ELA

MINIMIZATION PROBLEMS FOR
CERTAIN STRUCTURED MATRICES[*]

ZHONGYUN LIU[†], RUI RALHA[‡], YULIN ZHANG[‡], AND CARLA FERREIRA[‡]

Abstract. For given $Z, B \in \mathbb{C}^{n \times k}$, the problem of finding $A \in \mathbb{C}^{n \times n}$, in some prescribed class $\mathcal{W}$, that minimizes $\|AZ - B\|$ (Frobenius norm) has been considered by different authors for distinct classes $\mathcal{W}$. Here, this minimization problem is studied for two other classes, which include the symmetric Hamiltonian, symmetric skew-Hamiltonian, real orthogonal symplectic and unitary conjugate symplectic matrices. The problem of minimizing $\|A - \tilde{A}\|$, where $\tilde{A}$ is given and A is a solution of the previous problem, is also considered (as has been done by others, for different classes $\mathcal{W}$). The key idea of this contribution is the reduction of each one of the above minimization problems to two independent subproblems in orthogonal subspaces of $\mathbb{C}^{n \times n}$. This is possible due to the special structures under consideration. MATLAB codes are developed, and numerical results of some tests are presented.

Key words. Least-squares approximation, Centralizer of $\mathcal{J}$, Anticentralizer of $\mathcal{J}$, Moore-Penrose inverse.

AMS subject classifications. 65F20, 65F30.

1. Introduction. Structured matrices appear in many applications and, in general, exploiting the structure may enable the development of more accurate algorithms. Such algorithms may also be more economic both in computation and storage and, in the application where such problems do occur, the structured solutions may have a more precise physical meaning. To mention just a few works related to this topic, the notion of strongly stable algorithms was introduced in [1] and used in the context of solving Toeplitz linear systems, eigensolvers for matrices of several special structures have been given in [2], a chart of backward errors for structured eigenvalue problems was presented in [10].

ELA

Skew-Hamiltonian matrices and Hamiltonian matrices are two important classes of structured matrices. In this paper, we are interested in certain classes that include symmetric skew-Hamiltonian matrices and symmetric Hamiltonian matrices. Symmetric skew-Hamiltonian matrices arise in quantum mechanical problems with time reversal symmetry. Symmetric Hamiltonian matrices appear in response theory, more exactly in the study of closed shell Hartree-Fock wave functions, and also in solving algebraic Riccati equations in the context of continuous time linear quadratic optimal control problems. See [10] for references dealing with such applications.

We will be concerned with the following problems in certain structured classes $\mathcal{W}$ of matrices of size $n \times n$ (to be defined in the next section):

PROBLEM I. Given $Z, B \in \mathbb{C}^{n \times k}$, find

$$\sigma = \min_{A \in \mathcal{W}} \|AZ - B\|$$

and characterize the class

$$\mathcal{S} = \{A \in \mathcal{W} : \|AZ - B\| = \sigma\}.$$

In particular, find necessary and sufficient conditions for the existence of $A \in \mathcal{W}$ such that $AZ = B$ and give a general form for A.

PROBLEM II. Given $\widetilde{A} \in \mathbb{C}^{n \times n}$, find

$$\widetilde{\sigma} = \min_{A \in \mathcal{S}} \|\widetilde{A} - A\|$$

and find $A \in \mathcal{S}$ such that $\|\widetilde{A} - A\| = \widetilde{\sigma}$.

The above problems have been studied by many authors in various contexts and different classes of structured matrices were considered: centrosymmetric matrices [13], centrohermitian matrices [7], hermitian R-symmetric and hermitian R-skew symmetric matrices [11], matrices which satisfy $RAS = A$ and $RAS = -A$ for R and S nontrivial involutions, i.e., $R^2 = S^2 = I$ and $R, S \neq \pm I$ (Trench referred to these matrices as (R, S) symmetric and (R, S)-skew symmetric [12]). Some applications that lead to the above minimization problems are also described in these references. In [6], Problem II has also been considered for A in the set $\mathcal{W}$ of generalized K-centrohermitian matrices with prescribed spectra.

In Section 2, we present the classes of structured matrices to be studied and prove a fundamental theorem. Some more theoretical results are given in Section 3. The derivation of the explicit solutions for our minimization problems is presented in Section 4. In Section 5, we exhibit our structure-exploiting algorithms for computing those solutions. In Section 6, we present results of our numerical experiments and finish with some conclusions in Section 7.

ELA

2. The centralizer and anticentralizer of $\mathcal{J}$. Let $\mathcal{J} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$, where I is the identity matrix of order m. In [2], the sets

$$(2.1) \qquad \mathscr{C} = \left\{ A \in \mathbb{C}^{2m \times 2m} \mid A\mathcal{J} = \mathcal{J}A \right\} = \left\{ \begin{bmatrix} D & -E \\ E & D \end{bmatrix} \mid D, E \in \mathbb{C}^{m \times m} \right\}$$

and

$$(2.2) \qquad \mathscr{C}^{(a)} = \left\{ A \in \mathbb{C}^{2m \times 2m} \mid A\mathcal{J} = -\mathcal{J}A \right\} = \left\{ \begin{bmatrix} G & F \\ F & -G \end{bmatrix} \mid F, G \in \mathbb{C}^{m \times m} \right\}$$

are called the *centralizer* of $\mathcal{J}$ and the *anticentralizer* of $\mathcal{J}$, respectively. Also it was observed that orthogonal sympletic matrices and unitary conjugate sympletic matrices belong to $\mathscr{C}$. It can be seen that symmetric skew-Hamiltonian matrices also belong to $\mathscr{C}$ and symmetric Hamiltonian matrices belong to $\mathscr{C}^{(a)}$.

The next theorem shows the relation between spaces $\mathbb{C}^{2m \times 2m}$, $\mathscr{C}$ and $\mathscr{C}^{(a)}$. We use the Frobenius inner product, $\langle X, Y \rangle = trace(Y^H X)$ for matrices X and Y. An orthogonal direct sum will be denoted with the symbol $\oplus^\perp$.

LEMMA 2.1. *We have*

$$\mathbb{C}^{2m \times 2m} = \mathscr{C} \oplus^\perp \mathscr{C}^{(a)}.$$

Proof. First, we show that every matrix in $\mathbb{C}^{2m \times 2m}$ can be written as a sum of two matrices, one in $\mathscr{C}$ and the other one in $\mathscr{C}^{(a)}$. Given $A \in \mathbb{C}^{2m \times 2m}$, let

$$(2.3) \qquad A_c = \frac{A - \mathcal{J}A\mathcal{J}}{2} \qquad \text{and} \qquad A_s = \frac{A + \mathcal{J}A\mathcal{J}}{2}.$$

We have $A = A_c + A_s$ and

$$\mathcal{J}A_c\mathcal{J} = \frac{\mathcal{J}A\mathcal{J} - \mathcal{J}^2 A\mathcal{J}^2}{2} = \frac{\mathcal{J}A\mathcal{J} - A}{2} = -A_c,$$

$$\mathcal{J}A_s\mathcal{J} = \frac{\mathcal{J}A\mathcal{J} + \mathcal{J}^2 A\mathcal{J}^2}{2} = \frac{\mathcal{J}A\mathcal{J} + A}{2} = A_s,$$

that is, $A_c \in \mathscr{C}$ and $A_s \in \mathscr{C}^{(a)}$. This decomposition is unique since $\mathscr{C} \bigcap \mathscr{C}^{(a)} = \{O\}$. In fact, if $A \in \mathscr{C} \cap \mathscr{C}^{(a)}$, then $\mathcal{J}A\mathcal{J} = -A$ and $\mathcal{J}A\mathcal{J} = A$, which implies that $A = O$. So, we have a direct sum.

