STUDIA MATHEMATICA 179 (1) (2007)

Approximating real linear operators

by

Marko Huhtanen and Olavi Nevanlinna (Helsinki)

Abstract. A framework to extend the singular value decomposition of a matrix to a real linear operator $\mathcal{M} : \mathbb{C}^n \to \mathbb{C}^p$ is suggested. To this end real linear operators called operets are introduced, to have an appropriate generalization of rank-one matrices. Then, adopting the interpretation of the singular value decomposition of a matrix as providing its nearest small rank approximations, $\mathcal{M}$ is approximated with a sum of operets.

1. Introduction. In real linear matrix analysis we are concerned with operators $\mathcal{M} : \mathbb{C}^n \to \mathbb{C}^p$ defined as

$$z \mapsto \mathcal{M}z = Mz + M_{\#}\tau z \tag{1.1}$$

for a pair of matrices $M, M_{\#} \in \mathbb{C}^{p\times n}$. Here τ is the conjugation operator $\tau z = \overline{z}$ allowing us to use interchangeably the abbreviation $M + M_{\#}\tau$ for $\mathcal{M}$. If $M_{\#} = 0$, then we have the ordinary matrix-vector product (in which case the operator is $\mathbb{C}$-linear) while for $M = 0$ we are dealing with an antilinear operator. We regard the set of real linear operators as a vector space over $\mathbb{C}$ and denote it by $\mathcal{M}_{p,n}$ (the addition operation is obvious while the scalar multiplication in this paper is defined from the left). By extending the standard $\mathbb{C}$-linear matrix analysis, real linear matrix analysis yields new insights into matrix theory. For computational purposes the interplay between $\mathbb{R}$-linearity and $\mathbb{C}$-linearity provides novel approaches to solving linear algebra problems. For the background, applications and motivation of this study, see [2, 10, 8, 9] where basic decompositions and spectral theory were introduced.

In this paper we present a real linear framework extending the singular value decomposition. Interpreting the singular value decomposition of a matrix $M \in \mathbb{C}^{p\times n}$ as a tool to solve

$$\min_{\operatorname{rank}(M_j)\le j} \|M - M_j\| \quad \text{for } j = 1, \dots, \min\{p, n\}, \tag{1.2}$$

2000 *Mathematics Subject Classification*: 15A18, 15A04.

Key words and phrases: real linear matrix analysis, operet, singular value decomposition, matrix nearness problem, approximation number.

Research of M. Huhtanen partially supported by National Science Foundation grant DMS–0209437.

for a unitarily invariant norm $\|\cdot\|$, which is arguably the most important practical application of the SVD having its matrix analytic origins in [14, 1], we consider analogous nearness problems in the real linear case. For this purpose we introduce real linear operators called operets, to have an appropriate generalization of rank-one matrices. By approximating $\mathcal{M}$ with sums of operets we obtain a finite sequence of nearness problems conforming with the classical formulation (1.2) in the sense that if $\mathcal{M}$ is $\mathbb{C}$-linear, then the respective approximations coincide. Hence, we obtain an extension of the concept of singular values. Unlike the matrix case, our approximations depend on the unitarily invariant norm used. Here we concentrate on the operator and Frobenius norms. Presently the most tractable one appears to be the Frobenius norm for which we manage to solve the problem completely leading to an SVD-like, in a certain sense optimal, expansion of the real linear operator $\mathcal{M}$. The terms in this expansion can be computed in sequence. More notably, because the standard complex inner product is employed, this representation of $\mathcal{M}$ consumes less storage than what the real singular value decomposition of the corresponding matrix representation of $\mathcal{M}$ in $\mathbb{R}^{2p\times 2n}$ requires. By using this reasoning conversely, the proposed expansion has many obvious uses for real problems. To mention immediate applications, starting with a real-entry data matrix, the approach yields new criteria as well as algorithms to optimally compress information. Also ill-conditioned problems can be solved by using our representation of $\mathcal{M}$. As a further illustration, a Karhunen–Loève expansion is derived within the framework proposed.

The paper is organized as follows. In Section 2 we introduce operets and consider representations of a given real linear operator $\mathcal{M}$ as a sum of operets. In Section 3 we approximate $\mathcal{M}$ with sums of operets in the operator and Frobenius norms. The corresponding approximation numbers are defined. In the Frobenius norm the problem is solved completely.

2. Representing a real linear operator as a sum of operets. The main use of the singular value decomposition in matrix computations, and in infinite-dimensional operator theory in connection with compactness, is to provide an expansion of, as well as low rank approximations to, a linear operator; see, e.g., [6]. (For historical accounts from different points of view, see also [15, 7, 11]. For many uses of the SVD, see [3].) Taking this as our starting point, we consider expanding a given real linear operator $\mathcal{M} = M + M_{\#}\tau \in \mathcal{M}_{p,n}$ as a sum of simpler ones.

As a first approach, one could readily employ the real form of $\mathcal{M}$, by which we mean its equivalent representation via the matrix

$$A = \begin{bmatrix} \operatorname{Re}(M + M_{\#}) & -\operatorname{Im}(M - M_{\#}) \\ \operatorname{Im}(M + M_{\#}) & \operatorname{Re}(M - M_{\#}) \end{bmatrix} \in \mathbb{R}^{2p\times 2n}, \tag{2.1}$$

which, conversely, provides a means of rewriting a real-entry matrix in complex form as a real linear operator $\mathcal{M} : \mathbb{C}^n \to \mathbb{C}^p$. For more details, see [2]. In this manner, after rewriting the rank-one matrices appearing in the singular value decomposition of A in their complex form, $\mathcal{M}$ can be represented as a sum of $2\min\{p, n\}$ real linear operators

$$z \mapsto \frac{1}{2}\,(uv^*z + uv^T\overline{z}) = u\,\mathrm{Re}(v^*z) \quad \text{with } u \in \mathbb{C}^p,\ v \in \mathbb{C}^n; \tag{2.2}$$

see [2, Section 2.3]. Operators of this form are also encountered in connection with the real linear Householder transformations, employed, for instance, in computing the QR-factorization of a real linear operator [2, Section 2.2].

In spite of these uses, the complex forms of rank-one matrices from $\mathbb{R}^{2p\times 2n}$ are not sufficiently versatile tools for real linear matrix computations. A reason for this is that in the real-entry formulation A of $\mathcal{M}$, the complex structure of the vector spaces $\mathbb{C}^n$ and $\mathbb{C}^p$ gets ignored when they are regarded as vector spaces over $\mathbb{C}$. Therefore, to give an example, if $\mathcal{M}$ is $\mathbb{C}$-linear, the prescribed representation consisting of the terms (2.2) does not yield the standard complex singular value decomposition of the matrix M. Nor does it provide very useful information on the spectral properties of $\mathcal{M}$. Recall that the SVD of a square matrix $M \in \mathbb{C}^{n\times n}$ gives an upper bound, typically sharp, on the number of distinct eigenvalues of M through the number of its nonzero singular values.

Remark. Let $p = n$. Then the spectrum of a real linear operator $\mathcal{M}$, denoted by $\sigma(\mathcal{M})$, is the collection of those $\lambda \in \mathbb{C}$ for which $\lambda I - \mathcal{M}$ is not invertible. Being the zero set of a bivariate polynomial $p : \mathbb{R}^2 \to \mathbb{R}$ of degree $2n$, the spectrum is an algebraic plane curve of degree at most $2n$ [2, 10].

For more applicable tools, let us proceed by analogy by reviewing the $\mathbb{C}$-linear case. To this end, take a matrix $M \in \mathbb{C}^{p\times n}$ and fix a unit vector $v \in \mathbb{C}^n$. Then a natural operator to approximate M is the rank-one matrix uv^* with $u = Mv$. Correspondingly, we introduce the following generalization of rank-one matrices in the real linear case.

Definition 2.1. Let $u, u_\# \in \mathbb{C}^p$ and $v \in \mathbb{C}^n$. A real linear operator of the form

$$z \mapsto uv^*z + u_\#v^T\overline{z}$$

is called an *operet* ([1]).

([1]) Analogously to wavelets being thought of as "small" waves, we regard operets as small operators.

Equivalently, we denote an operet by $\mathcal{O} = (u + u_{\#}\tau)v^*$. (Here we use $\tau v^* = v^T\tau$.) In this manner we obtain a slightly more general structure than (2.2) that remains manageable enough to resemble scalar real linear operators in the following sense.

