$$\det(\lambda I_3 - (C - tI_3)^*(C - tI_3)) = \lambda^3 + p_2(t)\lambda^2 + p_1(t)\lambda + p_0(t),$$

where

$$\begin{aligned} p_2(t) &= -3t^2 - 5, \\ p_1(t) &= 3t^4 + 3t^2 - 2t + 4, \\ p_0(t) &= -t^6 + 2t^4 - t^2. \end{aligned}$$

Therefore, the condition in Lemma 3.2 is satisfied and $\sigma_\varepsilon(C) \neq -\sigma_\varepsilon(C)$.

ii) If either a_{11} or $a_{22} \neq 0$, then, applying a unitary similarity, we may assume that $a_{11} \neq 0$. Replacing A by $e^{i\theta}A$, we may assume that $a_{11} \in \mathbb{R}$. Then we may further assume that $a_{12} = 1$. Let $B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix} \oplus O_{n-3}$ and $C = [A, B]$. Then $C = C_1 \oplus 0_{n-3}$, where $C_1 = \begin{bmatrix} 0 & -1 & a_{11} \\ 0 & 0 & 0 \\ a_{11} & 1 & 0 \end{bmatrix}$. Then

$$\det(\lambda I_3 - (C_1 - tI_3)^*(C_1 - tI_3)) = \lambda^3 + p_2(t)\lambda^2 + p_1(t)\lambda + p_0(t),$$

where

$$\begin{aligned} p_2(t) &= -3t^2 - 2 - 2a_{11}^2, \\ p_1(t) &= 3t^4 + 2t^2 - 4a_{11}t + 2a_{11}^2 + a_{11}^4, \\ p_0(t) &= -t^6 + 2a_{11}^2 t^4 - a_{11}^4 t^2. \end{aligned}$$

Therefore, the condition in Lemma 3.2 is satisfied and $\sigma_\varepsilon(C) \neq -\sigma_\varepsilon(C)$.

The proof is complete. □

Theorem 3.4. *Suppose $n \geq 3$ and $\varepsilon > 0$. Then a surjective map $\phi : M_n \to M_n$ satisfies*

$$\sigma_\varepsilon([A, B]) = \sigma_\varepsilon([\phi(A), \phi(B)]) \qquad \textit{for all } A, B \in M_n$$

if and only if there exist $\mu \in \{1, -1\}$, a unitary matrix $U \in M_n$, and a set $\mathcal{T}$ of normal matrices with at most two distinct eigenvalues such that

$$\phi(A) = \begin{cases} \mu U\psi(A)U^* + \nu_A I & \text{if } A \in M_n \setminus \mathcal{T}, \\ -\mu U\psi(A)U^* + \nu_A I & \text{if } A \in \mathcal{T}, \end{cases} \tag{3.1}$$

where $\nu_A \in \mathbb{C}$ depends on A, and ψ is one of the maps: $A \mapsto A$, $A \mapsto iA^t$.

Proof. To prove the sufficiency, if ψ has the first form, then $\sigma_\varepsilon([A, B]) = \sigma_\varepsilon([\phi(A), \phi(B)]) = \mu_A\mu_B\sigma_\varepsilon([A, B])$ if none, one, or both of $A, B \in \mathcal{T}$ by Proposition 3.3. If ψ has the second form, then $\sigma_\varepsilon([A, B]) = \sigma_\varepsilon([\phi(A), \phi(B)]) = -\mu_A\mu_B\sigma_\varepsilon([A^t, B^t]) = \mu_A\mu_B\sigma_\varepsilon([A, B])$ if none, one, or both of $A, B \in \mathcal{T}$ by Proposition 3.3.

To prove the necessity, we may compose ϕ by a map of the form $X \mapsto VXV^*$ and adjust ν_X if necessary so that ϕ has the form $A \mapsto \mu_A\psi(A)$, where ψ is one of the maps $A \mapsto A, A \mapsto A^t, A \mapsto \overline{A}, A \mapsto A^*$. Focusing on rank one Hermitian matrices, we see that one of the following happens.

(1) For any rank one $A = xx^*$, $\phi(A) = \mu_A A$.
(2) For any rank one $A = xx^*$, $\phi(A) = \mu_A A^t$.

Suppose (2) holds. We may replace ϕ by the map $X \mapsto i\phi(X)^t$. Then the modified map will satisfy condition (1). Thus, we can focus on the case when (1) holds, and prove that ϕ has the asserted form with $\psi(X) = X$ for all $X \in M_n$.