Finally, let us see that $\mathscr{C} \perp \mathscr{C}^{(a)}$. If $X \in \mathscr{C}$ and $Y \in \mathscr{C}^{(a)}$, we have $\mathcal{J}X\mathcal{J} = -X$ and $\mathcal{J}Y\mathcal{J} = Y$. Since $\mathcal{J}^H\mathcal{J} = \mathcal{J}\mathcal{J}^H = I$, we have

$$\begin{aligned} \langle X, Y \rangle &= \langle -\mathcal{J}X\mathcal{J}, \mathcal{J}Y\mathcal{J} \rangle \\ &= - trace[(\mathcal{J}Y\mathcal{J})^H \mathcal{J}X\mathcal{J}] = - trace(\mathcal{J}^H Y^H X\mathcal{J}) \\ &= - trace(Y^H X) = -\langle X, Y \rangle, \end{aligned}$$

which means that $\langle X, Y \rangle = 0$ and, thus, $\mathscr{C} \perp \mathscr{C}^{(a)}$. $\square$

The minimization problems that we are tackling here have not been considered in [2]. As compared to the work in [12], for (R, S) symmetric and (R, S)-skew symmetric matrices, we emphasize that matrices in $\mathscr{C}$ and $\mathscr{C}^{(a)}$ do not fall into these classes as $\mathcal{J}$ is not an involution ($\mathcal{J}^2 \neq I$). The key idea of our contribution is the reduction of each one of the minimization problems under consideration to two independent subproblems in orthogonal subspaces of $\mathbb{C}^{2m \times 2m}$.

3. Preliminary results. The following result has been given in [2, Lemma 2.5].

LEMMA 3.1. *Let* $P \in \mathbb{C}^{2m \times 2m}$ *be the unitary matrix given by*

$$(3.1) \qquad P = \tfrac{\sqrt{2}}{2} \begin{bmatrix} I & I \\ iI & -iI \end{bmatrix},$$

where I *is the identity matrix of order* m. *We have:*

   (i) *if* $A \in \mathscr{C}$ *(as defined in (2.1)), then*

$$(3.2) \qquad P^H A P = \begin{bmatrix} M_1 & \\ & N_1 \end{bmatrix},$$

      where $M_1 = D - iE$ *and* $N_1 = D + iE$;

   (ii) *if* $A \in \mathscr{C}^{(a)}$ *(as defined in (2.2)), then*

$$(3.3) \qquad P^H A P = \begin{bmatrix} & M_2 \\ N_2 & \end{bmatrix},$$

      where $M_2 = G - iF$ *and* $N_2 = G + iF$.

Using Lemma 3.1 and Lemma 2.1, for any matrix $\tilde{A} \in \mathbb{C}^{2m \times 2m}$, we have

$$P^H \tilde{A} P = P^H (A_c + A_s) P = \begin{bmatrix} M_1 & \\ & N_1 \end{bmatrix} + \begin{bmatrix} & M_2 \\ N_2 & \end{bmatrix},$$

where A_c and A_s are as defined in (2.3). For the sake of convenience, we will use M to represent either M_1 or M_2 if there is no danger of ambiguity. The same applies to N, N_1 and N_2.

To tackle Problem I, we now define two new subspaces

$$(3.4) \qquad \mathcal{C}_{\mathcal{J}}^{2m \times k} = \{X \in \mathbb{C}^{2m \times k} \colon i\mathcal{J}X = X\} = \left\{ \begin{bmatrix} \widehat{X} \\ -i\widehat{X} \end{bmatrix} \mid \widehat{X} \in \mathbb{C}^{m \times k} \right\}$$

and

$$(3.5) \qquad \mathcal{S}_{\mathcal{J}}^{2m \times k} = \{Y \in \mathbb{C}^{2m \times k} \colon i\mathcal{J}Y = -Y\} = \left\{ \begin{bmatrix} \widehat{Y} \\ i\widehat{Y} \end{bmatrix} \mid \widehat{Y} \in \mathbb{C}^{m \times k} \right\}.$$

ELA

Any matrix $B \in \mathbb{C}^{2m \times k}$ can be expressed as the sum of one matrix in $\mathcal{C}_{\mathcal{J}}^{2m \times k}$ and another in $\mathcal{S}_{\mathcal{J}}^{2m \times k}$.

LEMMA 3.2. *We have*

$$\mathbb{C}^{2m \times k} = \mathcal{C}_{\mathcal{J}}^{2m \times k} \oplus^{\perp} \mathcal{S}_{\mathcal{J}}^{2m \times k}.$$

In fact, any matrix $B \in \mathbb{C}^{2m \times k}$ can be written as

$$B = \frac{B + i\mathcal{J}B}{2} + \frac{B - i\mathcal{J}B}{2},$$

where the first matrix in the right hand side belongs to $\mathcal{C}_{\mathcal{J}}^{2m \times k}$ and the second to $\mathcal{S}_{\mathcal{J}}^{2m \times k}$. The rest of the proof is similar to the one given for the previous result. This theorem plays a major role in Problem I. Furthermore, we have the following lemma.

LEMMA 3.3. *Let P be the unitary matrix defined in (3.1). If $X \in \mathcal{C}_{\mathcal{J}}^{2m \times k}$ and $Y \in \mathcal{S}_{\mathcal{J}}^{2m \times k}$, then*

(3.6) $$P^H X = \sqrt{2} \begin{bmatrix} O \\ \widehat{X} \end{bmatrix} \quad and \quad P^H Y = \sqrt{2} \begin{bmatrix} \widehat{Y} \\ O \end{bmatrix}.$$

The proof is straightforward if we perform the matrix products. It is also straightforward to prove the following result concerning the invariance (and swapping) of the subspaces $\mathcal{C}_{\mathcal{J}}^{2m \times k}$ and $\mathcal{S}_{\mathcal{J}}^{2m \times k}$ under $A \in \mathscr{C}$ (and $A \in \mathscr{C}^{(a)}$).

LEMMA 3.4. *Let $X \in \mathcal{C}_{\mathcal{J}}^{2m \times k}$ and $Y \in \mathcal{S}_{\mathcal{J}}^{2m \times k}$.*

1. *For $A \in \mathscr{C}$, we have $AX \in \mathcal{C}_{\mathcal{J}}^{2m \times k}$ and $AY \in \mathcal{S}_{\mathcal{J}}^{2m \times k}$.*
2. *For $A \in \mathscr{C}^{(a)}$, we have $AX \in \mathcal{S}_{\mathcal{J}}^{2m \times k}$ and $AY \in \mathcal{C}_{\mathcal{J}}^{2m \times k}$.*

4. The minimization problems. In this section, we recall two results concerning the minimization problem for the general case, that is, given matrices $Z \in \mathbb{C}^{q \times k}$ and $B \in \mathbb{C}^{p \times k}$, we seek $\min_{A \in \mathbb{C}^{p \times q}} \|AZ - B\|$, when there is no constraint on A. We also give a different interpretation of the known necessary and sufficient condition for such minimum to be zero. Then we turn to the problem with $A \in \mathscr{C}$ and $A \in \mathscr{C}^{(a)}$, exploiting the structure properties of these matrices.

4.1. General case. For a matrix $Z \in \mathbb{C}^{q \times k}$, we will use the standard notation $Z^{\dagger}$ for the Moore-Penrose inverse. We recall that throughout the paper we use the Frobenius norm, the norm induced by the Frobenius inner product.

LEMMA 4.1. [12, Lemma 2] *If $Z \in \mathbb{C}^{q \times k}$ and $B \in \mathbb{C}^{p \times k}$, then*

(4.1) $$\min_{A \in \mathbb{C}^{p \times q}} \|AZ - B\| = \left\| B \left(I - Z^{\dagger} Z \right) \right\|,$$

ELA

and this minimum is attained if and only if

$$(4.2) \qquad A = BZ^\dagger + K\left(I - ZZ^\dagger\right)$$

with $K \in \mathbb{C}^{p \times q}$ arbitrary. Moreover, $A_0 = BZ^\dagger$ is the unique matrix of this form with minimum norm.

In particular, we have $AZ = B$ if and only if $B\left(I - Z^\dagger Z\right) = O$ ([12, Lemma 3]). For $q \geq k$, if Z is full rank[1] then $Z^\dagger Z = I$ and the previous condition holds. More generally, we have the following result.

THEOREM 4.2. *Given Z in $\mathbb{C}^{q \times k}$ (with $q \geq k$) and B in $\mathbb{C}^{p \times k}$, we have that $B\left(I - Z^\dagger Z\right) = O$ if and only if the null space of Z is contained in the null space of B, that is, $null(Z) \subseteq null(B)$.*

Proof. To conclude that the given condition is necessary, note that from $B\left(I - Z^\dagger Z\right) = O$ we get that $Zx = 0$ implies $Bx = 0$. Conversely, assume that $null(Z) \subseteq null(B)$. Then, $row(Z)$ and $row(B)$ are the orthogonal complements of $null(Z)$ and $null(B)$, respectively, and satisfy $row(B) \subseteq row(Z)$ [5, Theorems 3.12 and 3.11]. Thus, there is a matrix M such that $B = MZ$ and, since $ZZ^\dagger Z = Z$,

$$B\left(I - Z^\dagger Z\right) = MZ - MZZ^\dagger Z = MZ - MZ = O. \quad \square$$

In the proofs of the results concerning Problem II, we will use the following.