PROPOSITION 2.2. *Let $p = n$ with $n > 1$. Then the spectrum of $\mathcal{O} = (u+u_{\#}\tau)v^*$ is the union of the origin and the circle of radius $|v^*u_{\#}|$ centered at v^*u.*

Proof. By taking z from the orthogonal complement of v, we can infer that the origin is in the spectrum. For the circle, we can consider $z \mapsto vu^*z + vu_{\#}{}^T\overline{z}$ which corresponds to the adjoint of this real linear operator when $\mathbb{C}^n$ is regarded as a vector space over $\mathbb{R}$. Now, take $z = e^{i\theta}v$ to have $e^{i\theta}v \mapsto e^{i\theta}v(u^*v+e^{-2i\theta}u_{\#}{}^T\overline{v})$. Hence, for this operator, $e^{i\theta}v$ is an eigenvector corresponding to the eigenvalue $u^*v + e^{-2i\theta}u_{\#}{}^T\overline{v}$. As θ varies, we get the circle claimed, after complex conjugating. ∎

We denote by $\|z\|$ the 2-norm of a vector z. The operator norm of a real linear operator $\mathcal{M}$ is defined as $\|\mathcal{M}\|_2 = \max_{z\neq 0}\|\mathcal{M}z\|/\|z\|$.

For an operet $\mathcal{O} = (u + u_{\#}\tau)v^*$ we have

$$\max_{\|z\|=1}\|uv^*z + u_{\#}v^T\overline{z}\|^2 = \max_{\theta\in[0,2\pi)}\|v\|^2(\|u\|^2 + 2\,\mathrm{Re}\,(e^{i\theta}u_{\#}{}^*u) + \|u_{\#}\|^2)$$

so that

$$\|\mathcal{O}\|_2 = (\|u\|^2 + 2|u_{\#}{}^*u| + \|u_{\#}\|^2)^{1/2}\|v\|. \tag{2.3}$$

By forming the sum of j operets we obtain an extension of the set of matrices of rank at most j with

$$z \mapsto UV^*z + U_{\#}V^T\overline{z}, \quad \text{where } U, U_{\#} \in \mathbb{C}^{p\times j} \text{ and } V \in \mathbb{C}^{n\times j}, \tag{2.4}$$

for which we also use the abbreviation $(U+U_{\#}\tau)V^*$. Without loss of generality, V can be assumed to have orthonormal columns after, for instance, QR-factorizing the original matrix V. Even this guarantees no uniqueness since $(U + U_{\#}\tau)V^* = (UW + U_{\#}\overline{W}\tau)(VW)^*$ for any unitary matrix $W \in \mathbb{C}^{j\times j}$. Obviously, if $U_{\#} = 0$, then we have a matrix of rank at most j.

REMARK. Assume again $p = n$. As an analogy of elementary matrices, for $U, U_{\#}, V \in \mathbb{C}^{n\times j}$, consider the real linear operator $\mathcal{M} = I+(U+U_{\#}\tau)V^*$. Real linear analogues of Gauss transformations yield operators of this type, with $j = 1$, in connection with the LU-factorization of a real linear operator [2, Section 2.2]. Assuming the inverse of $\mathcal{M}$ exists, it has the same structure $I + (\widehat{U} + \widehat{U}_{\#}\tau)V^*$ with $\widehat{U}, \widehat{U}_{\#} \in \mathbb{C}^{n\times j}$. Namely, write $\mathcal{M}^{-1} = N + N_{\#}\tau$ and consider the identity $\mathcal{M}\mathcal{M}^{-1} = I$. Then, with the help of the Sherman–Morrison formula, we get

$$N = (I + \widetilde{U}V^*)^{-1} = I + \widehat{U}V^*, \quad \text{where} \quad \widetilde{U} = U - U_{\#}V^T\,\overline{M^{-1}}\,\overline{U}_{\#}.$$

Hence

$$N_{\#} = -M^{-1}M_{\#}\overline{N} = -M^{-1}U_{\#}(I + V^T\overline{\widehat{U}})V^T,$$

as claimed. In the case $j = 1$ this gives us the following formulae.

EXAMPLE 1. If $u, u_{\#}, v \in \mathbb{C}^n$, then

$$\widehat{u} = \frac{v^T\overline{u}_{\#}u_{\#} - (1 + v^T\overline{u})u}{1 + v^*u + v^T\overline{u} + |v^*u|^2 - |v^*u_{\#}|^2},$$

$$\widehat{u}_{\#} = \frac{v^*u_{\#}u - (1 + v^*u)u_{\#}}{1 + v^*u + v^T\overline{u} + |v^*u|^2 - |v^*u_{\#}|^2}.$$

DEFINITION 2.3. The *adjoint* $\mathcal{M}^*$ of $\mathcal{M} = M + M_{\#}\tau$ is $M^* + {M_{\#}}^T\tau$.

Obviously, if $\mathcal{M} : \mathbb{C}^n \to \mathbb{C}^p$, then $\mathcal{M}^* : \mathbb{C}^p \to \mathbb{C}^n$. For $\mathcal{M}$ and $\mathcal{N}$ of appropriate size, we have the familiar relationship $(\mathcal{M}\mathcal{N})^* = \mathcal{N}^*\mathcal{M}^*$. In case $p = n$, $\mathcal{M}^*$ is invertible if and only if $\mathcal{M}$ is.

Denote by $(z, w) = w^*z$ the standard inner product on $\mathbb{C}^n$. Then a simple computation gives

$$(\mathcal{M}v, w) = (v, \mathcal{M}^*w) - 2i\,\mathrm{Im}\,(v, {M_{\#}}^T\overline{w}). \tag{2.5}$$

(Hence, if $(z, w)_{\mathbb{R}} = \mathrm{Re}\, w^*z$ denotes the standard real inner product on $\mathbb{C}^n$, we have the familiar relationship $(\mathcal{M}v, w)_{\mathbb{R}} = (v, \mathcal{M}^*w)_{\mathbb{R}}$.) We say that $\mathcal{M}$ is *self-adjoint* if $\mathcal{M}^* = \mathcal{M}$. This property has implications, to give an example, on the location of the spectrum of $\mathcal{M}$; see [10].

For a given real linear operator $\mathcal{M} = M + M_{\#}\tau$ we are interested in finding its representation as a sum of j operets with the smallest possible integer j. To this end, denote by $R(\mathcal{M}^*)$ the range of the adjoint of $\mathcal{M}$. It is an $\mathbb{R}$-linear subspace of $\mathbb{C}^n$ such that the smallest subspace containing $R(\mathcal{M}^*)$ is $(R(\mathcal{M}^*)^\perp)^\perp$. (Recall that we regard $\mathbb{C}^n$ as a vector space over $\mathbb{C}$ with the standard complex inner product.) Obviously $(R(\mathcal{M}^*)^\perp)^\perp \subset R(M^*) + R({M_{\#}}^T)$, the sum of the ranges of the matrices M and $M_{\#}$, so that its dimension is at most $\mathrm{rank}(M) + \mathrm{rank}(M_{\#})$.

PROPOSITION 2.4. *$\mathcal{M}$ can be represented as a sum of $k = \dim (R(\mathcal{M}^*)^\perp)^\perp$ operets. Moreover, k is the smallest such integer.*

Proof. Let $v_1, \ldots, v_k$ be an orthonormal basis of $(R(\mathcal{M}^*)^\perp)^\perp$ and let $u_j = Mv_j$ and $u_{\#j} = M_{\#}\overline{v}_j$ be the columns of U and $U_{\#}$, both from $\mathbb{C}^{p\times k}$, while $V = [v_1 \cdots v_k] \in \mathbb{C}^{n\times k}$. Take v from the orthogonal complement of $(R(\mathcal{M}^*)^\perp)^\perp$, i.e., from $R(\mathcal{M}^*)^\perp$. We show that $\mathcal{M}v = 0$. Indeed, by using (2.5) for a vector $w \in \mathbb{C}^p$, we have $(\mathcal{M}v, w) = -2i\,\mathrm{Im}\,(v, {M_{\#}}^T\overline{w})$. But $(\mathcal{M}v, w)$ being pure imaginary for any w forces $\mathcal{M}v = 0$. Hence $\mathcal{M} = (U + U_{\#}\tau)V^*$.