In the rest of the proof, we assume that (1) holds. Then we have either

$$\text{i) } \phi(A) = \mu_A A \text{ for all } A \in M_n, \quad \text{or} \quad \text{ii) } \phi(A) = \mu_A A^* \text{ for all } A \in M_n.$$

We will show that for some μ, we have $\mu_A = \mu$ for all $A \in M_n \setminus \mathcal{T}$ and $\mu_A = -\mu$ for all $A \in \mathcal{T}$ satisfying (3.1). Clearly, we need only consider non-scalar matrices.

Assertion 1. For every non-scalar matrix $A \in M_n$, $\mu_A \in \{-1, 1\}$.

To prove Assertion 1, let $A = xx^*$. If $B = yy^*$ such that $0 \neq [A, B]$, then $[A, B]$ is unitarily similar to $\mathrm{diag}(ai, -ai) \oplus O_{n-2}$ with $a = \sqrt{-\mathrm{tr}([A, B]^2)/2} > 0$ so that

$$\sigma_\varepsilon([A, B]) = D(-ai, \varepsilon) \cup D(0, \varepsilon) \cup D(ai, \varepsilon).$$

Because $\sigma_\varepsilon([\phi(A), \phi(B)]) = \mu_A\mu_B\sigma_\varepsilon([A, B])$, we see that $\mu_A\mu_B = \pm 1$.

Let $\mu = \mu_{E_{11}}$. Suppose $B = xx^*$ for a nonzero $x \in \mathbb{C}^n$. We can find $C = yy^*$ such that $[E_{11}, C] \neq 0$ and $[B, C] \neq 0$. Then $\mu\mu_C, \mu_B\mu_C \in \{1, -1\}$ so that $\mu\mu_C = \pm\mu_B\mu_C$. It follows that $\mu_B \in \{\mu, -\mu\}$.

Choose $B_j = x_j x_j^*$, $j = 1, 2$ so that $[E_{11}, B_1]$, $[E_{11}, B_2]$ and $[B_1, B_2] \neq 0$. Then

$$\mu\mu_{B_1},\ \mu\mu_{B_2},\ \mu_{B_2}\mu_{B_1} \in \{1, -1\}.$$

Hence, $\mu^2 \in \{-1, 1\}$. So we have either

$$\text{(a) } \mu^2 = -1 \ \Rightarrow\ \mu_B \in \{-i, i\} \text{ for all } B = xx^*, \quad \text{or}$$
$$\text{(b) } \mu^2 = 1 \ \Rightarrow\ \mu_B \in \{-1, 1\} \text{ for all } B = xx^*.$$

Next we will show that $\phi(A) = \mu_A A$ for all $A \in M_n$. Assume the contrary that $\phi(A) = \mu_A A^*$ for all $A \in M_n$. Let $B_1 = E_{11} + E_{13} + E_{31} + E_{33}$, $B_2 = E_{22} + E_{23} + E_{32} + E_{33}$ and $C = E_{11} + e^{i\pi/6} E_{22}$. Then

$$\sigma_\varepsilon([B_1, C]) = D(-i, \varepsilon) \cup D(i, \varepsilon) \cup D(0, \varepsilon)$$

and

$$\sigma_\varepsilon([\phi(B_1), \phi(C)]) = \mu_{B_1}\mu_C D(-i, \varepsilon) \cup D(i, \varepsilon) \cup D(0, \varepsilon).$$

Hence, $\mu_{B_1}\mu_C \in \{-1, 1\}$. By a direct computation,

$$\sigma_\varepsilon([B_2, C]) = D(-e^{-2\pi i/3}, \varepsilon) \cup D(e^{-2\pi i/3}, \varepsilon) \cup D(0, \varepsilon)$$

and

$$\sigma_\varepsilon([\phi(B_2), \phi(C)]) = \mu_{B_2}\mu_C \left(D(-e^{-\pi i/3}, \varepsilon) \cup D(e^{-\pi i/3}, \varepsilon) \cup D(0, \varepsilon) \right).$$

Since $\mu_{B_1} = \pm\mu_{B_2}$ and $\mu_{B_1}\mu_C \in \{-1, 1\}$, we have $\mu_{B_2}\mu_C \in \{-1, 1\}$. Hence, $\sigma_\varepsilon([\phi(B_2), \phi(C)]) \neq \sigma_\varepsilon([B_2, C])$, a contradiction. Therefore, we have $\phi(A) = \mu_A A$ for all $A \in M_n$.