LEMMA 4.3. [12, Lemma 4] *Suppose that $L \in \mathbb{C}^{p \times q}$ and $\Gamma \in \mathbb{C}^{q \times q}$ where $\Gamma^2 = \Gamma = \Gamma^H$. Then the unique matrix of the form $L - M\Gamma$ with minimum Frobenius norm is $L - L\Gamma = L(I - \Gamma)$.*

4.2. $A \in \mathscr{C}$ (centralizer of $\mathcal{J}$). We first consider Problem I.

Let $Z, B \in \mathbb{C}^{2m \times k}$ ($m \geq k$). Write $Z = Z_1 + Z_2$ with

$$(4.3) \qquad Z_1 = \frac{Z + i\mathcal{J}Z}{2}, \qquad Z_2 = \frac{Z - i\mathcal{J}Z}{2},$$

and $B = B_1 + B_2$ with

$$(4.4) \qquad B_1 = \frac{B + i\mathcal{J}B}{2}, \qquad B_2 = \frac{B - i\mathcal{J}B}{2},$$

where $Z_1, B_1 \in \mathcal{C}_\mathcal{J}^{2m \times k}$ and $Z_2, B_2 \in \mathcal{S}_\mathcal{J}^{2m \times k}$.

From Lemma 3.4, for $A \in \mathscr{C}$, $(AZ_1 - B_1) \in \mathcal{C}_\mathcal{J}^{2m \times k}$ and $(AZ_2 - B_2) \in \mathcal{S}_\mathcal{J}^{2m \times k}$. The orthogonality of these subspaces implies

$$\|AZ - B\|^2 = \|(AZ_1 - B_1) + (AZ_2 - B_2)\|^2 = \|AZ_1 - B_1\|^2 + \|AZ_2 - B_2\|^2.$$

---

[1]Since $q \geq k$, by *full rank* we mean *full column rank*.

$$\boxed{\textbf{ELA}}$$

According to (3.4) and (3.5),

$$(4.5) \qquad Z_1 = \begin{bmatrix} \widehat{Z}_1 \\ -i\widehat{Z}_1 \end{bmatrix}, \qquad Z_2 = \begin{bmatrix} \widehat{Z}_2 \\ i\widehat{Z}_2 \end{bmatrix}, \qquad B_1 = \begin{bmatrix} \widehat{B}_1 \\ -i\widehat{B}_1 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} \widehat{B}_2 \\ i\widehat{B}_2 \end{bmatrix},$$

where $\widehat{Z}_1, \widehat{Z}_2, \widehat{B}_1, \widehat{B}_2 \in \mathbb{C}^{m \times k}$.

For $A \in \mathscr{C}$, the invariance of the Frobenius norm under unitary multiplication, (3.2) and (3.6) imply that

$$\|AZ_1 - B_1\|^2 = \|P^H A P P^H Z_1 - P^H B_1\|^2 = 2\left\| \begin{bmatrix} M & \\ & N \end{bmatrix} \begin{bmatrix} O \\ \widehat{Z}_1 \end{bmatrix} - \begin{bmatrix} O \\ \widehat{B}_1 \end{bmatrix} \right\|^2$$
$$= 2\left\| N\widehat{Z}_1 - \widehat{B}_1 \right\|^2.$$

Similarly,

$$\|AZ_2 - B_2\|^2 = 2\left\| M\widehat{Z}_2 - \widehat{B}_2 \right\|^2.$$

Therefore, our problem is reduced to two independent subproblems of the same type in the orthogonal subspaces, one for the pair (Z_1, B_1) and the other for the pair (Z_2, B_2). Thus, with respect to Problem I, we have the following result.

THEOREM 4.4. *Let $Z, B \in \mathbb{C}^{2m \times k}$. Consider P as given in (3.1) and $\widehat{Z}_1, \widehat{Z}_2, \widehat{B}_1, \widehat{B}_2$ in (4.5). Then*

$$(4.6) \qquad \min_{A \in \mathscr{C}} \|AZ - B\| = \sqrt{2}\left(\left\| \widehat{B}_1\big(I - \widehat{Z}_1^\dagger \widehat{Z}_1\big) \right\|^2 + \left\| \widehat{B}_2\big(I - \widehat{Z}_2^\dagger \widehat{Z}_2\big) \right\|^2 \right)^{1/2},$$

and this minimum is attained if and only if

$$(4.7) \qquad A = P \begin{bmatrix} M & \\ & N \end{bmatrix} P^H = \tfrac{1}{2} \begin{bmatrix} M+N & -(M-N)i \\ (M-N)i & M+N \end{bmatrix},$$

where $M = \widehat{B}_2 \widehat{Z}_2^\dagger + K_2\big(I - \widehat{Z}_2 \widehat{Z}_2^\dagger\big)$ and $N = \widehat{B}_1 \widehat{Z}_1^\dagger + K_1\big(I - \widehat{Z}_1 \widehat{Z}_1^\dagger\big)$ with arbitrary $K_1, K_2 \in \mathbb{C}^{m \times m}$. Moreover, $A = P \begin{bmatrix} \widehat{B}_2 \widehat{Z}_2^\dagger & \\ & \widehat{B}_1 \widehat{Z}_1^\dagger \end{bmatrix} P^H$ is the unique matrix of this form with minimum norm.

The relations (4.6) and (4.7) follow from the application of Lemma 4.1 to $\|N\widehat{Z}_1 - \widehat{B}_1\|$ and $\|M\widehat{Z}_2 - \widehat{B}_2\|$.

Now we consider Problem II.

$$\boxed{\textbf{ELA}}$$

Let $\sigma_c = \min\limits_{A \in \mathscr{C}} \|AZ - B\|$ and $\mathcal{S}_c = \{A \in \mathscr{C} : \|AZ - B\| = \sigma_c\}$.

For $\widetilde{A} \in \mathbb{C}^{2m \times 2m}$, using Lemma 2.1, there exists a unique pair of matrices $\widetilde{A}_1 \in \mathscr{C}$, $\widetilde{A}_2 \in \mathscr{C}^{(a)}$ such that $\widetilde{A} = \widetilde{A}_1 + \widetilde{A}_2$ with

$$(4.8) \qquad \widetilde{A}_1 = \frac{\widetilde{A} - \mathcal{J}\widetilde{A}\mathcal{J}}{2} \quad \text{and} \quad \widetilde{A}_2 = \frac{\widetilde{A} + \mathcal{J}\widetilde{A}\mathcal{J}}{2}.$$

Using Lemma 3.1, we write

$$(4.9) \qquad P^H \widetilde{A}_1 P = \begin{bmatrix} M_1 & \\ & N_1 \end{bmatrix} \quad \text{and} \quad P^H \widetilde{A}_2 P = \begin{bmatrix} & M_2 \\ N_2 & \end{bmatrix}.$$

THEOREM 4.5. *Assume the conditions of Theorem 4.4. Let $\tilde{A} \in \mathbb{C}^{2m \times 2m}$ and M_1, N_1, M_2, N_2 given in (4.9). Then*
$$(4.10)$$
$$\min_{A \in \mathcal{S}_c} \|\widetilde{A} - A\| = \left(\|M_2\|^2 + \|N_2\|^2 + \left\| (M_1 \widehat{Z}_2 - \widehat{B}_2)\widehat{Z}_2^\dagger \right\|^2 + \left\| (N_1 \widehat{Z}_1 - \widehat{B}_1)\widehat{Z}_1^\dagger \right\|^2 \right)^{1/2},$$

and this minimum is attained if and only if

$$(4.11) \qquad A = \tfrac{1}{2} \begin{bmatrix} M_3 + N_3 & -(M_3 - N_3)i \\ (M_3 - N_3)i & M_3 + N_3 \end{bmatrix},$$

where $M_3 = \widehat{B}_2 \widehat{Z}_2^\dagger + M_1(I - \widehat{Z}_2 \widehat{Z}_2^\dagger)$ and $N_3 = \widehat{B}_1 \widehat{Z}_1^\dagger + N_1(I - \widehat{Z}_1 \widehat{Z}_1^\dagger)$.