We identify $V \in \mathbb{C}^{n\times(k-1)}$ with the subspace of $\mathbb{C}^n$ its columns span. Clearly, V cannot contain $R(\mathcal{M}^*)$. (Proof: Then $V \cap (R(\mathcal{M}^*)^\perp)^\perp$ would be

of dimension at most $k-1$ containing $R(\mathcal{M}^*)$.) Take $z \in V^\perp$ not in $R(\mathcal{M}^*)^\perp$. Then for some $w \in \mathbb{C}^p$ we have $0 \neq e^{-i\theta}(\mathcal{M}^* w, z) = (\mathcal{M}^* w, e^{i\theta} z) = (w, \mathcal{M} e^{i\theta} z) - 2i \operatorname{Im}(w, M_\# e^{-i\theta} \overline{z})$ by (2.5) and for any $\theta \in [0, 2\pi)$. Choose θ such that $\operatorname{Im}(w, M_\# e^{-i\theta} \overline{z}) = 0$. Hence $\mathcal{M}(I - P) e^{i\theta} z = \mathcal{M} e^{i\theta} z \neq 0$. ∎

Note that $\dim (R(\mathcal{M}^*)^\perp)^\perp$ need not equal $\dim (R(\mathcal{M})^\perp)^\perp$. For a simple illustration of this, consider a suitable operet.

Definition 2.5. $\dim (R(\mathcal{M}^*)^\perp)^\perp$ is the *right-rank* and $\dim (R(\mathcal{M})^\perp)^\perp$ the *left-rank* of a real linear operator $\mathcal{M}$.

To find the right-rank (resp. left-rank), compute the rank of the matrix $\left[\begin{smallmatrix} M \\ \overline{M}_\# \end{smallmatrix}\right]$ (resp. $\left[\begin{smallmatrix} M^* \\ M_\#^* \end{smallmatrix}\right]$); see Theorem 3.6 below.

We set $\mu(\mathcal{M}) = (\widetilde{k}, k) \in \mathbb{N}^2$ where $\widetilde{k}$ is the left-rank and k the right-rank of $\mathcal{M}$. We clearly have $|\widetilde{k} - k| \leq \min\{\widetilde{k}, k\}$. Obviously, if either $M = 0$ or $M_\# = 0$, then $\widetilde{k} = k$. For an operet $\mathcal{O}$ we have, generically, $\mu(\mathcal{O}) = (2, 1)$. For a nongeneric case, if $\mathcal{O}$ is the complex form (2.2) of a rank-one matrix from $\mathbb{R}^{2p \times 2n}$, then $\mu(\mathcal{O}) = (1, 1)$.

Example 2. Assume $M = \alpha M_\#$ with $\alpha \in \mathbb{C} \setminus \{0\}$. Then using the SVD of $M_\#$ gives $\min\{\widetilde{k}, k\} \leq \operatorname{rank}(M_\#)$.

Proposition 2.6. *Let $\mathcal{N}$ be a real linear operator and K a matrix of appropriate size such that $\mathcal{N}\mathcal{M}K$ is defined. Then*

$$\text{right-rank}(\mathcal{N}\mathcal{M}K) \leq \text{right-rank}(\mathcal{M}).$$

Proof. For $\mathcal{M} = (U + U_\# \tau) V^* : \mathbb{C}^n \to \mathbb{C}^p$ and $\mathcal{N} = N + N_\# \tau : \mathbb{C}^p \to \mathbb{C}^m$ we have

$$\mathcal{N}\mathcal{M} = (NU + N_\# \overline{U}_\# + (NU_\# + N_\# \overline{U})\tau) V^*. \tag{2.6}$$

For a product with a matrix $K \in \mathbb{C}^{n \times l}$ from the right, the rank of $V^* K$ is at most the rank of V^*. ∎

In case $p = n$, recall that the spectrum of $\mathcal{M}$ is an algebraic plane curve of degree at most $2n$. With $\mu(\mathcal{M})$ we can give a better estimate.

Theorem 2.7. *Let $p = n$. The spectrum of $\mathcal{M}$ with $\mu(\mathcal{M}) = (\widetilde{k}, k)$ is annihilated by a nonzero bivariate polynomial of degree at most $2(\min\{\widetilde{k}, k\} + 1)$.*

Proof. Assume $k \leq \widetilde{k}$, otherwise proceed with $\mathcal{M}^*$ and use the fact that $\overline{\sigma(\mathcal{M})}$ equals $\sigma(\mathcal{M}^*)$; see [2].

Since $\mathcal{M}$ is of the form (2.4) we have $\mathcal{M}^* = V(U^* + {U_\#}^T \tau)$ and thus the span of the columns of V is an invariant subspace of $\mathcal{M}^*$. Moreover, a nonzero $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathcal{M}^*$ if and only if λ is an eigenvalue of $\mathcal{M}^*$ restricted to the span of the columns of V. The characteristic bivariate

polynomial [2] of this restriction is of degree at most $2k$. To include the origin we multiply this by the monomial xy to have the claim. ∎

COROLLARY 2.8. *Let q be a polynomial. Then the spectrum of $q(\mathcal{M})$ is annihilated by a nonzero bivariate polynomial of degree at most $2(k+1)$.*

Proof. If $q(z) = \sum_{j=0}^{d} \alpha_j z^j$, then take $\widehat{q}(\mathcal{M}) = \sum_{j=1}^{d} \alpha_j \mathcal{M}^j$ whose spectrum is $\sigma(q(\mathcal{M}))$ translated by $-\alpha_0$. For $\widehat{q}(\mathcal{M})$ use (2.6) inductively to obtain the representation (2.4) with the same V as for $\mathcal{M}$. ∎

REMARK. When dealing with an $\mathbb{R}$-linear subspace $V_{\mathbb{R}}$ of $\mathbb{C}^n$, as above in the case of $R(\mathcal{M}^*)$, it is of interest to measure its "distance" from being a subspace of $\mathbb{C}^n$. To this end, let a real linear operator $\mathcal{P}$ with $\mathcal{P}^2 = \mathcal{P}$ and $\mathcal{P}^* = \mathcal{P}$ have range $V_{\mathbb{R}}$. Then $\mathcal{P}$ is unique and can be represented as $\mathcal{P}(z) = Q\,\mathrm{Re}(Q^*z)$ for some $Q \in \mathbb{C}^{n\times j}$ satisfying $\mathrm{Re}(Q^*Q) = I$. The norm, or the singular values of the antilinear part of $\mathcal{P}$ yield an appropriate measure equaling zero if and only if $V_{\mathbb{R}}$ is a subspace of $\mathbb{C}^n$.

3. Approximating a real linear operator with a sum of operets. In what follows we are interested in approximating a real linear operator $\mathcal{M}$ with a sum of j operets for $j \le$ right-rank$(\mathcal{M})$. For this purpose we employ norms on $\mathcal{M}_{p,n}$ that are unitarily invariant, i.e., satisfy $\|U\mathcal{M}V\| = \|\mathcal{M}\|$ for any unitary matrices $U \in \mathbb{C}^{p\times p}$ and $V \in \mathbb{C}^{n\times n}$. For instance, the operator norm is unitarily invariant.

Denote by $\mathbb{P}_j$ the set of orthogonal projectors on $\mathbb{C}^n$ of rank at most j. We look for an appropriate formulation of the nearness problems (1.2) in the real linear case. To have a concept that extends to infinite dimensions, for a given real linear operator $\mathcal{M} = M + M_{\#}\tau \in \mathcal{M}_{p,n}$ we consider the approximation problem

$$\min_{P\in\mathbb{P}_j} \|\mathcal{M}(I-P)\|_2 \tag{3.1}$$

where $\|\cdot\|_2$ is the operator norm. It is immediate that if $P = VV^*$ solves this, with V having orthonormal columns, then $\mathcal{M}P = (MV + M_{\#}\overline{V}\tau)V^*$ yields a best approximation to $M + M_{\#}\tau$ as a sum of j operets. In case $M_{\#} = 0$ we obtain a formulation equivalent to the classical nearness problem for the matrix M leading to its singular value decomposition. Therefore it seems natural to call the minimum value (3.1) the $(j+1)$th *approximation number* of $\mathcal{M}$ with respect to the operator norm and denote it by $\sigma_{2,j+1}(\mathcal{M})$ for $j = 0, 1, 2, \ldots, n-1$.

EXAMPLE 3. Let $p = n$. For the conjugation operator we have $\tau^* = \tau$ and $\sigma_{2,j}(\tau) = 1$ for $j = 1, \ldots, n$.

By considering the real formulations, we can infer that $\|\mathcal{M}\|_2 = \|\mathcal{M}^*\|_2$. Equivalently, $\sigma_{2,1}(\mathcal{M}) = \sigma_{2,1}(\mathcal{M}^*)$. However, even for $p = n$ the correspond-

ing approximation problem (3.1) for the adjoint of $\mathcal{M}$ yields different results in the sense that

$$\min_{P\in\mathbb{P}_j} \|\mathcal{M}^*(I-P)\|_2 \tag{3.2}$$

need not equal $\sigma_{2,j+1}(\mathcal{M})$ for $j \geq 1$. For example, consider a suitable operet.