For any non-scalar normal matrix B with spectral decomposition $\sum_{j=1}^n b_j x_j x_j^*$ with $b_1 \neq b_2$, let $C = yy^*$ with $y = x_1 + x_2$. Then $[B, C]$ is unitarily similar to $\mathrm{diag}(a, -a) \oplus O_{n-2}$. It follows that $\mu_B\mu_C \in \{1, -1\}$. Because $\mu_C\mu \in \{1, -1\}$, we see that $\mu_B \in \{\mu, -\mu\}$. Suppose B is non-normal. There is a unitary U such that $UBU^* = H + iG$, where $G = G^*$ is in diagonal form and $H = H^*$ has a nonzero $(1, 2)$ entry. Then for $C = UE_{11}U^*$, the matrix $[B, C]$ is unitarily similar $\mathrm{diag}(a, -a) \oplus O_{n-2}$. Again, we can conclude that $\mu_B = \pm\mu$. So, $\mu_B \in \{\mu, -\mu\}$ for every $B \in M_n$. Consequently, we have

$$\text{(c) } \mu_B \in \{-i, i\} \text{ for all } X \in M_n, \qquad \text{or} \qquad \text{(d) } \mu_B \in \{-1, 1\} \text{ for all } X \in M_n.$$

We claim that the condition (d) holds. To this end, let $D = \mathrm{diag}(1, -1) \oplus O_{n-2}$ and $B = E_{12}/2 + E_{23} + E_{31}$. Then $[D, B] = E_{12} - E_{23} - E_{31}$ is a unitary matrix with eigenvalues $\lambda_1 = 1$, $\lambda_2 = e^{i2\pi/3}$, $\lambda_3 = e^{i4\pi/3}$. Thus,

$$\sigma_\varepsilon([D, B]) = D(\lambda_1, \varepsilon) \cup D(\lambda_2, \varepsilon) \cup D(\lambda_3, \varepsilon).$$

We see that $\mu_B\mu_D = 1$ for such a matrix B. Similarly, if $C = -E_{21}/2 + iE_{32} - iE_{13}$, then $\mu_C\mu_D = 1$. Thus, $\mu_B = \mu_C$. Now, $[B, C] = (1 + i/4)E_{11} + (1 - i/4)E_{22} - 2E_{33}$. Then $\mu_B\mu_C\sigma_\varepsilon([B, C]) = \sigma_\varepsilon([B, C])$ will imply that $\mu_B\mu_C = 1$. Because, $\mu_B = \mu_C$, we see that $\mu_B = \mu_C \in \{-1, 1\}$. Hence, the condition (d) holds.

Assertion 2. There is $\mu \in \{1, -1\}$ such that $\mu_A = \mu$ if A is not a normal matrix with at most two distinct eigenvalues.

Proof. First we show that for any nonzero vectors x, y, f such that 1) $y, f \in x^\perp$, 2) $\{y, f\}$ is linearly independent and 3) $\mathrm{Re}(f^*y) \neq 0$, then the following holds.

$$\mu_{xf^*} = \mu_{yx^*} \tag{3.2}$$

Note that $C = [xf^*, yx^*] = (f^*y)xx^* - \|x\|^2 yf^*$ which has a matrix representation of the form

$$C = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & -\alpha & 0 \\ 0 & \beta & 0 \end{bmatrix} \oplus 0 = X \oplus 0$$

with $\alpha = f^*y\|x\|^2$, $\beta = \|x\|^2\sqrt{\|f\|^2\|y\|^2 - |f^*y|^2} \neq 0$. Then

$$\det(\lambda I_3 - (X - tI_3)^*(X - tI_3)) = \lambda^3 + p_2(t)\lambda^2 + p_1(t)\lambda + p_0(t),$$

where

$$\begin{aligned} p_2(t) &= -3t^2 - (2|\alpha|^2 + |\beta|^2), \\ p_1(t) &= 3t^4 + (4(\mathrm{Im}(\alpha))^2 + \beta^2)t^2 - 2\,\mathrm{Re}(\alpha)\beta^2 t + |\alpha|^2\left(|\alpha|^2 + \beta^2\right), \\ p_0(t) &= -t^6 + (\alpha^2 + \bar{\alpha}^2)t^4 - |\alpha|^4 t^2. \end{aligned}$$

Since $\mathrm{Re}(\alpha)$ and $\beta \neq 0$, the condition in Lemma 3.2 is satisfied. Therefore, $\sigma_\varepsilon(C) \neq -\sigma_\varepsilon(C)$. Since $\sigma_\varepsilon(C) = \mu_{xf^*}\mu_{yx^*}\sigma_\varepsilon(C)$, we have $\mu_{xf^*}\mu_{yx^*} = 1$, and thus $\mu_{xf^*} = \mu_{yx^*}$.