Proof. Let $\widetilde{A} = \widetilde{A}_1 + \widetilde{A}_2$, with $\widetilde{A}_1, \widetilde{A}_2$ given in (4.8), and $A \in \mathcal{S}_c$ given by (4.7). Since the Frobenius norm is invariant under unitary transformations, we have

$$\begin{aligned} \left\| \widetilde{A} - A \right\|^2 &= \left\| (\tilde{A}_1 - A) + \tilde{A}_2 \right\|^2 \\ &= \left\| P^H(\widetilde{A}_1 - A)P + P^H \widetilde{A}_2 P \right\|^2 \\ &= \left\| \begin{bmatrix} M_1 - \widehat{B}_2 \widehat{Z}_2^\dagger - K_2(I - \widehat{Z}_2 \widehat{Z}_2^\dagger) & M_2 \\ N_2 & N_1 - \widehat{B}_1 \widehat{Z}_1^\dagger - K_1(I - \widehat{Z}_1 \widehat{Z}_1^\dagger) \end{bmatrix} \right\|^2 \\ &= \|M_2\|^2 + \|N_2\|^2 + \left\| M_1 - \widehat{B}_2 \widehat{Z}_2^\dagger - K_2(I - \widehat{Z}_2 \widehat{Z}_2^\dagger) \right\|^2 + \\ &\quad + \left\| N_1 - \widehat{B}_1 \widehat{Z}_1^\dagger - K_1(I - \widehat{Z}_1 \widehat{Z}_1^\dagger) \right\|^2. \end{aligned}$$

Observe that only $K_1, K_2 \in \mathbb{C}^{m \times m}$ are free. Thus,

$$\begin{aligned} \min_{A \in \mathcal{S}_c} \left\| \widetilde{A} - A \right\|^2 &= \|M_2\|^2 + \|N_2\|^2 + \min_{K_2 \in \mathbb{C}^{m \times m}} \left\| M_1 - \widehat{B}_2 \widehat{Z}_2^\dagger - K_2(I - \widehat{Z}_2 \widehat{Z}_2^\dagger) \right\|^2 + \\ &\quad + \min_{K_1 \in \mathbb{C}^{m \times m}} \left\| N_1 - \widehat{B}_1 \widehat{Z}_1^\dagger - K_1(I - \widehat{Z}_1 \widehat{Z}_1^\dagger) \right\|^2. \end{aligned}$$

$$\boxed{\textbf{ELA}}$$

For $X \in \mathbb{C}^{m \times k}$, the properties of the Moore-Penrose inverse imply that $(I - XX^\dagger)^2 = (I - XX^\dagger) = (I - XX^\dagger)^H$. Thus, we can apply Lemma 4.3 to each of the two minimization problems on the right side of the previous equation. The minimum values are attained if and only if

$$K_2 = M_1 - \widehat{B}_2 \widehat{Z}_2^\dagger \quad \text{and} \quad K_1 = N_1 - \widehat{B}_1 \widehat{Z}_1^\dagger$$

which give

$$K_2\big(I - \widehat{Z}_2 \widehat{Z}_2^\dagger\big) = \big(M_1 - \widehat{B}_2 \widehat{Z}_2^\dagger\big)\big(I - \widehat{Z}_2 \widehat{Z}_2^\dagger\big) = M_1\big(I - \widehat{Z}_2 \widehat{Z}_2^\dagger\big) - \widehat{B}_2 \widehat{Z}_2^\dagger + \widehat{B}_2 \widehat{Z}_2^\dagger \widehat{Z}_2 \widehat{Z}_2^\dagger$$
$$= M_1\big(I - \widehat{Z}_2 \widehat{Z}_2^\dagger\big)$$

and

$$K_1\big(I - \widehat{Z}_1 \widehat{Z}_1^\dagger\big) = \big(N_1 - \widehat{B}_1 \widehat{Z}_1^\dagger\big)\big(I - \widehat{Z}_1 \widehat{Z}_1^\dagger\big) = N_1\big(I - \widehat{Z}_1 \widehat{Z}_1^\dagger\big).$$

Hence, the minimum values are

$$\left\| M_1 - \widehat{B}_2 \widehat{Z}_2^\dagger - M_1\big(I - \widehat{Z}_2 \widehat{Z}_2^\dagger\big) \right\|^2 = \left\| \big(M_1 \widehat{Z}_2 - \widehat{B}_2\big) \widehat{Z}_2^\dagger \right\|^2$$

and

$$\left\| N_1 - \widehat{B}_1 \widehat{Z}_1^\dagger - N_1\big(I - \widehat{Z}_1 \widehat{Z}_1^\dagger\big) \right\|^2 = \left\| \big(N_1 \widehat{Z}_1 - \widehat{B}_1\big) \widehat{Z}_1^\dagger \right\|^2.$$

This shows that $\min\limits_{A \in \mathcal{S}_c} \|\widetilde{A} - A\|$ is given by (4.10) and it is attained if and only if A takes the form (4.11). $\square$

4.3. $A \in \mathscr{C}^{(a)}$ (anticentralizer of $\mathcal{J}$). For $A \in \mathscr{C}^{(a)}$ the results are entirely analogous to the ones just presented for $A \in \mathscr{C}$ and therefore we will omit the proofs. In view of the fact that for $A \in \mathscr{C}^{(a)}$, $(AZ_2 - B_1) \in \mathcal{C}_{\mathcal{J}}^{2m \times k}$ and $(AZ_1 - B_2) \in \mathcal{S}_{\mathcal{J}}^{2m \times k}$ are orthogonal, we may use arguments similar to those used in the previous subsection to get the following results.

THEOREM 4.6. *Let $Z, B \in \mathbb{C}^{2m \times k}$. Consider P as given in (3.1) and $\widehat{Z}_1, \widehat{Z}_2, \widehat{B}_1, \widehat{B}_2$ in (4.5). Then*

$$(4.12) \qquad \min_{A \in \mathscr{C}^{(a)}} \|AZ - B\| = \sqrt{2}\left(\left\| \widehat{B}_2\big(I - \widehat{Z}_1^\dagger \widehat{Z}_1\big) \right\|^2 + \left\| \widehat{B}_1\big(I - \widehat{Z}_2^\dagger \widehat{Z}_2\big) \right\|^2 \right)^{1/2}$$

and this minimum is attained if and only if

$$(4.13) \qquad A = P \begin{bmatrix} & M \\ N & \end{bmatrix} P^H = \tfrac{1}{2}\begin{bmatrix} M + N & (M - N)i \\ (M - N)i & -(M + N) \end{bmatrix},$$

where $M = \widehat{B}_2 \widehat{Z}_1^\dagger + K_2\big(I - \widehat{Z}_1 \widehat{Z}_1^\dagger\big)$ and $N = \widehat{B}_1 \widehat{Z}_2^\dagger + K_1\big(I - \widehat{Z}_2 \widehat{Z}_2^\dagger\big)$ with arbitrary $K_1, K_2 \in \mathbb{C}^{m \times m}$. Moreover, $A = P \begin{bmatrix} & \widehat{B}_2 \widehat{Z}_1^\dagger \\ \widehat{B}_1 \widehat{Z}_2^\dagger & \end{bmatrix} P^H$ is the unique matrix of this form with minimum norm.

$$\boxed{\textbf{ELA}}$$

Let $\sigma_s = \min\limits_{A \in \mathscr{C}^{(a)}} \|AZ - B\|$ and $\mathcal{S}_s = \left\{ A \in \mathscr{C}^{(a)} \colon \|AZ - B\| = \sigma_s \right\}$.