The Frobenius norm, aside from the operator norm, is another classical unitarily invariant norm employed in connection with the singular value decomposition of a matrix. Let us show how to solve

$$\min_{P\in\mathbb{P}_j} \|\mathcal{M}(I-P)\|_{\mathrm{F}}, \tag{3.3}$$

where the Frobenius norm of a real linear operator $\mathcal{M} = M + M_{\#}\tau$ is defined by $\|\mathcal{M}\|_{\mathrm{F}} = (\|M\|_{\mathrm{F}}^2 + \|M_{\#}\|_{\mathrm{F}}^2)^{1/2}$. Clearly, this yields a unitarily invariant norm on $\mathcal{M}_{p,n}$. The corresponding inner product can be defined as $(\mathcal{M},\mathcal{N})_F = \mathrm{trace}(N^*M + {N_{\#}}^*M_{\#})$ for $\mathcal{M} = M + M_{\#}\tau$ and $\mathcal{N} = N + N_{\#}\tau$ in $\mathcal{M}_{p,n}$.

Since the Frobenius norm of a matrix is preserved under taking the complex conjugate, for any orthogonal projector $P = VV^*$ we obtain

$$\|\mathcal{M}(I-P)\|_{\mathrm{F}}^2 = \|M - MVV^*\|_{\mathrm{F}}^2 + \|\overline{M}_{\#} - \overline{M}_{\#}VV^*\|_{\mathrm{F}}^2. \tag{3.4}$$

Hence, to find $V \in \mathbb{C}^{n\times j}$ with orthonormal columns solving (3.3), we can employ the SVD of the matrix $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right] \in \mathbb{C}^{2p\times n}$. This implies that the singular values of $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$ are of interest for the corresponding real linear operator $\mathcal{M}$. In particular, since for $M_{\#} = 0$ (resp. $M = 0$) we obtain the singular value decomposition of the matrix M (resp. $M_{\#}$), we make the following definition.

Definition 3.1. The *approximation numbers of $M + M_{\#}\tau$ with respect to the Frobenius norm* are the singular values of the matrix $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$.

Equivalently, the approximation numbers with respect to the Frobenius norm equal the square roots of the eigenvalues of the $\mathbb{C}$-linear part of $\mathcal{M}^*\mathcal{M}$.

Example 4. The $\mathbb{C}$-linear part of $\mathcal{M}^*\mathcal{M}$ is the matrix $M^*M + {M_{\#}}^T\overline{M}_{\#}$. For a curious special case, assume that $M_{\#} = \kappa I$ with $\kappa \in \mathbb{C}$, i.e., M is "conjugate-translated". For such an $\mathcal{M}$ the behavior of the approximation numbers with respect to the Frobenius norm when κ varies is well understood.

Denote the approximation numbers of $\mathcal{M}$ with respect to the Frobenius norm by $\sigma_{\mathrm{F},j}(\mathcal{M})$, for $j = 1,\ldots,\min\{2p,n\}$. The number of nonzero approximation numbers equals the right-rank k of $\mathcal{M}$. Hence we have $k \leq \min\{\mathrm{rank}(M) + \mathrm{rank}(M_{\#}), n\}$. Moreover, $\|\mathcal{M}\|_{\mathrm{F}} = (\sum_{j=1}^{k} \sigma_{\mathrm{F},j}(\mathcal{M})^2)^{1/2}$. If $p = n$ and $\mathcal{M}$ has a zero approximation number, then $\mathcal{M}$ is obviously not invertible. The converse is not true; see Example 5 below.

The approximation problem (3.3) gives a good reason for using a particular basis of $\mathbb{C}^n$ once we take the full SVD $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right] = U\Sigma V^*$ of $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$. Then, by considering $\mathcal{M} = (MV + M_{\#}\overline{V}\tau)V^*$, we can represent $\mathcal{M}$ as the sum of operets as

$$\mathcal{M} = \sum_{l=1}^{k} (u_l + u_l^{\#}\tau)v_l^*, \tag{3.5}$$

where $V = [v_1 \cdots v_n]$ so that $[u_1 \cdots u_n] = MV = [Mv_1 \cdots Mv_n]$ while $[u_1^{\#} \cdots u_n^{\#}] = M_{\#}\overline{V} = [M_{\#}\overline{v}_1 \cdots M_{\#}\overline{v}_n]$. (This should be compared with (2.2).) From this representation we can recover the approximation numbers by noticing that

$$\sigma_{\mathrm{F},j}(\mathcal{M}) = (\|u_j\|^2 + \|u_j^{\#}\|^2)^{1/2}.$$

It follows that these approximation numbers can be interpreted to yield a refined average of the way $\mathcal{M}$ acts. More precisely, from (2.3) we obtain the operator norm of an operet. Similarly, the minimum of $z \mapsto \|uv^*z + u_{\#}v^T\overline{z}\|$ for z of unit length restricted to the span of v over $\mathbb{C}$ equals $(\|u\|^2 - 2|u_{\#}{}^*u| + \|u_{\#}\|^2)^{1/2}\|v\|$. Hence, squaring and taking the sum with (2.3) squared leads to vanishing of the inner product term. We can thus conclude that, with respect to this averaging, the representation (3.5) yields an optimally decreasing expansion of $\mathcal{M}$, analogously to the way the SVD of a matrix $M \in \mathbb{C}^{p\times n}$ can be interpreted to represent M as the sum of rank-one matrices decreasing in norm in an optimal way. This optimality is actually realized in terms of a norm; see (3.8) below.

Assume $n \le p$, otherwise consider $\mathcal{M}^*$. To represent $\mathcal{M}$ through the singular value decomposition of its real formulation A, we need $2n$ vectors from $\mathbb{C}^p$ and $2n$ vectors from $\mathbb{C}^n$; see (2.2). In all, this means storing $n(2p + 2n)$ complex numbers. It would take the same number of vectors to approximate the matrices M and $M_{\#}$ separately through their respective singular value decompositions (which could be used to give somewhat naive approximations to $\mathcal{M}$). The representation (3.5) takes at most $2n$ vectors from $\mathbb{C}^p$ and n vectors from $\mathbb{C}^n$ yielding thereby a more "compressed" optimal expansion of a real linear operator, as we need to store only $n(2p + n)$ complex numbers. These savings in storage were a reason for our way of defining an operet.

Just as for matrices, the full representation (3.5) of $\mathcal{M}$ may not be of interest to us. Rather, let us truncate the expansion up to the jth approximation number with respect to the Frobenius norm to have

$$\mathcal{M}_j = \sum_{l=1}^{j} (u_l + u_l^{\#}\tau)v_l^* = \mathcal{U}_j V_j^* \tag{3.6}$$

with $j < n$ (or rather $j \ll n$ in applications), where $\mathcal{U}_j = U_j + U_j^{\#}\tau$ with $\mathcal{U}_j : \mathbb{C}^j \to \mathbb{C}^p$ while the matrix V_j consists of the first j columns of V. In practice one might be interested in the pseudo-inverse of $\mathcal{M}_j$ to solve an overdetermined system $\mathcal{M}_j z = b$, with $b \in \mathbb{C}^p$, in the least squares sense. To this end, compute the real linear QR-factorization $\mathcal{U}_j = \mathcal{Q}\mathcal{R}$ of $\mathcal{U}_j$; see [2]. Then

$$\min_{z\in\mathbb{C}^n} \|\mathcal{Q}\mathcal{R}V_j^* z - b\| = \min_{z\in\mathbb{C}^n} \|\mathcal{R}V_j^* z - \mathcal{Q}^* b\| \tag{3.7}$$

so that, assuming the kernel of $\mathcal{U}_j$ is zero, $w \mapsto V_j\mathcal{R}^{-1}\mathcal{Q}^* w$ from $\mathbb{C}^p$ to $\mathbb{C}^n$ is the solution operator.

Another application of these ideas arises from data compression. If we are given a real-entry data matrix A, then it can be compressed, as an alternative to using the singular value decomposition of A, by employing the approximation (3.6) after rewriting A in its complex form $\mathcal{M}$. As a criterion for choosing j, use the magnitude of $\sigma_{\mathrm{F},j+1}(\mathcal{M})$.

The operator $\mathcal{Q}$ appearing in (3.7) is an isometry.

DEFINITION 3.2. Let $p=n$. An invertible $\mathcal{M}$ is an *isometry* if $\mathcal{M}^{-1}=\mathcal{M}^*$.