If xf^* and xu^* are rank one nilpotent and if $u \in f^\perp$, then (3.2) ensures that

$$\mu_{xf^*} = \mu_{(f+u)x^*} = \mu_{xu^*} = \mu_{u(x+f)^*} = \mu_{fu^*} = \mu_{(x+u)f^*} = \mu_{fx^*}.$$

So we have

$$\mu_{xf^*} = \mu_{xu^*} = \mu_{fx^*} \tag{3.3}$$

whenever the vectors x, f, u are pairwise orthogonal.

Next we show that

$$\mu_{xf^*} = \mu_{xu^*} \quad \text{for any nonzero vectors } f, u \in x^\perp. \tag{3.4}$$

Suppose f, u are nonzero vectors in $x^\perp$. If $u \in f^\perp$, the equality follows from (3.3). If $u = \lambda f$ for some nonzero scalar λ, taking $v \in \{x, f\}^\perp$ we have

$$\mu_{xf^*} = \mu_{xv^*} = \mu_{xu^*}.$$

If $u \notin f^\perp$ and the vectors u, f are linearly independent, then let $v = u - cf$, where $c = \dfrac{f^*u}{f^*f}$. Then $v \in \{x, f\}^\perp$ and $u^*v = u^*u - \dfrac{|f^*u|^2}{f^*f} \neq 0$. By (3.2) and (3.3), we have

$$\mu_{xu^*} = \mu_{vx^*} = \mu_{xf^*}.$$

Next, we show that $\mu_A = \mu_B$ for any rank one nilpotent matrices A, B. To this end, $A = xf^*$ and $B = yg^*$, taking unit vector $u \in \{x, y\}^\perp$ and using (3.4), we have

$$\mu_{xf^*} = \mu_{xu^*} = \mu_{yu^*} = \mu_{yg^*}.$$

By Proposition 3.3, if A is not a normal matrix with at most two distinct eigenvalues, then there is a rank one nilpotent B such that

$$-\sigma_\varepsilon([A, B]) \neq \sigma_\varepsilon([B, A]) = \mu_B\mu_A\sigma_\varepsilon([B, A]).$$

Thus, $\mu_A\mu_B = 1$, which implies $\mu_A = \mu_B$. The desired conclusion follows. $\square$

Acknowledgements

The research of the first author was partially supported by National Natural Science Foundation of China (No. 11271217). The research of the second and third authors was partially supported by USA National Science Foundation Grant (DMS 1007768/DMS 1007835) and Hong Kong General Research Fund (PolyU 502512/PolyU 502411). The second author is an honorary professor of the Shanghai University and an honorary professor of the University of Hong Kong.

References

[1] J.T. Chan, C.K. Li, N.S. Sze, Isometries for unitarily invariant norms, Linear Algebra Appl. 399 (2005) 53–70.
[2] J.T. Chan, C.K. Li, N.S. Sze, Mappings on matrices: invariance of functional values of matrix products, J. Aust. Math. Soc. 81 (2006) 165–184.
[3] W.S. Cheung, S. Fallat, C.K. Li, Multiplicative preservers on semigroups of matrices, Linear Algebra Appl. 355 (2002) 173–186.
[4] J. Cui, V. Forstall, C.K. Li, V. Yannello, Properties and preservers of the pseudospectrum, Linear Algebra Appl. 436 (2012) 316–325.
[5] J. Cui, C.K. Li, Y.T. Poon, Pseudospectra of special operators and pseudospectrum preservers, J. Math. Anal. Appl. (2014) 1261–1273.
[6] N. Guglielmi, M.L. Overton, Fast algorithms for the approximation of the pseudospectral abscissa and pseudospectral radius of a matrix, SIAM J. Matrix Anal. Appl. 32 (2011) 1166–1192.

[7] R. Guralnick, C.K. Li, Invertible preservers and algebraic groups III: preservers of unitary similarity (congruence) invariants and overgroups of some unitary subgroups, Linear Multilinear Algebra 43 (1997) 257–282.
[8] K.E. Gustafson, D.K.M. Rao, Numerical Range: The Field of Values of Linear Operators and Matrices, Universitext, Springer-Verlag, New York, 1997.
[9] R. Horn, C. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[10] H. Kestelman, Automorphisms of the field of complex numbers, Proc. Lond. Math. Soc. 53 (1951) 1–12.
[11] C.K. Li, E. Poon, N.S. Sze, Preservers for norms of Lie product, Oper. Matrices 3 (2009) 187–203.
[12] C.K. Li, N.K. Tsing, Linear operators preserving certain classes of functions on singular values of matrices, Linear Multilinear Algebra 26 (1990) 133–143.
[13] M. Marcus, M. Sandy, Conditions for the generalized numerical range to be real, Linear Algebra Appl. 71 (1985) 219–239.
[14] P. Šemrl, Non-linear commutativity preserving maps, Acta Sci. Math. (Szeged) 71 (2005) 781–819.