THEOREM 4.7. *Assume the conditions of Theorem 4.6. Let* $\tilde{A} \in \mathbb{C}^{2m \times 2m}$ *and* M_1, N_1, M_2, N_2 *given in* (4.9). *Then*

(4.14)

$$\min_{A \in \mathcal{S}_s} \|\tilde{A} - A\| = \left(\|M_1\|^2 + \|N_1\|^2 + \left\| (M_2 \widehat{Z}_1 - \widehat{B}_2) \widehat{Z}_1^\dagger \right\|^2 + \left\| (N_2 \widehat{Z}_2 - \widehat{B}_1) \widehat{Z}_2^\dagger \right\|^2 \right)^{1/2},$$

and this minimum is attained if and only if

(4.15)
$$A = \tfrac{1}{2} \begin{bmatrix} M_4 + N_4 & (M_4 - N_4)i \\ (M_4 - N_4)i & -(M_4 + N_4) \end{bmatrix},$$

where $M_4 = \widehat{B}_2 \widehat{Z}_1^\dagger + M_2 \left(I - \widehat{Z}_1 \widehat{Z}_1^\dagger \right)$ *and* $N_4 = \widehat{B}_1 \widehat{Z}_2^\dagger + N_2 \left(I - \widehat{Z}_2 \widehat{Z}_2^\dagger \right)$.

4.4. Further results on residuals. The conditions for the existence of solutions $A \in \mathscr{C}$ and $A \in \mathscr{C}^{(a)}$ such that $\|AZ - B\| = 0$ are clear from (4.6) and (4.12), respectively. So, for $Z = Z_1 + Z_2$ with $Z_1 \in \mathcal{C}_{\mathcal{J}}^{2m \times k}$ and $Z_2 \in \mathcal{S}_{\mathcal{J}}^{2m \times k}$, if Z_1 and Z_2 have full rank, a null residual is achieved with matrices given by (4.7) and (4.13). Now, it also follows from Theorem 4.2 that consistency will be achieved for $A \in \mathscr{C}$ even when Z_1 and/or Z_2 are not full rank if the null spaces of $\widehat{Z}_1$ and $\widehat{Z}_2$ are contained in the null spaces of $\widehat{B}_1$ and $\widehat{B}_2$, respectively. For $A \in \mathscr{C}^{(a)}$, the conditions are $null(\widehat{Z}_1) \subseteq null(\widehat{B}_2)$ and $null(\widehat{Z}_2) \subseteq null(\widehat{B}_1)$. Observing that $null(\widehat{Z}_1) = null(Z_1)$, $null(\widehat{Z}_2) = null(Z_2)$, $null(\widehat{B}_1) = null(B_1)$ and $null(\widehat{B}_2) = null(B_2)$, the previous conditions may be stated in terms of the null spaces of Z_1, Z_2, B_1 and B_2.

In general, of course, it may exist a matrix $A \in \mathbb{C}^{n \times n}$ that maps Z onto B and no solution be possible with $A \in \mathscr{C}$ or $A \in \mathscr{C}^{(a)}$. In particular, if Z is in $\mathcal{C}_{\mathcal{J}}^{2m \times k}$ and is full rank, that is, $Z_1 = Z$ and $Z_2 = O$, then no $A_c \in \mathscr{C}$ can map Z onto $B_2 \neq O$ and no $A_s \in \mathscr{C}^{(a)}$ can map Z onto $B_1 \neq O$. In this case, again from (4.6) and (4.12), we have

(4.16)
$$\min_{A \in \mathscr{C}} \|AZ - B\| = \sqrt{2} \|\widehat{B}_2\| = \|B_2\|$$

and

(4.17)
$$\min_{A \in \mathscr{C}^{(a)}} \|AZ - B\| = \sqrt{2} \|\widehat{B}_1\| = \|B_1\|.$$

Analogously, if Z is full rank and is in $\mathcal{S}_{\mathcal{J}}^{2m \times k}$, that is, $Z_1 = O$ and $Z_2 = Z$, then, from (4.12) we have

(4.18)
$$\min_{A \in \mathscr{C}} \|AZ - B\| = \|B_1\| \quad \text{and} \quad \min_{A \in \mathscr{C}^{(a)}} \|AZ - B\| = \|B_2\|.$$

ELA

Obviously, it is always the case that $rank(Z_1) \leq rank(Z)$ and $\mathrm{rank}(Z_2) \leq rank(Z)$. For the particular case of these ranks being equal, we have

THEOREM 4.8. *If Z, Z_1 and Z_2 have the same rank then $\min \|AZ - B\|$ is the same for $A \in \mathscr{C}$, $A \in \mathscr{C}^{(a)}$ or general A.*

Proof. From $Zv = 0 \Rightarrow Z_1 v = \frac{1}{2}(Z + iJZ)v = 0$ we conclude that $null(Z) \subseteq null(Z_1)$; this and the hypothesis $rank(Z_1) = rank(Z)$ allows us to state that those null spaces are the same. Similarly, we may say that $null(Z) = null(Z_2)$. Therefore, the right singular vectors corresponding to null singular values of Z, Z_1 (or $\widehat{Z}_1$) and Z_2 (or $\widehat{Z}_2$) are the same. This implies (see Section 5.5) that

$$(4.19) \qquad \left(I - Z^\dagger Z \right) = \left(I - \widehat{Z}_1^\dagger \widehat{Z}_1 \right) = \left(I - \widehat{Z}_2^\dagger \widehat{Z}_2 \right)$$

and the residuals expressed in (4.6) and (4.12) are the same. Now, for a general A we write

$$
\begin{aligned}
\|B(I - Z^\dagger Z)\|^2 &= \|P(B_1 + B_2)(I - Z^\dagger Z)\|^2 \\
&= 2 \left\| \begin{bmatrix} \widehat{B}_2 \\ \widehat{B}_1 \end{bmatrix} (I - Z^\dagger Z) \right\|^2 \\
&= 2 \left(\left\| \widehat{B}_2(I - Z^\dagger Z) \right\|^2 + \left\| \widehat{B}_1(I - Z^\dagger Z) \right\|^2 \right)
\end{aligned}
$$

and, again from (4.19) conclude that the residual is equal to those in (4.6) and (4.12). $\Box$

We end this section observing that when Z and B are real, the ranks of Z, Z_1 and Z_2 are always the same and the previous theorem applies, i.e., there are always real matrices in $\mathscr{C}$ and in $\mathscr{C}^{(a)}$ that attain the same minimum residual as a general matrix. This will be shown in Section 5.3.

5. Algorithms. In this section, we propose algorithms for computing the solutions of Problems I and II, first the case of $A \in \mathscr{C}$ and then the case of $A \in \mathscr{C}^{(a)}$. We also present an algorithm to compute the solutions of Problems I and II for an arbitrary matrix $A \in \mathbb{C}^{2m \times 2m}$, the unstructured minimization problem.

5.1. Unstructured minimization. An essential part of the computation of A in (4.2) and the minimum value in (4.1) is the full SVD computation of $Z \in \mathbb{C}^{q \times k}$, with $q \geq k$,

$$(5.1) \qquad Z = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^H,$$

where, for $rank(Z) = r < k$, $U_1 \in \mathbb{C}^{q \times r}$, $U_2 \in \mathbb{C}^{q \times (q-r)}$, $V_1 \in \mathbb{C}^{k \times r}$, $V_2 \in \mathbb{C}^{k \times (k-r)}$ have orthonormal columns and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$, with $\sigma_i > 0$, $i = 1, \ldots, r$, the

ELA

nonzero singular values of Z. When Z is full rank $(r = k)$, (5.1) gives

$$(5.2) \qquad Z = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma \\ O \end{bmatrix} V_1^H$$

if $q > k$, and gives place to

$$(5.3) \qquad Z = U_1 \Sigma V_1^H,$$

if $q = k$. Thus, in the full rank case, V_2 does not exist and, if $p = k$, U_2 does not exist either. The following expressions are always valid if we make the convention that in the full rank case $V_2 V_2^H$ and $V_2^H V_2$ are null matrices and, if $p = k$, $U_2 U_2^H$ is also a null matrix. This convention will be used throughout Section 5, including the structured cases.