For an isometry, Mz and $M_{\#}\overline{z}$ are orthogonal for any $z \in \mathbb{C}^n$. In particular, we have $u_j^{\#\,*} u_j = 0$ for $j = 1, \ldots, n$ in (3.5). Hence the approximation numbers of an isometry with respect to the Frobenius norm are equal to 1. The converse does not hold.

EXAMPLE 5. Assume $u, u_{\#}, v \in \mathbb{C}^2$ are such that u and $u_{\#}$ are orthonormal and v is of unit length and consider an operet $\mathcal{O} = (u + u_{\#}\tau)v^*$. Then $\sigma_{\mathrm{F},1}(\mathcal{O}) = \sqrt{2}$ and $\sigma_{\mathrm{F},2}(\mathcal{O}) = 0$ while for its adjoint $\mathcal{O}^* = vu^* + v\tau u_{\#}{}^*$ we have $\sigma_{\mathrm{F},1}(\mathcal{O}^*) = \sigma_{\mathrm{F},2}(\mathcal{O}^*) = 1$ so that $\mu(\mathcal{O}) = (2, 1)$. Note that $\mathcal{O}^*$ is not invertible.

In spite of the fact that individual approximation numbers with respect to the Frobenius norm are not preserved under taking the adjoint in general, we have the following proposition.

PROPOSITION 3.3. *Let $\mathcal{M} \in \mathcal{M}_{p,n}$ with $\mu(\mathcal{M}) = (\widetilde{k}, k)$. Then*

$$\sum_{j=1}^{k} \sigma_{\mathrm{F},j}(\mathcal{M})^2 = \sum_{j=1}^{\widetilde{k}} \sigma_{\mathrm{F},j}(\mathcal{M}^*)^2.$$

Proof. For the adjoint of $\mathcal{M}$ the approximation numbers with respect to the Frobenius norm are determined by the matrix $\left[\begin{smallmatrix} M \\ M_{\#} \end{smallmatrix}\right]$ whose Frobenius norm equals that of $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$. ∎

Being defined through the matrix $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$, the approximation numbers with respect to the Frobenius norm satisfy the corresponding maxmin and

minmax characterization (see, e.g., [5, Theorem 7.3.10]). For the sum of two real linear operators $\mathcal{M}$ and $\mathcal{N}$ these approximation numbers behave in the classical manner, i.e.,

$$\sigma_{\mathrm{F},i+j-1}(\mathcal{M}+\mathcal{N}) \le \sigma_{\mathrm{F},i}(\mathcal{M}) + \sigma_{\mathrm{F},j}(\mathcal{N}) \tag{3.8}$$

for $1 \le i, j \le \min\{2p, n\}$ and $i+j \le \min\{2p, n\}+1$. This follows immediately from [6, Theorem 3.3.16(a)]. Therefore, let the Ky Fan j-norms ([2]) with respect to the Frobenius norm on $\mathcal{M}_{p,n}$ be defined as

$$\sum_{k=1}^{j} \sigma_{\mathrm{F},k}(\mathcal{M}).$$

These are unitarily invariant norms but, unlike the case of the operator or Frobenius norm, the norm of the adjoint of $\mathcal{M}$ need not equal the norm of $\mathcal{M}$. The Ky Fan 1-norm is the most interesting since it takes the maximum average length over the images of one-dimensional subspaces of $\mathbb{C}^n$ as described in connection with the expansion (3.5).

Remark. Ky Fan j-norms are unitarily invariant but not isometrically invariant norms on $\mathcal{M}_{p,n}$. We say that a norm $\|\cdot\|$ is *isometrically invariant* if $\|\mathcal{U}\mathcal{M}\mathcal{W}\| = \|\mathcal{M}\|$ for any isometries $\mathcal{U}$ and $\mathcal{V}$. The operator and Frobenius norm are isometrically invariant.

In view of this, the approximation numbers are preserved under the following operation.

Proposition 3.4. *Let $\mathcal{M} \in \mathcal{M}_{p,n}$. If $\mathcal{W}$ is an isometry and W is a unitary matrix, then*

$$\sigma_{2,j}(\mathcal{M}) = \sigma_{2,j}(\mathcal{W}^*\mathcal{M}W) \quad \textit{and} \quad \sigma_{\mathrm{F},j}(\mathcal{M}) = \sigma_{\mathrm{F},j}(\mathcal{W}^*\mathcal{M}W)$$

for $j = 1, \dots, k$.

Proof. For the operator norm the claim follows from $\|\mathcal{W}^*\mathcal{M}W\|_2 = \|\mathcal{M}\|_2$. Hence, if P solves (3.1) then $\widehat{P} = W^*PW$ solves the problem for $\mathcal{W}^*\mathcal{M}W$.

For the Frobenius norm, it is clear that the singular values of $\left[\begin{smallmatrix} MW \\ M_{\#}W \end{smallmatrix}\right]$ equal the singular values of $\left[\begin{smallmatrix} M \\ M_{\#} \end{smallmatrix}\right]$ so let us consider $\mathcal{W}^*\mathcal{M}$ with $\mathcal{W}^* = U + U_{\#}\tau$. To this corresponds

$$\begin{bmatrix} UM + U_{\#}\overline{M}_{\#} \\ \overline{U}M_{\#} + \overline{U}_{\#}M \end{bmatrix} = \begin{bmatrix} N \\ \overline{N}_{\#} \end{bmatrix}.$$

By using $U^*U + {U_{\#}}^T\overline{U}_{\#} = I$ and $U^*U_{\#} + {U_{\#}}^T\overline{U} = 0$, we deduce that

$$v^*(N^*N + {N_{\#}}^T\overline{N}_{\#})v = v^*(M^*M + {M_{\#}}^T\overline{M}_{\#})v$$

([2]) The *Ky Fan j-norm* of a matrix M is defined by $\sum_{k=1}^{j} \sigma_{2,k}(M)$.

for any vector $v \in \mathbb{C}^n$. Hence the norms of $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} N \\ \overline{N}_{\#} \end{smallmatrix}\right]$ are equal. Since this is true for any $\mathcal{M}$, it holds for $\mathcal{M}(I-P)$, proving the claim. ∎

As an important special case, recall that τ is an isometry.

In this manner in the real linear case the "two-sided rotations" allow more freedom in multiplications from the left. (For two-sided rotations, see [6, Example 7.4.13].) Note that when acted upon by an isometry from the left, $\mathcal{M}$ can change drastically. As an extreme, consider the case of $\mathcal{W}^*\mathcal{M}$ becoming $\mathbb{C}$-linear.

PROPOSITION 3.5. *Let $N + N_{\#}\tau = UM + VM_{\#}\tau$ with unitary $U, V \in \mathbb{C}^{p\times p}$. Then $\sigma_{\mathrm{F},j}(\mathcal{N}) = \sigma_{\mathrm{F},j}(\mathcal{M})$ for $j = 1, \ldots, k$.*

Proof. The approximation numbers of $\mathcal{N}$ with respect to the Frobenius norm are obtained from the SVD of the matrix $\left[\begin{smallmatrix} UM \\ \overline{V}\overline{M}_{\#} \end{smallmatrix}\right] = \left[\begin{smallmatrix} U & 0 \\ 0 & \overline{V} \end{smallmatrix}\right]\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$. Since $\left[\begin{smallmatrix} U & 0 \\ 0 & \overline{V} \end{smallmatrix}\right]$ is unitary, we have the claim. ∎

The approximations given by (3.1) differ from those provided by (3.3). To see this with $P \in \mathbb{P}_1$ consider the following example.

EXAMPLE 6. Take $\mathcal{M} = \mathcal{O}^*$ on $\mathbb{C}^2$, where $\mathcal{O} = (\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]\tau)\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]^*$ is an operet. For the Frobenius norm we have a minimizing $P = v_1 v_1^* \in \mathbb{P}_1$ with $v_1 = (1/\sqrt{10+2\sqrt{5}})\left[\begin{smallmatrix} 1+\sqrt{5} \\ 2 \end{smallmatrix}\right]$ giving $\sigma_{\mathrm{F},2}(\mathcal{M}) = \sqrt{3-\sqrt{5}} \approx 0.874$. For the operator norm we have a minimizing $P = \widehat{v}_1\widehat{v}_1^* \in \mathbb{P}_1$ with $\widehat{v}_1 = (1/\sqrt{2})\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ giving $\sigma_{2,2}(\mathcal{M}) = 1$. It is also of interest to note that even if $\|\mathcal{M}\|_2 = \|\mathcal{M}v\| = \sqrt{10}$ with $v = (1/\sqrt{5})\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$, we have $\sigma_{2,2}(\mathcal{M}) < \|\mathcal{M}(I - vv^*)\|_2 = 1.6$. (In fact, $\|\mathcal{M}(I - v_1 v_1^*)\|_2 \approx 1.20$ so that even the Frobenius projector is better in the operator norm.) Thus, the classical existence proof for the SVD of a matrix (see, e.g., [6]) cannot be applied to find the approximation numbers with respect to the operator norm. For the expansion (3.5), taking $V = [v_1\, v_2]$ gives

$$\mathcal{M} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\left(\frac{1+\sqrt{5}}{10+2\sqrt{5}} + \frac{3+\sqrt{5}}{10+2\sqrt{5}}\tau\right)\begin{bmatrix} 1+\sqrt{5} \\ 2 \end{bmatrix}^* + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\left(\frac{-2}{10+2\sqrt{5}} + \frac{\sqrt{5}-1}{10+2\sqrt{5}}\tau\right)\begin{bmatrix} -2 \\ 1+\sqrt{5} \end{bmatrix}^*.$$

Regardless of this difference, we have the following bound for the approximation numbers in the operator and Frobenius norms.

THEOREM 3.6. *Let $\mathcal{M} \in \mathcal{M}_{p,n}$. Then*

$$\sigma_{\mathrm{F},j}(\mathcal{M}) \le \sigma_{2,j}(\mathcal{M}) \le \sqrt{2}\,\sigma_{\mathrm{F},j}(\mathcal{M}) \quad \text{for } j = 1, \ldots, k.$$

Proof. For the first inequality, choose an orthogonal projector $P = VV^*$ of rank at most j realizing $\min_{P\in\mathbb{P}_j} \|\mathcal{M}(I-P)\|_2$. For this P, assume $z \in \mathbb{C}^n$

of unit length yields the norm of $\left[\begin{smallmatrix} M(I-VV^*) \\ \overline{M}_{\#}(I-VV^*) \end{smallmatrix}\right]$. We may assume that

$$\operatorname{Re}\left(M_{\#}\overline{(I-VV^*)z}, M(I-VV^*)z\right) \geq 0 \tag{3.9}$$

after possibly multiplying z by $e^{i\theta}$ with $\theta \in [0, 2\pi)$. Then

$$\begin{aligned}
\sigma_{\mathrm{F},j+1}(\mathcal{M})^2 &\leq \|M(I-VV^*)z\|^2 + \|\overline{M}_{\#}(I-VV^*)z\|^2 \\
&\leq \|M(I-VV^*)z\|^2 + 2\operatorname{Re}\left(M_{\#}\overline{(I-VV^*)z}, M(I-VV^*)z\right) \\
&\quad + \|M_{\#}\overline{(I-VV^*)z}\|^2 \\
&= \|M(I-VV^*)z + M_{\#}\overline{(I-VV^*)z}\|^2 \\
&\leq \max_{w\in\mathbb{C}^n,\,\|w\|=1} \|M(I-VV^*)w + M_{\#}\overline{(I-VV^*)w}\|^2 = \sigma_{2,j+1}(\mathcal{M})^2.
\end{aligned}$$

For the second inequality, for any orthogonal projector $P = VV^*$ of rank at most j, take a vector $z \in \mathbb{C}^n$ of unit length realizing the norm $\|\mathcal{M}(I-P)\|_2$. Then $\|\mathcal{M}(I-P)\|_2$ squared can be bounded from above as

$$\begin{aligned}
&\|M(I-VV^*)z + M_{\#}\overline{(I-VV^*)z}\|^2 \\
&\qquad \leq (\|M(I-VV^*)z\| + \|\overline{M}_{\#}(I-VV^*)z\|)^2 \\
&\qquad \leq 2(\|M(I-VV^*)z\|^2 + \|\overline{M}_{\#}(I-VV^*)z\|^2).
\end{aligned}$$

Choose P to be the orthogonal projector yielding the $(j+1)$th singular value of the matrix $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$. For this choice we obviously have

$$\min_{P\in\mathbb{P}_j} \|\mathcal{M}(I-P)\|_2 \leq \|\mathcal{M}(I-P)\|_2 \tag{3.10}$$

and

$$\begin{aligned}
&\|M(I-VV^*)z\|^2 + \|\overline{M}_{\#}(I-VV^*)z\|^2 \\
&\quad \leq \max_{w\in\mathbb{C}^n,\,\|w\|=1} (\|M(I-VV^*)w\|^2 + \|\overline{M}_{\#}(I-VV^*)w\|^2) = \sigma_{\mathrm{F},j+1}(\mathcal{M})^2,
\end{aligned}$$

proving the claim. ∎

Both of these bounds are sharp. For the upper bound, see Example 7. Moreover, note that with (3.10) we obtain a tighter upper bound once we compute $\|\mathcal{M}(I-P)\|_2$ for P corresponding to the $(j+1)$th singular value of $\left[\begin{smallmatrix} M \\ \overline{M}_{\#} \end{smallmatrix}\right]$. Numerical experiments show that the improvement can be significant.

For the lower bound, it is the inner product term (3.9) that causes the possible increase. When it disappears, the nearness problems (3.1) and (3.3) are solved simultaneously.

COROLLARY 3.7. *Assume $M^*M_{\#}$ is skew-symmetric. Then*

$$\sigma_{\mathrm{F},j}(\mathcal{M}) = \sigma_{2,j}(\mathcal{M}) \quad \text{for } j = 1, \ldots, k.$$

Proof. The assumption is necessary and sufficient for $(Mv, M_{\#}\overline{v}) = 0$ to hold for any $v \in \mathbb{C}^n$. Consequently, the operator norm and the Ky Fan 1-norm with respect to the Frobenius norm coincide. ∎

See also Theorem 3.10 below.

For another bound on the operator norm, denote by $S(M) = \frac{1}{2}(M+M^T)$ the symmetric part of a square matrix $M \in \mathbb{C}^{n\times n}$. Recall that $(x, M\overline{z}) = (z, S(M)\overline{z})$ for any $z \in \mathbb{C}^n$ [10].

THEOREM 3.8. *Assume* $\mathcal{M} \in \mathcal{M}_{p,n}$ *and let* $F = \begin{bmatrix} M^*M \\ 2\,S(M^T\overline{M}_{\#}) \\ {M_{\#}}^T\overline{M}_{\#} \end{bmatrix}$. *Then*

$$\sigma_{2,j+1}(\mathcal{M}) \le 3^{1/4}\sigma_{j+1}(F)^{1/2}.$$

Proof. Take any orthogonal projector $P \in \mathbb{P}_j$ and a unit vector $z \in \mathbb{C}^n$. Compute

$$\begin{aligned}\|\mathcal{M}(I-P)z\|^2 &= (\mathcal{M}(I-P)z, \mathcal{M}(I-P)z) = (z, (I-P)M^*M(I-P)z)\\ &\quad +2\,\mathrm{Re}(z, (I-P)S(M^*M_{\#})(I-\overline{P})\overline{z}) + (\overline{z}, (I-\overline{P}){M_{\#}}^*M_{\#}(I-\overline{P})\overline{z}).\end{aligned}$$

The three terms on the right hand side can be bounded as follows:

$$\begin{aligned}(z, (I-P)M^*M(I-P)z) &\le \|M^*M(I-P)z\|,\\ |(z, (I-P)S(M^*M_{\#})(I-\overline{P})\overline{z})| &\le \|S(M^T\overline{M}_{\#})(I-P)z\|\\ (\overline{z}, (I-\overline{P}){M_{\#}}^*M_{\#}(I-\overline{P})\overline{z}) &\le \|{M_{\#}}^T\overline{M}_{\#}(I-P)z\|.\end{aligned}$$

Let then z be chosen such that $\|\mathcal{M}(I-P)\|_2 = \|\mathcal{M}(I-P)z\|$. Choosing $P \in \mathbb{P}_j$ so that

$$\|F(I-P)\|_2 = \sigma_{j+1}(F) \tag{3.11}$$

we hence obtain

$$\begin{aligned}\min_{\widehat{P}\in\mathbb{P}_j} \|\mathcal{M}(I-\widehat{P})\|_2^2 &\le \|\mathcal{M}(I-P)\|_2^2 = \|\mathcal{M}(I-P)z\|^2\\ &\le \|M^*M(I-P)z\| + \|2S(M^T\overline{M}_{\#})(I-P)z\| + \|{M_{\#}}^T\overline{M}_{\#}(I-P)z\|\\ &\le \sqrt{3}\max_{\|w\|_2=1}(\|M^*M(I-P)w\|^2 + \|2S(M^T\overline{M}_{\#})(I-P)w\|^2\\ &\quad + \|{M_{\#}}^T\overline{M}_{\#}(I-P)w\|^2)^{1/2}\end{aligned}$$

which equals $\sqrt{3}\,\sigma_{j+1}(F)$. Taking the square root completes the proof. ∎