The Moore-Penrose inverse of Z is

$$(5.4) \qquad Z^\dagger = V_1 \Sigma^{-1} U_1^H.$$

We have

$$(5.5) \qquad \left(I - Z^\dagger Z\right) = (I - V_1 \Sigma^{-1} U_1^H U_1 \Sigma V_1^H) = (I - V_1 V_1^H) = V_2 V_2^H$$

and

$$(5.6) \qquad \left(I - Z Z^\dagger\right) = \left(I - U_1 \Sigma V_1^H V_1 \Sigma^{-1} U_1^H\right) = (I - U_1 U_1^H) = U_2 U_2^H.$$

Entering (5.6) in (4.2), we get

$$(5.7) \qquad A = B V_1 \Sigma^{-1} U_1^H + K U_2 U_2^H,$$

with $K \in \mathbb{C}^{2m \times 2m}$ arbitrary. Using (5.5) in (4.1) we get

$$(5.8) \qquad \min_{A \in \mathbb{C}^{2m \times 2m}} \|AZ - B\| = \|B V_2 V_2^H\|.$$

Denoting this value by σ and defining

$$\mathcal{S} = \left\{ A \in \mathbb{C}^{2m \times 2m} : \|AZ - B\| = \sigma \right\},$$

in Problem II, for a given $\widetilde{A} \in \mathbb{C}^{2m \times 2m}$, we want to compute

$$\min_{A \in \mathcal{S}} \left\| \widetilde{A} - A \right\| = \min_{K \in \mathbb{C}^{2m \times 2m}} \| \widetilde{A} - B V_1 \Sigma^{-1} U_1^H - K U_2 U_2^H \|.$$

Since $\left(U_2 U_2^H\right)^2 = U_2 U_2^H = \left(U_2 U_2^H\right)^H$, we can apply Lemma 4.3 to conclude that the unique matrix K which gives the minimum norm is

$$K = \widetilde{A} - B V_1 \Sigma^{-1} U_1^H.$$

ELA

Thus, $\min_{A \in \mathcal{S}} \left\| \widetilde{A} - A \right\|$ is attained for

$$A = BV_1 \Sigma^{-1} U_1^H + \left(\widetilde{A} - BV_1 \Sigma^{-1} U_1^H \right) U_2 U_2^H = BV_1 \Sigma^{-1} U_1^H + \widetilde{A} U_2 U_2^H.$$

Notice that this is the expression in (5.7) with $K = \widetilde{A}$.

The following algorithm computes the solutions for Problem I (steps 1 and 2) and Problem II (steps 1 and 3) for an arbitrary matrix $A \in \mathbb{C}^{2m \times 2m}$.

ALGORITHM 1.

Given Z, B, K and $\widetilde{A}$, compute:

1. the full SVD of Z,

$$Z = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^H ;$$

2. $A = BV_1 \Sigma^{-1} U_1^H + KU_2 U_2^H$ and the minimum residual $\| (BV_2) V_2^H \|$, according to (5.8);
3. $A = BV_1 \Sigma^{-1} U_1^H + \widetilde{A} U_2 U_2^H$ and the minimum value $\left\| \widetilde{A} - A \right\|$.

To give an estimate of the number of complex arithmetic operations involved in the algorithm we will assume that m is significantly larger than k. In step 1 the number of flops for the complete SVD is about $16m^2k + 22k^3$ and $12mk^2 + 20k^3$ if U_2 is not required (using the R-SVD) [4, p. 254]. In step 2, the first term in the expression for A requires about $8m^2r$ and the second term is more expensive since it requires about $16m^3$ flops. Therefore, the solution A of minimal norm for unstructured Problem I, which occurs for the choice $K = O$, is the cheapest to compute, requiring a total of $12mk^2 + 8m^2r + 20k^3$ flops. The computation of the residual costs an extra $8mk(k - r)$ flops, approximately, which is rather cheap when r is close to k. A direct computation of $\|AZ - B\|$ is more expensive (it costs about $8m^2k$ flops) and in our numerical experiments we did not find significant differences in the size of the computed residuals.

The cost of computing A in step 3 is equal to the corresponding cost of step 2 when $K \neq O$. When Problem II is to be solved after Problem I we may reuse the first term in the expression for A in step 2, which is the same in both problems, and also the product $U_2 U_2^H$, if it has been computed before. The computation of the minimum value $\|\widetilde{A} - A\|$ costs about $8m^2$ flops.

5.2. Minimization for $A \in \mathscr{C}$ and $A \in \mathscr{C}^{(a)}$. Let $Z, B \in \mathbb{C}^{2m \times k}$ and P be given in (3.1). Consider $\widehat{Z}_1, \widehat{Z}_2, \widehat{B}_1, \widehat{B}_2$ given in (4.5). Let $r_1 = \mathrm{rank}(\widehat{Z}_1)$ and the full

$$\boxed{\textbf{ELA}}$$

SVD of $\widehat{Z}_1$ be

$$(5.9) \qquad \widehat{Z}_1 = \begin{bmatrix} \widehat{U}_1 & \widehat{U}_2 \end{bmatrix} \begin{bmatrix} \widehat{\Sigma} & O \\ O & O \end{bmatrix} \begin{bmatrix} \widehat{V}_1 & \widehat{V}_2 \end{bmatrix}^H = \widehat{U}_1 \widehat{\Sigma} \widehat{V}_1^H,$$

where $\widehat{U}_1 \in \mathbb{C}^{m \times r_1}$, $\widehat{U}_2 \in \mathbb{C}^{m \times (m-r_1)}$, $\widehat{V}_1 \in \mathbb{C}^{k \times r_1}$, $\widehat{V}_2 \in \mathbb{C}^{k \times (k-r_1)}$ have orthonormal columns and $\widehat{\Sigma} = \operatorname{diag}(\widehat{\sigma}_1, \ldots, \widehat{\sigma}_{r_1})$, with $\widehat{\sigma}_i > 0$, $i = 1, \ldots, r_1$, the nonzero singular values of $\widehat{Z}_1$. Analogously, let $r_2 = \operatorname{rank}(\widehat{Z}_2)$ and the SVD of $\widehat{Z}_2$ be

$$(5.10) \qquad \widehat{Z}_2 = \begin{bmatrix} \widetilde{U}_1 & \widetilde{U}_2 \end{bmatrix} \begin{bmatrix} \widetilde{\Sigma} & O \\ O & O \end{bmatrix} \begin{bmatrix} \widetilde{V}_1 & \widetilde{V}_2 \end{bmatrix}^H = \widetilde{U}_1 \widetilde{\Sigma} \widetilde{V}_1^H,$$

where $\widetilde{U}_1 \in \mathbb{C}^{m \times r_2}$, $\widetilde{U}_2 \in \mathbb{C}^{m \times (m-r_2)}$, $\widetilde{V}_1 \in \mathbb{C}^{k \times r_2}$, $\widetilde{V}_2 \in \mathbb{C}^{k \times (k-r_2)}$ have orthonormal columns and $\widetilde{\Sigma} = \operatorname{diag}(\widetilde{\sigma}_1, \ldots, \widetilde{\sigma}_r)$, with $\widetilde{\sigma}_i > 0$, $i = 1, \ldots, r_2$, the nonzero singular values of $\widehat{Z}_2$. Then, the Moore-Penrose inverses of $\widehat{Z}_1$ and $\widehat{Z}_2$ are

$$\widehat{Z}_1^\dagger = \widehat{V}_1 \widehat{\Sigma}^{-1} \widehat{U}_1^H \quad \text{and} \quad \widehat{Z}_2^\dagger = \widetilde{V}_1 \widetilde{\Sigma}^{-1} \widetilde{U}_1^H.$$

If both $\widehat{Z}_1^\dagger$ and $\widehat{Z}_2^\dagger$ are rank-deficient ($r_1, r_2 < k$), then we have

$$(5.11) \qquad \left(I - \widehat{Z}_1^\dagger \widehat{Z}_1 \right) = \widehat{V}_2 \widehat{V}_2^H, \quad \left(I - \widehat{Z}_2^\dagger \widehat{Z}_2 \right) = \widetilde{V}_2 \widetilde{V}_2^H$$

and, according to Theorem 4.4 and (5.11),

$$(5.12) \qquad \min_{A \in \mathscr{C}} \| AZ - B \| = \sqrt{2} \left(\left\| \widehat{B}_1 \widehat{V}_2 \widehat{V}_2^H \right\|^2 + \left\| \widehat{B}_2 \widetilde{V}_2 \widetilde{V}_2^H \right\|^2 \right)^{1/2}.$$

In the full rank case ($r_1 = r_2 = k$), $\left(I - \widehat{Z}_1^\dagger \widehat{Z}_1 \right) = \left(I - \widehat{Z}_2^\dagger \widehat{Z}_2 \right) = O$ and we have a null residual. In both cases,

$$(5.13) \qquad \left(I - \widehat{Z}_1 \widehat{Z}_1^\dagger \right) = \widehat{U}_2 \widehat{U}_2^H, \quad \left(I - \widehat{Z}_2 \widehat{Z}_2^\dagger \right) = \widetilde{U}_2 \widetilde{U}_2^H$$

and, using (4.7) in Theorem 4.4, this minimum is attained with

$$(5.14) \qquad A = \tfrac{1}{2} \begin{bmatrix} M + N & -(M-N)i \\ (M-N)i & M+N \end{bmatrix},$$

where

$$(5.15) \qquad M = \widehat{B}_2 \widetilde{V}_1 \widetilde{\Sigma}^{-1} \widetilde{U}_1^H + K_2 \widetilde{U}_2 \widetilde{U}_2^H \quad \text{and} \quad N = \widehat{B}_1 \widehat{V}_1 \widehat{\Sigma}^{-1} \widehat{U}_1^H + K_1 \widehat{U}_2 \widehat{U}_2^H,$$

with $K_1, K_2 \in \mathbb{C}^{m \times m}$ arbitrary.