Again this yields, with the projector P satisfying (3.11), a way to approximate $\mathcal{M}$, since

$$\|\mathcal{M}(I-P)\|_2 \le 3^{1/4}\sigma_{k+1}(F)^{1/2}.$$

Clearly we have

$$\sigma_{2,j}(\mathcal{M}) = \min_{w_1,\dots,w_{j-1}\in\mathbb{C}^n} \max_{z\neq 0,\, z\perp w_1,\dots,w_{j-1}} \frac{\|\mathcal{M}z\|}{\|z\|}$$

for any $\mathcal{M} \in \mathcal{M}_{p,n}$. Intriguingly, those $\mathcal{M} = M + M_{\#}\tau$ for which $M^*M_{\#}$ is skew-symmetric give rise to a family of real linear operators satisfying also the classical maxmin characterization, i.e.,

$$\max_{w_1,\dots,w_{n-j}\in\mathbb{C}^n} \min_{z\neq 0,\, z\perp w_1,\dots,w_{n-j}} \frac{\|\mathcal{M}z\|}{\|z\|} = \sigma_{2,j}(\mathcal{M}).$$

Since $M^*M_{\#}$ is skew-symmetric if and only if $\mathcal{M}^*\mathcal{M}$ is $\mathbb{C}$-linear, we conclude that such a skew-symmetry property is preserved in forming $\mathcal{W}^*\mathcal{M}K$, where $\mathcal{W}$ is an isometry and K a matrix of appropriate size. In this connection, the following characterization is of use.

PROPOSITION 3.9. *Let $p = n$. Then $\mathcal{M}^*\mathcal{M}$ is $\mathbb{C}$-linear if and only if $\mathcal{M} = \mathcal{W}DW^*$ for an isometry $\mathcal{W}$, a diagonal matrix D and a unitary matrix W.*

Proof. Sufficiency is clear, so let us prove the necessity. If $\mathcal{M}^*\mathcal{M}$ is $\mathbb{C}$-linear, then the SVD of $\mathcal{M}^*\mathcal{M} = W\Sigma W^*$ yields $\widehat{\mathcal{M}} = \mathcal{M}W$ so that $\widehat{\mathcal{M}}^*\widehat{\mathcal{M}}$ is a diagonal matrix with nonnegative entries. Considering the real forms, by using Lemma 3.11 below, we infer that $\widehat{\mathcal{M}} = \mathcal{W}\mathcal{D}$ for an isometry $\mathcal{W}$ and a real linear diagonal operator $\mathcal{D} = D + D_{\#}\tau$. Since $\mathcal{M}^*\mathcal{M}$ is $\mathbb{C}$-linear, either of the (j,j)-entries of D and $D_{\#}$ is necessarily zero for each j. Therefore $\mathcal{D} = \mathcal{U}\widehat{D}$ where D is a diagonal matrix and $\mathcal{U} = U + U_{\#}\tau$ is an isometry with diagonal $U, U_{\#} \in \mathbb{C}^{n\times n}$. ∎

Assume $p = n$. The problem of diagonalizing a bilinear form under orthogonal substitutions resulted in the earliest versions of the singular value decomposition; see [5, Chapter 3]. For an analogy, it is natural to look for a diagonal structure in the expansion (3.5). To this end, assume $W_1, W_2 \in \mathbb{C}^{n\times n}$ are unitary and suppose that

$$\mathcal{M} = W_1(D + D_{\#}\tau)W_2^* \tag{3.12}$$

with diagonal matrices $D = \operatorname{diag}(d_1,\dots,d_n)$ and $D_{\#} = \operatorname{diag}(d_1^{\#},\dots,d_n^{\#})$. Then the approximation numbers of $\mathcal{M}$ with respect to the Frobenius norm are $(|d_j|^2 + |d_j^{\#}|^2)^{1/2}$, after arranging these numbers in nonincreasing order. In this case solving the nearness problem (3.1) is also straightforward: to find V, choose those columns of W_2 that correspond to the k largest values of $|d_j| + |d_j^{\#}|$. For an illustration, consider the following case.

EXAMPLE 7. In [12] there appear infinite-dimensional operators to which correspond matrices $M = I$ and $M_{\#}$ with ${M_{\#}}^T = M_{\#}$. Clearly, the real linear operator $\mathcal{M} = I + M_{\#}\tau$ is unitarily diagonalizable, i.e., (3.12) holds

with $W_1 = W_2$. We have $\sigma_{\mathrm{F},j}(\mathcal{M}) = (1+\sigma_j(M_\#)^2)^{1/2}$, where the $\sigma_j(M_\#)$ are the singular values of the matrix $M_\#$. For the approximation numbers with respect to the operator norm we get $\sigma_{2,j}(\mathcal{M}) = 1+\sigma_j(M_\#)$. In particular, if $\sigma_j(M_\#) = 1$ for $j = 1, \ldots, n$, then we have a real linear operator illustrating that the upper bound of Theorem 3.6 is sharp.

If $\mathcal{M}$ can be decomposed according to (3.12), then its approximation numbers with respect to the operator and Frobenius norms coincide with those of $\mathcal{M}^*$.

To recognize this structure we can use the following conditions.

THEOREM 3.10. *Assume $p = n$. Then* (3.12) *holds for $\mathcal{M} = M + M_\#\tau$ if and only if $MM_\#^T$ and $M^*M_\#$ are symmetric.*

Proof. Since necessity is clear, let us prove sufficiency. Indeed, if $MM_\#^T$ and $M^*M_\#$ are symmetric, then so are $NN_\#^T$ and $N^*N_\#$ with $N = UMV^*$ and $N_\# = UM_\#V^T$ for any unitary matrices U and V. If $M = U(\Sigma_1 \oplus 0)V^*$ is the SVD of M, where Σ_1 contains the nonzero singular values of M, consider $\mathcal{N} = U^*\mathcal{M}V$. Since $NN_\#^T$ is symmetric we have $N_\# = \left[\begin{smallmatrix} B_1 & 0 \\ 0 & B_2 \end{smallmatrix}\right]$ where $B_2 \in \mathbb{C}^{k\times k}$ can be chosen freely corresponding to the k zero singular values of M. By using the SVD $B_2 = U_2\Sigma_2V_2^*$ of the B_2 block we obtain $\left[\begin{smallmatrix} \Sigma & 0 \\ 0 & 0 \end{smallmatrix}\right] + \left[\begin{smallmatrix} B_1 & 0 \\ 0 & \Sigma_2 \end{smallmatrix}\right]\tau$. Proceeding similarly with the B_1 block to reduce the problem further, we obtain two invertible blocks in the upper-left corners of the transformed operator. Hence we may assume that M and $M_\#$ are invertible.

For M and $M_\#$ invertible, let $MM_\#^T = U\widehat{D}U^T$ be the Tagaki decomposition of $MM_\#^T$ with $\widehat{D}$ having strictly positive diagonal entries. Let us consider $\mathcal{N} = U^*\mathcal{M}\overline{U}$ instead of $\mathcal{M}$. Since N is invertible, $N_\# = \widehat{D}N^{-T}$. As $N^*N_\#$ is symmetric, we obtain $\widehat{D}\overline{N}N^T = NN^*\widehat{D} = \overline{\overline{N}N^T}\widehat{D}$. Since the diagonal entries of $\widehat{D}$ are strictly positive, this forces NN^* to be a real (obviously Hermitian) block diagonal matrix, with blocks of size corresponding to the equaling entries of $\widehat{D}$. Take a real block unitary matrix V that diagonalizes NN^* and consider $\mathcal{L} = V^*\mathcal{N}\overline{V}$. Then LL^* is diagonal and hence by Lemma 3.11 below, $L = D_NW$ with a diagonal matrix D_N and a unitary matrix W. As $\widehat{D}$ and V commute, $N_\# = \widehat{D}N^{-T}$ gives $L_\# = \widehat{D}D_N^{-T}\overline{W}$. Hence, $\mathcal{L} = \mathcal{D}W$. ∎

LEMMA 3.11. *For $L \in \mathbb{C}^{n\times n}$ assume LL^* is diagonal. Then $L = DW$ with a diagonal matrix D with nonnegative entries and a unitary matrix W.*

Proof. Denote LL^* by D. Let $U\Sigma V^*$ be an SVD of L and apply, if needed, a permutation P such that $L = (UP^*)D^{1/2}(VP)^*$. Then $LL^* = UP^*D(UP^*)^* = D$. Hence UP^* commutes with D and therefore with $D^{1/2}$ as well, and we have $L = D^{1/2}UP^*(VP)^*$. ∎

If $\mathcal{M}$ satisfies the assumptions of Theorem 3.10, then obviously so do $\mathcal{M}^*$ and $\tau\mathcal{M}$ as well as $\overline{\mathcal{M}} = \overline{M} + \overline{M}_{\#}\tau$. In view of this, consider the following example.