Now we will consider the computation of the solution for problem II. Consider $\widetilde{A} \in \mathbb{C}^{2m \times 2m}$, $\widetilde{A}_1$, given in (4.8) and M_1 and N_1 given in (4.9). Then, according to Theorem 4.5, A is given by (5.14), where in (5.15) $K_2 = M_1$ and $K_1 = N_1$.

$$\boxed{\textbf{ELA}}$$

The following algorithm computes the solutions for Problem I (steps 1, 2 and 3) and Problem II (steps 1, 2, 4 and 5) for $A \in \mathscr{C}$.

ALGORITHM 2.

Given Z, B, K_1, K_2 and $\widetilde{A}$, compute:

1. $\widehat{Z}_1, \widehat{Z}_2, \widehat{B}_1$ and $\widehat{B}_2$, in (4.5);
2. the full SVD's of $\widehat{Z}_1$ and $\widehat{Z}_2$, in (5.9) and (5.10);
3. M and N in (5.15), for the given K_1 and K_2, A in (5.14) and the residual with (5.12);
4. M_1 and N_1 given in (4.9);
5. M, N and A (as in step 3) for $K_1 = M_1$ and $K_2 = N_1$ and the minimum value $\left\| \widetilde{A} - A \right\|$.

For $A \in \mathscr{C}^{(a)}$, the algorithm is similar to Algorithm 2 and we just point out the differences in the expressions to be used in this case. In step 3, we use

$$M = \widehat{B}_2 \widehat{V}_1 \widehat{\Sigma}^{-1} \widehat{U}_1^H + K_2 \widehat{U}_2 \widehat{U}_2^H,$$
$$N = \widehat{B}_1 \widetilde{V}_1 \widetilde{\Sigma}^{-1} \widetilde{U}_1^H + K_1 \widetilde{U}_2 \widetilde{U}_2^H,$$
$$A = \tfrac{1}{2} \begin{bmatrix} M + N & (M - N)i \\ (M - N)i & -(M + N) \end{bmatrix},$$
$$\min_{A \in \mathscr{C}^{(a)}} \|AZ - B\| = \sqrt{2}\left(\left\| \widehat{B}_2 \widehat{V}_2 \widehat{V}_2^H \right\|^2 + \left\| \widehat{B}_1 \widetilde{V}_2 \widetilde{V}_2^H \right\|^2 \right)^{1/2}.$$

In step 4, use (4.9) for M_2 and N_2. In step 5, compute M, N and A as in step 3 for $K_1 = N_2$ and $K_2 = M_2$.

The costs of steps 1 and 4 are relatively inexpensive. In step 2 the number of flops for each one of the two complete SVD's is about $4m^2k + 22k^3$ and $6mk^2 + 20k^3$ if $\widehat{U}_2$ and $\widetilde{U}_2$ are not required. In step 3, to compute M and N we need $2m^2r_1$ and $2m^2r_2$ flops, respectively, if $K_1 = K_2 = O$ which, as in the unstructured case, produce the minimal norm solution. Since $r_1 + r_2 \leq 2r$, the total cost to compute M and N, in this case, is not larger than $4m^2r$ flops and this is half the cost of computing A in the unstructured case. For general matrices K_1 and K_2, the computation of M and N requires in addition $4m^3$ extra flops in contrast to $16m^3$ flops in the general case. The explicit construction of A requires $2m^2$ flops. The computation of the residual costs an extra $4mk(k - r_1) + 4mk(k - r_2)$ flops, approximately. If $r_1 = r_2 = r$, this equal the corresponding cost in the unstructured case; however, if $r_1 \ll r$ or $r_2 \ll r$, this cost may be larger in the structured case (the extreme situation occurs when $r = k$ and r_1 or r_2 are equal to zero).

The cost of computing M, N and A in step 5 equals the corresponding cost in step 3 when $K_1, K_2 \neq O$. As in the unstructured case, when Problem II is to be solved after Problem I, we may reuse the first terms in the expressions for M and N in step 3, which are the same in both problems, and also the products $\widehat{U}_2 \widehat{U}_2^H$ and $\widetilde{U}_2 \widetilde{U}_2^H$, if computed before. The computation of the minimum value $\|\widetilde{A} - A\|$ costs about $8m^2$ flops.

In conclusion, to produce structured solutions for Problems I and II, our algorithm requires less arithmetic than the algorithm for the general case. As noted before, to compute the minimal norm solution for Problem I, our algorithm is about half as expensive as the general algorithm. For solutions with general matrices K, K_1 and K_2 (which is always the case in Problem II), our algorithm requires approximately $1/4$ of the number of flops of the general algorithm.

5.3. The real case. Although Algorithm 1 and Algorithm 2 have been presented for complex matrices, there are situations where one may be interested in producing a real solution A (we noted before that real orthogonal sympletic matrices belong to $\mathscr{C}$). When Z and B are real, Algorithm 1 produces necessarily real matrices A since all the arithmetic involved is real. This is not the case in Algorithm 2. From (5.14) it follows that A is real if and only if M and N in (5.15) are complex conjugate. This happens whenever Z and B are real and we take K_1 and K_2 in (5.15) with $K_1 = \overline{K}_2$. Because (Z_1, Z_2) and (B_1, B_2), in (4.3) and (4.4), respectively, are pairs of conjugate matrices, the same is true for $\widehat{Z}_1$ and $\widehat{Z}_2$ and the corresponding factors in their decompositions in (5.9) and (5.10), respectively. As a consequence of this, Algorithm 2 gets the following simplifications when Z and B are real: in step 2, one single complex SVD (either of $\widehat{Z}_1$ or $\widehat{Z}_2$) is required and in step 3 only one of M and N is computed. Furthermore, in Problem II, for a real $\widetilde{A}$, M_1 and N_1 in step 4 are also conjugate (see Lemma 3.1). The amount of computation is also reduced in step 5 similarly to step 3.

6. Numerical examples. We implemented our algorithms in MATLAB 7.5.0342 (R2007b), with rounding error unit equal to $2^{-53} \approx 1.1 \times 10^{-16}$. For Problem I, we computed the minimum norm solutions with null matrices $K, K1$ and $K2$. In Table 6.1 we summarize the numerical results obtained with our codes for 15 given pairs (Z, B). Here, as before, $A_c \in \mathscr{C}$, $A_s \in \mathscr{C}^{(a)}$ and A is a general matrix. The residuals $\mathcal{O}(10^{-12})$, $\mathcal{O}(10^{-13})$ and $\mathcal{O}(10^{-14})$ are due to rounding errors since they correspond to cases for which null residuals are expected, in accordance with Theorem 4.2 and our analysis in section 4.4.

In the first 3 cases, matrices Z and B are generated with MATLAB function rand. The produced matrix Z, as well as the projections Z_1 and Z_2 are full rank and consistency is always expected.

$$\boxed{\textbf{ELA}}$$

TABLE 6.1
Minimal residuals for problem I.

	m,k	$\|AZ-B\|$	$\|A_cZ-B\|$	$\|A_sZ-B\|$	$\|A\|$	$\|A_c\|$	$\|A_s\|$
1	$10,10$	4.9×10^{-14}	3.3×10^{-14}	3.2×10^{-14}	4.37	19.02	19.96
2	$100,20$	6.7×10^{-14}	1.1×10^{-13}	1.2×10^{-13}	4.73	7.09	7.06
3	$1000,200$	1.0×10^{-12}	1.2×10^{-12}	1.2×10^{-12}	18.26	44.57	45.00
4	$10,10$	6.8×10^{-14}	7.88	8.39	11.45	8.34	7.84
5	$100,20$	5.7×10^{-14}	36.39	36.40	7.01	4.93	4.99
6	$1000,200$	1.6×10^{-12}	365.42	365.14	22.38	15.82	15.83
7	$100,10$	1.6×10^{-14}	1.4×10^{-14}	26.14	2.47	2.47	0
8	$10,10$	3.80	3.80	3.80	3.56	8.45	8.76
9	$50,10$	8.31	18.82	18.83	4.76	3.38	3.35
10	$100,20$	11.76	11.76	11.76	4.60	6.85	6.79
11	$1000,200$	36.93	366.41	365.73	22.32	15.77	15.80
12	$20,20$	3.35	3.35	3.35	4.95	13.12	34.13
13	$50,10$	1.4×10^{-14}	18.03	17.96	3.76	2.64	2.68
14	$100,20$	6.4×10^{-14}	2.3×10^{-13}	2.5×10^{-13}	4.13	6.16	6.10
15	$100,20$	2.1×10^{-13}	32.55	32.67	6.08	4.22	4.38

In cases 4–7, Z has full rank and is either in the subspace $\mathcal{C}_{\mathcal{J}}^{2m\times k}$ ($Z_2 = O$) or in the subspace $\mathcal{S}_{\mathcal{J}}^{2m\times k}$ ($Z_1 = O$). Here, matrices Z were generated according to (4.5) from random ($m \times k$) matrices $\widehat{Z}_1$ and $\widehat{Z}_2$. The computed residuals are in accordance with (4.16) - (4.18). In particular, in case 7 we used $Z_2 = B_2 = O$ and Algorithm 1 produced the same $A_c \in \mathscr{C}$ as Algorithm 2. This is because the general solution with minimum norm in this case is necessarily in $\mathscr{C}$. With $A = \widetilde{A}_c + \widetilde{A}_s$, $\widetilde{A}_c \in \mathscr{C}$ and $\widetilde{A}_s \in \mathscr{C}^{(a)}$, we have

$$\|AZ - B\|^2 = \left\| \left(\widetilde{A}_c + \widetilde{A}_s \right)(Z_1 + Z_2) - (B_1 + B_2) \right\|^2$$

$$= \left\| \widetilde{A}_c Z_1 + \widetilde{A}_s Z_2 - B_1 \right\|^2 + \left\| \widetilde{A}_c Z_2 + \widetilde{A}_s Z_1 - B_2 \right\|^2$$

$$= \left\| \widetilde{A}_c Z_1 - B_1 \right\|^2 + \left\| \widetilde{A}_s Z_1 \right\|^2$$

and the minimum residual is attained with $\widetilde{A}_s = O$ and $A = \widetilde{A}_c$ that minimizes $\left\| \widetilde{A}_c Z_1 - B_1 \right\|$. That $A_s = O$ also minimizes $\|A_sZ - B\|^2$ in this case ($Z_2 = B_2 = O$) is a consequence of its expression in Theorem 4.6.

Cases 8–11 are examples of Z being rank-deficient (by making copies of random columns) and B full rank. The problem is never consistent since AZ can not be full rank.

ELA

In cases 12–15, Z and B are both rank deficient and the condition $null(Z) \subseteq null(B)$ in Theorem 4.2 only fails in case 12. In case 14, we also have $null(Z_1) \subseteq null(B_1)$ and $null(Z_2) \subseteq null(B_2)$. We adopted a simple method to generate rank deficient consistent problems. For instance, in case 13 the first 6 independent columns of Z and B have been generated randomly and then, for $i = 7, \ldots, 10$, the ith column of Z and B is computed using the same linear combination of the first six columns of Z and B, respectively.

In cases 8, 10, 12 and 14, the ranks of Z, Z_1 and Z_2 are the same and, in accordance with Theorem 4.8, we have equal residuals with A, A_c and A_s.

The next table shows results for Problem II with matrix $\widetilde{A}$ generated randomly.

TABLE 6.2
Minimal residuals for problem II.

	m, k	$\|\widetilde{A} - A\|$	$\|\widetilde{A} - A_c\|$	$\|\widetilde{A} - A_s\|$
1	$10, 10$	15.47	21.97	21.62
2	$50, 50$	92.89	86.68	88.11
3	$100, 10$	71.91	98.90	115.25
4	$100, 20$	141.96	153.67	153.66
5	$1000, 200$	758.97	1201.40	1001.82

When the smallest residual $\|AZ - B\|$ for unstructured A is also attained for structured A_c (A_s) it is obvious that $\|\widetilde{A} - A\| \leq \|\widetilde{A} - A_c\|$ and $\|\widetilde{A} - A\| \leq \|\widetilde{A} - A_s\|$ because the set of solutions A_c (A_s) is contained in the set of general solutions A. In case 2 the problem is consistent for A but not for A_c or A_s. In case 3, $\widetilde{A} \in \mathscr{C}$ but does not belong to the set of solutions A_c which give minimum residual $\|A_c Z - B\|$ (otherwise, $\|\widetilde{A} - A_c\|$ would be zero). In case 5, $\widetilde{A} \in \mathscr{C}^{(a)}$ and a similar observation can be made.

7. Conclusions. The problems addressed have been studied before by other authors for different types of structured matrices. We have extended the previous work to other structures which are exhibited by certain matrices that occur in applications. More precisely, for given complex matrices Z and B, we proposed algorithms to compute $A \in \mathscr{C}$ (centralizer of $\mathcal{J}$) or $A \in \mathscr{C}^{(a)}$ (anti-centralizer of $\mathcal{J}$) that minimize $\|AZ - B\|$. We proved several new theoretical results and illustrated them with numerical examples. Also, for such solutions A and for a given matrix $\tilde{A}$, we have considered the problem of minimizing $\|\tilde{A} - A\|$. The key idea of our work has been the use of special orthogonal subspaces that allow the decomposition of each one the minimization problems into two independent subproblems. As a consequence of this, our algorithms not only deliver solutions which are of the required type but they need less arithmetic than previous algorithms for unstructured matrices.

ELA

REFERENCES

[1]   J.R. Bunch. The weak and strong stability of algorithms in Numerical Linear Algebra. *Linear Algebra Appl.*, 88:49–60, 1987.

[2]   A. Bunse-Gerstner, R. Byers, and V. Mehrmann. A chart of numerical methods for structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 13:419–453, 1992.

[3]   J.W. Demmel. *Applied Numerical Linear Algebra*. SIAM, Philadelphia, 1997.

[4]   G.H. Golub and C.F. Van Loan. *Matrix Computations*, third edition. Johns Hopkins University Press, Baltimore, 1996.

[5]   Alan J. Laub. *Matrix Analysis for Scientists and Engineers*. SIAM, Philadelphia, 2005.

[6]   Z.Y. Liu and H. Fassbender. An inverse eigenvalue problem and an associated approximation problem for generalized K-centrohermitian matrices. *J. Comput. Appl. Math.*, 206(1):578–585, 2007.

[7]   Z. Liu, Z. Tian, and Y. Tan. Computing the least-square solutions for centrohermitian matrix problems. *Appl. Math. Comput.*, 174:566–577, 2006.

[8]   D.S. Mackey, N. Mackey, and F. Tisseur. Structured tools for structured matrices. *Electron. J. Linear Algebra*, 10:106–145, 2003.

[9]   D.S. Mackey, N. Mackey, and F. Tisseur. Structured factorizations in scalar product spaces. *SIAM J. Matrix Anal. Appl.*, 27:821–850, 2006.

[10] F. Tisseur. A chart of backward errors for singly and doubly structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 24:877–897, 2003.

[11] W.F. Trench. Hermitian, hermitian R-symmetric, and hermitian R-skew symmetric procrustes problems. *Linear Algebra Appl.*, 387:83–98, 2004.

[12] W.F. Trench. Minimization problems for (R,S)-symmetric matrices and (R,S)-skew symmetric matrices. *Linear Algebra Appl.*, 389:23–31, 2004.

[13] F. Zhou, L. Zhang, and Z. Hu. Least-square solution for inverse problems of centrosymmetric matrices. *Comput. Math. Appl.*, 45:1581–1589, 2003.