EXAMPLE 8. Assume $p = n$ and let $\mathcal{M} = \kappa I + M_{\#}\tau$ with $\kappa \in \mathbb{C}$. If $\kappa = 0$, there are no restrictions on $M_{\#}$ while for $\kappa \neq 0$ we must have a symmetric $M_{\#}$ for the assumptions of Theorem 3.10 to hold. (Note that the skew-symmetric case is understood by Corollary 3.7.)

Theorem 3.10 extends the singular value decomposition of a matrix in the sense that a real linear operator $\mathcal{M}$ with $M_{\#} = 0$ (resp. M) satisfies the assumptions and the assertion simply refers to the SVD of the matrix M (resp. $M_{\#}$). With circulant and Hankel matrices we can give interesting examples. To this end, let $P \in \mathbb{C}^{n\times n}$ denote the "backward identity" [6], i.e., the permutation matrix with ones on the diagonal joining the left lower corner with the right upper corner.

DEFINITION 3.12. $M \in \mathbb{C}^{n\times n}$ is a *circulant-Hankel matrix* if $M = PC$ for a circulant matrix $C \in \mathbb{C}^{n\times n}$.

A circulant-Hankel matrix has cyclically appearing antidiagonals.

EXAMPLE 9. If M is a circulant and $M_{\#}$ a circulant-Hankel matrix, then $\mathcal{M}$ is unitarily diagonalizable with $U = (1/\sqrt{n})F_n$, where F_n is the Fourier matrix [10]. Conversely, let M be a circulant-Hankel and $M_{\#}$ a circulant matrix. Then the corresponding $\mathcal{M}$ is, generically, not unitarily diagonalizable but does satisfy the assumptions of Theorem 3.10.

EXAMPLE 10. If we have an isometry satisfying the assumptions of Theorem 3.10, then the corresponding diagonal real linear operator $\mathcal{D} = D + D_{\#}\tau$ is an isometry. Hence, necessarily either $|d_{jj}| = 1$ and $|d^{\#}_{jj}| = 0$, or $|d_{jj}| = 0$ and $|d^{\#}_{jj}| = 1$ for the diagonal elements of D and $D_{\#}$ with $j = 1, \ldots, n$.

Observe though that for a case where the nearness problems (3.1) and (3.3) can be solved simultaneosly, the structure (3.12) is not the most general. According to Proposition 3.4, (3.1) and (3.3) can be solved if we manage to obtain a diagonal real linear operator $\mathcal{W}_1^*\mathcal{M}W_2$ with an isometry $\mathcal{W}_1$ and a unitary matrix W_2. (Recall that we can always have an upper triangular form $\mathcal{W}_1^*\mathcal{M}$ with an isometry $\mathcal{W}_1$; see [2, Section 2.2].) Clearly, $\mathcal{M}^*\mathcal{M}$ is then necessarily unitarily diagonalizable.

Let $H = \frac{1}{2}(M + M^*)$ and $K = \frac{1}{2i}(M - M^*)$ denote the Hermitian and skew-Hermitian parts of $M \in \mathbb{C}^{n\times n}$. With this notation, we have the following result.

THEOREM 3.13. *$\mathcal{M} = M + M_{\#}\tau$ is unitarily diagonalizable if and only if M is normal and $M_{\#}$, $HM_{\#}$ and $KM_{\#}$ are symmetric.*

Proof. Use [4, Corollary 5.3] together with the fact that due to normality the matrices H and K commute to deduce that there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^*\mathcal{M}U$ has diagonal linear and antilinear parts. ∎

We conclude this section with two remarks. First, let us illustrate how to form the proper orthogonal decomposition in our framework.

Example 11. The Karhunen–Loève expansion has a wide range of applications; see, e.g., [13] and the references therein. Assume that we have functions (for instance, a collection of signals or trajectories) $a_1(t), a_2(t), \ldots$ $\ldots, a_{2n}(t) \in \mathbb{R}^{2p}$, for $t \in [0,T]$, forming a parameter dependent matrix $A(t) = [a_1(t) \cdots a_{2n}(t)]$. In digital image processing $A(t)$ could represent a continuous sequence of digitalized images. In control theory $A(t)$ could be the transfer function. We take its complex formulation $\mathcal{M}(t)$ for which we want to find in a sense an optimal orthonormal set of $\mathbb{C}^n$ to approximately represent $\mathcal{M}$. To this end, we compute an orthogonal projector $P \in \mathbb{P}_j$ such that

$$E(\|\mathcal{M}(t)(I-P)\|^2) = \int_0^T \|\mathcal{M}(t)(I-P)\|^2\, dt$$

is minimal for a unitarily invariant norm $\|\cdot\|$. If we take the Frobenius norm, then it is straightforward that the correlation matrix (which is obviously Hermitian)

$$R = \int_0^T [M(t)^*\ M_\#(t)^T] \begin{bmatrix} M(t) \\ \overline{M_\#(t)} \end{bmatrix} dt$$

yields the desired P via its SVD.

As a final remark, there are applications where it is of interest to preserve a structure. For instance, for $p = n$, consider $M + M_\#\tau$ with $M^* = -M$ and ${M_\#}^T = M_\#$. To preserve this, one option is to consider the modified nearness problem

$$\min_{P\in\mathbb{P}_j} \|\mathcal{M} - P\mathcal{M}P\|$$

for a unitarily invariant norm. Then, for $P = VV^*$ solving this, the matrix

$$V(V^*MV + V^*M_\#\overline{V}\tau)V^*$$

approximates $\mathcal{M}$.

References

[1] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrica 1 (1936), 31–37.

[2] T. Eirola, M. Huhtanen and J. von Pfaler, *Solution methods for $\mathbb{R}$-linear problems in $\mathbb{C}^n$*, SIAM J. Matrix Anal. Appl. 25 (2004), 804–828.

[3] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed., John Hopkins Univ. Press, Baltimore and London, 1996.

[4] Y. P. Hong and R. A. Horn, *On simultaneous reduction of families of matrices to triangular or diagonal form by unitary congruences*, Linear Multilinear Algebra 17 (1985), 271–288.

[5] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1987.

[6] —, —, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1991.

[7] L. Hubert, J. Meulman and W. Heiser, *Two purposes for matrix factorization: a historical appraisal*, SIAM Rev. 42 (2000), 68–82.

[8] M. Huhtanen and O. Nevanlinna, *Real linear matrix analysis*, in: Banach Center Publ. 75, to appear.

[9] M. Huhtanen and O. Nevanlinna, *The real linear resolvent and cosolvent operators*, J. Operator Theory, to appear.

[10] M. Huhtanen and J. von Pfaler, *The real linear eigenvalue problem in* $\mathbb{C}^n$, Linear Algebra Appl. 394 (2005), 169–199.

[11] A. Pietsch, *Eigenvalues and s-numbers*, Cambridge Stud. Adv. Math. 13, Cambridge Univ. Press, Cambridge, 1987.

[12] M. Putinar and H. S. Shapiro, *The Friedrichs operator of a planar domain*, in: Complex Analysis, Operators, and Related Topics, Oper. Theory Adv. Appl. 113, Birkhäuser, Basel, 2000, 303–330.

[13] M. Rathinam and L. Petzold, *A new look at proper orthogonal decomposition*, SIAM J. Numer. Anal. 41 (2003), 1893–1925.

[14] E. Schmidt, *Zur Theorie der linearen und nichtlinearen Integralgleichungen. I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener*, Math. Ann. 63 (1907), 433–476.

[15] G. W. Stewart, *On the early history of the singular value decomposition*, SIAM Rev. 35 (1993), 551–566.

Institute of Mathematics
Helsinki University of Technology
Box 1100, FIN-02015, Finland
E-mail: Marko.Huhtanen@hut.fi
Olavi.Nevanlinna@hut.fi

Received November 9, 2005
Revised version January 12, 2007 (5801)